# Capacity of Steganographic Channels

Jeremiah J. Harmsen
Center for Image Processing Research
Electrical Computer and Systems Department
Rensselaer Polytechnic Institute, Troy, NY
harmsj@cipr.rpi.edu

William A. Pearlman
Center for Image Processing Research
Electrical Computer and Systems Department
Rensselaer Polytechnic Institute, Troy, NY
pearlman@cipr.rpi.edu

## ABSTRACT

An information-theoretic approach is used to determine the amount of information that may be safely transferred over a steganographic channel with a passive adversary. A steganographic channel, or stego-channel is a pair consisting of the channel transition probabilities and a detection function. When a message is sent it first encounters a distortion (due to the channel), then is subject to inspection by a passive adversary (using the detection function). This paper presents results on the amount of information that may be transferred over an arbitrary stego-channel with vanishing probabilities of error and detection.

## Keywords

Steganographic capacity, stego-channel, steganalysis, steganography, capacity, information theory, information spectrum

## 1. INTRODUCTION

### 1.1 Background

Shannon's pioneering work provides bounds on the amount of information that can be transmitted over a noisy channel. His results show that capacity is an intrinsic property of the channel itself. This work takes a similar viewpoint in seeking to find the amount of information that may be transferred over a stego-channel as seen in Figure 1.

The stego-channel is equivalent to the classic channel with the addition of the detection function. For the classic channel, a transmission is considered successful if the decoder properly determines which message the encoder has sent. In the stego-channel a transmission is successful not only if the decoder properly determines the sent message, but if the detection function is not triggered as well.

This additional constraint on the channel use leads to the fundamental view that *the capacity of a stego-channel is an intrinsic property of both the channel and the detection function.* That is, the properties of the detection function influence the capacity just as much as the noise in the channel.


*ACM Multimedia and Security Workshop* '05 New York, New York USA


### 1.2 Previous Work

There have been a number of applications of information theory to the steganographic capacity problem[10, 11]. These works give capacity results under distortion constraints on the hider as well as active adversary. The additional constraint that the stego-signal retain the same distribution as the cover-signal serves as the steganalysis detection function.

Somewhat less work exists exploring capacity with arbitrary detection functions. These works are written from a steganalysis perspective[1, 9] and accordingly give heavy consideration to the detection function.

This work differs from previous work in a number of aspects. Most notable is the use of information-spectrum methods that allow for the analysis of arbitrary detection algorithms. This eliminates the need to restrict interest to detection algorithms that operate on sample averages or behave consistently. Instead the detection functions may be instantaneous, that is, the properties of a detector for $n$ samples need not have any relation to the same detector for $n+1$ samples.

Another substantial difference is the presence of noise *before* the detector. This placement enables the modeling of common signal processing distortions such as compression, quantization, etc. The location of the noise adds complexity not only because of confusion at the decoder, but also a signal, carefully crafted to avoid detection, may be corrupted into one that will trigger the detector.

Finally, the consideration of a cover-signal and distortion constraint in the encoding function is omitted. This is due to the view that steganographic capacity is a property of the channel and the detection function. This viewpoint, along with the above differences, make a direct comparison to previous work somewhat difficult, although possible with a number of simplifications explored in Section 6.

## 2. PRELIMINARIES

### 2.1 Random Variables

Random variables are denoted by capital letters, e.g. $X$. Realizations of these random variables are denoted as lowercase letters, e.g. $x$. Each random variable is defined over a domain denoted with a script $\mathcal{X}$. A sequence of $n$ random variables is denoted with $X^n = (X_1, \ldots, X_n)$. Similarly, an $n$-length sequence of random variable realizations is denoted $\mathbf{x} = (x_1, \ldots, x_n) \in \mathcal{X}^n$. The probability of $X$ taking value $x \in \mathcal{X}$ is $P_X(x)$.

The space of all channel inputs (stego-signals) is denoted $\mathcal{X}$ and the space of channel outputs (corrupted stego-signals)

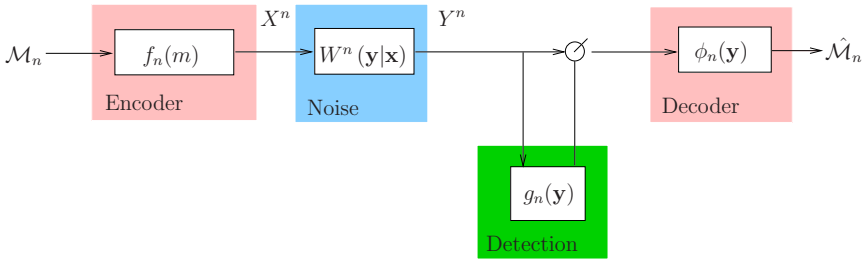


**Figure 1: Passive System with Noise**

is $\mathcal{Y}$. Similarly the space of $n$-length channel inputs and outputs are denoted as $\mathcal{X}^n$ and $\mathcal{Y}^n$ respectively.

## 2.2 Detection Function

*Definition 1.* The *steganalysis detection function* is a function $g_n : \mathcal{Y}^n \to \{0, 1\}$ that classifies a channel output into one of two categories: containing steganographic information, and not containing steganographic information.

The function is defined as follows for all $\mathbf{y} \in \mathcal{Y}^n$,

$$g_n(\mathbf{y}) = \begin{cases} 1, & \text{if } \mathbf{y} \text{ is steganographic} \\ 0, & \text{if } \mathbf{y} \text{ is not steganographic} \end{cases} \tag{1}$$

The specific type of function may be that of support vector machine or a Bayesian, etc. When no confusion can occur, $g$ is written in place of $g_n$.

*Definition 2.* A *detection function sequence* is denoted as,

$$\mathbf{g} := \{g_1, g_2, g_3, \ldots\}, \tag{2}$$

where $g_n : \mathcal{Y}^n \to \{0, 1\}$.

The set of all $n$-length detection functions is denoted $\mathcal{G}_n$.

## 2.3 Permissible Set

For any detection function $g_n$, the space of signals $\mathcal{Y}^n$ is split into the permissible set and the impermissible set, defined below.

*Definition 3.* The *permissible set*, $\mathcal{P}_{g_n} \subseteq \mathcal{Y}^n$, is the inverse image of 0 under $g_n$.

That is,

$$\mathcal{P}_{g_n} := g_n^{-1}(\{0\}) = \{\mathbf{y} \in \mathcal{Y}^n : g_n(\mathbf{y}) = 0\}. \tag{3}$$

The permissible set is the set of all channel outputs that the given detection function, $g_n$ will classify as non-steganographic.

As a detection function is completely specified by its permissible set, we may describe a detection function sequence as,

$$\mathbf{g} = \{\mathcal{P}_1, \mathcal{P}_2, \mathcal{P}_3, \ldots\},$$

where $\mathcal{P}_n \subseteq \mathcal{Y}^n$ is the permissible set for $g_n$.

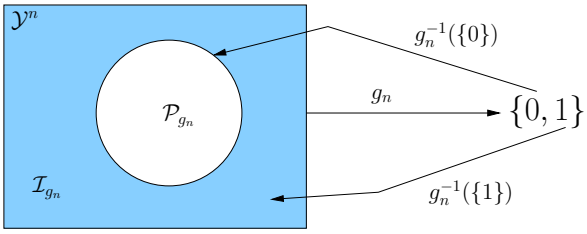


**Figure 2: Permissible and Impermissible Sets**

## 2.4 Impermissible Set

*Definition 4.* The *impermissible set*, $\mathcal{I}_{g_n} \subseteq \mathcal{Y}^n$, is the inverse image of 1 under $g_n$.

That is,

$$\mathcal{I}_{g_n} := g_n^{-1}(\{1\}) = \{\mathbf{y} \in \mathcal{Y}^n : g_n(\mathbf{y}) = 1\}. \tag{4}$$

For a given $g_n$ the impermissible set is the set of all channel outputs that $g_n$ will classify as steganographic.

Note that two sets are related as follows,

$$\mathcal{I}_g = \mathcal{Y}^n \setminus \mathcal{P}_g := \mathcal{P}_g^c. \tag{5}$$

The permissible set and impermissible set are shown in Figure 2.

## 2.5 Memoryless Detection Functions

*Definition 5.* A *memoryless detection function*, $\mathbf{g} = \{g_n\}_{n=1}^{\infty}$ is one where each $g_n$ is defined for $\mathbf{y} = (y_1, y_2, \ldots, y_n)$ as,

$$g_n(\mathbf{y}) = \begin{cases} 1, & \text{if } \exists i \in \{1, 2, \ldots, n\} \text{ such that } g(y_i) = 1 \\ 0, & \text{if } g(y_i) = 0 \ \forall\ i \in \{1, 2, \ldots, n\} \end{cases} \tag{6}$$

where $g \in \mathcal{G}_1$ is said to specify $g_n$ (and $\mathbf{g}$).

To denote a detection function sequence is memoryless the following notation will be used $\mathbf{g} = \{g\}$.

For a memoryless detection function $\mathbf{g} = \{g\}$, we have that,

$$\mathcal{P}_{g_n} = \prod_{i=1}^{n} \mathcal{P}_g. \tag{7}$$

That is, the permissible set of $g_n$ is defined by the $n$-dimensional product of $\mathcal{P}_g$.

## 2.6 Channels

The *channel* is denoted as $W^n$ where $W^n : \mathcal{Y}^n \times \mathcal{X}^n \to [0,1]$ and has the following property for all $\mathbf{x} \in \mathcal{X}^n$,

$$W^n\left(\mathcal{Y}^n|\mathbf{x}\right) := \sum_{\mathbf{y}\in\mathcal{Y}^n} W^n\left(\mathbf{y}|\mathbf{x}\right) = 1.$$

The channel represents the conditional probabilities of receiving $\mathbf{y} \in \mathcal{Y}^n$ when $\mathbf{x} \in \mathcal{X}^n$ is sent, that is $W^n\left(\mathbf{y}|\mathbf{x}\right) := P_{Y^n|X^n}(\mathbf{y}|\mathbf{x})$.

The random variable, $Y$ resulting from transmitting $X$ through the channel $W$ will be denoted as $Y = W(X)$ or the relation reinforced by noting $X \xrightarrow{W} Y$.

*Definition 6.* A *general channel* is a sequence,

$$\mathbf{W} := \{W^1, W^2, W^3, \ldots\},$$

where each $W^n$ is the $n$-length channel transition probability.

In the case where channel distortions act independently and identically on each input letter $x_i$, we say it is a *memoryless channel*. In this instance the $n$-length transition probabilities can be written as,

$$W^n\left(\mathbf{y}|\mathbf{x}\right) = \prod_{i=1}^{n} W(y_i|x_i), \tag{8}$$

where $W$ is said to define the channel. To denote a channel is memoryless and defined by $W$ we will write $\mathbf{W} = \{W\}$.

## 2.7 Encoder and Decoder

The purpose of the encoder and decoder is to transmit and receive information across a channel. The information to be transferred is assumed to be from a uniformly distributed message set denoted $\mathcal{M}_n$, with a cardinality of $M_n$.

The *encoding function* embeds a message into a stego-signal. That is, $f_n : \mathcal{M}_n \to \mathcal{X}^n$. The element of $\mathcal{X}^n$ that the $i$th message maps to is called the *codeword* for $i$ and is denoted, $\mathbf{u}_i$. That is,

$$f_n(i) = \mathbf{u}_i, \qquad i \in \{1, \ldots, M_n\}.$$

The collection of codewords, $\mathcal{C}_n = \{\mathbf{u}_1, \ldots, \mathbf{u}_{M_n}\}$ is called the *code*. The *rate* of an encoding function is given as,

$$R_n := \frac{1}{n}\log M_n.$$

The *decoding function*, $\phi_n : \mathcal{Y}^n \to \mathcal{M}_n$, maps a corrupted stego-signal to a message. The decoder is defined by a set of *decoding regions*: $\mathcal{D}_1, \ldots, \mathcal{D}_{M_n}$. The decoding regions are disjoint sets that cover $\mathcal{Y}^n$ and are defined such that,

$$\begin{aligned}\phi_n^{-1}(\{m\}) &= \mathcal{D}_m \\ &:= \left\{F \subseteq \mathcal{Y}^n : \phi_n(\mathbf{y}) = m,\ \forall\ \mathbf{y} \in F\right\},\end{aligned}$$

for $m = 1, \ldots, M_n$.

Next, two important terms are presented that allow for the analysis of steganographic systems. The first is the probability of error and the second is the probability of detection. In both cases they are calculated for a given code $\mathcal{C} = \{\mathbf{u}_1, \ldots, \mathbf{u}_{M_n}\}$, channel $W^n$, and impermissible set $\mathcal{I}_{g_n}$ (corresponding to some $g_n$).

The *probability of error* can be found as,

$$\epsilon_n = \frac{1}{M_n}\sum_{i=1}^{M_n} W^n\left(\mathcal{D}_i^c|\mathbf{u}_i\right). \tag{9}$$

Similarly the *probability of detection* is calculated as,

$$\delta_n = \frac{1}{M_n}\sum_{i=1}^{M_n} W^n\left(\mathcal{I}_{g_n}|\mathbf{u}_i\right). \tag{10}$$

## 2.8 Stego-Channel

*Definition 7.* A *steganographic channel* or *stego-channel* is a pair $(\mathbf{W}, \mathbf{g})$, where $\mathbf{W}$ is a general channel and $\mathbf{g}$ is a detection function sequence.

To reinforce the notion that a stego-channel is defined by a sequence of pairs we will typically write $(\mathbf{W}, \mathbf{g}) = \{(W^n, g_n)\}_{n=1}^{\infty}$.

*Definition 8.* A *discrete stego-channel* is one where at least one of the following holds: $|\mathcal{X}| < \infty$, $|\mathcal{Y}| < \infty$, or $|\mathcal{P}_{g_n}| < \infty\ \forall n$.

*Definition 9.* A *discrete memoryless stego-channel* (DMSC) is a stego-channel where,

1. $(\mathbf{W}, \mathbf{g})$ is discrete
2. $\mathbf{W}$ is memoryless
3. $\mathbf{g}$ is memoryless

A DMSC is said to be defined by the pair $(W, g)$ and will be denoted $(\mathbf{W}, \mathbf{g}) = \{(W, g)\}$.

## 2.9 Steganographic Capacity

*Definition 10.* An $(n, M_n, \epsilon_n, \delta_n)$-*code* (for a given stego-channel) consists of an encoder and decoder. The encoder and decoder are capable of transferring one of $M_n$ messages in $n$ uses of the channel with an average probability of error of $\epsilon_n$ and a probability of detection of $\delta_n$.

*Definition 11.* A rate $R$ is said to be *securely achievable* for a stego-channel $(\mathbf{W}, \mathbf{g}) = \{(W^n, g_n)\}_{n=1}^{\infty}$, if there exists a sequence of $(n, M_n, \epsilon_n, \delta_n)$-codes such that:

1. $\lim_{n\to\infty} \epsilon_n = 0$
2. $\liminf_{n\to\infty} \frac{1}{n}\log M_n \geq R$
3. $\lim_{n\to\infty} \delta_n = 0$

*Definition 12.* The *secure capacity* of a stego-channel $(\mathbf{W}, \mathbf{g})$ is denoted as $C(\mathbf{W}, \mathbf{g})$. This is defined as the supremum of all securely achievable rates for $(\mathbf{W}, \mathbf{g})$.

*Definition 13.* A rate $R$ is said to be $(\epsilon, \delta)$-*securely achievable* for a stego-channel $(\mathbf{W}, \mathbf{g}) = \{(W^n, g_n)\}_{n=1}^{\infty}$, if there exists a sequence of $(n, M_n, \epsilon_n, \delta_n)$-codes such that:

1. $\limsup_{n\to\infty} \epsilon_n \leq \epsilon$
2. $\liminf_{n\to\infty} \frac{1}{n}\log M_n \geq R$
3. $\limsup_{n\to\infty} \delta_n \leq \delta$

*Definition 14.* The $(\epsilon, \delta)$ *secure capacity* of a stego-channel $(\mathbf{W}, \mathbf{g})$ is denoted as $C(\epsilon, \delta|\mathbf{W}, \mathbf{g})$. This is defined as the supremum of all $(\epsilon, \delta)$-securely achievable rates for $(\mathbf{W}, \mathbf{g})$.

# 3. SECURE CAPACITY FORMULA

## 3.1 Information-Spectrum Methods

The information-spectrum method[4, 5, 6, 7, 13] is a generalization of information theory created to apply to systems where either the channel or its inputs are not necessarily ergodic or stationary. Its use is required in this work because the detection function is not assumed to have any ergodic or stationary properties.

The information-spectrum method uses the *general source* (or *general sequence*) defined as,

$$\mathbf{X} := \left\{ X^n = (X_1^{(n)}, X_2^{(n)}, \ldots, X_n^{(n)}) \right\}_{n=1}^{\infty}, \tag{11}$$

where each $X_m^{(n)}$ is a random variable defined over alphabet $\mathcal{X}$. It is important to note that the general source makes no assumptions about consistency, ergodicity, or stationarity.

The information-spectrum method also uses two novel quantities defined for sequences of random variables, called the lim sup and lim inf in probability.

The *limsup in probability* of a sequence of random variables, $\{Z_n\}_{n=1}^{\infty}$ is defined as,

$$\text{p-}\limsup Z_n := \inf \left\{ \alpha : \lim_{n\to\infty} \Pr\{Z_n > \alpha\} = 0 \right\}.$$

Similarly, the *liminf in probability* of a sequence of random variables, $\{Z_n\}_{n=1}^{\infty}$ is,

$$\text{p-}\liminf Z_n := \sup \left\{ \beta : \lim_{n\to\infty} \Pr\{Z_n < \beta\} = 0 \right\}.$$

The *spectral sup-entropy rate* of a general source $\mathbf{X} = \{X^n\}_{n=1}^{\infty}$ is defined as,

$$\overline{H}(\mathbf{X}) := \text{p-}\limsup_{n\to\infty} \frac{1}{n} \log \frac{1}{P_{X^n}(X^n)}. \tag{12}$$

Analogously, the *spectral inf-entropy rate* of a general source $\mathbf{X} = \{X^n\}_{n=1}^{\infty}$ is defined as,

$$\underline{H}(\mathbf{X}) := \text{p-}\liminf_{n\to\infty} \frac{1}{n} \log \frac{1}{P_{X^n}(X^n)}. \tag{13}$$

The spectral entropy rates are the information-spectrum counterparts to the Shannon entropy and share a number of its natural properties. For instance, $\overline{H}(\mathbf{X}) \geq \underline{H}(\mathbf{X}) \geq 0$ for any $\mathbf{X}$ [6, 13].

The *spectral sup-mutual information rate* for the pair of general sequences $(\mathbf{X}, \mathbf{Y}) = \{(X^n, Y^n)\}_{n=1}^{\infty}$ is defined as,

$$\overline{I}(\mathbf{X}; \mathbf{Y}) := \text{p-}\limsup_{n\to\infty} \frac{1}{n} i(X^n; Y^n), \tag{14}$$

where,

$$i(X^n; Y^n) := \log \frac{W^n(Y^n|X^n)}{P_{Y^n}(Y^n)}. \tag{15}$$

Likewise the *spectral inf-mutual information rate* for the pair of general sequences $(\mathbf{X}, \mathbf{Y}) = \{(X^n, Y^n)\}_{n=1}^{\infty}$ is defined as,

$$\underline{I}(\mathbf{X}; \mathbf{Y}) := \text{p-}\liminf_{n\to\infty} \frac{1}{n} i(X^n; Y^n). \tag{16}$$

The spectral mutual information rates share a number of properties with the classic mutual information such as $\underline{I}(\mathbf{X}; \mathbf{Y}) = \underline{I}(\mathbf{Y}; \mathbf{X}) \geq 0$ [6, 13]. A number of useful inequalities for the spectral mutual information are listed in Appendix A.

A rate $R$ (for a general channel $\mathbf{W}$) is said to be $\epsilon$-achievable if there exists a sequence of $(n, M_n, \epsilon_n)$-codes such that:

1. $\limsup_{n\to\infty} \epsilon_n \leq \epsilon$
2. $\liminf_{n\to\infty} \frac{1}{n} \log M_n \geq R$

The $\epsilon$-capacity of a general channel $\mathbf{W}$, denoted $C(\epsilon|\mathbf{W})$, is the supremum of $\epsilon$-achievable rates.

## 3.2 Secure distributions

*Definition 15.* An output distribution $\mathbf{Y} = \{Y^n\}_{n=1}^{\infty}$ is called a *$\delta$-secure output* if for a given detection function sequence $\mathbf{g} = \{g_n\}_{n=1}^{\infty}$,

$$\limsup_{n\to\infty} \Pr\{g_n(Y^n) = 1\} \leq \delta, \tag{17}$$

or either of the equivalent conditions,

$$\limsup_{n\to\infty} P_{Y^n}(\mathcal{I}_{g_n}) \leq \delta, \tag{18}$$

$$\liminf_{n\to\infty} P_{Y^n}(\mathcal{P}_{g_n}) \geq 1 - \delta. \tag{19}$$

The set of all general output sequences that are $\delta$-secure outputs is denoted $\mathcal{T}_\delta$, that is,

$$\mathcal{T}_\delta := \left\{ \mathbf{Y} = \{Y^n\}_{n=1}^{\infty} : \limsup_{n\to\infty} P_{Y^n}(\mathcal{I}_{g_n}) \leq \delta \right\}. \tag{20}$$

*Definition 16.* The set of all $\delta$-secure outputs for $\delta = 0$ is called the *secure output set* and denoted $\mathcal{T}_0$.

*Definition 17.* A general source $\mathbf{X} = \{X^n\}_{n=1}^{\infty}$ is called a *$\delta$-secure input* (for a given stego-channel) if the resulting output sequence, $\mathbf{Y}$, is a $\delta$-secure output, i.e. $\mathbf{W}(\mathbf{X}) = \mathbf{Y} \in \mathcal{T}_\delta$.

The set of all general sources that are $\delta$-secure inputs is denoted $\mathcal{S}_\delta$, that is,

$$\mathcal{S}_\delta := \left\{ \mathbf{X} = \{X^n\}_{n=1}^{\infty} : \limsup_{n\to\infty} \sum_{\mathbf{x}\in\mathcal{X}^n} W^n(\mathcal{I}_{g_n}|\mathbf{x}) P_{X^n}(\mathbf{x}) \leq \delta \right\}. \tag{21}$$

*Definition 18.* The set of all $\delta$-secure inputs for $\delta = 0$ is called the *secure input set* and denoted $\mathcal{S}_0$.

## 3.3 $(\epsilon, \delta)$-Channel Capacity

We are now prepared to derive the first fundamental result: the $(\epsilon, \delta)$-Channel Capacity. This capacity will make use of the following definition,

$$\begin{aligned} J(R|\mathbf{X}) &:= \limsup_{n\to\infty} \Pr\left\{ \frac{1}{n} i(X^n; Y^n) \leq R \right\} \\ &= \limsup_{n\to\infty} \Pr\left\{ \frac{1}{n} \log \frac{W^n(Y^n|X^n)}{P_{Y^n}(Y^n)} \leq R \right\}. \end{aligned}$$

THEOREM 1 ($(\epsilon, \delta)$-CHANNEL CAPACITY). *The $(\epsilon, \delta)$-channel capacity of a stego-channel $(\mathbf{W}, \mathbf{g})$ is given by,*

$$C(\epsilon, \delta|\mathbf{W}, \mathbf{g}) = \sup_{\mathbf{X}\in\mathcal{S}_\delta} \sup \{R : J(R|\mathbf{X}) \leq \epsilon\}, \tag{22}$$

*for any $0 \leq \epsilon < 1$ and $0 \leq \delta < 1$.*

PROOF. *From [6, 13] we have that for the general channel* $\mathbf{W}$ *the* $\epsilon$*-capacity is,*

$$C(\epsilon|\mathbf{W}) = \sup_{\mathbf{X}} \sup \left\{ R : J\left(R|\mathbf{X}\right) \leq \epsilon \right\}.$$

*In order for the probability of detection* $\delta_n \to 0$*, we must restrict the channel inputs to the* $\delta$*-secure input set, thus the* sup *is restricted to* $\mathbf{X} \in \mathcal{S}_\delta$*.* □

## 3.4 Secure Channel Capacity

The next result deals with a special case of $(\epsilon, \delta)$-capacity, namely the one where $\epsilon = \delta = 0$. The secure channel capacity is the maximum amount of information that may be sent over a channel with arbitrarily small probabilities of error and detection.

THEOREM 2 (SECURE CHANNEL CAPACITY). *The secure channel capacity of a stego-channel* $(\mathbf{W}, \mathbf{g})$ *is given by,*

$$C(\mathbf{W}, \mathbf{g}) = \sup_{\mathbf{X} \in \mathcal{S}_0} \underline{I}(\mathbf{X}; \mathbf{Y}). \tag{23}$$

PROOF. *We apply Theorem 1 with* $\epsilon = 0$ *and* $\delta = 0$*. This gives,*

$$\begin{aligned} C(\mathbf{W}, \mathbf{g}) &= C(0, 0|\mathbf{W}, \mathbf{g}) && \text{(24a)} \\ &= \sup_{\mathbf{X} \in \mathcal{S}_0} \sup \left\{ R : J\left(R|\mathbf{X}\right) \leq 0 \right\} && \text{(24b)} \\ &= \sup_{\mathbf{X} \in \mathcal{S}_0} \sup \left[ R : \limsup_{n \to \infty} \Pr \left\{ \frac{1}{n} i(X^n; Y^n) \leq R \right\} \leq 0 \right] && \text{(24c)} \\ &= \sup_{\mathbf{X} \in \mathcal{S}_0} \underline{I}(\mathbf{X}; \mathbf{Y}) && \text{(24d)} \end{aligned}$$

*Here the last line is due to the definitions of* $p$-$\liminf$ *and the spectral inf-information rate.* □

## 3.5 Strong Converse of Secure Capacity

A stego-channel $(\mathbf{W}, \mathbf{g})$ is said to satisfy the strong converse property if for any $R > C(\mathbf{W}, \mathbf{g})$, every $(n, M_n, \epsilon_n, \delta_n)$-code with,

$$\liminf_{n \to \infty} \frac{1}{n} \log M_n \geq R \qquad \text{and} \qquad \lim_{n \to \infty} \delta_n = 0,$$

we have,

$$\lim_{n \to \infty} \epsilon_n = 1.$$

THEOREM 3 (STRONG CONVERSE). *A stego-channel* $(\mathbf{W}, \mathbf{g})$ *satisfies the strong converse property if and only if,*

$$\sup_{\mathbf{X} \in \mathcal{S}_0} \underline{I}(\mathbf{X}; \mathbf{Y}) = \sup_{\mathbf{X} \in \mathcal{S}_0} \overline{I}(\mathbf{X}; \mathbf{Y}). \tag{25}$$

PROOF. *A general channel satisfies the strong converse if and only if[6, 13],*

$$\sup_{\mathbf{X}} \underline{I}(\mathbf{X}; \mathbf{Y}) = \sup_{\mathbf{X}} \overline{I}(\mathbf{X}; \mathbf{Y}). \tag{26}$$

*By our constraint we have that* $\delta_n \to 0$*, thus the* sup *is over the secure input set.* □

## 3.6 Maximum Bounds

We now derive a number of useful bounds on the spectral-entropy of an output sequence in relation to the permissible set. These bounds will then be used to prove general bounds for steganographic systems.

THEOREM 4 (SPECTRAL INF-ENTROPY BOUND). *For a discrete* $\mathbf{g} = \{\mathcal{P}_n\}_{n=1}^{\infty}$ *with corresponding secure output set* $\mathcal{T}_0$*,*

$$\sup_{\mathbf{Y} \in \mathcal{T}_0} \underline{H}(\mathbf{Y}) = \liminf_{n \to \infty} \frac{1}{n} \log |\mathcal{P}_n|. \tag{27}$$

PROOF. *Let* $\mathcal{U}(A)$ *represent the uniform distribution on a set* $A$*.*
*Since* $\mathbf{Y}^* = \{\mathcal{U}(\mathcal{P}_n)\}_{i=1}^{\infty} \in \mathcal{T}_0$ *we have,*

$$\begin{aligned} \sup_{\mathbf{Y} \in \mathcal{T}_0} \underline{H}(\mathbf{Y}) &\geq \underline{H}(\mathbf{Y}^*) && \text{(28a)} \\ &= \liminf_{n \to \infty} \frac{1}{n} \log |\mathcal{P}_n| && \text{(28b)} \end{aligned}$$

*Now assume there exists* $\overline{\mathbf{Y}} \in \mathcal{T}_0$ *with* $\overline{\mathbf{Y}} = \{\bar{Y}^n\}_{n=1}^{\infty}$*, such that,*

$$\underline{H}(\overline{\mathbf{Y}}) = \underline{H}(\mathbf{Y}^*) + 3\gamma, \tag{29}$$

*for any* $\gamma > 0$*.*
*This means that,*

$$\lim_{n \to \infty} \Pr \left\{ \frac{1}{n} \log \frac{1}{P_{\bar{Y}^n}(\bar{Y}^n)} < \underline{H}(\mathbf{Y}^*) + 2\gamma \right\} = 0 \tag{30}$$

*By* (28b) *we have* $\underline{H}(\mathbf{Y}^*) = \liminf_{n \to \infty} \frac{1}{n} \log |\mathcal{P}_n|$ *and from the definition of* $\liminf$ *we may find a subsequence indexed by* $k_n$ *such that,*

$$\underline{H}(\mathbf{Y}^*) + 2\gamma \geq \frac{1}{k_n} \log |\mathcal{P}_{k_n}| + \gamma. \tag{31a}$$

*For any* $k_n$ (31a) *holds and we have,*

$$\Pr \left\{ \frac{1}{k_n} \log \frac{1}{P_{\bar{Y}^{k_n}}(\bar{Y}^{k_n})} < \frac{1}{k_n} \log |\mathcal{P}_{k_n}| + \gamma \right\} \leq \Pr \left\{ \frac{1}{k_n} \log \frac{1}{P_{\bar{Y}^{k_n}}(\bar{Y}^{k_n})} < \underline{H}(\mathbf{Y}^*) + 2\gamma \right\}.$$

*Applying* (30) *to this result we have,*

$$\lim_{n \to \infty} \Pr \left\{ \frac{1}{k_n} \log \frac{1}{P_{\bar{Y}^{k_n}}(\bar{Y}^{k_n})} < \frac{1}{k_n} \log |\mathcal{P}_{k_n}| + \gamma \right\} = 0. \tag{32}$$

*Rearranging the inner term we have,*

$$\lim_{n \to \infty} \Pr \left\{ P_{\bar{Y}^{k_n}}(\bar{Y}^{k_n}) > \frac{e^{-k_n \gamma}}{|\mathcal{P}_{k_n}|} \right\} = 0. \tag{33}$$

*Thus given any* $\epsilon > 0$ *there exists* $n_0$ *such that when* $n > n_0$*,*

$$\Pr \left\{ P_{\bar{Y}^{k_n}}(\bar{Y}^{k_n}) > \frac{e^{-k_n \gamma}}{\mathcal{P}_{k_n}} \right\} < \epsilon. \tag{34}$$

*Let,*

$$A_{k_n} = \left\{ \mathbf{y} \in \mathcal{Y}^n : P_{\bar{Y}^{k_n}}(\bar{Y}^{k_n}) > \frac{e^{-k_n \gamma}}{|\mathcal{P}_{k_n}|} \right\}, \tag{35}$$

*so for all* $n > n_0$*,*

$$P_{\bar{Y}^{k_n}}(A_{k_n}) < \epsilon. \tag{36}$$

*For $n > n_0$ we may calculate the probability of the permissible set (for the subsequence) as,*

$$\begin{aligned}
P_{\bar{Y}^{k_n}}(\mathcal{P}_{k_n}) &= \sum_{\mathbf{y}\in\mathcal{P}_{k_n}} P_{\bar{Y}^{k_n}}(\mathbf{y}) \\
&= \sum_{\mathbf{y}\in\mathcal{P}_{k_n}\cap A_{k_n}^c} P_{\bar{Y}^{k_n}}(\mathbf{y}) + \sum_{\mathbf{y}\in\mathcal{P}_{k_n}\cap A_{k_n}} P_{\bar{Y}^{k_n}}(\mathbf{y}) \\
&\le \sum_{\mathbf{y}\in\mathcal{P}_{k_n}\cap A_{k_n}^c} \frac{e^{-k_n\gamma}}{|\mathcal{P}_{k_n}|} + \sum_{\mathbf{y}\in\mathcal{P}_{k_n}\cap A_{k_n}} P_{\bar{Y}^{k_n}}(\mathbf{y}) \\
&\le \sum_{\mathbf{y}\in\mathcal{P}_{k_n}} \frac{e^{-k_n\gamma}}{|\mathcal{P}_{k_n}|} + \sum_{\mathbf{y}\in\mathcal{P}_{k_n}\cap A_{k_n}} P_{\bar{Y}^{k_n}}(\mathbf{y}) \\
&= e^{-k_n\gamma} + \sum_{\mathbf{y}\in\mathcal{P}_{k_n}\cap A_{k_n}} P_{\bar{Y}^{k_n}}(\mathbf{y}) \\
&\le e^{-k_n\gamma} + \sum_{\mathbf{y}\in A_{k_n}} P_{\bar{Y}^{k_n}}(\mathbf{y}) \\
&< e^{-k_n\gamma} + \epsilon
\end{aligned}$$

*Thus for the subsequence,*

$$\limsup_{n\to\infty} P_{\bar{Y}^{k_n}}(\mathcal{P}_{k_n}) < \epsilon, \tag{37}$$

*for all $\epsilon > 0$ so clearly,*

$$\lim_{n\to\infty} P_{\bar{Y}^n}(\mathcal{P}_n) = 1, \tag{38}$$

*is impossible.*

*Thus from* (19) *we have a contradiction as the above implies that $\overline{\mathbf{Y}} \notin \mathcal{T}_0$.* □

THEOREM 5 (SPECTRAL SUP-ENTROPY BOUND). *For discrete $\mathbf{g} = \{\mathcal{P}_n\}_{n=1}^{\infty}$ with corresponding secure output set $\mathcal{T}_0$,*

$$\sup_{\mathbf{Y}\in\mathcal{T}_0} \overline{H}(\mathbf{Y}) = \limsup_{n\to\infty} \frac{1}{n}\log|\mathcal{P}_n|. \tag{39}$$

*The proof is similar to Theorem 4, and given in Appendix B.*

## 3.7 Capacity Bounds

This section presents a number of fundamental bounds on the secure capacity of a stego-channel based on the properties of that channel.

The first gives an upperbound based on the sup-entropy of the secure input set.

THEOREM 6 (INPUT SUP-ENTROPY BOUND). *For a stego-channel $(\mathbf{W},\mathbf{g})$ the secure capacity is bounded as,*

$$C(\mathbf{W},\mathbf{g}) \le \sup_{\mathbf{X}\in\mathcal{S}_0} \overline{H}(\mathbf{X}). \tag{40}$$

PROOF. *Using Theorem 14 and the property that $\overline{H}(\mathbf{X}|\mathbf{Y}) \ge 0$ we have,*

$$\begin{aligned}
C(\mathbf{W},\mathbf{g}) &= \sup_{\mathbf{X}\in\mathcal{S}_0} \underline{I}(\mathbf{X};\mathbf{Y}) && (41a)\\
&\le \sup_{\mathbf{X}\in\mathcal{S}_0} \left\{\overline{H}(\mathbf{X}) - \overline{H}(\mathbf{X}|\mathbf{Y})\right\} && (41b)\\
&\le \sup_{\mathbf{X}\in\mathcal{S}_0} \overline{H}(\mathbf{X}) && (41c)
\end{aligned}$$

□

The following corollary specializes the previous result with the restriction that the input alphabet is finite.

COROLLARY 1. *For a given stego-channel $(\mathbf{W},\mathbf{g})$ with a discrete input set ($|\mathcal{X}| < \infty$) the secure capacity is bounded from above as,*

$$C(\mathbf{W},\mathbf{g}) \le \log|\mathcal{X}|. \tag{42}$$

PROOF. *We make use of Theorem 6,*

$$\begin{aligned}
C(\mathbf{W},\mathbf{g}) &\le \sup_{\mathbf{X}\in\mathcal{S}_0} \overline{H}(\mathbf{X}) && (43a)\\
&\le \sup_{\mathbf{X}} \overline{H}(\mathbf{X}) && (43b)\\
&= \log|\mathcal{X}| && (43c)
\end{aligned}$$

□

The next theorem gives two upper bounds on the capacity based on the sup-entropy of the secure input and output sets.

THEOREM 7 (OUTPUT SUP-ENTROPY BOUNDS). *For a stego-channel $(\mathbf{W},\mathbf{g})$ the secure capacity is bounded as,*

$$\begin{aligned}
C(\mathbf{W},\mathbf{g}) &\le \sup_{\mathbf{X}\in\mathcal{S}_0} \overline{H}(\mathbf{Y}) && (44a)\\
&\le \sup_{\mathbf{Y}\in\mathcal{T}_0} \overline{H}(\mathbf{Y}) && (44b)
\end{aligned}$$

PROOF. *Using Theorem 14 and the property that $\overline{H}(\mathbf{Y}|\mathbf{X}) \ge 0$ we have,*

$$\begin{aligned}
C(\mathbf{W},\mathbf{g}) &= \sup_{\mathbf{X}\in\mathcal{S}_0} \underline{I}(\mathbf{X};\mathbf{Y}) && (45a)\\
&\le \sup_{\mathbf{X}\in\mathcal{S}_0} \left\{\overline{H}(\mathbf{Y}) - \overline{H}(\mathbf{Y}|\mathbf{X})\right\} && (45b)\\
&\le \sup_{\mathbf{X}\in\mathcal{S}_0} \overline{H}(\mathbf{Y}) && (45c)\\
&\le \sup_{\mathbf{Y}\in\mathcal{T}_0} \overline{H}(\mathbf{Y}) && (45d)
\end{aligned}$$

*Here the final line follows since if $\mathbf{X} \in \mathcal{S}_0$ then $\mathbf{Y} = \mathbf{W}(\mathbf{X}) \in \mathcal{T}_0$.* □

The next corollary specializes the above theorem when the permissible set is finite.

COROLLARY 2 (DISCRETE PERMISSIBLE SET BOUND). *For a given discrete stego-channel $(\mathbf{W},\mathbf{g}) = \{(W^n, \mathcal{P}_{g_n})\}_{n=1}^{\infty}$ the secure capacity is bounded from above as,*

$$C(\mathbf{W},\mathbf{g}) \le \limsup_{n\to\infty} \frac{1}{n}\log|\mathcal{P}_{g_n}| \tag{46}$$

PROOF. *Combining Theorem 5 and line* (44b) *of Theorem 7 gives the desired result.* □

The next theorem provides an intuitive result dealing with the capacity of two stego-channels having related detection functions.

THEOREM 8 (PERMISSIBLE SET RELATION). *For a given channel $\mathbf{W} = \{W^n\}_{n=1}^{\infty}$ and two detection functions $\mathbf{g} = \{g_n\}_{n=1}^{\infty}$ and $\mathbf{v} = \{v_n\}_{n=1}^{\infty}$, if $\mathcal{P}_{g_n} \subseteq \mathcal{P}_{v_n}$ for all but finitely many $n$, then,*

$$C(\mathbf{W},\mathbf{g}) \le C(\mathbf{W},\mathbf{v}). \tag{47}$$

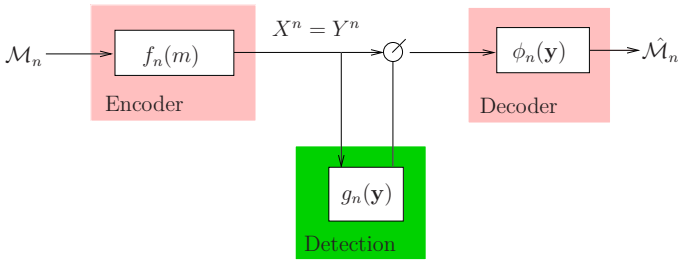


**Figure 3: Noiseless Passive System**

PROOF. *Let $\{f_n\}_{n=1}^{\infty}$ and $\{\phi_n\}_{n=1}^{\infty}$ be a sequence of encoding and decoding functions that achieves $C(\mathbf{W},\mathbf{g})$. Such a sequence exists by the definition of secure capacity. It suffices to show that this sequence is secure for $C(\mathbf{W},\mathbf{v})$.*

*The following definitions will be used for $i = 1,\ldots,M_n$,*

$$\mathbf{u}_i = f_n(i),$$
$$\mathcal{D}_i = \phi_n^{-1}(\{i\})\,..$$

*The probability of error for this sequence is given by* (9)*,*

$$\epsilon_n = \frac{1}{M_n}\sum_{i=1}^{M_n} W^n\left(\mathcal{D}_i^c|\mathbf{u}_i\right).$$

*Clearly, this value is independent of the permissible sets and if $\epsilon_n \to 0$ for the stego-channel $(\mathbf{W},\mathbf{g})$ then it also goes to zero for $(\mathbf{W},\mathbf{v})$.*

*Next we know that the probability of detection for $(\mathbf{W},\mathbf{g})$ is given by* (10)*,*

$$\delta_n^{\mathbf{g}} = \frac{1}{M_n}\sum_{i=1}^{M_n} W^n\left(\mathcal{I}_{g_n}|\mathbf{u}_i\right),$$

*and that $\delta_n^{\mathbf{g}} \to 0$.*

*Since $\mathcal{P}_{g_n} \subseteq \mathcal{P}_{v_n}$ for all $n > n_0$, we have that, $\mathcal{I}_{g_n} \supseteq \mathcal{I}_{v_n}$ if $n > n_0$ and thus,*

$$W^n\left(\mathcal{I}_{g_n}|\mathbf{x}\right) \geq W^n\left(\mathcal{I}_{v_n}|\mathbf{x}\right), \qquad \forall n > n_0, \mathbf{x} \in \mathcal{X}^n. \tag{48}$$

*Using this we may bound the probability of detection for $(\mathbf{W},\mathbf{v})$ and $n > n_0$ as,*

$$\delta_n^{\mathbf{v}} = \frac{1}{M_n}\sum_{i=1}^{M_n} W^n\left(\mathcal{I}_{v_n}|\mathbf{u}_i\right) \tag{49a}$$
$$\leq \frac{1}{M_n}\sum_{i=1}^{M_n} W^n\left(\mathcal{I}_{g_n}|\mathbf{u}_i\right) \tag{49b}$$
$$= \delta_n^{\mathbf{g}} \tag{49c}$$

*Since $\delta_n^{\mathbf{g}} \to 0$ we see that $\delta_n^{\mathbf{v}} \to 0$ as well.* □

# 4. NOISELESS CHANNELS

This section investigates the capacity of the system shown in Figure 3. The system is a special case of the general passive adversary shown in Figure 1 with the condition that there is no noise after the encoder.

This section finds the perfectly secure capacity of this system, and then derives a number of intuitive bounds relating to this capacity. We begin with a definition of noiseless channel.

## 4.1 Noiseless Channel

Consider the case where there is no distortion after the embedding, that is $Y^n = X^n$. If this is true, the channel transition probabilities reduce to,

$$W^n(\mathbf{y}|\mathbf{x}) = \begin{cases} 1, & \text{if } \mathbf{y} = \mathbf{x} \\ 0, & \text{if } \mathbf{y} \neq \mathbf{x} \end{cases} \tag{50}$$

In this case both the detection function and the decoder receive the stego-signal exactly as the encoder has constructed it.

If $W^n$ is noiseless for all $n$ then the general channel $\mathbf{W} = \{W^n\}_{n=1}^{\infty}$ is said to be a *noiseless channel.*

*Definition 19.* A stego-channel $(\mathbf{W},\mathbf{g})$ is said to be a *noiseless stego-channel* if $\mathbf{W}$ is noiseless.

## 4.2 Noiseless Channel Information-Spectrum Properties

This first Lemma shows that due to the noiseless property, the spectral mutual information rate and the spectral entropy rate coincide.

LEMMA 1. *For the noiseless channel $\mathbf{W}$, and any general source $\mathbf{X} = \{X^n\}_{n=1}^{\infty}$,*

$$\underline{I}(\mathbf{X};\mathbf{Y}) = \underline{H}(\mathbf{X}), \tag{51}$$
$$\overline{I}(\mathbf{X};\mathbf{Y}) = \overline{H}(\mathbf{X}). \tag{52}$$

PROOF. *Since $W^n(\mathbf{x}|\mathbf{x}) = 1$ for all $\mathbf{x} \in \mathcal{X}^n$ and $X^n = Y^n$ we see that,*

$$i(X^n;Y^n) = \log\frac{W^n(Y^n|X^n)}{P_{Y^n}(Y^n)} = \log\frac{1}{P_{X^n}(X^n)}. \tag{53}$$

*So the spectral inf-mutual information rate reduces to,*

$$\underline{I}(\mathbf{X};\mathbf{Y}) = p\text{-}\liminf_{n\to\infty}\frac{1}{n}i(X^n;Y^n)$$
$$= p\text{-}\liminf_{n\to\infty}\frac{1}{n}\log\frac{1}{P_{X^n}(X^n)} = \underline{H}(\mathbf{X}). \tag{54}$$

*Similarly for the spectral sup-mutual information rate,*

$$\overline{I}(\mathbf{X};\mathbf{Y}) = p\text{-}\limsup_{n\to\infty}\frac{1}{n}i(X^n;Y^n)$$
$$= p\text{-}\limsup_{n\to\infty}\frac{1}{n}\log\frac{1}{P_{X^n}(X^n)} = \overline{H}(\mathbf{X}). \tag{55}$$

□

The second useful Lemma is an extension of Theorems 4 and 5 under the noiseless conditions. It shows the relation between the supremum of the spectral entropy rate for secure-inputs and the permissible set size.

LEMMA 2. *For a discrete noiseless stego-channel defined by $\mathbf{g} = \{\mathcal{P}_{g_n}\}_{n=1}^{\infty}$,*

$$\sup_{\mathbf{X}\in\mathcal{S}_0}\underline{H}(\mathbf{X}) = \liminf_{n\to\infty}\frac{1}{n}\log|\mathcal{P}_{g_n}|, \tag{56}$$

$$\sup_{\mathbf{X}\in\mathcal{S}_0}\overline{H}(\mathbf{X}) = \limsup_{n\to\infty}\frac{1}{n}\log|\mathcal{P}_{g_n}|. \tag{57}$$

PROOF. *As the channel is noiseless we have that $\mathbf{X} = \mathbf{Y}$ and $\mathcal{S}_0 = \mathcal{T}_0$ so we may apply Theorems 4 and 5 for the desired results.* □

## 4.3 Secure Noiseless Capacity

The majority of the work in deriving a formula for the secure capacity of a discrete noiseless system has been done in Theorem 4 (Lemma 2) and Lemma 1. These results are combined with the general secure capacity formula of Theorem 2 in the following fundamental formula.

THEOREM 9 (SECURE NOISELESS CAPACITY). *For a discrete noiseless channel* $(\mathbf{W}, \mathbf{g}) = \{(W^n, g_n)\}_{n=1}^{\infty}$ *the secure channel capacity is given by,*

$$C(\mathbf{W}, \mathbf{g}) = \liminf_{n\to\infty} \frac{1}{n} \log |\mathcal{P}_{g_n}| . \tag{58}$$

PROOF. *By Theorem 2 we have,*

$$\begin{aligned} C(\mathbf{W}, \mathbf{g}) &= \sup_{\mathbf{X}\in\mathcal{S}_0} \underline{I}(\mathbf{X};\mathbf{Y}) && (59)\\ &= \sup_{\mathbf{X}\in\mathcal{S}_0} \underline{H}(\mathbf{X}) && (60)\\ &= \liminf_{n\to\infty} \frac{1}{n} \log |\mathcal{P}_{g_n}| && (61) \end{aligned}$$

*Where the final two lines are due to Lemma 1 and Lemma 2 respectively.* □

## 4.4 Strong Converse for Noiseless Channels

We now present a fundamental result for discrete noiseless channels regarding the strong converse property. It gives the necessary and sufficient conditions for a noiseless stego-channel to satisfy the strong converse property.

THEOREM 10 (NOISELESS STRONG CONVERSE). *A discrete noiseless stego-channel* $(\mathbf{W}, \mathbf{g})$ *satisfies the strong converse property if and only if,*

$$\lim_{n\to\infty} \frac{1}{n} \log |\mathcal{P}_{g_n}| , \tag{62}$$

*exists.*

*Furthermore the secure capacity is equal to this limit,*

$$C(\mathbf{W}, \mathbf{g}) = \lim_{n\to\infty} \frac{1}{n} \log |\mathcal{P}_{g_n}| . \tag{63}$$

PROOF. *First assume that the stego-channel satisfies the strong converse property. By Theorem 3 we have,*

$$\sup_{\mathbf{X}\in\mathcal{S}_0} \underline{I}(\mathbf{X};\mathbf{Y}) = \sup_{\mathbf{X}\in\mathcal{S}_0} \overline{I}(\mathbf{X};\mathbf{Y}) \tag{64}$$

*This gives,*

$$\begin{aligned} \sup_{\mathbf{X}\in\mathcal{S}_0} \underline{I}(\mathbf{X};\mathbf{Y}) &= \liminf_{n\to\infty} \frac{1}{n} \log |\mathcal{P}_{g_n}| && (65a)\\ &= \sup_{\mathbf{X}\in\mathcal{S}_0} \overline{I}(\mathbf{X};\mathbf{Y}) && (65b)\\ &= \sup_{\mathbf{X}\in\mathcal{S}_0} \overline{H}(\mathbf{X}) && (65c)\\ &= \limsup_{n\to\infty} \frac{1}{n} \log |\mathcal{P}_{g_n}| && (65d) \end{aligned}$$

*Here the equality of* (65a) *is due to Theorem 9,* (65b) *is by assumption,* (65c) *is by Lemma 1 and* (65d) *is from Lemma 2.*

*This shows that,*

$$\begin{aligned} C(\mathbf{W}, \mathbf{g}) &= \liminf_{n\to\infty} \frac{1}{n} \log |\mathcal{P}_{g_n}| \\ &= \limsup_{n\to\infty} \frac{1}{n} \log |\mathcal{P}_{g_n}| = \lim_{n\to\infty} \frac{1}{n} \log |\mathcal{P}_{g_n}| . \end{aligned}$$

*For the other direction assume that* $C(\mathbf{W}, \mathbf{g}) = \lim_{n\to\infty} \frac{1}{n} \log |\mathcal{P}_{g_n}|$ *thus we have,*

$$\begin{aligned} C(\mathbf{W}, \mathbf{g}) &= \sup_{\mathbf{X}\in\mathcal{S}_0} \underline{I}(\mathbf{X};\mathbf{Y}) && (66)\\ &= \lim_{n\to\infty} \frac{1}{n} \log |\mathcal{P}_{g_n}| && (67)\\ &= \limsup_{n\to\infty} \frac{1}{n} \log |\mathcal{P}_{g_n}| && (68)\\ &= \sup_{\mathbf{X}\in\mathcal{S}_0} \overline{H}(\mathbf{X}) && (69)\\ &= \sup_{\mathbf{X}\in\mathcal{S}_0} \overline{I}(\mathbf{X};\mathbf{Y}) && (70) \end{aligned}$$

*Where we have used Lemma 2 for line* (69) *and Lemma 1 for* (70).

*Thus,* $\sup_{\mathbf{X}\in\mathcal{S}_0} \underline{I}(\mathbf{X};\mathbf{Y}) = \sup_{\mathbf{X}\in\mathcal{S}_0} \overline{I}(\mathbf{X};\mathbf{Y})$ *and by Theorem 3 the stego-channel satisfies the strong-converse property.* □

## 4.5 Capacity of the Noiseless DMSC

In this section we briefly investigate the secure capacity of the discrete memoryless stego-channel.

THEOREM 11 (NOISELESS DMSC SECURE CAPACITY). *For a noiseless DMSC defined by* $\mathbf{g} = \{g\}$ *the secure capacity is equal to,*

$$C(\mathbf{W}, \mathbf{g}) = \log |\mathcal{P}_g| , \tag{71}$$

*and furthermore this stego-channel satisfies the strong converse.*

PROOF. *As the channel is noiseless and the input alphabet is finite we may use Theorem 9,*

$$C(\mathbf{W}, \mathbf{g}) = \liminf_{n\to\infty} \frac{1}{n} \log |\mathcal{P}_{g_n}| . \tag{72}$$

*Note that by* (7) *we have for all* $n$*,*

$$\begin{aligned} \frac{1}{n} \log |\mathcal{P}_{g_n}| &= \frac{1}{n} \log \left| \underbrace{\mathcal{P}_g \times \mathcal{P}_g \times \cdots \times \mathcal{P}_g}_{n} \right| \\ &= \frac{1}{n} \log |\mathcal{P}_g|^n \\ &= \log |\mathcal{P}_g| . \end{aligned}$$

*Thus,*

$$C(\mathbf{W}, \mathbf{g}) = \log |\mathcal{P}_g| .$$

*We also have that,*

$$C(\mathbf{W}, \mathbf{g}) = \liminf_{n\to\infty} \frac{1}{n} \log |\mathcal{P}_{g_n}| = \log |\mathcal{P}_g| = \lim_{n\to\infty} \frac{1}{n} \log |\mathcal{P}_{g_n}| , \tag{73}$$

*thus by Theorem 10 the stego-channel satisfies the strong converse.* □

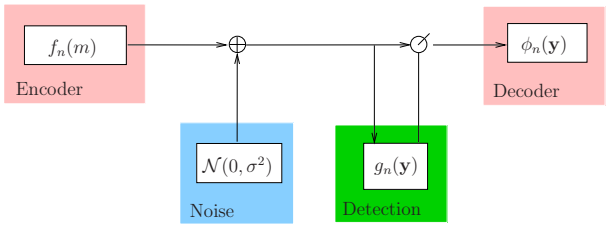


**Figure 4: AWGN Channel**

# 5. NOISY EXAMPLE

In this section we evaluate the capacity of particular stego-channel. The channel is the commonly studied additive white Gaussian noise one. The detector is motivated by the use of spread spectrum steganography[8] or more generally stochastic modulation[3].

## 5.1 Additive Gaussian Channel

The channel to be considered is called the additive white Gaussian noise channel and is shown in Figure 4. For a stego-signal, $\mathbf{x} = (x_1, \ldots, x_n)$, the corrupted stego-signal is given by,

$$\mathbf{y} = (x_1 + n_1, \ldots, x_n + n_n),$$

where each $n_i \sim \mathcal{N}(0, \sigma^2)$, and all are independent.

The transition probabilities of this channel are given by,

$$W^n(\mathbf{y}|\mathbf{x}) = \frac{1}{(2\pi\sigma^2)^{\frac{n}{2}}} \exp\left\{-\frac{1}{2\sigma^2}\sum_{i=1}^{n}(y_i - x_i)^2\right\}. \quad (74)$$

## 5.2 Variance Detector

In stochastic modulation a pseudo-noise is modulated by a message and added to the cover signal. This is done as the introduction of noise in signal processing applications is a common occurrence.

If the passive adversary has a knowledge of the distribution of the cover-signal and knows that the hider is using stochastic modulation, they also know that the variance of a cover-signal will differ from the variance of a stego-signal. If the passive adversary knows that the variance of the cover-distribution they could design a detector to trigger if the variance of a test signal is higher than this threshold.

For example, when testing the signal $\mathbf{y} = (y_1, \ldots, y_n)$ the variance detector operates as,

$$g_n(\mathbf{y}) = \begin{cases} 1, & \text{if } \frac{1}{n}\sum_{i=1}^{n} y_i^2 > c \\ 0, & \text{else} \end{cases} \quad (75)$$

Thus, if the empirical variance of a test signal is above a certain threshold, the signal is considered steganographic.

## 5.3 Channel Properties

Let $\mathbf{N} = \{N\}$ denote the channel noise where each $N$ is i.i.d. Gaussian with zero mean and variance $\sigma^2$. For a general source $\mathbf{X} = \{X^n\}_{n=1}^{\infty}$ define the general source of their sum as $\mathbf{N} + \mathbf{X} := \{N^n + X^n\}_{n=1}^{\infty}$.

Using this notation the corrupted stego-signal is expressed as $\mathbf{Y} = \mathbf{N} + \mathbf{X}$.

Since $\mathbf{N}$ is independent of $\mathbf{X}$ we have that,

$$\begin{aligned}
\overline{H}(\mathbf{Y}|\mathbf{X}) &= \text{p-}\limsup \frac{1}{n}\log\frac{1}{P_{Y^n|X^n}(Y^n|X^n)} \\
&= \text{p-}\limsup \frac{1}{n}\log\frac{1}{P_{Y^n|X^n}(N^n + X^n|X^n)} \\
&= \text{p-}\limsup \frac{1}{n}\log\frac{1}{P_{N^n}(N^n)} \\
&= \overline{H}(\mathbf{N}) \\
&= \text{p-}\limsup -\frac{1}{n}\log\sum_{i=1}^{n} P_N(N_i) \\
&= H(N) \\
&= \frac{1}{2}\log 2\pi e\sigma^2
\end{aligned}$$

This gives the following useful simplification,

$$\overline{H}(\mathbf{Y}|\mathbf{X}) = \overline{H}(\mathbf{N}) = H(N) = \frac{1}{2}\log 2\pi e\sigma^2. \quad (76)$$

## 5.4 Secure Capacity

We now derive the secure capacity of the above stego-channel.

THEOREM 12. *For the stego-channel* $(\mathbf{W}, \mathbf{g}) = \{(W^n, g_n)\}_{n=1}^{\infty}$ *with* $W^n$ *defined by* (74) *and* $g_n$ *defined by* (75) *the secure capacity is,*

$$C(\mathbf{W}, \mathbf{g}) = \frac{1}{2}\log\frac{c}{\sigma^2}. \quad (77)$$

PROOF. **Achievability:**
*Using Theorem 15 to show a lowerbound on the secure capacity,*

$$\begin{aligned}
C(\mathbf{W}, \mathbf{g}) &= \sup_{\mathbf{X}\in\mathcal{S}_0} \underline{I}(\mathbf{X}; \mathbf{Y}) && (78a)\\
&\geq \sup_{\mathbf{X}\in\mathcal{S}_0} \left\{\underline{H}(\mathbf{Y}) - \overline{H}(\mathbf{Y}|\mathbf{X})\right\} && (78b)\\
&= \sup_{\mathbf{X}\in\mathcal{S}_0} \left\{\underline{H}(\mathbf{Y}) - H(N)\right\} && (78c)\\
&= \sup_{\mathbf{X}\in\mathcal{S}_0} \left\{\underline{H}(\mathbf{Y}) - \frac{1}{2}\log\left(2\pi e\sigma^2\right)\right\} && (78d)\\
&= \sup_{\mathbf{X}\in\mathcal{S}_0} \underline{H}(\mathbf{Y}) - \frac{1}{2}\log\left(2\pi e\sigma^2\right) && (78e)
\end{aligned}$$

*Here line* (78d) *is due to the additive noise simplification of* (76).

*Now assume that* $\overline{\mathbf{X}}$ *is i.i.d. Gaussian with variance of* $c - \sigma^2$. *So for* $\overline{\mathbf{Y}} = \overline{\mathbf{X}} + \mathbf{N}$ *we have that* $\overline{\mathbf{Y}}$ *is i.i.d. Gaussian with variance* $c$. *From this we have that* $\overline{\mathbf{X}} \in \mathcal{S}_0$ *since,*

$$\Pr\left\{\frac{1}{n}\sum y_i^2 > c\right\} = \Pr\left\{\frac{1}{n}\sum (x_i + n_i)^2 > c\right\} \to 0. \quad (79)$$

*Also note that the inf-entropy of* $\overline{\mathbf{Y}}$ *is,*

$$\underline{H}(\overline{\mathbf{Y}}) = \underline{H}(\overline{\mathbf{X}} + \mathbf{N}) = \frac{1}{2}\log(2\pi ec). \quad (80)$$

*Applying this result to* (78e) *we see,*

$$\begin{aligned}
C(\mathbf{W}, \mathbf{g}) &\geq \underline{H}(\overline{\mathbf{Y}}) - \frac{1}{2}\log\left(2\pi e\sigma^2\right) && (81a)\\
&= \frac{1}{2}\log(2\pi ec) - \frac{1}{2}\log\left(2\pi e\sigma^2\right) && (81b)\\
&= \frac{1}{2}\log\frac{c}{\sigma^2} && (81c)
\end{aligned}$$

**Converse:**
*To find the upperbound we will make use of a number of simple lemmas:*

LEMMA *3. For* $Y^n = (Y_1^{(n)}, Y_2^{(n)}, \ldots, Y_n^{(n)})$ *let* $K_{ij}^{(n)} := EY_i^{(n)}Y_j^{(n)}$. *For the stego-channel defined above, if* $\mathbf{Y} = \{Y^n\}_{n=1}^{\infty} \in \mathcal{T}_0$ *we have for any* $\gamma > 0$ *there exists some* $n_0$ *such that for all* $n > n_0$,

$$\frac{1}{n}\sum_{i=1}^{n} K_{ii}^{(n)} < c + \gamma. \tag{82}$$

PROOF. *To show this assume that no such* $n_0$ *exists, thus we have a subsequence* $n_k$ *such that,*

$$\frac{1}{n_k}\sum_{i=1}^{n_k} K_{ii}^{(n_k)} \geq c + \gamma. \tag{83}$$

*This means that,*

$$\frac{1}{n_k}\sum_{i=1}^{n_k} K_{ii}^{(n_k)} = E\left\{\frac{1}{n_k}\sum_{i=1}^{n_k} y_i^2\right\} \geq c + \gamma,$$

*which in turn implies that,*

$$\Pr\left\{g_{n_k}(Y^{n_k}) = 0\right\} \to 0.$$

*This is a contradiction in that* $\mathbf{Y} = \{Y^n\}_{n=1}^{\infty} \in \mathcal{T}_0$. □

LEMMA *4. For any* $\mathbf{Y} \in \mathcal{T}_0$ *and any* $\epsilon > 0$ *we have,*

$$\liminf_{n\to\infty} \frac{1}{n} H(Y^n) < \frac{1}{2}\log 2\pi e c + \epsilon. \tag{84}$$

PROOF. *Let any* $\epsilon > 0$ *be given and choose* $\gamma > 0$ *such that,*

$$\gamma \leq c\left(e^{2\epsilon} - 1\right),$$

*this gives,*

$$\frac{1}{2}\log 2\pi e\left(c + \gamma\right) \leq \frac{1}{2}\log 2\pi e c + \epsilon. \tag{85}$$

*For all but a finite number of* $n$ *we have,*

$$\begin{aligned}
\frac{1}{n}H(Y^n) &\leq \frac{1}{2n}\log(2\pi e)^n \prod_{i=1}^{n} K_{ii}^{(n)} && (86)\\
&\leq \frac{1}{2n}\log(2\pi e)^n \left(\frac{1}{n}\sum_{i=1}^{n} K_{ii}^{(n)}\right)^n && (87)\\
&< \frac{1}{2n}\log(2\pi e)^n \left(c + \gamma\right)^n && (88)\\
&= \frac{1}{2}\log 2\pi e\left(c + \gamma\right) && (89)\\
&\leq \frac{1}{2}\log 2\pi e c + \epsilon && (90)
\end{aligned}$$

*where we have used the fact that for any* $Y^n = (Y_1^{(n)}, \ldots, Y_n^{(n)})$ *with* $K_{ij}^{(n)} = EY_i^{(n)}Y_j^{(n)}$ *we have [2, Chap. 9.6],*

$$H(Y^n) \leq \frac{1}{2}\log(2\pi e)^n \prod_{i=1}^{n} K_{ii}^{(n)}, \tag{91}$$

*as well as the arithmetic-geometric inequality in* (87). □

*We now show the upperbound:*

$$\begin{aligned}
C(\mathbf{W}, \mathbf{g}) &= \sup_{\mathbf{X}\in\mathcal{S}_0} \underline{I}(\mathbf{X}; \mathbf{Y}) && (92a)\\
&\leq \sup_{\mathbf{Y}\in\mathcal{S}_0} \underline{I}(\mathbf{X}; \mathbf{Y}) && (92b)\\
&\leq \sup_{\mathbf{Y}\in\mathcal{T}_0} \liminf_{n\to\infty} \frac{1}{n} I(X^n; Y^n) && (92c)\\
&= \sup_{\mathbf{Y}\in\mathcal{T}_0} \liminf_{n\to\infty} \frac{1}{n}\left\{H(Y^n) - H(Y^n|X^n)\right\} && (92d)\\
&= \sup_{\mathbf{Y}\in\mathcal{T}_0} \liminf_{n\to\infty} \left\{\frac{1}{n}H(Y^n)\right\} - H(N) && (92e)\\
&< \frac{1}{2}\log 2\pi e c + \epsilon - \frac{1}{2}\log 2\pi e\sigma^2 && (92f)\\
&= \frac{1}{2}\log\frac{c}{\sigma^2} + \epsilon && (92g)
\end{aligned}$$

*Here line* (92b) *is due to the fact that for any* $\mathbf{X} \in \mathcal{S}_0$, *we have* $\mathbf{Y} = \mathbf{W}(\mathbf{X}) \in \mathcal{T}_0$. *Next, line* (92c) *is from Theorem 16 and the inequality of* (92f) *is due to Lemma 4.*

*Thus combining* (81c) *and* (92g) *we have for any* $\epsilon > 0$,

$$\frac{1}{2}\log\frac{c}{\sigma^2} \leq C(\mathbf{W}, \mathbf{g}) < \frac{1}{2}\log\frac{c}{\sigma^2} + \epsilon,$$

*and we see that* $C(\mathbf{W}, \mathbf{g}) = \frac{1}{2}\log\frac{c}{\sigma^2}$.

# 6. COMPARISON TO PREVIOUS WORK

## 6.1 Cachin Perfect Security

In Cachin's definition of perfect security the cover-signal distribution and the stego-signal distribution are each required to be independent and identically distributed. This gives the following secure-input set,

$$\mathcal{S}_0 = \left\{\mathbf{X} = \{X\} : \lim_{n\to\infty} \frac{1}{n} D\left(S^n||X^n\right) = 0\right\}. \tag{93}$$

The i.i.d. property means that $D\left(S^n||X^n\right) = nD\left(S||X\right)$ so we see that the above is equivalent to,

$$\begin{aligned}
\mathcal{S}_0 &= \{\mathbf{X} = \{X\} : D\left(S||X\right) = 0\} && (94)\\
&= \{\mathbf{X} = \{X\} : P_S = P_X\} && (95)
\end{aligned}$$

Since Cachin's definition does not model noise, we may consider it as noiseless and apply Theorem 9,

$$C(\mathbf{W}, \mathbf{g}) = \sup_{\mathbf{X}\in\mathcal{S}_0} \underline{H}(\mathbf{X}) = H(S). \tag{96}$$

This result states that in a system that is perfectly secure (in Cachin's definition) the limit on the amount of information that may be transferred each channel use is equal to the entropy of the source. This is intuitive because in Cachin's definition the output distribution of the encoder is constrained to be equal to the cover distribution.

## 6.2 Empirical Distribution Detection Function

The *empirical distribution detection function* is motivated by the fact that the empirical distribution from a stationary memoryless source converges to the actual distribution of that source. Accordingly, if a test signals empirical distribution converges to the cover-signal distribution is it considered to be non-steganographic.

Assume that $P_S$ is a discrete distribution over the finite alphabet $\mathcal{S}$. Let a sequence, $\{s^n\}_{n=1}^{\infty}$ with each $s^n \in \mathcal{S}^n$ be

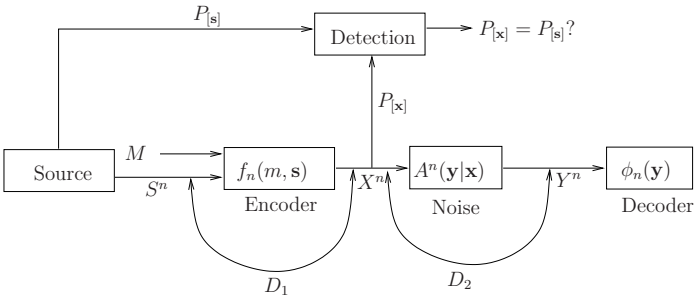


**Figure 5: Moulin Stego-channel**

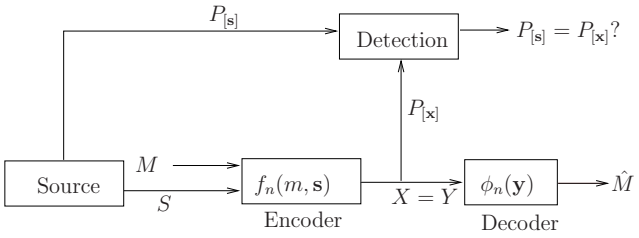


**Figure 6: Equivalent Stego-channel**

used to specify the detection function for a test signal $\mathbf{x}$ as,

$$g_n(\mathbf{x}) = \begin{cases} 0 \text{ if } P_{[s^n]} = P_{[\mathbf{x}]}, \\ 1 \text{ if } P_{[s^n]} \neq P_{[\mathbf{x}]}. \end{cases} \tag{97}$$

where $P_{[\mathbf{x}]}$ is the empirical distribution of $\mathbf{x}$.

The permissible set for $g_n$ is equal to the type class of $P_{[s^n]}$, i.e.,

$$\mathcal{P}_{g_n} = T(P_{[s^n]}) := \left\{\mathbf{x} \in \mathcal{X}^n : P_{[\mathbf{x}]} = P_{[s^n]}\right\}. \tag{98}$$

## 6.3 Moulin Steganographic Capacity

Moulin's formulation[11] of the stego-channel is shown in Figure 5. This is somewhat different than the formulation shown in Figure 1, most notable is the presence of distortion constraints and an active adversary. Additionally there is an absence of a distortion function prior to the detection function. Also in this model the detection function is fixed as the previously discussed empirical distribution detection function. The sequence of $s^n$ to specify the detection function is drawn i.i.d. as $S$. In order to have the two formulations coincide a number of simplifications are needed for each model.

For our model,

- The stego-channel is noiseless
- The detection function is the empirical distribution

For Moulin's model,

- Passive Adversary ($D_2 = 0$)
- No distortion constraint on encoder ($D_1 = \infty$)

These changes produce the stego-channel shown in Figure 6.

THEOREM 13. *For the stego-channel shown in Figure 6, the capacities of this work and Moulin's agree. That is,*

$$C(\mathbf{W}, \mathbf{g}) = C^{STEG}(\infty, 0) = H(S). \tag{99}$$

PROOF. *Since the channel is noiseless we may apply Theorem 9.*

$$\begin{aligned} C(\mathbf{W}, \mathbf{g}) &= \liminf_{n\to\infty} \frac{1}{n} \log |\mathcal{P}_{g_n}| && \text{(100a)} \\ &= \liminf_{n\to\infty} \frac{1}{n} \log |T(s^n)| && \text{(100b)} \\ &= H(S) && \text{(100c)} \end{aligned}$$

*Here we have used the fact that the permissible set for the empirical distribution detection function is the type class in* (100b)*. Additionally, by the Glivenko-Cantelli Theorem[12], $P_{[s^n]}(x) \to P_S(x)$ almost surely (here the convergence is uniform in $x$ as well). This allows for the use of the type class-entropy bound from Appendix C that provides the final result.*

*We now show Moulin's capacity is equal to this value. In the case of a passive adversary ($D_2 = 0$), the following is the capacity of the stego-channel[11],*

$$C^{STEG}(D_1, 0) = \sup_{Q' \in \mathcal{Q}'} H(X|S) \tag{101}$$

*where a $p \in \mathcal{Q}'$ is feasible if,*

$$\sum_{s,x} p(x|s) P_S(s) d(s, x) \leq D_1, \tag{102}$$

*and*

$$\sum_{s} p(x|s) P_S(s) = P_S(x). \tag{103}$$

*First we upper-bound the secure capacity as,*

$$\begin{aligned} C^{STEG}(\infty, 0) &= \sup_{p(x|s) \in \mathcal{Q}'} H(X|S) && \text{(104a)} \\ &\leq \sup_{p(x) \in \mathcal{Q}'} H(X) && \text{(104b)} \\ &= H(S) && \text{(104c)} \end{aligned}$$

*Where the final line comes from the requirement that if $p \in \mathcal{Q}'$ and $p(x|s) = p(x)$ then $p(x) = P_S(x)$ for all $x$, to satisfy* (103)*.*

*For the lower-bound we let Let $P_{\tilde{X}S}(x, s) = P_{\tilde{X}|S}(x|s)P_S(s) = P_S(x)P_S(s)$, i.e. $\tilde{X} \sim P_S$. This defines a feasible covert-channel as* (102) *is trivially satisfied (since $D_1 = \infty$) and* (103) *is as well since,*

$$\sum_{s} P_{\tilde{X}|S}(x|s) P_S(s) = \sum_{s} P_S(x) P_S(s) = P_S(x). \tag{105}$$

*This gives,*

$$\begin{aligned} C^{STEG}(\infty, 0) &= \sup_{p(x|s) \in \mathcal{Q}'} H(X|S) && \text{(106a)} \\ &\geq H(\tilde{X}|S) && \text{(106b)} \\ &= H(\tilde{X}) && \text{(106c)} \\ &= H(S) && \text{(106d)} \end{aligned}$$

*Here* (106c) *is because $\tilde{X}$ and $S$ are independent ($P_{\tilde{X}S}(x, s) = P_{\tilde{X}}(x)P_S(s)$).* □

## 7. CONCLUSIONS

A framework for evaluating the capacity of steganographic channels under a passive adversary has been introduced. The system considers a noise corrupting the signal before the detection function in order to model real-world distortions such as compression, quantization, etc.

Constraints on the encoder dealing with distortion and a cover-signal are not considered. Instead, the focus is to develop the theory necessary to analyze the interplay between the channel and detection function that results in the steganographic capacity.

The method uses an information-spectrum approach that allows for the analysis of arbitrary detection functions and channels. This provides machinery necessary to analyze a very broad range of steganographic channels.

In addition to offering insight into the limits of performance for steganographic algorithms, this formulation of capacity can be used to analyze a different, and fundamentally important, facet of steganalysis. While false alarms and missed signals have rightfully dominated the steganalysis literature, very little is known about the amount of information that can be sent past these algorithms. This work presents a theory to shed light onto this important quantity called steganographic capacity.

## 8. ACKNOWLEDGMENTS

The support of the Center for Integrated Transmission and Exploitation (CITE) and the Information Directorate of the Air Force Research Laboratory, Rome, NY is gratefully acknowledged.

The authors would also like to thank the reviewers for their valuable input.

## 9. REFERENCES

[1] R. Chandramouli and N. Memon. Steganography capacity: a steganalysis perspective. In *Proc. SPIE Electronic Imaging 5022*, Santa Clara, CA, Jan. 21–24, 2003.

[2] T. M. Cover and J. A. Thomas. *Elements of information theory*. Wiley-Interscience, 1991.

[3] J. Fridrich and M. Goljan. Digital image steganography using stochastic modulation. In *Proc. SPIE Electronic Imaging 5022*, Santa Clara, CA, Jan. 21–24, 2003.

[4] T. S. Han. An information-spectrum approach to source coding theorems with a fidelity criterion. *IEEE Trans. on Information Theory*, 43(4):1145–1164, July 1997.

[5] T. S. Han. Hypothesis testing with the general source. *IEEE Trans. on Information Theory*, 46(7):2415–2427, Nov. 2000.

[6] T. S. Han. *Information-Spectrum Methods in Information Theory*, volume 50 of *Applications of Mathematics*. Springer-Verilog, Berlin, Germany, 2003.

[7] T. S. Han and S. Verdú. Generalizing the Fano inequality. *IEEE Trans. on Information Theory*, 40(4):1247–1251, July 1994.

[8] L. M. Marvel, C. G. Boncelet, Jr, and C. T. Retter. Spread spectrum image steganography. *IEEE Trans. Image Processing*, 8(8):1075–1083, Aug. 1999.

[9] I. S. Moskowitz, L. Chang, and R. E. Newman. Capacity is the wrong paradigm. In *Proceedings of the 2002 workshop on New security paradigms*, pages 114–126. ACM Press, 2002.

[10] P. Moulin and J. A. O'Sullivan. Information-theoretic analysis of information hiding. *IEEE Transactions on Information Theory*, 49(3):563–593, Mar. 2003.

[11] P. Moulin and Y.Wang. New results on steganographic capacity. In *Proc. CISS Conference*, Princeton, NJ, Mar. 2004.

[12] S. I. Resnick. *A Probability Path*. Birkhäuser, Boston, MA, 2 edition, 2001.

[13] S. Verdú and T. S. Han. A general formula for channel capacity. *IEEE Trans. on Information Theory*, 40(4):1147–1157, July 1994.

## APPENDIX

## A. SPECTRAL INFORMATION PROPERTIES

The following two theorems are basic information-spectrum properties[6, Chap. 3.2] and Theorem 8 of [13].

THEOREM 14.

$$\underline{I}(\mathbf{X};\mathbf{Y}) \leq \overline{H}(\mathbf{Y}) - \overline{H}(\mathbf{Y}|\mathbf{X}) \tag{107}$$

*where,*

$$\overline{H}(\mathbf{Y}|\mathbf{X}) := p\text{-}\limsup_{n\to\infty} \frac{1}{n} \log \frac{1}{P_{Y^n|X^n}(Y^n|X^n)}. \tag{108}$$

THEOREM 15.

$$\underline{I}(\mathbf{X};\mathbf{Y}) \geq \underline{H}(\mathbf{Y}) - \overline{H}(\mathbf{Y}|\mathbf{X}). \tag{109}$$

The following relation is Theorem 3.5.2 of [6, Chap. 3.5] as well as Theorem 8 of [13],

THEOREM 16. *For* $\mathbf{X} = \{X^n\}_{n=1}^{\infty}$ *and* $\mathbf{Y} = \{Y^n\}_{n=1}^{\infty}$*, the following inequality holds,*

$$\underline{I}(\mathbf{X};\mathbf{Y}) \leq \liminf_{n\to\infty} \frac{1}{n} I(X^n;Y^n). \tag{110}$$

## B. THEOREM 5 PROOF

PROOF. Since $\mathbf{Y}^* = \{\mathcal{U}(\mathcal{P}_n)\}_{i=1}^{\infty} \in \mathcal{T}_0$ we have,

$$\sup_{\mathbf{Y}\in\mathcal{T}_0} \overline{H}(\mathbf{Y}) \geq \overline{H}(\mathbf{Y}^*) \tag{111a}$$

$$= \limsup_{n\to\infty} \frac{1}{n} \log |\mathcal{P}_n| \tag{111b}$$

Now assume there exists $\overline{\mathbf{Y}} \in \mathcal{T}_0$, with $\overline{\mathbf{Y}} = \{\bar{Y}^n\}_{n=1}^{\infty}$ such that,

$$\overline{H}(\overline{\mathbf{Y}}) = \overline{H}(\mathbf{Y}^*) + \frac{\gamma}{2}, \tag{112}$$

for any $\gamma > 0$.

This means that,

$$\lim_{n\to\infty} \Pr\left\{ \frac{1}{n} \log \frac{1}{P_{\bar{Y}^n}(\bar{Y}^n)} > \overline{H}(\overline{\mathbf{Y}}) + \frac{\gamma}{2} \right\} = 0 \tag{113}$$

Thus for some subsequence $k_n$ we have,

$$\frac{1}{k_n} \log |\mathcal{P}_{k_n}| + \gamma > \overline{H}(\overline{\mathbf{Y}}) + \frac{\gamma}{2} \tag{114}$$

and

$$\lim_{n\to\infty} \Pr\left\{ \frac{1}{k_n} \log \frac{1}{P_{\bar{Y}^{k_n}}(\bar{Y}^{k_n})} > \frac{1}{k_n} \log |\mathcal{P}_{k_n}| + \gamma \right\} = 0. \tag{115}$$

Rewriting we have,

$$\lim_{n\to\infty} \Pr\left\{ P_{\bar{Y}^{k_n}}(\bar{Y}^{k_n}) < \frac{e^{-k_n\gamma}}{|\mathcal{P}_{k_n}|} \right\} = 0 \tag{116}$$

Let,

$$A_{k_n} := \left\{ \mathbf{y} \in \mathcal{X}^n : P_{\bar{Y}^{k_n}}(\bar{Y}^{k_n}) < \frac{e^{-k_n\gamma}}{|\mathcal{P}_{k_n}|} \right\} \tag{117}$$

and given any $\epsilon > 0$ we may find $n_0$ so for $n > n_0$,

$$P_{\bar{Y}^{k_n}}(A_{k_n}) < \epsilon. \tag{118}$$

For $n > n_0$ the probability of the permissible set (in this subsequence) is,

$$\begin{aligned}
P_{\bar{Y}^{k_n}}(\mathcal{P}_{k_n}) &= \sum_{\mathbf{x}\in\mathcal{P}_{k_n}} P_{\bar{Y}^{k_n}}(\mathbf{y}) \\
&= \sum_{\mathbf{y}\in\mathcal{P}_{k_n}\cap A_{k_n}} P_{\bar{Y}^{k_n}}(\mathbf{y}) + \sum_{\mathbf{y}\in\mathcal{P}_{k_n}\cap A_{k_n}} P_{\bar{Y}^{k_n}}(\mathbf{y}) \\
&\le \frac{e^{-k_n\gamma}}{|\mathcal{P}_{k_n}|} \sum_{\mathbf{y}\in\mathcal{P}_{k_n}\cap A_{k_n}} + \sum_{\mathbf{y}\in\mathcal{P}_{k_n}\cap A_{k_n}} P_{\bar{Y}^{k_n}}(\mathbf{y}) \\
&\le e^{-k_n\gamma} + \sum_{\mathbf{y}\in\mathcal{P}_{k_n}\cap A_{k_n}} P_{\bar{Y}^{k_n}}(\mathbf{y}) \\
&\le e^{-k_n\gamma} + \sum_{\mathbf{y}\in A_{k_n}} P_{\bar{Y}^{k_n}}(\mathbf{y}) \\
&< e^{-k_n\gamma} + \epsilon
\end{aligned}$$

This gives,

$$\limsup_{n\to\infty} P_{\bar{Y}^{k_n}}(\mathcal{P}_{k_n}) < \epsilon, \tag{120}$$

for any $\epsilon > 0$. Since the subsequence above does not converge to 1 it is impossible for,

$$\lim_{n\to\infty} P_{\bar{Y}^n}(\mathcal{P}_n) = 1, \tag{121}$$

and by (19) we see $\overline{\mathbf{Y}} \notin \mathcal{T}_0$. □

## C. TYPE CLASS BOUNDS

THEOREM 17. *Let $(p_1, p_2, \ldots)$ be a sequence of types defined over the finite alphabet $\mathcal{X}$ where $p_n \in \mathcal{P}_n$. Assume this sequence satisfies the following:*

1. $p_n \to p$
2. $p_n \prec\prec p, \quad \forall n$

*Then,*

$$\liminf_{n\to\infty} \frac{1}{n} \log |T(p_n)| = H(p). \tag{122}$$

PROOF. *We first show,*

$$\liminf_{n\to\infty} \frac{1}{n} \log |T(p_n)| \ge H(p). \tag{123}$$

*A sharpening of Stirling's approximation states that,*

$$n! = \sqrt{2\pi} n^{n+\frac{1}{2}} e^{-n} e^{\lambda_n}$$

*with $\frac{1}{12n+1} < \lambda_n < \frac{1}{12}$.*

*Let the empirical distribution, $p_n$ be specified by $(n_1, \ldots, n_{K_n})$. That is, if we enumerate the outcomes as $(a_1, \ldots, a_{K_n})$ we have that,*

$$p_n(a_i) = \frac{n_i}{n}.$$

*By definition $\sum_{i=1}^{K_n} n_i = n$, and from the above condition of absolute continuity we have that $K_n \le s(p)$ for all $n$, where $s(p)$ is the support of the final distribution.*

$$\begin{aligned}
\log |T(p_n)| &= \log \binom{n!}{n_1!, n_2!, \ldots, n_{K_n}!} \\
&= \log \frac{n!}{\prod_{i=1}^{K_n} n_i!} \\
&= \log \frac{\sqrt{2\pi} n^{n+\frac{1}{2}} e^{-n} e^{\lambda_n}}{\prod_{i=1}^{K_n} \left( \sqrt{2\pi} n_i^{n_i+\frac{1}{2}} e^{-n_i} e^{\lambda_{n_i}} \right)} \\
&= \log \sqrt{2\pi} n^{n+\frac{1}{2}} e^{-n} e^{\lambda_n} \\
&\qquad - \sum_{i=1}^{K_n} \log \left( \sqrt{2\pi} n_i^{n_i+\frac{1}{2}} e^{-n_i} e^{\lambda_{n_i}} \right) \\
&= \log \sqrt{2\pi} n^{n+\frac{1}{2}} e^{\lambda_n} - \sum_{i=1}^{K_n} \log \left( \sqrt{2\pi} n_i^{n_i+\frac{1}{2}} e^{\lambda_{n_i}} \right) \\
&= n \log n - \sum_{i=1}^{K_n} n_i \log n_i + \log \sqrt{2\pi n} e^{\lambda_n} \\
&\qquad - \sum_{i=1}^{K_n} \log \left( \sqrt{2\pi n_i} e^{\lambda_{n_i}} \right) \\
&= nH(p_n) + \log \sqrt{2\pi n} e^{\lambda_n} - \sum_{i=1}^{K_n} \log \left( \sqrt{2\pi n_i} e^{\lambda_{n_i}} \right) \\
&\ge nH(p_n) - \sum_{i=1}^{K_n} \log \left( \sqrt{2\pi n} e^{\frac{1}{12}} \right) \\
&\ge nH(p_n) - K_n \log \left( \sqrt{2\pi n} e^{\frac{1}{12}} \right)
\end{aligned}$$

*This implies that,*

$$\frac{1}{n} \log |T(p_n)| \ge H(p_n) - \frac{s(p)}{n} \log \left( \sqrt{2\pi n} e^{\frac{1}{12}} \right).$$

*Taking the* lim inf *of each side,*

$$\liminf_{n\to\infty} \frac{1}{n} \log |T(p_n)| \ge \liminf_{n\to\infty} H(p_n) = H(p). \tag{124}$$

*Now we have from the type class upper-bound[2] that,*

$$|T(p_n)| \le e^{nH(p_n)}, \tag{125}$$

*so*

$$\frac{1}{n} \log |T(p_n)| \le H(p_n). \tag{126}$$

*Taking the* lim sup *of each side gives,*

$$\limsup_{n\to\infty} \frac{1}{n} \log |T(p_n)| \le \limsup_{n\to\infty} H(p_n) = H(p). \tag{127}$$

*Finally combining* (124) *and* (127) *we have,*

$$H(p) \geq \limsup_{n\to\infty} \frac{1}{n} \log |T(p_n)| \tag{128}$$

$$\geq \liminf_{n\to\infty} \frac{1}{n} \log |T(p_n)| \geq H(p), \tag{129}$$

*so,*

$$H(p) = \liminf_{n\to\infty} \frac{1}{n} \log |T(p_n)| \,. \tag{130}$$

□