Systematic Codes for Rank Modulation

Sarit Buzaglo[*], **Eitan Yaakobi**[*], **Tuvi Etzion**[*], and **Jehoshua Bruck**[†]

[*]Computer Science Department, Technion – Israel Institute of Technology, Haifa 32000, Israel
[†]Electrical Engineering Department, California Institute of Technology, Pasadena, CA 91125, U.S.A
{*sarahb, yaakobi, etzion*}@cs.technion.ac.il, *bruck@caltech.edu*

Abstract—The goal of this paper is to construct systematic error-correcting codes for permutations and multi-permutations in the Kendall's τ-metric. These codes are important in new applications such as rank modulation for flash memories. The construction is based on error-correcting codes for multi-permutations and a partition of the set of permutations into error-correcting codes. For a given large enough number of information symbols k, and for any integer t, we present a construction for $(k + r, k)$ systematic t-error-correcting codes, for permutations from S_{k+r}, with less redundancy symbols than the number of redundancy symbols in the codes of the known constructions. In particular, for a given t and for sufficiently large k we can obtain $r = t+1$. The same construction is also applied to obtain related systematic error-correcting codes for multi-permutations.

I. Introduction

Flash memory is one of the most widely used non-volatile technology. In flash memories, cells usually represent multiple levels, which correspond to the amount of electrons trapped in each cell. Currently, one of the main challenges in flash memory cells is to program each cell exactly to its designated level. In order to overcome this difficulty, the novel framework of *rank modulation codes* was introduced in [8]. In this setup, the information is carried by the relative values between the cells rather than by their absolute levels. Thus, every group of cells induces a permutation, which is derived by the ranking of the level of each cell in the group. There are several works which study the correction of errors under the setup of permutations for the rank modulation scheme; see e.g. [1], [9], [14], [15], [17], [18]. In all these works t-error-correcting codes were considered for the set S_n, which consists of all permutations on n elements, endowed with either the Kendall's τ-metric or the infinity metric. Recently, to improve the number of rewrites, the model of rank modulation was extended such that multiple cells can share the same ranking [5], [6]. Thus, the cells no longer determine permutations but rather multi-permutations, which are also known as permutations with repetitions. Error-correcting codes for multi-permutations subject to the Kendall's τ-metric were presented in [12] and also studied in [2].

The main goal of this paper is to construct *systematic* error-correcting codes for permutations. This concept for permutations was proposed in [17], [18]. In a systematic code $\mathcal{C}$ for permutations in S_n we have $k!$ codewords. Each permutation of S_k (on a given set of specific k symbols) is a sub-permutation of exactly one codeword of $\mathcal{C}$. In this paper we improve on some of the results in [17], [18]. Our construction of systematic error-correcting codes

for permutations is based on two ingredients. The first is a partition of S_k into t-error-correcting codes. The second is a code $\mathcal{C}_r$ for multi-permutations from the multi-set $\{0^k, k + 1, \ldots, k + r\}$ with minimum Kendall's τ-distance $2t$, whose size is the number of parts in the partition. Each code from the partition of S_k will be substituted into a different codeword of $\mathcal{C}_r$. This construction will be generalized to systematic codes for multi-permutations.

The rest of this work is organized as follows. Our construction is heavily based on error-correcting codes for multi-permutations. Hence, in Section II we define the basic concepts for multi-permutations and the Kendall's τ-metric. In Section III we will review and amend some of the known constructions of error-correcting codes for permutations and multi-permutations, using the Kendall's τ-metric. These concepts and constructions will be used in Section IV to obtain our main construction of systematic error-correcting codes for permutations. We will also perform some analysis for the number of redundancy symbols of these codes. We extend this construction to form systematic error-correcting codes for multi-permutations in Section V. We conclude in Section VI.

II. Basic Concepts

We denote by $[n]$ the set of n integers $\{1, 2, \ldots, n\}$. For two integers a, b, $a < b$, we denote by $[a, b]$ the set of $b - a + 1$ integers $[a, a + 1, a + 2, \ldots, b]$. Let S_n be the set of all permutations on $[n]$, and let $S([a, b])$ be the set of all permutations on $[a, b]$. A more general concept is *multi-permutations*, which is also known as permutations with repetitions. A *multi-set* $\mathcal{M} = \{v_1^{m_1}, v_2^{m_2}, \cdots, v_\ell^{m_\ell}\}$ is a collection of the elements $\{v_1, v_2, \ldots, v_\ell\}$ in which v_i appears m_i times, for each i, $1 \leq i \leq \ell$. The elements of $\{v_1, v_2, \ldots, v_\ell\}$ are called *ranks* while for every i, $1 \leq i \leq \ell$, the positive integer m_i is called the *multiplicity* of the ith rank. If $m_1 = m_2 = \cdots = m_\ell = m$ then $\mathcal{M}$ is called a *balanced multi-set*. A multi-permutation on the multi-set $\mathcal{M}$ is an ordering of all the elements of $\mathcal{M}$. Note, that a permutation is a special case of a multi-permutation. We denote a multi-permutation σ of length n by $\sigma = [\sigma(1), \sigma(2), \ldots, \sigma(n)]$, $n = \sum_{i=1}^{\ell} m_i$. For example, if $\mathcal{M} = \{1^2, 2^3, 3\}$, then $\sigma = [1, 2, 2, 1, 3, 2]$ is a multi-permutation on $\mathcal{M}$. We denote by $S(\mathcal{M})$ the set of all multi-permutations on $\mathcal{M}$. The size of $S(\mathcal{M})$ is equal to $\frac{n!}{\prod_{i=1}^{\ell} m_i!}$.

Given a multi-permutation $\sigma = [\sigma(1), \sigma(2), \ldots, \sigma(n)]$ from $S(\mathcal{M})$, an *adjacent transposition* is an exchange

of two distinct adjacent elements $\sigma(i), \sigma(i + 1)$, in σ, for some $1 \leq i \leq n - 1$. The result of such an adjacent transposition is the multi-permutation $[\sigma(1), \ldots, \sigma(i - 1), \sigma(i + 1), \sigma(i), \sigma(i + 2), \ldots, \sigma(n)]$. The Kendall's τ-distance between two multi-permutations $\sigma, \pi \in S(\mathcal{M})$ denoted by $d_K(\sigma, \pi)$ is the minimum number of adjacent transpositions required to obtain the multi-permutation π from the multi-permutation σ.

Example 1. *If $\sigma = [1, 1, 2, 2]$ and $\pi = [2, 1, 2, 1]$, then $d_K(\sigma, \pi) = 3$, since at least three adjacent transpositions are required to change the multi-permutation σ to π: $[1, 1, 2, 2] \rightarrow [1, 2, 1, 2] \rightarrow [2, 1, 1, 2] \rightarrow [2, 1, 2, 1]$.*

The Kendall's τ-metric was originally defined for permutations [4], [10]. For two permutations $\sigma, \pi \in S_n$ it is known [9], [11] that $d_K(\sigma, \pi)$ can be expressed as

$$d_K(\sigma, \pi) = |\{(i, j) : \sigma^{-1}(i) < \sigma^{-1}(j), \ \pi^{-1}(i) > \pi^{-1}(j)\}|.$$

For a multi-permutation $\sigma \in S(\mathcal{M})$, where $\mathcal{M} = \{v_1^{m_1}, v_2^{m_2}, \ldots, v_\ell^{m_\ell}\}$, we distinguish between appearances of the same rank in σ, by their positions in σ. We consider the increasing order of these positions. By abuse of notation we sometimes write $\sigma(j) = v_{i,r}$ and $j = \sigma^{-1}(v_{i,r})$ to indicate that the rth appearance of v_i is in the jth position in σ. The computation of the Kendall's τ-distance between two permutations can be generalized to two multi-permutations $\sigma, \pi \in S(\mathcal{M})$ as follows

$$d_K(\sigma, \pi) = \left| \left\{ ((i, r), (j, s)) \ : \ \begin{array}{l} \sigma^{-1}(v_{i,r}) < \sigma^{-1}(v_{j,s}) \\ \pi^{-1}(v_{i,r}) > \pi^{-1}(v_{j,s}) \end{array} \right\} \right|.$$

Let $n_0 = 0$ and for $1 \leq i \leq \ell$ let $n_i = \sum_{j=1}^{i} m_j$, which implies that $n = n_\ell$. For each i, $1 \leq i \leq \ell$, let θ_i be a permutation on $[n_{i-1} + 1, n_i]$ and let $\theta = (\theta_1, \theta_2, \ldots, \theta_\ell)$. We define a mapping $T_\theta : S(\mathcal{M}) \rightarrow S_n$, such that for every $\sigma \in S(\mathcal{M})$, $T_\theta(\sigma)$ is the permutation in S_n obtained as follows. For each i, $1 \leq i \leq \ell$, the permutation θ_i is substituted, in the same order, in the m_i positions in which the rank v_i appears in σ. More precisely, if $\sigma(j) = v_{i,r}$ then $(T_\theta(\sigma))(j) = \theta_i(r) = \theta(r + n_{i-1})$. For example, if $\theta = (\theta_1, \theta_2, \theta_3)$, where $\theta_1 = [1, 3, 2]$, $\theta_2 = [4, 5]$, and $\theta_3 = [8, 7, 6]$, then $T_\theta([1, 2, 1, 3, 3, 1, 3, 2]) = [1, 4, 3, 8, 7, 2, 6, 5])$. The mappings T_θ are useful for the computation of the Kendall's τ-distance between two multi-permutations since it is reduced to the computation of the Kendall's τ-distance on the corresponding permutations.

Lemma 1. *For every two multi-permutations $\sigma, \pi \in S(\mathcal{M})$ and $\theta = (\theta_1, \theta_2, \ldots, \theta_\ell)$ we have*

$$d_K(\sigma, \pi) = d_K(T_\theta(\sigma), T_\theta(\pi)).$$

Example 2. *If $\sigma = [1, 1, 2, 2]$, $\pi = [2, 1, 2, 1]$, $\theta = (\theta_1, \theta_2)$, where $\theta_1 = [2, 1]$ and $\theta_2 = [3, 4]$, then $d_K(\sigma, \pi) = 3$, and $d_K(T_\theta([1, 1, 2, 2]), T_\theta([2, 1, 2, 1])) = d_K([2, 1, 3, 4], [3, 2, 4, 1]) = 3$.*

Lemma 2. *Let $\sigma, \pi \in S(\mathcal{M})$ and let $\theta = (\theta_1, \theta_2, \ldots, \theta_\ell)$, $\eta = (\eta_1, \eta_2, \ldots, \eta_\ell)$, where $\theta_i, \eta_i \in S([n_{i-1} + 1, n_i])$, for each i, $1 \leq i \leq \ell$. Then*

$$d_K(T_\theta(\sigma), T_\eta(\pi)) \geq d_K(\sigma, \pi) + \sum_{i=1}^{\ell} d_K(\theta_i, \eta_i).$$

Another simple and important property of the Kendall's τ-metric on multi-permutations is presented in the following lemma.

Lemma 3. *If σ, π, and ρ, are three multi-permutations in $S(\mathcal{M})$, then $d_K(\sigma, \pi) + d_K(\pi, \rho) \equiv d_K(\sigma, \rho) \ (mod \ 2)$.*

III. ERROR-CORRECTING CODES

For the construction of systematic error-correcting codes for permutations and multi-permutations given in Sections IV and V we need general error-correcting codes for multi-permutations. In this section we discuss the constructions for such error-correcting codes for multi-permutations with the Kendall's τ-distance.

Such a construction was given in [12]. It is based on a metric embedding (mapping) of $S(\mathcal{M})$, where $\mathcal{M}$ is a balanced multi-set, into the metric space $\mathbb{Z}^{n-m}$, where m is the multiplicity of the ranks. The *Manhattan distance* (also called the L_1-distance) is used in $\mathbb{Z}^{n-m}$. This construction is a generalization of the constructions in [1], [9] for error-correcting codes for permutations.

Let $\mathbf{x}, \mathbf{y} \in \mathbb{Z}^N$, $\mathbf{x} = (x_1, x_2, \ldots, x_N)$, $\mathbf{y} = (y_1, y_2, \ldots, y_N)$. The Manhattan distance $d_M(\mathbf{x}, \mathbf{y})$ is defined by

$$d_M(\mathbf{x}, \mathbf{y}) \stackrel{\text{def}}{=} \sum_{i=1}^{N} |x_i - y_i|.$$

This metric embedding (mapping) is injective and for every two multi-permutations σ and π in $S(\mathcal{M})$, $d_K(\sigma, \pi)$ is greater or equal to the Manhattan distance between their images in $\mathbb{Z}^{n-m}$. These properties allow to construct error-correcting codes in $S(\mathcal{M})$ from error-correcting codes in the Manhattan metric over $\mathbb{Z}^{n-m}$.

We present a slightly modified version of this mapping. It will be defined on $S(\mathcal{M})$, where $\mathcal{M}$ is any multi-set, not necessarily a balanced multi-set. We will also restrict its range to its image, in order to obtain a bijective mapping. This is important for encoding purpose. We will show an encoding of $S(\mathcal{M})$, based on the enumerative encoding algorithm of Cover [3] in the full version of this paper.

A vector $\mathbf{x} = (x_1, x_2, \ldots, x_k) \in \mathbb{Z}^k$ is *monotone* if $x_1 \geq x_2 \geq \ldots \geq x_k$. For a set S of integers let $[S]^k$ be the set of all monotone vectors of length k over S. Let

$$[\mathbb{Z}]^{\mathcal{M}} \stackrel{\text{def}}{=} [\mathbb{Z}_{n_1+1}]^{m_2} \times [\mathbb{Z}_{n_2+1}]^{m_3} \times \ldots \times [\mathbb{Z}_{n_{\ell-1}+1}]^{m_\ell}.$$

The mapping $\psi : S(\mathcal{M}) \rightarrow [\mathbb{Z}]^{\mathcal{M}}$ is defined as follows. For every $\sigma \in S(\mathcal{M})$, $\psi(\sigma)$ is the vector $\mathbf{x} \in [\mathbb{Z}]^{\mathcal{M}}$, $\mathbf{x} = (\mathbf{x}_2, \mathbf{x}_3, \ldots, \mathbf{x}_\ell)$, where for each i, $2 \leq i \leq \ell$, $\mathbf{x}_i = (x_{i,1}, x_{i,2}, \ldots, x_{i,m_i})$, and for each s, $1 \leq s \leq m_i$,

$$x_{i,s} \stackrel{\text{def}}{=} |\{j \ : \ j > \sigma^{-1}(v_{i,s}) \wedge \sigma(j) < i\}|.$$

Namely, $x_{i,s}$ counts the number of ranks smaller than i which appear to the right of the sth appearance of i. For example, if $\sigma = [2,1,3,4,3,2,1,4]$ then $\psi(\sigma) = (\mathbf{x}_2, \mathbf{x}_3, \mathbf{x}_4) = ((2,1),(2,2),(3,0))$.

Lemma 4. *The mapping ψ is bijective.*

Lemma 5. *For any two multi-permutations $\sigma, \pi \in S(\mathcal{M})$ we have*

$$d_M(\psi(\sigma), \psi(\pi)) \leq d_K(\sigma, \pi).$$

Let $\mathbb{Z}_q^N$ be the set of all vectors of length N over the alphabet $\mathbb{Z}_q$. For every two vectors $\mathbf{x}, \mathbf{y} \in \mathbb{Z}_q^N$, the Lee distance $d_L(\mathbf{x}, \mathbf{y})$ is defined by

$$d_L(\mathbf{x}, \mathbf{y}) \overset{\text{def}}{=} \sum_{i=1}^N \min\{|x_i - y_i|, q - |x_i - y_i|\}.$$

Clearly, $d_M(\mathbf{x}, \mathbf{y}) \geq d_L(\mathbf{x}, \mathbf{y})$ for all $\mathbf{x}, \mathbf{y} \in \mathbb{Z}_q^N$. The set $[\mathbb{Z}]^{\mathcal{M}}$ is a subset of $\mathbb{Z}_q^{n-m_1}$, where $q > n_{\ell-1}$. Hence, $d_L(\psi(\sigma), \psi(\pi)) \leq d_K(\sigma, \pi)$ for every two multi-permutations $\sigma, \pi \in S(\mathcal{M})$. We are now in a position to present a construction which transfers codes with the Lee metric to codes with the Kendall's τ-metric. The related theorem is a slight generalization of the result in [12]. This construction will be a major component in our main construction of systematic codes, which is the primary goal of this paper.

Theorem 1. *If there exists a code $\mathcal{C}_L \subseteq \mathbb{Z}_q^{n-m_1}$, $q > n_{\ell-1}$, with minimum Lee distance d, then there exists a code $\mathcal{C}_K \subseteq S(\mathcal{M})$ with minimum Kendall's τ-distance at least d and of size $|\mathcal{C}_K| = |\mathcal{C}_L \cap [\mathbb{Z}]^{\mathcal{M}}|$.*

By Theorem 1, error-correcting codes in $S(\mathcal{M})$ with the Kendall's τ-metric can be constructed from error-correcting codes over $\mathbb{Z}_q^{n-m_1}$ in the Lee metric. Next, we present some of the known constructions of error-correcting codes in the Lee metric and use Theorem 1 to obtain error-correcting codes in $S(\mathcal{M})$ and to estimate the size of these codes. First, we consider single-error-correcting codes in the Lee metric. Golomb and Welch [7] presented the following construction of a perfect linear single-error-correcting code in the Lee metric.

Theorem 2. *For every positive integer N, the code*

$$\mathcal{C}_L = \left\{ \mathbf{x} \in \mathbb{Z}_{2N+1}^N \;:\; \sum_{i=1}^N i \cdot x_i \equiv 0 \; (mod\ 2N+1) \right\}$$

is a perfect linear single-error-correcting code in $\mathbb{Z}_{2N+1}^N$ with the Lee metric.

The construction in Theorem 2 was used in [9] to construct single-error-correcting codes for permutations with the Kendall's τ-distance. Combining this construction with Theorem 1 implies the following corollary.

Corollary 1. *There exists a single-error-correcting code $\mathcal{C}_K \subset S(\mathcal{M})$ of size $|\mathcal{C}_K| \geq \frac{|S(\mathcal{M})|}{2(n-m_1)+1}$.*

The following construction was first proposed by Varshamov and Tenengolts [16] (see also [1]) for codes which correct a single asymmetric error. Let $||\mathbf{x}||$ denote the Manhattan weight of $\mathbf{x}$.

Theorem 3. *Let $q \geq N$ and let $h_1, h_2, \ldots, h_N$ be integers, $0 < h_i < q$ for all $1 \leq i \leq N$. Assume that for every $\mathbf{e} \in \mathbb{Z}^N$ with $||\mathbf{e}|| \leq t$, the sums $\sum_{i=1}^N e_i \cdot h_i$ are all distinct modulo q. Then the code*

$$C = \left\{ \mathbf{x} \in \mathbb{Z}_q^N \mid \sum_{i=1}^N x_i \cdot h_i \equiv 0 \; (mod\ q) \right\}$$

is a linear t-error-correcting code in $\mathbb{Z}_q^N$ with the Lee metric.

In order to use the construction in Theorem 3 we need the following theorem of Barg and Mazumdar [1].

Theorem 4. *Let q be a power of a prime and $M = (q^{t+1} - 1)/(q-1)$. Let*

$$M_t = \begin{cases} t(t+1)M, & t \text{ is odd} \\ t(t+2)M, & t \text{ is even} \end{cases}$$

Then there exist integers $h_1, h_2, \ldots, h_{q+1}$ such that for all $\mathbf{e} \in \mathbb{Z}^{q+1}$, $||\mathbf{e}|| \leq t$, the sums $\sum_{i=1}^{q+1} e_i h_i$ are all distinct modulo M_t.

The construction in Theorem 3 of a t-error-correcting code in the Lee metric, combined with Theorem 4, was used in [1] to construct t-error-correcting codes for permutations with the Kendall's τ-metric, and also used in [12] to construct t-error-correcting codes with the Kendall's τ-metric for multi-permutations over a balanced multi-set. Other constructions of codes with the Kendall's τ-distance that might useful in this context can be found in [13]. By combining the construction in Theorems 1, 3, and 4 we obtain the following Corollary.

Corollary 2. *Let $M = ((n-m_1-1)^{t+1}-1)/(n-m_1-2)$, where $n-m_1-1$ is a power of a prime. There exists a t-error-correcting code $\mathcal{C} \subset S(\mathcal{M})$ in the Kendall's τ-metric, whose size satisfies*

$$|\mathcal{C}| \geq \begin{cases} \frac{|S(\mathcal{M})|}{t(t+1)M}, & t \text{ is odd} \\ \frac{|S(\mathcal{M})|}{t(t+2)M}, & t \text{ is even} \end{cases}$$

Now, after presenting the concepts and ideas in constructions of error-correcting codes for multi-permutations, we are ready to present our main results on systematic error-correcting codes for permutations and multi-permutations in the next two sections.

IV. SYSTEMATIC ECC FOR PERMUTATIONS

In this section we present systematic t-error-correcting codes for permutations. Let k, n be integers such that $n \geq k \geq 1$. For a permutation $\alpha \in S_n$, we define $\alpha_{\downarrow k}$ to be the permutation obtained from α by deleting all the elements of $\{k+1, k+2, \ldots, n\}$ from α. We also

define $\alpha_{k \mapsto 0}$ to be the multi-permutation obtained from α by replacing in α every element of $\{1, 2, \ldots, k\}$ by 0. For example, if $\alpha = [2, 5, 4, 1, 3, 6]$ and $k = 3$ then $\alpha_{\downarrow k} = [2, 1, 3]$ and $\alpha_{k \mapsto 0} = [0, 5, 4, 0, 0, 6]$. In [17], the authors define systematic codes in the following way. A code $\mathcal{C} \subseteq S_n$ is an (n, k) *systematic* code if for every $\sigma \in S_k$ there exists exactly one $\alpha \in \mathcal{C}$ such that $\alpha_{\downarrow k} = \sigma$, which implies that $|\mathcal{C}| = k!$. The number of *redundancy symbols* of an (n, k) systematic code is $r = n - k$.

Let r be a positive integer and let $\mathcal{M}_{k,r} \overset{\text{def}}{=} \{0^k, k+1, k+2, \ldots, k+r\}$. For every permutation $\sigma \in S_k$ and multi-permutation $\rho \in S(\mathcal{M}_{k,r})$, we define the permutation $\sigma * \rho$ to be the permutation in S_{k+r} obtained from ρ by replacing the k zeros in ρ by the k elements of $\{1, 2, \ldots, k\}$, in the same order as in σ. For example, if $k = 4$, $r = 3$, $\rho = [0, 6, 0, 0, 5, 7, 0]$, and $\sigma = [2, 4, 1, 3]$, then $\sigma * \rho = [2, 6, 4, 1, 5, 7, 3]$.

Lemma 6. *For every $\rho \in S(\mathcal{M}_{k,r})$ and $\sigma \in S_k$ we have*

1) $(\sigma * \rho)_{\downarrow k} = \sigma$.
2) $(\sigma * \rho)_{k \mapsto 0} = \rho$.

By Lemma 2 we have.

Lemma 7. *Let $\sigma, \pi \in S_k$ and $\rho_1, \rho_2 \in S(\mathcal{M}_{k,r})$. Then*

$$d_K(\sigma * \rho_1, \pi * \rho_2) \geq d_K(\sigma, \pi) + d_K(\rho_1, \rho_2) .$$

We are now in a position to present our construction for systematic error-correcting codes for permutations.

Theorem 5. *Let $h_1, h_2, \ldots, h_{k-1}$, and M_t, be integers such that for every $\mathbf{e} \in \mathbb{Z}^{k-1}$ with $||\mathbf{e}|| \leq t$, the sums $\sum_{i=1}^{k-1} e_i h_i$ are all distinct modulo M_t. Assume further that there exists a code $\mathcal{C}_r \subset S(\mathcal{M}_{k,r})$ with minimum Kendall's τ-distance $2t$ and of size $|\mathcal{C}_r| \geq M_t$. Let $\rho_0, \rho_1, \ldots, \rho_{M_t-1}$ be distinct multi-permutations in $\mathcal{C}_r$. Let $\mathcal{C}$ be the code in S_{k+r} defined as follows.*

$$\mathcal{C} = \{\sigma * \rho_j \; : \; \sigma \in S_k, \; \sum_{i=1}^{k-1}(\psi(\sigma))_{i+1} h_i \equiv j \bmod M_t\}.$$

Then the code $\mathcal{C}$ is a $(k+r, k)$ systematic t-error-correcting code.

Example 3. *Let k be an integer, let $r = 2$, and let $M_1 = 2(k-1)+1$. As in Theorem 2, for every $\mathbf{e} \in \mathbb{Z}^{k-1}$, $||\mathbf{e}|| \leq 1$, the sums $\sum_{i=1}^{k-1} e_i i$ are all distinct modulo M_1. For the construction, we need a code in $S(\mathcal{M}_{k,2})$ with minimum distance 2 and of size at least M_1. To this end, fix a multi-permutation $\rho \in S(\mathcal{M}_{k,2})$ and consider the codes $\mathcal{C}_2^e = \{\gamma \in S(\mathcal{M}_{k,2}) \; : \; d_K(\rho, \gamma) \equiv 0 \pmod{2}\}$ and $\mathcal{C}_2^o = \{\gamma \in S(\mathcal{M}_{k,2}) \; : \; d_K(\rho, \gamma) \equiv 1 \pmod{2}\}$. By Lemma 3, the minimum distance of both $\mathcal{C}_2^e$ and $\mathcal{C}_2^o$ is 2. Clearly, the size of either $\mathcal{C}_2^e$ or $\mathcal{C}_2^o$ is at least $\frac{|S(\mathcal{M}_{k,2})|}{2} = \frac{(k+2)!}{k! \cdot 2} = \frac{(k+2)(k+1)}{2}$. For all $k \geq 1$ we have that $\frac{(k+2)(k+1)}{2} \geq 2(k-1)+1$ and hence by Theorem 5 there exists a $(k+2, k)$ systematic single-error-correcting code.*

Example 4. *Let k be an integer such that $k-2$ is a power of a prime, let $r = 3$, and let $M_2 = 8((k-2)^3 - 1)/(k-3) = 8((k-2)^2 + k - 1)$. By Theorem 4, it follows that there exist $h_1, h_2, \ldots, h_{k-1}$ such that for all $\mathbf{e} \in \mathbb{Z}^{k-1}$, $||\mathbf{e}|| \leq 2$, the sums $\sum_{i=1}^{k-1} e_i h_i$ are all distinct modulo M_2. We have to show the existence of a code in $S(\mathcal{M}_{k,3})$ with minimum distance 4 and of size at least M_2. By Corollary 1, there exists a single-error-correcting code $\mathcal{C}_K \subset S(\mathcal{M}_{k,3})$ of size $|\mathcal{C}_K| \geq \frac{|S(\mathcal{M}_{k,3})|}{2 \cdot 3 + 1}$. We fix a multi-permutation $\rho \in S(\mathcal{M}_{k,3})$ and consider the codes $\mathcal{C}_3^e = \{\gamma \in \mathcal{C}_K \; : \; d_K(\rho, \gamma) \equiv 0 \pmod{2}\}$ and $\mathcal{C}_3^o = \{\gamma \in \mathcal{C}_K \; : \; d_K(\rho, \gamma) \equiv 1 \pmod{2}\}$. By Lemma 3, it follows that the minimum distance of the codes $\mathcal{C}_3^e$ and $\mathcal{C}_3^o$ is 4. One of these codes must be of size at least $\frac{|\mathcal{C}_K|}{2}$. If $\mathcal{C}_3$ is this code then $|\mathcal{C}_3| \geq \frac{|S(\mathcal{M}_{k,3})|}{14} = \frac{(k+3)!}{k! \cdot 14} = \frac{(k+3)(k+2)(k+1)}{14}$. For all $k \geq 113$ we have that $\frac{(k+3)(k+2)(k+1)}{14} \geq 8((k-2)^2 + k - 1)$ and hence by Theorem 5, if $k \geq 113$ such that $k-2$ is a power of a prime then there exists a $(k+3, k)$ systematic double-error-correcting code.*

Theorem 4, Theorem 5, and Corollary 2 lead to the following result which follows in the same lines as Example 3 and Example 4.

Corollary 3. *Let t be a power of a prime and let $r = t+1$. Then there exists an integer K_t such that for every integer $k \geq K_t$ for which $k-2$ is a power of a prime, there exists a $(k+r, k)$ systematic t-error-correcting code.*

In [17], [18] a construction of systematic $(k, k+2)$ single-error-correcting codes for permutations with two redundancy symbols was given. They have the same number of redundancy symbols as in Example 3. They construct (n, k) systematic t-error-correcting codes with at most $2t + 1$ redundancy symbols. If k and t have the same magnitude then for some parameters the codes of our construction have the same number of redundancy symbols, but for most parameters the number of redundancy symbols of the codes in our construction is considerably better. Our main theorem is stated as follows.

Theorem 6. *Let k be an integer such that $k-2$ is a power of a prime, let $t = k^\epsilon$ be a positive integer, and let $r = \lceil \mu t \rceil$, where $r - 1$ is a power of a prime. If k is large enough and if*

$$\begin{cases} \mu > 1 + \epsilon & \text{for} \quad 0 \leq \epsilon \leq 1 \\ \mu > 1 + \frac{1}{\epsilon} & \text{for} \quad 1 < \epsilon , \end{cases}$$

then there exists a $(k+r, k)$ systematic t-error-correcting code.

The conditions in Theorem 6 can be relaxed. If we use a power of a prime $k' - 2$, $k \leq k' \leq 2k$, in Theorem 5 and the integers $h_1, h_2, \ldots, h_{k-1}$ from the integers $h_1, h_2, \ldots, h_{k'-1}$ then we can omit the requirement that $k-2$ should be a power of a prime. Related arguments can be used to drop the requirement that $r - 1$ is a power of a prime. The related result is described in Section VI.

V. SYSTEMATIC ECC FOR MULTI-PERMUTATIONS

In this section we generalize the construction in Section IV to obtain systematic error-correcting codes for multi-permutations. In the most general definition of systematic codes for multi-permutations we have a multi-set $\mathcal{K}$ with k elements (with repetitions) serving as the information symbols and a multi-set $\mathcal{R}$ with r elements serving as the redundancy symbols. The intersection between $\mathcal{K}$ and $\mathcal{R}$ must be empty. The codewords are multi-permutations over the multi-set $\mathcal{K} \cup \mathcal{R}$. The number of codewords in the error-correcting code must be the number of distinct multi-permutations over the multi-set $\mathcal{K}$. In the systematic code $\mathcal{C}$ each multi-permutation over the multi-set $\mathcal{K}$, appears as a sub-multi-permutation of exactly one codeword from $\mathcal{C}$. The construction for systematic multi-permutations will be a direct generalization of the construction in Theorem 5. Instead of the set $\mathcal{M}_{k,r}$ we use the set $\mathcal{M}$ defined by $\mathcal{M} \stackrel{\mathrm{def}}{=} \{0^k\} \cup \mathcal{R}$, where 0 is a symbol which does not appear in $\mathcal{R}$. The size of the code $\mathcal{C}_r \subset S(\mathcal{M})$ is at least M_t (note, that the number of h_i's is smaller than $k-1$, unless $\mathcal{K}$ is a set rather than a multi-set, and hence M_t will be smaller).

The challenge for systematic permutations codes is to minimize the number of redundancy symbols of the codes. For systematic error-correcting codes for multi-permutations there is a tradeoff between the number of redundancy ranks and the magnitudes of their multiplicities. For example, in a systematic code for multi-permutations with only one redundancy rank, the multiplicity of the redundancy rank might be large. However, by allowing two redundancy ranks, the multiplicity of each redundancy rank should be smaller. The construction in Theorem 5 allows any desirable number of redundancy ranks.

Example 5. *Let $\mathcal{K} = \{1^{m_1}, 2^{m_2}, \ldots, \ell^{m_\ell}\}$ be a multi-set which consists of $k = \sum_{i=1}^{\ell} m_i$ information symbols, let $\mathcal{R} = \{\ell+1, \ell+1\}$ and $\mathcal{M} = \{0^k, \ell+1, \ell+1\}$. Let $M_1 = 2(k - m_1) + 1$. For every $\mathbf{e} \in \mathbb{Z}^{k-m_1}$, $\|\mathbf{e}\| \leq 1$, the sums $\sum_{i=1}^{k-m_1} e_i i$ are all distinct modulo M_1. For the construction, we need a code in $S(\mathcal{M})$ with minimum distance 2 and of size at least M_1. To this end, fix a multi-permutation $\rho \in S(\mathcal{M})$ and consider the codes $\mathcal{C}_2^e = \{\gamma \in S(\mathcal{M}) : d_K(\rho, \gamma) \equiv 0 \pmod 2\}$ and $\mathcal{C}_2^o = \{\gamma \in S(\mathcal{M}) : d_K(\rho, \gamma) \equiv 1 \pmod 2\}$. By Lemma 3 it follows that the minimum distance of both $\mathcal{C}_2^e$ and $\mathcal{C}_2^o$ is 2. Clearly, the size of either $\mathcal{C}_2^e$ or $\mathcal{C}_2^o$ is at least $\frac{|S(\mathcal{M})|}{2} = \frac{(k+2)!}{k! \cdot 2! \cdot 2} = \frac{(k+2)(k+1)}{4}$. For all $k \geq 1$ we have that $\frac{(k+2)(k+1)}{4} \geq 2(k - m_1) + 1$ and hence by Theorem 5 there exists a systematic single-error-correcting code in $S(\mathcal{K} \cup \mathcal{R})$.*

VI. CONCLUSION

We have considered constructions of systematic error-correcting codes over permutations and multi-permutations with the Kendall's τ-distance. The construction is based on error-correcting codes for multi-permutations. The main result is for a large enough integer k, a positive integer $t = k^\epsilon$, and $r = \lceil \mu t \rceil$. In this case, there exists a $(k+r, k)$ systematic t-error-correcting code if

$$\begin{cases} \mu > 1 + \epsilon & \text{for} \quad 0 \leq \epsilon \leq 1 \\ \mu > 1 + \frac{1}{\epsilon} & \text{for} \quad 1 < \epsilon \, . \end{cases}$$

REFERENCES

[1] A. Barg and A. Mazumdar, "Codes in permutations and error correction for rank modulation," *IEEE Trans. on Information Theory*, vol. 56, pp. 3158–3165, July 2010.
[2] S. Buzaglo, E. Yaakobi, T. Etzion, and J. Bruck, "Error-correcting codes for multipermutations," *Proc. IEEE Inter. Symp. on Information Theory*, pp. 724–728, Istanbul, Turkey, July 2013.
[3] T. M. Cover, "Enumerative source encoding," *IEEE Trans. on Information Theory*, vol. 19, pp. 73–77, January 1973.
[4] P. Diaconis and R. L. Graham, "Spearman's foortule as a measue of disarray," *J. Roy. Statis. Soc. B*, vol. 39, no. 2, pp. 262–268, 1977.
[5] E. En Gad, A. Jiang, and J. Bruck, "Trade-offs between instantaneous and total capacity in multi-cell flash memories," *Proc. IEEE Inter. Symp. on Inform. Theory*, pp. 990–994, Cambridge, MA, July 2012.
[6] E. En Gad, E. Yaakobi, A. Jiang, and J. Bruck, "Rank-modulation rewriting codes for flash memories," *Proc. IEEE Inter. Symp. on Information Theory*, pp. 704–708, Istanbul, Turkey, July 2013.
[7] S. W. Golomb and L. R. Welch, "Perfect codes in the Lee metric and the packing of polyminoes," *SIAM J. Appl. Math.*, vol. 18, pp. 302–317, January 1970.
[8] A. Jiang, R. Mateescu, M. Schwartz, and J. Bruck, "Rank modulation for flash memories," *IEEE Trans. on Information Theory*, vol. 55, pp. 2659–2673, June 2009.
[9] A. Jiang, M. Schwartz, and J. Bruck, "Correcting charge-constrained errors in the rank-modulation scheme," *IEEE Trans. on Information Theory*, vol. 56, pp. 2112–2120, May 2010.
[10] M. Kendall and J. D. Gibbons, *Rank Correlation Methods*, New York: Oxford Univ. Press, 1990.
[11] D. E. Knuth, *The Art of Computer Programming, Volume 3: Sorting and Searching*, Reading, MA: Addiaon-Wesley, 1998.
[12] F. Sala, R. Gabrys, and L. Dolecek, "Dynamic threshold schemes for multi-level non-volatile memories," *IEEE Trans. on Communications*, vol. 61, pp. 2624–2634, July 2012.
[13] A. Mazumdar, A. Barg and G. Zémor, "Construction of rank modulation codes," *IEEE Trans. on Information Theory*, vol. 59, pp. 1018–1029, February 2013.
[14] I. Tamo and M. Schwartz, "Correcting limited-magnitude errors in the rank-modulation scheme," *IEEE Trans. on Information Theory*, vol. 56, pp. 2551–2560, June 2010.
[15] I. Tamo and M. Schwartz, "On the labeling problem of permutation group codes under the infinity metric," *IEEE Trans. on Information Theory*, vol. 58, pp. 6595–6604, October 2012.
[16] R. R. Varshamov and G. M. Tenengolts, "A code for correcting a single asymmetric error," *Autom. Telemkh.*, vol. 26, pp. 288–292, 1965.
[17] H. Zhou, A. Jiang, and J. Bruck, "Systematic error-correction codes for rank modulation," *Proc. IEEE Inter. Symp. on Information Theory*, pp. 2978–2982, Cambridge, MA, July 2012.
[18] H. Zhou, M. Schwartz, A. Jiang, and J. Bruck, "Systematic error-correction codes for rank modulation," *arxiv.org/abs/1310.6817*.