Recursive Methods for Some Problems in Coding and Random Permutations

Ghurumuruhan Ganesan[✉]

Institute of Mathematical Sciences, HBNI, Chennai, India
gganesan82@gmail.com

Abstract. In this paper, we study three applications of recursion to problems in coding and random permutations. First, we consider locally recoverable codes with partial locality and use recursion to estimate the minimum distance of such codes. Next we consider weighted lattice representative codes and use recursive subadditive techniques to obtain convergence of the minimum code size. Finally, we obtain a recursive relation involving cycle moments in random permutations and as an illustration, evaluate recursions for the mean and variance.

Keywords: Locally recoverable codes · Partial locality · Minimum distance · Lattice identification codes · Minimum size · Random permutations · Cycle moments

1 Introduction

Recursive techniques are used quite frequently in coding to obtain bounds on code sizes. As a typical example, the Singleton bound [9] obtains bounds on sizes of $n-$length codes by reducing the problem to that of an $(n-1)-$length code. Similarly, recursive relations are also frequent in terms related to permutations like for example, Stirling numbers of the first kind [8]. In this paper, we study further applications of recursive methods to problems in coding and random permutations.

The paper is organized as follows: In Sect. 2, we consider locally recoverable codes with partial locality and estimate the minimum distance of such codes (Theorem 1) using iterations on subcodes. Next, in Sect. 3, we study lattice representative codes with weights and prove asymptotic convergence of the minimum size, using subadditive techniques (Theorem 2). Finally, in Sect. 4, we establish a recursion for cycle moments of random permutations (Theorem 3) and illustrate our result for the cases of mean and variance (Corollary 1).

2 Locally Recoverable Codes with Partial Locality

Locally recoverable codes for erasures have tremendous applications in distributed storage and retrieval [12] and it is therefore important to understand the

properties of such codes. Typically each erasure correction is performed using a locality set of small size and it is of interest to design codes capable of correcting multiple erasures simultaneously. Such codes are also known as locally repairable codes and storage-bandwidth tradeoff and construction of such codes has been well-studied; for an overview we refer to the recent survey [2]. For distinction, we refer to codes above as *fully* locally recoverable codes since *every* symbol position has a locality set of small size associated with it. In [7], bounds are obtained for the minimum distance of linear fully locally recoverable codes in terms of the size of the locality sets. Later [5] studied bounds on the minimum distance of non-linear systematic fully locally recoverable codes.

In this section, we study minimum distance of locally recoverable codes with partial locality. We assume that only a *subset* of symbol positions have locality set size at most r and obtain bounds on the minimum distance. Let $n \geq k \geq 1$ be integers and let $\mathcal{A}$ be a set of cardinality $\#\mathcal{A} = q$. A set $\mathcal{C} \subseteq \mathcal{A}^n$ of cardinality q^k is defined to be an $(n, k)-$code.

For a set $\mathcal{U} \subseteq \{1, 2, \ldots, n\}$ and an integer $j \in \{1, 2, \ldots, n\} \setminus \mathcal{U}$, we say that position j is determined by $\mathcal{U}$ if there exists a function g_j such that

$$c_j = g_j \left(c_i : i \in \mathcal{U} \right) =: g_j \left(\mathcal{U} \right) \tag{2.1}$$

for all codewords $\mathbf{c} = (c_1, \ldots, c_n) \in \mathcal{C}$. In words, the symbol at the jth position of any codeword can be determined from the symbols with positions in $\mathcal{U}$. The set $\mathcal{F}_\mathcal{U}$ of all positions determined by $\mathcal{U}$ is called the *reach* of $\mathcal{U}$. For integer $1 \leq w \leq n$, we define

$$L(w) = L(w, \mathcal{C}) := \max_{\mathcal{U}:\#\mathcal{U}=w} \#\mathcal{F}_\mathcal{U}. \tag{2.2}$$

For any $w \geq 1$ we have that $L(w, \mathcal{C}) \leq \Delta(w)$, where $\Delta(w) = q^w$ if $\mathcal{C}$ is a linear code and $\Delta(w) = q^{q^w}$ otherwise. We remark here that q^{q^w} is the *total* number of maps from $\mathcal{A}^w$ to $\mathcal{A}$.

Definition 1. *For integers $\tau, r \geq 2$ and a subset $\Theta \subseteq \{1, 2, \ldots, n\}$, we say that the code $\mathcal{C}$ has $(\Theta, \tau, r)-$local correction capability if for every subset $\mathcal{P} \subseteq \Theta$ of size τ, there exists a set $\mathcal{T}_\mathcal{P} \subseteq \{1, 2, \ldots, n\} \setminus \mathcal{P}$ of size at most r such that each position in $\mathcal{P}$ is determined by $\mathcal{T}_\mathcal{P}$.*

We define $\mathcal{T}_\mathcal{P}$ to be the $r-$locality set corresponding to the set $\mathcal{P}$. Also if $\Theta = \{1, 2, \ldots, n\}$, we say that $\mathcal{C}$ has $(\tau, r)-$local correction capability.

For example, consider the binary linear code $\mathcal{C}$ formed in the following way: For $k \geq 10$ and a word $(c_1, \ldots, c_k) \in \{0, 1\}^k$, let $d_{i_1, i_2, i_3} := c_{i_1} \oplus c_{i_2} \oplus c_{i_3}$ where $\oplus$ denotes addition modulo two. There are $\binom{k}{3} =: n - k - 1$ such terms $\{d_{i_1, i_2, i_3}\}$ which we relabel as $c_{k+1}, \ldots, c_{n-1}$. Finally we let $c_n := \oplus_{i=1}^{k} c_i$. We let $(c_1, \ldots, c_n)$ be the codeword corresponding to the word $(c_1, \ldots, c_k)$. The collection of codewords $\mathcal{C}$ has $(\Theta, \tau, r)-$local correction capability with $\Theta = \{1, 2, \ldots, n-1\}, \tau = 1$ and $r = 3$. For example to recover c_1, we use the relation

$$d_{1,2,3} \oplus d_{1,2,4} \oplus d_{1,3,4} = c_1.$$

In general, each bit $c_j, 1 \leq j \leq k$ can be recovered in a similar manner. Because $k \geq 10$, the bit c_n cannot be recovered by using any three bits of $\{c_1, \ldots, c_{n-1}\}$.

Let $\mathcal{C}$ be any code with (Θ, τ, r)−local correction capability. For words $\mathbf{x} = (x_1, \ldots, x_n)$ and $\mathbf{y} = (y_1, \ldots, y_n)$ in $\mathcal{C}$ we define the Hamming distance between $\mathbf{x}$ and $\mathbf{y}$ to be $d(\mathbf{x}, \mathbf{y}) := \sum_{i=1}^{n} \mathbb{1}(x_i \neq y_i)$, where $\mathbb{1}(.)$ denotes the indicator function. The minimum distance of $\mathcal{C}$ is then defined as $d(\mathcal{C}) := \min_{\mathbf{x}, \mathbf{y} \in \mathcal{C}} d(\mathbf{x}, \mathbf{y})$. We have the following result.

Theorem 1. *Let $\mathcal{C}$ be any (n, k)−code with (Θ, τ, r)−parallel correction capability and let $\theta = \#\Theta$. The minimum distance of $\mathcal{C}$ satisfies*

$$d(\mathcal{C}) \leq n - k + 1 - T \cdot \tau, \qquad (2.3)$$

where T is the largest integer t such that

$$t \cdot r \leq k - 1 + \theta - n \text{ and } t \cdot r + \Delta(t \cdot r) \leq \theta - \tau + 1. \qquad (2.4)$$

To obtain the bound (2.4), we proceed as in [5] and iteratively construct a sequence of codes with decreasing size, until no further reduction is possible. We use the pigeonhole principle at the end of each step and obtain the sufficient conditions that allow continuation of the iteration procedure. In the proof below, we see that the first estimate in (2.4) determines the maximum number of iterations the procedure can proceed before we run out of codewords to choose from and the second estimate in (2.4) determines the maximum number of iterations for which we are able to choose a "fresh" locality set.

Finally, we recall that $n - k + 1$ is the Singleton bound [9] and is the maximum possible minimum distance of an (n, k)−code. Therefore the parameter T is in some sense, the "cost" for requiring partial locality.

Proof of Theorem 1. Let $\mathcal{P}_1 \subseteq \Theta$ be any set of size τ and let $\mathcal{I}_1 := \mathcal{T}_{\mathcal{P}_1} = \{l_1, \ldots, l_{m_1}\}, m_1 \leq r$ be the corresponding locality set of cardinality at most r as defined in the paragraph following (2.2) that determines the value of the symbols in positions in $\mathcal{P}_1$. For $\mathbf{x} = (x_1, \ldots, x_{m_1}) \in \mathcal{A}^{m_1}$, let $\mathcal{C}(\mathbf{x}) = \mathcal{C}(\mathbf{x}, \mathcal{I}_1)$ be the set of codewords of $\mathcal{C}$ such that the symbol in position l_j equals x_j for $1 \leq j \leq m_1$.

The number of choices for $\mathbf{x}$ is at most q^{m_1} and there are q^k codewords in $\mathcal{C}$. Therefore by pigeonhole principle, there exists $\mathbf{x}_1$ such that

$$\#\mathcal{C}(\mathbf{x}_1) \geq \frac{\#\mathcal{C}}{q^{m_1}} = q^{k - m_1}. \qquad (2.5)$$

We set $\mathcal{C}_1 := \mathcal{C}(\mathbf{x}_1)$ and let $\mathcal{J}_1 := \mathcal{F}_{\mathcal{I}_1}$ be the reach of $\mathcal{I}_1$ (see (2.1)) with cardinality $\tau \leq \#\mathcal{J}_1 \leq \Delta(r)$. The first inequality is true since $\mathcal{P}_1 \subseteq \mathcal{J}_1$ and the second estimate follows from (2.1). By construction, all words in the code $\mathcal{C}_1$ have the same values in the symbol positions determined by $\mathcal{J}_1$; i.e., if $a = (a_1, \ldots, a_n)$ and $b = (b_1, \ldots, b_n)$ both belong to $\mathcal{C}_1$, then $a_j = b_j$ for all $j \in \mathcal{J}_1$.

We now repeat the above procedure with the code $\mathcal{C}_1$ assuming that $r + n - \theta < k$, where $\theta := \#\Theta$. If $\mathcal{R}_1 := \Theta \setminus (\mathcal{I}_1 \cup \mathcal{J}_1)$ is the set of positions not encountered in the first iteration then

$$\#\mathcal{R}_1 \geq \#\Theta - \#\mathcal{J}_1 - \#\mathcal{I}_1 \geq \theta - \Delta(r) - r \tag{2.6}$$

since $\#\mathcal{J}_1 \leq \Delta(r)$. For a set $\mathcal{P} \subseteq \mathcal{R}_1$ of size τ, let $\mathcal{I}(\mathcal{P}) := \mathcal{T}_{\mathcal{P}} \bigcap \mathcal{R}_1$ be the union of positions within the locality sets of the selected τ positions in $\mathcal{P}$, not encountered in the first iteration.

Suppose for every $\mathcal{P} \subseteq \mathcal{R}_1$, we have $\mathcal{I}(\mathcal{P}) = \emptyset$. This means that all symbols with positions in $\mathcal{R}_1$ can simply be determined by the symbol values with positions in $\mathcal{I}_1 \cup \mathcal{J}_1$. This in turn implies that the symbols with positions in $\mathcal{I}_1$ determine *all* the symbols with positions in Θ. Using $r + n - \theta < k$ we then get that the total number of words in the code $\mathcal{C}$ is at most

$$q^{\#\mathcal{I}_1} \cdot q^{n - \#\Theta} = q^{m_1 + n - \theta} \leq q^{r + n - \theta} < q^k,$$

a contradiction. Thus there exists $\mathcal{P}_2 \subseteq \mathcal{R}_1$ of size τ whose corresponding set $\mathcal{I}_2 := \mathcal{I}(\mathcal{P}_2)$ is not completely contained in $\mathcal{I}_1 \cup \mathcal{J}_1$.

Letting $1 \leq m_2 \leq r$ denote the cardinality of $\mathcal{I}_2$ and using the pigeonhole principle as before, we get a code $\mathcal{C}_2 \subseteq \mathcal{C}_1$ of size

$$\#\mathcal{C}_2 \geq \frac{\#\mathcal{C}_1}{q^{m_2}} \geq q^{k - m_1 - m_2} \geq q^{k - 2r}$$

and all of whose words have the same symbol values in the positions determined by $\mathcal{I}_1 \cup \mathcal{I}_2$. In the above, we use the estimate for $\mathcal{C}_1$ obtained in (2.5). As before, let $\mathcal{J}_2 \subseteq \{1, 2, \ldots, n\}$ be the set of positions of the codeword symbols determined by the set $\mathcal{I}_1 \cup \mathcal{I}_2$ so that $\mathcal{P}_1 \cup \mathcal{P}_2 \subseteq \mathcal{J}_2$. The set $\mathcal{I}_1 \cup \mathcal{I}_2$ has cardinality at most $2r$ and so we have from (2.2) that the reach $\mathcal{J}_2$ has cardinality

$$2\tau \leq \#\mathcal{P}_1 + \#\mathcal{P}_2 \leq \#\mathcal{J}_2 \leq \Delta(2r).$$

If $\mathcal{R}_2 := \Theta \setminus (\mathcal{I}_1 \cup \mathcal{I}_2 \cup \mathcal{J}_2)$, then $\#\mathcal{R}_2 \geq \theta - 2r - \Delta(2r)$. Continuing this way, after the end of t iterations, we have a code $\mathcal{C}_t$ of size

$$\#\mathcal{C}_t \geq q^{k - \sum_{j=1}^{t} \#\mathcal{I}_j} \geq q^{k - t \cdot r} \tag{2.7}$$

and a set $\mathcal{R}_t \subseteq \Theta$ of remaining positions not fixed so far, with cardinality

$$\#\mathcal{R}_t \geq \theta - \sum_{j=1}^{t} \#\mathcal{I}_j - \#\mathcal{J}_t \geq \theta - t \cdot r - \Delta(t \cdot r). \tag{2.8}$$

The above procedure can therefore be performed for at least T steps where T is the largest integer t such that

$$t \cdot r \leq k - 1 + \theta - n \text{ and } t \cdot r + \Delta(t \cdot r) \leq \theta - \tau + 1. \tag{2.9}$$

The first condition in (2.9) ensures that $k - T \cdot r \geq 1$ and so the code $\mathcal{C}_T$ has at least two codewords. The second condition in (2.9) ensures that the set $\mathcal{P}_j \subseteq \Theta$ of symbols we pick is at least τ and so

$$\#\mathcal{J}_j \geq \sum_{l=1}^{j} \#\mathcal{P}_l \geq j \cdot \tau \tag{2.10}$$

for each $1 \leq j \leq T$.

Since $\mathcal{C}_T \subseteq \mathcal{C}$, the minimum distance $d(\mathcal{C}_T)$ of $\mathcal{C}_T$ is at least the minimum distance $d(\mathcal{C})$ of $\mathcal{C}$. By definition we recall that the symbol values in positions determined by the set $\bigcup_{1 \leq j \leq T} \mathcal{I}_j \bigcup \mathcal{J}_T := \{1, 2, \ldots, n\} \setminus \mathcal{Q}_T$ is the same for all the words in $\mathcal{C}_T$. For every word $\mathbf{x} = (x_1, \ldots, x_n) \in \mathcal{C}_T$, we therefore let $\mathbf{x}_T = (x_i)_{i \in \mathcal{Q}_T}$ be the reduced word obtained by just considering the symbols in the remaining positions determined by $\mathcal{Q}_T$. Defining the reduced code $\mathcal{D}_T = \{\mathbf{x}_T : \mathbf{x} \in \mathcal{C}_T\}$ we then have that the minimum distance $d(\mathcal{D}_T) \geq d(\mathcal{C}_T) \geq d(\mathcal{C})$.

The length of the each word in $\mathcal{D}_T$ equals $n - \#\mathcal{Q}_T$ and so using the estimate for $\#\mathcal{D}_T = \#\mathcal{C}_T$ from (2.7) and the Singleton bound we have

$$d(\mathcal{D}_T) \leq \left(n - \sum_{j=1}^{T} \#\mathcal{I}_j - \#\mathcal{J}_T \right) - \left(k - \sum_{j=1}^{T} \#\mathcal{I}_j \right) + 1.$$

Thus $d(\mathcal{D}_T) \leq n - k + 1 - \#\mathcal{J}_T \leq n - k + 1 - T \cdot \tau$, by (2.10). Using the fact that $d(\mathcal{C}) \leq d(\mathcal{D}_T)$, we then get (2.3). $\qquad \square$

3   Lattice Representative Codes

Representative codes [10] (also known as hitting sets in some contexts) are important from both theoretical and application perspectives. In [11] the minimum size of hitting sets that intersect all combinatorial rectangles of a given volume are studied. Explicit constructions were described using expander graphs and random walks. Later [14] determined lower bounds for the hitting set size of combinatorial rectangles and also illustrated an application in approximation algorithms. Recently [4] used fractional perfect hash families to study construction of explicit hitting sets for combinatorial shapes.

In this section, we study lattice representative codes for *weighted* rectangles where each vertex is assigned a positive finite weight. We study the minimum size of a representative code that intersects *all* subsets of a given minimum weight. For integers $d, m \geq 1$ let $\mathcal{S} = \mathcal{S}(m) := \{1, 2, \ldots, m\}^d$. We refer to elements of $\mathcal{S}$ as points and for a point $v = (v_1, \ldots, v_d) \in \mathcal{S}$, we refer to v_i as the ith entry. For each point $v \in \mathcal{S}$, we assign a finite positive weight $w(v)$. For a set $\mathcal{U} \subseteq \mathcal{S}$ the corresponding weight $w(\mathcal{U}) := \sum_{v \in \mathcal{U}} w(v)$ is the sum of the weights of the points in $\mathcal{U}$. The size of $\mathcal{U}$ is the number of points in $\mathcal{U}$ and is denoted by $\#\mathcal{U}$.

Let

$$1 = \inf_{m} \min_{v \in \mathcal{S}(m)} w(v) \leq \sup_{m} \max_{v \in \mathcal{S}(m)} w(v) =: \beta < \infty \tag{3.1}$$

and for $0 < \epsilon < 1$ say that $\mathcal{B} \subseteq \mathcal{S}$ is an $\epsilon-$representative code or simply representative code if $\mathcal{B} \cap \mathcal{U} \neq \emptyset$ for any set $\mathcal{U} \subseteq \mathcal{S}$ of weight $w(\mathcal{U}) \geq \epsilon \cdot m^d$.

The following result obtains an estimate on the minimum size b_m of an $\epsilon-$representative code.

Theorem 2. *For any $0 < \epsilon < 1$ and $\beta \geq 1$ we have that*

$$m^d \cdot (1 - \epsilon) \leq b_m \leq m^d \cdot \left(1 - \frac{\epsilon}{\beta} \right) + 1. \tag{3.2}$$

Suppose the weight function satisfies the following monotonicity relation: If u and v are any two points of S differing only in the ith entry and $u_i > v_i$, then the weights $w(u) \leq w(v)$. We then have

$$\frac{b_m}{m^d} \longrightarrow \lambda \qquad (3.3)$$

as $m \to \infty$ where $1 - \epsilon \leq \lambda \leq 1 - \frac{\epsilon}{\beta}$.

Thus there exists a *fraction* of vertices in a large rectangle that hits *all* sets of a given minimum weight. Moreover, if the weight assignment is monotonic, then the scaled minimum representative code size converges to a positive constant strictly between 0 and 1.

An example of non-trivial weight assignment that satisfies the monotonicity relation is the following: Defining $w(1,1) := 2$ we iteratively assign the weight of each vertex in the set $\{1, 2, \ldots, i+1\}^d \setminus \{1, 2, \ldots, i\}^d$ as $1 + \frac{1}{i}$. The conditions in Theorem 2 are then satisfied with $\beta = 2$.

Proof of Theorem 2. We begin with the proof of (3.2). Throughout we assume that $d = 2$ and an analogous analysis holds for general d. If $\mathcal{F}$ is any representative code of $\mathcal{S}$, then by definition, the weight of the set $\mathcal{S} \setminus \mathcal{F}$ is at most ϵm^2 and since the weight of each vertex is at least one (see (3.1)), we get that the number of points in $\mathcal{S} \setminus \mathcal{F}$ is at most ϵm^2. This implies that the size of $\mathcal{F}$ is at least $(1 - \epsilon)m^2$ and so $b_m \geq (1 - \epsilon)m^2$.

To find an upper bound on b_m, we let $\mathcal{T} \subseteq \mathcal{S}$ be any "critical" set such that the weight of $\mathcal{T}$ is at most $\epsilon m^2 - 1$ and the weight of $\mathcal{T} \cup \{v\}$ for any point $v \in \mathcal{S} \setminus \mathcal{T}$ is at least ϵm^2. The set $\mathcal{S} \setminus \mathcal{T}$ is then a representative code of $\mathcal{S}$ and since the weight of any point is at most β (see (3.1)), we get that the number of vertices in $\mathcal{T}$ is at least $\frac{\epsilon}{\beta} \cdot m^2 - 1$. This in turn implies that $b_m \leq m^2 \left(1 - \frac{\epsilon}{\beta}\right) + 1$. This proves (3.2).

To prove (3.3), we use a subsequence argument analogous to the proof of Fekete's lemma. For integers $m \geq r \geq 1$, we let $m = k \cdot r + s$, where $k \geq 1$ and $0 \leq s \leq r - 1$ are integers and split $\{1, 2, \ldots, m\}^2$ into four sets

$$\mathcal{S}_1 := \{1, 2, \ldots, kr\}^2, \quad \mathcal{S}_2 := \{1, 2, \ldots, kr\} \times \{kr + 1, kr + 2, \ldots, kr + s\},$$

$$\mathcal{S}_3 := \{kr + 1, \ldots, kr + s\} \times \{1, 2, \ldots, kr\} \text{ and } \mathcal{S}_4 := \{kr + 1, \ldots, kr + s\}^2.$$

Thus $\mathcal{S}_2$ is essentially a "rotated" version of $\mathcal{S}_3$. For $1 \leq i \leq 4$, let $\mathcal{G}_i(\epsilon)$ be a representative code of $\mathcal{S}_i$ and let $\mathcal{R}$ be any set in $\{1, 2, \ldots, m\}^2$ of weight $w(\mathcal{R}) \geq \epsilon m^2 = \epsilon(kr + s)^2$. We first see that $\bigcup_{i=1}^{4} \mathcal{G}_i(\epsilon)$ is a representative code of $\{1, 2, \ldots, m\}^2$. Indeed if $\mathcal{R}_i = \mathcal{R} \cap \mathcal{S}_i$, then using the fact that $w(\mathcal{R}) = \sum_{i=1}^{4} w(\mathcal{R}_i)$, we get that either $w(\mathcal{R}_1) \geq \epsilon(kr)^2$ or $w(\mathcal{R}_2) \geq \epsilon krs$ or $w(\mathcal{R}_3) \geq \epsilon krs$ or $w(\mathcal{R}_4) \geq \epsilon s^2$. Consequently, we must have that $\mathcal{R} \cap \bigcup_{i=1}^{4} \mathcal{G}_i(\epsilon) \neq \emptyset$.

If $b^{(i)}$ denotes the minimum size of a representative code of $\mathcal{S}_i$, then from the discussion above we get

$$b_{kr+s} \leq \sum_{i=1}^{4} b^{(i)} \leq b^{(1)} + 2krs + s^2 \leq b^{(1)} + (2k + 1)r^2 \qquad (3.4)$$

where the second inequality in (3.4) follows from the trivial estimate that the size of any representative code of $\mathcal{R}_i$ is at most the total number of points in $\mathcal{R}_i$ and the final inequality in (3.4) follows from the fact that $s \leq r$.

To estimate $b^{(1)}$, we split $\mathcal{S}_1 := \{1, 2, \ldots, kr\}^2$ into k^2 disjoint rectangles $\mathcal{T}_i$, $1 \leq i \leq k^2$ each containing r^2 points with $\mathcal{T}_1 = \{1, 2, \ldots, r\}^2$. If c_i denotes the minimum size of a representative code of $\mathcal{T}_i$, then using the weight monotonicity relation, we get that $c_i \leq c_1$. To see this is true suppose $\mathcal{T}_2 = \{1, 2, \ldots, r\} \times \{r+1, \ldots, r+2r\}$ so that $\mathcal{T}_2 = \mathcal{T}_1 + (r, 0)$ is obtained by translation of $\mathcal{T}_1$. If $\mathcal{U} \subset \mathcal{T}_2$ is any set of weight at least ϵr^2 then $\mathcal{U} - (r, 0) \subseteq \mathcal{T}_1$ also has weight at least ϵr^2, by the weight monotonicity relation. Consequently if $\mathcal{W}_1$ is a representative code of $\mathcal{T}_1$, then $\mathcal{W}_1 + (r, 0)$ is a representative code of $\mathcal{T}_2$. Thus $c_2 \leq c_1$ and the proof of general c_i is analogous.

From (3.4) and the discussion in the above paragraph, we get that $b_m = b_{kr+s} \leq k^2 b_r + (2k+1)r^2$ and so

$$\frac{b_m}{m^2} \leq \frac{k^2 b_r + (r^2(2k+1))}{m^2} = \left(\frac{kr}{kr+s}\right)^2 \left(\frac{b_r}{r^2} + \frac{2k+1}{k^2}\right).$$

If $m \to \infty$ with r fixed, then $k = k(m, r) = \frac{m-s}{r} \geq \frac{m-r}{r} \to \infty$ as well and so $\frac{kr}{kr+s} \to 1$. This in turn implies that

$$\limsup_m \frac{b_m}{m^2} \leq \limsup_m \left(\frac{b_r}{r^2} + \frac{2k+1}{k^2}\right) = \frac{b_r}{r^2}. \tag{3.5}$$

Since $r \geq 1$ is arbitrary we get from (3.5) that

$$\limsup_m \frac{b_m}{m^2} = \liminf_r \frac{b_r}{r^2} = \inf_r \frac{b_r}{r^2} =: \lambda.$$

Also, the bounds for λ follow from (3.2). $\qquad\square$

4 Random Permutations

Random permutations and applications are frequently encountered in computing problems and it is of interest to study the cycle properties of a randomly chosen permutation. The papers [6], [13] studied limiting distributions for the convergence of the number of cycles and cycles lengths of a uniform random permutation, after suitable renormalization. Later [1] used Poisson approximation and estimates on the total variation distance to study the convergence of the overall cycle structure to a process of independent Poisson random variables. Recently [3] have used probability generating functions to study convergence of number of cycles of uniform random permutations conditioned not to have large cycles, scaled and centred, to the Gaussian distribution.

From the combinatorial aspect, Stirling numbers of the first kind and generating functions have been used to study random permutation statistics. Using the Flajolet-Sedgewick theorem it is possible to enumerate permutations with

constraints [8]. In this section, we use conditioning to obtain a recursive relation involving cycle moments of random permutations. As an illustration, we compute recursive relation involving the mean and the variance of the number of cycles in a uniformly random permutation.

We begin with a couple of definitions. A permutation π of $\{1, 2, \ldots, n\}$ is a bijective map $\pi : \{1, 2, \ldots, n\} \rightarrow \{1, 2, \ldots, n\}$. The total number of possible permutations of $\{1, 2, \ldots, n\}$ is therefore

$$n! := n \cdot (n-1) \cdots 2 \cdot 1.$$

A *cycle* of length k in a permutation π is a $k-$tuple $(i_1, \ldots, i_k)$ such that $\pi(i_j) = i_{j+1}$ for $1 \leq j \leq k-1$ and $\pi(i_k) = i_1$. Every number in $\{1, 2, \ldots, n\}$ belongs to some cycle of π and this provides an alternate representation of π; for example $(1345)(267)(89)$ is the cycle representation of the permutation π on $\{1, 2, \ldots, 9\}$ satisfying $\pi(1) = 3, \pi(3) = 4, \pi(4) = 5, \pi(5) = 1, \pi(2) = 6, \pi(6) = 7, \pi(7) = 2, \pi(8) = 9, \pi(9) = 8$.

Let Π denote a uniformly chosen random permutation of $\{1, 2, \ldots, n\}$ defined on the probability space $(\Omega_n, \mathcal{F}_n, \mathbb{P}_n)$ so that

$$\mathbb{P}_n(\Pi = \pi) = \frac{1}{n!}$$

for any deterministic permutation π. Let $N_n = N_n(\Pi)$ be the random number of cycles in Π and for integers $n, s \geq 1$, set $\mu_{0,s} := 0$ and $\mu_{n,s} := \mathbb{E}N_n^s$. We have the following result.

Theorem 3. *For integers $n, s \geq 1$ we have*

$$\mu_{n,s} = 1 + \frac{1}{n} \sum_{r=1}^{s} \sum_{j=1}^{n-1} \binom{s}{r} \mu_{j,r}, \tag{4.1}$$

where $\binom{s}{r} = \frac{s!}{r!(s-r)!}$ is the Binomial coefficient.

From the recursive structure of equation (4.1), we then have that $\mu_{n,s}$ could be computed using the previous values $\{\mu_{j,r}\}_{j \leq n-1, r \leq s}$.

As a Corollary of Theorem 3 we have the following recursive relations for the mean and variance of N_n.

Corollary 1. *The mean $\mu_n := \mu_{n,1}$ satisfies $\mu_1 = 1$ and the recursive equation*

$$\mu_n = 1 + \frac{1}{n} \sum_{i=1}^{n-1} \mu_i \tag{4.2}$$

for $n \geq 2$. The sequence $H_n := \sum_{j=1}^{n} \frac{1}{j}$ is the unique sequence satisfying (4.2).

The variance $v_n = var(N_n) := \mu_{n,2} - \mu_{n,1}^2$ satisfies $v_1 = 0$ and the recursive equation

$$v_n = 1 + \frac{1}{n} \sum_{i=1}^{n-1} v_i - \frac{H_n}{n}. \tag{4.3}$$

The sequence $M_n := H_n - \sum_{i=1}^{n} \frac{1}{i^2}$ is the unique sequence satisfying (4.3).

Using (4.2), (4.3) and the recursive relation (4.1), we could similarly compute higher order moments.

We prove Theorem 3 and Corollary 1 in that order.

Proof of Theorem 3. To obtain the desired recursive relation, we condition on the length of the first cycle and study the number of cycles in the remaining set of elements.

Let $\mathcal{S}_1$ denote the cycle of the random permutation Π containing the number 1 and let $L_1 = \#\mathcal{S}_1$ be the length of $\mathcal{S}_1$ so that $\mathcal{S}_1$ is an L_1-tuple. If $L_1 = k \leq n-1$, then Π induces a permutation $\sigma : \{1, 2, \ldots, n-k\} \to \{1, 2, \ldots, n-k\}$ on the remaining $n-k$ numbers $\{1, 2, \ldots, n\} \setminus \mathcal{S}_1$ in the following way. Arrange the numbers in $\{1, 2, \ldots, n\} \setminus \mathcal{S}_1$ in increasing order $j_1 < j_2 < \ldots < j_{n-k}$ and suppose that $\pi(j_l) = m_l$ for $1 \leq l \leq n-k$. The induced permutation σ then satisfies $m_l = j_{\sigma(l)}$ for $1 \leq l \leq n-k$.

Conditional on $L_1 = k$ we now see that σ is uniformly distributed in the sense that for any deterministic permutation $\sigma_0 : \{1, 2, \ldots, n-k\} \to \{1, 2, \ldots, n-k\}$ we have

$$\mathbb{P}_n\left(\sigma = \sigma_0 | L_1 = k\right) = \mathbb{P}_{n-k}(\sigma_0) = \frac{1}{(n-k)!}. \tag{4.4}$$

To see (4.4) is true, we first write

$$\mathbb{P}_n\left(\sigma = \sigma_0 | L_1 = k\right) = \frac{\mathbb{P}_n\left(\{\sigma = \sigma_0\} \cap \{L_1 = k\}\right)}{\mathbb{P}_n(L_1 = k)}. \tag{4.5}$$

If $k = 1$, then the numerator in the right side of (4.5) is $\frac{1}{n!}$. Moreover, if the first cycle simply consists of the single element 1, then the remaining $n-1$ numbers can be arranged in $(n-1)!$ ways and so $\mathbb{P}_n(L_1 = 1) = \frac{(n-1)!}{n!}$. Thus (4.4) is true for $k = 1$.

For $2 \leq k \leq n-1$, we have from (4.5) that $\mathbb{P}_n\left(\sigma = \sigma_0 | L_1 = k\right)$ equals

$$\frac{\sum_{(i_1, \ldots, i_{k-1})} \mathbb{P}_n\left(\{\sigma = \sigma_0\} \cap \{\mathcal{S}_1 = (1, i_1, \ldots, i_{k-1})\}\right)}{\sum_{(i_1, \ldots, i_{k-1})} \mathbb{P}_n(\mathcal{S}_1 = (1, i_1, \ldots, i_{k-1}))}$$

where the summation is over all $k-1$ tuples $(i_1, \ldots, i_{k-1})$ containing distinct elements. For any $(1, i_1, \ldots, i_{k-1})$, the term

$$\mathbb{P}_n\left(\{\sigma = \sigma_0\} \cap \{\mathcal{S}_1 = (1, i_1, \ldots, i_{k-1})\}\right) = \frac{1}{n!} \tag{4.6}$$

and

$$\mathbb{P}_n(\mathcal{S}_1 = (1, i_1, \ldots, i_{k-1})) = \frac{(n-k)!}{n!} \tag{4.7}$$

since there are $(n-k)!$ ways to permute the remaining $n-k$ elements of the set $\{2, \ldots, n\} \setminus \{i_1, \ldots, i_{k-1}\}$. Substituting (4.6) and (4.7) into (4), we get (4.4).

Summing (4.7) over all $k-1$ tuples with distinct entries (for which there are $(n-1) \cdot (n-2) \cdots (n-k+1)$ choices), we also get that $\mathbb{P}_n(L_1 = \#\mathcal{S}_1 = k) = \frac{1}{n}$.

From the discussion in the previous paragraph, we get that the above relation holds for all $1 \leq k \leq n$. Thus

$$\mu_{n,s} = \mathbb{E}_n N_n^s = \sum_{k=1}^{n} \mathbb{E}_n(N_n^s | L_1 = k) \mathbb{P}_n(L_1 = k) = \frac{1}{n} \sum_{k=1}^{n} \mathbb{E}_n(N_n^s | L_1 = k). \quad (4.8)$$

If $k = n$ then $N_n = 1$ and if $1 \leq k \leq n-1$, then $N_n = 1 + M_n$, where M_n is the number of cycles in the induced permutation σ. Therefore we get from (4.8) that

$$\mu_{n,s} = \frac{1}{n} + \frac{1}{n} \sum_{k=1}^{n-1} \mathbb{E}_n \left((1 + M_n)^s | L_1 = k \right). \quad (4.9)$$

Using the conditional distribution equivalence (4.4), we have for $1 \leq k \leq n-1$ that $\mathbb{E}_n \left((1 + M_n)^s | L_1 = k \right)$ equals

$$\mathbb{E}_{n-k}(1 + N_{n-k})^s = 1 + \sum_{r=1}^{s} \binom{s}{r} \mathbb{E}_{n-k} N_{n-k}^r, \quad (4.10)$$

by the Binomial expansion. Substituting (4.10) into (4.9) we get (4.1). □

Proof of Corollary 1. We begin with the proof of (4.2). Setting $s = 1$ in (4.1) and $\mu_n = \mu_{n,1}$, we get that $\mu_1 = 1$ and for $n \geq 2$, we get that μ_n satisfies (4.2). We first see by induction that H_n as defined in Proposition 1 satisfies (4.2). For $n = 2$, this statement is true and suppose H_l satisfies (4.2) for $1 \leq l \leq n-1$. For $l = n$, the right side of (4.2) evaluated with $\mu_i = H_i$ equals

$$1 + \frac{1}{n} \sum_{i=1}^{n-1} \sum_{j=1}^{i} \frac{1}{j} = 1 + \frac{1}{n} \sum_{j=1}^{n-1} \sum_{i=j}^{n-1} \frac{1}{j} = 1 + \frac{1}{n} \sum_{j=1}^{n-1} \frac{n-j}{j}, \quad (4.11)$$

by interchanging the order of summation in the second equality. The final term in (4.11) equals H_n and this proves the induction step.

Suppose now that $\{b_n\}$ is some sequence satisfying (4.2) with $b_1 = 1$ and let $u_n = b_n - H_n$ denote the difference. The sequence $\{u_n\}$ satisfies $u_1 = 0$ and $u_n = \frac{1}{n} \sum_{i=1}^{n-1} u_i$ for all $n \geq 2$. Thus $u_2 = \frac{u_1}{2} = 0$ and iteratively, we get $u_n = u_2 = 0$ for all $n \geq 2$. Thus H_n is the unique sequence satisfying (4.2).

We now obtain the variance estimate as follows. Letting $d_n := \mu_{n,2}$ and $\mu_n := \mu_{n,1} = H_n$, we get from (4.1) that

$$d_n = 1 + \frac{1}{n} \sum_{i=1}^{n-1} (d_i + 2\mu_i) = 2\mu_n - 1 + \frac{1}{n} \sum_{i=1}^{n-1} d_i, \quad (4.12)$$

since $\frac{1}{n} \sum_{i=1}^{n-1} \mu_i = \mu_n - 1$ (see (4.2)). From (4.12) we get that $v_n = d_n - \mu_n^2$ equals

$$v_n = \frac{1}{n} \sum_{i=1}^{n-1} (d_i - \mu_i^2) + \frac{1}{n} \sum_{i=1}^{n-1} \mu_i^2 - (\mu_n - 1)^2$$

$$= \frac{1}{n} \sum_{i=1}^{n-1} v_i + \frac{1}{n} \sum_{i=1}^{n-1} \mu_i^2 - (\mu_n - 1)^2.$$

It only remains to see that $\frac{1}{n}\sum_{i=1}^{n-1}\mu_i^2 - (\mu_n - 1)^2 = 1 - \frac{H_n}{n}$ and for that we use $\mu_i = H_i = \sum_{j=1}^{i}\frac{1}{j}$ (see (4.2)) to first get that $\frac{1}{n}\sum_{i=1}^{n-1}\mu_i^2$ equals

$$\frac{1}{n}\sum_{i=1}^{n-1}\sum_{j_1=1}^{i}\sum_{j_2=1}^{i}\frac{1}{j_1\cdot j_2} = \frac{1}{n}\sum_{j_1=1}^{n-1}\sum_{j_2=1}^{n-1}\sum_{i=\max(j_1,j_2)}^{n-1}\frac{1}{j_1\cdot j_2}$$

$$\frac{1}{n}\sum_{j_1=1}^{n-1}\sum_{j_2=1}^{n-1}\frac{(n-\max(j_1,j_2))}{j_1\cdot j_2} = \frac{1}{n}\sum_{j_1=1}^{n-1}\Delta(j_1),\tag{4.13}$$

where $\Delta(j_1) = \sum_{j_2=1}^{j_1}\frac{n-j_1}{j_1\cdot j_2} + \sum_{j_2=j_1+1}^{n-1}\frac{n-j_2}{j_1\cdot j_2}$ equals

$$\sum_{j_2=1}^{n-1}\frac{n}{j_1\cdot j_2} - \sum_{j_2=1}^{j_1}\frac{1}{j_2} - \sum_{j_2=j_1+1}^{n-1}\frac{1}{j_1}.$$

Thus $\frac{1}{n}\sum_{i=1}^{n-1}\mu_i^2$ equals

$$\sum_{j_1=1}^{n-1}\sum_{j_2=1}^{n-1}\frac{1}{j_1\cdot j_2} - \frac{1}{n}\sum_{j_1=1}^{n-1}\sum_{j_2=1}^{j_1}\frac{1}{j_2} - \frac{1}{n}\sum_{j_1=1}^{n-1}\sum_{j_2=j_1+1}^{n-1}\frac{1}{j_1}.\tag{4.14}$$

The first term in (4.14) is

$$\sum_{j_1=1}^{n-1}\sum_{j_2=1}^{n-1}\frac{1}{j_1\cdot j_2} = H_{n-1}^2 = \left(H_n - \frac{1}{n}\right)^2$$

and the second term in (4.14) is

$$\frac{1}{n}\sum_{j_1=1}^{n-1}\sum_{j_2=1}^{j_1}\frac{1}{j_2} = \frac{1}{n}\sum_{j_1=1}^{n-1}H_{j_1} = H_n - 1$$

using the fact that $\mu_n = H_n$ satisfies (4.2). The third term in (4.14) equals

$$\frac{1}{n}\sum_{j_1=1}^{n-1}\frac{n-1-j_1}{j_1} = \left(\frac{n-1}{n}\right)\left(H_n - 1 - \frac{1}{n}\right)$$

after rearrangement of terms. Substituting these three expressions into (4.14), we get that $\frac{1}{n}\sum_{i=1}^{n-1}\mu_i^2$ equals $1 + (H_n - 1)^2 - \frac{H_n}{n}$, which is what we wanted to prove. Finally, arguing as before, we also have that M_n is the unique sequence satisfying (4.3). □

References

1. Arratia, R., Tavaré, S.: The cycle structure of random permutations. Ann. Probab. **20**, 1567–1591 (1992)
2. Balaji, S.B., Krishnan, M.N., Vajha, M., Ramkumar, V., Sasidharan, B., Kumar, P.V.: Erasure coding for distributed storage: an overview. Sci. China Inf. Sci. **61** (2018)
3. Betz, V., Schäfer, H.: The number of cycles in random permutations without long cycles is asymptotically Gaussian, ALEA. Lat. Am. J. Probab. Stat. **14**, 427–444 (2017)
4. Bhaskara, A., Desai, D., Srinivasan, S.: Optimal hitting sets for combinatorial shapes. Theory Comput. **9**, 441–470 (2013)
5. Forbes, M., Yekhanin, S.: On the locality of codeword symbols in non-linear codes. Discrete Math. **324**, 78–84 (2014)
6. Gončarov, V.: On the field of combinatory analysis. Am. Math. Soc. Trans. **19**, 1–46 (1962)
7. Gopalan, P., Huang, C., Simitci, H., Yekhanin, S.: On the locality of codeword symbols. IEEE Trans. Inf. Theory **58**, 6925–6934 (2012)
8. Graham, R., Knuth D., Patashnik, O.: Concrete Mathematics. Addison-Wesley, Boston (1989)
9. Huffman, W.C., Pless, V.: Fundamentals of Error Correcting Codes. Cambridge University Press, Cambridge (2003)
10. Karpovsky, M.G., Chakrabarty, K., Levitin, L.B.: On a new class of codes for identifying vertices in graphs. IEEE Trans. Inf. Theory **44**, 599–611 (1998)
11. Linial, N., Luby, M., Saks, M., Zuckerman, D.: Efficient construction of a small hitting set for combinatorial rectangles in high dimension. Combinatorica **17**, 215–234 (1997)
12. Rashmi, K.V., Shah, N.B., Kumar, P.V.: Optimal exact-regenerating codes for distributed storage at the MSR and MBR points via a product-matrix construction. IEEE Trans. Inf. Theory **57**, 5227–5239 (2011)
13. Shepp, L.A., Lloyd, S.P.: Ordered cycle lengths in a random permutation. Trans. Am. Math. Soc. **121**, 340–357 (1966)
14. Sunil Chandran, L.: A lower bound for the hitting set size for combinatorial rectangles and an application. Inf. Process. Lett. **86**, 75–78 (2003)