Graph Modeling for Quadratic Assignment Problems Associated with the Hypercube

Hans Mittelmann*, Jiming Peng† and Xiaolin Wu**

*Department of Mathematics, Arizona State University, Tempe, AZ 85287-1804, USA
†Department of Industrial and Enterprise System Engineering , University of Illinois at Urbana-Champaign. Urbana, IL, 61801
**Department of Electrical & Computer Engineering, McMaster University, Ontario, Canada

Abstract. In the paper we consider the quadratic assignment problem arising from channel coding in communications where one coefficient matrix is the adjacency matrix of a hypercube in a finite dimensional space. By using the geometric structure of the hypercube, we first show that there exist at least n different optimal solutions to the underlying QAPs. Moreover, the inherent symmetries in the associated hypercube allow us to obtain partial information regarding the optimal solutions and thus shrink the search space and improve all the existing QAP solvers for the underlying QAPs.

Secondly, we use graph modeling technique to derive a new integer linear program (ILP) models for the underlying QAPs. The new ILP model has $n(n-1)$ binary variables and $O(n^3\log(n))$ linear constraints. This yields the smallest known number of binary variables for the ILP reformulation of QAPs. Various relaxations of the new ILP model are obtained based on the graphical characterization of the hypercube, and the lower bounds provided by the LP relaxations of the new model are analyzed and compared with what provided by several classical LP relaxations of QAPs in the literature.

Keywords: Quadratic Assignment Problem (QAP),Integer Linear Program (ILP), Graph Modeling, Relaxation, Lower Bound
PACS: 02.60.Pn

1. INTRODUCTION

The standard quadratic assignment problem takes the following form

$$\min_{X\in\Pi} \mathrm{Tr}\big(\mathrm{XAX^T B}\big) \tag{1}$$

where $A,B\in\Re^{n\times n}$, and Π is the set of permutation matrices. This problem was first introduced by Koopmans and Beckmann [27] for facility location. The model covers many scenarios arising from various applications such as in chip design [22], image processing [38], and keyboard design [9]. For more applications of QAPs, we refer to the survey paper [10] where many interesting QAPs from numerous fields are listed.

In this work we focus on a special class of QAPs where the matrix B is the adjacency matrix of a hypercube in space $\Re^d$ for a positive integer d. The problem arises from channel coding in communication where the purpose is to minimize the total channel distortion caused by the channel noise [6, 33]. Let us consider a given source alphabet (*binary codebook*) of fixed length (d)

$$\mathscr{C}=\{c_1,\cdots,c_n\}, n=z^d.$$

Since the the communication system is imperfect, there is a certain probability that the received message is different from the transmitted message. If we use the so-called memoryless binary-symmetric channel (BSC), the conditional probability of decoding c_j when transmitting c_i is

$$p(j/i)=q^{\delta(c_i,c_j)}(1-q)^{d-\delta(c_i,c_j)},$$

where q is the error bit rate, and $\delta(c_i,c_j)$ is the Hamming distance between c_i and c_j. Suppose that the codeword c_i occurs with a probability P_i and the assignment of a binary code to each codeword in $\mathscr{C}$ is obtained by a permutation $i\to\pi(i), i\in\{1,\cdots,n\}$. The problem of minimizing the channel distortion is defined by

$$\min_{\pi\in\Pi}\sum_{i=1}^{n} P_i \sum_{j=1}^{n} p(\pi(j)/\pi(i))\delta(c_i,c_j). \tag{2}$$

One can easily see that the above problem can be cast as a special case of problem (1) where the matrix B is the Hamming distance matrix, and A is the probability matrix. When the error bit rate is very small, the distortion corresponding to a large Hamming distance can be ignored. In such a case we can approximate problem (2) by using a simplified matrix B, derived from the adjacency matrix of a hypercube in $\Re^d$.

The problem (2) has been studied by many experts in the communication community. For example, in [33], Potter and Chiang showed that the problem is NP-hard in general. However, if the probability matrix A has also a very special structure such as the case of Harper code [23], then it can be solved in polynomial time. Numerous algorithms for general QAPs have been applied to solve the problem. For more engineering background and solving techniques for problem (2) we refer to [6, 33].

Since most existing solvers for problem (2) are based on algorithms for general QAPs, in what follows we give a brief review on these algorithms. It is known that QAPs are among the hardest discrete optimization problems. For example, a QAP with $n = 30$ is typically recognized as a great challenge from a computational perspective. The paper [5] was probably the first one where the optimal solutions to QAPs of size $n = 30$ were reported. A popular technique for finding the exact solution of the QAPs is branch and bound (B&B), which has been employed in most existing methods for the QAPs. In a typical B&B approach, we need to solve a relaxation of the original QAP whose solutions can further provide a lower bound for QAPs and help us in the process. There are several relaxations for QAPs. One relaxation is the well-known Gilmore-Lawler bound (GLB) which can be found via solving a linear programming problem [18]. Though the GLB bound can be obtained very fast, the quality of the GLB bound deteriorates fast as n increases. Other bounds include the bounds based on eigenvalues of the coefficient matrices [15], bounds based on convex quadratic programming [4] and semidefinite programming [41]. It has been observed that all these bounds are tighter than the GLB, but the computational expense to obtain these bounds is also much higher than that of the GLB. This prevents the B&B algorithm based on more complex relaxations from solving large scale QAPs with reasonable effort.

In this paper, we are particularly interested in methods based on ILP reformulations of the original QAP. There are several different ILP approaches for the QAP in the literature based on linearizing the quadratic form of the objective function. For example, Lawler [28] introduced n^4 additional binary variables and reformulated the QAP as an ILP with $n^4 + n^2$ binary variables and $n^4 + 2n + 1$ constraints. Kaufman and Broeckx [26] introduced n^2 new variables and reformulate the original QAP as an ILP with $O(n^2)$ constraints. Most ILP approaches for the QAPs are closely related to the following mixed ILP formulation

$$\begin{array}{lll} \min & \sum_{i=1}^n \sum_{j=1}^n \sum_{k=1}^n \sum_{l=1}^n a_{ik} b_{jl} y_{ijkl} & \\ s.t. & \sum_{i=1}^n x_{ij} = 1, & \forall\, j \in \{1,\cdots,n\}; \\ & \sum_{j=1}^n x_{ij} = 1, & \forall\, i \in \{1,\cdots,n\}; \\ & x_{ij} \in \{0,1\}, & \forall i,j \in \{1,\cdots,n\}, \\ & \sum_{i=1}^n y_{ijkl} = x_{kl}, & \forall\, j,k,l \in \{1,\cdots,n\}; \\ & \sum_{j=1}^n y_{ijkl} = x_{kl}, & \forall\, i,k,l \in \{1,\cdots,n\}; \\ & y_{ijkl} = y_{klij}, y_{ijkl} \geq 0, & \forall i,j,k,l \in \{1,\cdots,n\}, \end{array} \tag{3}$$

which was proposed by Adams and Johnson [1] based on the linearization techniques introduced by Adams and Sherali [2, 3]. A similar model was proposed in [16] where the authors imposed the constraints $y_{ijij} = x_{ij}$. Noting the transformation $y_{ijkl} = x_{ij}x_{kl}$ where x is the original assignment matrix, Hahn and Grant [20] introduced extra constraints $y_{ijil} = 0, j \neq l$. In [21], the authors further lift model (3) to higher dimension, with n^6 variables based on the technique in [3] and proposed algorithms based on Lagrangian relaxation to attack the lifted mixed ILP in higher dimension. It has been reported that the lower bound provided by the level-2 reformulation model in [21] is the tightest [30].

Note that if we directly apply the above algorithms for general QAPs to the special case in this work, then we might not be able to explore the specific structure of the associate hypercube to improve the algorithm. We also note that the idea of exploring the structure of the problem has been employed in the study of QAPs. For example, it has been proved that several special classes of QAPs can be solved in polynomial time [7, 8]. In [25], Karisch and Rendl used the triangle decomposition for QAPs with rectangular grids to design an algorithm. In particular, the papers [17, 13] used the idea of graph modeling to derive lower bounds for QAPs.

The distinction between our ILP formulation and the existing ILP formulations of QAPs lies in the linearization process. While most existing ILP formulations of QAPs perform the linearization on the quadratic objective function $\mathrm{Tr}\left(\mathrm{AXBX^T}\right)$, we try to characterize the set X^TBX (or XAX^T) in a certain space, by using the structure of the matrices

A or B. As an initial step in our new approach, we also had a closer look at all the problems in the QAP library [10]. Not to our surprise, many problems in the QAP library have well-structured coefficient matrices. For example, one matrix in the problem series Chr# (introduced in [12]) is the adjacency matrix of a weighted tree. The matrix in problems Nug# [32], Scr# [36] and Sko# [37] is the Manhattan distance matrix of rectangular grids. It is reasonable to expect that if we use the special structure of the QAPs, more efficient models and algorithms could be developed.

One of the major contributions of this work is the existence of at least n different optimal solutions to the underlying QAPs. In particular, we show that for any fixed index pair (i, j), there exists a optimal permutation matrix X^* satisfying $x^*_{ij} = 1$. This observation can help us to restrict the search space and speed up all the existing QAP solvers when applied to the underlying problems in the present work. It should be pointed out that due to the discrete feature of constraint set involved in QAPs, the existence of multiple optimal solutions to general QAPs is not a big surprise. For example, in [31], Mautor and Roucairol observed that for several special classes of QAPs associated with some graphes, the symmetries in the graph can lead to the multiplicity of optimal solutions. Further, the authors of [31] used the symmetries in the associated graphes to improve the B&B approach.

The second major contribution is the introduction of a new ILP model for the index assignment problem based on the geometric feature of the hypercube. Several LP relaxations of the new model, derived from the graphical characterization of the hypercube are suggested. The lower bounds provided by these LP relaxations of the new model are analyzed and compared with what provided by several classical LP relaxations of QAPs.

The paper is organized as follows. In Section 2, we first describe the basic geometric structure of the hypercube. Then, we use the geometric structure of the hypercube to explore the symmetries in the solution set of the underlying QAP. In Section 3, we discuss how to characterize a hypercube via its adjacency matrix. In Section 4, we reformulate the original QAP as an equivalent ILP. Various LP relaxations of the new ILP model based on the graphical features of the hypercube are derived. In particular, we construct a simple LP with n^2 variables, which can be viewed as a relaxation of the original problem based, can provide a stronger bound than the LP relaxation of model (3). Based on such a observation and the graphical characterization of the hypercube, we then lift the simple LP model to higher dimension. In Section 4 we report some numerical experiments that show that our new model can provide a tighter lower bound than most existing lower bounds in the literature. In Section 5 we conclude our paper with some remarks.

2. NON-UNIQUENESS OF OPTIMAL SOLUTIONS TO QAPS ASSOCIATED WITH HYPERCUBE

In this section, we first describe several basic geometric properties of a hypercube in space $\Re^d$. Then we use these geometric properties of the hypercube to explore the symmetries in the solution set of the associated QAPs.

First we note that corresponding to every vertex v is a binary code c_v of length d. By the definition of the hypercube, the *degree* of every vertex in the hypercube is d, i.e., $deg(v_i) = d$. Let us also define the distance between two vertices v_i and v_j to be the Hamming distance between two binary codes corresponding to the two vertices, i.e.,

$$\delta(v_i, v_j) = \sum_{k=1}^{d} |c^k_{v_i} - c^k_{v_j}|.$$

We next present several technical results regarding the distance functions that will be used later on to reformulate the QAP associated with a hypercube as an ILP. The following result is an immediate consequence of the above definition of the distance function.

Lemma 2.1 *Suppose that v_1, v_2, v_3 are three different vertices of a hypercube. There exists a unique index $l \in \{0, 1, \cdots, \lfloor |\delta(v_1, v_2) - \delta(v_2, v_3)|/2 \rfloor\}$ such that*

$$\delta(v_1, v_3) = |\delta(v_1, v_2) - \delta(v_2, v_3)| + 2l.$$

In particular, if $\delta(v_1, v_2) = \delta(v_2, v_3) = 1$ or $\delta(v_1, v_2) = \delta(v_2, v_3) = d - 1$, then

$$\delta(v_1, v_3) = 2.$$

We next consider a vertex pair whose distance is k. We have

Lemma 2.2 *Suppose that v_i, v_j are two vertices of the hypercube satisfying $\delta(v_i, v_j) = k$ with $1 < k \leq d$, then there exist exactly k different vertices $v_{i_1}, v_{i_2}, \cdots, v_{i_k}$ such that*

$$\delta(v_i, v_{i_1}) = \delta(v_i, v_{i_2}) = \cdots = \delta(v_i, v_{i_k}) = k - 1,$$
$$\delta(v_j, v_{i_1}) = \delta(v_j, v_{i_2}) = \cdots = \delta(v_j, v_{i_k}) = 1.$$

Moreover, for every vertex v of the hypercube, we have

$$|S(v,l)| = C(d,l), \forall l = 1, \cdots, d,$$

where

$$S(v,l) = \{v' \in V : \delta(v, v') = l\},$$

$|S(.,.)|$ denotes the cardinality of the set, and $C(d,l)$ is the combinatorial function.

Proof: Without loss of generality, we can assume that v_i is the origin of the space $\Re^d$, while v_j is the vertex whose first k elements are 1 and the others are 0. Let us choose $v_{k_i} = (1, \cdots, 1, 0, 1, \cdots, 1, 0, \cdots, 0)^T$ to be the vector whose first k elements are 1 except its i-th element, which has value 0. One can easily verify that our choice satisfies the relations in the first conclusion of the lemma. The second conclusion is a direct consequence of the definition of the hypercube. q.e.d.

As a direct consequence of Lemma 2.2, we have

Corollary 2.3 *For any vertex v of the hypercube, there exists a unique vertex v' (called the complement of v) such that $\delta(v, v') = d$. Moreover, for any two vertex pair v_1, v_2, we have*

$$\delta(v_1, v_2) = \delta(v'_1, v'_2).$$

For convenience, we also introduce the complement of a code c.

Definition 2.4 *let c be a binary code of length d. We call the binary code c' the complement of c if the Hamming distance between c and c' equal d, i.e., every component in c' is generated by switching 0 to 1 and verse vice for the component of c at the same position.*

We now discuss the variety of the optimal solutions of the underlying QAP in this paper. We have

Theorem 2.5 *Suppose that B is the adjacency matrix of a hypercube in a suitable space. Then, for any coefficient matrix A, there exist at least n different optimal solutions to problem (1).*

Proof: Let us assume that X^* is an optimal solution of problem (1). Let us define $B^* = (X^*)^T B X^*$. Then problem (1) can be equivalently stated as

$$\min_{X \in \mathscr{X}} \mathrm{Tr}\big(\mathrm{AX^T B^* X}\big). \tag{4}$$

The optimality of X^* implies that the identify matrix I is an optimal solution of problem (4). Now let us recall the fact that B^* is the adjacency matrix of a hypercube, i.e., for any index i, j, there is a corresponding vertex pair v_i, v_j such that $b^*_{ij} = 1$ if and only if $\delta(v_i, v_j) = 1$. Now recall Corollary 2.3, we know that for every index i, there exists another index i' such that $\delta(v_i, v_{i'}) = d$. Moreover, by Corollary 2.3, for any vertex pair (v_i, v_j), we have $\delta(v_i, v_j) = \delta(v_{i'}, v_{j'})$. This implies, if we choose the values of the elements of P in (4) by the following rule

$$x_{ij} = \begin{cases} 1 & \text{if } j = i'; \\ 0 & \text{otherwise;} \end{cases}$$

then it must hold $X^T B^* X = B^*$.

Now let us recall the fact that for every vertex v, there is an associated binary code c_v of d digits. Let I be an index set $I \subseteq \{1, 2, \cdots, d\}$ and its complement $\bar{I}$ defined by $\bar{I} = \{1, 2, \cdots, d\} - I$. We can separate the code corresponding to every vertex into two parts c_v^I and $c_v^{\bar{I}}$ respects to the index sets I and $\bar{I}$, i.e.,

$$c_v = c_v^I \oplus c_v^{\bar{I}}.$$

It is straightforward to see that for any two vertices v_i, v_j, we have

$$\delta(v_i, v_j) = \delta(c^I_{v_i}, c^I_{v_j}) + \delta(c^{\bar{I}}_{v_i}, c^{\bar{I}}_{v_j}).$$

Next we introduce the following transformation

$$c_v = c^I_v \oplus c^{\bar{I}}_v \Longrightarrow c^I_{v'} \oplus c^{\bar{I}}_v, \tag{5}$$

where $c^I_{v'}$ is the complement of c^I_v. By applying Corollary 2.3, one can easily verify that for any vertex pair (v_i, v_j), the following relation holds

$$\delta(v_i, v_j) = \delta(c^I_{v_i}, c^I_{v_j}) + \delta(c^{\bar{I}}_{v_i}, c^{\bar{I}}_{v_j}) = \delta(c^I_{v'_i}, c^I_{v'_j}) + \delta(c^{\bar{I}}_{v_i}, c^{\bar{I}}_{v_j}).$$

Let X be the corresponding permutation matrix induced by the transformation (5). Our above discussion implies that

$$X^T B^* X = B^*.$$

If we perform the above permutation for every index set $I \subseteq \{1, \cdots, d\}$, one can easily see that we will obtain $2^d = n$ different permutation matrices such that

$$X^T B^* X^T = B^*.$$

This shows that problem (4) has at least n different optimal solutions. q.e.d.

We next discuss how to use the conclusions from Theorem 2.5 to restrict the search space of the underlying QAP.

Theorem 2.6 *Suppose that B is the adjacency matrix of a hypercube in a suitable space. Then, for any coefficient matrix A and any fixed index pair (i, j), there exists a global optimal solution to problem (1) that satisfies $x_{ij} = 1$.*

Proof: We first prove the special case with fixed index pair $i = j = 1$. Without loss of generality, we can assume that B is the adjacency matrix of the hypercube corresponding to the natural labeling, i.e., v_i corresponds to the binary coding of number $i-1$. Suppose that P^* is a global optimal solution of problem (1) and v is the first labeled vertex in the final solution. Let c_v be the binary code of vertex v and $I \subseteq \{1, \cdots, d\}$ the index set corresponding to the components of c_v with value 1. Now we can apply the transformation (5) to this particular index set I. Let us denote the corresponding induced permutation matrix by P. From (5), we can conclude that

$$v \Longrightarrow v_1,$$

which implies

$$[X^* X]_{11} = 1.$$

The case for general fixed index pairs follows similarly. This finishes the proof of the theorem. q.e.d.

We remark the fact that there exist multiple optimal solutions to QAPs is not surprising due to the discrete structure of the constrained set. However, to the best of the authors' knowledge, our result is the first one that shows the number of different optimal solutions is at least n, and such a conclusion holds independent of coefficient matrix A. This partially explains why the classical B&B approach for QAPs might take a long process to locate a global optimum. On the other hand, though the constructed optimal permutation matrices for problem (4) are completely different, the corresponding adjacency matrix of the hypercube remains the same. This indicates that if we can use the geometric properties of the hypercube to derive an optimization model equivalent to model (1), then it is possible that the new optimization model won't have multiple optimal solutions. This will definitely help to speed up the B&B approach applied to the new model.

3. A NEW CHARACTERIZATION OF THE HYPERCUBE

In this section we characterize the hypercube via its adjacency matrix. For a given vertex pair (v, v') of a graph, we define the distance between v and v' as the length of the shortest path that links v and v'. For convenience, we introduce the following definition:

Definition 3.1 *For a given graph $G=(V,E)$, we define the indicator cubic matrix $W_{ij}^l \in \Re^{n\times n\times d}$ by*

$$w_{ij}^l = \begin{cases} 1 & \delta(v_i,v_j)=l; \\ 0 & \delta(v_i,v_j)\neq l \end{cases}$$

We also denote by W^l the square matrix $(W^l)_{ij}=w_{ij}^l$ for $l=1,\cdots,d$.

Based on Definition 3.1, the matrix W^1 is precisely the adjacency matrix of the graph G.

We next explore the properties of the indicator matrix W of a hypercube. Combining Lemma 2.1, Lemma 2.2 and Definition 3.1, we have

Proposition 3.2 *Let $W\in\Re^{n\times n\times d}$ be the cubic matrix defined by (3.1) and G be a hypercube in $\Re^d$. Then one has*

$$\sum_{l=1}^{d} w_{ij}^l = 1, \qquad \forall i\neq j\in\{1,\cdots,n\}, \tag{6}$$

$$\sum_{j=1}^{n} w_{ij}^l = C(d,l), \qquad \forall i\in\{1,\cdots,n\}, l\in\{1,\cdots,d\}. \tag{7}$$

From the second conclusion of Lemma 2.1, we can derive the following

$$w_{ik}^1 + w_{kj}^1 + w_{ij}^1 \leq 2, \forall i,j,k\in\{1,\cdots,n\}$$

which is also called anti-triangular inequalities in graph theory. By using the first conclusion from Lemma 2.1, we can further derive the following enhanced version of the anti-triangular inequalities

$$\begin{aligned} \sum_{\text{l is odd}} w_{ik}^l + \sum_{\text{l is odd}} w_{kj}^l + \sum_{\text{l is odd}} w_{ij}^l \leq 2 \qquad \forall i,j,k\in\{1,\cdots,n\}; \\ \sum_{\text{l is even}} w_{ik}^l + \sum_{\text{l is even}} w_{kj}^l + \sum_{\text{l is odd}} w_{ij}^l \leq 2 \qquad \forall i,j,k\in\{1,\cdots,n\}. \end{aligned}$$

For any $X\in\Re^{n\times n}$, let us denote $X_{i.}$ the i-th column of X. We have

Theorem 3.3 *Let $W\in\Re^{n\times n\times d}$ be the cubic matrix defined by (3.1) and G be a hypercube. Then it satisfies the following relations*

$$(W_{i.}^1)^T W_{j.}^1 = \begin{cases} 2 & \delta(v_i,v_j)=2; \\ 0 & \delta(v_i,v_j)\neq 2 \end{cases} \qquad \forall\, i\neq j. \tag{8}$$

Proof: In order to prove the relation (8), we recall

$$(W_{i.}^1)^T W_{j.}^1 = \sum_{k=1}^{n} w_{ik}^1 w_{jk}^1.$$

If $\delta(v_i,v_j)=2$, it follows from Lemma 2.2 immediately that $(W_{i.}^1)^T W_{j.}^1=2$. If $\delta(v_i,v_j)\neq 2$, then for every $k\in\{1,\cdots,n\}$, it must hold $w_{ik}^1 w_{jk}^1=0$. Because otherwise we have $w_{ik}^1=w_{jk}^1=1$, which implies $\delta(v_i,v_j)=2$ by Lemma 2.1. This contradicts to the assumption that $\delta(v_i,v_j)\neq 2$. This concludes the proof of the theorem. q.e.d.

When G is a hypercube, the relation (8) established a one-to-one map between w_{ij}^2 and the row (or column) pairs of W^1. Unfortunately, such a relation does not exist between w_{ij}^l and the row (or column) pairs from W^1 and W^{l-1} when $l\geq 3$. This can be seen from the following simple example. Suppose that $w_{12}^1=w_{23}^1=w_{34}^1=1$, then $w_{13}^2=1$. Therefore, $(W_{1.}^2)^T W_{2.}^1>0$. However, we have $w_{12}^1=1, w_{12}^3=0$. Nevertheless, we have

Theorem 3.4 *Let $W\in\Re^{n\times n\times d}$ be the cubic matrix defined by (3.1). Then $\delta(v_i,v_j)=l$ if and only if $(W_{i.}^{l-1})^T W_{j.}^1>0$ and $\sum_{k=1}^{l-1} w_{ij}^k=0$.*

Proof: The necessary part of the theorem follows from Definition 3.1. It remains to prove the sufficient part of the theorem. Suppose that $(W_{i.}^{l-1})^T W_{j.}^1 > 0$. Then there exists a path from v_i to v_j whose length is at most l. Because $\sum_{k=1}^{l-1} w_{ij}^k = 0$, it must hold $\delta(v_i, v_j) = l$. q.e.d.

We next discuss how to characterize a hypercube in $\Re^d$ by making use of its adjacency matrix. The new characterization is based on the results in [29] where the authors characterized the hypercube in a certain space based on the relations between distinct adjacent edges of a graph. For notational convenience, we introduce the following definition.

Definition 3.5 *A graph $G = (V, E)$ is called an exact 4-cycle graph if each pair of distinct adjacent edges lies in exactly one 4-cycle.*

An exact 4-cycle graph shares many common interest properties as the hypercube. In particular, Laborde et'al [29] observed the following interesting link between an exact 4-cycle graph and the hypercube in a certain space.

Theorem 3.6 *[29] Suppose that a graph $G = (V, E)$ is a connected exact 4-cycle graph. If the degree of every vertex of G is d and $|V| = 2^d$, then G is a hypercube.*

Though the above theorem provides a complete geometric characterization of the hypercube in a certain space, it does not provide an algebraic characterization of the hypercube. We next show how to characterize the so-called exact 4-cycle relation in term of the adjacency matrix of a graph. It is worthwhile mentioning that from a viewpoint of computation, the approach based on the relations among different vertices of a graph is more promising since the number of vertices is much smaller than that of the edges.

Recall that the relation (8) describes the relation whether the distance between two vertices is 2. It follows from Definition 3.1 that

$$2w_{ij}^2 = (W_{i.}^1)^T W_{j.}^1, \quad for\ i \neq j = 1, \cdots, n.$$

Moreover, from Definition 3.1, we can easily see that

$$w_{ij}^1 + w_{ij}^2 \leq 1.$$

The above two relations can be used to define the so-called 4-cycle relation of a graph as shown by the following theorem.

Lemma 3.7 *Let $W \in \Re^{n \times n \times d}$ be the indicator matrix of a graph. If it satisfies the following relations*

$$(W_{i.}^1)^T W_{j.}^1 = 2w_{ij}^2, \quad w_{ij}^2 \in \{0, 1\}, \forall\ i \neq j \in \{1, \cdots, n\}; \tag{9}$$

$$w_{ij}^1 + w_{ij}^2 \leq 1, \qquad \forall\ i \neq j = 1, \cdots, n. \tag{10}$$

then G is an exact 4-cycle graph.

Proof: Without loss of generality, we can assume that (v_1, v_2) and (v_2, v_3) are two distinct adjacent edges of G. We therefore have $w_{12}^1 = w_{23}^1 = 1$. This implies $w_{13}^2 = 1$ and thus $(W_{1.}^1)^T W_{3.}^1 = 2$. It follows that there exists another index $k \in \{1, \cdots, n\}$ such that $w_{1k}^1 = w_{k3}^1 = 1$ and the set of four points $\{v_1, v_2, v_3, v_k\}$ forms precisely the unique 4-cycle that contains (v_1, v_2) and (v_2, v_3). q.e.d.

Before closing this section, we present the main theorem in this section which gives a complete characterization of a hypercube in term of its indicator matrix W.

Theorem 3.8 *Let $W \in \Re^{n \times n \times d}$ be the indicator matrix of a graph with $n = 2^d$. If it satisfies the following relations*

C1

$$\sum_{j=1}^{n} w_{ij}^l = C(d, l), \quad \forall i \in \{1, \cdots, n\}, l \in \{1, \cdots, d\};$$

C2 The matrices W^1 and W^2 satisfy the relation (9);

C3 For $l \geq 3$,

$$w_{ij}^l = 1 \iff (W_{i.}^{l-1})^T W_{j.}^1 > 0\ and\ \sum_{m=1}^{l-1} w_{ij}^m = 0;$$

C4

$$\sum_{l=1}^{d} w_{ij}^l = 1, \quad i \neq j \in \{1, \cdots, n\},$$

then G is a hypercube in $\Re^d$.

Proof: The condition C1 ensures that the degree of every vertex in G is d. The second and forth conditions guarantee that two distinct adjacent edges of G lie precisely in a 4-cycle as shown by Lemma 3.7. The third and forth conditions ensure that the graph is connected since every pair of vertices is connected. Applying Theorem 3.6, we can conclude the proof. q.e.d.

Though Theorem 3.8 gives a new characterization of the hypercube in terms of its indicator matrix W, it is desirable to further reduce the number of constraints in the characterization. In what follows we discuss how to remove some constraints in Theorem 3.8.

First we observe that Condition C2 implies that

$$\sum_{j=1}^{n} w_{ij}^2 = C(d,2), \quad \forall i \in \{1, \cdots, n\}.$$

Secondly, we can relax the conditions in the last two items of Theorem 3.8 to the following

C3.1 For $l \in \{3, \cdots, \lceil \frac{d}{2} \rceil\}$,

$$w_{ij}^l = 1 \Longleftrightarrow (W_{i.}^{l-1})^T W_{j.}^1 > 0 \text{ and } \sum_{m=1}^{l-1} w_{ij}^m = 0;$$

C4.1

$$\sum_{l=1}^{\lceil \frac{d}{2} \rceil} w_{ij}^l \leq 1, \quad i \neq j \in \{1, \cdots, n\}$$

We have

Theorem 3.9 *Let $W \in \Re^{n \times n \times d}$ be the indicator matrix of a graph with $n = 2^d$. If it satisfies the relations C1, C2, C3.1 and C4.1, then G is a hypercube in $\Re^d$.*

Proof: Suppose that G is a graph satisfying the assumptions of the theorem. The conditions C2 and C4.1 ensure that two distinct adjacent edges of G lie precisely in a 4-cycle as shown by Lemma 3.7. For any $v \in V$, let us define

$$\mathscr{S}_1(v,l) = \{v' \in V : \delta(v,v') \leq l\}.$$

It follows immediately that

$$|\mathscr{S}_1(v, \lceil \frac{d}{2} \rceil)| = 1 + \sum_{l=1}^{\lceil \frac{d}{2} \rceil} C(d,l) > \frac{n}{2}.$$

Now let us consider any vertex pair $v, v' \in V$. We therefore have

$$|\mathscr{S}_1(v, \lceil \frac{d}{2} \rceil)| + |\mathscr{S}_1(v, \lceil \frac{d}{2} \rceil)| > n.$$

The above relation implies that there exists at least one vertex $\bar{v}$ in the intersection of the two sets $\mathscr{S}_1(v, \lceil \frac{d}{2} \rceil)$ and $\mathscr{S}_1(v', \lceil \frac{d}{2} \rceil)$. From the definition of $\mathscr{S}_1(v,.)$, we can conclude that there exists a path from v to v', i.e., the graph is connected. The theorem follows immediately from Theorem 3.6. q.e.d.

4. NEW ILP FORMULATIONS AND RELAXATIONS

In this section, we describe our new ILP model for the QAPs associated with the hypercube in $\Re^d$. The new model can be viewed as a combination of the geometric structure of the hypercube and the linear linearization technique for ILPs [3].

We start with the following simple LP model

$$\begin{array}{lll} \min & \mathrm{Tr(AZ)} & \\ s.t. & z_{ii}=0, Z=Z^T, z_{ij}\in[0,1], & \forall i\neq j\in\{1,\cdots,n\}; \\ & \sum_{j=1}^n z_{ij}=d, & \forall i\in\{1,\cdots,n\}. \end{array} \tag{11}$$

The above problem can be viewed as a special case of the transportation problem and thus it has an integer optimal solution. It is also easy to see that the graph that satisfies the constraints in the above model is a graph whose vertices have the same degree (d). Thus we can cast model (11) as a relaxation of the QAPs associated with a hypercube. Nevertheless, we have

Theorem 4.1 *Let Z^* be the optimal solution of the LP problem (11) and (X^*,Y^*) be the optimal solution of the LP relaxation of problem (3). If the matrix A is nonnegative and B is the adjacency matrix of the hypercube in $\Re^d$, then it must hold*

$$\sum_{i=1}^n\sum_{j=1}^n\sum_{k=1}^n\sum_{l=1}^n a_{ik}b_{jl}y^*_{ijkl}\leq \mathrm{Tr(AZ^*)}.$$

Proof: To prove the theorem, we first show that from any optimal solution of the LP problem (11), we can construct a feasible solution for the LP relaxation of problem (3). Suppose that Z^* is the optimal solution of the relaxed version of problem (11). Let us define

$$\begin{aligned} x_{ij} &= \frac{1}{d}z^*_{ij}, \quad i,j=\{1,\cdots,n\}; \\ y_{ijkl} &= x_{ij}x_{kl}, \quad i,j,k,l\in\{1,\cdots,n\}. \end{aligned}$$

Then it is easy to verify that (X,Y) is a feasible solution of the LP relaxation of problem (3). Let E be the matrix whose elements have value 1. From the definition of X and Y, we have

$$\mathrm{Tr(AZ^*)}=\mathrm{Tr}\big(\mathrm{XAX^TE}\big)\geq\mathrm{Tr}\big(\mathrm{XAX^TB}\big)\geq\mathrm{Tr}\big(\mathrm{X^*A(X^*)^TB}\big),$$

where the first inequality follows from the fact that $b_{ij}\leq 1$ for $i,j\in\{1,\cdots,n\}$. This finishes the proof of the theorem. q.e.d.

Theorem 4.1 shows that compared with the classical ILP approaches based on model (3), much stronger lower bound can be derived by using the geometric structure of the hypercube for the underlying QAPs in the present work. We next present a result showing that the bound provided by the LP model (11) is weaker than the bound provided by the model in [20].

Theorem 4.2 *Let Z^* be the optimal solution of the LP problem (11) and (X^*,Y^*) be the optimal solution of the LP relaxation of the model in [20]. If the matrix A is nonnegative and B is the adjacency matrix of the hypercube in $\Re^d$, then it must hold*

$$\sum_{i=1}^n\sum_{j=1}^n\sum_{k=1}^n\sum_{l=1}^n a_{ik}b_{jl}y^*_{ijkl}\geq \mathrm{Tr(AZ^*)}.$$

Proof: To prove the theorem, it suffices to construct a feasible solution to the LP model (11) from a solution of the LP relaxation of the model in [20]. Let (X^*,Y^*) be such a solution, we define

$$z_{ij}=\sum_{k,l=1}^n y^*_{ikjl}b_{kl}.$$

Since $b_{kl}\in\{0,1\},\forall k,l\in\{1,\cdots,n\}$, it follows immediately that

$$0\leq z_{ij}\leq\sum_{k,l=1}^n y^*_{ikjl}=\sum_{l=1}^n x_{jl}=1.$$

Moreover, because $y^*_{ikil} = 0$ for all $k \neq l$ and $b_{ii} = 0$, we can claim that $z_{ii} = 0$ for all $i \in \{1, \cdots, n\}$. On the other hand, because B is the adjacency matrix of the hypercube in $\Re^d$, for any fixed $j \in \{1, \cdots, n\}$, we have

$$\begin{aligned}\sum_{i}^{n} z_{ij} &= \sum_{i=1}^{n}\sum_{k,l=1}^{n} y^*_{ikjl} b_{kl} = \sum_{k,l}^{1} x^*_{jl} b_{kl} \\ &= \sum_{l=1}^{n} x^*_{jl} \sum_{k=1}^{n} b_{kl} = d \sum_{l=1}^{n} x^*_{jl} = d.\end{aligned}$$

Recall that in the model (3), we have $y^*_{ikjl} = y^*_{jlik}$ and B is symmetric. It follows immediately $z_{ij} = z_{ji}, \forall i, j \in \{1, \cdots, n\}$. From the above discussion we can conclude that Z is a feasible solution of the LP model (11). This finishes the proof of the theorem. q.e.d.

We can further add more constraints to problem (11) to close the gap between the constraint polyhedral set and the hypercube. For example, we can impose the condition that there is no triangles in graph. This will lead to the following model

$$\begin{array}{lll} \min & \mathrm{Tr(AZ)} & \\ s.t. & z_{ii} = 0, Z = Z^T, z_{ij} \in [0,1], & \forall i \neq j \in \{1, \cdots, n\}; \\ & \sum_{j=1}^{n} z_{ij} = d, & \forall i \in \{1, \cdots, n\} \\ & z_{ij} + z_{jk} + z_{ki} \leq 2, & \forall i \neq j \neq k \in \{1, \cdots, n\}. \end{array}$$

We now discuss how to add more constraints to model (11) so that the new model defines precisely a hypercube. Let us recall that when the coefficient matrix B in (1) is the adjacency matrix of a hypercube in $\Re^d$, for any permutation matrix P, the matrix $P^T BP$ remains an adjacency matrix of the hypercube with different numbering of the vertices. From Theorem 3.8, we can characterize such an adjacency matrix via the indicator matrix W. We recall that the conditions C1 and C4.1 can be represented by linear constraints. In what follows we discuss how to represent the conditions C2 and C3.1 by using linear equations and inequalities.

In the sequel we show how to replace the nonlinear relation (9) by a set of linear constraints. Since all the variables w^1_{ij} are binary, $(W^1_{i.})^T W^1_{j.} = 2 = 2w^2_{ij}$ if and only if there exist two indices k, l such that

$$w^1_{ik} = w^1_{jk} = w^1_{il} = w^1_{jl} = 1.$$

Let us define [1]

$$t^1_{ijk} = \min\{w^1_{ik}, w^1_{jk}\}, \quad k = 1, \cdots, n. \tag{12}$$

We can rewrite (9) as

$$\sum_{k=1}^{n} t^1_{ijk} = 2w^2_{ij}.$$

The relation (12) is still nonlinear. However, since all the elements w^1_{ij} are binary, it is easy to see that it can be replaced by the following relations

$$\begin{aligned} t^1_{ijk} &\leq \min\{w^1_{ik}, w^1_{jk}\}, \qquad k = 1, \cdots, n, \\ w^1_{ik} + w^1_{jk} &\leq 1 + t^1_{ijk}, \qquad k = 1, \cdots, n. \end{aligned}$$

We argue that the binary requirement on w^2_{ij} in (9) can be waived if we impose the following linear constraints $1 \geq w^2_{ij} \geq t^1_{ijk} \geq 0$ for all $i \neq j, k \in \{1, \cdots, n\}$. We need to consider only two cases. Suppose that $(W^1_{i.})^T W^1_{j.} > 0$. Then it must hold $\max_{k=1,\cdots,n} t^1_{ijk} = 1$, and thus $w^2_{ij} = 1$. If $(W^1_{i.})^T W^1_{j.} = 0$, then $w^2_{ij} = 0$. This in turn guarantees the condition (9).

---

[1] The artificial variables t^1_{ijk} can also be derived via the linearization technique in [3]. To see this, let us lift model (11) to higher dimension through the transformation $y_{ijkl} = x_{ij} x_{kl}$. We then use the geometric structure of the hypercube to impose constraints on the variables in the lifted space. In our special case, we have $t^1_{ijk} = y_{ikkj}$.

We now progress to characterize the relation in condition C3.1 by a set of linear constraints. For $l \in \{2,\cdots,d\}$, we introduce new artificial variables t^l_{ijk} satisfying the following relations

$$t^l_{ijk} \le \min\{w^l_{ik}, w^1_{jk}, w^{l+1}_{ij}\}, \qquad k = 1,\cdots,n, \tag{13}$$

$$w^l_{ik} + w^1_{jk} \le 1 + t^l_{ijk} + t^{l-1}_{ikj}, \qquad k = 1,\cdots,n, \tag{14}$$

One can further verify that $w^{l+1}_{ij} = 1$ if and only if $t^l_{ijk} = 1$ and $\sum_{m=1}^l w^m_{ij} = 0$. Note that w^{l+1}_{ij} is binary automatically if the linear equalities (13)-(14) are satisfied and $w^1_{ij},\cdots,w^l_{ij}$ are binary. This gives a complete characterization of condition C3.1 in Theorem 3.9. We thus have the following result.

Theorem 4.3 *Suppose that the matrix B in problem (1) is the adjacency matrix of a hypercube in $\Re^d$ and $n = 2^d$. Then the QAP (1) can be equivalently cast as the following ILP:*

$$\begin{array}{lrl}
\min & \mathrm{Tr}\big(AW^1\big) & \\
s.t. & w^1_{ii} = 0, w^1_{ij} = w^1_{ji} \in \{0,1\}, & \forall\, i,j \in \{1,\cdots,n\};\\
& w^l_{ii} = 0, w^l_{ij} = w^l_{ji}, w^l_{ij} \in [0,1], & \forall i \ne j \in \{1,\cdots,n\}, l \in \{2,\cdots,d\},\\
& t^0_{ijk} = 0, u^0_{ijk} = 0, & \forall\, i \ne j, k \in \{1,\cdots,n\}, l \in \{1,\cdots,d\};\\
& t^l_{ijk} + t^{l-1}_{ikj} \le \min\{w^l_{ik}, w^1_{jk}\}, & \forall i,j,k \in \{1,\cdots,n\}, l \in \{1,\cdots,d-1\},\\
& 0 \le t^l_{ijk} \le w^{l+1}_{ij}, & \forall i,j,k \in \{1,\cdots,n\}, l \in \{1,\cdots,d-1\},\\
& w^l_{ik} + w^1_{jk} \le 1 + t^l_{ijk} + t^{l-1}_{ikj}, & \forall i,j,k \in \{1,\cdots,n\}, l \in \{1,\cdots,d-1\},\\
& (l+1)w^{l+1}_{ij} = \sum_{k=1}^n t^l_{ijk}, & \forall i \ne j \in \{1,\cdots,n\}, l = \{1,\cdots,d-1\},\\
& u_{ijk} \le \min\{w^{d-1}_{ik}, w^{d-1}_{jk}\}, & \forall i \ne j, k \in \{1,\cdots,n\},\\
& t^1_{ijk} + u_{ijk} \le w^2_{ij}, & \forall i,j,k \in \{1,\cdots,n\},\\
& w^{d-1}_{ik} + w^{d-1}_{kj} \le 1 + u_{ijk}, & \forall i \ne j, k \in \{1,\cdots,n\};\\
& \sum_{k=1}^n u_{ijk} = 2w^2_{ij}, & \forall i \ne j \in \{1,\cdots,n\};\\
& \sum_{j=1,j\ne i}^n t^l_{ijk} = (d-l)w^l_{ik}, & \forall\, i \ne k \in \{1,\cdots,n\}, l \in \{1,\cdots,d-1\};\\
& \sum_{i=1,i\ne j}^n t^l_{ijk} = C(d-1,l)w^1_{kj}, & \forall\, j \ne k \in \{1,\cdots,n\};\\
& \sum_{j=1,j\ne i}^n u_{ijk} = (d-1)w^{d-1}_{ik}, & \forall\, i \ne k \in \{1,\cdots,n\};\\
& \sum_{i=1,i\ne j}^n u_{ijk} = (d-1)w^{d-1}_{kj}, & \forall\, j \ne k \in \{1,\cdots,n\};\\
& \sum_{l=1}^{d-1}\left(t^l_{ijk} + t^l_{ikj}\right) = w^1_{jk}, & \forall i \ne j, i \ne k \in \{1,\cdots,n\},\\
& \sum_{l=1}^d w^l_{ij} = 1, & \forall i \ne j \in \{1,\cdots,n\}.
\end{array} \tag{15}$$

Proof: By the discussion preceding the theorem, we can conclude that if W is a feasible solution of problem (15), then W^1 must be an adjacency matrix of the hypercube in $\Re^d$. This finishes the proof of the theorem. q.e.d. .

It is straightforward to verify that the above problem has $\frac{1}{2}n(n-1)$ binary variables and roughly $O(n^3 \log n)$ linear constraints.

An alternative model is the following

$$\begin{array}{lrl}
\min & \mathrm{Tr}\big(AW^1\big) & \\
s.t. & w^1_{ii} = 0, w^1_{ij} = w^1_{ji} \in \{0,1\}, & \forall\, i,j \in \{1,\cdots,n\};\\
& w^l_{ii} = 0, w^l_{ij} = w^l_{ji}, w^l_{ij} \in [0,1], & \forall i \ne j \in \{1,\cdots,n\}, l \in \{2,\cdots,d\},\\
& w^0_{ii} = 1, w^0_{ij} = 0, w^{d+1}_{ij} = 0, & \forall\, i \ne j, k \in \{1,\cdots,n\}, l \in \{1,\cdots,d\};\\
& 0 \le t_{ijk} \le \min\{w^1_{ik}, w^1_{jk},\}, & \forall i \ne j, k \in \{1,\cdots,n\},\\
& w^1_{ik} + w^1_{jk} \le 1 + t_{ijk}, & \forall i \ne j, k \in \{1,\cdots,n\},
\end{array} \tag{16}$$

$$
\begin{aligned}
0 \le u_{ijk} \le \min\{w_{ik}^{d-1}, w_{jk}^{d-1}, \}, &\quad \forall i \neq j, k \in \{1, \cdots, n\}, \\
w_{ik}^{d-1} + w_{kj}^{d-1} \le 1 + u_{ijk}, &\quad \forall i \neq j, k \in \{1, \cdots, n\}; \\
t_{ijk} + u_{ijk} \le w_{ij}^2, &\quad \forall i \neq j, k \in \{1, \cdots, n\}; \\
\textstyle\sum_{k=1}^n t_{ijk} = 2w_{ij}^2, &\quad \forall i \neq j \in \{1, \cdots, n\}, l = \{1, \cdots, d-1\}, \\
\textstyle\sum_{k=1}^n u_{ijk} = 2w_{ij}^2, &\quad \forall i \neq j \in \{1, \cdots, n\}; \\
\textstyle\sum_{j=1, j\neq i}^n t_{ijk} = (d-1)w_{ik}^1, &\quad \forall\ i \neq k \in \{1, \cdots, n\}; \\
\textstyle\sum_{i=1, i\neq j}^n t_{ijk} = (d-1)w_{kj}^1, &\quad \forall\ j \neq k \in \{1, \cdots, n\}; \\
\textstyle\sum_{j=1, j\neq i}^n u_{ijk} = (d-1)w_{ik}^{d-1}, &\quad \forall\ i \neq k \in \{1, \cdots, n\}; \\
\textstyle\sum_{i=1, i\neq j}^n u_{ijk} = (d-1)w_{kj}^{d-1}, &\quad \forall\ j \neq k \in \{1, \cdots, n\}; \\
w_{ik}^l + w_{kj}^1 \le 1 + w_{ij}^{l+1} + w_{ij}^{l-1}, &\quad i \neq j, k \in \{1, \cdots, n\}, l \in \{2, \cdots, d\}, \\
w_{ik}^{d-1} + w_{kj}^{d-l} \le 1 + w_{ij}^{l-1} + w_{ij}^{l+1}, &\quad i \neq j, k \in \{1, \cdots, n\}, l \in \{2, \cdots, d\}, \\
\textstyle\sum_{j=1}^n w_{ij}^l = C(d, l) &\quad \forall i \in \{1, \cdots, n\}, l \in \{1, \cdots, d\}, \\
\textstyle\sum_{l=1}^d w_{ij}^l = 1, &\quad \forall i \neq j \in \{1, \cdots, n\}.
\end{aligned}
$$

5. NUMERICAL EXPERIMENTS

In this section, we report some numerical results based on our model. As pointed out in Section 2, the theoretical results in our work can indeed be used to improve most QAP solvers for the underlying problems. Therefore, in this section we only compare the lower bounds provided by various LP relaxations. For a fair comparison, we list the results based on our two models (16) and (15), model (3) (denoted by A&J), and the model in [20] (G&H98)[2] and their improved versions (A&J+ and G&H98+) with additional constraints (6)-(7), the GL[18] bound and the bound by the projection method proposed in [19, 6] (denoted by HRW).

We examine the performance of these models for several different choices of the matrix A. The first one is the matrix used in [39] defined by

$$
a_{ij} = \frac{\Delta^3 \sqrt{1-\rho^2}}{(2\pi\sigma^6)^{\frac{3}{2}}} \sum_{l=1}^n \exp^{-\frac{\delta^2}{2\sigma^2}((1-\rho^2)(i-n_1)^2 + (j-n_1-\rho*(i-n_1))^2 + (l-n_1-\rho*(i-n_1))^2)},
$$

where $n_1 = \frac{n+1}{2}$, and Δ is the step size to quantize the source, σ is the variance of the Gaussian Markov source with zero mean and ρ its correlation coefficient. In our experiment, we set the step size $\Delta = 0.4, \sigma = 1$ and $\rho = 0.1, 0.9$ respectively. These two different choices (denoted by Eng1 and Eng9) of ρ represent the scenarios of the source with dense correlation and non-dense correlation. We also include one example from vector quantization (denoted by VQ) provided by the last author of this work.

The other test problems include the so-called Harper code [23] with

$$
a_{ij} = |i - j|.
$$

and a random positive semi-definite matrix A. We point out that due to the special structure of the matrix A in Harper code, the optimal code can be found in polynomial time.

Our experiments were done on an 2.4GHz 64-bit AMD Opteron. We used CPLEX-10.2 with the barrier method to solve the LP relaxation. To see the quality of these lower bound, we also list the best solutions that have been found with a heuristics. In the second row for each model, we also list the CPU time in seconds to obtain the bound. The results for $d = 4$ resp. $d = 5$ are summarized in Table 1 resp. Table 2.

---

[2] We did not test the high-order lift model in [21] because it requires a great amount of memory. On the other hand, though very strong lower bounds have been obtained in [21], these bounds are obtained based on the dual Lagrangian method combined with some heuristics, which is quite different from the approach in the present paper.

TABLE 1. Bounds for QAPs associated with hypercube in $\Re^4$

	Harper	Random	Eng1	Eng9	VQ
Best Sol	240	64.344	0.0823832	0.0702545	19.7672
(15)	224.87 1592	64.313 1469	0.0823832 1168	0.0592348 2921	18.8316 2199
(16)	215.64 78	64.0187 143	0.0823832 95	0.0576037 109	17.8121 153
A&J	74 78	62.8863 89	0.0749003 109	0.0479947 88	7.07798 87
A&J +	122.02 82	63.4405 64	0.0771238 65	0.0548599 88	9.8353 70
GH98	120 76	62.8863 77	0.0749003 76	0.0479947 67	11.2419 75
GH98 +	160.47 62	63.4405 48	0.0771238 43	0.0548599 58	13.7575 60
GL	106 1	36.2555 1	0.0149548 1	0.0231322 1	10.2583 1
HRW	197.55 1	50.3012 1	0.0483 1	0.0458 1	15.8599 1

TABLE 2. Bounds for QAPs associated with hypercube in $\Re^5$

	Harper	Random	Eng1	Eng9	VQ
Best Sol	992	438.51	.8397e-1	.12402	20.487
(16)	650	292	.1355e-3	.2978e-3	5.5576
A&J	218	285.30	.6566e-4	.1138e-3	2.8149
A&J +	354.76	288.54	.8094e-4	.1927e-3	3.2097
GH98	324	285.30	.6566e-4	.1138e-3	3.9697
GH98 +	450.71	288.54	.8094e-4	.1927e-3	4.4460
GL	304	107.21	.1022e-4	.2424e-4	2.11940
HRW	783.03	211.96	-.2506	-.1533	.5013

As we can see from both Table 1 and 2, the bound by model (15) is the strongest, followed by model (16). However, it requires more memory and takes longer time to solve the LP relaxation of model (15). When $d = 5$, we run out of memory for model (15). The CPU time for solving the LP relaxation of model (16) is slightly longer than that for A&J and GH98 with a comparable memory usage. Not surprising to see that the improved versions (A&J+ and GH98+) require less CPU time, but result in better bounds than their original counterparts (A&J and GH98). We also note that both the GL and HRW bounds are very easy to compute. However, the GL bound is much weaker than other lower bounds in general. We also note that the HRW model can find very good bounds for the Harper code and the random problem very efficiently. However, as we can see from Table 2, for Eng1 and Eng9, the lower bound when $d = 5$ turns out be useless. Overall, the numerical effort is very comparable as for $d = 4$ and since our emphasis is on bounds, we do not list CPU times for $d = 5$.

6. CONCLUSIONS

In this paper, we have studied the QAPs associated with a hypercube in a finite space. By using the geometric structure of the hypercube, we show that there exist at least n different optimal permutations. Moreover, the symmetries in the hypercube allow us to restrict the search space and thus improve many existing QAP solvers for the underlying problem. Numerical experiments confirm our theoretical conclusions.

New ILP models for the underlying QAP are suggested. Compared with the existing ILP formulations for the underlying problem, our new model has the smallest number of binary variables and its LP relaxations can provide tighter lower bounds as verified by numerical experiments.

There are several different ways to extend our results. First of all, the LP relaxation of our new models still have $O(n^3 \log(n))$ constraints, which implies the model can not be applied to large-scale problems. One possible way to deal with this issue is to use relaxation based on other convex optimization models such as the LP over the second-order cone. To see this, let us recall that the adjacency matrix of a hypercube satisfies the condition

$$(W^1_{i.})^T W^1_{j.} \in \{0,2\}, \quad \forall\, i \neq j$$

which can be relaxed to

$$(W^1_{i.})^T W^1_{j.} \leq 2, \quad \forall\, i \neq j,$$

which equals to

$$\|W^1_{i.} + W^1_{j.}\|^2 \leq 2d + 4, \quad \forall\, i \neq j.$$

In this way, we can relax (15) to a LP over the second-order cone. More study is needed to explore whether such an approach can lead to computational advantages.

Another way to extend our approach is to develop new ILP models for other QAPs with structured coefficient matrices such as the adjacency matrix of a tree and Manhattan distances of rectangular grids.

REFERENCES

1. W.P. Adams and T.A. Johnson, "Improved Linear Programming-Based Lower Bounds for the Quadratic Assignment Problem", In *P.M. Pardalos and H. Wolkowicz, editors, Quadratic Assignment and Related Problems*, DIMACS 25 Series in Discrete Mathematics and Theoretical Computer Science, 1994, pp. 43-75.
2. W.P. Adams and H.D. Sherali, "A tight linearization and an algorithm for zero-one quadrtic programming problems", *Management Science*, **32**, 1274-1290 (1986).
3. W.P. Adams and H.D. Sherali, "A hierarchy of relaxations between continuous and convex hull representations for zero-one programming problems". *SIAM J. Discrete Applied Mathematics* **52**, 83-106 (1990).
4. K.M. Anstreicher and N.W. Brixius, "A new bound for the quadratic assignment problem based on convex quadratic programming", *Mathematical Programming* **80**, 341-357 (2001).
5. K.M. Anstreicher, N.W. Brixius, J.-P. Goux, and J.Linderoth. "Solving large quadratic assignment problems on computational grids", *Mathematical Programming* **91**, 563-588 (2002).
6. G. Ben-David and D. Malah, "Bounds on the performance of vector-quantizers under channel errors". *IEEE Transactions on Information Theory*, **51**, 2227–2235 (2005).
7. R. E. Burkard, "Efficiently Solvable Special Cases of Hard Combinatorial Optimization Problems", *Mathematical Programming* **79**, 55–69 (1997).
8. R. E. Burkard, E. Çela, G. Rote and G.J. Woeginger, "The Quadratic Assignment Problem with a Monotone Anti-Monge Matrix and a Symmetric Toeplitz Matrix: Easy and Hard Cases", *Mathematical Programming*, **82**, 125–158 (1998).
9. R.E. Burkard and J. Offermann. "Entwurf von Schreibmaschinentastaturen mittels quadratischer Zuordnungsprobleme". *Zeitschrift für Operations Research* **21**, B121-B132 (1977).
10. R.E. Burkard, E. Çela, S.E. Karisch and F. Rendl, "QAPLIB-A quadratic assignment problem library". *J. Global Optimization*, **10**, 391–403 (1997).
11. N.T. Cheng and N.K. Kingsburg, "Robust zero-redundancy vector quantization for noisy channels". In *Proceedings of the International Conference on Communications*, Boston, 1989, pp. 1338–1342.
12. N. Christofides and E. Benavent, "An exact algorithm for the quadratic assignment problem". *Operations Research*, **37**, 760–768, (1989).
13. N. Christofides and M. Gerrard, "A graph theoretical analysis of bounds for the quadratic assignment problem". In *P. Hansen eds: Studies on Graphs and Discrete programming*, North Holland, 1981, pp. 61-68.
14. Z. Drezner, "Lower bounds based on linear programming for the quadratic assignment problem". *Comput. Optim. & Applic.,* **4**, 159–169 (1995).
15. G. Finke, R.E. Burkard, and F. Rendl, "Quadratic assignment problems". *Annals of Discrete Mathematics*, **31**, 61–82 (1987).
16. A.M. Frieze and J. Yadegar, "On the quadratic assignment problem", *Discrete Applied Mathematics*, **5**, 89–98 (1983).
17. J.W. Gavett and N.V. Plyter, "The optimal assignment of facilities to locations by branch-and-bound". *Operations Research*, **14**, 210-232 (1966).
18. P.C. Gilmore, "Optimal and suboptimal algorithms for the quadratic assignment problem. *SIAM Journal on Applied Mathematics*, **10**, 305–313 (1962).
19. S.W. Hadley, F. Rendl, and H. Wolkowicz, "A new lower bound via projection for the quadratic assignment problem". *Mathematics of Operations Research*, **17**, 727-739 (1992).

20. P. Hahn. and T. Grant, "Lower bounds for the quadratic assignment problem based upon a dual formulation". *Operations Research*, **46**, 912–922 (1998).
21. P. Hahn, W.L. Hightower, T.A. Johnson, M. Guignard-Spielberg and C. Roucairol, "A Lower Bound for the Quadratic Assignment Problem Based on a Level-2 Reformulation- Linearization Technique", *University of Pennsylvania, Department of Systems Engineering*, Technical Report, 2001.
22. M. Hannan and J.M. Kurtzberg, "Placement techniques". In *Design Automation of Digital Systems: Theory and Techniques*, M.A. Breuer Ed., vol(1), Prentice-hall: Englewood Cliffs, 1972, pp. 213–282.
23. L.H. Harper, "Optimal assignments of numbers to vertices". *J. of the Society for Industrial and Applied Mathematics*, **12** , 131–135 (1964).
24. S.E. Karisch, E. Çela, J. Clausen. and T. Espersen, "A dual framework for lower bounds of the quadratic assignment problem based on linearization". *Computing*, **63**, 351-403 (1999).
25. S.E. Karisch and F. Rendl, "Lower bounds for the quadratic assignment problem via triangle decompositions". *Mathematical Programming*, **71**, 137-152 (1995).
26. L. Kaufman and F. Broeckx, "An algorithm for the quadratic assignment problem using Benders decomposition", *European Journal of Operational Research*, **2**, 204–211 (1978).
27. T. Koopmans and M. Beckmann. "Assignment problems and the location of economic activities", *Econometrica*, **25**, 53–76 (1957).
28. E. Lawler, "The quadratic assignment problem". *Management Science*, **9**, 586-599 (1963).
29. J.M. Laborde, S. Prakash and R. Hebbare, "Another characterization of hypercubes". *Discrete Mathematics*, **39**, 161–166 (1982).
30. E.M. Loiola, N.M. Maia de Abreu, P.O. Boaventura-Netto, P. Hahn, and T. Querido, "A survey for the quadratic assignment problem". *European J. Oper. Res.*, **176**, 657–690 (2007).
31. T. Mautor and C. Roucairol. "A new exact algorithm for the solution of quadratic assignment problems". *Discrete Applied Mathematics*, **55**, 281–293 (1994).
32. C.E. Nugent, T.E. Vollman, and J. Ruml, "An experimental comparison of techniques for the assignment of facilities to locations". *Operations Research*, **16**, 150–173 (1968).
33. L.C. Potter and D.M. Chiang, "Minimax nonredundant channel coding". *IEEE Transactions on Communications*, **43**, 804–811 (1995).
34. K.G. Ramakrishnan, M.G.C. Resende, B. Ramachandran. and J.F. Pekny, "Tight QAP bounds via linear programming". In: *Combinatorial and Global Optimization*, Eds by P.M. Pardalos, A. Migdalas, and R.E. Burkard, World Scientific Publishing, Singapore, 2002, pp. 297–303.
35. M.G.C. Resende, K.G. Ramakarishnan and Z. Drezner, "Computing lower bounds for the quadratic assignment problem with an interior point algorithm for linear programming". *Operations Research*, **43**, 781–791 (1995).
36. M. Scriabin and R.C. Vergin, "Comparison of computer algorithms and visual based methods for plant layout". *Management Science*, **22**, 172–187 (1975).
37. J. Skorin-Kapov, "Tabu search applied to the quadratic assingnment problem". *ORSA Journal on Computing*, **2**, 33–45 (1990).
38. E.D. Taillard, "Comparison of iterative searches for the quadratic assingnment problem". *Location Science*, **3**, 87-105 (1995).
39. X. Wang, X. Wu and S. Dumitrescu, "On optimal index assignment for MAP decoding of Markov sequences", Technical report, *Department of Electrical and Computer Engineering*, McMaster University, Ontario, Canada, 2006.
40. K. Zeger and A. Gersho, "Pseudo-Gray Code". *IEEE Trans. on Communications*, **38**, 2147–2158 (1990).
41. Q. Zhao, S.E. Karisch, F. Rendl, and H. Wolkowicz, "Semidefinite programming relaxations for the quadratic assignment problem". *J. Combinatorial Optimization*, **2**, 71–109 (1998).