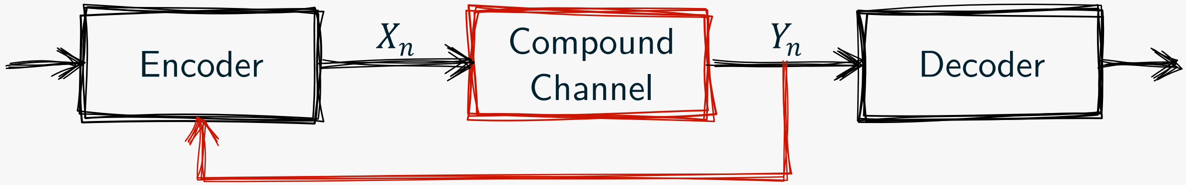
Error Exponents of Compound Channel with Feedback

Aditya Mahajan and Sekhar Tatikonda
Yale University

ISIT 2010

Problem Setup



Compound channel

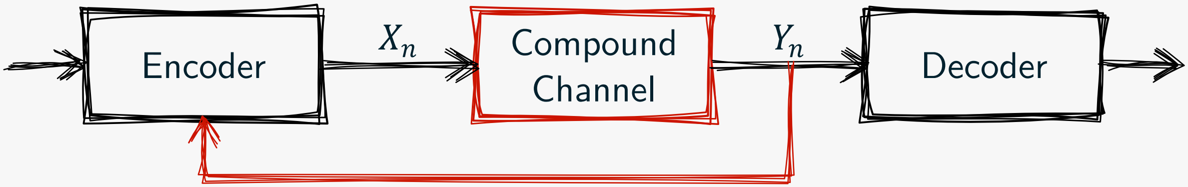
- Channel is memoryless

$$\mathbb{P}(Y_n \mid X^n, Y^{n-1}) = Q_\circ(Y_n \mid X_n)$$

- Channel is not known completely

$$Q_\circ \in \boldsymbol{\mathcal{Q}} := \{Q_1, Q_2, \ldots, Q_L\} \qquad \text{(Family of DMCs)}$$

Problem Setup



- Compound channel
  - Channel is memoryless

$$\mathbb{P}(Y_n \mid X^n, Y^{n-1}) = Q_\circ(Y_n \mid X_n)$$

  - Channel is not known completely

$$Q_\circ \in \boldsymbol{\mathcal{Q}} := \{Q_1, Q_2, \ldots, Q_L\} \qquad \text{(Family of DMCs)}$$

- Variable length communication
  - Capacity? (easy)
  - Error Exponents? (this talk)

Opportunistic Capacity

- Notation

  For a (variable length) coding scheme S

  - $P_\ell^{(S)}$ = Prob of error when $Q_\circ = Q_\ell$
  - $R_\ell^{(S)}$ = Rate when $Q_\circ = Q_\ell$
  - $\tau^{(S)}$ = Communication length (stopping time on $\{Y_n\}_{n\in\mathbb{N}}$)

Opportunistic Capacity

Notation

For a (variable length) coding scheme S

- $P_\ell^{(S)}$ = Prob of error when $Q_\circ = Q_\ell$
- $R_\ell^{(S)}$ = Rate when $Q_\circ = Q_\ell$
- $\tau^{(S)}$ = Communication length (stopping time on $\{Y_n\}_{n\in\mathbb{N}}$)

Opportunistic Achievability (main idea)

A rate $(R_1, R_2, \ldots, R_L)$ is achievable if $\exists$ a sequence of coding schemes $\{S_n\}_{n\in\mathbb{N}}$ such that for any $\varepsilon > 0$, $\exists n_\circ(\varepsilon)$ so that for all $n > n_\circ(\varepsilon)$

$$P_\ell^{(S_n)} < \varepsilon \quad \text{and} \quad R_\ell^{(S_n)} > R_\ell - \varepsilon \quad \text{for all } \ell = 1, 2, \ldots, L$$

Opportunistic Capacity

Notation

For a (variable length) coding scheme S

- $P_\ell^{(S)}$ = Prob of error when $Q_\circ = Q_\ell$
- $R_\ell^{(S)}$ = Rate when $Q_\circ = Q_\ell$
- $\tau^{(S)}$ = Communication length (stopping time on $\{Y_n\}_{n\in\mathbb{N}}$)

Opportunistic Achievability (main idea)

A rate $(R_1, R_2, \ldots, R_L)$ is achievable if $\exists$ a sequence of coding schemes $\{S_n\}_{n\in\mathbb{N}}$ such that for any $\varepsilon > 0$, $\exists n_\circ(\varepsilon)$ so that for all $n > n_\circ(\varepsilon)$

$$P_\ell^{(S_n)} < \varepsilon \quad \text{and} \quad R_\ell^{(S_n)} > R_\ell - \varepsilon \quad \text{for all } \ell = 1, 2, \ldots, L$$

Opportunistic Capacity: Union of all achievable rates.

$$\mathcal{C}_F(\mathcal{Q}) = \{(R_1, R_2, \ldots, R_L) : 0 \le R_\ell \le C_{Q_\ell},\ \ell = 1, 2, \ldots, L\}$$

where C_{Q_ℓ} = capacity of DMC Q_ℓ.

Opportunistic Capacity

- **Notation**

  For a (variable length) coding scheme S

  - $P_\ell^{(S)}$ = Prob of error when $Q_\circ = Q_\ell$
  - $R_\ell^{(S)}$ = Rate when $Q_\circ = Q_\ell$

Opportunistic capacity is a region

$\{S_n\}_{n\in\mathbb{N}}$ such that for any $\varepsilon > 0$, $\exists n_\circ(\varepsilon)$ so that for all $n > n_\circ(\varepsilon)$

$$P_\ell^{(S_n)} < \varepsilon \quad \text{and} \quad R_\ell^{(S_n)} > R_\ell - \varepsilon \quad \text{for all } \ell = 1, 2, \ldots, L$$

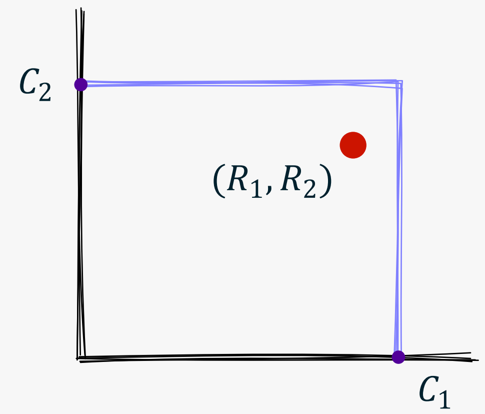
- **Opportunistic Capacity: Union of all achievable rates.**

$$\mathcal{C}_F(\mathcal{Q}) = \{(R_1, R_2, \ldots, R_L) : 0 \le R_\ell \le C_{Q_\ell},\ \ell = 1, 2, \ldots, L\}$$

where C_{Q_ℓ} = capacity of DMC Q_ℓ.

Error Exponent Region

- Suppose a sequence $\{S_n\}_{n\in\mathbb{N}}$ of coding schemes that achieves a rate vector $\boldsymbol{R} = (R_1, R_2, \ldots, R_L)$

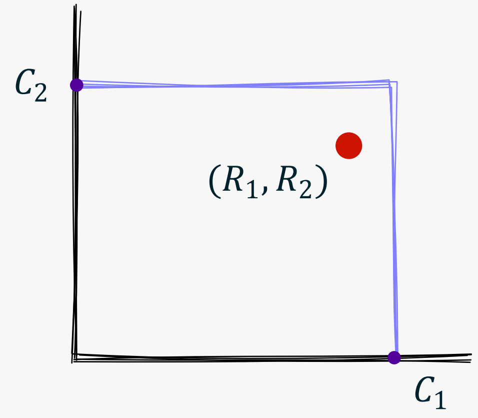


Error Exponent Region

- Suppose a sequence $\{S_n\}_{n\in\mathbb{N}}$ of coding schemes that achieves a rate vector $\boldsymbol{R} = (R_1, R_2, \dots, R_L)$

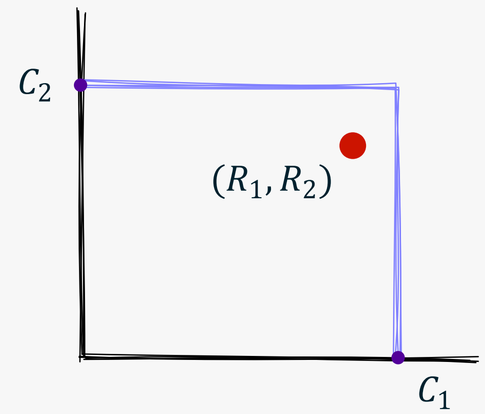
- Error exponent $\boldsymbol{E} = (E_1, E_2, \dots, E_L)$ of this scheme

$$E_\ell = \lim_{n\to\infty} -\frac{\log P_\ell^{(S_n)}}{\mathbb{E}_\ell[\tau^{(S_n)}]}$$



Error Exponent Region

- Suppose a sequence $\{S_n\}_{n\in\mathbb{N}}$ of coding schemes that achieves a rate vector $\boldsymbol{R} = (R_1, R_2, \ldots, R_L)$

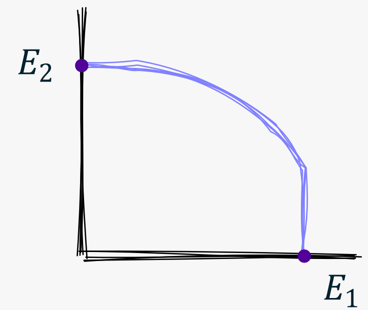


- Error exponent $\boldsymbol{E} = (E_1, E_2, \ldots, E_L)$ of this scheme

$$E_\ell = \lim_{n\to\infty} -\frac{\log P_\ell^{(S_n)}}{\mathbb{E}_\ell[\tau^{(S_n)}]}$$

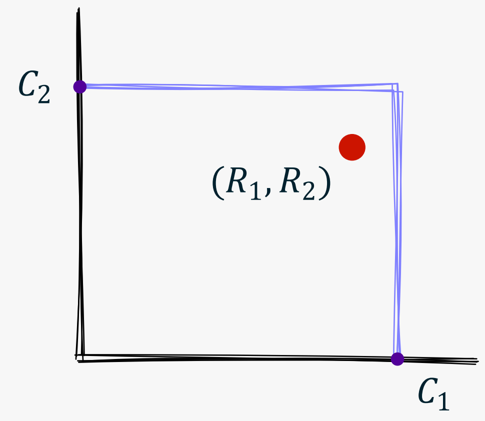
- Error exponent region (EER) $\boldsymbol{\mathcal{E}}(\boldsymbol{R})$

  Union of error exponent over all choices of coding schemes



Error Exponent Region

- Suppose a sequence $\{S_n\}_{n\in\mathbb{N}}$ of coding schemes that achieves a rate vector $\boldsymbol{R} = (R_1, R_2, \ldots, R_L)$

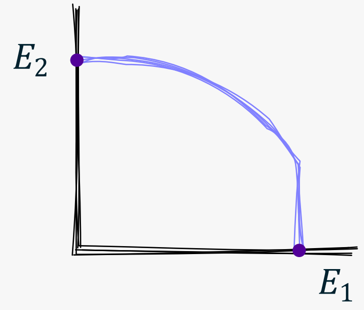


- Error exponent $\boldsymbol{E} = (E_1, E_2, \ldots, E_L)$ of this scheme

$$E_\ell = \lim_{n\to\infty} -\frac{\log P_\ell^{(S_n)}}{\mathbb{E}_\ell[\tau^{(S_n)}]}$$

- Error exponent region (EER) $\boldsymbol{\mathcal{E}}(\boldsymbol{R})$

  Union of error exponent over all choices of coding schemes



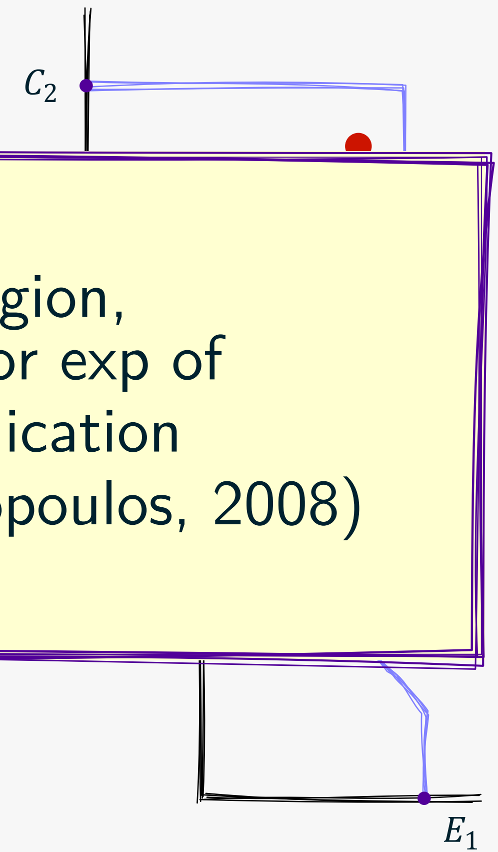
- Reliability ≡ Pareoto frontier of EER

Error Exponent Region

- Suppose a sequence $\{S_n\}_{n\in\mathbb{N}}$ of coding schemes that achieves a rate vector $\boldsymbol{R} = (R_1, R_2, \ldots, R_L)$

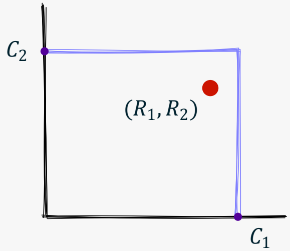
Since capacity is a region,
error exp behave like error exp of
multi-terminal communication
(cf. Weng, Pradhan, Anastasopoulos, 2008)

schemes



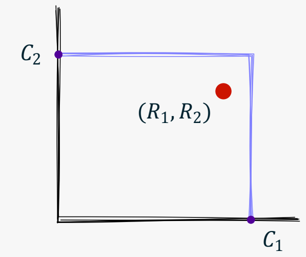
- Reliability ≡ Pareoto frontier of EER

Flavor of the results

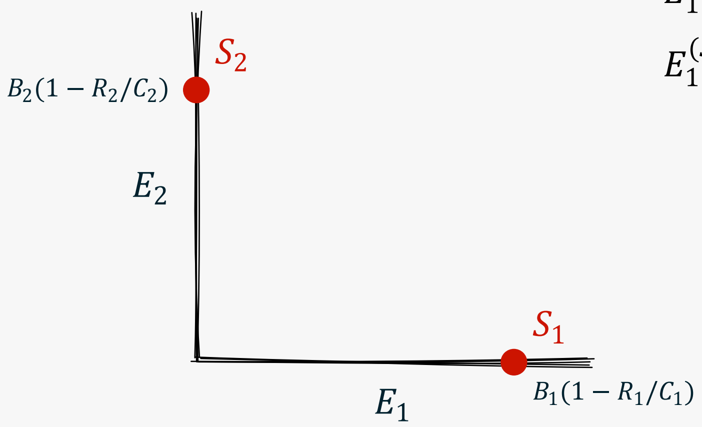


- Suppose $\boldsymbol{Q} = \{Q_1, Q_2\}$ and $\boldsymbol{R} = (R_1, R_2) \leq (C_1, C_2)$.

Flavor of the results



- Suppose $\boldsymbol{Q} = \{Q_1, Q_2\}$ and $\boldsymbol{R} = (R_1, R_2) \leq (C_1, C_2)$.
- Let $S_\ell = \{S_{\ell,n}\}_{n\in\mathbb{N}}$ achieve Burnashev exponent for DMC Q_ℓ.

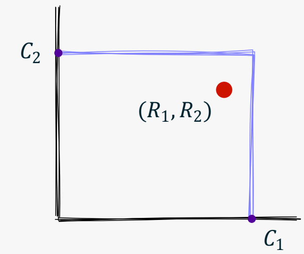


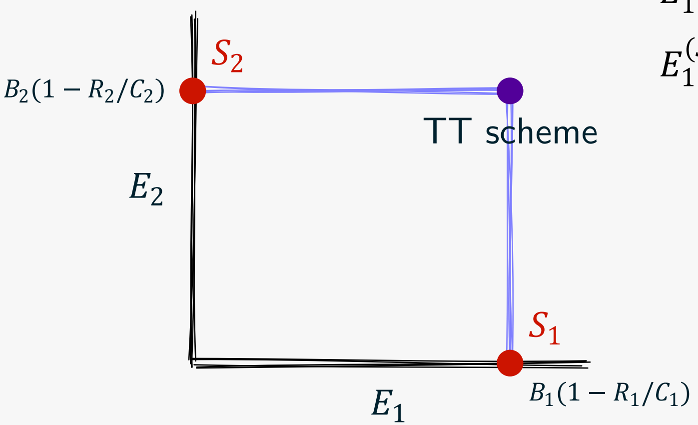
$E_1^{(S_1)} = B_1(1 - R_1/C_1)$ $\quad$ $E_2^{(S_1)} = \text{small}$

$E_1^{(S_2)} = \text{small}$ $\quad$ $E_2^{(S_2)} = B_2(1 - R_2/C_2)$

Flavor of the results

- Suppose $\boldsymbol{Q} = \{Q_1, Q_2\}$ and $\boldsymbol{R} = (R_1, R_2) \leq (C_1, C_2)$.
- Let $S_\ell = \{S_{\ell,n}\}_{n\in\mathbb{N}}$ achieve Burnashev exponent for DMC Q_ℓ.





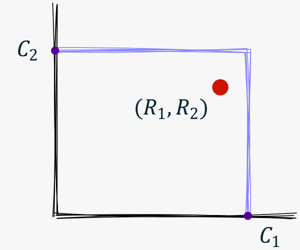
$E_1^{(S_1)} = B_1(1 - R_1/C_1)$ $\qquad E_2^{(S_1)} = \text{small}$

$E_1^{(S_2)} = \text{small}$ $\qquad E_2^{(S_2)} = B_2(1 - R_2/C_2)$

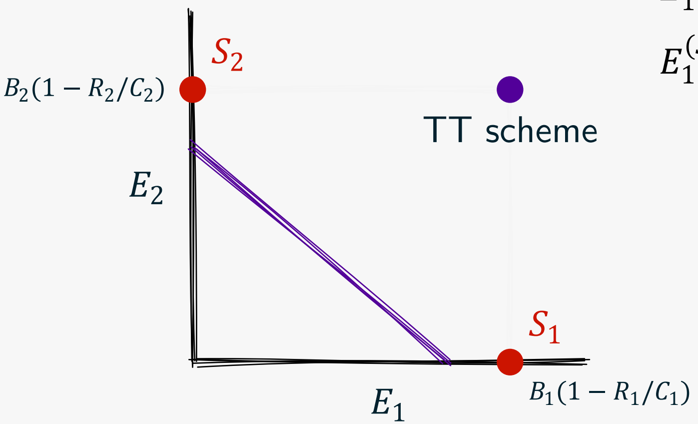
Tchamkerten and Telatar, 2006

- Conditions for universally achieving Burnashev exponent
- Restricted to $R_\ell/C_\ell = \text{constant}$

Flavor of the results



- Suppose $\boldsymbol{Q} = \{Q_1, Q_2\}$ and $\boldsymbol{R} = (R_1, R_2) \leq (C_1, C_2)$.
- Let $S_\ell = \{S_{\ell,n}\}_{n\in\mathbb{N}}$ achieve Burnashev exponent for DMC Q_ℓ.



$E_1^{(S_1)} = B_1(1 - R_1/C_1)$ $\quad E_2^{(S_1)} =$ small

$E_1^{(S_2)} =$ small $\quad E_2^{(S_2)} = B_2(1 - R_2/C_2)$

Tchamkerten and Telatar, 2006

- Conditions for universally achieving Burnashev exponent
- Restricted to $R_\ell/C_\ell =$ constant

- Our result: Propose a simple training based scheme such that

$$E_\ell \geq \lambda_{\boldsymbol{Q}} B_\ell \left(1 - \frac{R_\ell}{C_\ell}\right), \qquad \ell = 1, 2$$

Outline

- Literature overview
  - Burnashev exponent, Yamamoto-Itoh scheme
  - Tchamkerten-Telatar result
- Variable length coding scheme

  Transmit a compound message
- An achievable scheme

  A training based variation of Yamamoto-Itoh scheme
- Example

  $\{\mathrm{BSC}_p, \mathrm{BSC}_{1-p}\}$. Non-obvious channel estimation

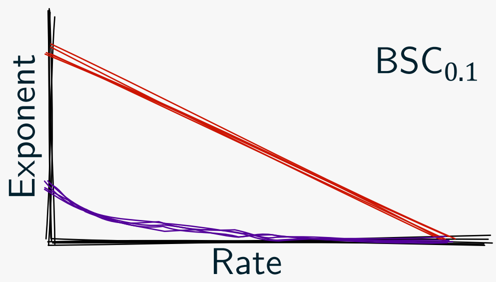
Literature Overview

DMC with Feedback

- Variable length coding significantly boosts the error exponents [Burnashev, 1976]

$$E = B_Q \left(1 - \frac{R}{C}\right)$$

where $B_Q = \max_{x_A, x_R \in \mathcal{X}} D(Q(\cdot|x_A) \| Q(\cdot|x_R))$

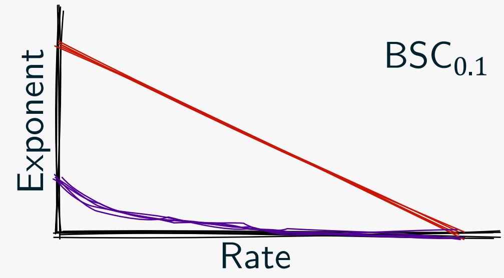


DMC with Feedback

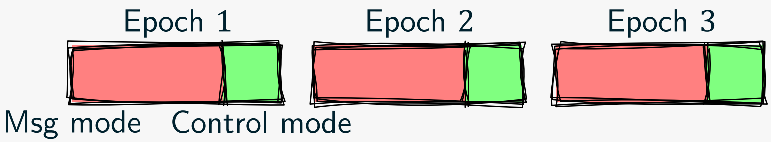
- Variable length coding significantly boosts the error exponents [Burnashev, 1976]

$$E = B_Q\left(1 - \frac{R}{C}\right)$$

where $B_Q = \max_{x_A, x_R \in \mathcal{X}} D(Q(\cdot|x_A) \| Q(\cdot|x_R))$



- A simple, two phase, coding scheme achieves Burnashev exponent [Yamamoto Itoh, 1979]

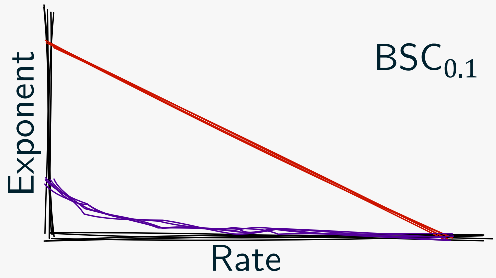


DMC with Feedback

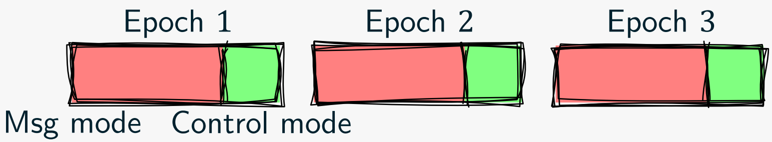
- Variable length coding significantly boosts the error exponents [Burnashev, 1976]

$$E = B_Q \left(1 - \frac{R}{C}\right)$$

where $B_Q = \max_{x_A, x_R \in \mathcal{X}} D(Q(\cdot|x_A) \| Q(\cdot|x_R))$



- A simple, two phase, coding scheme achieves Burnashev exponent [Yamamoto Itoh, 1979]



- Any scheme that achieves Burnashev exponent must have a control phase [Berlin *et. al*, 2009]

Compound Channel with Feedback

Compound Channel with Feedback

- Training based schemes have poor error exponent

[Tchamkerten Telatar, 2006a]

The slope of the error exponent goes to zero at rates near capacity

Compound Channel with Feedback

- Training based schemes have poor error exponent

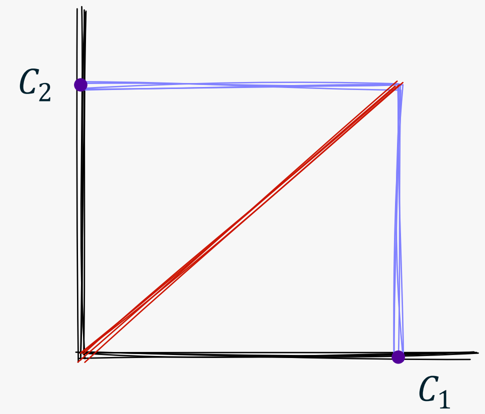
  [Tchamkerten Telatar, 2006a]

  The slope of the error exponent goes to zero at rates near capacity
- Burnashev exponent can be achieved universally

  [Tchamkerten Telatar, 2006b]

  The compound family is not too diverse

$$\Delta(\mathcal{Q}) = 0$$



Compound Channel with Feedback

- Training based schemes have poor error exponent

  [Tchamkerten Telatar, 2006a]

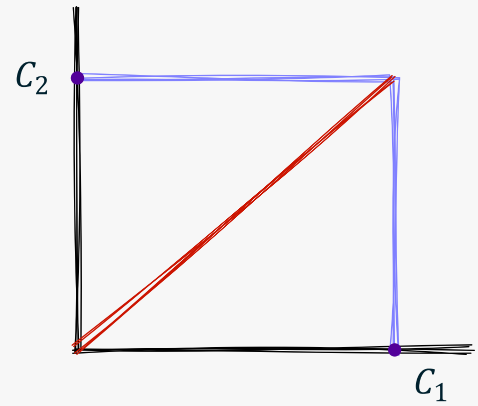
  The slope of the error exponent goes to zero at rates near capacity

- Burnashev exponent can be achieved universally

  [Tchamkerten Telatar, 2006b]

  The compound family is not too diverse

  $$\Delta(\boldsymbol{\mathcal{Q}}) = 0$$



- Positive example: $\boldsymbol{\mathcal{B}} = \{\mathrm{BSC}_p : p \in [0, 1/2)\}$ and $\boldsymbol{\mathcal{Z}} = \{Z_p : p \in [0, 1]\}$

Compound Channel with Feedback

- Training based schemes have poor error exponent

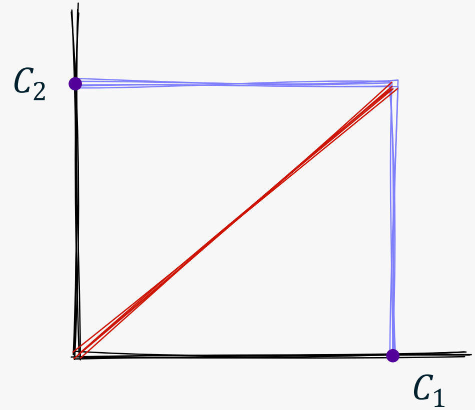
  [Tchamkerten Telatar, 2006a]

  The slope of the error exponent goes to zero at rates near capacity
- Burnashev exponent can be achieved universally

  [Tchamkerten Telatar, 2006b]

  The compound family is not too diverse

  $$\Delta(\boldsymbol{Q}) = 0$$



- Positive example: $\boldsymbol{\mathcal{B}} = \{\mathrm{BSC}_p : p \in [0, 1/2)\}$ and $\boldsymbol{\mathcal{Z}} = \{Z_p : p \in [0, 1]\}$
- Negative example: $Q_p = \{\mathrm{BSC}_p, \mathrm{BSC}_{1-p}\}$, p known.

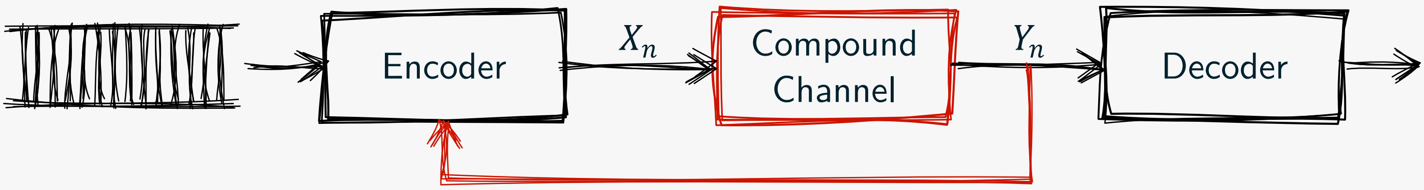
Variable length coding over compound channel with feedback

Compound Channel with Feedback

- Adapting transmission rate
  - Vary communication length
  - Vary the message size

Compound Channel with Feedback

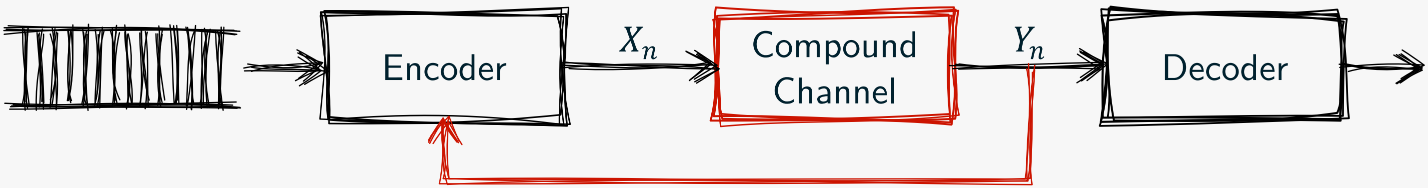
- Adapting transmission rate
  - Vary communication length
  - Vary the message size



- Transmit $\log_2 M_\ell$ bits reliably when channel $Q_\circ = Q_\ell$, $\ell = 1, 2, \ldots, L$

Compound Channel with Feedback

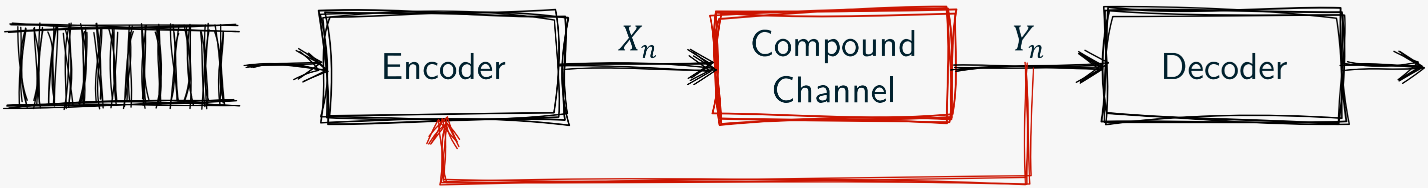
- Adapting transmission rate
  - Vary communication length
  - Vary the message size



- Transmit $\log_2 M_\ell$ bits reliably when channel $Q_\circ = Q_\ell$, $\ell = 1, 2, \ldots, L$
- Compound message $\boldsymbol{W} = (W_1, W_2, \ldots, W_L)$ of size $\boldsymbol{M} = (M_1, M_2, \ldots, M_L)$

Compound Channel with Feedback

- Adapting transmission rate
  - Vary communication length
  - Vary the message size

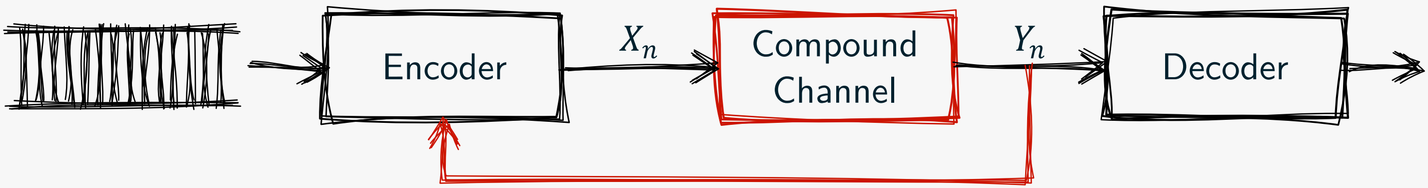


- Transmit $\log_2 M_\ell$ bits reliably when channel $Q_\circ = Q_\ell$, $\ell = 1, 2, \ldots, L$
- Compound message $\boldsymbol{W} = (W_1, W_2, \ldots, W_L)$ of size $\boldsymbol{M} = (M_1, M_2, \ldots, M_L)$
- Encoding scheme

$$X_1 = f_1(\boldsymbol{W}), \quad X_2 = f_2(\boldsymbol{W}, Y_1), \quad X_3 = f_3(\boldsymbol{W}, Y_1, Y_2), \quad \cdots$$

Compound Channel with Feedback

- Adapting transmission rate
  - Vary communication length
  - Vary the message size



- Transmit $\log_2 M_\ell$ bits reliably when channel $Q_\circ = Q_\ell$, $\ell = 1, 2, \ldots, L$
- Compound message $\boldsymbol{W} = (W_1, W_2, \ldots, W_L)$ of size $\boldsymbol{M} = (M_1, M_2, \ldots, M_L)$
- Encoding scheme

$$X_1 = f_1(\boldsymbol{W}), \quad X_2 = f_2(\boldsymbol{W}, Y_1), \quad X_3 = f_3(\boldsymbol{W}, Y_1, Y_2), \quad \cdots$$

- Decoding scheme: at stopping time τ, choose

$$(\hat{L}, \hat{W}) = g_\tau(Y_1, Y_2, \ldots, Y_\tau)$$

The coding scheme

Coding scheme is a tuple $(\boldsymbol{M}, \boldsymbol{f}, \boldsymbol{g}, \tau)$

- Compound message size
  $\boldsymbol{M} = (M_1, M_2, \dots, M_L)$

- Compound message alphabet
  $\boldsymbol{\mathcal{M}} = \prod_{\ell=1}^{L} \{1, 2, \dots, M_\ell\}$

- Encoding strategy $\boldsymbol{f} = (f_1, f_2, \dots)$

  $f_t : \boldsymbol{\mathcal{M}} \times \boldsymbol{\mathcal{Y}}^{t-1} \mapsto \boldsymbol{\mathcal{X}}$

- Decoding strategy $\boldsymbol{g} = (g_1, g_2, \dots)$

  $$g_t : \boldsymbol{\mathcal{Y}}^{t-1} \mapsto \bigcup_{\ell=1}^{L} \{(\ell, 1), (\ell, 2), \dots, (\ell, M_\ell)\}$$

- Stopping time τ with respect to filtration $\{2^{\boldsymbol{\mathcal{Y}}^t}, t \in \mathbb{N}\}$

The coding scheme

Coding scheme is a tuple $(\boldsymbol{M}, \boldsymbol{f}, \boldsymbol{g}, \tau)$

- Compound message size $\boldsymbol{M} = (M_1, M_2, \ldots, M_L)$
- Compound message alphabet
  $$\mathcal{M} = \prod_{\ell=1}^{L} \{1, 2, \ldots, M_\ell\}$$
- Encoding strategy $\boldsymbol{f} = (f_1, f_2, \ldots)$
  $$f_t : \mathcal{M} \times \mathcal{Y}^{t-1} \mapsto \mathcal{X}$$
- Decoding strategy $\boldsymbol{g} = (g_1, g_2, \ldots)$
  $$g_t : \mathcal{Y}^{t-1} \mapsto \bigcup_{\ell=1}^{L} \{(\ell, 1), (\ell, 2), \ldots, (\ell, M_\ell)\}$$
- Stopping time τ with respect to filtration $\{2^{\mathcal{Y}^t}, t \in \mathbb{N}\}$

- Rate $\boldsymbol{R} = (R_1, R_2, \ldots, R_L)$
  $$R_\ell = \frac{\mathbb{E}_\ell[\log_2 M_{\hat{L}}]}{\mathbb{E}_\ell[\tau]}$$
- Probability of error $\boldsymbol{P} = (P_1, P_2, \ldots, P_L)$
  $$P_\ell = \mathbb{P}_\ell(\hat{W} \neq W_{\hat{L}})$$

Capacity and Error Exponents

- Achievable Rate: A rate $\boldsymbol{R} = (R_1, R_2, \ldots, R_L)$ is achievable if there exists a sequence of coding schemes $(\boldsymbol{M}^{(n)}, \boldsymbol{f}^{(n)}, \boldsymbol{g}^{(n)}, \tau^{(n)})$, $n \in \mathbb{N}$, such that

  1. $\lim_{n\to\infty} \mathbb{E}_\ell[\tau^{(n)}] = \infty$, $\ell = 1, 2, \ldots, L$
  2. For any $\varepsilon > 0$, $\exists n_\circ(\varepsilon)$ s.t. $\forall n \geq n_\circ(\varepsilon)$

$$P_\ell^{(n)} < \varepsilon \quad \text{and} \quad R_\ell^{(n)} \geq R_\ell - \varepsilon, \qquad \ell = 1, 2, \ldots, L$$

- Capacity: Union of all achievable rates

Capacity and Error Exponents

- Achievable Rate: A rate $\boldsymbol{R} = (R_1, R_2, \ldots, R_L)$ is achievable if there exists a sequence of coding schemes $(\boldsymbol{M}^{(n)}, \boldsymbol{f}^{(n)}, \boldsymbol{g}^{(n)}, \tau^{(n)})$, $n \in \mathbb{N}$, such that

  1. $\lim_{n\to\infty} \mathbb{E}_\ell[\tau^{(n)}] = \infty$, $\ell = 1, 2, \ldots, L$
  2. For any $\varepsilon > 0$, $\exists n_\circ(\varepsilon)$ s.t. $\forall n \geq n_\circ(\varepsilon)$

$$P_\ell^{(n)} < \varepsilon \quad \text{and} \quad R_\ell^{(n)} \geq R_\ell - \varepsilon, \qquad \ell = 1, 2, \ldots, L$$

- Capacity: Union of all achievable rates

- Error Exponent: Given a sequence of coding schemes $(\boldsymbol{M}^{(n)}, \boldsymbol{f}^{(n)}, \boldsymbol{g}^{(n)}, \tau^{(n)})$, that achieve a rate $\boldsymbol{R}$, the error exponent is given by

$$E_\ell = \lim_{n\to\infty} -\frac{\log P_\ell^{(n)}}{\mathbb{E}_\ell[\tau^{(n)}]}$$

- Error Exponent Region: Union of all error exponents.

- Reliability: Pareoto frontier of Error Exponent region

Main Result

- Opportunistic Capacity

$$\mathcal{C}_F(\mathcal{Q}) = \{(R_1, \ldots, R_L) : 0 \leq R_\ell < C_\ell, \ell = 1, \ldots, L\}$$

where C_ℓ is the capacity of DMC Q_ℓ.

Main Result

- Opportunistic Capacity

$$\mathcal{C}_F(\mathcal{Q}) = \{(R_1, \ldots, R_L) : 0 \le R_\ell < C_\ell, \ell = 1, \ldots, L\}$$

where C_ℓ is the capacity of DMC Q_ℓ.

- Training based inner bound on Error Exponent Region

  - Let T_ℓ^c be the exponent of the channel estimation error when the channel is Q_ℓ. For any channel estimation scheme, $(T_1^c, \ldots, T_L^c) \in \mathcal{T}^\star$.

Main Result

- Opportunistic Capacity

$$\mathcal{C}_F(\mathcal{Q}) = \{(R_1, \ldots, R_L) : 0 \le R_\ell < C_\ell, \ell = 1, \ldots, L\}$$

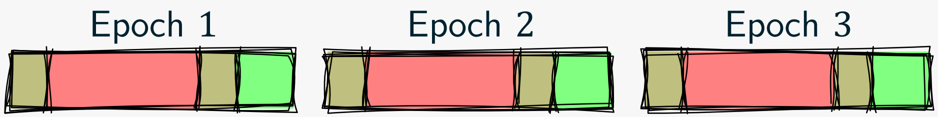
where C_ℓ is the capacity of DMC Q_ℓ.

- Training based inner bound on Error Exponent Region
  - Let T^c_ℓ be the exponent of the channel estimation error when the channel is Q_ℓ. For any channel estimation scheme, $(T^c_1, \ldots, T^c_L) \in \mathcal{T}^\star$.
  - At rate $\boldsymbol{R} = (R_1, \ldots, R_L)$, the error exponent is

$$E_\ell \ge \frac{T^c_\ell}{T^c_\ell + B_{Q_\ell}} B_{Q_\ell} \left(1 - \frac{R_\ell}{C_\ell}\right)$$

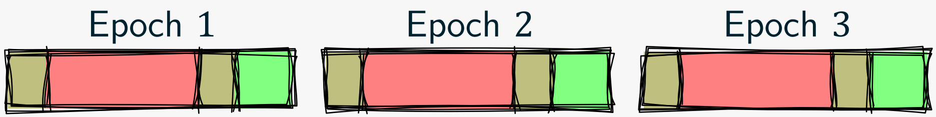
where $B_{Q_\ell} = \max_{x_A, x_R \in \mathcal{X}} D(Q_\ell(\cdot|x_A) \| Q_\ell(\cdot|x_R))$

Achievable Scheme for $(R_1, \ldots, R_L)$



Parameterized by $n \in \mathbb{N}$. Multiple epochs. Each epoch has four phases

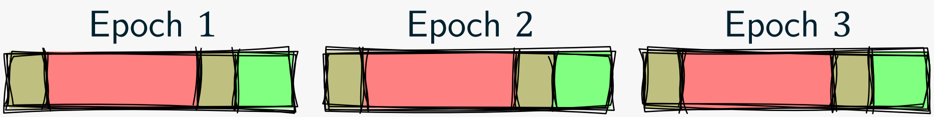
Achievable Scheme for $(R_1, \ldots, R_L)$



Parameterized by $n \in \mathbb{N}$. Multiple epochs. Each epoch has four phases

1. Training phase: length $\beta_1^{(n)} n$.
   Channel estimate $\hat{L}_m^{(k,n)}$.
2. Message phase: length $\beta_2^{(n)}(\hat{L}_m) n$
3. Retraining phase: length $\beta_3^{(n)}$.
   Channel estimate $\hat{L}_c^{(k,n)}$
4. Control phase: length $\beta_4^{(n)}(\hat{L}_c) n$.

Achievable Scheme for $(R_1, \ldots, R_L)$



Parameterized by $n \in \mathbb{N}$. Multiple epochs. Each epoch has four phases

1. Training phase: length $\beta_1^{(n)} n$. Channel estimate $\hat{L}_m^{(k,n)}$.
2. Message phase: length $\beta_2^{(n)}(\hat{L}_m) n$
3. Retraining phase: length $\beta_3^{(n)}$. Channel estimate $\hat{L}_c^{(k,n)}$
4. Control phase: length $\beta_4^{(n)}(\hat{L}_c) n$.

- Estimation rules: $\hat{\theta}_m^{(n)}$ for phase one, $\hat{\theta}_c^{(n)}$ for phase three.
- Training exponents: $(T_1^m, \ldots, T_L^m)$ and $(T_1^c, \ldots, T_L^c)$ respectively

$$T_\ell = \mathbb{P}_\ell(\text{Channel estimation is wrong})$$

Choice of parameters

Let $\kappa_\ell = \frac{T_\ell^c}{B_{Q_\ell}}$ and $\gamma_\ell = \frac{R_\ell}{C_\ell}$

Choice of parameters

Let $\kappa_\ell = \frac{T_\ell^c}{B_{Q_\ell}}$ and $\gamma_\ell = \frac{R_\ell}{C_\ell}$

- Choose a reference channel $Q^\star$

$$\alpha_\ell = \frac{(1+\kappa_\ell)}{(1-\gamma_\ell)} \frac{(1-\gamma^\star)}{(1+\kappa^\star)}$$

Choice of parameters

Let $\kappa_\ell = \frac{T_\ell^c}{B_{Q_\ell}}$ and $\gamma_\ell = \frac{R_\ell}{C_\ell}$

- Choose a reference channel $Q^\star$

$$\alpha_\ell = \frac{(1+\kappa_\ell)}{(1-\gamma_\ell)}\frac{(1-\gamma^\star)}{(1+\kappa^\star)}$$

- Choice of parameters

$$\beta_1^{(n)} \downarrow 0, \qquad \beta_2^{(n)}(\ell) \downarrow \alpha_\ell \gamma_\ell,$$

$$\beta_3^{(n)} \uparrow \frac{(1-\gamma^\star)}{(1+\kappa^\star)}, \qquad \beta_4^{(n)}(\ell) \uparrow \kappa_\ell \beta_3^{(n)},$$

Choice of parameters

Let $\kappa_\ell = \frac{T_\ell^c}{B_{Q_\ell}}$ and $\gamma_\ell = \frac{R_\ell}{C_\ell}$

- Choose a reference channel $Q^\star$

$$\alpha_\ell = \frac{(1+\kappa_\ell)}{(1-\gamma_\ell)} \frac{(1-\gamma^\star)}{(1+\kappa^\star)}$$

- Choice of parameters

$$\beta_1^{(n)} \downarrow 0, \qquad \beta_2^{(n)}(\ell) \downarrow \alpha_\ell \gamma_\ell,$$

$$\beta_3^{(n)} \uparrow \underbrace{\frac{(1-\gamma^\star)}{(1+\kappa^\star)}}_{=\alpha_\ell \frac{(1-\gamma_\ell)}{(1+\gamma_\ell)}}, \qquad \beta_4^{(n)}(\ell) \uparrow \kappa_\ell \beta_3^{(n)},$$

Choice of parameters

Let $\kappa_\ell = \frac{T_\ell^c}{B_{Q_\ell}}$ and $\gamma_\ell = \frac{R_\ell}{C_\ell}$

- Choose a reference channel $Q^\star$

$$\alpha_\ell = \frac{(1+\kappa_\ell)}{(1-\gamma_\ell)}\frac{(1-\gamma^\star)}{(1+\kappa^\star)}$$

- Choice of parameters

$$\beta_1^{(n)} \downarrow 0, \qquad \beta_2^{(n)}(\ell) \downarrow \alpha_\ell\gamma_\ell, \qquad \mathbb{E}_\ell[\beta_1^{(n)} + \beta_2^{(n)}(\ell)] \downarrow \alpha_\ell\gamma_\ell$$

$$\beta_3^{(n)} \uparrow \underbrace{\frac{(1-\gamma^\star)}{(1+\kappa^\star)}}_{=\alpha_\ell\frac{(1-\gamma_\ell)}{(1+\gamma_\ell)}}, \qquad \beta_4^{(n)}(\ell) \uparrow \kappa_\ell\beta_3^{(n)}, \qquad \mathbb{E}_\ell[\beta_3^{(n)} + \beta_3^{(n)}(\ell)] \uparrow \alpha_\ell(1-\gamma_\ell)$$

Choice of parameters

Let $\kappa_\ell = \frac{T_\ell^c}{B_{Q_\ell}}$ and $\gamma_\ell = \frac{R_\ell}{C_\ell}$

- Choose a reference channel $Q^\star$

$$\alpha_\ell = \frac{(1+\kappa_\ell)}{(1-\gamma_\ell)} \frac{(1-\gamma^\star)}{(1+\kappa^\star)}$$

- Choice of parameters

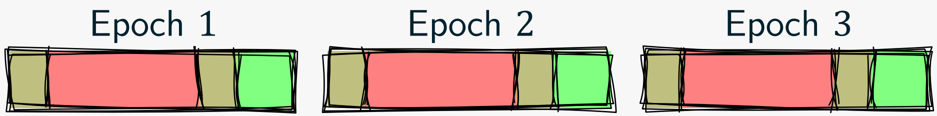
$$\beta_1^{(n)} \downarrow 0, \qquad \beta_2^{(n)}(\ell) \downarrow \alpha_\ell \gamma_\ell, \qquad \mathbb{E}_\ell[\beta_1^{(n)} + \beta_2^{(n)}(\ell)] \downarrow \alpha_\ell \gamma_\ell$$

$$\beta_3^{(n)} \uparrow \underbrace{\frac{(1-\gamma^\star)}{(1+\kappa^\star)}}_{=\alpha_\ell \frac{(1-\gamma_\ell)}{(1+\gamma_\ell)}}, \qquad \beta_4^{(n)}(\ell) \uparrow \kappa_\ell \beta_3^{(n)}, \qquad \mathbb{E}_\ell[\beta_3^{(n)} + \beta_3^{(n)}(\ell)] \uparrow \alpha_\ell (1-\gamma_\ell)$$

- Expected length of an epoch

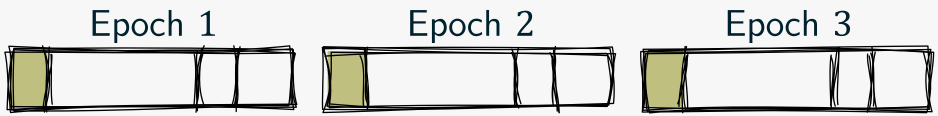
$$\mathbb{E}_\ell[\beta_1^{(n)} + \beta_2^{(n)}(\ell) + \beta_3^{(n)} + \beta_4^{(n)}(\ell)]n = \alpha_\ell n$$

Achievable Scheme for $(R_1, \ldots, R_L)$



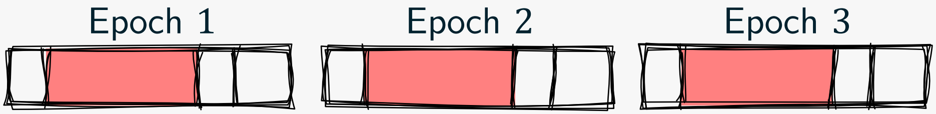
- Training phase
- Message phase
- Retraining phase
- Control phase

Achievable Scheme for $(R_1, \ldots, R_L)$



- Training phase
  - Send training seq of length $\beta_1^{(n)} n$.
  - Channel estimate $\hat{L}_m^{(k,n)}$ using estimation rule $\hat{\theta}_m^{(n)}$
- Message phase
- Retraining phase
- Control phase

Achievable Scheme for $(R_1, \ldots, R_L)$

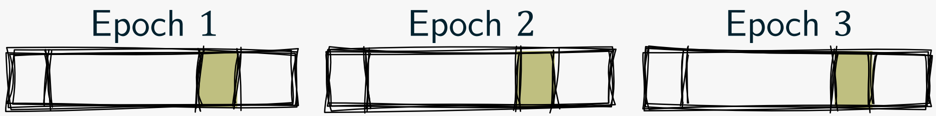


- Training phase
- Message phase

  Encoder and decoder agree on L codebooks

  - Codebook ℓ: No feedback comm over DMC Q_ℓ
    block length $\beta_2^{(n)}(\ell)$; size $M_\ell(n) = \lfloor 2^{n\alpha_\ell \gamma_\ell C_\ell} \rfloor$ (rate $\approx C_\ell$)
  - Use codebook $\hat{L}_m^{(k,n)}$
- Retraining phase
- Control phase

Achievable Scheme for $(R_1, \ldots, R_L)$

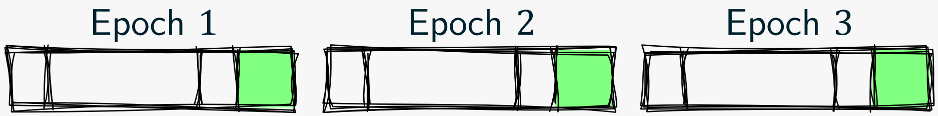


- Training phase
- Message phase
- Retraining phase
  - Send training seq of length $\beta_3^{(n)} n$.
  - Channel estimate $\hat{L}_c^{(k,n)}$ using estimation rule $\hat{\theta}_c^{(n)}$

> $\hat{L}_c^{(k,n)}$ only depends on training sequence of phase 3 of epoch k

- Control phase

Achievable Scheme for $(R_1, \ldots, R_L)$



- Training phase
- Message phase
- Retraining phase
- Control phase

Encoder and decoder agree upon

- L pairs of input symbols $(x_A(\ell), x_R(\ell))$ that are arg max of

$$\max_{x_A, x_R} D(Q_\ell(\cdot|x_A) \| Q_\ell(\cdot|x_R))$$

- L decoding regions $\boldsymbol{\mathcal{A}}_\ell \subset \boldsymbol{\mathcal{Y}}^{\beta_4^{(n)}(\ell)n}$ that optimally distinguish

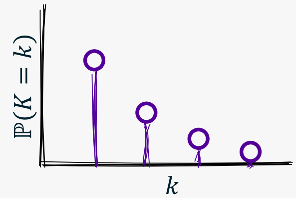
$$\underbrace{(x_A(\ell), \ldots, x_A(\ell))}_{\beta_4^{(n)}(\ell)\text{times}} \quad \text{from} \quad \underbrace{(x_R(\ell), \ldots, x_R(\ell))}_{\beta_4^{(n)}(\ell)\text{times}} \quad \text{over DMC } Q_\ell$$

- Send ACCEPT or REJECT symbols for channel $\hat{L}_c^{(k,n)}$

Performance Analysis

- Number of epochs $K(n)$

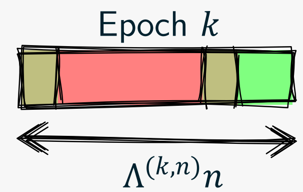
$$\mathbb{P}_\ell(K^{(n)} = k) = p_\ell(n)(1 - p_\ell(n))^n, \quad \lim_{n\to\infty} p_\ell(n) = 1$$



- Relative length of an epoch

$$\Lambda^{(k,n)} = \beta_1^{(n)} + \beta_2^{(n)}(\hat{L}_m^{k,n)}) + \beta_3^{(n)} + \beta_4^{(n)}(\hat{L}_c^{(k,n)})$$

$$\mathbb{E}_\ell[\Lambda^{(k,n)}] \to \alpha_\ell$$



- Expected communication length

$$\mathbb{E}_\ell[\tau^{(n)}] = \mathbb{E}_\ell[K^{(n)}\Lambda^{(K^{(n)},n)}n] \approx \alpha_\ell n$$

Performance Analysis

- Rate

$$\lim_{n\to\infty} \frac{\mathbb{E}_\ell[\log M_{\hat{L}_m^{(k,n)}}(n)]}{\mathbb{E}_\ell[\tau^{(n)}]}$$

$$M_\ell(n) = \lfloor 2^{n\alpha_\ell \gamma_\ell C_\ell} \rfloor$$

Performance Analysis

- Rate

$$\lim_{n\to\infty} \frac{\mathbb{E}_\ell[\log M_{\hat{L}_m^{(k,n)}}(n)]}{\mathbb{E}_\ell[\tau^{(n)}]} \approx \frac{n\alpha_\ell\gamma_\ell C_\ell}{\alpha_\ell n} = \gamma_\ell C_\ell$$

$$M_\ell(n) = \lfloor 2^{n\alpha_\ell\gamma_\ell C_\ell} \rfloor$$

Performance Analysis

- Rate

$$\lim_{n\to\infty} \frac{\mathbb{E}_\ell[\log M_{\hat{L}_m^{(k,n)}}(n)]}{\mathbb{E}_\ell[\tau^{(n)}]} \approx \frac{n\alpha_\ell\gamma_\ell C_\ell}{\alpha_\ell n} = \gamma_\ell C_\ell = R_\ell$$

$$M_\ell(n) = \lfloor 2^{n\alpha_\ell\gamma_\ell C_\ell} \rfloor$$

Performance Analysis

- Rate

$$\lim_{n\to\infty} \frac{\mathbb{E}_\ell[\log M_{\hat{L}_m^{(k,n)}}(n)]}{\mathbb{E}_\ell[\tau^{(n)}]} \approx \frac{n\alpha_\ell\gamma_\ell C_\ell}{\alpha_\ell n} = \gamma_\ell C_\ell = R_\ell$$

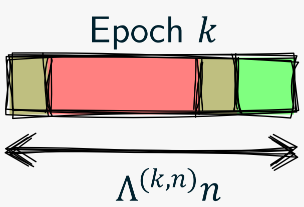
$$M_\ell(n) = \lfloor 2^{n\alpha_\ell\gamma_\ell C_\ell} \rfloor$$

- Error Exponent

$$E_\ell = \lim_{n\to\infty} -\frac{\log P_\ell^{(n)}}{\mathbb{E}_\ell[\tau^{(n)}]}$$

$$\approx \lim_{n\to\infty} -\frac{\text{log(decoding error)} + \text{log(hypothesis testing error)}}{\alpha_\ell n}$$

$$\geq \lim_{n\to\infty} -\frac{\text{log(hypothesis testing error)}}{\alpha_\ell n}$$



Performance Analysis

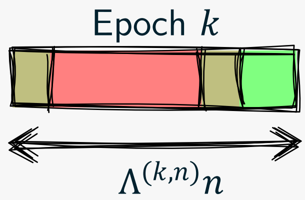
- Rate

$$\lim_{n\to\infty} \frac{\mathbb{E}_\ell[\log M_{\hat{L}_m^{(k,n)}}(n)]}{\mathbb{E}_\ell[\tau^{(n)}]} \approx \frac{n\alpha_\ell\gamma_\ell C_\ell}{\alpha_\ell n} = \gamma_\ell C_\ell = R_\ell$$

$$M_\ell(n) = \lfloor 2^{n\alpha_\ell\gamma_\ell C_\ell} \rfloor$$

- Error Exponent

$$\begin{aligned} E_\ell &= \lim_{n\to\infty} -\frac{\log P_\ell^{(n)}}{\mathbb{E}_\ell[\tau^{(n)}]} \\ &\approx \lim_{n\to\infty} -\frac{\text{log(decoding error)} + \text{log(hypothesis testing error)}}{\alpha_\ell n} \\ &\geq \lim_{n\to\infty} -\frac{\text{log(hypothesis testing error)}}{\alpha_\ell n} \\ &\approx -\log\left(e^{-\beta_3^{(n)} n(\square)} + e^{-\beta_4^{(n)}(\ell) n(\square)}\right) \Big/ (\alpha_\ell n) \end{aligned}$$



Performance Analysis

- Rate

$$M_\ell(n) = \lfloor 2^{n\alpha_\ell \gamma_\ell C_\ell} \rfloor$$

$$\lim_{n\to\infty} \frac{\mathbb{E}_\ell[\log M_{\hat{L}_m^{(k,n)}}(n)]}{\mathbb{E}_\ell[\tau^{(n)}]} \approx \frac{n\alpha_\ell\gamma_\ell C_\ell}{\alpha_\ell n} = \gamma_\ell C_\ell = R_\ell$$

- Error Exponent

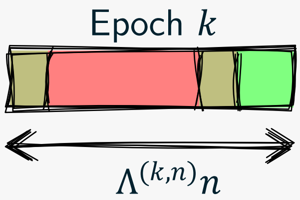
$$E_\ell = \lim_{n\to\infty} -\frac{\log P_\ell^{(n)}}{\mathbb{E}_\ell[\tau^{(n)}]}$$

$$\approx \lim_{n\to\infty} -\frac{\log(\text{decoding error}) + \log(\text{ hypothesis testing error})}{\alpha_\ell n}$$

$$\geq \lim_{n\to\infty} -\frac{\log(\text{ hypothesis testing error})}{\alpha_\ell n}$$

$$\approx -\log\left(e^{-\beta_3^{(n)} n(\square)} + e^{-\beta_4^{(n)}(\ell) n(\square)}\right)\Big/(\alpha_\ell n)$$

$$\geq \frac{(\square)\cdot(\square)}{(\square)+(\square)} \frac{\beta_3^{(n)} + \beta_4^{(n)}(\ell)}{\alpha_\ell}$$



Performance Analysis

- Rate

$$M_\ell(n) = \lfloor 2^{n\alpha_\ell\gamma_\ell C_\ell} \rfloor$$

$$\lim_{n\to\infty} \frac{\mathbb{E}_\ell[\log M_{\hat{L}_m^{(k,n)}}(n)]}{\mathbb{E}_\ell[\tau^{(n)}]} \approx \frac{n\alpha_\ell\gamma_\ell C_\ell}{\alpha_\ell n} = \gamma_\ell C_\ell = {\color{red}R_\ell}$$

- Error Exponent

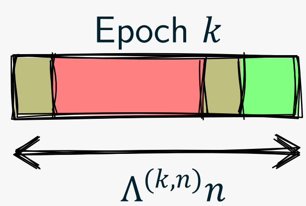
$$E_\ell = \lim_{n\to\infty} -\frac{\log P_\ell^{(n)}}{\mathbb{E}_\ell[\tau^{(n)}]}$$

$$\approx \lim_{n\to\infty} -\frac{\text{log(decoding error)} + \text{log(hypothesis testing error)}}{\alpha_\ell n}$$

$$\geq \lim_{n\to\infty} -\frac{\text{log(hypothesis testing error)}}{\alpha_\ell n}$$

$$\approx -\log\left(e^{-\beta_3^{(n)} n \left({\color{olive}\blacksquare}\right)} + e^{-\beta_4^{(n)}(\ell) n \left({\color{green}\blacksquare}\right)}\right) \Big/ (\alpha_\ell n)$$

$$\geq \frac{\left({\color{olive}\blacksquare}\right)\cdot\left({\color{green}\blacksquare}\right)}{\left({\color{olive}\blacksquare}\right)+\left({\color{green}\blacksquare}\right)} \frac{\beta_3^{(n)} + \beta_4^{(n)}(\ell)}{\alpha_\ell} = \frac{{\color{olive}T_\ell^c} \cdot {\color{green}B_{Q_\ell}}}{{\color{olive}T_\ell^c} + {\color{green}B_{Q_\ell}}} {\color{red}\left(1 - \frac{R_\ell}{C_\ell}\right)}$$



Performance Analysis

- Rate

$$\lim_{n\to\infty} \frac{\mathbb{E}_\ell[\log M_{\hat{L}_m^{(k,n)}}(n)]}{\mathbb{E}_\ell[\tau^{(n)}]} \approx \frac{n\alpha_\ell\gamma_\ell C_\ell}{\alpha_\ell n} = \gamma_\ell C_\ell = R_\ell$$

$$M_\ell(n) = \lfloor 2^{n\alpha_\ell\gamma_\ell C_\ell} \rfloor$$

- Error Exponent

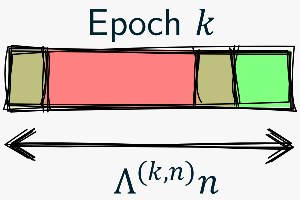
$$E_\ell = \lim_{n\to\infty} -\frac{\log P_\ell^{(n)}}{\mathbb{E}_\ell[\tau^{(n)}]}$$

$$\approx \lim_{n\to\infty} -\frac{\text{log(decoding error)} + \text{log(hypothesis testing error)}}{\alpha_\ell n}$$

$$\geq \lim_{n\to\infty} -\frac{\text{log(hypothesis testing error)}}{\alpha_\ell n}$$

$$\approx -\log\left(e^{-\beta_3^{(n)} n(\blacksquare)} + e^{-\beta_4^{(n)}(\ell) n(\blacksquare)}\right) \Big/ (\alpha_\ell n)$$

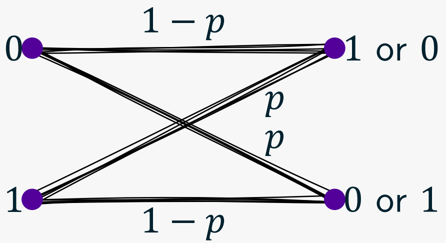
$$\geq \frac{(\blacksquare)\cdot(\blacksquare)}{(\blacksquare)+(\blacksquare)} \frac{\beta_3^{(n)} + \beta_4^{(n)}(\ell)}{\alpha_\ell} = \frac{T_\ell^c \cdot B_{Q_\ell}}{T_\ell^c + B_{Q_\ell}}\left(1 - \frac{R_\ell}{C_\ell}\right) = \lambda_{\boldsymbol{Q}} B_{Q_\ell}\left(1 - \frac{R_\ell}{C_\ell}\right)$$



An Example

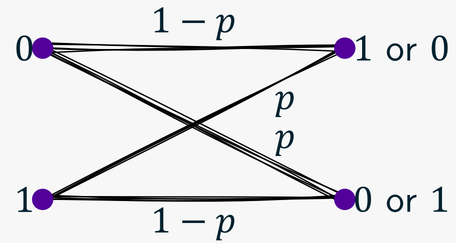
- $\boldsymbol{Q} = \{\mathsf{BSC}_p, \mathsf{BSC}_{1-p}\}$, p known
- Slope of Burnashev Exponent:

$$B_p = B_{1-p} = D(p \| 1-p)$$



An Example

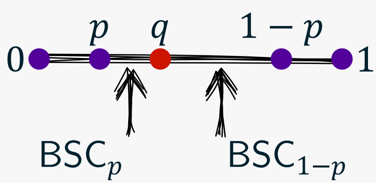
- $\mathcal{Q} = \{\mathrm{BSC}_p, \mathrm{BSC}_{1-p}\}$, p known



- Slope of Burnashev Exponent:

$$B_p = B_{1-p} = D(p \| 1 - p)$$

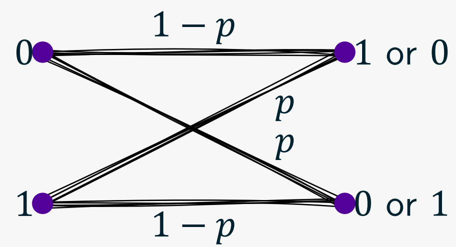
- Channel Estimation
  - Transmit all zero sequence
  - Freq of ones $< q$: estimate BSC_p



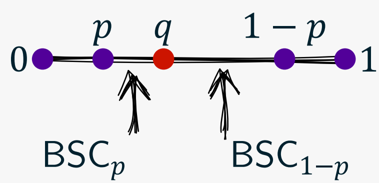
An Example

- $\boldsymbol{Q} = \{\mathrm{BSC}_p, \mathrm{BSC}_{1-p}\}$, p known
- Slope of Burnashev Exponent:

$$B_p = B_{1-p} = D(p\|1-p)$$



- Channel Estimation
  - Transmit all zero sequence
  - Freq of ones $< q$: estimate BSC_p



- Exponent of training error

$$T_p = D(p\|q), \quad T_{1-p} = D(1-p\|q)$$

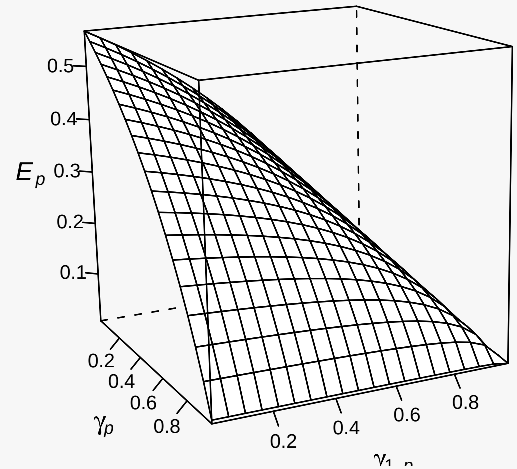
Performance evaluation

Rate $\boldsymbol{R} = (R_p, R_{1-p})$. Let $\gamma = R/C$.

- Error exponents

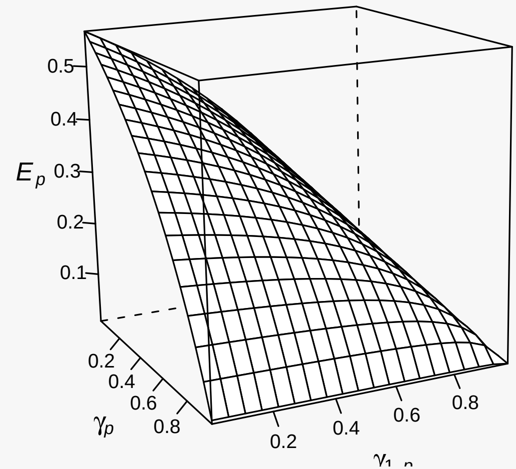
$$E_p \geq \frac{D(q|p) \cdot D(p|1-p)}{D(q|p) + D(p|1-p)}(1-\gamma_p)$$

$$E_{1-p} \geq \frac{D(q|1-p) \cdot D(p|1-p)}{D(q|1-p) + D(p|1-p)}(1-\gamma_{1-p})$$



Performance evaluation

Rate $\boldsymbol{R} = (R_p, R_{1-p})$. Let $\gamma = R/C$.



- **Error exponents**

$$E_p \geq \frac{D(q|p) \cdot D(p|1-p)}{D(q|p) + D(p|1-p)}(1-\gamma_p)$$

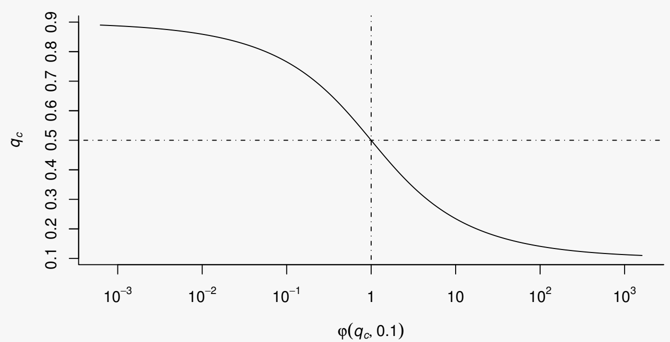
$$E_{1-p} \geq \frac{D(q|1-p) \cdot D(p|1-p)}{D(q|1-p) + D(p|1-p)}(1-\gamma_{1-p})$$

- **Optimal threshold q**

Choose q such that $E_p = E_{1-p}$: solve for q in

$$\varphi(q,p) = \frac{(1-\gamma_p)}{(1-\gamma_{1-p})}$$

where $\varphi(q,p)$ is appropriately defined



Conclusion

- Contributions
  - Defining opportunistic capacity and corresponding error exponent regions for compound channels with feedback.
  - A simple and easy to implement coding scheme whose error exponents are within a multiplicative factor of the best possible error exponents
  - In the presence of feedback, training based schemes can lead to reasonable performance

Conclusion

- Contributions
  - Defining opportunistic capacity and corresponding error exponent regions for compound channels with feedback.
  - A simple and easy to implement coding scheme whose error exponents are within a multiplicative factor of the best possible error exponents
  - In the presence of feedback, training based schemes can lead to reasonable performance
- Future directions
  - Channels defined over continuous families and continuous alphabets
  - Upper bound on error exponents

Thank you