ПРОБЛЕМЫ ПЕРЕДАЧИ ИНФОРМАЦИИ

Том 56 2020 Вып. 3

УДК 621.391.1 : 519.72

 Б. Накибоглу

ГРАНИЦА СФЕРИЧЕСКОЙ УПАКОВКИ ДЛЯ КАНАЛОВ БЕЗ ПАМЯТИ[1]

Границы сферической упаковки с полиномиальными по длине блока коэффициентами при экспоненте выводятся с помощью метода Августина для кодов в двух семействах каналов без памяти: (возможно, нестационарных) каналов без памяти с (возможно, многими) аддитивными ограничениями по стоимости и стационарных каналов без памяти с выпуклыми ограничениями на композицию (т.е. эмпирическое распределение, тип) входных кодовых слов. Получен также вариант границы Галлагера, показывающий, что эти границы сферической упаковки точны в смысле экспоненциальной по длине блока скорости убывания вероятности ошибки при довольно широких предположениях.

Ключевые слова: метод Августина, экспонента сферической упаковки, экспонента ошибки, функция надежности, каналы без памяти, гауссовские каналы, пуассоновские каналы.

DOI: 10.31857/S0555292320030018

§ 1. Введение

Большинство известных доказательств границы сферической упаковки (ГСУ) относятся либо к стационарным каналам с конечными множествами входов [2–14], либо к стационарным каналам со специальной структурой шума, например, с пуассоновским или гауссовским шумом [15–22]. Доказательства ГСУ, основанные на методе Августина, являются исключениями: в работах [23–25] не предполагается ни конечность множества входов, ни специальная структура шума, ни стационарность канала. Однако в [23; 24, § 31; 25] установлена ГСУ для каналов-произведений, а не для каналов без памяти, поэтому доказательства ГСУ для кодов с ограничениями по композиции[2] в стационарных каналах [9–20], в том числе для важного специального случая кодов с ограничениями по стоимости [15–20], не охвачены работами [23; 24, § 31; 25]. В [24, § 36] Августин доказал ГСУ для (возможно, нестационарных) каналов без памяти с ограничениями по стоимости в предположении ограниченной функции стоимости. Условия в [24, теорема 36.6] позволили охватить все ранее рассмотренные модели [2–22], кроме гауссовских [17–20].

Теорема 2, представленная в § 3, устанавливает ГСУ при условиях, полностью охватывающих все модели, изучавшиеся в [2–25], используя результаты работы [26], в которой были исследованы августиновские информационные величины. Мы применяем работу [26] и августиновские информационные величины аналогично тому, как в [25] применялась работа [27] и информационные величины Реньи. Для каналов-

[1] Работа была частично представлена на Международном симпозиуме IEEE по теории информации 2017 г. [1].

[2] Как указано в [12, с. 183], ГСУ для кодов с постоянной композицией содержалась в [9] с неполным доказательством. Первое полное доказательство ГСУ для кодов с постоянной композицией было дано в [10].

произведений в [25, теорема 2] были улучшены предыдущие результаты Августина из [23; 24, § 31], а именно установлена ГСУ с коэффициентом при экспоненте, полиномиальным по длине блока n, в предположении, что пропускные способности Реньи порядка $1/2$ каналов-компонент имеют порядок $O(\ln n)$. Для каналов без памяти с ограничениями по стоимости теорема 2 улучшает коэффициент, установленный в [24, теорема 36.6], аналогичным образом: с $e^{-O(\sqrt{n})}$ до $e^{-O(\ln n)}$. Коэффициент в теореме 2, однако, уступает коэффициентам, полученным в [3–6] для различных симметричных каналов, в [13] для кодов с постоянной композицией в дискретных стационарных каналах-произведениях, в [17] для стационарных гауссовских каналов, и в [21, 22] для некоторых некогерентных каналов с замиранием. Определение оптимального коэффициента в духе [3–5] остается открытой задачей в общем случае[3]. Как и [24, теорема 36.6], теорема 2 справедлива и для нестационарных каналов, но в отличие от [24, теорема 36.6] в теореме 2 функции стоимости не предполагаются ограниченными.

В большинстве ранее полученных ГСУ [2–22] предполагалась стационарность канала. Имея стационарный канал-произведение, можно построить стационарный канал без памяти, налагая ограничения на композицию (т.е. тип, эмпирическое распределение) входных кодовых слов. Ограничения по стоимости можно рассматривать как выпуклый частный случай этих более общих ограничений по композиции. Такая интерпретация, рассматриваемая вместе с техникой выбрасывания с учетом композиции, является одним из основных мотивов к изучению кодов с постоянной композицией. Однако техника выбрасывания с учетом композиции полезна только тогда, когда множество входов канала конечно. Тем не менее, если множество ограничений на композицию кодовых слов выпукло, можно получить ГСУ с полиномиальным коэффициентом, используя августиновские информационные величины (см. теорему 1 в § 3). Доказательство теоремы 1 использует августиновский центр множества ограничений, а не августиновское среднее наиболее частой композиции кода. Отметим, что наиболее частая композиция кода может даже состоять из не более одного кодового слова, когда множество входов бесконечно. Условия теоремы 1 достаточно общие, чтобы охватить условия всех известных нам предыдущих доказательств ГСУ для каналов без памяти, кроме тех, что основаны на методе Августина [23–25]. Теоремы 1 и 2 дают асимптотические ГСУ, но в их доказательстве используются неасимптотические ГСУ, представленные в леммах 9 и 10.

Из ГСУ следует, что экспоненциальная скорость убывания оптимальной вероятности ошибки с ростом длины блока, т.е. функция надежности, или экспонента ошибки, ограничена сверху экспонентой сферической упаковки (ЭСУ). Для рассматриваемых каналов без памяти августиновский вариант границы Галлагера показывает, что ЭСУ также ограничивает функцию надежности снизу при условии, что разрешено списочное декодирование. Августиновский вариант границы Галлагера представлен в п. 2.4. Но чтобы показать, что ЭСУ является нижней границей функции надежности при списочном декодировании, можно также использовать стандартные результаты, такие как [29, 30], с небольшими изменениями. Таким образом, августиновский вариант интересен нам не тем, что из него вытекают результаты о функции надежности, а тем, каким образом они из него вытекают. Отличительной особенностью августиновского варианта является то, что он устанавливает результаты о достижимости через информацию Августина, а не информацию Реньи, используемой в стандартной форме границы Галлагера [29]. Для этого августиновский вариант опирается на свойство августиновского среднего как неподвижной точки (формула (7)). Стоит отметить, что в [30] неявно использовалось это утверждение о неподвижной точке, хотя и иным способом.

[3] В работе [28] выведена улучшенная ГСУ (оптимальная в смысле коэффициента в неособых случаях) для всех случаев, рассмотренных в [3–6, 13, 17, 21, 22], с использованием августиновских информационных величин и результатов работы [26].

Прежде чем приступить к детальному обсуждению вопроса, укажем на одну тонкость при выводе ГСУ, на которую обычно не обращают внимания. В [31] утверждалось, что там доказана ГСУ для произвольных стационарных каналов-произведений без использования каких-либо рассуждений с постоянной композицией[4]. Однако установленная в [31, теорема 19] верхняя граница на функцию надежности строго больше ЭСУ во многих каналах. Это было показано численно в [13, с. 1594 и Приложение A]. Аналитическое подтверждение этого наблюдения приведено в Приложении A настоящей статьи. Проблематичным шагом в [31] было применение техники множителей Лагранжа (см. [13, сноска 8]). Доказательство теоремы 19 в [31] опирается на теорему 16 из [31], которая верна лишь для множителя Лагранжа s, соответствующего распределению на входе p, такому что $E_{\mathrm{sp}}(R, W) = E_{\mathrm{sp}}(R, W, p)$. Однако для произвольного распределения на входе p соответствующий множитель Лагранжа может быть, а может и не быть равным множителю для оптимального распределения на входе p. Именно по этой причине верхняя граница на функцию надежности, установленная в [31, теорема 19], не равна ЭСУ в общем случае вопреки тому, что затем утверждалось в [32, лемма 1] и [33, теорема 10.1.4]. Коротко говоря, доказательство теоремы 19 в [31] по умолчанию подразумевает минимаксное равенство, неверное в общем случае. Для стационарных каналов без памяти с конечными алфавитами на входе эту трудность можно преодолеть, используя рассуждения с постоянной композицией. Однако в этом случае доказательство, приведенное в [31], становится просто повторением доказательства из [10]. Недавно в [34] был предложен вывод ГСУ для стационарных каналов с одним ограничением по стоимости, использующий подход из [31]. Однако, как и в [31], доказательство в [34] подразумевает минимаксное равенство, неверное в общем случае. В частности, в [34, формула (26)] утверждается, что Q^n не зависит от $\boldsymbol{x}_m$. Чтобы утверждать это, следует добавить в [34, формулы (25) и (26)] дополнительный супремум по $\boldsymbol{x}_m$ в качестве самой внутренней оптимизации. Но с этим дополнительным супремумом станет неверным объяснение, приведенное в [34, с. 931]. С учетом сказанного в Приложении A не верится, что доказательство из [34] можно спасти без привлечения существенных новых идей, таких как техника выбрасывания с учетом композиции аналогично [10] или выбрасывания с учетом стоимости кодовых слов аналогично [18]. Коротко говоря, ни в [31], ни в [34] нет успешного доказательства ГСУ для стационарных каналов без памяти, даже для случая конечного множества входов.

В оставшейся части этого параграфа мы введем обозначения, опишем модель канала и определим коды со списочным декодированием. В § 2 вначале приведем краткие сведения о дивергенции Реньи, августиновских информационных величинах и ЭСУ, а затем выведем августиновский вариант границы Галлагера. В § 3 вначале сформулируем наши главные асимптотические результаты, т.е. ГСУ в теоремах 1 и 2, а затем выведем неасимптотические ГСУ, из которых они следуют. В § 4 мы выведем ЭСУ для отдельных гауссовских и пуассоновских каналов и подтвердим эквивалентность определения из § 2 и ранее использованных для этих каналов. В § 5 мы вкратце обсудим, почему метод Августина работает, и сравним наши результаты с результатами Августина из [24], а также обсудим применения метода Августина и августиновских информационных величин в родственных задачах.

1.1. Обозначения и соглашения. Скалярное произведение двух векторов μ и q в $\mathbb{R}^{\ell}$, т.е. $\sum_{i=1}^{\ell} \mu^i q^i$, обозначается через $\mu \cdot q$. Вектор из всех единиц размерности ℓ будем обозначать через $\mathbb{1}$ для любого $\ell \in \mathbb{Z}_+$, размерность ℓ будет ясна из контекста. Замыкание, внутреннюю область и выпуклую оболочку множества $\mathcal{S}$ будем

[4] В [31, с. 413] сказано: “Важной особенностью нижней границы, которую мы выведем, является то, что не делается никаких предположений о постоянной композиции кодовых слов, даже в качестве промежуточного шага”.

обозначать через $\operatorname{cl}\mathcal{S}$, $\operatorname{int}\mathcal{S}$ и $\operatorname{ch}\mathcal{S}$ соответственно; соответствующая топология или структура векторного пространства будет ясна из контекста.

Для произвольного множества $\mathcal{Y}$ обозначим через $\mathcal{P}(\mathcal{Y})$ множество всех вероятностных мер, не равных нулю лишь на конечном числе элементов из $\mathcal{Y}$. Для любой $p \in \mathcal{P}(\mathcal{Y})$ назовем множество всех $y \in \mathcal{Y}$, для которых $p(y) > 0$, носителем p и будем обозначать его через $\operatorname{supp} p$. Для любого измеримого пространства $(\mathcal{Y}, \mathscr{Y})$ обозначим множество всех вероятностных мер на нем через $\mathcal{P}(\mathscr{Y})$, а множество всех конечных мер – через $\mathcal{M}^+(\mathscr{Y})$. Интеграл от измеримой функции f по мере μ будем обозначать через $\int f\,\mu(dy)$ или $\int f(y)\,\mu(dy)$. Интегралы по мере Лебега на вещественной прямой обозначаем через $\int f\,dy$ или $\int f(y)\,dy$. Если μ – вероятностная мера, то интеграл от f по мере μ будем называть математическим ожиданием f и обозначать через $\mathbf{E}_\mu[f]$ или $\mathbf{E}_\mu[f(y)]$.

Некоторые символы будут использоваться в различных смыслах, однако конкретный смысл будет всегда ясен из контекста. Декартово произведение множеств [35, с. 38] будем обозначать знаком $\times$. Через $|\cdot|$ обозначаем абсолютную величину вещественных чисел и мощность множеств. Знаком $\leqslant$ обозначается как обычное отношение для вещественных чисел, так и соответствующее поточечное неравенство для функций и векторов. Для двух мер μ и q на измеримом пространстве $(\mathcal{Y}, \mathscr{Y})$ пишем $\mu \leqslant q$, если $\mu(\mathcal{E}) \leqslant q(\mathcal{E})$ для всех $\mathcal{E} \in \mathscr{Y}$. Произведение топологий [35, с. 38], σ-алгебр [35, с. 118] и мер [35, теорема 4.4.4] будем обозначать знаком $\otimes$. Используем краткие обозначения $\mathcal{X}_1^n$ для декартова произведения множеств $\mathcal{X}_1, \dots, \mathcal{X}_n$ и $\mathscr{Y}_1^n$ для произведения σ-алгебр $\mathscr{Y}_1, \dots, \mathscr{Y}_n$.

1.2. Модель канала. *Канал* W – это функция из *множества входов* $\mathcal{X}$ в множество всех вероятностных мер на *пространстве выходов* $(\mathcal{Y}, \mathscr{Y})$:

$$W\colon \mathcal{X} \to \mathcal{P}(\mathscr{Y}). \tag{1}$$

Множество $\mathcal{Y}$ называется *множеством выходов*, а $\mathscr{Y}$ – *σ-алгеброй выходных событий*. Множество всех каналов из множества входов $\mathcal{X}$ в пространство выходов $(\mathcal{Y}, \mathscr{Y})$ обозначим через $\mathcal{P}(\mathscr{Y}\,|\,\mathcal{X})$. Для любых $p \in \mathcal{P}(\mathcal{X})$ и $W \in \mathcal{P}(\mathscr{Y}\,|\,\mathcal{X})$ вероятностная мера с маргинальным распределением на $\mathcal{X}$, равным p, и условным распределением при условии x, равным $W(x)$, обозначается через $p \circledast W$. Структуры, описанной в (1), самой по себе в общем случае недостаточно для существования и единственности меры $p \circledast W$ с требуемыми свойствами для всех $p \in \mathcal{P}(\mathscr{X})$. Существование и единственность такой меры $p \circledast W$ гарантировано для всех $p \in \mathcal{P}(\mathscr{X})$, если W – функция вероятностей переходов из $(\mathcal{X}, \mathscr{X})$ в $(\mathcal{Y}, \mathscr{Y})$, т.е. элемент $\mathcal{P}(\mathscr{Y}\,|\,\mathscr{X})$, а не $\mathcal{P}(\mathscr{Y}\,|\,\mathcal{X})$.

Канал W называется *дискретным каналом*, если как $\mathcal{X}$, так и $\mathscr{Y}$ – конечные множества. Для любого $n \in \mathbb{Z}_+$ и любых каналов $W_t\colon \mathcal{X}_t \to \mathcal{P}(\mathscr{Y}_t)$, $t \in \{1, \dots, n\}$, *канал-произведение* $W_{[1,n]}\colon \mathcal{X}_1^n \to \mathcal{P}(\mathscr{Y}_1^n)$ *длины* n определяется следующим образом:

$$W_{[1,n]}(x_1^n) = \bigotimes_{t=1}^{n} W_t(x_t), \quad \forall x_1^n \in \mathcal{X}_1^n.$$

Канал $U\colon \mathcal{Z} \to \mathcal{P}(\mathscr{Y}_1^n)$ называется *каналом длины n без памяти*, если существует канал-произведение $W_{[1,n]}$, для которого $U(z) = W_{[1,n]}(z)$ для всех $z \in \mathcal{Z}$ и при этом $\mathcal{Z} \subset \mathcal{X}_1^n$. Канал-произведение является *стационарным*, если $W_t = W$ при всех $t \in \{1, \dots, n\}$ для некоторого $W\colon \mathcal{X} \to \mathcal{P}(\mathscr{Y})$. Для такого канала обозначим через $\varUpsilon(x)$, где $\varUpsilon(x) \in \mathcal{P}(\mathcal{X})$, композицию (т.е. эмпирическое распределение, тип) каждого из $x_1^n \in \mathcal{X}_1^n$.

Для любого $\ell \in \mathbb{Z}_+$ ℓ-мерная *функция стоимости* ρ – это функция из множества входов в $\mathbb{R}^\ell$, ограниченная снизу, т.е. функция вида $\rho\colon \mathcal{X} \to \mathbb{R}^\ell_{\geqslant z}$ для некоторого

$z \in \mathbb{R}$. Без ограничения общности будем предполагать[5], что

$$\inf_{x\in\mathcal{X}} \rho^i(x) \geqslant 0, \quad \forall i \in \{1, \ldots, \ell\}.$$

Множество всех ограничений по стоимости, которые выполняются хотя бы для одного элемента из $\mathcal{X}$, обозначим через $\Gamma_\rho^{\mathrm{ex}}$, а множество всех ограничений по стоимости, выполненных хотя бы для одного элемента из $\mathcal{P}(\mathcal{X})$, – через Γ_ρ:

$$\Gamma_\rho^{\mathrm{ex}} \triangleq \left\{\varrho \in \mathbb{R}_{\geqslant 0}^{\ell} : \exists x \in \mathcal{X}, \text{ такой что } \rho(x) \leqslant \varrho\right\},$$
$$\Gamma_\rho \triangleq \left\{\varrho \in \mathbb{R}_{\geqslant 0}^{\ell} : \exists p \in \mathcal{P}(\mathcal{X}), \text{ такой что } \sum_x p(x)\rho(x) \leqslant \varrho\right\}.$$

Тогда оба множества $\Gamma_\rho^{\mathrm{ex}}$ и Γ_ρ имеют непустые внутренние области, причем Γ_ρ является выпуклой оболочкой $\Gamma_\rho^{\mathrm{ex}}$, т.е. $\Gamma_\rho = \operatorname{ch} \Gamma_\rho^{\mathrm{ex}}$.

Функция стоимости на канале-произведении называется аддитивной, если ее можно представить в виде суммы функций стоимости на каналах-компонентах. Для заданных $W_t \colon \mathcal{X}_t \to \mathcal{P}(\mathcal{Y}_t)$ и $\rho_t \colon \mathcal{X}_t \to \mathbb{R}_{\geqslant 0}^{\ell}$, $t \in \{1, \ldots, n\}$, обозначим результирующую аддитивную функцию стоимости на $\mathcal{X}_1^n$ для канала $W_{[1,n]}$ через $\rho_{[1,n]}$, т.е.

$$\rho_{[1,n]}(x_1^n) = \sum_{t=1}^n \rho_t(x_t), \quad \forall x_1^n \in \mathcal{X}_1^n.$$

1.3. Коды со списочным декодированием. Пара (Ψ, Θ) называется (M, L)-*кодом* в канале $W \colon \mathcal{X} \to \mathcal{P}(\mathcal{Y})$, если

- *функция кодирования* Ψ – функция из множества сообщений $\mathcal{M} \triangleq \{1, 2, \ldots, M\}$ в множество входов $\mathcal{X}$;
- *функция декодирования* Θ – измеримая функция из пространства выходов $(\mathcal{Y}, \mathcal{Y})$ в множество $\widehat{\mathcal{M}} \triangleq \{\mathcal{L} : \mathcal{L} \subset \mathcal{M} \text{ и } |\mathcal{L}| = L\}$.

Для заданного (M, L)-кода (Ψ, Θ) в канале $W \colon \mathcal{X} \to \mathcal{P}(\mathcal{Y})$ *условная вероятность ошибки* P_{e}^m для $m \in \mathcal{M}$ и *средняя вероятность ошибки* P_{e} определяются как

$$P_{\mathrm{e}}^m \triangleq \mathbf{E}_{W(\Psi(m))}[\mathbb{1}_{\{m \notin \Theta(y)\}}],$$
$$P_{\mathrm{e}} \triangleq \frac{1}{M} \sum_{m\in\mathcal{M}} P_{\mathrm{e}}^m.$$

Будем говорить, что функция кодирования Ψ, и следовательно, соответствующий код, удовлетворяют ограничению по стоимости ϱ, если $\bigvee_{m\in\mathcal{M}} \rho(\Psi(m)) \leqslant \varrho$. Будем говорить, что функция кодирования Ψ, и следовательно, соответствующий код, на стационарном канале-произведении удовлетворяют ограничению на эмпирическое распределение $\mathcal{A} \subset \mathcal{P}(\mathcal{X})$, если композиции всех кодовых слов принадлежат $\mathcal{A}$, т.е. если $\Upsilon(\Psi(m)) \in \mathcal{A}$ для всех $m \in \mathcal{M}$.

§ 2. Предварительные сведения

Дивергенция Реньи, скошенные вероятностные меры и каналы, а также августиновские информационные величины являются ключевыми понятиями в нашем последующем анализе. Эти понятия вводятся в пп. 2.1 и 2.2, а более подробное обсуждение можно найти в [26,36]. В п. 2.3 мы определяем ЭСУ и выводим ее хорошо

[5] Августин в [24, § 33] накладывал дополнительное условие $\bigvee_{x\in\mathcal{X}} \rho(x) \leqslant \mathbb{1}$, однако это условие исключает из рассмотрения некоторые важные случаи, такие как гауссовские каналы.

известные свойства для нашей общей модели канала, а в п. 2.4 выводится августиновский вариант границы Галлагера.

2.1. Дивергенции Реньи и скошенные вероятностные меры и каналы.

Определение 1. Для любых $\alpha \in \mathbb{R}_+$ и $w, q \in \mathcal{M}^+(\mathcal{Y})$ *дивергенцией Реньи порядка α между w и q* называется величина

$$D_\alpha(w \,\|\, q) \triangleq \begin{cases} \dfrac{1}{\alpha-1} \ln \displaystyle\int \left(\frac{dw}{d\nu}\right)^{\alpha} \left(\frac{dq}{d\nu}\right)^{1-\alpha} \nu(dy), & \alpha \neq 1, \\ \displaystyle\int \frac{dw}{d\nu} \left[\ln \frac{dw}{d\nu} - \ln \frac{dq}{d\nu}\right] \nu(dy), & \alpha = 1, \end{cases}$$

где ν – любая мера, такая что $w \prec \nu$ и $q \prec \nu$.

За исчерпывающим обзором сведений о дивергенции Реньи, используемых на протяжении всей статьи, отсылаем читателя к [36]. Заметим, что дивергенция Реньи порядка 1 – это дивергенция Кульбака – Лейблера. Дивергенцию Реньи для других порядков также можно охарактеризовать через дивергенцию Кульбака – Лейблера (см. [36, теорема 30]). Эта характеризация связана с еще одним ключевым для нас понятием, а именно понятием скошенной вероятностной меры.

Определение 2. Для любого $\alpha \in \mathbb{R}_+$ и любых вероятностных мер $w, q \in \mathcal{P}(\mathcal{Y})$, таких что $D_\alpha(w \,\|\, q) < \infty$, *скошенной вероятностной мерой w_α^q порядка α* называется

$$\frac{dw_\alpha^q}{d\nu} \triangleq e^{(1-\alpha)D_\alpha(w \| q)} \left(\frac{dw}{d\nu}\right)^{\alpha} \left(\frac{dq}{d\nu}\right)^{1-\alpha}. \tag{2}$$

Условная дивергенция Реньи и скошенный канал – прямые обобщения понятий дивергенции Реньи и скошенной вероятностной меры, которые позволят нам кратко записывать некоторые соотношения на протяжении дальнейшего анализа.

Определение 3. Для любых $\alpha \in \mathbb{R}_+$, $W\colon \mathcal{X} \to \mathcal{P}(\mathcal{Y})$, $Q\colon \mathcal{X} \to \mathcal{P}(\mathcal{Y})$ и $p \in \mathcal{P}(\mathcal{X})$ *условной дивергенцией Реньи порядка α для распределения на входе p* называется

$$D_\alpha(W \,\|\, Q \,|\, p) \triangleq \sum_{x \in \mathcal{X}} p(x) D_\alpha(W(x) \,\|\, Q(x)).$$

Если существует $q \in \mathcal{P}(\mathcal{Y})$, такая что $Q(x) = q$ для всех $x \in \mathcal{X}$, то будем обозначать $D_\alpha(W \,\|\, Q \,|\, p)$ через $D_\alpha(W \,\|\, q \,|\, p)$.

Определение 4. Для любых $\alpha \in \mathbb{R}_+$, $W\colon \mathcal{X} \to \mathcal{P}(\mathcal{Y})$ и $Q\colon \mathcal{X} \to \mathcal{P}(\mathcal{Y})$ *скошенным каналом W_α^Q порядка α* называется функция из $\{x : D_\alpha(W(x) \,\|\, Q(x)) < \infty\}$ в $\mathcal{P}(\mathcal{Y})$, такая что

$$\frac{dW_\alpha^Q(x)}{d\nu} \triangleq e^{(1-\alpha)D_\alpha(W(x) \| Q(x))} \left(\frac{dW(x)}{d\nu}\right)^{\alpha} \left(\frac{dQ(x)}{d\nu}\right)^{1-\alpha}. \tag{3}$$

Если существует $q \in \mathcal{P}(\mathcal{Y})$, такая что $Q(x) = q$ для всех $x \in \mathcal{X}$, будем обозначать W_α^Q через W_α^q.

Следующий оператор $\mathrm{T}_{\alpha,p}(\cdot)$ неявно рассматривался Фано в [9, гл. 9], Арутюняном в [10] и Полтыревым в [30], а в явном виде Августином в [24, § 34], но во всех четырех источниках лишь для порядков, меньших единицы.

Определение 5. Для любых $\alpha \in \mathbb{R}_+$, $W\colon \mathcal{X} \to \mathcal{P}(\mathcal{Y})$ и $p \in \mathcal{P}(\mathcal{X})$ оператор Августина $\mathrm{T}_{\alpha,p}(\cdot)\colon \mathcal{Q}_{\alpha,p} \to \mathcal{P}(\mathcal{Y})$ порядка α для распределения на входе p определяется как

$$\mathrm{T}_{\alpha,p}(q) \triangleq \sum_x p(x) W_\alpha^q(x), \quad \forall q \in \mathcal{Q}_{\alpha,p}, \tag{4}$$

где $\mathcal{Q}_{\alpha,p} \triangleq \{q \in \mathcal{P}(\mathcal{Y}) :\ D_\alpha(W \,\|\, q \,|\, p) < \infty\}$, а скошенный канал W_α^q определен в (3).

2.2. Августиновские информационные величины.

О п р е д е л е н и е 6. Для любых $\alpha \in \mathbb{R}_+$, $W\colon \mathcal{X} \to \mathcal{P}(\mathcal{Y})$ и $p \in \mathcal{P}(\mathcal{X})$ *информацией Августина порядка α для распределения на входе p* называется величина

$$I_\alpha(p; W) \triangleq \inf_{q \in \mathcal{P}(\mathcal{Y})} D_\alpha(W \,\|\, q \,|\, p). \tag{5}$$

Инфимум в (5) достигается на единственной вероятностной мере, которая обозначается через $q_{\alpha,p}$ и называется *августиновским средним порядка α для распределения на входе p*. При этом для августиновского среднего порядка α справедливы следующие тождества:

$$\begin{aligned} &D_{1\vee\alpha}(q_{\alpha,p} \,\|\, q) \geqslant D_\alpha(W \,\|\, q \,|\, p) - I_\alpha(p; W) \geqslant D_{1\wedge\alpha}(q_{\alpha,p} \,\|\, q), \\ &\forall q \in \mathcal{P}(\mathcal{Y}),\ \alpha \in \mathbb{R}_+, \end{aligned} \tag{6}$$

$$\mathrm{T}_{\alpha,p}(q_{\alpha,p}) = q_{\alpha,p}, \quad \forall \alpha \in \mathbb{R}_+. \tag{7}$$

Эти свойства установлены в [26, лемма 13(b)–(d)]; ранее они были получены Августином [24, лемма 34.2] для порядков, меньших единицы. На протяжении всей статьи мы будем ссылаться на утверждения об августиновских информационных величинах из [26]. Более подробный обзор предшествующих результатов об августиновских информационных величинах можно также найти в [26].

О п р е д е л е н и е 7. Для любых $\alpha \in \mathbb{R}_+$, $W\colon \mathcal{X} \to \mathcal{P}(\mathcal{Y})$ и $\mathcal{A} \subset \mathcal{P}(\mathcal{X})$ *августиновской пропускной способностью порядка α канала W для множества ограничений $\mathcal{A}$* называется величина

$$C_{\alpha,W,\mathcal{A}} \triangleq \sup_{p \in \mathcal{A}} I_\alpha(p; W).$$

Когда множество ограничений $\mathcal{A}$ совпадает со всем $\mathcal{P}(\mathcal{X})$, августиновская пропускная способность порядка α обозначается через $C_{\alpha,W}$, т.е. $C_{\alpha,W} \triangleq C_{\alpha,W,\mathcal{P}(\mathcal{X})}$.

Используя определения информации Августина и августиновской пропускной способности, получаем следующее выражение для $C_{\alpha,W,\mathcal{A}}$:

$$C_{\alpha,W,\mathcal{A}} = \sup_{p \in \mathcal{A}} \inf_{q \in \mathcal{P}(\mathcal{Y})} D_\alpha(W \,\|\, q \,|\, p).$$

Если множество $\mathcal{A}$ выпукло, то согласно [26, теорема 1] порядок супремума и инфимума можно поменять:

$$\sup_{p \in \mathcal{A}} \inf_{q \in \mathcal{P}(\mathcal{Y})} D_\alpha(W \,\|\, q \,|\, p) = \inf_{q \in \mathcal{P}(\mathcal{Y})} \sup_{p \in \mathcal{A}} D_\alpha(W \,\|\, q \,|\, p). \tag{8}$$

Если к тому же $C_{\alpha,W,\mathcal{A}}$ конечна, то из [26, теорема 1] следует, что существует единственная вероятностная мер $q_{\alpha,W,\mathcal{A}}$, называемая *августиновским центром порядка α канала W для множества ограничений $\mathcal{A}$*, такая что

$$C_{\alpha,W,\mathcal{A}} = \sup_{p \in \mathcal{A}} D_\alpha(W \,\|\, q_{\alpha,W,\mathcal{A}} \,|\, p).$$

Множество всех вероятностных мер, удовлетворяющих ограничению по стоимости ϱ, будем обозначать через $\mathcal{A}(\varrho)$, т.е.

$$\mathcal{A}(\varrho) \triangleq \{p \in \mathcal{P}(\mathcal{X}) : \mathbf{E}_p[\rho] \leqslant \varrho\}.$$

Допуская некоторую вольность в обозначениях, множества ограничений, определяемые через ограничения по стоимости, будем обозначать через $C_{\alpha,W,\varrho}$ вместо $C_{\alpha,W,\mathcal{A}(\varrho)}$. Чтобы иметь возможность применять технику выпуклого сопряжения без существенных видоизменений, продолжим определение августиновской пропускной способности на недопустимые значения ограничений по стоимости, т.е. на ϱ, не лежащие в Γ_ρ:

$$C_{\alpha,W,\varrho} \triangleq \begin{cases} \sup\limits_{p\in\mathcal{A}(\varrho)} I_\alpha(p;W), & \text{если } \varrho \in \Gamma_\rho, \\ -\infty, & \text{если } \varrho \in \mathbb{R}^{\ell}_{\geqslant 0} \setminus \Gamma_\rho, \end{cases} \qquad \forall \alpha \in \mathbb{R}_+.$$

Чтобы получить характеризацию $C_{\alpha,W,\varrho}$ с помощью техники выпуклого сопряжения, вначале определим понятия информации и пропускной способности Августина – Лежандра. Эти понятия были впервые введены в [1, раздел III-A] и [26, п. 5.2] как обобщения аналогичных понятий из [12, гл. 8].

Определение 8. Для любых $\alpha \in \mathbb{R}_+$, канала W вида $W : \mathcal{X} \to \mathcal{P}(\mathcal{Y})$ с функцией стоимости $\rho : \mathcal{X} \to \mathbb{R}^{\ell}_{\geqslant 0}$, $p \in \mathcal{P}(\mathcal{X})$ и $\lambda \in \mathbb{R}^{\ell}_{\geqslant 0}$ *информацией Августина – Лежандра порядка α для распределения на входе p и множителя Лагранжа λ* называется величина

$$I^{\lambda}_{\alpha}(p;W) \triangleq I_\alpha(p;W) - \lambda \cdot \mathbf{E}_p[\rho].$$

Определение 9. Для любых $\alpha \in \mathbb{R}_+$, канала W вида $W : \mathcal{X} \to \mathcal{P}(\mathcal{Y})$ с функцией стоимости $\rho : \mathcal{X} \to \mathbb{R}^{\ell}_{\geqslant 0}$ и $\lambda \in \mathbb{R}^{\ell}_{\geqslant 0}$ *пропускной способностью Августина – Лежандра порядка α для множителя Лагранжа λ* называется величина

$$C^{\lambda}_{\alpha,W} \triangleq \sup_{p\in\mathcal{P}(\mathcal{X})} I^{\lambda}_{\alpha}(p;W).$$

С точностью до некоторых перемен знака функция $C^{\lambda}_{\alpha,W}$ является выпукло сопряженной к $C_{\alpha,W,\varrho}$ в силу аналогичного соотношения между $I^{\lambda}_{\alpha}(p;W)$ и $I_\alpha(p;W)$ (см. [26, формулы (72)–(74) и (76)]):

$$C^{\lambda}_{\alpha,W} = \sup_{\varrho\geqslant 0} C_{\alpha,W,\varrho} - \lambda \cdot \varrho, \quad \forall \lambda \in \mathbb{R}^{\ell}_{\geqslant 0}.$$

Тогда $C_{\alpha,W,\varrho}$ можно выразить через $C^{\lambda}_{\alpha,W}$, по крайней мере, для внутренних точек Γ_ρ:

$$C_{\alpha,W,\varrho} = \inf_{\lambda\geqslant 0} C^{\lambda}_{\alpha,W} + \lambda \cdot \varrho.$$

Кроме того, согласно [26, лемма 29] существует непустое выпуклое компактное множество величин $\lambda_{\alpha,W,\varrho}$, таких что $C_{\alpha,W,\varrho} = C^{\lambda_{\alpha,W,\varrho}}_{\alpha,W} + \lambda_{\alpha,W,\varrho} \cdot \varrho$, при условии, что $C_{\alpha,W,\varrho}$ конечна.

С другой стороны, используя определения величин $I_\alpha(p;W)$, $I^{\lambda}_{\alpha}(p;W)$ и $C^{\lambda}_{\alpha,W}$, получаем следующее выражение для $C^{\lambda}_{\alpha,W}$:

$$C^{\lambda}_{\alpha,W} = \sup_{p\in\mathcal{P}(\mathcal{X})} \inf_{q\in\mathcal{P}(\mathcal{Y})} D_\alpha(W \,\|\, q \,|\, p) - \lambda \cdot \mathbf{E}_p[\rho].$$

Для функции $C_{\alpha,W}^{\lambda}$ выполняется минимаксное соотношение, аналогичное (8) (см. [26, теорема 2]). Впрочем, это соотношение лучше объяснить с использованием понятия радиуса Августина – Лежандра, определяемого следующим образом.

Определение 10. Для любых $\alpha \in \mathbb{R}_+$, канала $W : \mathcal{X} \to \mathcal{P}(\mathcal{Y})$ с функцией стоимости $\rho \colon \mathcal{X} \to \mathbb{R}_{\geqslant 0}^{\ell}$ и $\lambda \in \mathbb{R}_{\geqslant 0}^{\ell}$ *радиусом Августина – Лежандра порядка* α *канала* W *для множителя Лагранжа* λ называется величина

$$S_{\alpha,W}^{\lambda} \triangleq \inf_{q \in \mathcal{P}(\mathcal{Y})} \sup_{x \in \mathcal{X}} D_{\alpha}(W(x) \,\|\, q) - \lambda \cdot \rho(x).$$

Тогда согласно [26, теорема 2] для любых $\alpha \in \mathbb{R}_+$, $W \colon \mathcal{X} \to \mathcal{P}(\mathcal{Y})$ с $\rho \colon \mathcal{X} \to \mathbb{R}_{\geqslant 0}^{\ell}$ и $\lambda \in \mathbb{R}_{\geqslant 0}^{\ell}$ справедливо равенство

$$C_{\alpha,W}^{\lambda} = S_{\alpha,W}^{\lambda}. \tag{9}$$

Если к тому же $C_{\alpha,W}^{\lambda}$ конечна, то существует единственная вероятностная мера $q_{\alpha,W}^{\lambda} \in \mathcal{P}(\mathcal{Y})$, называемая *центром Августина – Лежандра порядка* α *канала* W *для множителя Лагранжа* λ, такая что

$$C_{\alpha,W}^{\lambda} = \sup_{x \in \mathcal{X}} D_{\alpha}(W(x) \,\|\, q_{\alpha,W}^{\lambda}) - \lambda \cdot \rho(x).$$

Информационные величины Августина – Лежандра определяются стандартным применением техники выпуклого сопряжения. Однако начиная с [29, теоремы 8 и 10], т.е. с вариантов границы Галлагера с ограничениями по стоимости, обычным инструментом применения техники выпуклого сопряжения для вычисления экспоненты ошибки вместо информационных величин Августина – Лежандра становятся информационные величины Реньи – Галлагера (см., например, [18–20]). Краткое обсуждение информационных величин Реньи – Галлагера содержится в Приложении В, более подробно см. в [26].

2.3. Экспонента сферической упаковки.

Определение 11. Для любых $W : \mathcal{X} \to \mathcal{P}(\mathcal{Y})$, $\mathcal{A} \subset \mathcal{P}(\mathcal{X})$ и $R \in \mathbb{R}_{\geqslant 0}$ экспонентой сферической упаковки называется величина

$$E_{\mathrm{sp}}(R, W, \mathcal{A}) \triangleq \sup_{\alpha \in (0,1)} \frac{1-\alpha}{\alpha} \left(C_{\alpha,W,\mathcal{A}} - R\right). \tag{10}$$

ЭСУ для случая $\mathcal{A} = \mathcal{P}(\mathcal{X})$ будем обозначать через $E_{\mathrm{sp}}(R, W)$. Кроме того, допуская некоторую вольность в обозначениях, ЭСУ в случае $\mathcal{A} = \{p\}$ будем обозначать через $E_{\mathrm{sp}}(R, W, p)$, а в случае $\mathcal{A} = \{p : \mathbf{E}_p[\rho] \leqslant \varrho\}$ – через $E_{\mathrm{sp}}(R, W, \varrho)$.

Лемма 1. *Для любых* $W \colon \mathcal{X} \to \mathcal{P}(\mathcal{Y})$ *и* $\mathcal{A} \subset \mathcal{P}(\mathcal{X})$ *функция* $E_{\mathrm{sp}}(R, W, \mathcal{A})$ *не возрастает и выпукла по* R *на* $\mathbb{R}_{\geqslant 0}$, *конечна на* $(C_{0^+,W,\mathcal{A}}, \infty)$ *и непрерывна на* $[C_{0^+,W,\mathcal{A}}, \infty)$, *где* $C_{0^+,W,\mathcal{A}} = \lim_{\alpha \downarrow 0} C_{\alpha,W,\mathcal{A}}$. *В частности,*

$$E_{\mathrm{sp}}(R, W, \mathcal{A}) = \\ = \begin{cases} \infty, & R < C_{0^+,W,\mathcal{A}}, \\ \sup\limits_{\alpha \in (0,1)} \dfrac{1-\alpha}{\alpha} \left(C_{\alpha,W,\mathcal{A}} - R\right), & R = C_{0^+,W,\mathcal{A}}, \\ \sup\limits_{\alpha \in [\phi,1)} \dfrac{1-\alpha}{\alpha} \left(C_{\alpha,W,\mathcal{A}} - R\right), & R = C_{\phi,W,\mathcal{A}} \text{ для некоторого } \phi \in (0,1), \\ 0, & R \geqslant C_{1,W,\mathcal{A}}. \end{cases} \tag{11}$$

Лемма 1 вытекает из свойств непрерывности и монотонности функции $C_{\alpha,W,\mathcal{A}}$, установленных в [26, лемма 23]; ее доказательство приведено в Приложении С. Доказательство леммы 1 аналогично доказательству утверждения [25, лемма 13], которое опирается на [25, лемма 8], а не на [26, лемма 23].

Можно выразить $E_{\text{sp}}(R, W, \mathcal{A})$ через $E_{\text{sp}}(R, W, p)$, используя определения величин $C_{\alpha,W,\mathcal{A}}$, $E_{\text{sp}}(R, W, \mathcal{A})$ и $E_{\text{sp}}(R, W, p)$:

$$E_{\text{sp}}(R, W, \mathcal{A}) = \sup_{\alpha \in (0,1)} \sup_{p \in \mathcal{A}} \frac{1-\alpha}{\alpha} \left(I_\alpha(p; W) - R\right) =$$

$$= \sup_{p \in \mathcal{A}} \sup_{\alpha \in (0,1)} \frac{1-\alpha}{\alpha} \left(I_\alpha(p; W) - R\right) = \sup_{p \in \mathcal{A}} E_{\text{sp}}(R, W, p). \tag{12}$$

Утверждение леммы 1 для $E_{\text{sp}}(R, W, p)$ выполнено по определению, но его можно значительно усилить для значений R в интервале $(\lim_{\alpha \downarrow 0} I_\alpha(p; W), I_1(p; W)]$, используя простейшие свойства информации Августина.

Лемма 2. *Пусть $W : \mathcal{X} \to \mathcal{P}(\mathcal{Y})$ и $p \in \mathcal{P}(\mathcal{X})$ таковы, что $I_{0+}(p; W) \neq I_1(p; W)$, где $I_{0+}(p; W) \triangleq \lim_{\alpha \downarrow 0} I_\alpha(p; W)$. Тогда для любой скорости $R \in (I_{0+}(p; W), I_1(p; W)]$ существует единственный порядок $\alpha^* \in (0, 1]$, такой что*

$$R = I_1(p; W_{\alpha^*}^{q_{\alpha^*,p}}). \tag{13}$$

Порядки α^*, определенные в (13), образуют возрастающую непрерывную биективную функцию скорости R из $(I_{0+}(p; W), I_1(p; W)]$ в $(0, 1]$, такую что

$$E_{\text{sp}}(R, W, p) = D_1(W_{\alpha^*}^{q_{\alpha^*,p}} \,\|\, W \,|\, p), \tag{14}$$

$$\frac{\partial}{\partial R} E_{\text{sp}}(R, W, p) = \frac{\alpha^* - 1}{\alpha^*}. \tag{15}$$

Таким образом, $E_{\text{sp}}(R, W, p)$ конечна, выпукла, непрерывно дифференцируема, возрастает по R на $(I_{0+}(p; W), I_1(p; W))$ и удовлетворяет соотношению

$$E_{\text{sp}}(I_{0+}(p; W), W, p) = \lim_{\alpha \downarrow 0} D_1(W_\alpha^{q_{\alpha,p}} \,\|\, W \,|\, p). \tag{16}$$

Кроме того, если $E_{\text{sp}}(I_{0+}(p; W), W, p)$ конечна, то существует канал $V : \mathcal{X} \to \mathcal{P}(\mathcal{Y})$, для которого $I_1(p; V) = I_{0+}(p; W)$ и $D_1(V \,\|\, W \,|\, p) = E_{\text{sp}}(I_{0+}(p; W), W, p)$.

Доказательство. Заметим, что $I_1(p; W_\alpha^{q_{\alpha,p}})$ – возрастающая и непрерывная функция порядка α согласно [26, лемма 17(a),(f)], поскольку $I_{0+}(p; W) \neq I_1(p; W)$ по условию. Кроме того, $\lim_{\alpha \downarrow 0} I_1(p; W_\alpha^{q_{\alpha,p}}) = I_{0+}(p; W)$ в силу [26, лемма 17(g)]. Тогда по теореме о промежуточном значении [37, теорема 4.23] существует единственный порядок α^*, удовлетворяющий условию (13). Функция, определенная в (13), – возрастающая непрерывная биективная функция скорости R из $(I_{0+}(p; W), I_1(p; W)]$ в $(0, 1]$ как обратная к возрастающей непрерывной биективной функции $\alpha \rightsquigarrow I_1(p; W_\alpha^{q_{\alpha,p}})$ из $(0, 1]$ в $(I_{0+}(p; W), I_1(p; W)]$.

С другой стороны, $I_\alpha(p; W)$ непрерывно дифференцируема по α в силу [26, лемма 17(e)]; тогда из [26, формулы (35) и (46)] получаем

$$\frac{\partial}{\partial \alpha} \frac{1-\alpha}{\alpha} \left(I_\alpha(p; W) - R\right) = \frac{1}{\alpha^2} \left(R - I_1(p; W_\alpha^{q_{\alpha,p}})\right). \tag{17}$$

Следовательно, для любой $R \in (I_{0+}(p; W), I_1(p; W)]$ супремум в определении величины $E_{\text{sp}}(R, W, p)$ достигается на порядке α^*, удовлетворяющем (13). Тогда (14) вытекает из [26, формула (35)]. Кроме того, для любых $R \in (I_{0+}(p; W), I_1(p; W))$

и $\overline{R} \geqslant 0$ имеем

$$\begin{aligned} E_{\mathrm{sp}}(\overline{R}, W, p) &\geqslant \frac{1-\alpha^*(R)}{\alpha^*(R)}(I_{\alpha^*}(p;W) - \overline{R}) = \\ &= E_{\mathrm{sp}}(R, W, p) + \frac{1-\alpha^*(R)}{\alpha^*(R)}(R - \overline{R}). \end{aligned} \tag{18}$$

Рассуждая аналогично и меняя ролями R и $\overline{R}$, получаем, что для любых $\overline{R} \in \in (I_{0^+}(p;W), I_1(p;W))$ и $R \geqslant 0$

$$E_{\mathrm{sp}}(R, W, p) \geqslant E_{\mathrm{sp}}(\overline{R}, W, p) + \frac{1-\alpha^*(\overline{R})}{\alpha^*(\overline{R})}(\overline{R} - R). \tag{19}$$

Так как α^* возрастает и непрерывна по скорости, из (18) и (19) получаем (15) для всех $R \in (I_{0^+}(p;W), I_1(p;W)]$.

В случае $R = I_{0^+}(p;W)$ заметим, что $\dfrac{1-\alpha}{\alpha}(I_\alpha(p;W) - R)$ убывает по α на $(0,1)$ в силу (17) и [26, лемма 17(f),(g)]. Таким образом,

$$E_{\mathrm{sp}}(I_{0^+}(p;W), W, p) = \lim_{\alpha \downarrow 0} \frac{1-\alpha}{\alpha}(I_\alpha(p;W) - I_{0^+}(p;W)).$$

Тогда (16) вытекает из теоремы о среднем [37, теорема 5.10] и [26, формула (46)]. Кроме того, если $E_{\mathrm{sp}}(I_{0^+}(p;W), W, p) = \gamma$ для $\gamma \in \mathbb{R}_+$, то $D_1(W_\alpha^{q_{\alpha,p}}(x) \,\|\, W(x)) \leqslant \dfrac{\gamma}{p(x)}$ в силу неотрицательности дивергенции Реньи. Отсюда

$$\int G\left(\frac{dW_\alpha^{q_{\alpha,p}}(x)}{dW(x)}\right) W(dy \,|\, x) \leqslant \frac{\gamma}{p(x)} + \frac{1}{e} + 1$$

для $G(\tau) = \tau \mathbb{1}_{\{0 \leqslant \tau < e\}} + \tau \ln \tau \mathbb{1}_{\{\tau \geqslant e\}}$, поскольку $\tau \ln \tau \geqslant -1/e$. Тогда элементы семейства $\left\{\dfrac{dW_\alpha^{q_{\alpha,p}}(x)}{dW(x)}\right\}_{\alpha \in (0,1)}$ равномерно $W(x)$-интегрируемы в силу необходимого и достаточного условия равномерной интегрируемости Валле–Пуссена [38, теорема 4.5.9]. Таким образом, любая последовательность элементов $\left\{W_\alpha^{q_{\alpha,p}}(x)\right\}_{\alpha \in (0,1)}$ имеет сходящуюся подпоследовательность в топологии сходимости на множествах [38, теорема 4.7.25]. Для каждого $x \in \operatorname{supp} p$ пусть $V(x)$ – предельная точка вышеуказанной подпоследовательности последовательности $\left\{W_{1/\kappa}^{q_{1/\kappa,p}}(x)\right\}_{\kappa \in \mathbb{Z}_+}$. Тогда из соотношений (13), (14) и полунепрерывности снизу дивергенции Реньи по ее аргументам в топологии сходимости на множествах [36, теорема 15] следует, что $I_1(p;V) \leqslant I_{0^+}(p;W)$ и $D_1(V \,\|\, W \,|\, p) \leqslant \gamma$. С другой стороны в силу определения $E_{\mathrm{sp}}(R, W, p)$ и [26, лемма 13(e)] имеем

$$\begin{aligned} E_{\mathrm{sp}}(R, W, p) &= \sup_{\alpha \in (0,1)} \inf_{V \in \mathcal{P}(\mathcal{Y}|\mathcal{X})} D_1(V \,\|\, W \,|\, p) + \frac{1-\alpha}{\alpha}(I_1(p;V) - R) \leqslant \\ &\leqslant \sup_{\alpha \in (0,1)} D_1(V \,\|\, W \,|\, p) + \frac{1-\alpha}{\alpha}(I_1(p;V) - R) = \begin{cases} D_1(V \,\|\, W \,|\, p), & R \geqslant I_1(p;V), \\ \infty, & R < I_1(p;V). \end{cases} \end{aligned}$$

Таким образом, $I_1(p;V)$ не может быть меньше, чем $I_{0^+}(p;W)$, так как $E_{\mathrm{sp}}(R, W, p)$ равна бесконечности при всех $R < I_{0^+}(p;W)$. Значит, $I_1(p;V) = I_{0^+}(p;W)$. Следовательно, $D_1(V \,\|\, W \,|\, p)$ не может быть меньше, чем γ, поскольку $E_{\mathrm{sp}}(I_{0^+}(p;W), W, p) = = \gamma$. Отсюда $D_1(V \,\|\, W \,|\, p) = \gamma$. ▲

Лемма 2 дает простое объяснение альтернативному выражению для $E_{\mathrm{sp}}(R, W, p)$, известному как ЭСУ в форме Арутюняна [10].

Л е м м а 3. *Для любых* $W\colon \mathcal{X} \to \mathcal{P}(\mathcal{Y})$, $p \in \mathcal{P}(\mathcal{X})$ *и* $R \in \mathbb{R}_+$

$$E_{\text{sp}}(R, W, p) = \inf_{V\colon I_1(p;V) \leqslant R} D_1(V \,\|\, W \,|\, p). \tag{20}$$

Д о к а з а т е л ь с т в о. Если $R \in [I_1(p;W), \infty)$, то равенство (20) выполнено, так как выражение в правой части (20) равно нулю, что получается подстановкой $V = W$ с учетом неотрицательности дивергенции Реньи.

С другой стороны, в силу определения $E_{\text{sp}}(R, W, p)$ (см. [26, лемма 13(e)]) и неравенства максимина-минимакса имеем

$$\begin{aligned}
&E_{\text{sp}}(R, W, p) = \sup_{\alpha \in (0,1)} \inf_{V \in \mathcal{P}(\mathcal{Y}|\mathcal{X})} D_1(V \,\|\, W \,|\, p) + \frac{1-\alpha}{\alpha}(I_1(p;V) - R) \leqslant \\
&\leqslant \inf_{V \in \mathcal{P}(\mathcal{Y}|\mathcal{X})} \sup_{\alpha \in (0,1)} D_1(V \,\|\, W \,|\, p) + \frac{1-\alpha}{\alpha}(I_1(p;V) - R) = \\
&= \inf_{V\colon I_1(p;V) \leqslant R} D_1(V \,\|\, W \,|\, p).
\end{aligned}$$

Таким образом, (20) имеет место, когда $E_{\text{sp}}(R, W, p)$ равна бесконечности, т.е. для всех $R \in [0, I_{0^+}(p;W))$ и, возможно, для $R = I_{0^+}(p;W)$, а также когда $E_{\text{sp}}(R, W, p)$ конечна в силу леммы 2. ▲

Из выражения в форме Арутюняна вытекает следующее достаточное условие оптимальности порядка α в определении ЭСУ, данном в (10).

Л е м м а 4. *Для любых* $W\colon \mathcal{X} \to \mathcal{P}(\mathcal{Y})$, $\mathcal{A} \subset \mathcal{P}(\mathcal{X})$ *и* $R \in (C_{0^+,W,\mathcal{A}}, C_{1,W,\mathcal{A}})$, *если существуют* $\alpha^* \in (0,1)$ *и функция* V_p *меры* p *из* $\mathcal{A}$ *в* $\mathcal{P}(\mathcal{Y}\,|\,\mathcal{X})$, *удовлетворяющие двум неравенствам*

$$D_1(V_p \,\|\, q_{\alpha^*,W,\mathcal{A}} \,|\, p) \leqslant R, \quad \forall p \in \mathcal{A}, \tag{21}$$

$$D_1(V_p \,\|\, W \,|\, p) \leqslant \frac{1-\alpha^*}{\alpha^*}(C_{\alpha^*,W,\mathcal{A}} - R), \quad \forall p \in \mathcal{A}, \tag{22}$$

то $E_{\text{sp}}(R, W, \mathcal{A}) = \dfrac{1-\alpha^*}{\alpha^*}(C_{\alpha^*,W,\mathcal{A}} - R)$.

Для некоторых каналов $V_p = W_{\alpha^*}^{q_{\alpha^*,W,\mathcal{A}}}$ удовлетворяет как (21), так и (22); для таких каналов значение ЭСУ можно определить с помощью леммы 4. Однако для произвольных канала, скорости и соответствующего оптимального порядка α^* в (10) функция V_p, удовлетворяющая как (21), так и (22), может и не существовать, как, например, для Z-канала, обсуждаемого в Приложении A. Если бы достаточное условие оптимальности порядка α^*, указанное в (21) и (22), было также и необходимым, то доказательство Блейхута в [31] было бы верным; это, однако, не так в общем случае, как показано в Приложении A. Следует отметить, что для каналов, удовлетворяющих необходимому условию, указанному в (21) и (22), можно выводить ГСУ с помощью подхода из [31].

Д о к а з а т е л ь с т в о л е м м ы 4. Заметим, что в силу (6)

$$D_1(V_p \,\|\, q_{\alpha^*,W,\mathcal{A}} \,|\, p) = I_1(p;V_p) + D_1\left(\sum_x p(x) V_p(x) \,\|\, q_{\alpha^*,W,\mathcal{A}}\right).$$

Таким образом, $I_1(p;V_p) \leqslant R$ для всех $p \in \mathcal{A}$, поскольку дивергенция Реньи неотрицательна. Тогда $E_{\text{sp}}(R, W, p) \leqslant \dfrac{1-\alpha^*}{\alpha^*}(C_{\alpha^*,W,\mathcal{A}} - R)$ для всех $p \in \mathcal{A}$ по лемме 3. Тогда $E_{\text{sp}}(R, W, \mathcal{A}) \leqslant \dfrac{1-\alpha^*}{\alpha^*}(C_{\alpha^*,W,\mathcal{A}} - R)$ в силу (12). С другой стороны, $E_{\text{sp}}(R, W, \mathcal{A}) \geqslant$ $\geqslant \dfrac{1-\alpha^*}{\alpha^*}(C_{\alpha^*,W,\mathcal{A}} - R)$ по определению. Итак, $E_{\text{sp}}(R, W, \mathcal{A}) = \dfrac{1-\alpha^*}{\alpha^*}(C_{\alpha^*,W,\mathcal{A}} - R)$. ▲

2.4. Августиновский вариант границы Галлагера. Экспонента сферической упаковки является верхней границей экспоненциальной скорости убывания оптимальной вероятности ошибки с ростом длины блока, т.е. функции надежности, для каналов без памяти, удовлетворяющих довольно широким предположениям, как при списочном декодировании, так и без него, в силу ГСУ, приведенных в теоремах 1 и 2 в § 3. При списочном декодировании ЭСУ является также и нижней границей экспоненциальной скорости убывания оптимальной вероятности ошибки по длине блока [8, упражнение 5.20; 12, упражнение 10.28; 39]. Это последнее наблюдение можно проверить с помощью стандартных результатов, таких как [29, 30], хотя и с небольшими изменениями. Мы же проверим его, используя вариант границы Галлагера в терминах информации Августина. Напомним, что обычно граница Галлагера выводится через информацию Реньи, а не информацию Августина. Ключевую роль в нашем доказательстве играет свойство неподвижной точки (7). Это вариант границы Галлагера мы называем августиновским, так как, насколько нам известно, именно Августин впервые получил результат о достижимости [24, лемма 36.1] с помощью свойства неподвижной точки, описанного в (7).

Лемма 5. *Для любых $M, L \in \mathbb{Z}_+$, таких что $L < M$, любых $W\colon \mathcal{X} \to \mathcal{P}(\mathcal{Y})$, $p \in \mathcal{P}(\mathcal{X})$, $\mathcal{B} \subset \mathcal{X}$ и $\alpha \in \left[\frac{1}{1+L}, 1\right)$ существует (M, L)-код с функцией кодирования вида $\Psi\colon \mathcal{M} \to \mathcal{B}$, для которого*

$$\ln P_{\mathrm{e}} \leqslant \frac{\alpha - 1}{\alpha}\left[\tau - \ln\frac{(M-1)e}{L}\right] - \frac{\ln p(\mathcal{B})}{\alpha},$$

где $\tau = \inf_{x\in\mathcal{B}} D_\alpha(W(x) \,\|\, q_{\alpha,p})$.

Мы не даем отдельного доказательства леммы 5, поскольку она является частным случаем приведенной ниже леммы 6 с множителем Лагранжа, равным нулю. Прежде чем сформулировать лемму 6, укажем на немедленное следствие леммы 5 для кодов с постоянной композицией в стационарных каналах без памяти. Напомним, что для любой фиксированной композиции p при длине блока n вероятность множества всех последовательностей с композицией p (множества $\mathcal{T}_{p,n}$) для независимой и одинаково распределенной выборки (т.е. $p_t = p$ для всех t) согласно [12, с. 26] удовлетворяет следующему равенству для некоторого $\xi \in [0, 1]$:

$$\left(\bigotimes_{t=1}^{n} p_t\right)(\mathcal{T}_{p,n}) = e^{-\xi\frac{|\operatorname{supp} p|}{12\ln 2}}(2\pi n)^{-\frac{|\operatorname{supp} p|-1}{2}}\sqrt{\prod_{x\colon p(x)>0}\frac{1}{p(x)}}.$$

Тогда, учитывая равенство $q_{\alpha,\bigotimes_{t=1}^{n} p_t} = \bigotimes_{t=1}^{n} q_{\alpha,p_t}$, установленное в [26, лемма 14], и полагая $\mathcal{B}$ равным $\mathcal{T}_{p,n}$, получаем

Следствие 1. *Для любых $n \in \mathbb{Z}_+$, $\phi \in (0, 1)$, семейства $\{W_t\}_{t\in\mathbb{Z}_+}$, такого что $W_t = W$ для некоторого $W\colon \mathcal{X} \to \mathcal{P}(\mathcal{Y})$, вероятностной меры $p \in \mathcal{P}(\mathcal{X})$, такой что $np(x) \in \mathbb{Z}_{\geqslant 0}$ для всех $x \in \mathcal{X}$, целых чисел M, L, таких что $\frac{1}{L+1} < \eta$ и $\frac{1}{n}\ln\frac{M}{L} = I_\eta(p; W)$ для некоторого $\eta \in [\phi, 1)$, существует (M, L)-код в канале $\bigotimes_{t=1}^{n} W_t$ с функцией кодирования вида $\Psi\colon \mathcal{M} \to \mathcal{T}_{p,n}$, для которого*

$$\ln P_{\mathrm{e}} \leqslant -nE_{\mathrm{sp}}\Bigl(\frac{1}{n}\ln\frac{M}{L}, W, p\Bigr) + \\ + \frac{1}{\eta}\left[1 - \eta + \frac{|\operatorname{supp} p|}{12\ln 2} + \frac{|\operatorname{supp} p| - 1}{2}\ln(2\pi n) + \frac{1}{2}\sum_{x\colon p(x)>0}\ln p(x)\right].$$

Лемма 6. *Для любых $\ell, M, L \in \mathbb{Z}_+$, таких что $L < M$, любых $W\colon \mathcal{X} \to \mathcal{P}(\mathcal{Y})$, $\rho\colon \mathcal{X} \to \mathbb{R}^{\ell}_{\geqslant 0}$, $p \in \mathcal{P}(\mathcal{X})$, $\mathcal{B} \subset \mathcal{X}$ и $\alpha \in \left[\frac{1}{1+L}, 1\right)$ существует (M, L)-код с функцией кодирования вида $\Psi\colon \mathcal{M} \to \mathcal{B}$, для которого*

$$\ln P_{\mathrm{e}} \leqslant \frac{\alpha - 1}{\alpha}\left[\tau + \left(\inf_{x\in\mathcal{B}} \lambda \cdot \rho(x)\right) - \ln\frac{(M-1)e}{L}\right] - \frac{\ln p(\mathcal{B})}{\alpha}, \tag{23}$$

где $\tau = \inf\limits_{x\in\mathcal{B}} D_\alpha(W(x) \,\|\, q_{\alpha,p}) - \lambda \cdot \rho(x)$.

Следует отметить, что можно получить аналогичную границу через информацию Реньи – Галлагера, см. лемму 12 в Приложении В. Прежде чем перейти к доказательству леммы 6, вкратце обсудим ее следствия. Для каналов без памяти с ограничениями по стоимости с аддитивными функциями стоимости можно применить лемму 6 для $\lambda = 0$ и распределений-произведений на входе, удовлетворяющих ограничению по стоимости $\varrho - \epsilon$. Тогда, применяя слабый закон больших чисел с учетом [26, леммы 14, 27 и 28], можно сделать вывод, что функция надежности ограничена снизу экспонентой сферической упаковки при довольно широких предположениях. В случае $\ell = 1$ этот результат можно усилить, применяя теорему Берри – Эссеена и [25, лемма 19] при соответствующих предположениях. Если получить результат, аналогичный [25, лемма 19], для сумм случайных векторов, то следующее утверждение можно будет немедленно обобщить на случай $\ell > 1$.

Следствие 2. *Пусть $n \in \mathbb{Z}_+$, $n \geqslant 21$, $\phi \in (0,1)$, $W_{[1,n]}\colon \mathcal{X}_1^n \to \mathcal{P}(\mathcal{Y}_1^n)$ – канал-произведение длины n с аддитивной функцией стоимости $\rho_{[1,n]}\colon \mathcal{X}_1^n \to \mathbb{R}_{\geqslant 0}$, такой что $\inf\limits_{x_t\in\mathcal{X}_t} \rho_t(x_t) = 0$ для всех $t \in [1,n]$ и*

$$C_{\alpha,\widetilde{W}_{[1,n]},n\varrho} \geqslant C_{\alpha,W_{[1,n]},n\varrho} - \epsilon, \quad \forall\alpha \in [\phi, 1),\ \varrho \in \left[\widetilde{\varrho}, \widetilde{\varrho} + \frac{9\varsigma}{n}\right], \tag{24}$$

для некоторых $\widetilde{\varrho} \in \mathbb{R}_+$, $\varsigma \in \mathbb{R}_{\geqslant 0}$, $\epsilon \in \mathbb{R}_{\geqslant 0}$ и канала-произведения $\widetilde{W}_{[1,n]}\colon \widetilde{\mathcal{X}}_1^n \to \mathcal{P}(\mathcal{Y}_1^n)$ с множествами входов $\widetilde{\mathcal{X}}_t = \{x_t \in \mathcal{X}_t : \rho_t(x_t) \leqslant \varsigma\}$, такого что $\widetilde{W}_{[1,n]}(x_1^n) = W_{[1,n]}(x_1^n)$ для всех $x_1^n \in \widetilde{\mathcal{X}}_1^n$. Тогда для любых $\delta \in (0, \widetilde{\varrho})$, $\varrho \in \left(\widetilde{\varrho} + \frac{3e\varsigma}{n}, \widetilde{\varrho} + \frac{9\varsigma}{n}\right)$ и целых чисел M, L, таких что $\frac{1}{L+1} < \eta$ и $\ln\frac{M}{L} = C_{\eta,W_{[1,n]},n\varrho}$ для некоторого $\eta \in [\phi, 1)$, существует (M, L)-код с функцией кодирования вида $\Psi\colon \mathcal{M} \to \left\{x_1^n \in \widetilde{\mathcal{X}}_1^n : \rho_{[1,n]}(x_1^n) \leqslant n\varrho\right\}$, для которого

$$\begin{aligned}\ln P_{\mathrm{e}} &\leqslant -E_{\mathrm{sp}}\left(\ln\frac{M}{L}, W_{[1,n]}, n\varrho\right) + \\ &+ \frac{1-\eta}{\eta}\left(\frac{C_{\eta,W_{[1,n]},n\varrho}}{n}\frac{6\varsigma e}{\delta} + 2\epsilon + 1\right) + \frac{\ln 4n}{2\eta} + \frac{1}{n}.\end{aligned} \tag{25}$$

Заметим, что если бы область значений ϱ не зависела от ς в (24), а интервал значений α был бы компактным подмножеством в $(0,1)$, то можно было бы найти ς для любого положительного ϵ с помощью теоремы Дини [35, теорема 2.4.10], используя конструкцию, аналогичную той, что была использована в доказательстве [27, лемма 15(g)], поскольку $C_{\alpha,W,\varrho}$ непрерывна по (α, ϱ) на $(0,1) \times \mathbb{R}_+$ согласно [26, лемма 27(c)]. Это, однако, не так, и могут существовать каналы-произведения, для которых неравенство (24) не выполнено ни для какой тройки $(\epsilon, \varsigma, \widetilde{\varrho})$. Кроме того, даже когда условия следствия 2 выполнены, точность границы (25) зависит от значений констант, удовлетворяющих этим условиям. Тем не менее, легко видеть, что если либо все функции стоимости каналов-компонент ограничены, либо $C_{\phi,W_{[1,n]},n\varrho}$ имеет порядок $\Theta(n)$ (т.е. растет линейно по n), а условие (24) выполнено для ϵ и ς, имеющих порядок $O(\ln n)$ (т.е. растущих не более чем линейно по $\ln n$), то коэффициент в (25) равен $e^{-O(\ln n)}$, т.е. убывает полиномиально по n. Доказательство следствия 2 дано в Приложении С.

Доказательство леммы 6. Установим существование кода с требуемыми свойствами методом случайного кодирования. Для любых $p \in \mathcal{P}(\mathcal{X})$ и $\mathcal{B} \subset \mathcal{X}$, таких что $p(\mathcal{B}) > 0$, определим $p_{\mathcal{B}} \in \mathcal{P}(\mathcal{X})$ как

$$p_{\mathcal{B}}(x) \triangleq \frac{\mathbb{1}_{\{x \in \mathcal{B}\}} p(x)}{p(\mathcal{B})}, \quad \forall x \in \mathcal{X}. \tag{26}$$

Рассмотрим ансамбль кодов, в котором распределения сообщений по элементам $\mathcal{B}$ совместно независимы и $\Psi(m) = x$ с вероятностью $p_{\mathcal{B}}(x)$ для всех m из множества сообщений. Декодер выбирает L сообщений с наибольшим отношением $\dfrac{f_{\Psi(m)}}{h_{\Psi(m)}}$ для функций f_x и h_x, определяемых следующим образом:

$$f_x(y) \triangleq \frac{dW(x)}{dq}(y), \quad \forall x \in \mathcal{X},\ y \in \mathcal{Y}, \tag{27}$$

$$h_x \triangleq e^{\frac{\alpha-1}{\alpha}\lambda \cdot \rho(x)}, \quad \forall x \in \mathcal{X}, \tag{28}$$

где мера $q \in \mathcal{P}(\mathcal{Y})$, которую мы выберем позже, такова, что $q_{\alpha,p} \prec q$. В случае равенства значений декодер выбирает сообщение (или сообщения) с наименьшим индексом (наименьшими индексами). Чтобы оценить математическое ожидание вероятности ошибки по ансамблю, рассмотрим математическое ожидание условной вероятности ошибки сообщения с наибольшим индексом. Ошибка возникает только тогда, когда отношения $\dfrac{f}{h}$, соответствующие L или более сообщениям, по крайней мере столь же велики, как $\dfrac{f}{h}$, соответствующее переданному сообщению. Вероятность этого можно оценить, используя зависящее от y вспомогательное пороговое значение $\gamma(y)$, и получить следующую оценку математического ожидания P_{e} по ансамблю[6]:

$$\begin{aligned}\mathbf{E}[P_{\mathrm{e}}] \leqslant &\sum_x p_{\mathcal{B}}(x)\, \mathbf{E}_q\left[\mathbb{1}_{\{\frac{f_x}{h_x} \leqslant \gamma\}} f_x\right] + \\ &+ \binom{M-1}{L} \sum_x p_{\mathcal{B}}(x)\, \mathbf{E}_q\left[\mathbb{1}_{\{\frac{f_x}{h_x} > \gamma\}} \left[\sum_z p_{\mathcal{B}}(z) \mathbb{1}_{\{\frac{f_z}{h_z} \geqslant \frac{f_x}{h_x}\}}\right]^L f_x\right].\end{aligned} \tag{29}$$

Наложим ограничения $\alpha \leqslant 1$, чтобы оценить первое слагаемое в этой сумме, а также $\alpha \geqslant \dfrac{1}{1+L}$, чтобы оценить второе:

$$\mathbf{E}_q\left[\mathbb{1}_{\{\frac{f_x}{h_x} \leqslant \gamma\}} f_x\right] \leqslant h_x\, \mathbf{E}_q\left[\mathbb{1}_{\{\frac{f_x}{h_x} \leqslant \gamma\}} \left(\frac{f_x}{h_x}\right)^{\alpha} \gamma^{1-\alpha}\right],$$

$$\begin{aligned}&\mathbf{E}_q\left[\mathbb{1}_{\{\frac{f_x}{h_x} > \gamma\}} \left[\sum_z p_{\mathcal{B}}(z) \mathbb{1}_{\{\frac{f_z}{h_z} \geqslant \frac{f_x}{h_x}\}}\right]^L f_x\right] \leqslant \\ &\leqslant h_x\, \mathbf{E}_q\left[\mathbb{1}_{\{\frac{f_x}{h_x} > \gamma\}} \left[\sum_z p_{\mathcal{B}}(z) \left(\frac{f_z}{h_z}\right)^{\alpha}\right]^L \left(\frac{f_x}{h_x}\right)^{1-L\alpha}\right] \leqslant \\ &\leqslant h_x\, \mathbf{E}_q\left[\mathbb{1}_{\{\frac{f_x}{h_x} > \gamma\}} \left[\sum_z p_{\mathcal{B}}(z) \left(\frac{f_z}{h_z}\right)^{\alpha}\right]^L \left(\frac{f_x}{h_x}\right)^{\alpha} \gamma^{1-\alpha-L\alpha}\right].\end{aligned}$$

---

[6] Заметим, что $\sum\limits_{t=L}^{M} \binom{M}{t} s^t (1-s)^{M-t} = \binom{M}{L} s^L \sum\limits_{t=0}^{M-L} \frac{L!\,(M-L)!}{(L+t)!\,(M-L-t)} s^t (1-s)^{M-L-t} \leqslant \binom{M}{L} s^L$ для всех $s \in [0,1]$.

Полагая $\gamma = \left[\sum_x p_{\mathcal{B}}(x)\Big(\frac{f_x}{h_x}\Big)^{\alpha}\right]^{\frac{1}{\alpha}} \left[\binom{M-1}{L}\right]^{\frac{1}{L\alpha}}$, получаем

$$\mathbf{E}[P_{\mathrm{e}}] \leqslant \sum_x p_{\mathcal{B}}(x) h_x\, \mathbf{E}_q \left[\left(\frac{f_x}{h_x}\right)^{\alpha} \gamma^{1-\alpha}\right] =$$

$$= \mathbf{E}_q \left[\left[\sum_x p_{\mathcal{B}}(x) \left(\frac{f_x}{h_x}\right)^{\alpha}\right]^{\frac{1}{\alpha}}\right] \left(\sup_{x\in\mathcal{B}} h_x\right) \left[\binom{M-1}{L}\right]^{\frac{1-\alpha}{L\alpha}}.$$

С другой стороны, используя формулу Стирлинга $\sqrt{2\pi n}(n/e)^n \leqslant n! \leqslant e\sqrt{n}(n/e)^n$ и неравенство $\ln z \leqslant z-1$, получаем

$$\frac{1}{L}\ln\binom{M-1}{L} \leqslant \frac{1}{L}\ln\frac{e\sqrt{M-1}}{2\pi\sqrt{L(M-1-L)}} + \ln\frac{M-1}{L} +$$
$$+ \frac{M-1-L}{L}\ln\left(1+\frac{L}{M-1-L}\right) \leqslant \ln\frac{M-1}{L} + 1.$$

Отсюда вытекает следующая оценка математического ожидания P_{e} по ансамблю:

$$\ln \mathbf{E}[P_{\mathrm{e}}] \leqslant \ln \mathbf{E}_q \left[\left(\sum_x p_{\mathcal{B}}(x) e^{(1-\alpha)\lambda\cdot\rho(x)} \left(\frac{dW(x)}{dq}\right)^{\alpha}\right)^{\frac{1}{\alpha}}\right] +$$
$$+ \frac{\alpha-1}{\alpha}\left[\left(\inf_{x\in\mathcal{B}} \lambda\cdot\rho(x)\right) - \ln\frac{(M-1)e}{L}\right]. \tag{30}$$

Полагая $q = q_{\alpha,p}$ и используя вначале определение скошенного канала, данное в (3), а затем определение τ, получаем

$$\ln \mathbf{E}[P_{\mathrm{e}}] \leqslant \ln \mathbf{E}_{q_{\alpha,p}} \left[\left(\sum_x p_{\mathcal{B}}(x) e^{(\alpha-1)[D_\alpha(W(x)\,\|\,q_{\alpha,p}) - \lambda\cdot\rho(x)]} \frac{dW_\alpha^{q_{\alpha,p}}(x)}{dq_{\alpha,p}}\right)^{\frac{1}{\alpha}}\right] +$$
$$+ \frac{\alpha-1}{\alpha}\left[\left(\inf_{x\in\mathcal{B}} \lambda\cdot\rho(x)\right) - \ln\frac{(M-1)e}{L}\right] \leqslant \ln \mathbf{E}_{q_{\alpha,p}} \left[\left(\sum_x p_{\mathcal{B}}(x) \frac{dW_\alpha^{q_{\alpha,p}}(x)}{dq_{\alpha,p}}\right)^{\frac{1}{\alpha}}\right] +$$
$$+ \frac{\alpha-1}{\alpha}\left[\tau + \left(\inf_{x\in\mathcal{B}} \lambda\cdot\rho(x)\right) - \ln\frac{(M-1)e}{L}\right].$$

Так как существует код с вероятностью ошибки P_{e}, меньшей или равной $\mathbf{E}[P_{\mathrm{e}}]$, то существование кода с функцией кодирования вида $\Psi\colon \mathcal{M} \to \mathcal{B}$, удовлетворяющего условию (23), следует из (4), (7) и (26). ▲

§ 3. Границы сферической упаковки для каналов без памяти

Предположение 1. *Максимальное значение* $C_{1/2,U_{\mathcal{A}}^{(t)}}$ *для целых* t, *меньших или равных* n, *имеет порядок* $O(\ln n)$, *т.е.*

$$\exists n_0 \in \mathbb{Z}_+,\ K \in \mathbb{R}_+, \quad \textit{такие что} \quad \max_{t:\, t\leqslant n} C_{1/2,U_{\mathcal{A}}^{(t)}} \leqslant K\ln(n), \quad \forall n \geqslant n_0,$$

где канал $U_{\mathcal{A}}^{(t)}\colon \mathcal{X}_{\mathcal{A}}^{(t)} \to \mathcal{P}(\mathcal{Y})$ *таков, что* $U_{\mathcal{A}}^{(t)}(x) = W(x)$ *для всех* $x \in \mathcal{X}_{\mathcal{A}}^{(t)}$, *а* $\mathcal{X}_{\mathcal{A}}^{(t)}$ *определяется для любых* $t \in \mathbb{Z}_+$ *и* $\mathcal{A} \subset \mathcal{P}(\mathcal{X})$ *как*

$$\mathcal{X}_{\mathcal{A}}^{(t)} \triangleq \{x \in \mathcal{X} :\ \exists p \in \mathcal{A},\ \textit{такая что}\ p(z)t \in \mathbb{Z},\ \forall z \in \mathcal{X},\ \textit{таких что}\ p(x) > 0\}.$$

Теорема 1. *Пусть* $\{W_t\}_{t\in\mathbb{Z}_+}$ *– стационарная последовательность каналов, такая что* $W_t = W$ *для всех* $t \in \mathbb{Z}_+$, $\mathcal{A} \subset \mathcal{P}(\mathcal{X})$ *– выпуклое множество ограничений, удовлетворяющее предположению 1, а* ε, α_0, α_1 *– положительные параметры, такие что* $0 < \alpha_0 < \alpha_1 < 1$. *Тогда для любой последовательности кодов* $\left\{(\Psi^{(n)}, \Theta^{(n)})\right\}_{n\in\mathbb{Z}_+}$ *на каналах-произведениях* $\left\{W_{[1,n]}\right\}_{n\in\mathbb{Z}_+}$, *такой что*

$$C_{\alpha_1,W,\mathcal{A}} \geqslant \frac{1}{n}\ln\frac{M_n}{L_n} \geqslant C_{\alpha_0,W,\mathcal{A}} + \frac{\ln n}{n}[\ln(\ln n)]^{\varepsilon}, \quad \forall n \geqslant n_0, \tag{31}$$

и $\Upsilon(\Psi^{(n)}(m)) \in \mathcal{A}$ *для всех* $m \in \mathcal{M}^{(n)}$, *существуют* $\tau \in \mathbb{R}_+$ *и* $n_1 \geqslant n_0$, *такие что*

$$P_{\mathrm{e}}^{(n)} \geqslant n^{-\tau} e^{-nE_{\mathrm{sp}}\left(\frac{1}{n}\ln\frac{M_n}{L_n}, W, \mathcal{A}\right)}, \quad \forall n \geqslant n_1. \tag{32}$$

Следует отметить, что условия теоремы 1 выполнены для гауссовских и пуассоновских моделей, которые изучались в [15–22]; таким образом, из теоремы 1 следуют асимптотические ГСУ для таких каналов. Неасимптотический аналог этого утверждения представлен в п. 3.3 между леммой 9 и следствием 3.

Для кодов в каналах без памяти с ограничениями по стоимости можно отказаться от условия стационарности, используя технику выпуклого сопряжения (см. [24, теорема 36.6]). Теорема 2 усиливает [24, теорема 36.6], избавляясь от условия ограниченности функции стоимости и устанавливая ГСУ с коэффициентом, полиномиальным по длине блока n, вместо коэффициента вида $e^{-O(\sqrt{n})}$.

Предположение 2. *Максимальное значение* $C_{1/2,W_t,n\varrho}$ *для целых* t, *меньших или равных* n, *имеет порядок* $O(\ln n)$:

$$\exists n_0 \in \mathbb{Z}_+,\ K \in \mathbb{R}_+, \quad \textit{такие что} \quad \max_{t:\, t\leqslant n} C_{1/2,W_t,n\varrho} \leqslant K\ln(n), \quad \forall n \geqslant n_0.$$

Теорема 2. *Пусть* $\{(W_t, \rho_t)\}_{t\in\mathbb{Z}_+}$ *– последовательность каналов с соответствующими функциями стоимости, удовлетворяющая предположению 2, а* ε, α_0, α_1 *– положительные параметры, такие что* $0 < \alpha_0 < \alpha_1 < 1$. *Тогда в каналах-произведениях* $\left\{W_{[1,n]}\right\}_{n\in\mathbb{Z}_+}$ *с аддитивными функциями стоимости* $\left\{\rho_{[1,n]}\right\}_{n\in\mathbb{Z}_+}$ *вида* $\rho_{[1,n]}(x_1^n) = \sum\limits_{t=1}^{n}\rho_t(x_t)$, *для любой последовательности кодов* $\left\{(\Psi^{(n)}, \Theta^{(n)})\right\}_{n\in\mathbb{Z}_+}$, *такой что*

$$C_{\alpha_1,W_{[1,n]},n\varrho} \geqslant \ln\frac{M_n}{L_n} \geqslant C_{\alpha_0,W_{[1,n]},n\varrho} + \varepsilon(\ln n)^2, \quad \forall n \geqslant n_0, \tag{33}$$

и $\rho_{[1,n]}(\Psi^{(n)}(m)) \leqslant n\varrho$ *для всех* $m \in \mathcal{M}^{(n)}$ *при ограничении по стоимости* ϱ *на одно обращение к каналу, таком что* $n\varrho \in \operatorname{int}\Gamma_{\rho_{[1,n]}}$, *существуют* $\tau \in \mathbb{R}_+$ *и* $n_1 \geqslant n_0$, *для которых*

$$P_{\mathrm{e}}^{(n)} \geqslant n^{-\tau} e^{-E_{\mathrm{sp}}\left(\ln\frac{M_n}{L_n}, W_{[1,n]}, n\varrho\right)}, \quad \forall n \geqslant n_1. \tag{34}$$

Чтобы установить асимптотическую ГСУ с полиномиальным коэффициентом в теореме 2, мы предполагали, что августиновская пропускная способность порядка 1/2 с ограничениями по стоимости растет не быстрее чем логарифмически по стоимости. Хотя это условие и выполнено для большинства содержательных случаев, включая различные гауссовские и пуассоновские каналы, изучавшиеся в [15–22], а также многие их нестационарные варианты, существуют и каналы, для которых это условие не выполнено, см. [26, пример 1]. Таким образом, может иметь смысл избавиться от условия логарифмического роста по ограничению по стоимости. Этого можно добиться, если в доказательстве теоремы 2 использовать лемму 11 вместо леммы 10.

Наша основная цель в этом параграфе – доказать две асимптотические ГСУ, сформулированные в теоремах 1 и 2, которые и представляют собой основной результат настоящей статьи. Для этого вначале в п. 3.1 мы изложим результат о невозможности для задачи проверки гипотез с независимыми выборками с помощью теоремы Берри–Эссеена. Затем в п. 3.2 введем понятия усредненной августиновской пропускной способности и усредненной ЭСУ. Мы выведем неасимптотические ГСУ (хотя и в параметрическом виде) через эти усредненные величины для кодов с ограничениями по композиции в стационарных каналах без памяти в п. 3.3 и для кодов с ограничениями по стоимости в (возможно, нестационарных) каналах без памяти в п. 3.4. Достаточно очевидный вывод этих асимптотических ГСУ из неасимптотических с помощью леммы 8 из п. 3.2 представлен в Приложении C.

3.1. Результат о невозможности для задачи проверки гипотез.

Лемма 7. *Для любых* $\kappa \geqslant 3$, $\alpha \in (0,1)$, $n \in \mathbb{Z}_+$ *и произведений мер* $w, q \in \mathcal{P}(\mathcal{Y}_1^n)$ *вида* $w = \bigotimes_{t=1}^{n} w_t$ *и* $q = \bigotimes_{t=1}^{n} q_t$ *положим* $\xi_t \triangleq \ln \frac{dw_{t,\mathrm{ac}}}{dq_t} - \mathbf{E}_{w_\alpha^q}\left[\ln \frac{dw_{t,\mathrm{ac}}}{dq_t}\right]$, *где* $w_{t,\mathrm{ac}}$ – *компонента* w_t, *абсолютно непрерывная по* q_t *и такая, что*

$$g_\kappa \triangleq \left(\sum_{t=1}^{n} \mathbf{E}_{w_\alpha^q}\left[|\xi_t|^\kappa\right]\right)^{1/\kappa}.$$

Тогда любое $\mathcal{E} \in \mathcal{Y}$, *такое что* $q(\mathcal{E}) \leqslant \frac{1}{4\sqrt{n}} e^{-D_1(w_\alpha^q \,\|\, q) - \alpha 3 g_\kappa}$, *также удовлетворяет условию*

$$w(\mathcal{Y}_1^n \setminus \mathcal{E}) \geqslant \frac{1}{4\sqrt{n}} e^{-D_1(w_\alpha^q \,\|\, w) - (1-\alpha) 3 g_\kappa}. \tag{35}$$

Лемма 7 дает результат о невозможности для задачи проверки гипотез с независимыми выборками в духе [7, теорема 5]. Однако вместо неравенства Чебышева в ней применяется теорема Берри–Эссеена, использованная в [25, лемма 19].

Доказательство леммы 7. Положим $\mathcal{E}_0 = \left\{y : \left|\sum_{t=1}^{n} \xi_t\right| \leqslant 3 g_\kappa\right\}$; тогда

$$\left|\ln \frac{dw_\alpha^q}{dq}(y_1^n) - D_1(w_\alpha^q \,\|\, q)\right| \leqslant \alpha 3 g_\kappa, \quad \forall y_1^n \in \mathcal{E}_0,$$
$$\left|\ln \frac{dw_\alpha^q}{dw}(y_1^n) - D_1(w_\alpha^q \,\|\, w)\right| \leqslant (1-\alpha) 3 g_\kappa, \quad \forall y_1^n \in \mathcal{E}_0.$$

Отсюда

$$w_\alpha^q(\mathcal{E} \cap \mathcal{E}_0) \leqslant q(\mathcal{E} \cap \mathcal{E}_0) e^{D_1(w_\alpha^q \,\|\, q) + \alpha 3 g_\kappa}, \tag{36}$$
$$w_\alpha^q(\mathcal{E} \cap \mathcal{E}_0) \leqslant w(\mathcal{E} \cap \mathcal{E}_0) e^{D_1(w_\alpha^q \,\|\, w) + (1-\alpha) 3 g_\kappa}. \tag{37}$$

С другой стороны, в силу определения $\mathcal{E}_0$ и [25, лемма 19] имеем

$$\mathbf{P}_{w_\alpha^q}[\mathcal{E}_0] \geqslant \frac{1}{2\sqrt{n}}. \tag{38}$$

Неравенство (35) следует из (36)–(38). ▲

3.2. Усреднение по Августину. Наши неасимптотические ГСУ выражаются через усредненную августиновскую пропускную способность и усредненную ЭСУ, которые определяются для всех $\epsilon \in (0,1)$ и $R \in \mathbb{R}_{\geqslant 0}$ следующим образом:

$$\widetilde{C}^{\epsilon}_{\alpha,W,\mathcal{A}} \triangleq \frac{1}{\epsilon} \int\limits_{\alpha-\epsilon\alpha}^{\alpha+\epsilon(1-\alpha)} \left[1 \vee \left(\frac{\alpha}{1-\alpha}\frac{1-\eta}{\eta}\right)\right] C_{\eta,W,\mathcal{A}}\, d\eta, \tag{39}$$

$$\widetilde{E}^{\epsilon}_{\mathrm{sp}}(R,W,\mathcal{A}) \triangleq \sup_{\alpha\in(0,1)} \frac{1-\alpha}{\alpha}\big(\widetilde{C}^{\epsilon}_{\alpha,W,\mathcal{A}} - R\big). \tag{40}$$

Заметим, что $\lim\limits_{\epsilon\downarrow 0} \widetilde{C}^{\epsilon}_{\alpha,W,\mathcal{A}} = C_{\alpha,W,\mathcal{A}}$ для всех $\alpha \in (0,1)$, поскольку $C_{\alpha,W,\mathcal{A}}$ непрерывна по α на $(0,1)$ согласно [26, лемма 23(d)]. Кроме того, используя монотонность функций $C_{\alpha,W,\mathcal{A}}$ и $\frac{1-\alpha}{\alpha}C_{\alpha,W,\mathcal{A}}$ по α (см. [26, лемма 23(a),(b)]), можно показать, что сходимость на компактных подмножествах $(0,1)$ равномерна. Тем не менее, $C_{\alpha,W,\mathcal{A}}$ нельзя аппроксимировать величиной $\widetilde{C}^{\epsilon}_{\alpha,W,\mathcal{A}}$ равномерно на всем интервале $(0,1)$, поскольку $\lim\limits_{\alpha\uparrow 1} \widetilde{C}^{\epsilon}_{\alpha,W,\mathcal{A}} = \infty$ для всех положительных ϵ, когда $C_{1/2,W,\mathcal{A}}$ положительна, даже если $C_{1,W,\mathcal{A}}$ конечна. Этот факт следует из монотонности функций $C_{\alpha,W,\mathcal{A}}$ и $\frac{1-\alpha}{\alpha}C_{\alpha,W,\mathcal{A}}$ по α:

$$\widetilde{C}^{\epsilon}_{\alpha,W,\mathcal{A}} \geqslant \frac{1}{\epsilon} \int\limits_{\alpha-\epsilon\alpha}^{\alpha-\frac{\epsilon}{2}\alpha} \frac{\alpha}{1-\alpha}\frac{1-\eta}{\eta} C_{\eta,W,\mathcal{A}}\, d\eta \geqslant$$

$$\geqslant \frac{\alpha}{2-\epsilon}\left(1+\frac{\alpha\epsilon}{2(1-\alpha)}\right) C_{\alpha-\epsilon\alpha,W,\mathcal{A}}.$$

Как функция скорости, усредненная ЭСУ равномерно сходится к ЭСУ на любом компактном множестве скоростей, меньших августиновской пропускной способности порядка 1, как показывает следующая

Лемма 8. *Для любых $W\colon \mathcal{X} \to \mathcal{P}(\mathcal{Y})$ и $\mathcal{A} \subset \mathcal{P}(\mathcal{X})$, таких что $C_{1/2,W,\mathcal{A}} \in \mathbb{R}_+$, и любых $\phi \in (0,1)$, $R \in [C_{\phi,W,\mathcal{A}}, \infty)$ и $\epsilon \in (0,\phi)$*

$$0 \leqslant \widetilde{E}^{\epsilon}_{\mathrm{sp}}(R,W,\mathcal{A}) - E_{\mathrm{sp}}(R,W,\mathcal{A}) \leqslant \frac{\epsilon}{1-\epsilon}\frac{R \vee E_{\mathrm{sp}}(R,W,\mathcal{A})}{\phi} \leqslant \tag{41}$$

$$\leqslant \frac{\epsilon}{1-\epsilon}\frac{R}{\phi^2}. \tag{42}$$

Лемма 8 вытекает из монотонности функций $C_{\alpha,W,\mathcal{A}}$ и $\frac{1-\alpha}{\alpha}C_{\alpha,W,\mathcal{A}}$ по α, установленной в [26, лемма 23(a),(b)]; доказательство приведено в Приложении C. Доказательство леммы 8 практически совпадает с доказательством леммы 15 из [25], которая устанавливает тот же самый результат для случая $\mathcal{A} = \mathcal{P}(\mathcal{X})$.

3.3. Неасимптотические границы сферической упаковки для кодов с ограничениями по композиции. Коды с ограничениями по композиции в стационарных каналах без памяти активно изучались, но лишь для случая конечного множества входов. Чаще всего предполагается, что все кодовые слова имеют в точности одинаковую композицию. Если это не так, обычно применяется метод выбрасывания с учетом композиций кодовых слов, при котором оставляются слова с наиболее частой композицией. Однако такое выбрасывание приводит к нетривиальному результату только тогда, когда множество входов конечно. Границы, получаемые при анализе наиболее частой композиции в таких случаях, можно выводить через информацию Августина и августиновское среднее. Напротив того, в следующей лемме используются понятия августиновских пропускной способности и центра, а также усреднения по Августину, чтобы избежать предположений о конечности множества входов или о композиции, имеющей не менее чем полиномиальную по длине блока долю кодовых слов.

Лемма 9. *Пусть* $n \in \mathbb{Z}_+$, $W\colon \mathcal{X} \to \mathcal{P}(\mathcal{Y})$, $\mathcal{A}$ – *выпуклое подмножество* $\mathcal{P}(\mathcal{X})$, κ, ϕ, ϵ – *положительные параметры, такие что* $\kappa \geqslant 3$, $\phi < 1$, $\epsilon < 1$, *и пусть*

$$\gamma \triangleq 3\sqrt[\kappa]{3n}\left(\left[C_{1/2,U_{\mathcal{A}}^{(n)}} + (1-\phi)\ln(1+n)\right] \vee \kappa\right). \tag{43}$$

Если $W_t = W$ *для всех* $t \leqslant n$, *а* $M, L \in \mathbb{Z}_+$ *таковы, что* $\ln\dfrac{M}{L} > n\widetilde{C}^{\epsilon}_{\phi,W,\mathcal{A}} + \dfrac{\gamma}{1-\phi} + \ln\dfrac{8e^3n^{1{,}5}}{\epsilon}$, *то для любого* (M, L)-*кода в канале* $W_{[1,n]}$, *такого что* $\Upsilon(\Psi(m)) \in \mathcal{A}$ *для всех* $m \in \mathcal{M}$, *справедливо*

$$P_{\mathrm{e}} \geqslant \left(\frac{\epsilon e^{-2\gamma}}{8e^3n^{1{,}5}}\right)^{1/\phi} e^{-n\widetilde{E}^{\epsilon}_{\mathrm{sp}}\left(\frac{1}{n}\ln\frac{M}{L},W,\mathcal{A}\right)}. \tag{44}$$

Отсутствие в (44) члена, учитывающего скорость кода, можно считать достоинством данной формулы, но это возможно только за счет наличия коэффициента, при котором граница становится тривиальной, т.е. равной нулю, когда ϕ стремится к нулю. Слегка изменив рассуждения в доказательстве, можно получить следующую альтернативную границу:

$$P_{\mathrm{e}} \geqslant \frac{\epsilon e^{-2\gamma}}{8n^{1{,}5}} e^{-n\widetilde{E}^{\epsilon}_{\mathrm{sp}}(R,W,\mathcal{A})}, \quad R = \frac{1}{n}\ln\frac{M}{L} - \frac{1}{n}\ln\frac{8e^3n^{1{,}5}}{\epsilon e^{-2\gamma}}. \tag{45}$$

С учетом соотношения $C_{\alpha,U^{(n)}_{\mathcal{A}(\varrho)}} \leqslant C_{\alpha,W,n\varrho}$ из леммы 9 и формулы (45) вытекают приведенные ниже следствие 3 и формула (46).

Следствие 3. *Пусть* $n, \ell \in \mathbb{Z}_+$, $W\colon \mathcal{X} \to \mathcal{P}(\mathcal{Y})$, $\rho\colon \mathcal{X} \to \mathbb{R}^{\ell}_{\geqslant 0}$, $\varrho \in \Gamma_\rho$, κ, ϕ, ϵ – *положительные параметры, такие что* $\kappa \geqslant 3$, $\phi < 1$, $\epsilon < 1$, *и пусть*

$$\gamma \triangleq 3\sqrt[\kappa]{3n}\left(\left[C_{1/2,W,n\varrho} + (1-\phi)\ln(1+n)\right] \vee \kappa\right).$$

Если $W_t = W$ *для всех* $t \leqslant n$, *а* $M, L \in \mathbb{Z}_+$ *таковы, что* $\ln\dfrac{M}{L} > n\widetilde{C}^{\epsilon}_{\phi,W,\varrho} + \dfrac{\gamma}{1-\phi} + \ln\dfrac{8e^3n^{1{,}5}}{\epsilon}$, *то для любого* (M, L)-*кода в канале* $W_{[1,n]}$, *такого что* $\sum\limits_{t=1}^{n}\rho(\Psi_t(m)) \leqslant n\varrho$ *для всех* $m \in \mathcal{M}$, *справедливо*

$$P_{\mathrm{e}} \geqslant \left(\frac{\epsilon e^{-2\gamma}}{8e^3n^{1{,}5}}\right)^{1/\phi} e^{-n\widetilde{E}^{\epsilon}_{\mathrm{sp}}\left(\frac{1}{n}\ln\frac{M}{L},W,\varrho\right)}.$$

Слегка изменив рассуждения, можно получить следующую альтернативную границу:

$$P_{\mathrm{e}} \geqslant \frac{\epsilon e^{-2\gamma}}{8n^{1{,}5}} e^{-n\widetilde{E}^{\epsilon}_{\mathrm{sp}}(R,W,\varrho)}, \quad R = \frac{1}{n}\ln\frac{M}{L} - \frac{1}{n}\ln\frac{8e^3n^{1{,}5}}{\epsilon e^{-2\gamma}}. \tag{46}$$

Доказательство леммы 9 и формулы (45). Августиновский центр $q_{\alpha,W,\mathcal{A}}$ непрерывен по α на $(0,1)$ в топологии полной вариации на $\mathcal{P}(\mathcal{Y})$ согласно [26, леммы 23(d) и 24]. Таким образом, $q_{\cdot,W,\mathcal{A}}$ – это вероятность перехода из $((0,1), \mathcal{B}((0,1)))$ в $(\mathcal{Y}, \mathscr{Y})$. Для каждого $t \leqslant n$ определим усредненный центр $q^{\epsilon}_{\alpha,t}$ как маргинальное распределение на $\mathcal{Y}$ для вероятностной меры $u_{\alpha,\epsilon} \circledast q_{\cdot,W,\mathcal{A}}$, где $u_{\alpha,\epsilon}$ –

равномерное распределение вероятностей на $(\alpha-\epsilon\alpha, \alpha+\epsilon(1-\alpha))$:

$$q_{\alpha,t}^{\epsilon} \triangleq \frac{1}{\epsilon}\int\limits_{\alpha-\alpha\epsilon}^{\alpha+(1-\alpha)\epsilon} q_{\eta,W,\mathcal{A}}\, d\eta, \quad \forall t\in\{1,\ldots,n\}. \tag{47}$$

Определим $q_{\alpha,t}\in\mathcal{P}(\mathcal{Y}_t)$ и $q_\alpha\in\mathcal{P}(\mathcal{Y}_1^n)$ как

$$q_{\alpha,t} \triangleq \frac{n}{n+1}q_{\alpha,t}^{\epsilon} + \frac{1}{n+1}q_{1/2,U_{\mathcal{A}}^{(n)}}, \quad \forall t\in\{1,\ldots,n\},$$
$$q_\alpha \triangleq \bigotimes_{t=1}^{n} q_{\alpha,t}.$$

Для краткости обозначим вероятностную меру $\bigotimes\limits_{t=1}^{n} W(\Psi_t(m))$, порожденную $\Psi(m)$, через w^m. Имеем

$$\begin{aligned}
&D_\alpha(w^m \,\|\, q_\alpha) \overset{(a)}{=} \sum_{t=1}^{n} D_\alpha\bigl(W(\Psi_t(m)) \,\|\, q_{\alpha,t}\bigr) \overset{(b)}{\leqslant} \\
&\overset{(b)}{\leqslant} \sum_{t=1}^{n}\left(\ln\frac{n+1}{n} + D_\alpha(W(\Psi_t(m)) \,\|\, q_{\alpha,t}^{\epsilon})\right) \overset{(c)}{\leqslant} 1 + \sum_{t=1}^{n} D_\alpha\bigl(W(\Psi_t(m)) \,\|\, q_{\alpha,t}^{\epsilon}\bigr) \overset{(d)}{\leqslant} \\
&\overset{(d)}{\leqslant} 1 + \sum_{t=1}^{n}\frac{1}{\epsilon}\int\limits_{\alpha(1-\epsilon)}^{\alpha+(1-\alpha)\epsilon} D_\alpha\bigl(W(\Psi_t(m)) \,\|\, q_{\eta,W,\mathcal{A}}\bigr)\, d\eta \overset{(e)}{=} \\
&\overset{(e)}{=} 1 + \frac{n}{\epsilon}\int\limits_{\alpha(1-\epsilon)}^{\alpha+(1-\alpha)\epsilon} D_\alpha\bigl(W \,\|\, q_{\eta,W,\mathcal{A}} \,|\, \Upsilon(\Psi(m))\bigr)\, d\eta \overset{(f)}{\leqslant} \\
&\overset{(f)}{\leqslant} 1 + \frac{n}{\epsilon}\int\limits_{\alpha(1-\epsilon)}^{\alpha+(1-\alpha)\epsilon} \left(1 \vee \frac{1-\eta}{\eta}\frac{\alpha}{1-\alpha}\right) D_\eta\bigl(W \,\|\, q_{\eta,W,\mathcal{A}} \,|\, \Upsilon(\Psi(m))\bigr) \overset{(g)}{\leqslant} \\
&\overset{(g)}{\leqslant} 1 + \frac{n}{\epsilon}\int\limits_{\alpha(1-\epsilon)}^{\alpha+(1-\alpha)\epsilon} \left(1 \vee \frac{1-\eta}{\eta}\frac{\alpha}{1-\alpha}\right) C_{\eta,W,\mathcal{A}}\, d\eta \overset{(h)}{=} 1 + n\widetilde{C}_{\alpha,W,\mathcal{A}}^{\epsilon},
\end{aligned} \tag{48}$$

где (a) справедливо в силу [36, теорема 28], (b) – в силу [26, лемма 1], так как $\dfrac{n}{n+1}q_{\alpha,t}^{\epsilon} \leqslant q_{\alpha,t}$, (c) – в силу $\ln\tau \leqslant \tau - 1$, (d) – в силу неравенства Йенсена и [36, теорема 12], (e) – согласно определению 3 и определению композиции Υ, (f) – в силу [36, теорема 3 и предложение 2], (g) – в силу [26, теорема 1], так как $\Upsilon(\Psi(m))\in\mathcal{A}$, и (h) – по определению величины $\widetilde{C}_{\alpha,W,\mathcal{A}}^{\epsilon}$. Пусть v_α^m – скошенная вероятностная мера порядка α между w^m и q_α, определенная в (2). Тогда v_α^m имеет вид $v_\alpha^m = \bigotimes\limits_{t=1}^{n} W_\alpha^{q_{\alpha,t}}(\Psi_t(m))$ в силу структуры w^m и q_α как произведений мер. Введем случайные величины

$$\xi_{\alpha,t}^m \triangleq \ln\frac{d[W(\Psi_t(m))]_{\mathrm{ac}}}{dq_{\alpha,t}} - \mathbf{E}_{v_\alpha^m}\left[\ln\frac{d[W(\Psi_t(m))]_{\mathrm{ac}}}{dq_{\alpha,t}}\right],$$
$$\xi_\alpha^m \triangleq \sum_{t=1}^{n}\xi_{\alpha,t}^m,$$

где $[W(\Psi_t(m))]_{\mathrm{ac}}$ – компонента $W(\Psi_t(m))$, абсолютно непрерывная по $q_{\alpha,t}$. Тогда для всех $\kappa \in \mathbb{R}_+$ и $\alpha \in (0,1)$

$$\begin{aligned}
&\mathbf{E}_{v_\alpha^m}\left[|\xi_{\alpha,t}^m|^\kappa\right]^{1/\kappa} \overset{(a)}{\leqslant} 3^{1/\kappa}\frac{[(1-\alpha)D_\alpha(W(\Psi_t(m))\,\|\,q_{\alpha,t})]\vee\kappa}{\alpha(1-\alpha)} \overset{(b)}{\leqslant}\\
&\overset{(b)}{\leqslant} 3^{1/\kappa}\frac{\Big[(1-\alpha)D_\alpha\big(W(\Psi_t(m))\,\|\,q_{1/2,U_{\mathcal{A}}^{(n)}}\big)+(1-\alpha)\ln(1+n)\Big]\vee\kappa}{\alpha(1-\alpha)} \overset{(c)}{\leqslant}\\
&\overset{(c)}{\leqslant} 3^{1/\kappa}\frac{\Big[D_{1/2}\big(W(\Psi_t(m))\,\|\,q_{1/2,U_{\mathcal{A}}^{(n)}}\big)+(1-\alpha)\ln(1+n)\Big]\vee\kappa}{\alpha(1-\alpha)} \overset{(d)}{\leqslant}\\
&\overset{(d)}{\leqslant} 3^{1/\kappa}\frac{\Big[C_{1/2,U_{\mathcal{A}}^{(n)}}+(1-\alpha)\ln(1+n)\Big]\vee\kappa}{\alpha(1-\alpha)},
\end{aligned}$$

где (a) справедливо в силу [25, лемма 17], (b) – в силу [26, лемма 1], так как $\frac{q_{1/2,U_{\mathcal{A}}^{(n)}}}{n+1} \leqslant$ $\leqslant q_{\alpha,t}$, (c) – в силу [36, теорема 3 и предложение 2], а (d) – в силу [26, теорема 1], так как $\Psi_t(m) \in \mathcal{X}_{\mathcal{A}}^{(n)}$. Тогда, используя определение γ, данное в (43), получаем

$$\left[\sum_{t=1}^{n}\mathbf{E}_{v_\alpha^m}\left[|\xi_{\alpha,t}^m|^\kappa\right]\right]^{1/\kappa} \leqslant \frac{\gamma}{3\alpha(1-\alpha)}, \quad \forall\alpha\in[\phi,1). \tag{49}$$

С другой стороны, из [36, теорема 30] следует

$$D_1(v_\alpha^m\,\|\,q_\alpha) = D_\alpha(w^m\,\|\,q_\alpha) - \frac{\alpha}{1-\alpha}D_1(v_\alpha^m\,\|\,w^m), \quad \forall m\in\mathcal{M},\ \alpha\in(0,1). \tag{50}$$

Таким образом, можно оценить $D_1(v_\alpha^m\,\|\,q_\alpha)$, пользуясь неотрицательностью дивергенции Реньи (см. [36, теорема 8]) и соотношением (48):

$$0 \leqslant D_1(v_\alpha^m\,\|\,q_\alpha) \leqslant 1 + n\widetilde{C}_{\alpha,W,\mathcal{A}}^{\epsilon}.$$

Отсюда

$$\begin{aligned}
&\lim_{\alpha\downarrow\phi} D_1(v_\alpha^m\,\|\,q_\alpha) + \frac{\gamma}{1-\alpha} < \ln\frac{M}{L}\frac{\epsilon}{8e^2n^{1,5}},\\
&\lim_{\alpha\uparrow 1} D_1(v_\alpha^m\,\|\,q_\alpha) + \frac{\gamma}{1-\alpha} = \infty.
\end{aligned}$$

Функция $D_1(v_\alpha^m\,\|\,q_\alpha)$ непрерывна по α в силу [25, лемма 16], поскольку q_α непрерывна по α в топологии полной вариации[7] на $\mathcal{P}(\mathcal{Y})$. Тогда по теореме о среднем [37, теорема 4.23] для каждого $m \in \mathcal{M}$ существует $\alpha_m \in (\phi,1)$, такое что

$$\left.\left(D_1(v_\alpha^m\,\|\,q_\alpha) + \frac{\gamma}{1-\alpha}\right)\right|_{\alpha=\alpha_m} = \ln\frac{M}{L}\frac{\epsilon}{8e^2n^{1,5}}. \tag{51}$$

Для любого $K \in \mathbb{Z}_+$ в $(0,1)$ найдется замкнутый подынтервал длины $1/K$, в который попали $\left\lceil\frac{M}{K}\right\rceil$ или более таких α_m. Пусть этот интервал имеет вид $[\eta,\eta+1/K]$,

[7] В частности, $\|q_\alpha - q_\eta\| \leqslant \sqrt{8\ln\frac{\epsilon}{\epsilon-(1-\epsilon)|\alpha-\eta|}}$, так как $\|q_\alpha - q_\eta\| \leqslant \sqrt{4D_{1/2}(q_\alpha\,\|\,q_\eta)}$ согласно [36, теорема 31], $D_{1/2}(q_\alpha\,\|\,q_\eta) = nD_{1/2}(q_{\alpha,t}\,\|\,q_{\eta,t})$ в силу [36, теорема 28] и определения q_α, $D_{1/2}(q_{\alpha,t}\,\|\,q_{\eta,t}) \leqslant 2\ln\frac{2}{2-\|q_{\alpha,t}-q_{\eta,t}\|}$ согласно [26, формула (9)], и $\|q_{\alpha,t}-q_{\eta,t}\| \leqslant 2\frac{1-\epsilon}{\epsilon}|\eta-\alpha|$ по определению q_α.

и положим

$$\widetilde{q} \triangleq \bigotimes_{t=1}^{n} \widetilde{q}_t, \quad \widetilde{q}_t = \frac{n}{n+1} q^{\widetilde{\epsilon}}_{\widetilde{\alpha},t} + \frac{1}{n+1} q_{1/2,U_{\mathcal{A}}^{(n)}},$$

где

$$\widetilde{\epsilon} = \frac{1}{K} + \epsilon\Big(1 - \frac{1}{K}\Big), \quad \widetilde{\alpha} = \frac{1-\epsilon}{1-\widetilde{\epsilon}}\eta.$$

Тогда для всех $\alpha \in \Big[\eta, \eta + \dfrac{1}{K}\Big]$ согласно определению усредненного центра $q^{\epsilon}_{\alpha,t}$, данному в (47), получаем

$$q_{\alpha,t} \leqslant \frac{\widetilde{\epsilon}}{\epsilon}\widetilde{q}_t, \quad q_{\alpha} \leqslant \Big(\frac{\widetilde{\epsilon}}{\epsilon}\Big)^n \widetilde{q}.$$

По крайней мере половина сообщений, для которых $\alpha_m \in [\eta, \eta + 1/K]$, т.е. по крайней мере $\left\lceil \frac{1}{2} \left\lceil \frac{M}{K} \right\rceil \right\rceil$ сообщений, удовлетворяют условию $\widetilde{q}(\mathcal{E}_m) \leqslant 2\dfrac{L}{\lceil M/K \rceil}$ в силу неравенства Маркова, так как $\sum\limits_{m \in \widetilde{\mathcal{M}}} \widetilde{q}(\mathcal{E}_m) \leqslant L$ по определению списочного декодирования, где $\widetilde{\mathcal{M}} = \{m : \ \alpha_m \in [\eta, \eta + 1/K]\}$. Тогда по крайней мере $\left\lceil \frac{1}{2} \left\lceil \frac{M}{K} \right\rceil \right\rceil$ сообщений, для которых $\alpha_m \in [\eta, \eta + 1/K]$, удовлетворяют условию

$$q_{\alpha_m}(\mathcal{E}_m) \leqslant \frac{2LK}{M}\left(1 + \frac{1}{K}\frac{1-\epsilon}{\epsilon}\right)^n \leqslant \frac{2LK}{M}\left(1 + \frac{1}{K\epsilon}\right)^n.$$

Заметим, что $\dfrac{n}{\epsilon} > 1$, так как мы предположили, что $\epsilon < 1$ и $n \geqslant 1$. Полагая $K = \left\lfloor \dfrac{n}{\epsilon} \right\rfloor$ и применяя соотношение $(1+\tau)^{1/\tau} < e$ и (51), получаем

$$q_{\alpha_m}(\mathcal{E}_m) \leqslant \frac{2LK}{M}\left(1 + \frac{1}{K\epsilon}\right)^{K\epsilon\frac{n}{K\epsilon}} \leqslant \frac{2L}{M}\frac{n}{\epsilon}e^2 \leqslant \frac{1}{4\sqrt{n}} e^{-D_1\left(v^m_{\alpha_m} \| q_{\alpha_m}\right) - \frac{\gamma}{1-\alpha_m}}.$$

Тогда из неравенства (49) и леммы 7 следует, что

$$P^m_{\mathrm{e}} \geqslant \frac{1}{4\sqrt{n}} e^{-D_1\left(v^m_{\alpha_m} \| w^m\right) - \frac{\gamma}{\alpha_m}}. \tag{52}$$

Из (48) и (50)–(52) получаем

$$P^m_{\mathrm{e}} \geqslant \frac{1}{4\sqrt{n}} \left(\frac{\epsilon}{8e^2 n^{1,5}}\frac{M}{L}\right)^{\frac{1-\alpha_m}{\alpha_m}} e^{-\frac{1-\alpha_m}{\alpha_m}\left(1 + n\widetilde{C}^{\epsilon}_{\alpha_m,W}\right) - \frac{2\gamma}{\alpha_m}}. \tag{53}$$

Следовательно, для всех m, удовлетворяющих (51), в силу определения $\widetilde{E}^{\epsilon}_{\mathrm{sp}}(R, W, \mathcal{A})$, данного в (40), справедливо

$$P^m_{\mathrm{e}} \geqslant \frac{e^{-2\gamma/\phi}}{4\sqrt{n}} \left(\frac{\epsilon}{8e^3 n^{1,5}}\right)^{\frac{1-\phi}{\phi}} e^{-n\widetilde{E}^{\epsilon}_{\mathrm{sp}}(R,W,\mathcal{A})}, \quad R = \frac{1}{n}\ln\frac{M}{L}.$$

Поскольку таких сообщений не менее $\left\lceil \frac{1}{2} \left\lceil \frac{M}{K} \right\rceil \right\rceil$, а по построению $\left\lceil \frac{1}{2} \left\lceil \frac{M}{K} \right\rceil \right\rceil \geqslant \dfrac{M\epsilon}{2n}$, получаем нижнюю границу (44).

Заметим, что $\dfrac{\epsilon e^{-2\gamma}}{8e^3 n^{1,5}} < 1$, так как $\epsilon \in (0,1)$, $n \geqslant 1$ и $\gamma \geqslant 0$. Таким образом, коэффициент в (44) стремится к нулю при ϕ, стремящемся к нулю. Чтобы избежать этого, можно изменить рассуждения после неравенства (53), добавляя ошибку аппроксимации к скорости в члене с усредненной ЭСУ: Для всех m, удовлетворяющих (51), в силу определения $\widetilde{E}^{\epsilon}_{\rm sp}(R, W, \varrho)$, данного в (40), и неравенства (53) справедливо

$$P_{\rm e}^{m} \geqslant \frac{e^{-2\gamma}}{4\sqrt{n}} e^{-n\widetilde{E}^{\epsilon}_{\rm sp}(R,W)}, \quad R = \frac{1}{n}\ln\frac{M}{L} - \frac{2\gamma}{n} - \frac{1}{n}\ln\frac{8e^3 n^{1,5}}{\epsilon}.$$

Поскольку таких сообщений не менее $\left\lceil \frac{1}{2}\left\lceil \frac{M}{K}\right\rceil\right\rceil$, а по построению $\left\lceil \frac{1}{2}\left\lceil \frac{M}{K}\right\rceil\right\rceil \geqslant \frac{M\epsilon}{2n}$, получаем нижнюю границу (45). ▲

3.4. Неасимптотические границы сферической упаковки для кодов с ограничениями по стоимости. В п. 3.3 нас в основном интересовали стационарные каналы без памяти с выпуклыми ограничениями по композиции, а каналы без памяти с ограничениями по стоимости, рассмотренные в следствии 3, возникли просто как побочный результат. Однако, как будет показано далее, для каналов без памяти с ограничениями по стоимости можно установить ГСУ даже для нестационарных каналов, используя технику выпуклого сопряжения. Единственным недостатком использования техники выпуклого сопряжения будет то, что мы сможем установить ГСУ для ограничений по стоимости, лежащих в $\operatorname{int}\Gamma_\rho$, а не в Γ_ρ.

Лемма 10. *Пусть $n \in \mathbb{Z}_+$, $W_{[1,n]} \colon \mathcal{X}_1^n \to \mathcal{P}(\mathcal{Y}_1^n)$ – канал-произведение длины n с аддитивной функцией стоимости $\rho_{[1,n]} \colon \mathcal{X}_1^n \to \mathbb{R}^{\ell}_{\geqslant 0}$, такой что $\rho_{[1,n]}(x_1^n) = \sum\limits_{t=1}^{n} \rho_t(x_t)$ для некоторого $\rho_t \colon \mathcal{X}_t \to \mathbb{R}^{\ell}_{\geqslant 0}$, ϱ – ограничение по стоимости на одно обращение к каналу, такое что $n\varrho \in \operatorname{int}\Gamma_{\rho_{[1,n]}}$, а κ, ϕ, ϵ – произвольные положительные параметры, такие что $\kappa \geqslant 3$, $\phi < 1$, $\epsilon < 1$, и пусть*

$$\gamma \triangleq 3\left[3\sum_{t=1}^{n}\left(\left[C_{1/2,W_t,n\varrho} + (1-\phi)\ln(1+n)\right] \vee \kappa\right)^{\kappa}\right]^{1/\kappa}. \tag{54}$$

Если M и L – целые числа, такие что $\ln\frac{M}{L} > \widetilde{C}^{\epsilon}_{\phi,W_{[1,n]},n\varrho} + \frac{\gamma}{1-\phi} + \ln\frac{8e^3n^{1,5}}{\epsilon}$, то для любого (M,L)-кода в канале $W_{[1,n]}$, такого что $\bigvee\limits_{m=1}^{M} \rho_{[1,n]}(\Psi(m)) \leqslant n\varrho$, справедливо

$$P_{\rm e} \geqslant \left(\frac{\epsilon e^{-2\gamma}}{8e^3 n^{1,5}}\right)^{1/\phi} e^{-\widetilde{E}^{\epsilon}_{\rm sp}\left(\ln\frac{M}{L}, W_{[1,n]}, n\varrho\right)}. \tag{55}$$

Граница (55) не содержит члена, учитывающего скорость в усредненной ЭСУ, но она становится тривиальной, т.е. равной нулю, когда ϕ стремится к нулю. Слегка изменяя рассуждения в доказательстве, можно получить следующую альтернативную границу:

$$P_{\rm e} \geqslant \frac{\epsilon e^{-2\gamma}}{8n^{1,5}} e^{-\widetilde{E}^{\epsilon}_{\rm sp}(R, W_{[1,n]}, n\varrho)}, \quad R = \ln\frac{M}{L} - \ln\frac{8e^3n^{1,5}}{\epsilon e^{-2\gamma}}. \tag{56}$$

Границы в лемме 10 и (56) достаточно хороши для доказательства асимптотической ГСУ в теореме 2. Однако для более точного подсчета величину γ можно улучшить. Получающееся выражение, однако, включает в себя дополнительную оптимизацию.

Лемма 11. *Лемма 10 и граница* (56) *остаются справедливыми, если определение γ, данное в* (54), *заменить на следующее*:

$$\gamma \triangleq \max_{\substack{\{\varrho_t\} \colon \varrho_t \geqslant 0\ \forall t, \\ \sum_{t=1}^{n} \varrho_t \leqslant n\varrho}} 3\left[3\sum_{t=1}^{n}\left(\left[C^{\lambda}_{1/2,W_t} + \lambda\cdot\varrho_t + (1-\phi)\ln(1+n)\right] \vee \kappa\right)^{\kappa}\right]^{1/\kappa}, \tag{57}$$

где $\lambda = \lambda_{1/2,W_{[1,n]},n\varrho}$. *Если* $W_t = W$ *и* $\rho_t = \rho$ *для всех* $t \in [1,n]$, *то для* γ, *определенного в* (57), *справедливо*

$$\gamma = \max_{\substack{\{\varrho_t\}:\ \varrho_t \geqslant 0\ \forall t,\\ \sum\limits_{t=1}^{n} \varrho_t \leqslant n\varrho}} 3\left[3\sum_{t=1}^{n}\left(\left[C_{1/2,W,\varrho} + \lambda\cdot(\varrho_t-\varrho) + (1-\phi)\ln(1+n)\right]\vee\kappa\right)^{\kappa}\right]^{1/\kappa}. \quad (58)$$

Доказательство леммы 10 и формулы (56). Для каждого $\alpha \in (0,1)$ в силу [26, лемма 29(c)] существует $\lambda_\alpha \in \mathbb{R}^{\ell}_{\geqslant 0}$, такое что

$$C_{\alpha,W_{[1,n]},n\varrho} = C^{\lambda_\alpha}_{\alpha,W_{[1,n]}} + \lambda_\alpha\cdot\varrho n, \quad (59)$$

поскольку $n\varrho \in \operatorname{int}\Gamma_{\rho_{[1,n]}}$. Следовательно, $q_{\alpha,W_{[1,n]},n\varrho} = q^{\lambda_\alpha}_{\alpha,W_{[1,n]}}$ согласно [26, лемма 31]. При этом $q^{\lambda_\alpha}_{\alpha,W_{[1,n]}}$ имеет вид $q^{\lambda_\alpha}_{\alpha,W_{[1,n]}} = \bigotimes_{t=1}^{n} q^{\lambda_\alpha}_{\alpha,W_t}$ согласно [26, лемма 32]. Тогда

$$\|q^{\lambda_\alpha}_{\alpha,W_t} - q^{\lambda_\phi}_{\phi,W_t}\|^2 \overset{(a)}{\leqslant} 4D_{1/2}\big(q^{\lambda_\alpha}_{\alpha,W_t}\,\|\,q^{\lambda_\phi}_{\phi,W_t}\big) \overset{(b)}{\leqslant} 4D_{1/2}\big(q^{\lambda_\alpha}_{\alpha,W_{[1,n]}}\,\|\,q^{\lambda_\phi}_{\phi,W_{[1,n]}}\big) \overset{(c)}{=}$$
$$\overset{(c)}{=} 4D_{1/2}\big(q_{\alpha,W_{[1,n]},n\varrho}\,\|\,q_{\phi,W_{[1,n]},n\varrho}\big) \overset{(d)}{\leqslant} 8\ln\frac{2}{2-\|q_{\alpha,W_{[1,n]},\varrho} - q_{\phi,W_{[1,n]},\varrho}\|},$$

где (a) справедливо в силу [36, теорема 31], (b) – в силу [36, теоремы 8 и 28], (c) – поскольку $q_{\alpha,W_{[1,n]},n\varrho} = q^{\lambda_\alpha}_{\alpha,W_{[1,n]}}$ $\forall\alpha\in(0,1)$, и (d) – в силу [26, формула (9)]. С другой стороны, августиновский центр $q_{\alpha,W_{[1,n]},n\varrho}$ непрерывен по α на $(0,1)$ в топологии полной вариации на $\mathcal{P}(\mathcal{Y}_1^n)$ согласно [26, леммы 23(d) и 24]. Таким образом, $q^{\lambda_\alpha}_{\alpha,W_t}$ также непрерывна по α в топологии полной вариации на $\mathcal{P}(\mathcal{Y}_t)$. Тогда $q^{\lambda_\cdot}_{\cdot,W_t}$ – вероятность перехода из $((0,1),\mathcal{B}((0,1)))$ в $(\mathcal{Y}_t,\mathcal{Y}_t)$. Определим $q^{\epsilon}_{\alpha,t}$ как маргинальное распределение на $\mathcal{Y}_t$ для вероятностной меры $u_{\alpha,\epsilon} \circledast q^{\lambda_\cdot,W,\varrho}_{\cdot,W_t}$, где $u_{\alpha,\epsilon}$ – равномерное распределение вероятностей на $(\alpha-\epsilon\alpha,\alpha+\epsilon(1-\alpha))$:

$$q^{\epsilon}_{\alpha,t} = \frac{1}{\epsilon}\int\limits_{\alpha-\alpha\epsilon}^{\alpha+(1-\alpha)\epsilon} q^{\lambda_\eta}_{\eta,W_t}\,d\eta, \quad \forall t\in\{1,\ldots,n\}.$$

Положим

$$q_{\alpha,t} \triangleq \frac{n}{n+1}q^{\epsilon}_{\alpha,t} + \frac{1}{n+1}q_{1/2,W_t,n\varrho}, \quad \forall t\in\{1,\ldots,n\},$$
$$q_\alpha \triangleq \bigotimes_{t=1}^{n} q_{\alpha,t}.$$

Для краткости обозначим вероятностную меру $\bigotimes_{t=1}^{n} W_t(\Psi_t(m))$, порождаемую $\Psi(m)$, через w^m. Имеем

$$D_\alpha(w^m\,\|\,q_\alpha) \overset{(a)}{=} \sum_{t=1}^{n} D_\alpha(W_t(\Psi_t(m))\,\|\,q_{\alpha,t}) \overset{(b)}{\leqslant}$$
$$\overset{(b)}{\leqslant} \sum_{t=1}^{n}\left(\ln\frac{n+1}{n} + D_\alpha\big(W_t(\Psi_t(m))\,\|\,q^{\epsilon}_{\alpha,t}\big)\right) \overset{(c)}{\leqslant} 1 + \sum_{t=1}^{n} D_\alpha\big(W_t(\Psi_t(m))\,\|\,q^{\epsilon}_{\alpha,t}\big) \overset{(d)}{\leqslant}$$

$$\overset{(d)}{\leqslant} 1+\sum_{t=1}^{n}\frac{1}{\epsilon}\int\limits_{\alpha(1-\epsilon)}^{\alpha+(1-\alpha)\epsilon} D_\alpha\big(W_t(\Psi_t(m))\,\|\,q_{\eta,W_t}^{\lambda_\eta}\big)\,d\eta \overset{(e)}{=}$$

$$\overset{(e)}{=} 1+\frac{1}{\epsilon}\int\limits_{\alpha(1-\epsilon)}^{\alpha+(1-\alpha)\epsilon} D_\alpha\big(w^m\,\|\,q_{\eta,W_{[1,n]}}^{\lambda_\eta}\big)\,d\eta \overset{(f)}{\leqslant}$$

$$\overset{(f)}{\leqslant} 1+\frac{1}{\epsilon}\int\limits_{\alpha(1-\epsilon)}^{\alpha+(1-\alpha)\epsilon}\left(1\vee\frac{1-\eta}{\eta}\frac{\alpha}{1-\alpha}\right) D_\eta\big(w^m\,\|\,q_{\eta,W_{[1,n]}}^{\lambda_\eta}\big)\,d\eta \overset{(g)}{\leqslant}$$

$$\overset{(g)}{\leqslant} 1+\frac{1}{\epsilon}\int\limits_{\alpha(1-\epsilon)}^{\alpha+(1-\alpha)\epsilon}\left(1\vee\frac{1-\eta}{\eta}\frac{\alpha}{1-\alpha}\right)\big(C_{\eta,W}^{\lambda_\eta}+\lambda_\eta\cdot\varrho n\big)\,d\eta \overset{(h)}{=} 1+\widetilde{C}_{\alpha,W,\varrho}^{\epsilon}, \tag{60}$$

где (a) выполнено в силу [36, теорема 28], (b) – в силу [26, лемма 1], так как $\frac{n}{n+1}q_{\alpha,t}^{\epsilon}\leqslant q_{\alpha,t}$, (c) – в силу оценки $\ln\tau\leqslant\tau-1$, (d) – в силу неравенства Йенсена и [36, теорема 12], (e) – в силу [36, теорема 28] и [26, лемма 32], (f) – в силу [36, теорема 3 и предложение 2], (g) – в силу [26, теорема 2], так как $\rho_{[1,n]}(\Psi(m))\leqslant n\varrho$, и (h) – в силу (39) и (59). Пусть v_α^m – скошенная вероятностная мера порядка α между w^m и q_α, определенная в (2). Тогда v_α^m – вероятностная мера вида $v_\alpha^m = \bigotimes_{t=1}^{n}[W_t]_\alpha^{q_{\alpha,t}}(\Psi_t(m))$ в силу структуры w^m и q_α как произведений мер. Введем случайные величины

$$\xi_{\alpha,t}^m \triangleq \ln\frac{d[W_t(\Psi_t(m))]_{\mathrm{ac}}}{dq_{\alpha,t}} - \mathbf{E}_{v_\alpha^m}\left[\ln\frac{d[W_t(\Psi_t(m))]_{\mathrm{ac}}}{dq_{\alpha,t}}\right],$$

$$\xi_\alpha^m \triangleq \sum_{t=1}^{n}\xi_{\alpha,t}^m,$$

где $[W_t(\Psi_t(m))]_{\mathrm{ac}}$ – компонента $W_t(\Psi_t(m))$, абсолютно непрерывная по $q_{\alpha,t}$. Тогда для всех $\kappa\in\mathbb{R}_+$ и $\alpha\in(0,1)$

$$\mathbf{E}_{v_\alpha^m}[|\xi_{\alpha,t}^m|^\kappa]^{1/\kappa} \overset{(a)}{\leqslant} 3^{1/\kappa}\frac{[(1-\alpha)D_\alpha(W_t(\Psi_t(m))\,\|\,q_{\alpha,t})]\vee\kappa}{\alpha(1-\alpha)} \overset{(b)}{\leqslant}$$

$$\overset{(b)}{\leqslant} 3^{1/\kappa}\frac{\big[(1-\alpha)D_\alpha\big(W_t(\Psi_t(m))\,\|\,q_{1/2,W_t,n\varrho}\big)+(1-\alpha)\ln(1+n)\big]\vee\kappa}{\alpha(1-\alpha)} \overset{(c)}{\leqslant}$$

$$\overset{(c)}{\leqslant} 3^{1/\kappa}\frac{\big[D_{1/2}\big(W_t(\Psi_t(m))\,\|\,q_{1/2,W_t,n\varrho}\big)+(1-\alpha)\ln(1+n)\big]\vee\kappa}{\alpha(1-\alpha)} \overset{(d)}{\leqslant}$$

$$\overset{(d)}{\leqslant} 3^{1/\kappa}\frac{\big[C_{1/2,W_t,n\varrho}+(1-\alpha)\ln(1+n)\big]\vee\kappa}{\alpha(1-\alpha)},$$

где (a) справедливо в силу [25, лемма 17], (b) – в силу [26, лемма 1], так как $\frac{q_{1/2,W_t,n\varrho}}{n+1}\leqslant q_{\alpha,t}$, (c) – в силу [36, теорема 3 и предложение 2], и (d) – в силу [26, теорема 1], так как $\rho_t(\Psi_t(m))\leqslant n\varrho$. Тогда, используя определение γ, данное в (54), получаем

$$\left[\sum_{t=1}^{n}\mathbf{E}_{v_\alpha^m}\left[|\xi_{\alpha,t}^m|^\kappa\right]\right]^{1/\kappa}\leqslant\frac{\gamma}{3\alpha(1-\alpha)},\quad \forall\alpha\in[\phi,1). \tag{61}$$

Оставшаяся часть доказательства аналогична доказательству леммы 9 и границы (45) после формулы (49). Единственные отличия, заслуживающие упоминания, состоят в том, что вместо неравенств (48) и (49) используются (60) и (61), а вместо $\widetilde{q}_t = \dfrac{n}{n+1} q^{\widetilde{\epsilon}}_{\alpha,t} + \dfrac{1}{n+1} q_{1/2,U_{\mathcal{A}}^{(n)}}$ используется $\widetilde{q}_t = \dfrac{n}{n+1} q^{\widetilde{\epsilon}}_{\alpha,t} + \dfrac{1}{n+1} q_{1/2,W_t,\varrho}$. ▲

Доказательство леммы 11 совпадает с доказательством леммы 10 и границы (56) за исключением определения $q_{\alpha,t}$ и оценки для $\mathbf{E}_{v^m_\alpha}\left[|\xi^m_{\alpha,t}|^\kappa\right]^{1/\kappa}$. В частности, мы полагаем $q_{\alpha,t} \in \mathcal{P}(\mathcal{Y}_t)$ равной $\dfrac{n}{n+1} q^{\epsilon}_{\alpha,t} + \dfrac{1}{n+1} q^{\lambda_{1/2}}_{1/2,W_t}$ для всех $t \leqslant n$ и оцениваем $\mathbf{E}_{v^m_\alpha}\left[|\xi^m_{\alpha,t}|^\kappa\right]^{1/\kappa}$ следующим образом:

$$
\begin{aligned}
&\mathbf{E}_{v^m_\alpha}\left[|\xi^m_{\alpha,t}|^\kappa\right]^{1/\kappa} \overset{(a)}{\leqslant} 3^{1/\kappa} \frac{[(1-\alpha) D_\alpha(W_t(\Psi_t(m)) \,\|\, q_{\alpha,t})] \vee \kappa}{\alpha(1-\alpha)} \overset{(b)}{\leqslant} \\
&\overset{(b)}{\leqslant} 3^{1/\kappa} \frac{\left[(1-\alpha) D_\alpha\big(W_t(\Psi_t(m)) \,\|\, q^{\lambda_{1/2}}_{1/2,W_t}\big) + (1-\alpha)\ln(1+n)\right] \vee \kappa}{\alpha(1-\alpha)} \overset{(c)}{\leqslant} \\
&\overset{(c)}{\leqslant} 3^{1/\kappa} \frac{\left[D_{1/2}\big(W_t(\Psi_t(m)) \,\|\, q^{\lambda_{1/2}}_{1/2,W_t}\big) + (1-\alpha)\ln(1+n)\right] \vee \kappa}{\alpha(1-\alpha)} \overset{(d)}{\leqslant} \\
&\overset{(d)}{\leqslant} 3^{1/\kappa} \frac{\left[C^{\lambda_{1/2}}_{1/2,W_t} + \lambda_{1/2} \cdot \rho_t(\Psi_t(m)) + (1-\alpha)\ln(1+n)\right] \vee \kappa}{\alpha(1-\alpha)},
\end{aligned}
$$

где (a) справедливо в силу [25, лемма 17], (b) – в силу [26, лемма 1], так как $\dfrac{q^{\lambda_{1/2}}_{1/2,W_t}}{n+1} \leqslant$ $\leqslant q_{\alpha,t}$, (c) – в силу [36, теорема 3 и предложение 2], а (d) – в силу [26, теорема 2]. Тогда, используя определение γ, данное в (57), получаем

$$
\left[\sum_{t=1}^{n} \mathbf{E}_{v^m_\alpha}\left[|\xi^m_{\alpha,t}|^\kappa\right]\right]^{1/\kappa} \leqslant \frac{\gamma}{3\alpha(1-\alpha)}, \quad \forall \alpha \in [\phi, 1).
$$

Оставшаяся часть доказательства совпадает с доказательством леммы 10 и границы (56).

Если $W_t = W$ и $\rho_t = \rho$ для всех $t \in [1,n]$, то $C_{\alpha,W_{[1,n]},n\varrho} = nC_{\alpha,W,\varrho}$ для всех $\varrho \in \Gamma_\rho$ и $C^\lambda_{\alpha,W_{[1,n]}} = nC^\lambda_{\alpha,W}$ для всех $\lambda \in \mathbb{R}^\ell_{\geqslant 0}$ согласно [26, леммы 28 и 32]. Тогда (58) следует из равенства $C_{\alpha,W,\varrho} = C^{\lambda_{\alpha,W,\varrho}}_{\alpha,W} + \lambda_{\alpha,W,\varrho} \cdot \varrho$, установленного в [26, лемма 29(c)]. ▲

§ 4. Примеры

Из результатов п. 2.4 и § 3 можно сделать вывод, что ЭСУ управляет экспоненциальной скоростью убывания вероятности ошибки для кодов со списочным декодированием в каналах без памяти при довольно широких предположениях. Однако вычисление самой ЭСУ – это отдельный вопрос, весьма существенный с практической точки зрения. В этом параграфе мы выведем ЭСУ для различных гауссовских и пуассоновских каналов и покажем, что для этих каналов можно получать ее в параметрическом виде, аналогично тому, как сделано в лемме 2 для $E_{\mathrm{sp}}(R, W, p)$. Мы считаем, что эти параметрические выражения проще и интуитивно понятнее, чем обычно используемые равносильные выражения, полученные ранее.

4.1. Гауссовские каналы. Обозначим плотность распределения гауссовской случайной величины с нулевым средним и дисперсией σ^2 через φ_{σ^2}, т.е.

$$
\varphi_{\sigma^2}(z) \triangleq \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{z^2}{2\sigma^2}}, \quad \forall z \in \mathbb{R}.
$$

Допуская некоторую вольность в обозначениях, соответствующую вероятностную меру на $\mathcal{B}(\mathbb{R})$ будем также обозначать через φ_{σ^2}.

Пример 1 (скалярный гауссовский канал). Пусть W – скалярный гауссовский канал с дисперсией шума σ^2, и пусть соответствующая функция стоимости ρ квадратична:

$$W(\mathcal{E}\,|\,x) = \int\limits_{\mathcal{E}} \varphi_{\sigma^2}(y-x)\,dy, \quad \forall \mathcal{E} \in \mathcal{B}(\mathbb{R}),$$
$$\rho(x) = x^2, \quad \forall x \in \mathbb{R}.$$

Августиновские пропускная способность и центр с ограничениями по стоимости для этого канала вычислены в [26, пример 4]:

$$C_{\alpha,W,\varrho} = \\ = \begin{cases} \dfrac{\alpha\varrho}{2(\alpha\theta_{\alpha,\sigma,\varrho}+(1-\alpha)\sigma^2)} + \dfrac{1}{\alpha-1}\ln\dfrac{(\theta_{\alpha,\sigma,\varrho})^{\alpha/2}\sigma^{(1-\alpha)}}{\sqrt{\alpha\theta_{\alpha,\sigma,\varrho}+(1-\alpha)\sigma^2}}, & \alpha \in \mathbb{R}_+ \setminus \{1\}, \\ \dfrac{1}{2}\ln\Big(1+\dfrac{\varrho}{\sigma^2}\Big), & \alpha = 1, \end{cases} \tag{62}$$

$$q_{\alpha,W,\varrho} = \varphi_{\theta_{\alpha,\sigma,\varrho}}, \tag{63}$$

$$\theta_{\alpha,\sigma,\varrho} \triangleq \sigma^2 + \frac{\varrho}{2} - \frac{\sigma^2}{2\alpha} + \sqrt{\Big(\frac{\varrho}{2} - \frac{\sigma^2}{2\alpha}\Big)^2 + \varrho\sigma^2}. \tag{64}$$

Следует отметить, что $C_{\alpha,W,\varrho} = I_\alpha(\varphi_\varrho; W)$ и $q_{\alpha,W,\varrho} = q_{\alpha,\varphi_\varrho}$ для всех положительных порядков α, т.е. гауссовское распределение с нулевым средним и дисперсией ϱ является оптимальным распределением на входе для всех порядков. Таким образом, $E_{\mathrm{sp}}(R,W,\varrho) = E_{\mathrm{sp}}(R,W,\varphi_\varrho)$.

ЭСУ скалярного гауссовского канала можно охарактеризовать с помощью леммы 4. Для этого вначале заметим, что для любого $\theta > 0$ и соответствующей гауссовской вероятностной меры φ_θ скошенный канал $W_\alpha^{\varphi_\theta}$ порядка α, определенный в (3), имеет вид

$$W_\alpha^{\varphi_\theta}(\mathcal{E}\,|\,x) = \int\limits_{\mathcal{E}} \varphi_{\frac{\sigma^2\theta}{\alpha\theta+(1-\alpha)\sigma^2}}\Big(y - \frac{\alpha\theta}{\alpha\theta+(1-\alpha)\sigma^2}x\Big)\,dy, \quad \forall \mathcal{E} \in \mathcal{B}(\mathbb{R}). \tag{65}$$

Так как $\theta_{\alpha,\sigma,\varrho}$ является корнем уравнения $\theta^2 - \theta\Big[\varrho + \Big(2-\dfrac{1}{\alpha}\Big)\sigma^2\Big] + \Big(1-\dfrac{1}{\alpha}\Big)\sigma^4 = 0$ относительно неизвестной θ согласно [26, формулы (132) и (133)], то используя [26, формула (131)], можно непосредственной подстановкой убедиться, что

$$D_1\big(W_\alpha^{\varphi_{\theta_{\alpha,\sigma,\varrho}}} \,\|\, \varphi_{\theta_{\alpha,\sigma,\varrho}} \,|\, p\big) = \\ = \frac{\alpha^2\theta_{\alpha,\sigma,\varrho}}{2(\alpha\theta_{\alpha,\sigma,\varrho}+(1-\alpha)\sigma^2)^2}(\mathbf{E}_p[\rho]-\varrho) + \frac{1}{2}\ln\frac{\alpha\theta_{\alpha,\sigma,\varrho}+(1-\alpha)\sigma^2}{\sigma^2}, \tag{66}$$

$$D_1\big(W_\alpha^{\varphi_{\theta_{\alpha,\sigma,\varrho}}} \,\|\, W \,|\, p\big) = \frac{(1-\alpha)^2\sigma^2}{2(\alpha\theta_{\alpha,\sigma,\varrho}+(1-\alpha)\sigma^2)^2}(\mathbf{E}_p[\rho]-\varrho) + \\ + \frac{(1-\alpha)\varrho}{2(\alpha\theta_{\alpha,\sigma,\varrho}+(1-\alpha)\sigma^2)} + \frac{1}{2}\ln\frac{\alpha\theta_{\alpha,\sigma,\varrho}+(1-\alpha)\sigma^2}{\theta_{\alpha,\sigma,\varrho}}. \tag{67}$$

Таким образом, для любого $\alpha^* \in (0,1)$ канал $V_p = W_{\alpha^*}^{\varphi_{\theta_{\alpha^*,\sigma,\varrho}}}$ удовлетворяет условиям (21) и (22) леммы 4 для $R = \dfrac{1}{2}\ln\dfrac{\alpha^*\theta_{\alpha,\sigma,\varrho}+(1-\alpha^*)\sigma^2}{\sigma^2}$ в силу (62) и ограничения

$\mathbf{E}_p[\rho] \leqslant \varrho$. Кроме того, функция $f(\alpha) \triangleq \frac{1}{2} \ln \frac{\alpha\theta_{\alpha,\sigma,\varrho} + (1-\alpha)\sigma^2}{\sigma^2}$ непрерывна и возрастает по α, причем $\lim_{\alpha \downarrow 0} f(\alpha) = 0$ и $f(1) = C_{1,W,\varrho}$. Таким образом, ЭСУ можно записать в следующем параметрическом виде через $\alpha \in [0,1]$ для всех скоростей из $[0, C_{1,W,\varrho}]$:

$$R = \frac{1}{2} \ln \frac{\alpha\theta_{\alpha,\sigma,\varrho} + (1-\alpha)\sigma^2}{\sigma^2}, \tag{68}$$

$$E_{\mathrm{sp}}(R, W, \varrho) = \frac{(1-\alpha)\varrho}{2(\alpha\theta_{\alpha,\sigma,\varrho} + (1-\alpha)\sigma^2)} + \frac{1}{2} \ln \frac{\alpha\theta_{\alpha,\sigma,\varrho} + (1-\alpha)\sigma^2}{\theta_{\alpha,\sigma,\varrho}}. \tag{69}$$

С помощью (66) и (67) можно выразить как скорость, так и ЭСУ через скошенный канал $W_\alpha^{\varphi_{\theta_{\alpha,\sigma,\varrho}}}$. С другой стороны, $\varphi_{\theta_{\alpha,\sigma,\varrho}}$ является распределением на выходе для распределения на входе φ_ϱ в канале $W_\alpha^{\varphi_{\theta_{\alpha,\sigma,\varrho}}}$, поскольку $\varphi_{\theta_{\alpha,\sigma,\varrho}}$ является также августиновским средним $q_{\alpha,\varphi_\varrho}$ для распределения на входе φ_ϱ, удовлетворяющего свойству неподвижной точки $\mathrm{T}_{\alpha,\varphi_\varrho}(q_{\alpha,\varphi_\varrho}) = q_{\alpha,\varphi_\varrho}$. Таким образом, используя (66) и (67), можно переписать (68) и (69) следующим образом:

$$R = I_1\bigl(\varphi_\varrho; W_\alpha^{\varphi_{\theta_{\alpha,\sigma,\varrho}}}\bigr), \tag{70}$$

$$E_{\mathrm{sp}}(R, W, \varrho) = D_1\bigl(W_\alpha^{\varphi_{\theta_{\alpha,\sigma,\varrho}}} \,\|\, W \,|\, \varphi_\varrho\bigr). \tag{71}$$

Чтобы получить выражение для ЭСУ, не зависящее от $\theta_{\alpha,\sigma,\varrho}$ явно, заметим вначале, что из (64) и (68) следует

$$\alpha = \frac{e^{2R}-1}{2}\left(\sqrt{1 + \frac{4\sigma^2}{\varrho}\frac{e^{2R}}{e^{2R}-1}} - 1\right). \tag{72}$$

С другой стороны, $\varphi_{\theta_{\alpha,\sigma,\varrho}}$ является распределением на выходе в канале $W_\alpha^{\varphi_{\theta_{\alpha,\sigma,\varrho}}}$ при распределении на входе φ_ϱ. Таким образом, из (65) получаем

$$\frac{\sigma^2\theta_{\alpha,\sigma,\varrho}}{\alpha\theta_{\alpha,\sigma,\varrho} + (1-\alpha)\sigma^2} + \left(\frac{\alpha\theta_{\alpha,\sigma,\varrho}}{\alpha\theta_{\alpha,\sigma,\varrho} + (1-\alpha)\sigma^2}\right)^2 \varrho = \theta_{\alpha,\sigma,\varrho}.$$

Значит,

$$\frac{\sigma^2}{\theta_{\alpha,\sigma,\varrho}} = 1 - \frac{\alpha\varrho}{\alpha\theta_{\alpha,\sigma,\varrho} + (1-\alpha)\sigma^2} = 1 - \frac{\varrho\alpha}{\sigma^2 e^{2R}}, \tag{73}$$

где последнее равенство следует из (68).

Используя в (69) вначале (68) и (73), а затем (72), получаем следующее выражение для ЭСУ:

$$\begin{aligned} E_{\mathrm{sp}}(R, W, \varrho) &= \frac{1}{2}\frac{(1-\alpha)\varrho}{\sigma^2 e^{2R}} + R + \frac{1}{2}\ln\Bigl(1 - \frac{\alpha\varrho}{\sigma^2 e^{2R}}\Bigr) = \\ &= \frac{\varrho}{4\sigma^2}\left[1 + \frac{1}{e^{2R}} - \Bigl(1 - \frac{1}{e^{2R}}\Bigr)\sqrt{1 + \frac{4\sigma^2}{\varrho}\frac{e^{2R}}{e^{2R}-1}}\right] + \\ &+ \frac{1}{2}\ln\left[e^{2R} - \frac{\varrho}{\sigma^2}\left(\frac{e^{2R}-1}{2}\right)\left(\sqrt{1 + \frac{4\sigma^2}{\varrho}\frac{e^{2R}}{e^{2R}-1}} - 1\right)\right]. \end{aligned} \tag{74}$$

Выражение для ЭСУ, полученное в (74), равносильно формуле (7.4.33) из [8].

Параметрическую характеризацию (68), (69) можно также получить с помощью более прямого подхода, используя дифференцируемость $\theta_{\alpha,\sigma,\varrho}$ и $C_{\alpha,W,\varrho}$ по α. В частности, поскольку $\theta_{\alpha,\sigma,\varrho}$ – корень уравнения

$$\theta^2 - \theta\Big[\varrho + \Big(2 - \frac{1}{\alpha}\Big)\sigma^2\Big] + \Big(1 - \frac{1}{\alpha}\Big)\sigma^4 = 0$$

относительно неизвестной θ, получаем следующее выражение в замкнутом виде для производной августиновской пропускной способности по порядку:

$$\begin{aligned}
&\frac{\partial}{\partial\alpha}C_{\alpha,W,\varrho} = \frac{1}{2(1-\alpha)^2}\left[\frac{(1-\alpha)\varrho}{\alpha\theta_{\alpha,\sigma,\varrho}+(1-\alpha)\sigma^2} + \ln\frac{\alpha\theta_{\alpha,\sigma,\varrho}+(1-\alpha)\sigma^2}{\theta_{\alpha,\sigma,\varrho}}\right] + \\
&+ \frac{\theta^2_{\alpha,\sigma,\varrho} - \theta_{\alpha,\sigma,\varrho}\Big[\varrho + \Big(2-\frac{1}{\alpha}\Big)\sigma^2\Big] + \Big(1-\frac{1}{\alpha}\Big)\sigma^4}{2(\alpha\theta_{\alpha,\sigma,\varrho}+(1-\alpha)\sigma^2)^2}\left[\frac{\alpha}{1-\alpha} + \frac{\alpha^2}{\theta_{\alpha,\sigma,\varrho}}\frac{\partial}{\partial\alpha}\theta_{\alpha,\sigma,\varrho}\right] = \\
&= \frac{1}{2(1-\alpha)^2}\left[\frac{(1-\alpha)\varrho}{\alpha\theta_{\alpha,\sigma,\varrho}+(1-\alpha)\sigma^2} + \ln\frac{\alpha\theta_{\alpha,\sigma,\varrho}+(1-\alpha)\sigma^2}{\theta_{\alpha,\sigma,\varrho}}\right].
\end{aligned} \tag{75}$$

Используя вначале (75), а затем (62), получаем

$$\begin{aligned}
&\frac{d}{d\alpha}\frac{1-\alpha}{\alpha}(C_{\alpha,W,\varrho} - R) = \\
&= \frac{1-\alpha}{\alpha}\left(\frac{1}{2(1-\alpha)^2}\left[\frac{(1-\alpha)\varrho}{\alpha\theta_{\alpha,\sigma,\varrho}+(1-\alpha)\sigma^2} + \ln\frac{\alpha\theta_{\alpha,\sigma,\varrho}+(1-\alpha)\sigma^2}{\theta_{\alpha,\sigma,\varrho}}\right]\right) - \\
&- \frac{1}{\alpha^2}(C_{\alpha,W,\varrho} - R) = \frac{1}{\alpha^2}\left(R - \frac{1}{2}\ln\frac{\alpha\theta_{\alpha,\sigma,\varrho}+(1-\alpha)\sigma^2}{\sigma^2}\right).
\end{aligned}$$

Теперь проверка знака производной приводит к параметрическому выражению, приведенному в (68) и (69), в силу (62).

Пример 2 (параллельные гауссовские каналы). Пусть $W_{[1,n]}$ – произведение скалярных гауссовских каналов с дисперсией шума σ_i^2 для $i \in \{1,\dots,n\}$, и пусть функция стоимости $\rho_{[1,n]}$ аддитивна и квадратична, т.е.

$$W_{[1,n]}(\mathcal{E}\,|\,x_1^n) = \int\limits_{\mathcal{E}}\left[\prod_{i=1}^n \varphi_{\sigma_i^2}(y_i - x_i)\right]dy_1^n, \quad \forall\mathcal{E}\in\mathcal{B}(\mathbb{R}^n),$$

$$\rho_{[1,n]}(x_1^n) = \sum_{i=1}^n x_i^2, \quad \forall x_1^n \in \mathbb{R}^n.$$

Августиновские пропускная способность и центр канала $W_{[1,n]}$ с ограничениями по стоимости были вычислены в [26, пример 5]:

$$C_{\alpha,W_{[1,n]},\varrho} = \sum_{i=1}^n C_{\alpha,W_i,\varrho_{\alpha,i}}, \tag{76}$$

$$q_{\alpha,W_{[1,n]},\varrho} = \bigotimes_{i=1}^n \varphi_{\theta_{\alpha,\sigma_i,\varrho_{\alpha,i}}}, \tag{77}$$

$$\varrho_{\alpha,i} = \frac{\big|\alpha - 2\sigma_i^2\lambda_\alpha\big|^+}{2\lambda_\alpha\big(\alpha + 2(\alpha-1)\sigma_i^2\lambda_\alpha\big)}, \tag{78}$$

где $\theta_{\alpha,\sigma,\varrho}$ определено в (64), а λ_α однозначно определяется[8] условием $\sum\limits_i \varrho_{\alpha,i} = \varrho$. Кроме того, $\theta_{\alpha,\sigma_i,\varrho_{\alpha,i}}$ можно выразить через σ_i и λ_α, не используя $\varrho_{\alpha,i}$ явно, следующим образом:

$$\theta_{\alpha,\sigma_i,\varrho_{\alpha,i}} = \sigma_i^2 + \left|\frac{1}{2\lambda_\alpha} - \frac{\sigma_i^2}{\alpha}\right|^+. \tag{79}$$

С другой стороны, $\left.\frac{d}{d\varrho_i} C_{\alpha,W_i,\varrho_i}\right|_{\varrho_i=\varrho_{\alpha,i}} = \lambda_\alpha$ для всех i с положительными $\varrho_{\alpha,i}$, и $\left.\frac{d}{d\varrho_i} C_{\alpha,W_i,\varrho_i}\right|_{\varrho_i=\varrho_{\alpha,i}} \leqslant \lambda_\alpha$ для всех i. Вычисляя производную сложной функции с учетом (75) и (62), получаем

$$\begin{aligned}
&\frac{d}{d\alpha}\frac{1-\alpha}{\alpha}\big(C_{\alpha,W_{[1,n]},\varrho} - R\big) = \\
&= \frac{1-\alpha}{\alpha}\sum_{i=1}^n \left[\left.\frac{\partial}{\partial\phi} C_{\phi,W_i,\varrho_{\alpha,i}}\right|_{\phi=\alpha} + \left.\frac{\partial}{\partial\varrho} C_{\alpha,W_i,\varrho}\right|_{\varrho=\varrho_{\alpha,i}} \frac{d}{d\alpha}\varrho_{\alpha,i}\right] - \frac{1}{\alpha^2}\big(C_{\alpha,W_{[1,n]},\varrho} - R\big) = \\
&= \frac{1}{\alpha^2}\left[R - \frac{1}{2}\sum_{i=1}^n \ln\frac{\alpha\theta_{\alpha,\sigma_i,\varrho_{\alpha,i}} + (1-\alpha)\sigma_i^2}{\sigma_i^2}\right] + \frac{1-\alpha}{\alpha}\lambda_\alpha \sum_{i=1}^n \frac{d}{d\alpha}\varrho_{\alpha,i} = \\
&= \frac{1}{\alpha^2}\left[R - \frac{1}{2}\sum_{i=1}^n \ln\frac{\alpha\theta_{\alpha,\sigma_i,\varrho_{\alpha,i}} + (1-\alpha)\sigma_i^2}{\sigma_i^2}\right].
\end{aligned}$$

Итак, получаем следующее параметрическое выражение для $E_{\mathrm{sp}}(R, W_{[1,n]}, \varrho)$ через α для всех $R \in [0, C_{\alpha,W_{[1,n]},\varrho}]$:

$$R = \frac{1}{2}\sum_{i=1}^n \ln\frac{\alpha\theta_{\alpha,\sigma_i,\varrho_{\alpha,i}} + (1-\alpha)\sigma_i^2}{\sigma_i^2}, \tag{80}$$

$$E_{\mathrm{sp}}(R, W_{[1,n]}, \varrho) = \frac{1}{2}\sum_{i=1}^n \left[\frac{(1-\alpha)\varrho_{\alpha,i}}{\alpha\theta_{\alpha,\sigma_i,\varrho_{\alpha,i}} + (1-\alpha)\sigma_i^2} + \ln\frac{\alpha\theta_{\alpha,\sigma_i,\varrho_{\alpha,i}} + (1-\alpha)\sigma_i^2}{\theta_{\alpha,\sigma_i,\varrho_{\alpha,i}}}\right]. \tag{81}$$

Таким образом, $E_{\mathrm{sp}}(R, W_{[1,n]}, \varrho) = D_1(V_\alpha \,\|\, W_{[1,n]} \,|\, \varPhi_\alpha)$ при $R = I_1(\varPhi_\alpha; V_\alpha)$, где распределение на входе $\varPhi_\alpha$ – гауссовское распределение с нулевым средним с диагональной ковариационной матрицей с собственными значениями $\varrho_{\alpha,1}, \ldots, \varrho_{\alpha,n}$, а V_α – скошенный канал порядка α между $W_{[1,n]}$ и $q_{\alpha,\varPhi_\alpha}$.

Если $\sigma_i^2 \geqslant \frac{\alpha}{2\lambda_\alpha}$ для некоторого i, то $\varrho_{\alpha,i} = 0$, и соответствующие члены в суммах (80) и (81) нулевые. В [19] с помощью этого наблюдения получено альтернативное параметрическое выражение для ЭСУ. Чтобы вывести его, отметим вначале, что

$$\alpha\theta_{\alpha,\sigma_i,\varrho_{\alpha,i}} + (1-\alpha)\sigma_i^2 = \frac{\alpha}{2\lambda_\alpha} \vee \sigma_i^2, \quad \forall i \in \{1, \ldots, n\}$$

согласно (79). Поэтому из (79)–(81) и ограничения $\sum\limits_i \varrho_{\alpha,i} = \varrho$ вытекает следующее параметрическое выражение через $N = \frac{\alpha}{2\lambda_\alpha}$, равносильное выражению из [18] (см.

[8] Ограничение $\sum\limits_i \varrho_{\alpha,i} = \varrho$ задает λ_α однозначно, так как выражение в правой части (78) – невозрастающая функция λ_α.

[8, формулы (7.5.28), (7.5.32), и (7.5.34)], [18, с. 294], [19, формула (20)]):

$$R = \frac{1}{2} \sum_{i\colon \sigma_i^2 \leqslant N} \ln \frac{N}{\sigma_i^2}, \tag{82}$$

$$\varrho = \frac{1}{\alpha} \sum_{i=1}^{n} \frac{\left|N - \sigma_i^2\right|^+}{1 + (\alpha - 1)\dfrac{\sigma_i^2}{N}}, \tag{83}$$

$$E_{\mathrm{sp}}(R, W_{[1,n]}, \varrho) = \frac{(1-\alpha)\varrho}{2N} + \frac{1}{2} \sum_{i\colon \sigma_i^2 \leqslant N} \ln \frac{\alpha N}{N - (1-\alpha)\sigma_i^2}. \tag{84}$$

Заметим, что в этих выражениях не требуется знать значение λ_α или использовать равенство $N = \dfrac{\alpha}{2\lambda_\alpha}$. Как отмечено в [18,19], чтобы вычислить $E_{\mathrm{sp}}(R, W_{[1,n]}, \varrho)$, можно сперва определить N с помощью (82), а затем α с помощью (83).

4.2. Пуассоновские каналы. Пусть $T \in \mathbb{R}_+$ и $a, b \in \mathbb{R}_{\geqslant 0}$, причем $a \leqslant b$. Тогда для пуассоновского канала $\varLambda\colon \mathcal{F} \to \mathcal{P}(\mathcal{Y})$ множество входов $\mathcal{F}$ – это множество всех измеримых функций вида $f\colon (0,T] \to [a,b]$, множество выходов $\mathcal{Y}$ – множество всех неубывающих непрерывных справа функций на $(0,T]$ с целыми значениями, σ-алгебра выходных событий $\mathcal{Y}$ – борелевская σ-алгебра для топологии, порожденной метрикой Скорохода на $\mathcal{Y}$, а $\varLambda(f)$ – пуассоновский точечный процесс с детерминированной функцией интенсивности f для всех $f \in \mathcal{F}$. Допуская некоторую вольность в обозначениях, пуассоновский процесс с постоянной интенсивностью γ будем обозначать через $\varLambda(\gamma)$. Функция стоимости $\rho\colon \mathcal{F} \to \mathbb{R}_{\geqslant 0}$ имеет вид

$$\rho(f) \triangleq \frac{1}{T} \int_0^T f(t)\, dt.$$

В [27, раздел V-C] вычислены пропускные способности и центры Реньи (при отсутствии ограничений) для различных пуассоновских каналов. Эти выражения равны соответствующим августиновским пропускным способностям и центрам, поскольку $C^0_{\alpha,W} = C^{\mathrm{g}0}_{\alpha,W}$ для любого W и $q^0_{\alpha,W} = q^{\mathrm{g}0}_{\alpha,W}$ для любого W с конечной $C^0_{\alpha,W}$ согласно [26, теоремы 2 и 3].

Пример 3 (пуассоновские каналы с заданной средней интенсивностью). В [27, пример 9] рассмотрены каналы $\varLambda^\varrho\colon \mathcal{F}^\varrho \to \mathcal{P}(\mathcal{Y})$, где $\mathcal{F}^\varrho = \{f \in \mathcal{F} : \rho(f) = \varrho\}$ и $\varLambda^\varrho(f) = \varLambda(f)$ для всех $f \in \mathcal{F}^\varrho$. Пропускная способность и центр Реньи канала $\varLambda^\varrho$ – а следовательно, его августиновские пропускная способность и центр – приведены в [27, формулы (73) и (74)]:

$$C_{\alpha,\varLambda^\varrho} = \begin{cases} \dfrac{\alpha}{\alpha-1}(\zeta_{\alpha,\varrho} - \varrho)T, & \alpha \neq 1, \\ \left(\dfrac{\varrho - a}{b-a} b \ln \dfrac{b}{\varrho} + \dfrac{b-\varrho}{b-a} a \ln \dfrac{a}{\varrho}\right) T, & \alpha = 1, \end{cases} \tag{85}$$

$$q_{\alpha,\varLambda^\varrho} = \varLambda(\zeta_{\alpha,\varrho}), \tag{86}$$

$$\zeta_{\alpha,\varrho} \triangleq \left(\frac{\varrho - a}{b-a} b^\alpha + \frac{b-\varrho}{b-a} a^\alpha\right)^{1/\alpha}. \tag{87}$$

Так как выражение для $C_{\alpha,\varLambda^\varrho}$ дифференцируемо по α, то вычисляя знак производной, получаем следующее параметрическое выражение для ЭСУ:

$$R = \left(\frac{\varrho - a}{b-a} b^\alpha \zeta_{\alpha,\varrho}^{1-\alpha} \ln \frac{b^\alpha \zeta_{\alpha,\varrho}^{1-\alpha}}{\zeta_{\alpha,\varrho}} + \frac{b-\varrho}{b-a} a^\alpha \zeta_{\alpha,\varrho}^{1-\alpha} \ln \frac{a^\alpha \zeta_{\alpha,\varrho}^{1-\alpha}}{\zeta_{\alpha,\varrho}}\right) T, \tag{88}$$

$$E_{\mathrm{sp}}(R, \Lambda^\varrho) = \\ = \left(\varrho - \zeta_{\alpha,\varrho} + \frac{\varrho - a}{b-a} b^\alpha \zeta_{\alpha,\varrho}^{1-\alpha} \ln \frac{b^\alpha \zeta_{\alpha,\varrho}^{1-\alpha}}{b} + \frac{b-\varrho}{b-a} a^\alpha \zeta_{\alpha,\varrho}^{1-\alpha} \ln \frac{a^\alpha \zeta_{\alpha,\varrho}^{1-\alpha}}{a}\right) T. \tag{89}$$

Имеется альтернативное параметрическое выражение, которое гораздо легче представить, используя скошенные каналы. Чтобы его вывести, заметим вначале, что из [27, формулы (66) и (68)] и определения скошенного канала, данного в (3), следует, что

$$\Lambda_\alpha^{\Lambda(g)}(f) = \Lambda\big(f^\alpha g^{(1-\alpha)}\big). \tag{90}$$

Тогда, используя [27, формула (68)] и (86)–(89), получаем

$$R = D_1\big(\Lambda_\alpha^{q_{\alpha,\Lambda^\varrho}}(f_{\mathrm{opt}}) \,\|\, q_{\alpha,\Lambda^\varrho}\big), \tag{91}$$

$$E_{\mathrm{sp}}(R, \Lambda^\varrho) = D_1\big(\Lambda_\alpha^{q_{\alpha,\Lambda^\varrho}}(f_{\mathrm{opt}}) \,\|\, \Lambda(f_{\mathrm{opt}})\big), \tag{92}$$

где f_{opt} – любая функция из $\mathcal{F}^\varrho$ со значениями $\{a, b\}$, т.е. любая функция вида $f_{\mathrm{opt}}\colon (0, T] \to \{a, b\}$, такая что $\rho(f_{\mathrm{opt}}) = \varrho$.

П р и м е р 4 (пуассоновские каналы с ограничениями на среднюю интенсивность). Убедимся вначале, что августиновские пропускная способность $C_{\alpha,\Lambda,\varrho}$ и центр $q_{\alpha,\Lambda,\varrho}$ при наличии ограничений имеют вид[9]

$$C_{\alpha,\Lambda,\varrho} = C_{\alpha,\Lambda^{\varrho\wedge\varrho_\alpha}}, \tag{93}$$

$$q_{\alpha,\Lambda,\varrho} = q_{\alpha,\Lambda^{\varrho\wedge\varrho_\alpha}}, \tag{94}$$

где $C_{\alpha,\Lambda^\varrho}$ и $q_{\alpha,\Lambda^\varrho}$ определены в (85)–(87), а ϱ_α – убывающая функция порядка α, имеющая вид

$$\varrho_\alpha \triangleq \begin{cases} \alpha^{\frac{\alpha}{1-\alpha}} \left(\dfrac{b-a}{b^\alpha - a^\alpha}\right)^{\frac{1}{1-\alpha}} + \dfrac{ab^\alpha - ba^\alpha}{b^\alpha - a^\alpha}, & \alpha \neq 1, \\ e^{-1} b^{\frac{b}{b-a}} a^{-\frac{a}{b-a}}, & \alpha = 1. \end{cases} \tag{95}$$

Чтобы доказать равенства (93) и (94), заметим вначале, что $C_{\alpha,\Lambda^{\varrho\wedge\varrho_\alpha}} \leqslant C_{\alpha,\Lambda,\varrho}$, так как $C_{\alpha,\Lambda^{\varrho\wedge\varrho_\alpha}} = C_{\alpha,\Lambda,\mathcal{P}(\mathcal{F}^{\varrho\wedge\varrho_\alpha})}$ и $\mathcal{P}(\mathcal{F}^{\varrho\wedge\varrho_\alpha}) \subset \mathcal{A}(\varrho)$, где $\mathcal{A}(\varrho) = \{p \in \mathcal{P}(\mathcal{F}) : \mathbf{E}_p[\rho(f)] \leqslant \varrho\}$. С другой стороны, применяя вначале [27, формула (82)], а затем [27, формула (76)], получаем

$$D_\alpha\big(\Lambda(f) \,\|\, \Lambda(\zeta_{\alpha,\varrho\wedge\varrho_\alpha})\big) \leqslant \frac{b - \rho(f)}{b-a} D_\alpha\big(\Lambda(a) \,\|\, \Lambda(\zeta_{\alpha,\varrho\wedge\varrho_\alpha})\big) + \\ + \frac{\rho(f) - a}{b-a} D_\alpha\big(\Lambda(b) \,\|\, \Lambda(\zeta_{\alpha,\varrho\wedge\varrho_\alpha})\big) = C_{\alpha,\Lambda^{\varrho\wedge\varrho_\alpha}} + \\ + \frac{\varrho\wedge\varrho_\alpha - \rho(f)}{b-a} \left[D_\alpha\big(\Lambda(a) \,\|\, \Lambda(\zeta_{\alpha,\varrho\wedge\varrho_\alpha})\big) - D_\alpha\big(\Lambda(b) \,\|\, \Lambda(\zeta_{\alpha,\varrho\wedge\varrho_\alpha})\big)\right], \quad \forall f \in \mathcal{F}.$$

---

[9] Заметим, что $C_{\alpha,\Lambda,\varrho} = C_{\alpha,\Lambda^{\leqslant\varrho}}$ и $q_{\alpha,\Lambda,\varrho} = q_{\alpha,\Lambda^{\leqslant\varrho}}$ для пуассоновского канала $\Lambda^{\leqslant\varrho}\colon \{f \in \mathcal{F} : \rho(f) \leqslant \varrho\} \to \mathcal{P}(\mathcal{Y})$, рассмотренного в [27, пример 10].

Поскольку $D_\alpha(\Lambda(a) \,\|\, \Lambda(\zeta_{\alpha,\varrho_\alpha})) = D_\alpha(\Lambda(b) \,\|\, \Lambda(\zeta_{\alpha,\varrho_\alpha}))$ в силу [27, формулы (83) и (84)], получаем

$$D_\alpha\big(\Lambda(f) \,\|\, \Lambda(\zeta_{\alpha,\varrho\wedge\varrho_\alpha})\big) \leqslant C_{\alpha,\Lambda^{\varrho\wedge\varrho_\alpha}} + \\ + \mathbb{1}_{\{\varrho<\varrho_\alpha\}} \frac{\varrho - \rho(f)}{b-a} \left[D_\alpha\big(\Lambda(a) \,\|\, \Lambda(\zeta_{\alpha,\varrho})\big) - D_\alpha\big(\Lambda(b) \,\|\, \Lambda(\zeta_{\alpha,\varrho})\big)\right], \quad \forall f \in \mathcal{F}.$$

С другой стороны, $D_\alpha(\Lambda(a) \,\|\, \Lambda(\zeta_{\alpha,\varrho})) \leqslant D_\alpha(\Lambda(b) \,\|\, \Lambda(\zeta_{\alpha,\varrho}))$ для всех $\varrho \leqslant \varrho_\alpha$ согласно [27, формула (83)], и следовательно,

$$D_\alpha\big(\Lambda(f) \,\|\, \Lambda(\zeta_{\alpha,\varrho\wedge\varrho_\alpha}) \,|\, p\big) \leqslant C_{\alpha,\Lambda^{\varrho\wedge\varrho_\alpha}}, \quad \forall p:\ \mathbf{E}_p[\rho] \leqslant \varrho.$$

Таким образом, соотношения (93) и (94) вытекают из [26, лемма 25], поскольку $\Lambda(\zeta_{\alpha,\varrho\wedge\varrho_\alpha}) = q_{\alpha,\Lambda^{\varrho\wedge\varrho_\alpha}}$ согласно (86).

Так как выражение для $C_{\alpha,\Lambda,\varrho}$, данное в (93), дифференцируемо по α, то вычисляя знак производной, можно найти оптимальный порядок для ЭСУ, заданной в (10). В результате получаем следующее параметрическое выражение:

$$R = \left(\frac{\varrho\wedge\varrho_\alpha - a}{b-a} b^\alpha \zeta_{\alpha,\varrho\wedge\varrho_\alpha}^{1-\alpha} \ln \frac{b^\alpha \zeta_{\alpha,\varrho\wedge\varrho_\alpha}^{1-\alpha}}{\zeta_{\alpha,\varrho\wedge\varrho_\alpha}} + \right. \\ \left. + \frac{b - \varrho\wedge\varrho_\alpha}{b-a} a^\alpha \zeta_{\alpha,\varrho\wedge\varrho_\alpha}^{1-\alpha} \ln \frac{a^\alpha \zeta_{\alpha,\varrho\wedge\varrho_\alpha}^{1-\alpha}}{\zeta_{\alpha,\varrho\wedge\varrho_\alpha}}\right) T, \tag{96}$$

$$E_{\mathrm{sp}}(R,\Lambda,\varrho) = \left(\varrho\wedge\varrho_\alpha - \zeta_{\alpha,\varrho\wedge\varrho_\alpha} + \frac{\varrho\wedge\varrho_\alpha - a}{b-a} b^\alpha \zeta_{\alpha,\varrho\wedge\varrho_\alpha}^{1-\alpha} \ln \frac{b^\alpha \zeta_{\alpha,\varrho\wedge\varrho_\alpha}^{1-\alpha}}{b} + \right. \\ \left. + \frac{b - \varrho\wedge\varrho_\alpha}{b-a} a^\alpha \zeta_{\alpha,\varrho\wedge\varrho_\alpha}^{1-\alpha} \ln \frac{a^\alpha \zeta_{\alpha,\varrho\wedge\varrho_\alpha}^{1-\alpha}}{a}\right) T. \tag{97}$$

С учетом [27, формула (68)], (86), (90), (94), (96) и (97) получаем следующую параметрическую характеризацию:

$$R = D_1\big(\Lambda_\alpha^{q_{\alpha,\Lambda,\varrho}}(f_\alpha) \,\|\, q_{\alpha,\Lambda,\varrho}\big), \tag{98}$$

$$E_{\mathrm{sp}}(R,\Lambda,\varrho) = D_1\big(\Lambda_\alpha^{q_{\alpha,\Lambda,\varrho}}(f_\alpha) \,\|\, \Lambda(f_\alpha)\big), \tag{99}$$

где f_α – любая функция из $\mathcal{F}^{\varrho\wedge\varrho_\alpha}$ со значениями $\{a,b\}$, т.е. любая функция вида $f_\alpha\colon (0,T] \to \{a,b\}$, такая что $\rho(f_\alpha) = \varrho\wedge\varrho_\alpha$.

§ 5. Обсуждение

Мы применили метод Августина для вывода ГСУ для двух семейств каналов без памяти. Для стационарных каналов без памяти с выпуклыми ограничениями по композиции новым в методе Августина является следующее наблюдение:

$$\lim_{\phi\to\alpha} \sup_{p\in\mathcal{A}} D_\alpha(W \,\|\, q_{\phi,W,\mathcal{A}} \,|\, p) = C_{\alpha,W,\mathcal{A}}, \quad \forall \alpha \in (0,1). \tag{100}$$

Заметим, что результаты, полученные для выпуклых стационарных каналов без памяти с ограничениями по композиции справедливы также и для стационарных каналов без памяти с ограничениями по стоимости, так как любое ограничение по стоимости для стационарного канала без памяти можно также представить в виде выпуклого ограничения по композиции. Для нестационарных каналов без памяти с ограничениями по стоимости мы применяли соотношение (100) в сочетании с техникой выпуклого сопряжения. Теорема 2 улучшает аналогичный результат Августина, т.е. [24, теорема 36.6], благодаря членам, учитывающим ошибку аппроксимации. Коэффициент при экспоненте в теореме 2 имеет вид $e^{-O(\ln n)}$, а не $e^{-O(\sqrt{n})}$, как в [24, теорема 36.6]. Кроме того, в отличие от [24, теорема 36.6], в теореме 2 функции стоимости не предполагаются ограниченными, благодаря чему эта теорема справедлива и для гауссовских моделей, рассматриваемых в [17–20].

ГСУ для классически-квантовых каналов была установлена в [40]. После этого прорыва интерес к ГСУ для классически-квантовых каналов возобновился [41–44]. Однако метод Августина не применялся ни к каким квантовым теоретико-информационным моделям. Успешное использование метода Августина позволило бы, по меньшей мере, избавиться от предположений стационарности и конечности множества входов.

Августиновский вариант границы Галлагера, обсуждаемый в п. 2.4, известен мало. Нашей основной целью в п. 2.4 было изложить этот подход в простейшем виде. Так, при оценке $\mathbf{E}[P_{\mathrm{e}}]$ мы довольствовались переходом от (29) к (30). При более тщательном анализе, оценивая уклонение случайной величины, равной информации Августина – Лежандра порядка α, т.е. $\ln \dfrac{f_x}{h_x}$ для $q = q_{\alpha,p}$, можно получить более точные оценки, аналогичные оценкам из [45–47]. Граница случайного кодирования для классически-квантовых каналов была получена в [48]. В [42, с. 5606] было высказано предположение, что тот факт, что рассуждения в [48] опираются на коды с независимыми одинаково распределенными символами, может препятствовать модификации доказательства на коды с постоянной композицией, т.е. с ограничениями по композиции. Мы считаем, что августиновский вариант границы Галлагера может оказаться полезным для преодоления этой трудности.

В качестве стороннего замечания укажем, что из результатов п. 2.4 и § 3 следует, что ЭСУ является функцией надежности для определенных каналов с замиранием, т.е. каналов с состояниями, при условии, что разрешено списочное декодирование, даже в нестационарном случае. В частности, как каналы с быстрым замиранием и отсутствием информации о состоянии (т.е. со статистической информацией о состоянии), так и каналы с быстрым замиранием и информацией о состоянии, известной только на приемном конце, являются каналами без памяти с ограничениями по стоимости. Таким образом, для каналов, рассмотренных в [20], а также для каналов, рассмотренных в [49, § 4], функция надежности при списочном декодировании равна ЭСУ. Это также справедливо и для моделей с ограничением по мощности на каждую антенну, рассматриваемых в [50–52], поскольку такие каналы являются каналами без памяти с ограничениями по стоимости, хотя и со многими ограничениями. Однако, как отмечено в § 4, определение ЭСУ для таких каналов является отдельным вопросом.

Оптимальный коэффициент при экспоненте в ГСУ был известен для отдельных каналов с самых истоков теории информации (см., например, [3, 4, 17]). В последние годы интерес к установлению таких точных ГСУ при различных предположениях симметрии возобновился [5, 6, 13, 21, 22]. Параметрические выражения для ЭСУ, данные в (13) и (14), использовались в [14] для вывода подобных границ для кодов с постоянной композицией. Как показано в [28], используя аналогичные параметрические характеризации, улучшенную ГСУ можно получить в каждом случае из рассмотренных в [3–6, 13, 17, 21, 22]. Это является одной из причин, по которой мы представили параметрические выражения, данные в (70), (71), (91), (92), (98) и (99). Оказывается (см. [53]), что с помощью аналогичных параметрических характеризаций при соответствующих предположениях симметрии можно также улучшать сильные обращения теоремы кодирования в смысле коэффициентов при экспоненте.

В работе Блейхута [31] выводилась нижняя граница вероятности ошибки кодов без использования рассуждений с постоянной композицией. Блейхут утверждал, что экспоненциальная скорость убывания этой границы равна ЭСУ – см. [31, теорема 19]. В других публикациях Блейхута также утверждалось равенство вышеуказанных экспонент и ЭСУ, см. [32, лемма 1] и [33, теорема 10.1.4]. Мы покажем, что для для Z-канала экспонента в границе Блейхута равна бесконечности для любых скоростей, меньших пропускной способности канала. Следовательно, и [31, теорема 19], и [32, лемма 1], и [33, теорема 10.1.4] неверны. Что более важно, мы покажем, что даже наилучшая граница, которую можно получить методом Блейхута, строго хуже, чем ГСУ, в смысле экспоненциальной скорости убывания с ростом длины блока.

Для любых $W\colon \mathcal{X} \to \mathcal{P}(\mathcal{Y})$ и $R \in [0, C_{1,W}]$ положим

$$G(R,W,p,q) \triangleq \inf_{V\colon D_1(V \,\|\, q \,|\, p) \leqslant R} D_1(V \,\|\, W \,|\, p), \quad \forall p \in \mathcal{P}(\mathcal{X}),\ q \in \mathcal{P}(\mathcal{Y}).$$

Поскольку $D_1(V \,\|\, (pV) \,|\, p) = I_1(p;V)$, где $(pV) = \sum_x p(x)V(x)$, из альтернативного выражения для $E_{\mathrm{sp}}(R,W)$, данного в (20), следует, что

$$E_{\mathrm{sp}}(R,W) = \sup_{p \in \mathcal{P}(\mathcal{X})} \inf_{q \in \mathcal{P}(\mathcal{Y})} G(R,W,p,q), \quad \forall R \in [0, C_{1,W}]. \tag{A.1}$$

Таким образом, из неравенства максимина-минимакса вытекает

$$E_{\mathrm{sp}}(R,W) \leqslant \inf_{q \in \mathcal{P}(\mathcal{Y})} \sup_{p \in \mathcal{P}(\mathcal{X})} G(R,W,p,q), \quad \forall R \in [0, C_{1,W}].$$

Начальная часть доказательства теоремы 19 в [31] устанавливает следующую границу на вероятность ошибки для кодов в стационарном канале-произведении с каналами-компонентами W, у которых множество входов $\mathcal{X}$ и множество выходов $\mathcal{Y}$ конечны:

$$P_{\mathrm{e}}^{\max,n} \geqslant O(1) e^{-o(n) - n \sup\limits_{p \in \mathcal{P}(\mathcal{X})} G(R,W,p,q)}, \quad \forall q \in \mathcal{P}(\mathcal{Y}). \tag{A.2}$$

Во второй части доказательства [31, теорема 19] утверждается, что величина $\sup\limits_{p \in \mathcal{P}(\mathcal{X})} G(R,W,p,q_{\alpha_R,p_R})$ равна $E_{\mathrm{sp}}(R,W)$ для некоторой пары (α_R, p_R), удовлетворяющей следующим равенствам[10]:

$$E_{\mathrm{sp}}(R,W) = E_{\mathrm{sp}}(R,W,p_R) = \frac{1-\alpha_R}{\alpha_R}(I_{\alpha_R}(p_R;W) - R).$$

С учетом равенства (A.1) во второй части доказательства [31, теорема 19] утверждается, что

$$\sup_{p \in \mathcal{P}(\mathcal{X})} G(R,W,p,q_{\alpha_R,p_R}) \stackrel{?}{=} \sup_{p \in \mathcal{P}(\mathcal{X})} \inf_{q \in \mathcal{P}(\mathcal{Y})} G(R,W,p,q). \tag{A.3}$$

Заметим, что с учетом (A.1) из [32, лемма 1] и [33, теорема 10.1.4] также вытекает это равенство. Чтобы опровергнуть равенство (A.3), рассмотрим Z-канал, т.е. дискретный канал с множеством входов $\mathcal{X} = \{1,2\}$ и множеством выходов $\mathcal{Y} = \{a,b\}$, такой что

$$W = \begin{bmatrix} 1 & 0 \\ \varepsilon & 1-\varepsilon \end{bmatrix}.$$

[10] В явном виде Блейхут упоминает только первое равенство.

Вычислим августиновские пропускную способность и центр порядка α с помощью тождеств $C_{\alpha,W} = D_\alpha(W(1) \,\|\, q_{\alpha,W})$ и $C_{\alpha,W} = D_\alpha(W(2) \,\|\, q_{\alpha,W})$:

$$C_{\alpha,W} = \ln\left(1 + \left(\frac{1-\varepsilon^\alpha}{(1-\varepsilon)^\alpha}\right)^{\frac{1}{1-\alpha}}\right), \tag{A.4}$$

$$q_{\alpha,W}(a) = \frac{(1-\varepsilon)^{\frac{\alpha}{1-\alpha}}}{(1-\varepsilon)^{\frac{\alpha}{1-\alpha}} + (1-\varepsilon^\alpha)^{\frac{1}{1-\alpha}}}. \tag{A.5}$$

Тогда для $q = q_{\alpha,W}$ скошенный канал W_α^q, определенный в (3), имеет вид

$$W_\alpha^{q_{\alpha,W}} = \begin{bmatrix} 1 & 0 \\ \varepsilon^\alpha & 1-\varepsilon^\alpha \end{bmatrix}.$$

Кроме того, непосредственной подстановкой можно убедиться, что августиновский центр $q_{\alpha,W}$ порядка α является неподвижной точкой оператора Августина порядка α, определенного в (4), для априорного распределения $p_{\alpha,W}$, такого что

$$p_{\alpha,W}(1) = \frac{(1-\varepsilon)^{\frac{\alpha}{1-\alpha}} - \varepsilon^\alpha(1-\varepsilon^\alpha)^{\frac{\alpha}{1-\alpha}}}{(1-\varepsilon)^{\frac{\alpha}{1-\alpha}} + (1-\varepsilon^\alpha)^{\frac{1}{1-\alpha}}}.$$

Таким образом, августиновский центр $q_{\alpha,W}$ порядка α равен августиновскому среднему порядка α с априорным распределением $p_{\alpha,W}$, т.е. $q_{\alpha,W} = q_{\alpha,p_{\alpha,W}}$, в силу леммы [26, лемма 13(c),(d)] и равенства $I_\alpha(p_{\alpha,W}; W) = D_\alpha(W \,\|\, q_{\alpha,W} \,|\, p_{\alpha,W})$. Следовательно, также $I_\alpha(p_{\alpha,W}; W) = C_{\alpha,W}$.

Заметим, что функция $C_{\alpha,W}$, данная в (A.4), дифференцируема по α, причем

$$\frac{\partial}{\partial\alpha} C_{\alpha,W} = \frac{q_\alpha(b)}{(1-\alpha)^2}\left(\ln\frac{1-\varepsilon^\alpha}{1-\varepsilon} + \frac{\varepsilon^\alpha}{1-\varepsilon^\alpha}\ln\frac{\varepsilon^\alpha}{\varepsilon}\right) = \\ = \frac{1}{(1-\alpha)^2} D_1\big(W_\alpha^{q_{\alpha,W}} \,\|\, W \,|\, p_{\alpha,W}\big).$$

Тогда, применяя вначале тождество $I_\alpha(p_{\alpha,W}; W) = C_{\alpha,W}$, а затем [26, формула (35)], получаем

$$\frac{\partial}{\partial\alpha}\frac{1-\alpha}{\alpha}(C_{\alpha,W} - R) = \frac{1}{\alpha^2}\left(R - C_{\alpha,W} + (1-\alpha)\alpha\frac{\partial}{\partial\alpha}C_{\alpha,W}\right) = \\ = \frac{1}{\alpha^2}\left(R - I_\alpha(p_{\alpha,W}; W) + \frac{\alpha}{1-\alpha}D_1\big(W_\alpha^{q_{\alpha,W}} \,\|\, W \,|\, p_{\alpha,W}\big)\right) = \\ = \frac{1}{\alpha^2}\left(R - I_1\big(p_{\alpha,W}; W_\alpha^{q_{\alpha,W}}\big)\right).$$

Можно проверить численно, что $I_1(p_{\alpha,W}; W_\alpha^{q_{\alpha,W}})$ – возрастающая по α функция для любого $\varepsilon \in (0,1)$. Таким образом, можно выразить скорость и соответствующую ЭСУ в следующем параметрическом виде:

$$R(\alpha) = I_1\big(p_{\alpha,W}; W_\alpha^{q_{\alpha,W}}\big), \quad E_{\mathrm{sp}}(R(\alpha), W) = D_1\big(W_\alpha^{q_{\alpha,W}} \,\|\, W \,|\, p_{\alpha,W}\big).$$

Таким образом, для $R = R(\phi)$ получаем $(\alpha_R, p_R) = (\phi, p_{\phi,W})$ и $q_{\alpha_R,p_R} = q_{\phi,W}$. Тогда

$$\sup_{p\in\mathcal{P}(\mathcal{X})} G(R, W, p, q_{\alpha_R,p_R}) \geqslant G(R, W, p, q_{\alpha_R,W})\big|_{p:\, p(1)=1} = \\ = \inf_{v:\, D_1(v \| q_{\alpha_R,W}) \leqslant R} D_1(v \,\|\, W(1)).$$

С другой стороны, $v(b) > 0$ для всех v, таких что $D_1(v \,\|\, q_{\alpha_R,W}) \leqslant R$, поскольку если $v(b) = 0$ и $C_{\alpha_R,W} > R$, то $D_1(v \,\|\, q_{\alpha_R,W}) = \ln \dfrac{1}{q_{\alpha_R,W}(a)} = C_{\alpha_R,W}$. Кроме того, $D_1(v \,\|\, W(1)) = \infty$, если $v(b) > 0$. Итак,

$$\inf_{v\colon D_1(v \| q_{\alpha_R,W}) \leqslant R} D_1(v \,\|\, W(1)) = \infty.$$

Таким образом, граница, установленная в первой части доказательства [31, теорема 19], т.е. (A.2), тривиальна для $q = q_{\alpha_R,p_R}$. Кроме того, и [31, теорема 19], и [32, лемма 1], и [33, теорема 10.1.4] неверны, поскольку $E_{\rm sp}(R,W) < \infty$ для всех $R \in [0, C_{1,W}]$ в случае Z-каналов[11].

Выбирая другие q, из (A.2) можно получать нетривиальные границы, однако из них также нельзя вывести ГСУ. Если $q(a) < e^{-R}$, то $\sup\limits_{p \in \mathcal{P}(\mathcal{X})} G(R,W,p,q_{\alpha_R,p_R}) = \infty$ в силу приведенных выше аргументов. В случае $q(a) \geqslant e^{-R}$ получаем следующее:

$$\begin{aligned}
&\sup_{p\in\mathcal{P}(\mathcal{X})} G(R,W,p,q) \geqslant \inf_{V\colon D_1(V\|q|p_R)\leqslant R} D_1(V \,\|\, W \,|\, p_R) \geqslant \\
&\geqslant \inf_V D_1(V \,\|\, W \,|\, p_R) + \frac{1-\alpha_R}{\alpha_R}(D_1(V \,\|\, q \,|\, p_R) - R) \overset{(a)}{=} \\
&\overset{(a)}{=} \frac{1-\alpha_R}{\alpha_R}(D_{\alpha_R}(W \,\|\, q \,|\, p_R) - R) \overset{(b)}{\geqslant} \frac{1-\alpha_R}{\alpha_R}(I_{\alpha_R}(p_R;W) + D_{\alpha_R}(q_{\alpha_R,p_R} \,\|\, q) - R) = \\
&= E_{\rm sp}(R,W) + \frac{1-\alpha_R}{\alpha_R} D_{\alpha_R}(q_{\alpha_R,p_R} \,\|\, q) \overset{(c)}{\geqslant} E_{\rm sp}(R,W) + \frac{1-\alpha_R}{2}\|q - q_{\alpha_R,p_R}\|^2,
\end{aligned}$$

где (a) справедливо в силу [36, теорема 30], (b) – в силу [26, лемма 13(c)], а (c) – в силу [36, теорема 31]. С другой стороны, $q_{\alpha_R,p_R}(a) < e^{-R}$, поскольку $q_{\alpha_R,p_R} = q_{\alpha_R,W}$ и $C_{\alpha_R,W} > R$. Итак,

$$\begin{aligned}
&\inf_{q\in\mathcal{P}(\mathcal{Y})} \sup_{p\in\mathcal{P}(\mathcal{X})} G(R,W,p,q) \geqslant E_{\rm sp}(R,W) + 2(1-\alpha_R)(e^{-R} - q_{\alpha_R,W}(a))^2 = \\
&= \sup_{p\in\mathcal{P}(\mathcal{X})} \inf_{q\in\mathcal{P}(\mathcal{Y})} G(R,W,p,q) + 2(1-\alpha_R)(e^{-R} - q_{\alpha_R,W}(a))^2.
\end{aligned}$$

Следовательно, методом Блейхута в том виде, в каком как он приведен в [31], вывести ГСУ невозможно. Когда множество входов конечно, эту трудность можно преодолеть, используя технику выбрасывания с учетом композиции. Но такой подход применялся Арутюняном в [10] еще до работы [31].

ПРИЛОЖЕНИЕ B: *ИНФОРМАЦИОННЫЕ ВЕЛИЧИНЫ РЕНЬИ – ГАЛЛАГЕРА*

Информационные величины Реньи – Галлагера порядка 1 равны соответствующим информационным величинам Августина – Лежандра порядка 1 по определению. Поэтому в дальнейшем обсуждении мы ограничиваемся порядками, отличными от единицы.

Определение 12. Для любых $\alpha \in \mathbb{R}_+ \setminus \{1\}$, канала $W\colon \mathcal{X} \to \mathcal{P}(\mathcal{Y})$ с функцией стоимости $\rho\colon \mathcal{X} \to \mathbb{R}^{\ell}_{\geqslant 0}$, $p \in \mathcal{P}(\mathcal{X})$ и $\lambda \in \mathbb{R}^{\ell}_{\geqslant 0}$ *информацией Реньи – Галлагера поряд-*

[11] В недавней работе [54] отмечалось, что метод Блейхута можно использовать для вывода ГСУ, если выполнено минимаксное равенство [54, формула (3.63)]. Таким образом, из нашего анализа вытекает, что равенство [54, формула (3.63)] в общем случае выполнено не всегда. Этот факт можно также вывести из отсутствия минимаксного равенства для $G(R,W,p,q)$, не опираясь на результаты работы [54].

ка α для распределения на входе p и множителя Лагранжа λ называется величина

$$I_\alpha^{\mathrm{g}\lambda}(p;W) \triangleq \inf_{q\in\mathcal{P}(\mathcal{Y})} D_\alpha\big(p \circledast We^{\frac{1-\alpha}{\alpha}\lambda\cdot\rho} \,\|\, p\otimes q\big).$$

Если λ – нулевой вектор, то информация Реньи – Галлагера равна информации Реньи. Как и для информации Реньи, для информации Реньи – Галлагера имеется выражение в замкнутом виде через среднюю меру, достигающую инфимума в этом определении.

Определение 13. Для любых $\alpha\in\mathbb{R}_+$, канала $W\colon \mathcal{X}\to\mathcal{P}(\mathcal{Y})$ с функцией стоимости $\rho\colon \mathcal{X}\to\mathbb{R}^\ell_{\geqslant 0}$, $p\in\mathcal{P}(\mathcal{X})$ и $\lambda\in\mathbb{R}^\ell_{\geqslant 0}$ *средней мерой порядка α для распределения на входе p и множителя Лагранжа λ* называется

$$\frac{d\mu_{\alpha,p}^\lambda}{d\nu} \triangleq \left[\sum_x p(x)e^{(1-\alpha)\lambda\cdot\rho(x)}\left(\frac{dW(x)}{d\nu}\right)^\alpha\right]^{\frac{1}{\alpha}}. \tag{B.1}$$

Средним Реньи – Галлагера порядка α для распределения на входе p и множителя Лагранжа λ называется

$$q_{\alpha,p}^{\mathrm{g}\lambda} \triangleq \frac{\mu_{\alpha,p}^\lambda}{\|\mu_{\alpha,p}^\lambda\|}.$$

Как $\mu_{\alpha,p}^\lambda$, так и $q_{\alpha,p}^{\mathrm{g}\lambda}$ зависят от множителя Лагранжа λ при $\alpha\in\mathbb{R}_+\setminus\{1\}$. Кроме того, непосредственной подстановкой можно убедиться, что

$$D_\alpha\big(p \circledast We^{\frac{1-\alpha}{\alpha}\lambda\cdot\rho} \,\|\, p\otimes q\big) = I_\alpha^{\mathrm{g}\lambda}(p;W) + D_\alpha\big(q_{\alpha,p}^{\mathrm{g}\lambda} \,\|\, q\big), \quad \alpha\in\mathbb{R}_+\setminus\{1\}.$$

Тогда в силу [26, лемма 2] имеем

$$I_\alpha^{\mathrm{g}\lambda}(p;W) = D_\alpha\big(p \circledast We^{\frac{1-\alpha}{\alpha}\lambda\cdot\rho} \,\|\, p\otimes q_{\alpha,W}^{\mathrm{g}\lambda}\big) = \frac{\alpha}{\alpha-1}\ln\|\mu_{\alpha,p}^\lambda\|, \quad \alpha\in\mathbb{R}_+\setminus\{1\}. \tag{B.2}$$

Используя определения информации Августина – Лежандра и информации Реньи – Галлагера, с учетом неравенства Йенсена и вогнутости функции натурального логарифма получаем

$$I_\alpha^\lambda(p;W) \geqslant I_\alpha^{\mathrm{g}\lambda}(p;W), \quad \alpha\in(0,1], \tag{B.3}$$

$$I_\alpha^\lambda(p;W) \leqslant I_\alpha^{\mathrm{g}\lambda}(p;W), \quad \alpha\in[1,\infty). \tag{B.4}$$

Определение 14. Для любых $\alpha\in\mathbb{R}_+$, канала $W\colon \mathcal{X}\to\mathcal{P}(\mathcal{Y})$ с функцией стоимости $\rho\colon \mathcal{X}\to\mathbb{R}^\ell_{\geqslant 0}$ и $\lambda\in\mathbb{R}^\ell_{\geqslant 0}$ *пропускной способностью Реньи – Галлагера порядка α для множителя Лагранжа λ* называется величина

$$C_{\alpha,W}^{\mathrm{g}\lambda} \triangleq \sup_{p\in\mathcal{P}(\mathcal{X})} I_\alpha^{\mathrm{g}\lambda}(p;W). \tag{B.5}$$

Несмотря на то, что неравенства (B.3) и (B.4) являются строгими для большинства распределений, согласно [26, теорема 3] справедливо равенство

$$C_{\alpha,W}^{\mathrm{g}\lambda} = S_{\alpha,W}^\lambda. \tag{B.6}$$

Таким образом, $C_{\alpha,W}^{\mathrm{g}\lambda} = C_{\alpha,W}^\lambda$ в силу (9). По этой причине при определении оптимальной производительности можно использовать любое из этих семейств. Следующая лемма по существу является переформулировкой теоремы 8 из [29], благодаря которой стало популярно использование информационных величин Реньи – Галлагера в задачах с ограничениями по стоимости (см. [18–20]).

Лемма 12. *Для любых* $\ell, M, L \in \mathbb{Z}_+$, *таких что* $L < M$, $W\colon \mathcal{X} \to \mathcal{P}(\mathcal{Y})$, $\rho\colon \mathcal{X} \to \mathbb{R}^{\ell}_{\geqslant 0}$, $p \in \mathcal{P}(\mathcal{X})$, $\mathcal{B} \subset \mathcal{X}$ *и* $\alpha \in \left[\frac{1}{1+L}, 1\right)$ *существует* (M,L)*-код с функцией кодирования вида* $\Psi\colon \mathcal{M} \to \mathcal{B}$, *для которого*

$$\ln P_{\mathrm{e}} \leqslant \frac{\alpha - 1}{\alpha}\left[I^{\mathrm{g}\lambda}_{\alpha}(p;W) + \Big(\inf_{x\in\mathcal{B}} \lambda\cdot\rho(x)\Big) - \ln\frac{(M-1)e}{L}\right] - \frac{\ln p(\mathcal{B})}{\alpha}. \tag{B.7}$$

Доказательство леммы 12. Следуем доказательству леммы 6 вплоть до соотношения (30). В силу (26) и (30) имеем

$$\begin{aligned}
\ln \mathbf{E}[P_{\mathrm{e}}] &\leqslant \ln \mathbf{E}_q\left[\left(\sum_{x\in\mathcal{B}} \frac{p(x)}{p(\mathcal{B})} e^{(1-\alpha)\lambda\cdot\rho(x)}\left(\frac{dW(x)}{dq}\right)^{\alpha}\right)^{\frac{1}{\alpha}}\right] + \\
&+ \frac{\alpha-1}{\alpha}\left[\Big(\inf_{x\in\mathcal{B}} \lambda\cdot\rho(x)\Big) - \ln\frac{(M-1)e}{L}\right] \leqslant \\
&\leqslant \ln \mathbf{E}_q\left[\left(\sum_{x\in\mathcal{X}} p(x) e^{(1-\alpha)\lambda\cdot\rho(x)}\left(\frac{dW(x)}{dq}\right)^{\alpha}\right)^{\frac{1}{\alpha}}\right] + \\
&+ \frac{\alpha-1}{\alpha}\left[\Big(\inf_{x\in\mathcal{B}} \lambda\cdot\rho(x)\Big) - \ln\frac{(M-1)e}{L}\right] - \frac{\ln p(\mathcal{B})}{\alpha}.
\end{aligned}$$

Поскольку существует код с P_{e}, меньшей или равной $\mathbf{E}[P_{\mathrm{e}}]$, существование кода с функцией кодирования вида $\Psi\colon \mathcal{M} \to \mathcal{B}$, удовлетворяющего условию (B.7), вытекает из (B.1) и (B.2). ▲

ПРИЛОЖЕНИЕ C: ОПУЩЕННЫЕ ДОКАЗАТЕЛЬСТВА

Доказательство леммы 1. Функция $E_{\mathrm{sp}}(R, W, \mathcal{A})$ выпукла по R, поскольку поточечный супремум семейства выпуклых функций является выпуклым, а функция $\frac{1-\alpha}{\alpha}(C_{\alpha,W,\mathcal{A}} - R)$ выпукла по R для любого $\alpha \in (0,1)$. Аналогично доказывается, что $E_{\mathrm{sp}}(R, W, \mathcal{A})$ не возрастает по R. Непрерывность и конечность устанавливаются далее при доказательстве соотношения (11).

Напомним, что $C_{\alpha,W,\mathcal{A}}$ – неубывающая функция порядка α согласно [26, лемма 23(a)].

- Если $C_{0^+,W,\mathcal{A}} = \infty$, то $C_{1/2,W} = \infty$, и $E_{\mathrm{sp}}(R, W, \mathcal{A}) = \infty$ для всех $R \in \mathbb{R}_{\geqslant 0}$. С другой стороны, $R < C_{0^+,W,\mathcal{A}}$ для всех $R \in \mathbb{R}_{\geqslant 0}$. Следовательно, (11) выполнено, а утверждения о непрерывности и конечности $E_{\mathrm{sp}}(R, W, \mathcal{A})$ бессодержательны.
- Если $C_{0^+,W,\mathcal{A}} < \infty$ и $C_{0^+,W,\mathcal{A}} = C_{1,W,\mathcal{A}}$, то $E_{\mathrm{sp}}(R, W, \mathcal{A}) = \infty$ для всех $R \in [0, C_{1,W,\mathcal{A}})$, и $E_{\mathrm{sp}}(R, W, \mathcal{A}) = 0$ для всех $R \geqslant C_{1,W,\mathcal{A}}$. Следовательно, (11) выполнено, и утверждения о непрерывности и конечности $E_{\mathrm{sp}}(R, W, \mathcal{A})$ справедливы.
- Если $C_{0^+,W,\mathcal{A}} < \infty$ и $C_{0^+,W,\mathcal{A}} \neq C_{1,W,\mathcal{A}}$, то $E_{\mathrm{sp}}(R, W, \mathcal{A}) = \infty$ для всех $R \in [0, C_{0^+,W,\mathcal{A}})$. Для $R \geqslant C_{0^+,W,\mathcal{A}}$ из неотрицательности $\frac{1-\alpha}{\alpha}(C_{\alpha,W,\mathcal{A}} - R)$ следуют ограничения (11).

В силу (11) величина $E_{\mathrm{sp}}(R, W, \mathcal{A})$ конечна для всех $R > \lim_{\phi\downarrow 0} C_{\phi,W,\mathcal{A}}$. Таким образом, $E_{\mathrm{sp}}(R, W, \mathcal{A})$ непрерывна на $(C_{0^+,W,\mathcal{A}}, \infty)$ в силу [35, теорема 6.3.3]. Чтобы продолжить непрерывность на $[C_{0^+,W,\mathcal{A}}, \infty)$, заметим, что $\frac{1-\alpha}{\alpha}(C_{\alpha,W,\mathcal{A}} - R)$ не убывает и непрерывна по R для каждого $\alpha \in (0,1)$. Таким образом, $E_{\mathrm{sp}}(R,W,\mathcal{A})$ не возрастает и полунепрерывна снизу по R. Следовательно, $E_{\mathrm{sp}}(R,W,\mathcal{A})$ непрерывна справа, а значит, и в точке $R = C_{0^+,W,\mathcal{A}}$. ▲

Доказательство следствия 2. По лемме 1 для некоторого $\alpha \in (\eta,1)$ справедливо неравенство $\frac{1-\alpha}{\alpha}\Big(C_{\alpha,W_{[1,n]},n\varrho} - \ln\frac{M}{L}\Big) \geqslant E_{\mathrm{sp}}\Big(\ln\frac{M}{L}, W_{[1,n]}, n\varrho\Big) - \frac{1}{n}$. Существует $p \in \mathcal{P}(\widetilde{\mathcal{X}}_1^n)$ вида $p = \bigotimes_{t=1}^{n} p_t$, такая что $\mathbf{E}_p[\rho_{[1,n]}] \leqslant n(\widetilde{\varrho} - \delta)$. Положим $\widehat{\varrho} \triangleq \varrho - \frac{3e\varsigma}{n}$. В силу [26, леммы 14 и 28] существует $\widetilde{p} \in \mathcal{P}(\widetilde{\mathcal{X}}_1^n)$ вида $\widetilde{p} = \bigotimes_{t=1}^{n} \widetilde{p}_t$, такая что $I_\alpha\big(\widetilde{p};\widetilde{W}_{[1,n]}\big) \geqslant C_{\alpha,\widetilde{W}_{[1,n]},n\widehat{\varrho}} - \epsilon$ и $\mathbf{E}_{\widetilde{p}}[\rho_{[1,n]}] \leqslant n\widehat{\varrho}$. Пусть $\widehat{W}_{[1,n]} \colon \widehat{\mathcal{X}}_1^n \to \mathcal{P}(\mathcal{Y}_1^n)$ – канал-произведение, такой что $\widehat{W}_{[1,n]}(x_1^n) = W_{[1,n]}(x_1^n)$ для всех $x_1^n \in \widehat{\mathcal{X}}_1^n$ и $\widehat{\mathcal{X}}_t = \operatorname{supp}\widetilde{p}_t \cup \operatorname{supp} p_t$ для всех $t \in \{1,\ldots,n\}$. Тогда $C_{\alpha,\widehat{W}_{[1,n]},n\widehat{\varrho}} \geqslant C_{\alpha,\widetilde{W}_{[1,n]},n\widehat{\varrho}} - \epsilon$ и $\widetilde{\varrho} \leqslant \widehat{\varrho}$ по построению. Так как $\widehat{\mathcal{X}}_1^n$ конечно, то в силу [26, лемма 19] существует $\widehat{p} \in \mathcal{P}(\widehat{\mathcal{X}}_1^n)$, такая что $I_\alpha\big(\widehat{p};\widehat{W}_{[1,n]}\big) = C_{\alpha,\widehat{W}_{[1,n]},n\widehat{\varrho}}$ и $\mathbf{E}_{\widehat{p}}[\rho_{[1,n]}] \leqslant n\widehat{\varrho}$. Кроме того, так как функция стоимости $\rho_{[1,n]}$ аддитивна, то в силу [26, лемма 14] без ограничения общности можно считать, что $\widehat{p}$ имеет вид $\widehat{p} = \bigotimes_{t=1}^{n} \widehat{p}_t$.

Заметим, что $\mathbf{E}_{\widehat{p}_t}\big[|\rho_t - \mathbf{E}_{\widehat{p}_t}[\rho_t]|^\kappa\big]^{1/\kappa} \leqslant \varsigma$ для всех $\kappa \in \mathbb{R}_+$ и $t \in \{1,\ldots,n\}$, поскольку ρ_t положительно и меньше ς с вероятностью единица относительно $\widehat{p}_t$. Тогда, применяя [25, лемма 19] при $\kappa = \ln n$, получаем

$$\widehat{p}\big(\big|\rho_{[1,n]}(x) - \mathbf{E}_{\widehat{p}}[\rho_{[1,n]}]\big| < 3\varsigma e\big) \geqslant \frac{1}{2\sqrt{n}}.$$

С другой стороны, в силу [26, лемма 29(c)] существует $\widehat{\lambda} = \lambda_{\alpha,\widehat{W}_{[1,n]},n\widehat{\varrho}}$, для которого $C_{\alpha,\widehat{W}_{[1,n]},n\widehat{\varrho}} = C^{\widehat{\lambda}}_{\alpha,\widehat{W}_{[1,n]}} + \widehat{\lambda}n\widehat{\varrho}$, поскольку по построению $n\widehat{\varrho}$ принадлежит внутренней области допустимых ограничений по стоимости $\widehat{W}_{[1,n]}$. При этом $I_\alpha^{\widehat{\lambda}}\big(\widehat{p};\widehat{W}_{[1,n]}\big) = C^{\widehat{\lambda}}_{\alpha,\widehat{W}_{[1,n]}}$ в силу [26, лемма 29(d)], так как $I_\alpha\big(\widehat{p};\widehat{W}_{[1,n]}\big) = C_{\alpha,\widehat{W}_{[1,n]},n\widehat{\varrho}}$ и $\mathbf{E}_{\widehat{p}}[\rho_{[1,n]}] \leqslant n\widehat{\varrho}$. В результате получаем, что $D_\alpha\big(W_{[1,n]}(x_1^n)\,\|\,q_{\alpha,\widehat{p}}\big) - \widehat{\lambda}\rho_{[1,n]}(x_1^n) = C^{\widehat{\lambda}}_{\alpha,\widehat{W}_{[1,n]}}$ для всех $x_1^n \in \operatorname{supp}\widehat{p}$ и $\widehat{\lambda}n\widehat{\varrho} = \widehat{\lambda}\,\mathbf{E}_{\widehat{p}}[\rho_{[1,n]}]$. Применяя лемму 6 при $\mathcal{B} = \big\{x_1^n \in \operatorname{supp}\widehat{p} : \big|\rho_{[1,n]}(x_1^n) - \mathbf{E}_{\widehat{p}}[\rho_{[1,n]}]\big| < 3\varsigma e\big\}$, заключаем, что существует (M,L)-код, для которого

$$\begin{aligned}\ln P_{\mathrm{e}} &\leqslant \frac{\alpha-1}{\alpha}\left[C^{\widehat{\lambda}}_{\alpha,\widehat{W}_{[1,n]}} + \inf_{x_1^n \in \mathcal{B}} \widehat{\lambda}\rho_{[1,n]}(x_1^n) - \ln\frac{(M-1)e}{L}\right] + \frac{\ln 4n}{2\alpha} \leqslant \\ &\leqslant \frac{\alpha-1}{\alpha}\left[C_{\alpha,\widehat{W}_{[1,n]},n\widehat{\varrho}} - 3\varsigma e\widehat{\lambda} - \ln\frac{(M-1)e}{L}\right] + \frac{\ln 4n}{2\alpha}.\end{aligned}$$

Так как $C_{\alpha,\widehat{W}_{[1,n]},n\widehat{\varrho}} + \widehat{\lambda}n(\varrho - \widehat{\varrho}) \geqslant C_{\alpha,\widehat{W}_{[1,n]},n\varrho}$ согласно [26, лемма 27(b)], получаем

$$\ln P_{\mathrm{e}} \leqslant -E_{\mathrm{sp}}\Big(\ln\frac{M}{L}, W_{[1,n]}, \varrho\Big) + \frac{1}{n} + \frac{1-\alpha}{\alpha}\big(6\varsigma e\widehat{\lambda} + 2\epsilon + 1\big) + \frac{\ln 4n}{2\alpha}.$$

Теперь (25) выполнено, поскольку $\frac{1-\alpha}{\alpha}C_{\alpha,W_{[1,n]},n\varrho}$ не возрастает по α согласно [26, лемма 23(b)] при условии, что $\widehat{\lambda} \leqslant \frac{C_{\alpha,W_{[1,n]},n\varrho}}{n\delta}$. Чтобы увидеть, почему это условие выполнено, заметим вначале, что $\widehat{\varrho} \leqslant \varrho$ по определению, и поэтому $C_{\alpha,\widehat{W}_{[1,n]},n\widehat{\varrho}} \leqslant C_{\alpha,\widehat{W}_{[1,n]},n\varrho}$. При этом

$$C_{\alpha,\widehat{W}_{[1,n]},n\widehat{\varrho}} = \inf_{\lambda\geqslant\widehat{\lambda}} C^{\lambda}_{\alpha,\widehat{W}_{[1,n]}} + \lambda n\widehat{\varrho} > n\delta\widehat{\lambda} + \inf_{\lambda\geqslant\widehat{\lambda}} C^{\lambda}_{\alpha,\widehat{W}_{[1,n]}} + n\lambda(\widehat{\varrho}-\delta) \geqslant$$
$$\geqslant n\delta\widehat{\lambda} + C_{\alpha,\widehat{W}_{[1,n]},n(\widehat{\varrho}-\delta)} \geqslant n\delta\widehat{\lambda}. \quad \blacktriangle$$

Доказательство леммы 8. Согласно (39) и [26, лемма 23(a),(b)] имеем $C_{\alpha,W,\mathcal{A}} \leqslant \widetilde{C}^{\epsilon}_{\alpha,W,\mathcal{A}}$. Тогда в силу выражения для $E_{\rm sp}(R,W,\mathcal{A})$, данного в (11), и определения $\widetilde{E}^{\epsilon}_{\rm sp}(R,W,\mathcal{A})$, данного в (40), имеем

$$E_{\rm sp}(R,W,\mathcal{A}) \leqslant \widetilde{E}^{\epsilon}_{\rm sp}(R,W,\mathcal{A}), \quad \forall R \in \mathbb{R}_{\geqslant 0}. \tag{C.1}$$

Будем оценивать $\widetilde{E}^{\epsilon}_{\rm sp}(R,W,\mathcal{A}) - E_{\rm sp}(R,W,\mathcal{A})$ сверху при $R \in [C_{\phi,W,\mathcal{A}},\infty)$. Во-первых,

$$\frac{1-\alpha}{\alpha}\big(\widetilde{C}^{\epsilon}_{\alpha,W,\mathcal{A}} - R\big) = \frac{1}{\epsilon}\int\limits_{\alpha-\epsilon\alpha}^{\alpha+\epsilon(1-\alpha)} \left(\frac{1-\alpha}{\alpha} \vee \frac{1-\eta}{\eta}\right) C_{\eta,W,\mathcal{A}}\, d\eta - \frac{1-\alpha}{\alpha}R =$$
$$= \frac{1}{\epsilon}\frac{1-\alpha}{\alpha}\int\limits_{\alpha}^{\alpha+\epsilon(1-\alpha)} (C_{\eta,W,\mathcal{A}} - R)\, d\eta +$$
$$+ \frac{1}{\epsilon}\int\limits_{\alpha-\epsilon\alpha}^{\alpha} \frac{1-\eta}{\eta}(C_{\eta,W,\mathcal{A}} - R)\, d\eta + \frac{R}{\epsilon}\int\limits_{\alpha-\epsilon\alpha}^{\alpha} \frac{\alpha-\eta}{\eta\alpha}\, d\eta \leqslant$$
$$\leqslant \frac{1}{\epsilon}\frac{1-\alpha}{\alpha}\int\limits_{\alpha}^{\alpha+\epsilon(1-\alpha)} (C_{\eta,W,\mathcal{A}} - R)\, d\eta + \frac{1}{\epsilon}\int\limits_{\alpha-\epsilon\alpha}^{\alpha} \frac{1-\eta}{\eta}(C_{\eta,W,\mathcal{A}} - R)\, d\eta + \frac{\epsilon}{1-\epsilon}R. \tag{C.2}$$

Для оценки $\widetilde{E}^{\epsilon}_{\rm sp}(R,W,\mathcal{A})$ оценим выражение в (C.2) на двух интервалах значений α по отдельности.

Чтобы оценить (C.2) для $\alpha \in [\phi,1)$, используем тот факт, что $\sup\limits_{\eta\in(0,1)} \frac{1-\eta}{\eta}(C_{\eta,W,\mathcal{A}} - R) = E_{\rm sp}(R,W,\mathcal{A})$. Имеем

$$\frac{1-\alpha}{\alpha}\big(\widetilde{C}^{\epsilon}_{\alpha,W,\mathcal{A}} - R\big) \leqslant$$
$$\leqslant \frac{1}{\epsilon}\frac{1-\alpha}{\alpha}\int\limits_{\alpha}^{\alpha+\epsilon(1-\alpha)} (C_{\eta,W,\mathcal{A}} - R) d\eta + \frac{1}{\epsilon}\int\limits_{\alpha-\epsilon\alpha}^{\alpha} \frac{1-\eta}{\eta}(C_{\eta,W,\mathcal{A}} - R) d\eta + \frac{\epsilon}{1-\epsilon}R \leqslant$$
$$\leqslant \frac{1}{\epsilon}\frac{1-\alpha}{\alpha}\int\limits_{\alpha}^{\alpha+\epsilon(1-\alpha)} \frac{\eta}{1-\eta}E_{\rm sp}(R,W,\mathcal{A}) d\eta + \frac{1}{\epsilon}\int\limits_{\alpha-\epsilon\alpha}^{\alpha} E_{\rm sp}(R,W,\mathcal{A}) d\eta + \frac{\epsilon}{1-\epsilon}R \leqslant$$
$$\leqslant E_{\rm sp}(R,W,\mathcal{A}) + \frac{\epsilon}{1-\epsilon}\frac{1-\alpha}{\alpha}E_{\rm sp}(R,W,\mathcal{A}) + \frac{\epsilon}{1-\epsilon}R.$$

Так как $(1-\alpha)E_{\rm sp}(R,W,\mathcal{A}) + \alpha R \leqslant (R \vee E_{\rm sp}(R,W,\mathcal{A}))$ для всех $\alpha \in (0,1)$, то

$$\frac{1-\alpha}{\alpha}\big(\widetilde{C}^{\epsilon}_{\alpha,W,\mathcal{A}} - R\big) \leqslant E_{\rm sp}(R,W,\mathcal{A}) + \frac{\epsilon}{1-\epsilon}\frac{R \vee E_{\rm sp}(R,W,\mathcal{A})}{\phi}, \quad \alpha \in [\phi,1). \tag{C.3}$$

Чтобы оценить выражение в (C.2) для $\alpha \in (0,\phi]$, напомним, что $C_{\alpha,W,\mathcal{A}}$ не убывает по α согласно [26, лемма 23(a)]. Таким образом, для любой $R \geqslant C_{\phi,W,\mathcal{A}}$ имеем

$$
\begin{aligned}
&\frac{1-\alpha}{\alpha}\big(\widetilde{C}^{\epsilon}_{\alpha,W,\mathcal{A}}-R\big)\leqslant\\
&\leqslant\frac{1}{\epsilon}\frac{1-\alpha}{\alpha}\int\limits_{\phi}^{\alpha+\epsilon(1-\alpha)}\frac{\eta}{1-\eta}E_{\rm sp}(R,W,\mathcal{A})\,d\eta\,\mathbb{1}_{\{\alpha\in[\frac{\phi-\epsilon}{1-\epsilon},\phi]\}}+\frac{\epsilon}{1-\epsilon}R\leqslant\\
&\leqslant\frac{1}{\epsilon}\frac{\alpha+\epsilon(1-\alpha)}{\alpha(1-\epsilon)}\int\limits_{\phi}^{\alpha+\epsilon(1-\alpha)}E_{\rm sp}(R,W,\mathcal{A})\,d\eta\,\mathbb{1}_{\{\alpha\in[\frac{\phi-\epsilon}{1-\epsilon},\phi]\}}+\frac{\epsilon}{1-\epsilon}R=\\
&=\left[\frac{\alpha(1-\epsilon)+2\epsilon-\phi}{\epsilon}-\frac{\phi-\epsilon}{\alpha(1-\epsilon)}\right]E_{\rm sp}(R,W,\mathcal{A})\mathbb{1}_{\{\alpha\in[\frac{\phi-\epsilon}{1-\epsilon},\phi]\}}+\frac{\epsilon}{1-\epsilon}R\leqslant\\
&\leqslant(1-\phi)E_{\rm sp}(R,W,\mathcal{A})+\frac{\epsilon}{1-\epsilon}\frac{(1-\phi)E_{\rm sp}(R,W,\mathcal{A})+\phi R}{\phi},\quad \alpha\in(0,\phi]. \qquad \text{(C.4)}
\end{aligned}
$$

Теперь (41) следует из (C.1), (C.3) и (C.4).

С другой стороны, $C_{\alpha,W,\mathcal{A}}$ не убывает, а $\frac{1-\alpha}{\alpha}C_{\alpha,W,\mathcal{A}}$ не возрастает по α согласно [26, лемма 23(a),(b)]. Тогда в силу выражения для $E_{\rm sp}(R,W,\mathcal{A})$, данного в (11), имеем $E_{\rm sp}(R,W,\mathcal{A})\leqslant\frac{1-\phi}{\phi}R$ для всех $R\in[C_{\phi,W,\mathcal{A}},\infty)$. Следовательно,

$$
R\vee E_{\rm sp}(R,W,\mathcal{A})\leqslant R/\phi,\quad \forall R\in[C_{\phi,W,\mathcal{A}},\infty). \qquad \text{(C.5)}
$$

Теперь (42) следует из (41) и (C.5). ▲

Доказательство теоремы 1. Будем доказывать теорему 1, используя леммы 8 и 9. Мы можем выбирать различные значения κ и ϵ для различных значений n при условии, что выполняются условия лемм 8 и 9.

В силу предположения 1 существуют $K\in[1,\infty)$ и $n_0\in\mathbb{Z}_+$, такие что

$$
\max_{t:\,t\leqslant n}C_{1/2,U^{(t)}_{\mathcal{A}}}\leqslant K\ln(n),\quad \forall n\geqslant n_0.
$$

Положим $\kappa_n=K\ln(1+n)$. Тогда

$$
\gamma_n\leqslant 40(K+1)\ln(1+n),\quad \forall n\geqslant n_0. \qquad \text{(C.6)}
$$

Функция $C_{\alpha,W,\mathcal{A}}$ не убывает по α в силу [26, лемма 23(a)], а $\frac{1-\alpha}{\alpha}C_{\alpha,W,\mathcal{A}}$ не возрастает по α на интервале $(0,1)$ в силу [26, лемма 23(b)]. Таким образом, мы можем оценить $\widetilde{C}^{\epsilon}_{\alpha,W,\mathcal{A}}$, используя ее определение, данное в (39):

$$
\begin{aligned}
&\widetilde{C}^{\epsilon}_{\alpha,W,\mathcal{A}}=\frac{1}{\epsilon}\int\limits_{\alpha-\epsilon\alpha}^{\alpha+\epsilon(1-\alpha)}\left[1\vee\left(\frac{\alpha}{1-\alpha}\frac{1-\eta}{\eta}\right)\right]C_{\eta,W,\mathcal{A}}\,d\eta\leqslant\\
&\leqslant\frac{1}{\epsilon}\int\limits_{\alpha-\epsilon\alpha}^{\alpha+\epsilon(1-\alpha)}\left[1\vee\left(\frac{\alpha}{1-\alpha}\frac{1-\eta}{\eta}\right)\right]\left[1\vee\left(\frac{1-\alpha}{\alpha}\frac{\eta}{1-\eta}\right)\right]C_{\alpha,W,\mathcal{A}}\,d\eta=\\
&=\frac{C_{\alpha,W,\mathcal{A}}}{\epsilon}\int\limits_{\alpha-\epsilon\alpha}^{\alpha+\epsilon(1-\alpha)}\left[\left(\frac{\alpha}{1-\alpha}\frac{1-\eta}{\eta}\right)\vee\left(\frac{1-\alpha}{\alpha}\frac{\eta}{1-\eta}\right)\right]d\eta\leqslant\\
&\leqslant\left(1+\frac{\epsilon}{1-\epsilon}\frac{\alpha^2+(1-\alpha)^2}{\alpha(1-\alpha)}\right)C_{\alpha,W,\mathcal{A}}.
\end{aligned}
$$

Тогда при $\epsilon_n = 1/n$ из оценки (С.6) следует, что

$$n\widetilde{C}^{\epsilon}_{\phi,W,\mathcal{A}} + \frac{\gamma_n}{1-\phi} + \ln\frac{8e^3n^{1,5}}{\epsilon} \leqslant$$
$$\leqslant nC_{\phi,W,\mathcal{A}} + \frac{n}{n-1}\frac{C_{\phi,W,\mathcal{A}}}{\phi(1-\phi)} + \frac{40(K+1)\ln(1+n)}{1-\phi} + 3\ln(2en), \quad \forall n \geqslant n_0. \qquad (C.7)$$

Таким образом, из условий теоремы 1 вытекают условия леммы 9 при всех достаточно больших n. Значит, для всех достаточно больших n из леммы 9 и оценки (С.6) следует, что

$$P_{\mathrm{e}} \geqslant \left(\frac{(1+n)^{-80(K+1)}}{8e^3n^{2,5}}\right)^{1/\phi} e^{-n\widetilde{E}^{\epsilon_n}_{\mathrm{sp}}\left(\frac{1}{n}\ln\frac{M_n}{L_n},W,\mathcal{A}\right)}. \qquad (C.8)$$

С другой стороны, из леммы 8, соотношения (31) и монотонности $C_{\alpha,W,\mathcal{A}}$ по α следует, что для достаточно больших n

$$\widetilde{E}^{\epsilon_n}_{\mathrm{sp}}\left(\frac{1}{n}\ln\frac{M_n}{L_n},W,\mathcal{A}\right) \leqslant E_{\mathrm{sp}}\left(\frac{1}{n}\ln\frac{M_n}{L_n},W,\mathcal{A}\right) + \frac{C_{\alpha_1,W,\mathcal{A}}}{(n-1)\phi^2}. \qquad (C.9)$$

Теперь (32) следует из (С.8) и (С.9). ▲

Доказательство теоремы 2. Будем доказывать теорему 2, используя леммы 8 и 10. Мы можем выбирать различные значения κ и ϵ для различных значений n при условии, что выполняются условия лемм 8 и 10.

В силу предположения 2 существуют $K \in [1,\infty)$ и $n_0 \in \mathbb{Z}_+$, такие что

$$\max_{t:\, t\leqslant n} C_{1/2,W_t,n\varrho} \leqslant K\ln n, \quad \forall n \geqslant n_0. \qquad (C.10)$$

Положим $\kappa_n = K\ln(1+n)$. Тогда

$$\gamma_n \leqslant 40(K+1)\ln(1+n), \quad \forall n \geqslant n_0. \qquad (C.11)$$

Функция $C_{\alpha,W,\varrho}$ не убывает по α в силу [26, лемма 23(a)], а $\frac{1-\alpha}{\alpha}C_{\alpha,W,\varrho}$ не возрастает по α на интервале $(0,1)$ в силу [26, лемма 23(b)]. Таким образом, мы можем оценить $\widetilde{C}^{\epsilon}_{\alpha,W,\varrho}$, используя ее определение, данное в (39):

$$\widetilde{C}^{\epsilon}_{\alpha,W_{[1,n]},n\varrho} \leqslant \left(1 + \frac{\epsilon}{1-\epsilon}\frac{\alpha^2+(1-\alpha)^2}{\alpha(1-\alpha)}\right) C_{\alpha,W_{[1,n]},n\varrho}.$$

Тогда при $\epsilon_n = \frac{1}{n}$ из оценок (С.10) и (С.11), а также [26, леммы 23(a),(b), 27(a) и 28] следует, что

$$\widetilde{C}^{\epsilon}_{\alpha,W_{[1,n]},n\varrho} + \frac{\gamma_n}{1-\phi} + \ln\frac{8e^3n^{1,5}}{\epsilon} \leqslant C_{\alpha,W_{[1,n]},n\varrho} + \frac{n}{n-1}\frac{K\ln(n)}{\phi(1-\phi)} +$$
$$+ \frac{40(K+1)\ln(1+n)}{1-\phi} + 3\ln(2en), \quad \forall n \geqslant n_0. \qquad (C.12)$$

Таким образом, из условий теоремы 2 вытекают условия леммы 10 при всех достаточно больших n. Значит, для всех достаточно больших n из леммы 10 и оценки (С.11) следует, что

$$P_{\mathrm{e}} \geqslant \left(\frac{(1+n)^{-80(K+1)}}{8e^3n^{2,5}}\right)^{1/\phi} e^{-\widetilde{E}^{\epsilon_n}_{\mathrm{sp}}\left(\ln\frac{M_n}{L_n},W_{[1,n]},n\varrho\right)}. \qquad (C.13)$$

С другой стороны, из леммы 8, соотношения (33) и монотонности $C_{\alpha,W,\varrho}$ по α следует, что для достаточно больших n

$$\widetilde{E}_{\mathrm{sp}}^{\epsilon_n}\Big(\ln\frac{M_n}{L_n}, W_{[1,n]}, n\varrho\Big) \leqslant E_{\mathrm{sp}}\Big(\ln\frac{M_n}{L_n}, W_{[1,n]}, n\varrho\Big) + \frac{C_{\alpha_1,W_{[1,n]},n\varrho}}{(n-1)\phi^2}. \tag{C.14}$$

Заметим, что $C_{\alpha,W_{[1,n]},n\varrho}$ не убывает по α согласно [26, лемма 23(a)], а функция $\frac{1-\alpha}{\alpha}C_{\alpha,W_{[1,n]},n\varrho}$ не возрастает по α на интервале $(0,1)$ согласно [26, лемма 23(b)]. Таким образом,

$$C_{\alpha_1,W_{[1,n]},n\varrho} \leqslant \big(\frac{\alpha_1}{1-\alpha_1} \vee 1\big)C_{1/2,W_{[1,n]},n\varrho}. \tag{C.15}$$

Так как $C_{\alpha,W,\varrho}$ не убывает по ϱ согласно [26, лемма 27(a)], то из соотношения (C.10) и [26, лемма 28] следует, что для всех достаточно больших n

$$C_{1/2,W_{[1,n]},n\varrho} \leqslant Kn\ln n. \tag{C.16}$$

Для достаточно больших n из оценок (C.14), (C.15) и (C.16) вытекает, что

$$\widetilde{E}_{\mathrm{sp}}^{\epsilon_n}\Big(\ln\frac{M_n}{L_n}, W_{[1,n]}, n\varrho\Big) \leqslant E_{\mathrm{sp}}\Big(\ln\frac{M_n}{L_n}, W_{[1,n]}, n\varrho\Big) + \frac{2}{\phi^2}\Big(\frac{\alpha_1}{1-\alpha_1} \vee 1\Big)K\ln n. \tag{C.17}$$

Теперь (34) следует из (C.13) и (C.17). ▲

Автор выражает благодарность Фатьме и Мехмету Накибоглу за гостеприимство, без которого эта работа не была бы возможной. Автор также благодарит М. Далаи за указание на неявное утверждение Фано о неподвижной точке в [9], Г. Васкеса-Вилара за указание на работу Полтырева [30] о границе случайного кодирования и за предложения по улучшению статьи, В. Яна за указание на [54, формула (3.63)] и ее отношение к подходу Блейхута, а также рецензента за предложения по улучшению статьи.

СПИСОК ЛИТЕРАТУРЫ

1. *Nakiboğlu B.* The Augustin Center and the Sphere Packing Bound for Memoryless Channels // Proc. 2017 IEEE Int. Sympos. on Information Theory (ISIT'2017). Aachen, Germany. June 25–30, 2017. P. 1401–1405.
2. *Elias P.* Coding for Noisy Channels // IRE Conv. Rec. 1955. V. 4. P. 37–46. Reprinted in: Key Papers in the Development of Information Theory. New York: IEEE Press, 1974. P. 48–55.
3. *Elias P.* Coding for Two Noisy Channels // Information Theory: Third London Symposium (London, England, Sept. 1955). London: Butterworth Sci., 1956. P. 61–74.
4. *Добрушин Р.Л.* Асимптотические оценки вероятности ошибки при передаче сообщения по дискретному каналу связи без памяти с симметрической матрицей вероятностей перехода // Теория вероятн. и ее примен. 1962. V. 7. № 3. P. 283–311.
5. *Altuğ Y., Wagner A.B.* On Exact Asymptotics of the Error Probability in Channel Coding: Symmetric Channels, `arXiv:1908.11419 [cs.IT]`, 2019.
6. *Altuğ Y., Wagner A.B.* Refinement of the Sphere Packing Bound for Symmetric Channels // Proc. 49th Annu. Allerton Conf. on Communication, Control, and Computing. Monticello, IL, USA. Sept. 27–30, 2011. P. 30–37.
7. *Shannon C.E., Gallager R.G., Berlekamp E.R.* Lower Bounds to Error Probability for Coding on Discrete Memoryless Channels. I // Inform. Control. 1967. V. 10. № 1. P. 65–103.
8. *Gallager R.G.* Information Theory and Reliable Communication. New York: John Wiley & Sons, 1968.
9. *Fano R.M.* Transmission of Information: A Statistical Theory of Communications. New York: M.I.T. Press, 1961.

10. *Арутюнян Е.А.* Оценки экспоненты вероятности ошибки для полунепрерывного канала без памяти // Пробл. передачи информ. 1968. Т. 4. № 4. С. 37–48.
11. *Omura J.K.* A Lower Bounding Method for Channel and Source Coding Probabilities // Inform. Control. 1975. V. 27. № 2. P. 148–177.
12. *Csiszár I., Körner J.* Information Theory: Coding Theorems for Discrete Memoryless Systems. Cambridge, UK: Cambridge Univ. Press, 2011.
13. *Altuğ Y., Wagner A.B.* Refinement of the Sphere-Packing Bound: Asymmetric Channels // IEEE Trans. Inform. Theory. 2014. V. 60. № 3. P. 1592–1614.
14. *Nakiboğlu B.* A Simple Derivation of the Refined SPB for the Constant Composition Codes // Proc. 2019 IEEE Int. Symp. on Information Theory (ISIT'2019). Paris, France. July 7–12, 2019. P. 2659–2663.
15. *Wyner A.D.* Capacity and Error Exponent for the Direct Detection Photon Channel. II // IEEE Trans. Inform. Theory. 1988. V. 34. № 6. P. 1462–1471.
16. *Бурнашев М.В., Кутоянц Ю.А.* О границе сферической упаковки, пропускной способности и о подобных результатах для пуассоновского канала // Пробл. передачи информ. 1999. Т. 35. № 2. С. 3–22.
17. *Shannon C.E.* Probability of Error for Optimal Codes in a Gaussian Channel // Bell System Tech. J. 1959. V. 38. № 3. P. 611–656.
18. *Ebert P.M.* Error Bounds for Gaussian Noise Channels // Quarterly Progress Rep. № 77. Research Lab. of Electronics, MIT. Cambridge, MA, USA. 1965. P. 292–302. Available at `https://dspace.mit.edu/handle/1721.1/55609`.
19. *Ebert P.M.* Error Bounds for Parallel Communication Channels // Tech. Rep. № 448. Research Lab. of Electronics, MIT. Cambridge, MA, USA. 1966. Available at `https://dspace.mit.edu/handle/1721.1/4295`.
20. *Richters J.S.* Communication over Fading Dispersive Channels. Tech. Rep. № 464. Research Lab. of Electronics, MIT. Cambridge, MA, USA. 1967. Available at `https://dspace.mit.edu/handle/1721.1/4279`.
21. *Lancho Serrano A.* Fundamental Limits of Short-Packet Wireless Communications: Ph.D. Thesis. Departamento de Teoría de la Señal y Comunicaciones, Univ. Carlos III de Madrid, Spain. 2019. Available at `https://e-archivo.uc3m.es/handle/10016/29596`.
22. *Lancho A., Östman J., Durisi G., Koch T., Vazquez-Vilar G.* Saddlepoint Approximations for Short-Packet Wireless Communications // IEEE Trans. Wireless Commun. 2020. V. 19. № 7. P. 4831–4846.
23. *Augustin U.* Error Estimates for Low Rate Codes // Z. Wahrsch. Verw. Gebiete. 1969. V. 14. № 1. P. 61–88.
24. *Augustin U.* Noisy Channels // Habilitation Thesis. Univ. Erlangen-Nürnberg, Germany, 1978. Avaliable at `http://bit.ly/2ID8h7m`.
25. *Nakiboğlu B.* The Sphere Packing Bound via Augustin's Method // IEEE Trans. Inform. Theory. 2019. V. 65. № 2. P. 816–840.
26. *Накибоглу Б.* Августиновская пропускная способность и августиновский центр // Пробл. передачи информ. 2019. Т. 55. № 4. С. 3–51.
27. *Nakiboğlu B.* The Rényi Capacity and Center // IEEE Trans. Inform. Theory. 2019. V. 65. № 2. P. 841–860.
28. *Nakiboğlu B.* A Simple Derivation of the Refined Sphere Packing Bound under Certain Symmetry Hypotheses // Turkish J. Math. 2020. V. 44. № 3. P. 919–948.
29. *Gallager R.G.* A Simple Derivation of the Coding Theorem and Some Applications // IEEE Trans. Inform. Theory. 1965. V. 11. № 1. P. 3–18.
30. *Полтырев Г.Ш.* Границы случайного кодирования для дискретных каналов без памяти // Пробл. передачи информ. 1982. Т. 18. № 1. С. 12–26.
31. *Blahut R.E.* Hypothesis Testing and Information Theory // IEEE Trans. Inform. Theory. 1974. V. 20. № 4. P. 405–417.
32. *Blahut R.E.* Information Bounds of the Fano–Kullback Type // IEEE Trans. Inform. Theory. 1976. V. 22. № 4. P. 410–421.
33. *Blahut R.E.* Principles and Practice of Information Theory. Reading, MA: Addison-Wesley, 1987.

34. *Vazquez-Vilar G., Martinez A., Guillén i Fàbregas A.* A Derivation of the Cost-Constrained Sphere-Packing Exponent // Proc. 2015 IEEE Int. Sympos. on Information Theory (ISIT'2015). Hong Kong, China. June 14–19, 2015. P. 929–933.
35. *Dudley R.M.* Real Analysis and Probability. Cambridge: Cambridge Univ. Press, 2002.
36. *van Erven T., Harremoës P.* Rényi Divergence and Kullback–Leibler Divergence // IEEE Trans. Inform. Theory. 2014. V. 60. № 7. P. 3797–3820.
37. *Rudin W.* Principles of Mathematical Analysis. New York: McGraw-Hill, 1976.
38. *Bogachev V.I.* Measure Theory. Berlin: Springer, 2007.
39. *Elias P.* List Decoding for Noisy Channels // Tech. Rep. № 335. Research Lab. of Electronics, MIT. Cambridge, MA, USA. 1957. Available at `https://dspace.mit.edu/handle/1721.1/4484`.
40. *Dalai M.* Lower Bounds on the Probability of Error for Classical and Classical-Quantum Channels // IEEE Trans. Inform. Theory. 2013. V. 59. № 12. P. 8027–8056.
41. *Dalai M.* Some Remarks on Classical and Classical-Quantum Sphere Packing Bounds: Rényi vs. Kullback–Leibler // Entropy. 2017. V. 19. № 7. P. 355 (11 pp.).
42. *Dalai M., Winter A.* Constant Compositions in the Sphere Packing Bound for Classical-Quantum Channels // IEEE Trans. Inform. Theory. 2017. V. 63. № 9. P. 5603–5617.
43. *Cheng H.-C., Hsieh M.H.* Moderate Deviation Analysis for Classical-Quantum Channels and Quantum Hypothesis Testing // IEEE Trans. Inform. Theory. 2018. V. 64. № 2. P. 1385–1403.
44. *Cheng H.-C., Hsieh M.H., Tomamichel M.* Quantum Sphere-Packing Bounds with Polynomial Prefactors // IEEE Trans. Inform. Theory. 2019. V. 65. № 5. P. 2872–2898.
45. *Altuğ Y., Wagner A.B.* Refinement of the Random Coding Bound // IEEE Trans. Inform. Theory. 2014. V. 60. № 10. P. 6005–6023.
46. *Scarlett J., Martinez A., Guillén i Fàbregas A.* Mismatched Decoding: Error Exponents, Second-Order Rates and Saddlepoint Approximations // IEEE Trans. Inform. Theory. 2014. V. 60. № 5. P. 2647–2666.
47. *Honda J.* Exact Asymptotics for the Random Coding Error Probability // Proc. 2015 IEEE Int. Sympos. on Information Theory (ISIT'2015). Hong Kong, China. June 14–19, 2015. P. 91–95.
48. *Бурнашев М.В., Холево А.С.* О функции надежности квантового канала связи // Пробл. передачи информ. 1998. Т. 34. № 2. С. 3–15.
49. *Telatar E.* Capacity of Multi-antenna Gaussian Channels // Eur. Trans. Telecommun. 1999. V. 10. № 6. P. 585–595.
50. *Vu M.* MISO Capacity with Per-Antenna Power Constraint // IEEE Trans. Commun. 2011. V. 59. № 5. P. 1268–1274.
51. *Cao P.L., Oechtering T.J., Schaefer R.F., Skoglund M.* Optimal Transmit Strategy for MISO Channels with Joint Sum and Per-Antenna Power Constraints // IEEE Trans. Signal Process. 2016. V. 64. № 16. P. 4296–4306.
52. *Loyka S.* The Capacity of Gaussian MIMO Channels under Total and Per-Antenna Power Constraints // IEEE Trans. Commun. 2017. V. 65. № 3. P. 1035–1043.
53. *Cheng H.-C., Nakiboğlu B.* Refined Strong Converse for the Constant Composition Codes, `arXiv:2002.11414 [cs.IT]`, 2020.
54. *Yang W.* Fading Channels: Capacity and Channel Coding Rate in the Finite-Blocklength Regime: Ph.D. Thesis. Communication Systems Group, Dept. of Signals and Systems, Chalmers Univ. of Technology, Gothenburg, Sweden, 2015.

Накибоглу Барыш
Средневосточный технический университет,
Анкара, Турция
`bnakib@metu.edu.tr`

Поступила в редакцию
18.04.2018
После доработки
06.03.2020
Принята к публикации
29.04.2020