

Санкт-Петербургский государственный университет

# Кодирование случайных множеств в булевой модели с переменной интенсивностью

Дипломная работа  
студентки 4 курса  
направление «Математика»  
01.03.01.  
группы 16.Б01—мм  
очной формы обучения  
Давыденковой Марии Сергеевны

Научный руководитель:  
д.ф.-м.н. профессор  
Лифшиц Михаил Анатольевич

Санкт-Петербург  
2020 год

# Содержание

<b>1 Введение</b>	<b>2</b>
1.1 Постановка задачи . . . . .	2
1.2 Полученные результаты . . . . .	3
<b>2 Нижние оценки для вероятности больших уклонений</b>	<b>3</b>
2.1 Постоянный радиус . . . . .	3
2.2 Радиус с плотностью . . . . .	6
<b>3 Верхние оценки для вероятности больших уклонений</b>	<b>7</b>

## 1 Введение

### 1.1 Постановка задачи

Пусть  $S$  – случайный элемент некоторого метрического пространства  $(X, dist)$ . Средней ошибкой дискретизации  $S$  называется величина

$$D^{(q)}(r) := \inf_{\#C \leq e^r} \mathbb{E} \min_{A \in C} dist(S, A), \quad r > 0.$$

Скорость ее убывания при стремлении  $r$  к бесконечности характеризует сложность распределения  $S$ . Общие свойства величины  $D^{(q)}(r)$  изучены в [6, 4, 7]. В последние два десятилетия ошибки дискретизации исследовались, в основном, для траекторий случайных процессов, рассматриваемых как случайный элемент функционального пространства, см., например, [1, 5].

В работе [2] изучалась ошибка дискретизации для случайного множества, рассматриваемого как случайный элемент пространства компактов, снабжённого метрикой Хаусдорфа  $d_H(A, B) := \max(\sup_{a \in A} \inf_{b \in B} \|a - b\|, \sup_{b \in B} \inf_{a \in A} \|a - b\|)$ . В качестве случайного множества была взята стандартная Булева модель (Boolean model, см. [3, 8]), которая устроена следующим образом.

Рассмотрим куб  $[0, a]^d$  в  $\mathbb{R}^d$ ,  $d \geq 1$ , и случайный набор шаров с центрами в этом кубе, определенный следующим образом: пусть центры шаров  $\xi_i$  — случайные величины, распределенные равномерно в  $[0, a]^d$ , радиусы  $R_i$  — некоторые одинаково распределенные неотрицательные случайные величины, а количество шаров  $N$  — пуассоновская случайная величина с параметром  $a^d \lambda$ , где  $\lambda = \lambda(a)$  — положительный параметр, зависящий от  $a$ . Все эти случайные величины независимы. Отметим, что при таком построении множество  $(\xi_i)_{i \leq N}$  — это пуассоновский точечный процесс с интенсивностью  $\lambda$ .

Обозначим через  $B(\xi_i, R_i)$  шар с центром в  $\xi_i$  радиуса  $R_i$  (пока что мы не фиксируем норму в  $\mathbb{R}^d$ ). Мы будем рассматривать “картинку”, образованную этим набором шаров:

$$S_a = \bigcup_{i=1}^N B(\xi_i, R_i) \cap [0, a]^d.$$

Определенная таким образом картинка называется *Булевой моделью случайного множества*. Заметим, что свойства  $S_a$  будут зависеть от рассматриваемой нормы.

В работе [2] исследовался случай  $a = 1$  с постоянной интенсивностью  $\lambda$ . Результаты основаны на изучении вероятностей больших уклонений величины

$$K_1 = \min\{r \geq 1 | \exists i_1, \dots, i_r \subset \{1, \dots, N\} : S_1 = \bigcup_{l=1}^r B(\xi_{i_l}, R_{i_l}) \cap [0, a]^d\},$$

называемой *минимальным числом видимых шаров*, то есть на нахождении асимптотики  $\mathbb{P}[K_1 \geq n]$  при  $n$ , стремящемся к бесконечности.

В настоящей работе некоторые из этих оценок распространяются на случай, когда параметр  $a$  стремится к бесконечности, причем  $n \gg a^d \lambda$ .

Заметим, что тривиальная оценка, вытекающая из свойств пуассоновского распределения, такова:

$$\begin{aligned} \mathbb{P}[K_a \geq n] &\leq \mathbb{P}[N \geq n] \sim \mathbb{P}[N = n] = \\ &= \frac{(a^d \lambda)^n}{n!} e^{-a^d \lambda} \sim \left( \frac{ea^d \lambda}{n} \right)^n e^{-a^d \lambda} \sqrt{2\pi n} = \exp(-n \log n + n \log(a^d \lambda) + O(n)) = \\ &= \exp((-n \log n + dn \log a + n \log \lambda + O(n))), \quad n \rightarrow \infty, a \rightarrow \infty. \end{aligned}$$

Условие  $n \gg a^d \lambda$  гарантирует, что вероятность больших уклонений стремится к нулю, так как  $-n \log(n/a^d \lambda)$  стремится к минус бесконечности с ростом  $n$  и  $a$ .

## 1.2 Полученные результаты

Получены (пока что :) ) более точные оценки вероятности больших уклонений для  $\ell_1$ - и  $\ell_2$ -нормы. А именно, что существует число  $\beta > 1$ , для которого

$$\mathbb{P}[K_a \geq n] = \exp((- \beta \cdot n \log n + \beta \cdot dn \log a + n \log \lambda + O(n))), \quad n \rightarrow \infty, a \rightarrow \infty.$$

Число  $\beta$  зависит от размерности  $d$ , используемой нормы и распределения радиусов  $R_i$ .

## 2 Нижние оценки для вероятности больших уклонений

### 2.1 Постоянный радиус

Рассмотрим случай, когда радиусы – это константа  $r > 0$  п.н.

**Теорема 1.** Пусть  $R_1 \equiv r > 0$  п.н., где шары берутся в  $\ell_1$ -норме,  $d \geq 2$ , и  $n \gg a^d \lambda$ . Тогда

$$\mathbb{P}[K_a \geq n] \geq \exp \left( - \left( 1 + \frac{1}{d-1} \right) n \log n + \left( 1 + \frac{1}{d-1} \right) dn \log a + n \log \lambda + O(n) \right), \quad n, a \rightarrow \infty.$$

*Доказательство.* Рассмотрим набор ячеек:

$$\prod_{m=1}^{d-1} \left[ \frac{a^{d/(d-1)}(k_m + 1/4)}{(8rn)^{1/(d-1)}}, \frac{a^{d/(d-1)}(k_m + 3/4)}{(8rn)^{1/(d-1)}} \right] \times \\ \times \left[ l \left( \frac{c_1 a^{d/(d-1)}}{(rn)^{1/(d-1)}} + 2r \right), l \left( \frac{c_1 a^{d/(d-1)}}{(rn)^{1/(d-1)}} + 2r \right) + \frac{c_1 a^{d/(d-1)}}{(rn)^{1/(d-1)}} \right],$$

где  $k_m \in \{0, \dots, \lfloor (8rn/a)^{1/(d-1)} \rfloor - 1\}$ ,  $l \in \{0, \dots, \lfloor a/4r \rfloor\}$ ,  $c_1 = 2^{-(2+3/(d-1))}$ .

Все эти ячейки лежат в кубе  $[0, a]^d$ . По первым  $d-1$  координатам это очевидно, проверим по последней. Действительно,

$$\frac{a}{4r} \cdot \left( \frac{c_1 a^{d/(d-1)}}{(rn)^{1/(d-1)}} + 2r \right) + \frac{c_1 a^{d/(d-1)}}{(rn)^{1/(d-1)}} = \frac{c_1 a^{d/(d-1)}}{r(rn)^{1/(d-1)}} \cdot \left( \frac{a}{4} + r \right) + \frac{a}{2} < a$$

при достаточно больших  $a$  и  $n$ .

Заметим, что если центры шаров  $B(\xi_i, r)$ ,  $B(\xi_j, r)$  лежат в разных “рядах”, то есть  $\xi_i^{(d)}$  и  $\xi_j^{(d)}$  лежат в разных выбранных нами интервалах, то эти шары не пересекаются, так как расстояние между “рядами” хотя бы  $2r$ . Если же центры лежат в одном “ряду”, то в каждом из шаров есть точка, не покрытая другим шаром. Действительно, рассмотрим точку  $x_i = \xi_i + (0, \dots, 0, r)$ . Заметим, что для достаточно больших  $n$  и  $a$  эта точка действительно лежит в кубе  $[0, a]^d$ . Итак, при  $j \neq i$

$$\|x_i - \xi_j\|_1 = |\xi_i^{(d)} + r - \xi_j^{(d)}| + \sum_{m=1}^{d-1} |\xi_i^{(m)} - \xi_j^{(m)}| \geq \\ \geq r - |\xi_i^{(d)} - \xi_j^{(d)}| + \sum_{m=1}^{d-1} |\xi_i^{(m)} - \xi_j^{(m)}| \geq r - \frac{c_1 a^{d/(d-1)}}{(rn)^{1/(d-1)}} + \frac{1/2 \cdot a^{d/(d-1)}}{(8rn)^{1/(d-1)}} > r.$$

Мы построили порядка  $8rn/a \cdot a/4r = 2n$  непересекающихся ячеек, поэтому можно выбрать  $n$  из них. Назовем их  $V_1, \dots, V_n$  и рассмотрим событие

$$E = \{N = n\} \cap \bigcup_{\pi - \text{перестановка } \{1, \dots, n\}} \{\xi_i \in V_{\pi(i)}, i = 1, \dots, n\}.$$

Заметим, что событие  $E$  влечет событие  $\{K \geq n\}$  в силу свойств выбранных ячеек, показанных выше.

Следовательно,

$$\mathbb{P}[K_a \geq n] \geq \mathbb{P}[E] = \frac{(a^d \lambda)^n}{n!} e^{-a^d \lambda} \cdot n! \cdot \left( \frac{1}{a^d} \cdot \left( \frac{1/2 \cdot a^{d/(d-1)}}{(8rn)^{1/(d-1)}} \right)^{d-1} \cdot \frac{c_1 a^{d/(d-1)}}{(rn)^{1/(d-1)}} \right)^n = \\ = \exp \left( dn \log a + n \log \lambda - a^d \lambda - \left( 1 + \frac{1}{d-1} \right) n \log n + \frac{d}{d-1} n \log a + O(n) \right) = \\ = \exp \left( - \left( 1 + \frac{1}{d-1} \right) n \log n + \left( 1 + \frac{1}{d-1} \right) dn \log a + n \log \lambda + O(n) \right).$$

□

**Теорема 2.** Пусть  $R_1 \equiv r > 0$  п.н., где шары берутся в  $\ell_2$ -норме,  $d \geq 2$ , и  $n \gg a^d \lambda$ . Тогда

$$\mathbb{P}[K_a \geq n] \geq \exp \left( - \left( 1 + \frac{2}{d-1} \right) n \log n + \left( 1 + \frac{2}{d-1} \right) dn \log a + O(n) \right), \quad n, a \rightarrow \infty.$$

Доказательство теоремы аналогично случаю  $\ell_1$ -нормы, и отличается только размером ячеек.

*Доказательство.* Рассмотрим набор ячеек:

$$\prod_{m=1}^{d-1} \left[ \frac{a^{d/(d-1)}(k_m + 1/4)}{(8rn)^{1/(d-1)}}, \frac{a^{d/(d-1)}(k_m + 3/4)}{(8rn)^{1/(d-1)}} \right] \times \\ \times \left[ l \left( \frac{c_2 a^{2+2/(d-1)}}{r(rn)^{2/(d-1)}} + 2r \right), l \left( \frac{c_2 a^{2+2/(d-1)}}{r(rn)^{2/(d-1)}} + 2r \right) + \frac{c_2 a^{2+2/(d-1)}}{r(rn)^{2/(d-1)}} \right],$$

где  $k_m \in \{0, \dots, \lfloor (8rn/a)^{1/(d-1)} \rfloor - 1\}$ ,  $l \in \{0, \dots, \lfloor a/4r \rfloor\}$ ,  $c_2 = 2^{-(4+6/(d-1))}$ .

Все эти ячейки лежат в кубе  $[0, a]^d$ . По первым  $d-1$  координатам это очевидно, проверим по последней. Действительно,

$$\frac{a}{4r} \left( \frac{c_2 a^{2+2/(d-1)}}{r(rn)^{2/(d-1)}} + 2r \right) + \frac{c_2 a^{2+2/(d-1)}}{r(rn)^{2/(d-1)}} = \frac{c_2 a^{2+2/(d-1)}}{r^2(rn)^{2/(d-1)}} \cdot \left( \frac{a}{4} + r \right) + \frac{a}{2} < a$$

при достаточно больших  $a$  и  $n$ .

Заметим, что если центры шаров  $B(\xi_i, r)$ ,  $B(\xi_j, r)$  лежат в разных “рядах”, то есть  $\xi_i^{(d)}$  и  $\xi_j^{(d)}$  лежат в разных выбранных нами интервалах, то эти шары не пересекаются, так как расстояние между “рядами” хотя бы  $2r$ . Если же центры лежат в одном “ряду”, то в каждом из шаров есть точка, не покрытая другим шаром. Действительно, рассмотрим точку  $x_i = \xi_i + (0, \dots, 0, r)$ . Заметим, что для достаточно больших  $n$  и  $a$  эта точка действительно лежит в кубе  $[0, a]^d$ . Итак, при  $j \neq i$

$$\|x_i - \xi_j\|_2^2 = |\xi_i^{(d)} + r - \xi_j^{(d)}|^2 + \sum_{m=1}^{d-1} |\xi_i^{(m)} - \xi_j^{(m)}|^2 \geq \\ \geq r^2 - 2r|\xi_i^{(d)} - \xi_j^{(d)}| + (\xi_i^{(d)} - \xi_j^{(d)})^2 + \sum_{m=1}^{d-1} |\xi_i^{(m)} - \xi_j^{(m)}|^2 \geq \\ \geq r^2 - \frac{2c_2 a^{2+2/(d-1)}}{(rn)^{2/(d-1)}} + \left( \frac{1/2 a^{d/(d-1)}}{(8rn)^{1/(d-1)}} \right)^2 \geq r^2.$$

Мы построили порядка  $8rn/a \cdot a/4r = 2n$  непересекающихся ячеек, поэтому можно выбрать  $n$  из них. Назовем их  $V_1, \dots, V_n$  и рассмотрим событие

$$E = \{N = n\} \cap \bigcup_{\pi - \text{перестановка } \{1, \dots, n\}} \{\xi_i \in V_{\pi(i)}, i = 1, \dots, n\}.$$

Заметим, что событие  $E$  влечет событие  $\{K \geq n\}$  в силу свойств выбранных ячеек, показанных выше.

Следовательно,

$$\begin{aligned} \mathbb{P}[K_a \geq n] &\geq \mathbb{P}[E] = \frac{(a^d \lambda)^n}{n!} e^{-a^d \lambda} \cdot n! \cdot \left( \frac{1}{a^d} \cdot \left( \frac{1/2 \cdot a^{d/(d-1)}}{(8rn)^{1/(d-1)}} \right)^{d-1} \cdot \frac{c_2 a^{2+2/(d-1)}}{r(rn)^{2/(d-1)}} \right)^n = \\ &= \exp \left( dn \log a + n \log \lambda - a^d \lambda - \left( 2 + \frac{2}{d-1} \right) n \log n + \left( 2 + \frac{2}{d-1} \right) n \log a + O(n) \right) = \\ &= \exp \left( - \left( 1 + \frac{2}{d-1} \right) n \log n + \left( 1 + \frac{2}{d-1} \right) dn \log a + n \log \lambda + O(n) \right). \end{aligned}$$

□

## 2.2 Радиус с плотностью

**Теорема 3.** Пусть распределение  $R_1$  имеет плотность  $p$  относительно меры Лебега, причем  $p(z) \geq cz^{\alpha-1}$  для малых  $z$  и некоторых  $c, \alpha > 0$ ,  $d \geq 1$ , и  $n \gg a^d \lambda$ . Тогда

$$\mathbb{P}[K_a \geq n] \geq \exp \left( - (1 + \alpha/d) n \log n + (1 + \alpha/d) dn \log a + n \log \lambda + O(n) \right), \quad n, a \rightarrow \infty.$$

*Замечание.* Утверждение теоремы верно для произвольной нормы в  $\mathbb{R}^d$ .

*Доказательство.* Рассмотрим набор ячеек:

$$\prod_{m=1}^d \left[ \frac{a(k_m + 1/4)}{(2n)^{1/d}}, \frac{a(k_m + 3/4)}{(2n)^{1/d}} \right],$$

где  $k_m \in \{0, \dots, \lfloor (2n)^{1/d} \rfloor - 1\}$ .

Мы построили порядка  $2n$  непересекающихся ячеек, поэтому можно выбрать  $n$  из них. Назовем их  $V_1, \dots, V_n$  и рассмотрим событие

$$E = \{N = n\} \cap \bigcup_{\pi - \text{перестановка } \{1, \dots, n\}} \{\xi_i \in V_{\pi(i)}, R_i \in [c_1 a n^{-1/d}, c_2 a n^{-1/d}], i = 1, \dots, n\},$$

где  $c_2 > c_1 > 0$  – некоторые константы, которые могут зависеть от нормы и размерности, причем  $c_2$  выбирается так, чтобы для  $i \neq j$  шары  $B(\xi_i, R_i), B(\xi_j, R_j)$  не пересекались. В этом случае событие  $E$  влечет  $[K_a \geq n]$ .

Следовательно,

$$\begin{aligned}
\mathbb{P}[K \geq n] &\geq \mathbb{P}[E] = \frac{(a^d \lambda)^n}{n!} e^{-a^d \lambda} \cdot n! \cdot \left( \frac{1}{a^d} \left( \frac{1/2a}{(2n)^{1/d}} \right)^d \cdot \int_{c_1 a n^{-1/d}}^{c_2 a n^{-1/d}} p(z) dz \right)^n \geq \\
&\geq (a^d \lambda)^n e^{-a^d \lambda} \cdot \left( \frac{1}{a^d} \left( \frac{1/2a}{(2n)^{1/d}} \right)^d \cdot c \int_{c_1 a n^{-1/d}}^{c_2 a n^{-1/d}} z^{\alpha-1} dz \right)^n = \\
&= (a^d \lambda)^n e^{-a^d \lambda} \cdot \left( \frac{1}{a^d} \left( \frac{1/2a}{(2n)^{1/d}} \right)^d \cdot c(c_2^\alpha - c_1^\alpha)/\alpha \cdot a^\alpha n^{-\alpha/d} \right)^n = \\
&= \exp(-(1 + \alpha/d)n \log n + (1 + \alpha/d)dn \log a + n \log \lambda + O(n)).
\end{aligned}$$

□

### 3 Верхние оценки для вероятности больших уклонений

**Теорема 4.** Пусть  $R_1 \equiv \frac{r}{a}$  н.н., где  $r < 1$ , шары берутся в  $l_1$ -норме, и  $d \geq 2$ . Тогда

$$\mathbb{P}[K_a \geq n] \leq \exp \left( \left( - \left( 1 + \frac{1}{d-1} \right) n \log n + \left( 1 + \frac{1}{d-1} \right) n \log a^d \right) (1 + o(1)) \right), n, a \rightarrow \infty.$$

*Доказательство. Шаг 1. Объединение шаров в группы.* Пусть

$$S = \bigcup_{i=1}^{K_a} B(\xi_i, R_1) \cap [0, 1]^d$$

обозначает неубывающее представление картинки  $S$ . Тогда для каждого числа  $i \leq K_a$  существует точка  $\nu_i \in B(\xi_i, R_1) \cap [0, 1]^d$ , которая не лежит ни в каком другом шаре  $B(\xi_j, R_1), j \neq i$ . Зафиксируем такие  $\nu_i, i = 1, 2, \dots, K_a$ . Обозначим  $\Delta_i := \nu_i - \xi_i$ .

Объединим шары в группы  $J_0, J_1^+, J_1^-, \dots, J_d^+, J_d^-$  следующим образом. Определим  $J_0 := \{i: \|\Delta_i\|_1 \leq R_1/2\}$ . Оценим мощность этого множества. Для любых  $i, j \in J_0$  выполнено:

$$R_1 < \|\nu_j - \xi_i\| \leq \|\nu_j - \nu_i\| + \|\nu_i - \xi_i\| \leq \|\nu_j - \nu_i\| + \|\Delta_i\| \leq \|\nu_j - \nu_i\| + R_1/2.$$

То есть  $\|\nu_j - \nu_i\| > R_1/2$ , а значит,  $\#J_0 \leq c_1(R_1)^{-d} = c_1 r^{-d} a^d$ , где константа  $c_1 = 2^d / \text{vol}_d B(0, 1)$ .

Теперь пусть  $i \notin J_0$ . Значит,

$$\|\Delta_i\|_\infty \geq c_2 \|\Delta_i\|_1 > c_2 R_1/2,$$

где  $c_2$  — константа, зависящая только от нормы. Поэтому  $i$  принадлежит одному из  $2d$  множеств вида:

$$J_m^+ := \{i: \Delta_i^{(m)} > c_2 R_1/2\}, \quad J_m^- := \{i: \Delta_i^{(m)} < -c_2 R_1/2\}.$$

*Шаг 2. Оценка расстояний между центрами.* Пусть  $\sigma: \mathbb{R}^d \rightarrow \mathbb{R}^d$  – проекция, определенная таким образом:

$$\sigma x := (x^{(1)}, \dots, x^{(d-1)}, 0).$$

**Лемма 1** (Лемма 20 из статьи.). Пусть  $i, j \in J_d^+$ ,  $i \neq j$ , и пусть  $c_3 := c_2 r / 2$ . Тогда

$$|\xi_i^{(d)} - \xi_j^{(d)}| \notin [|\sigma \xi_i - \sigma \xi_j|_1, c_3/a].$$

*Доказательство.* См. статью. □

*Шаг 3. Подсчет ячеек, содержащих центры шаров.* Зафиксируем большое число  $A > 0$  и покроем куб  $[0, 1]^d$  следующим набором ячеек:

$$V_{\bar{k}, k_d} := \prod_{m=1}^d \left[ \frac{A k_m}{(n a^{-1})^{1/(d-1)}}, \frac{A(k_m + 1)}{(n a^{-1})^{1/(d-1)}} \right],$$

где  $k_m \in \{0, \dots, \lfloor A^{-1}(n a^{-1})^{1/(d-1)} \rfloor\}$  для  $1 \leq m \leq d$  и мульти-индекса  $\bar{k} := (k_1, \dots, k_{d-1})$ .

Зафиксируем некоторый индекс  $\bar{k}$  и оценим количество ячеек, содержащих центры шаров:

$$N(\bar{k}, d, +) := \#\{k: \xi_i \in V_{\bar{k}, k} \text{ для некоторого } i \in J_d^+\}.$$

Заметим, что если  $\xi_i \in V_{\bar{k}, \kappa_i}$  и  $\xi_j \in V_{\bar{k}, \kappa_j}$  для некоторых  $i, j \in J_d^+$ , то

$$\|\sigma \xi_i - \sigma \xi_j\|_1 \leq (d-1)A(n a^{-1})^{1/(d-1)}.$$

Поэтому по лемме

$$|\xi_i^{(d)} - \xi_j^{(d)}| \notin [(d-1)A(n a^{-1})^{1/(d-1)}, c_3/a].$$

**Замечание.** При достаточно больших  $a, n$  этот интервал определен (то есть левый конец меньше правого).

Исходя из этого, разобьем  $[0, 1]$  на  $\lceil a c_3^{-1} \rceil$  частей длины не более  $c_3/a$ , и заметим, что если  $\xi_i^{(d)}, \xi_j^{(d)}$  лежат в одной части, то  $|\xi_i^{(d)} - \xi_j^{(d)}| \leq (d-1)A(n a^{-1})^{1/(d-1)}$ . Поэтому тогда  $|\kappa_i - \kappa_j| \leq d$ . Отсюда получаем искомую оценку:

$$N(\bar{k}, d, +) \leq d \lceil a c_3^{-1} \rceil =: a c_4.$$

Теперь мы можем получить общее число ячеек, содержащих центры:

$$\sum_{\bar{k}} \sum_{m=1}^d (N(\bar{k}, m, +) + N(\bar{k}, m, -)) \leq (2d) \cdot (a c_4) \cdot (n a^{-1} A^{-(d-1)}) =: \frac{c_5 n}{A^{d-1}}.$$

**Замечание.** Размер ячеек специально выбирался так, чтобы общее количество ячеек, содержащих центры, не зависело от  $a$ . Иначе в финальном рассуждении мы не смогли бы избавиться от громоздкого множителя, взяв достаточно большое  $A$ . Чтобы все было хорошо (хотя бы с членом  $n \log n$  в финальной асимптотике), надо



чтобы этот множитель был  $O(n)$ . Из всех таких степеней  $a$  я выбрала ту, которая дает наилучшую оценку.

Пусть  $\mathcal{U}$  – семейство всех возможных объединений из  $\left\lfloor \frac{c_5 n}{A^{d-1}} \right\rfloor$  ячеек. Их количество – число способов выбрать  $\left\lfloor \frac{c_5 n}{A^{d-1}} \right\rfloor$  ячеек из  $\left\lfloor \frac{n^{d/(d-1)}}{a^{d/(d-1)A^d}} \right\rfloor$ . Поэтому можно выписать следующую простую оценку:

$$\begin{aligned} \#\mathcal{U} &\leq \exp \left( \frac{c_5 n}{A^{d-1}} \log \left\lfloor \frac{n^{d/(d-1)}}{a^{d/(d-1)A^d}} \right\rfloor \right) = \\ &= \exp \left( \left( \frac{c_5 d}{(d-1)A^{d-1}} n \log n + \frac{c_5}{(d-1)A^{d-1}} n \log a^d \right) (1 + o(1)) \right). \end{aligned}$$

Для каждого  $U \in \mathcal{U}$  объем можно оценить так:

$$\text{vol}_d(U) \leq \frac{c_5 n}{A^{d-1}} \left( \frac{A a^{1/(d-1)}}{n^{1/(d-1)}} \right)^d = \frac{A c_5 a^{d/(d-1)}}{n^{1/(d-1)}}.$$

*Шаг 4. Оценка вероятности.* Напомним, что

$$K_a = \#J_0 + \# \left( \bigcup_{m=1}^d (J_m^+ \cup J_m^-) \right) =: K^{(0)} + K^{(\pm)}.$$

Заметим, что для некоторого случайного множества  $U \in \mathcal{U}$  выполнено:

$$N_U := \#\{i: \xi_i \in U\} \geq K^{(\pm)}.$$

Пусть  $c_6 := c_1 r^{-d}$ . Тогда, как мы помним,  $K^{(0)} \leq c_6 a^d$ .

Поэтому, пользуясь тем, что случайная величина  $N_U$  имеет распределение Пуассона с математическим ожиданием  $a^d \lambda \text{vol}_d(U)$ , получаем следующую оценку:

$$\begin{aligned} \mathbb{P}[K_a \geq n] &\leq \mathbb{P}[K^{(\pm)} \geq n - c_6 a^d] \leq \sum_{U \in \mathcal{U}} \mathbb{P}[N_U \geq n - c_6 a^d] \leq \#\mathcal{U} \cdot \max_{U \in \mathcal{U}} \mathbb{P}[N_U \geq n - c_6 a^d] \leq \\ &\leq \exp \left( \left( \frac{c_5 d}{(d-1)A^{d-1}} n \log n + \frac{c_5}{(d-1)A^{d-1}} n \log a^d \right) (1 + o(1)) \right) \left( \frac{a^d \lambda \text{vol}_d(U) e}{n - c_6 a^d} \right)^{n - c_6 a^d} \leq \\ &\leq \exp \left( \left( \frac{c_5 d}{(d-1)A^{d-1}} n \log n + \frac{c_5}{(d-1)A^{d-1}} n \log a^d \right) (1 + o(1)) \right) \cdot \\ &\quad \cdot \left( \frac{(a^d \lambda A c_5 a^{d/(d-1)} e) n^{-1/(d-1)}}{n - c_6 a^d} \right)^{n - c_6 a^d} = \\ &= \exp \left( \left( \left( \frac{c_5 d}{(d-1)A^{d-1}} - 1 - \frac{1}{d-1} \right) n \log n + \right. \right. \\ &\quad \left. \left. + \left( \frac{c_5}{(d-1)A^{d-1}} + 1 + \frac{1}{d-1} \right) n \log a^d \right) (1 + o(1)) \right). \end{aligned}$$

Здесь использовался тот факт, что  $a^d = o(n)$ , а также следующее равенство:

$$n \log(n - c_6 a^d) = n \log(n - c_6 a^d) - n \log n + n \log n = n \log(1 - c_6 a^d/n) + n \log n = n \log n + o(n).$$

Так как константа  $A$  может быть выбрана сколь угодно большой, получаем:

$$\mathbb{P}[K_a \geq n] \leq \exp \left( \left( - \left( 1 + \frac{1}{d-1} \right) n \log n + \left( 1 + \frac{1}{d-1} \right) n \log a^d \right) (1 + o(1)) \right).$$

□

## Список литературы

- [1] F. Aurzada, S. Dereich, M. Scheutzw, C. Vormoor, High resolution quantization and entropy coding of jump processes, J. Complexity 25 (2) (2009) 163–187.
- [2] F. Aurzada, M. Lifshits, How complex is a random picture? Journal of Complexity 53 (2019) 133–161.
- [3] S.N. Chiu, D. Stoyan, W.S. Kendall, J. Mecke, Stochastic Geometry and its Applications, Wiley Series in Probability and Statistics, third ed., John Wiley & Sons, Ltd., Chichester, 2013, p. xxvi+544.
- [4] T.M. Cover, J.A. Thomas, Elements of Information Theory, second ed., Wiley-Interscience [John Wiley & Sons], Hoboken, NJ, 2006, p. xxiv+748.
- [5] S. Dereich, The coding complexity of diffusion processes under supremum norm distortion, Stochastic Process. Appl. 118 (6) (2008) 917–937.
- [6] S. Graf, H. Luschgy, Foundations of Quantization for Probability Distributions, Lecture Notes in Mathematics, vol. 1730, Springer-Verlag, Berlin, 2000, p. x+230.
- [7] A.N. Kolmogorov, Three approaches to the quantitative definition of information, Int. J. Comput. Math. 2 (1968) 157–168.
- [8] R. Schneider, W. Weil, Stochastic and Integral Geometry, Probability and its Applications (New York), Springer-Verlag, Berlin, 2008, p. xii+693.