ФГБОУ ВО
Московский государственный университет имени М. В. Ломоносова
Механико-математический факультет

*На правах рукописи*

**ЩУКИН Владислав Юрьевич**

# Дизъюнктивные коды со списочным декодированием

01.01.05 – теория вероятностей и математическая статистика

АВТОРЕФЕРАТ
диссертации на соискание ученой степени
кандидата физико–математических наук

Москва – 2016

Работа выполнена на кафедре теории вероятностей
механико–математического факультета ФГБОУ ВО
«Московский государственный университет имени М.В. Ломоносова»

**Научный руководитель:**
доктор физико-математических наук, профессор
Дьячков Аркадий Георгиевич

**Официальные оппоненты:**
доктор физико-математических наук, профессор
Шоломов Лев Абрамович, главный научный сотрудник
лаборатории «Математические методы анализа и синтеза сложных систем»
Института системного анализа ФГУ «Федеральный исследовательский центр
«Информатика и управление» РАН»,

кандидат физико-математических наук, старший научный сотрудник
Лебедев Владимир Сергеевич, старший научный сотрудник Добрушинской
лаборатории ФГБУН «Институт проблем передачи информации имени
А.А. Харкевича РАН»

**Ведущая организация:**
ФГБУН «Институт вычислительных технологий Сибирского отделения РАН»

Защита диссертации состоится «17» марта 2017 г. в $16^{45}$ на заседании диссертационного совета Д 501.001.85 на базе МГУ имени М.В. Ломоносова по адресу: 119234, Москва, ГСП–1, Ленинские горы, д. 1, МГУ имени М. В. Ломоносова, механико-математический факультет, аудитория 16–24.

С диссертацией можно ознакомиться в Фундаментальной библиотеке
ФГБОУ ВО «Московский государственный университет имени
М.В. Ломоносова» (Москва, Ломоносовский проспект, д. 27, сектор А,
$8^{\text{й}}$ этаж) и на сайтах механико-математического факультета:
http://mech.math.msu.su/∼snark/index.cgi,
http://istina.msu.ru/dissertations/38221419.

Автореферат разослан «  » февраля 2017 г.

Ученый секретарь диссертационного
совета Д 501.001.85 на базе МГУ,
доктор физико–математических наук,
профессор — Власов Виктор Валентинович

# Общая характеристика работы

## Актуальность темы

Тематика данной диссертации лежит на стыке теории вероятностей, теории информации и комбинаторной теории кодирования. Основным объектом изучения являются семейства кодов, обслуживающие канал множественного доступа.

Математическая модель *канала множественного доступа* (КМД)[1] представляет из себя дискретный канал без памяти, имеющий $s$ входов и один выход. На входы поступают $q$-ичные символы $x_1, x_2, \ldots, x_s$ из алфавита $\{0, 1, \ldots, q-1\}$, а на выходе воспроизводится символ $y = f(x_1, x_2, \ldots, x_s)$. Возникает задача построения кодов для передачи сообщений через КМД и методов для их восстановления. Через КМД посимвольно передаются последовательности $q$-ичных символов длины $N$, представляющие из себя закодированные сообщения. Требуется возможность восстанавливать передаваемые сообщения по последовательности длины $N$ на выходе. Причем для фиксированного количества всех различных сообщений $t$ длина кодовых слов $N$ должна быть минимальной.

Особое место занимает модель *дизъюнктивного канала множественного доступа*[2], в которой $q = 2$, а функция $f$ представляет из себя дизъюнктивную сумму аргументов, т.е. принимает значение 0, если все (двоичные) сигналы на входе равны 0, и принимает значение 1 иначе. Благодаря значительному разнообразию приложений (групповое тестирование[3], поиск файлов в системах хранения[4], совокупные цифровые подписи[5] и другие) коды для дизъюнктивного КМД достаточно хорошо изучены.

В наиболее актуальной для приложений ситуации некоторые отправители могут молчать, т.е. число передаваемых сообщений $\leq s$. Для того, чтобы восстановить исходные кодовые слова по их дизъюнктивной сумме, необходимо и достаточно, чтобы дизъюнктивные суммы всех различных подмножеств кодовых слов мощности $\leq s$ отличались. Восстановление передаваемых кодовых слов путем перебора всех подмножеств мощности $\leq s$ и отысканием подмножества с дизъюнктивной суммой, совпадающей с выходом канала,

---

[1] *Чисар И.*, *Кернер Я.*, Теория информации. Теоремы кодирования для дискретных систем без памяти. Мир, Москва, 1985.

[2] *Дьячков А.Г.*, *Рыков В.В.*, Применение кодов для канала с множественным доступом в системе связи АЛОХА, *Тр. VI Всесоюзной школы-семинара по вычислительным сетям. Москва - Винница.*, Т. 4, С. 18-24, 1981.

[3] *Dorfman R.*, The Detection of Defective Members of Large Populations, *Ann. Math. Statist.*, vol. 14, no. 4, pp. 436-440, 1943.

[4] *Kautz W.H.*, *Singleton R.C.*, Nonrandom Binary Superimposed Codes, *IEEE Trans. Inform. Theory*, vol. 10, no. 4, pp. 363-377, 1964.

[5] *Hartung G.*, *Kaidel B.*, *Koch A.*, *Koch J.*, *Rupp A.*, Fault-Tolerant Aggregate Signatures, *Public-Key Cryptography – PKC 2016: 19th IACR International Conference on Practice and Theory in Public-Key Cryptography, Taipei, Taiwan, Proceedings, Part I*, 2016.

будем называть *переборным декодированием*[6], а удовлетворяющие данному свойству коды – *дизъюнктивными s-планами*[4,7].

Будем говорить, что двоичный столбец $\boldsymbol{u}$ *покрывает* двоичный столбец $\boldsymbol{v}$, если дизъюнктивная сумма $\boldsymbol{u}$ и $\boldsymbol{v}$ не отличается от $\boldsymbol{u}$, т.е. $\boldsymbol{u} \bigvee \boldsymbol{v} = \boldsymbol{u}$. Так называемое *пофакторное декодирование*[6], требующее немного большей длины кодовых слов при том же количестве всех различных сообщений, однако имеющее более простой и быстрый алгоритм восстановления сообщений, основано на *дизъюнктивных s-кодах*[4]. По определению, дизъюнктивная сумма любых $s$ кодовых слов дизъюнктивного $s$-кода не покрывает постороннего кодового слова. Таким образом, пофакторное декодирование заключается в поиске тех кодовых слов, которые покрываются выходом КМД.

Дизъюнктивные $s$-коды и $s$-планы были введены У. Каутсом и Р. Синглетоном в 1964 году в основополагающей статье[4], где также получены первые нетривиальные свойства и описан ряд прикладных задач. *Асимптотической скоростью* дизъюнктивных $s$-кодов ($s$-планов) будем называть величину

$$R(s) = \lim_{N\to\infty} \frac{\log_2 t(s,N)}{N} \quad \left( R(\le s) = \lim_{N\to\infty} \frac{\log_2 \tilde{t}(s,N)}{N} \right),$$

где $t(s,N)$ ($\tilde{t}(s,N)$) означает максимальный объем дизъюнктивных $s$-кодов ($s$-планов) длины $N$. В частности, в работе[4] показано

$$R(s) \le R(\le s) \le R(s-1), \quad R(\le s) \le \frac{1}{s}.$$

В случае $s = 2$ в 1982 году П. Эрдеш и др.[8] доказали оценки, из которых следуют неравенства

$$0.182 \le R(2) \le 0.322. \tag{1}$$

Эти неравенства представляют из себя наилучшие известные нижнюю и верхнюю границы для $R(2)$ и в настоящее время. В том же 1982 году А.Г. Дьячков и В.В. Рыков[9] иным методом вывели верхнюю границу, которая в случае $s = 2$ совпадает с правой частью (1), а при $s \to \infty$ асимптотически эквивалентна неравенству

$$R(s) \le \frac{2\log_2 s}{s^2}(1 + o(1)), \quad s \to \infty.$$

---

[6] *Фрейдлина В.Л.*, Об одной задаче планирования отсеивающих экспериментов, *Теор. вер. и ее приложения*, Т. 20, № 1, С. 100-114, 1975.

[7] *D'yachkov A.G.*, *Rykov V.V.*, A Survey of Superimposed Code Theory, *Problems of Control and Inform. Theory*, vol. 12, no. 4, pp. 229-242, 1983.

[8] *Erdos P.*, *Frankl P.*, *Furedi Z.*, Families of Finite Sets in Which No Set Is Covered by the Union of Two Others, *J. Combin. Theory, Ser. A*, vol. 33, no. 2, pp. 158-166, 1982.

[9] *Дьячков А.Г.*, *Рыков В.В.*, Границы длины дизъюнктивных кодов, *Пробл. передачи информ.*, Т. 18, № 3, С. 7-13, 1982.

Нижняя граница скорости $R(s)$ была получена в 1983 году А.Г. Дьячковым и В.В. Рыковым в работе[7], из которой следует асимптотическое неравенство

$$R(s) \geq \frac{\log_2 e}{es^2}(1+o(1)) = \frac{0.5307}{s^2}(1+o(1)), \quad s \to \infty. \tag{2}$$

Немного позже, в 1985 году, П. Эрдешем и др.[10] независимо от[7] выведена нижняя граница, которая для больших значений параметра $s$ ведет себя следующим образом:

$$R(s) \geq \frac{\log_2 e}{4s^2}(1+o(1)) = \frac{0.3607}{s^2}(1+o(1)), \quad s \to \infty.$$

А.Г. Дьячков и др.[11] впоследствие улучшили результат 1983 года[7] и в 1989 году новым методом доказали нижнюю границу для скорости $R(s)$, из которой при $s = 2$ следует левая часть неравенства (1), а при $s \to \infty$ асимптотическое неравенство:

$$R(s) \geq \frac{1}{s^2 \log_2 e}(1+o(1)) = \frac{0.6931}{s^2}(1+o(1)), \quad s \to \infty. \tag{3}$$

Для многих приложений достаточно более слабого свойства кода. Двоичный код называется *дизъюнктивным кодом со списочным декодированием силы $s$ с объемом списка $L$* (СД $s_L$-кодом), если дизъюнктивная сумма любых $s$ кодовых слов покрывает не более $L-1$ других кодовых слов. Таким образом, если кодировать сообщения с помощью СД $s_L$-кода, при пофакторном декодировании получим список кодовых слов, среди которых будут присутствовать все $\leq s$ переданных и $\leq L-1$ лишних кодовых слов. СД $s_L$-коды были введены в 1981 году А.Г. Дьячковым и В.В. Рыковым в работе[2], где рассматривалось использование таких кодов при передаче информации через дизъюнктивный КМД в системе связи АЛОХА. В работе П.А. Виленкина 1998 года[12] приведены некоторые конструкции СД $s_L$-кодов, а также рассмотрено их применение при построении двухступенчатых процедур групповых проверок.

Аналогичным образом введем *асимптотическую скорость* СД $s_L$-кодов:

$$R_L(s) = \lim_{N\to\infty} \frac{\log_2 t(s,L,N)}{N},$$

где через $t(s,L,N)$ обозначен максимальный объем СД $s_L$-кодов длины $N$. В 1983 году А.Г. Дьячков и В.В. Рыков[7] получили нижнюю и верхние границы

[10] *Erdos P., Frankl P., Furedi Z.*, Families of Finite Sets in Which No Set Is Covered by the Union of $r$ others, *Israel J. Math.*, vol. 51, no. 1, pp. 79-89, 1985.

[11] *D'yachkov A.G., Rykov V.V., Rashad A.M.* Superimposed Distance Codes, *Problems of Control and Inform. Theory*, vol. 18, no. 4, pp. 237-250, 1989.

[12] *Vilenkin P.A.*, On Constructions of List-Decoding Superimposed Codes, *Proc. 6th Int. Workshop on Algebraic and Combinatorial Coding Theory (ACCT-6), Pskov, Russia*, pp. 228-231, 1998.

скорости $R_L(s)$, из которых, в частности, следуют неравенства:

$$R_L(s) \le \frac{1}{s}, \quad \text{при любых натуральных } s \text{ и } L;$$

$$\frac{L\log_2 e}{es^2}(1+o(1)) \le R_L(s) \le \frac{2L^2\log_2 s}{s^2}(1+o(1)), \quad \text{при } s \to \infty.$$

Отметим, что (2) является частным случаем данной нижней границы для $R_L(s)$. Позднее, в 2003 году А.Г. Дьяков[13] получил новую нижнюю границу на скорость $R_L(s)$, из которой следует асимптотическое неравенство

$$R(s) \ge \frac{L}{s^2\log_2 e}(1+o(1)), \quad s \to \infty,$$

обобщающее (3). А в 2005 году А. Де Бонис и др.[14] улучшили верхнюю границу на скорость $R_L(s)$ для достаточно больших значений параметра $L$ и получили асимптотическое неравенство:

$$R_L(s) \le \frac{8L\log_2 s}{s^2}(1+o(1)), \quad s \to \infty.$$

Также для приложений допустимо использование алгоритмов, которые предполагают незначительную ошибку при восстановлении кодовых слов, переданных через КМД. В связи с этим, Э. Макула и др.[15] ввели понятие *почти дизъюнктивных* $(s,\varepsilon)$*-кодов*. По определению, у такого кода доля множеств из $s$ кодовых слов, дизъюнктивная сумма которых покрывает хотя бы одно другое кодовое слово, не превосходит $\varepsilon$. Таким образом, при использовании почти дизъюнктивных $(s,\varepsilon)$-кодов для кодирования сообщений, передаваемых через дизъюнктивный КМД, пофакторное декодирование не восстанавливает переданные кодовые слова с вероятностью $\le \varepsilon$.

В работе[16] приведены конструкции почти дизъюнктивных $(s,\varepsilon)$-кодов, основанные на укороченных кодах Рида-Соломона. Асимптотическое поведение ошибки $\varepsilon$ для данных конструкций посчитано в 2013 году Л.А. Бассалыго и В.В. Рыковым[17], откуда следует существование почти дизъюнктивных $(s,\varepsilon)$-кодов длины $N$ и объема $t$, таких что:

$$\frac{\log_2 t}{N} = \frac{\ln 2}{s}(1+o(1)), \quad s^2 = \Theta(N), \quad \varepsilon \to 0, \quad \text{при } N \to \infty. \qquad (4)$$

---

[13] *D'yachkov A.G.* Lectures on Designing Screening Experiments, *Lecture Note Series 10*, Combinatorial and Computational Mathematics Center, Pohang University of Science and Technology (POSTECH), Korea Republic, Feb. 2003 (survey, 112 pages). http://arxiv.org/pdf/1401.7505

[14] *De Bonis A.*, *Gasieniec L.*, *Vaccaro U.*, Optimal Two-Stage Algorithms for Group Testing Problems, *SIAM J. Comput.*, vol. 34, no. 5, pp. 1253-1270, 2005.

[15] *Macula A.J.*, *Rykov V.V.*, *Yekhanin S.*, Trivial two-stage group testing for complexes using almost disjunct matrices, *Discret. Appl. Math.*, vol. 137, no. 1, pp. 97-107, 2004.

[16] *D'yachkov A.G.*, *Macula A.J.*, *Rykov V.V.*, New constructions of superimposed codes, *IEEE Trans. Inform. Theory*, vol. 46, no. 1, pp. 284-290, 2000.

[17] *Бассалыго Л.А.*, *Рыков В.В.*, Гиперканал множественного доступа, *Пробл. передачи информ.*, Т. 49, № 4, С. 3-12, 2013.

Обобщением почти дизъюнктивных $(s,\varepsilon)$-кодов, а также СД $s_L$-кодов, выступают *почти дизъюнктивные* $(s_L,\varepsilon)$*-коды со списочным декодированием* (СД $(s_L,\varepsilon)$-коды), для которых доля множеств из $s$ кодовых слов, дизъюнктивная сумма которых покрывает $\geq L$ других кодовых слов, не превосходит $\varepsilon$. Под *экспонентой ошибки* $\mathbf{E}_L(s,R)$ будем подразумевать максимальный показатель экспоненциального убывания ошибки $\varepsilon$ для последовательности СД $(s_L,\varepsilon)$-кодов возрастающей длины и фиксированной скорости $R$, а под *пропускной способностью* $C_L(s)$ – верхнюю грань значений скорости $R$, для которых экспонента ошибки $\mathbf{E}_L(s,R)$ положительна. В недавней работе[18] И.В. Воробьевым доказана нижняя граница для пропускной способности:

$$C_L(s) \geq \frac{\ln 2}{s}, \quad s \to \infty.$$

Вызывает интерес, что данная нижняя граница для фиксированного значения параметра $s$ совпадает со скоростью конструкций (4), где, однако, параметр $s$ зависит от длины кода.

Рассмотрим теперь ситуацию, когда число кодовых слов поступивших на входы дизъюнктивного КМД неизвестно (и не удовлетворяет условию $\leq s$). Наиболее наглядно такой случай представляется на примере модели группового тестирования с неизвестным числом дефектов. Предположим, что задано множество из $t$ элементов, среди которых присутствует неизвестное число $s_{un}$, $0 \leq s_{un} \leq t$, дефектных элементов. Требуется найти все дефекты за минимальное число групповых тестов, где под групповым тестом подразумевается некоторое подмножество элементов, а результат теста равен 1, если хотя бы один дефектный элемент попал в тестируемое множество, и 0 – иначе. Процедуру группового тестирования называют адаптивной, если каждый следующий тест строится, исходя из результатов предыдущих, и неадаптивной, если все тесты формируются изначально и могут проводиться одновременно. Различают также $k$-ступенчатые процедуры групповых проверок, для которых формирование тестов на $l$-й ступени, $1 \leq l \leq k$, основывается на результатах тестов на ступенях $1, 2, \ldots, l-1$. В случае неадаптивного алгоритма $N$ тестов представляют в виде двоичной матрицы из $N$ строк и $t$ столбцов, в которой каждая строка сопоставляется некоторому тесту, каждый столбец – некоторому элементу, а на пересечении $i$-й строки и $j$-го столбца стоит 1 тогда и только тогда, когда $j$-й элемент включен в $i$-й тест. Очевидно, что столбец результатов тестов равен дизъюнктивной сумме столбцов, соответствующих дефектным элементам.

Большинство разработанных неадаптивных алгоритмов группового тестирования предполагают ограничение на количество дефектов: $s_{un} \leq s$, однако в случае отсутствия такого предположения возникает необходимость прибегнуть к $k$-ступенчатым процедурам групповых проверок. В 2002 году Т. Бергер

[18] *Дьячков А.Г., Воробьев И.В., Полянский Н.А., Щукин В.Ю*, Почти дизъюнктивные коды со списочным декодированием, *Пробл. передачи информ.*, Т. 51, № 2, С. 27-49, 2015.

и В.И. Левенштейн[19] рассматривали применение *двухступенчатых процедур восстановления* в модели группового тестирования, в которой каждый элемент является дефектным с вероятностью $p$. В такой процедуре на первом этапе применяется некоторое количество неадаптивных тестов, а на второй ступени по одиночке проверяется каждый элемент из числа тех, что остались подозрительными после первой ступени. Для некоторых случаев зависимости вероятности $p$ от общего количества элементов $t$ в работе[19] получены как верхние, так и нижние границы для асимптотики математического ожидания общего количества тестов $N$. Например,

$$\frac{\log_2 e}{4}\frac{(\ln t)^2}{\ln\ln t}(1+o(1)) \le E[N] \le \frac{(\ln t)^2}{\ln\ln t}(1+o(1)), \quad p(t) = \frac{1}{t}(1+o(1)), \quad t \to \infty.$$

Отметим, что для случая $p = 1/t^{-\beta}$, $0 < \beta < 1$, верхние и нижние границы для $E[N]$ были улучшены в работе[20].

Другой подход к задаче группового тестирования с неизвестным числом дефектных элементов был предложен в 2010 году П. Дамашке и А.Ш. Мухаммадом[21]. Для начала с помощью групповых тестов производится оценивание количества дефектов, а затем, на основе полученной оценки, применяется один из известных алгоритмов для поиска ограниченного числа дефектов. В статье[21] авторы построили случайную конструкцию неадаптивной процедуры групповых проверок, с помощью которой за $G(\varepsilon, c)\log_2 t$ тестов определяется статистика $\hat{s}$, удовлетворяющая следующим условиям: вероятность $\Pr\{\hat{s} < s_{un}\}$ ограничена сверху параметром $\varepsilon \ll 1$, а математическое ожидание величины $\hat{s}/s_{un}$ ограничено сверху параметром $c > 1$. Отметим, что указанный результат является *универсальным*, то есть не зависит от распределения множества дефектных элементов. Адаптивный алгоритм для получения похожей оценки для количества дефектных элементов описан в статье[22].

Рассмотрим теперь другую модель КМД. КМД называется $q$-ичным *гиперканалом множественного доступа*[17,23] (ГМД), если при поступлении на его $s$ входов $q$-ичных символов $x_1, x_2, \ldots, x_s$, на выходе получим *гиперсумму* этих символов, т.е. множество всех поступивших на входы символов:

$$\bigcup_{k=1}^{s}\{x_k\}.$$

[19] *Berger T.*, *Levenshtein V.I.*, Asymptotic Efficiency of Two-Stage Disjunctive Testing, *IEEE Trans. Inform. Theory*, vol. 48, no. 7, pp. 1741-1749, 2002.

[20] *Mezard M.*, *Toninelli C.*, Group Testing With Random Pools: Optimal Two-Stage Algorithms, *IEEE Trans. Inform. Theory*, vol. 57, no. 3, pp. 1736-1745, 2011.

[21] *Damaschke P.*, *Muhammad A.S.*, Competitive group testing and learning hidden vertex covers with minimum adaptivity, *Discrete Math. Algorithm. Appl.*, vol. 2, no. 3, pp. 291-311, 2010.

[22] *Falahatgar M.*, *Jafarpour A.*, *Orlitsky A.*, *Pichapati V.*, *Suresh A.T.*, Estimating the Number of Defectives with Group Testing, *Proc. IEEE Int. Symp. Inf. Theory*, Barcelona, pp. 1376-1380, 2016.

[23] *Chang S.-C.*, *Wolf J.*, On the T-user M-frequency noiseless multiple-access channel with and without intensity information, *IEEE Trans. Inform. Theory*, vol. 27, no. 1, pp. 41-48, 1981.

Отметим, что данная модель КМД имеет различные названия в научной литературе, причем, зачастую, наблюдается отсутствие ссылок на раннее вышедшие работы. Так, в статье[23] ГМД называют A channel.

Последовательности, символами которых являются подмножества, будем называть *гиперсловами*. По аналогии с дизъюнктивной моделью КМД введем следующие понятия. Гиперслово $\boldsymbol{u}$ *подчиняет* гиперслово $\boldsymbol{v}$, если гиперсумма $\boldsymbol{u}$ и $\boldsymbol{v}$ не отличается от $\boldsymbol{u}$. $q$-ичный код называется *$s$-гиперкодом*, если гиперсумма произвольных $s$ кодовых слов не подчиняет другого кодового слова; *$s$-гиперпланом*, если гиперсумма произвольных $\leq s$ кодовых слов отличается от гиперсуммы любого другого набора из $\leq s$ кодовых слов. В англоязычной литературе $s$-гиперкод и $s$-гиперплан обычно называют $s$-frameproof code и $s$-separable code, соответственно. *Асимптотической скоростью* $q$-ичных $s$-гиперкодов ($q$-ичных $s$-гиперпланов) будем называть величину

$$R^{(q)}(s) = \lim_{N\to\infty} \frac{\log_q t^{(q)}(s,N)}{N} \quad \left( R^{(q)}(\leq s) = \lim_{N\to\infty} \frac{\log_2 \tilde{t}^{(q)}(s,N)}{N} \right),$$

где $t^{(q)}(s,N)$ ($\tilde{t}^{(q)}(s,N)$) означает максимальный объем $q$-ичных $s$-гиперкодов ($q$-ичных $s$-гиперпланов) длины $N$.

Мотивированные возможностью использования для защиты авторских прав на цифровую продукцию от недобросовестного распространения, $q$-ичные $s$-гиперкоды были введены Д. Боне и Д. Шоу в 1998 году[24]. Отметим, что двоичные $s$-гиперкоды, или *симметричные дизъюнктивные $s$-коды*, рассматривались и ранее[2]. А. Хан Винк и С. Мартиросян в 2000 году[25] и С. Блэкберн в 2003 году[26] независимо построили некоторые конструкции $s$-гиперкодов, основанные на укороченных кодах Рида-Соломона. Примечательно, что в недавних статьях[27,28] Т. Ван Чунг и М. Базрафшан доказали оптимальность данных конструкций.

Также в работах[25,26,29] независимо получена верхняя граница на скорость $s$-гиперкодов:

$$R^{(q)}(s) \leq \frac{1}{s}. \tag{5}$$

В недавней работе[30] 2014 года Ч. Шенгуэн и др. построили новую верхнюю

[24] *Boneh D.*, *Shaw J.*, Collusion-secure fingerprinting for digital data, *IEEE Trans. Inform. Theory*, vol. 44, no. 5, pp. 1897-1905, 1998.

[25] *Vinck A.J. Han*, *Martirossian S.*, On Superimposed Codes, Numbers, Information and Complexity, Springer US, pp. 325-331, 2000.

[26] *Blackburn S.R.*, Frameproof codes, *SIAM J. Discrete Math.*, vol. 16, no. 3, pp. 499-510, 2003.

[27] *Bazrafshan M.*, *van Trung T.*, Improved bounds for separating hash families, *Des. Codes Cryptogr.*, vol. 69, no. 3, pp. 369-382, 2013.

[28] *van Trung T.*, A tight bound for frameproof codes viewed in terms of separating hash families, *Des. Codes Cryptogr.*, vol. 72, no. 3, pp. 713-718, 2014.

[29] *Cohen G.D.*, *Schaathun H.G.*, Asymptotic overview on separating codes, *Tech. Report 248*, Department of Informatics, University of Bergen, Bergen, Norway, 2003.

[30] *Shangguan C.*, *Wang X.*, *Ge G.*, *Miao Y.*, New Bounds For Frameproof Codes, Preprint, 2014. http://arxiv.org/pdf/1411.5782v1

границу, которая улучшает (5) для больших значений параметра $s$ и порождает асимптотическое неравенство

$$R^{(q)}(s) \le \frac{4(q-1)\log_q s}{s^2}(1+o(1)), \quad s \to \infty.$$

В 2008 году Д. Стинсон и др.[31] вероятностным методом построили нижнюю границу на скорость $R^{(q)}(s)$:

$$R^{(q)}(s) \ge \frac{1}{s}\log_q \left\lceil \frac{q^s}{q^s-(q-1)^s} \right\rceil, \quad q \ge 2, \quad s \ge 2, \tag{6}$$

асимптотическое поведение которой при $s \to \infty$ описывается выражением $O\left(\frac{1}{s}\left(1-\frac{1}{q}\right)^s\right)$. Оптимизируя метод случайного кодирования из[31], авторы уже упоминаемой работы[30] улучшили данную нижнюю границу (6) для больших значений параметра $s$ и получили асимптотическое неравенство

$$R^{(q)}(s) \ge \frac{q-1}{s^2 e \ln q}(1+o(1)), \quad s \to \infty.$$

Отметим, что нижняя граница (6) достигает верхнюю границу (5) при росте $q$, откуда следует

$$\lim_{q\to\infty} R^{(q)}(s) = \frac{1}{s}.$$

Доказательство этого результата получено и ранее[29] и опирается на конструкциях $q$-ичных $s$-гиперкодов, основанных на алгебро-геометрических кодах.

В 2011 году М. Ченг и Ин Мяо[32] ввели $s$-гиперпланы в контексте защиты авторских прав на цифровую продукцию и идентификации недобросовестных пользователей, а также установили следующие соотношения между скоростями $s$-гиперкодов и $s$-гиперпланов:

$$R^{(q)}(s) \le R^{(q)}(\le s) \le R^{(q)}(s-1). \tag{7}$$

Границы для скорости $s$-гиперпланов, вытекающие из предыдущего неравенства (7) и известных границ для $R^{(q)}(s)$, были улучшены в недавних работах[33,34] для частного случая $s=2$, где получено неравенство $R^{(q)}(2) \le 2/3$.

Наряду с СД $s_L$-кодами, описанными выше, рассматривают *$q$-ичные гиперкоды со списочным декодированием силы $s$ с объемом списка $L$* (кратко, $q$-ичные СД $s_L$-гиперкоды). По определению, гиперсумма произвольных

[31] *Stinson D.R., Wei R., Chen K.*, On generalized separating hash families, *J. Combin. Theory, Ser. A*, vol. 115, no. 1, pp. 105-120, 2008.

[32] *Cheng M., Miao Y.*, On Anti-Collusion Codes and Detection Algorithms for Multimedia Fingerprinting, *IEEE Trans. Inform. Theory*, vol. 57, no. 7, pp. 4843-4851, 2011.

[33] *Gao F., Ge G.*, New Bounds on Separable Codes for Multimedia Fingerprinting, *IEEE Trans. Inform. Theory*, vol. 60, no. 9, pp. 5257-5262, 2014.

[34] *Blackburn S.R.*, Probabilistic Existence Results for Separable Codes, *IEEE Trans. Inform. Theory*, vol. 61, no. 11, pp. 5822-5827, 2015.

$s$ кодовых слов $q$-ичного СД $s_L$-гиперкода подчиняет не более $L-1$ других кодовых слов. По аналогии с рассмотренными ранее семействами кодов, вводится *асимптотическая скорость* СД $s_L$-гиперкодов, обозначим ее через $R_L^{(q)}(s)$. Для двоичного случая в 1989 году А. Рашад[35] построил нижнюю границу случайного кодирования для скорости $R_L^{(2)}(s)$, используя ансамбль кодов с независимыми одинаково распределенными двоичными компонентами кодовых слов. Асимптотическое поведение данной границы при $s \to \infty$ описывается неравенством:

$$R_L^{(2)}(s) \geq \frac{L \log_2 e}{e s^2}(1 + o(1)), \quad s \to \infty.$$

### Цель работы

Целью диссертационной работы является разработка теоретико-вероятностных и комбинаторных методов для построения более точных оценок асимптотики важных теоретико-информационных характеристик кодов, обслуживающих канал множественного доступа.

### Научная новизна

Все результаты, представленные в диссертации, являются новыми. В работе впервые исследуется задача сравнения количества дефектных элементов с заданной константой в дизъюнктивной модели группового тестирования. Также в работе впервые рассматриваются $q$-ичные СД $s_L$-гиперкоды с мощностью алфавита $q > 2$ и объемом списка $L > 1$.

### Основные результаты работы

Основные результаты, полученные в диссертации, перечислены ниже.

1. Установлена новая нижняя граница для асимптотической скорости $R_L(s)$ дизъюнктивных кодов со списочным декодированием, которая является наилучшей и в случае $L > 1$ превосходит наилучшую ранее известную нижнюю границу, полученную А.Г. Дьячковым[13].

2. Впервые получена нижняя граница для экспоненты ошибки $\mathbf{E}_L(s, R)$ почти дизъюнктивных кодов со списочным декодированием.

3. Впервые получена верхняя граница для пропускной способности $C_L(s)$ почти дизъюнктивных кодов со списочным декодированием.

[35] *Rashad A.M.*, On Symmetrical Superimposed Codes, *J. Inf. Process. Cybern EIK 29*, vol. 7, pp. 337-341, 1989.

4. Предложен нестандартный алгоритм декодирования результатов дизъюнктивных групповых тестов в задаче сравнения количества дефектных элементов с заданной константой, который по сравнению со стандартными алгоритмами позволяет использовать меньшее количество тестов для получения ответа с малой вероятностью ошибки. Для нового алгоритма получена нижняя граница на экспоненту вероятности ошибки.

5. Доказаны новые нижние границы для асимптотической скорости $R_L^{(q)}(s)$ гиперкодов со списочным декодированием, которые улучшают ранее известные нижние границы, полученные Д. Стинсоном и др.[31], Ч. Шенгуэном и др.[30], А. Рашадом[35].

6. Выведена новая верхняя граница на асимптотическую скорость $R_L^{(q)}(s)$ гиперкодов со списочным декодированием, которая является наилучшей для достаточно больших значений параметра $s$ и улучшает ранее известную верхнюю границу, доказанную Ч. Шэнгуэном и др.[30].

### Основные методы исследования

В работе используются вероятностные методы, в частности метод случайного кодирования для ансамбля равновесных кодов. Для вычисления логарифмических асимптотик вероятностей больших уклонений применяются классические методы выпуклого анализа и аналитические методы. В работе также используются методы комбинаторной теории кодирования.

### Теоретическая и практическая ценность работы

Результаты диссертации носят теоретический характер. Они могут быть полезны специалистам в области теории вероятностей, комбинаторной теории кодирования и теории информации.

### Апробация диссертации

Результаты диссертации неоднократно докладывались автором на следующих научно-исследовательских семинарах.

1. Спецсеминар “Экстремальная комбинаторика и случайные структуры” в 2016 г., кафедра теории вероятностей, мехмат, МГУ.

2. Спецсеминар “Проблемы современной теории информации” в 2013–2016 гг., кафедра теории вероятностей, мехмат, МГУ.

3. Семинар по теории кодирования под рук. Л.А. Бассалыго в 2013–2016 гг., ИППИ РАН.

4. Семинар по дискретной математике под рук. М.В Вялого и С.П. Тарасова в 2016 г., ВЦ РАН.

Результаты диссертации докладывались автором на следующих конференциях.

1. Конференция "*Ломоносов–2013*", Москва, 2013.
2. 14th International Workshop "*Algebraic and Combinatorial Coding Theory*", Svetlogorsk, Russia, 2014.
3. Ninth International Workshop on Coding and Cryptography, Paris, France, 2015.
4. IEEE International Symposium on Information Theory, Hong Kong, China, 2015.
5. Конференция "*Ломоносов–2016*", Москва, 2016.
6. 15th International Workshop "*Algebraic and Combinatorial Coding Theory*", Albena, Bulgaria, 2016.

### Публикации

Основные результаты настоящей диссертации опубликованы в работах [1]-[13], представленных в конце списка литературы. Среди них 6 работ [1]-[6] в журналах из перечня ВАК и 7 работ [7]-[13] в рецензируемых трудах международных конференций.

### Структура и объем диссертации

Диссертация состоит из введения, четырех глав, заключения и списка литературы, который включает 56 наименований. Объем диссертации составляет 77 страниц.

### Краткое содержание диссертации

Во **введении** определены основные объекты исследования, представлен краткий исторический обзор результатов, а также приведено краткое содержание данной диссертации.

В **главе 1** рассматриваются коды для дизъюнктивного канала множественного доступа.

В разделе 1.1 введены используемые в главе 1 обозначения и определения.

Следующий раздел 1.2 посвящен нижним границам скорости дизъюнктивных СД $s_L$-кодов. С помощью метода случайного кодирования получена следующая теорема, устанавливающая наилучшие нижние границы для скорости СД $s_L$-кодов.

**Теорема 1** (Граница случайного кодирования $\underline{R}_L(s)$)**.** *Имеют место следующие три утверждения.*
**1.** *Для скорости СД $s_L$-кодов справедливо неравенство*

$$R_L(s) \geq \underline{R}_L(s) \triangleq \frac{1}{s+L-1} \max_{0<Q<1} A_L(s,Q),$$

$$A_L(s,Q) \triangleq \log_2 \frac{Q}{1-y} - sK(Q,1-y) - L\,K\left(Q, \frac{1-y}{1-y^s}\right), \quad s \geq 2,\ L \geq 1,$$

*где используется обозначение расстояния Кульбака:*

$$K(a,b) \triangleq a \cdot \log_2 \frac{a}{b} + (1-a) \cdot \log_2 \frac{1-a}{1-b}, \quad 0 < a, b < 1,$$

*а параметр $y$, $1-Q \leq y < 1$, определяется как единственный корень уравнения*

$$y = 1 - Q + Qy^s \left[1 - \left(\frac{y-y^s}{1-y^s}\right)^L\right], \quad 1-Q \leq y < 1.$$

**2.** *При фиксированном $L = 1, 2, \ldots$ и $s \to \infty$ асимптотика границы случайного кодирования имеет вид*

$$\underline{R}_L(s) = \frac{L}{s^2 \log_2 e}(1 + o(1)), \quad s \to \infty.$$

**3.** *При фиксированном $s = 2, 3, \ldots$ и $L \to \infty$ существует предел*

$$\underline{R}_\infty(s) \triangleq \lim_{L\to\infty} \underline{R}_L(s) = \log_2 \left[\frac{(s-1)^{s-1}}{s^s} + 1\right].$$

*Если $s \to \infty$, то данный предел $\underline{R}_\infty(s) = \frac{\log_2 e}{e \cdot s}(1 + o(1)) = \frac{0{,}5307}{s}(1 + o(1))$.*

При доказательстве используется ансамбль $E(N,t,Q)$, $0 < Q < 1$, равновесных двоичных кодов длины $N$ и объема $t$, для которых кодовые слова выбираются независимо и равновероятно из множества всех двоичных кодовых слов фиксированных длины $N$ и веса $\lfloor QN \rfloor$. При фиксированном объеме кода $t$ рассматривается вероятность $P(s,L,Q,N)$ *плохого события*: дизъюнктивная сумма фиксированного множества из $s$ кодовых слов покрывает дизъюнктивную сумму фиксированного множества из $L$ кодовых слов. Далее, через $P(s,L,Q,N)$ оценивается математическое ожидание количества кодовых слов, после удаления которых код становится СД $s_L$-кодом. Данная оценка приводит к нижней границе на скорость СД $s_L$-кодов:

$$R_L(s) \geq \underline{R}_L(s) \triangleq \frac{1}{s+L-1} \max_{0<Q<1} \overline{\lim_{N\to\infty}} \frac{-\log_2 P(s,L,Q,N)}{N}.$$

Используя терминологию типов последовательностей, задача нахождения логарифмической асимптотики вероятности $P(s, L, Q, N)$ сводится к поиску минимума функционала

$$F(\tau, Q) \triangleq \sum_{\boldsymbol{a}\in\{0,1\}^s} \tau(\boldsymbol{a}) \log_2 [\tau(\boldsymbol{a})] - (1 - \tau(\boldsymbol{0}))\, L \cdot h\left(\frac{Q}{1-\tau(\boldsymbol{0})}\right) + (s + L)h(Q),$$
$$h(Q) \triangleq -Q \log_2 Q - (1 - Q) \log_2(1 - Q),$$

в области с линейными ограничениями на распределение $\{\tau(\boldsymbol{a})\}$, $\boldsymbol{a} \in \{0, 1\}^s$. Экстремальная задача решается стандартным методом множителей Лагранжа.

В разделе 1.3 сформулированы наилучшие верхние границы на скорость СД $s_L$-кодов.

В следующей **главе 2** рассматриваются почти дизъюнктивные коды со списочным декодированием. Основные определения вводятся в разделе 2.1. Под ошибкой $\varepsilon_L(s, R, N)$ будем подразумевать минимизированную по всем кодам $X$ длины $N$ и объема $t = \lfloor 2^{RN} \rfloor$ долю $s$-множеств кодовых слов, дизъюнктивная сумма которых покрывает $\geq L$ посторонних слов кода $X$. Экспонентой ошибки $\mathbf{E}_L(s, R)$ почти дизъюнктивных СД $s_L$-кодов и пропускной способностью $C_L(s)$ почти дизъюнктивных СД $s_L$-кодов назовем

$$\mathbf{E}_L(s, R) \triangleq \varlimsup_{N\to\infty} \frac{-\log_2 \varepsilon_L(s, R, N)}{N}, \quad R > 0, \text{ и}$$

$$C_L(s) \triangleq \sup\{\, R \,:\, \mathbf{E}_L(s, R) > 0\},$$

соответственно.

В разделе 2.2 с помощью метода случайного кодирования на ансамбле $E(N, t, Q)$, $0 < Q < 1$, равновесных двоичных кодов выведена нижняя граница на экспоненту ошибки почти дизъюнктивных СД $s_L$-кодов. Для начала, определим положительную часть функции:

$$[x]^+ \triangleq \begin{cases} x, & \text{если } x \geq 0, \\ 0, & \text{если } x < 0, \end{cases}$$

а через $\underline{C}(s)$ обозначим нижнюю границу на пропускную способность $C_L(s)$, полученную И.В. Воробьевым в работе[18]:

$$C_L(s) \geq \underline{C}(s) \triangleq \max_{0<Q<1} \left\{ h(Q) - [1 - (1 - Q)^s]\, h\left(\frac{Q}{1-(1-Q)^s}\right)\right\}.$$

**Теорема 2** (Нижняя граница для $\mathbf{E}_L(s, R)$)**.** *Справедливы 3 утверждения.*
**1.** *Величина* $\mathbf{E}_L(s, R)$ *удовлетворяет неравенству:*

$$\mathbf{E}_L(s, R) \geq \underline{\mathbf{E}}_L(s, R) \triangleq \max_{0<Q<1} E_L(s, R, Q), \quad s \geq 2,\ L \geq 1,\ R > 0,$$

$$E_L(s,R,Q) \triangleq \min_{Q\le q\le \min\{1,sQ\}} \left\{\mathcal{A}(s,Q,q) + L\cdot[h(Q) - q\cdot h(Q/q) - R]^+\right\},$$

*где функция* $\mathcal{A}(s,Q,q)$, $Q<q<\min\{1,sQ\}$, *определена следующим образом:*

$$\mathcal{A}(s,Q,q) \triangleq (1-q)\log_2(1-q) + q\log_2\left[\frac{Qy^s}{1-y}\right] + sQ\log_2\frac{1-y}{y} + sh(Q),$$

*а число* $y$ *в определении* $\mathcal{A}(s,Q,q)$ *является единственным решением уравнения*

$$q = Q\frac{1-y^s}{1-y}, \quad 0<y<1.$$

**2.** *Для любых* $s\ge 2$ *и* $L\ge 1$ *нижняя граница* $\underline{\mathbf{E}}_L(s,R)$, *является* $\cup$*-выпуклой функцией параметра* $R>0$. *При* $0<R<\underline{C}(s)$ *справедливо неравенство* $\underline{\mathbf{E}}_L(s,R)>0$. *Если* $R\ge\underline{C}(s)$, *то* $\underline{\mathbf{E}}_L(s,R)=0$. *Кроме того, существует такое число* $\underline{R}_L^{(cr)}(s)$, $0\le\underline{R}_L^{(cr)}(s)<\underline{C}(s)$, *что*

$$\underline{\mathbf{E}}_L(s,R) = (s+L-1)\underline{R}_L(s) - LR, \quad при \quad 0\le R\le \underline{R}_L^{(cr)}(s), \tag{8}$$

*и*

$$\underline{\mathbf{E}}_L(s,R) > (s+L-1)\underline{R}_L(s) - LR, \quad при \quad R> \underline{R}_L^{(cr)}(s),$$

*где граница случайного кодирования* $\underline{R}_L(s)$ *определена в теореме 1. Причем, прямая, задаваемая* (8), *является касательной к функции* $\underline{\mathbf{E}}_L(s,R)$ *в точке* $R=\underline{R}_L^{(cr)}(s)$.

**3.** *При любых* $s\ge 2$, $L\ge 1$ *и* $R\ge R_L^{(cr)}(s)$ *справедливо равенство* $\underline{\mathbf{E}}_L(s,R) = \underline{\mathbf{E}}_{L+1}(s,R)$, *а последовательности скорости* $R_L^{(cr)}(s)$ *и соответствующего значения нижней границы экспоненты ошибки имеют пределы:*

$$\lim_{L\to\infty}\underline{R}_L^{(cr)}(s) = \underline{R}_\infty(s),$$
$$\lim_{L\to\infty}\underline{\mathbf{E}}_L(s,\underline{R}_L^{(cr)}(s)) = (s-1)\underline{R}_\infty(s),$$

*где скорость* $\underline{R}_\infty(s)$ *определена в теореме 1.*

При доказательстве теоремы 2 рассматривается случайный код $X$ длины $N$ и объема $t$, распределение компонент которого подчиняется ансамблю $E(N,t,Q)$. Очевидные соображения, приводят к нижней границе на экспоненту ошибки почти дизъюнктивных СД $s_L$-кодов:

$$\mathbf{E}_L(s,R) \ge \varlimsup_{N\to\infty}\frac{-\log_2 P(s,L,Q,N,t)}{N},$$

где $P(s,L,Q,N,t)$ – вероятность того, что дизъюнктивная сумма фиксированного $s$-множества кодовых слов покрывает $L$ других кодовых слов кода $X$.

Затем, вероятность $P(s, L, Q, N, t)$ выражается по формуле полной вероятности для следующей группы несовместных событий: $B_k \triangleq \{$ вес дизъюнктивной суммы $s$ фиксированных кодовых слов равен $k\ \}, 0 \le k \le N$. Логарифмическая асимптотика вероятности события $B_{\lfloor qN \rfloor}$, $0 < q < 1$, вычисляется по аналогии с доказательством теоремы 1, использованием терминологии типов последовательностей и решением экстремальной задачи.

Следующий раздел 2.3 содержит теорему о верхней границе для пропускной способности почти дизъюнктивных СД $s_L$-кодов, доказательство которой основано на комбинаторных соображениях.

**Теорема 3** (Верхняя граница для $C_L(s)$)**.** *Справедливо неравенство*

$$C_L(s) \le 1/s, \qquad s \ge 1, \quad L \ge 1.$$

В **главе 3** рассматривается задача проверки гипотез о количестве отправителей сообщений через дизъюнктивный канал множественного доступа. Пусть объем кода, а соответственно и количество всех различных сообщений, равен $t$. Через $\mathcal{S}_{un}, \mathcal{S}_{un} \subset [t]$, обозначим неизвестное множество передаваемых сообщений. Рассматривается задача проверки гипотезы $\{H_0 : |\mathcal{S}_{un}| \le s\}$ против альтернативы $\{H_1 : |\mathcal{S}_{un}| \ge s+1\}$ для некоторой фиксированной константы $s$.

В разделе 3.1 приведены основные определения, причем описанная выше задача формулируется для дизъюнктивной модели группового тестирования. В такой постановке исследуемая величина $|\mathcal{S}_{un}|$ обозначает количество дефектов среди множества из $t$ элементов, а альтернатива $H_1$ является событием превышения количества дефектных элементов заданного порога $s$. В этом же разделе задача безошибочной проверки гипотезы $H_0$ против $H_1$ с помощью неадаптивных групповых тестов сведена к исследованию дизъюнктивных кодов и применению пофакторного декодирования для проверки гипотез:

**Предложение 1.** *Результаты неадаптивных групповых тестов, заданных кодом $X$, позволяют безошибочно проверить гипотезу $H_0$ против альтернативы $H_1$ в том и только том случае, если код $X$ является СД $s_1$-кодом.*

Далее рассматривается вероятностная постановка задачи, в которой для любого закона распределения с равновероятными множествами дефектов одинакового объема допустима лишь незначительная ошибка при проверке гипотез. Введен важный алгоритм, называемый *пороговым декодированием*:

$$\begin{cases} \text{принять } \{H_0 : |\mathcal{S}_{un}| \le s\}, & \text{если } |\boldsymbol{x}(\mathcal{S}_{un})| \le \lfloor \tau N \rfloor, \\ \text{принять } \{H_1 : |\mathcal{S}_{un}| \ge s+1\}, & \text{если } |\boldsymbol{x}(\mathcal{S}_{un})| \ge \lfloor \tau N \rfloor + 1, \end{cases}$$

где через $|\boldsymbol{x}(\mathcal{S}_{un})|$ обозначено количество тестов с положительным результатом, $N$ – общее количество тестов, а $\tau$, $0 < \tau < 1$, – заданная константа. *Экспонентой ошибки* $\mathbf{E}_s(\tau, R)$ для порогового декодирования называется

максимальный показатель экспоненциального убывания ошибок первого и второго рода при росте длины кодов $N$ и при фиксированной скорости кодов $R$.

В разделе 3.2 с помощью метода случайного кодирования на ансамбле равновесных двоичных кодов доказана следующая теорема.

**Теорема 4** (Нижняя граница для $\mathbf{E}_s(\tau, R)$)**.** *Справедливы два утверждения.*
**1.** *Экспонента ошибки порогового критерия удовлетворяет неравенству* $\mathbf{E}_s(\tau, R) \geq \underline{\mathbf{E}}_s(\tau)$ *где функция* $\underline{\mathbf{E}}_s(\tau)$ *не зависит от параметра* $R$ *и определяется как*

$$\underline{\mathbf{E}}_s(\tau) \triangleq \max_{1-(1-\tau)^{1/(s+1)}<Q<1-(1-\tau)^{1/s}} \min\left\{\mathcal{A}'(s, Q, \tau), \mathcal{A}(s+1, Q, \tau)\right\} > 0,$$

$$\mathcal{A}'(s, Q, \tau) \triangleq \begin{cases} \mathcal{A}(s, Q, \tau), & \text{если } Q \leq \tau \leq sQ, \\ \infty, & \text{иначе,} \end{cases}$$

*а функция* $\mathcal{A}(s, Q, \tau)$ *задана в формулировке теоремы 2.*
**2.** *При* $s \to \infty$ *оптимальное значение* $\underline{\mathbf{E}}_s(\tau)$ *удовлетворяет неравенству:*

$$\underline{\mathbf{E}}_{\mathrm{Thr}}(s) \triangleq \max_{0<\tau<1} \underline{\mathbf{E}}_s(\tau) \geq \frac{\log_2 e}{4s^2}(1 + o(1)), \quad s \to \infty.$$

Доказанная ранее теорема 3 фактически означает, что экспонента ошибки при проверке гипотез с помощью традиционного пофакторного декодирования обращается в нуль, как только скорость кодов превышает $1/s$. Поэтому нижняя граница для экспоненты ошибки $\mathbf{E}_s(\tau, R)$ из теоремы 4 устанавливает превосходство порогового декодирования над традиционным пофакторным декодированием в задаче проверки гипотезы $H_0$ против альтернативы $H_1$ для больших значений скорости кода $R$.

В разделе 3.3 приведены результаты моделирования проверки гипотез, которые также показывают, что для большой скорости кодов пороговое декодирование имеет преимущество перед пофакторным декодированием.

Заключительная **глава 4** посвящена кодам для гиперканала множественного доступа.

В разделе 4.1 введены основные обозначения и определения.

Раздел 4.2 посвящен разработке нижних границ для скорости СД $s_L$-гиперкодов. Первая теорема устанавливает нижнюю границу в общем случае $q \geq 2$, которая улучшает ранее известные нижние границы для скорости $R_L^{(q)}(s)$.

**Теорема 5** (Нижняя граница для $R_L^{(q)}(s)$)**.** *Справедливы 4 утверждения.*
**1.** *При любых фиксированных* $q \geq 2$, $s \geq 2$ *и* $L \geq 1$ *справедлива нижняя граница:*

$$R_L^{(q)}(s) \geq \underline{R}_L^{(q)}(s) \triangleq \max_{q' \geq q} \frac{-\log_q P(q', s, L)}{(s + L - 1)k(q, q')}, \quad \text{где}$$

$$P(q,s,L) \triangleq \sum_{m=1}^{\min\{q,s\}} \binom{q}{m} \left(\frac{m}{q}\right)^L \sum_{k=0}^{m} (-1)^k \binom{m}{k} \left(\frac{m-k}{q}\right)^s,$$

$$k(q,q') \triangleq \begin{cases} 1, & \text{при} \quad q = q', \\ \lceil \frac{q'}{q-1} \rceil, & \text{иначе.} \end{cases}$$

**2.** *При любых фиксированных $q \geq 2$ и $L \geq 1$ выполнено асимптотическое равенство*

$$\underline{R}_L^{(q)}(s) = \frac{L(q-1)\log_q e}{s^2(\log_2 e)^2}(1+o(1)), \quad s \to \infty.$$

**3.** *При любых фиксированных $q \geq 2$ и $s \geq 2$ существует предел $\underline{R}_\infty^{(q)}(s) \triangleq \lim_{L\to\infty} \underline{R}_L^{(q)}(s)$, причем при $s \to \infty$ выполнено асимптотическое равенство*

$$\underline{R}_\infty^{(q)}(s) = \frac{(q-1)\log_q e}{es}, \quad s \to \infty.$$

**4.** *При любых фиксированных $s \geq 2$ и $L \geq 1$ существует предел*

$$\lim_{q\to\infty} \underline{R}_L^{(q)}(s) = \frac{L}{s+L-1}.$$

Первая часть доказательства теоремы 5 основана на методе случайного кодирования на ансамбле $q$-ичных кодов длины $N$ и объема $t$, каждая $q$-ичная компонента которых выбирается независимо и равновероятно из множества всех $q$-ичных символов $\{0,1,\ldots,q-1\}$. Рассуждения, аналогичные проведенным при доказательстве теоремы 1, приводят к следующей нижней границе на скорость $q$-ичных СД $s_L$-гиперкодов:

$$R_L^{(q)}(s) \geq \frac{-\log_q P(q,s,L)}{s+L-1}, \tag{9}$$

где $P(q,s,L)$ – это вероятность события $\left\{\{\xi_1,\ldots,\xi_L\} \subset \{\xi_{L+1},\ldots,\xi_{L+s}\}\right\}$ для независимых и равномерно распределенных на множестве $\{0,1,\ldots,q-1\}$ случайный величин $\xi_1,\xi_2,\ldots,\xi_{L+s}$.

Вторая и наиболее значимая часть доказательства теоремы 5 заключается в построении $q$-ичного СД $s_L$-гиперкода из $q'$-ичного СД $s_L$-гиперкода, $q' > q$, и оптимизации нижней границы (9) по всем $q' \geq q$. Связь между скоростями $R_L^{(q)}(s)$ для различных объемов алфавита $q$ описывает предложение 2, доказанное в разделе 4.1.

**Предложение 2.** *Для любых целых чисел $q' > q \geq 2$, $s \geq 2$ и $L \geq 1$ выполнено неравенство*

$$R_L^{(q)}(s) \geq \frac{R_L^{(q')}(s)}{\lceil q'/(q-1) \rceil \log_{q'} q}.$$

Вторая теорема из раздела 4.2 служит уточнением нижних границ из теоремы 5 для случая $q = 2$.

**Теорема 6** (Нижняя граница для $R_L^{(2)}(s)$)**.** *Справедливы* 3 *утверждения.*
**1.** *При любых фиксированных* $s \geq 2$ *и* $L \geq 1$ *выполнено неравенство*

$$R_L^{(2)}(s) \geq \underline{R}_L^*(s) \triangleq \max_{0<Q\leq 1/2} \left( h(Q) + \frac{B_L(s,Q)}{s+L-1} \right),$$
$$B_L(s,Q) \triangleq Q \log_2 \left[ \frac{p(1-z)}{p(1-z)+q(1-z)} \right] + (1-Q) \log_2 \left[ \frac{p(z)}{p(z)+q(z)} \right],$$
$$p(z) \triangleq p_L(s,z) = z^s (z-z^s)^L,$$
$$q(z) \triangleq q_L(s,z) = (z-z^s)(1-z^s-(1-z)^s)^L,$$

*где параметр* $z \in (0,1)$ *в формуле для* $B_L(s,Q)$ *определяется как единственный корень уравнения*

$$Q(p(z)+q(z)) = (1-Q)(p(1-z)+q(1-z)).$$

**2.** *При фиксированном* $L = 1, 2, \ldots$ *и* $s \to \infty$ *справедливо асимптотическое неравенство*

$$\underline{R}_L^*(s) \geq \frac{L}{s^2 \log_2 e}(1+o(1)), \quad s \to \infty.$$

**3.** *При фиксированном* $s = 2, 3, \ldots$ *существует предел*

$$\underline{R}_\infty^*(s) \triangleq \lim_{L\to\infty} \underline{R}_L^*(s) = \log_2 \left[ \frac{(s-1)^{s-1}}{s^s} + 1 \right].$$

*Если* $s \to \infty$*, то*

$$\underline{R}_\infty^*(s) = \frac{\log_2 e}{es}(1+o(1)) = \frac{0,5307\ldots}{s}(1+o(1)).$$

Нижняя граница в теореме 6 получена с помощью метода случайного кодирования на ансамбле $E(N,t,Q)$ равновесных двоичных кодов. Ход доказательства теоремы 6 аналогичен доказательству нижней границы для скорости СД $s_L$-кодов в теореме 1.

В разделе 4.3 установлены соотношения между скоростями $q$-ичных СД $s_L$-гиперкодов и СД $s_L$-кодов. Подстановка верхних границ для $R_L(s)$ из раздела 1.3 позволяет получить верхнюю границу для скорости $R_L^{(q)}(s)$, улучшающую ранее известные верхние границы для больших значений параметра $s$.

**Теорема 7** (Соотношения между скоростями $R_L^{(q)}(s)$ и $R_L(s)$)**.** *Справедливы два утверждения.*

**1.** *Для фиксированных параметров $q \geq 2$, $s \geq 2$ и $L \geq 1$ скорости $R_L^{(q)}(s)$ и $R_L(s)$ удовлетворяют соотношению*

$$R_L^{(q)}(s) \leq \min\left\{ \frac{q}{\log_2 q} R_L(s),\ \frac{q-1}{\log_2 q} R_L(s-1) \right\}.$$

**2.** *При любых фиксированных $q \geq 2$, $L \geq 1$ и $s \to \infty$ скорость $q$-ичных СД $s_L$-гиперкодов удовлетворяет неравенству*

$$R_L^{(q)}(s) \leq \frac{2L(q-1)\log_q s}{s^2}(1 + o(1)), \quad s \to \infty.$$

Принимая во внимание очевидное неравенство $R_L(s) \leq R_L^{(2)}(s)$, получаем довольно важное следствие из теоремы 7, означающее, что при больших значениях параметра $s$ преимущество по скорости двоичных СД $s_L$-гиперкодов над СД $s_L$-кодами исчезает.

**Следствие 1.** *При фиксированном $L \geq 1$ и $s \to \infty$ имеет место асимптотическое равенство*

$$R_L^{(2)}(s) = R_L(s)(1 + o(1)), \quad s \to \infty.$$

В разделе 4.4 приведено детальное сравнение полученных в данной диссертации границ для скорости $R_L^{(q)}(s)$ с ранее известными, а также представлены таблицы численных значений наилучших границ скорости $R_L^{(q)}(s)$ для некоторых наборов параметров $q$, $s$ и $L$.

В разделе 4.5 рассматривается приложение $q$-ичных СД $s_L$-гиперкодов (а также СД $s_L$-кодов), связанное с кодированием недоопределенных данных. Оригинальный метод кодирования, полученный в работе[36], расширен на общий случай $q \geq 2$ и, что наиболее важно, на случай $L \geq 1$. Использование кодов при $L > 1$ позволяет построить более компактный метод кодирования, не увеличивая при этом асимптотические расходы на память и временную сложность алгоритма декодирования.

## Заключение

В настоящей диссертационной работе доказаны новые границы для асимптотической скорости некоторых семейств комбинаторных кодов. В частности, доказана новая нижняя граница асимптотической скорости $R_L(s)$ дизъюнктивных кодов со списочным декодированием, которая улучшает ранее известные нижние границы при $L > 1$. Также получены новые нижние и верхние

[36] *Шоломов Л.А.*, Двоичные представления недоопределенных данных и дизъюнктивные коды, *Прикладная дискретная математика*, № 1, С. 17-33, 2013.

границы асимптотической скорости $R_L^{(q)}(s)$ гиперкодов со списочным декодированием, превосходящие ранее известные границы для широкого набора параметров. Тем не менее, между наилучшими нижними и верхними границами как скорости $R_L(s)$, так и скорости $R_L^{(q)}(s)$, остался существенный разрыв в асимптотике при $s \to \infty$. Асимптотика наилучших нижних границ равна $O(1/s^2)$, а верхних – $O(\ln s/s^2)$. Нахождение главного члена асимптотики скорости $R_L(s)$ (и $R_L^{(q)}(s)$) представляет из себя интересную задачу для дальнейших исследований.

В работе также получена и исследована нижняя граница для экспоненты ошибки почти дизъюнктивных кодов со списочным декодированием, которая обосновывает существование двухступенчатой процедуры групповых проверок в дизъюнктивной модели поиска дефектов с заданным значением экспоненциального убывания ошибки при росте числа элементов. Для почти дизъюнктивных кодов со списочным декодированием доказана верхняя граница на пропускную способность.

Разработан новый алгоритм для проверки, что количество дефектных элементов превышает заданный порог в дизъюнктивной модели группового тестирования. Этот алгоритм, в отличие от ранее известных, обеспечивает меньшую ошибку проверки при проведении неадаптивных групповых тестов для достаточно большого количества элементов. Остается открытой задача построения нижней границы на минимальное количество тестов при заданной ошибке.

**Благодарности**

Автор выражает глубокую благодарность своему научному руководителю д.ф.-м.н., профессору Дьячкову Аркадию Георгиевичу за постановку интересных задач, обсуждение результатов и постоянное внимание к работе. Автор благодарен Воробьеву Илье Викторовичу и Полянскому Никите Андреевичу за ценные замечания и многочисленные детальные обсуждения, а также слушателям и докладчикам семинара по теории кодирования в ИППИ РАН за полезные замечания и проявленную заинтересованность в результатах работы.

# Работы автора по теме диссертации

[1] *Дьячков А.Г.*, *Воробьев И.В.*, *Полянский Н.А.*, *Щукин В.Ю.*, Границы скорости дизъюнктивных кодов, *Пробл. передачи информ.*, Т. 50, № 1, С. 31-63, 2014. [Дьячкову А.Г. принадлежат постановка задачи, теорема 3, предложения 1-3; Воробьеву И.В. – теоремы 1 и 6; Полянскому Н.А. – теоремы 2, 4 и 5; Щукину В.Ю. – теорема 7]

[2] *Дьячков А.Г.*, *Воробьев И.В.*, *Полянский Н.А.*, *Щукин В.Ю*, Почти дизъюнктивные коды со списочным декодированием, *Пробл. передачи информ.*, Т. 51, № 2, С. 27-49, 2015. [Дьячкову А.Г. принадлежат постановка задачи и предложение 1; Воробьеву И.В. – предложение 2, пункты 1 (о пропускной способности) и 2 теоремы 4; Полянскому Н.А. – пример 1, предложение 3 и теорема 1; Щукину В.Ю. – теорема 2, пункты 1 (об экспоненте ошибки) и 3 теоремы 4]

[3] *Щукин В.Ю*, Списочное декодирование для гиперканала множественного доступа, *Пробл. передачи информ.*, Т. 52, № 4, С. 14-30, 2016.

[4] *D'yachkov A.G.*, *Vorobyev I.V.*, *Polyanskii N.A.*, *Shchukin V.Yu.*, Symmetric Disjunctive List-Decoding Codes, *Des. Codes Cryptogr.*, http://dx.doi.org/10.1007/s10623-016-0278-4, pp. 1-19, 2016. [Дьячкову А.Г. принадлежат постановка задачи и предложение 4; Воробьеву И.В. – пункт 1 теоремы 1 и следствие 3; Полянскому Н.А. – предложение 1; Щукину В.Ю. – следствия 4 и 5, пункты 2 и 3 теоремы 1, теоремы 2 и 3]

[5] *D'yachkov A.G.*, *Vorobyev I.V.*, *Polyanskii N.A.*, *Shchukin V.Yu.*, Almost Cover-Free Codes and Designs, *Des. Codes Cryptogr.*, http://dx.doi.org/10.1007/s10623-016-0279-3, pp. 1-17, 2016. [Дьячкову А.Г. принадлежат постановка задачи и предложение 1; Воробьеву И.В. – пример 1 и теорема 1; Полянскому Н.А. – теоремы 2 и 3; Щукину В.Ю. – пример 2, предложения 2 и 3]

[6] *D'yachkov A.G.*, *Vorobyev I.V.*, *Polyanskii N.A.*, *Shchukin V.Yu.*, Cover-Free Codes and Separating System Codes, *Des. Codes Cryptogr.*, http://dx.doi.org/10.1007/s10623-016-0265-9, pp. 1-13, 2016. [Дьячкову А.Г. принадлежит постановка задачи; Воробьеву И.В. – теорема 1, пункты 1, 2 и 4 теоремы 2, теорема 3, леммы 3 и 4, пункты 2 и 3 теоремы 4; Полянскому Н.А. – леммы 1 и 2, пункт 1 теоремы 4; Щукину В.Ю. – пункты 3 и 5 теоремы 2]

[7] *D'yachkov A.G.*, *Vorobyev I.V.*, *Polyanskii N.A.*, *Shchukin V.Yu.*, Bounds on the Rate of Superimposed Codes, *Proc. IEEE Int. Symp. Inf. Theory*, Honolulu, pp. 2341-2345, 2014. [Дьячкову А.Г. принадлежат постановка задачи, теорема 3 и предложения 1-3; Воробьеву И.В. – теоремы 1 и 6; Полянскому Н.А. – теоремы 2, 4 и 5; Щукину В.Ю. – теорема 7]

[8] *D'yachkov A.G.*, *Vorobyev I.V.*, *Polyanskii N.A.*, *Shchukin V.Yu.*, Symmetric Disjunctive List-Decoding Codes, *Proc. IEEE Int. Symp. Inf. Theory*, Hong Kong, pp. 2236-2240, 2015. [Дьячкову А.Г. принадлежат постановка задачи и предложение 3; Воробьеву И.В. – пункт 1 теоремы 2, следствие 1'; Полянскому Н.А. – теорема 1; Щукину В.Ю. – следствия 2' и 3', пункты 2 и 3 теоремы 2, теорема 3]

[9] *D'yachkov A.G.*, *Vorobyev I.V.*, *Polyanskii N.A.*, *Shchukin V.Yu.*, Almost Cover-Free Codes and Designs, *Proc. IEEE Int. Symp. Inf. Theory*, Hong Kong, pp. 2899-2903, 2015. [Дьячкову А.Г. принадлежат постановка задачи и предложение 1; Воробьеву И.В. – пример 1 и теорема 1; Полянскому Н.А. – теоремы 2 и 3; Щукину В.Ю. – пример 2, предложения 2 и 3]

[10] *D'yachkov A.G.*, *Vorobyev I.V.*, *Polyanskii N.A.*, *Shchukin V.Yu.*, Cover-Free Codes and Separating System Codes, *Proc. IEEE Int. Symp. Inf. Theory*, Hong Kong, pp. 2894-2898, 2015. [Дьячкову А.Г. принадлежат постановка задачи и предложения 1-2; Воробьеву И.В. – предложение 3, теоремы 1 и 1', пункты 1-3 теоремы 2, лемма 1; Полянскому Н.А. – лемма 2, пункты 4 и 6 теоремы 2; Щукину В.Ю. – лемма 3, пункт 5 теоремы 2]

[11] *D'yachkov A.G.*, *Vorobyev I.V.*, *Polyanskii N.A.*, *Shchukin V.Yu.*, On a Hypergraph Approach to Multistage Group Testing Problems, *Proc. IEEE Int. Symp. Inf. Theory*, Barcelona, pp. 1183-1187, 2016. [Дьячкову А.Г. принадлежат постановка задачи, разделы 1 и 2; Воробьеву И.В. – разделы 4 и 5; Полянскому Н.А. – раздел 3; Щукину В.Ю. – раздел 6]

[12] *D'yachkov A.G.*, *Vorobyev I.V.*, *Polyanskii N.A.*, *Shchukin V.Yu.*, On Multistage Learning a Hidden Hypergraph, *Proc. IEEE Int. Symp. Inf. Theory*, Barcelona, pp. 1178-1182, 2016. [Дьячкову А.Г. принадлежит постановка задачи; Воробьеву И.В. – теорема 2; Полянскому Н.А. – теорема 3; Щукину В.Ю. – теорема 1]

[13] *D'yachkov A.G.*, *Vorobyev I.V.*, *Polyanskii N.A.*, *Shchukin V.Yu.*, Threshold Decoding for Disjunctive Group Testing, *Proc. 15th Int. Workshop on Algebraic and Combinatorial Coding Theory*, Albena, pp. 151-156, 2016. [Дьячкову А.Г. принадлежит постановка задачи; Воробьеву И.В. – раздел 3; Полянскому Н.А. – таблица 1; Щукину В.Ю. – теорема 3]