

Bounds for Permutation Rate-Distortion

Farzad Farnoud (Hassanzadeh)

Electrical Engineering
California Institute of Technology
Pasadena, CA 91125, U.S.A.
farnoud@caltech.edu

Moshe Schwartz

Electrical and Computer Engineering
Ben-Gurion University of the Negev
Beer Sheva 8410501, Israel
schwartz@ee.bgu.ac.il

Jehoshua Bruck

Electrical Engineering
California Institute of Technology
Pasadena, CA 91125, U.S.A.
bruck@paradise.caltech.edu

Abstract—We study the rate-distortion relationship in the set of permutations endowed with the Kendall τ -metric and the Chebyshev metric. Our study is motivated by the application of permutation rate-distortion to the average-case and worst-case distortion analysis of algorithms for ranking with incomplete information and approximate sorting algorithms. For the Kendall τ -metric we provide bounds for small, medium, and large distortion regimes, while for the Chebyshev metric we present bounds that are valid for all distortions and are especially accurate for small distortions. In addition, for the Chebyshev metric, we provide a construction for covering codes.

I. INTRODUCTION

In the analysis of sorting and ranking algorithms, it is often assumed that complete information is available, that is, the answer to *every* question of the form “is $x > y$?” can be found, either by query or computation. A standard and straightforward result in this setting is that, on average, one needs at least $\log_2 n!$ pairwise comparisons to sort a randomly-chosen permutation of length n . In practice, however, it is usually the case that only partial information is available. One example is the learning-to-rank problem, where the solutions to pairwise comparisons are learned from data, which may be incomplete, or in big-data settings, where the number of items may be so large as to make it impractical to query every pairwise comparison. It may also be the case that only an approximately-sorted list is required, and thus one does not seek the solutions to all pairwise comparisons. In such cases, the question that arises is what is the quality of a ranking obtained from incomplete data, or an approximately-sorted list.

One approach to quantify the quality of an algorithm that ranks with incomplete data is to find the relationship between the number of comparisons and the average, or worst-case, quality of the output ranking, as measured via a metric on the space of permutations. To explain, consider a deterministic algorithm for ranking n items that makes nR queries and outputs a ranking of length n . Suppose that the true ranking is π . The information about π is available to the algorithm only through the queries it makes. Since the algorithm is deterministic, the output, denoted as $f(\pi)$, is uniquely determined by π . The “distortion” of this output can be measured with a metric d as $d(\pi, f(\pi))$. The goal is to find the relationships between R and $d(\pi, f(\pi))$ when π is chosen at random and when it is chosen by an adversary.

This work was supported in part by an NSF grant CIF-1218005, and a U.S.-Israel Binational Science Foundation (BSF) grant 2010075.

A general way to quantify the best possible performance by such an algorithm is to use the rate-distortion theory on the space of permutations. In this context, the codebook is the set $\{f(\pi) : \pi \in \mathcal{S}_n\}$, where \mathcal{S}_n is the set of permutations of length n , and the rate is determined by the number of queries. For a given rate, no algorithm can have smaller distortion than what is dictated by rate-distortion.

With this motivation, we study rate-distortion in the space of permutations under the Kendall τ -metric and the Chebyshev metric. Previous work on this topic includes [17], which studies permutation rate-distortion with respect to the Kendall τ -metric and the ℓ_1 -metric of inversion vectors, and [7] which considers Spearman’s footrule. Our results on the Kendall τ -metric improve upon those presented in [17]. In particular, for the small distortion regime, as defined later in the paper, we eliminate the gap between the lower bound and the upper bound given in [17]; for the large distortion regime, we provide a stronger lower bound; and for the medium distortion regime, we provide upper and lower bounds with error terms.

Our study includes both worst-case and average-case distortions as both measures are frequently used in the analysis of algorithms. We also note that permutation rate-distortion results can also be applied to lossy compression of permutations, e.g., rank-modulation signals [8]. Finally, we also present covering codes for the Chebyshev metric, where covering codes for the Kendall τ -metric were already presented in [17]. The codes are the covering analog of the error-correcting codes already presented in [1], [9], [11], [16].

The rest of the paper is organized as follows. In Section II, we present preliminaries and notation. Section III contains non-asymptotic results valid for both metrics under study. Finally, Section IV and Section V focus on the Kendall τ -metric and the Chebyshev metric, respectively. Due to lack of space, some of the proofs are omitted or shortened. These can be found in the full version of the paper, available on arXiv [6].

II. PRELIMINARIES AND DEFINITIONS

For a nonnegative integer n , let $[n]$ denote the set $\{1, \dots, n\}$, and let \mathcal{S}_n denote the set of permutations of $[n]$. We denote a permutation $\sigma \in \mathcal{S}_n$ as $\sigma = [\sigma_1, \sigma_2, \dots, \sigma_n]$, where the permutation sets $\sigma(i) = \sigma_i$. We also denote the identity permutation by $\text{Id} = [1, 2, \dots, n]$.

The Kendall τ -distance between two permutations $\pi, \sigma \in \mathcal{S}_n$ is the number of transpositions of adjacent elements needed to transform π into σ , and is denoted by $d_K(\pi, \sigma)$. In contrast,

the Chebyshev distance between π and σ is defined as

$$d_C(\pi, \sigma) = \max_{i \in [n]} |\pi(i) - \sigma(i)|.$$

Additionally, let $d(\pi, \sigma)$ denote a generic distance measure between π and σ .

Both d_K and d_C are invariant; the former is left-invariant and the latter is right-invariant [5]. In other words, for all $f, g, h \in \mathcal{S}_n$, we have $d_K(f, g) = d_K(hf, hg)$ and $d_C(f, g) = d_C(fh, gh)$. Hence, the size of the ball of a given radius in either metric does not depend on its center. The size of a ball of radius r with respect to d_K , d_C , and d , is given, respectively, by $B_K(r)$, $B_C(r)$, and $B(r)$. The dependence of the size of the ball on n is implicit.

A code C is a subset $C \subseteq \mathcal{S}_n$. For a code C and a permutation $\pi \in \mathcal{S}_n$, let $d(\pi, C) = \min_{\sigma \in C} d(\pi, \sigma)$ be the (minimal) distance between π and C .

We use $\hat{M}(D)$ to denote the minimum number of codewords required for a worst-case distortion D . That is, $\hat{M}(D)$ is the size of the smallest code C such that for all $\pi \in \mathcal{S}_n$, we have $d(\pi, C) \leq D$. Similarly, let $\bar{M}(D)$ denote the minimum number of codewords required for an average distortion D under the uniform distribution on \mathcal{S}_n , that is, the size of the smallest code C such that $\frac{1}{n!} \sum_{\pi \in \mathcal{S}_n} d(\pi, C) \leq D$. Note that $\bar{M}(D) \leq \hat{M}(D)$. In what follows, we assume that the distortion D is an integer. For worst-case distortion, this assumption does not lead to a loss of generality as the metrics under study are integer valued.

We also define

$$\begin{aligned} \hat{R}(D) &= \frac{1}{n} \lg \hat{M}(D), & \bar{R}(D) &= \frac{1}{n} \lg \bar{M}(D), \\ \hat{A}(D) &= \frac{1}{n} \lg \frac{\hat{M}(D)}{n!}, & \bar{A}(D) &= \frac{1}{n} \lg \frac{\bar{M}(D)}{n!}, \end{aligned}$$

where we use \lg as a shorthand for \log_2 . It is clear that $\hat{R}(D) = \hat{A}(D) + \lg n! / n$, and that a similar relationship holds between $\bar{R}(D)$ and $\bar{A}(D)$. The reason for defining \hat{A} and \bar{A} is that they sometimes lead to simpler expressions compared to \hat{R} and \bar{R} . Furthermore, \hat{A} (resp. \bar{A}) can be interpreted as the difference between the number of bits per symbol required to identify a codeword in a code of size \hat{M} (resp. \bar{M}) and the number of bits per symbol required to identify a permutation in \mathcal{S}_n .

Throughout the paper, for \hat{M} , \bar{M} , \hat{A} , \bar{A} , \hat{R} , and \bar{R} , subscripts K and C denote that the subscripted quantity corresponds to the Kendall τ -metric and the Chebyshev metric, respectively. Lack of subscripts indicates that the result is valid for both metrics.

In the sequel, the following inequalities [4] will be useful,

$$\frac{2^{nH(p)}}{\sqrt{8np(1-p)}} \leq \binom{n}{pn} \leq \frac{2^{nH(p)}}{\sqrt{2\pi np(1-p)}}, \quad (1)$$

where $H(\cdot)$ is the binary entropy function, and $0 < p < 1$. Furthermore, to denote $\lim_{x \rightarrow \infty} \frac{f(x)}{g(x)} = 1$, we use

$$f(x) \sim g(x) \text{ as } x \rightarrow \infty,$$

or $f \sim g$ if doing so does not cause ambiguity.

III. NON-ASYMPTOTIC BOUNDS

In this section, we derive non-asymptotic bounds, that is, bounds that are valid for all positive integers n and D . The results in this section apply to both the Kendall τ -distance and the Chebyshev distance as well as any other invariant distance on permutations.

The next lemma gives two basic lower bounds for $\hat{M}(D)$ and $\bar{M}(D)$.

Lemma 1. For all $n, D \in \mathbb{N}$,

$$\hat{M}(D) \geq \frac{n!}{B(D)}, \quad \bar{M}(D) > \frac{n!}{B(D)(D+1)}.$$

Proof: Since the first inequality is well known and its proof is clear, we only prove the second one. Fix n and D . Consider a code $C \subseteq \mathcal{S}_n$ of size M and suppose the average distortion of this code is at most D . There are at most $MB(D)$ permutations π such that $d(\pi, C) \leq D$ and at least $n! - MB(D)$ permutations π such that $d(\pi, C) \geq D + 1$. Hence, $D > (D+1)(1 - MB(D)/n!)$. The second inequality then follows. ■

In the next lemma, we use a simple probabilistic argument to give an upper bound on $\hat{M}(D)$.

Lemma 2. For all $n, D \in \mathbb{N}$, $\hat{M}(D) \leq \lceil n! \ln n! / B(D) \rceil$.

Proof: Suppose that a sequence of M permutations, π_1, \dots, π_M , is drawn by choosing each π_i i.i.d. with uniform distribution over \mathcal{S}_n . Denote $C = \{\pi_1, \dots, \pi_M\} \subseteq \mathcal{S}_n$. The probability P_f that there exists $\sigma \in \mathcal{S}_n$ with $d(\sigma, C) > D$ is bounded by

$$\begin{aligned} P_f &\leq \sum_{\sigma \in \mathcal{S}_n} P(\forall i : d(\pi_i, \sigma) > D) = n!(1 - B(D)/n!)^M \\ &< n!e^{-MB(D)/n!} = e^{\ln n! - MB(D)/n!}. \end{aligned}$$

Let $M = \lceil n! \ln n! / B(D) \rceil$ so that $P_f < 1$. Hence, a code of size M exists with worst-case distortion D . ■

The following theorem by Stein [15], which can be used to obtain existence results for covering codes (see, e.g., [4]), enables us to improve the above upper bound. We use a simplified version of this theorem, which is sufficient for our purpose.

Theorem 3. [15] Consider a set X , with $|X| = N$, and a family $\{A_i\}_{i=1}^N$ of subsets of X . Suppose there is an integer Q such that $|A_i| = Q$ for all i and that each element of X is in Q of the sets A_i . Then there is subfamily of $\{A_i\}_{i=1}^N$, containing at most $(N/Q)(1 + \ln Q)$ sets, that covers X .

In our context X is \mathcal{S}_n , A_i are the balls of radius D centered at each permutation, $N = n!$ and $Q = B(D)$. Hence, the theorem implies that

$$\hat{M}(D) \leq \frac{n!}{B(D)} (1 + \ln B(D)).$$

The following theorem summarizes the results of this section.

Theorem 4. For all $n, D \in \mathbb{N}$,

$$\frac{n!}{B(D)} \leq \hat{M}(D) \leq \frac{n!}{B(D)} (1 + \ln B(D)), \quad (2)$$

$$\frac{n!}{B(D)(D+1)} < \bar{M}(D) \leq \hat{M}(D). \quad (3)$$

IV. THE KENDALL τ -METRIC

The goal of this section is to consider the rate-distortion relationship for the permutation space endowed by the Kendall τ -metric. First, we find non-asymptotic upper and lower bounds on the size of the ball in the Kendall τ -metric. Then, in the following subsections, we consider asymptotic bounds for small, medium, and large distortion regimes. Throughout this section, we assume $1 \leq D < \frac{1}{2}\binom{n}{2}$ and $n \geq 1$. Note that D is upper bounded by $\binom{n}{2}$, and the case of $\frac{1}{2}\binom{n}{2} \leq D \leq \binom{n}{2}$ leads to trivial codes, e.g., $\{\text{Id}, [n, n-1, \dots, 1]\}$ and $\{\text{Id}\}$.

A. Non-asymptotic Results

Let \mathbb{X}_n be the set of integer vectors $x = x_1, x_2, \dots, x_n$ of length n such that $0 \leq x_i \leq i-1$ for $i \in [n]$. It is well known (for example, see [9]) that there is a bijection between \mathbb{X}_n and \mathbb{S}_n such that for corresponding elements $x \in \mathbb{X}_n$ and $\pi \in \mathbb{S}_n$, we have $d_K(\pi, \text{Id}) = \sum_{i=2}^n x_i$. Hence

$$B_K(r) = \left| \left\{ x \in \mathbb{X}_n : \sum_{i=2}^n x_i \leq r \right\} \right|, \quad (4)$$

for $1 \leq r \leq \binom{n}{2}$. Thus, the number of nonnegative integer solutions to the equation $\sum_{i=2}^n x_i \leq r$ is at least $B_K(r)$, i.e.,

$$B_K(r) \leq \binom{r+n-1}{r}. \quad (5)$$

Furthermore, for $\delta \in \mathbb{Q}$, $\delta \geq 0$, such that δn is an integer, it can also be shown that

$$B_K(\delta n) \geq [1+\delta]![1+\delta]^{n-[1+\delta]}, \quad (6)$$

by noting the fact that the right-hand side of (6) counts the elements of \mathbb{X}_n such that

$$\begin{cases} 0 \leq x_i \leq i-1, & \text{for } i \leq [1+\delta], \\ 0 \leq x_i \leq [\delta], & \text{for } i > [1+\delta], \end{cases}$$

and that $\sum_{i \leq [1+\delta]} (i-1) + (n - [1+\delta])[\delta] \leq [\delta]n \leq \delta n$.

Next we find a lower bound on $B_K(r)$ with $r < n$. Let $I(n, r)$ denote the number of permutations in \mathbb{S}_n that are at distance r from the identity. We have [2, p. 51]

$$I(n, r) = \binom{n+r-1}{r} + \sum_{j=1}^{\infty} (-1)^j f_j,$$

where $f_j = \binom{n+r-(u_j-j)-1}{r-(u_j-j)} + \binom{n+r-u_j-1}{r-u_j}$, and $u_j = (3j^2 + j)/2$. It can be shown that $\binom{n+r-1}{r} - f_1 \geq \frac{1}{4}\binom{n+r-1}{r}$ and that, for $j \geq 2$, we have $f_j \geq f_{j+1}$. Thus, for $r < n$,

$$B_K(r) \geq I(n, r) \geq \frac{1}{4} \binom{n+r-1}{r}. \quad (7)$$

In the next two theorems, we use the aforementioned bounds on $B_K(r)$ to derive lower and upper bounds on $\hat{A}_K(D)$ and $\bar{A}_K(D)$.

Theorem 5. For all $n, D \in \mathbb{N}$, and $\delta = D/n$,

$$\begin{aligned} \hat{A}_K(D) &\geq -\lg \frac{(1+\delta)^{1+\delta}}{\delta^\delta}, \\ \bar{A}_K(D) &\geq -\lg \frac{(1+\delta)^{1+\delta}}{\delta^\delta} - \frac{\lg n}{n}. \end{aligned}$$

Proof: Using (5), (1), and the fact that $\delta \geq 1/n$, we can show $B_K(D) \leq 2^{n(1+\delta)H(\frac{1}{1+\delta})}$. The first result then follows from (2). The proof of the second result is in essence similar but uses (3). ■

To obtain the next theorem, the lower bounds given in (6) and (7) are used. We omit a detailed proof.

Theorem 6. Assume $n, D \in \mathbb{N}$, and let $\delta = D/n$. We have

$$\bar{A}_K(D) \leq \hat{A}_K(D) \leq -\lg \frac{(1+\delta)^{1+\delta}}{\delta^\delta} + \frac{3 \lg n + 12}{2n},$$

for $\delta < 1$, and

$$\bar{A}_K(D) \leq \hat{A}_K(D) \leq -\lg[1+\delta] + \frac{1}{n} \lg \left(ne^{[1+\delta]} \ln[1+\delta] \right),$$

for $\delta \geq 1$.

B. Small Distortion

In this subsection, we consider small distortion, that is, $D = O(n)$. First, suppose $D < n$, or equivalently, $\delta = D/n < 1$. The next lemma follows from Theorems 5 and 6.

Lemma 7. For $\delta = D/n < 1$, we have that

$$\hat{A}_K(D) = -\lg \frac{(1+\delta)^{1+\delta}}{\delta^\delta} + O\left(\frac{\lg n}{n}\right), \quad (8)$$

and that $\bar{A}_K(D)$ satisfies the same equation.

The next theorem, gives the asymptotic size of the ball $B_K(D)$ where $D = \Theta(n)$.

Theorem 8. Let $n = n(r) = \frac{r}{c} + O(1)$ for a constant $c > 0$. Then

$$B_K(r) \sim K_c \binom{n+r-1}{r} \quad (9)$$

as $r, n \rightarrow \infty$, where K_c is a positive constant that depends on c .

We omit the proof of the theorem due to its length and only mention that it relies on a probabilistic transform, namely, Theorem 3.1 of [12].

One can use (9), Theorem 4, and the definitions of \hat{A}_K and \bar{A}_K to obtain the following lemma.

Lemma 9. For a constant $c > 0$ and $D = cn + O(1)$, we have

$$\hat{A}_K(cn + O(1)) = -\lg \frac{(1+c)^{1+c}}{c^c} + O\left(\frac{\lg n}{n}\right). \quad (10)$$

Furthermore, $\bar{A}_K(cn + O(1))$ satisfies the same equation.

The results given in (8) and (10) are given as lower bounds in [17, Equation (14)]. We have thus shown that these lower bounds in fact match the quantity under study. Furthermore, we have shown that $\bar{A}_K(D)$ satisfies the same relations.

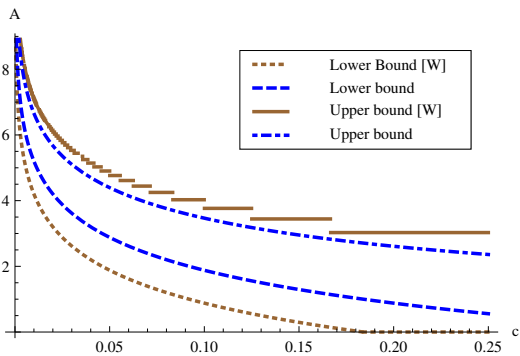


Figure 1. Bounds on $\hat{A}_K(D) + \lg n$ for $D = cn^2 + O(n)$ where the error terms are ignored. The bounds denoted by [W] are those from [17].

C. Medium Distortion

We next consider the medium distortion regime, that is, $D = cn^{1+\alpha} + O(n)$ for constants $c > 0$ and $0 < \alpha < 1$. For this case, from [17], we have $\hat{A}_K(D) \sim -\lg n^\alpha$. We improve upon this result by providing upper and lower bounds with error terms. Using Theorems 5 and 6, one can show the following lemma.

Lemma 10. For $D = cn^{1+\alpha} + O(n)$, where α and c are constants such that $0 < \alpha < 1$ and $c > 0$, we have

$$\begin{aligned} -\lg(ecn^\alpha) + O(n^{-\alpha}) &\leq \hat{A}_K(D) \\ &\leq -\lg(cn^\alpha) + O(n^{-\alpha} + n^{\alpha-1}) \end{aligned}$$

D. Large Distortion

In the large distortion regime, we have $D = cn^2 + O(n)$ and $\delta = cn + O(1)$. The following lemma can be proved using Theorems 5 and 6.

Lemma 11. Suppose $D = cn^2 + O(n)$ for a constant $0 < c < \frac{1}{2}$. We have

$$\begin{aligned} -\lg(ecn) + O(n^{-1}) &\leq \hat{A}_K(D) \leq \\ &-\lg(ecn) + (1+c)\lg e + O(n^{-1}\lg n). \end{aligned}$$

From [17], it follows that $-\lg(ecn) - 1 + O(n^{-1}\lg n) \leq \hat{A}_K(D) \leq -\lg \frac{n}{e^{\lceil 1/(2c) \rceil}} + O(n^{-1}\lg n)$. These bounds and those of Lemma 11 are compared in Figure 1, where we added the term $\lg n$ to remove dependence on n .

V. THE CHEBYSHEV METRIC

We now turn to consider the rate-distortion function for the permutation space under the Chebyshev metric. We start by stating lower and upper bounds on the size of the ball in the Chebyshev metric, and then construct covering codes.

A. Bounds

For an $n \times n$ matrix A , the permanent of $A = (A_{i,j})$ is defined as,

$$\text{per}(A) = \sum_{\pi \in \mathcal{S}_n} \prod_{i=1}^n A_{i,\pi(i)}.$$

It is well known [10], [14] that $B_C(r)$ can be expressed as the permanent of the $n \times n$ binary matrix A for which

$$A_{i,j} = \begin{cases} 1 & |i-j| \leq r \\ 0 & \text{otherwise.} \end{cases} \quad (11)$$

According to Brégman's Theorem (see [3]), for any $n \times n$ binary matrix A with r_i 1's in the i -th row

$$\text{per}(A) \leq \prod_{i=1}^n (r_i!)^{\frac{1}{r_i}}.$$

Using this bound we can state the following lemma (partially given in [10] and extended in [16]).

Lemma 12. [16] For all $0 \leq r \leq n-1$,

$$B_C(r) \leq \begin{cases} ((2r+1)!)^{\frac{n-2r}{2r+1}} \prod_{i=r+1}^{2r} (i!)^{\frac{2}{i}}, & 0 \leq r \leq \frac{n-1}{2}, \\ (n!)^{\frac{2r+2-n}{n}} \prod_{i=r+1}^{n-1} (i!)^{\frac{2}{i}}, & \frac{n-1}{2} \leq r \leq n-1. \end{cases}$$

Lemma 13. For all $0 \leq r \leq n-1$,

$$B_C(r) \geq \begin{cases} \frac{(2r+1)^n n!}{2^{2r} n^n}, & 0 \leq r \leq \frac{n-1}{2}, \\ \frac{n!}{2^{2(n-r)}}, & \frac{n-1}{2} \leq r \leq n-1. \end{cases}$$

Proof: The first case was already proved in [10]. Thus, only the second claim requires proof, so suppose that $(n-1)/2 \leq r \leq n-1$. The proof follows the same lines as the one appearing in [10]. Let A be defined as in (11), and let B be an $n \times n$ matrix with

$$B_{i,j} = \begin{cases} 2, & i+j \leq n-r, \\ 2, & i+j \geq n+r+2, \\ A_{i,j}, & \text{otherwise.} \end{cases}$$

We observe that B/n is doubly stochastic. It follows that

$$B_C(r) = \text{per}(A) \geq \frac{\text{per}(B)}{2^{2(n-r)}} \geq \frac{n^n \text{per}(B/n)}{2^{2(n-r)}} \geq \frac{n!}{2^{2(n-r)}},$$

where the last inequality follows from Van der Waerden's Theorem [13]. ■

Theorem 14. Let $n \in \mathbb{N}$, and let $0 < \delta < 1$ be a constant rational number such that $D = \delta n$ is an integer. Then

$$\hat{R}_C(D) \geq \begin{cases} \lg \frac{1}{2\delta} + 2\delta \lg \frac{e}{2} + O(\lg n/n), & 0 < \delta \leq \frac{1}{2} \\ 2\delta \lg \delta + 2(1-\delta) \lg e + O(\lg n/n), & \frac{1}{2} \leq \delta \leq 1 \end{cases}$$

and

$$\hat{R}_C(D) \leq \begin{cases} \lg \frac{1}{2\delta} + 2\delta + O(\lg n/n), & 0 < \delta \leq \frac{1}{2} \\ 2(1-\delta) + O(\lg n/n), & \frac{1}{2} \leq \delta \leq 1 \end{cases}$$

Furthermore, the same bounds also hold for $\bar{R}_C(D)$.

Proof: (outline) To prove the lower bound for $\hat{R}_C(D)$, we first use Lemma 12 to find an asymptotic upper bound on $B_C(D)$. Then, from Theorem 4, which states that $\hat{M}_C(D) \geq n!/B_C(D)$, we find a lower bound on $\hat{R}_C(D) = \lg \hat{M}_C(D)/n$.

To prove the upper bound for $\hat{R}_C(D)$, from Lemma 13, we find a lower bound on $B_C(D)$. The result then follows from Lemma 2, stating that $\hat{M}_C(D) \leq n! \ln n! / B_C(D)$, and the fact that $\hat{R}_C(D) = \lg \hat{M}_C(D)/n$.

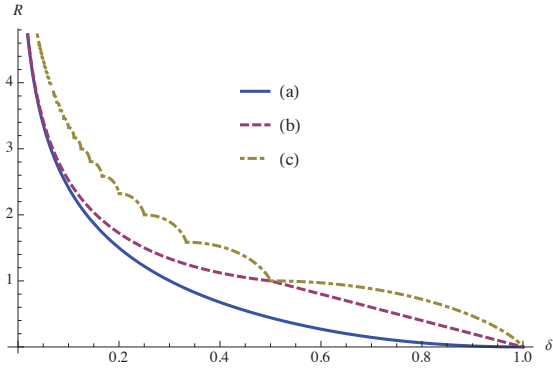


Figure 2. Rate-distortion in the Chebyshev metric: The lower and upper bounds of Theorem 14, (a) and (b), and the rate of the code construction, given in Theorem 17, (c).

The proof of the lower bound for $\bar{R}_C(D)$ is similar to that of $\hat{R}_C(D)$ except that we use $\bar{M}(D) > n!/(B(D)(D+1))$ from Theorem 4. The proof of the upper bound for $\bar{R}_C(D)$ follows from the fact that $\bar{R}_C(D) \leq \hat{R}_C(D)$. ■

B. Code Construction

Let $A = \{a_1, a_2, \dots, a_m\} \subseteq [n]$ be a subset of indices, $a_1 < a_2 < \dots < a_m$. For any permutation $\sigma \in \mathcal{S}_n$ we define $\sigma|_A$ to be the permutation in \mathcal{S}_m that preserves the relative order of the sequence $\sigma(a_1), \sigma(a_2), \dots, \sigma(a_m)$. Intuitively, to compute $\sigma|_A$ we keep only the *coordinates* of σ from A , and then relabel the entries to $[m]$ while keeping relative order. In a similar fashion we define $\sigma^A = (\sigma^{-1}|_A)^{-1}$. Intuitively, to calculate σ^A we keep only the *values* of σ from A , and then relabel the entries to $[m]$ while keeping relative order.

Example 15. Let $n = 6$ and consider the permutation $\sigma = [6, 1, 3, 5, 2, 4]$. We take $A = \{3, 5, 6\}$. We then have $\sigma|_A = [2, 1, 3]$, since we keep positions 3, 5, and 6, of σ , giving us $[3, 2, 4]$, and then relabel these to get $[2, 1, 3]$. Similarly, we have $\sigma^A = [3, 1, 2]$, since we keep the values 3, 5, and 6, of σ , giving us $[6, 3, 5]$, and then relabel these to get $[3, 1, 2]$. □

Construction 1. Let n and d be positive integers, $1 \leq d \leq n-1$. Furthermore, we define the sets

$$A_i = \{i(d+1) + j \mid 1 \leq j \leq d+1\} \cap [n],$$

for all $0 \leq i \leq \lfloor (n-1)/(d+1) \rfloor$. We now construct the code C defined by

$$C = \left\{ \sigma \in \mathcal{S}_n \mid \sigma^{A_i} = \text{Id for all } i \right\}.$$

We note that this construction may be seen as a dual of the construction given in [17].

Theorem 16. Let n and d be positive integers, $1 \leq d \leq n-1$. Then the code $C \subseteq \mathcal{S}_n$ of Construction 1 has covering radius exactly d and size

$$M = \frac{n!}{(d+1)!^{\lfloor n/(d+1) \rfloor} (n \bmod (d+1))!}.$$

Proof: Due to lack of space, we only prove the fact that the code $C \subseteq \mathcal{S}_n$ of Construction 1 has covering radius (at most) d .

Let $\sigma \in \mathcal{S}_n$ be any permutation. We let I_i denote the indices in which the elements of A_i appear in σ . Let us now construct a new permutation σ' in which the elements of A_i appear in indices I_i , but they sorted in ascending order. Thus $\sigma'^{A_i} = \text{Id}$, for all i , and so σ' is a codeword in C .

We observe that if $\sigma(j) \in A_i$, then $\sigma'(j) \in A_i$ as well. It follows that $|\sigma(j) - \sigma'(j)| \leq d$ and so $d_C(\sigma, \sigma') \leq d$. ■

The code construction has the following asymptotic form:

Theorem 17. The code from Construction 1 has the following asymptotic rate,

$$R = H\left(\delta \left\lfloor \frac{1}{\delta} \right\rfloor\right) + \delta \left\lfloor \frac{1}{\delta} \right\rfloor \lg \left\lfloor \frac{1}{\delta} \right\rfloor,$$

where H is the binary entropy function.

The bounds given in Theorem 14 and the rate of the code construction, given in Theorem 17, are shown in Figure 2.

REFERENCES

- [1] A. Barg and A. Mazumdar, "Codes in permutations and error correction for rank modulation," *IEEE Trans. Inform. Theory*, vol. 56, no. 7, pp. 3158–3165, Jul. 2010.
- [2] M. Bóna, *Combinatorics of permutations*. CRC Press, 2012.
- [3] L. M. Brégman, "Some properties of nonnegative matrices and their permanents," *Soviet Math. Dokl.*, vol. 14, pp. 945–949, 1973.
- [4] G. Cohen, I. Honkala, S. Litsyn, and A. Lobstein, *Covering Codes*. North-Holland, 1997.
- [5] M. Deza and H. Huang, "Metrics on permutations, a survey," *J. Comb. Inf. Sys. Sci.*, vol. 23, pp. 173–185, 1998.
- [6] F. Farnoud, M. Schwartz, and J. Bruck, "Rate-distortion for ranking with incomplete information," *arXiv preprint: http://arxiv.org/abs/1401.3093*, 2014.
- [7] J. Giesen, E. Schuberth, and M. Stojakovi, "Approximate sorting," in *LATIN 2006: Theoretical Informatics*, no. 3887. Springer, Jan. 2006, pp. 524–531.
- [8] A. Jiang, R. Mateescu, M. Schwartz, and J. Bruck, "Rank modulation for flash memories," *IEEE Trans. Inform. Theory*, vol. 55, no. 6, pp. 2659–2673, Jun. 2009.
- [9] A. Jiang, M. Schwartz, and J. Bruck, "Correcting charge-constrained errors in the rank-modulation scheme," *IEEE Trans. Inform. Theory*, vol. 56, no. 5, pp. 2112–2120, May 2010.
- [10] T. Kløve, "Spheres of permutations under the infinity norm – permutations with limited displacement," University of Bergen, Bergen, Norway, Tech. Rep. 376, Nov. 2008.
- [11] A. Mazumdar, A. Barg, and G. Zémor, "Constructions of rank modulation codes," *IEEE Trans. Inform. Theory*, vol. 59, no. 2, pp. 1018–1029, Feb. 2013.
- [12] O. Milenkovic and K. J. Compton, "Probabilistic transforms for combinatorial urn models," *Combinatorics, Probability, and Computing*, vol. 13, no. 4-5, pp. 645–675, Jul. 2004.
- [13] H. Minc, "Permanents," in *Encyclopedia of Mathematics and its Applications*. Cambridge University Press, 1978, vol. 6.
- [14] M. Schwartz, "Efficiently computing the permanent and Hafnian of some banded Toeplitz matrices," *Linear Algebra and its Applications*, vol. 430, no. 4, pp. 1364–1374, Feb. 2009.
- [15] S. Stein, "Two combinatorial covering theorems," *Journal of Combinatorial Theory, Series A*, vol. 16, no. 3, pp. 391–397, 1974.
- [16] I. Tamo and M. Schwartz, "Correcting limited-magnitude errors in the rank-modulation scheme," *IEEE Trans. Inform. Theory*, vol. 56, no. 6, pp. 2551–2560, Jun. 2010.
- [17] D. Wang, A. Mazumdar, and G. W. Wornell, "A rate distortion theory for permutation spaces," in *Proc. IEEE Int. Symp. Information Theory (ISIT), Istanbul, Turkey*, Jul. 2013, pp. 2562–2566.