Error Exponents of Compound Channel with Feedback

Aditya Mahajan and Sekhar Tatikonda
Yale University

ISIT 2010
Problem Setup

- **Compound channel**
  - Channel is memoryless
    \[ \mathbb{P}(Y_n \mid X^n, Y^{n-1}) = Q_o(Y_n \mid X_n) \]
  - Channel is not known completely
    \[ Q_o \in \mathcal{Q} := \{Q_1, Q_2, \ldots, Q_L\} \quad \text{(Family of DMCs)} \]
Problem Setup

Compound channel

- Channel is memoryless
  \[ \mathbb{P}(Y_n | X^n, Y^{n-1}) = Q_\circ(Y_n | X_n) \]
- Channel is not known completely
  \[ Q_\circ \in \mathcal{Q} := \{Q_1, Q_2, \ldots, Q_L\} \] (Family of DMCs)

Variable length communication

- Capacity? (easy)
- Error Exponents? (this talk)
Opportunistic Capacity

**Notation**

For a *(variable length)* coding scheme $S$

- $P^{(S)}_{\ell}$ = Prob of error when $Q_\circ = Q_\ell$
- $R^{(S)}_{\ell}$ = 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 $S$

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

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 $S$

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

- **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.**

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

  where $C_{Q_{\ell}} = \text{capacity of DMC } Q_{\ell}$. 
Opportunistic Capacity

- **Notation**
  
  For a (variable length) coding scheme $S$
  
  $P^{(S)}_\ell = \text{Prob of error when } Q_\circ = Q_\ell$
  
  $R^{(S)}_\circ = \text{Rate when } Q_\circ = Q_\circ$

  Opportunistic capacity is a region

  \[ \bigcup_{n \in \mathbb{N}} \text{ such that for any } \varepsilon > 0, \exists n_\circ(\varepsilon) \text{ so that for all } n > n_\circ(\varepsilon) \]
  \[ P^{(S_n)}_\ell < \varepsilon \quad \text{and} \quad R^{(S_n)}_\ell > R_\ell - \varepsilon \quad \text{for all } \ell = 1, 2, \ldots, L \]

  - **Opportunistic Capacity**: Union of all achievable rates.
    
    $\mathcal{C}_F(Q) = \{(R_1, R_2, \ldots, R_L) : 0 \leq R_\ell \leq C_{Q_\ell}, \ell = 1, 2, \ldots, L\}$
    
    where $C_{Q_\ell} = \text{capacity of DMC } Q_\ell$. 
Suppose a sequence \( \{S_n\}_{n \in \mathbb{N}} \) of coding schemes that achieves a rate vector \( \mathbf{R} = (R_1, R_2, \ldots, R_L) \)
Suppose a sequence \( \{S_n\}_{n \in \mathbb{N}} \) of coding schemes that achieves a rate vector \( \mathbf{R} = (R_1, R_2, \ldots, R_L) \).

Error exponent \( \mathbf{E} = (E_1, E_2, \ldots, E_L) \) of this scheme:

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

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

- Error exponent \( E = (E_1, E_2, \ldots, E_L) \) 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) \( \mathcal{E}(R) \)

Union of error exponent over all choices of coding schemes
Suppose a sequence \( \{S_n\}_{n \in \mathbb{N}} \) of coding schemes that achieves a rate vector \( R = (R_1, R_2, \ldots, R_L) \)

Error exponent \( E = (E_1, E_2, \ldots, E_L) \) of this scheme

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

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

Union of error exponent over all choices of coding schemes

Reliability \( \equiv \) Pareto frontier of EER
Suppose a sequence \( \{S_n\}_{n \in \mathbb{N}} \) of coding schemes that achieves a rate vector \( R = (R_1, R_2, \ldots, R_L) \).

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

Reliability \( \equiv \) Pareto frontier of EER.
Suppose $Q = \{Q_1, Q_2\}$ and $R = (R_1, R_2) \leq (C_1, C_2)$. 
Flavor of the results

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

\[ E_1^{(S_1)} = B_1(1 - R_1/C_1) \quad E_2^{(S_1)} = \text{small} \]
\[ E_1^{(S_2)} = \text{small} \quad E_2^{(S_2)} = B_2(1 - R_2/C_2) \]
Flavor of the results

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

\[ E_1^{(S_1)} = B_1(1 - R_1/C_1) \]
\[ E_1^{(S_2)} = \text{small} \]
\[ E_2^{(S_1)} = \text{small} \]
\[ E_2^{(S_2)} = B_2(1 - R_2/C_2) \]

Tchamkerten and Telatar, 2006

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

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

Let $\mathcal{P} = \{Q_1, Q_2\}$ and $\mathbf{R} = (R_1, R_2) \leq (C_1, C_2)$.

Let $S_\ell = \{S_{\ell,n}\}_{n \in \mathbb{N}}$ achieve Burnashev exponent for DMC $Q_\ell$.

\[ E_{1}^{(S_1)} = B_1(1 - R_1/C_1) \quad \text{and} \quad E_{2}^{(S_1)} = \text{small} \]
\[ E_{1}^{(S_2)} = \text{small} \quad \text{and} \quad E_{2}^{(S_2)} = B_2(1 - R_2/C_2) \]

Tchamkerten and Telatar, 2006

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

Our result: Propose a simple training based scheme such that

\[ E_\ell \geq \lambda Q B_\ell \left(1 - \frac{R_\ell}{C_\ell}\right), \quad \ell = 1, 2 \]
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

  \( \{ \text{BSC}_p, \text{BSC}_{1-p} \} \). Non-obvious channel estimation
Literature Overview
Variable length coding significantly boosts the error exponents [Burnashev, 1976]

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

where \( B_Q = \max_{x_A, x_R \in \mathcal{X}} D(Q(\cdot|x_A) \parallel Q(\cdot|x_R)) \)
Variable length coding significantly boosts the error exponents \[ E = B_Q \left( 1 - \frac{R}{C} \right) \]

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 \[ [\text{Yamamoto Itoh, 1979}] \]
DMC with Feedback

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

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

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

- 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
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
  \[ \text{[Tchamkerten Telatar, 2006a]} \]
  The slope of the error exponent goes to zero at rates near capacity

- Burnashev exponent can be achieved universally
  \[ \text{[Tchamkerten Telatar, 2006b]} \]
  The compound family is not too diverse
  \[ \Delta(Q) = 0 \]
Training based schemes have poor error exponent

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

Burnashev exponent can be achieved universally

The compound family is not too diverse

\[ \Delta(\mathcal{Q}) = 0 \]

Positive example: \( \mathcal{B} = \{ \text{BSC}_p : p \in [0, 1/2) \} \) and \( \mathcal{Z} = \{ Z_p : p \in [0, 1] \} \)
Compound Channel with Feedback

- Training based schemes have poor error exponent
  
  The slope of the error exponent goes to zero at rates near capacity

- Burnashev exponent can be achieved universally

  \[ \Delta(\mathcal{Q}) = 0 \]

  The compound family is not too diverse

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

- Negative example: \( Q_p = \{ \text{BSC}_p, \text{BSC}_{1-p} \}, p \) 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 $Q_o = Q_\ell, \ell = 1, 2, \ldots, L$
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 $\mathbf{W} = (W_1, W_2, \ldots, W_L)$ of size $\mathbf{M} = (M_1, M_2, \ldots, M_L)$
Compound Channel with Feedback

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

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

\[ X_1 = f_1(\mathbf{W}), \quad X_2 = f_2(\mathbf{W}, Y_1), \quad X_3 = f_3(\mathbf{W}, Y_1, Y_2), \quad \ldots \]
Compound Channel with Feedback

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

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

- Compound message $W = (W_1, W_2, \ldots, W_L)$ of size $M = (M_1, M_2, \ldots, M_L)$

- Encoding scheme
  \[ X_1 = f_1(W), \quad X_2 = f_2(W, Y_1), \quad X_3 = f_3(W, Y_1, Y_2), \quad \ldots \]

- Decoding scheme: at stopping time $\tau$, choose
  \[ (\hat{L}, \hat{W}) = g_{\tau}(Y_1, Y_2, \ldots, Y_{\tau}) \]
The coding scheme

Coding scheme is a tuple \((M, f, g, \tau)\)

- Compound message size
  \(M = (M_1, M_2, \ldots, M_L)\)

- Compound message alphabet
  \(\mathcal{M} = \prod_{\ell=1}^{L} \{1, 2, \ldots, M_{\ell}\}\)

- Encoding strategy \(f = (f_1, f_2, \ldots)\)
  \(f_t : \mathcal{M} \times \mathcal{Y}^{t-1} \mapsto \mathcal{X}\)

- Decoding strategy \(g = (g_1, g_2, \ldots)\)
  \(g_t : \mathcal{Y}^{t-1} \mapsto \bigcup_{\ell=1}^{L} \{(\ell, 1), (\ell, 2), \ldots, (\ell, M_{\ell})\}\)

- Stopping time \(\tau\) with respect to filtration \(\{2\mathcal{Y}^t, t \in \mathbb{N}\}\)
The coding scheme

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

- **Compound message size**
  \[ \mathcal{M} = (M_1, M_2, \ldots, M_L) \]

- **Compound message alphabet**
  \[ \mathcal{M} = \prod_{\ell=1}^{L} \{1, 2, \ldots, M_\ell\} \]

- **Encoding strategy** \(f = (f_1, f_2, \ldots)\)
  \[ f_t : \mathcal{M} \times \mathcal{Y}^{t-1} \mapsto \mathcal{X} \]

- **Decoding strategy** \(g = (g_1, g_2, \ldots)\)
  \[ g_t : \mathcal{Y}^{t-1} \mapsto \bigcup_{\ell=1}^{L} \{(\ell, 1), (\ell, 2), \ldots, (\ell, M_\ell)\} \]

- **Stopping time** \(\tau\) with respect to filtration \(\{2\mathcal{Y}_t, t \in \mathbb{N}\}\)

- **Rate** \(\mathbf{R} = (R_1, R_2, \ldots, R_L)\)
  \[ R_\ell = \frac{\mathbb{E}_\ell[\log_2 M_{\hat{L}}]}{\mathbb{E}_\ell[\tau]} \]

- **Probability of error** \(\mathbf{P} = (P_1, P_2, \ldots, P_L)\)
  \[ P_\ell = \mathbb{P}_\ell(\hat{W} \neq W_{\hat{L}}) \]
Achievable Rate: A rate \( \mathbf{R} = (R_1, R_2, \ldots, R_L) \) is achievable if there exists a sequence of coding schemes \((M^{(n)}, f^{(n)}, g^{(n)}, \tau^{(n)})\), \(n \in \mathbb{N}\), 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. \( \forall n \geq n_\circ(\varepsilon) \)

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

Capacity: Union of all achievable rates
Capacity and Error Exponents

- **Achievable Rate:** A rate $R = (R_1, R_2, \ldots, R_L)$ is achievable if there exists a sequence of coding schemes $(M^{(n)}, f^{(n)}, g^{(n)}, \tau^{(n)})$, $n \in \mathbb{N}$, 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. $\forall n \geq n_\circ(\varepsilon)$
     $$P_\ell^{(n)} < \varepsilon \quad \text{and} \quad R_\ell^{(n)} \geq R_\ell - \varepsilon, \quad \ell = 1, 2, \ldots, L$$

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

- **Error Exponent:** Given a sequence of coding schemes $(M^{(n)}, f^{(n)}, g^{(n)}, \tau^{(n)})$, that achieve a rate $R$, 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:** Pareto frontier of Error Exponent region
Main Result

- Opportunistic Capacity

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

where \( C_\ell \) is the capacity of DMC \( Q_\ell \).
Main Result

- **Opportunistic Capacity**

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

  where \(C_\ell\) is the capacity of DMC \(Q_\ell\).

- **Training based inner bound on Error Exponent Region**

  - Let \(T_\ell^c\) be the exponent of the channel estimation error when the channel is \(Q_\ell\). For any channel estimation scheme, \((T_1^c, \ldots, T_L^c) \in \mathcal{I}^*\).
Main Result

- **Opportunistic Capacity**

  \[ C_F(Q) = \{(R_1, \ldots, R_L) : 0 \leq R_\ell < C_\ell, \ell = 1, \ldots, L\} \]

  where \( C_\ell \) is the capacity of DMC \( Q_\ell \).

- **Training based inner bound on Error Exponent Region**

  - Let \( T_\ell^c \) be the exponent of the channel estimation error when the channel is \( Q_\ell \). For any channel estimation scheme, \((T_1^c, \ldots, T_L^c) \in \mathcal{T}^*\).

  - At rate \( R = (R_1, \ldots, R_L) \), the error exponent is

  \[
  E_\ell \geq \frac{T_\ell^c}{T_\ell^c + B_{Q_\ell}} B_{Q_\ell} \left( 1 - \frac{R_\ell}{C_\ell} \right)
  \]

  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 \(\hat{L}_r^{(k,n)}\)

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 $\kappa_\ell = \frac{T^c_\ell}{B_{Q_\ell}}$ and $\gamma_\ell = \frac{R_\ell}{C_\ell}$
Choice of parameters

Let \( \kappa_\ell = \frac{T^c_\ell}{B_{Q_\ell}} \) and \( \gamma_\ell = \frac{R_\ell}{C_\ell} \)

Choose a reference channel \( Q^* \)

\[
\alpha_\ell = \frac{(1 + \kappa_\ell)(1 - \gamma^*)}{(1 - \gamma_\ell)(1 + \kappa^*)}
\]
Choice of parameters

Let \( \kappa_\ell = \frac{T^c_\ell}{B_{Q_\ell}} \) and \( \gamma_\ell = \frac{R_\ell}{C_\ell} \)

- Choose a reference channel \( Q^* \)

\[
\alpha_\ell = \frac{(1 + \kappa_\ell)(1 - \gamma^*)}{(1 - \gamma_\ell)(1 + \kappa^*)}
\]

- Choice of parameters

\[
\beta_1^{(n)} \downarrow 0, \quad \beta_2^{(n)}(\ell) \downarrow \alpha_\ell \gamma_\ell, \\
\beta_3^{(n)} \uparrow \frac{(1 - \gamma^*)}{(1 + \kappa^*)}, \quad \beta_4^{(n)}(\ell) \uparrow \kappa_\ell \beta_3^{(n)}
\]
Choice of parameters

Let \( \kappa_\ell = \frac{T^c_\ell}{B_{Q_\ell}} \) and \( \gamma_\ell = \frac{R_\ell}{C_\ell} \)

- Choose a reference channel \( Q^* \)

\[
\alpha_\ell = \frac{(1 + \kappa_\ell)(1 - \gamma^*)}{(1 - \gamma_\ell)(1 + \kappa^*)}
\]

- Choice of parameters

\[
\beta_1^{(n)} \downarrow 0, \quad \beta_2^{(n)}(\ell) \downarrow \alpha_\ell \gamma_\ell, \\
\beta_3^{(n)} \uparrow \frac{(1 - \gamma^*)}{(1 + \kappa^*)}, \quad \beta_4^{(n)}(\ell) \uparrow \kappa_\ell \beta_3^{(n)}
\]

\[
= \alpha_\ell \frac{(1 - \gamma_\ell)}{(1 + \gamma_\ell)}
\]
**Choice of parameters**

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

- Choose a reference channel $Q^*$

\[
\alpha_\ell = \frac{(1 + \kappa_\ell)(1 - \gamma^*)}{(1 - \gamma_\ell)(1 + \kappa^*)}
\]

- Choice of parameters

\[
\begin{align*}
\beta_1^{(n)} & \downarrow 0, & \beta_2^{(n)}(\ell) & \downarrow \alpha_\ell \gamma_\ell, & \mathbb{E}_\ell [\beta_1^{(n)} + \beta_2^{(n)}(\ell)] & \downarrow \alpha_\ell \gamma_\ell \\
\beta_3^{(n)} & \uparrow \frac{(1 - \gamma^*)}{(1 + \kappa^*)}, & \beta_4^{(n)}(\ell) & \uparrow \kappa_\ell \beta_3^{(n)}, & \mathbb{E}_\ell [\beta_3^{(n)} + \beta_3^{(n)}(\ell)] & \uparrow \alpha_\ell (1 - \gamma_\ell)
\end{align*}
\]
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 $Q^*$

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

Choice of parameters

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

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

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 \(\ell\): No feedback comm over DMC \(Q_\ell\)
    - block length \(\beta_2^{(n)}(\ell)\); size \(M_\ell(n) = \lfloor 2^{n\alpha_\ell n_\ell C_\ell} \rfloor\) (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
  \[
  \max_{x_A, x_R} D(Q_\ell(\cdot|x_A) \parallel Q_\ell(\cdot|x_R))
  \]

- \(L\) decoding regions \(\mathcal{A}_\ell \subseteq \mathcal{Y}^{(n)}(\ell)^n\) that optimally distinguish
  \[
  \frac{(x_A(\ell), \ldots, x_A(\ell))}{\beta_4^{(n)}(\ell)\text{times}} \quad \text{from} \quad \frac{(x_R(\ell), \ldots, x_R(\ell))}{\beta_4^{(n)}(\ell)\text{times}}
  \]
  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
  \[ \Lambda^{(k,n)} = \beta_1^{(n)} + \beta_2^{(n)}(\hat{L}_m^{(k,n)}) + \beta_3^{(n)}(\hat{L}_c^{(k,n)}) \]
  \[ \mathbb{E}_\ell[\Lambda^{(k,n)}] \to \alpha_\ell \]

- 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

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

\[ M_\ell(n) = [2^{n\alpha_\ell \gamma_\ell C_\ell}] \]
Performance Analysis

Rate

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

\[M_\ell(n) = [2^{n\alpha_\ell \gamma_\ell C_\ell}]\]
Performance Analysis

■ Rate
\[
\lim_{n \to \infty} \frac{\mathbb{E}_\ell[\log M_{L_m^{(k,n)}}(n)]}{\mathbb{E}_\ell[\tau^{(n)}]} \approx \frac{n\alpha \gamma C}{\alpha n} = \gamma C = R_\ell
\]

■ Error Exponent
\[
E_\ell = \lim_{n \to \infty} \frac{-\log P_\ell^{(n)}}{\mathbb{E}_\ell[\tau^{(n)}]}
\approx \lim_{n \to \infty} -\frac{\log(\text{decoding error}) + \log(\text{hypothesis testing error})}{\alpha n}
\geq \lim_{n \to \infty} -\frac{\log(\text{hypothesis testing error})}{\alpha n}
\]

\[
M_\ell(n) = [2^{n\alpha \gamma C}]
\]
Performance Analysis

- Rate

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

- Error Exponent

\[
E_\ell = \lim_{n \to \infty} - \frac{\log P_\ell^{(n)}}{\mathbb{E}_\ell[\tau^{(n)}]}
\approx \lim_{n \to \infty} - \frac{\log(\text{decoding error}) + \log(\text{hypothesis testing error})}{\alpha_\ell n}
\]

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

\[
\approx -\log \left( e^{-\beta_3^{(n)}(k,n)} + e^{-\beta_4^{(n)}(\ell)n} \right) / (\alpha_\ell n)
\]

\[
M_\ell(n) = [2^{n\alpha_\ell \gamma_\ell C_\ell}]
\]
Performance Analysis

■ Rate

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

■ Error Exponent

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

\[M_\ell(n) = [2^{n\alpha_\ell \gamma_\ell C_\ell}]\]
Performance Analysis

Rate

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

Error Exponent

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

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

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

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

\[
\geq \frac{\beta_3^{(n)}(\square) \cdot (\square) + \beta_4^{(n)}(\ell)\square}{\square + (\square)} \frac{\beta_3^{(n)}(\square) + \beta_4^{(n)}(\ell)}{\alpha_\ell} = \frac{T_\ell^c \cdot B_{Q_\ell}}{T_\ell^c + B_{Q_\ell}} \left( 1 - \frac{R_\ell}{C_\ell} \right)
\]
Performance Analysis

- **Rate**

\[
\lim_{n \to \infty} \frac{\mathbb{E}_\ell[\log M_{\Lambda_m^{(k,n)}}(n)]}{\mathbb{E}_\ell[\tau^{(n)}]} \approx \frac{n\alpha \gamma \ell \epsilon}{\alpha \ell n} = \gamma \ell \epsilon = R \ell
\]

- **Error Exponent**

\[
E = \lim_{n \to \infty} -\log P^{(n)}_\ell - \frac{\log(\text{decoding error} + \log(\text{hypothesis testing error}))}{\mathbb{E}_\ell[\tau^{(n)}]} \leq \lim_{n \to \infty} -\log(\text{hypothesis testing error})/\alpha \ell n
\]

\[
= -\log \left( e^{-\beta_3^{(n)}(k)} + e^{-\beta_4^{(n)}(k)} \right) / (\alpha \ell n)
\]

\[
= T \ell^c \cdot B Q \ell \left( 1 - \frac{R \ell}{C \ell} \right) = \lambda Q B Q \ell \left( 1 - \frac{R \ell}{C \ell} \right)
\]
An Example

- $Q = \{\text{BSC}_p, \text{BSC}_{1-p}\}, \ p \ \text{known}$
- Slope of Burnashev Exponent:
  \begin{equation}
  B_p = B_{1-p} = D(p \| 1 - p)
  \end{equation}
An Example

- \( Q = \{\text{BSC}_p, \text{BSC}_{1-p}\}, \) \( p \) 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 \( \text{BSC}_p \)
An Example

- \( Q = \{ \text{BSC}_p, \text{BSC}_{1-p} \}, \) \( p \) 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 \( \text{BSC}_p \)

- Exponent of training error
  \[
  T_p = D(p \| q), \quad T_{1-p} = D(1 - p \| q)
  \]
Performance evaluation

Rate $\mathbf{R} = (R_p, R_{1-p})$. 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 $\mathbf{R} = (R_p, R_{1-p})$. 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