Erasure and Undetected Error Probabilities in the Moderate Deviations Regime

Masahito Hayashi[†] and **Vincent Y. F. Tan**[‡]

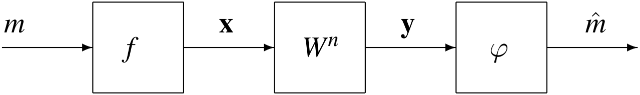
[†] Nagoya University and Center for Quantum Technologies, NUS

[‡] Electrical and Computer Engineering, National University of Singapore (NUS)

International Symposium on Information Theory (2015)

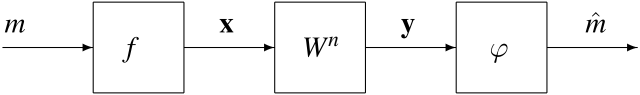
Full version: `http://arxiv.org/abs/1407.0142`

Channel Coding with the Erasure Option



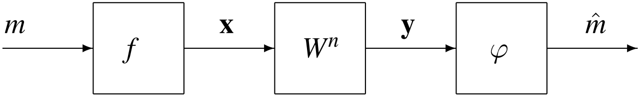
- An (n, M_n)-erasure code $\mathcal{C}_n$ consists of an encoder $f_n : \{1, \ldots, M_n\} \to \mathcal{X}^n$ and a decoder $\varphi : \mathcal{Y}^n \to \{1, \ldots, M_n\} \cup \{0\}$.

Channel Coding with the Erasure Option



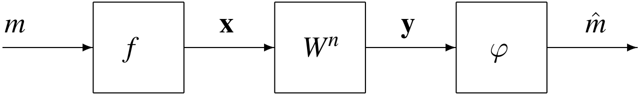
- An (n, M_n)-erasure code $\mathcal{C}_n$ consists of an encoder $f_n : \{1, \ldots, M_n\} \to \mathcal{X}^n$ and a decoder $\varphi : \mathcal{Y}^n \to \{1, \ldots, M_n\} \cup \{0\}$.
- Decoding region $\mathcal{D}_m = \varphi^{-1}(\mathbf{y})$ where $m = 0, 1, \ldots, M_n$

Channel Coding with the Erasure Option



- An (n, M_n)-erasure code $\mathcal{C}_n$ consists of an encoder $f_n : \{1, \ldots, M_n\} \to \mathcal{X}^n$ and a decoder $\varphi : \mathcal{Y}^n \to \{1, \ldots, M_n\} \cup \{0\}$.
- Decoding region $\mathcal{D}_m = \varphi^{-1}(\mathbf{y})$ where $m = 0, 1, \ldots, M_n$
- If $\mathbf{y} \in \mathcal{D}_m$, declare m was sent; if $\mathbf{y} \in \mathcal{D}_0$, declare an erasure

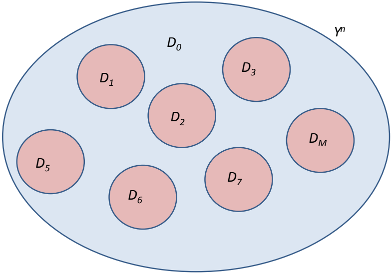
Channel Coding with the Erasure Option



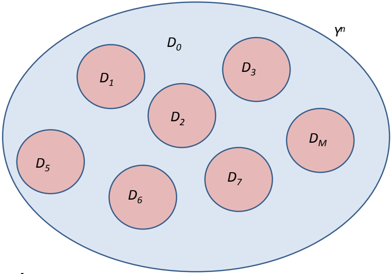
- An (n, M_n)-erasure code $\mathcal{C}_n$ consists of an encoder $f_n : \{1, \ldots, M_n\} \to \mathcal{X}^n$ and a decoder $\varphi : \mathcal{Y}^n \to \{1, \ldots, M_n\} \cup \{0\}$.
- Decoding region $\mathcal{D}_m = \varphi^{-1}(\mathbf{y})$ where $m = 0, 1, \ldots, M_n$
- If $\mathbf{y} \in \mathcal{D}_m$, declare m was sent; if $\mathbf{y} \in \mathcal{D}_0$, declare an erasure
- We assume W^n is a DMC, i.e.,

$$W^n(\mathbf{y} \,|\, \mathbf{x}) = \prod_{i=1}^{n} W(y_i \,|\, x_i)$$

Channel Coding with the Erasure Option : Illustration



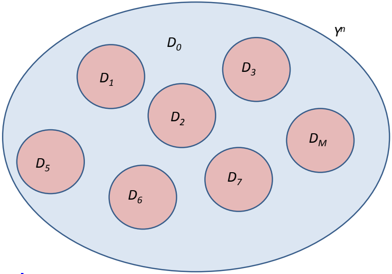
Channel Coding with the Erasure Option : Illustration



- Probability of total error

$$\Pr(\mathcal{E}_1 \,|\, \mathcal{C}_n) = \frac{1}{M_n} \sum_{m=1}^{M_n} \sum_{\mathbf{y} \in \mathcal{D}_m^c} W^n(\mathbf{y} \,|\, f_n(m)).$$

Channel Coding with the Erasure Option : Illustration



- Probability of total error

$$\Pr(\mathcal{E}_1 \,|\, \mathcal{C}_n) = \frac{1}{M_n} \sum_{m=1}^{M_n} \sum_{\mathbf{y} \in \mathcal{D}_m^c} W^n(\mathbf{y} \,|\, f_n(m)).$$

- Probability of undetected error

$$\Pr(\mathcal{E}_2 \,|\, \mathcal{C}_n) = \frac{1}{M_n} \sum_{m=1}^{M_n} \sum_{\mathbf{y} \in \mathcal{D}_m} \sum_{m' \neq m} W^n(\mathbf{y} \,|\, f_n(m')).$$

Some Historical Remarks on Erasure Decoding

- Forney (1968) derived exponential error bounds for erasure and list decoding using Gallager-type bounding techniques

Some Historical Remarks on Erasure Decoding

- Forney (1968) derived exponential error bounds for erasure and list decoding using Gallager-type bounding techniques

- Moulin (2009) presented improved universal error exponents over those presented in Csiszár-Körner

Some Historical Remarks on Erasure Decoding

- Forney (1968) derived exponential error bounds for erasure and list decoding using Gallager-type bounding techniques
- Moulin (2009) presented improved universal error exponents over those presented in Csiszár-Körner
- Merhav (2008) derived error exponents based on the type class enumerator method that are at least as good as Forney's
- Somekh-Baruch and Merhav (2011) showed that the the type class enumerator method is ensemble-tight

Some Historical Remarks on Erasure Decoding

- Forney (1968) derived exponential error bounds for erasure and list decoding using Gallager-type bounding techniques
- Moulin (2009) presented improved universal error exponents over those presented in Csiszár-Körner
- Merhav (2008) derived error exponents based on the type class enumerator method that are at least as good as Forney's
- Somekh-Baruch and Merhav (2011) showed that the the type class enumerator method is ensemble-tight
- Tan-Moulin (2014) considered non-vanishing total and undetected error version of this problem.

Various Asymptotic Regimes

- We consider the message size to be

$$\log M_n = nC - an^{1-t}, \quad 0 < t \leq 1/2$$

Various Asymptotic Regimes

- We consider the message size to be

$$\log M_n = nC - an^{1-t}, \quad 0 < t \le 1/2$$

- When $t = 1/2$, this is the normal approximation regime and

$$\epsilon_n^* \approx \Phi\left(-\frac{a}{\sqrt{V}}\right), \qquad \text{non-vanishing}$$

Strassen (1962), Hayashi (2009) and PPV (2010).

Various Asymptotic Regimes

- We consider the message size to be

$$\log M_n = nC - an^{1-t}, \quad 0 < t \le 1/2$$

- When $t = 1/2$, this is the normal approximation regime and

$$\epsilon_n^* \approx \Phi\left(-\frac{a}{\sqrt{V}}\right), \qquad \text{non-vanishing}$$

Strassen (1962), Hayashi (2009) and PPV (2010).

- When $0 < t < 1/2$, this is the moderate deviations regime and

$$\epsilon_n^* \approx \exp\left(-n^{1-2t}\frac{a^2}{2V}\right), \qquad \text{sub-exponential}$$

Altuğ-Wagner (2014) and Polyanskiy-Verdú (2010)

Our Contribution : Two Asymptotic Regimes

- Moderate Deviations Regime with $0 < t < 1/2$:

$$\Pr(\mathcal{E}_1 \,|\, \mathcal{C}_n) \approx \exp\left(-\Theta(n^{1-2t})\right), \qquad \Pr(\mathcal{E}_2 \,|\, \mathcal{C}_n) \approx \exp\left(-\Theta(n^{1-t})\right)$$

Our Contribution : Two Asymptotic Regimes

- Moderate Deviations Regime with $0 < t < 1/2$:

$$\Pr(\mathcal{E}_1 \,|\, \mathcal{C}_n) \approx \exp\left(-\Theta(n^{1-2t})\right), \qquad \Pr(\mathcal{E}_2 \,|\, \mathcal{C}_n) \approx \exp\left(-\Theta(n^{1-t})\right)$$

- Mixed Regime with $t = 1/2$:

$$\Pr(\mathcal{E}_1 \,|\, \mathcal{C}_n) \approx \Phi\left(\ldots\right), \qquad \Pr(\mathcal{E}_2 \,|\, \mathcal{C}_n) \approx \exp\left(-\Theta(n^{1/2})\right)$$

Our Contribution : Two Asymptotic Regimes

- Moderate Deviations Regime with $0 < t < 1/2$:

$$\Pr(\mathcal{E}_1 \,|\, \mathcal{C}_n) \approx \exp\left(-\Theta(n^{1-2t})\right), \qquad \Pr(\mathcal{E}_2 \,|\, \mathcal{C}_n) \approx \exp\left(-\Theta(n^{1-t})\right)$$

- Mixed Regime with $t = 1/2$:

$$\Pr(\mathcal{E}_1 \,|\, \mathcal{C}_n) \approx \Phi\left(\ldots\right), \qquad \Pr(\mathcal{E}_2 \,|\, \mathcal{C}_n) \approx \exp\left(-\Theta(n^{1/2})\right)$$

- Similar to Somekh-Baruch and Merhav (2011), we seek ensemble-tight results, i.e., we seek asymptotic equalities for

$$\mathbb{E}_{\mathcal{C}_n}[\Pr(\mathcal{E}_1 \,|\, \mathcal{C}_n)], \quad \text{and} \quad \mathbb{E}_{\mathcal{C}_n}[\Pr(\mathcal{E}_2 \,|\, \mathcal{C}_n)]$$

where $\mathbb{E}_{\mathcal{C}_n}[\cdot]$ denotes expectation over a random codebook.

Preliminary Quantities

- Mutual information

$$I(P,W) = \sum_x P(x) \sum_y W(y|x) \log \frac{W(y|x)}{PW(y)}$$

- Conditional information variance

$$V(P,W) = \sum_x P(x) \sum_y W(y|x) \left[\log \frac{W(y|x)}{PW(y)} - D(W(\cdot|x) \| PW) \right]^2$$

Preliminary Quantities

- Mutual information

$$I(P,W) = \sum_x P(x) \sum_y W(y|x) \log \frac{W(y|x)}{PW(y)}$$

- Conditional information variance

$$V(P,W) = \sum_x P(x) \sum_y W(y|x) \left[\log \frac{W(y|x)}{PW(y)} - D(W(\cdot|x) \| PW) \right]^2$$

- Minimum and maximum conditional information variances

$$V_{\min}(W) = \min_{P:I(P,W)=C} V(P,W), \qquad V_{\max}(W) = \max_{P:I(P,W)=C} V(P,W)$$

Preliminary Quantities

- Mutual information

$$I(P,W) = \sum_x P(x) \sum_y W(y|x) \log \frac{W(y|x)}{PW(y)}$$

- Conditional information variance

$$V(P,W) = \sum_x P(x) \sum_y W(y|x) \left[\log \frac{W(y|x)}{PW(y)} - D(W(\cdot|x) \| PW) \right]^2$$

- Minimum and maximum conditional information variances

$$V_{\min}(W) = \min_{P: I(P,W)=C} V(P,W), \qquad V_{\max}(W) = \max_{P: I(P,W)=C} V(P,W)$$

- Always assume $V_{\min}(W) > 0$.

Direct Results Using Information Spectrum : Moderate

Theorem (Moderate Deviations Direct)

Let $0 < t < 1/2$ *and* $a > b > 0$. *Let*

$$\log M_n = nC - an^{1-t}$$

Direct Results Using Information Spectrum : Moderate

Theorem (Moderate Deviations Direct)

Let $0 < t < 1/2$ *and* $a > b > 0$. *Let*

$$\log M_n = nC - an^{1-t}$$

There exists a sequence of codebooks $\mathcal{C}_n$ *with* M_n *codewords such that*

$$\lim_{n\to\infty} -\frac{1}{n^{1-2t}} \log \Pr(\mathcal{E}_1 \,|\, \mathcal{C}_n) = \frac{(a-b)^2}{2V_{\min}(W)}$$

$$\liminf_{n\to\infty} -\frac{1}{n^{1-t}} \log \Pr(\mathcal{E}_2 \,|\, \mathcal{C}_n) \geq b$$

Direct Results Using Information Spectrum : Moderate

Theorem (Moderate Deviations Direct)

Let $0 < t < 1/2$ *and* $a > b > 0$. *Let*

$$\log M_n = nC - an^{1-t}$$

There exists a sequence of codebooks $\mathcal{C}_n$ *with* M_n *codewords such that*

$$\lim_{n\to\infty} -\frac{1}{n^{1-2t}} \log \Pr(\mathcal{E}_1 \,|\, \mathcal{C}_n) = \frac{(a-b)^2}{2V_{\min}(W)}$$

$$\liminf_{n\to\infty} -\frac{1}{n^{1-t}} \log \Pr(\mathcal{E}_2 \,|\, \mathcal{C}_n) \geq b$$

Decoding regions are reminiscent of information spectrum analysis:

$$\tilde{\mathcal{D}}_m := \left\{ \mathbf{y} : \log \frac{W^n(\mathbf{y}|\mathbf{x}_m)}{(PW)^n(\mathbf{y})} \geq \log M_n + bn^{1-t} \right\}$$

Remarks on Direct Results : Moderate

Total $\Pr(\mathcal{E}_1 \,|\, \mathcal{C}_n) \approx \exp\left(-n^{1-2t} \frac{(a-b)^2}{2V_{\min}(W)} \right)$

Undetected $\Pr(\mathcal{E}_2 \,|\, \mathcal{C}_n) \lesssim \exp\left(-n^{1-t} b \right)$

- Undetected error probability $\ll$ Total error probability

Remarks on Direct Results : Moderate

$$\text{Total} \qquad \Pr(\mathcal{E}_1 \,|\, \mathcal{C}_n) \approx \exp\left(-n^{1-2t}\frac{(a-b)^2}{2V_{\min}(W)}\right)$$

$$\text{Undetected} \qquad \Pr(\mathcal{E}_2 \,|\, \mathcal{C}_n) \lesssim \exp\left(-n^{1-t}b\right)$$

- Undetected error probability $\ll$ Total error probability
- Note $a > b$. If a increases, fewer number of codewords and so total error $\Pr(\mathcal{E}_1 \,|\, \mathcal{C}_n)$ decays faster.

Remarks on Direct Results : Moderate

$$\text{Total} \qquad \Pr(\mathcal{E}_1 \,|\, \mathcal{C}_n) \approx \exp\left(-n^{1-2t}\frac{(a-b)^2}{2V_{\min}(W)}\right)$$

$$\text{Undetected} \qquad \Pr(\mathcal{E}_2 \,|\, \mathcal{C}_n) \lesssim \exp\left(-n^{1-t}b\right)$$

- Undetected error probability $\ll$ Total error probability
- Note $a > b$. If a increases, fewer number of codewords and so total error $\Pr(\mathcal{E}_1 \,|\, \mathcal{C}_n)$ decays faster.
- If b increases, decoding regions for $m \geq 1$

$$\tilde{\mathcal{D}}_m := \left\{\mathbf{y} : \log\frac{W^n(\mathbf{y}|\mathbf{x}_m)}{(PW)^n(\mathbf{y})} \geq \log M_n + bn^{1-t}\right\}$$

  become smaller

Remarks on Direct Results : Moderate

$$\text{Total} \qquad \Pr(\mathcal{E}_1 \,|\, \mathcal{C}_n) \approx \exp\left(-n^{1-2t}\frac{(a-b)^2}{2V_{\min}(W)}\right)$$

$$\text{Undetected} \qquad \Pr(\mathcal{E}_2 \,|\, \mathcal{C}_n) \lesssim \exp\left(-n^{1-t} b\right)$$

- Undetected error probability $\ll$ Total error probability
- Note $a > b$. If a increases, fewer number of codewords and so total error $\Pr(\mathcal{E}_1 \,|\, \mathcal{C}_n)$ decays faster.
- If b increases, decoding regions for $m \geq 1$

$$\tilde{\mathcal{D}}_m := \left\{ \mathbf{y} : \log \frac{W^n(\mathbf{y}|\mathbf{x}_m)}{(PW)^n(\mathbf{y})} \geq \log M_n + bn^{1-t} \right\}$$

  become smaller
- Undetected error $\Pr(\mathcal{E}_2 \,|\, \mathcal{C}_n)$ decays faster.

Direct Results Using Information Spectrum : Mixed

Theorem (Mixed Direct)

Let

$$\log M_n = nC - a\sqrt{n}.$$

Direct Results Using Information Spectrum : Mixed

Theorem (Mixed Direct)

Let

$$\log M_n = nC - a\sqrt{n}.$$

There exists a sequence of codebooks $\mathcal{C}_n$ *with* M_n *codewords such that*

$$\Pr(\mathcal{E}_1 \,|\, \mathcal{C}_n) \to \begin{cases} \Phi\left(\frac{b-a}{\sqrt{V_{\max}(W)}}\right) & a \le 0 \\ \Phi\left(\frac{b-a}{\sqrt{V_{\min}(W)}}\right) & a > 0 \end{cases}$$

Direct Results Using Information Spectrum : Mixed

Theorem (Mixed Direct)

Let

$$\log M_n = nC - a\sqrt{n}.$$

There exists a sequence of codebooks $\mathcal{C}_n$ *with* M_n *codewords such that*

$$\Pr(\mathcal{E}_1 \,|\, \mathcal{C}_n) \to \begin{cases} \Phi\left(\frac{b-a}{\sqrt{V_{\max}(W)}}\right) & a \le 0 \\ \Phi\left(\frac{b-a}{\sqrt{V_{\min}(W)}}\right) & a > 0 \end{cases}$$

and

$$\Pr(\mathcal{E}_2 \,|\, \mathcal{C}_n) \lesssim \exp\left(-\sqrt{n}\, b\right)$$

Ensemble Tightness?

- Are the results tight? Examine $\mathbb{E}_{\mathcal{C}_n}[\Pr(\mathcal{E}_j \,|\, \mathcal{C}_n)], j = 1, 2$.

Ensemble Tightness?

- Are the results tight? Examine $\mathbb{E}_{\mathcal{C}_n}[\Pr(\mathcal{E}_j \,|\, \mathcal{C}_n)], j = 1, 2$.
- Restrict attention to additive DMCs, i.e.,

$$Y_i = X_i \oplus Z_i$$

and $Z_i \sim P(\cdot) \in \mathcal{P}(\mathbb{F}_q)$, some noise distribution.

Ensemble Tightness?

- Are the results tight? Examine $\mathbb{E}_{\mathcal{C}_n}[\Pr(\mathcal{E}_j \,|\, \mathcal{C}_n)], j = 1, 2$.
- Restrict attention to additive DMCs, i.e.,

$$Y_i = X_i \oplus Z_i$$

  and $Z_i \sim P(\cdot) \in \mathcal{P}(\mathbb{F}_q)$, some noise distribution.
- Furthermore $\{Z_i\}_{i \geq 1}$ are i.i.d.

Ensemble Tightness?

- Are the results tight? Examine $\mathbb{E}_{\mathcal{C}_n}[\Pr(\mathcal{E}_j \,|\, \mathcal{C}_n)], j = 1, 2$.
- Restrict attention to additive DMCs, i.e.,

$$Y_i = X_i \oplus Z_i$$

and $Z_i \sim P(\cdot) \in \mathcal{P}(\mathbb{F}_q)$, some noise distribution.

- Furthermore $\{Z_i\}_{i \geq 1}$ are i.i.d.
- Varentropy

$$V_{\min}(W) = V_{\max}(W) = V(P) = \mathrm{var}\left[\log \frac{1}{P(Z)}\right]$$

Ensemble Tightness?

- Are the results tight? Examine $\mathbb{E}_{\mathcal{C}_n}[\Pr(\mathcal{E}_j \mid \mathcal{C}_n)], j = 1, 2$.
- Restrict attention to additive DMCs, i.e.,

$$Y_i = X_i \oplus Z_i$$

and $Z_i \sim P(\cdot) \in \mathcal{P}(\mathbb{F}_q)$, some noise distribution.

- Furthermore $\{Z_i\}_{i \ge 1}$ are i.i.d.
- Varentropy

$$V_{\min}(W) = V_{\max}(W) = V(P) = \mathrm{var}\left[\log \frac{1}{P(Z)}\right]$$

- Also need a stronger decoder; c.f. Forney (1968)

$$\mathcal{D}_m := \left\{ \mathbf{y} : \frac{W^n(\mathbf{y}|\mathbf{x}_m)}{\sum_{m' \neq m} W^n(\mathbf{y}|\mathbf{x}_{m'})} \ge \exp(nT_n) \right\}$$

T_n : Blocklength-varying threshold depending on regime.

Ensemble Tightness

Theorem (Moderate Deviations Regime)

Let $0 < t < 1/2$ *and* $a > b > 0$. *Let*

$$\log M_n = nC - an^{1-t}.$$

Ensemble Tightness

Theorem (Moderate Deviations Regime)

Let $0 < t < 1/2$ *and* $a > b > 0$*. Let*

$$\log M_n = nC - an^{1-t}.$$

Consider a sequence of random codebooks $\mathcal{C}_n$ *with* M_n *codewords each drawn uniformly at random from* $\mathbb{F}_q^n$*.*

Ensemble Tightness

Theorem (Moderate Deviations Regime)

Let $0 < t < 1/2$ *and* $a > b > 0$. *Let*

$$\log M_n = nC - an^{1-t}.$$

Consider a sequence of random codebooks $\mathcal{C}_n$ *with* M_n *codewords each drawn uniformly at random from* $\mathbb{F}_q^n$. *Choose the threshold as*

$$T_n = \frac{b}{n^t},$$

Ensemble Tightness

Theorem (Moderate Deviations Regime)

Let $0 < t < 1/2$ and $a > b > 0$. Let

$$\log M_n = nC - an^{1-t}.$$

Consider a sequence of random codebooks $\mathcal{C}_n$ with M_n codewords each drawn uniformly at random from $\mathbb{F}_q^n$. Choose the threshold as

$$T_n = \frac{b}{n^t},$$

Then,

$$\lim_{n\to\infty} -\frac{1}{n^{1-2t}} \log \mathbb{E}_{\mathcal{C}_n}[\Pr(\mathcal{E}_1 \,|\, \mathcal{C}_n)] = \frac{(a-b)^2}{2V(P)}$$

$$\lim_{n\to\infty} -\frac{1}{n^{1-t}} \log \mathbb{E}_{\mathcal{C}_n}[\Pr(\mathcal{E}_2 \,|\, \mathcal{C}_n)] = b$$

Remarks on Ensemble Tight Result

- Similar result for mixed regime, i.e.,

$$\mathbb{E}_{\mathcal{C}_n}[\Pr(\mathcal{E}_1 \,|\, \mathcal{C}_n)] \approx \Phi\left(\frac{b-a}{\sqrt{V(P)}}\right)$$

$$\mathbb{E}_{\mathcal{C}_n}[\Pr(\mathcal{E}_2 \,|\, \mathcal{C}_n)] \approx \exp\left(-\sqrt{n}\, b\right)$$

Remarks on Ensemble Tight Result

- Similar result for mixed regime, i.e.,

$$\mathbb{E}_{\mathcal{C}_n}[\Pr(\mathcal{E}_1 \,|\, \mathcal{C}_n)] \approx \Phi\left(\frac{b-a}{\sqrt{V(P)}}\right)$$

$$\mathbb{E}_{\mathcal{C}_n}[\Pr(\mathcal{E}_2 \,|\, \mathcal{C}_n)] \approx \exp\left(-\sqrt{n}\, b\right)$$

- Observe that with $T_n = \frac{b}{n^t}$,

$$\mathcal{D}_m := \left\{ \mathbf{y} : \log \frac{W^n(\mathbf{y}|\mathbf{x}_m)}{\sum_{m' \neq m} W^n(\mathbf{y}|\mathbf{x}_{m'})} \geq b n^{1-t} \right\}$$

Remarks on Ensemble Tight Result

- Similar result for mixed regime, i.e.,

$$\mathbb{E}_{\mathcal{C}_n}[\Pr(\mathcal{E}_1 \,|\, \mathcal{C}_n)] \approx \Phi\left(\frac{b-a}{\sqrt{V(P)}}\right)$$
$$\mathbb{E}_{\mathcal{C}_n}[\Pr(\mathcal{E}_2 \,|\, \mathcal{C}_n)] \approx \exp\left(-\sqrt{n}\, b\right)$$

- Observe that with $T_n = \frac{b}{n^t}$,

$$\mathcal{D}_m := \left\{ \mathbf{y} : \log \frac{W^n(\mathbf{y}|\mathbf{x}_m)}{\sum_{m' \neq m} W^n(\mathbf{y}|\mathbf{x}_{m'})} \geq b n^{1-t} \right\}$$

- As b increases towards a, decoding regions become smaller; erasure probability higher; undetected probability smaller

Main Analysis Technique

- Recall Forney's optimum decoding rule

$$\mathcal{D}_m := \left\{ \mathbf{y} : \frac{W^n(\mathbf{y}|\mathbf{x}_m)}{\sum_{m' \neq m} W^n(\mathbf{y}|\mathbf{x}_{m'})} \geq \exp(nT_n) \right\}$$

Main Analysis Technique

- Recall Forney's optimum decoding rule

$$\mathcal{D}_m := \left\{ \mathbf{y} : \frac{W^n(\mathbf{y}|\mathbf{x}_m)}{\sum_{m' \neq m} W^n(\mathbf{y}|\mathbf{x}_{m'})} \geq \exp(nT_n) \right\}$$

- Suppose $m = 1$ is sent. We should analyze behavior of the random variable

$$F_n := \log \left(\sum_{m' \neq 1} W^n(Y^n|X^n_{m'}) \right) - \log W^n(Y^n|X^n_1)$$

Main Analysis Technique

- Recall Forney's optimum decoding rule

$$\mathcal{D}_m := \left\{ \mathbf{y} : \frac{W^n(\mathbf{y}|\mathbf{x}_m)}{\sum_{m' \neq m} W^n(\mathbf{y}|\mathbf{x}_{m'})} \geq \exp(nT_n) \right\}$$

- Suppose $m = 1$ is sent. We should analyze behavior of the random variable

$$F_n := \log \left(\sum_{m' \neq 1} W^n(Y^n|X^n_{m'}) \right) - \log W^n(Y^n|X^n_1)$$

- Suffices to understand the cumulant generating function

$$\phi_n(s) := \log \mathbb{E}\left[\exp(sF_n)\right]$$

Here s doesn't have to be a constant; can vary with n

Cumulant Generating Function Asymptotics

Lemma

Consider $\phi_n(s) := \log \mathbb{E}\left[\exp(sF_n)\right]$. *Let* $0 < t \le 1/2$. *We have*

$$\phi_n\left(\frac{u}{n^t}\right) = \left(-au + u^2 \frac{V(P)}{2}\right) n^{1-2t} + o(1)$$

for any constant $u > 0$.

Cumulant Generating Function Asymptotics

Lemma

Consider $\phi_n(s) := \log \mathbb{E}\left[\exp(sF_n)\right]$. *Let* $0 < t \le 1/2$. *We have*

$$\phi_n\left(\frac{u}{n^t}\right) = \left(-au + u^2\frac{V(P)}{2}\right) n^{1-2t} + o(1)$$

for any constant $u > 0$.

- Type class enumerator method (cf. Merhav)
- Concentration bounds

Cumulant Generating Function Asymptotics

Lemma

Consider $\phi_n(s) := \log \mathbb{E}\left[\exp(sF_n)\right]$*. Let* $0 < t \le 1/2$*. We have*

$$\phi_n\left(\frac{u}{n^t}\right) = \left(-au + u^2 \frac{V(P)}{2}\right) n^{1-2t} + o(1)$$

for any constant $u > 0$*.*

- Type class enumerator method (cf. Merhav)
- Concentration bounds
- For example, if $t = 1/2$, we have

$$\lim_{n\to\infty} \phi_n\left(\frac{u}{\sqrt{n}}\right) = -au + u^2 \frac{V(P)}{2} \quad \Rightarrow \quad \text{Gaussian}$$

Modified Gärtner-Ellis Theorem

- Exponent of expected total error probability $\mathbb{E}_{\mathcal{C}_n}[\Pr(\mathcal{E}_1 \,|\, \mathcal{C}_n)]$ is

$$-\log \mathbb{E}_{\mathcal{C}_n}[\Pr(\mathcal{E}_1 \,|\, \mathcal{C}_n)] = -\log \Pr(F_n > -bn^{1-t})$$

Modified Gärtner-Ellis Theorem

- Exponent of expected total error probability $\mathbb{E}_{\mathcal{C}_n}[\Pr(\mathcal{E}_1 \,|\, \mathcal{C}_n)]$ is

$$-\log \mathbb{E}_{\mathcal{C}_n}[\Pr(\mathcal{E}_1 \,|\, \mathcal{C}_n)] = -\log \Pr(F_n > -bn^{1-t})$$

- Asymptotic behavior of the cumulant generating function of F_n

$$\lim_{n\to\infty} \frac{1}{n^{1-2t}} \phi_n\left(\frac{u}{n^t}\right) = -au + u^2 \frac{V(P)}{2}$$

Modified Gärtner-Ellis Theorem

- Exponent of expected total error probability $\mathbb{E}_{\mathcal{C}_n}[\Pr(\mathcal{E}_1 \,|\, \mathcal{C}_n)]$ is

$$-\log \mathbb{E}_{\mathcal{C}_n}[\Pr(\mathcal{E}_1 \,|\, \mathcal{C}_n)] = -\log \Pr(F_n > -bn^{1-t})$$

- Asymptotic behavior of the cumulant generating function of F_n

$$\lim_{n\to\infty} \frac{1}{n^{1-2t}} \phi_n\left(\frac{u}{n^t}\right) = -au + u^2 \frac{V(P)}{2}$$

- By a suitably modified of the Gärtner-Ellis theorem

$$-\log \mathbb{E}_{\mathcal{C}_n}[\Pr(\mathcal{E}_1 \,|\, \mathcal{C}_n)] = \frac{(a-b)^2}{2V(P)} n^{1-2t} + o(n^{1-2t}).$$

Modified Gärtner-Ellis Theorem

- Exponent of expected total error probability $\mathbb{E}_{\mathcal{C}_n}[\Pr(\mathcal{E}_1 \,|\, \mathcal{C}_n)]$ is

$$-\log \mathbb{E}_{\mathcal{C}_n}[\Pr(\mathcal{E}_1 \,|\, \mathcal{C}_n)] = -\log \Pr(F_n > -bn^{1-t})$$

- Asymptotic behavior of the cumulant generating function of F_n

$$\lim_{n\to\infty} \frac{1}{n^{1-2t}} \phi_n\left(\frac{u}{n^t}\right) = -au + u^2 \frac{V(P)}{2}$$

- By a suitably modified of the Gärtner-Ellis theorem

$$-\log \mathbb{E}_{\mathcal{C}_n}[\Pr(\mathcal{E}_1 \,|\, \mathcal{C}_n)] = \frac{(a-b)^2}{2V(P)} n^{1-2t} + o(n^{1-2t}).$$

- For $\mathbb{E}_{\mathcal{C}_n}[\Pr(\mathcal{E}_2 \,|\, \mathcal{C}_n)]$, we need a change-of-measure and shifted form of Gärtner-Ellis theorem.

Conclusion and Further Work

- We analyzed the total and undetected errors whose scalings are

$$\Pr(\mathcal{E}_1 \,|\, \mathcal{C}_n) \approx \exp\left(-\Theta(n^{1-2t})\right), \qquad \Pr(\mathcal{E}_2 \,|\, \mathcal{C}_n) \approx \exp\left(-\Theta(n^{1-t})\right)$$

$$\Pr(\mathcal{E}_1 \,|\, \mathcal{C}_n) \approx \Phi\left(\ldots\right), \qquad \Pr(\mathcal{E}_2 \,|\, \mathcal{C}_n) \approx \exp\left(-\Theta(n^{1/2})\right)$$

Conclusion and Further Work

- We analyzed the total and undetected errors whose scalings are

$$\Pr(\mathcal{E}_1 \,|\, \mathcal{C}_n) \approx \exp\left(-\Theta(n^{1-2t})\right), \qquad \Pr(\mathcal{E}_2 \,|\, \mathcal{C}_n) \approx \exp\left(-\Theta(n^{1-t})\right)$$

$$\Pr(\mathcal{E}_1 \,|\, \mathcal{C}_n) \approx \Phi\left(\dots\right), \qquad \Pr(\mathcal{E}_2 \,|\, \mathcal{C}_n) \approx \exp\left(-\Theta(n^{1/2})\right)$$

- Ensemble-tight results for additive DMCs.

Conclusion and Further Work

- We analyzed the total and undetected errors whose scalings are

$$\Pr(\mathcal{E}_1 \,|\, \mathcal{C}_n) \approx \exp\left(-\Theta(n^{1-2t})\right), \qquad \Pr(\mathcal{E}_2 \,|\, \mathcal{C}_n) \approx \exp\left(-\Theta(n^{1-t})\right)$$

$$\Pr(\mathcal{E}_1 \,|\, \mathcal{C}_n) \approx \Phi\left(\ldots\right), \qquad \Pr(\mathcal{E}_2 \,|\, \mathcal{C}_n) \approx \exp\left(-\Theta(n^{1/2})\right)$$

- Ensemble-tight results for additive DMCs.
- Relax assumption of additivity in future; Need to probably make Assumption 1 of Merhav (2008).

Conclusion and Further Work

- We analyzed the total and undetected errors whose scalings are

$$\Pr(\mathcal{E}_1 \,|\, \mathcal{C}_n) \approx \exp\left(-\Theta(n^{1-2t})\right), \qquad \Pr(\mathcal{E}_2 \,|\, \mathcal{C}_n) \approx \exp\left(-\Theta(n^{1-t})\right)$$

$$\Pr(\mathcal{E}_1 \,|\, \mathcal{C}_n) \approx \Phi\left(\ldots\right), \qquad \Pr(\mathcal{E}_2 \,|\, \mathcal{C}_n) \approx \exp\left(-\Theta(n^{1/2})\right)$$

- Ensemble-tight results for additive DMCs.
- Relax assumption of additivity in future; Need to probably make Assumption 1 of Merhav (2008).
- Higher-order terms in the asymptotic expansions of

$$-\log \mathbb{E}_{\mathcal{C}_n}[\Pr(\mathcal{E}_1 \,|\, \mathcal{C}_n)], \qquad \text{and} \qquad -\log \mathbb{E}_{\mathcal{C}_n}[\Pr(\mathcal{E}_2 \,|\, \mathcal{C}_n)].$$

Conclusion and Further Work

- We analyzed the total and undetected errors whose scalings are

$$\Pr(\mathcal{E}_1 \,|\, \mathcal{C}_n) \approx \exp\left(-\Theta(n^{1-2t})\right), \qquad \Pr(\mathcal{E}_2 \,|\, \mathcal{C}_n) \approx \exp\left(-\Theta(n^{1-t})\right)$$

$$\Pr(\mathcal{E}_1 \,|\, \mathcal{C}_n) \approx \Phi\left(\dots\right), \qquad \Pr(\mathcal{E}_2 \,|\, \mathcal{C}_n) \approx \exp\left(-\Theta(n^{1/2})\right)$$

- Ensemble-tight results for additive DMCs.
- Relax assumption of additivity in future; Need to probably make Assumption 1 of Merhav (2008).
- Higher-order terms in the asymptotic expansions of

$$-\log \mathbb{E}_{\mathcal{C}_n}[\Pr(\mathcal{E}_1 \,|\, \mathcal{C}_n)], \qquad \text{and} \qquad -\log \mathbb{E}_{\mathcal{C}_n}[\Pr(\mathcal{E}_2 \,|\, \mathcal{C}_n)].$$

- Full version: `http://arxiv.org/abs/1407.0142`.