

THESIS FOR THE DEGREE OF DOCTOR OF PHILOSOPHY

# Topics in convex and mixed binary linear optimization

Emil Gustavsson

**CHALMERS**



**UNIVERSITY OF GOTHENBURG**

Department of Mathematical Sciences,  
Chalmers University of Technology and University of Gothenburg  
SE-412 96 Göteborg, Sweden

Göteborg, 2015

Topics in convex and mixed binary linear optimization

Emil Gustavsson

ISBN 978-91-628-9409-2 (printed version)

ISBN 978-91-628-9410-8 (electronic version)

© Emil Gustavsson, 2015

Department of Mathematical Sciences

Chalmers University of Technology and University of Gothenburg

SE-412 96 Göteborg

Sweden

Telephone +46 (0)31 772 1000

Printed in Göteborg, Sweden 2015

# Topics in convex and mixed binary linear optimization

Emil Gustavsson

Department of Mathematical Sciences,  
Chalmers University of Technology and University of Gothenburg

## Abstract

This thesis concerns theory, algorithms, and applications for two problem classes within the realm of mathematical optimization; *convex optimization* and *mixed binary linear optimization*. To the thesis is appended five papers containing its main contributions.

In the first paper a subgradient optimization method is applied to the Lagrangian dual of a general convex and (possibly) nonsmooth optimization problem. The classic dual subgradient method produces primal solutions that are, however, neither optimal nor feasible. Yet, convergence to the set of optimal primal solutions can be obtained by constructing a class of *ergodic sequences* of the Lagrangian subproblem solutions. We generalize previous convergence results for such ergodic sequences by proposing a new set of rules for choosing the convexity weights defining the sequences. Numerical results indicate that by applying our new set of rules primal feasible solutions of higher quality than those created by the previously developed rules are achieved.

The second paper analyzes the properties of a subgradient method when applied to the *Lagrangian dual* of an infeasible convex program. The primal-dual pair of programs corresponding to an associated homogeneous dual function is shown to be in turn associated with a saddle-point problem, in which the primal part amounts to finding a solution such that the Euclidean norm of the infeasibility in the relaxed constraints is minimized. Convergence results for a conditional dual subgradient optimization method applied to the Lagrangian dual problem is presented. The sequence of ergodic primal iterates is shown to converge to the set of solutions to the primal part of the associated saddle-point problem.

The third paper applies a dual subgradient method to a general *mixed binary linear program* (MBLP). The resulting sequence of primal ergodic iterates is shown to converge to the set of solutions to a convexified version of the original MBLP, and three procedures for utilizing the primal ergodic iterates for constructing feasible solutions to the MBLP are proposed: a Lagrangian heuristic, the construction of a so-called core problem, and a framework for utilizing the ergodic primal iterates within a branch-and-bound algorithm. Numerical results for samples of uncapacitated facility location problems and set covering problems indicate that the proposed procedures are practically useful for solving structured MBLPs.

In the fourth paper, the *preventive maintenance scheduling problem with interval costs* is studied. This problem considers the scheduling of maintenance of the components in a multi-component system with the objective to minimize the sum of the set-up and interval costs for the system over a finite time period. The problem is shown to be NP-hard, and an MBLP model is introduced and utilized in three case studies from the railway, aircraft, and wind power industries.

In the fifth paper an MBLP model for the optimal scheduling of tamping operations on ballasted rail tracks is introduced. The objective is to minimize the total maintenance costs while maintaining an acceptable condition on the ballasted tracks. The model is thoroughly analyzed and the scheduling problem considered is shown to be NP-hard. A computational study shows that the total cost for maintenance can be reduced by up to 10% as compared with the best policy investigated.

**Keywords:** subgradient methods, Lagrangian dual, recovery of primal solutions, inconsistent convex programs, ergodic sequences, convex optimization, mixed binary linear optimization, maintenance scheduling, preventive maintenance, deterioration cost

### Appended papers:

**Paper I:** Gustavsson, E., Patriksson, M., Strömberg, A.-B., *Primal convergence from dual subgradient methods for convex optimization*, *Mathematical Programming* **150**(2) (2015), pp. 365–390.

**Paper II:** Önnheim, M., Gustavsson, E., Strömberg, A.-B., Patriksson, M., Larsson, T., *Ergodic, primal convergence in dual subgradient schemes for convex programming, II—the case of inconsistent primal problems*, submitted to *Mathematical Programming*.

**Paper III:** Gustavsson, E., Larsson, T., Patriksson, M., Strömberg, A.-B., *Recovery of primal solutions from dual subgradient methods for mixed binary linear programs*, preprint.

**Paper IV:** Gustavsson, E., Patriksson, M., Strömberg, A.-B., Wojciechowski, A., Önnheim, M., *Preventive maintenance scheduling of multi-component systems with interval costs*, *Computers & Industrial Engineering* **76** (2014), pp. 390–400.

**Paper V:** Gustavsson, E., *Scheduling tamping operations on railway tracks using mixed integer linear programming*, *EURO Journal on Transportation and Logistics* **4**, Issue 1: *Transportation Infrastructure Management* (2015), pp. 97–112.

### Publications not included in the thesis:

Andréasson, N., Evgrafov, A., Patriksson, M., with Gustavsson, E. and Önnheim, M., *Introduction to Continuous Optimization: Second Edition*, Studentlitteratur, Lund, 2013.



## Acknowledgements

I would like to start by expressing my sincerest gratitude towards my supervisor Ann-Brith Strömberg and my co-supervisors Michael Patriksson and Anders Ekberg for their guidance through the work of this thesis. Without your invaluable input this thesis would not have been accomplished. Thanks especially to Ann-Brith for all the late nights you have spent reading my manuscripts.

Many thanks to all past and present members of the optimization group for sharing a common passion for optimizing stuff and for making our corridor a welcoming place. All my co-workers at the Department of Mathematical Sciences deserve a sincere thank you for creating the creative and stimulating working environment that the department offers. Thanks also to the members of the research project MU27 at CHARMEC (Chalmers Railway Mechanics) for a great research collaboration.

Finally, I wish to thank Magdalena, all my friends, my parents, and my brother for their continuous support and unconditional love during all these years.

Emil Gustavsson  
Göteborg, April 2015





# Contents

<b>1</b>	<b>Introduction</b>	<b>1</b>
1.1	Outline . . . . .	1
<b>2</b>	<b>Convex optimization</b>	<b>2</b>
2.1	Problem properties . . . . .	2
2.1.1	Optimality conditions . . . . .	4
2.1.2	Lagrangian duality . . . . .	5
2.1.3	Non-coordinability of non-strictly convex problems . . . . .	7
2.2	Solution procedures . . . . .	7
2.2.1	Subgradient methods . . . . .	8
2.2.2	Dual subgradient methods . . . . .	9
2.2.3	Cutting-plane methods . . . . .	10
2.2.4	Bundle methods . . . . .	11
2.3	Applications and modelling . . . . .	11
<b>3</b>	<b>Mixed binary linear optimization</b>	<b>13</b>
3.1	Problem properties . . . . .	13
3.2	The importance of good modelling . . . . .	16
3.3	Solution procedures . . . . .	17
3.3.1	Branch-and-bound methods . . . . .	17
3.3.2	Dual subgradient methods and Lagrangian heuristics . . . . .	18
3.4	Applications to maintenance scheduling . . . . .	19
3.5	Other applications . . . . .	20
<b>4</b>	<b>Summary and conclusions</b>	<b>21</b>
4.1	Contributions of the thesis . . . . .	21
4.2	Future research . . . . .	21
<b>5</b>	<b>Summary of the appended papers</b>	<b>23</b>
5.1	<b>Paper I:</b> Primal convergence from dual subgradient methods for convex optimization . . . . .	23
5.2	<b>Paper II:</b> Ergodic, primal convergence in dual subgradient schemes for convex programming, II—the case of inconsistent primal problems . . . . .	24
5.3	<b>Paper III:</b> Recovery of primal solutions from dual subgradient methods for mixed binary linear programs . . . . .	24
5.4	<b>Paper IV:</b> Preventive maintenance scheduling of multi-component systems with interval costs . . . . .	25
5.5	<b>Paper V:</b> Scheduling tamping operations on railway tracks using mixed integer linear programming . . . . .	26



# 1 Introduction

This thesis covers two kinds of problems within *mathematical optimization*; *convex programs* (CPs) and *mixed binary linear programs* (MBLPs). Mathematical optimization can be described as the scientific discipline in which one tries to find the best decision from a set of available alternatives in a given quantitative context. In order to define what is meant by a "best decision", the notion of an *objective function* is needed. An objective function  $f : X \mapsto \mathbb{R}$  is a function from a set  $X$  to the set of real numbers and which defines the *objective value*  $f(x)$  of a *decision*  $x \in X$ . We say that  $x$  is *feasible* if  $x \in X$ . The notion of a *best* decision is then defined as a feasible decision which possesses either the maximum or the minimum possible objective value. For the case when the best decision is defined to be the one with minimum objective value, an *optimization problem* is defined as the problem to

$$\text{minimize } f(x), \tag{1a}$$

$$\text{subject to } x \in X. \tag{1b}$$

The problem (1) is thus the problem of finding a feasible decision  $x \in X$  with the minimal objective value  $f(x)$ . Mathematical optimization is the discipline in which problems as the one in (1) are analyzed from both a computational and a theoretical point of view.

An optimization problem as the one stated in (1) is classified according to the characteristics of the feasible set  $X$  and the objective function  $f$ . For example, a *linear program* (LP) is an optimization problem where *a*) the set  $X$  can be described by a finite number of affine inequalities, and *b*) the objective function  $f$  is linear. A *non-linear program* (NLP) is intuitively a problem where either the objective function or some of the constraint functions defining the feasible set are non-linear. Papers I and II analyze the case when the problem (1) is a CP, i.e., when the feasible set  $X$  is a convex set and when the objective function  $f$  is a convex function on the set  $X$ . Papers III, IV, and V deal with the case when the problem (1) is an MBLP, i.e., when the objective function  $f$  is linear, a subset of the variables are restricted to be binary, and all the variables are restricted to a polyhedron.

## 1.1 Outline

The following sections provide a short introduction to the areas of convex optimization and mixed binary linear optimization. In Section 2 the concept of a CP is introduced and thoroughly analyzed. The mathematical properties of CPs are presented and four appropriate solution methods are described. The section also presents applications which are described in more detail in Papers I and II. Section 3 introduces the mathematical properties of and solution procedures for MBLPs. Also here applications described and analyzed in Papers III, IV, and V are presented. In Section 4 a summary with the main contributions of the thesis is included together with some current and future research ideas. Finally, Section 5 summarizes the appended papers in an informal style.

## 2 Convex optimization

A convex program (CP) is essentially an optimization problem in which one wishes to minimize (maximize) a convex (concave) function over a convex set. A general form of a CP is the problem to find

$$f^* := \text{infimum } f(\mathbf{x}), \quad (2a)$$

$$\text{subject to } g_i(\mathbf{x}) \leq 0, \quad i = 1, \dots, m, \quad (2b)$$

$$\mathbf{x} \in X, \quad (2c)$$

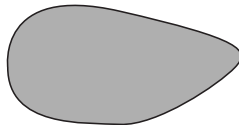
where  $X$  is a convex set,  $g_i : \mathbb{R}^n \mapsto \mathbb{R}$ ,  $i = 1, \dots, m$ , are convex functions on  $X$ , and  $f : \mathbb{R}^n \mapsto \mathbb{R}$  is a convex function over the set  $C := \{\mathbf{x} \in X \mid g_i(\mathbf{x}) \leq 0, i = 1, \dots, m\}$ . Let  $X^*$  denote the set of optimal solutions to the problem (2). The reason for not including the inequalities  $g_i(\mathbf{x}) \leq 0$ ,  $i = 1, \dots, m$ , in the definition of the set  $X$  will be apparent when we discuss the notion of Lagrangian duality in Section 2.1.2.

The convexity of the objective function and of the feasible set make powerful tools of convex analysis applicable. As we will see in Section 2.1, the existence of subgradients to  $f$ , the notion of necessary and sufficient conditions for optimality, and the important theory of duality follow quite easily from results in convex analysis. Furthermore, these results also provide powerful solution methods for CPs, as we will see in Section 2.2. In Section 2.3 applications and examples of CPs which have been studied in Papers I and II are described and analyzed.

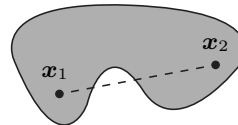
### 2.1 Problem properties

When considering a CP, many problem properties follow directly from the convexity of the feasible set and the convexity of the objective function. In order to describe the most important properties we need the following two definitions:

- A set  $C \subseteq \mathbb{R}^n$  is a *convex set* if the whole line segment joining any two points in the set also belongs to the set (see Figure 1). Mathematically, this means that if  $\mathbf{x}_1$  and  $\mathbf{x}_2$  are two points in  $C$ , then the point  $\lambda\mathbf{x}_1 + (1 - \lambda)\mathbf{x}_2$  must also belong to  $C$  for any  $\lambda \in [0, 1]$ .



(a) a convex set



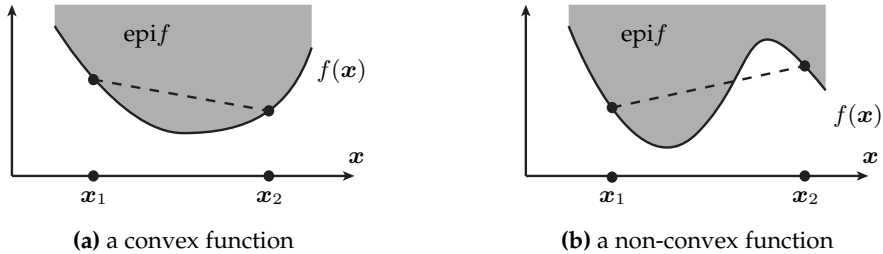
(b) a non-convex set

**Figure 1:** Illustrations of a convex and a non-convex set.

- A function  $f : \mathbb{R}^n \mapsto \mathbb{R}$  is *convex* on the convex set  $C$  if the inequality

$$f(\lambda\mathbf{x}_1 + (1 - \lambda)\mathbf{x}_2) \leq \lambda f(\mathbf{x}_1) + (1 - \lambda)f(\mathbf{x}_2),$$

holds for any  $x_1, x_2 \in C$  and any  $\lambda \in [0, 1]$ . A geometric interpretation of the convexity of a function  $f$  is that the line segment between any two points on the graph of  $f$  lies above the graph (see Figure 2). This is equivalent to saying that the *epigraph*,  $\text{epi}f := \{(x, \alpha) \in \mathbb{R}^{n+1} \mid f(x) \leq \alpha\}$  is a convex set (Bazaraa et al., 1993, Theorem 3.2.2).



**Figure 2:** Illustrations of a convex and a non-convex function, and their respective epigraphs.

According to the definitions above, there are strong connections between convex sets and convex functions; a function is convex if and only if its epigraph is a convex set. Using this property one can easily show the existence of *subgradients* of convex functions. But to establish this result, we need to define the notion of a supporting hyperplane.

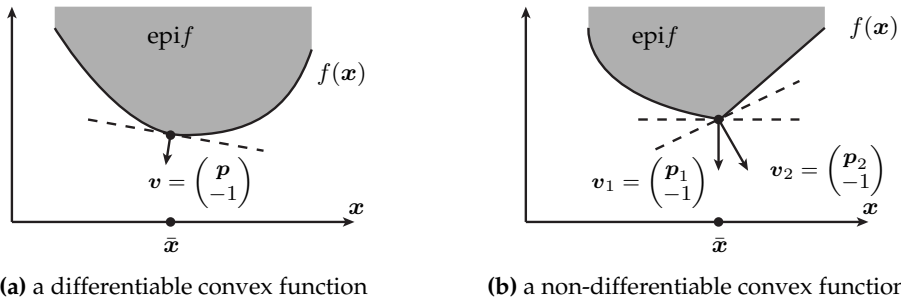
Let  $C \subseteq \mathbb{R}^n$  be a convex set and let  $\bar{x} \in \mathbb{R}^n$  be a point on the boundary of  $C$ . A vector  $v \in \mathbb{R}^n$  is said to define a *supporting hyperplane* of  $C$  at  $\bar{x}$  if  $v^T(x - \bar{x}) \leq 0$  for all  $x \in C$ . In geometrical terms, a supporting hyperplane is such that *a*) the set  $C$  is contained in one of the half-spaces defined by the hyperplane, and *b*) at least one boundary point of  $C$  is contained in the hyperplane. A well-known result from convex analysis is that if  $C$  is a convex set and  $\bar{x}$  is a point on the boundary of  $C$ , then there exists a supporting hyperplane of  $C$  which contains  $\bar{x}$  (Bazaraa et al., 1993, Theorem 2.4.7). Applying this result to the epigraph of a convex function (which we know from the above definition is a convex set) we obtain the following result (Bazaraa et al., 1993, 3.2.5). Let  $f : \mathbb{R}^n \mapsto \mathbb{R}$  be a convex function and let  $\bar{x} \in \mathbb{R}^n$ . Then there exists a vector  $p \in \mathbb{R}^n$  such that

$$f(x) \geq f(\bar{x}) + p^T(x - \bar{x}), \quad \text{for all } x \in \mathbb{R}^n. \quad (3)$$

We call vectors  $p \in \mathbb{R}^n$  satisfying (3) *subgradients* of  $f$  at  $\bar{x}$ . The set of subgradients of  $f$  at  $\bar{x}$  is called the *subdifferential*, denoted by  $\partial f(\bar{x})$ . Geometrically, a subgradient is a vector defining a supporting hyperplane to the epigraph of the function  $f$  containing the point  $\bar{x}$ .

The name "subgradient" stems from the fact that the notion is a generalization of the well-known concept of gradients. Whenever a convex function  $f$  is *differentiable* in a point  $x$ , the equivalence  $\partial f(x) = \{\nabla f(x)\}$  holds (Bazaraa et al., 1993, Lemma 3.3.2), which means that when considering differentiable functions, the notion of

subgradients is equivalent to the notion of gradients. In Figure 3(a), the unique subgradient of  $f$  at the point  $\bar{x}$  is illustrated for a differentiable convex function, and in Figure 3(b), two subgradients of  $f$  at  $\bar{x}$  are illustrated for a non-differentiable convex function.



**Figure 3:** Illustration of subgradients and their corresponding supporting hyperplanes of convex functions and their epigraphs. In figure (a) the unique supporting hyperplane  $v = (p^T, -1)^T$  of  $\text{epi} f$  at the point  $\bar{x}$  is illustrated. The vector  $p$  is the unique subgradient of the differentiable function  $f$  at  $\bar{x}$ . In figure (b) two supporting hyperplanes,  $v_1$  and  $v_2$ , of  $\text{epi} f$  at  $\bar{x}$  are illustrated. The vectors  $p_1$  and  $p_2$ , respectively, are the corresponding subgradients to the non-differentiable function  $f$  at the point  $\bar{x}$ .

For a general optimization problem as the one defined in (1), there are two notions of optimality of a point; *local optimality* and *global optimality*. A point  $\bar{x}$  is said to be a *global optimum* to the problem (1) if  $f(\bar{x}) \leq f(x)$  for all points  $x \in X$ , i.e., if the point has the lowest objective value among to all feasible points to the problem. A *local optimum* is, on the other hand, a point which has lowest objective value among all feasible points within a small neighborhood surrounding the point. Mathematically, a point  $\bar{x}$  is a local optimum if there exists a  $\varepsilon > 0$  such that  $f(\bar{x}) \leq f(x)$  for all  $x \in X \cap \{x \in \mathbb{R}^n \mid \|x - \bar{x}\| < \varepsilon\}$ . Consider the following fundamental results regarding CPs (Bazaraa et al., 1993, Theorem 3.4.2):

- If  $x^*$  is a local optimum to (2), then it is also a global optimum to (2), and
- the set of global optima to (2) is a convex set.

The fact that any local optimum to a CP is also a global optimum is one of the most crucial result regarding convex optimization. Many algorithms for solving optimization problems aim at finding local optima (e.g., steepest descent, Newton's method, and interior-point methods), which means that, for CPs they are also suitable for finding global optima.

### 2.1.1 Optimality conditions

Optimality conditions are necessary and/or sufficient criteria that are fulfilled by optimal solutions to optimization problems. Depending on the assumptions made

for the objective function and the feasible set, different optimality conditions must be stated. Consider problem (2) where the feasible set is denoted by  $C := \{\mathbf{x} \in X \mid g_i(\mathbf{x}) \leq 0, i = 1, \dots, m\}$ , and it is assumed that  $f$  is a convex function and that  $C$  is a convex set. Using these assumptions we can formulate the following necessary and sufficient optimality condition (Bazaraa et al., 1993, Theorem 3.4.3).

- A point  $\bar{\mathbf{x}} \in C$  is an optimal solution to (2) if and only if

$$\exists \mathbf{p} \in \partial f(\bar{\mathbf{x}}) \quad \text{such that} \quad \mathbf{p}^T(\mathbf{x} - \bar{\mathbf{x}}) \leq 0 \quad \text{for all} \quad \mathbf{x} \in C. \quad (4)$$

Geometrically, the condition (4) states that a point is a minimum if and only if, for any point  $\mathbf{x} \in C$ , the angle between a subgradient of  $f$  at  $\bar{\mathbf{x}}$  and the vector from  $\bar{\mathbf{x}}$  to  $\mathbf{x}$  is at least 90 degrees. For the unconstrained case, i.e., when  $C = \mathbb{R}^n$ , the condition (4) reduces to the criterion that  $\mathbf{0} \in \partial f(\bar{\mathbf{x}})$  (Bazaraa et al., 1993, Theorem 3.4.3, Corollary 1). The condition (4) is fairly useless as a means to verify the optimality of a candidate solution since it requires verifying that an inequality holds for, in the general case, an infinite number of points. But without adding any further assumptions on, e.g., the differentiability of the objective and constraint functions, we can not present any optimality conditions that are more useful or more easily verifiable before introducing the concept of Lagrangian duality, which is the topic for the next section.

### 2.1.2 Lagrangian duality

In convex optimization Lagrangian duality is a very important concept. By constructing a so-called *dual problem* associated with the problem (2), one can, as we will see, easily obtain lower bounds on the optimal objective value  $f^*$ . Consider the problem (2) where we only assume that  $X$  is a convex set, and that  $f$  and  $g_i$ ,  $i = 1, \dots, m$  are convex functions. Many optimization problems have the characteristic that the feasible set can be described as the intersection of two sets  $X$  and  $\{\mathbf{x} \in \mathbb{R}^n \mid g_i(\mathbf{x}) \leq 0, i = 1, \dots, m\}$ , where  $X$  possesses some nice structure making optimization problems on it efficiently solvable, but where the addition of the constraints  $g_i(\mathbf{x}) \leq 0, i = 1, \dots, m$ , destroys this structure. One example of an optimization problem having this structure is the *resource allocation problem* (e.g., Bretthauer and Shetty (1995), Patriksson (2008)) where  $X$  describes so-called *box constraints* for the variables and  $g_i(\mathbf{x}) \leq 0, i = 1, \dots, m$ , describe resource constraints.

Assume that when removing the constraints  $g_i(\mathbf{x}) \leq 0, i = 1, \dots, m$ , from the problem (2), the resulting problem can be solved efficiently. The constraints are, however, still essential for the problem definition, so they should of course not be entirely removed. Instead, an associated optimization problem in which each constraint is replaced by a penalty term in the objective function is constructed. Let the *Lagrangian dual function*,  $\theta : \mathbb{R}^m \mapsto \mathbb{R}$ , be defined as

$$\theta(\mathbf{u}) := \min_{\mathbf{x} \in X} \left[ f(\mathbf{x}) + \sum_{i=1}^m u_i g_i(\mathbf{x}) \right], \quad \mathbf{u} \in \mathbb{R}^m. \quad (5)$$

A well-known result, which actually holds even when the functions  $f$  and  $g_i$ ,  $i = 1, \dots, m$  are non-convex and/or the set  $X$  is non-convex, is that the function  $\theta$  is concave on  $\mathbb{R}_+^m$  (Bazaraa et al., 1993, Proposition 5.1.12). Let the set of solutions to the *subproblem*, defined in (5), at  $\mathbf{u} \in \mathbb{R}^m$  be denoted by  $X(\mathbf{u}) \subseteq X$ .

One of the most important properties of the Lagrangian dual function is described by the *weak duality theorem* which says the following. For any  $\mathbf{x} \in \mathbb{R}^n$  such that the constraints (2b)–(2c) hold and any  $\mathbf{u} \in \mathbb{R}_+^m$ , the inequality  $\theta(\mathbf{u}) \leq f(\mathbf{x})$  holds (Bazaraa et al., 1993, Theorem 6.2.1). For any non-negative  $\mathbf{u} \in \mathbb{R}^m$ , the dual function value  $\theta(\mathbf{u})$  is thus a lower bound on the objective value  $f(\mathbf{x})$  for all points  $\mathbf{x} \in \mathbb{R}^n$  that are feasible in (2). It also follows from the weak duality theorem that  $\theta(\mathbf{u}) \leq f^*$  for all  $\mathbf{u} \in \mathbb{R}_+^m$ , meaning that  $\theta(\mathbf{u})$  is a lower bound on the optimal value of the CP in (2). Note that the convexity assumption of neither the objective function  $f$  nor of the feasible set are required for the weak duality theorem to hold. To obtain an as good lower bound as possible, let us construct the *Lagrangian dual problem* as the problem to find

$$\theta^* := \sup_{\mathbf{u} \in \mathbb{R}_+^m} \theta(\mathbf{u}). \quad (6)$$

Since the dual function  $\theta : \mathbb{R}^m \mapsto \mathbb{R}$  is concave and the set  $\mathbb{R}_+^m$  is a convex set, the Lagrangian dual problem (6) is a CP; we denote its solution set by  $U^* \in \mathbb{R}_+^m$ .

It clearly holds, by the weak duality theorem, that  $\theta^* \leq f^*$ . In order to obtain a stronger result, we need to make the assumption that the set  $\{\mathbf{x} \in X \mid g_i(\mathbf{x}) < 0, i = 1, \dots, m\}$  is non-empty, i.e., *Slater's constraint qualification*. With this assumption the *strong duality theorem*, which states that  $\theta^* = f^*$ , can be deduced (Bazaraa et al., 1993, Theorem 6.2.4). This theorem further states the global optimality conditions that the inclusions  $\mathbf{x}^* \in X^*$  and  $\mathbf{u} \in U^*$  hold if and only if

$$\mathbf{u}^* \in \mathbb{R}_+^m, \quad (7a)$$

$$\mathbf{x}^* \in X(\mathbf{u}^*), \quad (7b)$$

$$g_i(\mathbf{x}^*) \leq 0, \quad i = 1, \dots, m, \quad (7c)$$

$$u_i^* g_i(\mathbf{x}^*) = 0, \quad i = 1, \dots, m. \quad (7d)$$

Hence, the strong duality theorem guarantees that the optimal value to the Lagrangian dual problem (6) equals that of the original CP (2). It also provides optimality conditions for the primal-dual pair of programs: the primal-dual solution pair  $(\mathbf{x}^*, \mathbf{u}^*) \in \mathbb{R}^n \times \mathbb{R}^m$  is optimal if and only if it holds that  $\mathbf{x}^*$  and  $\mathbf{u}^*$  are feasible in their respective problems (conditions (7a) and (7c)), that  $\mathbf{x}$  is a solution to the subproblem defined in (5) at  $\mathbf{u}$  (condition (7b)), and that the *complementarity condition*  $u_i^* g_i(\mathbf{x}^*) = 0$  holds for  $i = 1, \dots, m$  (condition (7d)). The optimality conditions (7a)–(7d) are the basis for the dual subgradient method which is described in Section 2.2.2.

When the original CP (2) is infeasible, i.e., when the set  $\{\mathbf{x} \in X \mid g_i(\mathbf{x}) \leq 0, i = 1, \dots, m\} = \emptyset$ , by convention, the optimal value  $f^*$  is defined to be  $+\infty$ . How does this affect the dual problem (6)? In Paper II we show that if the set  $X$  is bounded then it holds that also  $\theta^* = +\infty$ , meaning that the dual problem is unbounded.



### 2.1.3 Non-coordinability of non-strictly convex problems

The optimality conditions in (7) provides us with verifiable optimality criteria for candidate solution pairs  $(\mathbf{x}^*, \mathbf{u}^*) \in \mathbb{R}^n \times \mathbb{R}^m$ . But what if we only have a dual optimal solution  $\mathbf{u}^*$  at hand. Can we then easily construct an optimal solution to the primal problem (2)? Unfortunately, in general, the answer is no. The reason for this inconvenience is that the dual function is typically nonsmooth whenever the objective and constraint functions are non-strictly convex. This implies that an optimal primal solution is a (typically) nontrivial convex combination of the *extreme solutions* to the subproblem (3) at the optimal dual point  $\mathbf{u}^*$ . So when using methods that identify extreme solutions (as, e.g., the simplex method) for solving the subproblem (3), the solution obtained is, in general, non-optimal in the original problem. It may even be infeasible in the original problem. Within linear programming this property is referred to as the *non-coordinability* phenomenon (Dirickx and Jennergren, 1979, Chapter 3).

## 2.2 Solution procedures

There is no general analytical formula for the solution of convex optimization problems. However, there does exist some very effective methods for solving them. As stated by Rockafellar (1993) "... the great watershed in optimization isn't between linearity and nonlinearity, but convexity and nonconvexity." It appears that the convexity of a problem is inherently favorable when trying to solve it; one of the fundamental reasons is the fact that any local optima is also a global one.

Depending on the characteristics of the objective function and the feasible set, various solution methods are applicable. In this section four different procedures are presented. In Section 2.2.1, an iterative procedure—in which steps are taken in subgradient directions—is presented; for this procedure to be convergent only convexity of the optimization problem needs to be assumed. Then, in Section 2.2.2, Slater's constraint qualification is assumed, meaning that the strong duality theorem is applicable and that a Lagrangian dual approach can be used. In Sections 2.2.3 and 2.2.4, two methods based on the notion of *cutting planes* are introduced.

There are many other solution methods for CPs, for which more assumptions on the problem characteristics are made. Assuming differentiability of the objective and constraint functions leads to the famous KKT optimality conditions (Bazaraa et al., 1993, Theorem 4.2.13). These conditions constitute the basis for *interior point methods*, which have become increasingly popular because of their efficiency for solving structured convex programs (e.g., Nesterov et al. (1994), Potra and Wright (2000)). When the objective function is differentiable and convex and the feasible set is a polyhedron the *Frank-Wolfe method* (Frank and Wolfe, 1956) and the *simplicial decomposition algorithm* (Von Hohenbalken, 1977) are suitable. For a comprehensive study of algorithms for convex programming, see for example Boyd and Vandenberghe (2004) and Ben-Tal and Nemirovski (2001).

### 2.2.1 Subgradient methods

When convexity of the objective function  $f$  and of the feasible set  $C$  constitute the only assumptions on the problem (2), all available solution methods are based on the notion of subgradients. As stated in Section 2.1, the convexity ensures that subgradients of the objective function  $f$  exist on its entire domain. We will for simplicity assume that  $X$  is closed and bounded, i.e., compact, for the remainder of this section. The compactness ensures that the set  $X^*$  is non-empty, i.e., that there exists at least one optimal solution to the problem (2).

Shor (1967) developed the *subgradient method* which is an iterative method for minimizing convex functions. When applying the method to a constrained CP as the one in (2), the algorithm is called the *projected subgradient method* and is defined as follows. Starting at some initial feasible point  $\mathbf{x}^0 \in C$ , the iterates  $\mathbf{x}^t$ ,  $t \geq 1$ , are computed according to

$$\mathbf{x}^{t+1} = \text{proj}_C(\mathbf{x}^t - \alpha_t \mathbf{p}^t), \quad t = 0, 1, \dots, \quad (8)$$

where  $\mathbf{p}^t$  is a subgradient to  $f$  at  $\mathbf{x}^t$ ,  $\alpha_t > 0$  denotes the step length chosen at iteration  $t$ , and  $\text{proj}_C$  denotes the Euclidean projection onto  $C$ .

The convergence of the method defined by (8) to an optimal solution to the CP in (2) can be shown under specific rules for the computation of the step lengths. For some early convergence results regarding the subgradient method, see Ermol'ev (1966). Polyak (1967) showed that when utilizing step lengths fulfilling the so-called *divergent series* conditions

$$\lim_{t \rightarrow \infty} \alpha_t = 0 \quad \text{and} \quad \sum_{t=0}^{\infty} \alpha_t = \infty, \quad (9)$$

it holds that  $\lim_{t \rightarrow \infty} f(\mathbf{x}^t) = f^*$ , meaning that the objective values of the iterates converge to the optimal objective value of (2). Polyak also showed that the distance between  $\mathbf{x}^t$  and the solution set  $X^*$  tends to zero as  $t \rightarrow \infty$ . Proofs of these results are found in Shor (1985, Theorem 2.2). Shepilov (1976) further showed that when assuming that the step lengths also fulfill the condition

$$\sum_{t=0}^{\infty} \alpha_t^2 < \infty, \quad (10)$$

it holds that  $\lim_{t \rightarrow \infty} \mathbf{x}^t = \mathbf{x}^* \in X^*$ , i.e., that the sequence of iterates converges to a specific optimal solution.

So with fairly weak assumptions on properties of the objective function and of the feasible set, we can provide an iterative solution method for which the iterates are guaranteed to converge towards the solution set to the CP defined in (2). The three main drawbacks of the method are that:

- a) it can be difficult to find subgradients to  $f$ ,
- b) the projection onto the set  $C$  might not be easily computed, and

c) its convergence speed may be slow.

In the next sub-section the subgradient method is instead applied to the Lagrangian dual problem defined in (6). As we will see, for the application of the subgradient method to the dual problem, the the points *a*) and *b*) above will not pose any problems.

## 2.2.2 Dual subgradient methods

The Lagrangian dual problem defined in (6) is a CP, implying that the projected subgradient method described in Section 2.2.1 can be employed. For this particular problem, the method starts in an initial dual point  $\mathbf{u}^0 \in \mathbb{R}_+^m$  and then computes iterates  $\mathbf{u}^t, t \geq 1$ , according to

$$\mathbf{u}^{t+1} = [\mathbf{u}^t - \alpha_t \mathbf{g}(\mathbf{x}^t)]_+, \quad t = 0, 1, \dots, \quad (11)$$

where  $\mathbf{x}^t \in X(\mathbf{u}^t)$  solves the subproblem defined in (5) at  $\mathbf{u} = \mathbf{u}^t$ , which implies that  $\mathbf{g}(\mathbf{x}^t) := (g_1(\mathbf{x}^t), \dots, g_m(\mathbf{x}^t))^T \in \mathbb{R}^n$  is a subgradient to  $\theta$  at  $\mathbf{u}^t$  (Bazaraa et al., 1993, Theorem 6.3.4). Further,  $\alpha_t$  is the step length chosen at iteration  $t$  and  $[\cdot]_+$  denotes the Euclidean projection of a point onto the non-negative orthant.

In each iteration,  $t$ , the subproblem, defined in (5), at  $\mathbf{u} = \mathbf{u}^t$  has to be solved, generating a solution  $\mathbf{x}^t \in X(\mathbf{u}^t)$ . This solution automatically provides a subgradient  $\mathbf{g}(\mathbf{x}^t)$  to  $\theta$  at the point  $\mathbf{u}^t$ ; this implies that the drawback *a*) for the subgradient method described in Section 2.2.1 is not an issue. Furthermore, since the feasible set of the Lagrangian dual problem is the non-negative orthant, the Euclidean projection needed in each iteration can be efficiently performed, which implies that neither drawback *b*) is an issue.

We now assume that Slater's constraint qualification holds (i.e., that the set  $\{\mathbf{x} \in X \mid g_i(\mathbf{x}) < 0, i = 1, \dots, m\}$  is nonempty). This implies that strong duality holds, i.e., that  $\theta^* = f^*$ . By applying the convergence results stated in Section 2.2.1 it follows that by choosing step lengths fulfilling (9) and (10), one obtains convergence of the dual iterates  $\mathbf{u}^t$  to an optimal solution to the Lagrangian dual problem (6).

However, our aim is still to solve the original CP. Can we hope that the subproblem solutions  $\mathbf{x}^t \in X(\mathbf{u}^t)$  generated in the subgradient method converge to an optimal solution to the problem (2)? Unfortunately, the answer is no. In general, we can not even guarantee that any of the subproblem solutions found during the subgradient method will be feasible in (2); the reason being non-coordinability (see Section 2.1.3). In order to obtain convergence to an optimal solution to (2), we will utilize the notion of *ergodic sequences*. In each iteration of the method (11), an ergodic iterate,  $\bar{\mathbf{x}}^t$ , is defined as a convex combination of the subproblem solutions found so far, according to

$$\bar{\mathbf{x}}^t = \sum_{s=0}^{t-1} \mu_s^t \mathbf{x}^s, \quad \text{where } \mu_s^t \geq 0, \quad s = 0, \dots, t-1, \quad \sum_{s=0}^{t-1} \mu_s^t = 1, \quad t = 1, 2, \dots, \quad (12)$$

where  $\mathbf{x}^s \in X(\mathbf{u}^s)$  is the subproblem solution found at iteration  $s$ . The scalars  $\mu_s^t$  are called as *convexity weights*. The idea of creating convex combinations of subproblem solutions was first presented by Shor (1967), who showed convergence of the ergodic sequence,  $\bar{\mathbf{x}}^t$ , towards the optimal set  $X^*$  when employing suitable step lengths and convexity weights for the special case when the CP (2) is an LP. Larsson and Liu (1997) and Sherali and Choi (1996) developed these ideas further and proposed more general rules for choosing the convexity weights defining the ergodic sequences. Larsson et al. (1997) extended the results to the more general case when (2) is a CP. In Paper I we extend these convergence results further to include more general rules for computing the convexity weights.

In Paper II we analyze the case when the original CP is infeasible, i.e., when  $\{\mathbf{x} \in X \mid g_i(\mathbf{x}) \leq 0, i = 1, \dots, m\} = \emptyset$ . As described in the end of Section 2.1.2, the dual problem is then unbounded, meaning that  $\theta^* = +\infty$ . We show that when employing the subgradient method (11) on the unbounded dual problem, the sequence of dual iterates,  $\{\mathbf{u}^t\}$ , diverge in the direction of steepest ascent of the Lagrangian dual function. We also show that the sequence of ergodic iterates,  $\{\bar{\mathbf{x}}^t\}$ , converges to a point for which the Euclidean norm of the infeasibility in the relaxed constraints is minimized, i.e., to a point in the set  $\operatorname{argmin}_{\mathbf{x} \in X} \{\|[\mathbf{g}(\mathbf{x})]_+\|\}$ .

### 2.2.3 Cutting-plane methods

The simplicity of the subgradient method described in Section 2.2.1 comes at the price of ignoring past information. The information obtained when solving the subproblem defined in (5) can be used not only to determine search directions, but also to build a model of the function  $f$  itself.

The basic idea of *cutting plane methods* is the construction of a piecewise-linear approximation of the objective function. Each subgradient found during the course of the cutting plane method can in turn be used to improve the model of the objective function. The cutting plane method for the problem CP in (2) can be described as follows. Let  $\mathbf{x}^0, \dots, \mathbf{x}^{t-1}$  be the iterates found up to iteration  $t$  of the method. Then a lower approximation of  $f$  at iteration  $t$  is given by the piecewise-linear function  $\hat{f}_t : \mathbb{R}^n \mapsto \mathbb{R}$ , defined by

$$\hat{f}_t(\mathbf{x}) := \max_{s=0, \dots, t-1} \{f(\mathbf{x}^s) + (\mathbf{p}^s)^\top (\mathbf{x} - \mathbf{x}^s)\}, \quad (13)$$

where  $\mathbf{p}^s$  is a subgradient to  $f$  at  $\mathbf{x}^s$ . By construction it holds that  $\hat{f}_t(\mathbf{x}) \leq f(\mathbf{x})$ , for all  $\mathbf{x} \in C$ , meaning that the model is an underestimate of the objective function. The next iterate produced by the cutting plane method is then obtained by solving the problem to minimize  $\hat{f}_t$  over the feasible set  $C$ . Note that the function  $\hat{f}_t$  can be modelled by a linear objective function and linear constraints, meaning that whenever the feasible set  $C$  is a polyhedron, the next iterate can be found by solving an LP.

One advantage of cutting-plane methods is that they provide lower bounds on the optimal objective value during the course of the method, implying that natural stopping criteria can be utilized. A drawback is that the method sometimes

possesses a zig-zagging behavior of the iterates (Hiriart-Urruty and Lemaréchal, 1993). This drawback is the motivation for the class of methods presented next. For thorough analyses of cutting plane methods, see Kelley (1960) and Goffin and Vial (2002).

### 2.2.4 Bundle methods

*Bundle methods* can be viewed as stabilizations of the cutting plane method. An extra point  $\hat{\mathbf{x}}^{t-1}$ , called the *center*, is added to the information provided in iteration  $t$  of the method. The piecewise-linear approximation (13) is still used as a model for the objective function  $f$ , but no longer do we solve the problem of minimizing  $\hat{f}_t$  over the feasible set  $C$  to obtain the next iterate. Instead, the next iterate is computed by finding

$$\mathbf{x}^t \in \operatorname{argmin}_{\mathbf{x} \in C} \left\{ \hat{f}_t(\mathbf{x}) + \frac{\mu_t}{2} \|\mathbf{x} - \hat{\mathbf{x}}^{t-1}\|^2 \right\}. \quad (14)$$

Then the center is updated to be the new iterate if the iterate has improved the objective value sufficiently as compared to the objective value of the previous center, i.e., if  $f(\mathbf{x}^t) \leq f(\hat{\mathbf{x}}^{t-1}) - \delta_t$ , for some  $\delta_t > 0$ .

The quadratic term in the relation (14) stabilizes the cutting plane method. It will make the next iterate closer to the current center, as compared to the cutting plane method, by avoiding drastic movements. The role of the parameter  $\mu_t$  is to control the trade-off between minimizing the model  $\hat{f}_t$  and staying close to the center  $\hat{\mathbf{x}}^{t-1}$  which is known to have low objective value.

Depending on how the parameters  $\mu_t$  and  $\delta_t$  are updated, different versions of the bundle method are defined. For comprehensive description of bundle methods, see Lemaréchal et al. (1995) and Kiwiel (1990).

## 2.3 Applications and modelling

Convex optimization problems appear in many application areas. We next describe an assortment of such applications.

If a problem can be formulated as a convex optimization problem, then it can in most cases be solved efficiently. The challenge is in recognizing when and how a specific problem can be formulated as a CP. Examples of optimization problems which are non-convex in their natural form are so-called *geometric programs* (GPs); problems for which a posynomial  $f$  should be minimized subject to the constraints  $g_i(\mathbf{x}) \leq 1$ , where  $g_i$  is a posynomial, for  $i = 1, \dots, m$ . Even though neither the objective nor the constraint functions are convex, a general GP can still be translated into a CP by a change of variables and transformation of the objective and constraint functions (Boyd et al. (2007), Ecker (1980)).

Many problems may, of course, not be formulated as CPs. Convex optimization can, however, still play a role in solution courses for such problems. One idea is to combine convex optimization with local optimization methods. If the initial optimization problem is non-convex, one first tries to approximate the problem by a

CP. The solution of the CP (which we assume can be found efficiently) can then be used as a starting point for a local optimization method applied to the original non-convex optimization problem. Another idea is to use convex optimization for obtaining cheaply computable lower bounds on the optimal value of the non-convex optimization problem. This can be obtained by, for example, solving the Lagrangian dual problem, which is a CP even when the original problem is not.

Convex optimization is often used within approximation and fitting. In many approximation problems the aim is to construct a model that best fits some observed data and prior information. The variables in the optimization problem represent the parameters in the statistical model, while the constraints represent prior information regarding the parameters (for example nonnegativity). The objective function is often composed by measures of the differences between the observed data and the values predicted by the model. Often, the measure used is some norm of the vector of predicted errors; since all norms are convex functions this implies that the objective function is convex. For a thorough analysis of how convex optimization is utilized in the area of data fitting and model approximation, see Boyd and Vandenberghe (2004, Chapter 6.1).

In Paper I we describe the *nonlinear multicommodity flow problem* (NMFP; e.g., Bertsekas (1979)) in detail and evaluate the dual subgradient method on a sample of test instances. The NMFP is a CP where the aim is to send flows of multiple commodities between several pairs of source and sink nodes in a network. The objective function is convex and the feasible set is polyhedral. The NMFP has applications within, for example, traffic assignment (Patriksson, 1994) and multi-agent planning systems (Wellman, 1993).

Other problem classes being special cases of convex programs, and for which specialized solutions methods have been developed, include *second order cone programs* (SOCPs) and *semidefinite programs* (SDPs). SOCPs are generalizations of linear and quadratic programs that allow for affine combination of variables to be constrained inside second-order cones (for a comprehensive analysis, see Lobo et al. (1998)). In an SDP a linear function is to be minimized over the intersection of the cone of positive semidefinite matrices with an affine space (for a comprehensive analysis, see Vandenberghe and Boyd (1996)). Since both SOCPs and SDPs are efficiently solvable by *interior point methods*, they are becoming increasingly popular in areas such as combinatorial optimization (e.g., Alizadeh (1995), Goemans (1997)) and control theory (e.g., Parrilo and Lall (2003), Yao et al. (2001)).

### 3 Mixed binary linear optimization

This section provides a short introduction to the area of mixed binary linear optimization. For comprehensive analyses, see, for example, Nemhauser and Wolsey (1989) and Wolsey (1998). A *mixed binary linear program* (MBLP) can be expressed as the problem to find

$$z^* := \underset{\mathbf{x}}{\text{infimum}} \quad \mathbf{c}^\top \mathbf{x}, \quad (15a)$$

$$\text{subject to} \quad A\mathbf{x} \geq \mathbf{b}, \quad (15b)$$

$$x_i \in \{0, 1\}, \quad i \in \mathcal{I}, \quad (15c)$$

where  $\mathbf{c} \in \mathbb{R}^n$ ,  $A \in \mathbb{R}^{m \times n}$ ,  $\mathbf{b} \in \mathbb{R}^m$ , and  $\mathcal{I} \subseteq \{1, \dots, n\}$ . The set  $\mathcal{I}$  denotes the set of indices of variables in the optimization problem (15) which are required to be binary. These variables typically represent so-called either/or decisions, meaning that, for example, you either buy a machine or not. The name MBLP stems from the fact that relaxing the binary requirements (15c) yields an LP, which denotes an optimization problem composed by affine constraints and a linear objective function.

In Section 3.1 the important mathematical properties of MBLPs are presented and analyzed, in Section 3.2 the importance of modelling with mixed binary linear optimization is discussed, and in Section 3.3 a selection of solution procedures for MBLPs are presented. Section 3.4 describes how mixed binary linear optimization can be used for scheduling maintenance operations, and in Section 3.5 a number of other applications of MBLPs are presented.

#### 3.1 Problem properties

Consider the following small example, in which the objective is to find

$$z^* := \text{minimum} \quad -x_1 - 3x_2, \quad (16a)$$

$$\text{subject to} \quad x_1 + x_2 \leq 2, \quad (16b)$$

$$2x_2 \leq 3, \quad (16c)$$

$$x_2 \geq 0, \quad (16d)$$

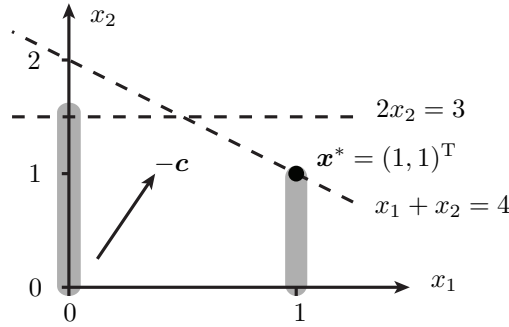
$$x_1 \in \{0, 1\}. \quad (16e)$$

Relating to the general definition of an MBLP in (15), it holds that

$$\mathbf{c} = (-1, -3)^\top, \quad A = \begin{pmatrix} -1 & 0 & 0 \\ -1 & -2 & 1 \end{pmatrix}^\top, \quad \mathbf{b} = (-2, -3, 0)^\top, \quad \text{and} \quad \mathcal{I} = \{1\}.$$

The problem (16) is illustrated in Figure 4, where the line segments marked in grey represent the feasible set. The optimal solution to the problem (16) is  $\mathbf{x}^* = (1, 1)^\top$  with objective value  $z^* = -4$ .

We first note that the feasible set of the MBLP defined in (16) is non-convex. In general, the MBLP defined in (15) is non-convex whenever  $\mathcal{I} \neq \emptyset$  and the set defined



**Figure 4:** Illustration of the MBLP defined in (16). The grey area represent the feasible set defined by constraints (16b)–(16e). The optimal solution is  $\mathbf{x}^* = (1, 1)^T$ .

by the constraints (15b) is not a singleton (if  $\mathcal{I} = \emptyset$ , then the problem is an LP), meaning that the theory presented in Section 2 may not be utilized here. A general MBLP is NP-hard, which basically means that there does not exist any algorithm which can solve the MBLP within a time bounded by a polynomial in the size of the problem. For comprehensive analyses and rigorous definitions of computational complexity in general and NP-hardness in particular, see Papadimitriou (1994) and Garey and Johnson (1979).

As described in Section 2.3, one important aspect of convex optimization is its application to non-convex optimization problems. By constructing and solving a CP which is a relaxation of the original (non-convex) optimization problem, one can obtain a lower bound on the optimal objective value of the original problem. One relaxation of the MBLP in (15) is achieved by replacing the binary requirements (15c) by continuous relaxations, i.e., by requiring that  $x_i \in [0, 1]$ ,  $i \in \mathcal{I}$ . The resulting problem is denoted the LP *relaxation* of the MBLP and is defined as the problem to find

$$z_{\text{LP}}^* := \infimum_{\mathbf{x}} \quad \mathbf{c}^T \mathbf{x}, \quad (17a)$$

$$\text{subject to} \quad A\mathbf{x} \geq \mathbf{b}, \quad (17b)$$

$$x_i \in [0, 1], \quad i \in \mathcal{I}. \quad (17c)$$

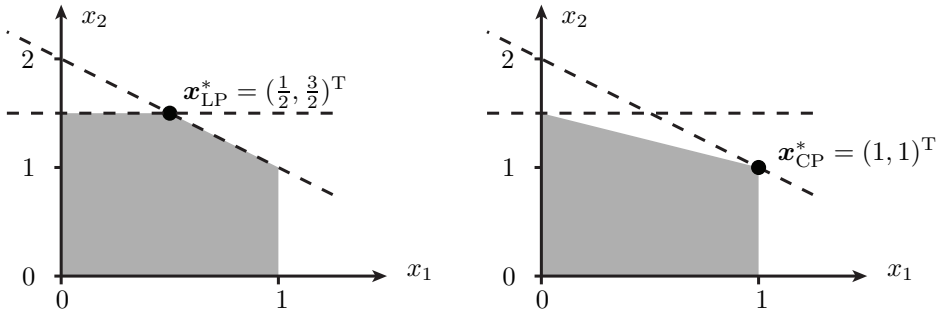
Clearly, since the problem (17) is a relaxation of the problem (15), it holds that  $z_{\text{LP}}^* \leq z^*$ . The LP relaxation of the example in (16) has the optimal solution  $\mathbf{x}_{\text{LP}}^* = (\frac{1}{2}, \frac{3}{2})^T$  with objective value  $z_{\text{LP}}^* = -5$  (see Figure 5a).

Another relaxation of a MBLP is the so called *convex relaxation*, which is the problem obtained when replacing the constraints (15b) and (15c) with the convex hull of the set defined by the constraints, i.e., the problem to find

$$z_{\text{CP}}^* := \infimum_{\mathbf{x}} \quad \mathbf{c}^T \mathbf{x}, \quad (18a)$$

$$\text{subject to} \quad \mathbf{x} \in \text{conv} \left( \mathbf{x} \in \mathbb{R}^n \mid \begin{array}{l} A\mathbf{x} \geq \mathbf{b}, \\ x_i \in \{0, 1\}, \quad i \in \mathcal{I}. \end{array} \right) \quad (18b)$$





(a) The LP relaxation of (16). The grey area illustrates the set  $\{\mathbf{x} \in \mathbb{R}^n \mid A\mathbf{x} \geq \mathbf{b}, x_i \in [0, 1], i \in \mathcal{I}\}$ . The optimal solution to the relaxed problem is  $\mathbf{x}_{LP}^* = (\frac{1}{2}, \frac{3}{2})^T$  with objective value  $z_{LP}^* = -5$ .

(b) The convex relaxation of (16). The grey area illustrates the set  $\text{conv}(\mathbf{x} \in \mathbb{R}^n \mid A\mathbf{x} \geq \mathbf{b}, x_i \in \{0, 1\}, i \in \mathcal{I})$ . The optimal solution to the relaxed problem is  $\mathbf{x}_{CP}^* = (1, 1)^T$  with objective value  $z_{CP}^* = -4$ .

**Figure 5:** Illustrations of the LP relaxation and the convex relaxation of the MBLP (16).

The problem (18) results in an LP since the feasible set described by (18b) is a polyhedron. It actually holds that  $z_{CP}^* = z^*$  (Nemhauser and Wolsey, 1989, Theorem 6.3 of Section I.4.6). The convex relaxation of the example in (16) is illustrated in Figure 5b. For this particular example it holds that the solution set (being the singleton  $(1, 1)^T$ ) to the convex relaxation equals that of the original MBLP. In general, it only holds that the solution set to the original MBLP is a subset of the solution set to the convex relaxation (Nemhauser and Wolsey, 1989, Theorem 6.3 of Section I.4.6). For an example where the two solution sets are not equal, consider the adjustment of example (15) where the objective function is replaced by  $-x_1 - 2x_2$ .

In general, the convex relaxation is always *stronger* than the LP relaxation in the sense that the feasible set of the convex relaxation is contained in the feasible set of the LP relaxation. Hence, the optimal objective value of the convex relaxation is always at least as high as that of the LP relaxation. To solve the relaxed problems, one first needs to formulate them as LPs, which requires the feasible sets to be described by affine inequalities and equalities. The feasible set of the LP relaxation can easily be described in this way by just replacing the constraints (15c) with the affine constraints  $x_i \in [0, 1]$ ,  $i \in \mathcal{I}$ . But for the convex relaxation, there does not exist any efficient general approach to construct the feasible set using only affine constraints; it is possible by generating exponentially many  $\{0, \frac{1}{2}\}$ -Chvátal-Gomory cuts, but this is not classified as computationally efficient (Gentile et al., 2006). This is the reason why the LP relaxation is often the basis for many methods for solving MBLPs, as we will see in Section 3.3. But first a few notes regarding the importance of good modelling.

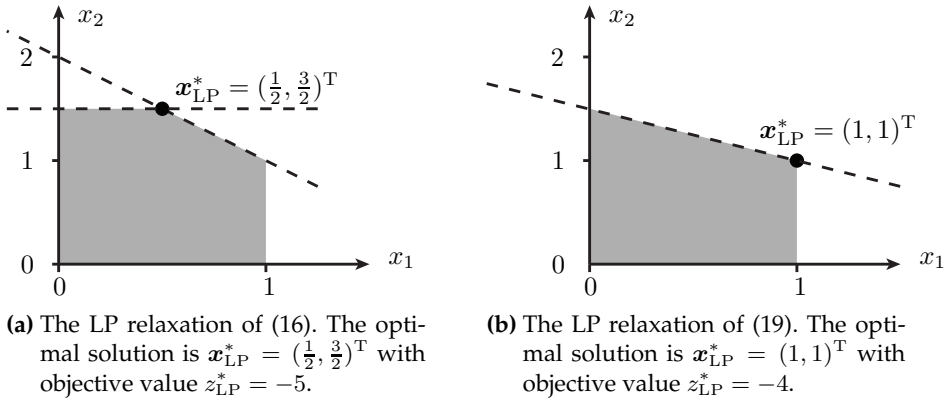


Figure 6: Illustration of the LP relaxations of the two equivalent problems (16) and (19).

### 3.2 The importance of good modelling

MBLPs can be formulated in many ways and the model selected has a vital impact on the solvability of the problem. Two models for a MBLP can be equivalent in the sense that they possess the same feasible set, while their LP relaxations, however, possess different feasible regions. Consider the following model

$$\text{minimize } -x_1 - 3x_2, \quad (19a)$$

$$\text{subject to } x_1 + 2x_2 \leq 3, \quad (19b)$$

$$x_2 \geq 0, \quad (19c)$$

$$x_1 \in \{0, 1\}. \quad (19d)$$

The problem modelled in (19) is equivalent to that modelled in (16), because they possess the same feasible set and the same objective function. However, their LP relaxations are not equivalent, as illustrated in Figure 6. For the model (19), it further holds that the LP relaxation and the convex relaxation are equivalent.

The difference between the optimal objective values of an MBLP and its LP relaxation is called the *integrality gap*. The smaller the integrality gap is, the better (or stronger) the model is. For the problem defined in (16), we would then say that the model (19) is a stronger formulation than the one in (16).

The importance of modelling when considering MBLPs cannot be understated. As Nemhauser and Wolsey (1989, p. 14) states: "In integer programming, formulating a 'good' model is of crucial importance to solving the model". The concepts of strong and weak formulations of models of MBLPs are discussed in Papers IV and V.

### 3.3 Solution procedures

In general MBLPs are much harder to solve than LPs; the main reason is that the convexity of the LP is lost when adding binary requirements to some of the variables. For LPs, the KKT optimality conditions give rise to efficient solution methods such as the simplex method (e.g., Murty (1983)) or interior point methods (e.g., Nesterov et al. (1994) and Potra and Wright (2000)). For MBLPs, such conditions cannot be formulated, so instead the methods need to rely on other properties of the problem.

The solution procedures that exist for MBLPs are characterized as *enumeration methods*, *relaxation methods*, *cutting plane methods*, or *heuristics*. For a comprehensive study of methods for solving MBLPs, see Lodi (2010).

One simple (and naive) approach to solve the MBLP in (15) is to enumerate all possible combinations of binary values for the binary variables  $x_i, i \in \mathcal{I}$ , and solve the corresponding LP (which is obtained when the binary variables are fixed) for each such combination. This approach, which is an enumeration approach, would require the solution of  $2^{|\mathcal{I}|}$  LPs which, for many practical applications, is highly intractable. In Section 3.3.1 a more efficient enumeration method, called the *branch-and-bound method*, is described. Relaxation methods use relaxations which are also simplifications of the problem formulation, in the sense that the relaxed problem is computationally less expensive than the original one. In Section 3.3.2 such a method, based on the concept of Lagrangian relaxation (described in Section 2.1.2), is presented. Cutting plane methods are based on the repeated solution of LP relaxations of the problem. In each new iteration, one or several constraints are added to the model in order to construct a stronger linear model of the problem. Section 2.2.3 describes in more detail the ideas of cutting plane methods. Heuristics are methods, which can be based on very simple rules or solutions to complicated optimization problems, or something in between, in order to obtain feasible solutions to the MBLP (for studies regarding heuristics for MBLPs, see Balas et al. (2004) and Lokketangen and Glover (1998)). There are, in general, no guarantees that a heuristic will find an optimal solution to the problem. However, heuristics are often fast and easily implemented, and can in many cases produce near-optimal solutions.

#### 3.3.1 Branch-and-bound methods

Branch-and-bound methods are enumeration methods based on the idea to partition the feasible region of the problem into sub-regions and solve a relaxed problem over each sub-region. The relaxation used is in most cases the LP relaxation, which means that a branch-and-bound method relies heavily on efficient solution methods for LPs.

Consider the problem (15) and its LP relaxation (17), obtained by relaxing the binary requirements (15c). A simple version of the branch-and-bound method can be described as follows. Construct the LP relaxation (17) and define a *root node* corresponding to this LP. Solve the LP relaxation to obtain a solution  $\mathbf{x}_{LP}^*$ . If this solution is feasible in (15), i.e., if all the variables  $x_i, i \in \mathcal{I}$ , possess binary values, then terminate. Otherwise, choose a variable,  $x_j$  say, for which  $x_j^*$  is fractional, and introduce

two new problems, in which the constraints  $x_j = 0$  and  $x_j = 1$ , respectively, are included in the model. This procedure is denoted as “branching on the variable  $x_j$ ”. The two new LPs correspond to two new nodes in the branch-and-bound *tree*. Solve the LPs corresponding to these two nodes and branch further (as done in the root node) if the solutions are not feasible in the original problem. Prune a node whenever *a*) the optimal objective value of the LP is higher than that of the best found feasible solution to (15), or *b*) the LP is infeasible. The procedure continues branching on variables in the tree until all leaves (the nodes which have not branched further) either have been pruned or possess solutions which are feasible in (15). For a more elaborate description of the branch-and-bound method and an introduction of important concepts within the framework, such as *presolve*, *branching rules*, and *cutting planes*, see Lodi (2010).

It should be noted that, in the worst case, the branch-and-bound method may reduce to a complete enumeration scheme (described in the beginning of Section 3.3). The success of the method relies to a large extent on how strong (in the sense of its LP relaxation) the model defining the problem is. A stronger model leads to better lower bounds which lead to earlier (and thus much more efficient) pruning of the nodes.

The branch-and-bound method for general integer linear programs was introduced by Land and Doig (1960) and Dakin (1965). The method has since then become increasingly popular as an exact solution procedure for MBLPs and several solvers exist, both commercial ones such as *Gurobi*, *CPLEX*, and *EXPRESS*, and open-source ones such as *SCIP*, *GLPK*, and *COIN-OR*.

### 3.3.2 Dual subgradient methods and Lagrangian heuristics

For some MBLPs, enumeration methods such as the branch-and-bound method may not be suitable. This can occur when, for example, the size of the problem is too large, which leads to an intractably large branch-and-bound tree. For such problems one has to rely on relaxation methods and/or heuristics. The following method is a hybrid between the two.

When presenting the theory of Lagrangian duality in Section 2.1.2, no additional assumptions were needed for the weak duality theorem to hold. The weak duality theorem states that the value of the Lagrangian dual function at a  $\mathbf{u} \in \mathbb{R}_+^m$  underestimates the optimal objective value of the original primal problem. So by constructing the Lagrangian dual function  $\theta$  corresponding to the MBLP in (15) by relaxing a subset of the constraints (15b), lower bounds on the optimal objective value of (15) are obtained by evaluating the dual function at a  $\mathbf{u} \in \mathbb{R}_+^m$ . Since the MBLP is not convex, we can, however, not expect the strong duality theorem is applicable.

In Paper III we propose to utilize the dual subgradient method described in Section 2.2.2 on the MBLP in (15). When the original problem is a CP it holds that the sequence of ergodic iterates  $\{\bar{\mathbf{x}}^t\}$  converge towards the solution set to the original problem. For the case when the problem is an MBLP, such strong convergence results do not hold. It holds, however, that the ergodic sequence converges to the solution set to a convexified version of the original MBLP. The convexified version is the

problem, in which the constraints that are not Lagrangian relaxed are replaced by the convex hull of the corresponding feasible set. When the model representing the problem is strong, it holds that the solution set to the convexified version lies close to the solution set to the original problem. So by utilizing the ergodic iterates as starting points in local search procedures, this results in heuristics which hopefully provide solutions of good quality.

As stated at the beginning of this section, there exist a variety of solution methods for MBLPs, and here only a subset have been presented. We now continue in the next two sections with some applications where mixed binary linear optimization has been utilized.

### 3.4 Applications to maintenance scheduling

*Maintenance* can be described as a set of activities performed to ensure that a system stays operational. To schedule maintenance activities means to specify at what times the different maintenance activities should be performed. Such decisions are often mathematically modelled using binary variables, denoting whether maintenance should be performed or not at a specific point in time. This means that optimization models occurring within maintenance scheduling often are MBLPs. The purpose of this section is to give an introduction to maintenance scheduling, which is the subject of Papers IV and V.

Maintenance activities are often characterized as either *preventive maintenance* (PM) or *corrective maintenance* (CM). PM denotes the scheduled maintenance activities performed in order to prevent failures of the system, while CM denotes the activities performed after a failure has occurred and in order to restore the system to an operational state. A second important characterization of maintenance activities is defined by the type of system considered—whether it is a single- or a multi-component system. Most of the research on maintenance scheduling until the 1990s focused on single-component systems; for a survey of that field, see Wang (2002). In the present work, however, the focus is on maintenance of multi-component systems with dependencies between the components. When dependencies are negligible (or neglected), one can apply single-component models for individual components, otherwise, one must consider the entire system in order to optimize the scheduling activities. Dependencies are categorized as either *economic*, *structural*, or *stochastic*. Positive (negative) economic dependencies imply that maintenance simultaneously performed on several components is less (more) expensive than when it is performed on the same set of components but at different points in time. Structural dependencies can imply, for instance, that maintenance on one component enforces the removal of another component. Stochastic dependencies arise, for instance, when the degradation of one component is correlated with that of another component. For comprehensive overviews of maintenance scheduling optimization in general, see Nicolai and Dekker (2008), Pham and Wang (2000), and Pintelon and

Gelders (1991).

In Paper IV we introduce the *preventive maintenance scheduling problem with interval costs* (PMSPIC), which is the problem of scheduling PM of a multi-component system with positive economic dependencies over a finite discretized time horizon. An MBLP model for the PMSPIC is introduced and solved using commercial optimization software. The polyhedral properties of the model is thoroughly analyzed and the computational complexity of the problem is discussed.

In Paper V the problem of *scheduling tamping operations on ballasted rail tracks* is analyzed and an MBLP model is introduced. The system considered is a set of rail track sections, which constitute the components of the system, and the maintenance activities considered are characterized as PM. The binary variables in the model denote whether or not maintenance should be performed on a given track section at a specific point in time, while the continuous variables describe the conditions of the track sections over time. Also here positive economic dependencies between the components exist, since set-up costs are paid each time any tamping is being performed on the system.

### 3.5 Other applications

There are numerous other applications for which MBLPs are utilized. In this section three problem classes are presented; *covering problems*, *flow problems*, and problems with *disjunctive constraints*.

Covering problems are optimization problems where the aim is to find a combinatorial structure of minimum cost which covers another structure. The most studied covering problem is the *set covering problem*, in which the aim is to choose a subset of columns in a 0-1-matrix such that each row is covered by at least one of the chosen columns. For a comprehensive analysis of the set covering problem, see, e.g., Caprara et al. (2000). In Paper III the dual subgradient method described in Section 3.3.2 is evaluated on a sample of set covering instances. Other covering problems include the *vertex cover problem* (e.g., Chen et al. (2001)) and the *edge cover problem* (e.g., Horton and Kilakos (1993)).

In flow problems the aim is to find the cheapest possible way of sending flow through a network. One example of a flow problem is the *fixed-charge network flow problem* (e.g., Kim and Pardalos (1999)), in which a fixed charge needs to be paid for sending any flow through an arc in the network. The binary variables in the model represent whether the arcs are opened or not, and the continuous variables represent the amount of flow sent on each arc.

In the usual statement of optimization problems, it is assumed that all of the constraints must be satisfied for a point to be called feasible. But in some applications, only a subset of the constraints need to be fulfilled for a solution to be acceptable. In such cases, we say that the constraints are *disjunctive*. Disjunctive constraints arise naturally in scheduling problems where several jobs need to be processed on a machine and where the order of the jobs is not specified. Then disjunctive constraints of the type either "job  $i$  precedes job  $j$  or vice versa" occur. By introducing binary

variables representing whether or not a constraint is fulfilled, the disjunctive constraints can be incorporated into an MBLP formulation. For more comprehensive analysis, see Balas (1998).

## 4 Summary and conclusions

The sections above are short introductions to the areas of convex and mixed binary linear optimization. For more elaborate descriptions of the notions introduced, see the references proposed in the text. The continuation of this section is as follows. In Section 4.1 the main contributions of the papers appended in the thesis are summarized, and in Section 4.2 some current and future research ideas are discussed.

### 4.1 Contributions of the thesis

This section presents a short summary of the main contributions of the thesis within the areas convex and mixed binary linear optimization.

Within convex optimization the main contributions are the following. Convergence results regarding sequences of ergodic iterates obtained from a dual subgradient method are generalized from the case when the primal problem is an LP to the general case when it is a CP. A new set of rules for constructing the sequences is introduced and evaluated on a sample of nonlinear multicommodity flow problems; the computational results clearly demonstrate that the new rules outperform previously proposed ones (Paper I). Theoretical analyses and convergence results for both the primal and dual sequences of iterates obtained in a dual subgradient method are developed for the special case when the primal problem is infeasible. The sequence of primal ergodic iterates is shown to converge to the set of points possessing minimum infeasibility in the relaxed constraints (Paper II).

In mixed binary linear programming the contributions are the following. The employment of a dual subgradient method on a general MBLP is analyzed and convergence results for the sequences of ergodic iterates constructed are developed (Paper III). A thorough analysis of the complexity and polyhedral properties of an MBLP model of the preventive maintenance scheduling problem with interval costs is presented, and numerical results on three case studies demonstrate the usefulness of the model (Paper IV). A MBLP model for the problem of scheduling tamping operations on ballasted tracks is extended by including economic dependencies between the track segments. The complexity and polyhedral properties are analyzed and the results from a numerical study show that the extended model can be solved more efficiently as compared with the previously proposed one (Paper V).

### 4.2 Future research

The idea of constructing sequences of ergodic iterates from the subproblem solutions obtained in a dual subgradient scheme is evaluated on a sample of nonlinear multicommodity flow problem instances in Paper I. The computational results

demonstrate that the proposed set of rules for the convexity weights defining the sequences outperforms previously proposed ones on the test instances. One natural continuation of this is to evaluate the performance of the procedure on other problems for which subgradient schemes have proven successful, for example within the fields network design (e.g., Balakrishnan et al. (1989) and Frangioni and Gendron (2009)) and traffic assignment problems (e.g., Patriksson (1994)). Another extension of the results regarding the dual subgradient method is to analyze the convergence of the sequence of ergodic iterates when allowing for inexact solutions of the subproblems; solutions providing  $\varepsilon$ -subgradients to the dual objective function.

We are currently investigating the possibility of utilizing the sequences of ergodic iterates obtained from the dual subgradient method applied to a MBLP within a branch-and-bound method. The ergodic iterates are then used as starting points in heuristics for finding feasible solutions, as well as providing input to branching decisions in the branch-and-bound nodes. A similar approach is employed by Görtz and Klose (2012), but without our proposed set of rules for computing the convexity weights defining the sequences.

Another future research plan is to analyze the usefulness of both the PMSPIC introduced in Paper IV and the scheduling problem introduced in Paper V when considering real life case studies. Numerical results show that the proposed models can reduce the total maintenance costs considerably, compared to using simple heuristics for constructing the maintenance schedules. But in order to validate the models further, more realistic case studies need to be set up and analyzed.



## 5 Summary of the appended papers

This section contains a summary of the papers appended to the thesis. The emphasis is on presenting the theory and results in a less formal style than in the papers themselves.

### 5.1 Paper I: Primal convergence from dual subgradient methods for convex optimization

(coauthored with Michael Patriksson and Ann-Brith Strömberg)

Subgradient optimization methods applied to Lagrangian dual formulations of structured optimization problems are often able to quickly identify near-optimal solutions to the dual problem. They do, however, not automatically provide primal feasible or primal optimal solutions. In this paper, we consider a convex (and possibly non-smooth) optimization problem, which is Lagrangian relaxed. A subgradient optimization method is applied to the corresponding dual problem and convergence to a dual optimal solution is ensured by employing divergent series step lengths. We create ergodic primal iterates as convex combinations of subproblem solutions found during the previous subgradient iterations; convergence of these to a primal optimal solution is established through suitable rules for the convexity weights.

The main result presented in this paper is the generalization of the convergence results for a class of convexity weight rules developed by Sherali and Choi (1996) from linear programs to general convex and nonsmooth optimization problems. Larsson et al. (1999) established convergence for two convexity weight rules for general convex programs. These rules are shown to be special cases of the class of convexity weight rules developed in this paper.

We further propose a new set of rules for choosing the convexity weights defining the ergodic primal sequences. The rules proposed are such that—in contrast to the previously proposed rules—more weight is assigned to later subproblem solutions than to earlier ones. In order to evaluate the computational efficiency of the new convexity weight rules, we apply the subgradient method to a set of nonlinear multicommodity network flow problems. The computational results are convincing and show that the new rules create ergodic sequences that converge to the solution set considerably faster than those created by previously developed rules.

**My contributions to the paper:** Performed the major part of the theoretical development, the writing, and the numerical tests.

The paper is published in *Mathematical Programming* **150**:365–390 (2015). Results from the paper have been presented at the 21st International Symposium on Mathematical Programming (ISMP 2012), Berlin, Germany, 2012, and at the conference Nonlinear optimization: A bridge from theory to practice, Erice, Italy, 2013.

## 5.2 Paper II: Ergodic, primal convergence in dual subgradient schemes for convex programming, II—the case of inconsistent primal problems

(coauthored with Torbjörn Larsson, Michael Patriksson, Ann-Brith Strömberg, and Magnus Önnheim)

We employ the dual subgradient method described in Section 2.2.2 on an infeasible convex optimization problem (convex in the sense that the objective and constraint functions are convex, and that the ground set  $X$  is also convex). The dual iterates constructed by the method are shown to diverge in a steepest feasible ascent direction of the Lagrangian dual function.

An interesting question is whether or not the sequence of ergodic iterates still yield relevant information regarding the primal program? We answer this question in the affirmative and show that the primal–dual pair of programs is associated with a saddle-point problem, in which the primal part amounts to finding a solution in the primal space such that the Euclidean norm of the infeasibility in the relaxed constraints is minimized. We show that the sequences of ergodic iterates obtained in the subgradient method converge to a solution with minimum infeasibility; for linear programs, convergence to the subset of such solutions in which the primal objective is at minimum is also achieved.

**My contributions to the paper:** Taken an important part in the theoretical development as well as in the writing of the paper.

The paper has been submitted to *Mathematical Programming, Series A*. Results from the paper have been presented at the 21st International Symposium on Mathematical Programming (ISMP 2012), Berlin, Germany, 2012, and at the Optimization and Systems Theory Seminar, Department of Mathematics, KTH, Stockholm, Sweden, 2012.

## 5.3 Paper III: Recovery of primal solutions from dual subgradient methods for mixed binary linear programs

(coauthored with Torbjörn Larsson, Michael Patriksson, and Ann-Brith Strömberg)

The main contribution of this paper is the description of how ergodic primal iterates can be utilized for constructing feasible solutions to mixed binary linear programs (MBLP). We show that, when employing a subgradient method to the Lagrangian dual of an MBLP, the sequence of ergodic primal iterates converges to the set of solutions to a convexified counterpart of the original MBLP. For problems possessing a small duality gap, one expects that the set of solutions to the convexified version lies close to that of the original problem. We argue, therefore, that it is useful to utilize

the information contained in the sequence of ergodic primal iterates for constructing feasible solutions to the original MBLP.

We propose three procedures for using these ergodic primal iterates to obtain feasible solutions. The first procedure is a Lagrangian heuristic, which utilizes the information contained in the ergodic sequences to construct feasible solutions to the original problem. The second procedure constructs a core problem, in which some of the variables in the original problem have been assigned fixed values. The third procedure introduces a framework for utilizing the dual subgradient method within a branch-and-bound method.

The Lagrangian heuristic is applied to a sample of uncapacitated facility location problem instances; the numerical results show that the proposed heuristics clearly outperforms the traditional Lagrangian heuristic, in which only the latest subproblem solution is utilized for constructing a feasible solution. The core problem procedure is evaluated on a sample of set covering instances; the numerical results show that the size of the core problems created are considerably smaller than the original problems while the solutions obtained are still in many cases optimal in the original problems.

**My contributions to the paper:** Performed the major part of the writing, the numerical tests, and took part in the theoretical development.

## 5.4 Paper IV: Preventive maintenance scheduling of multi-component systems with interval costs

(coauthored with Michael Patriksson, Ann-Brith Strömberg,  
Adam Wojciechowski, and Magnus Önnheim)

In this paper we consider the preventive maintenance (PM) scheduling of a multi-component system with economic dependencies. We model the corresponding decision problem by scheduling a set of PM activities for a system consisting of a set  $\mathcal{N}$  of components over a discretized time horizon represented by the discrete set  $\mathcal{T}$ . A maintenance occasion cost  $d_t$  arises if PM on any component in the system is performed at time  $t \in \mathcal{T}$ . An interval (or deterioration) cost  $c_{st}^i$  has to be paid if PM is performed on component  $i \in \mathcal{N}$  at the times  $s$  and  $t$ , but not in-between.

We introduce the PM scheduling problem with interval costs (PMSPIC), which is that to schedule the PM of the system such that the sum of all the maintenance occasion and interval costs is minimized. A binary linear program (BLP) is developed for the PMSPIC; it was originally introduced for the *dynamic joint replenishment problem* (DJRP) in Joneja (1990).

The contributions of Paper III include theoretical analyses of the complexity and polyhedral properties of the BLP. We show that the integrality restrictions on a subset of the variables in the BLP mode can be relaxed. We also establish that the necessary inequalities define facets of the convex hull of the set of feasible solutions to the BLP model.

Three case studies originating from the railway, aircraft, and wind industries, respectively, demonstrate the practical usefulness of the model. In the case study from the railway industry, the interval costs correspond to costs incurred by the degradation of the rail. In the aircraft and wind turbine cases, the interval costs correspond to the expected corrective maintenance (CM) costs for components exhibiting stochastic failures. Using the model can reduce the maintenance costs by up to 15% compared to the best simple policies.

**My contributions to the paper:** Took a considerable part in the theoretical development, the writing, and performed numerical tests.

The paper is published in *Computers & Industrial Engineering* 76:390–400 (2014). Results from the paper have been presented at the 21st International Symposium on Mathematical Programming (ISMP 2012), Berlin, Germany, 2012, and the Swedish Operations Research Conference and 6th Nordic Optimization Symposium (SOAK /NOS6 2013), Chalmers, Göteborg, Sweden, 2013.

The results were developed within the project "MU26: Optimum inspection and maintenance of rails and wheels" at CHARMEC (Chalmers Railway Mechanics).

## 5.5 Paper V: Scheduling tamping operations on railway tracks using mixed integer linear programming

The problem of scheduling tamping operations on ballasted rail tracks is considered in this article. The objective of the problem is to minimize the total maintenance costs, while retaining an acceptable condition of the geometry of the track. A degradation model is employed to describe the degradation of the geometry of the track segments, and an MBLP model for the scheduling problem is presented and analyzed. The proposed model extends the one proposed by Vale et al. (2011) by including economic dependencies between the rail track segments, as well as by employing a more general objective function to be minimized in the model.

The model is validated through a numerical study comparing the performance of the model with that of simple policies for constructing PM schedules. Using the proposed model reduces the maintenance costs by up to 10 % as compared with the best policy investigated.

The paper is published in *EURO Journal on Transportation and Logistics* 4:97–112, Issue 1: Transportation Infrastructure Management (2015). Results from the paper have been presented at the SIAM Conference on Optimization, San Diego, USA, 2014, and at the conference Industrial Efficiency, SICS, Stockholm, Sweden, 2014.

The results were developed within the project "MU26: Optimum inspection and maintenance of rails and wheels" at CHARMEC (Chalmers Railway Mechanics).

## References

- ALIZADEH, F. (1995) "Interior point methods in semidefinite programming with applications to combinatorial optimization" *SIAM Journal on Optimization* 5:13–51.
- BALAKRISHNAN, A., MAGNANTI, T., and WONG, R. (1989) "A dual-ascent procedure for large-scale uncapacitated network design" *Operations Research* 37:716–740.
- BALAS, E. (1998) "Disjunctive programming: Properties of the convex hull of feasible points" *Discrete Applied Mathematics* 89:3–44.
- BALAS, E., SCHMIETA, S., and WALLACE, C. (2004) "Pivot and shift—a mixed integer programming heuristic" *Discrete Optimization* 1:3–12.
- BAZARAA, M. S., SHERALI, H. D., and SHETTY, C. M. (1993) *Nonlinear Programming. Theory and Algorithms* Wiley, New York, NY, USA 2nd edition.
- BEN-TAL, A. and NEMIROVSKI, A. (2001) *Lectures on Modern Convex Optimization: Analysis, Algorithms, and Engineering Applications* SIAM, Philadelphia, PA, USA.
- BERTSEKAS, D. P. (1979) "Algorithms for nonlinear multicommodity network flow problems" in "International Symposium on Systems Optimization and Analysis," (pages 210–224) Springer-Verlag, Berlin, Germany.
- BOYD, S., KIM, S.-J., VANDENBERGHE, L., and HASSIBI, A. (2007) "A tutorial on geometric programming" *Optimization and engineering* 8:67–127.
- BOYD, S. and VANDENBERGHE, L. (2004) *Convex Optimization* Cambridge University Press, Cambridge, UK.
- BRETTTHAUER, K. M. and SHETTY, B. (1995) "The nonlinear resource allocation problem" *Operations Research* 43:670–683.
- CAPRARA, A., TOTH, P., and FISCHETTI, M. (2000) "Algorithms for the set covering problem" *Annals of Operations Research* 98:353–371.
- CHEN, J., KANJ, I. A., and JIA, W. (2001) "Vertex cover: further observations and further improvements" *Journal of Algorithms* 41:280–301.
- DAKIN, R. J. (1965) "A tree-search algorithm for mixed integer programming problems" *The Computer Journal* 8:250–255.
- DIRICKX, Y. and JENNERGREN, L. (1979) *Systems Analysis by Multi-Level Methods* Wiley, New York, NY, USA.
- ECKER, J. G. (1980) "Geometric programming: Methods, computations and applications" *SIAM Review* 22:338–362.
- ERMOL'EV, Y. M. (1966) "Methods of solution of nonlinear extremal problems" *Cybernetics* 2:1–14.
- FRANGIONI, A. and GENDRON, B. (2009) "0-1 reformulations of the multicommodity capacitated network design problem" *Discrete Applied Mathematics* 157:1229–1241.

- FRANK, M. and WOLFE, P. (1956) "An algorithm for quadratic programming" *Naval Research Logistics Quarterly* **3**:95–110.
- GAREY, M. R. and JOHNSON, D. S. (1979) *Computers and Intractability: A Guide to the Theory of NP-Completeness* W.H. Freeman and Company, New York, NY, USA.
- GENTILE, C., VENTURA, P., and WEISMANTEL, R. (2006) "Mod-2 cuts generation yields the convex hull of bounded integer feasible sets" *SIAM Journal on Discrete Mathematics* **20**:913–919.
- GOEMANS, M. X. (1997) "Semidefinite programming in combinatorial optimization" *Mathematical Programming* **79**:143–161.
- GOFFIN, J.-L. and VIAL, J.-P. (2002) "Convex nondifferentiable optimization: A survey focused on the analytic center cutting plane method" *Optimization Methods and Software* **17**:805–867.
- GÖRTZ, S. and KLOSE, A. (2012) "A simple but usually fast branch-and-bound algorithm for the capacitated facility location problem" *INFORMS Journal on Computing* **24**:597–610.
- HIRIART-URRUTY, J.-B. and LEMARÉCHAL, C. (1993) "Acceleration of the cutting-plane algorithm: Primal forms of bundle methods" in "Convex Analysis and Minimization Algorithms II," volume 306 of *Grundlehren der mathematischen Wissenschaften* (pages 275–330) Springer Berlin Heidelberg, Springer-Verlag, Berlin, Germany.
- HORTON, J. D. and KILAKOS, K. (1993) "Minimum edge dominating sets" *SIAM Journal on Discrete Mathematics* **6**:375–387.
- JONEJA, D. (1990) "The joint replenishment problem: new heuristics and worst case performance bounds" *Operations Research* **38**:711–723.
- KELLEY, J. E., JR (1960) "The cutting-plane method for solving convex programs" *Journal of the Society for Industrial & Applied Mathematics* **8**:703–712.
- KIM, D. and PARDALOS, P. M. (1999) "A solution approach to the fixed charge network flow problem using a dynamic slope scaling procedure" *Operations Research Letters* **24**:195–203.
- KIWIEL, K. C. (1990) "Proximity control in bundle methods for convex nondifferentiable minimization" *Mathematical Programming* **46**:105–122.
- LAND, A. H. and DOIG, A. G. (1960) "An automatic method of solving discrete programming problems" *Econometrica: Journal of the Econometric Society* **1**:497–520.
- LARSSON, T. and LIU, Z. (1997) "A Lagrangean relaxation scheme for structured linear programs with application to multicommodity network flows" *Optimization* **40**:247–284.
- LARSSON, T., LIU, Z.-W., and PATRIKSSON, M. (1997) "A dual scheme for traffic assignment problems" *Optimization* **42**:323–358.
- LARSSON, T., PATRIKSSON, M., and STRÖMBERG, A.-B. (1999) "Ergodic, primal convergence in dual subgradient schemes for convex programming" *Mathematical Programming* **86**:283–312.

- LEMARÉCHAL, C., NEMIROVSKII, A., and NESTEROV, Y. (1995) "New variants of bundle methods" *Mathematical Programming* **69**:111–147.
- LOBO, M. S., VANDENBERGHE, L., BOYD, S., and LEBRET, H. (1998) "Applications of second-order cone programming" *Linear Algebra and its Applications* **284**:193–228.
- LODI, A. (2010) "Mixed integer programming computation" in M. Jünger, T. M. Liebling, D. Naddef, G. L. Nemhauser, W. R. Pulleyblank, G. Reinelt, G. Rinaldi, and L. A. Wolsey (editors) "50 Years of Integer Programming 1958–2008," (pages 619–645) Springer-Verlag, Berlin, Germany.
- LOKKETANGEN, A. and GLOVER, F. (1998) "Solving zero-one mixed integer programming problems using tabu search" *European Journal of Operational Research* **106**:624–658.
- MURTY, K. G. (1983) *Linear Programming* John Wiley & Sons, New York, NY, USA.
- NEMHAUSER, G. L. and WOLSEY, L. A. (1989) *Integer and combinatorial optimization* John Wiley & Sons, New York, NY, USA.
- NESTEROV, Y., NEMIROVSKII, A., and YE, Y. (1994) *Interior-point polynomial algorithms in convex programming* SIAM, Philadelphia, PA, USA.
- NICOLAI, R. P. and DEKKER, R. (2008) "Optimal maintenance of multi-component systems: A review" in K. A. H. Kobbacy and D. N. P. Murthy (editors) "Complex System Maintenance Handbook," (pages 263–286) Springer Series in Reliability Engineering, Springer-Verlag, Berlin, Germany.
- PAPADIMITRIOU, C. H. (1994) *Computational Complexity* Addison-Wesley, Boston, MA, USA.
- PARRILO, P. A. and LALL, S. (2003) "Semidefinite programming relaxations and algebraic optimization in control" *European Journal of Control* **9**:307–321.
- PATRIKSSON, M. (1994) *The Traffic Assignment Problem: Models and Methods* Topics in Transportation series, VSP, Utrecht, The Netherlands.
- PATRIKSSON, M. (2008) "A survey on the continuous nonlinear resource allocation problem" *European Journal of Operational Research* **185**:1–46.
- PHAM, H. and WANG, H. (2000) "Optimal  $(\tau, T)$  opportunistic maintenance of a  $k$ -out-of- $n$ :G system with imperfect PM and partial failure" *Naval Research Logistics* **47**:223–239.
- PINTELON, L. and GELDERS, L. F. (1991) "Maintenance management decision making" *European Journal of Operational Research* **58**:301–317.
- POLYAK, B. (1967) "A general method of solving extremum problems" *Soviet Mathematics Doklady* **8**:593–597.
- POTRA, F. A. and WRIGHT, S. J. (2000) "Interior-point methods" *Journal of Computational and Applied Mathematics* **124**:281–302.
- ROCKAFELLAR, R. T. (1993) "Lagrange multipliers and optimality" *SIAM Review* **35**:183–238.
- SHEPILOV, M. (1976) "Method of the generalized gradient for finding the absolute minimum of a convex function" *Cybernetics and Systems Analysis* **12**:547–553.

- SHERALI, H. D. and CHOI, G. (1996) "Recovery of primal solutions when using subgradient optimization methods to solve Lagrangian duals of linear programs" *Operations Research Letters* **19**:105–113.
- SHOR, N. (1985) *Minimization Methods for Non-Differentiable Functions* Springer-Verlag, Berlin, Germany.
- SHOR, N. Z. (1967) "On the structure of algorithms for numerical solution of problems of optimal planning and design" Ph.D. thesis, Cybernetics Institute, Academy of Sciences of the Ukrainian SSR, Kiev, Russia.
- VALE, C., RIBEIRO, I. M., and CALÇADA, R. (2011) "Integer programming to optimize tamping in railway tracks as preventive maintenance" *Journal of Transportation Engineering* **138**:123–131.
- VANDENBERGHE, L. and BOYD, S. (1996) "Semidefinite programming" *SIAM Review* **38**:49–95.
- VON HOHENBALKEN, B. (1977) "Simplicial decomposition in nonlinear programming algorithms" *Mathematical Programming* **13**:49–68.
- WANG, H. (2002) "A survey of maintenance policies of deteriorating systems" *European Journal of Operational Research* **139**:469–489.
- WELLMAN, M. (1993) "A market-oriented programming environment and its application to distributed multicommodity flow problems" *Journal of Artificial Intelligence Research* **1**:1–23.
- WOLSEY, L. A. (1998) *Integer programming* John Wiley & Sons, New York, NY, USA.
- YAO, D. D., ZHANG, S., and ZHOU, X. Y. (2001) "Stochastic linear-quadratic control via semidefinite programming" *SIAM Journal on Control and Optimization* **40**:801–823.