A General Formula for Compound Channel Capacity

Sergey Loyka, Charalambos D. Charalambous

University of Ottawa, University of Cyprus

ETH Zurich (May 2015), ISIT-15
1. Introduction
2. Channel Model
3. Definitions
4. Capacity for a given channel state
5. General Formula for Compound Channel Capacity
6. Uniform Compound Channel
7. Examples & mistakes
8. Conclusion
Introduction

- Impact of channel state information (CSI) on capacity/system design
- Real world: measurement/modeling uncertainty
- Wireless channels
  - dynamic nature
  - estimation error
  - feedback limitations
- Uncertain CSI: deterministic vs. stochastic models
Introduction

- Impact of channel state information (CSI) on capacity/system design
- Real world: measurement/modeling uncertainty
  - Wireless channels
    - dynamic nature
    - estimation error
    - feedback limitations
- Uncertain CSI: deterministic vs. stochastic models
Introduction

- Impact of channel state information (CSI) on capacity/system design
- Real world: measurement/modeling uncertainty
- Wireless channels
  - dynamic nature
  - estimation error
  - feedback limitations
- Uncertain CSI: deterministic vs. stochastic models
Introduction

- Impact of channel state information (CSI) on capacity/system design
- Real world: measurement/modeling uncertainty
- Wireless channels
  - dynamic nature
  - estimation error
  - feedback limitations
- Uncertain CSI: deterministic vs. stochastic models
Introduction

- Impact of channel uncertainty: extensive studies since 1950s

- Several approaches
  - compound channel model
  - mixed/composite channel
  - arbitrary varying channel
  - CDI instead of CSI

---


Impact of channel uncertainty: extensive studies since 1950s\textsuperscript{12}

Several approaches

- compound channel model
- mixed/composite channel
- arbitrary varying channel
- CDI instead of CSI


Introduction

- **Compound channel/capacity**
  - Prior work: information-stable channels (e.g. stationary ergodic)
  - Many/all channels are non-stationary, non-ergodic
    - modulation-induced channels
    - quasi-static fading channels
    - how to measure ergodicity ????
  - This work: information-unstable channels
    - no ergodicity/stationarity is required
    - via information density/spectrum\(^3\)


Introduction

- **Compound channel/capacity**
- **Prior work: information-stable channels (e.g. stationary ergodic)**
  - Many/all channels are non-stationary, non-ergodic
    - modulation-induced channels
    - quasi-static fading channels
    - how to measure ergodicity ???
  - This work: information-unstable channels
    - no ergodicity/stationarity is required
    - via information density/spectrum$^{34}$

---


Introduction

- Compound channel/capacity
- Prior work: information-stable channels (e.g. stationary ergodic)
- Many/all channels are non-stationary, non-ergodic
  - modulation-induced channels
  - quasi-static fading channels
  - how to measure ergodicity ???
- This work: information-unstable channels
  - no ergodicity/stationarity is required
  - via information density/spectrum\(^3\)


Introduction

- Compound channel/capacity
- Prior work: information-stable channels (e.g. stationary ergodic)
- Many/all channels are non-stationary, non-ergodic
  - modulation-induced channels
  - quasi-static fading channels
  - how to measure ergodicity ???
- This work: information-unstable channels
  - no ergodicity/stationarity is required
  - via information density/spectrum

---

Compound Channel Model I

- a channel is selected from a given uncertainty set
- stays fixed during entire transmission
- design a single code good for any channel in the set
- largest achievable rate = compound capacity
- No CSI at Tx, complete CSI at Rx
Compound Channel Model II

- $X^n = \{X_1...X_n\}$ - (random) sequence of $n$ input symbols
- $Y^n$ - corresponding output sequence;
- $s \in S$ - channel state (which may also be a sequence)
- $p(x^n), p_s(y^n)$ - input/output distributions under state $s$.
- $p_s(y^n|x^n)$ - channel transition probability;
- information density ("unexpected" mutual information):

$$i(x^n; y^n, s) = \ln \frac{p_s(x^n, y^n)}{p(x^n)p_s(y^n)} = i(x^n; y^n|s)$$
Compound Channel Model III

- If information-stable:
  \[
  \frac{1}{n} i(x^n; y^n|s) \rightarrow I(X; Y|s) \text{ as } n \rightarrow \infty \tag{2}
  \]

- If information-unstable:
  \[
  \frac{1}{n} i(x^n; y^n|s) \rightarrow \text{RV} \tag{3}
  \]
Definitions I

- \((n, r_n, \varepsilon_{ns})\)-code:
  - \(n\) - the block length
  - \(\varepsilon_{ns}\) - the error probability (for channel state \(s\))
  - \(r_n = \ln M_n/n\) - the code rate
  - \(M_n\) - the number of codewords

- compound error probability:
  \[
  \varepsilon_n = \sup_{s \in S} \varepsilon_{ns}
  \]  
  \(\text{(4)}\)

- achievable rate \(R\): \(\exists (n, r_n, \varepsilon_{ns})\)-code such that
  \[
  \lim_{n \to \infty} \sup \varepsilon_n = 0, \quad \lim_{n \to \infty} \inf r_n \geq R
  \]  
  \(\text{(5)}\)

- compound channel capacity \(C\):
  \[
  C = \sup_{R} \{R : R \text{ is achievable}\}
  \]  
  \(\text{(6)}\)
Capacity for a given channel state

- Very general - information-unstable channels
- Given channel state $s$, the capacity is [Verdu-Han'94] $^5$

$$\begin{align*}
C(s) &= \sup_{p(x)} I(X; Y|s) \\
&= \sup_{p(x)} \left( R : \lim_{n \to \infty} \Pr \{ n^{-1} i(X^n; Y^n|s) \leq R \} = 0 \right)
\end{align*}$$

where $I(X; Y|s)$ is the inf-information rate

$$
I(X; Y|s) = \sup_{R} \left\{ R : \lim_{n \to \infty} \Pr \{ n^{-1} i(X^n; Y^n|s) \leq R \} = 0 \right\}
$$

- Proof: via Feinstein and Verdu-Han Lemmas

---

Lemma (Feinstein)

For arbitrary input $X^n$, any $r_n$ and a given channel state $s$, there exists a code satisfying the following inequality,

$$\varepsilon_{ns} \leq \Pr \left\{ n^{-1} i(X^n; Y^n | s) \leq r_n + \gamma \right\} + e^{-\gamma n} \tag{9}$$

for any $\gamma > 0$.

Lemma (Verdu-Han’94)

Every $(n, r_n, \varepsilon_{ns})$-code satisfies the following inequality,

$$\varepsilon_{ns} \geq \Pr \left\{ n^{-1} i(X^n; Y^n | s) \leq r_n - \gamma \right\} - e^{-\gamma n} \tag{10}$$

for any $\gamma > 0$, where $X^n$ is uniformly distributed over all codewords and $Y^n$ is the corresponding channel output under channel state $s$. 
Capacity for a given channel state

Lemma (Feinstein)

For arbitrary input $X^n$, any $r_n$ and a given channel state $s$, there exists a code satisfying the following inequality,

$$\varepsilon_{ns} \leq \Pr \left\{ n^{-1} i(X^n; Y^n|s) \leq r_n + \gamma \right\} + e^{-\gamma n} \quad (9)$$

for any $\gamma > 0$.

Lemma (Verdu-Han’94)

Every $(n, r_n, \varepsilon_{ns})$-code satisfies the following inequality,

$$\varepsilon_{ns} \geq \Pr \left\{ n^{-1} i(X^n; Y^n|s) \leq r_n - \gamma \right\} - e^{-\gamma n} \quad (10)$$

for any $\gamma > 0$, where $X^n$ is uniformly distributed over all codewords and $Y^n$ is the corresponding channel output under channel state $s$. 
Theorem

Consider a general compound channel; the Rx knows $s \in S$, but not the Tx; the Tx knows the (arbitrary) $S$. The compound channel capacity is

$$C_c = \sup_{p(x)} I(X; Y) \tag{11}$$

where $I(X; Y) = \sup_R \{R \in \Omega\}$ is the compound inf-information rate,

$$\Omega = \left\{ R : \lim_{n \to \infty} \sup_{s \in S} \Pr \left\{ \frac{1}{n} i(X^n; Y^n|s) \leq R \right\} = 0 \right\} \tag{12}$$

Proof:

via Feinstein and Verdu-Han Lemmas extended to compound setting
Consider a general compound channel; the Rx knows $s \in S$, but not the Tx; the Tx knows the (arbitrary) $S$. The compound channel capacity is

$$C_c = \sup_{p(x)} I(X;Y)$$ \hspace{1cm} (11)

where $I(X;Y) = \sup_R \{R \in \Omega\}$ is the compound inf-information rate,

$$\Omega = \left\{ R : \lim_{n \to \infty} \sup_{s \in S} \Pr \left\{ \frac{1}{n} i(X^n;Y^n|s) \leq R \right\} = 0 \right\}$$ \hspace{1cm} (12)

Proof:

via Feinstein and Verdu-Han Lemmas extended to compound setting
Compound Lemmas

Lemma (compound Feinstein)

For any $X^n$, $S$ and $r_n$, there exists a $(n, r_n, \varepsilon_n)$-code (where the codewords are independent of channel state $s$), such that, for any $\gamma > 0$,

$$
\varepsilon_n \leq \sup_{s \in S} \Pr \left\{ \frac{1}{n} I(X^n; Y^n | s) \leq r_n + \gamma \right\} + e^{-\gamma n}
$$

(13)

Lemma (compound Verdu-Han)

Every $(n, r_n, \varepsilon_n)$-code satisfies

$$
\varepsilon_n \geq \sup_{s \in S} \Pr \left\{ \frac{1}{n} I(X^n; Y^n | s) \leq r_n - \gamma \right\} - e^{-\gamma n}
$$

(14)

for any $\gamma > 0$, where $X^n$ is uniformly distributed over all codewords and $Y^n$ is the corresponding channel output under channel state $s$. 
Lemma (compound Feinstein)

For any $X^n$, $S$ and $r_n$, there exists a $(n, r_n, \varepsilon_n)$-code (where the codewords are independent of channel state $s$), such that, for any $\gamma > 0$,

$$\varepsilon_n \leq \sup_{s \in S} \Pr \left\{ n^{-1} i(X^n; Y^n|s) \leq r_n + \gamma \right\} + e^{-\gamma n}$$

(13)

Lemma (compound Verdu-Han)

Every $(n, r_n, \varepsilon_n)$-code satisfies

$$\varepsilon_n \geq \sup_{s \in S} \Pr \left\{ n^{-1} i(X^n; Y^n|s) \leq r_n - \gamma \right\} - e^{-\gamma n}$$

(14)

for any $\gamma > 0$, where $X^n$ is uniformly distributed over all codewords and $Y^n$ is the corresponding channel output under channel state $s$. 
Remark

$I(X, Y)$ is an extension of $I(X, Y|s)$ to the compound channel setting, not $\inf_s I(X, Y|s)$, in the general case.

The relationship between $I(X, Y)$ and $\inf_s I(X, Y|s)$ is established below.

Proposition

The following inequality holds for a general compound channel

$$I(X, Y) \leq I(X, Y) = \inf_s I(X, Y|s)$$  \hspace{1cm} (15)

Strict inequality can be shown by examples.
Remark

$I(X, Y)$ is an extension of $I(X, Y|s)$ to the compound channel setting, not $\inf_s I(X, Y|s)$, in the general case.

The relationship between $I(X, Y)$ and $\inf_s I(X, Y|s)$ is established below.

Proposition

The following inequality holds for a general compound channel

$$I(X, Y) \leq I(X, Y) = \inf_s I(X, Y|s)$$

(15)

Strict inequality can be shown by examples.
When does the equality hold?

**Definition**

A compound channel is uniform if

\[
\Pr \left\{ n^{-1} i(X^n; Y^n|s) \leq I(X, Y) - \gamma \right\} \to 0 \quad \forall \gamma > 0 \quad (16)
\]

uniformly in \( s \in S \) as \( n \to \infty \).

**Proposition**

The following equality holds for a uniform compound channel

\[
\underline{I}(X, Y) = I(X, Y) = \inf_s I(X, Y|s) \quad (17)
\]
When does the equality hold?

**Definition**

A compound channel is uniform if

$$ \Pr \left\{ n^{-1} i(X^n; Y^n | s) \leq I(X, Y) - \gamma \right\} \to 0 \ \forall \gamma > 0 $$

uniformly in $s \in S$ as $n \to \infty$.  

**Proposition**

The following equality holds for a uniform compound channel

$$ I(X, Y) = I(X, Y) = \inf_s I(X, Y | s) $$
Uniform Compound Channel

**Theorem**

The capacity of a uniform compound channel:

\[
C_c = \sup_{p(x)} \inf_{s \in S} I(X; Y|s)
\]  

(18)

Does not hold for non-uniform channels, contrary to what Theorem 3.3.5 in [Han’03]\(^6\) claims.

---

Examples & mistakes

- Binary non-stationary channel with memory:

\[ p_s(y^n|x^n) = p_s(y^n) \text{ if } n \leq s \]
\[ = \text{BSC}(0) \text{ if } n > s \]  \hspace{1cm} (19)

- Models the noise coherence time \( \tau = s \in S = \{1, 2, \ldots\} \).

- It follows that

\[ I(X; Y) = 0 < I(X; Y) = \inf_s I(X; Y|s) = \ln 2 \] \hspace{1cm} (21)

\[ C_c = \sup_{p(x)} I(X; Y) = 0 < \ln 2 = \sup_{p(x)} \inf_{s \in S} I(X; Y|s) \] \hspace{1cm} (22)

- Mistake: Theorem 3.3.5 in [Han’03] claims \( C_c = \ln 2 \).
Mistake: Theorem 3.3.5 in [Han’03] claims $C_c = \ln 2$.

Reason: improper error probability definition for compound channel:

$$\lim_{n \to \infty} \varepsilon_{ns} = 0 \quad \forall s \Rightarrow \lim_{n \to \infty} \sup_s \varepsilon_{ns} = 0 \tag{23}$$

In general,

$$\sup_s \lim_{n \to \infty} \varepsilon_{ns} \leq \lim_{n \to \infty} \sup_s \varepsilon_{ns} \tag{24}$$

In our example,

$$0 = \sup_s \lim_{n \to \infty} \varepsilon_{ns} < \lim_{n \to \infty} \sup_s \varepsilon_{ns} = 1 \tag{25}$$
Examples & mistakes

- Binary compound channel with additive noise:

\[ Y_k = X_k + Z_{ks}, \quad Z^n_s = \{ w_1, w_2, \ldots w_s, 0, 0 \ldots 0 \} \]  

(26)

and \( w_1 \ldots w_s \) are i.i.d. equiprobable.

- Its capacity

\[ C_c = \sup_{p(x)} I(X; Y) = 0 < \ln 2 = \sup_{p(x)} \inf_{s \in S} I(X; Y|s) \]  

(27)

- Why \( C_c = 0 \)?

- For any \( n \), does not matter how large, there are always channel states \( s \geq n \) for which the channel is BSC(1/2), i.e. useless.

- The standard sup – inf expression falls short of the channel capacity because this compound channel is not uniform.
Conclusion

- CSI uncertainty - compound channel
- Real world - information-unstable channels
- General formula for compound channel capacity
- Uniform compound channels
- Examples & mistakes