香港中文大學
The Chinese University of Hong Kong

Chapter 11
Continuous-Valued Channels

The Chinese University of Hong Kong

Preamble

Preamble

- In a physical communication system, the input and output of a channel often take continuous real values.

Preamble

- In a physical communication system, the input and output of a channel often take continuous real values.
- A waveform channel is one such that transmission is in continuous time.

Preamble

- In a physical communication system, the input and output of a channel often take continuous real values.
- A waveform channel is one such that transmission is in continuous time.
- At the physical layer, we need to deal with channels such that the values taken are continuous and transmission is in continuous time.

11.1 Discrete-Time Channel

Definition 11.1 (Continuous Channel I) Let $f(y|x)$ be a conditional pdf defined for all x, where

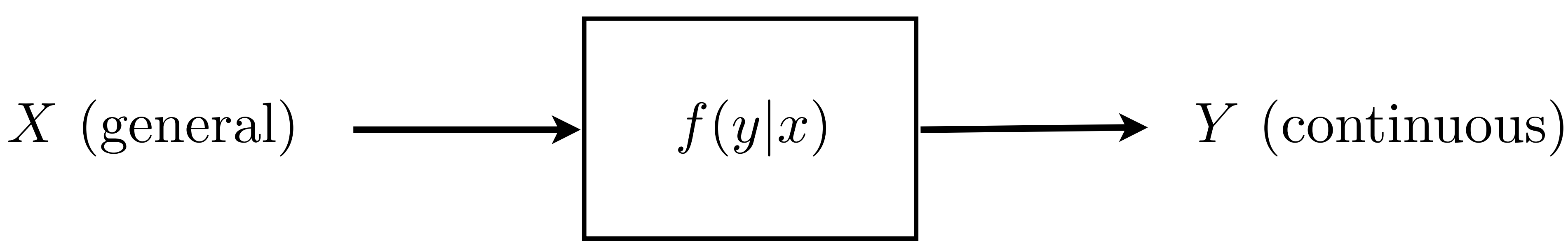
$$-\int_{\mathcal{S}_Y(x)} f(y|x) \log f(y|x) dy < \infty$$

for all x. A (discrete-time) continuous channel $f(y|x)$ is a system with input random variable X and output random variable Y such that Y is related to X through $f(y|x)$.

Definition 11.1 (Continuous Channel I) Let $f(y|x)$ be a conditional pdf defined for all x, where

$$-\int_{\mathcal{S}_Y(x)} f(y|x) \log f(y|x) dy < \infty$$

for all x. A (discrete-time) continuous channel $f(y|x)$ is a system with input random variable X and output random variable Y such that Y is related to X through $f(y|x)$.

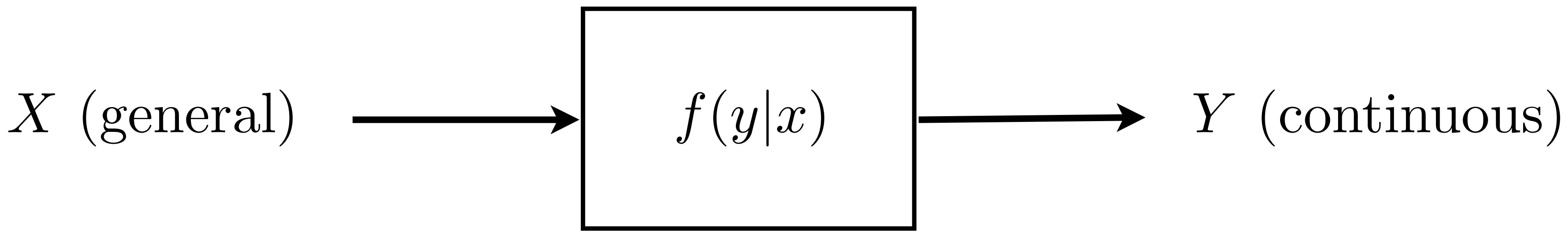


Definition 11.1 (Continuous Channel I) Let $f(y|x)$ be a conditional pdf defined for all x, where

$$-\int_{\mathcal{S}_Y(x)} f(y|x) \log f(y|x) dy < \infty$$

for all x. A (discrete-time) continuous channel $f(y|x)$ is a system with input random variable X and output random variable Y such that Y is related to X through $f(y|x)$.

Remark The integral in Definition 11.1 is precisely the conditional differential entropy $h(Y|X = x)$, which is required to be finite.

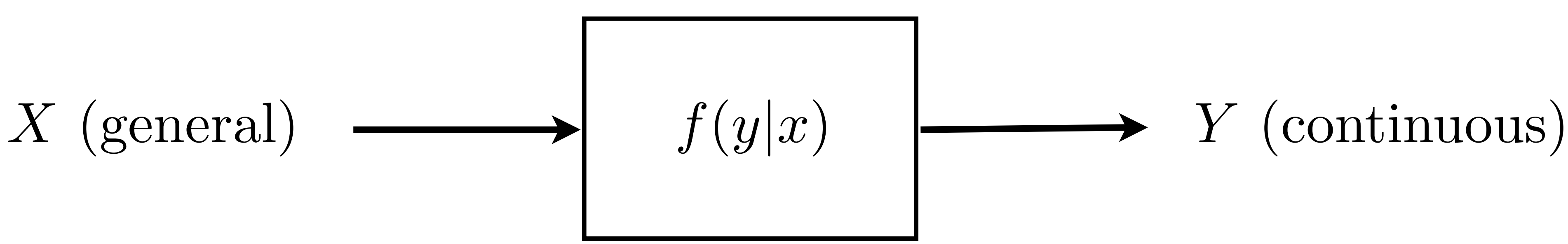


Definition 11.1 (Continuous Channel I) Let $f(y|x)$ be a conditional pdf defined for all x, where

$$h(Y|X=x) = -\int_{\mathcal{S}_Y(x)} f(y|x)\log f(y|x)dy < \infty$$

for all x. A (discrete-time) continuous channel $f(y|x)$ is a system with input random variable X and output random variable Y such that Y is related to X through $f(y|x)$.

Remark The integral in Definition 11.1 is precisely the conditional differential entropy $h(Y|X=x)$, which is required to be finite.

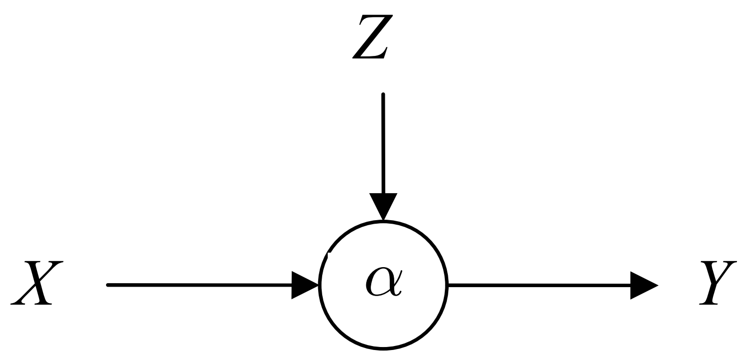


Definition 11.2 (Continuous Channel II) Let $\alpha : \Re \times \Re \to \Re$, and Z be a real random variable, called the noise variable. A (discrete-time) continuous channel (α, Z) is a system with a real input and a real output. For any input random variable X, the noise random variable Z is independent of X, and the output random variable Y is given by

$$Y = \alpha(X, Z).$$

Definition 11.2 (Continuous Channel II) Let $\alpha : \Re \times \Re \to \Re$, and Z be a real random variable, called the noise variable. A (discrete-time) continuous channel (α, Z) is a system with a real input and a real output. For any input random variable X, the noise random variable Z is independent of X, and the output random variable Y is given by

$$Y = \alpha(X, Z).$$



Definition 11.3 Two continuous channels $f(y|x)$ and (α, Z) are equivalent if for every input distribution $F(x)$,

$$\Pr\{X \leq x, \alpha(X, Z) \leq y\} = \int_{-\infty}^{x} \int_{-\infty}^{y} f_{Y|X}(v|u) dv \, dF_X(u)$$

for all x and y.

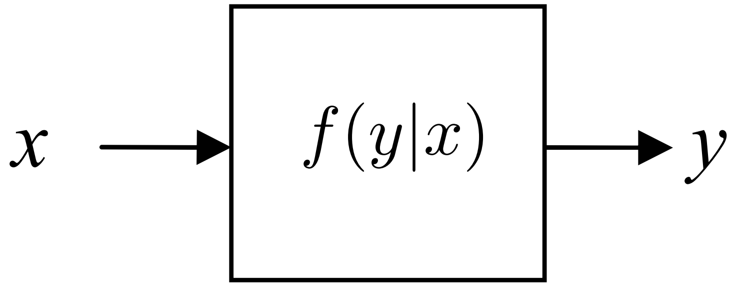
Definition 11.3 Two continuous channels $f(y|x)$ and (α, Z) are equivalent if for every input distribution $F(x)$,

$$\Pr\{X \le x, \alpha(X,Z) \le y\} = \int_{-\infty}^{x} \int_{-\infty}^{y} f_{Y|X}(v|u)\,dv\,dF_X(u)$$

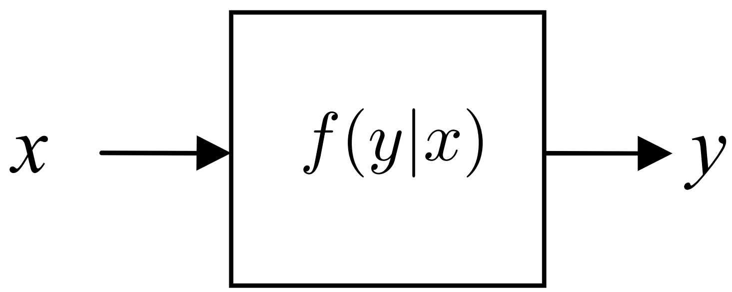
$F_{XY}(X \le x, Y \le y)$

for all x and y.

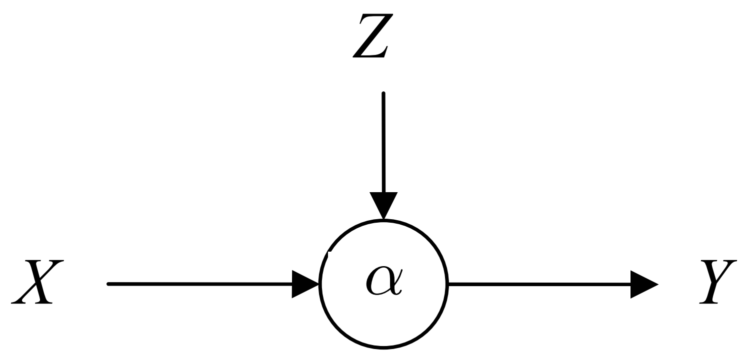
Definition 11.1



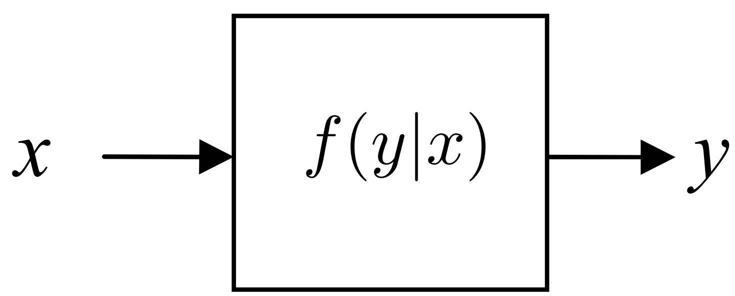
Definition 11.1



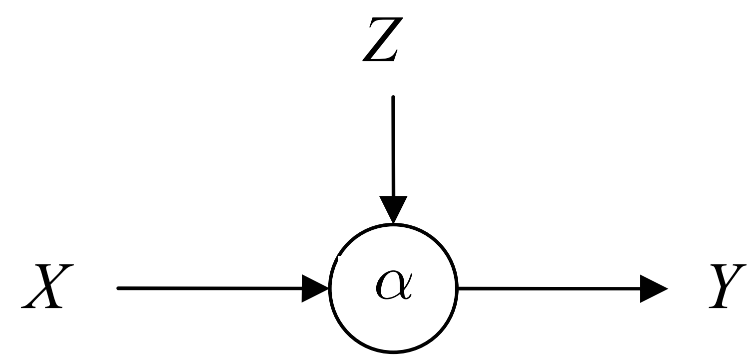
Definition 11.2



Definition 11.1

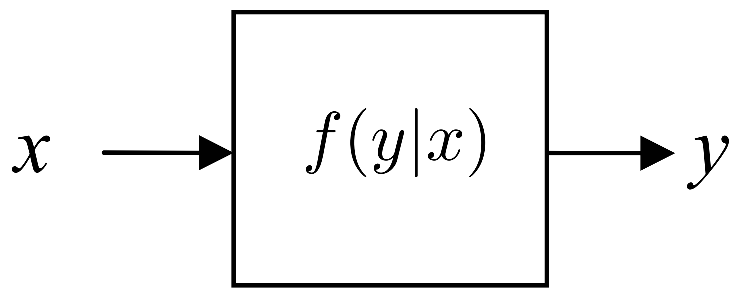


Definition 11.2

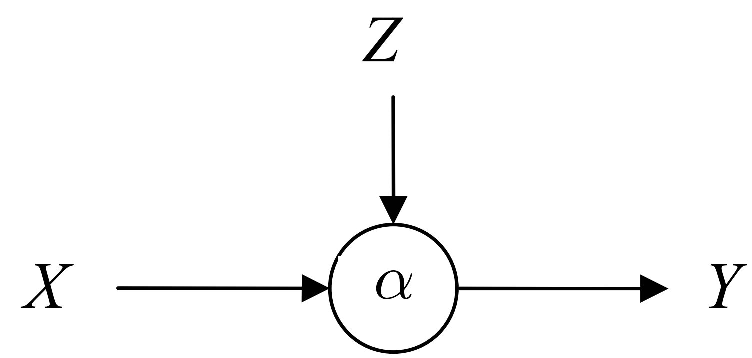


- Definition 11.2 is more general than Definition 11.1 because the former does not require the existence of $f(y|x)$.

Definition 11.1



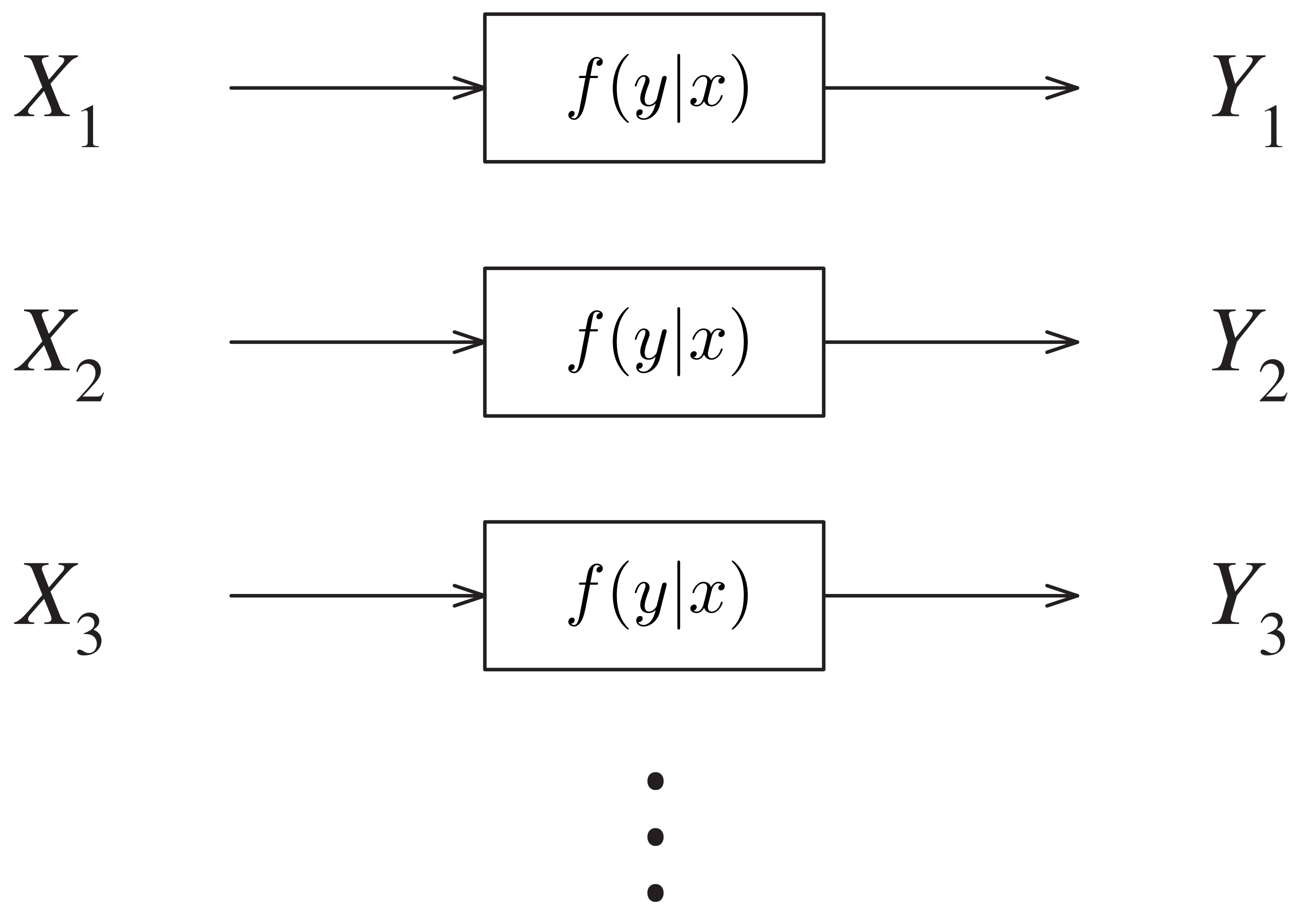
Definition 11.2



- Definition 11.2 is more general than Definition 11.1 because the former does not require the existence of $f(y|x)$.
- We confine our discussion to channels defined by Definition 11.1.

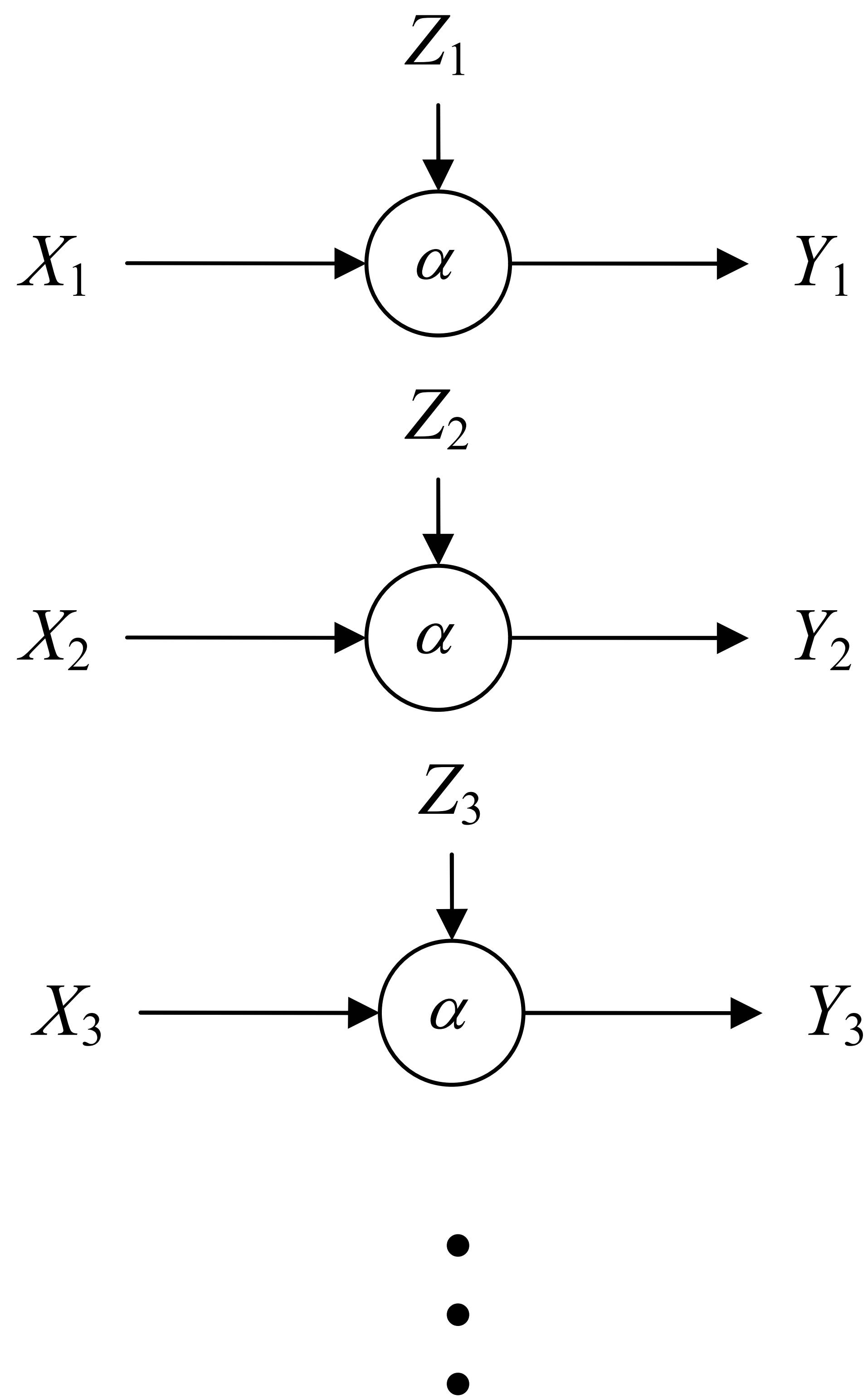
Definition 11.4 (CMC I) A continuous memoryless channel (CMC) $f(y|x)$ is a sequence of replicates of a generic continuous channel $f(y|x)$. These continuous channels are indexed by a discrete-time index i, where $i \geq 1$, with the ith channel being available for transmission at time i. Transmission through a channel is assumed to be instantaneous. Let X_i and Y_i be respectively the input and the output of the CMC at time i, and let T_{i-} denote all the random variables that are generated in the system before X_i. The Markov chain $T_{i-} \to X_i \to Y_i$ holds, and

$$\Pr\{X_i \leq x, Y_i \leq y\} = \int_{-\infty}^{x} \int_{-\infty}^{y} f_{Y|X}(v|u) dv \, dF_{X_i}(u).$$


X_1
$f(y|x)$
Y_1
X_2
$f(y|x)$
Y_2
X_3
$f(y|x)$
Y_3

Definition 11.5 (CMC II) A continuous memoryless channel (α, Z) is a sequence of replicates of a generic continuous channel (α, Z). These continuous channels are indexed by a discrete-time index i, where $i \geq 1$, with the ith channel being available for transmission at time i. Transmission through a channel is assumed to be instantaneous. Let X_i and Y_i be respectively the input and the output of the CMC at time i, and let T_{i-} denote all the random variables that are generated in the system before X_i. The noise variable Z_i for the transmission at time i is a copy of the generic noise variable Z, and is independent of (X_i, T_{i-}). The output of the CMC at time i is given by

$$Y_i = \alpha(X_i, Z_i).$$


Z_1
X_1
α
Y_1
Z_2
X_2
α
Y_2
Z_3
X_3
α
Y_3

Definition 11.6 Let κ be a real function. An average input constraint (κ, P) for a CMC is the requirement that for any codeword $(x_1, x_2, \cdots, x_n)$ transmitted over the channel,

$$\frac{1}{n}\sum_{i=1}^{n} \kappa(x_i) \leq P.$$

Definition 11.6 Let κ be a real function. An average input constraint (κ, P) for a CMC is the requirement that for any codeword $(x_1, x_2, \cdots, x_n)$ transmitted over the channel,

$$\frac{1}{n} \sum_{i=1}^{n} \kappa(x_i) \leq P.$$

Remarks

Definition 11.6 Let κ be a real function. An average input constraint (κ, P) for a CMC is the requirement that for any codeword $(x_1, x_2, \cdots, x_n)$ transmitted over the channel,

$$\frac{1}{n}\sum_{i=1}^{n}\kappa(x_i) \leq P.$$

Remarks

1. For a fixed value of x, $\kappa(x)$ is the "cost" for transmitting x.

Definition 11.6 Let κ be a real function. An average input constraint (κ, P) for a CMC is the requirement that for any codeword $(x_1, x_2, \cdots, x_n)$ transmitted over the channel,

$$\frac{1}{n} \sum_{i=1}^{n} \kappa(x_i) \leq P.$$

Remarks

1. For a fixed value of x, $\kappa(x)$ is the "cost" for transmitting x.
2. For example, if $\kappa(x) = x^2$, then $\kappa(x)$ is the energy, and P is the power.

Definition 11.6 Let κ be a real function. An average input constraint (κ, P) for a CMC is the requirement that for any codeword $(x_1, x_2, \cdots, x_n)$ transmitted over the channel,

$$\frac{1}{n}\sum_{i=1}^{n}\kappa(x_i) \leq P.$$

Remarks

1. For a fixed value of x, $\kappa(x)$ is the "cost" for transmitting x.

2. For example, if $\kappa(x) = x^2$, then $\kappa(x)$ is the energy, and P is the power.

Definition 11.7 The capacity of a continuous memoryless channel $f(y|x)$ with input constraint (κ, P) is defined as

$$C(P) = \sup_{F(x):E\kappa(X)\leq P} I(X;Y).$$

Theorem 11.8 $C(P)$ is non-decreasing, concave, and left-continuous.

Theorem 11.8 $C(P)$ is non-decreasing, concave, and left-continuous.

Ideas of Proof

Theorem 11.8 $C(P)$ is non-decreasing, concave, and left-continuous.

Ideas of Proof

1. Non-decreasing – immediate.

Theorem 11.8 $C(P)$ is non-decreasing, concave, and left-continuous.

Ideas of Proof

1. Non-decreasing – immediate.

2. Concave – a consequence of the the concavity of mutual information with respect to the input distribution.

Theorem 11.8 $C(P)$ is non-decreasing, concave, and left-continuous.

Ideas of Proof

1. Non-decreasing – immediate.
2. Concave – a consequence of the the concavity of mutual information with respect to the input distribution.
3. Left-continuous – a consequence of concavity.

Theorem 11.8 $C(P)$ is non-decreasing, concave, and left-continuous.

Ideas of Proof

1. Non-decreasing – immediate.
2. Concave – a consequence of the the concavity of mutual information with respect to the input distribution.
3. Left-continuous – a consequence of concavity.

Remarks

Theorem 11.8 $C(P)$ is non-decreasing, concave, and left-continuous.

Ideas of Proof

1. Non-decreasing – immediate.
2. Concave – a consequence of the the concavity of mutual information with respect to the input distribution.
3. Left-continuous – a consequence of concavity.

Remarks

1. $C(P)$ is also right-continous (a consequence of concavity) but requires a separate proof.

Theorem 11.8 $C(P)$ is non-decreasing, concave, and left-continuous.

Ideas of Proof

1. Non-decreasing – immediate.
2. Concave – a consequence of the the concavity of mutual information with respect to the input distribution.
3. Left-continuous – a consequence of concavity.

Remarks

1. $C(P)$ is also right-continous (a consequence of concavity) but requires a separate proof.
2. This property of $C(P)$ is not used in this chapter.

Theorem 11.8

$$C(P) = \sup_{F(x):E\kappa(X)\leq P} I(X;Y) \tag{1}$$

is non-decreasing, concave, and left-continuous.

Proof

Theorem 11.8

$$C(P) = \sup_{F(x):E\kappa(X)\leq P} I(X;Y) \tag{1}$$

is non-decreasing, concave, and left-continuous.

Proof

<u>$C(P)$ is non-decreasing</u>

Theorem 11.8

$$C(P) = \sup_{F(x): E\kappa(X) \le P} I(X;Y) \tag{1}$$

is non-decreasing, concave, and left-continuous.

Proof

<u>$C(P)$ is non-decreasing</u>

Note that in (1), the supremum is taken over a larger set for a larger P.

Theorem 11.8

$$C(P) = \sup_{F(x): E\kappa(X) \leq P} I(X;Y) \tag{1}$$

is non-decreasing, concave, and left-continuous.

Proof

$C(P)$ is non-decreasing

Note that in (1), the supremum is taken over a larger set for a larger P.

Theorem 11.8

$$C(P) = \sup_{F(x):E\kappa(X)\leq P} I(X;Y) \qquad (1)$$

is non-decreasing, concave, and left-continuous.

Proof

$C(P)$ is non-decreasing

Note that in (1), the supremum is taken over a larger set for a larger P.

$C(P)$ is concave

Theorem 11.8

$$C(P) = \sup_{F(x): E\kappa(X) \leq P} I(X;Y) \tag{1}$$

is non-decreasing, concave, and left-continuous.

Proof

$C(P)$ is non-decreasing

Note that in (1), the supremum is taken over a larger set for a larger P.

$C(P)$ is concave

1. Let $j = 1, 2$. For an input distribution $F_j(x)$, denote the corresponding input and output random variables by X_j and Y_j, respectively.

Theorem 11.8

$$C(P) = \sup_{F(x): E\kappa(X) \le P} I(X;Y) \tag{1}$$

is non-decreasing, concave, and left-continuous.

Proof

$C(P)$ is non-decreasing

Note that in (1), the supremum is taken over a larger set for a larger P.

$C(P)$ is concave

1. Let $j = 1, 2$. For an input distribution $F_j(x)$, denote the corresponding input and output random variables by X_j and Y_j, respectively.

2. Then for any P_j, for all $\epsilon > 0$, there exists $F_j(x)$ such that

Theorem 11.8

$$C(P) = \sup_{F(x):E\kappa(X)\leq P} I(X;Y) \tag{1}$$

is non-decreasing, concave, and left-continuous.

Proof

$C(P)$ is non-decreasing

Note that in (1), the supremum is taken over a larger set for a larger P.

$C(P)$ is concave

1. Let $j = 1, 2$. For an input distribution $F_j(x)$, denote the corresponding input and output random variables by X_j and Y_j, respectively.

2. Then for any P_j, for all $\epsilon > 0$, there exists $F_j(x)$ such that

$$E\kappa(X_j) \leq P_j$$

Theorem 11.8

$$C(P) = \sup_{F(x): E\kappa(X) \leq P} \underline{I(X;Y)} \qquad (1)$$

is non-decreasing, concave, and left-continuous.

Proof

$C(P)$ is non-decreasing

Note that in (1), the supremum is taken over a larger set for a larger P.

$C(P)$ is concave

1. Let $j = 1, 2$. For an input distribution $F_j(x)$, denote the corresponding input and output random variables by X_j and Y_j, respectively.

2. Then for any P_j, for all $\epsilon > 0$, there exists $F_j(x)$ such that

$$E\kappa(X_j) \leq P_j$$

and

$$I(X_j;Y_j) \geq C(P_j) - \epsilon.$$

Theorem 11.8

$$C(P) = \sup_{F(x): E\kappa(X) \le P} I(X;Y) \tag{1}$$

is non-decreasing, concave, and left-continuous.

Proof

$C(P)$ is non-decreasing

Note that in (1), the supremum is taken over a larger set for a larger P.

$C(P)$ is concave

1. Let $j = 1, 2$. For an input distribution $F_j(x)$, denote the corresponding input and output random variables by X_j and Y_j, respectively.

2. Then for any P_j, for all $\epsilon > 0$, there exists $F_j(x)$ such that

$$E\kappa(X_j) \le P_j$$

and

$$I(X_j; Y_j) \ge C(P_j) - \epsilon.$$

Theorem 11.8

$$C(P) = \sup_{F(x): E\kappa(X) \le P} I(X;Y) \tag{1}$$

is non-decreasing, concave, and left-continuous.

Proof

$C(P)$ is non-decreasing

Note that in (1), the supremum is taken over a larger set for a larger P.

$C(P)$ is concave

1. Let $j = 1, 2$. For an input distribution $F_j(x)$, denote the corresponding input and output random variables by X_j and Y_j, respectively.

2. Then for any P_j, for all $\epsilon > 0$, there exists $F_j(x)$ such that

$$E\kappa(X_j) \le P_j$$

and

$$I(X_j; Y_j) \ge C(P_j) - \epsilon.$$

Theorem 11.8

$$C(P) = \sup_{F(x): E\kappa(X) \leq P} I(X;Y) \qquad (1)$$

is non-decreasing, concave, and left-continuous.

Proof

<u>$C(P)$ is non-decreasing</u>

Note that in (1), the supremum is taken over a larger set for a larger P.

<u>$C(P)$ is concave</u>

1. Let $j = 1, 2$. For an input distribution $F_j(x)$, denote the corresponding input and output random variables by X_j and Y_j, respectively.

2. Then for any P_j, for all $\epsilon > 0$, there exists $F_j(x)$ such that

$$E\kappa(X_j) \leq P_j$$

and

$$I(X_j; Y_j) \geq C(P_j) - \epsilon.$$

3. For $0 \leq \lambda \leq 1$, let $\bar{\lambda} = 1 - \lambda$ and define the random variable

$$X^{(\lambda)} \sim \lambda F_1(x) + \bar{\lambda} F_2(x).$$

Theorem 11.8

$$C(P) = \sup_{F(x): E\kappa(X) \le P} I(X;Y) \qquad (1)$$

is non-decreasing, concave, and left-continuous.

Proof

<u>$C(P)$ is non-decreasing</u>

Note that in (1), the supremum is taken over a larger set for a larger P.

<u>$C(P)$ is concave</u>

1. Let $j = 1, 2$. For an input distribution $F_j(x)$, denote the corresponding input and output random variables by X_j and Y_j, respectively.

2. Then for any P_j, for all $\epsilon > 0$, there exists $F_j(x)$ such that

$$E\kappa(X_j) \le P_j$$

and

$$I(X_j; Y_j) \ge C(P_j) - \epsilon.$$

3. For $0 \le \lambda \le 1$, let $\bar{\lambda} = 1 - \lambda$ and define the random variable

$$X^{(\lambda)} \sim \lambda F_1(x) + \bar{\lambda} F_2(x).$$

Then

$$E\kappa(X^{(\lambda)}) = \lambda E\kappa(X_1) + \bar{\lambda} E\kappa(X_2)$$

Theorem 11.8

$$C(P) = \sup_{F(x): E\kappa(X) \le P} I(X;Y) \qquad (1)$$

is non-decreasing, concave, and left-continuous.

Proof

$C(P)$ is non-decreasing

Note that in (1), the supremum is taken over a larger set for a larger P.

$C(P)$ is concave

1. Let $j = 1, 2$. For an input distribution $F_j(x)$, denote the corresponding input and output random variables by X_j and Y_j, respectively.

2. Then for any P_j, for all $\epsilon > 0$, there exists $F_j(x)$ such that

$$E\kappa(X_j) \le P_j$$

and

$$I(X_j; Y_j) \ge C(P_j) - \epsilon.$$

3. For $0 \le \lambda \le 1$, let $\bar{\lambda} = 1 - \lambda$ and define the random variable

$$X^{(\lambda)} \sim \lambda F_1(x) + \bar{\lambda} F_2(x).$$

Then

$$E\kappa(X^{(\lambda)}) = \lambda E\kappa(X_1) + \bar{\lambda} E\kappa(X_2) \le \lambda P_1 + \bar{\lambda} P_2.$$

Theorem 11.8

$$C(P) = \sup_{F(x): E\kappa(X) \le P} I(X;Y) \quad (1)$$

is non-decreasing, concave, and left-continuous.

Proof

$C(P)$ is non-decreasing

Note that in (1), the supremum is taken over a larger set for a larger P.

$C(P)$ is concave

1. Let $j = 1, 2$. For an input distribution $F_j(x)$, denote the corresponding input and output random variables by X_j and Y_j, respectively.

2. Then for any P_j, for all $\epsilon > 0$, there exists $F_j(x)$ such that

$$E\kappa(X_j) \le P_j$$

and

$$I(X_j; Y_j) \ge C(P_j) - \epsilon.$$

3. For $0 \le \lambda \le 1$, let $\bar{\lambda} = 1 - \lambda$ and define the random variable

$$X^{(\lambda)} \sim \lambda F_1(x) + \bar{\lambda} F_2(x).$$

Then

$$E\kappa(X^{(\lambda)}) = \lambda E\kappa(X_1) + \bar{\lambda} E\kappa(X_2) \le \lambda P_1 + \bar{\lambda} P_2.$$

4. By the concavity of mutual information with respect to the input distribution, we have

Theorem 11.8

$$C(P) = \sup_{F(x):E\kappa(X)\leq P} I(X;Y) \qquad (1)$$

is non-decreasing, concave, and left-continuous.

Proof

$C(P)$ is non-decreasing

Note that in (1), the supremum is taken over a larger set for a larger P.

$C(P)$ is concave

1. Let $j = 1, 2$. For an input distribution $F_j(x)$, denote the corresponding input and output random variables by X_j and Y_j, respectively.

2. Then for any P_j, for all $\epsilon > 0$, there exists $F_j(x)$ such that

$$E\kappa(X_j) \leq P_j$$

and

$$I(X_j;Y_j) \geq C(P_j) - \epsilon.$$

3. For $0 \leq \lambda \leq 1$, let $\bar{\lambda} = 1 - \lambda$ and define the random variable

$$X^{(\lambda)} \sim \lambda F_1(x) + \bar{\lambda} F_2(x).$$

Then

$$E\kappa(X^{(\lambda)}) = \lambda E\kappa(X_1) + \bar{\lambda} E\kappa(X_2) \leq \lambda P_1 + \bar{\lambda} P_2.$$

4. By the concavity of mutual information with respect to the input distribution, we have

$$I(X^{(\lambda)};Y^{(\lambda)}) \quad \geq \quad \lambda I(X_1;Y_1) + \bar{\lambda} I(X_2;Y_2)$$

Theorem 11.8

$$C(P) = \sup_{F(x): E\kappa(X) \le P} I(X;Y) \tag{1}$$

is non-decreasing, concave, and left-continuous.

Proof

<u>$C(P)$ is non-decreasing</u>

Note that in (1), the supremum is taken over a larger set for a larger P.

<u>$C(P)$ is concave</u>

1. Let $j = 1, 2$. For an input distribution $F_j(x)$, denote the corresponding input and output random variables by X_j and Y_j, respectively.

2. Then for any P_j, for all $\epsilon > 0$, there exists $F_j(x)$ such that

$$E\kappa(X_j) \le P_j$$

and

$$I(X_j; Y_j) \ge C(P_j) - \epsilon.$$

3. For $0 \le \lambda \le 1$, let $\bar{\lambda} = 1 - \lambda$ and define the random variable

$$X^{(\lambda)} \sim \lambda F_1(x) + \bar{\lambda} F_2(x).$$

Then

$$E\kappa(X^{(\lambda)}) = \lambda E\kappa(X_1) + \bar{\lambda} E\kappa(X_2) \le \lambda P_1 + \bar{\lambda} P_2.$$

4. By the concavity of mutual information with respect to the input distribution, we have

$$\begin{aligned} I(X^{(\lambda)}; Y^{(\lambda)}) &\ge \lambda I(X_1; Y_1) + \bar{\lambda} I(X_2; Y_2) \\ &\ge \lambda (C(P_1) - \epsilon) + \bar{\lambda}(C(P_2) - \epsilon) \end{aligned}$$

Theorem 11.8

$$C(P) = \sup_{F(x): E\kappa(X) \leq P} I(X;Y) \qquad (1)$$

is non-decreasing, concave, and left-continuous.

Proof

$C(P)$ is non-decreasing

Note that in (1), the supremum is taken over a larger set for a larger P.

$C(P)$ is concave

1. Let $j = 1, 2$. For an input distribution $F_j(x)$, denote the corresponding input and output random variables by X_j and Y_j, respectively.

2. Then for any P_j, for all $\epsilon > 0$, there exists $F_j(x)$ such that

$$E\kappa(X_j) \leq P_j$$

and

$$I(X_j; Y_j) \geq C(P_j) - \epsilon.$$

3. For $0 \leq \lambda \leq 1$, let $\bar{\lambda} = 1 - \lambda$ and define the random variable

$$X^{(\lambda)} \sim \lambda F_1(x) + \bar{\lambda} F_2(x).$$

Then

$$E\kappa(X^{(\lambda)}) = \lambda E\kappa(X_1) + \bar{\lambda} E\kappa(X_2) \leq \lambda P_1 + \bar{\lambda} P_2.$$

4. By the concavity of mutual information with respect to the input distribution, we have

$$\begin{aligned} I(X^{(\lambda)}; Y^{(\lambda)}) &\geq \lambda I(X_1; Y_1) + \bar{\lambda} I(X_2; Y_2) \\ &\geq \lambda (C(P_1) - \epsilon) + \bar{\lambda}(C(P_2) - \epsilon) \\ &= \lambda C(P_1) + \bar{\lambda} C(P_2) - \epsilon. \end{aligned}$$

Theorem 11.8

$$C(P) = \sup_{F(x): E\kappa(X) \le P} I(X;Y) \qquad (1)$$

is non-decreasing, concave, and left-continuous.

Proof

$C(P)$ is non-decreasing

Note that in (1), the supremum is taken over a larger set for a larger P.

$C(P)$ is concave

1. Let $j = 1, 2$. For an input distribution $F_j(x)$, denote the corresponding input and output random variables by X_j and Y_j, respectively.

2. Then for any P_j, for all $\epsilon > 0$, there exists $F_j(x)$ such that

$$E\kappa(X_j) \le P_j$$

and

$$I(X_j; Y_j) \ge C(P_j) - \epsilon.$$

3. For $0 \le \lambda \le 1$, let $\bar{\lambda} = 1 - \lambda$ and define the random variable

$$X^{(\lambda)} \sim \lambda F_1(x) + \bar{\lambda} F_2(x).$$

Then

$$E\kappa(X^{(\lambda)}) = \lambda E\kappa(X_1) + \bar{\lambda} E\kappa(X_2) \le \lambda P_1 + \bar{\lambda} P_2.$$

4. By the concavity of mutual information with respect to the input distribution, we have

$$\begin{aligned} I(X^{(\lambda)}; Y^{(\lambda)}) &\ge \lambda I(X_1; Y_1) + \bar{\lambda} I(X_2; Y_2) \\ &\ge \lambda (C(P_1) - \epsilon) + \bar{\lambda}(C(P_2) - \epsilon) \\ &= \lambda C(P_1) + \bar{\lambda} C(P_2) - \epsilon. \end{aligned}$$

Theorem 11.8

$$C(P) = \sup_{F(x): E\kappa(X) \leq P} I(X;Y) \tag{1}$$

is non-decreasing, concave, and left-continuous.

Proof

$C(P)$ is non-decreasing

Note that in (1), the supremum is taken over a larger set for a larger P.

$C(P)$ is concave

1. Let $j = 1, 2$. For an input distribution $F_j(x)$, denote the corresponding input and output random variables by X_j and Y_j, respectively.

2. Then for any P_j, for all $\epsilon > 0$, there exists $F_j(x)$ such that

$$E\kappa(X_j) \leq P_j$$

and

$$I(X_j;Y_j) \geq C(P_j) - \epsilon.$$

3. For $0 \leq \lambda \leq 1$, let $\bar{\lambda} = 1 - \lambda$ and define the random variable

$$X^{(\lambda)} \sim \lambda F_1(x) + \bar{\lambda} F_2(x).$$

Then

$$E\kappa(X^{(\lambda)}) = \lambda E\kappa(X_1) + \bar{\lambda} E\kappa(X_2) \leq \lambda P_1 + \bar{\lambda} P_2.$$

4. By the concavity of mutual information with respect to the input distribution, we have

$$\begin{aligned} I(X^{(\lambda)};Y^{(\lambda)}) &\geq \lambda I(X_1;Y_1) + \bar{\lambda} I(X_2;Y_2) \\ &\geq \lambda(C(P_1) - \epsilon) + \bar{\lambda}(C(P_2) - \epsilon) \\ &= \lambda C(P_1) + \bar{\lambda} C(P_2) - \epsilon. \end{aligned}$$

Theorem 11.8

$$C(P) = \sup_{F(x): E\kappa(X) \le P} I(X;Y) \qquad (1)$$

is non-decreasing, concave, and left-continuous.

Proof

$C(P)$ is non-decreasing

Note that in (1), the supremum is taken over a larger set for a larger P.

$C(P)$ is concave

1. Let $j = 1, 2$. For an input distribution $F_j(x)$, denote the corresponding input and output random variables by X_j and Y_j, respectively.

2. Then for any P_j, for all $\epsilon > 0$, there exists $F_j(x)$ such that

$$E\kappa(X_j) \le P_j$$

and

$$I(X_j; Y_j) \ge C(P_j) - \epsilon.$$

3. For $0 \le \lambda \le 1$, let $\bar{\lambda} = 1 - \lambda$ and define the random variable

$$X^{(\lambda)} \sim \lambda F_1(x) + \bar{\lambda} F_2(x).$$

Then

$$E\kappa(X^{(\lambda)}) = \lambda E\kappa(X_1) + \bar{\lambda} E\kappa(X_2) \le \lambda P_1 + \bar{\lambda} P_2.$$

4. By the concavity of mutual information with respect to the input distribution, we have

$$\begin{aligned} I(X^{(\lambda)}; Y^{(\lambda)}) &\ge \lambda I(X_1; Y_1) + \bar{\lambda} I(X_2; Y_2) \\ &\ge \lambda (C(P_1) - \epsilon) + \bar{\lambda}(C(P_2) - \epsilon) \\ &= \lambda C(P_1) + \bar{\lambda} C(P_2) - \epsilon. \end{aligned}$$

Theorem 11.8

$$C(P) = \sup_{F(x):E\kappa(X)\le P} I(X;Y) \quad (1)$$

is non-decreasing, concave, and left-continuous.

Proof

$C(P)$ is non-decreasing

Note that in (1), the supremum is taken over a larger set for a larger P.

$C(P)$ is concave

1. Let $j = 1, 2$. For an input distribution $F_j(x)$, denote the corresponding input and output random variables by X_j and Y_j, respectively.

2. Then for any P_j, for all $\epsilon > 0$, there exists $F_j(x)$ such that

$$E\kappa(X_j) \le P_j$$

and

$$I(X_j; Y_j) \ge C(P_j) - \epsilon.$$

3. For $0 \le \lambda \le 1$, let $\bar{\lambda} = 1 - \lambda$ and define the random variable

$$X^{(\lambda)} \sim \lambda F_1(x) + \bar{\lambda} F_2(x).$$

Then

$$E\kappa(X^{(\lambda)}) = \lambda E\kappa(X_1) + \bar{\lambda} E\kappa(X_2) \le \lambda P_1 + \bar{\lambda} P_2.$$

4. By the concavity of mutual information with respect to the input distribution, we have

$$\begin{aligned} I(X^{(\lambda)}; Y^{(\lambda)}) &\ge \lambda I(X_1; Y_1) + \bar{\lambda} I(X_2; Y_2) \\ &\ge \lambda (C(P_1) - \epsilon) + \bar{\lambda}(C(P_2) - \epsilon) \\ &= \lambda C(P_1) + \bar{\lambda} C(P_2) - \epsilon. \end{aligned}$$

5. Then

$C(\lambda P_1 + \bar{\lambda} P_2)$

Theorem 11.8

$$C(P) = \sup_{F(x): E\kappa(X) \le P} I(X;Y) \tag{1}$$

is non-decreasing, concave, and left-continuous.

Proof

$C(P)$ is non-decreasing

Note that in (1), the supremum is taken over a larger set for a larger P.

$C(P)$ is concave

1. Let $j = 1, 2$. For an input distribution $F_j(x)$, denote the corresponding input and output random variables by X_j and Y_j, respectively.

2. Then for any P_j, for all $\epsilon > 0$, there exists $F_j(x)$ such that

$$E\kappa(X_j) \le P_j$$

and

$$I(X_j; Y_j) \ge C(P_j) - \epsilon.$$

3. For $0 \le \lambda \le 1$, let $\bar{\lambda} = 1 - \lambda$ and define the random variable

$$X^{(\lambda)} \sim \lambda F_1(x) + \bar{\lambda} F_2(x).$$

Then

$$E\kappa(X^{(\lambda)}) = \lambda E\kappa(X_1) + \bar{\lambda} E\kappa(X_2) \le \lambda P_1 + \bar{\lambda} P_2.$$

4. By the concavity of mutual information with respect to the input distribution, we have

$$\begin{aligned} I(X^{(\lambda)}; Y^{(\lambda)}) &\ge \lambda I(X_1; Y_1) + \bar{\lambda} I(X_2; Y_2) \\ &\ge \lambda (C(P_1) - \epsilon) + \bar{\lambda}(C(P_2) - \epsilon) \\ &= \lambda C(P_1) + \bar{\lambda} C(P_2) - \epsilon. \end{aligned}$$

5. Then

$$C(\lambda P_1 + \bar{\lambda} P_2) \ge I(X^{(\lambda)}; Y^{(\lambda)})$$

Theorem 11.8

$$C(P) = \sup_{F(x): E\kappa(X) \le P} I(X;Y) \quad (1)$$

is non-decreasing, concave, and left-continuous.

Proof

<u>$C(P)$ is non-decreasing</u>

Note that in (1), the supremum is taken over a larger set for a larger P.

<u>$C(P)$ is concave</u>

1. Let $j = 1, 2$. For an input distribution $F_j(x)$, denote the corresponding input and output random variables by X_j and Y_j, respectively.

2. Then for any P_j, for all $\epsilon > 0$, there exists $F_j(x)$ such that

$$E\kappa(X_j) \le P_j$$

and

$$I(X_j; Y_j) \ge C(P_j) - \epsilon.$$

3. For $0 \le \lambda \le 1$, let $\bar{\lambda} = 1 - \lambda$ and define the random variable

$$X^{(\lambda)} \sim \lambda F_1(x) + \bar{\lambda} F_2(x).$$

Then

$$E\kappa(X^{(\lambda)}) = \lambda E\kappa(X_1) + \bar{\lambda} E\kappa(X_2) \le \lambda P_1 + \bar{\lambda} P_2.$$

4. By the concavity of mutual information with respect to the input distribution, we have

$$\begin{aligned} I(X^{(\lambda)}; Y^{(\lambda)}) &\ge \lambda I(X_1; Y_1) + \bar{\lambda} I(X_2; Y_2) \\ &\ge \lambda (C(P_1) - \epsilon) + \bar{\lambda}(C(P_2) - \epsilon) \\ &= \lambda C(P_1) + \bar{\lambda} C(P_2) - \epsilon. \end{aligned}$$

5. Then

$$C(\lambda P_1 + \bar{\lambda} P_2) \ge I(X^{(\lambda)}; Y^{(\lambda)}) \ge \lambda C(P_1) + \bar{\lambda} C(P_2) - \epsilon.$$

Theorem 11.8

$$C(P) = \sup_{F(x):E\kappa(X)\leq P} I(X;Y) \qquad (1)$$

is non-decreasing, concave, and left-continuous.

Proof

<u>$C(P)$ is non-decreasing</u>

Note that in (1), the supremum is taken over a larger set for a larger P.

<u>$C(P)$ is concave</u>

1. Let $j = 1, 2$. For an input distribution $F_j(x)$, denote the corresponding input and output random variables by X_j and Y_j, respectively.

2. Then for any P_j, for all $\epsilon > 0$, there exists $F_j(x)$ such that

$$E\kappa(X_j) \leq P_j$$

and

$$I(X_j; Y_j) \geq C(P_j) - \epsilon.$$

3. For $0 \leq \lambda \leq 1$, let $\bar{\lambda} = 1 - \lambda$ and define the random variable

$$X^{(\lambda)} \sim \lambda F_1(x) + \bar{\lambda} F_2(x).$$

Then

$$E\kappa(X^{(\lambda)}) = \lambda E\kappa(X_1) + \bar{\lambda} E\kappa(X_2) \leq \lambda P_1 + \bar{\lambda} P_2.$$

4. By the concavity of mutual information with respect to the input distribution, we have

$$\begin{aligned} I(X^{(\lambda)}; Y^{(\lambda)}) &\geq \lambda I(X_1; Y_1) + \bar{\lambda} I(X_2; Y_2) \\ &\geq \lambda(C(P_1) - \epsilon) + \bar{\lambda}(C(P_2) - \epsilon) \\ &= \lambda C(P_1) + \bar{\lambda} C(P_2) - \epsilon. \end{aligned}$$

5. Then

$$C(\lambda P_1 + \bar{\lambda} P_2) \geq I(X^{(\lambda)}; Y^{(\lambda)}) \geq \lambda C(P_1) + \bar{\lambda} C(P_2) - \epsilon.$$

6. Letting $\epsilon \to 0$, we have

Theorem 11.8

$$C(P) = \sup_{F(x):E\kappa(X)\le P} I(X;Y) \tag{1}$$

is non-decreasing, concave, and left-continuous.

Proof

<u>$C(P)$ is non-decreasing</u>

Note that in (1), the supremum is taken over a larger set for a larger P.

<u>$C(P)$ is concave</u>

1. Let $j = 1, 2$. For an input distribution $F_j(x)$, denote the corresponding input and output random variables by X_j and Y_j, respectively.

2. Then for any P_j, for all $\epsilon > 0$, there exists $F_j(x)$ such that

$$E\kappa(X_j) \le P_j$$

and

$$I(X_j; Y_j) \ge C(P_j) - \epsilon.$$

3. For $0 \le \lambda \le 1$, let $\bar{\lambda} = 1 - \lambda$ and define the random variable

$$X^{(\lambda)} \sim \lambda F_1(x) + \bar{\lambda} F_2(x).$$

Then

$$E\kappa(X^{(\lambda)}) = \lambda E\kappa(X_1) + \bar{\lambda} E\kappa(X_2) \le \lambda P_1 + \bar{\lambda} P_2.$$

4. By the concavity of mutual information with respect to the input distribution, we have

$$\begin{aligned} I(X^{(\lambda)}; Y^{(\lambda)}) &\ge \lambda I(X_1; Y_1) + \bar{\lambda} I(X_2; Y_2) \\ &\ge \lambda(C(P_1) - \epsilon) + \bar{\lambda}(C(P_2) - \epsilon) \\ &= \lambda C(P_1) + \bar{\lambda} C(P_2) - \epsilon. \end{aligned}$$

5. Then

$$C(\lambda P_1 + \bar{\lambda} P_2) \ge I(X^{(\lambda)}; Y^{(\lambda)}) \ge \lambda C(P_1) + \bar{\lambda} C(P_2) - \epsilon.$$

6. Letting $\epsilon \to 0$, we have

$$C(\lambda P_1 + \bar{\lambda} P_2) \ge \lambda C(P_1) + \bar{\lambda} C(P_2), \tag{2}$$

Theorem 11.8

$$C(P) = \sup_{F(x):E\kappa(X)\leq P} I(X;Y) \tag{1}$$

is non-decreasing, concave, and left-continuous.

Proof

<u>$C(P)$ is non-decreasing</u>

Note that in (1), the supremum is taken over a larger set for a larger P.

<u>$C(P)$ is concave</u>

1. Let $j = 1, 2$. For an input distribution $F_j(x)$, denote the corresponding input and output random variables by X_j and Y_j, respectively.

2. Then for any P_j, for all $\epsilon > 0$, there exists $F_j(x)$ such that

$$E\kappa(X_j) \leq P_j$$

and

$$I(X_j; Y_j) \geq C(P_j) - \epsilon.$$

3. For $0 \leq \lambda \leq 1$, let $\bar{\lambda} = 1 - \lambda$ and define the random variable

$$X^{(\lambda)} \sim \lambda F_1(x) + \bar{\lambda} F_2(x).$$

Then

$$E\kappa(X^{(\lambda)}) = \lambda E\kappa(X_1) + \bar{\lambda} E\kappa(X_2) \leq \lambda P_1 + \bar{\lambda} P_2.$$

4. By the concavity of mutual information with respect to the input distribution, we have

$$\begin{aligned} I(X^{(\lambda)}; Y^{(\lambda)}) &\geq \lambda I(X_1; Y_1) + \bar{\lambda} I(X_2; Y_2) \\ &\geq \lambda(C(P_1) - \epsilon) + \bar{\lambda}(C(P_2) - \epsilon) \\ &= \lambda C(P_1) + \bar{\lambda} C(P_2) - \epsilon. \end{aligned}$$

5. Then

$$C(\lambda P_1 + \bar{\lambda} P_2) \geq I(X^{(\lambda)}; Y^{(\lambda)}) \geq \lambda C(P_1) + \bar{\lambda} C(P_2) - \cancel{\epsilon}.$$

6. Letting $\epsilon \to 0$, we have

$$C(\lambda P_1 + \bar{\lambda} P_2) \geq \lambda C(P_1) + \bar{\lambda} C(P_2), \tag{2}$$

proving that $C(P)$ is concave.

Theorem 11.8

$$C(P) = \sup_{F(x):E\kappa(X)\le P} I(X;Y) \tag{1}$$

is non-decreasing, concave, and left-continuous.

Proof

<u>$C(P)$ is non-decreasing</u>

Note that in (1), the supremum is taken over a larger set for a larger P.

<u>$C(P)$ is concave</u>

1. Let $j = 1, 2$. For an input distribution $F_j(x)$, denote the corresponding input and output random variables by X_j and Y_j, respectively.

2. Then for any P_j, for all $\epsilon > 0$, there exists $F_j(x)$ such that

$$E\kappa(X_j) \le P_j$$

and

$$I(X_j;Y_j) \ge C(P_j) - \epsilon.$$

3. For $0 \le \lambda \le 1$, let $\bar{\lambda} = 1 - \lambda$ and define the random variable

$$X^{(\lambda)} \sim \lambda F_1(x) + \bar{\lambda} F_2(x).$$

Then

$$E\kappa(X^{(\lambda)}) = \lambda E\kappa(X_1) + \bar{\lambda} E\kappa(X_2) \le \lambda P_1 + \bar{\lambda} P_2.$$

4. By the concavity of mutual information with respect to the input distribution, we have

$$\begin{aligned} I(X^{(\lambda)};Y^{(\lambda)}) &\ge \lambda I(X_1;Y_1) + \bar{\lambda} I(X_2;Y_2) \\ &\ge \lambda(C(P_1) - \epsilon) + \bar{\lambda}(C(P_2) - \epsilon) \\ &= \lambda C(P_1) + \bar{\lambda} C(P_2) - \epsilon. \end{aligned}$$

5. Then

$$C(\lambda P_1 + \bar{\lambda} P_2) \ge I(X^{(\lambda)};Y^{(\lambda)}) \ge \lambda C(P_1) + \bar{\lambda} C(P_2) - \epsilon.$$

6. Letting $\epsilon \to 0$, we have

$$C(\lambda P_1 + \bar{\lambda} P_2) \ge \lambda C(P_1) + \bar{\lambda} C(P_2), \tag{2}$$

proving that $C(P)$ is concave.

Theorem 11.8

$$C(P) = \sup_{F(x): E\kappa(X) \le P} I(X;Y) \quad (1)$$

is non-decreasing, concave, and left-continuous.

Proof

<u>$C(P)$ is non-decreasing</u>

Note that in (1), the supremum is taken over a larger set for a larger P.

<u>$C(P)$ is concave</u>

1. Let $j = 1, 2$. For an input distribution $F_j(x)$, denote the corresponding input and output random variables by X_j and Y_j, respectively.

2. Then for any P_j, for all $\epsilon > 0$, there exists $F_j(x)$ such that

$$E\kappa(X_j) \le P_j$$

and

$$I(X_j; Y_j) \ge C(P_j) - \epsilon.$$

3. For $0 \le \lambda \le 1$, let $\bar{\lambda} = 1 - \lambda$ and define the random variable

$$X^{(\lambda)} \sim \lambda F_1(x) + \bar{\lambda} F_2(x).$$

Then

$$E\kappa(X^{(\lambda)}) = \lambda E\kappa(X_1) + \bar{\lambda} E\kappa(X_2) \le \lambda P_1 + \bar{\lambda} P_2.$$

4. By the concavity of mutual information with respect to the input distribution, we have

$$\begin{aligned} I(X^{(\lambda)}; Y^{(\lambda)}) &\ge \lambda I(X_1; Y_1) + \bar{\lambda} I(X_2; Y_2) \\ &\ge \lambda(C(P_1) - \epsilon) + \bar{\lambda}(C(P_2) - \epsilon) \\ &= \lambda C(P_1) + \bar{\lambda} C(P_2) - \epsilon. \end{aligned}$$

5. Then

$$C(\lambda P_1 + \bar{\lambda} P_2) \ge I(X^{(\lambda)}; Y^{(\lambda)}) \ge \lambda C(P_1) + \bar{\lambda} C(P_2) - \epsilon.$$

6. Letting $\epsilon \to 0$, we have

$$C(\lambda P_1 + \bar{\lambda} P_2) \ge \lambda C(P_1) + \bar{\lambda} C(P_2), \quad (2)$$

proving that $C(P)$ is concave.

<u>$C(P)$ is left-continuous</u>

Theorem 11.8

$$C(P) = \sup_{F(x): E\kappa(X) \le P} I(X;Y) \qquad (1)$$

is non-decreasing, concave, and left-continuous.

Proof

<u>$C(P)$ is non-decreasing</u>

Note that in (1), the supremum is taken over a larger set for a larger P.

<u>$C(P)$ is concave</u>

1. Let $j = 1, 2$. For an input distribution $F_j(x)$, denote the corresponding input and output random variables by X_j and Y_j, respectively.

2. Then for any P_j, for all $\epsilon > 0$, there exists $F_j(x)$ such that

$$E\kappa(X_j) \le P_j$$

and

$$I(X_j; Y_j) \ge C(P_j) - \epsilon.$$

3. For $0 \le \lambda \le 1$, let $\bar{\lambda} = 1 - \lambda$ and define the random variable

$$X^{(\lambda)} \sim \lambda F_1(x) + \bar{\lambda} F_2(x).$$

Then

$$E\kappa(X^{(\lambda)}) = \lambda E\kappa(X_1) + \bar{\lambda} E\kappa(X_2) \le \lambda P_1 + \bar{\lambda} P_2.$$

4. By the concavity of mutual information with respect to the input distribution, we have

$$\begin{aligned} I(X^{(\lambda)}; Y^{(\lambda)}) &\ge \lambda I(X_1; Y_1) + \bar{\lambda} I(X_2; Y_2) \\ &\ge \lambda (C(P_1) - \epsilon) + \bar{\lambda}(C(P_2) - \epsilon) \\ &= \lambda C(P_1) + \bar{\lambda} C(P_2) - \epsilon. \end{aligned}$$

5. Then

$$C(\lambda P_1 + \bar{\lambda} P_2) \ge I(X^{(\lambda)}; Y^{(\lambda)}) \ge \lambda C(P_1) + \bar{\lambda} C(P_2) - \epsilon.$$

6. Letting $\epsilon \to 0$, we have

$$C(\lambda P_1 + \bar{\lambda} P_2) \ge \lambda C(P_1) + \bar{\lambda} C(P_2), \qquad (2)$$

proving that $C(P)$ is concave.

<u>$C(P)$ is left-continuous</u>

1. Let $P_1 < P_2$, so that $P_2 \ge \lambda P_1 + \bar{\lambda} P_2$. Since $C(P)$ is non-decreasing, we have

Theorem 11.8

$$C(P) = \sup_{F(x): E\kappa(X) \leq P} I(X;Y) \quad (1)$$

is non-decreasing, concave, and left-continuous.

Proof

<u>$C(P)$ is non-decreasing</u>

Note that in (1), the supremum is taken over a larger set for a larger P.

<u>$C(P)$ is concave</u>

1. Let $j = 1, 2$. For an input distribution $F_j(x)$, denote the corresponding input and output random variables by X_j and Y_j, respectively.

2. Then for any P_j, for all $\epsilon > 0$, there exists $F_j(x)$ such that

$$E\kappa(X_j) \leq P_j$$

and

$$I(X_j; Y_j) \geq C(P_j) - \epsilon.$$

3. For $0 \leq \lambda \leq 1$, let $\bar{\lambda} = 1 - \lambda$ and define the random variable

$$X^{(\lambda)} \sim \lambda F_1(x) + \bar{\lambda} F_2(x).$$

Then

$$E\kappa(X^{(\lambda)}) = \lambda E\kappa(X_1) + \bar{\lambda} E\kappa(X_2) \leq \lambda P_1 + \bar{\lambda} P_2.$$

4. By the concavity of mutual information with respect to the input distribution, we have

$$\begin{aligned} I(X^{(\lambda)}; Y^{(\lambda)}) &\geq \lambda I(X_1; Y_1) + \bar{\lambda} I(X_2; Y_2) \\ &\geq \lambda(C(P_1) - \epsilon) + \bar{\lambda}(C(P_2) - \epsilon) \\ &= \lambda C(P_1) + \bar{\lambda} C(P_2) - \epsilon. \end{aligned}$$

5. Then

$$C(\lambda P_1 + \bar{\lambda} P_2) \geq I(X^{(\lambda)}; Y^{(\lambda)}) \geq \lambda C(P_1) + \bar{\lambda} C(P_2) - \epsilon.$$

6. Letting $\epsilon \to 0$, we have

$$C(\lambda P_1 + \bar{\lambda} P_2) \geq \lambda C(P_1) + \bar{\lambda} C(P_2), \quad (2)$$

proving that $C(P)$ is concave.

<u>$C(P)$ is left-continuous</u>

1. Let $P_1 < P_2$, so that $P_2 \geq \lambda P_1 + \bar{\lambda} P_2$. Since $C(P)$ is non-decreasing, we have

$$C(P_2) \geq C(\lambda P_1 + \bar{\lambda} P_2)$$

Theorem 11.8

$$C(P) = \sup_{F(x): E\kappa(X) \le P} I(X;Y) \qquad (1)$$

is non-decreasing, concave, and left-continuous.

Proof

$C(P)$ is non-decreasing

Note that in (1), the supremum is taken over a larger set for a larger P.

$C(P)$ is concave

1. Let $j = 1, 2$. For an input distribution $F_j(x)$, denote the corresponding input and output random variables by X_j and Y_j, respectively.

2. Then for any P_j, for all $\epsilon > 0$, there exists $F_j(x)$ such that

$$E\kappa(X_j) \le P_j$$

and

$$I(X_j;Y_j) \ge C(P_j) - \epsilon.$$

3. For $0 \le \lambda \le 1$, let $\bar{\lambda} = 1 - \lambda$ and define the random variable

$$X^{(\lambda)} \sim \lambda F_1(x) + \bar{\lambda} F_2(x).$$

Then

$$E\kappa(X^{(\lambda)}) = \lambda E\kappa(X_1) + \bar{\lambda} E\kappa(X_2) \le \lambda P_1 + \bar{\lambda} P_2.$$

4. By the concavity of mutual information with respect to the input distribution, we have

$$\begin{aligned} I(X^{(\lambda)};Y^{(\lambda)}) &\ge \lambda I(X_1;Y_1) + \bar{\lambda} I(X_2;Y_2) \\ &\ge \lambda(C(P_1) - \epsilon) + \bar{\lambda}(C(P_2) - \epsilon) \\ &= \lambda C(P_1) + \bar{\lambda} C(P_2) - \epsilon. \end{aligned}$$

5. Then

$$C(\lambda P_1 + \bar{\lambda} P_2) \ge I(X^{(\lambda)};Y^{(\lambda)}) \ge \lambda C(P_1) + \bar{\lambda} C(P_2) - \epsilon.$$

6. Letting $\epsilon \to 0$, we have

$$C(\lambda P_1 + \bar{\lambda} P_2) \ge \lambda C(P_1) + \bar{\lambda} C(P_2), \qquad (2)$$

proving that $C(P)$ is concave.

$C(P)$ is left-continuous

1. Let $P_1 < P_2$, so that $P_2 \ge \lambda P_1 + \bar{\lambda} P_2$. Since $C(P)$ is non-decreasing, we have

$$C(P_2) \ge C(\lambda P_1 + \bar{\lambda} P_2) \ge \lambda C(P_1) + \bar{\lambda} C(P_2).$$

Theorem 11.8

$$C(P) = \sup_{F(x): E\kappa(X) \le P} I(X;Y) \tag{1}$$

is non-decreasing, concave, and left-continuous.

Proof

<u>$C(P)$ is non-decreasing</u>

Note that in (1), the supremum is taken over a larger set for a larger P.

<u>$C(P)$ is concave</u>

1. Let $j = 1, 2$. For an input distribution $F_j(x)$, denote the corresponding input and output random variables by X_j and Y_j, respectively.

2. Then for any P_j, for all $\epsilon > 0$, there exists $F_j(x)$ such that

$$E\kappa(X_j) \le P_j$$

and

$$I(X_j; Y_j) \ge C(P_j) - \epsilon.$$

3. For $0 \le \lambda \le 1$, let $\bar{\lambda} = 1 - \lambda$ and define the random variable

$$X^{(\lambda)} \sim \lambda F_1(x) + \bar{\lambda} F_2(x).$$

Then

$$E\kappa(X^{(\lambda)}) = \lambda E\kappa(X_1) + \bar{\lambda} E\kappa(X_2) \le \lambda P_1 + \bar{\lambda} P_2.$$

4. By the concavity of mutual information with respect to the input distribution, we have

$$\begin{aligned} I(X^{(\lambda)}; Y^{(\lambda)}) &\ge \lambda I(X_1; Y_1) + \bar{\lambda} I(X_2; Y_2) \\ &\ge \lambda (C(P_1) - \epsilon) + \bar{\lambda}(C(P_2) - \epsilon) \\ &= \lambda C(P_1) + \bar{\lambda} C(P_2) - \epsilon. \end{aligned}$$

5. Then

$$C(\lambda P_1 + \bar{\lambda} P_2) \ge I(X^{(\lambda)}; Y^{(\lambda)}) \ge \lambda C(P_1) + \bar{\lambda} C(P_2) - \epsilon.$$

6. Letting $\epsilon \to 0$, we have

$$C(\lambda P_1 + \bar{\lambda} P_2) \ge \lambda C(P_1) + \bar{\lambda} C(P_2), \tag{2}$$

proving that $C(P)$ is concave.

<u>$C(P)$ is left-continuous</u>

1. Let $P_1 < P_2$, so that $P_2 \ge \lambda P_1 + \bar{\lambda} P_2$. Since $C(P)$ is non-decreasing, we have

$$C(P_2) \ge C(\lambda P_1 + \bar{\lambda} P_2) \ge \lambda C(P_1) + \bar{\lambda} C(P_2).$$

2. Letting $\lambda \to 0$, we have

Theorem 11.8

$$C(P) = \sup_{F(x): E\kappa(X) \le P} I(X;Y) \quad (1)$$

is non-decreasing, concave, and left-continuous.

Proof

<u>$C(P)$ is non-decreasing</u>

Note that in (1), the supremum is taken over a larger set for a larger P.

<u>$C(P)$ is concave</u>

1. Let $j = 1, 2$. For an input distribution $F_j(x)$, denote the corresponding input and output random variables by X_j and Y_j, respectively.

2. Then for any P_j, for all $\epsilon > 0$, there exists $F_j(x)$ such that

$$E\kappa(X_j) \le P_j$$

and

$$I(X_j; Y_j) \ge C(P_j) - \epsilon.$$

3. For $0 \le \lambda \le 1$, let $\bar{\lambda} = 1 - \lambda$ and define the random variable

$$X^{(\lambda)} \sim \lambda F_1(x) + \bar{\lambda} F_2(x).$$

Then

$$E\kappa(X^{(\lambda)}) = \lambda E\kappa(X_1) + \bar{\lambda} E\kappa(X_2) \le \lambda P_1 + \bar{\lambda} P_2.$$

4. By the concavity of mutual information with respect to the input distribution, we have

$$\begin{aligned} I(X^{(\lambda)}; Y^{(\lambda)}) &\ge \lambda I(X_1;Y_1) + \bar{\lambda} I(X_2;Y_2) \\ &\ge \lambda(C(P_1) - \epsilon) + \bar{\lambda}(C(P_2) - \epsilon) \\ &= \lambda C(P_1) + \bar{\lambda} C(P_2) - \epsilon. \end{aligned}$$

5. Then

$$C(\lambda P_1 + \bar{\lambda} P_2) \ge I(X^{(\lambda)}; Y^{(\lambda)}) \ge \lambda C(P_1) + \bar{\lambda} C(P_2) - \epsilon.$$

6. Letting $\epsilon \to 0$, we have

$$C(\lambda P_1 + \bar{\lambda} P_2) \ge \lambda C(P_1) + \bar{\lambda} C(P_2), \quad (2)$$

proving that $C(P)$ is concave.

<u>$C(P)$ is left-continuous</u>

1. Let $P_1 < P_2$, so that $P_2 \ge \lambda P_1 + \bar{\lambda} P_2$. Since $C(P)$ is non-decreasing, we have

$$C(P_2) \ge C(\lambda P_1 + \bar{\lambda} P_2) \ge \lambda C(P_1) + \bar{\lambda} C(P_2).$$

2. Letting $\lambda \to 0$, we have

$$C(P_2) \ge \lim_{\lambda \to 0} C(\lambda P_1 + \bar{\lambda} P_2) \ge C(P_2),$$

Theorem 11.8

$$C(P) = \sup_{F(x):E\kappa(X)\le P} I(X;Y) \quad (1)$$

is non-decreasing, concave, and left-continuous.

Proof

<u>$C(P)$ is non-decreasing</u>

Note that in (1), the supremum is taken over a larger set for a larger P.

<u>$C(P)$ is concave</u>

1. Let $j = 1, 2$. For an input distribution $F_j(x)$, denote the corresponding input and output random variables by X_j and Y_j, respectively.

2. Then for any P_j, for all $\epsilon > 0$, there exists $F_j(x)$ such that

$$E\kappa(X_j) \le P_j$$

and

$$I(X_j;Y_j) \ge C(P_j) - \epsilon.$$

3. For $0 \le \lambda \le 1$, let $\bar{\lambda} = 1 - \lambda$ and define the random variable

$$X^{(\lambda)} \sim \lambda F_1(x) + \bar{\lambda} F_2(x).$$

Then

$$E\kappa(X^{(\lambda)}) = \lambda E\kappa(X_1) + \bar{\lambda} E\kappa(X_2) \le \lambda P_1 + \bar{\lambda} P_2.$$

4. By the concavity of mutual information with respect to the input distribution, we have

$$\begin{aligned} I(X^{(\lambda)};Y^{(\lambda)}) &\ge \lambda I(X_1;Y_1) + \bar{\lambda} I(X_2;Y_2) \\ &\ge \lambda (C(P_1) - \epsilon) + \bar{\lambda}(C(P_2) - \epsilon) \\ &= \lambda C(P_1) + \bar{\lambda} C(P_2) - \epsilon. \end{aligned}$$

5. Then

$$C(\lambda P_1 + \bar{\lambda} P_2) \ge I(X^{(\lambda)};Y^{(\lambda)}) \ge \lambda C(P_1) + \bar{\lambda} C(P_2) - \epsilon.$$

6. Letting $\epsilon \to 0$, we have

$$C(\lambda P_1 + \bar{\lambda} P_2) \ge \lambda C(P_1) + \bar{\lambda} C(P_2), \quad (2)$$

proving that $C(P)$ is concave.

<u>$C(P)$ is left-continuous</u>

1. Let $P_1 < P_2$, so that $P_2 \ge \lambda P_1 + \bar{\lambda} P_2$. Since $C(P)$ is non-decreasing, we have

$$C(P_2) \ge C(\lambda P_1 + \bar{\lambda} P_2) \ge \underline{\lambda C(P_1) + \bar{\lambda} C(P_2)}.$$

2. Letting $\lambda \to 0$, we have

$$C(P_2) \ge \lim_{\lambda \to 0} C(\lambda P_1 + \bar{\lambda} P_2) \ge C(P_2),$$

Theorem 11.8

$$C(P) = \sup_{F(x):E\kappa(X)\leq P} I(X;Y) \qquad (1)$$

is non-decreasing, concave, and left-continuous.

Proof

<u>$C(P)$ is non-decreasing</u>

Note that in (1), the supremum is taken over a larger set for a larger P.

<u>$C(P)$ is concave</u>

1. Let $j = 1, 2$. For an input distribution $F_j(x)$, denote the corresponding input and output random variables by X_j and Y_j, respectively.

2. Then for any P_j, for all $\epsilon > 0$, there exists $F_j(x)$ such that

$$E\kappa(X_j) \leq P_j$$

and

$$I(X_j;Y_j) \geq C(P_j) - \epsilon.$$

3. For $0 \leq \lambda \leq 1$, let $\bar{\lambda} = 1 - \lambda$ and define the random variable

$$X^{(\lambda)} \sim \lambda F_1(x) + \bar{\lambda} F_2(x).$$

Then

$$E\kappa(X^{(\lambda)}) = \lambda E\kappa(X_1) + \bar{\lambda} E\kappa(X_2) \leq \lambda P_1 + \bar{\lambda} P_2.$$

4. By the concavity of mutual information with respect to the input distribution, we have

$$\begin{aligned} I(X^{(\lambda)};Y^{(\lambda)}) &\geq \lambda I(X_1;Y_1) + \bar{\lambda} I(X_2;Y_2) \\ &\geq \lambda(C(P_1) - \epsilon) + \bar{\lambda}(C(P_2) - \epsilon) \\ &= \lambda C(P_1) + \bar{\lambda} C(P_2) - \epsilon. \end{aligned}$$

5. Then

$$C(\lambda P_1 + \bar{\lambda} P_2) \geq I(X^{(\lambda)};Y^{(\lambda)}) \geq \lambda C(P_1) + \bar{\lambda} C(P_2) - \epsilon.$$

6. Letting $\epsilon \to 0$, we have

$$C(\lambda P_1 + \bar{\lambda} P_2) \geq \lambda C(P_1) + \bar{\lambda} C(P_2), \qquad (2)$$

proving that $C(P)$ is concave.

<u>$C(P)$ is left-continuous</u>

1. Let $P_1 < P_2$, so that $P_2 \geq \lambda P_1 + \bar{\lambda} P_2$. Since $C(P)$ is non-decreasing, we have

$$C(P_2) \geq C(\lambda P_1 + \bar{\lambda} P_2) \geq \underline{\lambda C(P_1) + \bar{\lambda} C(P_2)}.$$

2. Letting $\lambda \to 0$, we have

$$C(P_2) \geq \lim_{\lambda \to 0} C(\lambda P_1 + \bar{\lambda} P_2) \geq \underline{C(P_2)},$$

Theorem 11.8

$$C(P) = \sup_{F(x): E\kappa(X) \le P} I(X;Y) \tag{1}$$

is non-decreasing, concave, and left-continuous.

Proof

<u>$C(P)$ is non-decreasing</u>

Note that in (1), the supremum is taken over a larger set for a larger P.

<u>$C(P)$ is concave</u>

1. Let $j = 1, 2$. For an input distribution $F_j(x)$, denote the corresponding input and output random variables by X_j and Y_j, respectively.

2. Then for any P_j, for all $\epsilon > 0$, there exists $F_j(x)$ such that

$$E\kappa(X_j) \le P_j$$

and

$$I(X_j; Y_j) \ge C(P_j) - \epsilon.$$

3. For $0 \le \lambda \le 1$, let $\bar{\lambda} = 1 - \lambda$ and define the random variable

$$X^{(\lambda)} \sim \lambda F_1(x) + \bar{\lambda} F_2(x).$$

Then

$$E\kappa(X^{(\lambda)}) = \lambda E\kappa(X_1) + \bar{\lambda} E\kappa(X_2) \le \lambda P_1 + \bar{\lambda} P_2.$$

4. By the concavity of mutual information with respect to the input distribution, we have

$$\begin{aligned} I(X^{(\lambda)}; Y^{(\lambda)}) &\ge \lambda I(X_1; Y_1) + \bar{\lambda} I(X_2; Y_2) \\ &\ge \lambda (C(P_1) - \epsilon) + \bar{\lambda}(C(P_2) - \epsilon) \\ &= \lambda C(P_1) + \bar{\lambda} C(P_2) - \epsilon. \end{aligned}$$

5. Then

$$C(\lambda P_1 + \bar{\lambda} P_2) \ge I(X^{(\lambda)}; Y^{(\lambda)}) \ge \lambda C(P_1) + \bar{\lambda} C(P_2) - \epsilon.$$

6. Letting $\epsilon \to 0$, we have

$$C(\lambda P_1 + \bar{\lambda} P_2) \ge \lambda C(P_1) + \bar{\lambda} C(P_2), \tag{2}$$

proving that $C(P)$ is concave.

<u>$C(P)$ is left-continuous</u>

1. Let $P_1 < P_2$, so that $P_2 \ge \lambda P_1 + \bar{\lambda} P_2$. Since $C(P)$ is non-decreasing, we have

$$C(P_2) \ge C(\lambda P_1 + \bar{\lambda} P_2) \ge \lambda C(P_1) + \bar{\lambda} C(P_2).$$

2. Letting $\lambda \to 0$, we have

$$C(P_2) \ge \lim_{\lambda \to 0} C(\lambda P_1 + \bar{\lambda} P_2) \ge C(P_2),$$

which implies

$$\lim_{\lambda \to 0} C(\lambda P_1 + \bar{\lambda} P_2) = C(P_2).$$

Theorem 11.8

$$C(P) = \sup_{F(x): E\kappa(X) \le P} I(X;Y) \quad (1)$$

is non-decreasing, concave, and left-continuous.

Proof

$C(P)$ is non-decreasing

Note that in (1), the supremum is taken over a larger set for a larger P.

$C(P)$ is concave

1. Let $j = 1, 2$. For an input distribution $F_j(x)$, denote the corresponding input and output random variables by X_j and Y_j, respectively.

2. Then for any P_j, for all $\epsilon > 0$, there exists $F_j(x)$ such that

$$E\kappa(X_j) \le P_j$$

and

$$I(X_j;Y_j) \ge C(P_j) - \epsilon.$$

3. For $0 \le \lambda \le 1$, let $\bar{\lambda} = 1 - \lambda$ and define the random variable

$$X^{(\lambda)} \sim \lambda F_1(x) + \bar{\lambda} F_2(x).$$

Then

$$E\kappa(X^{(\lambda)}) = \lambda E\kappa(X_1) + \bar{\lambda} E\kappa(X_2) \le \lambda P_1 + \bar{\lambda} P_2.$$

4. By the concavity of mutual information with respect to the input distribution, we have

$$\begin{aligned} I(X^{(\lambda)};Y^{(\lambda)}) &\ge \lambda I(X_1;Y_1) + \bar{\lambda} I(X_2;Y_2) \\ &\ge \lambda(C(P_1) - \epsilon) + \bar{\lambda}(C(P_2) - \epsilon) \\ &= \lambda C(P_1) + \bar{\lambda} C(P_2) - \epsilon. \end{aligned}$$

5. Then

$$C(\lambda P_1 + \bar{\lambda} P_2) \ge I(X^{(\lambda)};Y^{(\lambda)}) \ge \lambda C(P_1) + \bar{\lambda} C(P_2) - \epsilon.$$

6. Letting $\epsilon \to 0$, we have

$$C(\lambda P_1 + \bar{\lambda} P_2) \ge \lambda C(P_1) + \bar{\lambda} C(P_2), \quad (2)$$

proving that $C(P)$ is concave.

$C(P)$ is left-continuous

1. Let $P_1 < P_2$, so that $P_2 \ge \lambda P_1 + \bar{\lambda} P_2$. Since $C(P)$ is non-decreasing, we have

$$C(P_2) \ge C(\lambda P_1 + \bar{\lambda} P_2) \ge \lambda C(P_1) + \bar{\lambda} C(P_2).$$

2. Letting $\lambda \to 0$, we have

$$C(P_2) \ge \lim_{\lambda \to 0} C(\lambda P_1 + \bar{\lambda} P_2) \ge C(P_2),$$

which implies

$$\lim_{\lambda \to 0} C(\lambda P_1 + \bar{\lambda} P_2) = C(P_2).$$

3. Hence, we conclude that

$$\lim_{P \uparrow P_2} C(P) = C(P_2),$$

Theorem 11.8

$$C(P) = \sup_{F(x): E\kappa(X) \le P} I(X;Y) \tag{1}$$

is non-decreasing, concave, and left-continuous.

Proof

<u>$C(P)$ is non-decreasing</u>

Note that in (1), the supremum is taken over a larger set for a larger P.

<u>$C(P)$ is concave</u>

1. Let $j = 1, 2$. For an input distribution $F_j(x)$, denote the corresponding input and output random variables by X_j and Y_j, respectively.

2. Then for any P_j, for all $\epsilon > 0$, there exists $F_j(x)$ such that

$$E\kappa(X_j) \le P_j$$

and

$$I(X_j;Y_j) \ge C(P_j) - \epsilon.$$

3. For $0 \le \lambda \le 1$, let $\bar{\lambda} = 1 - \lambda$ and define the random variable

$$X^{(\lambda)} \sim \lambda F_1(x) + \bar{\lambda} F_2(x).$$

Then

$$E\kappa(X^{(\lambda)}) = \lambda E\kappa(X_1) + \bar{\lambda} E\kappa(X_2) \le \lambda P_1 + \bar{\lambda} P_2.$$

4. By the concavity of mutual information with respect to the input distribution, we have

$$\begin{aligned} I(X^{(\lambda)};Y^{(\lambda)}) &\ge \lambda I(X_1;Y_1) + \bar{\lambda} I(X_2;Y_2) \\ &\ge \lambda(C(P_1) - \epsilon) + \bar{\lambda}(C(P_2) - \epsilon) \\ &= \lambda C(P_1) + \bar{\lambda} C(P_2) - \epsilon. \end{aligned}$$

5. Then

$$C(\lambda P_1 + \bar{\lambda} P_2) \ge I(X^{(\lambda)};Y^{(\lambda)}) \ge \lambda C(P_1) + \bar{\lambda} C(P_2) - \epsilon.$$

6. Letting $\epsilon \to 0$, we have

$$C(\lambda P_1 + \bar{\lambda} P_2) \ge \lambda C(P_1) + \bar{\lambda} C(P_2), \tag{2}$$

proving that $C(P)$ is concave.

<u>$C(P)$ is left-continuous</u>

1. Let $P_1 < P_2$, so that $P_2 \ge \lambda P_1 + \bar{\lambda} P_2$. Since $C(P)$ is non-decreasing, we have

$$C(P_2) \ge C(\lambda P_1 + \bar{\lambda} P_2) \ge \lambda C(P_1) + \bar{\lambda} C(P_2).$$

2. Letting $\lambda \to 0$, we have

$$C(P_2) \ge \lim_{\lambda \to 0} C(\lambda P_1 + \bar{\lambda} P_2) \ge C(P_2),$$

which implies

$$\lim_{\lambda \to 0} C(\underbrace{\lambda P_1 + \bar{\lambda} P_2}_{P}) = C(P_2).$$

3. Hence, we conclude that

$$\lim_{P \uparrow P_2} C(P) = C(P_2),$$

Theorem 11.8

$$C(P) = \sup_{F(x): E\kappa(X) \leq P} I(X;Y) \tag{1}$$

is non-decreasing, concave, and left-continuous.

Proof

$C(P)$ is non-decreasing

Note that in (1), the supremum is taken over a larger set for a larger P.

$C(P)$ is concave

1. Let $j = 1, 2$. For an input distribution $F_j(x)$, denote the corresponding input and output random variables by X_j and Y_j, respectively.

2. Then for any P_j, for all $\epsilon > 0$, there exists $F_j(x)$ such that

$$E\kappa(X_j) \leq P_j$$

and

$$I(X_j; Y_j) \geq C(P_j) - \epsilon.$$

3. For $0 \leq \lambda \leq 1$, let $\bar{\lambda} = 1 - \lambda$ and define the random variable

$$X^{(\lambda)} \sim \lambda F_1(x) + \bar{\lambda} F_2(x).$$

Then

$$E\kappa(X^{(\lambda)}) = \lambda E\kappa(X_1) + \bar{\lambda} E\kappa(X_2) \leq \lambda P_1 + \bar{\lambda} P_2.$$

4. By the concavity of mutual information with respect to the input distribution, we have

$$\begin{aligned} I(X^{(\lambda)}; Y^{(\lambda)}) &\geq \lambda I(X_1; Y_1) + \bar{\lambda} I(X_2; Y_2) \\ &\geq \lambda (C(P_1) - \epsilon) + \bar{\lambda}(C(P_2) - \epsilon) \\ &= \lambda C(P_1) + \bar{\lambda} C(P_2) - \epsilon. \end{aligned}$$

5. Then

$$C(\lambda P_1 + \bar{\lambda} P_2) \geq I(X^{(\lambda)}; Y^{(\lambda)}) \geq \lambda C(P_1) + \bar{\lambda} C(P_2) - \epsilon.$$

6. Letting $\epsilon \to 0$, we have

$$C(\lambda P_1 + \bar{\lambda} P_2) \geq \lambda C(P_1) + \bar{\lambda} C(P_2), \tag{2}$$

proving that $C(P)$ is concave.

$C(P)$ is left-continuous

1. Let $P_1 < P_2$, so that $P_2 \geq \lambda P_1 + \bar{\lambda} P_2$. Since $C(P)$ is non-decreasing, we have

$$C(P_2) \geq C(\lambda P_1 + \bar{\lambda} P_2) \geq \lambda C(P_1) + \bar{\lambda} C(P_2).$$

2. Letting $\lambda \to 0$, we have

$$C(P_2) \geq \lim_{\lambda \to 0} C(\lambda P_1 + \bar{\lambda} P_2) \geq C(P_2),$$

which implies

$$\lim_{\lambda \to 0} C(\overbrace{\lambda P_1 + \bar{\lambda} P_2}^{P}) = C(P_2).$$

3. Hence, we conclude that

$$\lim_{P \uparrow P_2} C(P) = C(P_2),$$

i.e., $C(P)$ is left-continuous.