論文 / 著書情報
Article / Book Information

題目(和文)	探索空間拡張による二次制約非凸最適化問題に対する非凸緩和
Title(English)	Nonconvex relaxation for quadratically constrained nonconvex optimization problems by expanding the space of decision variables
著者(和文)	森耕平
Author(English)	
出典(和文)	学位:博士(工学), 学位授与機関:東京工業大学, 報告番号:甲第5173号, 授与年月日:2002年3月26日, 学位の種別:課程博士, 審査員:
Citation(English)	Degree:Doctor (Engineering), Conferring organization: Tokyo Institute of Technology, Report number:甲第5173号, Conferred date:2002/3/26, Degree Type:Course doctor, Examiner:
学位種別(和文)	博士論文
Type(English)	Doctoral Thesis

Nonconvex Relaxation for Quadratically Constrained Nonconvex Optimization Problems by Expanding the Space of Decision Variables

Kohei MORI

supervised by Professor Shinji HARA

Department of Computational Intelligence and Systems Science
Tokyo Institute of Technology

Acknowledgements

I would like to express my gratitude to my supervisor Professor Shinji Hara for his advice and continuing encouragement over the past five years.

I am also grateful to Professor Masakazu Kojima for his valuable comments. Advice in tuning the tabu search algorithms given by Dr. Katsuki Fujisawa of Kyoto University and comments on my paper given by Dr. Kouichi Taji of Osaka University are gratefully acknowledged. I also thank present and former members of Hara laboratory, Matsuno laboratory, Kojima laboratory, and many other individuals.

Abstract

Relaxation problems for NP-hard optimization problems are often adopted as a subroutine of optimization algorithms. In most of those situation, the relaxation problems are formulated and/or solved as convex optimization problems since it is generally required that they should be solved efficiently. In this thesis, we formulate and analyze a sort of relaxation problems. However, our relaxation problems are not convex in general.

The convex relaxation problems can be solved efficiently. However, they have large gaps to the original problem especially when the original problem is difficult. Therefore, the convex relaxation problems can be not useful at all in constructing efficient algorithms to solve the original problems.

On the other hand, nonconvex relaxation problems have at least a desirable property

> "They are able to have less gaps to the original problem in comparison with the convex relaxation problems,"

despite nonconvex relaxation problems are hardly considered as a set of tools for optimization. The observation implies that nice properties embedded in nonconvex relaxation problems by relaxing tend to provide tools directly useful for solving the original problems.

The contributions of this thesis are listed as follows.

- We formulate a sort of relaxation problems that have parameters representing the degree of relaxation. As a consequence, the relaxation problems include the original problem and the convex relaxation problem as two extreme cases. Therefore, the formulation allows us to consider trade-off between desirable properties in computation and similarity to the original problem.

- We reveal some fundamental properties of the proposed nonconvex relaxation problems. They do not require convexity for being derived.

- We establish some links between our analytical results and known optimization algorithms, and construct two heuristic algorithms. Additionally, we evaluate the performance of the proposed optimization algorithms in numerical experiments.

The class of the original problems in this thesis is minimization of a quadratic function under quadratic constraints. We formulate a sort of relaxation problems for the class of original problems by generalizing each real scalar variables of the original problem to a hypercomplex number. This formulation allows us to make a trade-off between tractability appears in the convex relaxation problems and the precise description of requirements from applications described as the original problems. We expect that we can construct new optimization algorithms and optimize the performance of optimization algorithms by adjusting the degree of relaxation.

Firstly, we especially pay attention to the existence of a "monotone path". Each monotone path is a continuous curve in the feasible region of the relaxation problem having the following property :

> If we change the decision variables along the curve, the objective value monotonically increases or decreases. Consequently, the decision variable can reach the global optimum of the original problem from any other feasible solution by changing the decision variables along the curve.

The existence of a monotone path does not require convexity of the relaxation problem. In consequence, the relaxation problem nearest to the original problem among relaxations, which guarantee the existence of a monotone path, is the relaxation problem in which each real variable is generalized to a ordinary complex number.

Secondly, we restrict the original problems to the maximum cut problems and derive detailed analytical results. Each element of a hypercomplex discrete space, which is a generalization of set of corners of the hypercube, is always on a monotone path. Moreover, the relaxation does not change the optimal value in spite the feasible region is expanded. Additionally, we can obtain a feasible solution of the original problem from that of the relaxation problem without increasing the value of the objective function. These results serve to generalize some known optimization algorithms including Goemans and Williamson's approximation algorithms for the maximum cut problems. We also show that any local optima of the nonconvex relaxation problem to the ordinary complex numbers gives a lower bound of the optimal value of the original problem for certain classes of problem instances.

Finally, we construct two optimization algorithms for the maximum cut problems based on our analytical results. One of the algorithms explicitly utilizes analytical results. The another is based on the existence of a monotone path and some conjectures. Computational experiments imply the availability of our approach. More concretely, the second proposed algorithm showed the performance better than the tabu search algorithm. The performance of the another proposed algorithm grew better as the dimension of hypercomplex number became higher.

Contents

Chapter 1

Introduction

In this chapter we explain the motivation and ideas of the research described in this thesis. Our objective is to provide a mathematical basis in constructing optimization algorithms for optimization problems by using relaxation problems, which are not convex in general and have less gaps to the original problems than the convex relaxation problems.

1.1 Background and Motivation

Many problems appear in engineering, economics, physics, and so on can be formulated as optimization problems. In spite the problems under consideration may be from manifestly different fields, common methods or ideas can be used in obtaining solutions for all of them.

Those optimization problems are generally formulated as

$$(\mathbf{P}) \quad \text{Minimize} \quad g(\boldsymbol{x}) \quad \text{subject to} \quad \boldsymbol{x} \in \mathcal{D}.$$

Here, the set $\mathcal{D}$ called the feasible region and the function $g : \mathcal{D} \to \mathbb{R}$ called the objective function are defined appropriately reflecting the requirements from applications.

Efficient methods or algorithms for solving **P** have been established for cases where g and $\mathcal{D}$ are convex such as linear programming [1, 2] and semidefinite programming [3, 4, 5], and cases where g and $\mathcal{D}$ have special combinatorial structures [6, 7] such as the problem of finding a minimal spanning tree of a graph and several subclasses of the maximum cut problems. In principle, such **P** can be solved by the following procedure :

For given $\bar{\boldsymbol{x}} \in \mathcal{D}$, solve the following optimization problem

$$(\mathbf{L}) \quad \text{Minimize} \quad g(\boldsymbol{x}) \quad \text{subject to} \quad \boldsymbol{x} \in \mathcal{D} \cap \mathcal{N}(\bar{\boldsymbol{x}})$$

repeatedly, where $\mathcal{N}(\bar{\boldsymbol{x}})$ is some neighborhood of $\bar{\boldsymbol{x}}$.

In other words, only the local information is required for solving **P**. On the other hand, in general, we need global information about g and $\mathcal{D}$ for solving **P**. Moreover, it is considered that there is no algorithm for solving general classes of **P** efficiently in the sense that the time complexity for solving **P** is bounded by a polynomial of the size of the data for describing **P**, since **P** is NP-hard [8].

To handle such hard problems, "algorithms called heuristics [9, 10, 11, 12, 13, 15]" and "relaxation problems for **P** [4, 16, 17, 18] " are often adopted. Heuristics can be understood as a variety of local search algorithms, and often are strategies or algorithms that try to find elements of $\mathcal{D}$ which are not necessarily optimal but are acceptable in real applications. The following are examples of such algorithms.

Local Search [9] : Solve **L** repeatedly until $\boldsymbol{x}$ converges ; the whole procedure is also carried out repeatedly.

Tabu Search [9, 12, 13, 14] : $\mathcal{N}(\cdot)$ is updated during the repetitive computation of **L** using the histories of $g(\boldsymbol{x})$ and $\boldsymbol{x}$.

Simulated Annealing [9, **11**] : **L** is carried out repeatedly with a probabilistic mechanism that gradually converges to a deterministic one.

Genetic Algorithms [**9, 10, 15**] : Solve **L**s in parallel , and $\boldsymbol{x}$ of each **L** are copied, modified, and/or deleted using information about whole $\boldsymbol{x}$s and $g(\boldsymbol{x})$s.

Relaxation problems are optimization problems of which some constraints in **P** are relaxed. In other words, the feasible region $\mathcal{D}$ of the original problem can be defined by

$$\mathcal{D} := \mathcal{D}' \cap \mathcal{X} \tag{1.1}$$

under an appropriate transformation of decision variables, where $\mathcal{D}'$ is the feasible region of the relaxation problem and $\boldsymbol{x} \in \mathcal{X}$ is the constraint relaxed in defining the relaxation problem.

Relaxation problems often are designed by transforming **P** in optimization problems that can be solved efficiently in numerical computation. They are often used to obtain algorithms with theoretically guaranteed performance [16, 17, 18, 19] and to reduce the effort in obtaining the optimal solution of **P** [4, 21, 22, 23, 33, 34, 44]. In such situations, being convex is often treated as if it is a part of definition of relaxation problems.

For relaxation problems, it is expected in usual that

- The optimal value of the relaxation problem is a good bound of the optimal value of the original problem.
- We can solve the relaxation problem efficiently in numerical computation.

From these reasons, almost all relaxation problems for NP-hard optimization problems are formulated and/or solved as convex optimization problems. The following properties are ensured in the convex optimization problem from its definition.

1. $ta + (1-t)b \in \mathcal{D}$ for $t \in [0,1]$ if $a, b \in \mathcal{D}$.

2. $g(ta + (1-t)b) \leq tg(a) + (1-t)g(b)$ for any $a, b \in \mathcal{D}$ and $t \in [0,1]$.

The definition of convex optimization problems ensures very strong properties. Namely, e.g., any local optimum of the problem is also a global optimum. Besides, polynomial time optimization algorithms are available for solving convex optimization problems [1, 3, 4, 5]. Thereby, strategies such as calculating the bound of the objective function and constructing polynomial time approximation algorithms become available strategies. Here, the solutions of convex relaxation problems are used as illustrated in the following :

(i) Let $\bar{X} \in \mathcal{D}'$ be a solution of the convex relaxation problem of $\mathbf{P}$ defined by

$$(\ \mathbf{P}'\ ) \quad \text{Minimize} \quad G(X) \quad \text{subject to} \quad X \in \mathcal{D}'.$$

Additionally, let $f : \mathcal{D}' \to \mathcal{D}$. Then, $\bar{x} = f(\bar{X})$ becomes a good approximation of the solution of $\mathbf{P}$.

(ii) Assume the feasible region $\mathcal{D}$ of $\mathbf{P}$ is represented as

$$\mathcal{D} = \mathcal{D}^1 \cup \mathcal{D}^2 \cup \cdots \cup \mathcal{D}^a$$

and $\mathcal{D}^i \cap \mathcal{D}^j = \emptyset$ for $i \neq j$. Let an $\bar{x}$ be in $\mathcal{D}^1$ and $\bar{X} \in \mathcal{D}'^2$ be an optimal solution of the convex relaxation problem for the original problem

$$(\ \mathbf{P}^2\ ) \quad \text{Minimize} \quad g(x) \quad \text{subject to} \quad x \in \mathcal{D}^2.$$

Then we can conclude that the solution of $\mathbf{P}$ is not in $\mathcal{D}^2$ if $g(\bar{x}) < G(\bar{X})$.

These observation allow to handle a variety of NP-hard problems by branch and bound methods. More concretely, the fact (ii) shrinks the feasible region and the approximation in (i) gives the better result.

On the contrary, in general, convex relaxation problems for difficult original problems have large gaps to the original problems that we actually want to solve. The gaps appear as, for example,

- The difference between the optimal value of the original problem and that of the convex relaxation problem.

This means that $g(\bar{x}) < G(\bar{X})$ hardly holds in (ii). Moreover, the gaps may appear as

- Theoretical guarantees on the approximation which do not match heuristics in real computation.

That is, an algorithm which has the best theoretical guarantee is often worse than heuristics when the theoretical guarantees are not vital.

In summary, the good properties found in convex relaxation problems is considered to be useless for solving the original problem because of the large gaps to the original problem.

We consider that it is useful to formulate a sort of relaxation problems in which minimal good properties are embedded by relaxing, and structures of the original problem are somewhat preserved. Nonconvex relaxation problems can be better than convex relaxation problems in the sense that it will have less gaps to the original problem compared to the convex relaxation problems. We expect the following for using nonconvex relaxation problems which are closer to the original problems :

- Generalizing or modifying optimization algorithms for the original problem will be available in constructing optimization algorithms based on nonconvex relaxation problems when they have structures resemble to the original one. Moreover, we will be able to enhance the performance of the algorithms if the relaxation problems have certain nice properties.
- If any local optima of the nonconvex relaxation problems provide a lower bound for the original problem under a formulation in which the relaxation problems connect the original problem and the convex relaxation problem continuously, then the bound of the optimal value never be worse than those given by the convex relaxations.

On these viewpoints, we define a sort of nonconvex relaxation problems for quadratically constrained minimization problems and analyze their fundamental properties.

The relaxation problems in this thesis are defined in such a way that each real variable of the original problem is expanded to the space consists of hypercomplex numbers [24]. The resulting relaxation problems are not convex in general, and contain the original problem and the semidefinite programming(SDP) relaxation problem as two extreme cases as we show later. This formulation allows us to make trade-off between tractability appears in the convex relaxation problem and the precise description of requirements from real applications described as the original problem.

The convex relaxation has been well studied. In particular, polynomial time approximation algorithms and techniques for obtaining a lower bound of the optimal value are known for many subclasses of **P** by using SDP relaxation developed in the last decade [16, 17, 18, 19]. Our approach in this thesis will connect such strategies using convex relaxation with strategies which handle **P** directly. On the other hand, some researches based on mathematical representations, which are partially similar to ours, are reported for combinatorial optimization [17, 18], associative memories [26, 27, 28], spin glasses [30], and communication [29]. However, the descriptions are mainly for representing a set of information as one variable. Therefore, those problems or approaches are different from ours.

1.2 Idea of Approach

The idea of our approach will be explained by using a simple example. Consider the following optimization problem :

$$\begin{array}{ll} \text{Minimize} & -x^*x - \frac{1}{2}(x + x^*) \\ \text{Subject to} & \mid x \mid \leq 1, \end{array}$$

where $\mathcal{X}$ in eq.(1.1) is set of real numbers $\mathbb{R}$ and * denotes the complex conjugate. If $x \in \mathbb{R}$, then we can really see that the local optima of the problem are $x = 1$ and $x = -1$. The unique global optimum is $x = 1$ as depicted in Figure 1.1.

Let us now consider a relaxation problem where $x \in \mathbb{C}$. The problem is written as

$$\begin{array}{lll} \text{Minimize} & : & -\left\{\Re(x)^2 + \Im(x)^2\right\} - \Re(x) \\ \text{Subject to} & : & \Re(x)^2 + \Im(x)^2 - 1 \leq 0. \end{array}$$

We can see that any local optimum of the relaxation problem must satisfy

$$\Re(x)^2 + \Im(x)^2 - 1 = 0,$$

since the Hessian matrix of the objective function with respect to the real and the imaginary part of x is written as

$$\begin{pmatrix} -2 & 0 \\ 0 & -2 \end{pmatrix}$$

; it is negative definite. Hence, the optimization problem is reduced to

$$\text{Minimize} \quad : \quad -1 - \cos\theta$$

by replacing $\Re(x)$ with $\cos\theta$. Consequently, the unique local optimum of the problem is $x = 1$, which is coincident with the global optimum of the original problem. For this reason, the decision variable can reach the global optimum of the original problem from any other feasible solution of the original problem without increasing the objective value.

Note that the relaxation problem where $x \in \mathbb{C}$ is still not convex. However the relaxation problem has a desirable property "The problem does not have any local optimum except for the global one." Hence, we can obtain the global optimum of the original problem by a decent type local search algorithm starting from an arbitrary feasible solution. On the other hand, the original problem is equivalent to minimizing under the constraint $x \in \{-1, +1\}$ since the objective function of the original problem is concave. Consider also another relaxation problem in which the feasible region is generalized to

$$x \in \{-1, +1, -\boldsymbol{j}, +\boldsymbol{j}\},$$

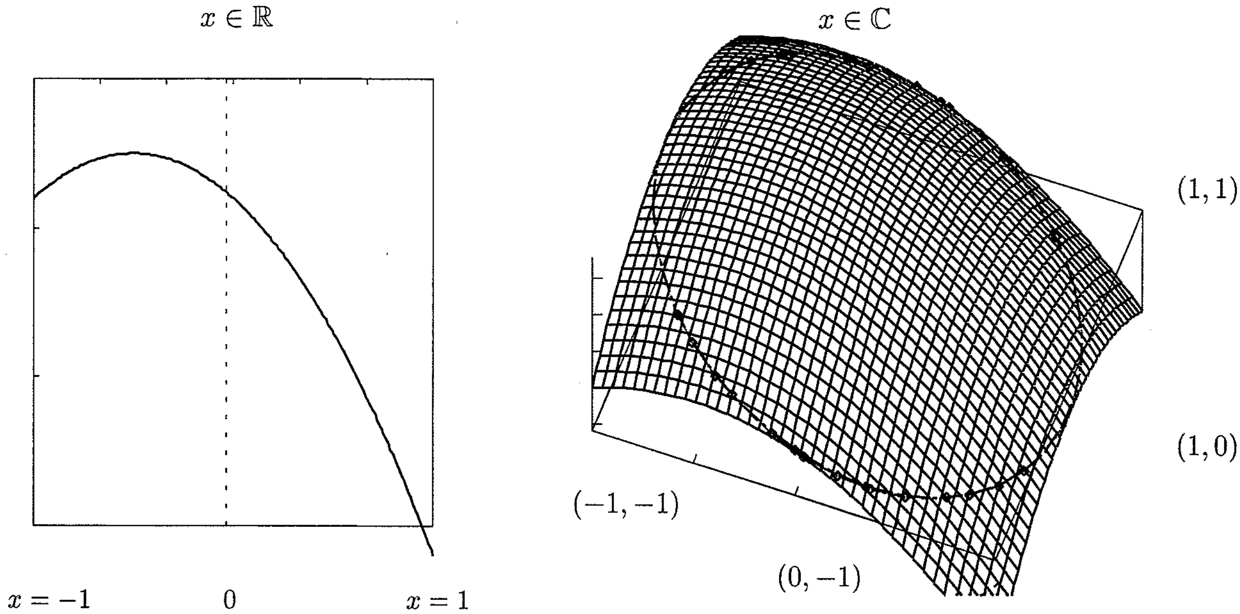

Figure 1.1: $x \in \mathbb{R}$(left) and $x \in \mathbb{C}$(right)

where j stands for imaginary unit. In this case, the optimal value of the relaxation problem coincides with that of the original problem. In other words, the relaxation problem is equivalent to the original problem in the sense that the optimal value does not change.

We can hardly expect that almost all nice properties derived from convexity to hold since the relaxation problems in this thesis are not convex in general. Nevertheless, we can expect some favorable properties to be revealed for certain classes of problems instances. Moreover, those properties will be useful for constructing optimization algorithms for solving the original problems since the gaps between the nonconvex relaxation problem and the original problem will be less in comparison with the convex relaxation.

1.3 Outline of Thesis

This thesis is divided into 6 chapters. Through the analysis of nonconvex relaxation problems, we show several nice properties we can obtain without relaxing the original problem to a convex problem. Based on the analytical results, we will construct optimization algorithms and confirm the effects of using the nonconvex relaxation through computational experiments.

The summaries of forthcoming chapters are given as follows.

Chapter 2 : Chapter 2 defines optimization problems to be analyzed in this thesis. The original problem is minimization of indefinite quadratic function under nonconvex quadratic constraints. The relaxation problems are defined in such a way that each real scalar decision variable of the original problem is generalized to a hypercomplex number. The

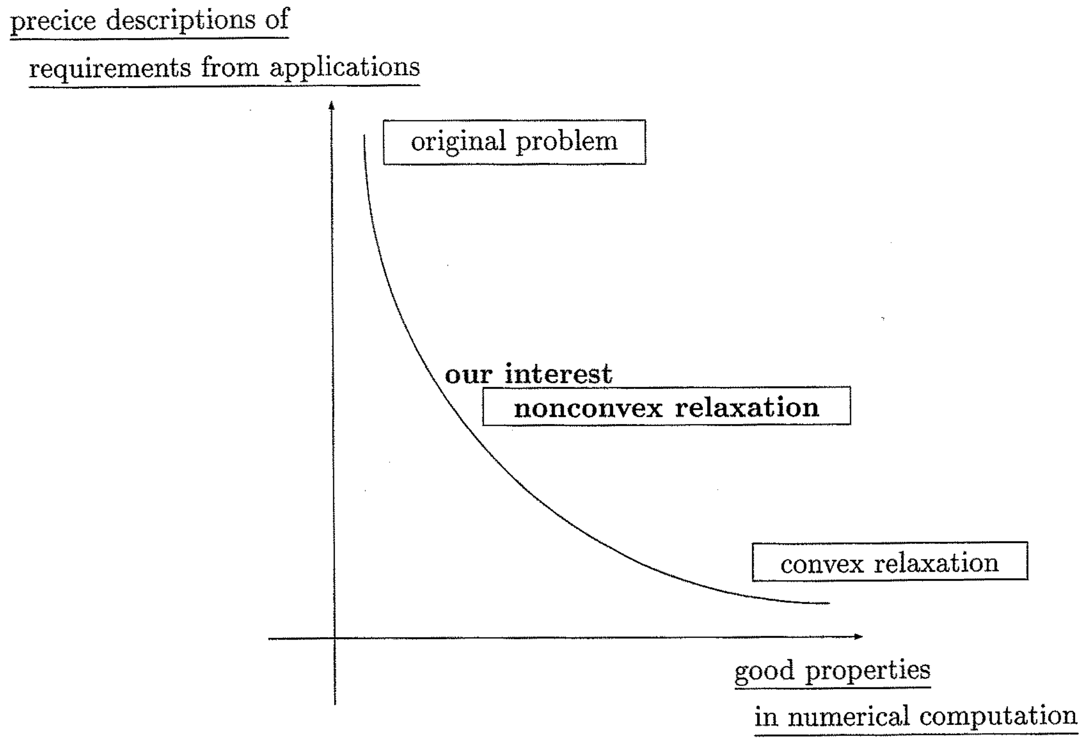

Figure 1.2: location of nonconvex relaxation problems

relaxation problems introduce a concept "degree of relaxation".

Chapter 3 : Chapters 3 and 4, which are the main parts of this thesis, are dedicated for describing analytical results [35, 36]. In chapter 3, we derive fundamental properties found in nonconvex relaxation problems of quadratically constrained original problems. The relaxation problems include the original problem and the convex relaxation problem in term of semidefinite programming as two extreme cases. Here, we especially pay attention to whether the decision variable can reach the global optimum of the original problem with changing the value of the objective function monotonically moving through the relaxed feasible region. We show conditions for the existence of a continuous curve $\boldsymbol{l}(t)$. We call the curve a monotone path, which is contained in the feasible region of the relaxation problem and satisfies

$$\boldsymbol{l}(0) = \boldsymbol{x} \ , \ \boldsymbol{l}(1) = \boldsymbol{y} \quad \text{and} \quad 1 \geq s > t \geq 0 \ \Leftrightarrow \ E(\boldsymbol{l}(s)) < E(\boldsymbol{l}(t)),$$

where $E(\cdot)$ is the objective function, and $\boldsymbol{x}$ and $\boldsymbol{y}$ are arbitrary feasible solutions of the original problem. We show the conditions about degree of relaxation for the existence of a monotone path. In consequence, the existence of a monotone path does not require convexity of the relaxation problem. We also derive some conditions for the local optimality of feasible solutions in the nonconvex relaxation problem of which the dimension of the expanded space is two.

Chapter 4 : Chapter 4 concentrates on a subclass of the the original problems where each real variable takes value $+1$ or -1. We investigate detailed properties of the nonconvex relaxation problems, and show that

- By applying the existence of a monotone path, some optimization algorithms designed for the original problems, which includes Geomans and Williamson's approximation algorithm for the maximum cut problems, can be performed in the space consists of hypercomplex numbers without changing the central ideas of the algorithms.
- For some classes of problem instances, any local optimum of a nonconvex relaxation problem provides a lower bound of the optimal value of the original problem.

Chapter 5 : Chapter 5 provides two optimization algorithms which exploit analytical results described in Chapters 3 and 4. We demonstrate some results of computational experiments for the maximum cut problems. We confirm that the performance of one of the proposed algorithms is better than the tabu search algorithm.

Chapter 6 : Chapter 6 summarizes the results in this thesis and mentions interesting future research topics.

1.4 Symbols and Notation

We use the following symbols and notation.

$\mathbb{Z}$	:	set of all integers
$\mathbb{R}$	:	set of all real numbers
$\mathbb{C}$	:	set of all complex numbers
$\mathbb{F}_M$	:	set of all M-dimensional hypercomplex numbers with a definition of multiplication
$\Re(a)$	:	real part of a complex or a hypercomplex number a
$\Im(a)$	:	imaginary part of a complex number a
$\boldsymbol{j}$	:	imaginary unit of complex numbers
a^*	:	complex conjugate of a complex or hypercomplex number a
$\mathcal{S}_N$	:	set of all $N \times N$ real symmetry matrices
$\mathcal{S}_N^0$	:	set of all $N \times N$ real symmetry matrices whose diagonal entries are all zero
$\boldsymbol{A}^{\mathrm{T}}$	:	transpose of a vector or a matrix $\boldsymbol{A}$
$\boldsymbol{e}_i$	:	the vector with ith entry 1, and 0 elsewhere
$\boldsymbol{E}_{ij}$	:	the matrix having i-jth entry 1, and 0 elsewhere
$\boldsymbol{E}_i(\phi)$	:	the diagonal matrix having i-ith entry $e^{\boldsymbol{j}\phi}$, and 1 at any other diagonal entries
$\boldsymbol{E}_i^-(\phi)$	:	the diagonal matrix with i-ith entry 1, and $e^{\boldsymbol{j}\phi}$ at any other diagonal entries
$\mathbf{0}$	:	the vector of all zero

$\boldsymbol{O}$	:	the square matrix of all zero
$\boldsymbol{A} \succeq \boldsymbol{B}$	:	a matrix $\boldsymbol{A} - \boldsymbol{B} \in \mathcal{S}_N$ is positive semidefinite

Matrices $\boldsymbol{W}$, $\boldsymbol{A}_i$, and $\bar{\boldsymbol{A}}_i$ and scalars M, $\mathcal{K}$ and ϵ are used for representing parameters or data of the original and the relaxation problems in whole parts of this thesis.

$\boldsymbol{W}$	:	coefficient matrix of the quadratic objective function
$\boldsymbol{A}_i$, $\bar{\boldsymbol{A}}_i$	:	coefficient matrices of the quadratic constraints
M	:	dimension of hypercomplex numbers in relaxation problems
ϵ	:	a positive parameter used for defining the feasible region of the relaxation problems
$\mathcal{K}$	:	a subset of natural numbers which defines feasible region of the original problems

Once they are defined, we will use them without stating explicitly.

Chapter 2

Problem Formulation

In this chapter, we formulate the class of original problems and their relaxation problems considered in this thesis. The original problem **QP** is minimization of indefinite quadratic function under nonconvex quadratic constraints. This class is NP-hard and can capture a lot of applications in wide disciplines. The class of relaxation problems $\mathbf{QP}_M^\epsilon$ is defined in such a way that each real scalar variable of the original problem is generalized to a M-dimensional hypercomplex number. The relaxation problems are not convex in general and have two parameters, namely, the dimension M of hypercomplex numbers and stretch ϵ of the feasible region in the directions of imaginary axes. Moreover, the relaxation problems include the original problem as a special case.

All original problems considered are minimizations of quadratic function under quadratic constraints. The terminologies "feasible solution", "local optimum", "global optimum", and so on are defined here for minimization **P** of the objective function $g(\boldsymbol{x}) \in \mathbb{R}$ in the feasible region $\mathcal{D}$. The notation $\|\boldsymbol{x}\|$ appears in definitions is some appropriate norm of $\boldsymbol{x}$.

Definition 2.1 *$\bar{\boldsymbol{x}}$ is a feasible solution of* **P** *if and only if $\bar{\boldsymbol{x}} \in \mathcal{D}$.*

Definition 2.2 *$\bar{\boldsymbol{x}} \in \mathcal{D}$ is a local optimum of* **P** *if and only if there exists a real constant $\delta > 0$ satisfying*

$$g(\bar{\boldsymbol{x}}) \le g(\boldsymbol{x}) \quad \text{for all} \quad \boldsymbol{x} \in \mathcal{D} \text{ such that } \|\boldsymbol{x} - \bar{\boldsymbol{x}}\| < \delta.$$

Definition 2.3 *$\bar{\boldsymbol{x}} \in \mathcal{D}$ is a strict local optimum of* **P** *if and only if there exists a real constant $\delta > 0$ satisfying*

$$g(\bar{\boldsymbol{x}}) < g(\boldsymbol{x}) \quad \text{for all } \boldsymbol{x} \in \mathcal{D} \text{ such that } 0 \neq \|\boldsymbol{x} - \bar{\boldsymbol{x}}\| < \delta.$$

Definition 2.4 *$\bar{\boldsymbol{x}} \in \mathcal{D}$ is a global optimum of* **P** *if and only if*

$$g(\bar{\boldsymbol{x}}) \le g(\boldsymbol{x}) \ ; \ \forall \boldsymbol{x} \in \mathcal{D}.$$

We also call such an $\bar{\boldsymbol{x}}$ a optimal solution of **P**.

Definition 2.5 *A real scalar α is the optimal value of* $\mathbf{P}$ *if and only if*

$$\alpha = g(\bar{\boldsymbol{x}}),$$

where $\bar{\boldsymbol{x}}$ is a global optimum of $\mathbf{P}$.

2.1 Original Problem

Let $\mathcal{K} \subset \{1, \cdots, N\}$, $\boldsymbol{W} \in \mathcal{S}_N$, $\boldsymbol{b}_i \in \mathbb{R}^{N-1}$, and $c_i \in \mathbb{R}$. In addition, let $\bar{\boldsymbol{A}}_i$ be defined by

$$\bar{\boldsymbol{A}}_i = \begin{pmatrix} c_i & \boldsymbol{b}_i^{\mathrm{T}} \\ \boldsymbol{b}_i & \boldsymbol{A}_i \end{pmatrix},$$

where $\boldsymbol{A}_i \in \mathcal{S}_{N-1}$. We describe a class of minimization problems as **QP**.

$$(\mathbf{QP}) \quad \begin{array}{lll} \text{Min.} & : & E(\boldsymbol{x}) = \boldsymbol{x}^{\mathrm{T}}\boldsymbol{W}\boldsymbol{x} \\ \text{Subj. to} & : & \boldsymbol{x} \in \mathcal{D}_{q1} = \left\{ \boldsymbol{x} \in \mathbb{R}^N \,\middle|\, \begin{array}{l} \boldsymbol{x}^{\mathrm{T}}\bar{\boldsymbol{A}}_i\boldsymbol{x} \leq 0 \ \text{ for } \ i = 1, \cdots, L \\ x_i \in \{-1, +1\} \ \text{ for } \ i \in \mathcal{K} \\ x_1 = 1 \end{array} \right\} \end{array}$$

The class **QP** is minimization of a quadratic function under quadratic constraints. In fact, the constraint $x_i \in \{-1, +1\}$ is redundant since it can be described by two quadratic inequality constraints. However, we use this description of **QP** for simplicity of later analyses.

The 0-1 quadratic programming(or the maximum cut problems) is reduced to $\mathbf{P}_1$, a subclass of **QP** with $\mathcal{K} = \emptyset$, by letting all diagonal entries of $\boldsymbol{W}$ be 0.

$$(\mathbf{P}_1) \quad \begin{array}{lll} \text{Min.} & : & E(\boldsymbol{x}) = \boldsymbol{x}^{\mathrm{T}}\boldsymbol{W}\boldsymbol{x} \\ \text{Subj. to} & : & \boldsymbol{x} \in \mathcal{D}_1 = \left\{ \boldsymbol{x} \in \mathbb{R} \,\middle|\, \begin{array}{l} x_i \in [-1, +1] \ \text{ for } \ i = 1, \cdots, N \\ x_1 = 1 \end{array} \right\} \end{array}$$

Of course, the problem also can be defined as the case where $\bar{\boldsymbol{A}}_i = \boldsymbol{O}$ and $\mathcal{K} = \{1, \cdots, N\}$. We begin analyses with the case the original problem is in **QP** and present detailed results for the special cases in chapter 4.

The class **QP** can capture a lot of applications, for example, optimal control problems of hybrid systems [46, 47] and many problems appear in planning and location [32]. Even if **QP** is restricted to $\mathbf{P}_1$, it covers problems appear in circuit layout design [31], statistical physics [30, 31], and so on. Furthermore, $\mathbf{P}_1$ is NP-hard in spite its simple outcome [39].

If $\mathcal{K} = \emptyset$, $\boldsymbol{A}_i \succeq \boldsymbol{O}$, and $\boldsymbol{W} \succeq \boldsymbol{O}$, then **QP** becomes a class of convex optimization problem. Therefore, **QP** can be solved efficiently by polynomial time interior point algorithms [3] in those cases. For some NP-hard subclasses of **QP**, optimization algorithms having average case time complexity which is not so worse are known [41, 42]. However, in general, no efficient optimization algorithm is known for the general class of **QP**, and the existence of a polynomial time algorithm is considered to be hopeless. Therefore, heuristics like tabu search, genetic algorithms, and simulated annealing are adopted in obtaining feasible solutions which are acceptable in real applications. As well as this, strategies like branch and bound method are selected to reduce the computational effort required to obtain the optimal solution of **QP** in many situations.

2.2 Nonconvex Relaxation Problem

Our formulation of relaxation problems is based on generalizing scalar decision variables to higher dimensional numbers. We use terminology "hypercomplex number" for calling the multi-dimensional numbers mainly because

- We want use properties such that a set of lower dimensional numbers can be defined as a subset of higher dimensional numbers.
- Most of our analyses are described in term of the ordinary complex numbers.

We can simplify the definition of hypercomplex numbers as far as it is in our analyses though there are freedoms in defining multiplication of imaginary units. Therefore, using the terminology "hypercomplex numbers" helps to simplify our formulation of relaxation problems. For the detail of hypercomplex numbers, see appendix A or reference [24].

2.2.1 Expanding the space of decision variables to the Space consists of Hypercomplex Numbers

At first, we generalize the real decision variable

$$\boldsymbol{x} = \begin{pmatrix} x_1, & \cdots, & x_N \end{pmatrix}^{\mathrm{T}} \in \mathbb{R}^N$$

of **QP** to the hypercomplex number

$$\begin{aligned} \boldsymbol{x} &= \begin{pmatrix} \mathbf{x}_1 & \cdots & \mathbf{x}_N \end{pmatrix}^{\mathrm{T}} \in \mathbb{F}_M^N \\ \mathbf{x}_i &= \begin{pmatrix} x_{i1} & \cdots & x_{iM} \end{pmatrix} \in \mathbb{F}_M, \end{aligned}$$

where each $\mathbf{x}_i$ is a hypercomplex number and x_{ij}s are real variables. We call x_{i1} the real part of the hypercomplex number $\mathbf{x}_i$ and represent the real part by $\Re(\mathbf{x}_i)$. We use notation $\boldsymbol{x}^{(i)}$ to represent a vector consists of $i-1$th imaginary part of $\boldsymbol{x} \in \mathbb{F}_M^N$. In other words, $\boldsymbol{x}^{(i)}$ is defined by

$$\boldsymbol{x}^{(i)} = \begin{pmatrix} x_{1i} & x_{2i} & \cdots & x_{Ni} \end{pmatrix}^{\mathrm{T}}. \tag{2.1}$$

In addition, the complex conjugate $\mathbf{x}_i^*$ of $\mathbf{x}_i \in \mathbb{F}_M$ is defined by

$$\mathbf{x}_i^* = \begin{pmatrix} x_{i1} & -x_{i2} & -x_{i3} & \cdots & -x_{iM} \end{pmatrix} \in \mathbb{F}_M.$$

In the same way, $\boldsymbol{x}^* \in \mathbb{F}_M^N$ is defined by

$$\boldsymbol{x}^* = \begin{pmatrix} \mathbf{x_1}^* & \mathbf{x_2}^* & \cdots & \mathbf{x_N}^* \end{pmatrix}^{\mathrm{T}} \in \mathbb{F}_M^N.$$

Secondly, we define some rules of multiplication of hypercomplex numbers. Let

$$\begin{aligned} \mathbf{x} &= \begin{pmatrix} x_1 & \cdots & x_M \end{pmatrix} \in \mathbb{F}_M \\ \mathbf{y} &= \begin{pmatrix} y_1 & \cdots & y_M \end{pmatrix} \in \mathbb{F}_M. \end{aligned}$$

The multiplication of hypercomplex numbers $\mathbf{x}$ and $\mathbf{y}$ in this thesis requires

$$\mathbf{xy} = x_1 y_1 + \mu(\mathbf{x}, \mathbf{y}) \in \mathbb{F}_M$$

and

$$\mathbf{x}^*\mathbf{x} = \sum_{i=1}^{M} x_i x_i \in \mathbb{R},$$

where, $\mu(\mathbf{x}, \mathbf{y})$ is the term depends on the definition of multiplication of the imaginary units. We also require

$$\mathbf{x}^*\mathbf{y} + \mathbf{y}^*\mathbf{x} \in \mathbb{R}.$$

Note that the real numbers, the ordinary complex numbers, and the quaternions satisfy the above properties. In addition, they are written as $\mathbb{F}_1 = \mathbb{R}$, $\mathbb{F}_2 = \mathbb{C}$, and $\mathbb{F}_4$, respectively.

Additionally, we use notation

$$< \mathbf{x}, \mathbf{y} >= \sum_{i=1}^{M} x_i y_i$$

for $\mathbf{x}, \mathbf{y} \in \mathbb{F}_M$. We also use notation like $\mathbf{x} = 1$, $x_i = 3$, $x_k = e^{\boldsymbol{j}t}$, $\mathbf{x} = \cos(\pi t) + \boldsymbol{j}\sin(\pi t)$ for cases such as $\mathbf{x} \in \mathbb{R}$ and $x_i \in \mathbb{C}$. Moreover, we introduce norms for $\mathbf{x} \in \mathbb{F}_M$ and $\boldsymbol{x} \in \mathbb{F}_M^N$, namely,

$$| \mathbf{x} |_\epsilon = \sqrt{x_1^2 + \frac{1}{\epsilon^2} \sum_{k=2}^{M} x_k^2} \tag{2.2}$$

$$\|\boldsymbol{x}\| = \sqrt{\sum_{i=2}^{N} | \mathbf{x}_i |_1^2}, \tag{2.3}$$

where ϵ is a positive constant. We will omit ϵ if $\epsilon = 1$ hereafter.

Finally, note that there are relations $\mathbb{R} \subset \mathbb{C} \subset \mathbb{F}_M (M \geq 2)$ under a consistent definition of multiplication. In other words, elements of $\mathbb{R}$, $\mathbb{C}$, $\mathbb{F}_3, \cdots$ are represented as

$$\begin{array}{ll} \begin{pmatrix} x_1 & 0 & 0 & 0 & \cdots & 0 \end{pmatrix} & \in \mathbb{R} \subset \mathbb{C}, \\ \begin{pmatrix} x_1 & x_2 & 0 & 0 & \cdots & 0 \end{pmatrix} & \in \mathbb{C} \subset \mathbb{F}_3, \\ \begin{pmatrix} x_1 & x_2 & x_3 & 0 & \cdots & 0 \end{pmatrix} & \in \mathbb{F}_3 \subset \mathbb{F}_4 \\ \vdots & \end{array} \tag{2.4}$$

2.2.2 Formulation of Nonconvex Relaxation Problem

Relaxation problems for **QP** to the space consists of hypercomplex numbers are defined as $\mathbf{QP}^{\epsilon}_{M}$ using the definition of hypercomplex numbers in the previous subsection.

($\mathbf{QP}^{\epsilon}_{M}$)	Min. :	$E(\boldsymbol{x}) = \boldsymbol{x}^{*\mathrm{T}}\boldsymbol{W}\boldsymbol{x}$
	Subj.to :	$\boldsymbol{x} \in \mathcal{D}^{\epsilon}_{qM}$

, where $\mathcal{D}^{\epsilon}_{qM}$is defined by

$$\mathcal{D}^{\epsilon}_{qM} = \left\{ \boldsymbol{x} \in \mathbb{F}^{N}_{M} \;\middle|\; \begin{array}{l} \boldsymbol{x}^{(1)\mathrm{T}}\bar{\boldsymbol{A}}_i\boldsymbol{x}^{(1)} + \dfrac{1}{\epsilon^2}\displaystyle\sum_{j=2}^{M}\boldsymbol{x}^{(j)\mathrm{T}}\bar{\boldsymbol{A}}_i\boldsymbol{x}^{(j)} \leq 0 \;\; \text{for} \;\; i = 1, \cdots, L \\ \mid \mathrm{x}_i \mid_{\epsilon} = 1 \;\; \text{for} \;\; i \in \mathcal{K} \\ \mathrm{x}_1 = 1 \end{array} \right\}.$$

The class $\mathbf{QP}^{\epsilon}_{M}$ is relaxation of **QP** since it coincides with **QP** if we impose constraints

$$\boldsymbol{x}^{(i)} = \boldsymbol{0} \quad i = 2, \cdots, M.$$

We can check this fact immediately from the definition of hypercomplex numbers. As well as this, the relation $\mathcal{D}^{\epsilon}_{qA} \subset \mathcal{D}^{\epsilon}_{qB}$ holds if $A \leq B$ from eq.(2.4). For this reason, M is considered to be representing the degree of relaxation of $\mathbf{QP}^{\epsilon}_{M}$.

In case the original problem is restricted to $\mathbf{P}_1$, then $\mathbf{QP}^{\epsilon}_{M}$ is represented simply as $\mathbf{P}^{\epsilon}_{M}$

($\mathbf{P}^{\epsilon}_{M}$)	Min. :	$E(\boldsymbol{x}) = \boldsymbol{x}^{*\mathrm{T}}\boldsymbol{W}\boldsymbol{x}$
	Subj. to :	$\boldsymbol{x} \in \mathcal{D}^{\epsilon}_{M} = \left\{ \boldsymbol{x} \in \mathbb{F}^{N}_{M} \;\middle|\; \begin{array}{l} \mid \mathrm{x}_i \mid_{\epsilon} \leq 1 \;\; \text{for} \;\; i = 1, \cdots, N \\ \mathrm{x}_1 = 1 \end{array} \right\}$

The class $\mathbf{P}^{\epsilon}_{M}$ is often used not only as examples for explaining properties of $\mathbf{QP}^{\epsilon}_{M}$, but also for investigating detailed properties in chapter 4.

The class $\mathbf{QP}^{\epsilon}_{M}$ is not convex in general. Hence, we cannot obtain a global optimum of $\mathbf{QP}^{\epsilon}_{M}$ in general as long we are using a simple decent type search algorithm. Therefore, strategies that requires a global optimum of the relaxation problem are not applicable immediately. For example, we cannot exploit $\mathbf{QP}^{\epsilon}_{M}$ in calculating a lower bound of the optimal value of **QP** in branch and bound method immediately.

Nevertheless, the formulation $\mathbf{QP}^{\epsilon}_{M}$ has other advantages. The crucial properties of this formulation of relaxation problems are listed as follows.

- The degree of relaxation is parametrized by M and ϵ.
- $\mathbf{QP}^{\epsilon}_{M}$ is not convex in general and includes the original problem as a special case.

These facts allow to consider relationship between the degree of relaxation and properties embedded by relaxing the feasible region. In other words, we will be able to take into account the trade-off between the precise description of the requirements from the real world applications and tractability in numerical computation. We consider that analyses according to this viewpoint will result in :

- maximization of the performance of known optimization algorithms by adjusting the degree of relaxation represented by M and ϵ
- constructing optimization algorithms which require some desirable properties of problems ; the properties are not found in the original problem

Our approach will gain significance especially for situations in which convex relaxation problems are too far from the original problem to be utilized in solving the original problems.

Chapter 3

Properties of Nonconvex Relaxation Problems

In this chapter, we describe fundamental properties of the relaxation problem $\mathbf{QP}^{\epsilon}_{M}$ defined in the previous chapter. First, we see some properties derived directly from the definition of $\mathbf{QP}^{\epsilon}_{M}$. The relaxation problem $\mathbf{QP}^{\epsilon}_{M}$ connects the original problem and the convex relaxation problem in term of semidefinite programming(SDP) continuously with parameter ϵ, and divide problems between the original problem and SDP relaxation problem into some grades derived by the parameter M. Next, we define a "monotone path" which plays a crucial role in this thesis. It is a continuous curve in the feasible region $\mathcal{D}^{\epsilon}_{qM}$. If we change the decision variables along the curve, the value of the objective function changes monotonically. The existence of a monotone path is considered to be one of the most fundamental properties in constructing optimization algorithms. We show conditions about M and ϵ for the existence of a monotone path. If $M \geq 2$ and $\epsilon = 1$, then a monotone path exists for any problem instance. We also show some conditions for local optimality of feasible solutions of the relaxation problem of which $M = 2$ and $\epsilon = 1$.

3.1 Relationship among the Original Problem, the Nonconvex Relaxation Problem and the Convex Relaxation Problem

As we have seen in the previous chapter, the feasible region of the relaxation problem to the space consists of higher dimensional hypercomplex numbers includes that of lower dimensional hypercomplex numbers for the same ϵ. Moreover, we have the following proposition in case $\mathcal{D}_{q1}$ is convex.

Proposition 3.1 *Let all $\boldsymbol{A}_i$s be positive semidefinite and let $\mathcal{K} = \emptyset$. Then $\mathcal{D}^{\epsilon}_{qM}$ is convex for any ϵ and M.*

Proof : Let $\boldsymbol{z}$ be a real vector whose entries are defined by

$$\boldsymbol{z} = \begin{pmatrix} x_{12} & x_{13} & \cdots & x_{1M} \ , & x_{22} & x_{23} & \cdots & x_{2M} \ , & x_{32} & x_{33} & \cdots & \cdots & \cdots & x_{NM} \end{pmatrix}^{\mathrm{T}}$$

for simplicity of notation. Then each quadratic constraint of $\mathbf{QP}^{\epsilon}_{M}$ about $\bar{\boldsymbol{A}}_i$ is written as

$$\boldsymbol{a}_i(\boldsymbol{z}) = \boldsymbol{z}^{\mathrm{T}}\boldsymbol{B}_i\boldsymbol{z} + \boldsymbol{f}_i^{\mathrm{T}}\boldsymbol{z} + \alpha_i \leq 0 \tag{3.1}$$

by substituting the constraint $\mathbf{x}_1 = 1$, where $\boldsymbol{B}_i \in \mathcal{S}_{M\times(N-1)}$ is a positive semidefinite block diagonal matrix whose diagonal entries are $\boldsymbol{A}_i$ or $\frac{1}{\epsilon^2}\boldsymbol{A}_i$. Therefore, $\boldsymbol{a}_i(\boldsymbol{z})$ is a convex function. Hence, its level set is convex. Consequently, their intersection $\mathcal{D}^{\epsilon}_{qM}$ is convex. ■

Moreover, the feasible region for large ϵ includes that for small ϵ if $\boldsymbol{A}_i$s are all positive semidefinite and $\mathcal{K} = \emptyset$. Hence, the following corollary is derived immediately from eq.(3.1).

Corollary 3.1 *Let* $0 < \gamma < \epsilon$ *and* $\mathcal{K} = \emptyset$, *and let all* $\boldsymbol{A}_i$*s be positive semidefinite. Then* $\mathcal{D}^{\gamma}_{qM} \subset \mathcal{D}^{\epsilon}_{qM}$ *holds.*

In this sense, the parameter M and ϵ represent the degree of relaxation. In addition, the degree is understood as inclusion of the feasible regions.

Next, we consider the case $M = N$ and $\epsilon = 1$ together with the SDP relaxation problem $\mathbf{QP_{sdp}}$.

($\mathbf{QP_{sdp}}$)	Min.	:	trace $(\boldsymbol{XW})$
	Subj. to	:	$\boldsymbol{X} \in \mathcal{D}_{sdp}$

The feasible region $\mathcal{D}_{sdp}$ is defined by

$$\mathcal{D}_{sdp} = \left\{ \boldsymbol{X} \in \mathcal{S}_N \;\middle|\; \begin{array}{l} \text{trace}\left(\boldsymbol{X}\bar{\boldsymbol{A}}_i\right) \leq 0 \;\; \text{for} \;\; i = 1, \cdots, L \\ X_{ii} = 1 \;\; \text{for} \;\; i \in \mathcal{K} \\ \boldsymbol{X} \succeq \boldsymbol{O} \\ X_{11} = 1 \end{array} \right\},$$

where X_{ij} denotes i-jth entry of the real symmetry matrix $\boldsymbol{X}$.

The constraint $\boldsymbol{X} \succeq \boldsymbol{O}$ yields that there are $\mathbf{x}_i$s satisfying

$$X_{ij} =< \mathbf{x}_i, \mathbf{x}_j >$$

for any i and j. Therefore the value of the objective function is coincident with $E(\boldsymbol{x})$ of $\mathbf{QP}_N$. Namely,

$$\begin{aligned} \text{trace}\,(\boldsymbol{XW}) &= \sum_{i=1}^{N}\sum_{j=1}^{N} w_{ij}X_{ij} = \sum_{i=1}^{N}\sum_{j=1}^{N} w_{ij} < \mathbf{x}_i, \mathbf{x}_j > = \sum_{i=1}^{N}\sum_{j=1}^{N} w_{ij} \sum_{k=1}^{N} x_{ik}x_{jk} \\ &= \sum_{k=1}^{N}\left(\sum_{i=1}^{N}\sum_{j=1}^{N} w_{ij}x_{ik}x_{jk}\right) = \sum_{k=1}^{N} \boldsymbol{x}^{(k)\mathrm{T}}\boldsymbol{W}\boldsymbol{x}^{(k)} \\ &= \boldsymbol{x}^{*\mathrm{T}}\boldsymbol{W}\boldsymbol{x} = E(\boldsymbol{x}) \end{aligned}$$

holds for $\epsilon = 1$ and $M = N$. In addition, we obtain

$$\text{trace}\left(\boldsymbol{X}\bar{\boldsymbol{A}}_i\right) = \sum_{j=1}^{M} \boldsymbol{x}^{(j)\mathrm{T}} \bar{\boldsymbol{A}}_i \boldsymbol{x}^{(j)} \tag{3.2}$$

in the same manner.

The class $\mathbf{QP_{sdp}}$ is a class of semidefinite programming problems which can be solved in polynomial time. Moreover, the transformation from positive semidefinite $\boldsymbol{X} \in \mathcal{S}_N$ to $\boldsymbol{x} \in \mathbb{F}_N^N$ also can be done in polynomial time. In this sense, $\mathbf{QP}_M^\epsilon$ with $M = N$ and $\epsilon = 1$ is equivalent to the convex relaxation problem $\mathbf{QP_{sdp}}$.

On the other hand, $\mathbf{QP}_M^\epsilon$ becomes $\mathbf{QP}$ when $M = 1$ as we have seen in the previous chapter. In addition, if $\boldsymbol{A}_i$s are all positive semidefinite and $\mathcal{D}_{qM}^\epsilon$ is bounded, then $\mathbf{QP}_M^\epsilon$ converges to $\mathbf{QP}$ as $\epsilon \to +0$ in the sense that $\boldsymbol{x}^{(j)} (j \neq 1)$s in the right hand side of eq.(3.2) must converge to $\mathbf{0}$ to keep $\boldsymbol{x}$ feasible. More precisely, we have the following proposition.

Proposition 3.2 *Let all $\boldsymbol{A}_j$s be positive semidefinite, $\mathcal{D}_{qM}^\epsilon$ be bounded for any $\epsilon > 0$, and $M \in \mathbb{N}$. In addition, let $\delta > 0$ be an arbitrary positive number. Then there exists an $\epsilon > 0$ such that*

$$\boldsymbol{x} \in \mathcal{D}_{qM}^\epsilon \textit{ implies } \|\boldsymbol{x}^{(i)}\| < \delta \textit{ for } i = 2, \cdots, M.$$

Proof : Let define

$$\boldsymbol{x}'^{(i)} = \begin{pmatrix} x_{2i} & x_{3i} & \cdots & x_{Ni} \end{pmatrix}^{\mathrm{T}} \in \mathbb{R}^{N-1} \quad (i = 1, \cdots, M).$$

Let $\lambda_M(\boldsymbol{A}_j)$ and $\lambda_m(\boldsymbol{A}_j)$ be the maximal and the minimal eigenvalues of $\boldsymbol{A}_j$, respectively. Then, there are positive constants K_{ij}s such that

$$\boldsymbol{x}'^{(i)\mathrm{T}} \boldsymbol{A}_j \boldsymbol{x}'^{(i)} < \epsilon^2 K_{ij} < \infty \quad (i = 2, \cdots, M)$$

for any $\boldsymbol{x} \in \mathcal{D}_{qM}^\epsilon$ and $j \in \{1, \cdots, L\}$.

When $\boldsymbol{x}'^{(i)}$ is not included in the eigenspace of $\lambda_m(\boldsymbol{A}_j)$ or all $\boldsymbol{A}_j$s are positive definite, there is a constant $\lambda^{i,j}$ such that $0 < \lambda^{i,j} \leq \lambda_M(\boldsymbol{A}_j)$ and

$$\boldsymbol{x}'^{(i)\mathrm{T}} \boldsymbol{A}_j \boldsymbol{x}'^{(i)} = \|\boldsymbol{x}'^{(i)}\|^2 \lambda^{i,j} = \|\boldsymbol{x}^{(i)}\|^2 \lambda^{i,j} < \epsilon^2 K_{ij} \quad (i = 2, \cdots, M).$$

Hence, there is an ϵ such that

$$\|\boldsymbol{x}^{(i)}\|^2 < \epsilon^2 K = \delta^2 \quad (i = 2, \cdots, M). \tag{3.3}$$

for any $\delta > 0$, where K is defined by

$$K = \max_{i,j(i \neq 1)} \frac{K_{ij}}{\lambda^{i,j}}.$$

In case $\boldsymbol{x}'^{(i)}$ is included in the eigenspace of $\lambda_m(\boldsymbol{A}_j)$ and $\boldsymbol{A}_j$ is not positive definite, similar discussions to derive eq.(3.3) must hold for at least one other constraint. Otherwise,

$$\boldsymbol{x}'^{(i)\mathrm{T}}\boldsymbol{A}_j\boldsymbol{x}'^{(i)} = (k\boldsymbol{x}'^{(i)\mathrm{T}})\boldsymbol{A}_j(k\boldsymbol{x}'^{(i)})$$

or $| k\mathbf{x}_i |_\epsilon = 1$ must hold for any $k \in \mathbb{R}$ and $j \in \{2, \cdots, L\}$ since the linear terms with respect to the imaginary parts vanish. This contradicts to the boundedness of $\mathcal{D}^\epsilon_{qM}$. ■

Note that proposition 3.2 does not require $\mathcal{K} = \emptyset$. This leads us to the conclusion that the formulation $\mathbf{QP}^\epsilon_M$ connects the original problem and the convex relaxation problem continuously by the parameter ϵ, and divides relaxation problems between two extreme case into some grades defined by M. The relationship among $\mathbf{QP}$, $\mathbf{QP}^\epsilon_M$, and $\mathbf{QP}_{\mathbf{sdp}}$ is illustrated in Fig.3.1.

At the end of this section, we mention to two extreme cases of problem instances. There are at least one problem instance of $\mathbf{QP}$ such that

- One of the nonconvex relaxation problem, which is not the original problem $\mathbf{QP}$, has exponentially many strict local optima.
- One of the nonconvex relaxation problem has infinitely many local optima even if the number of local optima of the original problem is finite.

The concrete examples of the extreme cases are left for the next chapter since the proof requires descriptions there.

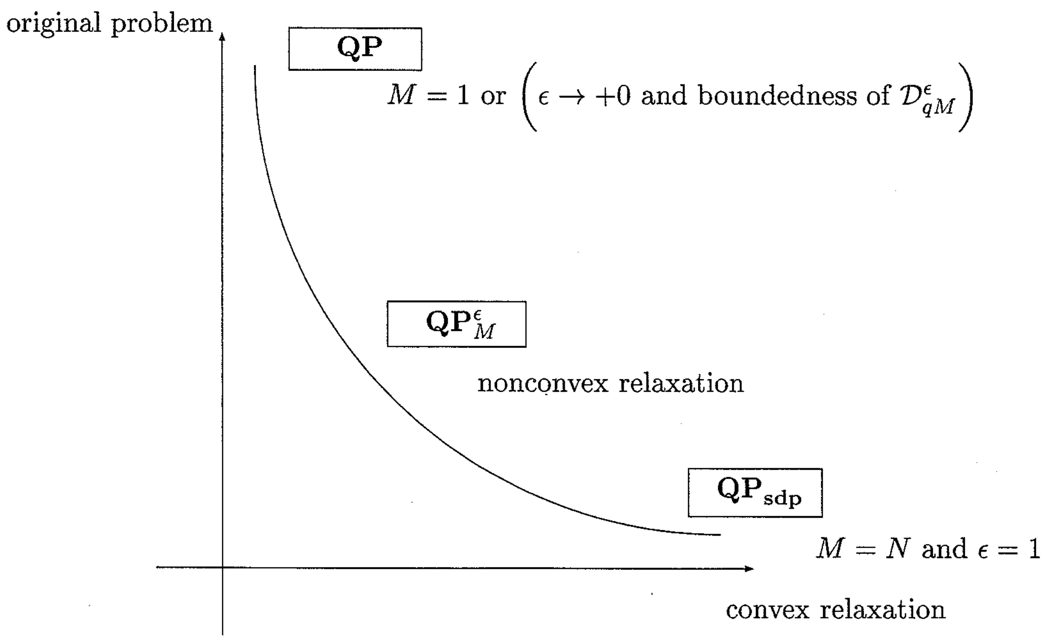

Figure 3.1: relationship among $\mathbf{QP}$, $\mathbf{QP}^\epsilon_M$, and $\mathbf{QP}_{\mathbf{sdp}}$

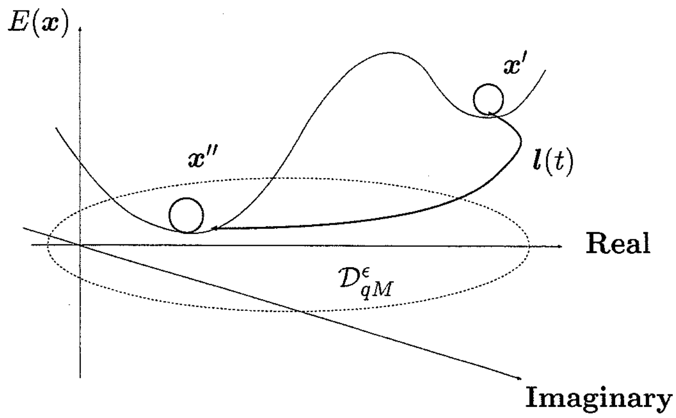

Figure 3.2: monotone path $l(t)$

3.2 Monotone Path

In this section, we define a "monotone path" which plays a crucial role in the rest of this thesis and analyze its properties.

We are interested in whether the decision variable x can move continuously to the global optimum of the original problem without increasing the value of the objective function $E(x)$ from an arbitrary feasible solution of the original problem **QP**. If it is impossible, then no decent type local search algorithm can find a global optimum of the original problem for some initial solutions and problem data. Immediately from the discussions in previous section, it is possible if $\mathbf{QP}^{\epsilon}_{M}$ is equivalent to the convex optimization problem $\mathbf{QP_{sdp}}$. On the contrary, there are many problem instances for which it is not possible to move if $\mathbf{QP}^{\epsilon}_{M}$ is equivalent to the original problem **QP**.

First, we define a monotone path as follows.

Definition 3.1 *Let $x, y \in \mathcal{D}_{q1}$ and $E(x) > E(y)$. Consider a continuous map $l : [0,1] \to \mathcal{D}^{\epsilon}_{qM}$ such that*

$$l(0) = x$$
$$l(1) = y$$
$$1 \geq s > t \geq 0 \;\Leftrightarrow\; E(l(s)) < E(l(t)).$$

We call

$$\left\{ l(t) \middle| t \in [0,1] \right\} \subset \mathcal{D}^{\epsilon}_{qM}$$

a monotone path between x and y.

If a monotone path exists for any $x \in \mathcal{D}_{q1}$ and $y \in \mathcal{D}_{q1}(E(y) \neq E(x))$, and if it is always easy to compute the monotone path, then obtaining a global optimum of **QP** seems to be easy.

In other words, we can find a global optimum of **QP** by starting from an arbitrary feasible solution of the original problem and selecting the direction of movement appropriately.

To analyze around the existence of a monotone path, let $\boldsymbol{x}', \boldsymbol{x}'' \in \mathcal{D}_{q1}$, $E(\boldsymbol{x}') \neq E(\boldsymbol{x}'')$, and

$$\boldsymbol{c} = \frac{\boldsymbol{x}' + \boldsymbol{x}''}{2} = \left(\frac{x_1' + x_1''}{2}, \frac{x_2' + x_2''}{2}, \dots, \frac{x_N' + x_N''}{2}\right)^{\mathrm{T}}$$

$$\boldsymbol{r} = \frac{\boldsymbol{x}' - \boldsymbol{x}''}{2} = \left(\frac{x_1' - x_1''}{2}, \frac{x_2' - x_2''}{2}, \dots, \frac{x_N' - x_N''}{2}\right)^{\mathrm{T}}.$$

In addition, define

$$\boldsymbol{l}(t) = \boldsymbol{c} + \boldsymbol{r}(\cos(\pi t) + \boldsymbol{j}\sin(\pi t)) \quad , \quad 0 \leq t \leq 1. \tag{3.4}$$

From eq.(3.4), we can show that there are real constants α and β such that

$$\boldsymbol{l}(t)^{*\mathrm{T}} \boldsymbol{K} \boldsymbol{l}(t) = \alpha + \beta \cos(\pi t) \tag{3.5}$$

for any $\boldsymbol{K} \in \mathcal{S}_N$. Immediately, we can say that $E(\boldsymbol{x})$ is monotone on $\boldsymbol{l}(t)$. Therefore, $\boldsymbol{l}(t)$ is a monotone path between $\boldsymbol{x}'$ and $\boldsymbol{x}''$ if it is contained in the feasible region. As a consequence, we have the following proposition.

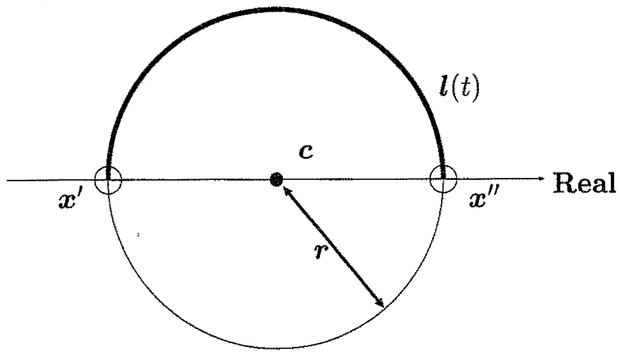

Figure 3.3: the monotone path defined by eq.(3.4)

Proposition 3.3 *Let $\boldsymbol{l}(t)$ be defined by eq.(3.4). Let $\epsilon = 1$, $M \geq 2$ and $\boldsymbol{x}', \boldsymbol{x}'' \in \mathcal{D}_{q1}$. Then*

$$\boldsymbol{l}(t) \in \mathcal{D}_{qM}^{\epsilon} \ , \ 0 \leq t \leq 1$$

holds. Additionally, if $\mathcal{D}_{q1}$ is convex, then $\epsilon \geq 1$ and $M \geq 2$ implies $\boldsymbol{l}(t) \in \mathcal{D}_{qM}^{\epsilon}$.

Proof : The left hand side of the quadratic inequality constraints are monotone on $\boldsymbol{l}(t)$ from eq.(3.5). In addition, the constraint $\mid \mathrm{x}_i \mid_{\epsilon} = 1$ holds for any t. Hence, $\epsilon = 1$ and $M \geq 2$ implies $\boldsymbol{l}(t) \in \mathcal{D}_{qM}^{\epsilon}$. From similar manner and from corollary 3.1, the relation $\boldsymbol{l}(t) \in \mathcal{D}_{qM}^{\epsilon}$ holds if $\boldsymbol{x}', \boldsymbol{x}'' \in \mathcal{D}_{q1}$ and $\epsilon \geq 1$. ■

Here, note that proposition 3.3 does not care about the problem data except for convexity of the feasible region $\mathcal{D}_{q1}$.

Proposition 3.3 is a sufficient condition about M and ϵ. Next, we show a necessary condition.

Proposition 3.4 *For any $N(N \geq 4)$, $\epsilon(0 < \epsilon < 1)$ and $M(M \geq 1)$, there exist $\bar{\boldsymbol{A}}_i$s, $\boldsymbol{W} \in \mathcal{S}_N$, and $\mathcal{K}$ such that there is no monotone path between certain two feasible solutions of* **QP** *having distinct objective values(even if $\boldsymbol{A}_i \succeq \boldsymbol{O}$).*

Proof : We will show an example for which no monotone path exists. Consider a matrix $\boldsymbol{W} \in \mathcal{S}_N^0$ having entries

$$\begin{aligned} & w_{12} = w_{21} = w_{34} = w_{43} = -1 \\ & w_{13} = w_{31} = w_{14} = w_{41} = w_{23} = w_{32} = w_{24} = w_{42} = \delta, \end{aligned} \tag{3.6}$$

and other off-diagonal entries $w_{ij} = \kappa$. In other words, define $\boldsymbol{W}$ by

$$\boldsymbol{W} = \begin{pmatrix} 0 & -1 & \delta & \delta & \kappa & \kappa & \cdots \\ -1 & 0 & \delta & \delta & \kappa & \kappa & \cdots \\ \delta & \delta & 0 & -1 & \kappa & \kappa & \cdots \\ \delta & \delta & -1 & 0 & \kappa & \kappa & \cdots \\ \kappa & \kappa & \kappa & \kappa & 0 & \kappa & \cdots \\ \kappa & \kappa & \kappa & \kappa & \kappa & 0 & \cdots \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \ddots \end{pmatrix},$$

where δ and κ are constants satisfying

$$0 < -\delta < \frac{1}{3} \text{ and } 0 < -\kappa < \frac{-\delta}{N^2}. \tag{3.7}$$

Let $\bar{\boldsymbol{A}}_i = \boldsymbol{O}$ and $\mathcal{K} = \{1, \cdots, N\}$ (or let $c_i = -1$, $\boldsymbol{b}_i = \boldsymbol{0}$, $\boldsymbol{A}_i = E_{ii}$ and $\mathcal{K} = \emptyset$).

For the above problem instance, the objective function is represented as

$$\begin{aligned} E(\boldsymbol{x}) &= -2 < \mathbf{x}_1, \mathbf{x}_2 > -2 < \mathbf{x}_3, \mathbf{x}_4 > \\ & \quad 2\delta(< \mathbf{x}_1, \mathbf{x}_3 > + < \mathbf{x}_1, \mathbf{x}_4 > + < \mathbf{x}_2, \mathbf{x}_3 > + < \mathbf{x}_2, \mathbf{x}_4 >) + \alpha(\boldsymbol{x}), \end{aligned}$$

where $| \alpha(\boldsymbol{x}) | < -\delta$ holds for any $\boldsymbol{x} \in \mathcal{D}_{qM}^{\epsilon}$ from definition of δ and κ. On the other hand, it is easy to show that

$$\boldsymbol{x}^1 = (\underbrace{1,1,1,1}_{4}, \underbrace{1,\cdots,1}_{N-4})^{\mathrm{T}} \text{ and } \boldsymbol{x}^2 = (\underbrace{1,1,-1,-1}_{4}, \underbrace{1,\cdots,1}_{N-4})^{\mathrm{T}}$$

are strict local optima of **QP**. The value of the objective function at $\boldsymbol{x}^1$ and $\boldsymbol{x}^2$ are represented as

$$\begin{aligned} E(\boldsymbol{x}^1) &= -4 + 8\delta + \alpha(\boldsymbol{x}^1) \\ E(\boldsymbol{x}^2) &= -4 - 8\delta + \alpha(\boldsymbol{x}^2). \end{aligned}$$

Moreover, $\boldsymbol{x}^1$ is the unique global optimum of the original problem.

From the constraint $\mathbf{x}_1 = 1$, we have $x_{1j} = 0 (j \neq 1)$. Therefore, the objective function is represented as

$$E(\boldsymbol{x}) = -2x_{21} - 2 < \mathbf{x}_3, \mathbf{x}_4 > +2\delta(x_{31} + x_{41} + < \mathbf{x}_2, \mathbf{x}_3 > + < \mathbf{x}_2, \mathbf{x}_4 >) + \alpha(\boldsymbol{x}).$$

Additionally, there exists a feasible solution $\bar{\boldsymbol{x}}$ satisfying $x_{31} = 0$ between $\boldsymbol{x}^1$ and $\boldsymbol{x}^2$ since $x_{31} = 1$ at $\boldsymbol{x}^1$ and $x_{31} = -1$ at $\boldsymbol{x}^2$. The objective function at $\bar{\boldsymbol{x}}$ is written as

$$\begin{aligned} E(\bar{\boldsymbol{x}}) &= -2\bar{x}_{21} - 2\left(\sum_{i=2}^{M} \bar{x}_{3i}\bar{x}_{4i}\right) + 2\delta\bar{x}_{41} + 2\delta\left(\sum_{i=2}^{M} \bar{x}_{2i}\bar{x}_{3i}\right) + 2\delta < \bar{\mathbf{x}}_2, \bar{\mathbf{x}}_4 > +\alpha(\bar{\boldsymbol{x}}) \\ &\geq -2 - 2\epsilon^2 + 2\delta + 2\delta\epsilon^2 + 2\delta + \alpha(\bar{\boldsymbol{x}}). \end{aligned} \tag{3.8}$$

Consequently, we have the following relation.

$$\begin{aligned} \beta &= E(\bar{\boldsymbol{x}}) - E(\boldsymbol{x}^2) \\ &\geq \{-2 - 2\epsilon^2 + 2\delta + 2\delta\epsilon^2 + 2\delta + \alpha(\bar{\boldsymbol{x}})\} - \{-4 + 8\delta + \alpha(\boldsymbol{x}^2)\} \\ &\geq 2 - 2\epsilon^2 + \delta(14 + 2\epsilon^2) \end{aligned} \tag{3.9}$$

For this reason, $\beta > 0$ implies

$$E(\bar{\boldsymbol{x}}) > E(\boldsymbol{x}^2) > E(\boldsymbol{x}^1) \tag{3.10}$$

and eq.(3.10) means that that $E(\boldsymbol{x})$ must increase at $\bar{\boldsymbol{x}}$ when $\boldsymbol{x}$ moves from $\boldsymbol{x}^1$ to $\boldsymbol{x}^2$.

If we define δ so as to satisfy

$$\delta > \frac{\epsilon^2 - 1}{7 + \epsilon^2}, \tag{3.11}$$

then eq.(3.10) holds.

As a conclusion, there is no monotone path between $\boldsymbol{x}^1$ and $\boldsymbol{x}^2$ if δ satisfies eq.(3.11). ■

From propositions 3.3 and 3.4, the relaxation problem which is the nearest to the original problem **QP** among relaxation problems that ensure the existence of a monotone path is $\mathbf{QP}_2$, that is, a relaxation problem in which $M = 2$ and $\epsilon = 1$. Here, note that $\mathbf{QP}_2$ is not convex. Therefore, convexity is not necessary for ensuring the existence of a monotone path.

3.3 Optimality Conditions

In this section, we derive conditions for local optimality of feasible solutions in the relaxation problem $\mathbf{QP}_2$. As we have seen in the previous section, the relaxation problem in which $\epsilon = 1$ and $M = 2$ is the critical case of our nonconvex relaxation problems in the criterion for the existence of a monotone path.

The conditions stated here utilize the fact that the feasible region is expanded towards the directions of imaginary axes.

Proposition 3.5 *If $\boldsymbol{x} \in \mathcal{D}_{q1}$ is a local optimum of* $\mathbf{QP}_2$*, then $\boldsymbol{x}$ is also a global optimum of the original problem* $\mathbf{QP}$.

Proof : Let $\boldsymbol{x} \in \mathcal{D}_{q1}$ be a local optimum of $\mathbf{QP}_2$ and let $\boldsymbol{y} \in \mathcal{D}_{q1}$ be any other feasible solution of $\mathbf{QP}$. Then we can define $\boldsymbol{l}(t)$ by eq.(3.4) such that $\boldsymbol{l}(0) = \boldsymbol{x}$ and $\boldsymbol{l}(1) = \boldsymbol{y}$.

The assumption that $\boldsymbol{x}$ is a local optimum of $\mathbf{QP}_2$ yields

$$\left.\frac{\mathrm{d}\boldsymbol{l}(t)}{\mathrm{d}t}\right|_{t=0} = 0 \quad \text{and} \quad \left.\frac{\mathrm{d}^2\boldsymbol{l}(t)}{\mathrm{d}t^2}\right|_{t=0} \geq 0.$$

In addition, $E(\boldsymbol{y}) \geq E(\boldsymbol{x})$ holds since $E(\boldsymbol{l}(t))$ is monotone in $t \in [0,1]$. In other words, the feasible solution $\boldsymbol{x}$ is a global optimum of $\mathbf{QP}$. ■

Corollary 3.2 *If* $\mathbf{QP}_2$ *has only one local optimum, then the local optimum is a global optimum of* $\mathbf{QP}$.

Proof : For any $\boldsymbol{A} \in \mathcal{S}_N$,

$$\boldsymbol{x}^{*\mathrm{T}}\boldsymbol{A}\boldsymbol{x} = \boldsymbol{x}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{x}^*$$

holds. Therefore, if the local optimum is unique, then $\boldsymbol{x}^* = \boldsymbol{x}$ must hold for the local optimum $\boldsymbol{x}$. In other words, it must be in $\mathbb{R}^N$. The rest of the proof comes immediately from proposition 3.5. ■

For some special instances, for example $\mathbf{P}_1$ with $\boldsymbol{W}$ without positive entries, we can make certain that the global optimum of $\mathbf{QP}$ is also a local optimum of $\mathbf{P}_2$. Consider an instance of $\mathbf{P}_2$ where all off-diagonal entries of $\boldsymbol{W} \in \mathcal{S}_{N0}$ are negative. Then

$$\bar{\boldsymbol{x}} = \begin{pmatrix} 1 & 1 & \cdots & 1 \end{pmatrix}^{\mathrm{T}}$$

is the unique global optimum and satisfies $w_{ij}x_ix_j = - \mid w_{ij} \mid$ for any i and j. Therefore, if $\|\boldsymbol{x}\| < N$, then $E(\boldsymbol{x})$ is greater than at $\bar{\boldsymbol{x}}$. On the other hand, if $x_i \notin \mathbb{R}$ for some i, then $\Re(w_{1i}x_1x_i) > - \mid w_{1i} \mid$ holds. Hence, each such instances of $\mathbf{P}_2$ has a local optimum which is also the unique global optimum of $\mathbf{P}_1$.

Proposition 3.5 can be understood as a necessary condition for local optimality of feasible solutions in $\mathcal{D}_{q1}(\subset \mathcal{D}_{q2})$. On the other hand, the following proposition is a necessary condition for the case $\boldsymbol{x} \notin \mathcal{D}_{q1}$.

Proposition 3.6 *Let $\boldsymbol{W}$ be defined by*

$$\boldsymbol{W} = \begin{pmatrix} \boldsymbol{W}_{11} & \boldsymbol{W}_{12} \\ \boldsymbol{W}_{12}^{\mathrm{T}} & \boldsymbol{W}_{22} \end{pmatrix} = \boldsymbol{W}^{\mathrm{T}},$$

where $\boldsymbol{W}_{11} \in \mathcal{S}_K$. Let $\boldsymbol{x} = (\boldsymbol{x}_1, \boldsymbol{x}_2)^{\mathrm{T}} \in \mathcal{D}_{q2}$ and $\boldsymbol{x}_1 \in \mathbb{C}^K$. In addition, assume that

$$\boldsymbol{y} = (\boldsymbol{x}_1, e^{\jmath\phi}\boldsymbol{x}_2)^{\mathrm{T}} \in \mathcal{D}_{q2} \tag{3.12}$$

for any real number ϕ satisfying that $\mid \phi \mid$ is sufficiently small. Then $\boldsymbol{x}$ is a local optimum of $\mathbf{QP}_2$ *only if*

$$A(\boldsymbol{x}) = \boldsymbol{x}_1^{*\mathrm{T}} \boldsymbol{W}_{12} \boldsymbol{x}_2 \leq 0 \quad \textit{for any} \quad K \in \{1, \cdots, N\}.$$

Here, note that $A(\boldsymbol{x})$ is not a real number in general.

Proof : The value of the objective function is written as

$$E(\boldsymbol{x}) = E_{11}(\boldsymbol{x}) + E_{12}(\boldsymbol{x}) + E_{22}(\boldsymbol{x}),$$

where $E_{11}(\boldsymbol{x}) = \boldsymbol{x}_1^{*\mathrm{T}} \boldsymbol{W}_{11} \boldsymbol{x}_1$, $E_{12}(\boldsymbol{x}) = 2\Re\left(\boldsymbol{x}_1^{*\mathrm{T}} \boldsymbol{W}_{12} \boldsymbol{x}_2\right)$, and $E_{22}(\boldsymbol{x}) = \boldsymbol{x}_2^{*\mathrm{T}} \boldsymbol{W}_{22} \boldsymbol{x}_2$.

Consider a feasible solution

$$\boldsymbol{y} = \begin{pmatrix} \boldsymbol{x}_1 & e^{j\phi} \boldsymbol{x}_2 \end{pmatrix}^{\mathrm{T}}.$$

Then $E_{11}(\boldsymbol{x}) = E_{11}(\boldsymbol{y})$ and $E_{22}(\boldsymbol{x}) = E_{22}(\boldsymbol{y})$ holds. The term $E_{12}(\boldsymbol{y})$ is represented as

$$\begin{aligned} E_{12}(\boldsymbol{y}) &= \boldsymbol{x}_1^{*\mathrm{T}} \boldsymbol{W}_{12} \boldsymbol{x}_2 e^{\boldsymbol{j}\phi} + \boldsymbol{x}_2^{*\mathrm{T}} \boldsymbol{W}_{12} \boldsymbol{x}_1 e^{-\boldsymbol{j}\phi} = Ce^{\boldsymbol{j}(\theta+\phi)} + Ce^{-\boldsymbol{j}(\theta+\phi)} \\ &= 2C\cos(\theta+\phi), \end{aligned}$$

where $A(\boldsymbol{x}) = \boldsymbol{x}_1^{*\mathrm{T}} \boldsymbol{W}_{12} \boldsymbol{x}_2 = Ce^{\boldsymbol{j}\theta}$ and $C \in \mathbb{R}$.

If $A(\boldsymbol{x}) \notin \mathbb{R}$, then $\theta \neq 0(\mathrm{mod}\pi)$ holds. In addition, this implies

$$\left.\frac{\mathrm{d}E_{12}(\boldsymbol{y})}{\mathrm{d}\phi}\right|_{\phi=0} \neq 0.$$

Therefore, there is a feasible solution $\boldsymbol{y}$ such that $E(\boldsymbol{x}) > E(\boldsymbol{y})$. Hence, $\boldsymbol{x}$ is not a local optimum.

In case $A(\boldsymbol{x}) \in \mathbb{R}$, the inequality $A(\boldsymbol{x}) > 0$ implies

$$\left.\frac{\mathrm{d}E_{12}(\boldsymbol{y})}{\mathrm{d}\phi}\right|_{\phi=0} = 0$$

and

$$\left.\frac{\mathrm{d}^2E_{12}(\boldsymbol{y})}{\mathrm{d}\phi^2}\right|_{\phi=0} < 0.$$

Hence, such an $\boldsymbol{x}$ is not a local optimum from the same reason.

Consequently, $\boldsymbol{x}$ is not a local optimum of $\mathbf{P}_2$ if $A(\boldsymbol{x}) \leq 0$ does not hold. ■

Note that sorting the order of indices of variables preserves proposition 3.6 for $\mathbf{P}_2$ since the assumption in eq.(3.12) is automatically satisfied. Therefore, we have the following corollary of proposition 3.6.

Corollary 3.3 *Let $\mathcal{T} \subset \mathcal{S}_N$ be the set of matrices having only one 1 in each column and each row, and having entry 0 elsewhere. In addition, let $\boldsymbol{W}'$ be defined by $\boldsymbol{T}^{-\mathrm{T}}\boldsymbol{W}\boldsymbol{T}^{-1}$ and $\boldsymbol{x}'$ be defined by $\boldsymbol{T}\boldsymbol{x}$, where $\boldsymbol{T} \in \mathcal{T}$. Moreover, we represent $\boldsymbol{W}'$ as*

$$\boldsymbol{W}' = \begin{pmatrix} \boldsymbol{W}'_{11} & \boldsymbol{W}'_{12} \\ \boldsymbol{W}'^{\mathrm{T}}_{12} & \boldsymbol{W}'_{22} \end{pmatrix} = \boldsymbol{W}'^{\mathrm{T}},$$

where $\boldsymbol{W}'_{11} \in \mathcal{S}_K$. Finally, let $\boldsymbol{x}' = (\boldsymbol{x}'_1, \boldsymbol{x}'_2)^{\mathrm{T}} \in \mathcal{D}_2$ and $\boldsymbol{x}'_1 \in \mathbb{C}^K$.

Then $\boldsymbol{x}$ is a local optimum of $\mathbf{P}_2$ only if

$$\boldsymbol{x}_1'^{*\mathrm{T}} \boldsymbol{W}'_{12} \boldsymbol{x}'_2 \leq 0 \quad \textit{for any} \quad K \in \{1, \cdots, N\} \quad \textit{and} \quad \boldsymbol{T} \in \mathcal{T}$$

Proposition 3.6 is not a sufficient condition for local optimality. Consider an instance of $\mathbf{P}_1$ and one of its local optima $\bar{\boldsymbol{x}}$ other than a global optimum. Then it is easy to show an example $\bar{\boldsymbol{x}}$ which satisfies the condition in proposition 3.6 and does not satisfy the condition in proposition 3.5. An example is represented as

$$\boldsymbol{W} = \begin{pmatrix} 0 & -10 & 1 & 1 \\ -10 & 0 & 1 & 1 \\ 1 & 1 & 0 & -10 \\ 1 & 1 & -10 & 0 \end{pmatrix}, \quad \bar{\boldsymbol{x}} = \begin{pmatrix} 1 & 1 & 1 & 1 \end{pmatrix}^{\mathrm{T}}.$$

On the other hand, corollary 3.3 requires conditions more than the first order KKT condition [38, 40] for $\mathbf{P}_2$. The first order KKT condition is equivalent to the presence of $\lambda_i \geq 0$ satisfying

$$2\boldsymbol{v} := 2\begin{pmatrix} \boldsymbol{W} & \boldsymbol{O} \\ \boldsymbol{O} & \boldsymbol{W} \end{pmatrix}\boldsymbol{z} + \lambda_1 \begin{pmatrix} 2\Re(x_1) \\ 0 \\ \vdots \\ 0 \\ 2\Im(x_1) \\ 0 \\ \vdots \\ 0 \end{pmatrix} + \lambda_2 \begin{pmatrix} 0 \\ 2\Re(x_2) \\ \vdots \\ 0 \\ 0 \\ 2\Im(x_2) \\ \vdots \\ 0 \end{pmatrix} + \cdots + \lambda_N \begin{pmatrix} 0 \\ \vdots \\ 0 \\ 2\Re(x_N) \\ 0 \\ \vdots \\ 0 \\ 2\Im(x_N) \end{pmatrix} = \mathbf{0},$$

where

$$\boldsymbol{z} = \begin{pmatrix} \boldsymbol{x}^{(1)} & \boldsymbol{x}^{(2)} \end{pmatrix}^{\mathrm{T}} \in \mathbb{R}^{2N}.$$

In fact, this condition is derived from the case of $K = 1$ in corollary 3.3. Let

$$u_i = \sum_{j=1}^{N} w_{ij} x_j.$$

Then $\Im(u_i) = v_{i+N}$ and $\Re(u_i) = v_i$ from definition of $\boldsymbol{v}$. Moreover, the case $K = 1$ of corollary 3.3 is written as

$$\angle u_i + \angle x_i^* = \angle u_i - \angle x_i = \pi.$$

Finally, this condition implies the presence of nonnegative λ_is.

Chapter 4

Relaxation Problem for 0-1 Quadratic Optimization

In this chapter we restrict our original problems to 0-1 quadratic optimization. For this class of original problems, we can derive more strong properties. One of the results enables us to generalize feasible region $\{-1,+1\}^N$ to discrete subset of the space of hypercomplex numbers without affecting the optimal value of the problem. As a consequence, this property leads us to generalizing some optimization algorithms which are designed directly for the original problem or exploit the convex relaxation. We demonstrate the effect theoretically through Goemans and Williamson's approximation algorithm for the maximum cut problems (The demonstration through experiments appears in chapter 5). In this chapter, we also show that for some classes of problem instances, any local optima of the relaxation problem $\mathbf{P}_2$ give a lower bound of the optimal value of the original problem in spite the relaxation problem is not convex.

4.1 Projection from Higher Dimensional Space to Lower Dimensional Space

The example of a monotone path defined by eq.(3.4) implies that if we know a feasible solution $\boldsymbol{x}$ of $\mathbf{P}_M$, which satisfies some conditions, then we can obtain $\bar{\boldsymbol{x}} \in \mathcal{D}_1$ satisfying $E(\bar{\boldsymbol{x}}) \leq E(\boldsymbol{x})$, immediately. Here, we show such a procedure.

Let define a set $\mathcal{D}_{Md} \subset \mathbb{F}_M$ by

$$\mathcal{D}_{Md} = \left\{ \mathbf{x} \in \mathbb{F}_M \;\middle|\; \mathbf{x} \in \{-1,0,+1\}^M \quad , \quad |\, \mathbf{x} \,| = 1 \right\}. \tag{4.1}$$

For example,

$$\mathcal{D}_{1d} = \{-1,+1\}$$

if $M = 1$, and

$$\mathcal{D}_{2d} = \{-1,+1,-\boldsymbol{j},+\boldsymbol{j}\}.$$

if $M = 2$.

For the case $M \neq 1$, let α and β be arbitrary natural numbers satisfying

$$1 \leq \alpha \leq M \ , \ 2 \leq \beta \leq M \quad \text{and} \quad \alpha \neq \beta.$$

Then we can define $\boldsymbol{x}', \boldsymbol{x}'' \in \mathcal{D}_{Md}^N$ in such a way that

$$\mathbf{x}_i' = (x_{i1}', \cdots, x_{iM}') \ , \ \mathbf{x}_i'' = (x_{i1}'', \cdots, x_{iM}'') \quad , \quad i \in \{1, \cdots, N\} \tag{4.2}$$

$$\boldsymbol{x}' = (\mathbf{x}_1', \cdots, \mathbf{x}_N')^{\mathrm{T}} \ , \ \boldsymbol{x}'' = (\mathbf{x}_1'', \cdots, \mathbf{x}_N'')^{\mathrm{T}} \tag{4.3}$$

$$x_{i\beta}' = x_{i\beta}'' = 0 \quad , \quad i \in \{1, \cdots, N\}$$

$$x_{i\alpha}' = \begin{cases} +1 & \text{if} \quad x_{i\beta} = -1 \\ -1 & \text{if} \quad x_{i\beta} = +1 \\ x_{i\alpha} & \text{if} \quad x_{i\beta} = 0 \end{cases} \quad , \quad i \in \{1, \cdots, N\}$$

$$x_{i\alpha}'' = \begin{cases} +1 & \text{if} \quad x_{i\beta} = +1 \\ -1 & \text{if} \quad x_{i\beta} = -1 \\ x_{i\alpha} & \text{if} \quad x_{i\beta} = 0 \end{cases} \quad , \quad i \in \{1, \cdots, N\}$$

$$x_{ij}' = x_{ij}'' = x_{ij} \quad , \quad \begin{cases} i = 1, \cdots, N \\ j \in \{1, \cdots, N\} \setminus \{\alpha, \beta\} \end{cases} \tag{4.4}$$

by using $\boldsymbol{x} \in \mathcal{D}_{Md}^N$. Then the following proposition holds.

Proposition 4.1 *Let $\boldsymbol{x}'$ and $\boldsymbol{x}''$ be defined from eq.(4.2) through eq.(4.4). Then*

$$E(\boldsymbol{x}) = \frac{1}{2}\Big[E(\boldsymbol{x}') + E(\boldsymbol{x}'')\Big] \tag{4.5}$$

holds for any α, β, and $\boldsymbol{x} \in \mathcal{D}_{Md}^N$.

Proof : Let identify

$$\boldsymbol{x}^{(\alpha)} = \begin{pmatrix} x_{1\alpha} & x_{2\alpha} & \cdots & x_{N\alpha} \end{pmatrix}^{\mathrm{T}}$$

as the real part $\boldsymbol{x}^{(1)}$ of $\boldsymbol{x}$ without loss of generality. By exchanging the indices of imaginary part, we can define $\boldsymbol{x}'$ and $\boldsymbol{x}''$ as the start point $\boldsymbol{l}(0)$ and the end point $\boldsymbol{l}(1)$ of the monotone path defined as $\boldsymbol{l}(t)$ in eq.(3.4) for any $\boldsymbol{x} \in \mathcal{D}_{Md}^N$. Then $\boldsymbol{x}$ coincides with the point $\boldsymbol{l}(1/2)$ on the monotone path. Therefore,

$$E(\boldsymbol{l}(0)) - E(\boldsymbol{l}(1/2)) = E(\boldsymbol{l}(1/2)) - E(\boldsymbol{l}(1)) \tag{4.6}$$

holds since $E(\boldsymbol{l}(t))$ is represented as eq.(3.4). Finally, eq.(4.6) is identical to eq.(4.5). ∎

Next, we construct a procedure that exploits proposition 4.1. Let $\mathbf{Prj}(\cdot) : \mathcal{D}_{Md}^N \to \mathcal{D}_{Md}^N$ be defined by

$$\mathbf{Prj}(\boldsymbol{x}) = \begin{cases} \boldsymbol{x}' & \text{if} \quad E(\boldsymbol{x}') \leq E(\boldsymbol{x}'') \\ \boldsymbol{x}'' & \text{else} \end{cases}, \tag{4.7}$$

and let $\mathbf{Shift}(\cdot) : \mathcal{D}_{Md}^N \to \mathcal{D}_{Md}^N$ be

$$\mathbf{Shift}(\boldsymbol{x}) = \left(\mathbf{Prj}(\boldsymbol{x})^{(1)}, \cdots, \mathbf{Prj}(\boldsymbol{x})^{(\beta-1)}, \mathbf{Prj}(\boldsymbol{x})^{(\beta+1)}, \cdots, \mathbf{Prj}(\boldsymbol{x})^{(M)}, \mathbf{Prj}(\boldsymbol{x})^{(\beta)}\right) \tag{4.8}$$

by using notation in eq.(2.1). Then $\mathbf{Shift}(\boldsymbol{x}) \in \mathcal{D}_{Md}^N$ can be represented as a hypercomplex number whose dimension is less than M since each entry of $\mathbf{Prj}(\boldsymbol{x})^{(\beta)}$ is zero.

For these reasons, by using given $\boldsymbol{x} \in \mathcal{D}_{Md}^N$, we can find a feasible solution $\mathbf{Shift}(\boldsymbol{x})$ where the value of the objective function is no more than at $E(\boldsymbol{x})$ and the minimal dimension of hypercomplex numbers necessary to represent it is less than M. That is,

- $E(\boldsymbol{x}) \geq E(\mathbf{Shift}(\boldsymbol{x}))$
- $\boldsymbol{x} \in \mathcal{D}_{Md}^N \Rightarrow \mathbf{Shift}(\boldsymbol{x}) \in \mathcal{D}_{M-1,d}^N$.

Moreover, we obtain a feasible solution of $\mathbf{P}_1$ by using $\mathbf{Shift}(\cdot)$ $M-1$ times. Furthermore, the definition of $\mathbf{Shift}(\cdot)$ implies that many optimization algorithms or procedures will be generalized from $\boldsymbol{x} \in \mathcal{D}_{1d}^N$ to $\boldsymbol{x} \in \mathcal{D}_{Md}^N$. We actually generalize an optimization algorithm in chapter 5.

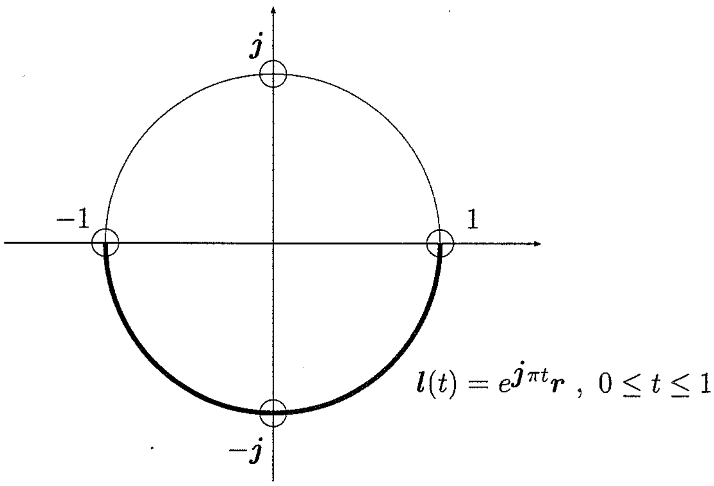

Figure 4.1: $\mathcal{D}_{2d}$ and monotone path $\boldsymbol{l}(t)$

Here note that procedures described from eq.(4.2) through eq.(4.8) can be done in $O(N^2M)$, that is, in polynomial time. This observation allows us to generalize a polynomial time approximation algorithm in the following section.

4.2 Improvement of Goemans and Williamson's Approximation Algorithm for the Maximum Cut Problems

In this section, we consider a polynomial time approximation algorithm for the maximum cut problems. Goemans and Williamson's approximation algorithm [16] is one of the approximation algorithms for the maximum cut problems. The approximation rate is still the best known one for the maximum cut problems though some generalizations for general classes of original problems are reported [17, 18, 19].

Here, we generalize Goemans and Williamson's algorithm to try to increase the approximation rate by using results in the previous section.

Note that we use "maximize" instead of "minimize" only in this section.

4.2.1 Goemans and Williamson's Approximation Algorithm

We describe the maximum cut problems as $\mathbf{MC_1}$.

($\mathbf{MC_1}$) Maximize $F(\boldsymbol{x}) = \frac{1}{4}\boldsymbol{x}^{\mathrm{T}}\boldsymbol{L}\boldsymbol{x}$, Subject to $\boldsymbol{x} \in \{-1, +1\}^N$

In $\mathbf{MC_1}$, $\boldsymbol{L}$ is a given real symmetry and positive semidefinite matrix. Its entries are defined in such a way that

- $\boldsymbol{W} = (w_{ij}) \in \mathcal{S}_N^0$: a given matrix whose each off-diagonal entry is nonnegative.
- If $i \neq j$, then i-jth entry of $\boldsymbol{L}$ is $-w_{ij}$.
- The i-ith entry of $\boldsymbol{L}$ equals to $\sum_{j=1}^{N} w_{ij}$.

If $\boldsymbol{W}$ is the adjacency matrix of a graph, then $\mathbf{MC_1}$ coincides with the original definition of the maximum cut problems, "*Find a subset V of vertices of a given graph which maximize the sum of weights of edges whose one endpoint is in V and the another is not*".

The approximation rate α of algorithm $(\mathbf{A})$ for a maximization problem is defined by

$$\alpha = \frac{(p - \underline{p})}{(\bar{p} - \underline{p})},$$

where p is the expected value of the objective function at a solution generated by $(\mathbf{A})$. The values $\bar{p}$ and $\underline{p}$ are supremum and infimum of the value of the objective function in the feasible region, respectively. For $\mathbf{MC_1}$, it is easy to show $\underline{p} = 0$ for any $\boldsymbol{L}$. We assume $\bar{p} \neq 0$; this is equivalent to $\|\boldsymbol{W}\| \neq 0$.

Goemans and Williamson's approximation algorithm achieves the rate $\alpha > 0.87856$ for any instances of $\mathbf{MC_1}$. In this algorithm, a relaxation problem $\mathbf{MC_{sdp}}$ is used as a subroutine.

$$
(\mathbf{MC_{sdp}}) \quad \begin{array}{ll} \text{Maximize} & \frac{1}{4}\text{trace}\,(\boldsymbol{X}\boldsymbol{L}) \\ \text{Subject to} & \begin{cases} \boldsymbol{X} = \boldsymbol{X}^{\mathrm{T}} \succeq \boldsymbol{O} \\ X_{ii} = 1 \quad \text{for} \quad i = 1, \cdots, N \end{cases} \end{array}
$$

The approximation algorithm is described as follows.

Goemans and Williamson's Approximation Algorithm

1. Let $\boldsymbol{X}$ be an optimal solution of $\mathbf{MC_{sdp}}$.
2. Generate N vectors $\boldsymbol{v}_1, \cdots, \boldsymbol{v}_N$ by

   $$\boldsymbol{X} = \boldsymbol{V}\boldsymbol{V}^{\mathrm{T}}$$

   $$\boldsymbol{V} = \begin{pmatrix} \boldsymbol{v}_1 & \cdots & \boldsymbol{v}_N \end{pmatrix}.$$

   The constraint $X_{ii} = 1$ guarantees that each $\boldsymbol{v}_i$ is on the origin centered unit sphere S.
3. Generate a vector $\boldsymbol{h}$ by uniform distribution on S.
4. Define feasible solution x_i of $\mathbf{MC_1}$ by a rule such that

   $$x_i = \begin{cases} 1 & \text{if} \quad \boldsymbol{v}_i^{\mathrm{T}}\boldsymbol{h} \geq 0 \\ -1 & \text{if} \quad \boldsymbol{v}_i^{\mathrm{T}}\boldsymbol{h} < 0 \end{cases}. \tag{4.9}$$

Then the expectation of the value of the objective function is given by

$$
\begin{aligned}
\mathrm{E}(F(\boldsymbol{x})) &= \mathrm{E}\left[\frac{1}{4}\sum_{i=1}^{N}\sum_{j=1}^{N} w_{ij}(1 - x_i x_j)\right] = \frac{1}{2}\sum_{i=1}^{N}\sum_{j=1}^{N} w_{ij}\mathrm{E}\left[\frac{1 - x_i x_j}{2}\right] \\
&= \frac{1}{2}\sum_{i=1}^{N}\sum_{j=1}^{N} w_{ij}\Pr\left[\mathrm{sign}(\boldsymbol{v}_i \cdot \boldsymbol{h}) \neq \mathrm{sign}(\boldsymbol{v}_j \cdot \boldsymbol{h})\right] \\
&= \frac{1}{2}\sum_{i=1}^{N}\sum_{j=1}^{N} w_{ij}\frac{\arccos(\boldsymbol{v}_i \cdot \boldsymbol{v}_j)}{\pi}.
\end{aligned}
$$

This is also written as

$$\mathrm{E}(F(\boldsymbol{x})) = \frac{1}{4}\sum_{i=1}^{N}\sum_{j=1}^{N} w_{ij} - \frac{1}{4}\sum_{i=1}^{N}\sum_{j=1}^{N} w_{ij}\left(\frac{\pi - 2\theta_{ij}}{\pi}\right),$$

where θ_{ij} is the angle between $\boldsymbol{v}_i$ and $\boldsymbol{v}_j$, that is, $\arccos(\boldsymbol{v}_i \cdot \boldsymbol{v}_j)$.

Note that the problem of maximizing $\mathrm{E}(F(\boldsymbol{x}))$ is a generalization of $\mathbf{MC_1}$ and is also a relaxation problem. In other words,

(i) If $(\boldsymbol{v}_1, \cdots, \boldsymbol{v}_N)$ is a optimum of $\mathbf{MC_1}$, then $\mathrm{E}(F(\boldsymbol{x}))$ is coincident with the optimal value $\bar{p}$.

(ii) The value $\mathrm{E}(F(\boldsymbol{x}))$ must be no more than $\bar{p}$ from the definition.

By the way, from the fact that

$$\frac{\arccos(y)}{\pi} \geq \frac{1}{2}\alpha(1-y) \;,\; \alpha > 0.8756$$

at $y \in [-1,1]$(see [16]), we have

$$\mathrm{E}(F(\boldsymbol{x})) \geq \frac{1}{4}\alpha \sum_{i=1}^{N}\sum_{j=1}^{N} w_{ij}(1 - \boldsymbol{v}_i \cdot \boldsymbol{v}_j) \tag{4.10}$$

in case $\boldsymbol{v}_i$s are generated by Cholesky factorization of the solution of $\mathbf{MC_{sdp}}$. Additionally, the problem of maximizing the right hand side of eq.(4.10) is equivalent to $\mathbf{MC_{sdp}}$. Therefore,

$$\mathrm{E}(F(\boldsymbol{x})) \geq \alpha \times (\text{ optimal value of } \mathbf{MC_{sdp}}) \geq \alpha\bar{p}$$

holds.

4.2.2 Improvement via Nonconvex Relaxation Approach

Let $\boldsymbol{v}_i$s be vectors on S. Let $\boldsymbol{h}_1$ and $\boldsymbol{h}_2$ be vectors generated by independent uniform distribution on S. Then define x_is by

$$x_i = \begin{cases} 1 & \text{if} \quad (\boldsymbol{v}_i^{\mathrm{T}}\boldsymbol{h}_1 \geq 0) \wedge (\boldsymbol{v}_i^{\mathrm{T}}\boldsymbol{h}_2 \geq 0) \\ \boldsymbol{j} & \text{if} \quad (\boldsymbol{v}_i^{\mathrm{T}}\boldsymbol{h}_1 \geq 0) \wedge (\boldsymbol{v}_i^{\mathrm{T}}\boldsymbol{h}_2 < 0) \\ -\boldsymbol{j} & \text{if} \quad (\boldsymbol{v}_i^{\mathrm{T}}\boldsymbol{h}_1 < 0) \wedge (\boldsymbol{v}_i^{\mathrm{T}}\boldsymbol{h}_2 \geq 0) \\ -1 & \text{if} \quad (\boldsymbol{v}_i^{\mathrm{T}}\boldsymbol{h}_1 < 0) \wedge (\boldsymbol{v}_i^{\mathrm{T}}\boldsymbol{h}_2 < 0) \end{cases} . \tag{4.11}$$

Here, the second and third cases in eq.(4.11) are caused by $\boldsymbol{h}_2$ which does not appear in the original algorithm.

Consider a problem $\mathbf{MC_2}$ together with eq.(4.11).

($\mathbf{MC_2}$) maximize $F(\boldsymbol{x}) = \frac{1}{4}\boldsymbol{x}^{*\mathrm{T}}\boldsymbol{L}\boldsymbol{x}$, s.t. $\boldsymbol{x} \in \{-1,+1,\boldsymbol{j},-\boldsymbol{j}\}^N = \mathcal{D}_{2d}^N$

Based on the above definition, we have the following proposition. This means that we obtain an algorithm whose approximation rate equals to Goemans and Williamson's one.

Proposition 4.2 *Let $\boldsymbol{v}_i$s in eq.(4.11) be generated from the solution of* $\mathbf{MC_{sdp}}$. *Let x_is be generated by eq.(4.11). Then the expectation of $F(\boldsymbol{x})$ is represented as*

$$\begin{aligned} \mathrm{E}(F(\boldsymbol{x})) &= \frac{1}{4}\sum_{i=1}^{N}\sum_{j=1}^{N} w_{ij} - \mathrm{E}\left(\frac{1}{4}\sum_{i=1}^{N}\sum_{j=1}^{N} w_{ij}x_i^* x_j\right) \\ &= \frac{1}{4}\sum_{i=1}^{N}\sum_{j=1}^{N} w_{ij} - \frac{1}{4}\sum_{i=1}^{N}\sum_{j=1}^{N} w_{ij}\left(\frac{\pi - 2\theta_{ij}}{\pi}\right). \end{aligned} \tag{4.12}$$

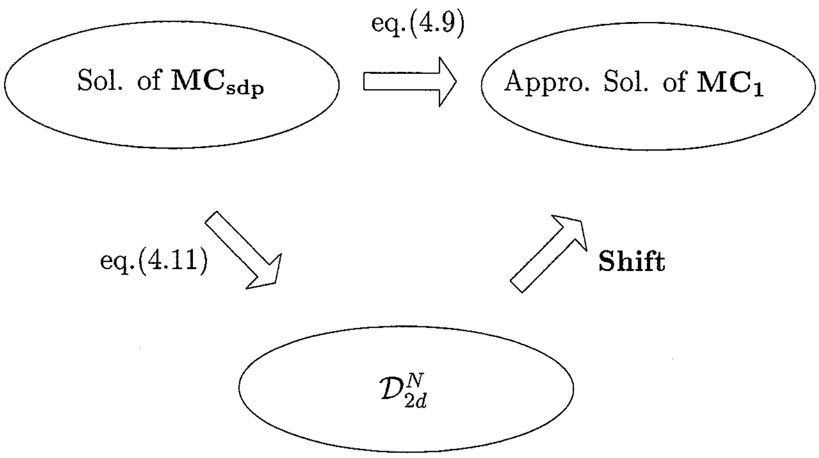

Figure 4.2: Goemans and Williamson's approximation algorithm

Note that eq.(4.12) is coincident with the expectation of Goemans and Williamson's approximation rate.

Proof : First, we have

$$\Pr\left[\mathrm{sign}(\boldsymbol{v}_i \cdot \boldsymbol{h}_k) = \mathrm{sign}(\boldsymbol{v}_j \cdot \boldsymbol{h}_k)\right] = \frac{\pi - \theta_{ij}}{\pi} \tag{4.13}$$

$$\Pr\left[\mathrm{sign}(\boldsymbol{v}_i \cdot \boldsymbol{h}_k) \neq \mathrm{sign}(\boldsymbol{v}_j \cdot \boldsymbol{h}_k)\right] = \frac{\theta_{ij}}{\pi} \tag{4.14}$$

for any i, j and k(see [16] for detail). Next, for $k = 1, 2$, we have

$$\Pr\left[\mathrm{sign}(\boldsymbol{v}_i \cdot \boldsymbol{h}_k) = 1\right] = \Pr\left[\mathrm{sign}(\boldsymbol{v}_i \cdot \boldsymbol{h}_k) = -1\right] = 0.5$$

from independency of $\boldsymbol{v}_i$ and $\boldsymbol{h}_k$. Hence, we have

$$\begin{aligned} \Pr\left[x_i^* x_j = 1\right] = \Pr\left[x_i = x_j\right] &= \left(\frac{\pi - \theta_{ij}}{\pi}\right)^2 \\ \Pr\left[x_i^* x_j = -1\right] = \Pr\left[x_i = -x_j\right] &= \left(\frac{\theta_{ij}}{\pi}\right)^2 \\ \Pr\left[x_i^* x_j = \boldsymbol{j}\right] = \Pr\left[x_i^* x_j = -\boldsymbol{j}\right] &= \left(\frac{\theta_{ij}}{\pi}\right)\left(\frac{\pi - \theta_{ij}}{\pi}\right) \end{aligned} \tag{4.15}$$

from eq.(4.11) and eq.(4.13). Therefore,

$$\begin{aligned} \mathrm{E}\left[\Re(w_{ij} x_i^* x_j)\right] &= w_{ij}\left\{1 \cdot \left(\frac{\pi - \theta_{ij}}{\pi}\right)^2 + (-1) \cdot \left(\frac{\theta_{ij}}{\pi}\right)^2\right\} \\ &= \frac{w_{ij}}{\pi^2}\left(\pi^2 - 2\pi\theta_{ij}\right) = \frac{w_{ij}}{\pi}\left(\pi - 2\theta_{ij}\right). \end{aligned}$$

holds. Finally, independency of summation of expectation yields eq.(4.12). ■

The proposition 4.2 means that changing the target of projection from $\mathcal{D}_{1d}^N$ to $\mathcal{D}_{2d}^N$ does not affect the expected value of the objective function. Until now, the projected solution is not a feasible solution of $\mathbf{MC_1}$. However, we can obtain a feasible solution of $\mathbf{MC_1}$ by using $\mathbf{Shift}(\cdot)$ and the value of objective function does not become worse as we have seen in proposition 4.1. In addition, the algorithm after generalization is also a polynomial time approximation algorithm since $\mathbf{Shift}(\cdot)$ can be done in polynomial time.

Proposition 4.3 yields the change of the expectation of the objective value caused by $\mathbf{Shift}(\cdot)$.

Proposition 4.3 *Let the inequality in eq.(4.7) be "$\geq$", here(since we are formulating the maximum cut problems as a maximization problem). Then* $F(\mathbf{Shift}(\boldsymbol{x})) - F(\boldsymbol{x})$ *is represented as*

$$F(\mathbf{Shift}(\boldsymbol{x})) - F(\boldsymbol{x}) = \frac{1}{2}\left|\Im\left(\sum_{i=1}^{N}\sum_{j=i+1}^{N} w_{ij}x_i^* x_j\right)\right| \tag{4.16}$$

Proof : Let $\boldsymbol{x}$ be given by eq.(4.11) and $\boldsymbol{L}$ be represented as $\boldsymbol{W}$ in proposition 3.6. In addition, let $\boldsymbol{x}_1 \in \{-1, +1\}^K$ and $\boldsymbol{x}_2 \in \{-\boldsymbol{j}, +\boldsymbol{j}\}^{N-K}$ without loss of generality. Then the objective function is represented as

$$F(\boldsymbol{x}) = \frac{1}{4}\left(\boldsymbol{x}_1^{*\mathrm{T}}\boldsymbol{L}_{11}\boldsymbol{x}_1 + \boldsymbol{x}_2^{*\mathrm{T}}\boldsymbol{L}_{22}\boldsymbol{x}_2 + \boldsymbol{x}_1^{*\mathrm{T}}\boldsymbol{L}_{12}\boldsymbol{x}_2 + \boldsymbol{x}_2^{*\mathrm{T}}\boldsymbol{L}_{12}^{\mathrm{T}}\boldsymbol{x}_1\right).$$

The increase of the objective function by $\mathbf{Shift}(\cdot)$ is represented as

$$\frac{1}{2}\left|\boldsymbol{x}_1^{*\mathrm{T}}\boldsymbol{L}_{12}\boldsymbol{x}_2\right|$$

since the operation $\mathbf{Prj}(\cdot)$ multiplies $\boldsymbol{j}$ or $-\boldsymbol{j}$ to $\boldsymbol{x}_2$. Hence, we have eq.(4.16). ■

Immediately from the result that eq.(4.16) is nonnegative, we have $\mathrm{E}(F(\mathbf{Shift}(\boldsymbol{x}))) \geq \alpha\bar{p}$.

In this thesis, we do not try to check whether

$$\mathrm{E}\left[F(\mathbf{Shift}(\boldsymbol{x}))\right] > \frac{1}{4}\sum_{i=1}^{N}\sum_{j=1}^{N} w_{ij} - \frac{1}{4}\sum_{i=1}^{N}\sum_{j=1}^{N} w_{ij}\frac{\pi - 2\theta_{ij}}{\pi}. \tag{4.17}$$

In other words, we do not check whether the approximation rate is a constant strictly greater than $\alpha = 0.878\cdots$ since any analysis with respect only to the effect of $\mathbf{Shift}(\cdot)$ does not work for checking eq.(4.17).

The value $\mathrm{E}\left[F(\mathbf{Shift}(\boldsymbol{x})) - F(\boldsymbol{x})\right]$ depends on $\boldsymbol{v}_i$s because we use eq.(4.11). For some $\boldsymbol{v}_i$s,

$$\mathrm{E}\left[F(\mathbf{Shift}(\boldsymbol{x})) - F(\boldsymbol{x})\right] = 0 \tag{4.18}$$

holds. The optimum of $\mathbf{MC_1}$ is an instance of such cases.

Note that eq.(4.18) does not mean

$$\min_{\boldsymbol{W}} \mathrm{E}\left[F(\boldsymbol{x})\right] = \min_{\boldsymbol{W}} \mathrm{E}\left[F(\mathbf{Shift}(\boldsymbol{x}))\right]$$

since $\mathrm{E}\left[F(\boldsymbol{x})\right]$ may be always larger than $\alpha\bar{p}$ whenever $F(\mathbf{Shift}(\boldsymbol{x})) - F(\boldsymbol{x})$ is small.

Note also that Goemans and Williamson's approximation algorithm is the best known polynomial time approximation algorithm in the sense of the presence of a theoretical guarantee about both of time complexity and approximation rate. On the contrary, it does not match to many heuristics [37] in numerical experiments. Hence, evaluating eq.(4.18) in numerical computation does not seem to be attractive. In our computational experiments, the modified version of Goemans and Williamson's algorithm did not match to heuristics at all ; the original one did not match, too.

On the other hand, of course, we can make M large in projection of eq.(4.11), and it may yield an algorithm having better performance in practice. However, we do not have theoretical guarantee except for $M = 1$ and $M = 2$ yet.

4.3 Lower Bounds and Local Optima

In this section, we show that any local optimum of $\mathbf{P}_2$ gives a lower bound of the optimal value of $\mathbf{P}_1$ for certain classes of $\boldsymbol{W}$ in spite $\mathbf{P}_2$ is not convex.

For general classes of $\boldsymbol{W}$, we do not know whether any local optima of $\mathbf{P}_2$ give a lower bound. However, if $\mathbf{P}_2$ always gives a lower bound, then we can expect of $\mathbf{P}_2$ to become a powerful tool for calculating a lower bound since

- obtaining a local optimum is often easy even if the problem is not convex.
- the lower bound given at a local optimum is never worse than that given at a global optimum. In addition, the lower bound by the global optimum of $\mathbf{P}_2$ is never worse than that of the convex relaxation problem $\mathbf{P}_N$.

4.3.1 Preliminaries

We describe lemmas used in the forthcoming discussions.

Let $\boldsymbol{u} := \boldsymbol{W}\boldsymbol{x}$, where each diagonal entry of $\boldsymbol{W} \in \mathcal{S}_N$ is zero. Define $\mathbf{P}_1^+$ and its relaxation problem $\mathbf{P}_2^+$ by

$$(\ \mathbf{P}_1^+\ )\qquad \text{Min. } \boldsymbol{x}^{\mathrm{T}}\boldsymbol{W}\boldsymbol{x}\ \ ,\ \ \text{s.t. } \boldsymbol{x} \in \mathcal{D}_1^+ = \left\{\boldsymbol{x} \in \mathbb{R}^N \;\middle|\; x_i \in [-R_i, +R_i] \text{ and } x_1 = R_1\ \right\}$$

$$(\ \mathbf{P}_2^+\ )\qquad \text{Min. } \boldsymbol{x}^{*\mathrm{T}}\boldsymbol{W}\boldsymbol{x}\ \ ,\ \ \text{s.t. } \boldsymbol{x} \in \mathcal{D}_2^+ = \left\{\boldsymbol{x} \in \mathbb{C}^N \;\middle|\; |\, x_i \,| \le R_i \text{ and } x_1 = R_1\ \right\}$$

, where R_is are positive constants. In addition, define $\mathrm{sign}_i(\cdot) : \mathbb{C}\backslash 0 \to R_i e^{j\theta}$ by

$$\mathrm{sign}_i(u) = \frac{R_i u}{|\, u\,|}.$$

In case $R_i = 1$, we will omit suffix i. For $\boldsymbol{u} = (\ u_1, \cdots, u_N\ )^{\mathrm{T}} \in \{\mathbb{C}\backslash 0\}^N$, $\mathbf{sign}(\boldsymbol{u})$ is defined by

$$\mathbf{sign}(\boldsymbol{u}) := (\ \mathrm{sign}_1(u_1), \cdots, \mathrm{sign}_N(u_N)\ )^{\mathrm{T}}.$$

At the end of this subsection, we will show that $\mathbf{sign}(\boldsymbol{u}) = -\boldsymbol{x}$ must hold for any local optimum of $\mathbf{P}_2^+$ under an assumption that there is a j such that $w_{ij} \neq 0$ for any i.

Let $\boldsymbol{T}$ be a diagonal matrix having diagonal entries 1 or -1, and $T_{11} = 1$. Consider instances of $\mathbf{P}_2^+$,

$$A) \quad \text{Min. } E(\boldsymbol{x}) = \boldsymbol{x}^{*\mathrm{T}}\boldsymbol{W}\boldsymbol{x} \quad , \quad \text{s.t. } \boldsymbol{x} \in \mathcal{D}_2$$

$$B) \quad \text{Min. } F(\boldsymbol{y}) = \boldsymbol{y}^{*\mathrm{T}}\boldsymbol{W}'\boldsymbol{y} \quad , \quad \text{s.t. } \boldsymbol{y} \in \mathcal{D}_2$$

$$C) \quad \text{Min. } G(\boldsymbol{z}) = \boldsymbol{z}^{*\mathrm{T}}\bar{\boldsymbol{W}}\boldsymbol{z} \quad , \quad \text{s.t. } \boldsymbol{z} \in \mathcal{D}_2^+$$

, where $\bar{\boldsymbol{W}}$ and $\boldsymbol{W}'$ is defined by $\bar{w}_{ij} = w_{ij}(R_iR_j)^{-1}$ and $\boldsymbol{W}' = \boldsymbol{TWT}$. In addition, let $z_i = R_ix_i$ and $\boldsymbol{y} = \boldsymbol{Tx}$.

Lemma 4.1 summarizes the relationship among A), B), and C).

Lemma 4.1 *The following statements hold.*

1. *The decision variable $\boldsymbol{z}$ is feasible for C) if and only if $\boldsymbol{x}$ is feasible for A).*
2. *The relation $E(\boldsymbol{x}) = F(\boldsymbol{y}) = G(\boldsymbol{z})$ holds for any $\boldsymbol{x} \in \mathcal{D}_2$.*
3. *Maps $\boldsymbol{g} : \mathcal{D}_2 \to \mathcal{D}_2$ defined by*

   $$\boldsymbol{y} = \boldsymbol{Tx}$$

   and $\boldsymbol{f} : \mathcal{D}_2 \to \mathcal{D}_2^+$ defined by

   $$\boldsymbol{z} = f(\boldsymbol{x}) = \mathbf{Diag}(R_1, \cdots, R_N)\boldsymbol{x}$$

   are homeomorphic, where $\mathbf{Diag}(\boldsymbol{a})$ is the diagonal matrix which takes vector $\boldsymbol{a}$ as its diagonal elements.
4. *The decision variable $\boldsymbol{y}$ is a local(global) optimum of B) if and only if $\boldsymbol{x}$ is a local(global) optimum of A).*
5. *The decision variable $\boldsymbol{z}$ is a local(global) optimum of C) if and only if $\boldsymbol{x}$ is a local(global) optimum of A).*

Proof : Immediately from simple arithmetic inspections. ■

Lemma 4.2 *Let $x, y, z \in \{\mathbb{C}\backslash 0\}$, $x + y \neq 0$ and $\angle x = \angle z (\mathrm{mod}\pi)$. In addition, assume*

$$\mathrm{sign}(z) = \pm\mathrm{sign}(x + y).$$

Then

$$\angle x = \angle y (\mathrm{mod}\pi)$$

holds.

Proof : Let $\angle x = \angle z = 0(\text{mod}\pi)$ without loss of generality. Assume $\angle x \neq \angle y(\text{mod}\pi)$. Immediately, we have $\Im(x) = 0$, $\Im(y) \neq 0$ and $\Im(x+y) \neq 0$. Therefore, $\angle z \neq \angle(x+y)(\text{mod}\pi)$. This means $\text{sign}(z) \neq \pm\text{sign}(x+y)$, and it contradicts to the assumption. ■

Lemma 4.3 *Let* $\boldsymbol{x} \in \mathcal{D}_2$ *and* $u_i \neq 0$. *Then* $\boldsymbol{x}$ *is a local optimum of* $\mathbf{P}_2$ *only if*

$$\text{sign}(u_i) = -x_i. \tag{4.19}$$

Proof : First, $u_N^* x_N \leq 0$ must hold from the case of $K = N-1$ in proposition 3.6. This result yields

$$x_N \neq 0 \quad \Rightarrow \quad \text{sign}(u_N) = -\text{sign}(x_N).$$

Thus, $\text{sign}(u_N) = -x_N$ holds for the case $x_N \neq 0$ since u_N is the gradient of $E(\boldsymbol{x})$ in the direction of x_N.

From the same reason, the case $x_N = 0$ need not be considered. Finally, the lemma follows from corollary 3.3. ■

Lemma 4.4 *Assume that* $w_{ij} \neq 0$ *for some* j. *Then* $u_i = 0$ *implies* $\boldsymbol{x}$ *is not a local optimum of* $\mathbf{P}_2$.

Proof : First, we consider the cases $i \neq 1$. Let $u_i = 0$. Then x_i does not contribute to u_i and $E(\boldsymbol{x})$ since $E(\boldsymbol{x})$ is multilinear. Hence, $u_i = 0$ holds at $\boldsymbol{x} + \delta \boldsymbol{e}_i$ for any complex number δ. Moreover, u_j changes as the decision variable moves from $\boldsymbol{x}$ to $\boldsymbol{x} + \delta \boldsymbol{e}_i$. From the same reason, $u_j \neq 0$ at $\boldsymbol{x} + \delta \boldsymbol{e}_i (\delta \neq 0)$. Additionally, u_j can change so as to break the relation in eq.(4.19) for some δ. Hence, the feasible solution $\boldsymbol{x} + \delta \boldsymbol{e}_i$ is not a local optimum. In addition, $E(\boldsymbol{x}) = E(\boldsymbol{x} + \delta \boldsymbol{e}_i)$ holds from $u_i = 0$ and multilinearity of $E(\boldsymbol{x})$. Hence, there exists a $\boldsymbol{z}$ such that $E(\boldsymbol{z}) < E(\boldsymbol{x})$ in the neighborhood of $\boldsymbol{x}$. Therefore, such an $\boldsymbol{x}$ is not a local optimum.

Next, we consider the case $i = 1$. If $u_1 = 0$, then any perturbation of x_1 does not contribute to $E(\boldsymbol{x})$. Hence, $E(\boldsymbol{E}_1(\phi)\boldsymbol{x}) = E(\boldsymbol{x})$ holds, and is identical to $E(\boldsymbol{x}) = E(\boldsymbol{E}_1^-(-\phi)\boldsymbol{x})$. Let $\bar{u}_j$ be the value of u_j at $\boldsymbol{E}_1^-(-\phi)\boldsymbol{x}$, that is, $\bar{u}_j$ is defined by

$$\bar{u}_j = w_{j1}x_1 + e^{-\boldsymbol{j}\phi} \sum_{k=2}^{N} w_{jk} x_k.$$

From the fact $w_{j1}x_1 \in \mathbb{R}\backslash 0$ and lemma 4.2,

$$\text{sign}(\bar{u}_j) \neq -e^{-\boldsymbol{j}\phi} x_j$$

holds for some ϕ if

$$\sum_{k=2}^{N} w_{jk} x_k \neq 0.$$

On the other hand, $\bar{u}_j = w_{j1}x_1 \in \mathbb{R}$ holds if

$$\sum_{k=2}^{N} w_{jk}x_k = 0.$$

Hence, $\boldsymbol{E}_1^-(-\phi)\boldsymbol{x}$ is not a local optimum of $\mathbf{P}_2$ for some ϕ from lemma 4.3. Finally, $\boldsymbol{x}$ is not a local optimum since ϕ can be arbitrary small. ∎

Based on the above lemmas, we have the following necessary condition for local optimality.

Proposition 4.4 *Assume that*

$$\forall i \ \ \exists j \quad \text{such that} \ \ w_{ij} \neq 0.$$

Then

$$\mathbf{sign}(\boldsymbol{u}) = -\boldsymbol{x}$$

must hold at any local optimum of $\mathbf{P}_2^+$.

Proof : Immediately from lemmas 4.1, 4.3 and 4.4. ∎

Note that the property in lemma 4.4 and proposition 4.4, i.e., $| x_i | = R_i$ at any local optimum, does not hold for the original problem $\mathbf{P}_1^+$. Consider the case

$$\boldsymbol{W} = \begin{pmatrix} 0 & 1 & 1 \\ 1 & 0 & 1 \\ 1 & 1 & 0 \end{pmatrix} \quad , \quad R_1 = R_2 = R_3 = 1.$$

Then $\boldsymbol{x} = (\ 1 \ \ -1 \ \ \ 0\ )^{\mathrm{T}}$ is a local(and global) optimum of $\mathbf{P}_1$.

4.3.2 Instances of Problem Data for which Local Optima Gives Lower Bounds

In this subsection, we show that any local optimum of $\mathbf{P}_2$ gives a lower bound for $\mathbf{P}_1$ for certain classes of problem instances even if the original and the relaxation problem have many local optima.

We use the same notation in the previous subsection such as $\mathbf{P}_1^+$ and R_i.

Lemma 4.5 *Let* $\boldsymbol{s} = (\pm R_1, \cdots, \pm R_N)^{\mathrm{T}} \in \{\mathbb{R}\backslash 0\}^N$ *and define* $\boldsymbol{W}$ *by*

$$\boldsymbol{W} = \boldsymbol{s}\boldsymbol{s}^{\mathrm{T}} - \boldsymbol{D}. \tag{4.20}$$

, where $\boldsymbol{D}$ *is a diagonal matrix whose i-th diagonal entry is* R_i^2*. Then*

$$r = -\sum_{i=1}^{N} R_i^2$$

is a lower bound of the optimal value of $\mathbf{P}_1$.

Proof : The inequality $\boldsymbol{x}^{\mathrm{T}}(\boldsymbol{W}+\boldsymbol{D})\boldsymbol{x} \geq 0$ holds since $\boldsymbol{s}\boldsymbol{s}^{\mathrm{T}}$ is positive semidefinite. Therefore, $\boldsymbol{x}^{\mathrm{T}}(\boldsymbol{W}+\boldsymbol{D})\boldsymbol{x} = E(\boldsymbol{x}) + \boldsymbol{x}^{\mathrm{T}}\boldsymbol{D}\boldsymbol{x} \geq 0$. Thus, $E(\boldsymbol{x}) \geq -\boldsymbol{x}^{\mathrm{T}}\boldsymbol{D}\boldsymbol{x} \geq r$ at $\boldsymbol{x} \in \mathcal{D}_1$. ■

Proposition 4.5 *The value of the objective function at any local optimum of* $\mathbf{P}_2^+$ *is a lower bound of the optimal value of* $\mathbf{P}_1^+$ *if* $\boldsymbol{W}$ *is defined by*

$$\boldsymbol{W} = k(\boldsymbol{s}\boldsymbol{s}^{\mathrm{T}} - \boldsymbol{I}) \ , \ k \in \mathbb{R} \ , \ \boldsymbol{s} \in \{1\}^N.$$

Proof : Let y be defined by

$$y = \sum_{i=1}^{N} x_i.$$

Case of $y \neq 0$: When $k = 0$, $E(\boldsymbol{x})$ always equals to 0. Otherwise,

$$\boldsymbol{x} = -\mathbf{sign}(\boldsymbol{u}) = -\mathbf{sign}\left(k(y\boldsymbol{s} - \boldsymbol{x})\right) \tag{4.21}$$

must hold at any local optimum from proposition 4.4. In addition, we have $y \in \mathbb{R}$ from lemma 4.2 and $x_1 = 1$. This implies $\boldsymbol{x} \in \mathbb{R}^N$ at any local optimum from lemma 4.2. Finally, from proposition 3.5, the local optimum is a global optimum of $\mathbf{P}_1^+$. Hence, it gives a lower bound of the optimal value of $\mathbf{P}_1^+$.

Case of $y = 0$: If $k < 0$, then $\boldsymbol{x}$ is not a local optimum of $\mathbf{P}_2^+$ from eq.(4.21) and proposition 4.4. If $k > 0$, then the objective value is represented as

$$E(\boldsymbol{x}) = -k\sum_{i=1}^{N} R_i^2,$$

and is a lower bound from lemmas 4.1 and 4.5. ■

Finally, we have proposition 4.6.

Proposition 4.6 *Let* $\boldsymbol{W}$ *be defined by eq.(4.20). Then any local optimum of* $\mathbf{P}_2$ *gives a lower bound of the optimal value of* $\mathbf{P}_1$.

Proof : Immediately from lemma 4.1 and proposition 4.5. ■

Next, we consider the benchmark problem given by Pardalos et.al. [44]. It can be represented in term of $\mathbf{P}_1$. The coefficient matrix of $E(\boldsymbol{x})$ is represented as

$$\boldsymbol{W} = \begin{pmatrix} 0 & -\boldsymbol{e}^{\mathrm{T}} \quad \boldsymbol{0}^{\mathrm{T}} \\ -\boldsymbol{e} & \multirow{2}{*}{$\boldsymbol{A}$} \\ \boldsymbol{0} & \end{pmatrix} \in \mathcal{S}_{N+1} \quad , \quad \boldsymbol{A} = \begin{pmatrix} \boldsymbol{W}_{11} & \boldsymbol{W}_{12} \\ \boldsymbol{W}_{12}^{\mathrm{T}} & \boldsymbol{W}_{22} \end{pmatrix} \tag{4.22}$$

, where $\boldsymbol{W}_{11} \in \mathcal{S}_{\frac{N}{2}}$, $\boldsymbol{e} \in \{1\}^{N/2}$, $\boldsymbol{0} \in \{0\}^{N/2}$, and N is even. The matrix $\boldsymbol{A}$ is real symmetry and has diagonal entries 0, and N elsewhere. The unique global optimum of this problem is

$$\boldsymbol{x} = (\underbrace{1, \cdots, 1}_{\frac{N}{2}+1}, \underbrace{-1, \cdots, -1}_{\frac{N}{2}})^{\mathrm{T}}$$

and each feasible solution, at which exactly $N/2$ x_is equal to -1 and others equal to 1, is a strict local optimum.

Proposition 4.7 *Let $\boldsymbol{W}$ in $\mathbf{P}_1$ be defined by eq.(4.22). Then $\boldsymbol{x}$ is a local optimum of $\mathbf{P}_2$ if and only if $\boldsymbol{x}$ is a global optimum of $\mathbf{P}_1$.*

Proof : Define sets of indices by

$$\mathcal{F} = \left\{2, \cdots, \frac{N}{2} + 1\right\} \text{ and } \mathcal{L} = \left\{\frac{N}{2} + 2, \cdots, N + 1\right\},$$

and let

$$\boldsymbol{x}_{\mathcal{F}} = (x_2, \cdots, x_{\frac{N}{2}+1})^{\mathrm{T}} \quad \text{and} \quad \boldsymbol{x}_{\mathcal{L}} = (x_{\frac{N}{2}+2}, \cdots, x_{N+1})^{\mathrm{T}}.$$

In addition, let y be defined by

$$y = N \sum_{i=2}^{N+1} x_i.$$

Then, $E(\boldsymbol{x})$ is represented as

$$E(\boldsymbol{x}) = -2\Re\left(\sum_{i \in \mathcal{F}} x_i\right) - N \sum_{i \in \mathcal{F}, \mathcal{L}} \mid x_i \mid^2 + \frac{\mid y \mid^2}{N}. \tag{4.23}$$

Additionally, we define $\boldsymbol{u}$ by $\boldsymbol{W}\boldsymbol{x}$. Then we have

$$u_1 = -\sum_{i \in \mathcal{F}} x_i \quad , \quad u_i = -1 + y - N x_i \ \ (i \in \mathcal{F}) \ , \text{ and } \ u_i = y - N x_i \ \ (i \in \mathcal{L}). \tag{4.24}$$

We split the rest of proof into three cases about y.

Case of $y(y-1) \neq 0$

We have

$$\angle x_i = \angle x_j = \angle(-1 + y)(\text{mod } \pi) \quad (i, j \in \mathcal{F}) \tag{4.25}$$

$$\angle x_i = \angle x_j = \angle y(\text{mod } \pi) \quad (i, j \in \mathcal{L}) \tag{4.26}$$

at any local optimum from proposition 4.4 and lemma 4.2. From eq.(4.25), proposition 4.4 and $x_1 = 1$, we have $\angle x_i = 0(\text{mod } \pi)$ for any $i \in \mathcal{F}$. Additionally, we have $\angle y = 0(\text{mod}\pi)$ from lemma 4.2. Therefore, $\angle x_i = 0(\text{mod}\pi)$ for all $i \in \mathcal{L}$ from eq.(4.26). This means $\boldsymbol{x} \in \mathcal{D}_1$. Hence, $\boldsymbol{x}$ is a global optimum of $\mathbf{P}_1$ from proposition 3.5 if such a local optimum exists.

Case of $y = 0$

If $y = 0$, then $u_i = -1 - Nx_i$ for all $i \in \mathcal{F}$. Hence, $x_i \in \mathbb{R}$ for all $i \in \mathcal{F}$ from proposition 4.4 and lemma 4.2. Then, immediately, we have

$$y - N\sum_{i\in\mathcal{F}} x_i = -N\sum_{i\in\mathcal{F}} x_i = N\sum_{i\in\mathcal{L}} x_i \in \mathbb{R}. \tag{4.27}$$

If $x_k = 1$ for all $k \in \mathcal{F}$, then $x_l = -1$ must hold for any $l \in \mathcal{L}$ from eq.(4.27). In this case, the local optimum is a global optimum. Hence, assume that an $\alpha \in \mathcal{F}$ exists such that $x_\alpha = -1$, hereafter.

Let $e^{\boldsymbol{j}\theta_i}$ denotes x_i. Consider an operation that changes all $x_i = e^{\boldsymbol{j}\theta_i}(i \in \mathcal{F})$ satisfying $x_i = -1$ to $e^{\boldsymbol{j}(\theta_i+\Delta)}$, and does not change other $x_i(i \in \mathcal{F})$s. Then, the change of $\sum_{i\in\mathcal{F}} x_i$ is represented by $p\{(1-\cos\Delta) - \boldsymbol{j}\sin\Delta\}$, where p is the number of $x_i(i \in \mathcal{F})$ with the value -1. The condition $y = 0$ is preserved by the operation when the following equality hold.

$$\begin{aligned} &\sum_{i\in\mathcal{L}} \{\cos(\theta_i+\delta_i) + \boldsymbol{j}\sin(\theta_i+\delta_i)\} - \sum_{i\in\mathcal{L}} (\cos\theta_i + \boldsymbol{j}\sin\theta_i) \\ = \;& -p\{(1-\cos\Delta) + \boldsymbol{j}\sin\Delta\} \end{aligned} \tag{4.28}$$

, where $e^{\boldsymbol{j}(\theta_i+\delta_i)}$ is the value of $x_i(i \in \mathcal{L})$ after the operation. The left hand side of eq.(4.28) is represented as

$$\boldsymbol{f}(\boldsymbol{\delta}) = \begin{pmatrix} \sum_{i\in\mathcal{L}} \cos\theta_i(\cos\delta_i - 1) - \sin\theta_i\sin\delta_i \\ \sum_{i\in\mathcal{L}} \sin\theta_i(\cos\delta_i - 1) + \cos\theta_i\sin\delta_i \end{pmatrix}$$

if we deal a complex number as a two dimensional vector. The partial derivative of $\boldsymbol{f}$ is

$$\frac{\partial \boldsymbol{f}}{\partial \delta_i} = \begin{pmatrix} -\cos\theta_i\sin\delta_i - \sin\theta_i\cos\delta_i \\ -\sin\theta_i\sin\delta_i + \cos\theta_i\cos\delta_i \end{pmatrix}.$$

The Jacobian matrix around $\boldsymbol{\delta} = \boldsymbol{0}$ is

$$\boldsymbol{J} = \begin{pmatrix} -\sin\theta_{\frac{N}{2}+2} & , & \cdots & , & -\sin\theta_{N+1} \\ \cos\theta_{\frac{N}{2}+2} & & & & \cos\theta_{N+1} \end{pmatrix}.$$

If $\text{rank}(\boldsymbol{J}) = 2$, there is a $\boldsymbol{\delta}$ satisfying eq.(4.28) around origin. Then the decision variable after the operation reduces $E(\boldsymbol{x})$ since $y = 0$ and $|x_i| = 1$ are preserved, and $-\sum_{i\in\mathcal{F}} \Re(x_i)$ decreases(see eq.(4.23)).

If $\text{rank}(\boldsymbol{J}) = 1$, then $\angle x_i = \angle x_j (i,j \in \mathcal{L}) (\text{mod}\pi)$ holds. Therefore, $x_i \in \mathbb{R}(i \in \mathcal{L})$ or $\sum_{i\in\mathcal{L}} x_i = 0$ holds from eq.(4.27). If $x_i \in \mathbb{R}(i \in \mathcal{L})$, the feasible solution is a global optimum of the original problem from proposition 3.5. If $\sum_{i\in\mathcal{L}} x_i = 0$, then $\sum_{i\in\mathcal{F}} x_i = 0$ from eq.(4.27). Hence, $u_1 = 0$, and the feasible solution is not a local optimum from lemma 4.4.

Case of $y = 1$

First, $u_i = 1 - Nx_i$ for all $i \in \mathcal{L}$. Thus, $\angle x_i = 0 (\text{mod } \pi)(i \in \mathcal{L})$ holds from proposition 4.4 and lemma 4.2. Hence, we have

$$y - N\sum_{i\in\mathcal{L}} x_i = 1 - N\sum_{i\in\mathcal{L}} x_i = N\sum_{i\in\mathcal{F}} x_i \in \mathbb{R}.$$

If $x_\alpha \in \mathbb{R}$ for all $\alpha \in \mathcal{F}$, then the local optimum is a global optimum from proposition.3.5. Hence assume that $\alpha \in \mathcal{F}$ such that $x_\alpha \notin \mathbb{R}$ exists, hereafter.

If $x_\alpha = -1$ for all $\alpha \in \mathcal{L}$, then y cannot be 1 from the definition of y. Hence, let there exist an $\alpha \in \mathcal{L}$ satisfying $x_\alpha = 1$.

Consider the operation that changes $x_i(i \in \mathcal{L})$ satisfying $x_i = 1$ to $e^{\jmath\theta_i + \Delta}$ and does not change other $x_i(i \in \mathcal{L})$s. Then the rest of proof is same as the case of $y = 0$. ■

4.4 Extreme Instances

In this section, we present examples touched in section 3.1.

4.4.1 A Case $\mathbf{QP}_2$ has infinite local optima

Here, we show an instance of problem data for which $\mathbf{QP}_2$ has infinitely many local optima.

Consider $\mathbf{P}_2$ and define $\boldsymbol{W}$ by

$$\boldsymbol{W} = 3\boldsymbol{I} - \sum_{i=1}^{3} \boldsymbol{v}^{(i)}\boldsymbol{v}^{(i)\mathrm{T}} = \begin{pmatrix} 0 & -1 & -1 & 1 \\ -1 & 0 & 1 & -1 \\ -1 & 1 & 0 & -1 \\ 1 & -1 & -1 & 0 \end{pmatrix},$$

where $\boldsymbol{v}_i$s are defined by

$$\boldsymbol{v}^{(1)} = \begin{pmatrix} 1 & 1 & 1 & 1 \end{pmatrix}^{\mathrm{T}}, \quad \boldsymbol{v}^{(2)} = \begin{pmatrix} 1 & 1 & -1 & -1 \end{pmatrix}^{\mathrm{T}}, \quad \boldsymbol{v}^{(1)} = \begin{pmatrix} 1 & -1 & 1 & -1 \end{pmatrix}^{\mathrm{T}}.$$

They are orthogonal one another. For this problem instance

$$\boldsymbol{x} = \begin{pmatrix} 1 & 1 & e^{\jmath\phi} & e^{\jmath\phi} \end{pmatrix}^{\mathrm{T}}$$

becomes a local optimum of $\mathbf{P}_2$ for any ϕ. This is proved as follows.

Let $\boldsymbol{v}_i(i = 1, \cdots, N)$ be vectors in $\mathbb{C}^N$ that are orthogonal one another. Additionally, let each of $\boldsymbol{v}_1, \cdots, \boldsymbol{v}_K$ be in $\{e^{\jmath\theta}\}^N$. Define $\boldsymbol{W}$ by

$$\boldsymbol{W} = K\boldsymbol{I} - \sum_{i=1}^{K} \boldsymbol{v}_i\boldsymbol{v}_i^{*\mathrm{T}}, \quad 1 \le K < N,$$

and consider a feasible solution represented as

$$x = \sum_{i=1}^{N} k_i v_i$$

and satisfies $\mid x_i \mid = 1$.

In this case, we have

$$\begin{aligned} u &= Wx = \left(MI - \sum_{i=1}^{K} v_i v_i^{*\mathrm{T}} \right) x = \left(MI - \sum_{i=1}^{K} v_i v_i^{*\mathrm{T}} \right) \sum_{i=1}^{N} k_i v_i \\ &= M \sum_{i=1}^{K} k_i v_i - \sum_{i=1}^{K} N k_i v_i. \end{aligned}$$

If $k_{K+1}, \cdots, k_N = 0$, then we have $u = (M - N)x$. In addition, in general, $E(x)$ is represented as

$$E(x) = x^{*\mathrm{T}} u = N(M - N) \sum_{i=1}^{K} k_i^* k_i.$$

On the other hand, $E(x)$ attains its minimum if and only if x can be represented as a linear combination of $v_1, \cdots, v_K$ since

$$x^* x = N \sum_{i=1}^{N} k_i^* k_i = N$$

holds.

Finally,

$$x = \begin{pmatrix} 1 & 1 & e^{j\theta} & e^{j\theta} \end{pmatrix}^{\mathrm{T}}$$

is a linear combination of v_1, v_2, and v_3. Namely,

$$\begin{pmatrix} 1 & 1 & e^{j\theta} & e^{j\theta} \end{pmatrix}^{\mathrm{T}} = \frac{(1 + e^{j\theta})}{2} v_1 + \frac{(1 - e^{j\theta})}{2} v_2 + 0 \cdot v_3.$$

Thus, $\mathbf{P}_2$ for this problem instance has infinitely many local optima.

Next, we show that any local optimum of the original problem is also a global optimum. If $\mid x_i \mid < 1$ at a local optimum, then

$$\frac{\partial E(x)}{\partial x_i} = 0$$

must hold. For this reason, $x_3 - x_4 = -1$, $x_2 - x_4 = 1$ and $x_2 + x_3 = 1$ must hold for any local optimum satisfying $\mid x_2 \mid < 1$, $\mid x_3 \mid < 1$ and $\mid x_3 \mid < 1$, respectively. This implies that

$$\begin{aligned} &\mid x_3 \mid \neq 1 \quad \text{or} \quad \mid x_4 \mid \neq 1 , \quad \text{if} \quad \mid x_2 \mid < 1 \\ &\mid x_2 \mid \neq 1 \quad \text{or} \quad \mid x_4 \mid \neq 1 , \quad \text{if} \quad \mid x_3 \mid < 1 \\ &\mid x_2 \mid \neq 1 \quad \text{or} \quad \mid x_3 \mid \neq 1 , \quad \text{if} \quad \mid x_4 \mid < 1. \end{aligned}$$

Moreover, $| x_i |< 1$ implies $\boldsymbol{x}$ is not a local optimum since each principal submatrix of the Hessian matrix of $E(\boldsymbol{x})$ is not positive semidefinite. Finally, the number of local optima is finite for $\mathbf{P}_1$since the number of feasible solutions satisfying

$$\begin{pmatrix} x_2 & x_3 & x_4 \end{pmatrix} \in \{\pm 1\}^3,$$

is finite.

4.4.2 A Case QP_2 has exponentially many strict local optima

We show an example for which $\mathbf{QP}_2$ has exponentially many strict local optima.

Consider $\mathbf{P}_2$ and define $\boldsymbol{W}$ by

$$\boldsymbol{W} = \begin{pmatrix} 0 & \boldsymbol{e}^T \\ \boldsymbol{e} & \bar{\boldsymbol{W}}. \end{pmatrix}$$

using a block diagonal matrix

$$\bar{\boldsymbol{W}} = \mathbf{Diag}\begin{pmatrix} \bar{\boldsymbol{W}}_{11} & \bar{\boldsymbol{W}}_{22} & \cdots & \bar{\boldsymbol{W}}_{aa} \end{pmatrix} \ , \ \bar{\boldsymbol{W}}_{ii} = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}.$$

Then $\mathbf{P}_2$ has 2^a local optima. This fact is proved as follows.

From proposition 4.4, we only have to consider the case $| x_i |= 1$.

First, define a vector $\boldsymbol{x}$ by

$$\boldsymbol{x} = \begin{pmatrix} 1 & \boldsymbol{y}_1 & \cdots & \boldsymbol{y}_a \end{pmatrix}^{\mathrm{T}}, \tag{4.29}$$

where

$$\boldsymbol{y}_i \in \left\{ \begin{pmatrix} -\frac{1}{2} - j\frac{\sqrt{3}}{2} & , & -\frac{1}{2} + j\frac{\sqrt{3}}{2} \end{pmatrix}^{\mathrm{T}}, \begin{pmatrix} -\frac{1}{2} + j\frac{\sqrt{3}}{2} & , & -\frac{1}{2} - j\frac{\sqrt{3}}{2} \end{pmatrix}^{\mathrm{T}} \right\}.$$

Then such an $\boldsymbol{x}$ becomes a stationary point of

$$T(\boldsymbol{\theta}) = \sum_{i=1}^{N}\sum_{j=1}^{N} w_{ij}\cos(\theta_i - \theta_j),$$

where $T(\boldsymbol{\theta})$ is defined by replacing x_i in $E(\boldsymbol{x})$ with $e^{j\theta_i}$. Moreover, there are 2^a such $\boldsymbol{x}$s from definition of $\boldsymbol{x}$ and $\boldsymbol{y}$.

On the other hand, Hessian matrix $\boldsymbol{H}(\boldsymbol{\theta})$ of $\boldsymbol{T}(\boldsymbol{\theta})$ has entries

$$\begin{aligned} & h_{ij} < 0 \quad \text{for any} \quad i, j \quad \text{such that} \quad w_{ij} = 1 \\ & h_{ij} = 0 \quad \text{for any} \quad i, j \quad \text{such that} \quad w_{ij} = 0 \\ & \sum_{j=1}^{N} h_{ij} = 0 \end{aligned}$$

at any $\boldsymbol{x}$ represented as eq.(4.29). Therefore, for such $\boldsymbol{H}(\boldsymbol{\theta})$,

$$\boldsymbol{z}^{\mathrm{T}}\boldsymbol{H}(\boldsymbol{\theta})\boldsymbol{z} = -\sum_{i=1}^{N}\sum_{j=1}^{N} h_{ij}(z_i - z_j)^2$$

holds. This means that $\boldsymbol{H}(\boldsymbol{\theta})$ is positive semidefinite. Therefore, $\boldsymbol{z}^{\mathrm{T}}\boldsymbol{H}\boldsymbol{z} = 0$ is satisfied only if $z_i = z_j$ for any i and j such that $w_{ij} \neq 0$.

From the fact that there are $\alpha, \beta, \cdots, \gamma$ such that

$$w_{i\alpha}w_{\alpha\beta}\cdots w_{\gamma j} \neq 0$$

for any $i \neq j$, the above condition holds only if

$$\boldsymbol{z} = k\boldsymbol{e}$$

Finally, the direction $\boldsymbol{z}$ of $\boldsymbol{\theta}$ is not feasible from the constraint $x_1 = 1$. Therefore, each $\boldsymbol{x}$ defined by eq.(4.29) is a strict local optimum.

4.4.3 Discussions

In this section, we showed propositions only for the special classes of $\boldsymbol{W}$. For general class of $\boldsymbol{W}$, we do not know whether any local optima of $\mathbf{P}_2$ gives a lower bound. If $\mathbf{P}_2$ always gives a lower bound, then $\mathbf{P}_2$ may become a powerful tool for calculating a lower bound since

- obtaining a local optimum is often easy even if the problem is not convex.
- The lower bound given at a local optimum is never worse than that given at a global optimum of $\mathbf{P}_2$. In addition, a lower bound by global optimum of $\mathbf{P}_2$ is never worse than that by convex relaxation problem $\mathbf{P}_N$.

Chapter 5

Optimization Algorithms

In this chapter, we consider optimization algorithms for the maximum cut problems based on analytical results in the previous chapters and some conjectures. Computational experiments show that the performance of one of the proposed algorithms is better than the tabu search's one. They also show the effects of expanding the space of decision variables.

5.1 Framework of Constructing Optimization Algorithms

This section will be dedicated to describe a framework or an idea of optimization algorithms which make use of a monotone path. The description of algorithms here does not work alone for **QP** as far as a monotone path cannot be calculated easily. However, for some cases, it is expected that the framework is available by using properties given by specifying the class of the original problems to $\mathbf{P}_1$.

If we can find and follow exactly the monotone path towards the global optimum, the global optimum of the original problem $\mathbf{P}_1$ will be found by a decent type search. Therefore, in principle, it is sufficient to follow the monotone path with $\epsilon = 1$. However, we do not know any algorithm to do so efficiently. Moreover, the existence of such algorithms seems to be hopeless for NP-hard problems.

The aim of this chapter is to imply and demonstrate the availability of our approach by

- applying the analytical results in constructing optimization algorithms
- showing the performance of proposed algorithms through experiments

Algorithms having structures similar to those for the original problems are expected to be available since the relaxation $\mathbf{QP}_2$ has less gaps to **QP**. Algorithms to be proposed are simple generalizations of a local search algorithm for the original problem. A local optimization algorithm for the relaxation problem $\mathbf{QP}_2^\epsilon$ is conceptually formulated as follows.

Consider the following optimization problem,

$$(\mathbf{LS}) \quad \text{Minimize}_{\boldsymbol{d}} \ G(\boldsymbol{d}) \quad , \quad \text{Subject to } \boldsymbol{x}+\boldsymbol{d} \in \mathcal{G} \cap \mathcal{D}_{q2}^{\epsilon}$$

The constraint $\boldsymbol{x}+\boldsymbol{d} \in \mathcal{G}$ and the objective function $G(\boldsymbol{d})$ specifies **LS**. By solving **LS** we can obtain a new feasible solution of $\mathbf{QP}_2^{\epsilon}$. After $\boldsymbol{x}$ converged to a point at which **LS** has an optimal solution $\boldsymbol{d} = \mathbf{0}$, making ϵ small yields tighter relaxation. From definition of $\mathbf{QP}_M^{\epsilon}$ and its fundamental properties described in chapter 3, the feasibility is kept at $\boldsymbol{x}^{(1)} + \boldsymbol{j}\epsilon\boldsymbol{x}^{(2)}$ if $\mathcal{D}_{q2}^{\epsilon}$ is convex. Moreover, if we can use properties like the operation **Shift**, then making M small also keeps feasibility. Moreover, in cases we use $\mathbf{Shift}(\cdot)$, the value of the objective function does not increase.

Though covering class **QP** by such an algorithm is hard since the class is too large compared to our analytical results, we expect for $\mathbf{P}_1$ to have some good effects from our approach. In other words, we expect that optimization algorithms analogous to that for the original problem will be applicable to the relaxation problem.

5.2 Optimization Algorithms

Here, we derive special cases of the local search algorithm **LS** for $\mathbf{P}_M^{\epsilon}$. They are generalizations of an algorithm such that

(alg.0 : simple coordinate optimization)

Solve **LS** repeatedly with $G(\boldsymbol{d}) = E(\boldsymbol{x}+\boldsymbol{d})$ and $\mathcal{G} = \{-1,+1\}^N$.

5.2.1 Convergence to Real Number(alg.1)

We showed in proposition 4.4 that any local optimum of $\mathbf{P}_2$ lies on edges of the feasible region. Therefore the most important information about local optima lies on the edges of $\mathcal{D}_2$. If we restrict our discussion within feasible solutions on edges of $\mathcal{D}_2^{\epsilon}$, the objective function is written as

$$\begin{aligned} E(\boldsymbol{x}) = T(\boldsymbol{\theta}) &= \sum_{i=1}^{N}\sum_{j=1}^{N} w_{ij}\left\{\cos\theta_i\cos\theta_j + \epsilon^2\sin\theta_i\sin\theta_j + \boldsymbol{j}\epsilon\cos\theta_i\sin\theta_j - \epsilon\sin\theta_i\cos\theta_j\right\} \\ &= \sum_{i=1}^{N}\sum_{j=1}^{N} w_{ij}\left\{\cos\theta_i\cos\theta_j + \epsilon^2\sin\theta_i\sin\theta_j\right\} \end{aligned}$$

by replacing variables x_i with $\cos\theta_i + \boldsymbol{j}\epsilon\sin\theta_i$. Immediately, we have

$$\frac{\partial T(\boldsymbol{\theta})}{\partial\theta_i} = 2\sum_{j=1}^{N} w_{ij}\left\{-\sin\theta_i\cos\theta_j + \epsilon^2\cos\theta_i\sin\theta_j\right\}$$

$$\frac{\partial^2 T(\boldsymbol{\theta})}{\partial\theta_i^2} = 2\sum_{j=1}^{N} w_{ij}\left\{-\cos\theta_i\cos\theta_j - \epsilon^2\sin\theta_i\sin\theta_j\right\}$$

$$\frac{\partial^2 T(\boldsymbol{\theta})}{\partial\theta_i \partial\theta_j} = 2w_{ij}\left\{\sin\theta_i \sin\theta_j + \epsilon^2 \cos\theta_i \cos\theta_j\right\} \ , \ (i \neq j),$$

where θ_i is not the argument of a complex number when $\epsilon \neq 1$. The argument of a complex variable x_i is $\arctan(\epsilon \sin\theta_i / \cos\theta_i)$.

Based on this description, local search algorithms like coordinate optimization and quasi-Newton method are applicable for $\mathbf{P}_2$. Moreover, the gradient and the Hessian matrix are computed by simply substituting θ_is. Moreover, the optimization problem has no constraint except for $\theta_1 = 0$. Let $\boldsymbol{H}(\boldsymbol{\theta})$ stand for Hessian matrix of $T(\boldsymbol{\theta})$. Then, local optimality of $\boldsymbol{x}\{-1,+1\}^N$ is equivalent to nonnegativity of diagonal entries of $\boldsymbol{H}(\boldsymbol{\theta})$ for $\mathbf{P}_1$. On the other hand, it is nearly equivalent to $\boldsymbol{x} = -\mathbf{sign}(\boldsymbol{u})$ and $H(\boldsymbol{\theta}) \succeq \boldsymbol{O}$ for $\mathbf{P}_2$. In addition, $-E(\boldsymbol{x})$ coincides with the trace of $\boldsymbol{H}(\boldsymbol{\theta})$ when $\epsilon = 1$. From the relation

$$E(\boldsymbol{x}) = -\mathrm{trace}\left(\boldsymbol{H}(\boldsymbol{\theta})\right)$$

and the condition $\boldsymbol{H}(\boldsymbol{\theta}) \succeq \boldsymbol{O}$ for local optimality, it is expected that a feasible solution $\boldsymbol{x}$ hardly become a local optimum of $\mathbf{P}_2$ when ϵ is not too small and $T(\boldsymbol{\theta})$ is greater than the optimal value of the original problem.

From this observation, we expect **alg.1** defined later will output a good feasible solution of $\mathbf{P}_1$ by keeping the decision variable away from local optima at which $E(\boldsymbol{x}) = -\mathrm{trace}\left(\boldsymbol{H}(\boldsymbol{\theta})\right)$ are large.

The algorithm(**alg.1**) we used in experiments is defined as **LS** in which $G(\boldsymbol{d}) = E(\boldsymbol{x} + \boldsymbol{d})$ and

$$\mathcal{G} = \left\{\boldsymbol{x} + \boldsymbol{d} \middle| d_i = 0 \ \text{ except for an } \ i \right\}$$

The fundamental structure of **alg.1** is understood as follows.

- add a bias to the search direction while **LS** is repeatedly solved.

In other words, we begin with the state in which the existence of a monotone path is guaranteed and make the decision variables converge to real axis. We expect $\boldsymbol{x}$ to follow monotone paths approximately.

The **for** loop begins at line 5) is the procedure for changing the decision variable $\boldsymbol{x}$ and search local optimum for fixed ϵ. At the **if** statement in line 10), the bias to the real axis is incremented. Third **if** statement appears at line 11) is a criterion to terminate iterations.

(alg.1 : convergence to real axis)

1) **given** $\boldsymbol{x} \in \mathcal{D}_2(|\ x_i\ | = 1)$, $\gamma > 0$, $R(1 > R > 0)$ and $\delta(1 > \delta > 0)$
2) $\epsilon := 1$
3) $\boldsymbol{x}_{prv} := \boldsymbol{x}$
4) **while** [1]

5) **for** $i(2 \leq i \leq N)$
6) $u := -\sum_j w_{ij} x_j$
7) $u := \Re[u] + \epsilon \Im[u]$
8) **if** $u \neq 0$, $x_i := u / \mid u \mid$.
9) **end for**
10) **if** $\|\boldsymbol{x} - \boldsymbol{x}_{prv}\| < \gamma$, $\epsilon := R \times \epsilon$.
11) **if** $\epsilon < \delta$, **break while**
12) $\boldsymbol{x}_{prv} := \boldsymbol{x}$
13) **end while**
14) **for** $i(2 \leq i \leq N)$
15) **if** $\Re[x_i] \geq 0$, $x_i := 1$. **else** $x_i := -1$.
16) **end for**

Note that if "given $\boldsymbol{x} \in \mathcal{D}_2$" is replaced with "given $\boldsymbol{x} \in \mathcal{D}_1$" in **alg.1**, then we obtain **alg.0**. Note also that the following proposition holds.

Proposition 5.1 *Let* $\epsilon = 1$*. In* **alg.1**, $E(\boldsymbol{x})$ *decreases at line 8) if* $\boldsymbol{x}$ *changes.*

Proof : Let l be the index of the variable to be updated at line 8). Let x_l and x_l' be the value before updating and after updating, respectively. The change dE of the objective value is represented as

$$dE = -\Re\left(x_l'^* u - x_l^* u\right).$$

From the rule

$$x_i := \frac{u}{\mid u \mid},$$

we have

$$\Re(x_l'^* u) = \mid u \mid > \Re(x_l^* u)$$

when $x_l' \neq x_l$. This means dE is negative. ∎

5.2.2 Local Search in Discrete Subset of Hypercomplex Numbers(alg.2)

The following algorithm(**alg.2**) utilizes proposition 4.1. It is written as

$$G(\boldsymbol{d}) = E(\boldsymbol{x} + \boldsymbol{d}) \quad , \quad \mathcal{G} = \left\{ \boldsymbol{d} \middle| \boldsymbol{x} + \boldsymbol{d} \in \mathcal{D}_{Md}^N \text{ and } d_j = 0 \text{ except for a } j \right\}$$

in term of $G(\cdot)$ and $\mathcal{G}$.

The following is a high level description of **alg.2**.

- a local search procedure is performed in $\mathcal{D}_{Md}^{N}$ repeatedly with decreasing the dimension of hypercomplex numbers by adopting proposition 4.1.

The **for** loop at line 5) is a coordinate optimization in $\mathcal{D}_{Md}^{N}$, where

$$\mathbf{u} = -\sum_{j=1}^{N} w_{ij}\mathbf{x}_j$$

is a decent direction of the objective function. The operator **Shift**$(\cdot)$ is same as defined in chapter 4.

(alg.2 : a local search in discrete hypercomplex numbers)

1) **given** $\boldsymbol{x} \in \mathcal{D}_{Md}^{N}$
2) $L := M$
3) $\boldsymbol{x}_{prv} := \boldsymbol{x}$
4) **while** [1]
5)   **for** $i (2 \leq i \leq N)$
6) $\mathbf{u} := -\sum_j w_{ij}\mathbf{x}_j$
7) $\mathbf{x}_i := \underline{\text{argmin}}_{\mathbf{y}} \mid \mathbf{u} - \mathbf{y} \mid$, s.t. $\mathbf{y} \in \mathcal{D}_{Ld}$
8)   **end for**
9)   **if** $\boldsymbol{x} == \boldsymbol{x}_{prv}$
10)   **if** $L == 1$, **then break while.**
11)   **else** $\boldsymbol{x} :=$ **Shift**$(\boldsymbol{x})$ and $L := L - 1$.
12)   **end if**
13) $\boldsymbol{x}_{prv} := \boldsymbol{x}$
14) **end while**

The operation $\underline{\arg\min}_{\mathbf{y}}$ in line 7) is defined as follows. Let $\mathcal{S}$ be the set of optimal solutions of the following optimization problem.

$$\min_{\mathbf{y}} \mid \mathbf{u} - \mathbf{y} \mid \quad \text{subject to} \quad \mathbf{y} \in \mathcal{D}_{Ld}$$

Let l be the index of the variable to be updated in line 7) and $\mathbf{x}_l$ be the value before updating. Then, we define

$$\underline{\arg\min}_{\mathbf{y}} \mid \mathbf{u} - \mathbf{y} \mid = \begin{cases} \text{an element of } \mathcal{S} & \text{if} \quad \mathbf{x}_l \notin \mathcal{S} \\ \mathbf{x}_l & \text{if} \quad \mathbf{x}_l \in \mathcal{S} \end{cases} .$$

The integer L in **alg.2** is the dimension of hypercomplex numbers, and the minimization at $\underline{\arg\min}_{\mathbf{y}}$ selects an element of $\mathcal{D}_{Ld}$. For **alg.2**, the following proposition holds.

Proposition 5.2 *In* **alg.2**, *$E(\boldsymbol{x})$ decreases at line 7) when $\boldsymbol{x}$ changes. In addition, $E(\boldsymbol{x})$ never increases until the algorithm terminates.*

Proof : Let the index of the hypercomplex variable to be updated be l. Let the feasible solution before updating be $\boldsymbol{x}$ and after updating be $\boldsymbol{x}'$. In addition, let the difference of the objective value be dE. Then,

$$dE = \left(\boldsymbol{x}'^{*\mathrm{T}}\boldsymbol{W}\boldsymbol{x}' - \boldsymbol{x}^{*\mathrm{T}}\boldsymbol{W}\boldsymbol{x}\right) = 2\Re\left(\sum_{j=1}^{N} w_{lj}\mathbf{x}_l'^{*}\mathbf{x}_j - \sum_{j=1}^{N} w_{lj}\mathbf{x}_l^{*}\mathbf{x}_j\right) = -\Re\left(\mathbf{x}_l'^{*}\mathbf{u} - \mathbf{x}_l^{*}\mathbf{u}\right)$$

, where $\mathbf{u} = \begin{pmatrix} u_1 & \cdots & u_N \end{pmatrix}$ and $\mathbf{x}_l = \begin{pmatrix} x_{l1} & \cdots & x_{lM} \end{pmatrix}$.

From the definition of $\mathcal{D}_{Md}$, only one of $x_{l1}, x_{l2}, \cdots, x_{lM}$ does not equal to 0. Therefore, we have

$$\frac{1}{2}dE = -u_q x_{lq}' + u_p x_{lp}$$

, where x_{lp}' stands for the nonzero entry of $\mathbf{x}_l'$ and x_{lq} is the nonzero entry of $\mathbf{x}_l$.

On the other hand, from $\mid \mathbf{u} - \mathbf{x}_l' \mid < \mid \mathbf{u} - \mathbf{x}_l \mid$, we have

$$(u_q - x_{lq}')^2 + u_p^2 < u_q^2 + (u_p - x_{lp})^2$$

if $\mathbf{x}' \notin \mathcal{S}$. Then

$$dE = -2u_q x_{lq}' + 2u_p x_{lp} < 0$$

holds since $x_{lq}', x_{lp} \in \{\pm 1\}$. Therefore, if $\mathbf{x}_l \notin \mathcal{S}$, then the value of the objective function decreases.

In case $\mathbf{x}_l \in \mathcal{S}$, the decision variable $\boldsymbol{x}$ does not change from the definition of $\underline{\arg\min}_{\mathbf{y}}$.

Finally, the fact that **Shift** does not make $E(\boldsymbol{x})$ larger follows the proposition. ■

5.3 Computational Experiments

We compared the performance of the proposed algorithms and some known algorithms in computational experiments. All algorithms used are implemented in C++ and compiled by g++ version 2.95.x [49]. The tabu search algorithm was designed based on [9, 14] and tuned through experiments.

The test problems are randomly generated 30 instances for each class of problems such as unweighted maximum cut(U_MC), weighted maximum cut(W_MC), and maximum cut with negative weight(N_MC). We formulated the maximum cut problems as minimization problems and set the offset term of the objective function to zero, that is, all w_{ii}s were set to 0. The elements of coefficient $\boldsymbol{W}$ of the objective function are generated by integral uniform distribution over $[0, 1000]$ for weighted maximum cut, and $[-1000, 1000]$ for maximum cut with negative weight. The sparsity of edges of the problem instances of unweighted maximum cut was set to be 0.03, namely,

$$\frac{\sum_{i=1}^{N}\sum_{j=1}^{N} w_{ij}}{N(N-1)} = 0.03.$$

Table 5.1: algorithms used in experiments

algorithm	implementation
tabu search	based on [9] and [14], and tuned through experiments
alg.0	alg.2 with $M = 1$
alg.1	$R = 0.95$, $\gamma = 1.00 \times 10^{-6}$ and $\delta = 0.1$
alg.2($M = 2$)	
alg.2($M = 3$)	selection of imaginary axes in **Shift** is fixed

Table 5.2: problem instances

class	dim. of $\boldsymbol{x}$	# of instances	entries of $\boldsymbol{W}$	sparsity
U_MC	1000	30	$\{0,1\}$	0.03
W_MC	1000	30	$[0,1000] \cap \mathbb{Z}$	-
N_MC	1000	30	$[-1000,1000] \cap \mathbb{Z}$	-

We performed the following procedure 20 times for each of the 5 algorithms and each of the 90 problem instances.

Procedure repeated 20 times for each instance of $\boldsymbol{W}$ and each algorithms

Run the algorithm repeatedly in 3 minutes on Ultra SPARC2(400MHz) by changing the initial value of the decision variable $\boldsymbol{x}$. After 3 minutes computation, output the minimal value of the objective function reported.

The time interval 3 minutes is determined as the time period in which **alg.1** terminates at least 3 times on our machine and implementation.

Figures 5.1(U_MC), 5.2(W_MC), and 5.3(N_MC) show the average of the outputs of 20 times repeated computations. Figures 5.4(U_MC), 5.5(W_MC), and 5.6(N_MC) show the best and worst case performance of the 20 outputs of **alg.1** and the tabu search. The vertical axis in each graph represents the value of the objective function and the horizontal axis represents the indices($0 \sim 29$) of the problem instances.

5.3.1 Performance in Fixed Time Interval

In figures 5.1, 5.2 and 5.3, the thin solid line indicates **alg.0** which is equivalent to the case $M = 1$ in **alg.2**, and also equivalent to the case initial solution is real in **alg.1**. The thin broken line represents the tabu search. The thick broken and dotted lines represent **alg.2** with $M = 2$ and $M = 3$, respectively. The thick solid line represents **alg.1**.

In figures 5.4, 5.5 and 5.6, the thick lines indicate **alg.1** and the thin lines indicate the tabu search. For each pair of lines, the upper, middle, and lower lines indicate the best, average(broken line depicted also in Figs. 5.1, 5.2 and 5.3), and worst case of outputs, respectively.

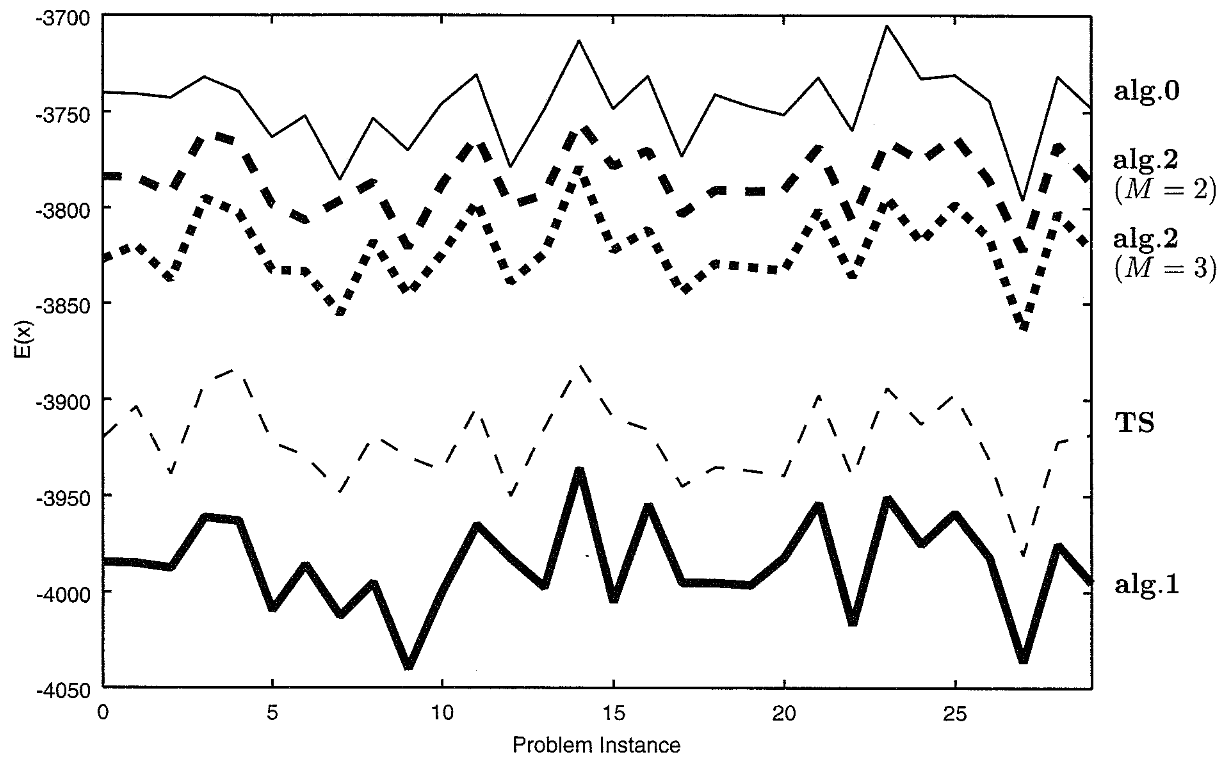

Figure 5.1: average performance : $w_{ij} \in \{0, 1\}$ (U_MC)

For almost all problem instances and criteria such as average case, the best case and the worst case, **alg.1** reported the best performance among the algorithms used. The worst case performance of **alg.1** nearly equals to the average performance of the tabu search.

By comparing **alg.0** with the case $M = 3$ of **alg.2**, we can see that the performance of the algorithms in same framework grew as the dimension of hypercomplex numbers became higher. Here, note that the outputs for evaluation are not about each termination of the algorithms, but about the fixed time interval. Therefore, the growth of time complexity caused by making M larger has been already taken into account as the number of terminations of the algorithms in 3 minutes.

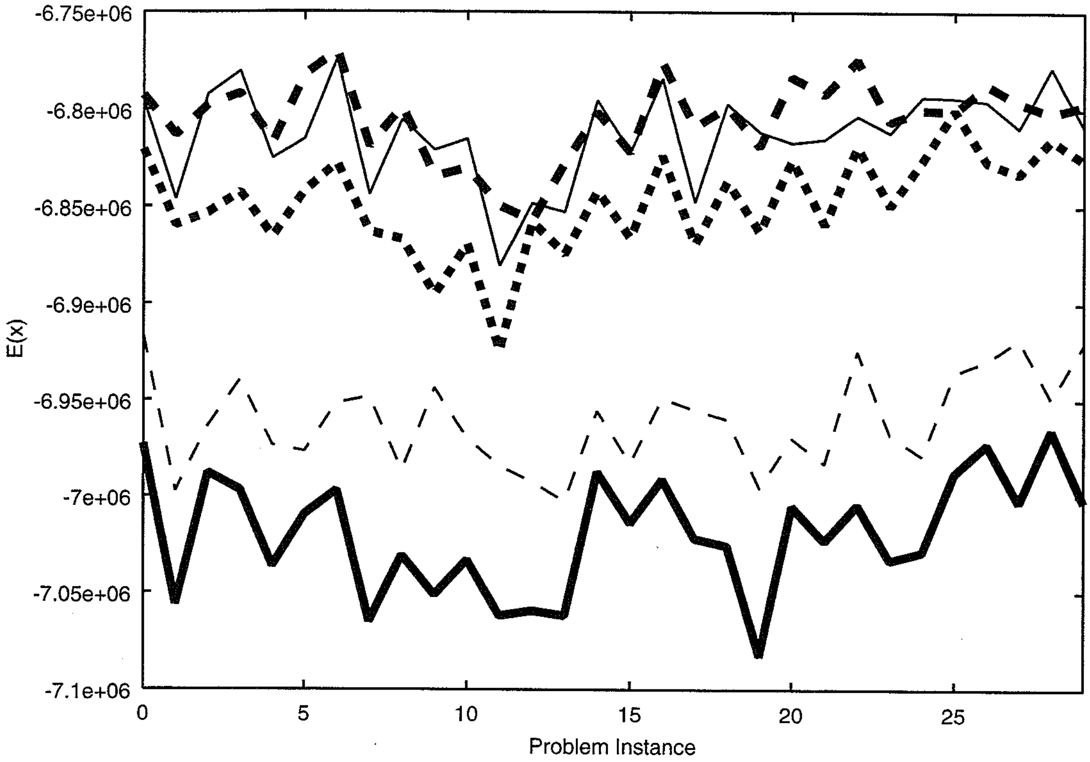

Figure 5.2: average performance : $w_{ij} \geq 0$(W_MC)

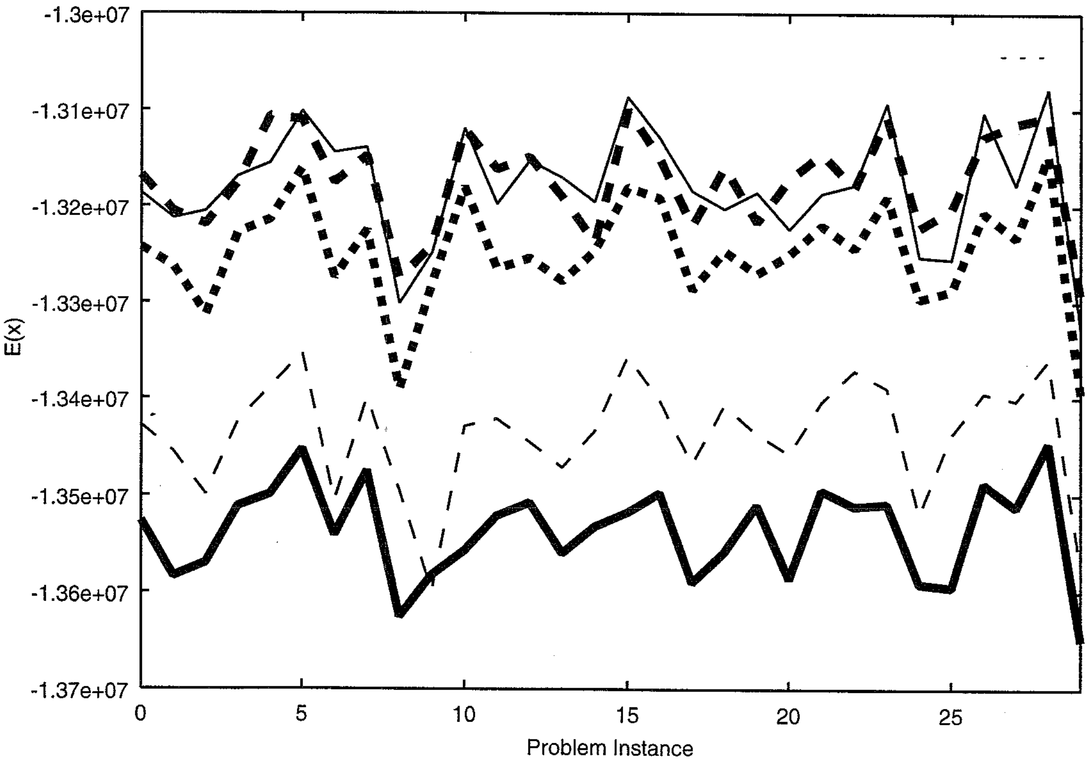

Figure 5.3: average performance : $w_{ij} \in \mathbb{R}$(N_MC)

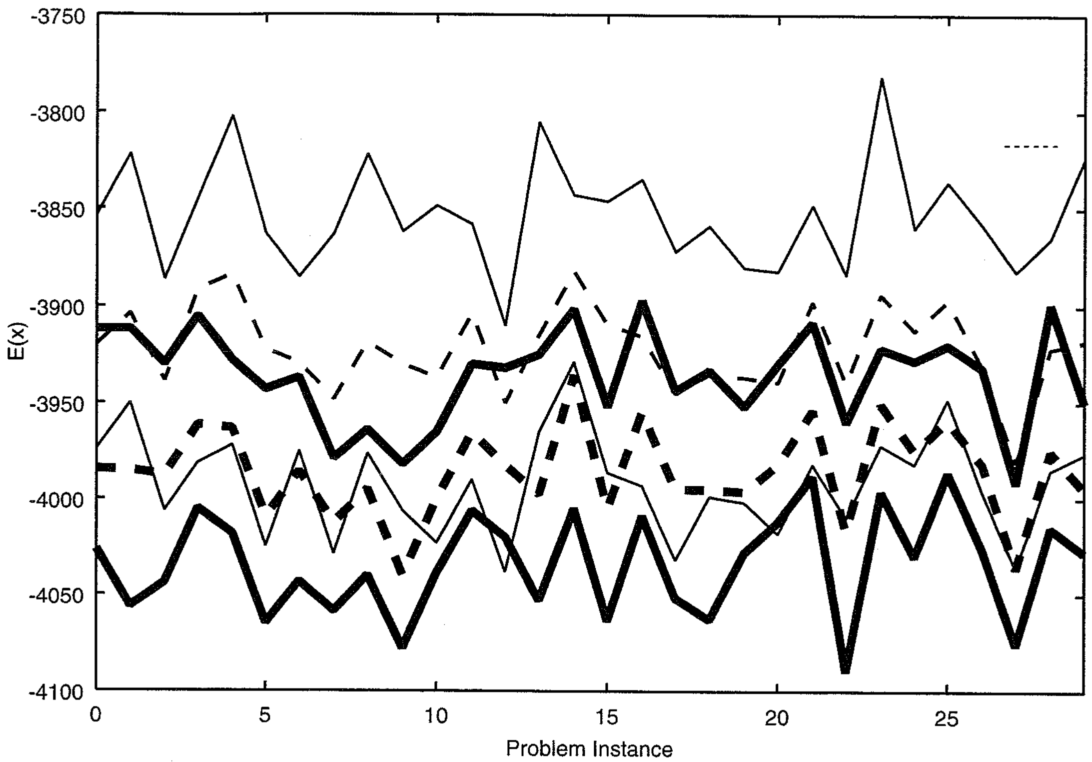

Figure 5.4: worst and best cases of tabu search and **alg.1** : $w_{ij} \in \{0, 1\}$ (U_MC)

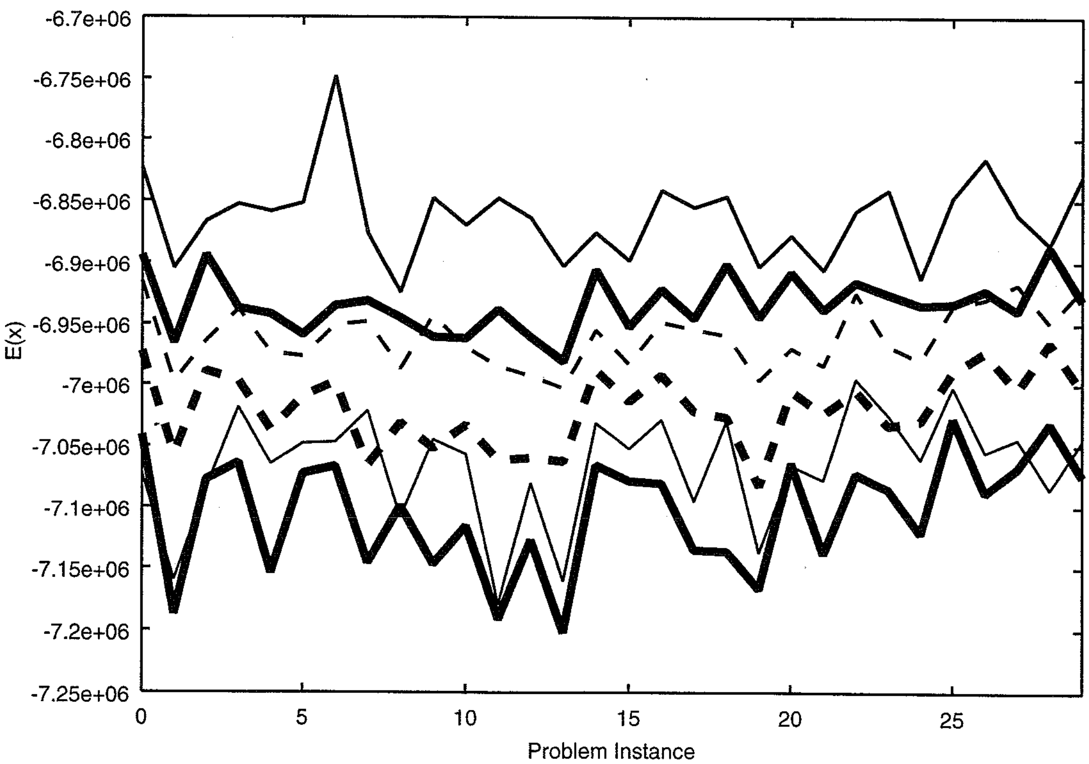

Figure 5.5: worst and best cases of tabu search and **alg.1** : $w_{ij} \geq 0$(W_MC)

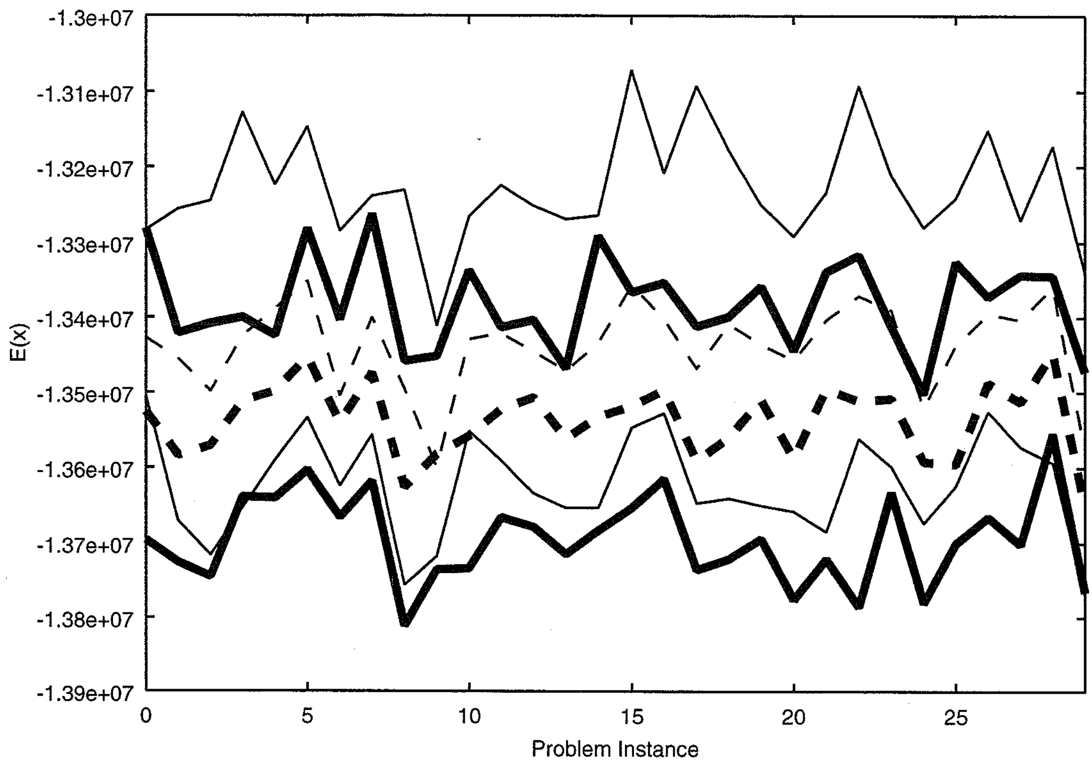

Figure 5.6: worst and best cases of tabu search and **alg.1** : $w_{ij} \in \mathbb{R}$(N_MC)

5.3.2 Performance in Each Termination of Algorithms

Figures 5.7, 5.8, and 5.9 are the average about each termination of **alg.2**, that is, not about the fixed time interval. Each of the upper, middle and lower lines indicates the case $M = 1$, $M = 2$ and $M = 3$, respectively. As we can see, the feasible solution output by one time termination became better as the dimension of hypercomplex numbers became higher.

5.3.3 Discussions

Algorithms evaluated in this chapter are prototypes of optimization algorithms which take advantage of the approach of nonconvex relaxation to the space consists of hypercomplex numbers.

In fact, there are many ways for evaluating the performance of algorithms. Moreover, our experiments are not about the general class **QP** but about $\mathbf{P}_1$. Nevertheless, the significance of these experimental results is on the following points.

- We applied a new approach that utilizes the nonconvex relaxation problems. In addition, the algorithms can be implemented.
- The effects of expanding the space of decision variables were observed through experiments.

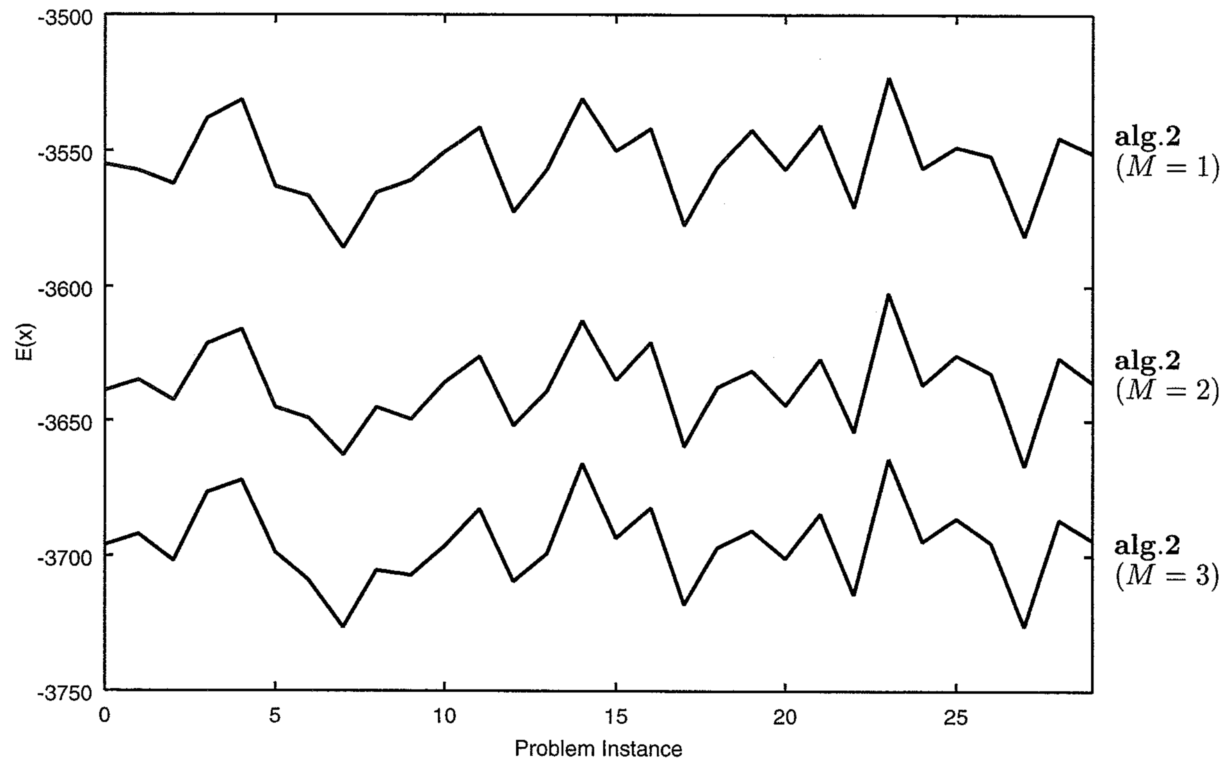

Figure 5.7: average per iteration : $w_{ij} \in \{0, 1\}$ (U_MC)

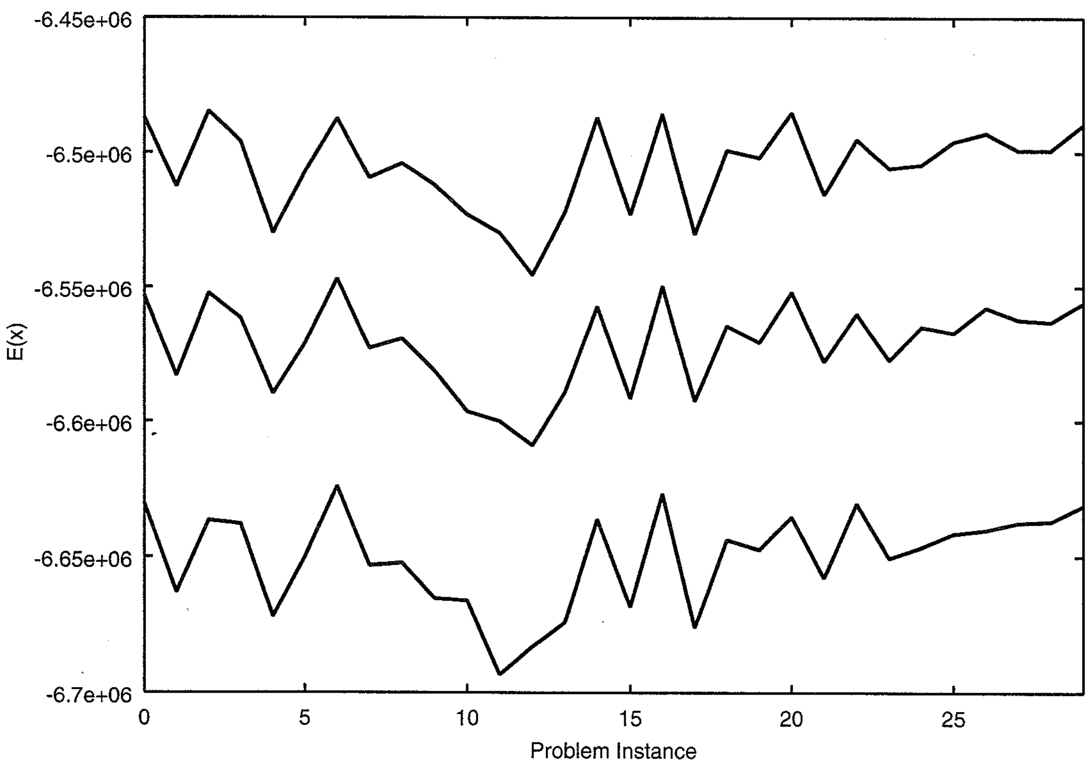

Figure 5.8: average per iteration : $w_{ij} \geq 0$(W_MC)

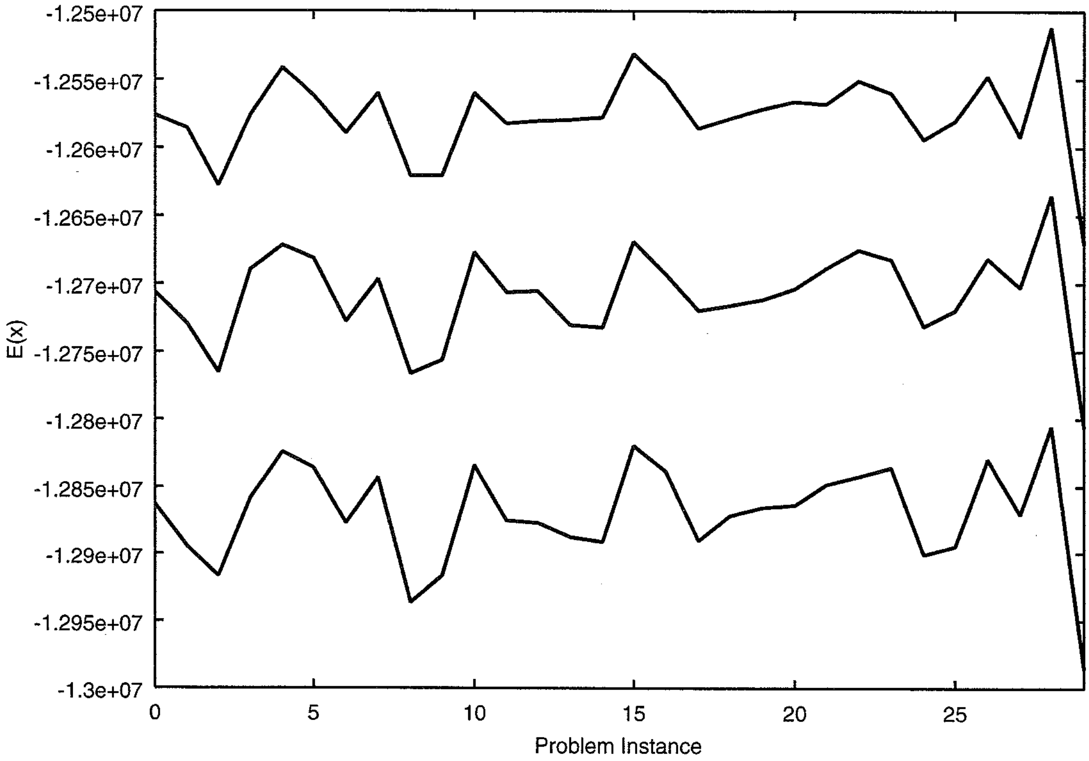

Figure 5.9: average per iteration : $w_{ij} \in \mathbb{R}$(N_MC)

- In our criteria, the performance of one of the proposed algorithms was better than a well-known algorithm which is considered to be one of the best strategies.

alg.1 utilizes the existence of a monotone path guaranteed in $\mathbf{P}_2$ and is also based on some conjectures or hopes. In our experiments, **alg.1** reported the performance better than the tabu search. The time interval three minutes may be advantageous to **alg.1**, for example, in the following sense.

> **alg.1** hardly terminates within one minute. It means that **alg.1** hardly output any acceptable feasible solution within one minute. On the other hand, the other algorithms output some feasible solution within one minutes for almost all cases.

However, the observation does not mean that the results of experiments have no sense even if we refer **alg.1** to be an algorithm which fully utilizes the all properties of our relaxation problem $\mathbf{P}_2$ since there are situations for which criteria similar to in our experiments are suitable. Moreover, **alg.1**'s performance was better than an algorithm that is considered to be one of the most effective heuristics. In addition, the performance of the proposed algorithms can be enhanced through analyses of $\mathbf{QP}^{\epsilon}_{M}$.

The performance of **alg.2** became better as the dimension M of the hypercomplex number became higher. The results imply that the operation **Shift**$(\cdot)$ has essential effects in **alg.2**.

Therefore, applications of **Shift**$(\cdot)$ to other optimization algorithms are expected to result in enhancements of the performance of some optimization algorithms. For example, Goemans and Williamson's approximation algorithm for the maximum cut problems is generalized through **Shift**$(\cdot)$ as we have seen in section 4.2.

Chapter 6

Conclusion

In this thesis, we considered nonconvex relaxation problems for quadratically constrained quadratic minimization problems, especially, for the maximum cut problems. The relaxation problems have the original problem and the SDP relaxation problem as two extreme cases. Moreover they lead us to consideration of trade-off between similarity to the original problem and desirable properties in numerical computation.

Based mainly on the existence of a monotone path, we analyzed properties embedded in the nonconvex relaxation problems. We showed that some nice properties are obtained without relaxing problem to a convex optimization problem. For example,

- The relaxation problem for the ordinary complex numbers ensures the existence of a monotone path for any instance of problem data and any feasible solutions of the original problem, which have distinct objective values.
- Based on the analytical results around the monotone path, we can generalize some optimization algorithms. The algorithms includes Goemans and Williamson's approximation algorithm for the maximum cut problems.
- Each local optimum of the complex valued relaxation problem gives a lower bound of the optimal value of the original problem for some classes of problem instances.

We also proposed two optimization algorithms using the nonconvex formulation of relaxation problems. Additionally, preliminary experiments implied availability of the algorithms and our approach.

Constructing optimization algorithms based on nonconvex relaxation problems for general classes of **QP** was not described in thesis, since the class **QP** is too large to construct optimization algorithms by applying our current results. However, further analyses in our approach will lead us to such algorithms.

The following problems have not been solved in this thesis.

- whether any local optima of $\mathbf{QP}_2$ give lower bound for **QP**
- whether the approximation rate of the modified version of Geomans and Williamson's approximation algorithm for the maximum cut problems is strictly greater than $0.878\cdots$

Additionally , there will be some critical degrees of relaxation in some sense other than the existence of a monotone path. For example,

- the degree where the relaxation problem becomes equivalent to the convex optimization problem
- the degree where any local optimum gives a lower bound for the original problem
- the degree where the performance of an optimization algorithm is maximized

Mathematical and numerical studies to reveal such properties will make our approach a set of new tools for constructing optimization algorithms.

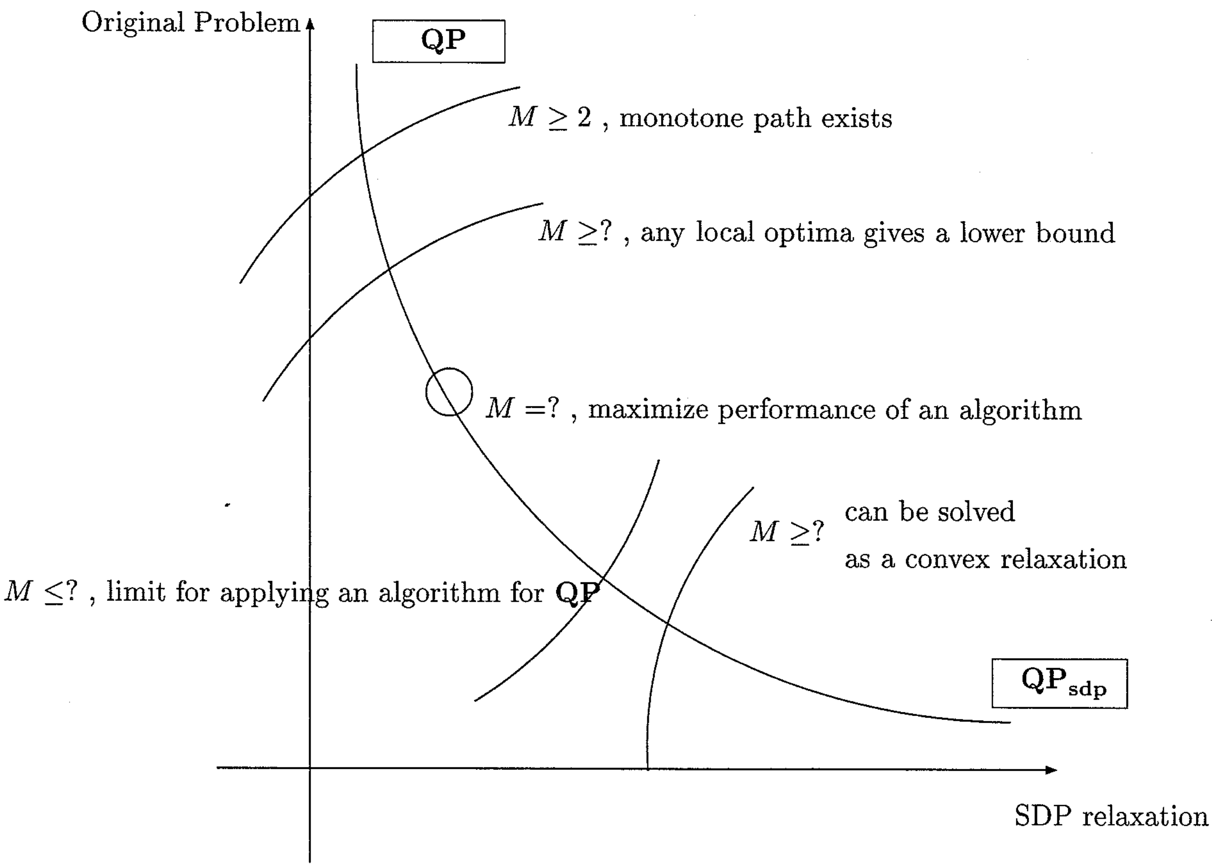

Figure 6.1: Open Problems

Bibliography

[1] N. Karmarker : "A New Polynomial-Time Algorithm for Linear Programming", *Combinatorica*, vol.4, pp.373-395(1984)

[2] 今野浩 : "線形計画法", 日科技連, (1987)

[3] Yu. Nesterov and A. Nemirovskii : "Interior-Point Polynomial Algorithms in Convex Programming", *SIAM Studies in Applied Mathematics*, SIAM (1994)

[4] W. F. Alizadeh : "Interior Point Methods in Semidefinite Programming with Application to Combinatorial Optimization", *SIAM Journal on Optimization*, Vol. 5, pp.13-51 (1995)

[5] L. Vandenberghe and S. Boyd : "Semidefinite Programming", *SIAM Review*, Vol.38, pp.49-95 (1996)

[6] 室田一雄 : "離散凸解析", 共立出版 (2001)

[7] C. Delorme and S. Poljak : "Laplacian Eigenvalues and the Maximum Cut Problem", *Mathematical Programming*, Vol.62, pp.557-574 (1993)

[8] M. Sipser 著 (渡辺治 鑑訳) : "Introduction to the Theory of Computation(計算理論の基礎)", 共立出版 (2000)

[9] 柳浦睦憲・茨木俊秀 : "組合せ最適化 — メタ戦略を中心として —", 朝倉書店 (2001)

[10] 三宮信夫・喜多一・玉置久・岩本貴司 : "遺伝アルゴリズムと最適化", 朝倉書店 (1998)

[11] S. Kirkpatrick and C. D. Gelatt and M. P. Vecchi : "Optimization by Simulated annealing", *Science*, Vol.220, pp.671-680 (1983)

[12] F. Glover : "Tabu Search I", *ORSA Journal on Computing*, vol.1, pp.190-206 (1989)

[13] F. Glover : "Tabu Search II", *ORSA Journal on Computing*, vol.2, pp.4-32 (1989)

[14] 藤沢克樹・久保幹夫・森戸晋 : "Tabu Search のグラフ分割問題への適用と実験的解析", 電気学会誌, Vol.114-C(4), pp.430-437 (1994)

[15] D. E. Goldberg : "Genetic Algorithms in Search, Optimization, and Machine Learning", *Addison-Wesley* (1998)

[16] M. X. Goemans and D. P. Williamson : "Improved Approximation Algorithms for Maximum Cut and Satisfiability Problems Using Semidefinite Programming", *Journal of the ACM*, Vol.42, pp.1115-1145 (1995)

[17] M. X. Goemans and D. P. Williamson : "Approximation Algorithms for Max-3-Cut and Other Problems via Complex Semidefinite Programming", *Proceedings of 33rd STOC*, pp.443-452 (2001)

[18] A. Frieze and M. Jerrum : "Improved Approximation Algorithms for MAX k-CUT and MAX BISECTION", *Algorithmica*, Vol.18, pp.67-81 (1997)

[19] Y. Ye : "Approximating Quadratic Programming with Bound and Quadratic Constraints", Working paper, department of management sciences, The University of Iowa (1997)

[20] D. Karger and R. Motwani and M. Sudan : "Approximate Graph Coloring by Semidefinite Programming" , *Journal of ACM*, vol.45, pp.246-265 (1998)

[21] Y. Yajima and T. Fujie : "A Polyhedral Approach for Nonconvex Quadratic Programming Problems with Box Constraints", Dept. of Math. and Comp. Sci. B-323, Tokyo Institute of Technology (1996)

[22] H. Fujioka and K. Hoshijima : "Bounds for the BMI Eigenvalue Problem", *Transactions of SICE* Vol.33, pp.616-621 (1997)

[23] M. Fukuda and M. Kojima : "Branch-and-Cut Algorithms for the Bilinear Matrix Inequality Eigenvalue Problem", *Computational Optimization and Application*, Vol.19, pp.79-105 (2001)

[24] I. L. Kantor and A. S. Solodovnikov : "Hypercomplex Numbers — An Elementary Introduction to Algebras — ", Springer-Verlag (1989)

[25] I. L. Kantor and A. S. Solodovnikov(浅野 監訳) : "超複素数入門— 多元環へのアプローチ, — ", 森北出版 (1999)

[26] A. J. Noest : "A Associative Memory in Sparse Phasor Neural Networks", *Europhisics Letters*, Vol.6, pp.469-474 (1988)

[27] M. Agu and K. Yamanaka and H. Takahashi : "A Local Property of the Phasor Model of Neural Networks", *IEICE Transactions on Information and Systems*, Vol.E79, pp.1209-1211 (1996)

[28] S. Jankowski and A. Lozowski and J. Zurada : Complex-Valued Multistate Neural Associative Memory, *IEEE Transactions on Neural Networks*, Vol.7, pp.1491-1496 (1996)

[29] T. Miyajima and K. Yamanaka : "An Application of a Phasor Model with Resting States to Multiuser Detection", *Knowledge-based Intelligent Information Engineering Systems & Allied Technologies* , Part I, pp.571-575(2001)

[30] J. J. Hamilton : "Hypercomplex Numbers and the Description of Sipn States", *Journal of Mathematical Physics*, Vol.38, pp.4914-4928 (1997)

[31] F. Barahona and M. Grotschel and M. Junger and G. Reinelt : "An Application of Combinatorial Optimization to Statistical Physics and Circuit Layout Design", *Operations Research* Vol.36, No.3, pp.493-513 (1988)

[32] H. P. Williams : "Model Building in Mathematical Programming(4th edition)", John Wiley & Sons (1993)

[33] Q. Zhao and S. E. Karisch and F. Rendl and H. Wolkowicz : "Semidefinite Programming Relaxations for the Quadratic Assignment Problem", *Journal of Combinatorial Optimization*, vol.2, pp.71-110(1998)

[34] C. A. Floudas and V. Visweswaran : "Primal-Relaxed Dual Global Optimization Approach", *Journal of Optimization Theory and Applications*, Vol.78, pp.187-225 (1993)

[35] 森耕平・原辰次："探索空間拡張による二次形式の 0-1 最適化", システム制御情報学会論文誌, Vol.14, No.1 (2001)

[36] 森耕平・原辰次："二次制約非凸最小化問題への超複素数による非凸緩和 — その性質と最適化アルゴリズム", システム制御情報学会論文誌, Vol.14, No.10 (2001)

[37] 藤沢克樹："組合せ最適化問題に対する近似解法 — 半正定値緩和とメタヒューリスティックを中心として —", "http://is-mj.archi.kyoto-u.ac.jp/ fujisawa/research.html", *The 8th RAMP Symposium*, 日本オペレーションズ・リサーチ学会 (1996)

[38] 今野浩・山下浩："非線形計画法", 日科技連, (1978)

[39] T. Matsui : "NP-hardness of Linear Multiplicative Programming and Related Problems", *Journal of Global Optimization*, Vol.9, pp.113-119 (1996)

[40] J. Jahn : "Introduction to the Theory of Nonlinear Optimization(2nd Edition)", Springer (1996)

[41] H. Konno and T. Kuno : "Linear Mulipricative Programming", *Mathematical Programming*, Vol. 56, pp.51-64, (1992)

[42] H. Konno and P. T. Thach and H Tuy : "Optimization on Low Rank Nonconvex Structure", Kluwer (1997)

[43] "特集・大域最適化", オペレーションズ・リサーチ, Vol.44, pp.226-252 (1999)

[44] P. M. Pardalos and G. P. Rodgers : "Computational Aspects of a Branch and Bound Algorithm for Quadratic Zero-One Programming", *Computing*, Vol. 45, pp.131-144 (1990),

[45] Y. Nesterov: "Semidefinite Relaxation and Non-Convex Quadratic Optimization", *Optimization Methods and Software* Vol.12, pp.1-20 (1997)

[46] A. Bemporad and M. Morari : "Control of Systems Integrating Logic, Dynamics, and Constraints", *Automatica*, Vol.35, pp.407-427 (1999)

[47] A. Bemporad and F. Borrelli and M. Morari : "Optimal Controllers for Hybrid Systems : Stability and Piecewise Linear Explicit Form", *Proc. 39th IEEE Conference on Decision and Control*, vol.2, pp.1810-1815 (2000)

[48] R. Horst and H. Tuy : "Global Optimization", *Springer-Verlag* (1993)

[49] `GCC home page ; http://gcc.gnu.org/`

Appendix A

Hypercomplex Numbers

In this appendix, we summarize the definition of hypercomplex numbers based on [24]. Hypercomplex numbers are generalization of the ordinary complex numbers. Complex numbers, quaternions, and so on are special cases of hypercomplex numbers.

A.1 Definition

Let M be a fixed natural number and a_is be real numbers. We call symbols $\boldsymbol{j}_i$s imaginary units. Consider a set $\mathbb{F}_M$ of numbers represented as

$$a_0 + a_1\boldsymbol{j}_1 + a_2\boldsymbol{j}_2 + \cdots + a_M\boldsymbol{j}_{M-1}.$$

For $\boldsymbol{a}$ and $\boldsymbol{b}$ in $\mathbb{F}_M$, equality $\boldsymbol{a} = \boldsymbol{b}$ holds if and only if

$$a_i = b_i \ \text{ for all } \ i = 0, \cdots M-1.$$

We define addition $\boldsymbol{a} + \boldsymbol{b}$ and subtraction $\boldsymbol{a} - \boldsymbol{b}$ by

$$\begin{aligned} \boldsymbol{a} + \boldsymbol{b} &= a_0 + b_0 + \sum_{i=1}^{M-1} (a_i + b_i)\boldsymbol{j}_i \\ \boldsymbol{a} - \boldsymbol{b} &= a_0 - b_0 + \sum_{i=1}^{M-1} (a_i - b_i)\boldsymbol{j}_i. \end{aligned}$$

Multiplication for elements of $\mathbb{F}_M$ is defined as follows.

Let $p_{\alpha,\beta,i}$ be real constants, and define multiplication of imaginary unit by

$$\boldsymbol{j}_\alpha \boldsymbol{j}_\beta = p_{\alpha,\beta,0} + \sum_{i=1}^{M-1} p_{\alpha,\beta,i}\boldsymbol{j}_i. \tag{A.1}$$

Then, multiplication of $\boldsymbol{a} \in \mathbb{F}_M$ and $\boldsymbol{b} \in \mathbb{F}_M$ is defined by substituting eq.(A.1) into

$$\begin{aligned}\boldsymbol{ab} \;=\; & a_0b_0 + a_0b_1\boldsymbol{j}_1 + a_0b_2\boldsymbol{j}_2 + \cdots + a_0b_{M-1}\boldsymbol{j}_{M-1} \\ & + a_1b_0\boldsymbol{j}_1 + a_1b_1\boldsymbol{j}_1\boldsymbol{j}_1 + a_1b_2\boldsymbol{j}_1\boldsymbol{j}_2 + \cdots + a_1b_{M-1}\boldsymbol{j}_1\boldsymbol{j}_{M-1} \\ & + a_2b_0\boldsymbol{j}_2 + a_2b_1\boldsymbol{j}_1\boldsymbol{j}_2 + a_2b_2\boldsymbol{j}_2\boldsymbol{j}_2 + \cdots + a_2b_{M-1}\boldsymbol{j}_2\boldsymbol{j}_{M-1} \\ & \vdots \\ & + a_{M-1}b_0\boldsymbol{j}_{M-1} + a_{M-1}b_1\boldsymbol{j}_{M-1}\boldsymbol{j}_1 + a_{M-1}b_2\boldsymbol{j}_{M-1}\boldsymbol{j}_2 + \cdots \\ & \cdots + a_{M-1}b_{M-1}\boldsymbol{j}_{M-1}\boldsymbol{j}_{M-1}.\end{aligned} \tag{A.2}$$

Together with the above rule of arithmetic operations, we call an element of $\mathbb{F}_M$ a hypercomplex number. For example, an ordinary complex number is a special case of a hypercomplex number where $M = 2$, $p_{1,1,0} = -1$, and $p_{1,1,1} = 0$. Quaternion is defined as a special case where $M = 4$, and $p_{\alpha,\beta,i}$s are defined by

$$\begin{pmatrix} p_{1,1,0} & p_{1,1,1} & p_{1,1,2} & p_{1,1,3} \\ p_{1,2,0} & p_{1,2,1} & p_{1,2,2} & p_{1,2,3} \\ p_{1,3,0} & p_{1,3,1} & p_{1,3,2} & p_{1,3,3} \\ p_{2,1,0} & p_{2,1,1} & p_{2,1,2} & p_{2,1,3} \\ p_{2,2,0} & p_{2,2,1} & p_{2,2,2} & p_{2,2,3} \\ p_{2,3,0} & p_{2,3,1} & p_{2,3,2} & p_{2,3,3} \\ p_{3,1,0} & p_{3,1,1} & p_{3,1,2} & p_{3,1,3} \\ p_{3,2,0} & p_{3,2,1} & p_{3,2,2} & p_{3,2,3} \\ p_{3,3,0} & p_{3,3,1} & p_{3,3,2} & p_{3,3,3} \end{pmatrix} = \begin{pmatrix} -1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \\ -1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & -1 & 0 & 0 \\ -1 & 0 & 0 & 0 \end{pmatrix}.$$

A.2 Fundamental Properties

For any definitions of $p_{\alpha,\beta,i}$s, the following properties holds.

Proposition A.1 *Consider a case where $\boldsymbol{a} \in \mathbb{F}_M$ and $\boldsymbol{b} \in \mathbb{F}_M$ are represented by*

$$\boldsymbol{a} = a_0 + 0\boldsymbol{j}_1 + 0\boldsymbol{j}_2 + \cdots + 0\boldsymbol{j}_{M-1}$$

$$\boldsymbol{b} = b_0 + b_1 0\boldsymbol{j}_1 + b_2\boldsymbol{j}_2 + \cdots + b_{M-1}\boldsymbol{j}_{M-1}.$$

Then,

$$\boldsymbol{ab} = \boldsymbol{ba} = a_0b_0 + a_0b_1\boldsymbol{j}_1 + a_0b_2\boldsymbol{j}_2 + \cdots + a_0b_{M-1}\boldsymbol{j}_{M-1}$$

holds.

Proof : Immediately from the definition of multiplication. ■

Corollary A.1 *Multiplication of a hypercomplex number and a real number is commutative.*

Corollary A.2 *The real number* 1 *is an identity of multiplication.*

Proposition A.2 *Let* $\boldsymbol{a} \in \mathbb{F}_M$, $\boldsymbol{b} \in \mathbb{F}_M$, $\alpha \in \mathbb{R}$ *and* $\beta \in \mathbb{R}$. *Then, we have*

$$(\alpha\boldsymbol{a})(\beta\boldsymbol{b}) = (\alpha\beta)(\boldsymbol{a}\boldsymbol{b})$$

Proof : From corollary A.1 and the definition of multiplication. ■

Proposition A.3 *Let* $\boldsymbol{a} \in \mathbb{F}_M$, $\boldsymbol{b} \in \mathbb{F}_M$ *and* $\boldsymbol{c} \in \mathbb{F}_M$. *Then we have*

1. $\boldsymbol{a}(\boldsymbol{b}+\boldsymbol{c}) = \boldsymbol{a}\boldsymbol{b}+\boldsymbol{a}\boldsymbol{c}$
2. $(\boldsymbol{a}+\boldsymbol{b})\boldsymbol{c} = \boldsymbol{a}\boldsymbol{c}+\boldsymbol{b}\boldsymbol{c}$

Proof : From the definitions of addition and multiplication. ■

Proposition A.4 *Multiplication of hypercomplex numbers is associative if and only if* $p_{\alpha,\beta,i}$*s satisfy*

$$(\boldsymbol{j}_\alpha\boldsymbol{j}_\beta)\boldsymbol{j}_\gamma = \boldsymbol{j}_\alpha(\boldsymbol{j}_\beta\boldsymbol{j}_\gamma) \quad \textit{for all} \quad \alpha, \beta, \quad \textit{and} \quad \gamma. \tag{A.3}$$

Proof : If eq.(A.3) is satisfied, then multiplication becomes associative immediately from the definition of multiplication. If eq.(A.3) does not hold, then we have

$$(\boldsymbol{a}\boldsymbol{b})\boldsymbol{c} = (\boldsymbol{j}_1\boldsymbol{j}_2)\boldsymbol{j}_3 \neq \boldsymbol{j}_1(\boldsymbol{j}_2\boldsymbol{j}_3) = \boldsymbol{a}(\boldsymbol{b}\boldsymbol{c})$$

for the case

$$\boldsymbol{a} = \boldsymbol{j}_1 \ , \ \boldsymbol{b} = \boldsymbol{j}_2 \ , \ \boldsymbol{c} = \boldsymbol{j}_3.$$

Therefore, multiplication of hypercomplex numbers whose imaginary units do not satisfy eq.(A.3) is not associative. ■

In some definition of hypercomplex numbers, associativity of multiplication is dealt as a part of definition of hypercomplex numbers. But in this thesis and [24], associativity does not contained in the definition.

Proposition A.5 *Multiplication of hypercomplex numbers is commutative if and only if*

$$p_{\alpha,\beta,i} = p_{\beta,\alpha,i} \quad \textit{for all} \quad \alpha, \beta \quad \textit{and} \quad i. \tag{A.4}$$

Proof : If eq.(A.4) does not hold, then for $\boldsymbol{a} = \boldsymbol{j}_\alpha$ and $\boldsymbol{b} = \boldsymbol{j}_\beta$,

$$\boldsymbol{a}\boldsymbol{b} \neq \boldsymbol{b}\boldsymbol{a}$$

for some α and β from eq.(A.1).

If eq.(A.4) holds, then eq.(A.1) holds. Substituting this fact into eq.(A.2) yields $\boldsymbol{ab} = \boldsymbol{ba}$ for any $\boldsymbol{a} \in \mathbb{F}_M$ and $\boldsymbol{b} \in \mathbb{F}_M$. ■

For example, multiplication of hypercomplex numbers is commutative if

$$\boldsymbol{j}_\alpha \boldsymbol{j}_\beta = \begin{cases} 0 & \text{if } \alpha \neq \beta \\ -1 & \text{if } \alpha = \beta \end{cases} .$$

This definition of multiplication is consistent with the definition of hypercomplex number in chapter 2. In thesis, we described M dimensional hypercomplex number

$$a_0 + a_1 \boldsymbol{j}_1 + a_2 \boldsymbol{j}_2 + \cdots + a_{M-1} \boldsymbol{j}_{M-1}$$

as

$$(a_0, a_1, a_2, \cdots, a_{M-1})$$

for simplicity.

Finally, note that

- Commutativity and associativity of multiplication do not hold in general.
- Division cannot be defined except for cases $M = 1, 2, 4, 8$. In other words, for $\boldsymbol{u} \in \mathbb{F}_M$ and $\boldsymbol{v}(\boldsymbol{v} \neq 0) \in \mathbb{F}_M$, the following simultaneous equations about $\boldsymbol{x}$ cannot have unique solution for $M \neq 1, 2, 4, 8$ and for some $\boldsymbol{v}$s and $\boldsymbol{u}$s.

$$\left(\begin{array}{l} \boldsymbol{vx} = \boldsymbol{u} \\ \boldsymbol{xv} = \boldsymbol{u} \end{array} \right.$$