IEEE TRANSACTIONS ON
INFORMATION THEORY

IEEE Xplore® Digital Library
http://ieeexplore.ieee.org/Xplore

Theoretical Bounds and Constructions of Codes in the Generalized Cayley Metric

Siyi Yang, *Student Memeber, IEEE,* Clayton Schoeny, *Student Memeber, IEEE,* and Lara Dolecek, *Senior Membor, IEEE*

***Abstract*—Permutation codes have recently garnered substantial research interest due to their potential in various applications, including cloud storage systems, genome resequencing, and flash memories. In this paper, we study the theoretical bounds and constructions of permutation codes in the generalized Cayley metric. The generalized Cayley metric captures the number of generalized transposition errors in a permutation, and subsumes previously studied error types, including transpositions and translocations, without imposing restrictions on the lengths and positions of the translocated segments. Based on the so-called *breakpoint analysis* method proposed by Chee and Vu, we first present a coding framework that leads to order-optimal constructions, thus improving upon the existing constructions that are not order-optimal. We then use this framework to also develop an order-optimal coding scheme that is additionally explicit and systematic.**

***Index Terms*—Permutation codes, systematic permutation codes, generalized Cayley distance, block permutation distance, order-optimality.**

I. Introduction

GENERALIZED transposition errors are encountered in various applications, including cloud storage systems, genome resequencing, and flash memories. Cloud storage applications such as Dropbox, OneDrive, iTunes, Google play, etc., are becoming increasingly popular, since they help manage and synchronize data stored across different devices [2]. When items to be synchronized across are ordered, e.g., in a play list, changes on one device can be viewed as transpositions in the permutation on the other device. In DNA resequencing, released genomes consist of collections of unassembled contigs (a contig is an ordered list of genes in the corresponding genome [3]), whose organizations evolve over time by undergoing rearrangement operations. Gene order in a chromosome is subject to rearrangements including reversals, transpositions, translocations, block-interchanges, etc. [3], [4]. Generalized transpositions are also encountered in flash memories that utilize rank modulation, a representation in which cells store relative ranks of their charge levels as a permutation. Charge leakage across cells can then be viewed as a sequence of transpositions in the stored permutation. Errors encountered in the applications described above can be appropriately modeled by the generalized Cayley metric for permutation codes, introduced by Chee and Vu, that captures the number of generalized transpositions between two permutations [5].

Permutation codes in the Kendall-τ metric and the Ulam metric, along with codes in the Levenshtein metric have been recently actively studied, in [6]–[8], [9]–[11], and [12], [13], respectively. Generalized transposition errors subsume transpositions and translocations that the Kendall-τ metric and Ulam metric capture, and in particular no restrictions are imposed on the positions and lengths of the translocated segments as in these two metrics. Codes in the generalized Cayley metric were first studied in [5] using the *breakpoint analysis*, wherein a coding scheme is constructed based on permutation codes, previously introduced in [10], in the Ulam metric. Let N be the length of the codewords, and t be the maximum number of errors in the generalized Cayley metric. While the coding scheme proposed in [5] is explicitly constructive and implementable, the interleaving technique used inevitably incurs a noticeable redundancy of $\Theta\left(N\right)$, without even considering the number of errors that the code is able to correct. As we show later, the best possible redundancy of a length-N code that corrects t generalized transposition errors is $\Theta\left(t\log N\right)$. When t is $o(\frac{N}{\log N})$, the gap between the redundancy of the existing codes based on interleaving and the optimal redundancy increases with N, thus motivating the need to introduce other techniques that are not based on interleaving. We say a length-N code that corrects t generalized transposition errors is order-optimal if the redundancy is $\Theta\left(t\log N\right)$.

In order to obtain codes in the generalized Cayley metric that are order-optimal, we present a coding method that is not based on interleaving. The main idea of our coding scheme is to map each permutation of $\{1, 2, \cdots, N\}$ to a unique characteristic set in the Galois field $\mathbb{F}_q$, where q is a prime number such that $N^2 - N < q < 2N^2 - 2N$ and N is the codelength. We prove that the knowledge of the boundaries of the unaltered segments is sufficient for recovering the permutation from its modified version, obtained through generalized transpositions. We exploit the fact that the symmetric difference of the characteristic sets of two distinct permutations corresponds to these boundaries. Given that the

This paper was presented in part at the IEEE Information Theory Workshop, Kaohsiung, Taiwan, November, 2017 [1].

Siyi Yang is with the Department of Electrical and Computer Engineering, University of California, Los Angeles, CA 90095 USA (email: siyiyang@ucla.edu).

Clayton Schoeny is with the Department of Electrical and Computer Engineering, University of California, Los Angeles, CA 90095 USA (email: cschoeny@ucla.edu).

Lara Dolecek is with the Department of Electrical and Computer Engineering, University of California, Los Angeles, CA 90095 USA (email: dolecek@ee.ucla.edu).

Digital Object Identifier: 10.1109/TIT.2019.2899868

number of such boundaries is linearly upper bounded by the number of generalized transpositions, it is sufficient to find permutations with corresponding characteristic sets on $\mathbb{F}_q$ that have large enough set differences to ensure the desired error correction property. Our proposed method provides a sufficient condition for ensuring the lower bound on the cardinalities of these set differences, which in turn ensures a large enough minimum distance of the resulting code, while the code is order-optimal. Using this approach, we further develop a systematic scheme that is also order-optimal.

The rest of this paper is organized as follows. In Section II, we introduce the basic notation and properties for the generalized Cayley metric and the so-called block permutation metric, which is introduced for metric embedding. In Section III, we define the notion of error-correcting codes in these two metrics and derive useful upper and lower bounds on their optimal rates. We prove the optimal rate to be $1-\Theta\left(\frac{t}{N}\right)$, and use these results to guide the construction of order-optimal codes. In Section IV, we present a method for constructing permutation codes in the generalized Cayley metric. We assign to each permutation of length N a syndrome with elements chosen from a Galois field $\mathbb{F}_q$, where q is a prime number such that $N^2-N<q<2(N^2-N)$. We prove that the permutations with the same syndrome constitute a codebook, and we prove that the largest one is order-optimal. Based on this method, we then develop a construction for order-optimal systematic permutation codes in the generalized Cayley metric in Section V. In Section VI, we prove that the rates of our proposed codes are higher than those of existing codes based on interleaving, namely, our coding scheme is more rate efficient when N is sufficiently large and $t=o\left(\frac{N}{\log N}\right)$. Lastly, we conclude and summarize our main contributions in Section VII.

II. Measure of Distance

A. Notation

In this paper, we denote by $[N]$ the set $\{1,2,\cdots,N\}$. We let $\mathbb{S}_N$ represent the set of all permutations on $[N]$, where each permutation $\sigma : [N] \to [N]$ is a bijection between $[N]$ and itself. The symbol $\circ$ denotes the composition of functions. Specifically, $\sigma \circ \pi$ denotes the composition of two permutations $\sigma,\ \pi \in \mathbb{S}_N$, i.e., $(\sigma \circ \pi)(i) = \sigma(\pi(i))$, $\forall\ i \in [N]$. We assign a vector $(\sigma(1),\sigma(2),\cdots,\sigma(N))$ to each permutation[1] $\sigma \in \mathbb{S}_N$. Under this notation, we call $e=(1, 2,\cdots,N)$ the *identity permutation*. Additionally, σ^{-1} is the inverse permutation of σ. The subsequence of σ from position i to j, $i \le j$, is written as $\sigma[i;j] \triangleq (\sigma(i),\sigma(i+1),\cdots,\sigma(j))$. The symbol Δ refers to the symmetric difference of two sets. Let $\mathrm{GCD}(\cdot)$ and $\mathrm{LCM}(\cdot)$ be the greatest common divisor and the least common multiple, respectively. The symbol $\equiv$ denotes 'congruent modulo'.

B. Generalized Cayley Distance

A ***generalized transposition*** $\phi(i_1,j_1,i_2,j_2) \in \mathbb{S}_N$, where $i_1 \le j_1 < i_2 \le j_2 \in [N]$, refers to a permutation that is obtained from swapping two segments, $e[i_1,j_1]$ and $e[i_2,j_2]$, of the identity permutation [5],

$$\begin{aligned}\phi(i_1,j_1,i_2,j_2) \triangleq (&1,\cdots,i_1-1,i_2,\cdots,j_2,\\ &j_1+1,\cdots,i_2-1,i_1,\cdots,j_1,j_2+1,\cdots,N).\end{aligned} \tag{1}$$

Denote the set of all permutations that represent one generalized transposition on any permutation of length N by $\mathbb{T}_N$. For a given $\pi \in \mathbb{S}_N$ and $\phi(i_1,j_1,i_2,j_2) \in \mathbb{T}_N$, the permutation obtained from swapping the segments $\pi[i_1;j_1]$ and $\pi[i_2;j_2]$ is exactly $\pi \circ \phi$, i.e., the permutation,

$$\begin{aligned}&(\pi(1),\cdots,\pi(i_1-1),\pi(i_2),\cdots,\pi(j_2),\pi(j_1+1),\\ &\cdots,\pi(i_2-1),\pi(i_1),\cdots,\pi(j_1),\pi(j_2+1),\cdots,\pi(N)).\end{aligned} \tag{2}$$

Example 1. *Let* $\pi=(3,5,6,7,9,8,1,2,10,4) \in \mathbb{S}_{10}$. *Let the underlines mark the subsequences that are swapped by* $\phi(2,5,7,8) = \big(1,\underline{7,8},6,\underline{2,3,4,5},9,10\big)$. *Then, for* $\pi = \big(3,\underline{5,6,7,9},8,\underline{1,2},10,4\big)$, *we have:*

$$\pi \circ (\phi(2,5,7,8)) = \big(3,\underline{1,2},8,\underline{5,6,7,9},10,4\big).$$

Definition 1. (Generalized Cayley Distance, cf. [5]) *The* ***generalized Cayley distance*** $d_G(\pi_1,\pi_2)$ *is defined as the minimum number of generalized transpositions that are needed to obtain the permutation* π_2 *from* π_1*, i.e.,*

$$\begin{aligned}d_G(\pi_1,\pi_2) \triangleq \min_d\{\exists\ \phi_1,\phi_2,\cdots,\phi_d \in \mathbb{T}_N,\ \textit{s.t.},\\ \pi_2 = \pi_1 \circ \phi_1 \circ \phi_2 \cdots \circ \phi_d\}.\end{aligned} \tag{3}$$

Remark 1. (cf. [5]). *For all* $\pi_1,\pi_2,\pi_3 \in \mathbb{S}_N$*, the generalized Cayley distance* d_G *satisfies the following properties:*

1) *(Symmetry)* $d_G(\pi_2,\pi_1)=d_G(\pi_1,\pi_2)$.
2) *(Left-invariance)* $d_G(\pi_3 \circ \pi_1,\pi_3 \circ \pi_2)=d_G(\pi_1,\pi_2)$.
3) *(Triangle Inequality)* $d_G(\pi_1,\pi_3) \le d_G(\pi_1,\pi_2)+d_G(\pi_2, \pi_3)$.

Notice that the generalized Cayley distance d_G between two permutations is hard to compute, which makes it difficult to construct codes in the generalized Cayley metric. The common method to address the difficulty of specifying the distances between permutations is metric embedding, where one finds another metric that is computable and is of the same order of magnitude as the original metric. We therefore seek to construct codes under the new metric, the so-called *block permutation distance* to be introduced next, and use this construction to specify codes under d_G.

C. Block Permutation Distance

We say a permutation $\pi \in \mathbb{S}_N$ is ***minimal***[2] if and only if no consecutive elements in π are also consecutive elements in the identity permutation e, i.e.,

$$\forall\ 1 \le i < N,\ \pi(i+1) \ne \pi(i)+1. \tag{4}$$

The set of all minimal permutations of length N is denoted by $\mathbb{D}_N$. Next, we define the *block permutation distance* as follows.

[1] We note that this is different from the cycle notation typically used in algebra.

[2] We note that this is different from the usual notion of minimal permutation specified in group theory.

Definition 2. *The **block permutation distance** $d_B(\pi_1, \pi_2)$ between two permutations $\pi_1, \pi_2 \in \mathbb{S}_N$ is equal to d if*

$$\begin{aligned} \pi_1 &= (\psi_1, \psi_2, \cdots, \psi_{d+1}), \\ \pi_2 &= \left(\psi_{\sigma(1)}, \psi_{\sigma(2)}, \cdots, \psi_{\sigma(d+1)}\right), \end{aligned} \tag{5}$$

where $\sigma \in \mathbb{D}_{d+1}$, $\psi_k = \pi_1[i_{k-1}+1 : i_k]$ for some $0 = i_0 < i_1 \cdots < i_d < i_{d+1} = N$, and $1 \leq k \leq d+1$.

Note that the block permutation distance between permutations π_1 and π_2 is d if and only if $(d+1)$ is the minimum number of blocks the permutation π_1 needs to be divided into in order to obtain π_2 through a block-level permutation. Here by block-level permutation we refer to partitioning the original permutation π_1 into multiple blocks and permuting these blocks.

Example 2. *Let $\pi_1 = (3,5,6,7,9,8,1,2,10,4)$, $\pi_2 = (3,1,2,8,5,6,7,9,10,4)$. Define ψ_i, $1 \leq i \leq 4$, and σ as follows,*

$$\begin{aligned} &\psi_1 = (3), \psi_2 = (5,6,7,9), \psi_3 = (8), \psi_4 = (1,2), \\ &\psi_5 = (10,4), \sigma = (1,4,3,2,5). \end{aligned}$$

Then,

$$\begin{aligned} \pi_1 &= (\psi_1, \psi_2, \psi_3, \psi_4, \psi_5), \\ \pi_2 &= \left(\psi_{\sigma(1)}, \psi_{\sigma(2)}, \psi_{\sigma(3)}, \psi_{\sigma(4)}, \psi_{\sigma(5)}\right), \end{aligned} \tag{6}$$

and thus, $d_B(\pi_1, \pi_2) = 4$, since σ is minimal. This example is in accordance with Definition 2.

Lemma 1. *The block permutation distance d_B also satisfies the properties of symmetry and left-invariance, which are defined in* Remark 1.

Proof: We suppose $\pi_1, \pi_2 \in \mathbb{S}_N$ such that $d_B(\pi_1, \pi_2) = d$. Then, there exist $\sigma \in \mathbb{S}_{d+1}$, and $\psi_1, \psi_2, \cdots, \psi_{d+1}$ such that $\pi_1 = (\psi_1, \psi_2, \cdots, \psi_{d+1})$ and $\pi_2 = \left(\psi_{\sigma(1)}, \psi_{\sigma(2)}, \cdots, \psi_{\sigma(d+1)}\right)$.

To prove the symmetry property, we define $\psi_i' = \psi_{\sigma(i)}$ for $1 \leq i \leq d+1$, and $\sigma' = \sigma^{-1}$. Then, $\sigma' \in \mathbb{D}_{d+1}$, and

$$\begin{aligned} \pi_2 &= \left(\psi_1', \psi_2', \cdots, \psi_{d+1}'\right), \\ \pi_1 &= \left(\psi'_{\sigma'(1)}, \psi'_{\sigma'(2)}, \cdots, \psi'_{\sigma'(d+1)}\right), \end{aligned}$$

thus, $d_B(\pi_2, \pi_1) = d = d_B(\pi_1, \pi_2)$.

To prove the left-invariance property, suppose the length of ψ_i is l_i and let $\psi_i = (\psi_i(1), \psi_i(2), \cdots, \psi_i(l_i))$ for all $1 \leq i \leq d+1$. For a given $\pi_3 \in \mathbb{S}_N$, we define $\tilde{\psi}_i = (\pi_3(\psi_i(1)), \pi_3(\psi_i(2)), \cdots, \pi_3(\psi_i(l_i)))$, for $1 \leq i \leq d+1$. Then,

$$\begin{aligned} \pi_3 \circ \pi_1 &= \left(\tilde{\psi}_1, \tilde{\psi}_2, \cdots, \tilde{\psi}_{d+1}\right), \\ \pi_3 \circ \pi_2 &= \left(\tilde{\psi}_{\sigma(1)}, \tilde{\psi}_{\sigma(2)}, \cdots, \tilde{\psi}_{\sigma(d+1)}\right). \end{aligned}$$

Therefore, $d_B(\pi_3 \circ \pi_1, \pi_3 \circ \pi_2) = d = d_B(\pi_1, \pi_2)$. ■

Note that Definition 2 is an implicit representation of d_B. Next, we seek to characterize d_B explicitly.

Definition 3. *The **characteristic set** $A(\pi)$ for any $\pi \in \mathbb{S}_N$ is defined as the set of all consecutive pairs in π, i.e.,*

$$A(\pi) \triangleq \{(\pi(i), \pi(i+1)) \,|1 \leq i < N\}. \tag{7}$$

Recall that e refers to the identity permutation.

Definition 4. *The **block permutation weight** $w_B(\pi)$ is defined as the number of consecutive pairs in π that do not belong to $A(e)$ (w_B is exactly the number of so-called breakpoints in [5]), i.e.,*

$$w_B(\pi) \triangleq |A(\pi) \setminus A(e)|. \tag{8}$$

Lemma 2 and Remark 2 state explicit representations of the block permutation distance d_B by the characteristic set and the block permutation weight, respectively, and will be used later in the paper to establish our main result.

Lemma 2. *For all $\pi_1, \pi_2 \in \mathbb{S}_N$,*

$$d_B(\pi_1, \pi_2) = |A(\pi_2) \setminus A(\pi_1)| = |A(\pi_1) \setminus A(\pi_2)|. \tag{9}$$

Proof: The proof is in Appendix A. ■

Remark 2. *From Lemma 2 and Definition 4, it is obvious that*

$$w_B(\pi) = d_B(e, \pi) = d_B(\pi, e). \tag{10}$$

For all $\pi_1, \pi_2 \in \mathbb{S}_N$, it follows immediately from the left-invariance property of d_B and (8) that

$$d_B(\pi_1, \pi_2) = w_B\left(\pi_1^{-1} \circ \pi_2\right). \tag{11}$$

In Example 3, we show how to compute the block permutation distance of two permutations from their characteristic sets, as it is indicated in Lemma 2.

Example 3. *For π_1, π_2 specified in Example 2,*

$$\begin{aligned} A(\pi_1) = \{&(3,5), (5,6), (6,7), (7,9), (9,8), \\ &(8,1), (1,2), (2,10), (10,4)\}, \\ A(\pi_2) = \{&(3,1), (1,2), (2,8), (8,5), (5,6), \\ &(6,7), (7,9), (9,10), (10,4)\}. \end{aligned}$$

Therefore,

$$\begin{aligned} |A(\pi_1) \setminus A(\pi_2)| =& |\{(3,5), (9,8), (8,1), (2,10)\}| \\ =& 4 = d_B(\pi_1, \pi_2). \end{aligned}$$

This example is in accordance with Lemma 2.

D. Metric Embedding

The generalized Cayley distance is difficult to compute, whereas the block permutation distance can be computed efficiently. Therefore, it is easier to check whether two distinct candidate codewords in a codebook meet the minimum requirement on the block permutation distance, than it is to check whether they meet the minimum requirement on the generalized Cayley distance. In light of this observation, in the next section, we apply metric embedding to transform the problem of code design in d_G into that in d_B, which is easier to deal with, using the following results.

Lemma 3. *For all $\pi_1, \pi_2 \in \mathbb{S}_N$, the following inequality holds,*

$$w_B(\pi_1 \circ \pi_2) \leq w_B(\pi_1) + w_B(\pi_2). \tag{12}$$

Proof: The proof is in Appendix B. ■

Remark 3. *It follows immediately from equation (11) and Lemma 3 that the block permutation distance satisfies the triangle inequality, i.e., $\forall\ \pi_1, \pi_2, \pi_3 \in \mathbb{S}_N$,*

$$d_B(\pi_1, \pi_3) \leq d_B(\pi_1, \pi_2) + d_B(\pi_2, \pi_3). \tag{13}$$

From Lemma 3 and the definitions of the generalized Cayley metric and the block permutation metric, we observe the following relation between d_B and d_G. This result is used later in Section IV.

Lemma 4. *For all $\pi_1, \pi_2 \in \mathbb{S}_N$, the following inequality holds,*

$$d_G(\pi_1, \pi_2) \leq d_B(\pi_1, \pi_2) \leq 4 d_G(\pi_1, \pi_2). \tag{14}$$

Proof:

To prove the upper bound, we consider two arbitrary permutations π_1, $\pi_2 \in \mathbb{S}_N$, and let $k = d_G(\pi_1, \pi_2)$. We know from definitions of a generalized transposition and the block permutation weight that for any generalized transposition $\phi \in \mathbb{T}_N$ (recall that $\mathbb{T}_N$ is defined at the beginning of Section II-B as the set of all permutations that represent a generalized transposition in permutations of length N), the following inequality holds,

$$w_B(\phi) \leq 4. \tag{15}$$

From the definition of the generalized Cayley metric and $d_G(\pi_1, \pi_2) = k$, it follows that for some $\phi_1, \phi_2, \cdots, \phi_k \in \mathbb{T}_N$,

$$\pi_2 = \pi_1 \circ \phi_1 \circ \phi_2 \cdots \circ \phi_k.$$

Then, from Lemma 3 and (15),

$$\begin{aligned} d_B(\pi_1, \pi_2) =& w_B\left(\pi_1^{-1} \circ \pi_2\right) \\ =& w_B(\phi_1 \circ \phi_2 \circ \cdots \circ \phi_k) \\ \leq& \sum_{i=1}^{k} w_B(\phi_i) \\ \leq& 4k = 4 d_G(\pi_1, \pi_2). \end{aligned}$$

The upper bound is proved.

The lower bound is trivially attained when $\pi_1 = \pi_2$. When π_1 and π_2 are distinct, it follows that $d_B(\pi_1, \pi_2) = d$ for some positive integer d. Then, according to the definition of the block permutation distance, there exists a minimal permutation σ (minimal permutation is defined in Section II-C as a permutation where no consecutive elements in σ are also consecutive elements in the identity permutation) and a partition $\{\psi_i\}_{i=1}^{d+1}$ of π_1 such that, $\pi_1 = (\psi_1, \psi_2, \cdots, \psi_{d+1})$, and $\pi_2 = \left(\psi_{\sigma(1)}, \psi_{\sigma(2)}, \cdots, \psi_{\sigma(d+1)}\right)$.

Next, suppose l_0 is the smallest index l such that $\sigma(l) \neq l$, $1 \leq l \leq d+1$ (the assumption that $\pi_1 \neq \pi_2$ ensures the existence of l_0). Let $k_0 = \sigma^{-1}(l_0)$, then $k_0 > l_0$. Let ϕ_1 represent the generalized transposition that swaps the subsequences $\left(\psi_{\sigma(l_0)}, \psi_{\sigma(l_0+1)}, \cdots, \psi_{\sigma(k_0-1)}\right)$ and $\psi_{\sigma(k_0)} = \psi_{l_0}$ in π_2. Let $\pi_2^{(1)} = \pi_2 \circ \phi_1$ and $\sigma^{(1)} = (1, 2, \cdots, l_0, \sigma(l_0), \sigma(l_0+1), \cdots, \sigma(k_0-1), \sigma(k_0+1), \cdots, \sigma(d+1))$. Then,

$$\pi_2^{(1)} = \left(\psi_{\sigma^{(1)}(1)}, \psi_{\sigma^{(1)}(2)}, \cdots, \psi_{\sigma^{(1)}(d+1)}\right).$$

If $\pi_2^{(1)} = \pi_1$, then $\pi_1 = \pi_2 \circ \phi_1$. Otherwise let l_1 be the smallest index l such that $\sigma^{(1)}(l) \neq l$, $1 \leq l \leq d+1$, then $l_1 > l_0$ holds true.

Following this procedure, one can find a series of generalized transpositions $\phi_1, \phi_2, \cdots, \phi_m$, $1 \leq m \leq d$, sequentially, such that $\pi_2 \circ \phi_1 \circ \phi_2 \circ \cdots \circ \phi_m = \pi_1$. Suppose $\phi_1, \phi_2, \cdots, \phi_i$ are found for some i, $1 \leq i \leq d$. Let $\pi_2^{(i)} = \pi_2 \circ \phi_1 \circ \phi_2 \circ \cdots \circ \phi_i = \left(\psi_{\sigma^{(i)}(1)}, \psi_{\sigma^{(i)}(2)}, \cdots, \psi_{\sigma^{(i)}(d+1)}\right)$. If $\pi_2^{(i)} = \pi_1$, then $\pi_1 = \pi_2 \circ \phi_1 \circ \phi_2 \circ \cdots \circ \phi_i$, and we have established the desired composition. Otherwise, we let l_i be the smallest index such that $\sigma^{(i)}(l_i) \neq l_i$. Suppose $k_i = \left(\sigma^{(i)}\right)^{-1}(l_i)$, and it follows that $k_i > l_i$. Denote the generalized transposition that swaps the subsequences $\left(\psi_{\sigma^{(i)}(l_i)}, \psi_{\sigma^{(i)}(2)}, \cdots, \psi_{\sigma^{(i)}(k_i-1)}\right)$ and $\psi_{\sigma^{(i)}(k_i)} = \psi_{l_i}$ in $\pi_2^{(i)}$ by ϕ_{i+1}. Let $\pi_2^{(i+1)} = \pi_2^{(i)} \circ \phi_{i+1}$, and $\sigma^{(i+1)} = (1, 2, \cdots, l_i, \sigma^{(i)}(l_i), \sigma^{(i)}(l_i+1), \cdots, \sigma^{(i)}(k_i-1), \sigma^{(i)}(k_i+1), \cdots, \sigma^{(i)}(d+1))$. Then,

$$\pi_2^{(i+1)} = \left(\psi_{\sigma^{(i+1)}(1)}, \psi_{\sigma^{(i+1)}(2)}, \cdots, \psi_{\sigma^{(i+1)}(d+1)}\right).$$

Finally, one finds the smallest integer m such that $\pi_2^{(m)} = \pi_1$. In this procedure, $l_0, \cdots, l_{m-1}$ are obtained sequentially, where $1 < l_0 < l_1 < \cdots < l_{m-1}$. We also know that $l_{m-1} \leq d$, otherwise if $l_{m-1} = d+1$, then $\sigma^{(m-1)}(i) = i$ holds true for all $1 \leq i \leq d$, and $\sigma^{(m-1)}(d+1) \neq d+1$, which leads to a contradiction. Therefore, $d \geq l_{m-1} > \cdots > l_0 \geq 1$, which implies that $m \leq d$. Note that $\pi_1 = \pi_2 \circ \phi_1 \circ \cdots \circ \phi_m$, from which $d_G(\pi_1, \pi_2) \leq m \leq d = d_B(\pi_1, \pi_2)$ follows. The lower bound is proved. ■

III. Theoretical Bounds on the Code Rate

A subset $\mathcal{C}_G(N, t)$ of $\mathbb{S}_N$ is called a ***t-generalized Cayley code*** if it can correct t generalized transposition errors. Any t-generalized Cayley code has the minimum generalized Cayley distance $d_{G,min} \geq 2t+1$. Similarly, a subset $\mathcal{C}_B(N, t)$ of $\mathbb{S}_N$ is called a ***t-block permutation code*** if its minimum block permutation distance $d_{B,min} \geq 2t+1$. For any permutation code $\mathcal{C} \subset \mathbb{S}_N$, denote the rate of $\mathcal{C}$ by $R(\mathcal{C})$. Then, the following equation holds true,

$$R(\mathcal{C}) = \frac{\log|\mathcal{C}(N, t)|}{\log N!}. \tag{16}$$

In the remainder of this paper, the logarithm base is always 2 unless it is explicitly specified with a different base.

Let $\mathcal{C}_{G,opt}(N, t)$ and $\mathcal{C}_{B,opt}(N, t)$ be t-generalized Cayley codes and t-block permutation codes with the optimal rates, denoted by $R_{G,opt}(N, t)$ and $R_{B,opt}(N, t)$, respectively. We next derive the lower bounds and the upper bounds of $R_{G,opt}(N, t)$ and $R_{B,opt}(N, t)$.

For each $\pi \in \mathbb{S}_N$, we define the ***generalized Cayley ball*** $B_G(N, t, \pi)$ of radius t centered at π to be the set of all permutations in $\mathbb{S}_N$ that have a generalized Cayley distance from π not exceeding t. We know from the left-invariance property of d_G that the cardinality of $B_G(N, t, \pi)$ is independent of π; we denote $|B_G(N, t, \pi)|$ as $b_G(N, t)$. The ***block permutation ball*** $B_B(N, t, \pi)$ and the corresponding ball-size $b_B(N, t)$ are similarly defined.

We derive the lower and upper bounds of $b_B(N, t)$ and $b_G(N, t)$ in the following two lemmas, respectively. We build on these results and Lemma 7 to compute the bounds of the rates of optimal codes in d_G and d_B, proving that the optimal redundancy is $\Theta(\frac{t}{N})$ in both of the two metrics.

Lemma 5. *For all $N \in \mathbb{N}^*$, $t \le N - \sqrt{N} - 1$, $b_B(N,t)$ is bounded by the following inequality:*

$$\prod_{k=1}^{t}(N-k) \le b_B(N,t) \le \prod_{k=0}^{t}(N-k). \tag{17}$$

Proof: The proof is in Appendix C. ■

Lemma 6. *For all $N \in \mathbb{N}^*$, $t \le \min\{N - \sqrt{N} - 1, \frac{N-1}{4}\}$, $b_G(N,t)$ is bounded as follows:*

$$\prod_{k=1}^{t}(N-k) \le b_G(N,t) \le \prod_{k=0}^{4t}(N-k). \tag{18}$$

Proof: The proof is in Appendix D. ■

As the metrics d_B and d_G both satisfy the triangle inequality, the cardinalities of the optimal codes $\mathcal{C}_{B,opt}(N,t)$ and $\mathcal{C}_{G,opt}(N,t)$ are bounded as follows,

$$\begin{aligned} \frac{N!}{b_B(N,2t)} &\le |\mathcal{C}_{B,opt}(N,t)| \le \frac{N!}{b_B(N,t)}, \\ \frac{N!}{b_G(N,2t)} &\le |\mathcal{C}_{G,opt}(N,t)| \le \frac{N!}{b_G(N,t)}. \end{aligned} \tag{19}$$

According to [14, (1)-(2)], for all $N \in \mathbb{N}^*$,

$$N! = \sqrt{2\pi}N^{N+1/2}e^{-N} \cdot e^{r_N}, \tag{20}$$

where

$$\frac{1}{12N+1} < r_N < \frac{1}{12N}. \tag{21}$$

From (20) and (21), Lemma 7 follows.

Lemma 7. *For all $N \in \mathbb{N}^*$, it follows that*

$$\begin{aligned} (N+\frac{1}{2})\log N - (\log e)N &< \sum_{n=1}^{N}\log n \\ &< (N+\frac{1}{2})\log N - (\log e)N + 2. \end{aligned}$$

We now state the main result of this section.

Theorem 1. *For any $t, N \in \mathbb{N}^*$, $t \le \min\{N - \sqrt{N} - 1, \frac{N-1}{4}\}$ and $N \ge 9$, the optimal rates $R_{B,opt}(N,t)$, $R_{G,opt}(N,t)$ satisfy the following inequalities,*

$$\begin{aligned} 1 - c \cdot \frac{2t+1}{N} &\le R_{B,opt}(N,t) \le 1 - \frac{t}{N}, \\ 1 - c \cdot \frac{8t+1}{N} &\le R_{G,opt}(N,t) \le 1 - \frac{t}{N}, \end{aligned} \tag{22}$$

where $c = 1 + \frac{2\log e}{\log N}$.

Proof: From (16) and (19), it follows that

$$\begin{aligned} 1 - \frac{\log b_B(N,2t)}{\log N!} &\le R_{B,opt}(N,t) \le 1 - \frac{\log b_B(N,t)}{\log N!}, \\ 1 - \frac{\log b_G(N,2t)}{\log N!} &\le R_{G,opt}(N,t) \le 1 - \frac{\log b_G(N,t)}{\log N!}. \end{aligned} \tag{23}$$

By applying Lemma 5 and Lemma 7 to (23), when $\min\{N - \sqrt{N} - 1, \frac{N-1}{4}\} \ge t \ge 1$ and $N \ge 9$, it follows that

$$\begin{aligned} R_{B,opt}(N,t) &\ge 1 - \frac{\log\left[\prod_{k=0}^{2t}(N-k)\right]}{\log N!} \\ &> 1 - \frac{(2t+1)\log N}{(N+\frac{1}{2})\log N - (\log e)N} \\ &> 1 - \frac{(2t+1)\log N}{N(\log N - \log e)} \\ &> 1 - \frac{2t+1}{N}\left(1 + \frac{2\log e}{\log N}\right), \end{aligned} \tag{24}$$

and

$$\begin{aligned} &R_{B,opt}(N,t) \\ \le& 1 - \frac{\log\left[\prod_{k=1}^{t}(N-k)\right]}{\log N!} \\ =& 1 - \frac{\frac{1}{2}\sum_{k=1}^{t}(\log(N-k) + \log(N-t-1+k))}{\log N!} \\ \le& 1 - \frac{\frac{t}{2}\log\left((N-1)(N-t)\right)}{(N+\frac{1}{2})\log N - (\log e)N + 2} \\ \le& 1 - \frac{\frac{t}{2}\log\left((N-1)(N-\frac{N-1}{4})\right)}{(N+\frac{1}{2})\log N - (\log e)N + 2} \\ \le& 1 - \frac{\frac{t}{2}\log\left(\frac{N^2}{2}\right)}{(N+\frac{1}{2})\log N - (\log e)N + 2} \\ \le& 1 - \frac{t(\log N - \frac{1}{2})}{N\log N - \frac{1}{2}N} \\ =& 1 - \frac{t}{N}. \end{aligned} \tag{25}$$

Similarly, by applying Lemma 6 and Lemma 7 to (23), when $\min\{N - \sqrt{N} - 1, \frac{N-1}{4}\} \ge t \ge 1$ and $N \ge 9$, it follows that

$$\begin{aligned} R_{G,opt}(N,t) &\ge 1 - \frac{\log\left[\prod_{k=0}^{\min\{8t,N-1\}}(N-k)\right]}{\log N!} \\ &> 1 - \frac{(8t+1)\log N}{(N+\frac{1}{2})\log N - (\log e)N} \\ &> 1 - \frac{(8t+1)\log N}{N\log N - (\log e)N} \\ &> 1 - \frac{8t+1}{N}\left(1 + \frac{2\log e}{\log N}\right), \end{aligned} \tag{26}$$

and

$$\begin{aligned} R_{G,opt}(N,t) &\le 1 - \frac{\log\left[\prod_{k=1}^{t}(N-k)\right]}{\log N!} \\ &\le 1 - \frac{t}{N}. \end{aligned} \tag{27}$$

The theorem is proved. ■

Inequalities (24)-(27) indicate that $R = 1 - \Theta\left(\frac{t}{N}\right)$ is the rate of the t-generalized Cayley codes and the t-block permutation codes that are order-optimal.

IV. Non-Systematic Permutation Codes in the Generalized Cayley Metric

We studied the optimal rates of t-generalized Cayley Codes and t-block permutation codes in the previous section. We now seek constructions of order-optimal codes in these metrics. We know from Lemma 4 that any $4t$-block permutation code is also a t-generalized Cayley code. In the sequel, we thus focus on the construction of order-optimal t-block permutation codes, which is sufficient for obtaining order-optimal generalized Cayley codes.

In Section IV-A, we present a construction of order-optimal t-block permutation codes (Theorem 2). We then develop a decoding scheme for the proposed codes in Section IV-B.

A. Encoding Scheme

Denote the set of all ordered pairs of non-identical elements from $[N]$ by P; then $|P| = N^2 - N$. Suppose q is a prime number such that $q \geq |P|$. From *Bertrand's postulate* [15], one can always find a prime number q such that $|P| \leq q \leq 2|P|$.

Let $\upsilon : \; P \to \mathbb{F}_q$ be an arbitrary injection from P to $\mathbb{F}_q$, where $\mathbb{F}_q$ is a Galois field of order q. Let $\mathcal{P}(\mathbb{F}_q)$ represent the set of all the subsets of $\mathbb{F}_q$ with cardinality $N-1$. We define an injection $\nu : \mathbb{S}_N \to \mathcal{P}(\mathbb{F}_q)$ as follows:

$$\nu(\pi) \triangleq \{\upsilon(p) | p \in A(\pi)\}. \tag{28}$$

Then, ν is invertible, namely, one is able to compute π based on $\nu(\pi)$.

We then define a class of functions $\alpha^{(q,d)} : \; \mathbb{S}_N \to \mathbb{F}_q^{2d-1}$, as follows:

$$\alpha^{(q,d)}(\pi) \triangleq (\alpha_1, \alpha_2, \cdots, \alpha_{2d-1}), \tag{29}$$

where

$$\begin{cases} \alpha_1 & \equiv \sum\limits_{b \in \nu(\pi)} b & \mod q, \\ \alpha_2 & \equiv \sum\limits_{b \in \nu(\pi)} b^2 & \mod q, \\ & \vdots & \\ \alpha_{2d-1} & \equiv \sum\limits_{b \in \nu(\pi)} b^{2d-1} & \mod q. \end{cases} \tag{30}$$

The following Algorithm 1 describes the main steps of the proposed encoding scheme, the correctness of which can be verified by Lemma 8 and Theorem 2.

The following Lemma 8 states that the cardinality of the symmetric difference of $\nu(\pi_1), \nu(\pi_2)$ for any two distinct permutation $\pi_1, \pi_2 \in \mathbb{S}_N$ is greater than $2d$ if their syndromes $\alpha^{(q,d)}(\pi_1)$ and $\alpha^{(q,d)}(\pi_2)$ are identical. Therefore, their block permutation distance is greater than d based on Lemma 2. This lemma will be repeatedly used in the rest of the paper for the constructions of order-optimal permutation codes in the block permutation distance.

Lemma 8. *For all π_1, $\pi_2 \in \mathbb{S}_N$ such that $\pi_1 \neq \pi_2$, if $\alpha^{(q,d)}(\pi_1) = \alpha^{(q,d)}(\pi_2)$, then,*

$$|\nu(\pi_1) \Delta \nu(\pi_2)| > 2d. \tag{31}$$

Proof: The proof is in Appendix E. ■

Algorithm 1 Encoding Scheme

```
Input:
    Minimum block permutation distance: $2t+1$;
    Codelength: $N$;
    Alphabet size: $q$, where $q$ is a prime number such that $N^2 - N \leq q < 2(N^2 - N)$;
Output:
    Codebook $\mathcal{C}$ of a $t$-block permutation code;
For each $\pi \in \mathbb{S}_N$, compute $A(\pi)$, $\nu(\pi)$, and its syndrome $\alpha^{(q,2t)}(\pi)$ ($\alpha^{(q,2t)}(\pi) \in \mathbb{F}_q^{4t-1}$), sequentially, where $A(\pi)$, $\nu(\pi)$, $\alpha^{(q,d)}$ are defined in Definition 3, (28), (29) and (30), respectively;
For each $\boldsymbol{\alpha} \in \mathbb{F}_q^{4t-1}$, denote the set consisting of all permutations with the syndrome $\boldsymbol{\alpha}$ by $\mathcal{C}_{\boldsymbol{\alpha}}(N,t)$;
Find $\boldsymbol{\alpha}$ such that $\mathcal{C}_{\boldsymbol{\alpha}}(N,t)$ is of the maximum cardinality;
return $\mathcal{C} = \mathcal{C}_{\boldsymbol{\alpha}}(N,t)$.
```

Note that the function $\alpha^{(q,2t)}$ induces a map from $\mathbb{S}_N$ to $\mathbb{F}_q^{4t-1}$ and divides $\mathbb{S}_N$ into q^{4t-1} subsets based on their syndromes $\boldsymbol{\alpha} = (\alpha_1, \alpha_2, \cdots, \alpha_{4t-1})$. We next prove that each such subset is a t-block permutation code, which is stated as the following theorem.

Theorem 2. *For all $\boldsymbol{\alpha} \in \mathbb{F}_q^{4t-1}$, suppose:*

$$\mathcal{C}_{\boldsymbol{\alpha}}(N,t) = \{\pi | \pi \in \mathbb{S}_N, \; \alpha^{(q,2t)}(\pi) = \boldsymbol{\alpha}\}, \tag{32}$$

where $\alpha^{(q,2t)}$ is defined in (29) and (30). Then $\forall \; \pi_1, \pi_2 \in \mathcal{C}_{\boldsymbol{\alpha}}(N,t)$, $\pi_1 \neq \pi_2$, the following inequality holds,

$$d_B(\pi_1, \pi_2) \geq 2t + 1. \tag{33}$$

Proof: Let $d = 2t$ in Lemma 8 and Lemma 2. Then,

$$\begin{aligned} d_B(\pi_1, \pi_2) &= \frac{1}{2} |A(\pi_1) \Delta A(\pi_2)| \\ &= \frac{1}{2} |\nu(\pi_1) \Delta \nu(\pi_2)| \\ &> \frac{1}{2}(2 \cdot 2t) = 2t, \end{aligned} \tag{34}$$

where Δ refers to the symmetric difference of sets. ■

Example 4. *Suppose $N = 10$, $t = 2$, $q = 97 > 10^2 - 10$. Define $\upsilon(i,j)$ for all $i \neq j \in [10]$ as follows:*

$$\upsilon(i,j) = 10(i-1) + j - 1.$$

Let $\pi_1 = (10, 9, 8, 7, 6, 5, 4, 3, 2, 1)$, and $\pi_2 = (9, 6, 5, 8, 2, 4, 7, 3, 10, 1)$. Suppose $\boldsymbol{\alpha} = (83, 28, 80, 77, 40, 3, 88)$. Then,

$$\alpha^{(q,2t)}(\pi_1) = \alpha^{(q,2t)}(\pi_2) = \boldsymbol{\alpha}.$$

Observe that $d_B(\pi_1, \pi_2) = 8 > 4 = 2t$. This example is in accordance with Theorem 2.

Theorem 2 implies that $\{\mathcal{C}_{\boldsymbol{\alpha}}(N,t) \; : \; \boldsymbol{\alpha} \in \mathbb{F}_q^{4t-1}\}$ is a partition of $\mathbb{S}_N$, where each component $\mathcal{C}_{\boldsymbol{\alpha}}(N,t)$ is a t-block permutation code indexed by $\boldsymbol{\alpha}$. Suppose $\mathcal{C}_{\boldsymbol{\alpha}_{\max}}(N,t)$ is the one with the maximal cardinality, whose syndrome is $\boldsymbol{\alpha}_{\max}$. It follows from the *Pigeonhole Principle* that:

$$|\mathcal{C}_{\boldsymbol{\alpha}_{\max}}(N,t)| \geq \frac{N!}{|\mathbb{F}_q^{4t-1}|} = \frac{N!}{q^{4t-1}}. \tag{35}$$

Denote the rate of $\mathcal{C}_{\boldsymbol{\alpha}_{\max}}(N,t)$ by $R(\mathcal{C}_1)$. Given that $N^2-N=|P|\le q<2|P|=2N^2-2N<2N^2$, it follows from Lemma 7 that for $N>e^2$ (note that here e refers to the base of the natural logarithm),

$$\begin{aligned} R(\mathcal{C}_1) &\ge 1-\frac{4t\log q}{\log N!}>1-\frac{8t\log N+4t}{\log N!}\\ &>1-\frac{8t(\log N+\frac{1}{2})}{(N+\frac{1}{2})\log N-(\log e)N}\\ &>1-\frac{8t}{N}\left(\frac{\log N+\frac{1}{2}}{\log N-\log e}\right)\\ &=1-\frac{8t}{N}\left[1+\frac{\frac{1}{2}+\log e}{\log N}\left(1+\frac{\log e}{\log N-\log e}\right)\right]\\ &>1-\frac{8t}{N}\left(1+\frac{2\log e+1}{\log N}\right). \end{aligned} \tag{36}$$

Then, $\mathcal{C}_{\boldsymbol{\alpha}_{\max}}(N,t)$ is an order-optimal t-block permutation code.

B. Decoding Scheme

In Section IV-A, we map each permutation $\pi\in\mathbb{S}_N$ to a unique set $\nu(\pi)\in\mathcal{P}(\mathbb{F}_q)$ as defined in equation (28), where $N^2-N\le q\le 2N^2-2N$ and $\mathcal{P}(\mathbb{F}_q)$ represents the set consisting of all subsets of $\mathbb{F}_q$ with cardinality $N-1$. Suppose the transmitter sends $\pi\in\mathbb{S}_N$ and the receiver receives π', where $d_G(\pi,\pi')\le t$. In the decoding scheme, our objective is to compute $\nu(\pi)$ from the a priori $\boldsymbol{\alpha}$ and the received permutation π'. The strategy is, for each set $B\in\mathcal{P}(\mathbb{F}_q)$, map B to a polynomial $f(X;B)$ defined as follows:

$$f(X;B)\triangleq\prod_{b\in B}(X+b)\,. \tag{37}$$

We call $f(X;B)$ the ***characteristic function*** of set B. All the polynomials as well as the polynomial operations are defined on $\mathbb{F}_q$. Let a_i^B, $0\le i\le N-1$, represent the coefficients of X^{N-1-i} in $f(X;B)$. Then, $a_0^B=1$.

Given the a priori agreement on the codebook, i.e., the choice of $\boldsymbol{\alpha}$, and the received permutation π', the value of the first $4t$ coefficients of $f(X;B)$, $f(X;B')$ can be computed, where $B=\nu(\pi)$ and $B'=\nu(\pi')$, as we shall shortly show. We then use these coefficients to derive $\nu(\pi)$. This coding strategy bears resemblance to that proposed in [16], the key difference being that the coefficients of the polynomials we discussed are partially known, thus making our decoding scheme more complicated, whereas those in [16] are fully known.

Note that $a_i^B, 1\le i\le N-1$, in (38) is the i-th elementary symmetric polynomial of the elements in B. Also note that the i-th component $\alpha_i, 1\le i\le 4t-1$, of the value $\boldsymbol{\alpha}=\alpha^{(q,2t)}(\pi)$ is exactly the i-th power sum of the elements in $B=\nu(\pi)$. We know from *Newton's identities* [17] that there exists a bijection between the $(4t-1)$ power sums and the first $(4t-1)$ elementary symmetric polynomials of elements in B, as described below:

$$\begin{cases} a_0^B=1,\\ a_1^B=\alpha_1,\\ a_2^B=2^{-1}(a_1^B\alpha_1-\alpha_2),\\ a_3^B=3^{-1}(a_2^B\alpha_1-a_1^B\alpha_2+\alpha_3),\\ \vdots\\ a_{4t-1}^B=(4t-1)^{-1}(a_{4t-2}^B\alpha_1-a_{4t-3}^B\alpha_2+\cdots+\alpha_{4t-1}). \end{cases} \tag{38}$$

Denote a_i^B, $a_i^{B'}$ by a_i, a_i', $0\le i\le N-1$, respectively, for simplicity. Let $r(B)=(a_1,a_2,\cdots,a_{4t-1})$, $r(B')=(a_1',a_2',\cdots,a_{4t-1}')$. The receiver uses the a priori $\boldsymbol{\alpha}$ to compute $r(B)$ and to derive $r(B')$ from B', where $B=\nu(\pi)$ and $B'=\nu(\pi')$. Note that π can be computed from $B=\nu(\pi)$ since ν is an injection from $\mathbb{S}_N$ to $\mathcal{P}(\mathbb{F}_q)$. Thus the objective is to compute B from $r(B)$, $r(B')$, and B'.

Suppose $D_1=B\setminus B'$, $D_2=B'\setminus B$, $D_3=B\cap B'$. Let $f_1=f(X;B)$ and $f_2=f(X;B')$. Then,

$$\begin{aligned} g_1(X)&=\frac{f_1}{GCD(f_1,f_2)}=\prod_{b\in D_1}(X+b),\\ g_2(X)&=\frac{f_2}{GCD(f_1,f_2)}=\prod_{b\in D_2}(X+b),\\ g_3(X)&=GCD(f_1,f_2)=\prod_{b\in D_3}(X+b). \end{aligned} \tag{39}$$

Notice that g_1,g_2,g_3 uniquely determine f_1,f_2, so they are sufficient for computing π. We next seek to compute g_1,g_2,g_3 from $r(B)$ and $f_2=g_2\cdot g_3$, from which $f_1=g_1\cdot g_3$ can be determined. Let $(h_1,h_2)=(X^{t-k}g_2,X^{t-k}g_1)$, where $k=\deg g_1=\deg g_2=|D_1|=|D_2|\le t$. Then (h_1,h_2) satisfy $h_1\cdot f_1=h_2\cdot f_2$. We will also prove later in Theorem 3 that g_1,g_2,g_3 can be computed from an arbitrary nonzero solution (h_1,h_2) of $h_1\cdot f_1=h_2\cdot f_2$. Therefore, any nonzero solution to $h_1\cdot f_1=h_2\cdot f_2$ is sufficient for computing π. Also notice that the first $4t$ coefficients of $h_1\cdot f_1$ and $h_2\cdot f_2$ uniquely determine $r(B)$ and $r(B')$, respectively, by (38), if h_1,h_2 are known. In order to compute g_1,g_2,g_3, it is sufficient to find h_1 and h_2, both of degree t, such that the first $4t$ coefficients of $h_1\cdot f_1$ and that of $h_2\cdot f_2$ are equal, i.e., the following inequality holds,

$$\deg(h_1\cdot f_1-h_2\cdot f_2)<N-3t. \tag{40}$$

For each $\mathbf{c}\in\mathbb{F}_q^{2t}$, suppose

$$\mathbf{c}=\left(c_1,\cdots,c_t,-c_1',\cdots,-c_t'\right)^T, \tag{41}$$

and define the polynomials $h_1(\mathbf{c}),h_2(\mathbf{c})$ of degree t as follows,

$$\begin{aligned} h_1(\mathbf{c})&\triangleq X^t+c_1X^{t-1}+c_2X^{t-2}+\cdots+c_t,\\ h_2(\mathbf{c})&\triangleq X^t+c_1'X^{t-1}+c_2'X^{t-2}+\cdots+c_t'. \end{aligned} \tag{42}$$

Define

$$\mathbf{A} = \begin{pmatrix} 1 & 0 & \cdots & 0 & 1 & 0 & \cdots & 0 \\ a_1 & 1 & \ddots & \vdots & a'_1 & 1 & \ddots & \vdots \\ \vdots & \vdots & \ddots & 0 & \vdots & \vdots & \ddots & 0 \\ a_{t-1} & a_{t-2} & \cdots & 1 & a'_{t-1} & a'_{t-2} & \cdots & 1 \\ \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \ddots & \vdots \\ a_{4t-2} & a_{4t-3} & \cdots & a_{3t-1} & a'_{4t-2} & a'_{4t-3} & \cdots & a'_{3t-1} \end{pmatrix}, \tag{43}$$

and

$$\mathbf{b} = \left(a'_1, \cdots, a'_{4t-1}\right)^T - \left(a_1, \cdots, a_{4t-1}\right)^T. \tag{44}$$

The following Algorithm 2 describes the decoding algorithm of the code constructed in Section IV-A. The correctness of this algorithm is proved by Lemma 9 and Theorem 3.

Algorithm 2 Decoding Algorithm

```
Input:
    Syndrome: $\boldsymbol{\alpha}$;
    Received sequence: $\pi'$;
Output:
    Estimated codeword: $\hat{\pi}$;
Compute the coefficients $\{a'_i\}_{i=1}^{4t-1}$ of $f_2$ and $B'$ from $\pi'$;
Compute the coefficients of $\{a_i\}_{i=1}^{4t-1}$ of $f_1$ from $\boldsymbol{\alpha}$ by Newton's identities;
Compute $\mathbf{A}$ and $\mathbf{b}$ using (43) and (44);
Find a nonzero solution $\mathbf{c}$ to $\mathbf{Ac} = \mathbf{b}$, $\mathbf{c} = \left(c_1, \cdots, c_{2t}\right)^T$;
Compute $h_1 = X^t + c_1 X^{t-1} + c_2 X^{t-2} + \cdots + c_t$, $h_2 = X^t - c_{t+1} X^{t-1} - c_{t+2} X^{t-2} - \cdots - c_{2t}$;
Compute $h = \gcd(h_1, h_2)$, $v_1 = \frac{h_2}{h}$, $v_2 = \frac{h_1}{h}$;
Let the set of negative roots of $v_1$ and $v_2$ be $V_1$ and $V_2$, respectively;
Compute $\hat{\pi} = \nu^{-1}\left(V_1 \cup (B' \setminus V_2)\right)$, where $\nu$ is defined in (28);
return $\hat{\pi}$.
```

Lemma 9 presents an equivalent linear equation to find a solution that satisfies (40), and Theorem 3 shows how to compute π from this intermediate value.

Lemma 9. *Suppose $\mathbf{A} \in \mathbb{F}_q^{(4t-1)\times(2t)}$, $\mathbf{b} \in \mathbb{F}_q^{4t-1}$ are defined in (43) and (44), respectively. Consider the following equation defined on $\mathbb{F}_q$:*

$$\mathbf{Ac} = \mathbf{b}. \tag{45}$$

For any vector $\mathbf{c} \in \mathbb{F}_q^{2t}$, $\mathbf{c}$ is a nonzero solution to (45) if and only if $(h_1(\mathbf{c}), h_2(\mathbf{c}))$ is a nonzero solution to (40).

Proof: The proof is in Appendix F. ■

Theorem 3. *Let $\mathbf{c}$ be an arbitrary nonzero solution to (45), and $h_1 = h_1(\mathbf{c})$, $h_2 = h_2(\mathbf{c})$. Denote h, v_1, v_2 by the following equations,*

$$h = GCD(h_1, h_2), v_1 = \frac{h_2}{h},\ v_2 = \frac{h_1}{h}. \tag{46}$$

Suppose V_1, V_2 are the sets of the additive inverses of roots of v_1, v_2, respectively. Then π can be computed from the following equation:

$$\pi = \nu^{-1}\left(V_1 \cup (B' \setminus V_2)\right).$$

Recall $B' = \nu(\pi')$, where ν is defined in (28).

Proof: Note that $B = \nu(\pi)$ and ν is an injection, so we only need to prove that $B = V_1 \cup (B' \setminus V_2)$. From (39), it follows that

$$h_1 \cdot f_1 - h_2 \cdot f_2 = (h_1 \cdot g_1 - h_2 \cdot g_2) \cdot g_3,$$

where $\deg g_3 = |B \cap B'| \geq N - 1 - t$. From Lemma 9, (40) holds true, which means that $\deg(h_1 \cdot g_1 - h_2 \cdot g_2) \cdot g_3 = \deg(h_1 \cdot f_1 - h_2 \cdot f_2) < N - 3t$. If $h_1 \cdot f_1 \neq h_2 \cdot f_2$, then $N - t - 1 < N - 3t$ and thus $t = 0$, $h_1 = h_2 = 0$. Therefore for any nonzero pair of h_1 and h_2,

$$h_1 \cdot f_1 = h_2 \cdot f_2.$$

We know from (46) that

$$v_2 \cdot f_1 = v_1 \cdot f_2,$$

where $\text{GCD}(v_1, v_2) = 1$. Let $v_2 | f_2$ and $v_1 | f_1$. Then,

$$\frac{f_1}{v_1} = \frac{f_2}{v_2} = f.$$

Suppose V_3 is the set of the additive inverses of roots of f. Then $V_1 \cup V_3 = B$, $V_2 \cup V_3 = B'$, thus $B = V_1 \cup V_3 = V_1 \cup (B' \setminus V_2)$. ■

Note that V_1, V_2 computed in Theorem 3 are exactly identical to D_1, D_2 described before (39), respectively.

Example 5. *Suppose the sender transmits the permutation $\pi_1 = (2,4,7,3,5,1,8,6,9,10) \in \mathcal{C}_{\boldsymbol{\alpha}}(10,2)$, where $\boldsymbol{\alpha} = (16, 0, 86, 44, 61, 9, 49)$, and the receiver recives $\pi' = (8,6,9,10, 5,1,2,4,7,3) \in \mathbb{S}_{10}$. In the encoding scheme, $q = 97 > 10^2 - 10$, and for all $i, j \in [10]$, $i \neq j$,*

$$\upsilon(i,j) = 10(i-1) + j - 1.$$

The receiver applies Newton's identities *[17] to compute $r(B) = (16,31,0,42,54,94,59)$ from $\boldsymbol{\alpha}$, and then derives $r(B') = (80,64,83,10,72,22,26)$ from $B' = \nu(\pi') = \{75, 58,89,94,40,1,13,36,62\}$. Then*

$$\mathbf{A} = \begin{pmatrix} 1 & 16 & 31 & 0 & 42 & 54 & 94 \\ 0 & 1 & 16 & 31 & 0 & 42 & 54 \\ 1 & 80 & 64 & 83 & 10 & 72 & 22 \\ 0 & 1 & 80 & 64 & 83 & 10 & 72 \end{pmatrix}^T, \tag{47}$$
$$\mathbf{b} = \begin{pmatrix} 64 & 33 & 83 & 65 & 18 & 25 & 64 \end{pmatrix}^T.$$

Notice that $\mathbf{c} = \big(95,94,66,26\big)$ is a solution to $\mathbf{Ac} = \mathbf{b}$. Therefore $h_1 = X^2 + 95X + 94 = (X+1)(X+94)$, $h_2 = X^2 + 31X + 71 = (X+24)(X+7)$. The receiver then knows that $V_1 = \{24, 7\}$, $V_2 = \{1, 94\}$. Therefore $\nu(\pi) = B = V_1 \cup (B' \setminus V_2) = \{13,36,62,24,40,7,75,58,89\}$. It follows that $A(\pi) = \{(2,4),(4,7),(7,3),(3,5),(5,1),(1,8), (8,6),(6,9),(9,10)\}$. From the definition of the characteristic set in Definition 4, the receiver is able to decode π from $A(\pi)$ as $\hat{\pi} = (2,4,7,3,5,1,8,6,9,10)$.

V. Systematic Permutation Codes in the Generalized Cayley Metric

In this section, we discuss systematic permutation codes. Specifically, in Section V-A, we present an explicit coding scheme for systematic permutation codes in the generalized Cayley metric, and in Section V-B, we provide the decoding scheme for this construction. We refine our construction to ensure order-optimality, which we then discuss in Section V-C.

A. Encoding Scheme

Let messages be permutations in $\mathbb{S}_N$. In systematic permutation codes, the codewords are permutations of length $N+M$. We derive each codeword $\sigma \in \mathbb{S}_{N+M}$ from a message $\pi \in \mathbb{S}_N$ by sequentially inserting components $N+1, N+2, \cdots, N+M$ into π, in the positions specified by a sequence $S = (s_1, s_2, \cdots, s_M)$, where S is determined by the syndrome $\alpha^{(q,2t)}(\pi)$ defined in (29) and (30). Our key result is established in Theorem 4, where we present the construction of systematic permutation codes. We start the discussion by presenting a collection of definitions and lemmas to support our main result.

Definition 5. *For any permutation $\pi \in \mathbb{S}_N$ and the integer $i \in \mathbb{N}$, where $1 \leq s \leq N$, let $E(\pi, s)$ be a permutation in $\mathbb{S}_{N+1}$ derived by inserting the element $N+1$ after the element s in π, i.e.,*

$$E(\pi, s) \triangleq (\pi(1), \cdots, \pi(k), N+1, \pi(k+1), \cdots, \pi(N)),$$

where $k = \pi^{-1}(s)$. We call $E(\pi, s)$ the ***extension*** *of π on the* ***extension point*** *s.*

Consider a sequence $S = (s_1, s_2, \cdots, s_M)$, where $s_m \in [N]$ for all $1 \leq m \leq M$. The ***extension*** *$E(\pi, S)$ of π on the* ***extension sequence*** *S is a permutation in $\mathbb{S}_{N+M}$ derived from inserting the elements $N+1, \cdots, N+M$ sequentially after the elements $s_1, \cdots, s_M$ in π, i.e.,*

$$E(\pi, S) \triangleq E(E(\cdots E(E(\pi, s_1), s_2) \cdots, s_{M-1}), s_M).$$

Note that in Definition 5, the elements $s_1, \cdots, s_M$ in the extension sequence S are not necessarily distinct. If different symbols are sequentially inserted after the same element, then they are all placed right after this element in descending order, as shown in Example 6.

Example 6. *Suppose $\pi = (1, 4, 5, 7, 6, 2, 3)$, $I = (4, 1, 2, 2)$, then*

$$E(\pi, I) = (1, 9, 4, 8, 5, 7, 6, 2, 11, 10, 3).$$

Based on the definition of the extensions, Algorithm 3 describes the major steps of our encoding scheme. The correctness of this scheme is proved later by Lemma 10 and Theorem 4.

Definition 6 presents the notion of the *jump points* of the extensions of two permutations. Then Lemma 10 states that the block permutation distance between two extensions is strictly larger than that of their original permutations if and only if the extension point of one of them is a jump point. Based on this result, we further introduce the notion of *jump index* and *jump set* in Definition 7. As shown in Remark 4, the block permutation distance of two permutations in $\mathbb{S}_N$ is lower bounded by the sum of that of their extensions and the cardinality of the jump set.

Algorithm 3 Encoding Scheme

```
Input:
    Information sequence: $\pi \in \mathbb{S}_N$;
    Number of additional symbols: $K$;
    Minimum block permutation distance: $2t+1$;
Output:
    Codeword: $\sigma$ ($\sigma \in \mathbb{S}_{N+K}$);
    Compute the syndrome $\boldsymbol{\alpha} = \alpha^{(q,2t)}(\pi)$ of $\pi$, which is defined in (29);
    Compute the extension sequence $S = \varphi(\boldsymbol{\alpha})$, where $\varphi$ is a function such that the image of $\varphi$ is a $t$-auxiliary set of length $K$ in the range $[N]$, as defined in Definition 9;
    Compute $\sigma = E(\pi, S)$, according to Definition 5;
    return $\sigma$.
```

Definition 6. *Let $\pi_1, \pi_2 \in \mathbb{S}_N$, $s_1, s_2 \in [N]$. We note that for any $k \in [N]$, $\pi_i(k)$ refers to the k-th element of π_i, $i \in \{1, 2\}$. Suppose $E(\pi_1, s_1)$, $E(\pi_2, s_2)$ are two arbitrary extensions of π_1 and π_2, respectively, where $\pi_1, \pi_2 \in \mathbb{S}_N$, $\pi_1(k_1) = s_1$ and $\pi_2(k_2) = s_2$. Then s_1 is called a* ***jump point*** *of $E(\pi_1, s_1)$ with respect to $E(\pi_2, s_2)$, if $s_1 \neq s_2$ and at least one of the following conditions is satisfied:*

1) $k_1 = N$ or $k_2 = N$;
2) $k_1, k_2 < N$, and $\pi_1(k_1+1) \neq \pi_2(k_2+1)$.

Lemma 10. *Let $\pi_1, \pi_2 \in \mathbb{S}_N$, $s_1, s_2 \in [N]$. For any two extensions $E(\pi_1, s_1)$ and $E(\pi_2, s_2)$, if s_1 is a jump point of $E(\pi_1, s_1)$ with respect to $E(\pi_2, s_2)$, then*

$$d_B(E(\pi_1, s_1), E(\pi_2, s_2)) > d_B(\pi_1, \pi_2), \tag{48}$$

else

$$d_B(E(\pi_1, s_1), E(\pi_2, s_2)) = d_B(\pi_1, \pi_2). \tag{49}$$

Proof: The proof is in Appendix G. ■

In the following Example 7, we provide examples of jump points that satisfy the two conditions indicated in Definition 6. We also provide an example of an extension point that is not a jump point.

Example 7. *Suppose $\pi = (1, 5, 7, 2, 3, 6, 4)$, $\pi' = (2, 3, 1, 5, 7, 6, 4)$, $s_1 = 4$, $s'_1 = 5$, $s_2 = 5$, $s'_2 = 6$, $s_3 = 3$, $s'_3 = 7$. Then,*

$$\begin{aligned} \sigma_1 = E(\pi, s_1) &= (1, 5, 7, 2, 3, 6, 4, 8), \\ \sigma'_1 = E(\pi', s'_1) &= (2, 3, 1, 5, 8, 7, 6, 4), \\ \sigma_2 = E(\pi, s_2) &= (1, 5, 8, 7, 2, 3, 6, 4), \\ \sigma'_2 = E(\pi', s'_2) &= (2, 3, 1, 5, 7, 6, 8, 4), \\ \sigma_3 = E(\pi, s_3) &= (1, 5, 7, 2, 3, 8, 6, 4), \\ \sigma'_3 = E(\pi', s'_3) &= (2, 3, 1, 5, 7, 8, 6, 4). \end{aligned}$$

Given that $d_B(\pi, \pi') = 2$, we observe that

$d_B(\sigma_1, \sigma'_1) = 4 > d_B(\pi, \pi')$, and s_1 is a jump point;
$d_B(\sigma_2, \sigma'_2) = 5 > d_B(\pi, \pi')$, and s_2 is a jump point;
$d_B(\sigma_3, \sigma'_3) = 2 = d_B(\pi, \pi')$, and s_3 is not a jump point.

Notice that s_1 is a jump point that satisfies the first condition in Definition 6, and s_2 is a jump point that satisfies the second condition. This example is consistent with Lemma 10.

We know from Lemma 10 that the block permutation distance between the resulting codewords cannot be smaller than that of their original messages. Recall that Theorem 2 indicates that permutations with the same syndrome result in codewords having the block permutation distance of at least $2t+1$. Therefore, it suffices to show that the permutations with different syndromes are mapped to codewords that are sufficiently far apart under the block permutation distance; Lemma 11 establishes a property that ensures that this condition is satisfied. We then use this result in Theorem 4 to present the construction of systematic permutation codes.

Definition 7. *Let $\pi_1,\pi_2 \in \mathbb{S}_N$, $s_1,s_2 \in [N]$. Suppose $E(\pi_1, S_1)$ and $E(\pi_2,S_2)$ are extensions of π_1 and π_2 on extension sequences S_1 and S_2, respectively, where $\pi_1,\pi_2 \in \mathbb{S}_N$, $S_1 = (s_{1,1},s_{1,2},\cdots,s_{1,M})$ and $S_2 = (s_{2,1},s_{2,2},\cdots,s_{2,M})$. Then, for any $m \in [M]$, m is called a **jump index** of $E(\pi_1,S_1)$ and $E(\pi_2,S_2)$ if $s_{1,m}$ is a jump point of $E(E(\pi_1,J_{1,m-1}),s_{1,m})$ with respect to $E(E(\pi_2,J_{2,m-1}),s_{2,m})$, where $J_{1,m-1} = (s_{1,1},s_{1,2},\cdots,s_{1,m-1})$, $J_{2,m-1} = (s_{2,1},s_{2,2},\cdots,s_{2,m-1})$. Define the **jump set** $F(\pi_1,\pi_2,S_1,S_2)$ as the set of all jump indices of $E(\pi_1,S_1)$ and $E(\pi_2,S_2)$.*

Remark 4. *Let $\pi_1,\pi_2 \in \mathbb{S}_N$, $s_1,s_2 \in [N]$. For any extensions $E(\pi_1,S_1)$, $E(\pi_2,S_2)$ of π_1, π_2 on extension sequences S_1, S_2, respectively, it is obvious from Definition 7 and Lemma 10 that*

$$d_B(E(\pi_1,S_1),E(\pi_2,S_2)) \geq d_B(\pi_1,\pi_2)+|F(\pi_1,\pi_2,S_1,S_2)|. \tag{50}$$

Here $F(\pi_1,\pi_2,S_1,S_2)$ is the jump set defined in Definition 7.

In the following Example 8, we provide an example of how to identify the jump indices and compute the jump set. This example satisfies inequality (50).

Example 8. *Continuing with the values of π, π' specified in Example 7, let $S=(4,6,7)$ and $S'=(5,6,5)$. Then,*

$$\begin{aligned}
\sigma_0 = \pi &= (1,5,7,2,3,6,4)\,,\\
\sigma_0' = \pi' &= (2,3,1,5,7,6,4)\,,\\
\sigma_1 = E(\sigma_0,s_1) &= (1,5,7,2,3,6,4,8)\,,\\
\sigma_1' = E(\sigma_0',s_1') &= (2,3,1,5,8,7,6,4)\,,\\
\sigma_2 = E(\sigma_1,s_2) &= (1,5,7,2,3,6,9,4,8)\,,\\
\sigma_2' = E(\sigma_1',s_2') &= (2,3,1,5,8,7,6,9,4)\,,\\
\sigma_3 = E(\sigma_2,s_3) &= (1,5,7,10,2,3,6,9,4,8)\,,\\
\sigma_3' = E(\sigma_2',s_3') &= (2,3,1,5,10,8,7,6,9,4)\,.
\end{aligned}$$

It follows immediately that

$$\begin{aligned}
&d_B(\sigma_0,\sigma_0') = 2,\\
&d_B(\sigma_1,\sigma_1') = 4 > d_B(\sigma_0,\sigma_0'),\ \textit{and } 1 \textit{ is a jump index};\\
&d_B(\sigma_2,\sigma_2') = 4 = d_B(\sigma_1,\sigma_1'),\ \textit{and } 2 \textit{ is not a jump index};\\
&d_B(\sigma_3,\sigma_3') = 5 > d_B(\sigma_2,\sigma_2'),\ \textit{and } 3 \textit{ is a jump index.}
\end{aligned}$$

According to Definition 7, $F(\pi,\pi',S,S') = \{1,3\}$. Moreover, $d_B(\sigma_3,\sigma_3') = 5 > 4 = d_B(\pi,\pi') + |F(\pi,\pi',S,S')|$, which is in accordance with equation (50).

Next we prove in Lemma 11 that the right hand side of equation (50) can be lower bounded by the cardinality of the so-called *Hamming set*. The Hamming set of S_1 with respect to S_2 is defined in the following Definition 8. Based on this result, we present a construction of systematic t-block permutation codes in Theorem 4 with the help of a so-called *t-auxiliary set* that is defined in Definition 9.

Definition 8. *For any sequences $\mathbf{v}_1$, $\mathbf{v}_2$ of integers with length M, where $\mathbf{v}_1 = (v_{1,1},v_{1,2},\cdots,v_{1,M})$ and $\mathbf{v}_2 = (v_{2,1},v_{2,2},\cdots,v_{2,M})$, define the **Hamming set** of $\mathbf{v}_1$ with respect to $\mathbf{v}_2$ as follows,*

$$H(\mathbf{v}_1,\mathbf{v}_2) \triangleq \{v_{1,m} | v_{1,m} \neq v_{2,m}, m \in [M]\}. \tag{51}$$

We note that d_H refers to the Hamming distance throughout this paper.

Remark 5. *It is obvious that $d_H(\mathbf{v}_1,\mathbf{v}_2) \geq |H(\mathbf{v}_1,\mathbf{v}_2)|$. Additionally, for any three sequences $\mathbf{v}_1,\mathbf{v}_2,\mathbf{v}_3$ of integers, the following triangle inequality holds true:*

$$|H(\mathbf{v}_1,\mathbf{v}_3)| \leq |H(\mathbf{v}_1,\mathbf{v}_2)| + |H(\mathbf{v}_2,\mathbf{v}_3)|. \tag{52}$$

Lemma 11. *Let $\pi_1,\pi_2 \in \mathbb{S}_N$, $s_1,s_2 \in [N]$. For any extensions $E(\pi_1,S_1)$, $E(\pi_2,S_2)$ of π_1, π_2 on extension sequences S_1, S_2, respectively, it follows that*

$$d_B(E(\pi_1,S_1),E(\pi_2,S_2)) \geq |H(S_1,S_2)|. \tag{53}$$

Proof: The proof is in Appendix H. ■

Example 9. *Continuing on with the numerical values of π, π', S,S' as in Example 8, we conclude that, $H(S,S') = \{4, 7\}$, $m(4) = 1$, $m(7) = 3$. Then it follows that $d_B(\sigma,\sigma') = 5 > 2 = |H(S,S')|$, which is in accordance with the above Lemma 11.*

Definition 9. *Consider a set $\mathcal{A}(N,K,t) \subset [N]^K$. We call $\mathcal{A}(N,K,t)$ a **t-auxiliary set** of length K in range $[N]$ if for any $\mathbf{c}_1,\mathbf{c}_2 \in \mathcal{A}(N,K,t)$, $\mathbf{c}_1 \neq \mathbf{c}_2$, $|H(\mathbf{c}_1,\mathbf{c}_2)| \geq 2t+1$ holds.*

Theorem 4. *For any t-auxiliary set $\mathcal{A}(N,K,t)$ with cardinality that is no less than q^{4t-1}, suppose $\varphi:\ \alpha^{(q,2t)}(\mathbb{S}_N) \rightarrow \mathcal{A}(N,K,t)$ is an arbitrary injection, where q is a prime number such that $N^2-N < q < 2(N^2-N)$ and the syndrome $\alpha^{(q,2t)}$ is defined in (29) and (30). Then, the set $\mathcal{C}_B^{\mathrm{sys}}(N,K, t) = \{E(\pi,\varphi \circ \alpha^{(q,2t)}(\pi)) | \pi \in \mathbb{S}_N\}$ is a systematic t-block permutation code.*

Proof: It is clear by the choice of $E(\pi,S)$ that $\mathcal{C}_B^{\mathrm{sys}}(N, K,t)$ is systematic. For any two messages $\pi_1,\pi_2 \in \mathbb{S}_N$, denote their corresponding codewords by $\sigma_1 = E(\pi_1,\varphi \circ \alpha^{(q,2t)}(\pi_1))$ and $\sigma_2 = E(\pi_2,\varphi \circ \alpha^{(q,2t)}(\pi_2))$, respectively. Suppose $\boldsymbol{\alpha}_1 = \alpha^{(q,2t)}(\pi_1)$, $\boldsymbol{\alpha}_2 = \alpha^{(q,2t)}(\pi_2)$, $S_1 = \varphi(\boldsymbol{\alpha}_1)$ and $S_2 = \varphi(\boldsymbol{\alpha}_2)$. Then $\sigma_1 = E(\pi_1,S_1)$, $\sigma_2 = E(\pi_2,S_2)$. Consider the following two cases:

1) $\boldsymbol{\alpha}_1 = \boldsymbol{\alpha}_2$. According to Theorem 2, $d_B(\pi_1,\pi_2) > 2t$ in this case. Then Lemma 10 implies that $d_B(\sigma_1,\sigma_2) \geq d_B(\pi_1,\pi_2) \geq 2t+1$.
2) $\boldsymbol{\alpha}_1 \neq \boldsymbol{\alpha}_2$. In this case, $S_1,S_2 \in \mathcal{A}(N,K,t)$ and $S_1 \neq S_2$. Then from Definition 9, $|H(S_1,S_2)| \geq 2t+1$. Therefore, from Lemma 11, $d_B(\sigma_1,\sigma_2) \geq |H(S_1,S_2)| \geq 2t+1$.

From the above discussion, $d_B(\sigma_1,\sigma_2)\geq 2t+1$ is aways true, which means that $\mathcal{C}_B^{\text{sys}}(N,K,t)$ is indeed a systematic t-block permutation code. ■

B. Decoding Scheme

Based on the construction and the notation in Theorem 4, suppose the sender sends a codeword $\sigma = E(\pi, \varphi\circ\alpha^{(q,2t)}(\pi))$ through a noisy channel and the receiver receives a noisy version σ', where $d_B(\sigma,\sigma')\leq t$.

In this section, we prove in the forthcoming Lemma 12 that the extension sequence S of the codeword $E(\pi,S)$ is decodable given that $d_B(\sigma,\sigma')\leq t$, from which the syndrome, defined in (29) and (30), of the transmitted information π can be derived.

For convenience, we introduce the following definition of the *truncation* and use it throughout this subsection.

Definition 10. *For any permutation $\sigma\in\mathbb{S}_{N+1}$ and an integer $u\in[N+1]$, denote $T(\sigma,u)$ to be the sequence derived by removing the element u from σ, i.e.,*

$$T(\sigma,u)\triangleq(\sigma(1),\sigma(2),\cdots,\sigma(k-1),\sigma(k+1),\cdots,\sigma(N)), \tag{54}$$

where $k=\sigma^{-1}(u)$.

Then, for any permutation $\sigma\in\mathbb{S}_{N+M}$ and a set $U\subset[N+M]$, denote the ***truncation*** *$T(\sigma,U)$ of σ on set U to be the sequence derived by removing the elements contained in $U=\{u_1,u_2,\cdots,u_{|U|}\}$ from σ, i.e.,*

$$T(\sigma,U)\triangleq T(T(\cdots T(T(\sigma,u_1),u_2)\cdots,u_{|U|-1}),u_{|U|}). \tag{55}$$

Note that in Definition 10, the ordering of $u_1,\cdots,u_{|U|}$ has no impact on the value of $T(\sigma,U)$. The following is an example of the truncation of a permutation.

Example 10. *Suppose $\sigma=(1,4,5,2,3,9,8,6,7)$, $U=\{4,5,9\}$, then*

$$T(\sigma,U)=(1,2,3,8,6,7)\,.$$

The following Algorithm 4 describes the decoding algorithm of the code constructed in Theorem 4. The correctness of this algorithm is proved by Lemma 12.

Our decoding scheme has two major steps. Recall that $\alpha^{(q,2t)}$ is defined in (29) and (30) as the syndrome of π. The first step is to derive the syndrome $\hat{\boldsymbol{\alpha}}=\alpha^{(q,2t)}(\pi)$ of $\pi=T(\sigma,\{N+1,\cdots,N+K\})$, from the received permutation σ'. The second step is to apply Algorithm 2 to the pair of inputs, the syndrome $\hat{\boldsymbol{\alpha}}$ and the subsequence $\pi'=T(\sigma',\{N+1,\cdots,N+K\})$, and compute π.

Note that it is sufficient to compute the sequence S in order to derive the syndrome $\hat{\boldsymbol{\alpha}}$. Lemma 12 proves the sufficiency of obtaining the sequence S from S', where S is the extension sequence of π in σ, by showing that the cardinality of the Hamming set $H(S,S')$ does not exceed t, provided that $d_B(\sigma,\sigma')\leq t$. Therefore, from (52) and Definition 9, we are able to obtain an estimate $\hat{S}$ of S from S' since each t-auxiliary set $\mathcal{A}(N,K,t)$ has the property that the cardinalities of Hamming sets constructed from its pairwise distinct elements are at least $2t+1$. The syndrome $\hat{\boldsymbol{\alpha}}$ is then uniquely derived from $\hat{S}$.

Algorithm 4 Decoding Algorithm

```
Input:
    Received sequence: $\sigma'$;
    Number of additional symbols: $K$;
    Minimum block permutation distance: $2t+1$;
Output:
    Estimated information sequence: $\hat{\pi}$;
Compute $\pi'=T(\sigma',\{N+1,\cdots,N+K\})$, according to Definition 10;
Find $S'$ such that $\sigma'=E(\pi',S')$, where $E(\pi,S)$ is defined in Definition 5;
Find $\hat{S}\in\mathrm{Img}(\varphi)$ such that $H(\hat{S},S')\leq t$, where $H$ is defined in Definition 8, and $\varphi$ is specified in Theorem 4;
Compute $\hat{\boldsymbol{\alpha}}=\varphi^{-1}(\hat{S})$;
Let $\hat{\boldsymbol{\alpha}},\pi'$ be the inputs of Algorithm 2 and obtain $\hat{\pi}$;
return $\hat{\pi}$.
```

Lemma 12. *Consider an arbitrary $\sigma\in\mathcal{C}=\{E(\pi,\varphi\circ\alpha^{(q,2t)}(\pi))|\pi\in\mathbb{S}_N\}$, for $\mathcal{C}$ defined in Theorem 4 (then $\sigma\in\mathbb{S}_{N+K}$). Suppose there is a σ' such that $d_B(\sigma,\sigma')\leq t$. Let $S=\varphi\circ\alpha^{(q,2t)}(\pi)$ and $\pi'=T(\sigma',[N+1:N+K])$. Suppose σ' is the extension of π' on the extension sequence S', i.e., $\sigma'=E(\pi',S')$. Then,*

$$H(S,S')\leq t. \tag{56}$$

Proof: Suppose $S=(s_1,s_2,\cdots,s_K)$, $S'=(s'_1,s'_2,\cdots,s'_K)$. Then, according to Theorem 4, $S\in\mathcal{A}(N,K,t)$. Let $\mathcal{M}=\{m|s_m\neq s'_m, 1\leq m\leq K\}$. For all $m\in\mathcal{M}$, it follows from Definition 6 that there exist subsequences of σ,σ': $\mathbf{p}_m=(s_m,n_{k(m)},n_{k(m)-1},\cdots,n_1,N+m)$ and $\mathbf{p}'_m=(s'_m,n'_{k'(m)},n'_{k'(m)-1},\cdots,n'_1,N+m)$, where $k(m),k'(m)\in[K]$, $n_1,n_2,\cdots,n_{k(m)},n'_1,n'_2,\cdots,n'_{k(m)'}\in[N+1:N+K]$. Note that $s_m\neq s'_m$, which means that $(s_m,n_{k(m)},n_{k(m)-1},\cdots,n_1)\neq(s'_m,n'_{k'(m)},n'_{k'(m)-1},\cdots,n'_1)$. Let

$$i(m)=\min_{\substack{1\leq i\leq\min\{k(m),k'(m)\}\\ n_i\neq n'_i}} i.$$

Then $n_{i(m)}\neq n'_{i(m)}$ and $n_{i(m)-1}=n'_{i(m)-1}$, where we let $n_0=n'_0=N+m$ if $i(m)=1$.

Recall the notion of characteristic sets in Definition 3. We know that $(n_{i(m)},n_{i(m)-1})\in A(\sigma)$ and $(n'_{i(m)},n'_{i(m)-1})\in A(\sigma')$. These two conditions along with the fact that $n_{i(m)}\neq n'_{i(m)}$ and $n_{i(m)-1}=n'_{i(m)-1}$ imply that $(n_{i(m)},n_{i(m)-1})\in(A(\sigma)\setminus A(\sigma'))$ for all $m\in\mathcal{M}$. Notice that for all $s_m\in\{s_m: m\in\mathcal{M}\}=H(S,S')$, the associated subsequences $\mathbf{p}_m$ start with different s_m and they do not overlap, which indicates that the pairs $(n_{i(m)},n_{i(m)-1})$ are distinct. Then $|A(\sigma)\setminus A(\sigma')|\geq|H(S,S')|$, which is equivalent to $H(S,S')\leq d_B(\sigma,\sigma')\leq t$. ■

From Lemma 12, the receiver first computes $\pi'=T(\sigma',\{N+1,\cdots,N+K\})$ and derives the extension sequence S' such that $\sigma'=E(\pi',S')$. Then, the receiver decodes $\hat{S}=\varphi\circ\alpha^{(q,2t)}(\pi)\in\mathcal{A}(N,K,t)$ from S' such that $|H(S',\hat{S})|\leq t$

and derives $\hat{\boldsymbol{\alpha}}$ from $\hat{S}$. From Lemma 10, $d_B(\pi, \pi') \leq d_B(\sigma, \sigma') \leq t$ follows. Then, the receiver can apply Algorithm 2 to compute $\hat{\pi}$ from π' and $\hat{\boldsymbol{\alpha}}$ reliably. The decoding scheme for the systematic t-block permutation code $\mathcal{C}$ constructed in Theorem 4 is then complete.

C. Order-optimal Systematic t-Block Permutation Codes

Theorem 4 presents the construction of systematic t-block permutation codes with K redundant symbols based on a t-auxiliary set $\mathcal{A}(N, K, t)$. When N is sufficiently large and K is relatively small compared to N, the code rate is $1 - \Theta(\frac{K}{N})$, which is not necessarily order-optimal. In this section, based on the upcoming Lemma 13 and Theorem 5, we provide an explicit construction of a t-auxiliary set of length $K = 56t$ in Theorem 6, from which we are able to explicitly construct an order-optimal permutation code by Theorem 4.

Lemma 13. *For all $k, N \in \mathbb{N}^*$, $k > 3$, $N > k^2$, consider an arbitrary subset $Y \subset [k]$, where $|Y| = M < k$, $Y = \{i_1, i_2, \cdots, i_M\}$, then*

$$\mathrm{LCM}\left(N + i_1, N + i_2, \cdots, N + i_M\right) > N^{M - \frac{k}{2}}. \qquad (57)$$

Proof: The proof is in Appendix I. ■

Theorem 5. *For all $N, k, d \in \mathbb{N}^*$, $N > k^2$, $k > 3$, define a function $\beta^{(q,d,k)} : \mathbb{F}_q^d \rightarrow [N+1] \times [N+2] \times \cdots \times [N+k]$ as follows:*

$$\begin{aligned}\beta^{(q,d,k)}(\boldsymbol{x}) = &\left(\beta_1^{(q,d,k)}(\boldsymbol{x}), \beta_2^{(q,d,k)}(\boldsymbol{x}), \cdots, \beta_k^{(q,d,k)}(\boldsymbol{x})\right)\\ &\triangleq(\gamma(\boldsymbol{x}) \bmod (N+1), \gamma(\boldsymbol{x}) \bmod (N+2),\\ &\cdots, \gamma(\boldsymbol{x}) \bmod (N+k)),\end{aligned} \qquad (58)$$

where $\boldsymbol{x} = (x_1, x_2, \cdots, x_d) \in \mathbb{F}_q^d$, $\gamma(\boldsymbol{x}) \triangleq \sum_{i=1}^{d} x_i q^{i-1}$. Then $\forall$ $\boldsymbol{x}_1, \boldsymbol{x}_2 \in \mathbb{F}_q^d$, $\boldsymbol{x}_1 \neq \boldsymbol{x}_2$,

$$d_H(\beta^{(q,d,k)}(\boldsymbol{x}_1), \beta^{(q,d,k)}(\boldsymbol{x}_2)) > \frac{k}{2} - d(2 + \log_N 2). \qquad (59)$$

Proof: For arbitrary $\boldsymbol{x}_1, \boldsymbol{x}_2 \in \mathbb{F}_q^d$, $\boldsymbol{x}_1 \neq \boldsymbol{x}_2$, let $\beta^{(q,d,k)}(\boldsymbol{x}_1) = (\beta_{1,1}, \beta_{1,2}, \cdots, \beta_{1,k})$, $\beta^{(q,d,k)}(\boldsymbol{x}_2) = (\beta_{2,1}, \beta_{2,2}, \cdots, \beta_{2,k})$. Let $Z = \{i : \ \beta_{1,i} = \beta_{2,i}, \ 1 \leq i \leq d\}$, then $d_H(\beta^{(q,d,k)}(\boldsymbol{x}_1), \beta^{(q,d,k)}(\boldsymbol{x}_2)) = k - |Z| = k - M$, where $M = |Z|$.

Suppose $Z = \{i_1, i_2, \cdots, i_M\}$. Let $\gamma_1 = \gamma(\boldsymbol{x}_1)$, $\gamma_2 = \gamma(\boldsymbol{x}_2)$. According to the definition of $\beta^{(q,d,k)}$ in (58),

$$\begin{cases} \gamma_1 & \equiv & \gamma_2 & \bmod (N + i_1) \\ \gamma_1 & \equiv & \gamma_2 & \bmod (N + i_2) \\ & \vdots & & \\ \gamma_1 & \equiv & \gamma_2 & \bmod (N + i_M). \end{cases}$$

Then,

$$\gamma_1 \equiv \gamma_2 \mod \mathrm{LCM}\left(N + i_1, N + i_2, \cdots, N + i_M\right).$$

Given that $\boldsymbol{x}_1, \boldsymbol{x}_2 \in \mathbb{F}_q^d$, $\boldsymbol{x}_1 \neq \boldsymbol{x}_2$, then $\gamma_1 \neq \gamma_2$. From Lemma 13, it follows that

$$|\gamma_1 - \gamma_2| \geq \mathrm{LCM}\left(N + i_1, N + i_2, \cdots, N + i_M\right) > N^{M - \frac{k}{2}}. \qquad (60)$$

Moreover, the condition $\boldsymbol{x}_1, \boldsymbol{x}_2 \in \mathbb{F}_q^d$, $\boldsymbol{x}_1 \neq \boldsymbol{x}_2$ implies that $0 \leq \gamma_1, \gamma_2 < q^d$ and $\gamma_1 \neq \gamma_2$. Therefore,

$$|\gamma_1 - \gamma_2| < q^d. \qquad (61)$$

According to (60) and (61), $N^{M - \frac{k}{2}} < |\gamma_1 - \gamma_2| < q^d < (2N^2)^d$ is true, which means that $M - \frac{k}{2} < d(2 + \log_N 2)$. Therefore $M < \frac{k}{2} + d(2 + \log_N 2)$, and then

$$\begin{aligned} d_H(\boldsymbol{\beta}_1, \boldsymbol{\beta}_2) = k - M &> k - (\frac{k}{2} + d(2 + \log_N 2)) \\ &= \frac{k}{2} - d(2 + \log_N 2). \end{aligned}$$

The theorem is proved. ■

Example 11. *Let $k = 7$, $N = 50$, $d = 1$, $q = 2503$, $\boldsymbol{x}_1 = (280)$, $\boldsymbol{x}_2 = (1008)$, then $\gamma_1 = 280$, $\gamma_2 = 1008$, and*

$$\begin{aligned} \boldsymbol{\beta}_1 =& (280 \bmod 51, 280 \bmod 52, \cdots, 280 \bmod 57) \\ =& (25, 20, 15, 10, 5, 0, 52), \\ \boldsymbol{\beta}_2 =& (1008 \bmod 51, 1008 \bmod 52, \cdots, 1008 \bmod 57) \\ =& (39, 20, 1, 36, 18, 0, 39). \end{aligned}$$

Then $d_H(\boldsymbol{\beta}_1, \boldsymbol{\beta}_2) = 5 > \frac{k}{2} - d(2 + \log_N 2)$, which is in accordance with Theorem 5.

Based on Theorem 5, we provide an explicit construction of a t-auxiliary set $\mathcal{A}(N, 56t, t)$ in the following Theorem 6.

Theorem 6. *For all $N, k, t \in \mathbb{N}^*$, $k \geq 28t$, $k < \lfloor\sqrt{N} - \frac{1}{2}\rfloor$. Suppose $\mathbb{F}_q^{4t-1} = \{\boldsymbol{x}_1, \boldsymbol{x}_2, \cdots, \boldsymbol{x}_{q^{4t-1}}\}$, where q is a prime number such that $N^2 - N < q < 2N^2 - 2N$. For any $s \in \left[q^{4t-1}\right]$, suppose $\boldsymbol{x}_s = (x_1, x_2, \cdots, x_{4t-1})$, let $\mathbf{c}_s = (c_1, c_2, \cdots, c_{2k})$, $\beta^{(q,4t-1,k)}(\boldsymbol{x}_s) = (\beta_1, \beta_2, \cdots, \beta_k)$ for all $1 \leq i \leq k$, where $\mathbf{c}_s$ is defined as follows:*

$$\begin{cases} c_{2i} = (i-1)\lfloor\frac{N}{k}\rfloor + 1 + \left(\beta_i \mod \lfloor\frac{N}{k}\rfloor\right), \\ c_{2i-1} = (i-1)\lfloor\frac{N}{k}\rfloor + 1 + \left\lfloor\frac{\beta_i}{\lfloor\frac{N}{k}\rfloor}\right\rfloor. \end{cases} \qquad (62)$$

Then $\mathcal{A}(N, 2k, t) = \{\mathbf{c}_s : \ s \in \left[q^{4t-1}\right]\}$ is a t-auxiliary set with cardinality q^{4t-1}.

Proof: Without loss of generality, we prove the statement for $\boldsymbol{x}_1, \boldsymbol{x}_2 \in \mathbb{F}_q^{4t-1}$, $\boldsymbol{x}_1 \neq \boldsymbol{x}_2$, let $\boldsymbol{\beta}_1 = \beta^{(q,4t-1,k)}(\boldsymbol{x}_1)$, $\boldsymbol{\beta}_2 = \beta^{(q,4t-1,k)}(\boldsymbol{x}_2)$. Then, according to Theorem 5,

$$\begin{aligned} d_H(\boldsymbol{\beta}_1, \boldsymbol{\beta}_2) &> \frac{k}{2} - (4t-1)(2 + \log_N 2) \\ &> \frac{k}{2} - (12t - 3) > \frac{28t}{2} - 12t = 2t. \end{aligned}$$

In equation (62), let $m_i = (i-1)\lfloor\frac{N}{k}\rfloor + 1$. Notice that $(c_{2i-1} - m_i)\lfloor\frac{N}{k}\rfloor + (c_{2i} - m_i) = \beta_i$, for $1 \leq i \leq k$. Given $\beta_i \leq N + k$ for all $1 \leq i \leq k$, and $k < \lfloor\sqrt{N} - \frac{1}{2}\rfloor$, it follows that

$$\begin{aligned}\left\lfloor\frac{N}{k}\right\rfloor^2 &> \left(\frac{N}{k}-1\right)^2 \geq \left(\frac{N}{\sqrt{N}-\frac{3}{2}}-1\right)^2\\ &> \left(\sqrt{N}+\frac{3}{2}-1\right)^2 = \left(\sqrt{N}+\frac{1}{2}\right)^2\\ &> N+\sqrt{N} > N+k \geq \beta_i.\end{aligned}$$

Therefore, $(c_{2i-1}-m_i, c_{2i}-m_i)$ is exactly the $\lfloor\frac{N}{k}\rfloor$-ary representation of β_i, for all $1 \leq i \leq k$.

Suppose $\boldsymbol{\beta}_1 = (\beta_{1,1}, \beta_{1,2}, \cdots, \beta_{1,k})$ and $\boldsymbol{\beta}_2 = (\beta_{2,1}, \beta_{2,2}, \cdots, \beta_{2,k})$. Let $Y = \{i : \ \beta_{1,i} \neq \beta_{2,i}, \ 1 \leq i \leq k\}$, then $|Y| = d_H(\boldsymbol{\beta}_1, \boldsymbol{\beta}_2)$. Notice that for all $i \in Y$, $\beta_{1,i} \neq \beta_{2,i}$, then either $c_{1,2i-1}-m_i \neq c_{2,2i-1}-m_i$ or $c_{1,2i}-m_i \neq c_{2,2i}-m_i$, which means that

$$|H(\mathbf{c}_1, \mathbf{c}_2) \cap \{c_{1,2i-1}, c_{1,2i}\}| \geq 1, \ i \in Y. \tag{63}$$

Notice that $(i-1)\lfloor\frac{N}{k}\rfloor < c_{1,2i-1}, c_{1,2i} \leq i\lfloor\frac{N}{k}\rfloor$, and therefore,

$$\{c_{1,2i-1}, c_{1,2i}\} \cap \{c_{1,2i'-1}, c_{1,2i'}\} = \emptyset, \ \forall\ 1 \leq i < i' \leq k. \tag{64}$$

From (63) and (64),

$$\begin{aligned}|H(\mathbf{c}_1, \mathbf{c}_2)| &= \sum_{i=1}^{k} |H(\mathbf{c}_1, \mathbf{c}_2) \cap \{c_{1,2i-1}, c_{1,2i}\}|\\ &\geq \sum_{i \in Y} |H(\mathbf{c}_1, \mathbf{c}_2) \cap \{c_{1,2i-1}, c_{1,2i}\}|\\ &\geq \sum_{i \in Y} 1 = |Y| = d_H(\boldsymbol{\beta}_1, \boldsymbol{\beta}_2) > 2t.\end{aligned}$$

From Definition 9, $\mathcal{A}(N,k,t)$ is indeed a t-auxiliary set. ■

Remark 6. *Suppose we use $k = 28t$ in Theorem 6 to construct a t-auxiliary set $\mathcal{A}(N, 56t, t)$. Then the code $\mathcal{C}_B^{\mathrm{sys}}(N, 56t, t)$ constructed using Theorem 4 based on this $\mathcal{A}(N, 56t, t)$ is an order-optimal systematic t-block permutation code.*

VI. Comparison of Cardinality of the Codebooks

In Section IV, we constructed a t-generalized Cayley code $\mathcal{C}_G(N,t) = \mathcal{C}_{\boldsymbol{\alpha}}(N, 4t)$. Let the cardinality of $\mathcal{C}_G(N,t)$ be $A_G(N,t)$. In [5], a t-generalized Cayley code with cardinality $A_{\rho_g C}(N,t)$ was constructed. We next compare in Lemma 14 the logarithms of the cardinalities of these two codes, which reflects the redundancy in terms of bits. We show that the proposed scheme requires a smaller number of redundant bits than its counterpart presented in [5] for sufficiently large N and $t = o(\frac{N}{\log N})$.

Lemma 14. $\log|A_G(N,t)| > \log|A_{\rho_g C}(N,t)|$ *when* $t < \frac{N}{(16\log N+8)}$ *for sufficiently large* N.

Proof:

We know from [12, Appendix A] that:

$$\log|A_{\rho_g C}(N,t)| \leq N\log N - (2+\log e)N + O\left((\log N)^2\right). \tag{65}$$

Also,

$$\begin{aligned}&\log|A_G(N,t)|\\ &> \log N! - (16t(2\log N+1))\\ &> \left(N+\frac{1}{2}\right)\log N - (\log e)N - 16t(2\log N+1).\end{aligned} \tag{66}$$

Then,

$$\begin{aligned}&\log|A_G(N,t)| - \log|A_{\rho_g C}(N,t)|\\ &> \left(N+\frac{1}{2}\right)\log N - (\log e)N - 16t(2\log N+1)\\ &\quad - \left(N\log N - (2+\log e)N + O\left((\log N)^2\right)\right)\\ &= \frac{1}{2}\log N + 2N - 16t(2\log N+1) + O\left((\log N)^2\right)\end{aligned} \tag{67}$$

for sufficiently large N and $t < \frac{N}{(16\log N+8)}$.

From the above discussion, our proposed code in Section IV indeed has a higher rate than the interleaving-based code for sufficiently large N and $t = o\left(\frac{N}{\log N}\right)$. ■

Based on Remark 6 in Section V, we presented a construction of systematic t-generalized Cayley code $\mathcal{C}'_G(N,t) = \mathcal{C}_B^{\mathrm{sys}}(N, 56{\cdot}4t, 4t) = \mathcal{C}_B^{\mathrm{sys}}(N, 224t, 4t)$ with cardinality $A'_G(N, t)$.

In the next Lemma 15, we compare the logarithm of $A'_G(N, t)$ with that of $A_{\rho_g C}(N,t)$.

Lemma 15. $A'_G(N,t) > A_{\rho_g C}(N,t)$ *when* $t < \min\{\frac{N}{112\log N}, \frac{1}{112}\lfloor\sqrt{N}-\frac{1}{2}\rfloor\}$ *for sufficiently large* N.

Proof: We know from Lemma 7 that:

$$\log|A'_G(N,t)| > (N+\frac{1}{2})\log N - (\log e)N - 224t\log N. \tag{68}$$

Then it follows from (68) and (65) that

$$\begin{aligned}&\log|A'_G(N,t)| - \log|A_{\rho_g C}(N,t)|\\ &> \left(N+\frac{1}{2}\right)\log N - (\log e)N - 224t\log N\\ &\quad - \left(N\log N - (2+\log e)N + O\left((\log N)^2\right)\right)\\ &= \frac{1}{2}\log N + 2N - 224t\log N + O\left((\log N)^2\right).\end{aligned} \tag{69}$$

for sufficiently large N and $t < \min\{\frac{N}{112\log N}, \frac{1}{112}\lfloor\sqrt{N}-\frac{1}{2}\rfloor\}$.

From the above discussion, our proposed systematic code indeed has a higher rate than the interleaving-based code, for sufficiently large N and $t = o\left(\frac{N}{\log N}\right)$, in the generalized Cayley distance. ■

VII. Conclusion

The generalized Cayley metric is a distance measure that generalizes the Kendall-tau metric and the Ulam metric. Interleaving was previously shown to be convenient in constructions of permutation codes in the generalized Cayley metric. However, interleaving incurs a noticeable rate penalty such that the constructed permutation codes cannot be order-optimal. In this paper, we presented a framework for constructing order-optimal permutation codes that does not require interleaving.

Based on this framework, we then presented an explicit construction of systematic permutation codes from so-called extensions of permutations. We further provided a systematic construction that is order-optimal. Lastly, we proved that our proposed codes are more rate efficient than the existing coding schemes based on interleaving for sufficiently large N and $t = o\left(\frac{N}{\log N}\right)$.

Appendix A
Proof of Lemma 2

Lemma 2. *For all $\pi_1, \pi_2 \in \mathbb{S}_N$,*

$$d_B(\pi_1, \pi_2) = |A(\pi_2) \setminus A(\pi_1)| = |A(\pi_1) \setminus A(\pi_2)|.$$

Proof: According to the symmetry property of the block permutation distance, it is sufficient to prove $d_B(\pi_1, \pi_2) = |A(\pi_1) \setminus A(\pi_2)|$.

Suppose $\pi_1, \pi_2 \in \mathbb{S}_N$ such that $d_B(\pi_1, \pi_2) = d$. Then, there exists $\sigma \in \mathbb{S}_{d+1}$, $\psi_1, \psi_2, \cdots, \psi_{d+1}$, such that $\pi_1 = (\psi_1, \psi_2, \cdots, \psi_{d+1})$ and $\pi_2 = \left(\psi_{\sigma(1)}, \psi_{\sigma(2)}, \cdots, \psi_{\sigma(d+1)}\right)$. Suppose $\psi_k = \pi_1[i_{k-1} + 1 : i_k]$ for $1 \le k \le d+1$, where $0 = i_0 < i_1 \cdots < i_d < i_{d+1} = N$. Then $(\pi_1(i), \pi_1(i+1)) \in (A(\pi_1) \setminus A(\pi_2))$ if and only if $i \in \{i_1, \cdots, i_d\}$. Therefore, $|A(\pi_1) \setminus A(\pi_2)| = |\{i_1, \cdots, i_d\}| = d$. ■

Appendix B
Proof of Lemma 3

Lemma 3. *For all $\pi_1, \pi_2 \in \mathbb{S}_N$, the following inequality holds,*

$$w_B(\pi_1 \circ \pi_2) \le w_B(\pi_1) + w_B(\pi_2).$$

Proof:

For $\pi \in \mathbb{S}_N$, define $B(\pi)$ as follows,

$$B(\pi) \triangleq \{i | \pi(i+1) \ne \pi(i) + 1,\ 1 \le i < N\}.$$

Then, for all $i \in B(\pi)$, $(\pi(i), \pi(i+1)) \notin A(e)$. Therefore,

$$B(\pi) = \{i | (\pi(i), \pi(i+1)) \in (A(\pi) \setminus A(e)), 1 \le i < N\},$$

which indicates that

$$|B(\pi)| = |A(\pi) \setminus A(e)| = w_B(\pi). \tag{70}$$

Let $B_1 = B(\pi_1)$, $B_2 = B(\pi_2)$, $B_3 = B(\pi_1 \circ \pi_2)$. Then $\forall$ $i \in B_3$,

$$\pi_1(\pi_2(i+1)) \ne \pi_1(\pi_2(i)) + 1.$$

Therefore, i must satisfy at least one of the conditions below:

$$\begin{aligned} &\{\pi_2(i+1) \ne \pi_2(i) + 1\},\ \mathit{or} \\ &\{\pi_2(i) = k\ \mathit{and}\ \pi_1(k+1) \ne \pi_1(k) + 1\}. \end{aligned} \tag{71}$$

Equation (71) means that either $i \in B_2$ or $\pi_2(i) \in B_1$ is true for all $i \in B_3$. Then the function $f : (B_3 \setminus B_2) \to B_1$ specified by $f(i) \triangleq \pi_2(i)$ is an injection, which implies that

$$|B_3| = |B_3 \setminus B_2| + |B_3 \cap B_2| \le |B_1| + |B_2|. \tag{72}$$

Apply (70) to (72), we obtain the following inequality:

$$w_B(\pi_1 \circ \pi_2) \le w_B(\pi_1) + w_B(\pi_2).$$

■

Appendix C
Proof of Lemma 5

Lemma 5. *For all $N \in \mathbb{N}^*$, $t \le N - \sqrt{N} - 1$, $b_B(N,t)$ is bounded by the following inequality:*

$$\prod_{k=1}^{t} (N-k) \le b_B(N,t) \le \prod_{k=0}^{t} (N-k).$$

Proof: Denote the number of permutations of length N with block permutation weight m by $F(m)$, then $b_B(N,t) = \sum_{m=0}^{t} F(m)$.

We know that $F(0) = 1$, and from [18, equation (3)], for all $1 \le m \le t$,

$$F(m) = \binom{N-1}{m} m! \sum_{k=0}^{m} (-1)^{m-k} \frac{(k+1)}{(m-k)!}. \tag{73}$$

Let $a_k = \frac{(k+1)}{(m-k)!}$, $0 \le k \le m$, $1 \le m \le t$. Then, $m + 1 = a_m > a_{m-1} = m > a_{m-2} > \cdots > a_0 > 0$. Therefore, the following inequalities hold true,

$$a_{2k} - a_{2k-1} + \cdots + a_0 = a_0 + \sum_{i=1}^{k} (a_{2i} - a_{2i-1}) > 0,$$

$$a_{2k-1} - a_{2k-2} + \cdots - a_0 = \sum_{i=1}^{k} (a_{2i-1} - a_{2i-2}) > 0.$$

For $1 \le m \le t$, define A_m as follows,

$$A_m = \sum_{k=0}^{m} (-1)^{m-k} \frac{(k+1)}{(m-k)!}.$$

Then, $A_1 = 1$ and for $2 \le m \le t$,

$$\begin{aligned} A_m &= m + 1 - (a_{m-1} - a_{m-2} + \cdots + (-1)^{m-1} a_0) < m + 1, \\ A_m &= m + 1 - m + (a_{m-2} - a_{m-3} + \cdots + (-1)^m a_0) > 1. \end{aligned} \tag{74}$$

According to (73) and (74), for all $1 \le m \le t$,

$$\binom{N-1}{m} m! \le F(m) < \binom{N-1}{m} (m+1)!.$$

To derive the upper bound of the ballsize $b_B(N,t)$, we first find an upper bound of $F(m)$, $1 \le m \le t$, as follows,

$$F(m) \le \binom{N-1}{m} (m+1)! = (m+1) \cdot \prod_{k=1}^{m} (N-k).$$

For $t \le N - \sqrt{N} - 1$, it follows that $i \le N - \sqrt{N} - 1$ for all $1 \le i \le t$. Therefore, for all $1 \le i \le t$,

$$(N - i - 1)^2 \ge (N - (N - \sqrt{N}))^2 = N > i + 1.$$

Then,

$$\begin{aligned}
b_B(N,t) &= \sum_{i=0}^{t} F(i) \\
&\leq 1 + \sum_{i=1}^{t} (i+1) \cdot \prod_{k=1}^{i} (N-k) \\
&= 1 + \sum_{i=1}^{t} (N-(N-i-1)) \cdot \prod_{k=1}^{i} (N-k) \\
&= 1 + \sum_{i=1}^{t} \left(\prod_{k=0}^{i} (N-k) - \prod_{k=1}^{i+1} (N-k) \right) \\
&= \prod_{k=0}^{t} (N-k) - \sum_{i=2}^{t} \left(\prod_{k=1}^{i+1} (N-k) - \prod_{k=0}^{i-1} (N-k) \right) \\
&\quad - (N-1)(N-2) + 1 \\
&= \prod_{k=0}^{t} (N-k) - \sum_{i=2}^{t} \left(\prod_{k=1}^{i-1} (N-k) \right) \\
&\quad ((N-i)(N-i-1) - N) - ((N-1)(N-2) - 1) \\
&= \prod_{k=0}^{t} (N-k) - \sum_{i=2}^{t} \left(\prod_{k=1}^{i-1} (N-k) \right) \\
&\quad \left((N-i-1)^2 - i - 1\right) - ((N-1)(N-2) - 1) \\
&\leq \prod_{k=0}^{t} (N-k).
\end{aligned}$$

Similarly, for the lower bound, the following inequality holds true.

$$b_B(N,t) = \sum_{i=0}^{t} F(i) \geq 1 + \sum_{i=1}^{t} \prod_{k=1}^{i} (N-k) > \prod_{k=1}^{t} (N-k).$$

The lemma is proved. ■

Appendix D
Proof of Lemma 6

Lemma 6. *For all $N \in \mathbb{N}^*$, $t \leq \min\{N - \sqrt{N} - 1, \frac{N-1}{4}\}$, $b_G(N,t)$ is bounded as follows:*

$$\prod_{k=1}^{t} (N-k) \leq b_G(N,t) \leq \prod_{k=0}^{4t} (N-k).$$

Proof: The upper bound is obtained from replacing t by $4t$ in (17) and utilizing (14). Note that $\pi \in B_G(N,t,e)$ implies that $d_G(\pi,e) \leq t$. Then from (14), $d_B(\pi,e) \leq 4 d_G(\pi,e) \leq 4t$ holds true, which means that $\pi \in B_B(N,4t,e)$. Therefore, $B_G(N,t,e) \subseteq B_B(N,4t,e)$, which implies that $b_G(N,t) \leq b_B(N,4t)$. From (17) we will get the upper bound.

Similarly, (14) also implies that $B_B(N,t,e) \subseteq B_G(N,t,e)$, which means that $b_B(N,t) \leq b_G(N,t)$. From (17) the lower bound follows immediately. The lemma is proved. ■

Appendix E
Proof of Lemma 8

Lemma 8. *For all π_1, $\pi_2 \in \mathbb{S}_N$ such that $\pi_1 \neq \pi_2$, if $\alpha^{(q,d)}(\pi_1) = \alpha^{(q,d)}(\pi_2)$, then,*

$$|\nu(\pi_1) \Delta \nu(\pi_2)| > 2d.$$

Proof: Let $B_1 = \nu(\pi_1)$, $B_2 = \nu(\pi_2)$. We prove the statement by contradiction. If the lemma is not true, i.e., $|B_1 \Delta B_2| \leq 2d$, then $k = |D_1| = |D_2| \leq d$, where $D_1 = B_1 \setminus B_2$, $D_2 = B_2 \setminus B_1$. Suppose $D_1 = \{x_1, x_2, \cdots, x_k\}$, $D_2 = \{x_{k+1}, x_{k+2}, \cdots, x_{2k}\}$. Then, $\alpha^{(q,d)}(\pi_1) = \alpha^{(q,d)}(\pi_2)$ is equivalent to the following equations.

$$\left\{ \begin{array}{rcl}
x_1 + \cdots + x_k &=& x_{k+1} + \cdots + x_{2k}, \\
x_1^2 + \cdots + x_k^2 &=& x_{k+1}^2 + \cdots + x_{2k}^2, \\
&\vdots& \\
x_1^{2d-1} + \cdots + x_k^{2d-1} &=& x_{k+1}^{2d-1} + \cdots + x_{2k}^{2d-1}.
\end{array} \right. \tag{75}$$

From (75), it follows that

$$\begin{pmatrix}
1 & 1 & \cdots & 1 \\
x_1 & x_2 & \cdots & x_{2k} \\
x_1^2 & x_2^2 & \cdots & x_{2k}^2 \\
\vdots & \vdots & \ddots & \vdots \\
x_1^{2d-1} & x_2^{2d-1} & \cdots & x_{2k}^{2d-1}
\end{pmatrix} \mathbf{y} = \mathbf{0},$$

where $\mathbf{y} = (y_1, y_2, \cdots, y_{2k})^T$, and

$$y_i = \begin{cases} 1, & 1 \leq i \leq k, \\ -1, & k < i \leq 2k. \end{cases}$$

Given that $2k \leq 2d$, the above equation implies that

$$\begin{pmatrix}
1 & 1 & \cdots & 1 \\
x_1 & x_2 & \cdots & x_{2k} \\
x_1^2 & x_2^2 & \cdots & x_{2k}^2 \\
\vdots & \vdots & \ddots & \vdots \\
x_1^{2k-1} & x_2^{2k-1} & \cdots & x_{2k}^{2k-1}
\end{pmatrix} \mathbf{y} = \mathbf{0}. \tag{76}$$

Denote the Vandermonde matrix in equation (76) by $\mathbf{U}$. Then $\mathbf{y}$ is in the nullspace of $\mathbf{U}$. Therefore, $\mathbf{U}$ is singular, which implies that the determinant of $\mathbf{U}$ is equal to 0 in $\mathbb{F}_q$, i.e.,

$$0 = \det \mathbf{U} = \prod_{1 \leq i < j \leq 2k} (x_i - x_j). \tag{77}$$

As q is a divisor of 0, q should also be a divisor of the right hand side of equation (77), which implies that $\exists\ i \neq j \in [2k]$ such that $q | (x_i - x_j)$. Then $x_i = x_j$ on $\mathbb{F}_q$, and we must have $x_i \in D_1, x_j \in D_2$ or $x_i \in D_2, x_j \in D_1$, which implies that $x_i, x_j \in D_1 \cap D_2$, a contradiction. ■

Appendix F
Proof of Lemma 9

Lemma 9. *Suppose $\mathbf{A} \in \mathbb{F}_q^{(4t-1)\times(2t)}$, $\mathbf{b} \in \mathbb{F}_q^{4t-1}$ are defined in (43) and (44), respectively. Consider the following equation defined on $\mathbb{F}_q$:*

$$\mathbf{A}\mathbf{c} = \mathbf{b}.$$

For any vector $\mathbf{c} \in \mathbb{F}_q^{2t}$, $\mathbf{c}$ is a nonzero solution to (45) if and only if $(h_1(\mathbf{c}), h_2(\mathbf{c}))$ is a nonzero solution to (40).

Proof: Suppose

$$\begin{aligned}
f_1 &= X^{N-1} + a_1 X^{N-2} + \cdots + a_{4t-1} X^{N-4t} + g_1, \\
f_2 &= X^{N-1} + a_1' X^{N-2} + \cdots + a_{4t-1}' X^{N-4t} + g_2.
\end{aligned} \tag{78}$$

Additionally, suppose

$$\begin{aligned} h_1 \cdot f_1 &= X^{N+t-1} + s_{N+t-2}X^{N+t-2} + \cdots + s_0, \\ h_2 \cdot f_2 &= X^{N+t-1} + s'_{N+t-2}X^{N+t-2} + \cdots + s'_0. \end{aligned}$$

Then, from (78) and (42), it follows that

$$\begin{cases} s_{N+t-2} = a_1 + c_1, \\ s_{N+t-3} = a_2 + c_1 a_1 + c_2, \\ \vdots \\ s_{N-1} = a_t + c_1 a_{t-1} + \cdots + c_t, \\ \vdots \\ s_{N-3t} = a_{4t-1} + c_1 a_{4t-2} + c_2 a_{4t-3} + \cdots + c_t a_{3t-1}. \end{cases}$$

Similarly, we also have

$$\begin{cases} s'_{N+t-2} = a'_1 + c'_1, \\ s'_{N+t-3} = a'_2 + c'_1 a'_1 + c'_2, \\ \vdots \\ s'_{N-1} = a'_t + c'_1 a'_{t-1} + \cdots + c'_t, \\ \vdots \\ s'_{N-3t} = a'_{4t-1} + c'_1 a'_{4t-2} + c'_2 a'_{4t-3} + \cdots + c'_t a'_{3t-1}. \end{cases}$$

Then (40) is true iff $s_i = s'_i$ for all $N - 3t \leq i \leq N + t - 2$, which is equivalent to the following equation:

$$\begin{pmatrix} 1 & & & \\ a_1 & 1 & & \\ \vdots & \vdots & \ddots & \\ a_{t-1} & a_{t-2} & \cdots & 1 \\ \vdots & \vdots & \ddots & \vdots \\ a_{4t-2} & a_{4t-3} & \cdots & a_{3t-1} \end{pmatrix} \begin{pmatrix} c_1 \\ c_2 \\ \vdots \\ c_t \end{pmatrix} + \begin{pmatrix} a_1 \\ a_2 \\ \vdots \\ a_{4t-1} \end{pmatrix} = \\ \begin{pmatrix} 1 & & & \\ a'_1 & 1 & & \\ \vdots & \vdots & \ddots & \\ a'_{t-1} & a'_{t-2} & \cdots & 1 \\ \vdots & \vdots & \ddots & \vdots \\ a'_{4t-2} & a'_{4t-3} & \cdots & a'_{3t-1} \end{pmatrix} \begin{pmatrix} c'_1 \\ c'_2 \\ \vdots \\ c'_t \end{pmatrix} + \begin{pmatrix} a'_1 \\ a'_2 \\ \vdots \\ a'_{4t-1} \end{pmatrix}. \tag{79}$$

We note that (79) is equivalent to (45). ■

Appendix G
Proof of Lemma 10

Lemma 10. *Let $\pi_1, \pi_2 \in \mathbb{S}_N$, $s_1, s_2 \in [N]$. For any two extensions $E(\pi_1, s_1)$ and $E(\pi_2, s_2)$, if s_1 is a jump point of $E(\pi_1, s_1)$ with respect to $E(\pi_2, s_2)$, then*

$$d_B(E(\pi_1, s_1), E(\pi_2, s_2)) > d_B(\pi_1, \pi_2),$$

else

$$d_B(E(\pi_1, s_1), E(\pi_2, s_2)) = d_B(\pi_1, \pi_2).$$

Proof: Let $\sigma_1 = E(\pi_1, s_1)$ and $\sigma_2 = E(\pi_2, s_2)$. Recall the notion of *characteristic sets* in Definition 3. Suppose $A(\pi_1)$, $A(\pi_2)$, $A(\sigma_1)$, $A(\sigma_2)$ are the characteristic sets of π_1, π_2, σ_1, σ_2, respectively. According to Lemma 2,

$$\begin{aligned} d_B(\pi_1, \pi_2) &= |A(\pi_1) \setminus A(\pi_2)|, \\ d_B(\sigma_1, \sigma_2) &= |A(\sigma_1) \setminus A(\sigma_2)|. \end{aligned} \tag{80}$$

Let $k_1 = \pi_1^{-1}(s_1)$, $k_2 = \pi_2^{-1}(s_2)$, then $\pi_1(k_1) = s_1$ and $\pi_2(k_2) = s_2$. If $1 \leq k_1, k_2 < N$, let $\pi_1(k_1 + 1) = j_1$ and $\pi_2(k_2 + 1) = j_2$.

Suppose first s_1 is a jump point, then consider the following cases.

1) $s_1 \neq s_2$ and either $k_1 = N$ or $k_2 = N$.
   a) $k_1 = k_2 = N$. In this case, $A(\sigma_1) = A(\pi_1) \cup \{(s_1, N + 1)\}$, $A(\sigma_2) = A(\pi_2) \cup \{(s_2, N + 1)\}$. Therefore, $A(\sigma_1) \setminus A(\sigma_2) = (A(\pi_1) \setminus A(\pi_2)) \cup \{(s_1, N + 1)\}$. From (80), $d_B(\sigma_1, \sigma_2) = d_B(\pi_1, \pi_2) + 1$ follows.
   b) $k_1 = N \neq k_2$. In this case, $A(\sigma_1) = A(\pi_1) \cup \{(s_1, N + 1)\}$, $A(\sigma_2) = (A(\pi_2) \setminus \{(s_2, j_2)\}) \cup \{(s_2, N + 1), (N + 1, j_2)\}$. Therefore, $A(\sigma_1) \setminus A(\sigma_2) = (A(\pi_1) \setminus (A(\pi_2) \setminus \{(s_2, j_2)\})) \cup \{(s_1, N + 1)\}$, which means $((A(\pi_1) \setminus A(\pi_2)) \cup \{(s_1, N + 1)\}) \subseteq (A(\sigma_1) \setminus A(\sigma_2))$. From (80), it follows that $d_B(\sigma_1, \sigma_2) \geq d_B(\pi_1, \pi_2) + 1$.
   c) $k_2 = N \neq k_1$. Following the same logic in the previous case, $d_B(\sigma_1, \sigma_2) \geq d_B(\pi_1, \pi_2) + 1$ holds true.
2) $s_1 \neq s_2$, $k_1, k_2 \neq N$. Since s_1 is a jump point, $j_1 \neq j_2$.
   a) In this case, $A(\sigma_1) = (A(\pi_1) \setminus \{(s_1, j_1)\}) \cup \{(s_1, N + 1), (N + 1, j_1)\}$, $A(\sigma_2) = (A(\pi_2) \setminus \{(s_2, j_2)\}) \cup \{(s_2, N + 1), (N + 1, j_2)\}$. Therefore, the equation $(((A(\pi_1) \setminus A(p_2)) \setminus \{s_1, j_1\}) \cup \{(s_1, N + 1), (N + 1, j_1)\}) \subseteq (A(\sigma_1) \setminus A(\sigma_2))$ follows. From (80), $d_B(\sigma_1, \sigma_2) \geq d_B(\pi_1, \pi_2) + 1$.

If s_1 is not a jump point, then consider the following cases.

1) $s_1 = s_2$ and either $k_1 = N$ or $k_2 = N$.
   a) $k_1 = k_2 = N$. In this case, $A(\sigma_1) = A(\pi_1) \cup \{(s_1, N + 1)\}$, $A(\sigma_2) = A(\pi_2) \cup \{(s_1, N + 1)\}$. Therefore, $A(\sigma_1) \setminus A(\sigma_2) = A(\pi_1) \setminus A(\pi_2)$. From (80), $d_B(\sigma_1, \sigma_2) = d_B(\pi_1, \pi_2)$ follows.
   b) $k_1 = N \neq k_2$. In this case, $A(\sigma_1) = A(\pi_1) \cup \{(s_1, N + 1)\}$, $A(\sigma_2) = (A(\pi_2) \setminus \{(s_1, j_2)\}) \cup \{(s_1, N + 1), (N + 1, j_2)\}$. Therefore, $A(\sigma_1) \setminus A(\sigma_2) = A(\pi_1) \setminus (A(\pi_2) \setminus \{(s_1, j_2)\}) = A(\pi_1) \setminus A(\pi_2)$. From (80), it follows that $d_B(\sigma_1, \sigma_2) = d_B(\pi_1, \pi_2)$.
   c) $k_2 = N \neq k_1$. Follow the same logic in the previous case, $d_B(\sigma_1, \sigma_2) = d_B(\pi_1, \pi_2)$ holds true.
2) $k_1, k_2 \neq N$. Since s_1 is not a jump point, either $s_1 = s_2$ or $j_1 = j_2$ must be satisfied.
   a) $s_1 = s_2$ and $j_1 = j_2$. In this case, $A(\sigma_1) = (A(\pi_1) \setminus \{(s_1, j_1)\}) \cup \{(s_1, N + 1), (N + 1, j_1)\}$, $A(\sigma_2) = (A(\pi_2) \setminus \{(s_1, j_1)\}) \cup \{(s_1, N + 1), (N + 1, j_1)\}$. Therefore, $A(\sigma_1) \setminus A(\sigma_2) = A(\pi_1) \setminus A(\pi_2)$. From (80), $d_B(\sigma_1, \sigma_2) = d_B(\pi_1, \pi_2)$ follows.
   b) $s_1 = s_2$ and $j_1 \neq j_2$. In this case, $A(\sigma_1) = (A(\pi_1) \setminus \{(s_1, j_1)\}) \cup \{(s_1, N + 1), (N + 1, j_1)\}$, $A(\sigma_2) = (A(\pi_2) \setminus \{(s_1, j_2)\}) \cup \{(s_1, N + 1), (N + 1, j_2)\}$. Therefore, $A(\sigma_1) \setminus A(\sigma_2) =$

$((A(\pi_1) \setminus A(\pi_2)) \setminus \{(s_1, j_1)\}) \cup \{(N+1, j_1)\}$. From (80), it follows that $d_B(\sigma_1, \sigma_2) = d_B(\pi_1, \pi_2)$.

c) $s_1 \neq s_2$ and $j_1 = j_2$. Follow the same logic as indicated in the previous case, $d_B(\sigma_1, \sigma_2) = d_B(\pi_1, \pi_2)$ holds true.

The lemma is proved. ■

Appendix H
Proof of Lemma 11

Lemma 11. *Let $\pi_1, \pi_2 \in \mathbb{S}_N$, $s_1, s_2 \in [N]$. For any extensions $E(\pi_1, S_1)$, $E(\pi_2, S_2)$ of π_1, π_2 on extension sequences S_1, S_2, respectively, it follows that*

$$d_B(E(\pi_1, S_1), E(\pi_2, S_2)) \geq |H(S_1, S_2)|.$$

Proof: For all $i \in H(S_1, S_2)$, let

$$m(i) = \min\{m: \ s_{1,m} = i, \ s_{2,m} \neq i\}. \tag{81}$$

Suppose $J_{1,m(i)-1} = (s_{1,1}, s_{1,2}, \cdots, s_{1,m(i)-1})$, $J_{2,m(i)-1} = (s_{2,1}, s_{2,2}, \cdots, s_{2,m(i)-1})$. Let $\sigma_1^{m(i)-1} = E(\pi_1, J_{1,m(i)-1})$ and $\sigma_2^{m(i)-1} = E(\pi_2, J_{2,m(i)-1})$. Recall the definition of the *jump set* $F(\pi_1, \pi_2, S_1, S_2)$ in Definition 7. Consider the following two cases:

1) If $m(i) \in F(\pi_1, \pi_2, S_1, S_2)$, then $s_{1,m(i)} = i$ is a jump point of $E(\sigma_1^{m(i)-1}, s_{1,m(i)})$ with respect to $E(\sigma_2^{m(i)-1}, s_{2,m(i)})$.
2) If $m(i) \notin F(\pi_1, \pi_2, I_1, I_2)$, then i is not a jump point of $E(\sigma_1^{m(i)-1}, s_{1,m(i)})$ with respect to $E(\sigma_2^{m(i)-1}, s_{2,m(i)})$. Let $k_1' = (\sigma_1^{m(i)-1})^{-1}(s_{1,m(i)})$, $k_1 = \pi_1^{-1}(s_{1,m(i)})$, $k_2' = (\sigma_2^{m(i)-1})^{-1}(s_{2,m(i)})$, $k_2 = \pi_2^{-1}(s_{2,m(i)})$, then $\sigma_1^{m(i)-1}(k_1') = \pi_1(k_1) = s_{1,m(i)}$ and $\sigma_2^{m(i)-1}(k_2') = \pi_2(k_2) = s_{2,m(i)}$. Given that $s_{1,m(i)}$ is not a jump point and $s_{1,m(i)} \neq s_{2,m(i)}$, it follows from Definition 6 that k_1, $k_2 \neq N+m(i)-1$ and $\sigma_1^{m(i)-1}(k_1'+1) = \sigma_2^{m(i)-1}(k_2'+1)$ must be true. Let $j = \sigma_1^{m(i)-1}(k_1'+1) = \sigma_2^{m(i)-1}(k_2'+1)$. From (81), $\pi_1(k_1+1) = \pi_2(k_2+1) = j \in [N]$ holds, otherwise $N < j < N + m(i)$ is inserted after i in π_1 and is not inserted after i in π_2, a contradiction. Then $(i, j) \in A(\pi_1)$, $(s_{2,m(i)}, j) \in A(\pi_2)$ and $s_{2,m(i)} \neq i$. Therefore $(i, j) \in (A(\pi_1) \setminus A(\pi_2))$.

Suppose $J = \{i | m(i) \notin F(\pi_1, \pi_2, S_1, S_2), i \in H(S_1, S_2)\}$, then from the above discussion:

$$|F(\pi_1, \pi_2, S_1, S_2)| \geq |H(S_1, S_2) \setminus J|,$$
$$d_B(\pi_1, \pi_2) = |A(\pi_1) \setminus A(\pi_2)| \geq |J|.$$

And from Lemma 10, it follows that

$$\begin{aligned} d_B(E(\pi_1, S_1), E(\pi_2, S_2)) \geq & d_B(\pi_1, \pi_2) + |F(\pi_1, \pi_2, S_1, S_2)| \\ \geq & |H(S_1, S_2) \setminus J| + |J| \\ \geq & |H(S_1, S_2)|. \end{aligned}$$

The lemma is proved. ■

Appendix I
Proof of Lemma 13

Lemma 13. *For all $k, N \in \mathbb{N}^*$, $k > 3$, $N > k^2$, consider an arbitrary subset $Y \subset [k]$, where $|Y| = M < k$, $Y = \{i_1, i_2, \cdots, i_M\}$, then*

$$\operatorname{LCM}(N + i_1, N + i_2, \cdots, N + i_M) > N^{M - \frac{k}{2}}.$$

Proof: For all $r, n \in \mathbb{N}^*$, it follows from [19, equation (13)] that

$$g_r(n) = \operatorname{GCD}(r!, (n+r) g_{r-1}(n)), \tag{82}$$

where for all $r \in \mathbb{N}$, $n \in \mathbb{N}^*$,

$$g_r(n) = \frac{n(n+1)\cdots(n+r)}{\operatorname{LCM}(n, n+1, \cdots, n+r)}. \tag{83}$$

From (82) and (83), the following statement holds true,

$$g_r(n) | r!, \ \forall r, n \in \mathbb{N}^*, \tag{84}$$

which implies that

$$\frac{n(n+1)\cdots(n+r)}{\operatorname{LCM}(n, n+1, \cdots, n+r)} \leq r!. \tag{85}$$

Let $n = N+1$, $r = k-1$ in (85). Then, for all $N, k \in \mathbb{N}^*$,

$$\begin{aligned} & \operatorname{LCM}(N+1, N+2, \cdots, N+k) \\ \geq & \frac{(N+1)(N+2)\cdots(N+k)}{(k-1)!}. \end{aligned} \tag{86}$$

Let $[k] \setminus Y = \{j_1, j_2, \cdots, j_{k-M}\}$. Notice that

$$\begin{aligned} & \operatorname{LCM}(N+1, N+2, \cdots, N+k) \\ = & \operatorname{LCM}(\operatorname{LCM}(N + i_1, N + i_2, \cdots, N + i_M), \\ & \operatorname{LCM}(N + j_1, N + j_2, \cdots, N + j_{k-M})) \\ \leq & \left[\prod_{s=1}^{k-M} (N + j_s)\right] \operatorname{LCM}(N + i_1, N + i_2, \cdots, N + i_M). \end{aligned} \tag{87}$$

From equation (86) and (87),

$$\begin{aligned} & \operatorname{LCM}(N + i_1, N + i_2, \cdots, N + i_M) \\ \geq & \frac{\operatorname{LCM}(N+1, N+2, \cdots, N+k)}{\prod_{s=1}^{k-M} (N + j_s)} \\ \geq & \frac{(N+1)(N+2)\cdots(N+k)}{(k-1)! \prod_{s=1}^{k-M} (N + j_s)} = \frac{\prod_{s=1}^{M} (N + i_s)}{(k-1)!} > \frac{N^M}{k!}. \end{aligned}$$

From Lemma 7, for all $k > 3$ and $N > k^2$,

$$\frac{N^M}{k!} > \frac{N^M}{2^{(k+\frac{1}{2})\log k - k + 2}} > \frac{N^M 2^{k-2}}{k^{k+1}} \geq \frac{N^M}{k^k} > N^{M - \frac{k}{2}}.$$

The lemma is proved. ■

Acknowledgment

Research supported in part by NSF CAREER grant No.1150212 and NSF CIF grant No.1527130.

References

[1] S. Yang, C. Schoeny, and L. Dolecek, "Order-optimal permutation codes in the generalized Cayley metric," in *IEEE Information Theory Workshop*, Kaohsiung, Taiwan, Nov 2017, pp. 234–238.

[2] I. Dixon and G. Whittaker, Eds., *Storing your music in the iCloud*. Apress, 2015.

[3] R. Zeira and R. Shamir, "Sorting by cuts, joins and whole chromosome duplications," *Journal of Computational Biology*, vol. 24, pp. 127–137, 2017.

[4] K.-T. Chen, C.-L. Li, H.-T. Chiu, and C. L. Lu, "An efficient algorithm for one-sided block ordering problem under block-interchange distance," *Theoretical Computer Science*, vol. 609, pp. 296 – 305, 2016.

[5] Y. M. Chee and V. K. Vu, "Breakpoint analysis and permutation codes in generalized Kendall tau and Cayley metrics," in *Proc. IEEE Int. Symp. Inf. Theory*, Hawaii, USA, Jun. 2014, pp. 2959–2963.

[6] S. Buzaglo, E. Yaakobi, T. Etzion, and J. Bruck, "Systematic error-correcting codes for permutations and multi-permutations," *IEEE Trans. Inf. Theory*, vol. 62, no. 6, pp. 3113–3124, Jun. 2016.

[7] S. Buzaglo and T. Etzion, "Bounds on the size of permutation codes with the Kendall tau-metric," *IEEE Trans. Inf. Theory*, vol. 61, no. 6, pp. 3241–3250, Jun. 2015.

[8] Y. Zhang and G. Ge, "Snake-in-the-box codes for rank modulation under Kendall's τ-metric," *IEEE Trans. Inf. Theory*, vol. 62, no. 1, pp. 151–158, Jan. 2016.

[9] F. Farnoud and O. Milenkovic, "Multipermutation codes in the Ulam metric for nonvolatile memories," *IEEE J. Sel. Areas Commun.*, vol. 32, no. 5, pp. 919–932, May 2014.

[10] F. Farnoud, V. Skachek, and O. Milenkovic, "Error-correction in flash memories via codes in the Ulam metric," *IEEE Trans. Inf. Theory*, vol. 59, no. 5, pp. 3003–3020, May 2013.

[11] F. Göloğlu, J. Lember, A. E. Riet, and V. Skachek, "New bounds for permutation codes in Ulam metric," in *Proc. IEEE Int. Symp. Inf. Theory*, Hong Kong, Jun. 2015, pp. 1726–1730.

[12] R. Gabrys, E. Yaakobi, F. Farnoud, F. Sala, J. Bruck, and L. Dolecek, "Codes correcting erasures and deletions for rank modulation," *IEEE Trans. Inf. Theory*, vol. 62, no. 1, pp. 136–150, Jan. 2016.

[13] Y. M. Chee, V. K. Vu, and X. Zhang, "Permutation codes correcting a single burst deletion I: Unstable deletions," in *Proc. IEEE Int. Symp. Inf. Theory*, Hong Kong, Jun. 2015, pp. 1741–1745.

[14] H. Robbins, "A remark on Stirling's formula," *The American Mathematical Monthly*, vol. 62, no. 1, pp. 26–29, 1955. [Online]. Available: http://www.jstor.org/stable/2308012

[15] S. Ramanujan, "A proof of Bertrands postulate," *Journal of the Indian Mathematical Society*, vol. 11, no. 181-182, p. 27, 1919.

[16] L. Dolecek and V. Anantharam, "Repetition error correcting sets: Explicit constructions and prefixing methods," *SIAM J. Discrete Math.*, vol. 23, no. 4, pp. 2120–2146, Jan. 2010.

[17] D. Zeilberger, "A combinatorial proof of Newton's identities," *Discrete mathematics*, vol. 49, no. 3, p. 319, 1984.

[18] A. Myers, "Counting permutations by their rigid patterns," *J. Combin. Theory Ser. A*, vol. 99, no. 2, pp. 345–357, 2002.

[19] B. Farhi, "Nontrivial lower bounds for the least common multiple of some finite sequences of integers," *J. Number Theory*, vol. 125, no. 2, pp. 393–411, 2007.

Biographies

Siyi Yang (S'17) is a Ph.D. candidate in the Electrical and Computer Engineering department at the University of California, Los Angeles (UCLA). She received her B.S. degree in Electrical Engineering from the Tsinghua University, in 2016 and the M.S. degree in Electrical and Computer Engineering from the University of California, Los Angeles (UCLA) in 2018. Her research interests include design of error-correction codes for non-volatile memory and distributed storage.

Clayton Schoeny (S'09) is a data scientist working at Fair. He received his Ph.D. in the Electrical and Computer Engineering Department at the University of California, Los Angeles (UCLA) in 2018. He received his B.S. and M.S. degrees in Electrical Engineering from UCLA in 2012 and 2014, respectively. He is a recipient of the Henry Samueli Excellence in Teaching Award, the 2016 Qualcomm Innovation Fellowship, and the 2017 UCLA Dissertation Year Fellowship.

Lara Dolecek (S'05–M'10–SM'13) is a Full Professor with the Electrical and Computer Engineering Department and Mathematics Department (courtesy) at the University of California, Los Angeles (UCLA). She holds a B.S. (with honors), M.S., and Ph.D. degrees in Electrical Engineering and Computer Sciences, as well as an M.A. degree in Statistics, all from the University of California, Berkeley. She received the 2007 David J. Sakrison Memorial Prize for the most outstanding doctoral research in the Department of Electrical Engineering and Computer Sciences at UC Berkeley. Prior to joining UCLA, she was a postdoctoral researcher with the Laboratory for Information and Decision Systems at the Massachusetts Institute of Technology. She received IBM Faculty Award (2014), Northrop Grumman Excellence in Teaching Award (2013), Intel Early Career Faculty Award (2013), University of California Faculty Development Award (2013), Okawa Research Grant (2013), NSF CAREER Award (2012), and Hellman Fellowship Award (2011). With her research group and collaborators, she received numerous best paper awards. Her research interests span coding and information theory, graphical models, statistical methods, and algorithms, with applications to emerging systems for data storage and computing. She currently serves as an Associate Editor for IEEE Transactions on Communications. Prof. Dolecek has served as a consultant for a number of companies specializing in data communications and storage.