LP Decodable Permutation Codes based on Linearly Constrained Permutation Matrices

Tadashi WADAYAMA[†], Manabu HAGIWARA[††*]

[†] Department of Computer Science, Nagoya Institute of Technology
Gokiso-cho, Showa-ku, Nagoya 466-8555, Japan
Email: wadayama@nitech.ac.jp

[††] National Institute of Advanced Industrial Science and Technology,
Research Center for Information Security,
Akihabara-Daibiru, 1-18-13 Sotokanda, Chiyoda-ku, Tokyo, Japan,
Email: hagiwara.hagiwara@aist.go.jp

[*] Center for Research and Development Initiative, Chuo University,
1-13-27 Kasuga, Bunkyo-ku, Tokyo, Japan.

***Abstract*— A set of linearly constrained permutation matrices are proposed for constructing permutation codes. Making use of linear constraints imposed on the permutation matrices, we can formulate a soft-decoding problem for the proposed class of permutation codes as a linear programming (LP) problem. An efficient LP solver based on simplex methods or interior point methods can be employed for solving this LP problem. Two types of linear constraints are discussed; one is structured constraints and another is random constraints. The structured constraints based on block permutation matrices lead to an efficient encoding algorithm. On the other hand, the random constraints enable us to use probabilistic methods for analyzing several code properties such as the average cardinality and the average weight distribution.**

I. Introduction

The class of linear codes defined over a finite field is ubiquitously employed in digital equipments for achieving reliable communication and storage systems. For example, the class of codes includes practically important codes such as Reed-Solomon codes, BCH codes, and LDPC codes. The linearity of codes enables us to use efficient encoding and decoding algorithms based on their linear algebraic properties.

On the other hand, there are some classes of nonlinear codes which are interesting from both theoretical and practical points of view. The class of *permutation codes* is such a class of nonlinear codes.

The origin of permutation codes dates back to 60's. Slepian [17] proposed a class of simple permutation codes, which is called *permutation modulation*, and efficient soft decoding algorithms for them. The variant I code [17] is obtained by applying all the permutations to the initial vector

$$(\overbrace{\mu_1,\mu_1\ldots,\mu_1}^{n_1}\overbrace{\mu_2,\ldots,\mu_2}^{n_2}\cdots\overbrace{\mu_k,\mu_k\ldots,\mu_k}^{n_k}),$$

where μ_i is a real value and $n = n_1 + \cdots + n_k$. This work has been extended and further investigated by many researchers; Biglieri and Elia [19], Karlof [18], Ingemarsson [20] studied optimization of the initial vector of the permutation modulation. Berger et al. [21] discussed applications of permutation codes to source coding problems.

There is another thread of researches on a class of permutation codes of length n whose codeword contains exactly n-distinct symbols; i.e., any codeword can be obtained by applying a permutation to an initial vector, e.g., $(0, 1, \ldots, n-1)$.

Some fundamental properties of such permutation codes were discussed in Blake et al. [1], and Frankl and Deza [8]. Vinck [13] [14] proposed applications of permutation codes for power-line communication and this triggered subsequent works on permutation codes. Wadayama and Vinck [16] presented a multi-level construction of permutation codes with large minimum distance. Many constructions for permutation codes have been developed so far, including the construction given in [4] [6]. Especially, the idea of a distance-preserving map due to Vinck and Ferreira [15] had influence on the study of permutation codes such as subsequent works by Chang et al. [2] [3].

Recently, rank modulation codes for flash memory proposed by Jiang et al. [9] [10] produced renewed interest in permutation codes. For example, for flash memory coding, Kløve et al. gave a new construction for permutation codes based on Chebyshev Distance [11], which is an appropriate distance measure for flash memory coding. Barg and Mazumdar [24] also studied some fundamental bounds on permutation codes in terms of the Kendall tau distance.

In order to employ a permutation code in a practical application, efficient encoding and soft-decoding algorithms are crucial to achieve reliable communication over noisy channels, such as an AWGN channel. Nonlinearity of permutation codes prevents the use of conventional encoding and decoding techniques based on linear algebraic properties. Although much works on permutation codes have been conducted, an aspect of efficient soft-decoding has not been intensively discussed so far. Therefore, there is still room for further researches on permutation codes with efficient encoding and soft-decoding algorithms.

In this paper, a new class of permutation codes called *LP decodable permutation codes* is introduced. An LP decodable permutation code is obtained by applying permutation matrices satisfying certain linear constraints to an n-dimensional real initial vector.

It is well known that permutation matrices are vertices of the Birkhoff polytope [35], which is the set of doubly stochastic matrices. Thus, a set of linearly constrained permutation matrices can be expressed by a set of linear equalities and linear inequalities. This property leads to the main feature of this class of permutation codes: *LP-decodable property*. For this class of codes, a decoding problem can be formulated as a linear programming (LP) problem. This means that we can exploit efficient LP solvers based on simplex methods or interior point methods to decode LP decodable permutation codes.

Furthermore, for a combination of this class of codes and its LP decoding, the maximum likelihood (ML) certificate property can be proved as in the case of the LP decoding for LDPC codes [7]. This is due to the fact that the LP problem given in this paper is a relaxed problem of an ML decoding problem.

In general, the fundamental polytope [27] [7] employed for LP decoding of LDPC codes contains many fractional vertices which are major source of sub-optimality of LP decoding. The constraints corresponding to an LDPC matrix are defined based on $\mathbb{F}_2$-arithmetics. On the other hand, an LP decoder works on the real number field. This domain "mismatch" produces many undesirable fractional vertices on the fundamental polytope. One of motivations of this work is to establish a coding scheme without this mismatch. Namely, the LP decodable permutation codes are defined on the real number field and they are decoded by using an LP solver working on the real number field as well.

The organization of the paper is as follows. Section II introduces some definitions and notation required for discussion. Section III gives the definition of the LP decodable permutation codes and its decoding algorithm. Section IV provides analysis for decoding performance of LP decoding and ML decoding. Section V presents a class of block permutation codes which are easy to encode with a combinatorial algorithm. Section VI offers probabilistic analysis on the cardinality and weight distribution of random LP decodable permutation codes. Section VII gives a concluding summary.

II. Preliminaries

A. Notation and definition

In this paper, matrices are represented by capital letters and a vector is assumed to be a column vector. Let X be an $n \times n$ real matrix. The notation $X \ge 0$ means that every element in X is non-negative. The notation $\mathsf{vec}(X)$ represents a vectorization of X given by

$$\mathsf{vec}(X) \overset{\triangle}{=} (X_{1,1} \ \cdots X_{1,n} \ X_{2,1} \ \cdots X_{2,n}, X_{3,1} \cdots X_{n,n})^T .$$

The vector $\mathbf{1}$ is the all-one vector whose length is determined by the context. The norm $|| \cdot ||$ denotes the Euclidean norm given by $||x|| \overset{\triangle}{=} (x^T x)^{1/2}$. The trace function $\mathsf{trace}(X)$ returns the sum of the diagonal elements of X. The sets $\mathbb{R}$ and $\mathbb{Z}$ are the sets of real numbers and integers, respectively. The set $[\alpha, \beta]$ denotes the set of consecutive integers from $\alpha \in \mathbb{Z}$ to $\beta \in \mathbb{Z}$.

The symbol $\trianglelefteq$ is defined by

$$\begin{pmatrix} a_1 \\ \vdots \\ a_m \end{pmatrix} \trianglelefteq \begin{pmatrix} b_1 \\ \vdots \\ b_m \end{pmatrix} \Leftrightarrow \forall i \in [1,m], a_i \star_i b_i,$$

where $\star_i$ is either $=$ or $\le$. For simplicity, the notation $\trianglelefteq = (\star_1, \star_2, \ldots, \star_m)$ is used to define $\trianglelefteq$ (e.g., $\trianglelefteq = (\le, =, \le)$).

The next definition gives a class of matrices of crucial importance in this paper.

Definition 1 (Permutation matrix): An $n \times n$ binary real matrix $X \overset{\triangle}{=} (X_{i,j})_{i,j \in [1,n]} \in \{0,1\}^{n \times n}$ is called a *permutation matrix* if and only if

$$\forall i, j \in [1,n], \sum_{j' \in [1,n]} X_{i,j'} = 1, \sum_{i' \in [1,n]} X_{i',j} = 1. \qquad (1)$$

The set of $n \times n$ permutation matrices is denoted by Π_n. □
It is also known that an $n \times n$ binary matrix is a permutation matrix if and only if the Hamming weights of every column and every row are exactly 1. The cardinality of Π_n is $n!$.

Removing the binary constraint from the definition of the permutation matrices, we have the definition of doubly stochastic matrices.

Definition 2 (Doubly stochastic matrix): An $n \times n$ non-negative real matrix $X \overset{\triangle}{=} (X_{i,j})_{i,j \in [1,n]}$ is called a *doubly stochastic matrix* if and only if (1) holds. □

The following theorem for a double stochastic matrix implies that the set of doubly stochastic matrices is a convex polytope.

Theorem 1 (Birkhoff-von Neumann theorem [35] [36]): Every doubly stochastic matrix is a convex combination of permutation matrices.
A simple proof of Birkhoff-von Neumann theorem can be found in [39].

The set of $n \times n$ doubly stochastic matrices is a polytope called the *Birkhoff polytope* B_n [35], which is also known as perfect matching polytope. The Birkhoff polytope is a $(n-1)^2$-dimensional convex polytope with $n!$-vertices and n^2-facets [34]. The Birkhoff-von Neumann theorem implies that any vertex (i.e., extreme point) of the Birkhoff polytope is a permutation matrix.

B. LP decoding for permutation vectors

Assume that $s \in \mathbb{R}^n$, called the *initial vector*, is given[1]. The set of images of s by left action of $X \in \Pi_n$ is called the *permutation vectors* of s, which is given by

$$\Lambda(s) \overset{\triangle}{=} \{Xs \mid X \in \Pi_n\}. \qquad (2)$$

For example, if $s = (0,1,2)^T$, then $\Lambda(s)$ is given by

$$\Lambda(s) = \{(0,1,2),(0,2,1),(1,0,2),(1,2,0),(2,0,1)(2,1,0)\}.$$

[1] The elements in s are not necessarily distinct each other.

We here consider a situation such that a vector of $\Lambda(s)$ is transmitted to a receiver over an AWGN channel. In such a case, it is desirable to use an ML decoding algorithm to estimate the transmitted vector. The ML decoding rule can be describe as

$$\hat{x} = \arg \min_{x \in \Lambda(s)} ||y - x||^2, \tag{3}$$

where y is a received word.

The next theorem states that the ML decoding for $\Lambda(s)$ can be formulated as the following LP problem.

Theorem 2 (LP decoding and ML certificate property): Assume that a vector in $\Lambda(s)$ is transmitted over an AWGN channel and $y \in \mathbb{R}^n$ is received on the receiver side. Let X^* be the solution of the following LP problem:

$$\begin{array}{rcl} \multicolumn{3}{l}{\text{maximize } \mathsf{trace}(C^T X)} \\ \multicolumn{3}{c}{\text{subject to}} \\ X & \in & \mathbb{R}^{n \times n} \\ X\mathbf{1} & = & \mathbf{1} \\ \mathbf{1}^T X & = & \mathbf{1}^T \\ X & \ge & 0, \end{array} \tag{4}$$

where $C \overset{\triangle}{=} ys^T$. If X^* is integral, $\hat{x} = X^* s$ holds.

Proof: The linear constraints in the above LP problem implies that X is constrained to be a doubly stochastic matrix.

One the other hand, the ML decoding rule can be recast as follows:

$$\begin{aligned} \hat{x} &= \arg \min_{x \in \Lambda(s)} ||y - x||^2 \\ &= (\arg \min_{X \in \Pi_n} ||y - Xs||^2) s \\ &= (\arg \min_{X \in \Pi_n} (||y||^2 - 2y^T(Xs) + ||Xs||^2)) s \\ &= (\arg \max_{X \in \Pi_n} y^T X s) s = (\arg \max_{X \in \Pi_n} \mathsf{trace}(C^T X)) s, \end{aligned}$$

where $C = ys^T$. Note that

$$\mathsf{trace}(C^T X) = \sum_{i=1}^{n} \sum_{j=1}^{n} C_{i,j} X_{i,j}. \tag{5}$$

Since the vertices of the Birkhoff polytope is a permutation matrix, the ML decoding can be formulated as an integer LP (ILP) problem:

$$\begin{aligned} &\text{maximize } \mathsf{trace}(C^T X) \\ &\text{subject to } X \in B_n, \quad X \text{ is an integral matirx.} \end{aligned}$$

By removing the integral constraint (X is an integral matrix), we obtain the LP problem (4). If the solution of this LP problem is integral, it must coincide with the solution of the above ILP problem. ■

As we have seen, the feasible set of the above LP problem is the Birkhoff polytope. Thus, an output of the above LP is highly likely integral.

The following example illustrates an LP decoding procedure.

Example 1: Let $s \overset{\triangle}{=} (0,1)^T$. In this case, the set of permutation vectors becomes $\Lambda(s) = \{(0,1)^T, (1,0)^T\}$. Assume that $y = (0.9, 0.2)^T$ is received. In this case,

$$C = ys^T = \begin{pmatrix} 0.9 \\ 0.2 \end{pmatrix} (0 \ 1) = \begin{pmatrix} 0 & 0.9 \\ 0 & 0.2 \end{pmatrix}$$

is obtained. By letting

$$X = \begin{pmatrix} X_{1,1} & X_{1,2} \\ X_{2,1} & X_{2,2} \end{pmatrix},$$

we have the objective function

$$\mathsf{trace}\left(\begin{pmatrix} 0 & 0 \\ 0.9 & 0.2 \end{pmatrix} \begin{pmatrix} X_{1,1} & X_{1,2} \\ X_{2,1} & X_{2,2} \end{pmatrix} \right) = 0.9 X_{1,2} + 0.2 X_{2,2}.$$

As a result, the LP decoding problem is given by

$$\begin{aligned} &\text{maximize } 0.9X_{1,2} + 0.2X_{2,2} \text{ subject to} \\ &X_{1,1} + X_{1,2} = 1, \quad X_{2,1} + X_{2,2} = 1, \\ &X_{1,1} + X_{2,1} = 1, \quad X_{1,2} + X_{2,2} = 1 \\ &X_{1,1}, X_{1,2}, X_{2,1}, X_{2,2} \ge 0. \end{aligned}$$

The solution of the problem is

$$X^* = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix},$$

and then we have the estimated word $X^* s = (1,0)^T$. □

III. Linearly constrained permutation matrices and LP decodable permutation codes

It is natural to consider an extension of the LP decoding presented in the previous section. Additional linear constraints imposed on Π_n produce a restricted set of $\Lambda(s)$. A decoding problem of such a set can be formulated as an LP problem, as in the case of the ML decoding of $\Lambda(s)$.

A. Definitions

The next definition for linearly constrained permutations gives an LP-decodable subset of $\Lambda(s)$.

Definition 3 (linearly constrained permutation matrix): Let m, n be positive integers. Assume that $A \in \mathbb{Z}^{m \times n^2}$, $b \in \mathbb{Z}^m$ and $\unlhd$ are given. A set of *linearly constrained permutation matrices* is defined by

$$\Pi(A, b, \unlhd) \overset{\triangle}{=} \{X \in \Pi_n \mid A \ \mathsf{vec}(X) \unlhd b\}. \tag{6}$$

□

Note that $A \ \mathsf{vec}(X) \unlhd b$ formally represents additional m equalities and inequalities. These additional constraints provide a restriction on permutation matrices.

From the linearly constrained permutation matrices, LP decodable permutation codes are naturally defined as follows.

Definition 4 (LP decodable permutation code): Assume the same set up as in Definition 3. Suppose also that $s \in \mathbb{R}^n$ is given. The set of vectors $\Lambda(A, b, \unlhd, s)$ given by

$$\Lambda(A, b, \unlhd, s) \overset{\triangle}{=} \{Xs \in \mathbb{R}^n \mid X \in \Pi(A, b, \unlhd)\} \tag{7}$$

is called an LP decodable permutation code. □

If

$$X^{(1)}, X^{(2)}(X^{(1)} \neq X^{(2)}) \in \Pi(A, b, \trianglelefteq) \Rightarrow X^{(1)}s \neq X^{(2)}s \tag{8}$$

holds, then an LP decodable permutation code is said to be *non-singlar*. Namely, there is one-to-one correspondence between permutation matrices in $\Pi(A, b, \trianglelefteq)$ and codewords of $\Lambda(A, b, \trianglelefteq, s)$ if a code is non-singular. Note that a code may become singular if identical symbols exist in s.

The next example shows a case where an additional linear constraint imposes a restriction on permutation matrices.

Example 2: Consider the set of linearly constrained permutation matrices which consists of 4×4 permutation matrices satisfying the linear constraint $\mathsf{trace}(X) = 0$. The constraint implies that the diagonal elements of the permutation matrices are constrained to be zero. This means that such permutation matrices correspond to permutations without fixed points, which are called *derangements*. For $n = 4$, there are 9-derangement permutation matrices as follows:

$$\begin{pmatrix} 0100 \\ 1000 \\ 0001 \\ 0010 \end{pmatrix} \begin{pmatrix} 0100 \\ 0010 \\ 0001 \\ 1000 \end{pmatrix} \begin{pmatrix} 0100 \\ 0001 \\ 1000 \\ 0010 \end{pmatrix}$$
$$\begin{pmatrix} 0010 \\ 1000 \\ 0001 \\ 0100 \end{pmatrix} \begin{pmatrix} 0010 \\ 0001 \\ 1000 \\ 0100 \end{pmatrix} \begin{pmatrix} 0010 \\ 0001 \\ 0100 \\ 1000 \end{pmatrix}$$
$$\begin{pmatrix} 0001 \\ 1000 \\ 0100 \\ 0010 \end{pmatrix} \begin{pmatrix} 0001 \\ 0010 \\ 1000 \\ 0100 \end{pmatrix} \begin{pmatrix} 0001 \\ 0010 \\ 0100 \\ 1000 \end{pmatrix}.$$

In this case, the triple $(A, b, \trianglelefteq)$ is defined by

$$A = \mathsf{vec}(I), \quad b = 0, \quad \trianglelefteq = (=), \tag{9}$$

where I is the 4×4 identity matrix. Multiplying these matrices to the initial vector $s = (0, 1, 2, 3)^T$ from left, we immediately obtain the members of $\Lambda(A, b, \trianglelefteq, (0, 1, 2, 3)^T)$:

$$\begin{array}{lll} (1,0,3,2)^T, & (1,2,3,0)^T, & (1,3,0,2)^T, \\ (2,0,3,1)^T, & (2,3,0,1)^T, & (2,3,1,0)^T, \\ (3,0,1,2)^T, & (3,2,0,1)^T, & (3,2,1,0)^T. \end{array} \tag{10}$$

This code is thus non-singular. If the initial vector is

$$s = (0, 0, 0, 0)^T,$$

then the resulting code has the only codeword $(0, 0, 0, 0)$. In this case, the code is singular. □

B. LP decoding for LP decodable permutation codes

The LP decoding of $\Lambda(A, b, \trianglelefteq, s)$ is a natural extension of the LP decoding for $\Lambda(s)$. Assume that a vector in $\Lambda(A, b, \trianglelefteq, s)$ is transmitted over an AWGN channel and $y \in \mathbb{R}^n$ is given. The procedure for the LP decoding of $\Lambda(A, b, \trianglelefteq, s)$ is given as follows.

LP decoding for an LP decodable permutation code

1) Solve the following LP problem and let X^* be the solution.

$$\begin{array}{rcl} & \text{maximize } \mathsf{trace}(C^T X) & \\ & \text{subject to} & \\ X & \in & \mathbb{R}^{n \times n}, \\ X & \geq & 0, \\ X\mathbf{1} & = & \mathbf{1}, \\ \mathbf{1}^T X & = & \mathbf{1}^T, \\ A\, \mathsf{vec}(X) & \trianglelefteq & b, \end{array} \tag{11}$$

where $C = ys^T$.

2) Output $X^* s$ if X^* is integral. Otherwise, declare decoding failure.

C. Remarks

There are several remarks that should be made on the LP decoding for $\Lambda(A, b, \trianglelefteq, s)$.

The feasible set of (11) is a subset of the feasible set of (4). All the matrices in $\Pi(A, b, \trianglelefteq)$ are feasible and permutation matrices which do not belong to $\Pi(A, b, \trianglelefteq)$ are infeasible. This implies that all the integral points of the feasible set (11) coincide with $\Pi(A, b, \trianglelefteq)$.

The LP problem (11) is a relaxed problem of the ML decoding problem over AWGN channels:

$$\text{minimize } ||y - x||^2 \text{ subject to } x \in \Lambda(A, b, \trianglelefteq, s). \tag{12}$$

This can be easily shown, as in the case (4). As a consequence of the above properties on integral points and on the relaxation, it can be concluded that the LP decoding for $\Lambda(A, b, \trianglelefteq, s)$ has the ML-certificate property as well. Namely, if the output of LP decoding is not decoding failure (i.e., X^* is integral), the output is exactly the same as the solution of the minimum distance decoding problem (12).

The feasible set of the LP problem (11) is the intersection of the Birkhoff polytope and a (possibly unbounded) convex set defined by the additional constraints. The intersection becomes a polytope which is called a *code polytope*. The decoding performance of LP decoding is closely related to the code polytope given by the following definition.

Definition 5 (Code polytope): The polytope $\mathcal{P}(A, b, \trianglelefteq)$ defined by

$$\mathcal{P}(A, b, \trianglelefteq) \stackrel{\triangle}{=} B_n \cap \{X \in \mathbb{R}^{n \times n} \mid A\ \mathsf{vec}(X) \trianglelefteq b\} \tag{13}$$

is called the code polytope for $\Pi(A, b, \trianglelefteq)$, where B_n is the Birkhoff polytope corresponding to Π_n. □

Figure 1 illustrates a code polytope. It should be remarked that the set of integral vertices of the code polytope coincides with $\Pi(A, b, \trianglelefteq)$. Due to additional linear constraints $A\ \mathsf{vec}(X) \trianglelefteq b$, a code polytope may have some fractional vertices, which contain components of fractional number.

In an LP decoding process, these fractional vertices becomes a possible candidate of an LP solution. Thus, these fractional vertices can be considered as *pseudo permutation matrices* which degrade the decoding performance of the LP decoding.

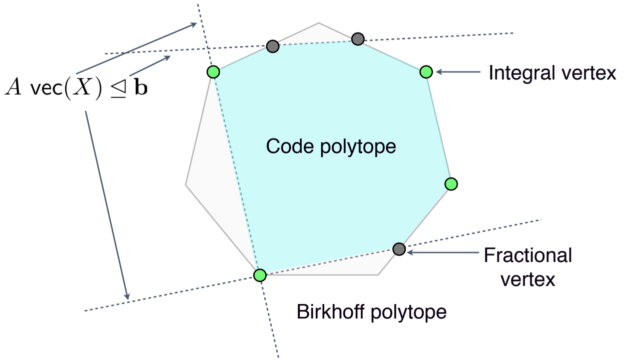


Fig. 1. Code polytope $\mathcal{P}(A, b, \trianglelefteq)$

IV. Analysis for decoding performance of LP decoding and ML decoding

In this section, upper bounds on decoding error probability for LP decoding and ML decoding are presented.

A. Upper bound on LP decoding error probability

An advantage of the LP formulation of a decoding algorithm is its simplicity for detailed decoding performance analysis. The geometrical properties of a code polytope is closely related to its decoding performance of the LP decoding. We can evaluate the block error probability of the proposed scheme with reasonable accuracy if we have enough information on a set of vertices of a code polytope. The bound presented in this section has close relationship to the pseudo codeword analysis on LDPC codes [5].

In this section, a set of parameters $A, b, \trianglelefteq, s$ are assumed to be given. Let V be the set of vertices of the code polytope $\mathcal{P}(A, b, \trianglelefteq, s)$. In general, V contains fractional vertices.

The next lemma gives bridge between a code polytope and corresponding decoding error probability.

Lemma 1 (Upper bound on block error rate for LPD): Assume that a codeword Xs is transmitted to a receiver via an AWGN channel, where $X \in \Pi(A, b, \trianglelefteq)$. The additive white Gaussian noise with mean 0 and variance σ^2 is assumed. The receiver uses the LP decoding algorithm presented in the previous section. In this case, the block error probability $P_{LP}(X)$ is upper bounded by

$$P_{LP}(X) \le \sum_{\tilde{X} \in V \backslash \{X\}} Q\left(\frac{||Xs||^2 - (\tilde{X}s)^T Xs}{\sigma ||\tilde{X}s - Xs||}\right), \tag{14}$$

where the Q-function is the tail probability of the normal Gaussian distribution, which is given by

$$Q(x) \triangleq \int_x^\infty \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{t^2}{2}\right) dt. \tag{15}$$

Proof: Let $y = Xs + z$, where z is an additive white Gaussian noise term. We first consider the pairwise block error probability $P_e(X, \tilde{X})$ between X and $\tilde{X} \in \Pi(A, b, \trianglelefteq)$, which is given by

$$P_e(X, \tilde{X}) \triangleq Prob[y^T \tilde{X}s \ge y^T Xs]. \tag{16}$$

Namely, $P_e(X, \tilde{X})$ is the probability such that $\tilde{X}s$ is more likely than Xs for a given y under the assumption that only $\tilde{X}$ and X are allowable permutation matrices.

The difference $y^T \tilde{X}s - y^T Xs$ can be transformed into

$$\begin{aligned} y^T \tilde{X}s - y^T Xs &= (Xs + z)^T (\tilde{X}s - Xs) \\ &= (\tilde{X}s - Xs)^T z + (\tilde{X}s - Xs)^T Xs \\ &= (\tilde{X}s - Xs)^T z \\ &- (||Xs||^2 - (\tilde{X}s)^T Xs). \end{aligned} \tag{17}$$

We thus have

$$Prob[y^T \tilde{X}s \ge y^T Xs] = Prob[a^T z \ge b], \tag{18}$$

where $a \in \mathbb{R}^n$ and $b \in \mathbb{R}$ are given by

$$a \triangleq \tilde{X}s - Xs, \tag{19}$$
$$b \triangleq ||Xs||^2 - (\tilde{X}s)^T Xs. \tag{20}$$

The left-hand side of $a^T z \ge b$ is a linear combination of Gaussian noises. The mean of $a^T z$ is zero and the variance is given by

$$Var[a^T z] = \sigma^2 ||a||^2. \tag{21}$$

The probability such that the Gaussian random variable $a^T z$ takes a value larger than or equal to b can be expressed as

$$\begin{aligned} P_e(X, \tilde{X}) &= Prob[a^T z \ge b] \\ &= Q\left(\frac{b}{\sigma ||a||}\right). \end{aligned} \tag{22}$$

Combining the union bound and this pairwise error probability, we immediately obtain the claim of this lemma. ■

The upper bound on decoding error probability in Lemma 1 naturally leads to a pseudo distance measure on $\mathbb{R}^{n \times n}$.

Definition 6 (Pseudo distance): The function

$$D_s(X, \tilde{X}) \triangleq \frac{||Xs||^2 - (\tilde{X}s)^T Xs}{||\tilde{X}s - Xs||} \tag{23}$$

is called the *pseudo distance* where $X, \tilde{X} \in \mathbb{R}^{n \times n}$ are doubly stochastic matrices. □

Note that $D_s(\cdot, \cdot)$ is not a distance function since it does not satisfy the axioms of distance. In terms of decoding error probability, geometry of the vertices of a code polytope should be established based on this pseudo distance.

For example, in high SNR regime, the *minimum pseudo distance*

$$\Delta_s \triangleq \min_{X \in \Pi(A, b, \trianglelefteq), \tilde{X} \in V, \tilde{X} \neq X} D_s(X, \tilde{X}) \tag{24}$$

is expected to be highly influential to the decoding error probability.

Example 3: Assume that $X \in \mathbb{R}^{3\times 3}$. Suppose the linear constraint $\mathtt{trace}(X) = 1$. In this case, the code polytope has the following 5-vertices:

$$\begin{aligned}
M^{(1)} &\triangleq \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix}, M^{(2)} \triangleq \begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix}, \\
M^{(3)} &\triangleq \begin{pmatrix} 0 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \end{pmatrix}, M^{(4)} \triangleq \begin{pmatrix} 1/3 & 0 & 2/3 \\ 2/3 & 1/3 & 0 \\ 0 & 2/3 & 1/3 \end{pmatrix}, \\
M^{(5)} &\triangleq \begin{pmatrix} 1/3 & 2/3 & 0 \\ 0 & 1/3 & 2/3 \\ 2/3 & 0 & 1/3 \end{pmatrix}. \qquad (25)
\end{aligned}$$

In this case, the set of vertices consists of 3-integral vertices and 2-fractional vertices. Let $s = (0, 1, 2)^T$. The pseudo distance distribution form $M^{(1)}$ is given by

$$\begin{aligned}
D_s(M^{(1)}, M^{(2)}) &= 1.388730 \\
D_s(M^{(1)}, M^{(3)}) &= 1.224745 \\
D_s(M^{(1)}, M^{(4)}) &= 1.224745 \\
D_s(M^{(1)}, M^{(5)}) &= 1.224745.
\end{aligned}$$

The minimum pseudo distance of this code polytope is $\Delta_s = 1.224745$. □

B. Upper bound on ML decoding error probability

Assume the same setting as in the previous subsection. In the case of ML decoding, we can neglect the effect of fractional vertices. Therefore, we obtain an upper bound on the ML block error probability

$$\begin{aligned}
P_{ML}(X) &\le \sum_{\tilde{X} \in \Pi(A,b,\trianglelefteq) \setminus \{X\}} Q\left(\frac{||Xs||^2 - (\tilde{X}s)^T Xs}{\sigma ||\tilde{X}s - Xs||} \right) \\
&= \sum_{\tilde{X} \in \Pi(A,b,\trianglelefteq) \setminus \{X\}} Q\left(\frac{||\tilde{X}s - Xs||}{2\sigma} \right) \qquad (26)
\end{aligned}$$

based on a similar argument. The above equality holds since $||Xs|| = ||\tilde{X}s||$ holds for any $\tilde{X} \in \Pi(A, b, \trianglelefteq)$. Note that this simplification cannot apply to $\tilde{X}$ if which $\tilde{X}$ is a fractional vertex. This is because the preservation of Euclidean norm does not hold in general for a doubly stochastic matrix. For example, we have

$$\left|\left| \begin{pmatrix} 1/3 & 2/3 & 0 \\ 0 & 1/3 & 2/3 \\ 2/3 & 0 & 1/3 \end{pmatrix} s \right|\right| = 1.9147 \ne ||s|| = \sqrt{5}, \qquad (27)$$

where $s = (0, 1, 2)^T$.

If $\Pi(A, b, \trianglelefteq)$ have a group structure under the matrix multiplication, the above upper bound can be further simplified as

$$P_{ML} \le \sum_{\tilde{X} \in \Pi(A,b,\trianglelefteq) \setminus \{I\}} Q\left(\frac{||\tilde{X}s - s||}{2\sigma} \right). \qquad (28)$$

It should be remarked that the second upper bound (28) is independent of the transmitted codeword. In order to prove the bound (28), it is sufficient to prove $\Pi(A, b, \trianglelefteq)$ is distance invariant with respect to the Euclidean distance.

In the following, the distance invariant property of $\Pi(A, b, \trianglelefteq)$ will be shown. Let us define the Euclidean distance enumerator by

$$W_X(Z) \triangleq \sum_{\tilde{X} \in \Pi(A,b,\trianglelefteq)} Z^{||Xs - \tilde{X}s||}. \qquad (29)$$

This enumerator has the information on distance distributions measured from the permutation matrix X.

The next lemma states that the Euclidean distance enumerator does not depend on the center point X if the linearly constrained permutation matrices have a group structure. This property can be regarded as a *distance invariance property* of permutation codes.

Lemma 2 (Distance invariance): If $\Pi(A, b, \trianglelefteq)$ forms a group under the matrix multiplication over $\mathbb{R}$, the equality

$$W_X(Z) = W(Z) \qquad (30)$$

holds for any $X \in \Pi(A, b, \trianglelefteq)$. The *pseudo weight enumerator* $W(Z)$ is defined by $W(Z) = W_I(Z)$ where I is the $n \times n$ identity matrix.

Proof: Since $\Pi(A, b, \trianglelefteq)$ forms a group, the inverse X^{-1} belongs to $\Pi(A, b, \trianglelefteq)$ as well. Since the inverse X^{-1} induces a symbol-wise permutation, it is evident that

$$||Xs - \tilde{X}s|| = ||X^{-1}Xs - X^{-1}\tilde{X}s|| = ||s - X^{-1}\tilde{X}s|| \qquad (31)$$

holds for any $X, \tilde{X} \in \Pi(A, b, \trianglelefteq)(X \ne \tilde{X})$. The Euclidean distance enumerator can be rewritten as

$$\begin{aligned}
W_X(Z) &= \sum_{\tilde{X} \in \Pi(A,b,\trianglelefteq)} Z^{||Xs - \tilde{X}s||} \\
&= \sum_{\tilde{X} \in \Pi(A,b,\trianglelefteq)} Z^{||s - X^{-1}\tilde{X}s||} \\
&= \sum_{X' \in \Pi(A,b,\trianglelefteq)} Z^{||s - X's||} = W(Z). \qquad (32)
\end{aligned}$$

The second equality is a consequence of Eq. (31). The last equality is due to the assumption that $\Pi(A, b, \trianglelefteq)$ forms a group. ■

Example 4: We have performed the following computer experiment for the following two codes:

1) LP decodable permutation code corresponding to the derangements of length 5. The additional linear constraint is $\mathtt{trace}(X) = 0$. A transmitted word $(1, 0, 4, 2, 3)^T$ is assumed. The code polytope has 44-vertices which are all integral vertices.
2) LP decodable permutation code of length 5 corresponding to an additional linear constraint $X_{1,1} + X_{5,5} = 1$. A transmitted word $(0, 4, 3, 2, 1)^T$ is assumed. The code polytope has 330-vertices. The set of vertices contains 36-integral vertices and 294-fractional vertices.

The AWGN channel with noise variance σ^2 is assumed. The signal-to-noise ratio is defined by $SNR = 10 \log_{10} \left(1/\sigma^2\right)$.

The LP decoding described in the previous section was employed for decoding.

Figure 2 presents the upper bounds and simulation results on block error probability of these permutation code. It is readily observed that the upper bounds presented in this section shows reasonable agreement with the simulation results.

The both codes have the same minimum pseudo distance 0.707107 and similar cardinalities (44 and 36) but the derangement code provides much better block error probabilities than those of the code with the constraint $X_{1,1}+X_{5,5}=1$. This is because the existence of fractional vertices (i.e., 294-fractional vertices) severely degrades the decoding performance of the code with the constraint $X_{1,1}+X_{5,5}=1$ compared with the derangement code. □

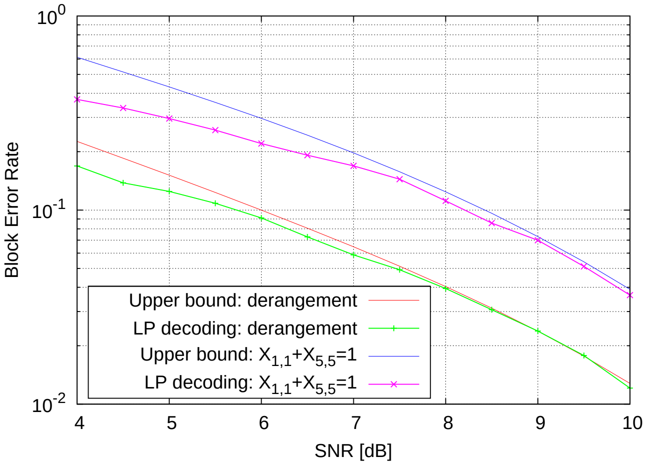


Fig. 2. Comparison of upper bounds and simulation results for LP decoding on block error probabilities $(n = 5)$

V. Structured permutation matrices

An efficient encoding algorithm is required for realizing a coded system employing the LP decodable permutation codes. A straightforward way to implement an encoding map is the use of a look-up table for converting a message vector to a permutation matrix. If the cardinality of $\Pi(A, b, \trianglelefteq)$ is not too large, a table look-up method for encoding may be practical. An evident disadvantage of the table look-up approach is that the use of this approach is limited to the case in which the code length is sufficiently short.

In this section, we discuss structured constraints which enable us to use efficient combinatorial algorithms for encoding. Especially, the focus is on a set of linearly constrained permutation matrices which forms a block-wise permutation group. The set of block permutation matrices plays a key role in this section.

A. Block permutation matrix

In this subsection, we prepare required definitions to handle block permutation matrices.

Suppose the situation where the set $[1,n]\times[1,n]$ is divided into mutually disjoint $\gamma\times\gamma$ square blocks of size $\nu\times\nu$ (i.e., $n=\gamma\nu$ holds). The square blocks are called *blocks* which is explicitly defined as follows.

Definition 7 (Block): For $k, b \in [1,\gamma]$, a *block* $B_{k,b}$ is defined by

$$B_{k,b} \triangleq \{(i,j)\in[1,n]^2 \mid \nu(k-1) < i \le \nu k, \nu(b-1) < j \le \nu b\}. \tag{33}$$

The indices k and b are called *block indices*. □

We further split a block into a set of *rectangle regions*.

Definition 8 (Rectangle region): The rectangle region $T^{(l)}_{k,b}$ is defined as

$$T^{(l)}_{k,b} \triangleq \{(x,y)\in B_{k,b} \mid y = \nu(b-1)+l\} \tag{34}$$

for $k,b\in[1,\gamma]$ and $l\in[1,\nu]$. The subscript k,b specifies the block where the rectangle region $T^{(l)}_{k,b}$ belongs to. The superscript $l\in[1,\nu]$, which is called a *subindex,* indicates the relative position in the block $B_{k,b}$. □

We are now ready to define a block permutation matrix which is the basis for realizing a block-wise permutation group.

Definition 9 (Block permutation matrix): Assume that a permutation matrix $X\in\Pi_n$ is given. If, for any $b\in[1,\gamma]$, there exists the unique block index k satisfying

$$X(B_{k,b}) \ne 0 \tag{35}$$

then X is called a *block permutation matrix*. The notation $X(B_{k,b})$ represents the sub-matrix of X corresponding to the block $B_{k,b}$. □

From this definition, it is apparent that a nonzero $X(B_{k,b})\in\{0,1\}^{\nu\times\nu}$ is a permutation matrix if X is a block permutation matrix. Furthermore, there exists the unique block index b satisfying $X(B_{k,b})\ne 0$ for any block index $k\in[1,\gamma]$. This equivalent statement can be obtained by exchanging the role of column and row in the above definition.

B. Linear constraints for block permutation matrix

In this subsection, a set of linear constraints providing a block permutation is discussed. For preparation, we need the definition of the skewed column set $U^{(l)}_{k,b}\subset[1,n]^2$ as follows.

Definition 10: (Skewed column set) For block indices $k,b\in[1,\gamma]$ and subindex $l\in[1,\nu]$, the *skewed column set* is defined by

$$U^{(l)}_{k,b} \triangleq T^{(l)}_{k,b} \cup \left(\bigcup_{k'\in[1,\gamma]\setminus\{k\}} T^{(l \bmod \nu)+1}_{k',b} \right). \tag{36}$$

□

Figure 3 illustrates the subsets of $[1,n]\times[1,n]$ appeared so far such as the blocks, the rectangle regions, and the skewed column set.

The next theorem presents a set of linear constraints characterizing block permutation matrices.

Theorem 3 (Characterization of block permutation matrix): Let $X\in\Pi_n$ be a permutation matrix. The permutation matirx

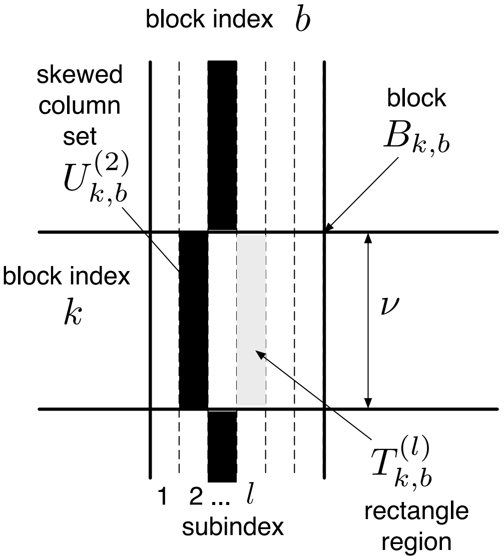


Fig. 3. Blocks, rectangle regions and skewed column set

X is a block permutation matrix if and only if

$$\sum_{(u,v)\in U^{(l)}_{k,b}} X_{u,v} = 1 \tag{37}$$

holds for any $b, k \in [1,\gamma], l \in [1,\nu]$.

Proof: In the first part of the proof, we will show that any block permutation matrix satisfies (37).

Assume that $k, b \in [1,\gamma]$ and $l \in [1,\nu]$ are arbitrary chosen. From the definition of the skewed column set $U^{(l)}_{k,b}$, the left-hand side of (37) can be rewritten as

$$\begin{aligned}\sum_{(u,v)\in U^{(l)}_{k,b}} X_{u,v} &= \sum_{(u,v)\in T^{(l)}_{k,b}} X_{u,v} \\ &+ \sum_{k'\in[1,\gamma]\setminus\{k\}} \left(\sum_{(u,v)\in T^{(l \bmod \nu)+1}_{k',b}} X_{u,v} \right).\end{aligned} \tag{38}$$

Recall that X is assumed to be a block permutation matrix. This means that there exists a unique block index $\kappa \in [1,\gamma]$ satisfying $X(B_{\kappa,b}) \neq 0$ for given block index b, and the sub-matrix $X(B_{\kappa,b})$ is a permutation matrix. If $k = \kappa$ holds, then

$$\sum_{(u,v)\in U^{(l)}_{k,b}} X_{u,v} = \sum_{(u,v)\in T^{(l)}_{k,b}} X_{u,v} = 1 \tag{39}$$

holds. Otherwise (i.e., $k \neq \kappa$), the equality

$$\sum_{(u,v)\in U^{(l)}_{k,b}} X_{u,v} = \sum_{(u,v)\in T^{(l \bmod \nu)+1}_{\kappa,b}} X_{u,v} = 1. \tag{40}$$

holds. Thus, it has been proved that (37) holds if X is a block permutation matrix.

We then move to the opposite direction; i.e., (37) implies that X is a block permutation matrix.

Assume that a block index $b \in [1,\gamma]$ and a subindex $l \in [1,\nu]$ are arbitrary chosen. Let $j = \nu(b-1)+l$. Since X is a permutation matrix, there exists the unique row index $i \in [1,n]$ satisfying $X_{i,j} = 1$. The block $B_{k,b}$ containing the set of indices (i,j) is uniquely determined because the blocks are mutually disjoint. Under this setting, it is clear that $X(B_{k,b}) \neq 0$ holds.

In the following, we will show that

$$k' \neq k \Rightarrow X(B_{k',b}) = 0. \tag{41}$$

From the definition of the block index k, It is clear that

$$\sum_{(u,v)\in T^{(l)}_{k,b}} X_{u,v} = 1 \tag{42}$$

holds. Combining Eq. (38) and Eq. (42), we immediately obtain

$$\sum_{k'\in[1,\gamma],k'\neq k} \left(\sum_{(u,v)\in T^{(l \bmod \nu)+1}_{k',b}} X_{u,v} \right) = 0. \tag{43}$$

This equality implies that

$$(u,v) \in \bigcup_{k'\in[1,\gamma]\setminus\{k\}} T^{(l \bmod \nu)+1}_{k',b} \Rightarrow X_{u,v} = 0. \tag{44}$$

Because X is a permutation matrix,

$$\sum_{(i,j)\in T^{(l \bmod \nu)+1}_{k,b}} = 1 \tag{45}$$

should be satisfied. Applying the same argument iteratively, we consequently have

$$(u,v) \in \bigcup_{k'\in[1,\gamma]\setminus\{k\}} \bigcup_{l'\in[1,\nu]} T^{(l')}_{k',b} \Rightarrow X_{u,v} = 0. \tag{46}$$

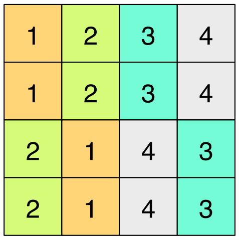
This statement is equivalent to $k' \neq k \Rightarrow X(B_{k',b}) = 0$. Due to the definition of the block permutation matrix, it has been proved that X should be a block permutation matrix. ■

The next example clarifies the linear constraints characterizing a 4×4 block permutation matrix.

Example 5: Let $n = 4, \nu = 2, \gamma = 2$. The necessary and sufficient condition for a permutation matrix $X \in \Pi_4$ being a block permutation matrix are as follows:

$$\begin{aligned} X_{1,1}+X_{2,1}+X_{3,2}+X_{4,2} &= 1 \\ X_{1,2}+X_{2,2}+X_{3,1}+X_{4,1} &= 1 \\ X_{1,3}+X_{2,3}+X_{3,4}+X_{4,4} &= 1 \\ X_{1,4}+X_{2,4}+X_{3,3}+X_{4,3} &= 1. \end{aligned} \tag{47}$$

Figure 4 illustrates the allocation of each linear constraint on a 4×4 array.



1	2	3	4
1	2	3	4
2	1	4	3
2	1	4	3

Fig. 4. Linear characterization of block permutation matrix ($n = 4, \nu = 2, \gamma = 2$): The variables with the same number label participate in a single linear equality.

C. Block permutation codes

Let us denote the set of block permutation matrices by

$$\Pi(n,\nu) \stackrel{\triangle}{=} \{X \in \Pi_n \mid X \text{ satisfies } (37)\}. \tag{48}$$

Note that we here employ a lighter notation $\Pi(n,\nu)$ instead of $\Pi(A,b,\trianglelefteq)$ since it explicitly express dependency on n and ν. It should be remarked that $\Pi(n,\nu)$ forms a group under matrix multiplication over $\mathbb{R}$.

The class of block permutation codes defined below is a class of LP decodable permutation codes.

Definition 11 (Block permutation code): Let n be a positive integer. A positive integer ν is a divisor of n. The initial vector s belongs to $\mathbb{R}^n$. The *block permutation code* $C(n,\nu,s)$ is defined by

$$C(n,\nu,s) \stackrel{\triangle}{=} \{Xs \in \mathbb{R}^n : X \in \Pi(n,\nu)\}. \tag{49}$$

□

A block permutation code can be decoded by LP decoding. We can use the linear equalities in Theorem 3 as the additional linear constraints $A \ \mathsf{vec}(X) \trianglelefteq b$ in LP decoding for LP decodable permutation codes. The size of the LP problem is the following. The number of variables is n^2 (i.e., $X_{i,j}$). The number of linear equality constraints is $2n + \gamma n$ because the number of linear equalities for X to be a doubly stochastic matrix is $2n$ and the number of additional linear equalities is $\gamma^2\nu = \gamma n$ for the block permutation constraint. The number of inequality constraints is n^2 corresponding to the constraint $X \geq 0$. In summary, the number of variables and constraints are bounded by $O(n^2)$. Therefore, if an interior point method is employed for solving this LP problem, we obtain a reasonably accurate solution in polynomial time.

D. Encoding of block permutation codes

A block permutation matrix induces a block-wise permutation if it applies to a vector. This block permutation structure supports a combinatorial encoding algorithm presented in this subsection.

We here write the initial vector s as

$$s = \begin{pmatrix} s_1 \\ \vdots \\ s_\gamma \end{pmatrix}, \tag{50}$$

where $s_1, s_2, \ldots, s_\gamma \in \mathbb{R}^\nu$. The next lemma explicitly indicates the block permutation structure of the block permutation codes.

Lemma 3 (Block permutation structure): Any codeword $Xs \in C(n,\nu,s)$ can be represented as

$$Xs = \begin{pmatrix} Q_1 s_{\sigma(1)} \\ \vdots \\ Q_\gamma s_{\sigma(\gamma)} \end{pmatrix}, \tag{51}$$

where $\sigma \in \mathcal{S}_\gamma$, $Q_1, \ldots, Q_\gamma \in \Pi_\nu$. The set $\mathcal{S}_\gamma$ is the set of bijections $\sigma : [1,\gamma] \to [1,\gamma]$.

Proof: Any codeword of $C(n,\nu,s)$ can be represented by Xs where X is a block permutation matrix. From the definition of the block permutation matrix, for any $b \in [1,\gamma]$, there exists a unique block index $k_b \in [1,\gamma]$ such that $X(B_{k_b,b})(k' \neq k_b)$ becomes a permutation matrix. We here introduce a map $\sigma : [1,\gamma] \to [1,\gamma]$ which is defined by $\sigma(k_b) = b$ for any $b \in [1,\gamma]$. Due to the definition of the block permutation matrix, σ is an injection. Furthermore, since σ is an injection between the sets of the identical cardinality, it should be a bijection.

On the basis of the above argument, it becomes clear that

$$Xs = \begin{pmatrix} X(B_{1,\sigma(1)})s_{\sigma(1)} \\ X(B_{2,\sigma(2)})s_{\sigma(2)} \\ \vdots \\ X(B_{\gamma,\sigma(\gamma)})s_{\sigma(\gamma)} \end{pmatrix} \tag{52}$$

holds for any $Xs \in C(n,\nu,s)$. By letting $Q_i \stackrel{\triangle}{=} X(B_{i,\sigma(i)})$, we have the claim of this lemma because Q_i is a permutation matrix. ■

It should be remarked that the number of codewords of $C(n,\nu,s)$ is $\gamma! \times (\nu!)^\gamma$ if $C(n,\nu,s)$ is non-singular. The block-wise permutation σ in (51) and the permutation corresponding to Q_i are called *inter-block and intra-block permutations*, respectively.

For a non-singular code $C(n,\nu,s)$, consider an encoding map $\phi : [1, \gamma! \times (\nu!)^\gamma] \to C(n,\nu,s)$. Note that the set $[1, \gamma! \times (\nu!)^\gamma]$ corresponds to the message space. Lemma 3 suggests that any efficient ranking algorithm (a bijection algorithm converting an integer into a permutation) for permutations can be used for encoding $C(n,\nu,s)$. For example, a simple ranking algorithm based on an inversion table is discussed in Sec. 5.1 of [25]. The encoding map ϕ is a bijection and it is invertible. This means that, from a codeword of $C(n,\nu,s)$, we can reconstruct the message corresponding to the codeword.

The details of an encoding algorithm for block permutation codes is given as follows.

Combinatorial encoding algorithm for $C(n,\nu,s)$

1) Convert a message $m \in [1, \gamma! \times (\nu!)^\gamma]$ into a message vector $(m_0, m_1, \ldots, m_\gamma) \in [1,\gamma!] \times [1,\nu!]^\gamma$.
2) By using a ranking algorithm for permutations denoted as a map F, compute permutations corresponding to the message vector:

$$(\sigma, Q_1, Q_2, \ldots, Q_\gamma) = F(m_0, m_1, \ldots, m_\gamma) \tag{53}$$

where $\sigma \in \mathcal{S}_\gamma, Q_i \in \Pi_\nu$.

3) Output
$$\begin{pmatrix} Q_1 s_{\sigma(1)} \\ \vdots \\ Q_\gamma s_{\sigma(\gamma)} \end{pmatrix} \tag{54}$$
as the codeword corresponding to the message m.

E. Minimum squared Euclidean distance of block permutation codes

In Section IV, we saw the minimum pseudo distance is one of most influential parameters for LP decoding performance. Unfortunately, the evaluation of the minimum pseudo distance is not a trivial problem. As a possible alternative, we here evaluate the minimum squared Euclidean distance of $C(n, \nu, s)$ defined by
$$d^2_{min} \triangleq \min_{x,y \in C(n,\nu,s)(x \neq y)} ||x - y||^2. \tag{55}$$
At least, we can say that decoding performance degrades even with an ML decoder if $C(n, \nu, s)$ has small d^2_{min}.

The block-wise permutation structure can be exploited for deriving a simple lower bound on the minimum squared Euclidean distance.

Theorem 4 (Minimum squared Euclidean distance): Let us define Δ_1^2 and Δ_2^2 by
$$\begin{aligned} \Delta_1^2 &\triangleq \min_{\sigma \in \mathcal{S}_\gamma} \min_{Q_1,Q_2 \in \Pi_\nu (Q_1 \neq Q_2)} ||Q_1(s_{\sigma(1)}) - Q_2(s_{\sigma(1)})||^2 \\ \Delta_2^2 &\triangleq \min_{\sigma \in \mathcal{S}_\gamma} \min_{Q_1,Q_2 \in \Pi_\nu} ||Q_1(s_{\sigma(1)}) - Q_2(s_{\sigma(2)})||^2. \end{aligned} \tag{56}$$
Assume that both Δ_1^2 and Δ_2^2 are positive for given n, ν, s. In such a case, $C(n, \nu, s)$ is non-singular and the minimum squared Euclidean distance of $C(n, \nu, s)$ is given by
$$d^2_{min} = \min\{\Delta_1^2, 2\Delta_2^2\}. \tag{57}$$

Proof: Assume that arbitrary two distict codewords $c^a, c^b \in C(n, \nu, s)(c^a \neq c^b)$ are given. Lemma 3 guarantees that c^a and c^b can be expressed as
$$c^a = \begin{pmatrix} Q_1^a s_{\sigma^a(1)} \\ \vdots \\ Q_\gamma^a s_{\sigma^a(\gamma)} \end{pmatrix}, \quad c^b = \begin{pmatrix} Q_1^b s_{\sigma^b(1)} \\ \vdots \\ Q_\gamma^b s_{\sigma^b(\gamma)} \end{pmatrix}, \tag{58}$$
where $Q_1^a, \ldots, Q_\gamma^a, Q_1^b, \ldots, Q_\gamma^b \in \Pi_\nu$, $\sigma^a, \sigma^b \in \mathcal{S}_\gamma$. The assumption $c^a \neq c^b$ implies
$$(\sigma^a, \pi_1^a, \ldots, \pi_\gamma^a) \neq (\sigma^b, \pi_1^b, \ldots, \pi_\gamma^b). \tag{59}$$
In the following, we will consider the following two cases.

1) Case A: $\sigma^a = \sigma^b$
In this case, there exists at least one block index $k' \in [1, \gamma]$ which satisfies $Q_{k'}^a \neq Q_{k'}^b$. The squared Euclidian distance between c^a and c^b can be lower bounded by
$$\begin{aligned} ||c^a - c^b||^2 &= \sum_{k=1}^{\gamma} ||Q_k^a s_{\sigma^a(k)} - Q_k^b s_{\sigma^b(k)}||^2 \\ &\geq ||Q_{k'}^a s_{\sigma^a(k')} - Q_{k'}^b s_{\sigma^b(k')}||^2 \\ &\geq \Delta_1^2. \end{aligned} \tag{60}$$

2) Case B: $\sigma^a \neq \sigma^b$
At least two block indices $\xi, \eta \in [1, \gamma]$ $(\xi \neq \eta)$ satisfies
$$\sigma^a(\xi) \neq \sigma^b(\xi), \quad \sigma^a(\eta) \neq \sigma^b(\eta). \tag{61}$$
This relations lead to the following lower bound:
$$\begin{aligned} ||c^a - c^b||^2 &= \sum_{k=1}^{\gamma} ||Q_k^a s_{\sigma^a(k)} - Q_k^b s_{\sigma^b(k)}||^2 \\ &\geq ||Q_\xi^a s_{\sigma^a(\xi)} - Q_\xi^b s_{\sigma^b(\xi)}||^2 \\ &+ ||Q_\eta^a s_{\sigma^a(\eta)} - Q_\eta^b s_{\sigma^b(\eta)}||^2 \\ &\geq 2\Delta_2^2. \end{aligned} \tag{62}$$

Due to inequalities (60) and (62), we have $d^2_{min} \geq \min\{\Delta_1^2, 2\Delta_2^2\}$. Thus, the assumption $\Delta_1^2 > 0, \Delta_2^2 > 0$ guarantees that $C(n, \nu, s)$ is non-singular. It is clear that there exists pairs of codewords satisfying (60) and (62) with equality. Thus, we obtain the claim of the theorem. ■

Example 6: Let $n = 6, \gamma = 3, \nu = 2$. The initial vector s is assumed to be $s = (4\ 1\ 5\ 2\ 6\ 3)^T$. In this case, we have
$$s_1 = \begin{pmatrix} 4 \\ 1 \end{pmatrix}, \quad s_2 = \begin{pmatrix} 5 \\ 2 \end{pmatrix}, \quad s_3 = \begin{pmatrix} 6 \\ 3 \end{pmatrix}$$
From the definition of Δ_1^2, Δ_2^2, we easily obtain $\Delta_1^2 = 18$, $\Delta_2^2 = 2$. Applying Theorem 4, we have $d^2_{min} = \min\{18, 2 \times 2\} = 4$. The number of codewords is $\gamma! \times (\nu!)^\gamma = 48$ □

It should be remarked that Δ_1^2 and Δ_2^2 have a simpler form:
$$\begin{aligned} \Delta_1^2 &= \min_{k \in [1,\gamma]} \min_{Q \in \Pi_\nu (Q \neq I)} ||s_k - Q(s_k)||^2 \\ \Delta_2^2 &= \min_{k,j \in [1,\gamma](k \neq j)} \min_{Q \in \Pi_\nu} ||s_k - Q(s_j)||^2. \end{aligned} \tag{63}$$
The derivation of this simpler form is based on the group structure of Π_ν and on the norm preservation property of permutation matrices.

The order of elements in s is highly influential to the minimum distance. The following corollary gives an appropriate ordering in terms of the minimum squared Euclidean distance. The idea is very similar to *set partitioning* for constructing a good trellis-coded modulation scheme due to Ungerboeck [23].

Corollary 1 (Set partitioning): Assume that the initial vector s is given by
$$\begin{aligned} s_1 &\triangleq (1 + (\nu - 1)\gamma, \ldots, 1 + 2\gamma, 1 + \gamma, 1)^T \\ s_2 &\triangleq (2 + (\nu - 1)\gamma, \ldots, 2 + 2\gamma, 2 + \gamma, 2)^T \\ &\vdots \\ s_\gamma &\triangleq (\gamma + (\nu - 1)\gamma, \ldots, \gamma + 2\gamma, \gamma + \gamma, \gamma)^T. \end{aligned}$$
In this case, we have $d^2_{min} = \min\{2\gamma^2, 2\nu\}$.

Proof: It is easy to see that $\Delta_1^2 = 2\gamma^2, \Delta_2^2 = \nu$. By using Theorem 4, we immediately have the claim of corollary. ■

F. An extension of block permutation codes

Theorem 3 states that any codeword of a block permutation code has the form in Eq. (51). A natural way to extend the block permutation codes is to introduce a restriction on the inter-block and the intra-block permutations. Namely, we can choose the inter-block and intra-block permutations as

$$\sigma \in \mathcal{T}_\gamma,\ Q_1, \ldots, Q_\gamma \in P_\nu,$$

where $\mathcal{T}_\gamma$ is a subset of $\mathcal{S}_\gamma$ and P_ν is a subset of Π_ν. This restriction reduces the number of codewords but it may improve the decoding performance. Combining linear constraints for block permutation matrices and additional linear constraints for $\mathcal{T}_\gamma$ and P_ν, it is expected that such an extended block permutation code is LP decodable.

A generalization of Theorem 4 is straightforward; we only need to replace Δ_1^2 and Δ_2^2 by

$$\begin{aligned}
\Delta_1^{\prime 2} &\overset{\triangle}{=} \min_{\sigma \in \mathcal{T}_\gamma} \min_{Q_1, Q_2 \in P_\nu (Q_1 \neq Q_2)} ||Q_1(s_{\sigma(1)}) - Q_2(s_{\sigma(1)})||^2 \\
\Delta_2^{\prime 2} &\overset{\triangle}{=} \min_{\sigma \in \mathcal{T}_\gamma} \min_{Q_1, Q_2 \in P_\nu} ||Q_1(s_{\sigma(1)}) - Q_2(s_{\sigma(2)})||^2 \qquad (64)
\end{aligned}$$

in order to obtain a generalized result.

Appropriate choice for $\mathcal{T}_\gamma$ and Π_ν to construct a good permutation code is an interesting open problem.

VI. Randomly constrained permutation matrices

In the previous section, we discussed a set of structured permutation matrices. Another possible choice for linear constraints is to generate them randomly. Such random linear constraints are amenable for probabilistic analysis and appears interesting from information theoretic view. In this section, we study a class of LP decodable permutation codes defined based on random constraints.

A. Sparse constraint matrix ensemble

Since the LP decodable permutation codes are non-linear codes, the cardinality of a given code cannot be determined directly from the constraints in general. In the following part of this section, we will analyze the cardinality of codes and their Hamming weight distributions. A sparse constraint matrix ensemble is assumed in the following analysis, which has a close relationship to the analysis on average weight distribution of LDPC ensembles [12].

Let S be the set of binary constraint matrices:

$$S \overset{\triangle}{=} \{A \in \{0,1\}^{m \times n^2} : \text{every row of } A \text{ contains } r\text{-ones}\}. \qquad (65)$$

We assign the uniform probability

$$P(A) \overset{\triangle}{=} \frac{1}{\binom{n^2}{r}^m} \qquad (66)$$

to each matrix in S. The pair (S, P) can be considered as an ensemble of matrices, which becomes the basis of the following probabilistic method.

Let α be a positive integer. In the following, we focus on the linearly constrained permutation code $\Lambda(A, \alpha\mathbf{1}, \trianglelefteq, s)$, where $A \in S$ and $\trianglelefteq = (\overbrace{\leq, \leq \ldots, \leq}^{m})$. The symbol $\mathbf{1}$ denotes the vector of length m whose entries are all ones. Extensions of the analysis for more general classes of LP decodable permutation codes are possible, but we here focus on the simplest class to explain the idea of the analysis. Throughout this section, we assume that components of the initial vector s differ each other.

B. Probabilistic analysis on average cardinality of codes

The number of codewords in $\Lambda(A, \alpha\mathbf{1}, \trianglelefteq, s)$ is given by

$$M(A) \overset{\triangle}{=} \sum_{X \in \Pi_n} \mathbb{I}[A\ \mathsf{vec}(X) \trianglelefteq \alpha\mathbf{1}], \qquad (67)$$

where $\mathbb{I}$ is the indicator function. The indicator function takes the value one when the given condition is true and otherwise gives the value zero. The next lemma gives the average cardinality of this code.

Lemma 4 (Average cardinality of codes): The average cardinality of $\Lambda(A, \alpha\mathbf{1}, \trianglelefteq, s)$ is given by

$$\mathsf{E}[M(A)] = n! \left(\frac{1}{\binom{n^2}{r}} \sum_{i=0}^{\alpha} \binom{n}{i} \binom{n^2 - n}{r - i} \right)^m, \qquad (68)$$

where the operator E denotes the expectation defined on (S, P).

Proof: From the definition of $M(A)$, the expectation of the cardinality $M(A)$ can be written as

$$\begin{aligned}
\mathsf{E}[M(A)] &= \sum_{A \in S} P(A) M(A) \\
&= \sum_{A \in S} P(A) \sum_{X \in \Pi_n} I[A\ \mathsf{vec}(X) \trianglelefteq \alpha\mathbf{1}]. \quad (69)
\end{aligned}$$

By changing the order of summation, we can further transform this into

$$\begin{aligned}
\mathsf{E}[M(A)] &= \sum_{X \in \Pi_n} \sum_{A \in S} P(A) I[A\ \mathsf{vec}(X) \trianglelefteq \alpha\mathbf{1}] \\
&= \frac{n!}{\binom{n^2}{r}^m} \sum_{A \in S} I[A\ \mathsf{vec}(X') \trianglelefteq \alpha\mathbf{1}], \qquad (70)
\end{aligned}$$

where X' is an arbitrary permutation matrix in Π_n. The last equality is due to the symmetry of the ensemble. Namely, this means that the quantity $\sum_{A \in S} I[A\ \mathsf{vec}(X') \trianglelefteq \alpha\mathbf{1}]$ does not depend on the choice of X'. The evaluation of $\sum_{A \in S} I[A\ \mathsf{vec}(X') \trianglelefteq \alpha\mathbf{1}]$ can be performed on the basis of the following combinatorial argument. It is evident that any $X' \in \Pi_n$ contains n-ones as its components. This implies that $x' \overset{\triangle}{=} \mathsf{vec}(X')$ is a binary vector of length n^2 with Hamming weight n. Let $I_1 \overset{\triangle}{=} \{i \in [1, n^2] \mid x'_i = 1\}$, where x'_i is the ith element of x'. Consider the first row of A, which is denoted by a^T. The relation $a^T x' \leq \alpha$ holds if and only if

$$|\{i \in I_1 \mid a_i = 1\}| \leq \alpha. \qquad (71)$$

The number of possible ways to choose such a vector a is given by

$$\sum_{i=0}^{\alpha} \binom{n}{i} \binom{n^2 - n}{r - i}.$$

The term $\binom{n}{i}\binom{n^2-n}{r-i}$ corresponds to the number of possible ways such that I_1 (of cardinality n) contains i-ones and the other indices (of cardinality n^2-n) contain $(r-i)$-ones. Since each row of A can be chosen independently, we consequently have

$$\sum_{A \in S} I[A\ \mathsf{vec}(X') \trianglelefteq \alpha \mathbf{1}] = \left(\sum_{i=0}^{\alpha} \binom{n}{i} \binom{n^2-n}{r-i} \right)^m . \quad (72)$$

Substituting (72) into (70), we immediately obtain the claim of the lemma. ■

C. Probabilistic analysis on weight distribution

The origin $o \overset{\triangle}{=} (o_1, \ldots, o_n)$ is an arbitrary permutation vector of length n; namely, $o \in \Lambda(s)$. The number of codewords of $\Lambda(A, \alpha\mathbf{1}, \trianglelefteq, s)$ with Hamming weight w is denoted by $L_w(A)$, where the Hamming weight $w_H(\cdot)$ is defined by

$$w_H(x) \overset{\triangle}{=} \sum_{i=1}^{n} \mathbb{I}[o_i \neq x_i], \quad (73)$$

where $x = (x_1, \ldots, x_n)$. This means the Hamming weight of x is equal to the Hamming distance between the origin and x. In other words, $L_w(A)$ is defined as

$$L_w(A) \overset{\triangle}{=} \sum_{x \in \Lambda(A, \alpha\mathbf{1}, \trianglelefteq, s)} \mathbb{I}[w_H(x) = w]. \quad (74)$$

The set $\{L_1(A), \ldots, L_n(A)\}$ is referred to as the weight distribution of $\Lambda(A, \alpha\mathbf{1}, \trianglelefteq, s)$.

The weight distribution indicates a geometric property of $\Lambda(A, \alpha\mathbf{1}, \trianglelefteq, s)$. The next lemma gives the ensemble average of the weight distribution.

Lemma 5: The average weight distribution of the linearly constrained permutation code $\Lambda(A, \alpha\mathbf{1}, \trianglelefteq, s)$ is given by

$$\mathsf{E}[L_w(A)] = \binom{n}{w} \left\lfloor \frac{w!+1}{e} \right\rfloor \left(\frac{1}{\binom{n^2}{r}} \sum_{i=0}^{\alpha} \binom{n}{i} \binom{n^2-n}{r-i} \right)^m . \quad (75)$$

Proof: The weight distribution $L_w(A)$ can also be expressed as

$$L_w(A) = \sum_{X \in Z_w(o)} \mathbb{I}[A\ \mathsf{vec}(X) \trianglelefteq \alpha\mathbf{1}], \quad (76)$$

where $Z_w(o)$ is defined by

$$Z_w(o) \overset{\triangle}{=} \{X \in \Pi_n : w_H(Xs) = w\}. \quad (77)$$

The expectation can be simplified as follows:

$$\begin{aligned} \mathsf{E}[L_w(A)] &= \sum_{A \in S} P(A) \sum_{X \in Z_w(o)} \mathbb{I}[A\ \mathsf{vec}(X) \trianglelefteq \alpha\mathbf{1}] \\ &= \frac{1}{\binom{n^2}{r}^m} \sum_{X \in Z_w(o)} \sum_{A \in S} \mathbb{I}[A\ \mathsf{vec}(X) \trianglelefteq \alpha\mathbf{1}] \\ &= \left(\frac{1}{\binom{n^2}{r}} \sum_{i=0}^{\alpha} \binom{n}{i} \binom{n^2-n}{r-i} \right)^m |Z_w(o)|. \end{aligned} \quad (78)$$

The last equality is due to the symmetry of the ensemble and equation (72).

The cardinality of $Z_w(o)$ is given by the following combinatorial argument. Let $x \in \Lambda(s)$ be an arbitrary vector satisfying $w_H(x) = w$. The index set I_{diff} is defined by $I_{diff}(x) = \{i \in [1, n] \mid o_i \neq x_i\}$. Let $T \subset [1, n]$ be an index set of cardinality w. The quantity $|\{x \in \Lambda(s) \mid T = I_{diff}(x)\}|$ is equal to the number of derangements of length w, which is known to be $\lfloor (w!+1)/e \rfloor$ [33] [39]. Note that the number of possible ways to choose T is $\binom{n}{w}$. Thus, we have the equality

$$|Z_w(o)| = \binom{n}{w} \left\lfloor \frac{w!+1}{e} \right\rfloor . \quad (79)$$

This completes the proof of the lemma. ■

VII. Conclusion

In this paper, a novel class of permutation codes, LP decodable permutation codes, is introduced. The LP decodable property is the main feature of this class of permutation codes.

The set of doubly stochastic matrices, i.e., the Birkhoff polytope, have $n!$ integral vertices which are permutation matrices. Additional linear constraints defines a code polytope which plays a fundamental role in the coding scheme presented in this paper. An LP decodable permutation code is the set of integral vertices of a code polytope.

In an LP decoding process, a certain linear objective function is maximized under the assumption that the feasible set is a code polytope. The decoding performance can be evaluated from geometrical properties of a code polytope.

The choice of additional linear constraints are crucial to construct good codes. In this paper, two approaches are discussed; namely, structured permutation matrices and randomly constrained permutation matrices.

The key result of Section V is that block permutation matrices can be characterized by a set of linear equalities. The group structure of the block permutation matrices is useful for efficient encoding. Bridging an algebraic property (group structure) and a geometric property (code polytope) appears a first step towards a novel paradigm for the study of permutation codes.

The random constraints discussed in Section VI enable us to use probabilistic methods for analyzing some sproperties of codes. The probabilistic methods [26] are very powerful tool for grasping the relation between the number of constraints and important code parameters such as the cardinality of a code.

Although the paper provides fundamental aspects of the LP decodable permutation codes, many problems remain still open. The following list is a part of open problems.

1) Construction of good block permutation codes; choice of an initial vector, block size, and inter-block and intra-block permutations.
2) Efficient algorithm for solving the LP problem arising in the LP decoding.
3) Permutation modulation for linear vector channels; Let H be a $n \times n$ real matrix. An ML decoding problem for

a linear vector channel can be formulated as

$$\text{minimize } ||y - Hx||^2 \text{ subject to } x \in \Lambda(A, b, \trianglelefteq, s). \quad (80)$$

As discussed in this paper, the decoding problem can be relaxed to a quadratic programming (QP) problem:

$$\text{minimize } ||y - Hx||^2 \text{ subject to } x \in \mathcal{P}(A, b, \trianglelefteq, s). \quad (81)$$

A QP-based decoding algorithm like [31] appears interesting for this problem.

4) Development of a theory for convex optimization over a set of permutation matrices; see also Appendix.

Further investigation on related topics may open an interdisciplinary research field among coding, combinatorial optimization and algebra.

Appendix

In this paper, we considered optimization problems defined on a set of permutation matrices. The first part of Appendix provides an abstract framework for this class of optimization problems and show a relaxation approach to solve them. The second part discusses convex relaxations of a set of permutation matrices.

A. *convex optimization over a set of permutation matrices*

Let $Q \subset \Pi_n$ be a set of permutation matrices. Assume that $f : \mathbb{R}^{n \times n} \to \mathbb{R}$ be a convex function. Our target problem is assumed to be

$$P_1 : \quad \text{minimize } f(X) \text{ subject } X \in Q. \quad (82)$$

In this setting, the feasible set becomes a discrete set. Thus, the problem P_1 can be considered as a combinatorial problem. Of course, in general, it is very hard to solve P_1. We here consider a relaxed problem of P_1:

$$P_2 : \quad \text{minimize } f(X) \text{ subject } X \in \mathsf{conv}(Q), \quad (83)$$

where $\mathsf{conv}(Q)$ denotes the convex hull of Q. Note that $f(X)$ is a convex function with respect to X and the feasible set $\mathsf{conv}(Q)$ is a convex set. This means that P_2 is a convex programming problem. Therefore, P_2 can be efficiently solved with an interior point method. In general, the solution of P_2 and P_1 are not equal. However, it is expected that the solution of P_2 is a good approximation of the solution of P_1.

In many cases, $\mathsf{conv}(Q)$ is not easy to handle. For example, in some cases, finding linear equalities and inequalities expressing $\mathsf{conv}(Q)$ is a non-trivial problem. In such cases, it is useful to consider a relaxation of $\mathsf{conv}(Q)$.

A relaxed polytope $\mathcal{P}$ for $\mathsf{conv}(Q)$ should satisfy the following conditions.

1) $\mathsf{conv}(Q) \subset \mathcal{P}$
2) The set of integral points in $\mathcal{P}$ coincides with Q.
3) The number of linear constraints expressing $\mathcal{P}$ is much smaller than that of $\mathsf{conv}(Q)$.

Base on a relaxed polytope $\mathcal{P}$, a relaxed problem

$$P_3 : \quad \text{minimize } f(X) \text{ subject } X \in \mathcal{P}$$

can be defined. Note that a code polytope defined in this paper does not coincide with the convex hull of an LP decodable permutation code in general. This means that the LP decoding can be considered as an instance of P_3.

B. *Convex relaxation of a set of permutation matrices*

We here discuss convex relaxations of some sets of permutation matrices such as the set of involutions, permutation matrices for cyclic group. Some results presented here give us insight for the convex relaxation of permutation matrices and may become a start point for further progress on structured LP decodable permutation codes.

1) Notation: We here refer the following constraints for doubly stochastic matrix

$$X1 = 1, \quad 1^T X = 1^T, \quad X \geq 0 \quad (84)$$

as the *basic constraints*. Let $\mathcal{A}$ be a set of linear constraints including the basic constraints and additional linear constraints (e.g., $\mathsf{trace}(X) = 0$). In this section, the relaxed polytope corresponding to $\mathcal{A}$ is denoted by

$$\mathcal{P}(\mathcal{A}) \overset{\triangle}{=} \{X \in \mathbb{R}^{n \times n} : X \text{ satisfies all constraints in } \mathcal{A}\}. \quad (85)$$

The set of permutation matrices satisfying the constraints in $\mathcal{A}$ is denoted by

$$\Pi(\mathcal{A}) \overset{\triangle}{=} \{X \in \Pi_n : X \text{ satisfies all constraints in } \mathcal{A}\}. \quad (86)$$

It is known that an LP solution is achieved on a vertex of a polytope in most cases. Therefore, it is desirable that the relaxed polytope $\mathcal{P}(\mathcal{A})$ coincides with the the convex hull of $\Pi(\mathcal{A})$ to avoid an fractional solution. If the equality $\mathcal{P}(\mathcal{A}) = \mathsf{conv}\,(\Pi(\mathcal{A}))$ holds, then the relaxed polytope (or code polytope) is said to be *tight*. In other words, a tight polytope, called a tight polytope in this paper, is a polytope whose vertices are all integral. In some cases, linearly dependent linear constraints are useful because such redundant constraints tighten the relaxation.

2) Convex relaxations for $n = 4$: Table I presents linear constraints for some sets of permutation matrices and their tightness of corresponding relaxed polytopes. In this table, it is assumed that $X \in \mathbb{R}^{4 \times 4}$. The tightness is numerically checked with the vertex enumeration program cdd based on double description method by K. Fukuda [32].

Some remarks on Table I are listed as follows.

1) Cyclic group of order 4: The cyclic permutation matrices of order 4 is given by the following additional linear constraints:

$$\begin{aligned} X_{1,1} &= X_{2,2},\ X_{2,2} = X_{3,3},\ X_{3,3} = X_{4,4} \\ X_{2,1} &= X_{3,2},\ X_{3,2} = X_{4,3},\ X_{4,3} = X_{1,4} \\ X_{3,1} &= X_{4,2},\ X_{4,2} = X_{1,3},\ X_{1,3} = X_{2,4} \\ X_{4,1} &= X_{1,2},\ X_{1,2} = X_{2,3},\ X_{2,3} = X_{3,4}. \end{aligned} \quad (87)$$

The arrangement of equalities in 4×4 array is depicted in Fig. 5. In a similar way as in the case $n = 4$, we can define the cyclic permutation matrices of order n. We

TABLE I

TIGHTNESS OF RELAXED POLYTOPES FOR SOME SETS OF PERMUTATION MATRICES ($n = 4$)

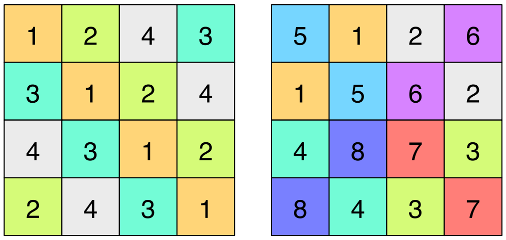
set of perm. matrices	additional constraints	tightness	$\#V$
symmetric group S_4	none	Y	24
cyclic group C_4	(87)	Y	4
derangement	$\mathsf{trace}(X) = 0$	Y	9
involution	$X = X^T$	N	14
transposition (1)	$\mathsf{trace}(X) = n - 2$	N	20
transposition (2)	$\mathsf{trace}(X) = n - 2$ $X = X^T$	Y	6
Klein four group	(89)	Y	4
dihedral group D_8	(90)	Y	8
2×2 block	constraints (47)	N	28
2×2 block	constraints (47) and (91)	Y	8

The column of tightness (Y/N) represents the relaxed polytope is tight (Y) or not (N). The column $\#V$ denotes the number of vertices on the relaxed polytope.

conjecture that this type of linear constraints for cyclic permutation matrices, i.e.,

$$\forall i, j \in [1, n], \quad X_{i,j} = X_{(i \bmod n)+1,(j \bmod n)+1}, \tag{88}$$

give a tight polytope. We confirmed that this linear constraints also give the tight polytope when $n = 5$.



1	2	4	3
3	1	2	4
4	3	1	2
2	4	3	1

Cyclic permutation matrix

5	1	2	6
1	5	6	2
4	8	7	3
8	4	3	7

Dihedral group with order 8

Fig. 5. Constraints for cyclic permutation matrices and dihedral group with order 8; The variables with the same number label constrained to be the same value.

2) Transposition: The permutation matrices satisfying the linear constraint $\mathsf{trace}(X) = n - 2$ exactly coincides with the set of transpositions (i.e., permutations of two elements). Note that the constraint $\mathsf{trace}(X) = n - 2$ does not give the tight polytope. Combining a redundant constraint $X = X^T$ (i.e., the involution constraint) to the trace constraint, the relaxed polytope becomes tight. This example indicates that redundant constraints are necessary for constructing a tight polytope in some cases.

3) Klein four-group: The Klein four-group is isomorphic to $\mathbb{Z}_2 \times \mathbb{Z}_2$. This group structure imposes block-wise diagonal structure on a permutation matrix. The convex hull is given by the basic constraints and the following constraints:

$$\begin{aligned} X_{2,1} &= X_{1,2}, \quad X_{4,3} = X_{3,4}, \\ X_{1,3} &= X_{1,4,} = X_{2,3} = X_{2,4} = 0, \\ X_{3,1} &= X_{3,2} = X_{4,1} = X_{4,2} = 0. \end{aligned} \tag{89}$$

4) Dihedral group D_8: A set of permutation matrices corresponding to the dihedral group with order 8 is given by the set of equalities (see also Fig. 5):

$$\begin{aligned} & X_{1,2} = X_{2,1}, X_{1,3} = X_{2,4}, X_{3,1} = X_{4,2}, \\ & X_{3,4} = X_{4,3}, X_{1,1} = X_{2,2}, X_{1,4} = X_{2,3}, \\ & \qquad X_{4,1} = X_{3,2}, X_{4,4} = X_{3,3}. \end{aligned} \tag{90}$$

Symmetry (90 degree rotation around the center) of the allocation of the equality constraints can be observed. It suggests that linear constraints for a set of permutation matrices forming a group may have this type of symmetric structure.

5) Block constraint: The linear constraints for block permutation matrices (47) introduced in Theorem 3 does not give the tight polytope in $n = 4$. However, combining (47) and a set of redundant constraints (i.e., 90 degree rotation of (47))

$$\begin{aligned} X_{1,1} + X_{1,2} + X_{2,3} + X_{2,4} &= 1 \\ X_{2,1} + X_{2,2} + X_{1,3} + X_{1,4} &= 1 \\ X_{3,1} + X_{3,2} + X_{4,3} + X_{4,4} &= 1 \\ X_{4,1} + X_{4,2} + X_{3,3} + X_{3,4} &= 1, \end{aligned} \tag{91}$$

we have the convex hull of 2×2 block permutation matrices. This case also shows importance of redundant constraints from the optimization perspective. From this result, it is expected that the LP decoding performance of block permutation codes might be improved by incorporating this redundant linear equalities.

We here presented convex relaxations of some sets of permutation matrices. Further exploration on this topic including rigorous proof of the above mentioned conjectures appears interesting not only from engineering point of view but also from mathematical point of view.

Acknowledgement

The first author appreciates Prof. Han Vinck for directing the author's attention to the field of permutation codes. The first author also wishes to thank Dr. Jun Muramatsu for inspiring discussions on permutations.

REFERENCES

[1] I. F. Blake, G. Cohen, and M. Deza, "Coding with permutations," Inform. Contr., vol. 43, pp. 1–19, 1979.

[2] J.C. Chang, R.J. Chen, T. Kløve and S.C. Tsai, "Distance-preserving mappings from binary vectors to permutations, " IEEE Transactions on Information Theory, vol. 49, no.4, pp.1054-1059, Apr. 2003.

[3] J.C. Chang, "Distance-increasing mappings from binary vectors to permutations," IEEE Transactions on Information Theory, vol.51, pp.359-363, Jan. 2005.

[4] C. J. Colbourn, T. Kløve, and A. C. H. Ling, "Permutation arrays for powerline communication and mutually orthogonal latin squares, " IEEE Transactions on Information Theory, vol. 54, No. 6, June, 2004.

[5] G.D. Forney, Jr., R. Koetter, F.R. Kschischang, and A. Reznik, “On the effective weights of pseudocodewords for codes defined on graphs with cycles, ” in Codes, Systems, and Graphical Models (B. Marcus and J. Rosenthal, eds.), vol. 123 of IMA Vol. Math. Appl., pp. 101-112, Springer Verlag, New York, Inc., 2001.

[6] C. Ding, F. W. Fu, T. Kløve and V. K. W. Wei, “Constructions of permutation arrays,” IEEE Transactions on Information Theory, vol. 48, no. 4, Apr. 2002.

[7] J. Feldman, “Decoding error-correcting codes via linear programming,” Massachusetts Institute of Technology, Ph. D. thesis, 2003.

[8] P. Frankl and M. Deza, “On the maximum number of permutations with given maximal and minimal distance,” J. Comb. Theory, Ser. A, vol. 22, pp. 352–360, 1977.

[9] A. Jiang, R. Mateescu, M. Schwartz, and J. Bruck, “Rank modulation for flash memories,” in Proc. IEEE Int. Symp. Information Theory, 2008.

[10] A. Jiang, M. Schwartz, and J. Bruck, “Error-correcting codes for rank modulation,” in Proc. IEEE Int. Symp. Information Theory, 2008.

[11] T. Kløve, T. Lin, S.-C. Tsai, and W. G. Tzeng, “Permutation arrays under the Chebyshev distance,” IEEE Transactions on Information Theory, vol. 56, no. 6, June 2010.

[12] S.Litsyn and V. Shevelev, “On ensembles of low-density parity-check codes: asymptotic distance distributions,” *IEEE Trans. Inform. Theory*, vol.48, pp.887–908, Apr. 2002.

[13] A. J. H. Vinck, “Coded modulation for powerline communications, ” AEÜ Int. J. Electron. Commun., vol. 54, pp. 45–49, Jan. 2000.

[14] A. J. H. Vinck, J. Häring, and T. Wadayama, “Coded M-FSK for powerline communications,” in Proc. IEEE Int. Symp. Information Theory, 2000.

[15] A. J. H. Vinck, and H.C. Ferreira, “Permutation trellis-codes, ” in Proc. IEEE Int. Symp. Information Theory, 2001.

[16] T. Wadayama and A.J.Han Vinck, “A multilevel construction of permutation codes,” IEICE Transactions on Fundamentals, vol.E84-A, no.10, pp.2518–2522, 2001.

[17] D. Slepian, “Permutation modulation” ,Proc. IEEE, pp. 228-236, 1965.

[18] J. Karlof, “Permutation codes for the Gaussian channel,” IEEE Trans. Inform. Theory, vol. 35, no. 4, pp. 726-732, July 1989.

[19] E. Biglieri and M. Elia, “Optimum permutation modulation codes and their asymptotic performance,” IEEE Trans. Inform. Theory, vol. IT-22, no. 6, Nov. 1976.

[20] I. Ingemarsson, “Optimized permutation modulation,” IEEE Trans. Inform. Theory, vol. 36, pp. 1098-1100, Sept. 1990.

[21] T. Berger, F. Jelinek, and J. K. Wolf, “Permutation codes for sources,” IEEE Trans. Inform. Theory, vol. IT-18, pp. 160-169, Jan. 1972.

[22] D. Slepian, “Group codes for the Gaussian channel,” Bell Syst. Tech. J., vol. 47, pp. 575-602, Apr. 1968.

[23] G. Ungerboeck, “Channel coding with multilevel/phase signals,” IEEE Trans. Itform. Theorv,vol.IT-28,Jan. 1982.

[24] A. Barg and A. Mazumdar, “Codes in permutations and error correction for rank modulation,” in Proc. IEEE Int. Symp. Information Theory, 2010.

[25] D. Knuth “The Art of Computer Programming Volume 3, ” Addison-Wesley, 1998.

[26] N. Alon and J. H. Spencer, “The Probabilistic Method, 3rd. ed.,” John Wiley & Sons, 2008.

[27] R. Koetter and P. O. Vontobel, “Graph covers and iterative decoding of finite-length codes”, in *Proc. 3rd Int. Symp. Turbo Codes and Related Topics*, Brest, France, Sep. 2003.

[28] D. P. Bertsekas, “Nonlinear programming, ” 2nd edition, Athena Scientific, 1999.

[29] S. Boyd and L. Vandenberghe, “Convex optimization,” Cambridge University Press, 2004.

[30] A. Schrijver, “Combinatorial optimization: polyhedra and efficiency,” Springer, 2003.

[31] T. Wadayama, “Interior point decoding for linear vector channels based on convex optimization,” IEEE Trans. Inform. Theory, pp.4905-4921, vol.56, no.10, Oct. (2010)

[32] K. Fukuda, “cdd and cddplus Homepage, ” `http://www.ifor.math.ethz.ch/~fukuda/cdd_home/`

[33] “The On-Line Encyclopedia of Integer Sequences, A000166”, `http://oeis.org/A000166`

[34] G. M. Ziegler, “Lectures on Polytopes, ” Springer-Verlag New York, 1995.

[35] G. Birkhoff, “Three observations on linear algebra,” Univ. Nac. Tacuman, Rev. Ser. A 5, 147?151, 1946.

[36] J. von Neumann, “A certain zero-sum two-person game equivalent to an optimal assignment problem,” Ann. Math. Studies 28, 5?12, 1953.

[37] A. Schrijver, “Combinatorial optimization, polyhedra and efficiency,” Springer-Verlag Berlin, 2003.

[38] M. Hassani, “Counting and computing by e, ”, arXiv:math/0606613v1, 2006

[39] G. Hurlbert, “A short proof of the Birkhoff-von Neumann Theorem, ” `http://mingus.la.asu.edu/~hurlbert/papers/SPBVNT.pdf`, 2008