Conference Paper

Second-Order Rate of Constant-Composition Codes for the Gel'fand-Pinsker Channel

Author(s): Scarlett, Jonathan

Publication Date: 2014

Permanent Link: https://doi.org/10.3929/ethz-a-010094561

Rights / License: In Copyright - Non-Commercial Use Permitted

This page was generated automatically upon download from the ETH Zurich Research Collection. For more information please consult the Terms of use.
Second-Order Rate of Constant-Composition Codes for the Gel’fand-Pinsker Channel

Jonathan Scarlett

Abstract—This paper presents an achievable second-order coding rate for the discrete memoryless Gel’fand-Pinsker channel. The result is obtained using constant-composition random coding, and by using an asymptotically negligible fraction of the block to transmit the type of the state sequence.

I. INTRODUCTION

In this paper, we present an achievable second-order coding rate [1]–[3] for channel coding with a random state known non-causally at the encoder, as studied by Gel’fand and Pinsker [4]. The alphabets of the input, output and state are denoted by $\mathcal{X}$, $\mathcal{Y}$ and $\mathcal{S}$ respectively, and each are assumed to be finite. The channel transition law is given by $W^n(y|x,s) \triangleq \prod_{i=1}^{n} W(y_i|x_i,s_i)$, where $n$ is the block length. The state sequence $S = (S_1, \cdots, S_n)$ is assumed to be independent and identically distributed (i.i.d.) according to a distribution $\pi(s)$. The capacity is given by [4]

$$C = \max_{\mathcal{U}, \mathcal{Q}_{U|S}: \phi(\cdot, \cdot), \pi(\cdot)} I(U;Y) - I(U;S),$$

where the mutual informations are with respect to

$$P_{SU^iY}(s, u, y) = \pi(s)\mathcal{Q}_{U|S}(u|s)W(y|\phi(u,s),s)$$

and the maximum is over all finite alphabets $\mathcal{U}$, conditional distributions $\mathcal{Q}_{U|S}$ and functions $\phi : \mathcal{U} \times \mathcal{S} \rightarrow \mathcal{X}$.

We say that a triplet $(n,M,\epsilon)$ is achievable if there exists a code with block length $n$ containing at least $M$ messages and yielding an average error probability not exceeding $\epsilon$, and we define $M^*(n,\epsilon) \triangleq \max \{ M : (n,M,\epsilon) \text{ is achievable} \}$. Letting $P_{XY}$, $P_{X}$, $P_Y$, etc. denote the marginals of (2), we define the information densities

$$i(u,s) \triangleq \log \frac{Q_{U|S}(u|s)}{P_U(u)}$$

$$i(u,y) \triangleq \log \frac{P_{Y|U}(y|u)}{P_Y(y)}$$

with a slight abuse of notation.

\textbf{Theorem 1.} Let $\mathcal{U}$, $\mathcal{Q}_{U|S}$ and $\phi(\cdot,\cdot)$ by any set of capacity-achieving parameters in (1), and let $P_{XY}$, $i(u,s)$ and $i(u,y)$ be as given in (2)–(4) under these parameters. If $\mathbb{E}[\text{Var}[i(U,Y) | U,S]] > 0$, then

$$\log M^*(n, \epsilon) \geq nC - \sqrt{nV}Q^{-1}(\epsilon) + O(\log n),$$

for $\epsilon \in (0,1)$, where

$$V \triangleq \mathbb{E}[\text{Var}[i(U,Y) | U,S]] + \text{Var}[\mathbb{E}[i(U,Y) - i(U,S) | S]]$$

$$= \mathbb{E}[i(U,Y) - i(U,S)].$$

\textbf{Proof:} We provide a number of preliminary results in Section II, and present the proof in Section III.

It should be noted that the equality in (7) holds under the capacity-achieving parameters, but more generally (7) is at least as high as (6), with strict inequality possible for suboptimal choices of $\mathcal{Q}_{U|S}$.

To our knowledge, the only previous result on the second-order asymptotics for the present problem is that of Watanabe \textit{et al.} [5] and Yassaee \textit{et al.} [6], who used i.i.d. random coding. In [7], we show that for $\epsilon < \frac{1}{2}$ our second-order term is at least as good as that of [5], [6], with strict improvement possible. Furthermore, we show in [7] that Theorem 1 recovers, as a special case, the dispersion for channels with i.i.d. state known at both the encoder and decoder, which was derived in [8].

\textbf{Notation:} Bold symbols are used for vectors and matrices (e.g. $x$), and the corresponding $i$-th entry of a vector is denoted with a subscript (e.g. $x_i$). The marginals of a joint distribution $P_{XY}$ are denoted by $P_X$ and $P_Y$. The empirical distribution (i.e. type [9, Ch. 2]) of a vector $x$ is denoted by $P_x$. The set of all types of length $n$ on an alphabet $\mathcal{X}$ is denoted by $P_n(\mathcal{X})$. The set of all sequences of length $n$ with a given type $P_X$ is denoted by $T^n(P_X)$, and similarly for joint types. We make use of the standard asymptotic notations $O(\cdot)$ and $o(\cdot)$.

II. PRELIMINARY RESULTS

In this section, we present a number of preliminary results which will prove useful in the proof of Theorem 1. We assume that $U$, $\mathcal{Q}_{U|S}$ and $\phi(\cdot,\cdot)$ achieve the capacity in (1).

A. A Genie-Aided Setting

We prove Theorem 1 by first proving the following result for a genie-aided setting.

\textbf{Theorem 2.} Theorem 1 holds true in the case that the empirical distribution $P_S$ of $S$ is known at the decoder.

To see that Theorem 2 implies Theorem 1, we use a technique which was proposed in [10]. We use the first $g(n) = K_0 \log(n + 1)$ symbols of the block to transmit the
wherever

\[ \sum_{u \in \mathcal{Q}} \log \frac{P_{\mathcal{Y}^n}(y|u)}{P_{\mathcal{Y}^n}(y)} - \log \frac{Q_{U|S}(u|s)}{P_{U}(u)} \right) \right), 

(16)

and

\[ \max_{P_{\mathcal{S}} \in \mathcal{P}_{\mathcal{S}}} |\Delta(P_{\mathcal{S}})| \leq \frac{K_1 \log n}{n} \tag{17} \]

for some constant \( K_1 \).

### III. Proof of Theorem 1

As stated above, it suffices to prove Theorem 2. Thus, we assume that the state type \( P_{\mathcal{S}} \) is known at the decoder.

1) Random-Coding Parameters: The parameters are the auxiliary alphabet \( \mathcal{U} \), input distribution \( Q_{U|S} \), function \( \phi: \mathcal{U} \times \mathcal{S} \rightarrow \mathcal{X} \), and number of auxiliary codewords \( L^{(P_{\mathcal{S}})} \) for each state type \( P_{\mathcal{S}} \in \mathcal{P}_{\mathcal{S}}(S) \). We assume that \( \mathcal{U}, Q_{U|S} \) and \( \phi \) are capacity-achieving.

2) Codebook Generation: For each state type \( P_{\mathcal{S}} \in \mathcal{P}_{\mathcal{S}}(S) \) and each message \( m \), we randomly generate an auxiliary codebook \( \{U^{(P_{\mathcal{S}})}(m, l)\}_{l=1}^{L} \), where each codeword is drawn independently according to the uniform distribution on the type class \( T^{(P_{\mathcal{S}})} \) (see (12)). Each auxiliary codebook is revealed to the encoder and decoder.

3) Encoding and Decoding: Given the state sequence \( S \in T^{(P_{\mathcal{S}})}(S) \) and message \( m \), the encoder sends

\[ \phi^{(P_{\mathcal{S}})}(U, S) \triangleq \{ \phi(U_1, S_1), \ldots, \phi(U_n, S_n) \}, \tag{18} \]

where \( U \) is an auxiliary codeword \( U^{(P_{\mathcal{S}})}(m, l) \) with \( l \) chosen such that \( (S, U) \in T^{(P_{\mathcal{S}})}(S_{\mathcal{U}^n}) \), with an error declared if no such auxiliary codeword exists. Given \( \pi \) and the state type \( P_{\mathcal{S}} \), the decoder estimates \( m \) according to the pair \( (m, l) \) whose corresponding sequence \( U^{(P_{\mathcal{S}})}(m, l) \) maximizes

\[ i^{(P_{\mathcal{S}})}(u, y) \triangleq \sum_{i=1}^{n} i^{(P_{\mathcal{S}})}(u_i, y_i), \tag{19} \]

where

\[ i^{(P_{\mathcal{S}})}(u_i, y_i) \triangleq \log \frac{P_{\mathcal{Y}^n}(y|u)}{P_{\mathcal{Y}^n}(y)} \tag{20} \]

with \( P^{(P_{\mathcal{S}})}_{\mathcal{U}S} \) defined in (11). It should be noted that \( P^{(P_{\mathcal{S}})}_{\mathcal{U}S} \) coincides with the distribution in (2), and hence \( i^{(P_{\mathcal{S}})}(u, y) \) coincides with (4).

We consider the events

\[ E_1 \triangleq \left\{ \text{No l yields } (S, U^{(P_{\mathcal{S}})}(m, l)) \in T^{(P_{\mathcal{S}})}(S_{\mathcal{U}^n}) \right\}, \tag{21} \]

\[ E_2 \triangleq \left\{ \text{Decoder chooses a message } \tilde{m} \neq m \right\}. \tag{22} \]

It follows from these definitions and (9) that the overall random-coding error probability \( \mathcal{P}_e \) satisfies

\[ \mathcal{P}_e \leq \sum_{P_{\mathcal{S}} \in \mathcal{P}_{\mathcal{S}}} \mathbb{P}[P_{\tilde{S}} = P_{\mathcal{S}}] \left[ \mathbb{P}[E_1 | P_{\tilde{S}} = P_{\mathcal{S}}] + \mathbb{P}[E_2 | P_{\tilde{S}} = P_{\mathcal{S}}, E_1] + O \left( \frac{1}{n^2} \right) \right]. \tag{23} \]
4) Analysis of $E_1$: We study the probability of $E_1$ conditioned on $S$ having a given type $P_S \in \tilde{P}_n$. Combining (13) with a standard property of types [12, Eq. (18)], each of the auxiliary codewords induces the joint type $P_{SU,n}$ with probability at least $p_0(n)^{-1}e^{-nI(P_S(U);S)}$, where $I(P_S(U);S)$ is defined in Section II-D, and $p_0(n)$ is polynomial in $n$. Since the codewords are independent, we have

$$
\mathbb{P}[E_1 | \tilde{P}_S = P_S] \leq (1 - p_0(n)^{-1}e^{-nI(P_S(U);S)})^{U_P}
\leq \exp \left( - p_0(n)^{-1}e^{-nI(P_S(U);S) - H_e(P_S)} \right),
$$

(24) where (25) follows using $1 - \alpha \leq e^{-\alpha}$ and defining

$$
R_e(P_S) = \frac{1}{n} \log L(P_S).
$$

Choosing

$$
R_e(P_S) = I(P_S(U);S) + K_2 \frac{\log n}{n}
$$

(27) with $K_2$ equal to one plus the degree of the polynomial $p_0(n)$, we obtain from (25) that

$$
\mathbb{P}[E_1 | P_S] \leq e^{-\psi n}
$$

(28) for some $\psi > 0$ and sufficiently large $n$.

5) Analysis of $E_2$: We study the probability of $E_2$ conditioned on $S$ having a given type $P_S \in \tilde{P}_n$, and also conditioned on $E_1$. By symmetry, all $(s, u) \in T^n(P_{SU,n})$ are equally likely, and hence the conditional distribution given $P_S = P_S$ and $E_1$ of the state sequence $S$, auxiliary codeword $U$, and received sequence $Y$ is given by

$$(S, U, Y) \sim P_{SU}^{(P_S)}(s, u)W^n(y|\phi^n(u, s), s),
$$

(29) where $P_{SU}^{(P_S)}$ is uniform on the type class:

$$
P_{SU}^{(P_S)}(s, u) = \frac{1}{T^n(P_{SU,n})} \mathbb{1}\{ (s, u) \in T^n(P_{SU,n}) \}.
$$

(30)

Let $P_{Y}^{(P_S)}(y) = \sum_{s,u} P_{SU}^{(P_S)}(s, u)W^n(y|\phi^n(u, s), s)$ be the corresponding output distribution. Using a standard change of measure from constant-composition to i.i.d. (e.g. see [9, Ch. 2]), we can easily show that

$$
P_{Y}^{(P_S)}(y) \leq p_1(n) \prod_{i=1}^{n} P_{Y}^{(P_S)}(y_i),
$$

(31) where $p_1(n)$ is polynomial in $n$.

Recall that the decoder maximizes $i_{n}(P_S)$, given in (19). Using a well-known threshold-based non-asymptotic bound [2], we have for any $\gamma(P_S)$ that

$$
\mathbb{P}[E_2 | \tilde{P}_S = P_S, E_1] \leq \frac{1}{T^n(P_{SU,n})} \mathbb{1}\{ (s, u) \in T^n(P_{SU,n}) \}
$$

$$
+ ML^{(P_S)}P_{Y}^{(P_S)}(y) \mathbb{1}\{ i_{n}(P_S(U, Y) > \gamma(P_S)) \},
$$

(32) where $U \sim P_{SU}^{(P_S)}$ independently of $(S, U, Y)$. Using the change of measure given in (31), we can apply standard steps (e.g. see [3]) to upper bound the second term in (32) by $p_2(n)ML^{(P_S)}e^{-\gamma(P_S)}$, where $p_2(n)$ is polynomial in $n$. We can ensure that this term is $O(\frac{1}{n})$ by choosing

$$
\gamma(P_S) = \log ML^{(P_S)} + K_3 \log n,
$$

where $K_3$ is one higher than the degree of $p_2(n)$. Under this choice, and defining $K_4 = K_2 + K_3$, we obtain from (27) and (32) that

$$
\mathbb{P}[E_2 | \tilde{P}_S = P_S] \leq \frac{1}{T^n(P_{SU,n})} \mathbb{1}\{ (s, u) \in T^n(P_{SU,n}) \}
$$

$$
+ nI(P_S(U);S) + K_4 \log n + O(\frac{1}{n}).
$$

(33)

6) Application of the Berry-Esseen Theorem: Combining (28) and (33), we have for all $P_S \in \tilde{P}_n$ that

$$
\mathbb{P}[E_1 \cap E_2 | \tilde{P}_S = P_S] \leq \mathbb{P} \left[ i_{n}(P_S(U, Y)) \leq \log M 
$$

$$
+ nI(P_S(U);S) + K_4 \log n + O(\frac{1}{n}).
$$

(34)

In order to apply the Berry-Esseen theorem to the right-hand side of (34), we first compute the mean and variance of $i_{n}(P_S(U, Y))$, defined according to (19) and (29). The required third moment can easily be uniformly bounded in terms of the alphabet sizes [13, Appendix D]. We will use the fact that, by the symmetry of the constant-composition distribution in (30), the statistics of $i_{n}(P_S(U, Y))$ are unchanged upon conditioning on $(S, U) = (s, u)$ for some $(s, u) \in T^n(P_{SU,n})$. Using the joint distribution $P_{SU}^{(P_S)}$ defined in (12), it follows that

$$
E[i_{n}(P_S(U, Y))] = \sum_{u,y} P_{SU}^{(P_S)}(u,y)i_{n}(P_S(u,y))
$$

(35)

$$
= nI(P_S(U);Y) + O(1),
$$

(36) where (35) follows by expanding the expectation as a sum from 1 to $n$, and (36) follows from (13) and the definitions of $i_{n}(P_S(u,y))$ and $I(P_S(U);Y)$. A similar argument yields

$$
\text{Var}[i_{n}(P_S(U, Y))] = nE \left[ \text{Var}[i_{n}(P_S(U, Y)) | U, S] \right] + O(1)
$$

(37)

$$
\leq nV(P_S) + O(1).
$$

(38)

It should be noted that $V(P_S)$ is bounded away from zero for $P_S \in \tilde{P}_n$ and sufficiently large $n$, since $V(\pi) > 0$ by assumption in Theorem 1. Furthermore, the $O(1)$ terms in (36) and (38) are uniform in $P_S \in \tilde{P}_n$.

Using the definition of $I(P_S)$ in (14), we choose

$$
\log M = nI(\pi) - K_4 \log n - \beta_n,
$$

(39) where $\beta_n$ will be specified later, and will behave as $O(\sqrt{n})$. Combining (34), (36), (38) and (39), we have

$$
\mathbb{P}[E_1 \cup E_2 | \tilde{P}_S = P_S] \leq \mathbb{P} \left[ i_{n}(P_S(U, Y)) \leq nI(\pi) + nI(P_S(U);S) - \beta_n \right] + O(\frac{1}{n}).
$$

(40)

$$
\leq Q \left( \frac{\beta_n + nI(P_S) - nI(\pi) + K_5}{\sqrt{nV(P_S) + K_6}} \right) + O\left(\frac{1}{\sqrt{n}}\right).
$$

(41)
statistics of $\{P_n\}(U, Y)$, applying the Berry-Esseen theorem for independent and non-identically distributed variables [14, Sec. XVI.5], and introducing the constants $K_5$ and $K_6$ to represent the uniform $O(1)$ terms in (36) and (38).

7) Averaging Over the State Type: Substituting (41) into (23), we have

$$p_e \leq \sum_{P_S \in P_n} P[\hat{P}_S = P_S] Q(\frac{\beta + nI(P_S) - nI(\pi)}{\sqrt{nV(P_S)}}) + O\left(\frac{1}{\sqrt{n}}\right).$$

where we have factored the constants $K_5$ and $K_6$ into the remainder term using standard Taylor expansions along with the assumption $\beta_n = O(\sqrt{n})$; see [7] for details. Analogously to [8, Lemmas 17-18], we simplify (42) using two lemmas.

**Lemma 1.** For any $\beta_n = O(\sqrt{n})$, we have

$$\sum_{P_S \in P_n} P[\hat{P}_S = P_S] Q(\frac{\beta_n + nI(P_S) - nI(\pi)}{\sqrt{nV(P_S)}}) \leq \sum_{P_S \in P_n} P[\hat{P}_S = P_S] Q(\frac{\beta_n + nI(P_S) - nI(\pi)}{\sqrt{nV(\pi)}}) + O\left(\frac{\log n}{\sqrt{n}}\right).$$

**Proof:** This follows using standard Taylor expansions along with the definition of $P_n$ in (8) and the fact that $V(P_S)$ is continuously differentiable at $P_S = \pi$; see [7].

**Lemma 2.** For any $\beta_n$, we have

$$\sum_{P_S \in P_n} P[\hat{P}_S = P_S] Q(\frac{\beta_n + nI(P_S) - nI(\pi)}{\sqrt{nV(\pi)}}) \leq Q(\frac{\beta_n}{\sqrt{nV(\pi)}}) + O\left(\frac{\log n}{\sqrt{n}}\right).$$

where $V$ is defined in (6).

**Proof:** Using the expansion of $I(P_S)$ in terms of $I(S)$ and $\Delta(P_S)$ given in (15), along with the property given in (17), we can easily show that the left-hand side of (44) is upper bounded by

$$\sum_{P_S \in P_n} P[\hat{P}_S = P_S] Q(\frac{\beta_n + nI(\pi) + n\tilde{I}(P_S)}{\sqrt{nV(\pi)}}) + O\left(\frac{\log n}{\sqrt{n}}\right).$$

Since $\tilde{I}(P_S)$ is written in the form $\sum_s P_S(s)\psi(s)$, a trivial generalization of [8, Lemma 18] gives

$$\sum_{P_S} P[\hat{P}_S = P_S] \left| \frac{\beta_n + nI(P_S) - nI(\pi)}{\sqrt{nV(\pi)}} \right| + O\left(\frac{1}{\sqrt{n}}\right).$$

where $V^*(\pi) \triangleq \text{Var}_\pi[\psi(S)]$. Using (16), we see that $\psi(S) = \mathbb{E}[\psi(S)]$ and it follows that $V(\pi) + V^*(\pi)$ is equal to $V$, defined in (6). The proof is concluded by expanding the summation in (45) to be over all types, and substituting (46).

Using (42) along with Lemmas 1 and 2, we have

$$p_e \leq Q(\frac{\beta_n}{\sqrt{nV}}) + O\left(\frac{\log n}{\sqrt{n}}\right).$$

Setting $\beta_n = \epsilon$ and solving for $\beta_n$, we obtain

$$\beta_n = \sqrt{nV}Q^{-1}(\epsilon) + O(\log n).$$

Consistent with (42) and Lemma 1, we have $\beta_n = O(\sqrt{n})$. Substituting (48) into (39) yields the desired result with $V$ of the form given in (6).

By analyzing the Karush-Kuhn-Tucker (KKT) corresponding to the maximization in (1), it can be shown that the equality in (7) holds under any $Q_{U|S}$ which maximizes the objective for a given pair $(U, \phi)$ [7]. Since the parameters are capacity-achieving by assumption, this completes the proof.

**Acknowledgment**

I would like to thank Vincent Tan for many helpful comments and suggestions.

**References**


