Rate-Distortion for Ranking with Incomplete Information

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 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 rankings, 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 relationship between R and $d(\pi, f(\pi))$ when π is chosen at random and when it is chosen by an adversary.

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 $\{f(\pi) : \pi \in \mathsf{S}_n\}$, where 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.

In this work we study rate distortion in the Kendall τ-metric, which counts the number of pairs that are ranked incorrectly, and the Chebyshev metric, which is the largest error in the rank of any item. 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 [2], [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.

II. PRELIMINARIES AND DEFINITIONS

For a nonnegative integer n, let $[n]$ denote the set $\{1, \ldots, n\}$, and let S_n denote the set of permutations of $[n]$. We denote a permutation $\sigma \in S_n$ as $\sigma = [\sigma_1, \sigma_2, \ldots, \sigma_n]$,

where the permutation sets $\sigma(i) = \sigma_i$. We also denote the identity permutation by $\mathrm{Id} = [1, 2, \ldots, n]$.

The Kendall τ-distance between two permutations $\pi, \sigma \in \mathbb{S}_n$ 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 [6]. 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 $C \subseteq \mathbb{S}_n$. For a code C and a permutation $\pi \in \mathbb{S}_n$, 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 $\pi \in \mathbb{S}_n$, we have $d(\pi, C) \leqslant D$. Similarly, let $\bar{M}(D)$ denote the minimum number of codewords required for an average distortion D under the uniform distribution on $\mathbb{S}_n$, that is, the size of the smallest code C such that

$$\frac{1}{n!} \sum_{\pi \in \mathbb{S}_n} d(\pi, C) \leqslant D.$$

Note that $\bar{M}(D) \leqslant \hat{M}(D)$. 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

$$\hat{R}(D) = \frac{1}{n} \lg \hat{M}(D), \qquad \bar{R}(D) = \frac{1}{n} \lg \bar{M}(D),$$

$$\hat{A}(D) = \frac{1}{n} \lg \frac{\hat{M}(D)}{n!}, \qquad \bar{A}(D) = \frac{1}{n} \lg \frac{\bar{M}(D)}{n!},$$

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 $\mathbb{S}_n$.

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 [5] will be useful,

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

$$\sqrt{2\pi n}(n/e)^n < n! < \sqrt{2\pi n}(n/e)^n e^{1/(12n)}, \qquad (2)$$

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

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

or if the variable x is clear from the context, we simply write $f \sim g$.

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 $n, D \in \mathbb{N}$,*

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

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 $C \subseteq \mathbb{S}_n$ of size M and suppose the average distortion of this codes is at most D. There are at most $MB(D)$ permutations π such that $d(\pi, C) \leqslant D$ and at least $n! - MB(D)$ permutations π such that $d(\pi, C) \geqslant D + 1$. 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 $n, D \in \mathbb{N}$, $\hat{M}(D) \leqslant \lceil n! \ln n!/B(D) \rceil$.*

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

$$P_f \leqslant \sum_{\sigma \in \mathbb{S}_n} P(\forall i : d(\pi_i, \sigma) > D) = n!(1 - B(D)/n!)^M$$

$$< n! e^{-MB(D)/n!} = e^{\ln n! - MB(D)/n!}.$$

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., [5]), 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 and a family $\{A_i\}_{i=1}^N$ of sets that cover X. Suppose there are integers N and Q such that, $|X| = N$, $|A_i| \leqslant Q$ for all i, and each element of X is*

in at least 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 cover X.

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

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

The following theorem summarizes the results of this section.

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

$$\frac{n!}{\mathsf{B}(D)} \leqslant \hat{\mathsf{M}}(D) \leqslant \frac{n!}{\mathsf{B}(D)}(1 + \ln \mathsf{B}(D)), \tag{3}$$

$$\frac{n!}{\mathsf{B}(D)(D+1)} < \bar{\mathsf{M}}(D) \leqslant \hat{\mathsf{M}}(D). \tag{4}$$

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 $1 \leqslant D < \frac{1}{2}\binom{n}{2}$ and $n \geqslant 1$. Note that the case of $D \geqslant \frac{1}{2}\binom{n}{2}$ leads to the trivial codes, e.g., $\{\mathrm{Id}, [n, n-1, \ldots, 1]\}$ and $\{\mathrm{Id}\}$.

A. Non-asymptotic Results

Let $\mathbb{X}_n$ be the set of integer vectors $x = x_1, x_2, \ldots, x_n$ of length n such that $0 \leqslant x_i \leqslant i - 1$ 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

$$\mathsf{d}_K(\pi, \mathrm{Id}) = \sum_{i=2}^n x_i.$$

Hence

$$\mathsf{B}_K(r) = \left| \left\{ x \in \mathbb{X}_n : \sum_{i=2}^n x_i \leqslant r \right\} \right|, \tag{5}$$

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

$$\mathsf{B}_K(r) \leqslant \binom{r+n-1}{r}. \tag{6}$$

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

$$\mathsf{B}_K(\delta n) \geqslant \lfloor 1 + \delta \rfloor! \lfloor 1 + \delta \rfloor^{n - \lfloor 1 + \delta \rfloor}, \tag{7}$$

by noting the fact that the right-hand side of (7) counts the elements of $\mathbb{X}_n$ such that

$$\begin{cases} 0 \leqslant x_i \leqslant i - 1, & \text{for } i \leqslant \lfloor 1 + \delta \rfloor, \\ 0 \leqslant x_i \leqslant \lfloor \delta \rfloor, & \text{for } i > \lfloor 1 + \delta \rfloor, \end{cases}$$

and that $\sum_{i \leqslant \lfloor 1+\delta \rfloor}(i-1) + (n - \lfloor 1 + \delta \rfloor)\lfloor \delta \rfloor \leqslant \lfloor \delta \rfloor n \leqslant \delta n$.

Next we find a lower bound on $\mathsf{B}_K(r)$ 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 [3, p. 51]

$$I(n, r) = \binom{n+r-1}{r} - \left(\binom{n+r-2}{r-1} + \binom{n+r-3}{r-2} \right)$$
$$+ \sum_{j=2}^{\infty}(-1)^j f_j,$$

where

$$f_j = \binom{n+r-(u_j-j)-1}{r-(u_j-j)} + \binom{n+r-u_j-1}{r-u_j},$$

and $u_j = (3j^2 + j)/2$. For $j \geqslant 2$, we have $f_j \geqslant f_{j+1}$. Thus, for $r < n$,

$$I(n,r) \geqslant \binom{n+r-1}{r}\left(1 - \frac{r}{n+r-1}\left(1 + \frac{r-1}{n+r-2}\right)\right)$$
$$\geqslant \frac{1}{4}\binom{n+r-1}{r}.$$

Hence, for $r < n$, we have

$$\mathsf{B}_K(r) \geqslant \frac{1}{4}\binom{n+r-1}{r}. \tag{8}$$

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

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

$$\hat{\mathsf{A}}(D) \geqslant -\lg \frac{(1+\delta)^{1+\delta}}{\delta^\delta},$$

$$\bar{\mathsf{A}}(D) \geqslant -\lg \frac{(1+\delta)^{1+\delta}}{\delta^\delta} - \frac{\lg n}{n}.$$

Proof: For the worst-case distortion, we have

$$\mathsf{B}_K(D) \overset{(a)}{\leqslant} \binom{n+\delta n-1}{\delta n} \leqslant \binom{(1+\delta)n}{\delta n}$$
$$\overset{(b)}{\leqslant} \frac{2^{n(1+\delta)H\left(\frac{1}{1+\delta}\right)}}{\sqrt{2\pi n\delta/(1+\delta)}} \overset{(c)}{\leqslant} 2^{n(1+\delta)H\left(\frac{1}{1+\delta}\right)},$$

where (a) follows from (6), (b) follows from (1), and (c) follows from the facts that $\delta \geqslant 1/n$ and $n \geqslant 1$. The first result then follows from (3).

For the case of average distortion, we proceed as follows:

$$\mathsf{B}_K(D)(D+1) \leqslant \mathsf{B}_K(\delta n)(\delta n + 1)$$
$$\leqslant \binom{n+\delta n-1}{\delta n}(\delta n + 1)$$
$$= \binom{n+\delta n}{\delta n}\frac{\delta n + 1}{1 + \delta}$$
$$\overset{(a)}{\leqslant} 2^{n(1+\delta)H\left(\frac{1}{1+\delta}\right)}\frac{\delta n + 1}{\sqrt{2\pi n\delta(1+\delta)}}$$
$$= 2^{n(1+\delta)H\left(\frac{1}{1+\delta}\right)}\sqrt{\frac{2\delta n}{\pi}}\frac{1 + 1/(\delta n)}{2\sqrt{n\delta(1+\delta)}}$$
$$\overset{(b)}{\leqslant} 2^{n(1+\delta)H\left(\frac{1}{1+\delta}\right)}\sqrt{2\delta n/\pi},$$

where (a) follows from (1) and (b) is proved as follows. The expression $\frac{1+1/(\delta n)}{2\sqrt{n\delta(1+\delta)}}$ is decreasing in δ for positive δ and so it is maximized by letting $\delta = 1/n$. Hence,

$$\frac{1+1/(\delta n)}{2\sqrt{n\delta(1+\delta)}} \leqslant \frac{1}{\sqrt{1+1/n}} \leqslant 1.$$

Now, using (4) leads to (a stronger version of) the statement in the theorem. ∎

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

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

for $\delta < 1$, and

$$\bar{A}(D) \leqslant \hat{A}(D) \leqslant -\lg\lfloor 1+\delta\rfloor + \frac{1}{n}\lg\left(ne^{\lfloor 1+\delta\rfloor}\ln\lfloor 1+\delta\rfloor\right),$$

for $\delta \geqslant 1$.

Proof: For $\delta < 1$, we have

$$B_K(D) = B_K(\delta n) \geqslant \frac{1}{4}\binom{n+\delta n - 1}{\delta n}$$

$$\geqslant \frac{n}{4(n+\delta n)}\binom{n+\delta n}{\delta n}$$

$$\geqslant \frac{1}{4(1+\delta)} \cdot \frac{2^{n(1+\delta)H\left(\frac{1}{1+\delta}\right)}}{\sqrt{8n\delta/(1+\delta)}}$$

$$= \frac{1}{4} \cdot \frac{2^{n(1+\delta)H\left(\frac{1}{1+\delta}\right)}}{\sqrt{8n\delta(1+\delta)}} \geqslant \frac{2^{n(1+\delta)H\left(\frac{1}{1+\delta}\right)}}{16\sqrt{n}},$$

where the first inequality follows from (8) and the last step follows from the fact that $\delta \leqslant 1$, and so $\delta(1+\delta) \leqslant 2$.

Since $\frac{1+\ln x}{x}$ is a decreasing function for $x \geqslant 1$, we can substitute the lower bound on $B_K(D)$ in (3) to obtain
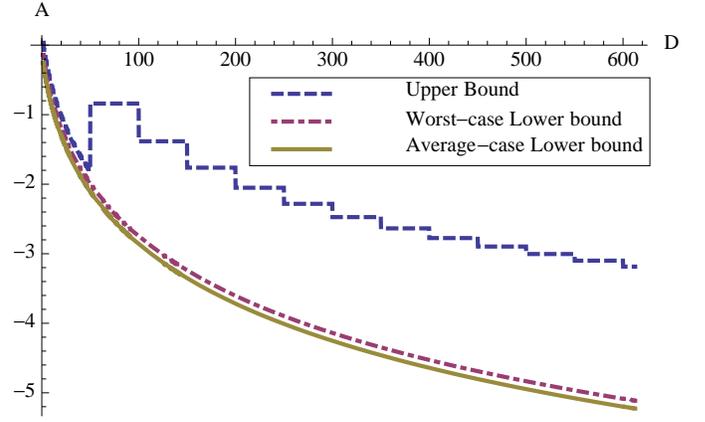
$$\hat{M}(D) \leqslant \frac{n!16\sqrt{n}}{2^{n(1+\delta)H\left(\frac{1}{1+\delta}\right)}}\ln\left(\frac{e2^{n(1+\delta)H\left(\frac{1}{1+\delta}\right)}}{16\sqrt{n}}\right)$$

$$\overset{(a)}{\leqslant} \frac{n!16n^{3/2}}{2^{n(1+\delta)H\left(\frac{1}{1+\delta}\right)}}(1+\delta)H\left(\frac{1}{1+\delta}\right)\ln 2$$

$$\overset{(b)}{\leqslant} \frac{n!64n^{3/2}}{2^{n(1+\delta)H\left(\frac{1}{1+\delta}\right)}},$$

where (a) follows from the fact $e \leqslant 16\sqrt{n}$ and (b) from the fact that for $\delta \leqslant 1$, we have $(1+\delta)H\left(\frac{1}{1+\delta}\right)\ln 2 \leqslant 2\ln 2 \leqslant 4$. Thus

$$\hat{A}(D) \leqslant -\lg\frac{(1+\delta)^{1+\delta}}{\delta^{\delta}} + \frac{3\lg n + 12}{2n}.$$

For $\delta \geqslant 1$, by (7) and (2) we have

$$B_K(D) = B_K(\delta n) \geqslant \lfloor 1+\delta\rfloor!\lfloor 1+\delta\rfloor^{n-\lfloor 1+\delta\rfloor} \geqslant \frac{\lfloor 1+\delta\rfloor^n}{e^{\lfloor 1+\delta\rfloor}},$$

A

Figure 1. Upper bound and lower bounds for $n = 50$ from Theorems 5 and 6.

implying

$$\hat{A}(D) \leqslant \frac{1}{n}\lg\frac{1+\ln B_K(\delta n)}{B_K(\delta n)}$$

$$\leqslant \frac{1}{n}\lg\frac{e^{\lfloor 1+\delta\rfloor}}{\lfloor 1+\delta\rfloor^n} + \frac{1}{n}\lg\left(1 + n\ln\lfloor 1+\delta\rfloor - \lfloor 1+\delta\rfloor\right)$$

$$\leqslant \frac{1}{n}\lg\frac{e^{\lfloor 1+\delta\rfloor}}{\lfloor 1+\delta\rfloor^n} + \frac{1}{n}\lg\left(n\ln\lfloor 1+\delta\rfloor\right)$$

$$\leqslant -\lg\lfloor 1+\delta\rfloor + \frac{1}{n}\lg\left(ne^{\lfloor 1+\delta\rfloor}\ln\lfloor 1+\delta\rfloor\right).$$

∎

The plots for the expressions given in Theorems 5 and 6 are given in Figure 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 Lemmas 5 and 6.

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

$$\hat{A}(D) = -\lg\frac{(1+\delta)^{1+\delta}}{\delta^{\delta}} + O\left(\frac{\lg n}{n}\right), \qquad (9)$$

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

Next, let us consider the case of $D = \Theta(n)$. From (5), it follows that

$$B_K(k) = \left[z^k\right]\frac{1}{1-z}\prod_{i=2}^{n}\frac{1-z^i}{1-z} = \left[z^k\right]\frac{\prod_{i=2}^{n}(1-z^i)}{(1-z)^n}.$$

Let

$$g(k,n) = \binom{n+k-1}{k}^{-1}B_K(k),$$

$$\gamma(z,n) = \sum_{i=0}^{\infty}\Gamma_i(n)z^i = \prod_{i=2}^{n}\left(1-z^i\right), \qquad (10)$$

and

$$f(z,n) = \sum_{i=0}^{\infty}F_i(n)z^i = \frac{1}{(1-z)^n},$$

where
$$F_i(n) = \binom{n+i-1}{i},$$

so that
$$g(k,n) = \frac{1}{F_k(n)} \left[z^k\right] f(z,n)\,\gamma(z,n).$$

We use the following theorem to find the asymptotics of $g(k,n)$ and $B_K(k)$ using the asymptotics of $\gamma(z,n)$ in Theorem 9.

Theorem 8. *[12, Theorem 3.1] Let $f(z,n)$ and $\gamma(z,n)$ be two functions with Taylor series for all n,*
$$f(z,n) = \sum_{i=0}^{\infty} F_i(n)\,z^i, \quad \gamma(z,n) = \sum_{i=0}^{\infty} \Gamma_i(n)\,z^i,$$

where $F_i(n) > 0$ for all sufficiently large n. Suppose
$$g(k,n) = \frac{1}{F_k(n)} \left[z^k\right] f(z,n)\,\gamma(z,n),$$

and let $n = n(k)$ be a function of k such that the limit $\rho = \lim_{k\to\infty} \frac{F_{k-1}(n(k))}{F_k(n(k))}$ exists. We have
$$g(k, n(k)) \sim \gamma(\rho, n(k)) \text{ as } k \to \infty,$$

provided that

1) *for all sufficiently large k and for all i,*
$$\left| \frac{\Gamma_i(n(k))}{\gamma(\rho, n(k))} \right| \le p_i,$$
   where $\sum_{i=0}^{\infty} p_i \rho^i < \infty$, and
2) *there exists a constant c, such that for all sufficiently large $i \le k$ and large k,*
$$\left| \frac{F_{k-i}(n(k))}{F_k(n(k))} \right| \le c\rho^i.$$

Theorem 9. *Let $n = n(k) = \frac{k}{c} + O(1)$ for a constant $c > 0$. Then*
$$B_K(k) \sim K_c \binom{n+k-1}{k} \tag{11}$$

as $k, n \to \infty$, where K_c is a positive constant equal to $\lim_{n\to\infty} \gamma(c/(1+c), n)$.

Proof: To prove the theorem, we use Theorem 8. To do this, we first let
$$\rho = \lim_{k\to\infty} \frac{\binom{n(k)+k-2}{k-1}}{\binom{n(k)+k-1}{k}} = \lim_{k\to\infty} \frac{k}{n(k)+k-1} = \frac{c}{1+c}.$$

We now turn our attention to Condition 1 of Theorem 8. First, we show that $\gamma(\rho, n(k))$ is bounded away from 0. We have
$$\ln \gamma(\rho, n(k)) \ge \sum_{i=2}^{\infty} \ln\left(1-\rho^i\right) \ge -\sum_{i=2}^{\infty} \frac{\rho^i}{1-\rho^i}$$
$$\ge -\sum_{i=2}^{\infty} \frac{\rho^i}{1-\rho} = -\frac{\rho^2}{(1-\rho)^2},$$

where the second inequality follows from the fact that
$$\ln(1-x) = -\sum_{i=1}^{\infty} \frac{x^i}{i} \ge -\sum_{i=1}^{\infty} x^i = \frac{x}{1-x},$$

for $0 < x < 1$. Hence,
$$\gamma(\rho, n(k)) \ge e^{-\left(\frac{\rho}{1-\rho}\right)^2} > 0.$$

To satisfy Condition 1 of Theorem 8, it thus suffices to find p_i' such that $|\Gamma_i(n(k))| \le p_i'$ and $\sum_{i=0}^{\infty} p_i' \rho^i < \infty$. For all positive integers m, we have
$$|\Gamma_i(m)| = \left| \left[z^i\right] \prod_{j=2}^{m} \left(1-z^j\right) \right| \le \left| \left[z^i\right] \prod_{j=2}^{m} \left(1+z^j\right) \right|$$
$$\le \left| \left[z^i\right] \prod_{j=1}^{\infty} \left(1+z^j\right) \right| < e^{\pi\sqrt{2/3}\sqrt{i}},$$

where the last inequality follows from the facts that $\prod_{j=1}^{\infty}\left(1+z^j\right)$ is the generating function for the number of partitions of a positive integer into distinct parts and that the number of partitions of a positive integer i is bounded by $e^{\pi\sqrt{2/3}\sqrt{i}}$ [1, p. 316].

We let $p_i' = e^{\pi\sqrt{2/3}\sqrt{i}}$ and apply the ratio test to the sum $\sum_{i=0}^{\infty} p_i' \rho^i$ to prove its convergence. Since
$$\lim_{i\to\infty} \left(p_i' \rho^i\right)^{1/i} = \lim_{i\to\infty} e^{\pi\sqrt{2/3}/\sqrt{i}}\rho < 1,$$

the sum converges and Condition 1 of Theorem 8 is satisfied. Hence,
$$\frac{B_K(k)}{\binom{n+k-1}{k}} \sim \gamma\left(\frac{c}{1+c}, n\right).$$

To complete the proof, we show that the limit $\lim_{n\to\infty} \gamma(c/(1+c), n)$ exists and is positive. This is evident as $\gamma(c/(1+c), n)$ is decreasing and, as shown before, bounded away from 0. ∎

For $D = cn + O(1)$, we have
$$\frac{1}{n} \lg B_K(D) = \frac{1}{n} \lg \binom{n+D-1}{D} + O\left(\frac{1}{n}\right)$$
$$= \frac{n+cn+O(1)}{n} H\left(\frac{c}{1+c} + O\left(\frac{1}{n}\right)\right)$$
$$\quad + O\left(\frac{\lg n}{n}\right)$$
$$= (1+c) H\left(\frac{c}{1+c}\right) + O\left(\frac{1}{n}\right) + O\left(\frac{\lg n}{n}\right)$$
$$= (1+c) H\left(\frac{c}{1+c}\right) + O\left(\frac{\lg n}{n}\right),$$

where we used (11) for the first step. Using (3), for $D = cn + O(1)$, we find
$$\hat{A}(D) \ge -\frac{1}{n} \lg B_K(D) = (1+c) H\left(\frac{c}{1+c}\right) + O\left(\frac{\lg n}{n}\right)$$

and

$$\hat{A}(D) \leqslant -\frac{1}{n} \lg B_K(D) + \frac{1}{n} \lg \left(1 + \ln B_K(D)\right)$$
$$= -(1+c)H\left(\frac{c}{1+c}\right) + O\left(\frac{\lg n}{n}\right)$$

The derivation for $\bar{A}(cn + O(1))$ is similar. We thus have the following lemma.

Lemma 10. *For a constant $c > 0$ and $D = cn + O(1)$, we have*

$$\hat{A}\left(cn + O\left(1\right)\right) = -\lg \frac{(1+c)^{1+c}}{c^c} + O\left(\frac{\lg n}{n}\right), \quad (12)$$

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

The results given in (9) and (12) 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 $\bar{A}(D)$ satisfies the same relations.

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

$$\hat{A}(D) \sim -\lg n^\alpha,$$

In this subsection, we improve upon this result by providing upper and lower bound with error terms.

Lemma 11. *For $D = cn^{1+\alpha} + O(n)$, where α and c are constants such that $0 < \alpha < 1$ and $c > 0$, we have*

$$-\lg (ecn^\alpha) + O\left(n^{-\alpha}\right) \leqslant \hat{A}(D)$$
$$\leqslant -\lg (cn^\alpha) + O\left(n^{-\alpha} + n^{\alpha-1}\right)$$
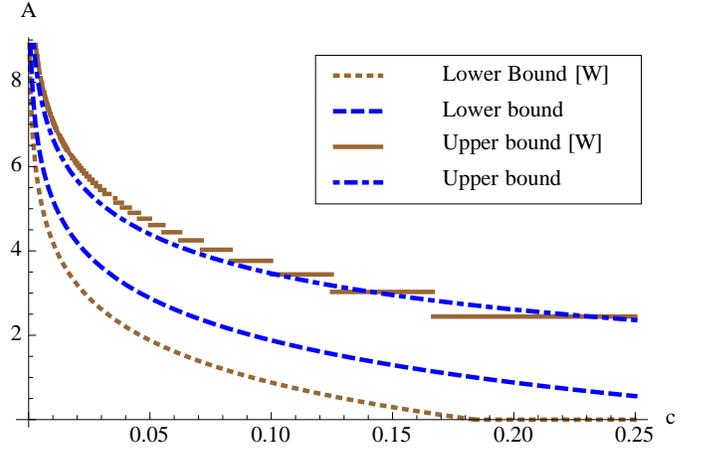
Proof: Note that from Theorem 5, we have

$$\hat{A}(D) \geqslant -\lg \frac{(1+\delta)^{1+\delta}}{\delta^\delta} = \lg \frac{1}{1+\delta} + \lg \left(1 + \frac{1}{\delta}\right)^{-\delta}$$
$$\geqslant -\lg (e(1+\delta)).$$

Let $\delta = D/n = cn^\alpha + O(1)$. We find

$$\hat{A}(D) \geqslant -\lg (e(1+\delta)) = -\lg e - \lg (cn^\alpha + O(1))$$
$$= -\lg (ecn^\alpha) + O\left(n^{-\alpha}\right).$$

On the other hand, from Theorem 6,

$$\hat{A}(D) \leqslant -\lg (cn^\alpha + O(1)) + \frac{1}{n} \lg e^{O(n^\alpha)}$$
$$= -\lg (cn^\alpha) + O\left(n^{-\alpha} + n^{\alpha-1}\right)$$

■

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

D. Large Distortion

In the large distortion regime, we have $D = cn^2 + O(n)$ and $\delta = cn + O(1)$.

Lemma 12. *Suppose $D = cn^2 + O(n)$ for a constant $0 < c < \frac{1}{2}$. We have*

$$-\lg (ecn) + O\left(\frac{1}{n}\right) \leqslant \hat{A}(D) \leqslant$$
$$-\lg (ecn) + (1+c)\lg e + O\left(\frac{\lg n}{n}\right).$$

Proof: Let $\delta = cn + O(1)$. Similar to the proof of the lower bound in Lemma 11, we have $\hat{A}(D) \geqslant -\lg (e(1+\delta))$, and thus

$$\hat{A}(D) \geqslant -\lg (ecn + O(1)) \geqslant -\lg (ecn) + O\left(\frac{1}{n}\right).$$

On the other hand, from Theorem 6,

$$\hat{A}(D) \leqslant -\lg (ecn) + (1+c)\lg e + O\left(\frac{\lg n}{n}\right).$$

■

From [17], we have

$$-\lg (ecn) - 1 + O\left(\frac{\lg n}{n}\right) \leqslant \hat{A}(D) \leqslant$$
$$-\lg \frac{n}{e \lceil 1/(2c)\rceil} + O\left(\frac{\lg n}{n}\right). \quad (13)$$

These bounds are compared in Figure 2, 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,

$$\text{per}(A) = \sum_{\pi \in S_n} \prod_{i=1}^{n} a_{i,\pi(i)}.$$

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

$$A_{i,j} = \begin{cases} 1 & |i - j| \leqslant r \\ 0 & \text{otherwise.} \end{cases} \quad (14)$$

According to Brégman's Theorem (see [4]), for any $n \times n$ binary matrix A with r_i 1's in the i-th row

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

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

Lemma 13. *[16] For all* $0 \leqslant r \leqslant n - 1$,

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

The following lower bound was given in [10].

Lemma 14. *[10] For all* $0 \leqslant r \leqslant \frac{n-1}{2}$,

$$\mathsf{B_C}(r) \geqslant \frac{(2r+1)^n}{2^{2r}} \frac{n!}{n^n}.$$

We extend this lemma to the full range of parameters.

Lemma 15. *For all* $0 \leqslant r \leqslant n - 1$,

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

Proof: Only the second claim requires proof, so suppose that $(n-1)/2 \leqslant r \leqslant n - 1$. The proof follows the same lines as the one appearing in [10]. Let A be defined as in (14), and let B be an $n \times n$ matrix with

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

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

$$\mathsf{B_C}(r) = \text{per}(A) \geqslant \frac{\text{per}(B)}{2^{2(n-r)}} \geqslant \frac{n^n}{2^{2(n-r)}} \text{per}\left(\frac{B}{n}\right)$$

$$\geqslant \frac{n!}{2^{2(n-r)}},$$

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

Theorem 16. *Let* $n \in \mathbb{N}$, *and let* $0 < \delta < 1$ *be a constant rational number such that* δn *is an integer. Then*

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

and

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

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

Proof: First, we prove the lower bound for $\hat{\mathsf{R}}_C(D)$ using Theorem 4, which states $\hat{\mathsf{M}}_C(D) \geqslant n!/\mathsf{B_C}(D)$, and Lemma 13. Let

$$T_1 = ((2D+1)!)^{(n-2D)/(2D+1)},$$

$$T_2 = \prod_{i=D+1}^{2D} (i!)^{2/i}.$$

We have

$$\lg T_1 = \frac{n - 2\delta n}{2\delta n + 1} \lg(2\delta n + 1)!$$

$$= \frac{n - 2\delta n}{2\delta n + 1}\left((2\delta n + 1)\lg\left(\frac{2\delta n + 1}{e}\right) + O(\lg n)\right)$$

$$= (n - 2\delta n)\lg\left(\frac{2\delta n + 1}{e}\right) + O(\lg n)$$

$$= (n - 2\delta n)\lg(2\delta n/e) + O(\lg n),$$

and

$$\lg T_2 = 2\sum_{i=\delta n+1}^{2\delta n} \frac{1}{i}\lg i! = 2\sum_{i=\delta n+1}^{2\delta n}\left(\lg\frac{i}{e} + O\left(\frac{\lg i}{i}\right)\right)$$

$$= 2\sum_{i=\delta n+1}^{2\delta n} \lg i - 2\delta n \lg e + O(\lg n)$$

$$= 2\lg\frac{(2\delta n)!}{(\delta n)!} - 2\delta n \lg e + O(\lg n)$$

$$= 2\delta n + 2\delta n \lg(2\delta n/e) - 2\delta n \lg e + O(\lg n).$$

From these expressions and Lemma 13, it follows that

$$\frac{1}{n}\lg\mathsf{B_C}(D) \leqslant \lg(2\delta n/e) + 2\delta\lg(2/e) + O(\lg n/n).$$

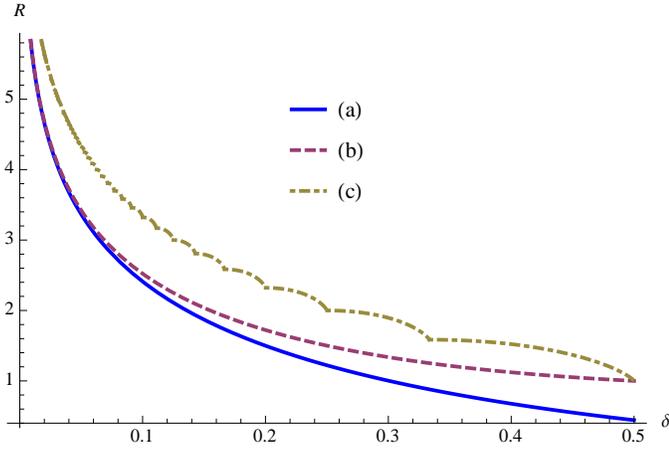The lower bound for $0 < \delta \leqslant 1/2$ then follows from Theorem 4. The proof of the lower bound for $1/2 < \delta \leqslant 1$ is similar.

Next, we prove the upper bound for $\hat{\mathsf{R}}_C(D)$. From Lemma 2, we have

$$\hat{\mathsf{M}}_C(D) \leqslant \frac{n! \ln n!}{\mathsf{B_C}(D)}.$$

Hence, for $0 \leqslant D \leqslant (n-1)/2$,

$$\hat{\mathsf{R}}_C(D) \leqslant \frac{1}{n}\lg\left(\frac{n! \ln n!}{\mathsf{B_C}(\delta n)}\right)$$

$$\leqslant \frac{1}{n}\lg\left(\frac{2^{2\delta n} n^n}{(2\delta n + 1)^n}\right) + O\left(\frac{\lg n}{n}\right)$$

$$\leqslant \lg\frac{1}{2\delta} + 2\delta + + O\left(\frac{\lg n}{n}\right) \quad (15)$$

where we have used Lemma 15 for the second inequality.

Figure 3. Rate-distortion in the Chebyshev metric: The lower and upper bounds of Theorem 16, (a) and (b), and the rate of the code construction, given in Theorem 19, (c).

Similarly, for $(n-1)/2 < D \leqslant n$,

$$\hat{R}_C(D) \leqslant \frac{1}{n}\lg 2^{2n(1-\delta)} + O\left(\frac{\lg n}{n}\right)$$

$$\leqslant 2(1-\delta) + O\left(\frac{\lg n}{n}\right). \tag{16}$$

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 $\bar{R}_C(D) \leqslant \hat{R}_C(D)$. ∎

B. Code Construction

Let $A = \{a_1, a_2, \ldots, a_m\} \subseteq [n]$ be a subset of indices, $a_1 < a_2 < \cdots < a_m$. For any permutation $\sigma \in S_n$ we define $\sigma|_A$ to be the permutation in S_m that preserves the relative order of the sequence $\sigma(a_1), \sigma(a_2), \ldots, \sigma(a_m)$. 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

$$\sigma|^A = \left(\sigma^{-1}|_A\right)^{-1}.$$

Intuitively, to calculate $\sigma|^A$ we keep only the *values* of σ from A, and then relabel the entries to $[m]$ while keeping relative order.

Example 17. 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, $1 \leqslant d \leqslant n - 1$. Furthermore, we define the sets

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

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

$$C = \left\{\sigma \in S_n \mid \sigma|^{A_i} = \mathrm{Id} \text{ for all } i\right\}.$$

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

Theorem 18. Let n and d be positive integers, $1 \leqslant d \leqslant n - 1$. Then the code $C \subseteq S_n$ 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: Let $\sigma \in S_n$ 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

$$\sigma'|^{A_i} = \mathrm{Id},$$

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

$$|\sigma(j) - \sigma'(j)| \leqslant d$$

and so

$$d_C(\sigma, \sigma') \leqslant d.$$

Finally, we contend the permutation $\sigma = [n, n-1, \ldots, 1]$ is at distance exactly d from the code C. We note we already know that there is a codeword $\sigma' \in C$ such that $d_C(\sigma, \sigma') \leqslant d$. We now show there is no closer codeword in C. Let us attempt to build such a permutation σ''. Consider $\sigma(n) = 1$. The value of $\sigma''(n)$ is in A_i for some i, and since σ'' is a codeword, $\sigma''(n)$ must be the largest in A_i. Thus

$$\sigma''(n) \in \{\max(A_i) \mid 1 \leqslant i \leqslant \lceil n/(d+1) \rceil\} \geqslant d+1.$$

It follows that

$$|\sigma''(n) - \sigma(n)| = d$$

and so

$$d_C(\sigma, \sigma'') \geqslant d.$$

∎

The code construction has the following asymptotic form:

Theorem 19. *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 16 and the rate of the code construction, given in Theorem 19, are shown in Figure 3.

REFERENCES

[1] T. M. Apostol, *Introduction to Analytic Number Theory*. Springer-Verlag, NY, 1976.

[2] 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.

[3] M. Bóna, *Combinatorics of permutations*. CRC Press, 2012.

[4] L. M. Brégman, "Some properties of nonnegative matrices and their permanents," *Soviet Math. Dokl.*, vol. 14, pp. 945–949, 1973.

[5] G. Cohen, I. Honkala, S. Litsyn, and A. Lobstein, *Covering Codes*. North-Holland, 1997.

[6] M. Deza and H. Huang, "Metrics on permutations, a survey," *J. Comb. Inf. Sys. Sci.*, vol. 23, pp. 173–185, 1998.

[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 *Proceedings of the 2013 IEEE International Symposium on Information Theory (ISIT2013), Istanbul, Turkey*, Jul. 2013, pp. 2562–2566.