

УДК 621.391 : 519.72 : 519.832

© 2021 г. А.С. Вора, А.А. Кулкарни¹**ТЕОРЕМЫ О МИНИМАКСЕ ДЛЯ СОВМЕСТНОГО КОДИРОВАНИЯ ИСТОЧНИКА И КАНАЛА С ПОТЕРЯМИ ПРИ КОНЕЧНОЙ ДЛИНЕ БЛОКА В ПРОИЗВОЛЬНО МЕНЯЮЩЕМСЯ КАНАЛЕ**

Задача о передаче информации при конечной длине блока в присутствии источника целенаправленных помех, мотивированная приложениями, связанными с безопасностью киберфизических систем, рассматривается как игра с нулевой суммой между командой, состоящей из кодера и декодера, и источником помех, где как этой команде, так и источнику помех разрешается использовать лишь локально рандомизированные стратегии. При использовании локально рандомизированных кодов задача для устанавливающей связь команды является невыпуклой, и поэтому в общем случае теорема о минимаксе для такой игры может не выполняться. Тем не менее показана справедливость *приближенных теорем о минимаксе* в том смысле, что минимакс и максимин для этой игры асимптотически сходятся друг к другу. В частности, для скоростей, строго меньших некоторого порогового значения, *обе* эти величины стремятся к нулю, а для скоростей, строго больших этого порога, – к единице. Затем доказана теорема о минимаксе *для вторых приближений*, т.е. показано, что для скоростей, стремящихся в точности к этому пороговому значению по определенному закону, как минимакс, так и максимин сходятся к некоторому постоянному значению, не равному ни нулю, ни единице. Ключевую роль для этих результатов играют полученные границы на минимакс и максимин игры при конечной длине блока и границы второго порядка, основанные на дисперсии.

Ключевые слова: произвольно меняющиеся каналы, игра с нулевой суммой, невыпуклость, стохастические коды.

DOI: 10.31857/S0555292321020017

§ 1. Введение

Киберфизические системы состоят из физических объектов, дистанционно управляемых по каналам связи. Примерами таких систем являются инфраструктура “умного города”, современные автомобили, электросети и атомные электростанции, где для передачи срочной информации используется вспомогательный кибернетический уровень [2]. Эффективная и безопасная передача информации через этот кибернетический уровень является основополагающим фактором для функционирования таких систем. Чтобы гарантировать надежную связь в парадигме блокового кодирования, обычно требуется все большее число обращений к каналу, что приводит к задержкам при передаче. Поскольку вышеуказанные системы чувствительны к задержкам, крайне важно исследовать эти системы в режиме конечной длины блока. С другой стороны, наличие кибернетического уровня в этих системах делает их уязвимыми для атак противника, которые могут иметь катастрофические последствия [3, 4]. С учетом важности этих двух факторов – ограниченности задержки и

¹ Предварительная версия настоящей статьи была представлена в [1].

возможности целенаправленных помех – в данной статье изучается задача прямой передачи с потерями при конечной длине блока в присутствии источника помех.

Мы рассматриваем постановку задачи, в которой передатчик и приемник осуществляют связь по каналу, состояние которого контролируется активным источником помех. Источник помех может изменять состояние канала *произвольным образом*, т.е. в любой момент передачи, с целью нарушить связь. Стандартная формулировка этой задачи рассматривает эту ситуацию только с точки зрения осуществляющей связь команды путем поиска стратегий кодирования, устойчивых к *любому* действию источника помех. Однако поскольку возможность применять стратегии имеют как осуществляющая связь команда, так и источник помех, будет логично принять нейтральную точку зрения, позволяющую и команде, и источнику помех действовать стратегически. С этой целью мы формулируем вышеуказанную проблему как игру с нулевой суммой, в которой команда, состоящая из передатчика и приемника, стремится минимизировать потери при передаче путем построения соответствующих кодов, а источник помех пытается максимизировать их, выбирая состояния канала.

Ключевым для наших результатов является определение допустимых стратегий для осуществляющей связь команды и источника помех, в частности, понятие рандомизации, которую мы допускаем. Мы предполагаем, что стратегией для осуществляющей связь команды, т.е. для пары кодер–декодер, является *стохастический код* или коды только с *локальной* рандомизацией. Как правило, для задачи передачи информации интерес представляет нахождение детерминированных кодов. В теоретико-игровой терминологии это соответствует *чистым* стратегиям для осуществляющей связь команды. Однако по причинам, которые мы обсудим далее, нахождение детерминированного кода для этой задачи требует анализа, выходящего за рамки настоящей статьи. Поэтому мы будем искать рандомизированные стратегии. Теория игр предоставляет два возможных типа рандомизации чистых стратегий (см., например, [5]). *Смешанная* стратегия – это случайный выбор чистой стратегии; в терминологии теории связи это соответствует случайно выбранной паре кодер–декодер, или *случайному коду*. Однако реализация такого кода требует *совместной* случайности между кодером и декодером, что может оказаться невыполнимым в киберфизических системах, где кодер и декодер децентрализованы, а канал является единственным средством связи. Другим типом рандомизации в теории игр являются *поведенческие* стратегии – при такой стратегии действие выбирается случайным образом с учетом информации, имеющейся на данный момент времени. В литературе по теории информации такие стратегии называются *стохастическими кодами*. При такой стратегии кодер и декодер применяют локальную рандомизацию, используя свой собственный личный источник случайности. В настоящей статье мы используем именно этот тип рандомизации для осуществляющей связь команды. Однако при этом для нас достаточно, чтобы локально случайную стратегию использовал только передатчик. Декодирование может по-прежнему быть детерминированной функцией. Таким образом, под стохастическим кодом в этой статье мы понимаем пару, состоящую из стохастического кодера и детерминированного декодера. Мы предполагаем, что источник помех может применять рандомизированный выбор последовательностей состояний канала.

К сожалению, при рассмотрении стохастических кодов в дальнейшем анализе возникают значительные трудности. Во-первых, сформулировав задачу как игру с нулевой суммой, ее приходится анализировать, используя понятие *седловой точки* [5]. Говорят, что игра имеет седловую точку, если минимальные потери по всем стратегиям источника помех, которые осуществляющая связь команда может понести в худшем случае (называемые *верхней ценой* игры), равны максимальным потерям по всем стратегиям осуществляющей связь команды, которые источник помех может понести в худшем случае (называемые *нижней ценой* игры). Однако

для каждого фиксированного действия источника помех задача оптимизации для осуществляющей связь команды при использовании стохастических кодов с необходимостью *невыпукла* из-за неклассической структуры информации [6]. Как следствие, седловая точка не обязательно существует для получающейся игры с нулевой суммой. Этот факт ставит под вопрос любой дальнейший теоретико-игровой анализ.

Кроме того, построение стохастического кода сопряжено с определенными неотъемлемыми трудностями. Для вычисления границы сверху на верхнюю цену игры требуется схема достижимости для совместного кодирования источника и канала. Однако для построения детерминированного кода для совместного кодирования источника и канала может потребоваться установление существования детерминированного кода в произвольно меняющемся канале при критерии максимальной вероятности ошибки. Эта последняя задача в общем случае не решена, а в некоторых случаях эквивалентна нахождению пропускной способности при нулевой ошибке для дискретного канала без памяти [7].

Наши основные результаты показывают, что, несмотря на вышеуказанную невыпуклость, для этой игры справедливы теоремы, близкие к минимаксным, в том смысле, что при возрастании длины блока верхняя и нижняя цена игры становятся сколь угодно близкими. В частности, мы показываем, что существует такой порог, что если асимптотическая скорость строго ниже него, то верхняя и нижняя цена стремятся к нулю, а для скоростей, строго превышающих порог, верхняя и нижняя цена стремятся к единице. Затем мы рассматриваем более тонкий режим, при котором асимптотическая скорость в точности равна порогу, но допускается отклонение от порога, *изменяющееся* в зависимости от длины блока по некоторому специальному закону. В этом случае верхняя и нижняя цена стремятся к одной и той же константе, не равной ни нулю, ни единице. Это показывает, что теорема о минимаксе справедлива даже в этом новом, более тонком режиме.

Верхняя цена игры соответствует совместному кодированию источника и канала с потерями в *произвольно меняющемся канале* (ПМК). С другой стороны, нижняя цена соответствует нахождению распределения вероятностей на последовательностях состояний, такого что минимальные потери в получающемся канале (не обязательно без памяти) максимальны. Наши результаты получены с помощью обращения теоремы кодирования для этой задачи, дающего границу снизу на нижнюю цену игры, и построения схемы достижимости для ПМК, дающей границу сверху на верхнюю цену. Эти границы приводят к границам второго порядка, основанным на дисперсии, для скорости передачи, которые вызвали значительный интерес в недавнем прошлом (см., например, [8, 9] и многочисленные последующие работы). В нашем контексте границы, основанные на дисперсии, дают вышеупомянутую более тонкую теорему о минимаксе (теорему “второго порядка”).

В нашей схеме достижимости используется стохастическое кодирование, доступное для совместного кодирования источника и канала. Отдельную схему кодирования для источника и канала в постановке задачи кодирования источника и канала без потерь можно построить как композицию кода для источника и кода для канала, построенных независимо друг от друга. Хотя эта схема достаточна для теорем о минимаксе с классическими скоростями передачи, она недостаточно точна для получения искомой минимаксной теоремы второго порядка. Мы разработали схему совместного кодирования источника и канала, используя идеи из [9, 10], дающую стохастический код, приводящий к желаемой теореме. Для построения этой схемы достижимости мы сперва получаем редуцированный случайный код, являющийся кодом с равномерным распределением на меньшем количестве детерминированных кодов. Используя детерминированный код для канала при критерии средней вероятности ошибки и вышеупомянутый редуцированный случайный код, мы строим совместный код для источника и канала, требующий только локальной случайности на

передающем конце, получая таким образом стохастический код. Для этой конструкции мы используем детерминированный код для канала, представленный в [11].

Для нижней граница используется метод ослабления, основанного на линейном программировании, (ЛП-ослабления) из [12], где было показано, что хотя задача прямой передачи (без помех) и невыпукла в пространстве стохастических кодов, она тем не менее обладает некоторой скрытой выпуклостью. В частности, было показано, что для больших длин блоков задачу можно сколь угодно точно аппроксимировать методом ЛП-ослабления. Это позволяет предположить, что в нашей задаче может быть справедлива *приближенная* теорема о минимаксе. Наши результаты подтверждают эту догадку. Они показывают, что ЛП-ослабление является точным даже в постановке задачи с преднамеренными помехами, расширяя тем самым список случаев, когда метод ЛП-ослабления дает точные результаты [12, 13].

Насколько нам известно, постановка задачи, рассматриваемая здесь, ранее не изучалась. Конечно, теоретико-игровая формулировка задачи передачи информации при наличии помех изучалась на самом деле и ранее, но в несколько ином смысле. Например, в [14, 15] рассматривалась взаимная информация как функция выигрыша и было доказано существование стратегий с седловой точкой. В теории управления изучались подобные задачи для непрерывного алфавита, например, в [16] рассматривался вопрос о передаче последовательности гауссовских случайных величин при наличии помех и формулировалась задача об игре с нулевой суммой со среднеквадратической ошибкой в качестве функции выигрыша. В контексте ПМК теоремы кодирования для гауссовских ПМК при ограничениях на пиковую и среднюю мощности на входе и для источника помехе были доказаны в [17]. В этих теоремах кодирования неявно предполагалась приближенная теорема о минимаксе для игры с нулевой суммой. Формулировка для игры с нулевой суммой, аналогичная нашей, обсуждалась также в [10, гл. 12], где рассматривались и формулировались различные случаи передачи по ПМК как игры с нулевой суммой. Наиболее близкой к нашей постановке является задача, изучавшаяся в [18, 19] вторым автором настоящей статьи, где рассматривалась только задача кодирования канала, а действие источника помех было фиксировано на протяжении всей передачи, что делало вопрос о верхней цене игры эквивалентным кодированию для *составного канала*; кроме того, в [18, 19] не рассматривались минимаксные теоремы второго порядка.

Совместное кодирование источника и канала без потерь в ПМК было рассмотрено нами в предварительной версии данной статьи, представленной на конференции [1]. Настоящая статья отличается от [1] в следующих аспектах. Здесь мы рассматриваем совместное кодирование источника и канала с потерями, в отличие от [1], где рассматривалось только кодирование без потерь. Кроме того, мы выводим основанные на дисперсии границы на скорость, открывающие дорогу к минимаксной теореме второго порядка. Для этого нам нужна характеристика дисперсии ПМК. Далее, для получения членов высшего порядка, учитывающих дисперсию, нам требуется более сложный совместный стохастический код для совместного кодирования источника и канала. Отдельные коды для источника и для канала, использованные в [1], достаточны для получения минимаксных теорем первого порядка. Однако такое раздельное кодирование не позволяет получить точные члены высшего порядка.

Статья имеет следующую структуру. Формулировка задачи приведена в § 3. Нижняя граница выводится в § 4, а верхняя – в § 5. Соответствующий асимптотический анализ производится в § 6, а § 7 содержит заключительные замечания.

§ 2. Предварительные сведения

2.1. Обозначения. Все случайные величины в этой статье дискретны и заданы на соответствующем вероятностном пространстве с мерой \mathbf{P} . Случайные величины обозначаются прописными буквами X , а их реализации – строчными буквами x ;

если это не оговаривается специально, они являются векторами, длины которых ясны из контекста. Ажурными буквами типа \mathbb{X}, \mathbb{Y} и т.д. обозначаются соответствующие однократные случайные величины, а пространства таких случайных величин обозначаются рукописными буквами \mathcal{X}, \mathcal{Y} и т.д. Множество всех распределений вероятностей на пространстве \mathcal{X} обозначается через $\mathcal{P}(\mathcal{X})$, а одно такое конкретное распределение – через $P_{\mathbb{X}}$.

Типом последовательности $x \in \mathcal{X}^n$ называется эмпирическое распределение $T_x \in \mathcal{P}(\mathcal{X})$ вида $T_x(\cdot) \equiv |\{i : x_i = \cdot\}|/n$. Совместный тип величин x, y обозначается через $T_{x,y}$. Через $\mathcal{P}_n(\mathcal{X}) \subseteq \mathcal{P}(\mathcal{X})$ обозначается множество всех типов последовательностей в \mathcal{X}^n . Множество последовательностей с типом P обозначается через $T(P)$.

Для любого выражения \mathcal{B} индикаторная функция $\mathbb{I}\{\mathcal{B}\}$ равна единице, когда выражение истинно, и нулю в противном случае. Вероятность события A по мере, индуцированной распределением P , обозначается через $P\{A\} := \sum_x \mathbb{I}_{x \in A} P(x)$. Дисперсия случайной величины $X \in \mathcal{X}$ обозначается через $\text{Var}(X) := \mathbf{E}[X - \mathbf{E}[X]]^2$. Для любых распределений P_X и $P_{Y|X}$ полагаем $(P_X \times P_{Y|X})(x, y) := P_X(x)P_{Y|X}(y|x)$ и $(P_X P_{Y|X})(y) := \sum_x P_X(x)P_{Y|X}(y|x)$. Функция, дополнительная к гауссовской функции распределения, обозначается через \mathbb{Q} (не путать с условными функциями распределения $Q_{X|Y}$ для случайных величин X, Y – такие функции будут вводиться и использоваться в соответствующих контекстах). Все экспоненты и логарифмы (\exp и \log) рассматриваются по основанию 2.

2.2. Произвольно меняющийся канал. Произвольно меняющийся канал (ПМК) был впервые введен в [20], где рассматривался канал, закон в котором может произвольным образом изменяться при каждой передаче. ПМК можно моделировать следующим образом. Рассмотрим семейство каналов $\mathbb{V} := \{P_{\mathbb{Y}|\mathbb{X}, \Theta} \in \mathcal{P}(\mathcal{Y}|\mathcal{X}), \Theta \in \mathcal{T}\}$ с одинаковыми пространствами входов и выходов, являющихся конечными множествами \mathcal{X} и \mathcal{Y} соответственно, где каждый канал индексируется параметром θ , называемым состоянием канала и выбираемым из конечного множества \mathcal{T} . Предполагается, что канал не имеет памяти. Тогда получаемый канал можно моделировать как следующий дискретный ПМК без памяти: вероятность получения последовательности $y \in \mathcal{Y}^n$ при передаче последовательности $x \in \mathcal{X}^n$ и при последовательности состояний $\theta \in \mathcal{T}^n$ равна $P_{Y|X, \Theta}(y|x, \theta) = \prod_{i=1}^n P_{\mathbb{Y}|\mathbb{X}, \Theta}(y_i|x_i, \theta_i)$.

Детерминированный код для канала – это пара функций (f_C, φ_C) вида

$$f_C: \mathcal{W} \rightarrow \mathcal{X}^n, \quad \varphi_C: \mathcal{Y}^n \rightarrow \mathcal{W},$$

где $\mathcal{W} = \{1, \dots, M\}$ для $M \in \mathbb{N}$. Вероятность ошибки для каждого сообщения $m \in \mathcal{W}$ и последовательности состояний $\theta \in \mathcal{T}^n$ при использовании детерминированного кода для канала (f_C, φ_C) имеет вид

$$e_{m, \theta}(f_C, \varphi_C) := \sum_{y \in \mathcal{Y}^n} \mathbb{I}\{\varphi(y) \neq m\} P_{Y|X, \Theta}(y|f_C(m), \theta).$$

Максимальная и средняя вероятности ошибки по всем сообщениям для такого кода определяются как

$$e(f_C, \varphi_C) := \max_{\theta \in \mathcal{T}^n} \max_{m \in \{1, \dots, M\}} e_{m, \theta}(f_C, \varphi_C), \quad \bar{e}(f_C, \varphi_C) := \max_{\theta \in \mathcal{T}^n} \frac{1}{M} \sum_{m=1}^M e_{m, \theta}(f_C, \varphi_C)$$

соответственно.

В отличие от обычного дискретного канала без памяти (ДКБП), пропускная способность ПМК зависит от критерия ошибки и от типов кодов, используемых кодером и декодером. Кроме того, в некоторых случаях для пропускной способности

ПМК не известно никакое выражение в замкнутом виде. В частности, в [7] было показано, что вычисление пропускной способности детерминированного кодирования для определенного класса ПМК при критерии максимальной вероятности ошибки сводится к вычислению пропускной способности ДКБП при нулевой ошибке, что является весьма сложной задачей. Однако пропускную способность ПМК при критериях как максимальной, так и средней вероятности ошибки можно вычислить, если кодеру и декодеру разрешено рандомизировать свои действия.

Стохастический код определяется как пара условных распределений $Q_{X|W} \in \mathcal{P}(\mathcal{X}^n | \mathcal{W})$, $Q_{\widehat{W}|Y} \in \mathcal{P}(\mathcal{W} | \mathcal{Y}^n)$. Максимальная и средняя вероятности ошибки для стохастического кода $(Q_{X|W}, Q_{\widehat{W}|Y})$ определяются как

$$\begin{aligned} e(Q_{X|W}, Q_{\widehat{W}|Y}) &:= \\ &:= \max_{\theta \in \mathcal{T}^n} \max_{w \in \{1, \dots, M\}} \sum_{x, y, \widehat{w}} \mathbb{I}\{\widehat{w} \neq w\} Q_{X|W}(x|w) P_{Y|X, \theta}(y|x, \theta) Q_{\widehat{W}|Y}(\widehat{w}|y), \\ \bar{e}(Q_{X|W}, Q_{\widehat{W}|Y}) &:= \\ &:= \max_{\theta \in \mathcal{T}^n} \frac{1}{M} \sum_{w=1}^M \sum_{x, y, \widehat{w}} \mathbb{I}\{\widehat{w} \neq w\} Q_{X|W}(x|w) P_{Y|X, \theta}(y|x, \theta) Q_{\widehat{W}|Y}(\widehat{w}|y). \end{aligned}$$

Пусть e_n – критерий вероятности ошибки (максимальной или средней). Пусть $R_n = \log M/n$ – скорость передачи. Скорость R называется достижимой, если существует последовательность (M, n) -кодов со скоростями R_n (детерминированных или случайных), такая что вероятность ошибки $e_n \rightarrow 0$ и $R_n \rightarrow R$ при n , стремящемся к бесконечности. Точная верхняя грань всех таких скоростей называется пропускной способностью канала и обозначается через C .

В случае ПМК пропускная способность C может оказаться нулевой для определенных типов кодов и критериев ошибки. Приведем условия, при которых пропускная способность ПМК положительна. ПМК называется несимметризуемым, если не существует распределения $P_{\Theta|\mathbb{X}} \in \mathcal{P}(\mathcal{T} | \mathcal{X})$, такого что $\forall x \in \mathcal{X}, x' \in \mathcal{X}, y \in \mathcal{Y}$ выполняется равенство

$$\sum_{\theta \in \mathcal{T}} P_{\Theta|\mathbb{X}}(\theta|x) P_{\mathbb{Y}|\mathbb{X}, \Theta}(y|x', \theta) = \sum_{\theta \in \mathcal{T}} P_{\Theta|\mathbb{X}}(\theta|x') P_{\mathbb{Y}|\mathbb{X}, \Theta}(y|x, \theta).$$

В настоящей статье все рассматриваемые ПМК предполагаются несимметризуемыми. Для несимметризуемого ПМК пропускная способность стохастического кодирования при критерии максимальной (и средней) вероятности ошибки имеет вид [21]

$$C = \max_{P_{\mathbb{X}} \in \mathcal{P}(\mathcal{X})} \min_{q_{\Theta} \in \mathcal{P}(\mathcal{T})} I(\mathbb{X}; \mathbb{Y}_{q_{\Theta}}), \quad (1)$$

где $I(\mathbb{X}; \mathbb{Y}_{q_{\Theta}})$ – взаимная информация между \mathbb{X} и $\mathbb{Y}_{q_{\Theta}}$, а $\mathbb{Y}_{q_{\Theta}}$ – выход усредненного канала $(q_{\Theta} P_{\mathbb{Y}|\mathbb{X}, \Theta}) := \sum_{\theta \in \mathcal{T}} q_{\Theta}(\theta) P_{\mathbb{Y}|\mathbb{X}, \Theta}(\cdot | \cdot, \theta)$ при входе \mathbb{X} . Более подробные сведения о ПМК содержатся в [10, гл. 12; 22].

2.3. Скорость как функция искажения. Пусть \mathcal{S} – пространство сообщений, и пусть $d_{\mathcal{S}}: \mathcal{S} \times \mathcal{S} \rightarrow [0, \infty)$ – функция искажения. Рассмотрим задачу, в которой требуется выразить наборы длины k из \mathcal{S}^k с помощью M сообщений таким образом, чтобы среднее искажение удовлетворяло неравенству $\mathbf{E}[d_{\mathcal{S}}(\mathbb{S}, \widehat{\mathbb{S}})] \leq \mathbf{d}$ для некоторого фиксированного уровня искажения $\mathbf{d} > 0$.

Детерминированный код для источника определяется как пара функций (f_S, φ_S) вида

$$f_S: S^k \rightarrow \mathcal{W}, \quad \varphi_S: \mathcal{W} \rightarrow S^k,$$

где $\mathcal{W} = \{1, \dots, M\}$ для $M \in \mathbb{N}$.

Искажением для последовательностей длины k называется величина $d(S, \widehat{S}) := \sum_{i=1}^k d_S(S_i, \widehat{S}_i)/k$. Скорость определяется как $R_k = \log M/k$. Скорость R называется достижимой для заданного уровня искажения \mathbf{d} , если существует код, такой что $\lim_{k \rightarrow \infty} \mathbf{E}[d(S, \widehat{S})] \leq \mathbf{d}$ и $R_k \rightarrow R$. Точная нижняя грань всех таких скоростей носит название “скорость как функция искажения” и имеет вид

$$R(\mathbf{d}) = \min_{P_{\mathbb{S}|\mathbb{S}}, \mathbf{E}[d_S(\mathbb{S}, \widehat{\mathbb{S}})] \leq \mathbf{d}} I(\mathbb{S}, \widehat{\mathbb{S}}), \quad (2)$$

где $\mathbb{S} \in \mathcal{S}$ имеет распределение $P_{\mathbb{S}} \in \mathcal{P}(\mathcal{S})$, а $\widehat{\mathbb{S}} \in \mathcal{S}$ – распределение $P_{\widehat{\mathbb{S}}|\mathbb{S}} \in \mathcal{P}(\mathcal{S}|\mathcal{S})$. Подробнее о теории скорости как функции искажения см. в [23, гл. 10].

2.4. Информационные величины. В этом пункте приводятся определения некоторых информационных величин, которые будут использоваться в настоящей статье. Информационная плотность определяется как

$$i_{\mathbb{X}; \mathbb{Y}}^{q_{\Theta}}(x; y) = \log \frac{(q_{\Theta} P_{\mathbb{Y}|\mathbb{X}, \Theta})(y|x)}{(P_{\mathbb{X}} q_{\Theta} P_{\mathbb{Y}|\mathbb{X}, \Theta})(y)}, \quad x \in \mathcal{X}, \quad y \in \mathcal{Y}, \quad (3)$$

где

$$(q_{\Theta} P_{\mathbb{Y}|\mathbb{X}, \Theta})(y|x) = \sum_{\theta \in \mathcal{T}} q_{\Theta}(\theta) P_{\mathbb{Y}|\mathbb{X}, \Theta}(y|x, \theta),$$

$$(P_{\mathbb{X}} q_{\Theta} P_{\mathbb{Y}|\mathbb{X}, \Theta})(y) = \sum_{x \in \mathcal{X}, \theta \in \mathcal{T}} P_{\mathbb{X}}(x) q_{\Theta}(\theta) P_{\mathbb{Y}|\mathbb{X}, \Theta}(y|x, \theta).$$

Информационная плотность для векторнозначных случайных величин (X, Y) обозначается через $i_{X; Y_q}$ и определяется аналогично. Определим \mathbf{d} -скошенную информацию как

$$j_{\mathbb{S}}(s, \mathbf{d}) = \log \frac{1}{\mathbf{E}[\exp(\lambda^* \mathbf{d} - \lambda^* d(s, \widehat{\mathbb{S}}))]}, \quad s \in \mathcal{S}, \quad (4)$$

где математическое ожидание вычисляется относительно безусловного распределения $P_{\widehat{\mathbb{S}}^*}$, на котором достигается минимум в (2), и $\lambda^* = -R'(\mathbf{d})$. Далее, \mathbf{d} -скошенная информация для векторнозначных случайных величин (S, \widehat{S}) определяется как

$$R_S(\mathbf{d}) = \inf_{P_{\widehat{S}|S}: \mathbf{E}[d(S, \widehat{S})] \leq \mathbf{d}} I(S; \widehat{S}), \quad (5)$$

где $d(s, \widehat{s})$ – функция искажения, определенная выше. Аналогично, \mathbf{d} -скошенная информация для случайных величин S определяется как

$$j_S(s, \mathbf{d}) := \log \frac{1}{\mathbf{E}[\exp(\lambda^* \mathbf{d} - \lambda^* d(s, \widehat{S}))]},$$

где математическое ожидание берется относительно распределения $P_{\widehat{S}^*}$, на котором достигается инфимум в (5), и $\lambda^* = -R'_S(\mathbf{d})$. Дальнейшие сведения о \mathbf{d} -скошенной информации можно найти в [24].

Определим следующее множество распределений, на которых достигается пропускная способность:

$$\Pi_{\Theta} = \left\{ q_{\Theta} \in \mathcal{P}(\mathcal{T}) : \max_{P_{\mathbb{X}}} I(\mathbb{X}; \mathbb{Y}_{q_{\Theta}}) = C \right\},$$

$$\Pi_{\mathbb{X}} = \left\{ P_{\mathbb{X}} \in \mathcal{P}(\mathcal{X}) : \min_{q_{\Theta}} I(\mathbb{X}; \mathbb{Y}_{q_{\Theta}}) = C \right\}.$$

Дисперсии источника и канала определяются следующим образом:

$$V_S = \text{Var}(j_S(\mathbb{S}, \mathbf{d})), \quad (6)$$

$$V_C^+ = \min_{P_{\mathbb{X}} \in \Pi_{\mathbb{X}}} \max_{q_{\Theta} \in \Pi_{\Theta}} \text{Var}(i_{\mathbb{X}; \mathbb{Y}_{q_{\Theta}}}(\mathbb{X}; \mathbb{Y})), \quad V_C^- = \max_{q_{\Theta} \in \Pi_{\Theta}} \min_{P_{\mathbb{X}} \in \Pi_{\mathbb{X}}} \text{Var}(i_{\mathbb{X}; \mathbb{Y}_{q_{\Theta}}}(\mathbb{X}; \mathbb{Y})), \quad (7)$$

где \mathbb{S} имеет распределение $P_{\mathbb{S}}$, а (\mathbb{X}, \mathbb{Y}) – распределение $P_{\mathbb{X}} \times (q_{\Theta} P_{\mathbb{Y}|\mathbb{X}, \Theta})$. Для вычисления асимптотики будем предполагать, что существует единственное распределение состояний $q_{\Theta}^* \in \Pi_{\Theta}$, на котором достигается пропускная способность. В этом случае определенные выше дисперсии канала равны между собой: $V_C^- = V_C^+ =: V_C$.

§ 3. Постановка задачи

Рассмотрим конечное семейство каналов $\mathbb{V} := \{P_{\mathbb{Y}|\mathbb{X}, \Theta = \theta} \in \mathcal{P}(\mathcal{Y}|\mathcal{X}), \theta \in \mathcal{T}\}$. Пусть \mathcal{S} – конечное пространство. Пусть требуется передать случайное исходное сообщение $S \in \mathcal{S}^k$, $k \in \mathbb{N}$, порождаемое независимо от других сообщений согласно одному и тому же распределению $P_{\mathbb{S}} \in \mathcal{P}(\mathcal{S})$, по этому семейству каналов, где источник помех может выбирать канал из множества \mathbb{V} в каждый момент передачи. Передатчик (кодер) кодирует сообщение S во входную последовательность $X \in \mathcal{X}^n$, $n \in \mathbb{N}$, согласно закону $Q_{X|S} \in \mathcal{P}(\mathcal{X}^n | \mathcal{S}^k)$, а декодер декодирует последовательность $Y \in \mathcal{Y}^n$ на выходе канала в сообщение $\hat{S} \in \mathcal{S}^k$ согласно закону $Q_{\hat{S}|Y} \in \mathcal{P}(\hat{\mathcal{S}}^k | \mathcal{Y}^n)$. Пара $(Q_{X|S}, Q_{\hat{S}|Y})$ называется стохастическим кодом. Говорят, что произошла ошибка, если искажение между декодированной последовательностью и исходной последовательностью источника превышает заранее определенный уровень \mathbf{d} , т.е. если $d(S, \hat{S}) := \sum_{i=1}^k d_S(\mathbb{S}_i, \hat{\mathbb{S}}_i)/k > \mathbf{d}$, где d_S – функция искажения, определенная в п. 2.3, а $\mathbf{d} \in [0, \infty)$ – максимальный допустимый уровень искажения. Источник помех выбирает каналы, которые будут использованы при передаче, путем выбора случайной последовательности состояний $\Theta \in \mathcal{T}^n$ с распределением $q \in \mathcal{P}(\mathcal{T}^n)$; через $\Theta \in \mathcal{T}$ обозначаем соответствующую однократную случайную величину. Мы предполагаем, что кодеру и декодеру не известны действия источника помех и что источник помех также не имеет никакой информации ни о кодере и декодере, ни об исходном сообщении.

Предполагается, что память в канале отсутствует. Таким образом, получающийся канал описывается уравнением $P_{Y|X, \Theta}(y|x, \theta) = \prod_{i=1}^n P_{Y|\mathbb{X}, \Theta}(y_i|x_i, \theta_i)$, задающим вероятность получения последовательности на выходе $y = (y_1, \dots, y_n)$ при последовательности на входе $x = (x_1, \dots, x_n)$ и последовательности состояний $\theta = (\theta_1, \dots, \theta_n)$. Скорость передачи в такой постановке определяется как $R = k/n$.

Вероятность ошибки имеет вид

$$\begin{aligned} \mathbf{P}(d(S, \hat{S}) > \mathbf{d}) &= \\ &= \sum_{s, x, y, \hat{s}, \theta} \mathbb{I}\{d(s, \hat{s}) > \mathbf{d}\} q(\theta) P_S(s) Q_{X|S}(x|s) P_{Y|X, \Theta}(y|x, \theta) Q_{\hat{S}|Y}(\hat{s}|y). \end{aligned} \quad (8)$$

Предполагается, что кодер и декодер стремятся минимизировать вероятности ошибки, выбирая для этого стохастические коды $(Q_{X|S}, Q_{\hat{S}|Y})$, а источник помех – максимизировать ее путем выбора распределения q . Таким образом, для каждой пары (k, n) получается игра с нулевой суммой между командой, состоящей из кодера и декодера, и источником помех, выигрышем в которой служит вероятность ошибки. Необходимые сведения об играх с нулевой суммой можно найти в [5, гл. 4].

Минимаксом, или *верхней ценой* игры, называется величина

$$\bar{\nu}(k, n) = \min_{Q_{X|S}, Q_{\hat{S}|Y}} \max_q \mathbf{P}(d(S, \hat{S}) > \mathbf{d})$$

$$\text{при условиях } Q_{X|S} \in \mathcal{P}(\mathcal{X}^n | \mathcal{S}^k), \quad Q_{\hat{S}|Y} \in \mathcal{P}(\mathcal{S}^k | \mathcal{Y}^n), \quad q \in \mathcal{P}(\mathcal{T}^n),$$

а максимином, или *нижней ценой* игры, – величина

$$\underline{\nu}(k, n) = \max_q \min_{Q_{X|S}, Q_{\hat{S}|Y}} \mathbf{P}(d(S, \hat{S}) > \mathbf{d})$$

$$\text{при условиях } Q_{X|S} \in \mathcal{P}(\mathcal{X}^n | \mathcal{S}^k), \quad Q_{\hat{S}|Y} \in \mathcal{P}(\mathcal{S}^k | \mathcal{Y}^n), \quad q \in \mathcal{P}(\mathcal{T}^n).$$

Очевидно, $\bar{\nu}(k, n) \geq \underline{\nu}(k, n)$.

Можно заметить, что задача о минимаксе – это задача совместного кодирования источника и канала с потерями при передаче по ПМК с помощью стохастических кодов, поскольку кодери и декодери нужно найти стохастические коды, минимизирующие вероятность ошибки в наихудшем случае. Далее, оптимальным выбором для источника помех будет детерминированная последовательность состояний, поскольку вероятность ошибки, задаваемая формулой (8), линейна по распределению q . В задаче совместного кодирования источника и канала при передаче по ДКБП без источника помех асимптотически стремящейся к нулю вероятности ошибки можно достигнуть для скоростей, меньших $C'/R(\mathbf{d})$, где C' – пропускная способность ДКБП, а $R(\mathbf{d})$ – скорость как функция искажения. При этом для скоростей, больших $C'/R(\mathbf{d})$, вероятность ошибки стремится к единице [25]. В настоящей статье мы покажем, что эти результаты распространяются и на нашу игру, где $\bar{\nu}(k, n)$ и $\underline{\nu}(k, n)$ приближаются друг к другу при $k, n \rightarrow \infty$, а значение, к которому они стремятся, зависит от асимптотического значения скорости k/n .

Основными результатами статьи являются следующие две теоремы о минимаксе. Как показывает первая из них, и верхняя, и нижняя цена игры стремятся к нулю, если $\lim_{k, n \rightarrow \infty} k/n < C/R(\mathbf{d})$.

Теорема 1. *Рассмотрим последовательность (k, n) , такую что $\lim_{k, n \rightarrow \infty} k/n < C/R(\mathbf{d})$. Тогда*

$$\lim_{k, n \rightarrow \infty} \underline{\nu}(k, n) = \lim_{k, n \rightarrow \infty} \bar{\nu}(k, n) = 0.$$

Если же $\lim_{k, n \rightarrow \infty} k/n > C/R(\mathbf{d})$, то и верхняя, и нижняя цена игры стремятся к единице, как показывает

Теорема 2. *Рассмотрим последовательность (k, n) , такую что $\lim_{k, n \rightarrow \infty} k/n > C/R(\mathbf{d})$. Тогда*

$$\lim_{k, n \rightarrow \infty} \underline{\nu}(k, n) = \lim_{k, n \rightarrow \infty} \bar{\nu}(k, n) = 1.$$

Далее рассмотрим более тонкий режим. Пусть последовательность (k, n) такова, что $k/n \rightarrow C/R(\mathbf{d})$ по некоторому специально выбранному закону. В частности,

пусть последовательность параметризована величиной $\rho \in \mathbb{R}$ следующим образом:

$$\frac{k}{n} = \frac{C}{R(\mathbf{d})} + \frac{\rho}{\sqrt{n}}. \quad (9)$$

Для такой последовательности имеет место следующий результат.

Теорема 3. Пусть $V_C^+ = V_C^- = V_C$. Тогда для последовательности (k, n) , удовлетворяющей (9), справедливы равенства

$$\lim_{k, n \rightarrow \infty} \underline{\nu}(k, n) = \lim_{k, n \rightarrow \infty} \bar{\nu}(k, n) = Q\left(\frac{-\rho R(\mathbf{d})}{\sqrt{V_C + \frac{C}{R(\mathbf{d})} V_S(\mathbf{d})}}\right).$$

Этот результат показывает, что в отличие от теорем 1 и 2 верхняя и нижняя цена игры стремятся к промежуточному значению между нулем и единицей. Это неэкстремальное значение достигается при выборе последовательности с пределом $C/R(\mathbf{d})$, являющимся пороговым значением для надежной передачи в постановке совместного кодирования источника и канала. Такое понимание более тонкой асимптотики возможно исключительно благодаря основанным на дисперсии границам высших порядков, которые выводятся в последующих параграфах.

§ 4. Нижняя граница на максимин

Приступая к выводу вышеуказанных результатов, установим границы на $\bar{\nu}(k, n)$ и $\underline{\nu}(k, n)$ при конечной длине блока. В этом параграфе мы выведем нижнюю границу на $\underline{\nu}(k, n)$ путем ослабления внутренней минимизации по $(Q_{X|S}, Q_{\hat{S}|Y})$ в задаче о максимине. Для каждого $q \in \mathcal{P}(\mathcal{T}^n)$ задачу минимизации можно представить в виде

$$\text{SC}(q): \quad \min_{Q_{X|S}, Q_{\hat{S}|Y}} \sum_{s, x, y, \hat{s}} \mathbb{I}\{d(s, \hat{s}) > \mathbf{d}\} Q(s, x, y, \hat{s})$$

$$\text{при условиях} \quad Q(s, x, y, \hat{s}) \equiv P_S(s) Q_{X|S}(x|s) P_{Y_q|X}(y|x) Q_{\hat{S}|Y}(\hat{s}|y),$$

$$\sum_x Q_{X|S}(x|s) = 1 \quad \forall s,$$

$$\sum_{\hat{s}} Q_{\hat{S}|Y}(\hat{s}|y) = 1 \quad \forall y,$$

$$Q_{X|S}(x|s), Q_{\hat{S}|Y}(\hat{s}|y) \geq 0 \quad \forall s, x, y, \hat{s},$$

$$\text{где } P_{Y_q|X}(y|x) := \sum_{\theta \in \mathcal{T}^n} q(\theta) P_{Y|X, \Theta}(y|x, \theta).$$

Эта задача невыпукла в пространстве распределений $(Q_{X|S}, Q_{\hat{S}|Y})$ [6]. Возможным подходом к таким задачам является получение выпуклого ослабления путем включения невыпуклой области допустимых решений в некоторое выпуклое множество. Рассмотрим ослабление, основанное на линейном программировании (ЛП-ослабление), которое было представлено в [12] и получено методом типа подъема и проекции.

В этом методе целевая функция и ограничения линеаризуются путем замены произведений $Q_{X|S}(x|s) Q_{\hat{S}|Y}(\hat{s}|y)$ на вспомогательную переменную $V(s, x, y, \hat{s})$. Затем добавляются ограничения, обеспечивающие тождественное равенство $V(s, x, y, \hat{s}) \equiv Q_{X|S}(x|s) Q_{\hat{S}|Y}(\hat{s}|y)$. Наконец, для получения ослабления задачи $\text{SC}(q)$ условие $V(s, x, y, \hat{s}) \equiv Q_{X|S}(x|s) Q_{\hat{S}|Y}(\hat{s}|y)$ опускается. Фактически невыпуклая область допустимых решений в задаче $\text{SC}(q)$ аппроксимируется многогранником, тем самым

задавая *поднятие* задачи в пространство большей размерности. Результатом этого являются следующая задача линейного программирования:

$$\begin{aligned}
\text{LP}(q): \quad & \min_{Q_{X|S}, Q_{\hat{S}|Y}, V} \sum_{s, x, y, \hat{s}} \mathbb{I}\{d(s, \hat{s}) > \mathbf{d}\} \bar{Q}(s, x, y, \hat{s}) \\
& \text{при условиях} \quad \bar{Q}(s, x, y, \hat{s}) \equiv P_S(s)P_{Y_q|X}(y|x)V(s, x, y, \hat{s}), \\
& \sum_x Q_{X|S}(x|s) = 1 : \gamma_q^S(s) \quad \forall s, \\
& \sum_{\hat{s}} Q_{\hat{S}|Y}(\hat{s}|y) = 1 : \gamma_q^C(y) \quad \forall y, \\
& \sum_x V(s, x, y, \hat{s}) - Q_{\hat{S}|Y}(\hat{s}|y) = 0 : \lambda_q^S(s, \hat{s}, y) \quad \forall s, \hat{s}, y, \\
& \sum_{\hat{s}} V(s, x, y, \hat{s}) - Q_{X|S}(x|s) = 0 : \lambda_q^C(x, s, y) \quad \forall x, s, y, \\
& Q_{X|S}(x|s), Q_{\hat{S}|Y}(\hat{s}|y) \geq 0 \quad \forall s, x, y, \hat{s}, \\
& V(s, x, y, \hat{s}) \geq 0 \quad \forall s, x, y, \hat{s},
\end{aligned}$$

где функции $\gamma_q^S: \mathcal{S}^k \rightarrow \mathbb{R}$, $\gamma_q^C: \mathcal{Y}^n \rightarrow \mathbb{R}$, $\lambda_q^S: \mathcal{S}^k \times \mathcal{S}^k \times \mathcal{Y}^n \rightarrow \mathbb{R}$ и $\lambda_q^C: \mathcal{S}^k \times \mathcal{X}^n \times \mathcal{Y}^n \rightarrow \mathbb{R}$ – множители Лагранжа.

Любая допустимая точка для задачи LP(q) задается тройкой $(Q_{X|S}, Q_{\hat{S}|Y}, V)$, а соответствующая допустимая точка задачи SC(q) получается проекцией этой тройки на пространство пар $(Q_{X|S}, Q_{\hat{S}|Y})$. Последовательно повторяя эту процедуру, получаем все более точные выпуклые реализации исходной задачи. Дальнейшие подробности о методе поднятия и проекции можно найти в [26, гл. 5].

Очевидно, справедливо неравенство

$$\text{OPT}(\text{SC}(q)) \geq \text{OPT}(\text{LP}(q)), \quad \forall q \in \mathcal{P}(\mathcal{T}^n).$$

Теперь сформулируем двойственную задачу для LP(q) и, используя слабую двойственность, оценим оптимальное значение для LP(q), получая тем самым границу на OPT(SC(q)). Соответствующая двойственная задача выглядит следующим образом:

$$\begin{aligned}
\text{DP}(q): \quad & \max_{\gamma_q^S, \gamma_q^C, \lambda_q^S, \lambda_q^C} \sum_s \gamma_q^S(s) + \sum_y \gamma_q^C(y) \\
& \text{при условиях} \quad \gamma_q^S(s) - \sum_y \lambda_q^C(x, s, y) \leq 0 \quad \forall x, s, \\
& \gamma_q^C(y) - \sum_s \lambda_q^S(s, \hat{s}, y) \leq 0 \quad \forall \hat{s}, y, \\
& \lambda_q^S(s, \hat{s}, y) + \lambda_q^C(x, s, y) \leq \Pi(s, x, y, \hat{s}) \quad \forall s, x, y, \hat{s},
\end{aligned}$$

где $\Pi(s, x, y, \hat{s}) \equiv \mathbb{I}\{d(s, \hat{s}) > \mathbf{d}\} P_S(s)P_{Y_q|X}(y|x)$. Из слабой двойственности следует, что $\text{OPT}(\text{SC}(q)) \geq \text{OPT}(\text{LP}(q)) = \text{OPT}(\text{DP}(q)) \forall q$, и согласно результатам [18] получаем

$$\begin{aligned}
\underline{\mu}(k, n) &= \max_q \text{OPT}(\text{SC}(q)) \geq \max_q \text{OPT}(\text{LP}(q)) = \\
&= \max_q \text{OPT}(\text{DP}(q)) \geq \text{FEAS}(\text{DP}(q)),
\end{aligned} \tag{10}$$

где $\text{FEAS}(\text{DP}(q))$ – значение целевой функции для $\text{DP}(q)$, вычисленное в допустимой точке. Таким образом, для вывода нижней границы на максимак и максимин нашей игры с нулевой суммой достаточно получить допустимое решение задачи $\text{DP}(q)$. В следующей теореме приводится одна такая конструкция и вычисляется целевая функция двойственной задачи $\text{DP}(q)$ для этого допустимого решения, давая тем самым нижнюю границу на максимин.

Теорема 4. *Значение $\underline{\nu}(k, n)$ ограничено снизу следующим образом:*

$$\begin{aligned} \underline{\nu}(k, n) &\geq \max_q \text{OPT}(\text{DP}(q)) \geq \\ &\geq \max_{q, P_{\bar{Y}_q}, U} \sup_{\gamma > 0} \left[\sum_s P_S(s) \min_x \left[\mathbf{P} \left(j_S(s, \mathbf{d}) - i_{X; \bar{Y}_q | U}(x; Y | U) \leq \gamma \right) + \right. \right. \\ &+ \exp(j_S(s, \mathbf{d}) - \gamma) \sum_{u=1}^U \sum_y P_{U|X}(u|x) P_{\bar{Y}_q | U}(y|u) \times \\ &\left. \left. \times \mathbb{I} \left\{ j_S(s, \mathbf{d}) - i_{X; \bar{Y}_q | U}(x; y|u) > \gamma \right\} \right] - \frac{U}{\exp(\gamma)} \right], \end{aligned} \quad (11)$$

где $U \in \mathcal{U} := \{1, \dots, U\}$, $U \in \mathbb{N}$, $P_{\bar{Y}_q} \in \mathcal{P}(\mathcal{Y}^n)$, $i_{X; \bar{Y}_q | U}(x; y|u) = \frac{\log P_{Y_q | X, U}(y|x, u)}{P_{\bar{Y}_q | U}(y|u)}$, а $j_S(s, \mathbf{d})$ – \mathbf{d} -скошенная информация, определенная в п. 2.4.

Доказательство. Определим случайную величину U , принимающую значения в множестве $\mathcal{U} := \{1, \dots, U\}$, такую что

$$P_{Y_q | X}(y|x) = \sum_{u=1}^U P_{U|X}(u|x) P_{Y_q | X, U}(y|x, u).$$

Рассмотрим следующие двойственные переменные для $\text{DP}(q)$:

$$\begin{aligned} \lambda_q^S(s, \hat{s}, y) &\equiv -\mathbb{I}\{d(s, \hat{s}) \leq \mathbf{d}\} \frac{P_S(s) \sum_{u=1}^U P_{\bar{Y}_q | U}(y|u)}{\exp(\gamma - j_S(s, \mathbf{d}))}, \\ \lambda_q^C(x, s, y) &\equiv P_S(s) \sum_{u=1}^U P_{U|X}(u|x) \min \left\{ P_{Y_q | X, U}(y|x, u), \frac{P_{\bar{Y}_q | U}(y|u)}{\exp(\gamma - j_S(s, \mathbf{d}))} \right\}, \\ \gamma_q^S(s) &\equiv \min_x \sum_y \lambda_q^C(x, s, y), \quad \gamma_q^C(y) \equiv -\exp(-\gamma) \sum_{u=1}^U P_{\bar{Y}_q | U}(y|u), \end{aligned}$$

где $\gamma > 0$, $P_{\bar{Y}_q | U}(y|u) := \sum_{\theta} q(\theta) P_{\bar{Y} | \Theta, U}(y|\theta, u)$, а $P_{\bar{Y} | \Theta, U}$ – произвольное распределение из $\mathcal{P}(\mathcal{Y}^n | \mathcal{T}^n, \mathcal{U})$.

Из доказательства теоремы 5.3 работы [12] следует, что такой выбор двойственных переменных является допустимым для задачи $\text{DP}(q)$. При этом целевая функция двойственной задачи имеет вид

$$\begin{aligned} &\sum_s \min_x \sum_y \lambda_q^C(x, s, y) + \sum_y -\exp(-\gamma) \sum_{u=1}^U P_{\bar{Y}_q | U}(y|u) \geq \\ &\geq \sum_s \min_x \sum_y P_S(s) \sum_{u=1}^U P_{U|X}(u|x) \min \left\{ P_{Y_q | X, U}(y|x, u), \frac{P_{\bar{Y}_q | U}(y|u)}{\exp(\gamma - j_S(s, \mathbf{d}))} \right\} - \end{aligned}$$

$$\begin{aligned}
& - \frac{\sum_y \sum_{u=1}^U P_{\bar{Y}_q|U}(y|u)}{\exp(\gamma)} = \sum_s P_S(s) \min_x \left[\sum_{u=1}^U \sum_y P_{U|X}(u|x) P_{Y_q|X,U}(y|x,u) \times \right. \\
& \times \mathbb{I} \left\{ \frac{P_{Y_q|X,U}(y|x,u)}{P_{\bar{Y}_q|U}(y|u)} \leq \frac{\exp(j_S(s, \mathbf{d}))}{\exp(\gamma)} \right\} + \frac{1}{\exp(\gamma - j_S(s, \mathbf{d}))} \sum_{u=1}^U \sum_y P_{U|X}(u|x) \times \\
& \times P_{\bar{Y}_q|U}(y|u) \mathbb{I} \left\{ \frac{P_{Y_q|X,U}(y|x,u)}{P_{\bar{Y}_q|U}(y|u)} > \frac{\exp(j_S(s, \mathbf{d}))}{\exp(\gamma)} \right\} \left. \right] - \frac{U}{\exp(\gamma)}.
\end{aligned}$$

Полагая $i_{X; \bar{Y}_q|U}(x; y|u) = \log P_{Y_q|X,U}(y|x,u) / P_{\bar{Y}_q|U}(y|u)$, получаем

$$\begin{aligned}
& \sum_s \min_x \sum_y \lambda_q^C(x, s, y) + \sum_y - \exp(-\gamma) \sum_{u=1}^U P_{\bar{Y}_q|U}(y|u) = \\
& = \sum_s P_S(s) \min_x \left[\mathbf{P} \left(i_{X; \bar{Y}_q|U}(x; Y|U) - j_S(s, \mathbf{d}) \leq -\gamma \right) + \right. \\
& + \exp(j_S(s, \mathbf{d}) - \gamma) \sum_{u=1}^U \sum_y P_{U|X}(u|x) P_{\bar{Y}_q|U}(y|u) \times \\
& \times \mathbb{I} \left\{ i_{X; \bar{Y}_q|U}(x; y|u) - j_S(s, \mathbf{d}) > -\gamma \right\} \left. \right] - \frac{U}{\exp(\gamma)}.
\end{aligned}$$

Переходя к супремуму по γ и максимуму по U , $P_{\bar{Y}_q}$ и q в этом равенстве, получаем выражение в правой части формулы (11). Требуемая граница вытекает из соотношений (10). \blacktriangle

ЛП-ослабление дает границу снизу на верхнюю и нижнюю цену игры. Таким образом, эта граница определяет также обращение теоремы кодирования для задачи совместного кодирования источника и канала с потерями в произвольном меняющемся канале. Эта нижняя граница справедлива для всех стратегий источника помех и не требует предположения о том, что стратегия является н.о.р.

Известно, что такая техника ослабления, основанного на линейном программировании, предложенная в [12], дает точные и улучшенные обращения теоремы кодирования для множества задач кодирования, как указано в [12, 13, 19]. Для больших длин блока этот метод приводит к достаточно близкому выпуклому ослаблению задачи кодирования, подчеркивая тем самым ее почти выпуклую природу. Получение границы снизу на вероятность ошибки основывается только на построении подходящих переменных в двойственной задаче, и поэтому метод ослабления представляет собой структурный подход к выводу границ.

Теперь перейдем к выводу верхней границы с помощью построения кода, достигающего пропускной способности, в задаче совместного кодирования источника и канала.

§ 5. Верхняя граница на минимакс

Для вывода верхней границы на величину минимакса $\bar{V}(k, n)$ построим стохастический код для задачи совместного кодирования источника и канала. Напомним, что стохастический код для совместного кодирования источника и канала – это пара распределений $(Q_{X|S}, Q_{\hat{S}|Y})$, где $Q_{X|S} \in \mathcal{P}(\mathcal{X}^n | \mathcal{S}^k)$ и $Q_{\hat{S}|Y} \in \mathcal{P}(\mathcal{S}^k | \mathcal{Y}^n)$.

В настоящей статье мы рассматриваем стохастический код со стохастическим кодированием и детерминированным декодированием. Для построения такого стоха-

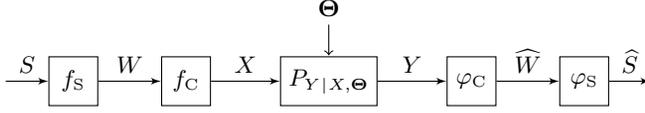


Рис. 1. Совместное кодирование источника и канала

стического кода будем рассматривать случайный код для совместного кодирования источника и канала, а также детерминированный код для канала. Применяя идеи редукции случайных кодов, построим другой случайный код с равномерным распределением на меньшем количестве кодов. Кодер случайным образом выбирает код из этого ансамбля для передачи сообщения, а затем использует детерминированный код для передачи декодеру номера этого кода из ансамбля. Декодер восстанавливает этот номер с помощью детерминированного кода, а затем использует код с этим номером для декодирования исходного сообщения.

Чтобы построить стохастический код для совместного кодирования источника и канала, вначале рассмотрим детерминированный код для канала, дающий гарантированную среднюю вероятность ошибки.

5.1. Детерминированный код для канала для заданной средней вероятности ошибки. Вначале рассмотрим детерминированный код для канала, не зависящий от кода, выбираемого для передачи сообщений. Пусть (f_c, φ_c) – код для канала, определенный в п. 2.2.

Известен следующий результат [11, теорема 1], обеспечивающий существование детерминированного кода для заданной средней вероятности ошибки.

Теорема 5. Пусть $P_X \in \mathcal{P}(\mathcal{X}^n)$, пусть $Z(x, \bar{x}, y) \in \{0, 1\}$ – функция, такая что

$$Z(x, \bar{x}, y)Z(\bar{x}, x, y) = 0 \quad \forall x \in \mathcal{X}^n, \bar{x} \in \mathcal{X}^n, y \in \mathcal{Y}^n, \quad (12)$$

и пусть $\mathcal{A} \subset \mathcal{X}^n \times \mathcal{Y}^n$ – типичное множество. Пусть $(X, \bar{X}, Y) \sim P_X \times P_X \times P_{Y|X, \Theta}$. Тогда существует детерминированный код (f_c, φ_c) для канала, такой что

$$\begin{aligned} \bar{e}(f_c, \varphi_c) \leq & \sqrt{\frac{2 \ln(3|\mathcal{T}|^n)}{M}} + \min_{P_X} \max_{\theta \in \mathcal{T}^n} \left(\mathbf{P}((X, Y) \notin \mathcal{A} | \theta) + \right. \\ & + 2 \log 3 |\mathcal{T}|^n \max_{\bar{x} \in \mathcal{X}^n} \mathbf{P}(Z(X, \bar{x}, Y) = 0, (X, Y) \in \mathcal{A} | \theta) + \\ & \left. + 2M \log e \mathbf{P}(Z(X, \bar{X}, Y) = 0, (X, Y) \in \mathcal{A} | \theta) \right). \end{aligned} \quad (13)$$

Таким образом, имеется детерминированный код для канала, дающий верхнюю границу на среднюю вероятность ошибки. Теперь построим случайный код для совместного кодирования источника и канала.

5.2. Случайный код для совместного кодирования источника и канала. Схема совместного кодирования источника и канала представлена на рис. 1. Пусть $W \in \mathcal{W} = \{1, \dots, M\}$ – случайная величина на выходе кодера, где $M \in \mathbb{N}$, а $\widehat{W} \in \mathcal{W}$ – случайная величина на входе декодера. Код для источника – это пара функций (f_S, φ_S) , определенная в п. 2.3. Код для канала – это пара функций (f_C, φ_C) , определенная в п. 2.2. Композиция этих двух кодов задает код (f, φ) для совместного кодирования источника и канала вида

$$f := (f_C \circ f_S): \mathcal{S}^k \rightarrow \mathcal{X}^n, \quad \varphi := (\varphi_S \circ \varphi_C): \mathcal{Y}^n \rightarrow \mathcal{S}^k.$$

Случайный код для совместного кодирования источника и канала – это пара случайных величин (F, Φ) , принимающих значения в множестве $\{(f, \varphi) \mid f: \mathcal{S}^k \rightarrow \mathcal{X}^n, \varphi: \mathcal{Y}^n \rightarrow \mathcal{S}^k\}$. Вероятность ошибки для детерминированного кода для совместного кодирования источника и канала имеет вид $e_{d,\theta}(f, \varphi) := \sum_{s,y} \mathbb{I}\{d(s, \varphi(y)) > \mathbf{d}\} P_{\mathcal{S}}(s) \times P_{Y|X,\Theta}(y|f(s), \theta)$. Соответственно, ошибка для случайного кода $(F, \Phi) \sim \psi$ для совместного кодирования источника и канала определяется следующим образом:

$$e_{d,\theta}(\psi) := \mathbf{E}[e_{d,\theta}(F, \Phi)], \quad e_{\mathbf{d}}(\psi) := \max_{\theta \in \mathcal{T}^n} \mathbf{E}[e_{d,\theta}(\psi)].$$

Теперь построим случайный код для совместного кодирования источника и канала подобно тому, как в [9] был построен детерминированный код для задачи совместного кодирования источника и канала для ДКБП без преднамеренных помех.

Теорема 6. *Существует случайный код ψ , такой что для его вероятности ошибки справедлива оценка сверху*

$$e_{\mathbf{d}}(\psi) \leq \min_{P_{L|S}} \max_{\theta \in \mathcal{T}^n} \left(\mathbf{E} \left[\exp \left(-|i_{X;Y_{q^*}}(X;Y) - \log L|^+ \right) \mid \theta \right] + \mathbf{E} \left[(1 - P_{\widehat{\mathcal{S}}}(\mathcal{B}_{\mathbf{d}}(S)))^L \right] \right),$$

где $(S, L, \widehat{\mathcal{S}}, X, Y) \sim P_{\mathcal{S}} \times P_{L|S} \times P_{\widehat{\mathcal{S}}} \times P_X \times P_{Y|X, \Theta=\theta}$,

$$i_{X;Y_{q^*}}(x; y) = \log \frac{(q^* P_{Y|X, \Theta})(y|x)}{(P_X q^* P_{Y|X, \Theta})(y)},$$

причем $(X, Y) \sim P_X \times P_{Y|X, \Theta=\theta}$, $q^*(\theta) = \prod_{i=1}^n q_{\Theta}^*(\theta_i)$, $q_{\Theta}^* \in \Pi_{\Theta}$ и $\mathcal{B}_{\mathbf{d}}(s) := \{\widehat{s} \in \mathcal{S}^k : d(s, \widehat{s}) \leq \mathbf{d}\}$.

Доказательство. Случайный код ψ будем строить следующим образом. Вначале построим случайный код $(F_{\mathcal{S}}, \Phi_{\mathcal{S}})$ для источника сообщений. Затем для конкретной реализации кода $(f_{\mathcal{S}}, \varphi_{\mathcal{S}})$ для источника построим случайный код $(F_{\mathcal{C}}, \Phi_{\mathcal{C}})$ для канала.

Кодирование для источника: Порождаются M кодовых слов из множества \mathcal{S}^k согласно распределению $P_{\widehat{\mathcal{S}}} \in \mathcal{P}(\mathcal{S}^k)$ каждое. Обозначим i -е кодовое слово через \widehat{S}_i . Далее для заданного исходного сообщения источника $s \in \mathcal{S}^k$ порождается случайная величина $L \sim P_{L|S=s} \in \mathcal{P}(\mathbb{N}|\mathcal{S}^k)$, где $L \leq M$. Кодер кодирует сообщение s как

$$F_{\mathcal{S}}(s) = \begin{cases} \min\{m, L\}, & d(s, \widehat{S}_m) \leq \mathbf{d} < \min_{i < m} d(s, \widehat{S}_i), \\ L, & \mathbf{d} < \min_{i=1, \dots, L} d(s, \widehat{S}_i), \end{cases}$$

где d – функция искажения, а \mathbf{d} – максимальный уровень искажения, определенные в п. 2.3.

Декодирование для источника: Декодер сообщений декодирует выход m декодера канала как $\Phi_{\mathcal{S}}(m) = \widehat{S}_m$.

Теперь для заданной реализации кода $(F_{\mathcal{S}}, \Phi_{\mathcal{S}}) = (f_{\mathcal{S}}, \varphi_{\mathcal{S}})$ для источника построим случайный код для канала.

Кодирование для канала: Определим M случайных величин $\{X_m\}_{m=1}^M$ с распределением $P_X \in \mathcal{P}(\mathcal{X}^n)$ каждая. Для заданного выхода m кодера источника кодер канала кодирует его как $F_{\mathcal{C}}(m) = X_m$.

Декодирование для канала: Для функции декодирования канала определим случайную величину $U \in \{1, \dots, M+1\}$ следующим образом:

$$U = \begin{cases} f_{\mathcal{S}}(S), & d(S, \varphi_{\mathcal{S}} \circ f_{\mathcal{S}}(S)) \leq \mathbf{d}, \\ M+1 & \text{в противном случае.} \end{cases}$$

Положим $P_{Y_{q^*}|X}(y|x) := \sum_{\theta \in \mathcal{T}^n} q^*(\theta) P_{Y|X, \Theta}(y|x, \theta)$, где $q^*(\theta) = \prod_{i=1}^n q_{\Theta}^*(\theta_i)$, $q_{\Theta}^* \in \Pi_{\Theta}$.

Для наблюдаемого выхода канала y декодер канала декодирует его следующим образом:

$$\Phi_C(y) = m \in \arg \max_{j \in \{1, \dots, M\}} P_{U|\hat{S}^M}(j|\hat{S}^M) P_{Y_{q^*}|X}(y|X_j),$$

где $P_{U|\hat{S}^M} \in \mathcal{P}(\{1, \dots, M+1\} | (\mathcal{S}^k)^M)$.

Пусть ψ – распределение, индуцированное распределением $P_{\hat{S}}$ кодовой книги источника и распределением P_X кодовой книги канала, на множестве кодов для совместного кодирования источника и канала $\{(f, \varphi) \mid f: \mathcal{S}^k \rightarrow \mathcal{X}^n, \varphi: \mathcal{Y}^n \rightarrow \mathcal{S}^k\}$. Теперь вычислим вероятность ошибки для этого случайного кода ψ при фиксированном $\theta \in \mathcal{T}^n$. Пусть $\hat{S}^M = (\hat{s}_1, \dots, \hat{s}_M) \in \mathcal{S}^k$ и $x^M = (x_1, \dots, x_M) \in \mathcal{X}^n$ – кодовые книги источника и канала. Вероятность ошибки оценивается сверху как

$$\begin{aligned} \mathbf{P}(d(S, \hat{S}) > \mathbf{d} | \hat{S}^M, x^M, \theta) &\leq \mathbf{P}(d(S, \varphi_S \circ f_S(S)) > \mathbf{d} | \hat{S}^M) + \\ &+ \mathbf{P}(\varphi_C(Y) \neq f_S(S) | \hat{S}^M, x^M, \theta, d(S, \varphi_S \circ f_S(S)) \leq \mathbf{d}), \end{aligned} \quad (14)$$

где первое слагаемое – это ошибка кодирования источника, а второе – ошибка декодирования в канале при отсутствии ошибки кодирования источника. Ошибку декодирования можно представить в виде $\mathbf{P}(d(S, \varphi_S \circ f_S(S)) > \mathbf{d} | \hat{S}^M) = \mathbf{P}(U > L | \hat{S}^M)$, так как $\{d(S, \varphi_S \circ f_S(S)) > \mathbf{d}\} = \{U > L\}$. Ошибку декодирования в канале можно вычислить следующим образом:

$$\begin{aligned} \mathbf{P}(\varphi_C(Y) \neq f_S(S) | \hat{S}^M, x^M, \theta, d(S, \varphi_S \circ f_S(S)) \leq \mathbf{d}) &= \\ = \mathbf{P}(\varphi_C(Y) \neq U | \hat{S}^M, x^M, \theta) &= \\ = \sum_{m=1}^M P_{U|\hat{S}^M}(m|\hat{S}^M) \mathbf{P}(\varphi_C(Y) \neq m | x^M, \theta) &= \\ = \sum_{m=1}^M P_{U|\hat{S}^M}(m|\hat{S}^M) \mathbf{P}\left(\bigcup_{m'=1}^M \frac{P_{U|\hat{S}^M}(m'|\hat{S}^M) P_{Y_{q^*}|X}(Y|X_{m'})}{P_{U|\hat{S}^M}(m|\hat{S}^M) P_{Y_{q^*}|X}(Y|X_m)} \geq 1 | x^M, \theta\right), \end{aligned} \quad (15)$$

где (15) следует из того факта, что $d(S, \varphi_S \circ f_S(S)) \leq \mathbf{d}$ влечет $f_S(S) = U$.

Усредняя по кодовым книгам (\hat{S}^M, x^M) , получаем

$$\sum_{\hat{S}^M, x^M} P_{\hat{S}^M}(\hat{S}^M) P_{X^M}(x^M) \mathbf{P}(d(S, \hat{S}) > \mathbf{d} | \hat{S}^M, x^M, \theta) = \mathbf{P}(d(S, \hat{S}) > \mathbf{d} | \theta) = e_{\mathbf{d}, \theta}(\psi),$$

где $P_{\hat{S}^M}(\hat{S}^M) = \prod_{i=1}^M P_{\hat{S}}(\hat{s}_i)$, $\hat{s}_i \in \mathcal{S}^k$, и $P_{X^M}(x^M) = \prod_{i=1}^M P_X(x_i)$, $x_i \in \mathcal{X}^n$. Тогда получаем $\mathbf{P}(d(S, \hat{S}) > \mathbf{d} | \theta) = e_{\mathbf{d}, \theta}(\psi)$, поскольку рассматриваемый случайный код индуцирован распределениями P_{X^M} и $P_{\hat{S}^M}$.

Подставляя это в (14), получаем

$$\begin{aligned} e_{\mathbf{d}, \theta}(\psi) &\leq \mathbf{P}(U > L) + \sum_{\hat{S}^M} P_{\hat{S}^M}(\hat{S}^M) \sum_{m=1}^M P_{U|\hat{S}^M}(m|\hat{S}^M) \times \\ &\times \mathbf{P}\left(\bigcup_{m'=1}^M \frac{P_{U|\hat{S}^M}(m'|\hat{S}^M) P_{Y_{q^*}|X}(Y|X_{m'})}{P_{U|\hat{S}^M}(m|\hat{S}^M) P_{Y_{q^*}|X}(Y|X_m)} \geq 1 | \theta\right). \end{aligned} \quad (16)$$

Повторяя рассуждения из доказательства [9, теорема 7], получаем

$$\begin{aligned} & \sum_{\widehat{s}^M} P_{\widehat{S}^M}(\widehat{s}^M) \sum_{m=1}^M P_{U|\widehat{S}^M}(m|\widehat{s}^M) \times \\ & \times \mathbf{P} \left(\bigcup_{m'=1}^M \frac{P_{U|\widehat{S}^M}(m'|\widehat{s}^M) P_{Y_{q^*}|X}(Y|X_{m'})}{P_{U|\widehat{S}^M}(m|\widehat{s}^M) P_{Y_{q^*}|X}(Y|X_m)} \geq 1 \mid \theta \right) \leq \\ & \leq \mathbf{E} \left[\exp(-|i_{X;Y_{q^*}}(X;Y) - \log L|^+) \mid \theta \right] \end{aligned}$$

и $\mathbf{P}(U > L) = \mathbf{E}[(1 - \mathbf{P}(\mathcal{B}_d(S)))^L]$. Подставляя это в (16), получаем

$$e_{d,\theta}(\psi) \leq \mathbf{E} \left[\exp(-|i_{X;Y_{q^*}}(X;Y) - \log L|^+ \mid \theta) \right] + \mathbf{E} \left[(1 - P_{\widehat{S}}(\mathcal{B}_d(S)))^L \right].$$

Переходя к минимуму по распределениям $P_{L|S}$ и максимуму по состояниям θ , получаем требуемую оценку. \blacktriangle

Теперь сформулируем более слабую оценку, полагая $L = \lfloor \gamma / P_{\widehat{S}}(\mathcal{B}_d(S)) \rfloor$, где $\gamma > 0$ выбирается произвольным образом согласно [9, теорема 8].

Теорема 7. *Существует случайный код ψ для совместного кодирования источника и канала, такой что*

$$e_d(\psi) \leq \inf_{\gamma > 0} \max_{\theta \in \mathcal{T}^n} \left(\mathbf{E} \left[\exp \left(- \left| i_{X;Y_{q^*}}(X;Y) - \frac{\gamma}{P_{\widehat{S}}(\mathcal{B}_d(S))} \right|^+ \right) \mid \theta \right] + e^{1-\gamma} \right),$$

где величины (S, \widehat{S}, X, Y) описаны в теореме 6.

Теперь, используя детерминированный код для канала и случайный код для совместного кодирования источника и канала, построим стохастический код для совместного кодирования источника и канала.

5.3. Стохастический код для совместного кодирования источника и канала и верхняя граница. Изложим подход к построению стохастического кода для совместного кодирования источника и канала. Для этого рассмотрим следующий вариант леммы о редукции случайного кода, приведенной в [10, гл. 4] (доказательство приведено в Приложении).

Лемма 1. *Пусть задано $\varepsilon > 0$. Рассмотрим случайный код $(F, \Phi) \sim \psi$, и пусть $e_d(\psi) < \varepsilon$. Пусть $K \in \mathbb{N}$ и ε' таковы, что*

$$\varepsilon' > \varepsilon + \sqrt{\frac{\log |\mathcal{T}^n|}{2K}}. \quad (17)$$

Тогда существуют K детерминированных кодов $(f_i, \varphi_i)_{i=1}^K$ для совместного кодирования источника и канала, таких что

$$\frac{1}{K} \sum_{i=1}^K e_{d,\theta}(f_i, \varphi_i) < \varepsilon' \quad \forall \theta \in \mathcal{T}^n. \quad (18)$$

Предпочтительно иметь ε' как можно более близким к ε . Для этого достаточно взять $K = n^2$ и $\varepsilon' = \varepsilon + \sqrt{(\log |\mathcal{T}^n| + \varepsilon_0) / 2K}$, где $\varepsilon_0 > 0$ сколь угодно мало. Таким образом можно добиться $\varepsilon' \sim \varepsilon$ при достаточно больших n .

Теперь построим код $(Q_{X|S}, \varphi)$ для совместного кодирования источника и канала, где $Q_{X|S}$ – стохастический кодер, а φ – детерминированный декодер. Конструкция та же, что в [10, теорема 12.13]. Рассматриваем K детерминированных

кодов $(f_i, \varphi_i)_{i=1}^K$, где $f_i: \mathcal{S}^k \rightarrow \mathcal{X}^n$ и $\varphi_i: \mathcal{Y}^n \rightarrow \mathcal{S}^k$ удовлетворяют условию (18). Случайно выбранный код из этого ансамбля используется для передачи сообщений из множества \mathcal{S}^k . Далее рассматриваем детерминированный код $(\hat{f}, \hat{\varphi})$, у которого $\hat{f}: \{1, \dots, K\} \rightarrow \mathcal{X}^{d_n}$ и $\hat{\varphi}: \mathcal{Y}^{d_n} \rightarrow \{1, \dots, K\}$ удовлетворяют условию (13), где d_n является функцией от n . Код $(\hat{f}, \hat{\varphi})$ используется для передачи номера i выбранного кода. Рассмотрим случайную величину $i \in \{1, \dots, K\}$, распределенную равномерно и не зависящую от остальных случайных величин. Для заданного $s \in \mathcal{S}^k$ кодер выбирает входную последовательность $X \in \mathcal{X}^{d_n+n}$ случайным образом как $X = (\hat{f}(i), f_i(s))$. Декодер $\varphi: \mathcal{Y}^{d_n+n} \rightarrow \mathcal{S}^k$ декодирует $y = (\hat{y}, \bar{y}) \in \mathcal{Y}^{d_n+n}$ как

$$\varphi(y) = \begin{cases} s, & \text{если } (\hat{\varphi}(\hat{y}), \varphi_i(\bar{y})) = (i, s) \text{ для некоторого } i, \\ 0 & \text{в противном случае,} \end{cases}$$

где $\hat{y} \in \mathcal{Y}^{d_n}$ и $\bar{y} \in \mathcal{Y}^n$.

Используя этот стохастический код для совместного кодирования источника и канала, оценим теперь нижнюю цену рассматриваемой игры с нулевой суммой. Отметим, что поскольку номер выбранного кода кодируется в виде последовательности длины d_n , то скорость передачи равна $k/(d_n + n)$.

Теорема 8. *Значение минимакса игры $\bar{v}(k, n)$ ограничено сверху следующим образом:*

$$\begin{aligned} \bar{v}(k, n) &= \min_{Q_{X|S}, Q_{\hat{S}|Y}} \max_q \mathbf{P}(d(S, \hat{S}) > \mathbf{d}) \leq \\ &\leq \sqrt{\frac{2 \ln(3|\mathcal{T}|^{d_n})}{K}} + \min_{P_{X_a}} \max_{\theta_a} \left[\mathbf{P}((X_a, Y_a) \notin \mathcal{A} | \theta_a) + \right. \\ &+ 2K \log e \mathbf{P}(Z(X_a, \bar{X}_a, Y_a) = 0, (X_a, Y_a) \in \mathcal{A} | \theta_a) + \\ &+ \left. \max_{\bar{x}_a \in \mathcal{X}^{d_n}} 2 \log 3 |\mathcal{T}|^{d_n} \mathbf{P}(Z(X_a, \bar{x}_a, Y_a) = 0, (X_a, Y_a) \in \mathcal{A} | \theta_a) \right] + \\ &+ \inf_{\gamma > 0} \max_{\theta_b} \left[\mathbf{E} \left[\exp \left(- \left| i_{X; Y_a^*}(X_b; Y_b) - \frac{\gamma}{P_{\hat{S}}(\mathcal{B}_d(S))} \right| \right)^+ \middle| \theta_b \right] + e^{1-\gamma} \right] + \\ &+ \sqrt{\frac{\log |\mathcal{T}^n| + \varepsilon_0}{2K}}, \end{aligned} \tag{19}$$

где $\mathcal{X}^{d_n} \ni \bar{X}_a \sim P_{X_a}$, $\mathcal{X}^{d_n} \times \mathcal{Y}^{d_n} \ni (X_a, Y_a) \sim P_{X_a} P_{Y|X, \Theta = \theta_a}$ для $\theta_a \in \mathcal{T}^{d_n}$, $\mathcal{X}^n \times \mathcal{Y}^n \ni (X_b, Y_b) \sim P_{X_b} P_{Y|X, \Theta = \theta_b}$ для $\theta_b \in \mathcal{T}^n$, а $P_{X_a} \in \mathcal{P}(\mathcal{X}^{d_n})$ и $P_{X_b} \in \mathcal{P}(\mathcal{X}^n)$ – произвольные распределения. Величина $\varepsilon_0 > 0$ постоянна.

Доказательство. Максимальную вероятность ошибки оценим следующим образом:

$$\begin{aligned} \max_{\theta \in \mathcal{T}^{n+d_n}} \mathbf{P}(d(S, \hat{S}) > \mathbf{d} | \Theta = \theta) &= \max_{\theta \in \mathcal{T}^{n+d_n}} \frac{1}{K} \sum_{i=1}^K \mathbf{P}(d(S, \hat{S}) > \mathbf{d} | i = i, \Theta = \theta) = \\ &= \max_{\theta \in \mathcal{T}^{n+d_n}} \left(\frac{1}{K} \sum_{i=1}^K \mathbf{P}(d(S, \hat{S}) > \mathbf{d}, \hat{\varphi}(Y_a) \neq i | i = i, \Theta = \theta) + \right. \\ &+ \left. \frac{1}{K} \sum_{i=1}^K \mathbf{P}(d(S, \hat{S}) > \mathbf{d}, \hat{\varphi}(Y_a) = i | i = i, \Theta = \theta) \right) \leq \\ &\leq \max_{\theta \in \mathcal{T}^{n+d_n}} \left(\frac{1}{K} \sum_{i=1}^K \mathbf{P}(\hat{\varphi}(Y_a) \neq i | i = i, \Theta = \theta) \right)' + \end{aligned}$$

$$\begin{aligned}
& + \frac{1}{K} \sum_{i=1}^K \mathbf{P}(d(S, \varphi_i(Y_b)) > \mathbf{d} \mid \mathbf{i} = i, \Theta = \theta) \Big) \leq \\
& \leq \max_{\theta_a \in \mathcal{T}^{d_n}} \frac{1}{K} \sum_{i, y_a} \mathbb{I}\{\widehat{\varphi}(y_a) \neq i\} P_{Y|X, \Theta}(y_a \mid \widehat{f}(i), \theta_a) + \\
& + \max_{\theta_b \in \mathcal{T}^n} \frac{1}{K} \sum_{i, s, y_b} \mathbb{I}\{d(s, \varphi_i(y_b)) > \mathbf{d}\} P_S(s) P_{Y|X, \Theta}(y_b \mid f_i(s), \theta_b) = \\
& = \max_{\theta_a \in \mathcal{T}^{d_n}} \frac{1}{K} \sum_{i, y_a} \mathbb{I}\{\widehat{\varphi}(y_a) \neq i\} P_{Y|X, \Theta}(y_a \mid \widehat{f}(i), \theta_a) + \max_{\theta_b \in \mathcal{T}^n} \frac{1}{K} \sum_{i=1}^K e_{\mathbf{d}, \theta_b}(f_i, \varphi_i).
\end{aligned}$$

Вначале оценим второе слагаемое. Возьмем $\varepsilon' = e_{\mathbf{d}}(\psi) + \sqrt{(\log |\mathcal{T}^n| + \varepsilon_0)/2K}$, где $\varepsilon_0 > 0$. Тогда по лемме 1 второе слагаемое можно оценить как

$$\begin{aligned}
& \frac{1}{K} \sum_i e_{\mathbf{d}, \theta}(f_i, \varphi_i) \leq e_{\mathbf{d}}(\psi) + \sqrt{\log |\mathcal{T}^n|/2K} + \varepsilon_0 \leq \\
& \leq \inf_{\gamma > 0} \max_{\theta_b} \left[\mathbf{E} \left[\exp \left(- \left| i_{X; Y_{q^*}}(X_b; Y_b) - \frac{\gamma}{P_{\widehat{S}}(\mathcal{B}_{\mathbf{d}}(S))} \right|^+ \right) \mid \theta_b \right] + e^{1-\gamma} \right] + \\
& + \sqrt{\frac{\log |\mathcal{T}^n| + \varepsilon_0}{2K}}. \tag{20}
\end{aligned}$$

Первый член оценим с помощью соотношения (13) из теоремы 5, и с учетом (20) получим

$$\begin{aligned}
& \max_{\theta \in \mathcal{T}^{n+d_n}} \mathbf{P}(d(S, \widehat{S}) > \mathbf{d} \mid \Theta = \theta) \leq \sqrt{\frac{2 \ln(3|\mathcal{T}^{d_n}|)}{K}} + \\
& + \min_{P_{X_a}} \max_{\theta_a} \left[\mathbf{P}((X_a, Y_a) \notin \mathcal{A} \mid \theta_a) + \right. \\
& + \max_{\bar{x}_a \in \mathcal{X}^{d_n}} 2 \log 3 |\mathcal{T}^{d_n}| \mathbf{P}(Z(X_a, \bar{x}_a, Y_a) = 0, (X_a, Y_a) \in \mathcal{A} \mid \theta_a) + \\
& + 2K \log e \mathbf{P}(Z(X_a, \bar{X}_a, Y_a) = 0, (X_a, Y_a) \in \mathcal{A} \mid \theta_a) \Big] + \\
& + \inf_{\gamma > 0} \max_{\theta_b} \left[\mathbf{E} \left[\exp \left(- \left| i_{X; Y_{q^*}}(X_b; Y_b) - \frac{\gamma}{P_{\widehat{S}}(\mathcal{B}_{\mathbf{d}}(S))} \right|^+ \right) \mid \theta_b \right] + e^{1-\gamma} \right] + \\
& + \sqrt{\frac{\log |\mathcal{T}^n| + \varepsilon_0}{2K}}. \tag{21}
\end{aligned}$$

Используя равенство $\max_{\theta \in \mathcal{T}^{n+d_n}} \mathbf{P}(d(S, \widehat{S}) > \mathbf{d} \mid \Theta = \theta) = \max_{q \in \mathcal{P}(\mathcal{T}^{d_n+n})} \mathbf{P}(d(S, \widehat{S}) > \mathbf{d})$, получаем требуемый результат. \blacktriangle

Верхнюю границу в рассматриваемой задаче мы получили путем построения стохастического кода для совместного кодирования источника и канала. В отличие от случайных кодов, для которых требуется совместная случайность у кодера и декодера, для этого стохастического кода нужна лишь локальная случайность у кодера. Такой тип кодов допустим в нашей задаче, поскольку мы не предполагаем наличия никакой среды для передачи между передатчиком и приемником, кроме самого канала.

Получив границы на нижнюю и верхнюю цену игры, перейдем к оценке скорости их сходимости для рассматриваемой игры с нулевой суммой.

§ 6. Асимптотика и теоремы о минимаксе

Теперь применим выведенные в предыдущих параграфах границы для вычисления пределов верхней и нижней цены игры. На протяжении этого параграфа будем предполагать, что существует единственное распределение состояний $q_{\Theta}^* \in \Pi_{\Theta}$, на котором достигается пропускная способность, благодаря чему $V_C^- = V_C^+ =: V_C$ и $V_C > 0$.

Пусть $(\mathbb{X}, \mathbb{Y}) \sim P_{\mathbb{X}} \times \sum_{\theta \in \mathcal{T}} q_{\Theta}(\theta) P_{\mathbb{Y}|\mathbb{X}, \Theta = \theta}$, где $q_{\Theta} \in \mathcal{P}(\mathcal{T})$. Далее, пусть для всех таких распределений $P_{\mathbb{X}} \in \mathcal{P}(\mathcal{X})$ выполнено неравенство

$$i_{\mathbb{X}; \mathbb{Y}, q_{\Theta}^*}(x; \mathbb{Y}) < \infty, \quad (22)$$

где

$$i_{\mathbb{X}; \mathbb{Y}, q_{\Theta}^*}(x; y) = \log \frac{(q_{\Theta}^* P_{\mathbb{Y}|\mathbb{X}, \Theta})(y | x)}{(P_{\mathbb{X}} q_{\Theta}^* P_{\mathbb{Y}|\mathbb{X}, \Theta})(y)}, \quad x \in \mathcal{X}, \quad y \in \mathcal{Y},$$

для $q_{\Theta}^* \in \Pi_{\Theta}$. Кроме того, предположим также, что

$$j_{\mathcal{S}}(s, \mathbf{d}) < \infty, \quad \forall s \in \mathcal{S}, \quad (23)$$

где $j_{\mathcal{S}}(s, \mathbf{d})$ определено в п. 2.4.

6.1. Асимптотика второго порядка для верхней границы. Здесь мы вычислим предел нижней границы с помощью теоремы 8. Вначале оценим члены в (19), отвечающие за среднюю вероятность ошибки.

Средняя ошибка. Пусть $P_{\mathbb{X}}^* \in \Pi_{\mathbb{X}}$, и положим $P_{\mathbb{X}}^*(x) := \prod_{i=1}^{d_n} P_{\mathbb{X}}^*(x_i)$ для $x \in \mathcal{X}^{d_n}$. Чтобы применить теорему 8, определим множество $\mathcal{A} \subseteq \mathcal{X}^{d_n} \times \mathcal{Y}^{d_n}$ следующим образом. Пусть $T_{\theta}^n(\theta') := \prod_{i=1}^{d_n} T_{\theta}(\theta'_i)$ для $T_{\theta} \in \mathcal{P}_n(\mathcal{T})$. Положим

$$i_{\mathbb{X}^*; \mathbb{Y}, T_{\theta}^n}(x; y) = \log \frac{(T_{\theta}^n P_{\mathbb{Y}|\mathbb{X}, \Theta})(y | x)}{(P_{\mathbb{X}}^* T_{\theta}^n P_{\mathbb{Y}|\mathbb{X}, \Theta})(y)}, \quad x \in \mathcal{X}^n, \quad y \in \mathcal{Y}^n.$$

Теперь зададим множество \mathcal{A} следующим образом:

$$\mathcal{A} = \left\{ (x, y) \in \mathcal{X}^{d_n} \times \mathcal{Y}^{d_n} : i_{\mathbb{X}^*; \mathbb{Y}, T_{\theta}^n}(x; y) > \gamma \text{ для некоторого } T_{\theta} \in \mathcal{P}_n(\mathcal{T}) \right\}. \quad (24)$$

Далее, пусть

$$V_0 := \sup_{q_{\Theta} \in \mathcal{P}(\mathcal{T})} \text{Var}(i_{\mathbb{X}^*; \mathbb{Y}, q_{\Theta}}(\mathbb{X}; \mathbb{Y})), \quad (25)$$

где $(\mathbb{X}, \mathbb{Y}) \sim P_{\mathbb{X}}^* \times \sum_{\theta \in \mathcal{T}} q_{\Theta}(\theta) P_{\mathbb{Y}|\mathbb{X}, \Theta = \theta}$.

Следующая лемма, которая вытекает из [11, лемма 6], дает необходимое условие несимметризуемости ПМК.

Лемма 2. Пусть $\mathbb{X}, \mathbb{X}' \sim P_{\mathbb{X}}$, и пусть $P_{\mathbb{X}}(x) > 0$ для всех $x \in \mathcal{X}$. Рассмотрим множество

$$D_{\eta} = \left\{ Q_{\mathbb{X}\mathbb{X}'\Theta\mathbb{Y}} \in \mathcal{P}(\mathcal{X} \times \mathcal{X} \times \mathcal{T} \times \mathcal{Y}) : D(Q_{\mathbb{X}\mathbb{X}'\Theta\mathbb{Y}} \| P_{\mathbb{X}} \times Q_{\mathbb{X}'\Theta} \times P_{\mathbb{Y}|\mathbb{X}, \Theta}) \leq \eta \right\},$$

где $(P_{\mathbb{X}} \times Q_{\mathbb{X}'\Theta} \times P_{\mathbb{Y}|\mathbb{X}, \Theta})(x, x', \theta, y) = P_{\mathbb{X}}(x) Q_{\mathbb{X}'\Theta}(x', \theta) P_{\mathbb{Y}|\mathbb{X}, \Theta}(y | x, \theta)$, $Q_{\mathbb{X}'\Theta}$ – маргинальное распределение для $Q_{\mathbb{X}\mathbb{X}'\Theta\mathbb{Y}}$, а $D(Q_{\mathbb{X}\mathbb{X}'\Theta\mathbb{Y}} \| P_{\mathbb{X}} \times Q_{\mathbb{X}'\Theta} \times P_{\mathbb{Y}|\mathbb{X}, \Theta})$ – относитель-

ная энтропия между $Q_{\mathcal{X}\mathcal{X}'\Theta\Upsilon}$ и $P_{\mathcal{X}} \times Q_{\mathcal{X}'\Theta} \times P_{\mathcal{Y}|\mathcal{X},\Theta}$. Положим

$$\eta^* = \inf \left\{ \eta : Q_{\mathcal{X}\mathcal{X}'\Theta\Upsilon} \in D_\eta, Q_{\mathcal{X}'\Theta\Upsilon} \in D_\eta \right. \\ \left. \text{для некоторого } Q_{\mathcal{X}\mathcal{X}'\Theta\Upsilon} \in \mathcal{P}(\mathcal{X} \times \mathcal{X}' \times \mathcal{T} \times \mathcal{T} \times \mathcal{Y}) \right\}.$$

Если ПМК несимметризуем, то $\eta^* > 0$.

Теперь с помощью этой леммы построим функцию $Z: \mathcal{X}^{d_n} \times \mathcal{X}'^{d_n} \times \mathcal{Y}^{d_n} \rightarrow \{0, 1\}$ следующим образом. Так как ПМК несимметризуем, то по лемме 2 имеем $\eta^* > 0$. Выберем η таким, что $0 < \eta < \eta^*$. Положим

$$Z(x_a, \bar{x}_a, y_a) = \begin{cases} 1, & \text{если } (x_a, y_a) \in \mathcal{A} \text{ и либо } (\bar{x}_a, y_a) \notin \mathcal{A}, \\ & \text{либо } \exists \theta_a \in \mathcal{T}^{d_n}, \text{ такое что } T_{x_a, \bar{x}_a, \theta_a, y_a} \in D_\eta, \\ 0 & \text{в противном случае,} \end{cases} \quad (26)$$

где, напомним, $T_{x_a, \bar{x}_a, \theta_a, y_a}$ – совместный тип для $x_a, \bar{x}_a, \theta_a, y_a$. Таким образом, функция Z удовлетворяет условию (12) и при этом $Z(x_a, \bar{x}_a, y_a)Z(\bar{x}_a, x_a, y_a) = 0 \forall x_a, \bar{x}_a \in \mathcal{X}^{d_n}, y_a \in \mathcal{Y}^{d_n}$, поскольку в противном случае найдутся $(x_a, \bar{x}_a, y_a) \in \mathcal{X}^{d_n} \times \mathcal{X}'^{d_n} \times \mathcal{Y}^{d_n}$, такие что $Z(x_a, \bar{x}_a, y_a)Z(\bar{x}_a, x_a, y_a) = 1$. Отсюда $(x_a, y_a) \in \mathcal{A}$ и $(\bar{x}_a, y_a) \in \mathcal{A}$, и значит, существуют совместные типы $T_{x_a, \bar{x}_a, \theta_a, y_a} \in D_\eta$ и $T_{\bar{x}_a, x_a, \theta'_a, y_a} \in D_\eta$ для некоторых $\theta_a, \theta'_a \in \mathcal{T}^{d_n}$. Таким образом, по определению η^* получаем $\eta^* \leq \eta$. Но это дает противоречие, поскольку η выбрано строго меньшим η^* . С помощью этой функции Z получаем следующую оценку сверху (доказательство приведено в Приложении).

Теорема 9. При вышеуказанном выборе \mathcal{A} , Z , K и d_n получаем

$$\begin{aligned} & \sqrt{\frac{2 \ln(3|\mathcal{T}|^{d_n})}{K}} + \min_{P_{X_a}} \max_{\theta_a \in \mathcal{T}^{d_n}} \left[\mathbf{P}((X_a, Y_a) \notin \mathcal{A} | \theta_a) + \right. \\ & + 2K \log e \mathbf{P}(Z(X_a, \bar{X}_a, Y_a) = 0, (X_a, Y_a) \in \mathcal{A} | \theta_a) + \\ & \left. + \max_{\bar{x}_a \in \mathcal{X}'^{d_n}} 2 \log 3 |\mathcal{T}|^{d_n} \mathbf{P}(Z(X_a, \bar{x}_a, Y_a) = 0, (X_a, Y_a) \in \mathcal{A} | \theta_a) \right] \leq \\ & \leq \sqrt{\frac{2 \ln(3|\mathcal{T}|^{d_n})}{K}} + \frac{V_0}{d_n \left(\delta - \frac{\log \sqrt{d_n}}{d_n} \right)^2} + \frac{2 \log e}{\sqrt{d_n}} + \\ & + \frac{2 \log 3 |\mathcal{T}|^{d_n} (d_n + 1)^{|\mathcal{X}'|^2 |\Theta| |\mathcal{Y}|}}{\exp(d_n(\eta))}. \end{aligned} \quad (27)$$

Ошибка для случайного кода. Оценим отвечающие за вероятность ошибки члены в (19), соответствующие случайному коду для совместного кодирования источника и канала (доказательство дано в Приложении).

Теорема 10. Вероятность ошибки для случайного кода в (19) оценивается как

$$\begin{aligned} & \inf_{\gamma > 0} \max_{\theta_b \in \mathcal{T}^n} \left[\mathbf{E} \left[\exp \left(- \left| i_{X; Y_{q^*}}(X_b; Y_b) - \log \frac{\gamma}{P_{\mathcal{S}}(B_a(S))} \right|^+ \right) | \theta_b \right] + e^{1-\gamma} \right] \leq \\ & \leq Q \left(\frac{nC - kR(\mathbf{d}) - \Gamma(k)}{\sqrt{nV_C + kV_S(\mathbf{d})}} \right) + \frac{B}{n+k} + \frac{K_0 + 2}{\sqrt{k}}, \end{aligned} \quad (28)$$

где $B, K_0, \varepsilon_0 > 0$ – константы, а $\Gamma(k) = \bar{c} \log k + c + \log \left(\frac{1}{2} \log k + 1 \right)$, $\bar{c}, c > 0$.

Объединяя результаты двух последних теорем, получаем следующий результат.

Теорема 11. Для верхней цены игры справедлива оценка

$$\begin{aligned} \bar{v}(k, n) \leq & Q\left(\frac{nC - kR(\mathbf{d}) - \Gamma(k)}{\sqrt{nV_C + kV_S(\mathbf{d})}}\right) + \frac{B}{\sqrt{n+k}} + \frac{K_0 + 2}{\sqrt{k}} + \sqrt{\frac{\log |\mathcal{T}^n| + \varepsilon_0}{2K}} + \\ & + \sqrt{\frac{2 \ln(3|\mathcal{T}|^{d_n})}{K}} + \frac{2 \log e}{\sqrt{d_n}} + \frac{V_0}{d_n \left(\delta - \frac{\log \sqrt{d_n}}{d_n}\right)^2} + \\ & + \frac{2 \log 3 |\mathcal{T}|^{d_n} (d_n + 1)^{|\mathcal{X}|^2 |\Theta| |\mathcal{Y}|}}{\exp(d_n(\eta))}, \end{aligned} \quad (29)$$

где $\Gamma(k)$, B , K_0 и ε_0 описаны в теореме 10.

Доказательство. Этот результат получается подстановкой в (19) оценки (27) из теоремы 9 и оценки (28) из теоремы 10. \blacktriangle

Теперь рассмотрим последовательность пар (k, n) , для которой верхняя и нижняя цена игры стремятся к нулю. Возьмем

$$K = c_0 n^2, \quad d_n = \left\lceil \frac{\log K}{C - \delta} \right\rceil, \quad (30)$$

где $c_0 \in \mathbb{N}$, а $\lceil \cdot \rceil$ – функция округления до ближайшего целого в большую сторону.

Границу второго порядка на скорость передачи описывает

Следствие 1 (граница достижимости второго порядка). Пусть $\varepsilon > 0$. Рассмотрим последовательность пар (k, n) , для которой

$$nC - kR(\mathbf{d}) \geq \sqrt{nV_C + kV_S(\mathbf{d})} Q^{-1}(\varepsilon) + O\left(\sqrt{\frac{n}{\log n}}\right).$$

Тогда $\underline{v}(k, n) \leq \bar{v}(k, n) \leq \varepsilon$.

Доказательство. Возьмем последовательность пар (k, n) , для которой выполнено $nC - kR(\mathbf{d}) \geq \sqrt{nV_C + kV_S(\mathbf{d})} Q^{-1}(\varepsilon - \Delta(k, n)) + \Gamma(k)$, где

$$\begin{aligned} \Delta(k, n) = & \frac{B}{\sqrt{n+k}} + \sqrt{\frac{2 \ln(3|\mathcal{T}|^{d_n})}{K}} + \frac{2 \log e}{\sqrt{d_n}} + \frac{V_0}{d_n \left(\delta - \frac{\log \sqrt{d_n}}{d_n}\right)^2} + \\ & + \frac{2 \log 3 |\mathcal{T}|^{d_n} (d_n + 1)^{|\mathcal{X}|^2 |\Theta| |\mathcal{Y}|}}{\exp(d_n(\eta))} + \sqrt{\frac{\log |\mathcal{T}^n| + \varepsilon_0}{2K}}, \end{aligned}$$

а K, d_n удовлетворяют (30). Подставляя это в (29), получаем

$$\begin{aligned} Q\left(\frac{nC - kR(\mathbf{d}) - \Gamma(k)}{\sqrt{nV_C + kV_S(\mathbf{d})}}\right) + \frac{B}{\sqrt{n+k}} + \sqrt{\frac{2 \ln(3|\mathcal{T}|^{d_n})}{K}} + \frac{2 \log e}{\sqrt{d_n}} + \\ + \frac{V_0}{d_n \left(\delta - \frac{\log \sqrt{d_n}}{d_n}\right)^2} + \frac{2 \log 3 |\mathcal{T}|^{d_n} (d_n + 1)^{|\mathcal{X}|^2 |\Theta| |\mathcal{Y}|}}{\exp(d_n(\eta))} + \sqrt{\frac{\log |\mathcal{T}^n| + \varepsilon_0}{2K}} \leq \varepsilon. \end{aligned}$$

Требуемый результат получается рассмотрением ряда Тейлора для $Q^{-1}(\varepsilon - \Delta(k, n))$. Остаточный член $O(\sqrt{n/\log n})$ возникает как произведение члена $O(1/\sqrt{\log n})$ из ряда Тейлора для $Q^{-1}(\varepsilon - \Delta(k, n))$ и члена $O(\sqrt{n})$. \blacktriangle

6.2. Асимптотика второго порядка для нижней границы. В этом пункте вычислим предел нижней границы на вероятность ошибки (доказательство приведено в Приложении).

Теорема 12. Для нижней цены игры $\underline{\nu}(k, n)$ справедлива оценка

$$\begin{aligned} \underline{\nu}(k, n) &\geq \max_{q, P_{\overline{Y}_q}, U} \sup_{\gamma > 0} \left[\sum_s P_S(s) \min_x \left[\mathbf{P}(j_S(s, \mathbf{d}) - i_{X; \overline{Y}_q | U}(x; Y | U) \leq \gamma) + \right. \right. \\ &+ \exp(j_S(s, \mathbf{d}) - \gamma) \sum_{u=1}^U \sum_y P_{U|X}(u|x) P_{\overline{Y}_q | U}(y|u) \times \\ &\left. \left. \times \mathbb{I}\{j_S(s, \mathbf{d}) - i_{X; \overline{Y}_q | U}(x; y | u) > \gamma\} \right] - \frac{U}{\exp(\gamma)} \right] \geq \\ &\geq Q \left(\frac{nC - kR(\mathbf{d}) + \gamma(n)}{\sqrt{nV_C + kV_S(\mathbf{d})} - K_3} \right) - \frac{K_1}{k} - \frac{K_2}{\sqrt{n}} - \frac{B'}{\sqrt{n+k}} - \frac{1}{\sqrt{n+1}}, \end{aligned} \quad (31)$$

где $K_1, K_2, K_3, B' > 0$ – константы, а $\gamma(n) = K_4 \log(n+1)$, $K_4 > 0$.

Справедливо также следующее утверждение, дающее границу второго порядка на скорость передачи.

Следствие 2 (обращение теоремы кодирования второго порядка). Пусть $\varepsilon > 0$. Рассмотрим последовательность пар (k, n) , такую что

$$nC - kR(\mathbf{d}) \leq \sqrt{nV_C + kV_S(\mathbf{d})} Q^{-1}(\varepsilon) - O(\log n).$$

Тогда $\overline{\nu}(k, n) \geq \underline{\nu}(k, n) \geq \varepsilon$.

Доказательство. Возьмем последовательность пар (k, n) , для которой выполнено $nC - kR(\mathbf{d}) + \gamma(n) \leq \sqrt{nV_C + kV_S(\mathbf{d})} Q^{-1}(\varepsilon + \Delta(k, n))$, где

$$\Delta(k, n) = \frac{K_1}{k} + \frac{K_2}{\sqrt{n}} + \frac{1}{\sqrt{n+1}} + \frac{B'}{\sqrt{n+k}}.$$

Подставляя это в (31), получаем

$$Q \left(\frac{nC - kR(\mathbf{d}) + \gamma(n)}{\sqrt{nV_C + kV_S(\mathbf{d})} - K_3} \right) - \frac{K_1}{k} - \frac{K_2}{\sqrt{n}} - \frac{1}{\sqrt{n+1}} - \frac{B'}{\sqrt{n+k}} \geq \varepsilon.$$

Требуемый результат получается рассмотрением ряда Тейлора для $Q^{-1}(\varepsilon + \Delta(k, n))$. Остаточный член $O(\log n)$ возникает за счет члена $\gamma(n)$. \blacktriangle

С использованием полученных границ опишем теперь дисперсию для заданных ε и $k, n \in \mathbb{N}$.

Теорема 13. Для $\varepsilon > 0$ и $k, n \in \mathbb{N}$ дисперсия для стохастического совместного кодирования источника и канала при передаче по ПМК имеет вид

$$\sqrt{V_C + \frac{k}{n} V_S} Q^{-1}(\varepsilon),$$

где V_C и V_S – соответствующие дисперсии канала и источника.

6.3. Доказательство минимаксных теорем. Теперь мы готовы к выводу асимптотических результатов, приведенных в § 3. Мы рассматриваем последовательности пар (k, n) и изучаем предельное поведение границ при конечной длине блока для таких последовательностей. Вначале докажем теорему 1.

Доказательство теоремы 1. Рассмотрим границу из теоремы 10. Выберем последовательность $(k, n) \uparrow \infty$, такую что $\lim k/n < C/R(\mathbf{d})$. Тогда

$$\lim_{k, n \rightarrow \infty} \mathbb{Q} \left(\sqrt{n} \left(\frac{C - \frac{k}{n}R(\mathbf{d}) + \frac{\Gamma(k)}{n}}{\sqrt{V_C + \frac{k}{n}V_S(\mathbf{d})}} \right) \right) = 0.$$

Далее, при $k, n \rightarrow \infty$ и для K и d_n , выбранных согласно (30), имеем

$$\begin{aligned} & \frac{B}{n+k} + \frac{K_0+2}{\sqrt{k}} + \sqrt{\frac{2 \ln(3|\mathcal{T}|^{d_n})}{K}} + \frac{2 \log e}{\sqrt{d_n}} + \frac{V_0}{d_n \left(\delta - \frac{\log \sqrt{d_n}}{d_n} \right)^2} + \\ & + \frac{2 \log 3|\mathcal{T}|^{d_n} (d_n+1)^{|\mathcal{X}|^2|\Theta||\mathcal{Y}|}}{\exp(d_n(\eta))} + \sqrt{\frac{\log |\mathcal{T}^n| + \varepsilon_0}{2K}} \rightarrow 0. \end{aligned}$$

Таким образом, $\overline{\nu}(k, n) \rightarrow 0$, и следовательно, $\underline{\nu}(k, n) \rightarrow 0$. \blacktriangle

Теперь перейдем к доказательству теоремы 2.

Доказательство теоремы 2. Рассмотрим границу из теоремы 12. Выберем последовательность $(k, n) \uparrow \infty$, такую что $\lim k/n > C/R(\mathbf{d})$. Тогда

$$\lim_{k, n \rightarrow \infty} \mathbb{Q} \left(\sqrt{n} \left(\frac{C - \frac{k}{n}R(\mathbf{d}) + \frac{\gamma(n)}{n}}{\sqrt{V_C + \frac{k}{n}V_S(\mathbf{d}) - \frac{K_3}{n}}} \right) \right) = 1.$$

Далее, при $k, n \rightarrow \infty$

$$-\frac{K_1}{k} - \frac{K_2}{\sqrt{n}} - \frac{1}{\sqrt{n+1}} - \frac{B'}{\sqrt{n+k}} \rightarrow 0.$$

Таким образом, $\underline{\nu}(k, n) \rightarrow 1$, и следовательно, $\overline{\nu}(k, n) \rightarrow 1$. \blacktriangle

Заметим, что совпадение верхней и нижней цены, указанных в теоремах 1 и 2, можно показать и без границ второго порядка (см., например, рассуждения в [1]). Следовательно, эти утверждения справедливы, даже когда распределение состояний, на котором достигается пропускная способность, не единственно.

Для доказательства теоремы 3 рассмотрим уточненное определение скорости передачи, отличающееся на $O(1/\sqrt{n})$ от $C/R(\mathbf{d})$. Напомним, что в (9) последовательность дана в виде

$$\frac{k}{n} = \frac{C}{R(\mathbf{d})} + \frac{\rho}{\sqrt{n}}, \tag{32}$$

где $\rho \in \mathbb{R}$ фиксировано. Выведенные нами неасимптотические формулы позволяют оценивать верхнюю и нижнюю цену игры при таком уточненном определении скорости. Для такой последовательности и верхняя, и нижняя цена теперь стремятся к одному и тому же значению, зависящему от ρ . В общем случае это значение не равно ни 0, ни 1. Такое совпадение в более тонкой асимптотике зависит от существования единственного распределения, на котором достигается пропускная способность, для источника помех. Мы обсудим это в следующем пункте.

Доказательство теоремы 3. Подставляя (32) в нижнюю оценку (31), получаем следующий гауссовский член в неравенстве:

$$Q\left(\frac{nC - kR(\mathbf{d}) + \gamma(n)}{\sqrt{nV_C + kV_S(\mathbf{d}) - K_3}}\right) = Q\left(\frac{-\rho R(\mathbf{d}) + \frac{\gamma(n)}{\sqrt{n}}}{\sqrt{V_C + \left(\frac{C}{R(\mathbf{d})} + \frac{\gamma(n)}{\sqrt{n}}\right)V_S(\mathbf{d}) - \frac{K_3}{n}}}\right).$$

Переходя к пределу при $k, n \rightarrow \infty$, получаем

$$\lim_{k, n \rightarrow \infty} \underline{\nu}(k, n) \geq Q\left(\frac{-\rho R(\mathbf{d})}{\sqrt{V_C + \frac{C}{R(\mathbf{d})}V_S(\mathbf{d})}}\right),$$

так как остальные члены в (31) асимптотически стремятся к нулю.

Аналогично, подставляя (32) в (29), получаем следующий гауссовский член в неравенстве:

$$Q\left(\frac{nC - kR(\mathbf{d}) - \Gamma(k)}{\sqrt{nV_C + kV_S(\mathbf{d})}}\right) = Q\left(\frac{-\rho R(\mathbf{d}) - \frac{\Gamma(k)}{\sqrt{n}}}{\sqrt{V_C + \left(\frac{C}{R(\mathbf{d})} + \frac{\rho}{\sqrt{n}}\right)V_S(\mathbf{d})}}\right).$$

Переходя к пределу при $k, n \rightarrow \infty$, получаем

$$\lim_{k, n \rightarrow \infty} \underline{\nu}(k, n) \geq Q\left(\frac{-\rho R(\mathbf{d})}{\sqrt{V_C + \frac{C}{R(\mathbf{d})}V_S(\mathbf{d})}}\right),$$

так как остальные члены в (29) асимптотически стремятся к нулю, что и завершает доказательство. \blacktriangle

6.4. Неединственность дисперсии канала. В общей постановке задачи передачи по ПМК может существовать несколько пар распределений $(P_{\mathbb{X}}, q_{\Theta})$, на которых достигается пропускная способность в (1), и поэтому множества $\Pi_{\mathbb{X}}$ и Π_{Θ} могут не состоять из одного элемента. Когда распределение на входе и распределение состояний, на которых достигается пропускная способность, не единственны, дисперсии канала V_C^+ и V_C^- , определенные в п. 2.4, могут быть не равны между собой. Они равны, когда любое одно из этих распределений (на входе или распределение состояний), на которых достигается пропускная способность, единственно.

В настоящей статье вычислены основанные на дисперсии границы для стохастического совместного кодирования источника и канала при передаче по ПМК в предположении, что распределение состояний, на котором достигается пропускная способность, единственно. Недавно в [11, 27] были получены границы второго порядка на скорость передачи по ПМК для детерминированных и случайных кодов соответственно. Было показано, что в случае неединственного распределения, на котором достигается пропускная способность, скорость достижимости и скорость в обращении теоремы кодирования не равны между собой. Таким образом, естественно предположить, что в случае неединственных распределений, на которых достигается пропускная способность, “более тонкая” теорема о минимаксе может не выполняться, и тем самым граница достижимости и обратная граница на скорость в задаче совместного кодирования источника и канала могут не совпадать. Подтверждение этого факта потребует улучшения стратегий для совместного кодирования источника и канала, что является предметом будущей работы.

§ 7. Заключение

Задача передачи информации при наличии целенаправленных помех рассмотрена в статье как игра с нулевой суммой между командой, состоящей из кодера и декодера, и источником помех, в которой эта команда старается минимизировать вероятность ошибки, а источник помех – максимизировать ее. Эта задача невыпукла в пространстве стратегий команды кодера–декодера, и поэтому теорема о минимаксе может не выполняться. Однако мы показали, что для этой игры справедлива приближенная теорема о минимаксе. Мы вывели верхнюю и нижнюю границы на значения минимакса и максимина игры при конечной длине блока и показали, что асимптотическая теорема о минимаксе имеет место, когда длина блока стремится к бесконечности. В частности, для скоростей выше порогового значения $C/R(\mathbf{d})$ верхняя и нижняя цена игры стремится к нулю, а для скоростей выше $C/R(\mathbf{d})$ эти величины стремятся к единице. Для скоростей, стремящихся в точности к $C/R(\mathbf{d})$ с отклонением не более $O(1/n^\delta)$, где $0 < \delta < 1$, эти величины стремятся к одной и той же константе при предположении технического характера о единственности распределений, на которых достигается пропускная способность.

Авторы благодарны рецензенту за внимательное прочтение первоначального варианта статьи и конструктивные замечания.

ПРИЛОЖЕНИЕ

Начнем со следующей центральной предельной теоремы, принадлежащей Берри и Эссеену (см. [8]).

Теорема 14 (ЦПТ Берри–Эссеена). Зафиксируем $n \in \mathbb{N}$. Пусть W_i – независимые случайные величины. Тогда для $t \in \mathbb{R}$ справедливо неравенство

$$\left| \mathbf{P} \left(\frac{1}{n} \sum_{i=1}^n W_i > D_n + t \sqrt{\frac{V_n}{n}} \right) - \mathbf{Q}(t) \right| \leq \frac{B_n}{\sqrt{n}},$$

где \mathbf{Q} – функция, дополнительная к гауссовской функции распределения,

$$\begin{aligned} D_n &= \frac{1}{n} \sum_{i=1}^n \mathbf{E}[W_i], & V_n &= \frac{1}{n} \sum_{i=1}^n \text{Var}[W_i], \\ A_n &= \frac{1}{n} \sum_{i=1}^n \mathbf{E}[|W_i - \mathbf{E}[W_i]|^3], & B_n &= \frac{c_0 A_n}{V_n^{3/2}}, \quad c_0 > 0. \end{aligned}$$

Доказательство леммы 1. Зафиксируем $\theta \in \mathcal{T}^n$ и рассмотрим K независимых копий случайного кода $\{F_i, \Phi_i\}_{i=1}^K$. Применяя неравенство Хёффдинга, получаем

$$\mathbf{P} \left(\frac{1}{K} \sum_i e_{\mathbf{d}, \theta}(F_i, \Phi_i) \geq \varepsilon' \right) \leq \exp \left(-2K(\varepsilon' - \mathbf{E}[e_{\mathbf{d}, \theta}(F, \Phi)])^2 \right).$$

Так как $e_{\mathbf{d}}(\psi) = \max_{\theta \in \mathcal{T}^n} \mathbf{E}[e_{\mathbf{d}, \theta}(F, \Phi)] < \varepsilon$, то $e_{\mathbf{d}, \theta}(F, \Phi) < \varepsilon \forall \theta \in \mathcal{T}^n$. Используя это неравенство и ε' из (18), получаем, что для всех $\theta \in \mathcal{T}^n$

$$\varepsilon' - \mathbf{E}[e_{\mathbf{d}, \theta}(F, \Phi)] > \varepsilon' - \varepsilon > \sqrt{\log |\mathcal{T}^n| / 2K},$$

откуда

$$\mathbf{P} \left(\frac{1}{K} \sum_i e_{\mathbf{d}, \theta}(F_i, \Phi_i) \geq \varepsilon' \right) \leq \exp \left(-2K^2(\varepsilon' - \mathbf{E}[e_{\mathbf{d}, \theta}(F, \Phi)])^2 \right) <$$

$$\langle \exp\left(-2K \frac{\log |\mathcal{T}^n|}{2K}\right) = \frac{1}{|\mathcal{T}^n|}.$$

Таким образом,

$$\begin{aligned} & \mathbf{P}\left(\frac{1}{K} \sum_i e_{d,\theta}(F_i, \Phi_i) < \varepsilon' \forall \theta \in \mathcal{T}^n\right) = \\ & = 1 - \mathbf{P}\left(\frac{1}{K} \sum_i e_{d,\theta}(F_i, \Phi_i) \geq \varepsilon' \text{ для некоторого } \theta \in \mathcal{T}^n\right) \geq \\ & \geq 1 - \sum_{\theta \in \mathcal{T}^n} \mathbf{P}\left(\frac{1}{K} \sum_i e_{d,\theta}(F_i, \Phi_i) \geq \varepsilon'\right) > 1 - \frac{|\mathcal{T}^n|}{|\mathcal{T}^n|} > 0. \end{aligned}$$

Следовательно, событие $\left\{\sum_i e_{d,\theta}(F_i, \Phi_i)/K < \varepsilon' \forall \theta \in \mathcal{T}^n\right\}$ имеет ненулевую вероятность, и поэтому существует K детерминированных кодов $\{f_i, \varphi_i\}_{i=1}^K$, таких что для всех $\theta \in \mathcal{T}^n$ справедливо

$$\frac{1}{K} \sum_i e_{d,\theta}(f_i, \varphi_i) < \varepsilon',$$

где $\varepsilon' > \varepsilon + \sqrt{\log |\mathcal{T}^n|/2K}$. \blacktriangle

Доказательство теоремы 9. Ослабим оценку в (19), выбирая $P_{X_a}(x_a) = \prod_{i=1}^{d_n} P_{\mathbb{X}}^*(x_i)$, $P_{\mathbb{X}}^* \in \Pi_{\mathbb{X}}$. Пусть \mathcal{A}, Z, K и d_n определены в (24), (26) и (30) соответственно. Положим $U_i := i_{\mathbb{X}^*; \mathbb{Y}_{T_\theta}}(X_{ai}; Y_{ai})$, где $(X_{ai}, Y_{ai}) \sim P_{\mathbb{X}}^* \times \sum_{\theta \in \mathcal{T}} T_\theta(\theta) P_{\mathbb{Y}|\mathbb{X}, \Theta=\theta} \forall i$. Так как память в канале отсутствует, то $i_{X^*; Y_{q^*}}(X_a; Y_a) = \sum_{i=1}^{d_n} U_i$. Таким образом, выбирая $\gamma = \log(\sqrt{d_n}K)$, получаем $\mathbf{P}(i_{X^*; Y_{q^*}}(X_a; Y_a) \leq \gamma) = \mathbf{P}\left(\sum_{i=1}^{d_n} U_i \leq \log(\sqrt{d_n}K)\right)$. Согласно [10, лемма 12.10] имеем $C \leq \mathbf{E}[i_{\mathbb{X}^*; \mathbb{Y}_{T_\theta}}(X_{ai}; Y_{ai})] = \mathbf{E}[U_i]$ для всех $T_\theta \in \mathcal{P}_n(\mathcal{T})$, и поэтому $d_n C \leq \sum_{i=1}^{d_n} \mathbf{E}[U_i]$. Кроме того, подставляя $\log K$ из условия (30), получаем

$$\begin{aligned} & \mathbf{P}\left(\sum_{i=1}^{d_n} U_i \leq \log \sqrt{d_n} + \log K\right) \leq \mathbf{P}\left(\sum_{i=1}^{d_n} U_i \leq \log \frac{\sqrt{d_n}}{\exp d_n \delta} + \sum_{i=1}^{d_n} \mathbf{E}[U_i]\right) \leq \\ & \leq \mathbf{P}\left(\left|\sum_{i=1}^{d_n} (U_i - \mathbf{E}[U_i])\right| \geq \log \frac{\exp d_n \delta}{\sqrt{d_n}}\right), \end{aligned}$$

где последнее неравенство вытекает из неравенства треугольника. Применяя неравенство Чебышева и переходя к супремуму по θ_a , получаем

$$\begin{aligned} & \sup_{\theta_a \in \mathcal{T}^{d_n}} \mathbf{P}\left(\left|\sum_{i=1}^{d_n} (U_i - \mathbf{E}[U_i])\right| \geq d_n \delta - \log \sqrt{d_n}\right) \leq \\ & \leq \sup_{\theta_a \in \mathcal{T}^{d_n}} \frac{d_n}{(d_n \delta - \log \sqrt{d_n})^2} \text{Var}(i_{\mathbb{X}^*; \mathbb{Y}_{T_{\theta_a}}}(\mathbb{X}_a; \mathbb{Y}_a)) \leq \\ & \leq \frac{V_0}{d_n \left(\delta - \frac{\log \sqrt{d_n}}{d_n}\right)^2}, \end{aligned} \tag{33}$$

где неравенство (33) следует из соотношения (25).

Оценки для остальных двух слагаемых $\mathbf{P}(Z(X_a, \bar{X}_a, Y_a) = 0, (X_a, Y_a) \in \mathcal{A} | \theta_a) \leq \leq 1/\sqrt{d_n}K$ и $\mathbf{P}(Z(X_a, \bar{x}_a, Y_a) = 0, (X_a, Y_a) \in \mathcal{A} | \bar{X}_a = \bar{x}_a, \theta_a) \leq (d_n + 1)^{|\mathcal{X}|^2|\Theta||\mathcal{Y}|} \times \times \exp(-d_n\eta)$ следуют из доказательства теоремы 4 работы [11]. Из (33) и этих оценок вытекает требуемая граница. \blacktriangle

Доказательство теоремы 10. Для вывода этой границы построим случайный код (F, Φ) для совместного кодирования источника и канала, т.е. распределение ψ на множестве кодов $\{(f, \varphi) \mid f: \mathcal{S}^k \rightarrow \mathcal{X}^n, \varphi: \mathcal{Y}^n \rightarrow \mathcal{S}^k\}$. По теореме 6 достаточно выбрать распределения P_X и $P_{\hat{S}}$. Пусть $P_{\mathbb{X}}^*$ – распределение из множества $\Pi_{\mathbb{X}}$. Положим $P_X(x) := \prod_{i=1}^n P_{\mathbb{X}}^*(x_i)$ для $x \in \mathcal{X}^n$. Далее, возьмем $P_{\hat{S}}(\hat{s}) = \prod_{i=1}^k P_{\hat{\mathbb{S}}}^*(\hat{s}_i)$, $\hat{s} \in \mathcal{S}^k$, где $P_{\hat{\mathbb{S}}}^*$ достигает оптимума в (2). Очевидно, что при таком выборе распределений

$$i_{X^*; Y_{q^*}}(X; Y) = \sum_{i=1}^n i_{\mathbb{X}^*; \mathbb{Y}_{q^*}}(X_i; Y_i), \quad j_S(S, \mathbf{d}) = \sum_{j=1}^k j_S(S_j, \mathbf{d}).$$

Рассмотрим слагаемые, составляющие вероятность ошибки случайного кода в теореме 8. Записывая максимизацию по $\theta_b \in \mathcal{T}^n$ как максимизацию по $q \in \mathcal{P}(\mathcal{T}^n)$, можно представить границу в виде

$$\max_{q \in \mathcal{P}(\mathcal{T}^n)} \left[\mathbf{E} \left[\exp \left(- \left| i_{X^*; Y_{q^*}}(X_b; Y_b) - \log \frac{\gamma}{P_{\hat{S}}(\mathcal{B}_{\mathbf{d}}(S))} \right|^+ \right) \right] + e^{1-\gamma} \right],$$

где $q(\theta) = \prod_{i=1}^n q_i(\theta_i)$ для $q_i \in \mathcal{P}(\mathcal{T})$, $\theta \in \mathcal{T}^n$. Пусть $h(X_b, Y_b, S) := \sum_{j=1}^n i_{\mathbb{X}^*; \mathbb{Y}_{q^*}}(X_{bj}; Y_{bj}) - - \log(\gamma/P_{\hat{S}}(\mathcal{B}_{\mathbf{d}}(S)))$. Далее, зададим множество

$$\mathcal{D} := \left\{ s \in \mathcal{S}^k : \log \frac{1}{P_{\hat{S}}(\mathcal{B}_{\mathbf{d}}(s))} \leq \sum_{i=1}^k j_S(s_i, \mathbf{d}) + \left(\bar{c} - \frac{1}{2} \right) \log k + c \right\},$$

где \bar{c}, c – константы, определенные в [9, лемма 5]. Определим случайную величину

$$W_\ell = W_\ell(n, k) := \begin{cases} i_{\mathbb{X}^*; \mathbb{Y}_{q^*}}(X_{b\ell}; Y_{b\ell}), & \text{если } \ell \leq n, \\ -j_S(S_{\ell-n}, \mathbf{d}), & \text{если } n < \ell \leq n+k. \end{cases} \quad (34)$$

Математическое ожидание $\mathbf{E}[\exp(-|h(X_b, Y_b, S)|^+)]$ можно представить в виде

$$\begin{aligned} & \mathbf{E}[\exp(-|h(X_b, Y_b, S)|^+) \mathbb{I}\{S \in \mathcal{D}\}] + \mathbf{E}[\exp(-|h(X_b, Y_b, S)|^+) \mathbb{I}\{S \notin \mathcal{D}\}] \leq \\ & \leq \mathbf{E} \left[\exp \left(- \left| \sum_{\ell=1}^{n+k} W_\ell - \log(k^{\bar{c}-\frac{1}{2}} \exp(c)\gamma) \right|^+ \right) \right] + \frac{K_0}{\sqrt{k}}, \end{aligned} \quad (35)$$

где первое слагаемое в (35) вытекает из определения множества \mathcal{D} , а второе получается из неравенства $\mathbf{E}[\exp(-|h(X_b, Y_b, S)|^+) \mathbb{I}\{S \notin \mathcal{D}\}] \leq \mathbf{P}(S \notin \mathcal{D})$ и леммы 5 работы [9].

Введем следующие моменты, которые будут использоваться для оценки первого слагаемого в (35) с помощью ЦПТ Берри – Ессеена:

$$\begin{aligned} D_{n+k}(q) &= \frac{1}{n+k} \sum_{\ell=1}^{n+k} \mathbf{E}[W_\ell], & V_{n+k}(q) &= \frac{1}{n+k} \sum_{\ell=1}^{n+k} \text{Var}[W_\ell], \\ A_{n+k}(q) &= \frac{1}{n+k} \sum_{\ell=1}^{n+k} \mathbf{E}[|W_\ell - \mathbf{E}[W_\ell]|^3], & B_{n+k}(q) &= \frac{c_0 A_{n+k}(q)}{V_{n+k}^{3/2}(q)}, \quad c_0 > 0. \end{aligned} \quad (36)$$

Отметим, что эти моменты вычисляются относительно распределения $P_X \times \times \sum_{\theta} q(\theta) P_{Y|X, \Theta=\theta} \times P_S$. Далее, определим следующее множество:

$$\mathcal{H} = \left\{ (x, y, s) \in \mathcal{X}^n \times \mathcal{Y}^n \times \mathcal{S}^k : \frac{1}{n+k} \sum_{\ell=1}^{n+k} W_\ell > D_{n+k}(q) - t_{k,n} \sqrt{\frac{V_{n+k}(q)}{n+k}} \right\},$$

где $t_{k,n} > 0$ будет выбрано позже. Для краткости введем следующие обозначения:

$$g(X_b, Y_b, S) = \sum_{\ell=1}^{n+k} W_\ell - \log(k^{\bar{c}-\frac{1}{2}} \exp(c)\gamma), \quad (37)$$

$$\Gamma_{n+k}(q) = (n+k) \left(D_{n+k}(q) - t_{k,n} \sqrt{\frac{V_{n+k}(q)}{n+k}} \right) - \log(k^{\bar{c}-\frac{1}{2}} \exp(c) \log \gamma). \quad (38)$$

Тогда член с экспонентой в (35) можно записать в виде $\mathbf{E}[\exp(-|g(X_b, Y_b, S)|^+)]$, что можно представить как

$$\begin{aligned} & \mathbf{E}[\exp(-|g(X_b, Y_b, S)|^+) \mathbb{I}\{(X_b, Y_b, S) \in \mathcal{H}\}] + \\ & + \mathbf{E}[\exp(-|g(X_b, Y_b, S)|^+) \mathbb{I}\{(X_b, Y_b, S) \notin \mathcal{H}\}] \leq \\ & \leq \mathbf{E}[\exp(-|\Gamma_{n+k}(q)|^+) \mathbb{I}\{(X_b, Y_b, S) \in \mathcal{H}\}] + \mathbf{E}[\mathbb{I}\{(X_b, Y_b, S) \notin \mathcal{H}\}], \end{aligned} \quad (39)$$

где первое слагаемое в (39) вытекает из определения множества \mathcal{H} , а второе получается с помощью неравенства $\exp(-|\cdot|^+) \leq 1$. Из этого с учетом (35) получаем следующую оценку:

$$\begin{aligned} & \max_{q \in \mathcal{P}(\mathcal{T}^n)} \left[\mathbf{E} \left[\exp \left(- \left| i_{X^*; Y_{q^*}}(X_b; Y_b) - \log \frac{\gamma}{P_{\hat{S}}(\mathcal{B}_d(S))} \right|^+ \right) \right] + e^{1-\gamma} \right] \leq \\ & \leq \max_{q \in \mathcal{P}(\mathcal{T}^n)} \mathbf{E}[\exp(-|\Gamma_{n+k}(q)|^+) \mathbb{I}\{(X_b, Y_b, S) \in \mathcal{H}\}] + \\ & + \max_{q \in \mathcal{P}(\mathcal{T}^n)} \mathbf{P}((X_b, Y_b, S) \notin \mathcal{H}) + e^{1-\gamma} + \frac{K_0}{\sqrt{k}}. \end{aligned} \quad (40)$$

Чтобы оценить сверху первое слагаемое в (40), вычислим максимум величины $\exp(-|\Gamma_{n+k}(q)|^+)$ по всем распределениям. Так как функция $\exp(-|\cdot|^+)$ убывает по $\Gamma_{n+k}(q)$, получаем

$$\max_{q \in \mathcal{P}(\mathcal{T}^n)} \exp(-|\Gamma_{n+k}(q)|^+) \leq \exp \left(- \left| \min_{q \in \mathcal{P}(\mathcal{T}^n)} \Gamma_{n+k}(q) \right|^+ \right).$$

Для вычисления последнего минимума рассмотрим следующие величины:

$$D_{n+k}(q) = \frac{1}{n+k} \sum_{i=1}^n \mathbf{E}'[i_{\mathbb{X}^*; \mathbb{Y}_{q_\Theta^*}}(\mathbb{X}_{bi}; \mathbb{Y}_{bi})] - \frac{k}{n+k} R(\mathbf{d}),$$

$$V_{n+k}(q) = \frac{1}{n+k} \sum_{i=1}^n \text{Var}(i_{\mathbb{X}^*, \mathbb{Y}_{q_\Theta^*}}(\mathbb{X}_{bi}; \mathbb{Y}_{bi})) + \frac{k}{n+k} V_S(\mathbf{d}),$$

где моменты берутся относительно распределения $P_{\mathbb{X}}^* \times \sum_{\theta \in \mathcal{T}} q_i(\theta) P_{\mathbb{Y}|\mathbb{X}, \Theta=\theta}$. Таким образом, этот минимум имеет вид

$$\begin{aligned} \min_{q \in \mathcal{P}(\mathcal{T}^n)} (n+k) \left(D_{n+k}(q) - t_{k,n} \sqrt{\frac{V_{n+k}(q)}{n+k}} \right) &= \\ &= nC - kR(\mathbf{d}) - (n+k) \max_{q_i \in \Pi_\Theta} t_{k,n} \sqrt{\frac{V_{n+k}(q)}{n+k}} + O(1), \end{aligned} \quad (41)$$

где максимум в (41) ограничен на множество Π_Θ согласно [8, леммы 63 и 64]. Так как $P_{\mathbb{X}}^* \in \Pi_{\mathbb{X}}$ и $q_\Theta^* \in \Pi_\Theta$, $|\Pi_\Theta| = 1$, то $\max_{q_i \in \Pi_\Theta} V_{n+k}(q) = nV_C$. Таким образом,

$$\begin{aligned} \min_{q \in \mathcal{P}(\mathcal{T}^n)} (n+k) \left(D_{n+k}(q) - t_{k,n} \sqrt{\frac{V_{n+k}(q)}{n+k}} \right) &= \\ &= nC - kR(\mathbf{d}) - t_{k,n} \sqrt{nV_C + kV_S(\mathbf{d})} + O(1), \end{aligned}$$

Выберем $t_{k,n}$ в виде $t_{k,n} = (nC - kR(\mathbf{d}) - \bar{c} \log k - \log \gamma - c) / \sqrt{nV_C + kV_S(\mathbf{d})}$. Тогда получаем $\min_{q \in \mathcal{P}(\mathcal{T}^n)} \Gamma_{n+k}(q) \geq \frac{1}{2} \log k + O(1)$. Подставляя это в (40), получаем следующую оценку сверху:

$$\begin{aligned} \max_{q \in \mathcal{P}(\mathcal{T}^n)} \frac{1}{\sqrt{k}} \mathbf{P}((X_b, Y_b, S) \in \mathcal{H}) + \max_{q \in \mathcal{P}(\mathcal{T}^n)} \mathbf{P}((X_b, Y_b, S) \notin \mathcal{H}) + e^{1-\gamma} + \frac{K_0}{\sqrt{k}} &\leq \\ \leq Q(t_{k,n}) + \max_{q \in \mathcal{P}(\mathcal{T}^n)} \frac{B_{n+k}(q)}{\sqrt{n+k}} + \frac{K_0 + 2}{\sqrt{k}}, \end{aligned} \quad (42)$$

где (42) получается в результате выбора $\gamma = (\log_e k)/2 + 1$ и оценивания вероятности $\mathbf{P}((X_b, Y_b, S) \notin \mathcal{H})$ с помощью ЦПТ Берри – Ессена.

Напомним, что $B_{n+k}(q)$ имеет вид $B_{n+k}(q) = c_0 A_{n+k}(q) / V_{n+k}^{3/2}(q)$. Предполагая, что $\min_{\ell \in \{1, \dots, n+k\}} \text{Var}[W_\ell] \neq 0$, можно оценить $B_{n+k}(q)$ с помощью определений $A_{n+k}(q)$ и $V_{n+k}(q)$ как

$$B_{n+k}(q) \leq \frac{c_0 \max_{\ell} \mathbf{E}[|W_\ell - \mathbf{E}[W_\ell]|^3]}{\left(\min_{\ell} \text{Var}[W_\ell] \right)^{3/2}}. \quad (43)$$

Из (22) и (23) следует, что правая часть конечна при всех q и не зависит от k, n . Таким образом, $\max_{q \in \mathcal{P}(\mathcal{T}^n)} B_{n+k}(q)$ является конечной константой. Подставляя указанные выше значения $t_{k,n}$, выбирая $\max_{q \in \mathcal{P}(\mathcal{T}^n)} B_{n+k}(q) \leq B$, где $B > 0$, и используя (42), получаем требуемую границу. \blacktriangle

Доказательство теоремы 12. В силу (11) для получения границы снизу на $\underline{\nu}(k, n)$ достаточно построить q , $P_{\hat{\mathcal{Y}}_q}$, случайную величину U и γ . Возьмем $q(\theta) = q^*(\theta) = \prod_{i=1}^n q_\Theta^*(\theta_i)$, где $q_\Theta^* \in \Pi_\Theta$. Пусть U – число типов в $\mathcal{P}_n(\mathcal{X})$, и пусть $\mathcal{U} = \{1, \dots, U\}$ – номера каждого из этих типов, т.е. для данной последовательности $x \in \mathcal{X}^n$ функция U задает ее тип, описываемый некоторым номером $u \in \mathcal{U}$. Далее,

определим $P_{\overline{Y}_q|U}$ как

$$P_{\overline{Y}_q|U}(y|u) = (P_X q^* P_{Y|X, \Theta})(y) = \sum_{x, \theta} P_X(x) q^*(\theta) P_{Y|X, \Theta}(y|x, \theta),$$

где $P_X(x) = \prod_{i=1}^n T_x(x_i)$, $x \in \mathcal{X}^n$, а $T_x \in \mathcal{P}_n(\mathcal{X})$ – тип, соответствующий номеру u .

Таким образом, $i_{X; \overline{Y}_q|U}(x; Y|u) = \sum_{i=1}^n i_{X; Y_{q^*}}(x_i; Y_i)$, где

$$i_{X; Y_{q^*}}(x'; y) := \log \frac{(q_{\Theta}^* P_{Y|X, \Theta})(y|x')}{(T_x q_{\Theta}^* P_{Y|X, \Theta})(y)}, \quad x' \in \mathcal{X}, \quad y \in \mathcal{Y},$$

$$\text{и } j_S(s, \mathbf{d}) = \sum_{j=1}^k j_S(s_j, \mathbf{d}).$$

Поскольку q выбираются как н.о.р. величины, по существу имеется канал, в котором $Y \sim \prod_{i=1}^n \sum_{\theta_i \in \mathcal{T}} q_{\Theta}^*(\theta_i) P_{Y|X=x_i, \Theta=\theta_i}$ при входе $x = (x_1, \dots, x_n)$. Таким образом, левая часть неравенства (31) представляет собой обращение теоремы кодирования для стандартного ДКБП без целенаправленных помех, где каналом является такой усредненный канал. Повторяя рассуждения из [9, Приложение С], приходим к следующему неравенству:

$$\begin{aligned} & \max_{q, P_{\overline{Y}_q}, U} \sup_{\gamma > 0} \left[\sum_s P_S(s) \min_x \left[\mathbf{P}(j_S(s, \mathbf{d}) - i_{X; \overline{Y}_q|U}(x; Y|U) \leq \gamma) + \right. \right. \\ & + \exp(j_S(s, \mathbf{d}) - \gamma) \sum_{u=1}^U \sum_y P_{U|X}(u|x) P_{\overline{Y}_q|U}(y|u) \times \\ & \left. \left. \times \mathbb{I}\{j_S(s, \mathbf{d}) - i_{X; \overline{Y}_q|U}(x; y|u) > \gamma\} \right] - \frac{U}{\exp(\gamma)} \right] \geq \\ & \geq \mathbf{P} \left(\sum_{i=1}^n i_{X; Y_{q_{\Theta}^*}}(x_i^*; Y_i) - \sum_{j=1}^k j_S(s_j, \mathbf{d}) \leq -\gamma \right) - \frac{K_1}{k} - \frac{K_2}{\sqrt{n}} - \\ & - (n+1)^{|\mathcal{X}|-1} \exp(-\gamma), \end{aligned} \quad (44)$$

где второе слагаемое в левой части ограничивается снизу нулем, K_1 и K_2 – некоторые константы, а $x^* = (x_1^*, \dots, x_n^*)$ – последовательность, тип которой T_{x^*} доставляет

$$\min_{T_x \in \mathcal{P}_n(\mathcal{X})} |T_x - P_{\mathbb{X}}^*|, \quad (45)$$

где $P_{\mathbb{X}}^* \in \Pi_{\mathbb{X}}$. Пусть

$$W_\ell = W_\ell(n, k) := \begin{cases} i_{X; Y_{q_{\Theta}^*}}(x_\ell^*; Y_\ell), & \text{если } \ell \leq n, \\ j_S(S_{n-\ell}, \mathbf{d}), & \text{если } n < \ell \leq n+k. \end{cases}$$

Определим следующие моменты случайной величины W_ℓ :

$$\begin{aligned} D_{n+k} &= \frac{1}{n+k} \sum_{\ell=1}^{n+k} \mathbf{E}[W_\ell], & V_{n+k} &= \frac{1}{n+k} \sum_{\ell=1}^{n+k} \text{Var}[W_\ell], \\ A_{n+k} &= \frac{1}{n+k} \sum_{\ell=1}^{n+k} \mathbf{E}[|W_\ell - \mathbf{E}[W_\ell]|^3], & B'_{n+k} &= \frac{c_0 A_{n+k}}{V_{n+k}^{3/2}}, \quad c_0 > 0. \end{aligned}$$

Согласно ЦПТ Берри – Ессеена имеем

$$\mathbb{P}\left(\sum_{\ell=1}^{n+k} W_{\ell} \leq -\gamma\right) \geq \mathbb{Q}\left(\frac{D_{n+k} + \frac{\gamma}{n+k}}{\sqrt{\frac{V_{n+k}}{n+k}}}\right) - \frac{B'_{n+k}}{\sqrt{n+k}}. \quad (46)$$

Справедливы также следующие неравенства из [9, Приложение С]:

$$D_{n+k} \leq \frac{n}{n+k}C - \frac{k}{n+k}R(\mathbf{d}), \quad (47)$$

$$V_{n+k} \geq \frac{n}{n+k}V_C + \frac{k}{n+k}V_S(\mathbf{d}) - \frac{K_3}{n+k}, \quad (48)$$

где $K_3 > 0$ – некоторая константа. Далее, с учетом (22) и (23) можно показать, что величина A_{n+k} ограничена, и поэтому B'_{n+k} ограничено некоторой константой $B' > 0$. Используя (47) и (48), получаем

$$\mathbb{Q}\left(\frac{D_{n+k} + \frac{\gamma}{n+k}}{\sqrt{\frac{V_{n+k}}{n+k}}}\right) \geq \mathbb{Q}\left(\frac{nC - kR(\mathbf{d}) + \gamma}{\sqrt{nV_C + kV_S(\mathbf{d}) - K_3}}\right).$$

Подставляя это в (46), выбирая $\gamma = (|\mathcal{X}| - 1/2) \log(n+1)$ и применяя (44), получаем требуемую границу. ▲

СПИСОК ЛИТЕРАТУРЫ

1. Vora A.S., Kulkarni A.A. A Minimax Theorem for Finite Blocklength Joint Source-Channel Coding over an AVC // 2019 National Conf. on Communications (NCC 2019). Bangalore, India. Feb. 20–23, 2019. P. 1–6. <https://doi.org/10.1109/NCC.2019.8732205>
2. Humayed A., Lin J., Li F., Luo B. Cyber-Physical Systems Security—A Survey // IEEE Internet of Things J. 2017. V. 4. № 6. P. 1802–1831. <https://doi.org/10.1109/JIOT.2017.2703172>
3. Slay J., Miller M. Lessons Learned from the Maroochy Water Breach // Critical Infrastructure Protection (Proc. Int. Conf. ICCIP'2007. Hanover, NH, USA. Mar. 19–21, 2007). Boston: Springer, 2008. https://doi.org/10.1007/978-0-387-75462-8_6
4. Langner R. Stuxnet: Dissecting a Cyberwarfare Weapon // IEEE Secur. Priv. 2011. V. 9. № 3. P. 49–51. <https://doi.org/10.1109/MSP.2011.67>
5. Maschler M., Solan E., Zamir S. Game Theory. Cambridge: Cambridge Univ. Press, 2013.
6. Kulkarni A.A., Coleman T.P. An Optimizer's Approach to Stochastic Control Problems with Nonclassical Information Structures // IEEE Trans. Autom. Control. 2015. V. 60. № 4. P. 937–949. <https://doi.org/10.1109/TAC.2014.2362596>
7. Ahlswede R. A Note on the Existence of the Weak Capacity for Channels with Arbitrarily Varying Channel Probability Functions and Its Relation to Shannon's Zero Error Capacity // Ann. Math. Statist. 1970. V. 41. № 3. P. 1027–1033. <https://doi.org/10.1214/aoms/1177696979>
8. Polyanskiy Y., Poor H.V., Verdú S. Channel Coding Rate in the Finite Blocklength Regime // IEEE Trans. Inform. Theory. 2010. V. 56. № 5. P. 2307–2359. <https://doi.org/10.1109/TIT.2010.2043769>
9. Kostina V., Verdú S. Lossy Joint Source-Channel Coding in the Finite Blocklength Regime // IEEE Trans. Inform. Theory. 2013. V. 59. № 5. P. 2545–2575. <https://doi.org/10.1109/TIT.2013.2238657>
10. Csizár I., Körner J. Information Theory: Coding Theorems for Discrete Memoryless Systems. Cambridge, UK: Cambridge Univ. Press, 2011.

11. *Kosut O., Kliewer J.* Finite Blocklength and Dispersion Bounds for the Arbitrarily-Varying Channel // Proc. 2018 IEEE Int. Symp. on Information Theory (ISIT'2018). Vail, CO, USA. June 17–22, 2018. P. 2007–2011. <https://doi.org/10.1109/ISIT.2018.8437724>
12. *Jose S.T., Kulkarni A.A.* Linear Programming-Based Converses for Finite Blocklength Lossy Joint Source-Channel Coding // IEEE Trans. Inform. Theory. 2017. V. 63. № 11. P. 7066–7094. <https://doi.org/10.1109/TIT.2017.2738634>
13. *Jose S.T., Kulkarni A.A.* Improved Finite Blocklength Converses for Slepian–Wolf Coding via Linear Programming // IEEE Trans. Inform. Theory. 2019. V. 65. № 4. P. 2423–2441. <https://doi.org/10.1109/TIT.2018.2873623>
14. *Borden J.M., Mason D.M., McEliece R.J.* Some Information Theoretic Saddlepoints // SIAM J. Control Optim. 1985. V. 23. № 1. P. 129–143. <https://doi.org/10.1137/0323011>
15. *Hegde M.V., Stark W.E., Teneketzis D.* On the Capacity of Channels with Unknown Interference // IEEE Trans. Inform. Theory. 1989. V. 35. № 4. P. 770–783. <https://doi.org/10.1109/18.32154>
16. *Başar T., Wu Y.-W.* A Complete Characterization of Minimax and Maximin Encoder-Decoder Policies for Communication Channels with Incomplete Statistical Description // IEEE Trans. Inform. Theory. 1985. V. 31. № 4. P. 482–489. <https://doi.org/10.1109/TIT.1985.1057076>
17. *Hughes B., Narayan P.* Gaussian Arbitrarily Varying Channels // IEEE Trans. Inform. Theory. 1987. V. 33. № 2. P. 267–284. <https://doi.org/10.1109/TIT.1987.1057288>
18. *Jose S.T., Kulkarni A.A.* On a Game Between a Delay-Constrained Communication System and a Finite State Jammer // Proc. 2018 IEEE Conf. on Decision and Control (CDC'2018). Miami, FL, USA. Dec. 17–19, 2018. P. 5063–5068. <https://doi.org/10.1109/CDC.2018.8618987>
19. *Jose S.T., Kulkarni A.A.* Shannon Meets von Neumann: A Minimax Theorem for Channel Coding in the Presence of a Jammer // IEEE Trans. Inform. Theory. 2020. V. 66. № 5. P. 2842–2859. <https://doi.org/10.1109/TIT.2020.2971682>
20. *Blackwell D., Breiman L., Thomasian A.J.* The Capacities of Certain Channel Classes under Random Coding // Ann. Math. Statist. 1960. V. 31. № 3. P. 558–567. <https://doi.org/10.1214/aoms/1177705783>
21. *Ahlsvede R.* Elimination of Correlation in Random Codes for Arbitrarily Varying Channels // Z. Wahrsch. Verw. Gebiete. 1978. V. 44. № 2. P. 159–175. <https://doi.org/10.1007/BF00533053>
22. *Lapidoth A., Narayan P.* Reliable Communication under Channel Uncertainty // IEEE Trans. Inform. Theory. 1998. V. 44. № 6. P. 2148–2177. <https://doi.org/10.1109/18.720535>
23. *Cover T.M., Thomas J.A.* Elements of Information Theory. Hoboken, NJ: Wiley, 2012.
24. *Kostina V., Verdú S.* Fixed-Length Lossy Compression in the Finite Blocklength Regime // IEEE Trans. Inform. Theory. 2012. V. 58. № 6. P. 3309–3338. <https://doi.org/10.1109/TIT.2012.2186786>
25. *Shannon C.E.* Coding Theorems for a Discrete Source with a Fidelity Criterion // IRE Nat. Conv. Rec. 1959. Part 4. P. 142–163.
26. *Conforti M., Cornuéjols G., Zambelli G.* Integer Programming. New York: Springer, 2014.
27. *Kosut O., Kliewer J.* Dispersion of the Discrete Arbitrarily-Varying Channel with Limited Shared Randomness // Proc. 2017 IEEE Int. Symp. on Information Theory (ISIT'2017). Aachen, Germany. June 25–30, 2017. P. 1242–1246. <https://doi.org/10.1109/ISIT.2017.8006727>

Вора Анудж Самиркумар
Кулкарни Анкур Ачот
 Индийский технологический институт Бомбея,
 Мумбаи, Индия
 anujvora@iitb.ac.in
 kulkarni.ankur@iitb.ac.in

Поступила в редакцию
 20.06.2019
 После доработки
 17.12.2020
 Принята к публикации
 05.03.2021