Upper Bounds of Channel Capacity for Bipolar Transmission Over Gauss-Markov Fading Channel

Prepared by Ming-Chun Chiang

Advisory by Prof. Po-Ning Chen

In Partial Fulfillment of the Requirements

For the Degree of Master of Science

Department of Communications Engineering

National Chiao Tung University

Hsinchu, Taiwan 300, R.O.C.

E-mail: mcchiang.cm92g@nctu.edu.tw

June, 2005

Acknowledgements

I am deeply grateful to my advisor, Prof. Po-Ning Chen, for his encouragement, support and brilliant guidance throughout this research. He always provides many constructive advices and analysis for us in philosophy and research. I greatly appreciate that he is always so patient with us that we can enjoy the research without any pressure. I feel very proud of being his master student.

I would like to thank Prof. Yunghsiang S. Han for his enthusiastic discussion, valuable suggestions and innovative ideas.

I have to thank dear lab-mates, Meng-Yu Wu, Chin-Hui Cheng, Ya-Ting Cho, Chia-Lung Wu, Hsin-Lin Hsieh, Shao-Yu Tseng and Chia-Wei Chang. They often give me precious suggestions and are concerned with me, while I am confused and depressed. Most importantly, they help me to understand myself more than ever. I wish that these two memorable years will always make me bear in mind forever.

I would like to give special thanks to my dear friends, Po-Jen Chang, Yu-Shin Chen and Yu-Sheng Lee and Shih-Wei Yeh. They not only talk with me about how to get along with people with patience and sincerity, but also encourage me to live a cheerful life. I really feel lucky to know them.

In the end, I would like to dedicate this thesis to my family, especially my mother and brother (Ming-Tsun), for their love, support and encouragement.

Contents

List of Tables

List of Figures

Abstract

In this thesis, we focus on the capacity for the time-varying Gauss-Markov fading channels. We first remark on four different definitions of channel capacities according to whether the transmitter and the receiver have or have not the channel state information (CSI). We then provide detailed derivations for the channel transition probability of the Gauss-Markov channels. As the true capacity formula for blind-CSI in both transmitter and receiver is hard to obtain, we derive its independent upper bound instead, and establish a close-form expression of the independent bound for any memory order M. Discussions are finally given by numerical evaluation of the independent bounds.

Chapter 1

Introduction

1.1 Problem formulation

Achieving high data-rate transmission at a highly mobile environment is a research challenge in wireless communications. On the one hand, the signal transmitted in air often propagates at a multipath environment so that inter-symbol interferences (ISIs) are introduced to the received signals. On the other hand, fast mobility in time makes these ISIs generally time-varying in nature, which greatly enforce the difficulty of channel estimation at the receiver.

Perhaps, the simplest stochastic model for a time-varying channel is the Gauss-Markov [4, 5]. It defines the time-varying ISIs through a discrete-time finite impulse response (FIR) miniature. The question that this thesis aims at is that what the *capacity* of a time-varying channel, like Gauss-Markov, is. The understanding of this quantity helps the researchers to be fully understood of the gap between the existing technology and the underlying limit.

There have been several publications investigating the capacity of fading channels in the literatures. The capacity of the flat Rayleigh fading channel has been studied in [7, 10] under the assumption that the state of channel fading is perfectly known to both the transmitter and the receiver. While neither the transmitter nor the receiver knows the channel state information (CSI), investigation of the capacity of memoryless Rayleigh fading channels can be found in [1]. However, seldom publications have been emerged in the capacity study of Gauss-Markov channels.

In [9], the authors addressed in the Introduction Section that perfect and imperfect CSI could have some effect on the capacity quantity. As a consequence, there can be four definitions of channel capacity according to the transmiter/receiver with/without CSI: namely, the capacity when both the transmitter and the receiver knows perfect CSI, the capacity when only the transmitter has perfect CSI, the capacity when only the receiver is perfectly CSI-aware, and the capacity when CSI is unknown to both the transmitter and the receiver. In this thesis, we will remark on these four definitions of capacity, and then turn to the derivations of bounds for the last one.

1.2 Objective of the research

After defining four definitions of channel capacity, we wish to evaluate them based on the Gauss-Markov fading channel model. Unfortunately, the problem of finding the channel input statistics that maximizes the channel input-output mutual information is beyond our management at this stage. Thus, we turn to the determination of good upper bounds for capacities. With the availability of capaicty upper bounds, performance lower bounds to bit error rates (BERs) can be obtained by means of the rate-distortion theorem and the joint source-channel coding theorem [6]. One can then evaluates the performance lower bound numerically in comparison with the simulations of his developed coding scheme. Details will be addressed in subsequent chapters.

1.3 Organization of thesis

This thesis is organized as follows. In Chapter 2, four kinds of definitions of fading channel capacities are introduced, and their operational meanings in practical communication systems are addressed. Chapter 3 introduces the system model. In Chapter 4, upper bounds of the blind-CSI capacity are derived for different channel memory orders, followed by the numerical presentation of their respective performance lower bounds. Chapter 5 concludes this thesis.

Chapter 2

The Definition of Channel Capacity

2.1 Capacity definition for memoryless additive channel

Let $X_1, X_2, \cdots$ and $Y_1, Y_2, \cdots$ denote the input and output sequences of the channel. In addition, denote the noise by $N_1, N_2, \cdots$. Then, a memoryless additive channel could be defined by:

$$Y_i = X_i + N_i, \quad i = 1, 2, \cdots, \tag{2.1}$$

where $\{X_i\}_{i=1}^{\infty}$ and $\{N_i\}_{i=1}^{\infty}$ are independent random variables, and are also independent to each other. If $\{N_i\}_{i=1}^{\infty}$ is assumed to be Gaussian distributed with white power spectral density (PSD), then a memoryless additive white Gaussian noise (AWGN) channel is established.

For discrete input X and discrete output Y, the mutual information can be written as:

$$\begin{aligned} I(X;Y) &\triangleq \sum_{x\in\mathcal{X}}\sum_{y\in\mathcal{Y}} P_{X,Y}(x,y)\cdot\log\left(\frac{P_{X,Y}(x,y)}{P_X(x)\cdot P_Y(y)}\right) \\ &= \sum_{x\in\mathcal{X}}\sum_{y\in\mathcal{Y}} P_X(x)P_{Y|X}(y|x)\cdot\log\left(\frac{P_{Y|X}(y|x)}{\sum_{x'\in\mathcal{X}} P_X(x')P_{Y|X}(y|x')}\right) \\ &= \sum_{x\in\mathcal{X}}\sum_{y\in\mathcal{Y}} P_X(x)P_N(y-x)\cdot\log\left(\frac{P_N(y-x)}{\displaystyle\sum_{x'\in\mathcal{X}} P_X(x')P_N(y-x')}\right). \end{aligned} \tag{2.2}$$

Based on this definition, the capacity-cost function for memoryless additive channel with identically distributed $\{X_i, N_i\}_{i=1}^{\infty}$ subject to input average power constraint $E[X^2] \le S$ is given by:

$$\begin{aligned} C(S) &= \max_{\{P_X\,:\,E[X^2]\le S\}} I(X;Y) \\ &= \max_{\{P_X\,:\,E[X^2]\le S\}} \sum_{x\in\mathcal{X}}\sum_{y\in\mathcal{Y}} P_X(x)P_N(y-x)\cdot\log\left(\frac{P_N(y-x)}{\displaystyle\sum_{x'\in\mathcal{X}} P_X(x')P_N(y-x')}\right). \end{aligned} \tag{2.3}$$

2.2 Capacity for time-invariant flat fading additive channel

The channel model for time-invariant flat fading additive channel is defined as:

$$Y_i = H \cdot X_i + N_i, \quad i = 1, 2, \cdots \tag{2.4}$$

where H is a time-invariant random variable, independent of $\{X_i\}_{i=1}^{\infty}$ and $\{N_i\}_{i=1}^{\infty}$. Then, the capacity-cost function for time-invariant flat fading channel given $[H = h]$ is equal to:

$$\begin{aligned}
C_h(S) &= \max_{\{P_X\,:\,E[X^2]\le S\}} I\left(X;Y|h\right) \\
&= \max_{\{P_X\,:\,E[X^2]\le S\}} \sum_{x\in\mathcal{X}}\sum_{y\in\mathcal{Y}} P_{X,Y|H}(x,y|h)\cdot\log\frac{P_{X,Y|H}(x,y|h)}{P_{X|H}(x|h)P_{Y|H}(y|h)} \\
&= \max_{\{P_X\,:\,E[X^2]\le S\}} \sum_{x\in\mathcal{X}}\sum_{y\in\mathcal{Y}} P_{X|H}(x|h)P_{Y|X,H}(y|x,h)\cdot\log\frac{P_{X|H}(x|h)P_{Y|X,H}(y|x,h)}{P_{X|H}(x|h)\sum_{x'\in\mathcal{X}}P_{X,Y|H}(x',y|h)} \\
&= \max_{\{P_X\,:\,E[X^2]\le S\}} \sum_{x\in\mathcal{X}}\sum_{y\in\mathcal{Y}} P_X(x)P_N(y-hx)\cdot\log\frac{P_N(y-hx)}{\sum_{x'\in\mathcal{X}}P_X(x')P_N(y-hx')}, \qquad (2.5)
\end{aligned}$$

where $P_{X|H}$ is replaced by P_X since X is independent of H.

Usually, the average signal-to-noise ratio (SNR) for this channel is given by:

$$\text{SNR} = \frac{E[E[H^2X^2|H]]}{E[N^2]} = E[H^2]\frac{E[X^2]}{E[N^2]}.$$

Therefore, researchers will sometimes fix $E[H^2] = 1$, and varies $E[N^2]$ to examine the system performance of their coding scheme over such a channel.

For continuous channel output alphabet, same derivations can give that:

$$C_h(S) = \max_{\{P_X\,:\,E[X^2]\le S\}} \sum_{x\in\mathcal{X}} P_X(x)\int_{\mathcal{Y}} p_N(y-Hx)\cdot\log\frac{p_N(y-Hx)}{\sum_{x'\in\mathcal{X}}P_X(x')p_N(y-Hx')}dy. \qquad (2.6)$$

Note that throughout the thesis, we will use the convention that uppercase $P_X(\cdot)$ and lowercase $p_X(\cdot)$ denote the probability mass function (pmf) and probability density function (pdf) of random variable X, respectively.

2.3 Definition of average capacity $\overline{C(S)}$ for time-invariant flat fading additive channel

Some researchers focus on the average capacity for fading channel, which is defined as [7]:

$$\overline{C(S)} \triangleq \int_{\mathcal{H}} p_H(h) \cdot C_h(S) dh. \tag{2.7}$$

The operational meaning of $\overline{C(S)}$ is that it is the underlying limit for a system in which both the transmitter and the receiver have perfect information about the channel state H. Hence, no matter what H is, the transmitter will always employ the best code that can achieve $C_h(S)$, and the receiver will use the best decoder corresponding the perfectly estimated $H = h$.

It can be easily seen that Eq. (2.7) can be rewritten as:

$$\overline{C(S)} = \int_{\mathcal{H}} p_H(h) \cdot \left[\max_{\{P_X \,:\, E[X^2] \le S\}} I(X;Y|h) \right] dh. \tag{2.8}$$

For clarity, let us give an exemplified computation for $\overline{C(S)}$.

[**Example**] Suppose that N is Gaussian distributed with zero mean and variance σ^2. Let the channel input alphabet $\mathcal{X}$ and output alphabet $\mathcal{Y}$ be the real line. Then

$$C_h(S) = \max_{\{P_X \,:\, E[X^2] \le S\}} I(X;Y|h) = \frac{1}{2} \cdot \log\left(1 + \frac{h^2 \cdot S}{\sigma^2}\right). \tag{2.9}$$

By assuming that H is Rayleigh distributed with $E[H^2] = \sigma_H^2$, we obtain that:

$$\overline{C(S)} = \int_{\Re} \frac{h}{\sigma_H^2} \cdot \exp\left(-\frac{h^2}{2\sigma_H^2}\right) \cdot \frac{1}{2} \log\left(1 + \frac{h^2 \cdot S}{\sigma^2}\right) dh \tag{2.10}$$

□

2.4 Definition of capacity $C(S)$ for time-invariant flat fading additive channel

In the previous section, $\overline{C(S)}$ is the channel capacity given that the transmitter and the receiver have perfect knowledge of channel state H. In situation where both the transmitter and the receiver are unknown of the channel state, the capacity-cost function should be given by:

$$\begin{aligned}
&C(S) \\
\triangleq\ & \max_{\{P_X\ :\ E[X^2]\le S\}} I\left(X;Y\right) \\
=\ & \max_{\{P_X\ :\ E[X^2]\le S\}} \left[h\left(Y\right)-h\left(Y|X\right)\right] \\
=\ & \max_{\{P_X\ :\ E[X^2]\le S\}} \sum_{x\in\mathcal{X}} \int_{\mathcal{Y}} p_{X,Y}\left(x,y\right)\cdot\log\left(\frac{p_{X,Y}\left(x,y\right)}{P_X\left(x\right)\cdot p_Y\left(y\right)}\right)dy \\
=\ & \max_{\{P_X\ :\ E[X^2]\le S\}} \int_{\mathcal{H}} p_H(h) \sum_{x\in\mathcal{X}} \int_{\mathcal{Y}} p_{X,Y|H}(x,y|h)\cdot\log\left(\frac{p_{Y|X}\left(y|x\right)}{p_Y\left(y\right)}\right)dydh \\
=\ & \max_{\{P_X\ :\ E[X^2]\le S\}} \int_{\mathcal{H}} p_H(h) \sum_{x\in\mathcal{X}} P_X(x) \int_{\mathcal{Y}} p_{Y|X,H}(y|x,h)\cdot\log\left(\frac{\int_{\mathcal{H}} p_H(h)p_{Y|X,H}\left(y|x,h\right)}{\int_{\mathcal{H}} p_H(h)p_{Y|H}\left(y|h\right)}\right)dydh.
\end{aligned} \tag{2.11}$$

2.5 Remarks on four definitions of channel capacities for time-invariant flat fading additive channel

The previous sections have introduced two definitions of channel capacities, namely $C(S)$ and $\overline{C(S)}$, for time-invariant flat fading additive channel. In fact, we can define four kinds of capacities according to different assumptions on the knowledge that the transmitter and the receiver have. Note that $C(S)$ corresponds to that both the transmitter and the receiver are unaware of the channel state, while $\overline{C(S)}$ is the capacity under the assumption of perfect

CSI knowledge to both the transmitter and the receiver. Their formulas are listed below.

$$C(S) \triangleq \max_{\{P_X\ :\ E[X^2]\leq S\}} \int_{\mathcal{H}} p_H(h) \sum_{x\in\mathcal{X}} P_X(x) \int_{\mathcal{Y}} p_{Y|X,H}(y|x,h) \cdot \log\left(\frac{p_{Y|X}(y)}{p_Y(y)}\right) dydh \tag{2.12}$$

$$\overline{C(S)} \triangleq \int_{\mathcal{H}} p_H(h) \max_{\{P_X\ :\ E[X^2]\leq S\}} \sum_{x\in\mathcal{X}} P_X(x) \int_{\mathcal{Y}} p_{Y|X,H}(y|x,h) \cdot \log\left(\frac{p_{Y|X,H}(y|x,h)}{p_{Y|H}(y|h)}\right) dydh \tag{2.13}$$

Now, if only the receiver knows the channel state, the transmitter cannot vary its encoding rule according the channel state. Hence, there can be only one maximization input statistics in the channel capacity formula, and the capacity formula is refined to:

$$C^{(R)}(S) \triangleq \max_{\{P_X\ :\ E[X^2]\leq S\}} \int_{\mathcal{H}} p_H(h) \sum_{x\in\mathcal{X}} P_X(x) \int_{\mathcal{Y}} p_{Y|X,H}(y|x,h) \cdot \log\left(\frac{p_{Y|X,H}(y|x,h)}{p_{Y|H}(y|h)}\right) dydh. \tag{2.14}$$

On the other hand, if only the transmitter is aware of the CSI, the capacity formula will become:

$$\overline{C^{(T)}(S)} \triangleq \int_{\mathcal{H}} p_H(h) \max_{\{P_X\ :\ E[X^2]\leq S\}} \sum_{x\in\mathcal{X}} P_X(x) \int_{\mathcal{Y}} p_{Y|X,H}(y|x,h) \cdot \log\left(\frac{p_{Y|X}(y|x)}{p_Y(y)}\right) dydh \tag{2.15}$$

In general, $C(S) \leq \overline{C^{(T)}(S)} \leq \overline{C(S)}$ and $C(S) \leq C^{(R)}(S) \leq \overline{C(S)}$.

In concept, if a perfect estimate of H is available to the receiver, then the receiver can surely take advantage of the knowledge of $p_{X,Y|H}$ and $p_{Y|H}$ at the decoding process. If, however, the receiver knows nothing about H, it can only use the average counterparts of $p_{X,Y|H}$ and $p_{Y|H}$ in its decoding process, which are exactly:

$$p_Y(y) = \int_{\mathcal{H}} p_H(h) \cdot p_{Y|H}(y|h)dh \quad \text{and} \quad p_{X,Y}(x,y) = \int_{\mathcal{H}} p_H(h) \cdot p_{X,Y|H}(y,x|h)dh. \tag{2.16}$$

This explains why (2.13) and (2.14) use $p_{Y|X,H}$ and $p_{Y|H}$ in the logarithm term, but (2.12) and (2.15) use $p_{Y|X}$ and p_Y instead.

Table 2.1: The operational characteristics of four definitions of capacities.

Capacity-cost function	TX Knows CSI?	RX Knows CSI?
$C(S)$	No	No
$\overline{C^{(T)}(S)}$	Yes	No
$C^{(R)}(S)$	No	Yes
$\overline{C(S)}$	Yes	Yes

In addition, if the transmitter has full knowledge of CSI, the encoding rule can vary according to H; hence, the maximization operation shall be inside the integral with respect p_H. In case the transmitter has no control of CSI, the transmitter can only fix the encoding rule regardless of the CSI, and hence, the maximization operation shall be placed outside the integral with respect to p_H.

These four definitions of capacities are summarized in Tab. 2.1.

2.6 Capacities for discrete input transmitted over time-invariant flat fading additive channel

In our research, we assume antipodal transmission with input alphabet $\{-s, +s\}$, where s can be any real number. Therefore, the power constraint on the input can be simplified to $E[X^2] = s^2 \leq S$ in the four definitions of capacities, and (2.12)–(2.15) can be re-formulated as that:

$$
\begin{aligned}
C(S) &\triangleq \max_{P_X \in \mathcal{P}_b(S)} \int_{\mathcal{H}} p_{\mathcal{H}}(h) \sum_{x \in \mathcal{X}} P_X(x) \int_{\mathcal{Y}} p_{Y|X,H}(y|x,h) \cdot \log\left(\frac{p_{Y|X}(y|x)}{p_Y(y)}\right) dydh \\
\overline{C^{(T)}(S)} &\triangleq \int_{\mathcal{H}} p_{\mathcal{H}}(h) \max_{P_X \in \mathcal{P}_b(S)} \sum_{x \in \mathcal{X}} P_X(x) \int_{\mathcal{Y}} p_{Y|X,H}(y|x,h) \cdot \log\left(\frac{p_{Y|X}(y|x)}{p_Y(y)}\right) dydh \\
C^{(R)}(S) &\triangleq \max_{P_X \in \mathcal{P}_b(S)} \int_{\mathcal{H}} p_{\mathcal{H}}(h) \sum_{x \in \mathcal{X}} P_X(x) \int_{\mathcal{Y}} p_{Y|X,H}(y|x,h) \cdot \log\left(\frac{p_{Y|X,H}(y|x,h)}{p_{Y|H}(y|h)}\right) dydh \\
\overline{C(S)} &\triangleq \int_{\mathcal{H}} p_{\mathcal{H}}(h) \max_{P_X \in \mathcal{P}_b(S)} \sum_{x \in \mathcal{X}} P_X(x) \int_{\mathcal{Y}} p_{Y|X,H}(y|x,h) \cdot \log\left(\frac{p_{Y|X,H}(y|x,h)}{p_{Y|H}(y|h)}\right) dydh,
\end{aligned}
$$

where $\mathcal{P}_b(S) \triangleq \{P_X \ : \ X \in \{-s, +s\} \text{ for some real } s \text{ with } s^2 \leq S\}$.

Chapter 3

System Model

3.1 Data model

In our system, we assume that binary phase shift keying (BPSK) signaling is transmitted as channel input. The probability of channel input is defined as $P_X(s) = p$ and $P_X(-s) = 1-p$, where $p \in [0, 1]$.

3.2 Channel model

A frequency-selective fast fading channel can be modelled by:

$$y_k = \boldsymbol{x}_k^T \boldsymbol{h}_k + z_k = \begin{bmatrix} x_k & x_{k-1} & \cdots & x_{k-M+1} \end{bmatrix} \begin{bmatrix} h_{k,1} \\ h_{k,2} \\ \vdots \\ h_{k,M} \end{bmatrix} + z_k, \quad k = 1, 2, \cdots, n \tag{3.1}$$

where $\boldsymbol{x}_k = \begin{bmatrix} x_k & x_{k-1} & \cdots & x_{k-M+1} \end{bmatrix}^T$ is the channel input vector consisting of the current input and the previous $(M-1)$ inputs, M is the time spread or temporal channel memory, $\boldsymbol{h}_k = \begin{bmatrix} h_{k,1} & h_{k,2} & \cdots & h_{k,M} \end{bmatrix}^T$ is a complex column vector containing the channel impulse response coefficients at time k, and z_k is the complex memoryless Gaussian noise at time k with variance $E[z_k z_k^*] = \sigma_z^2$.

Let $\boldsymbol{y} = [y_1, \cdots, y_n]^T$, $\boldsymbol{x} = [x_{2-M}, \cdots, x_n]^T$ and $\boldsymbol{z} = [z_1, \cdots, z_n]^T$ denote the received vector at the channel output, transmitted data at the channel input and complex additive noises, respectively. Then, (3.1) can be re-formulated as:

$$\boldsymbol{y} = \boldsymbol{H}\boldsymbol{x} + \boldsymbol{z}, \tag{3.2}$$

where

$$\boldsymbol{H} = \begin{bmatrix} h_{1,M} & \cdots & h_{1,1} & 0 & \cdots & 0 & 0 & 0 & 0 \\ 0 & h_{2,M} & \cdots & h_{2,1} & \cdots & 0 & 0 & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & 0 & \cdots & h_{n-1,M} & \cdots & h_{n-1,1} & 0 \\ 0 & 0 & 0 & 0 & \cdots & 0 & h_{n,M} & \cdots & h_{n,1} \end{bmatrix}. \tag{3.3}$$

Taking $M = 3$ as an example, we have:

$$y_k = \boldsymbol{x}_k^T \boldsymbol{h}_k + z_k = \begin{bmatrix} x_k & x_{k-1} & x_{k-2} \end{bmatrix} \begin{bmatrix} h_{k,1} \\ h_{k,2} \\ h_{k,3} \end{bmatrix} + z_k, \quad k = 1, \cdots, n, \tag{3.4}$$

which can be reduced to the matrix form as:

$$\begin{bmatrix} y_1 \\ y_2 \\ y_3 \\ \vdots \\ y_n \end{bmatrix} = \begin{bmatrix} h_{1,3} & h_{1,2} & h_{1,1} & 0 & \cdots & 0 & 0 & 0 & 0 \\ 0 & h_{2,3} & h_{2,2} & h_{2,1} & \cdots & 0 & 0 & 0 & 0 \\ 0 & 0 & h_{3,3} & h_{3,2} & \cdots & \vdots & \vdots & \vdots & \vdots \\ \vdots & \vdots & \vdots & \vdots & \ddots & h_{n-1,3} & h_{n-1,2} & h_{n-1,1} & 0 \\ 0 & 0 & 0 & 0 & \cdots & 0 & h_{n,3} & h_{n,2} & h_{n,1} \end{bmatrix} \begin{bmatrix} x_{-1} \\ x_0 \\ x_1 \\ \vdots \\ x_n \end{bmatrix} + \begin{bmatrix} z_1 \\ z_2 \\ z_3 \\ \vdots \\ z_n \end{bmatrix}. \tag{3.5}$$

According to (3.2), the pdf of the received vector $\boldsymbol{y}$ given $\boldsymbol{x}$ and $\boldsymbol{H}$ is equal to:

$$f\left(\boldsymbol{y}|\boldsymbol{x}, \boldsymbol{H}\right) = \frac{1}{\left(\pi\sigma_z^2\right)^n} \prod_{k=1}^{n} \exp\left(-\frac{\left|y_k - \boldsymbol{x}_k^T \boldsymbol{h}_k\right|^2}{\sigma_z^2}\right). \tag{3.6}$$

3.3 The statistics of channel state

As aforementioned, if the channel fading is known to the receiver, channel capacity can be evaluated according to either (2.13) or (2.14), depending on whether the transmitter has the information of channel fading or not. In this thesis, we focus more on the channel capacity at the situation that both the transmitter and the receiver are unaware of (and hence do not need to estimate) the channel state. In such case, we need to compute $f(\boldsymbol{y}|\boldsymbol{x})$. In principle,

$$f(\boldsymbol{y}|\boldsymbol{x}) = \int_{\mathcal{H}} f(\boldsymbol{y}|\boldsymbol{x}, \boldsymbol{H})\, f(\boldsymbol{H})\, d\boldsymbol{H}. \tag{3.7}$$

Hence, it remains to define $f(\boldsymbol{H})$, in addition to (3.6), to establish $f(\boldsymbol{y}|\boldsymbol{x})$.

A frequently used fading statistics is the Gauss-Markov. It defines the statistics of the fading through a recursive first-order Markovian equation as:

$$\boldsymbol{h}_k = \alpha \boldsymbol{h}_{k-1} + \boldsymbol{v}_k, \tag{3.8}$$

where $\boldsymbol{v}_k$ is complex, white, Gaussian distributed with mean $\boldsymbol{d}$ and covariance matrix $\boldsymbol{C}$, and α is a complex-valued scaling constant. The complex-valued constant α is a first-order Markov factor usually chosen according to $|\alpha| = e^{-\omega T}$, where T is the system sampling period and ω/π is the Doppler spread [4]. Note that the initial channel coefficient $\boldsymbol{h}_0$ is assumed to be perfectly estimated such that $\boldsymbol{h}_0$ is treated as a known constant. Based on the definition in (3.8), $f(\boldsymbol{h}_k|\boldsymbol{h}_{k-1})$ is complex Gaussian distributed with mean $(\alpha\boldsymbol{h}_{k-1} + \boldsymbol{d})$ and covariance matrix $\boldsymbol{C}$. We can therefore express $f(\boldsymbol{H})$ as:

$$\begin{aligned} f(\boldsymbol{H}) &= f(\boldsymbol{h}_1)\prod_{k=2}^{n} f(\boldsymbol{h}_k|\boldsymbol{h}_{k-1}) \\ &= \frac{1}{|\pi\boldsymbol{C}|^n}\prod_{k=1}^{n} \exp\left\{-\left(\boldsymbol{h}_k - \alpha\boldsymbol{h}_{k-1} - \boldsymbol{d}\right)^H \boldsymbol{C}^{-1}\left(\boldsymbol{h}_k - \alpha\boldsymbol{h}_{k-1} - \boldsymbol{d}\right)\right\}. \end{aligned} \tag{3.9}$$

3.4 The channel transition probability of Gauss-Markov fading channel

Substituting (3.6) and (3.9) into (3.7), we obtain the closed form of probability distribution of $f(\boldsymbol{y}|\boldsymbol{x})$ in Gauss-Markov fading as:

$$
\begin{aligned}
f(\boldsymbol{y}|\boldsymbol{x}) &= \frac{1}{(\pi\sigma_z^2)^n |\pi\boldsymbol{C}|^n} \int\limits_{\boldsymbol{\mathcal{H}}} \prod_{k=1}^{n} \left[\exp\left(-\frac{\left|y_k - \boldsymbol{x}_k^T \boldsymbol{h}_k\right|^2}{\sigma_z^2} \right) \right. \\
&\qquad \left. \exp\left(-(\boldsymbol{h}_k - \alpha\boldsymbol{h}_{k-1} - \boldsymbol{d})^H \boldsymbol{C}^{-1} (\boldsymbol{h}_k - \alpha\boldsymbol{h}_{k-1} - \boldsymbol{d}) \right) \right] d\boldsymbol{H} \\
&= \frac{1}{(\pi\sigma_z^2)^n |\pi\boldsymbol{C}|^n} \left[\prod_{k=1}^{n} |\pi\boldsymbol{G}_k| \exp\left(-\frac{|y_k|^2}{\sigma_z^2} \right) \right] \\
&\quad \times \left[\prod_{k=1}^{n-1} \exp\left[(\boldsymbol{q}_k - \alpha^*\boldsymbol{C}^{-1}\boldsymbol{d})^H \boldsymbol{G}_k (\boldsymbol{q}_k - \alpha^*\boldsymbol{C}^{-1}\boldsymbol{d}) - \boldsymbol{d}^H\boldsymbol{C}^{-1}\boldsymbol{d} \right] \right] \\
&\quad \times \exp\left[\boldsymbol{q}_n^H \boldsymbol{G}_n \boldsymbol{q}_n - (\alpha\boldsymbol{h}_0 + \boldsymbol{d})^H \boldsymbol{C}^{-1} (\alpha\boldsymbol{h}_0 + \boldsymbol{d}) \right],
\end{aligned} \tag{3.10}
$$

where

$$
\boldsymbol{G}_k = \begin{cases}
\left(\dfrac{\boldsymbol{x}_1^*\boldsymbol{x}_1^T}{\sigma_z^2} + (1+|\alpha|^2)\boldsymbol{C}^{-1} \right)^{-1}, & \text{if } k = 1 \\
\left(\dfrac{\boldsymbol{x}_k^*\boldsymbol{x}_k^T}{\sigma_z^2} + (1+|\alpha|^2)\boldsymbol{C}^{-1} - |\alpha|^2\boldsymbol{C}^{-1}\boldsymbol{G}_{k-1}\boldsymbol{C}^{-1} \right)^{-1}, & \text{if } 1 < k < n \\
\left(\dfrac{\boldsymbol{x}_n^*\boldsymbol{x}_n^T}{\sigma_z^2} + \boldsymbol{C}^{-1} - |\alpha|^2\boldsymbol{C}^{-1}\boldsymbol{G}_{n-1}\boldsymbol{C}^{-1} \right)^{-1}, & \text{if } k = n
\end{cases} \tag{3.11}
$$

and

$$
\boldsymbol{q}_k = \begin{cases}
\dfrac{y_1\boldsymbol{x}_1^*}{\sigma_z^2} + \boldsymbol{C}^{-1}\boldsymbol{d} + \alpha\boldsymbol{C}^{-1}\boldsymbol{h}_0, & \text{if } k = 1 \\
\dfrac{y_k\boldsymbol{x}_k^*}{\sigma_z^2} + \boldsymbol{C}^{-1}\boldsymbol{d} + \alpha\boldsymbol{C}^{-1}\boldsymbol{G}_{k-1}\boldsymbol{q}_{k-1} - |\alpha|^2\boldsymbol{C}^{-1}\boldsymbol{G}_{k-1}\boldsymbol{C}^{-1}\boldsymbol{d}, & \text{if } 1 < k \leq n
\end{cases} \tag{3.12}
$$

The detail to derive the above result is described in Appendix A.

Chapter 4

Upper Bounds of the Blind-CSI Capacity for Gauss-Markov Fading Channel

4.1 Independent bound

In this section, we focus on the independent bound of the channel capacity defined similarly in (2.11), i.e. the capacity without knowing the channel fading both at the transmitter and at the receiver. Here, we presume that the channel is reset every n symbols. Therefore,

$$C(S) \triangleq \frac{1}{n} \max_{\{P_{\boldsymbol{x}} \,:\, \frac{1}{n}\mathrm{tr}(E[\boldsymbol{x}^H \boldsymbol{x}]) \le S\}} I\left(\boldsymbol{x}; \boldsymbol{y}\right), \tag{4.1}$$

where, as defined in the previous chapter, $\boldsymbol{x} = [x_{2-M}, \cdots, x_n]^T$, and "tr(·)" denotes the traverse of a matrix. Since $x_{2-M}, \cdots, x_0$ are usually assumed deterministic zero (hence, they consume no power), and are nothing to do with the information transmission, we will abuse notation $\boldsymbol{x}$ in this chapter as $\boldsymbol{x} = [x_1, x_2, \cdots, x_n]^T$ without ambiguity. Thus, we can equivalently replace $\boldsymbol{x}$ by X^n to result in:

$$C(S) \triangleq \frac{1}{n} \max_{\{P_{X^n} \,:\, \frac{1}{n}\mathrm{tr}(\boldsymbol{\Lambda}_X) \le S\}} I\left(X^n; Y^n\right), \tag{4.2}$$

where $\mathbf{\Lambda}_X$ is the expectation matrix of

$$\begin{bmatrix} X_1^* & X_2^* & \cdots & X_n^* \end{bmatrix} \begin{bmatrix} X_1 \\ X_2 \\ \vdots \\ X_n \end{bmatrix}.$$

By elementary information theory operation [3],

$$\begin{aligned} \mathrm{I}\left(Y^n;X^n\right) &= \int_{\mathcal{Y}^n}\int_{\mathcal{X}^n} p_{X^n,Y^n}\left(x^n,y^n\right)\log\left[\frac{p_{X^n,Y^n}\left(x^n,y^n\right)}{p_{Y^n}\left(y^n\right)p_{X^n}\left(x^n\right)}\right]dx^n dy^n \\ &= \int_{\mathcal{Y}^n}\int_{\mathcal{X}^n} p_{X^n,Y^n}\left(x^n,y^n\right)\log\left[\frac{p_{Y^n|X^n}\left(y^n|x^n\right)}{p_{Y^n}\left(y^n\right)}\right]dx^n dy^n \\ &= h(Y^n|X^n) - h(X^n), \end{aligned} \tag{4.3}$$

where $h(\cdot)$ represents the differential entropy operation. For notational convenience, we will respectively abbreviate $p_{X^n,Y^n}(x^n, y^n)$ and $p_{Y^n}(y^n)$ by $f(\boldsymbol{x}, \boldsymbol{y})$ and $f(\boldsymbol{x})$ as we did in the previous chapter.

Although in the previous chapter, the closed form expression for $f(\boldsymbol{y}|\boldsymbol{x})$ is established, it is still hard to evaluate the capacity in (4.2). An upper bound based on the simple information-theoretical independent bound, however, can be easily obtained. The independent bound for mutual information is given by:

$$I\left(X^n;Y^n\right) \le I\left(X_1;Y_1\right) + \cdots + I\left(X_n;Y_n\right). \tag{4.4}$$

We then derive:

$$\begin{aligned} & \max_{\{P_{\boldsymbol{X}}:\frac{1}{n}\sum_{i=1}^n E[X_i^2]\le S\}} I\left(X^n;Y^n\right) \\ \le & \max_{\{P_{\boldsymbol{X}}:\frac{1}{n}\sum_{i=1}^n E[X_i^2]\le S\}} \left[I\left(X_1;Y_1\right) + \cdots + I\left(X_n;Y_n\right)\right] \\ = & \max_{\{P_{\boldsymbol{X}}:(\forall\ i)E[X_i^2]\le S\}} \left[I\left(X_1;Y_1\right) + \cdots + I\left(X_n;Y_n\right)\right] \qquad (4.5) \\ \le & \max_{\{P_{\boldsymbol{X}}:(\forall\ i)E[X_i^2]\le S\}} I\left(X_1;Y_1\right) + \cdots + \max_{\{P_{\boldsymbol{X}}:(\forall\ i)E[X_i^2]\le S\}} I\left(X_n;Y_n\right), \qquad (4.6) \end{aligned}$$

where (4.5) holds since in our system setting, every $E[X_i^2]$ is equal to s^2 due to $x_i \in \{-s, +s\}$ for every i. Note that $\sum_{i=1}^n E[X_i^2] = \mathrm{tr}\left(\mathbf{\Lambda}_X\right)$.

Let $C_k(S)$ denote the maximum of mutual information $I(X_k;Y_k)$ under input power constraints $E[X_k^2] \leq S$ for every k. Then,

$$C(S) \leq \frac{1}{n}\left[C_1(S) + \cdots + C_n(S)\right]. \tag{4.7}$$

4.2 Derivation of $f(y_k|\boldsymbol{x}_k)$ and $f(y_k)$

In order to evaluate the independent bound of channel capacity, $f(y_k|\boldsymbol{x}_k)$ and $f(y_k)$ have to be obtained first. The approach to obtain them can be described as follows.

First of all, we derive:

$$\begin{aligned}
f\left(y_k|\boldsymbol{x},\boldsymbol{h}_1,\cdots,\boldsymbol{h}_n\right) &\triangleq \int_{\mathcal{C}}\cdots\int_{\mathcal{C}} f\left(\boldsymbol{y}|\boldsymbol{x},\boldsymbol{h}_1,\cdots,\boldsymbol{h}_n\right)dy_1\cdots dy_{k-1}dy_{k+1}\cdots dy_n \\
&= \frac{1}{(\pi\sigma_z^2)^n}\int_{\mathcal{C}}\cdots\int_{\mathcal{C}}\prod_{m=1}^{n}\exp\left(-\frac{\left|y_m-\boldsymbol{x}_m^T\boldsymbol{h}_m\right|^2}{\sigma_z^2}\right)dy_1\cdots dy_{k-1}dy_{k+1}\cdots dy_n \\
&= \frac{1}{\pi\sigma_z^2}\exp\left(-\frac{\left|y_k-\boldsymbol{x}_k^T\boldsymbol{h}_k\right|^2}{\sigma_z^2}\right) \\
&= f\left(y_k|\boldsymbol{x}_k,\boldsymbol{h}_k\right),
\end{aligned} \tag{4.8}$$

where $\mathcal{C}$ denote the set of all complex numbers. Then, we notice that:

$$\begin{aligned}
f\left(y_k|\boldsymbol{x}\right) &= \int_{\mathcal{C}^M}\cdots\int_{\mathcal{C}^M} f\left(y_k|\boldsymbol{x},\boldsymbol{h}_1,\cdots,\boldsymbol{h}_n\right)f\left(\boldsymbol{h}_n|\boldsymbol{h}_{n-1}\right)\cdots f\left(\boldsymbol{h}_2|\boldsymbol{h}_1\right)f\left(\boldsymbol{h}_1\right)d\boldsymbol{h}_1\cdots d\boldsymbol{h}_n \\
&= \int_{\mathcal{C}^M}\cdots\int_{\mathcal{C}^M} f\left(y_k|\boldsymbol{x}_k,\boldsymbol{h}_k\right)f\left(\boldsymbol{h}_n|\boldsymbol{h}_{n-1}\right)\cdots f\left(\boldsymbol{h}_2|\boldsymbol{h}_1\right)f\left(\boldsymbol{h}_1\right)d\boldsymbol{h}_1\cdots d\boldsymbol{h}_n \\
&= \int_{\mathcal{C}^M} f\left(y_k|\boldsymbol{x}_k,\boldsymbol{h}_k\right)\left[\int_{\mathcal{C}^M} f\left(\boldsymbol{h}_k|\boldsymbol{h}_{k-1}\right)\cdots\int_{\mathcal{C}^M} f\left(\boldsymbol{h}_2|\boldsymbol{h}_1\right)f\left(\boldsymbol{h}_1\right)d\boldsymbol{h}_1\cdots d\boldsymbol{h}_{k-1}\right]d\boldsymbol{h}_k \\
&= \int_{\mathcal{C}^M} f\left(y_k|\boldsymbol{x}_k,\boldsymbol{h}_k\right)f\left(\boldsymbol{h}_k\right)d\boldsymbol{h}_k
\end{aligned} \tag{4.9}$$

It remains to find $f(\boldsymbol{h}_k)$ for $k=1,\cdots,n$.

By the system model,

$$f(\boldsymbol{h}_1) = \frac{1}{|\pi\boldsymbol{C}|}\exp\Big(-\left(\boldsymbol{h}_1-\alpha\boldsymbol{h}_0-\boldsymbol{d}\right)^H\boldsymbol{C}^{-1}\left(\boldsymbol{h}_1-\alpha\boldsymbol{h}_0-\boldsymbol{d}\right)\Big). \tag{4.10}$$

Thus, $\boldsymbol{h}_1$ is complex Gaussian distributed with mean $\boldsymbol{d}_1 \triangleq \alpha \boldsymbol{h}_0 + \boldsymbol{d}$ and covariance matrix $\boldsymbol{D}_1 \triangleq \boldsymbol{C}$. Next,

$$
\begin{aligned}
f(\boldsymbol{h}_2) &= \int_{\mathcal{C}^M} f(\boldsymbol{h}_2|\boldsymbol{h}_1) f(\boldsymbol{h}_1)\, d\boldsymbol{h}_1 \\
&= \frac{1}{|\pi \boldsymbol{C}|\,|\pi \boldsymbol{D}_1|} \int_{\mathcal{C}^M} \exp\Big(-(\boldsymbol{h}_2 - \alpha \boldsymbol{h}_1 - \boldsymbol{d})^H \boldsymbol{C}^{-1} (\boldsymbol{h}_2 - \alpha \boldsymbol{h}_1 - \boldsymbol{d})\Big) \\
&\qquad \exp\Big(-(\boldsymbol{h}_1 - \boldsymbol{d}_1)^H \boldsymbol{D}_1^{-1} (\boldsymbol{h}_1 - \boldsymbol{d}_1)\Big)\, d\boldsymbol{h}_1.
\end{aligned} \tag{4.11}
$$

The negative exponent inside the integral in (4.11) is:

$$
\begin{aligned}
& (\boldsymbol{h}_2 - \alpha \boldsymbol{h}_1 - \boldsymbol{d})^H \boldsymbol{C}^{-1} (\boldsymbol{h}_2 - \alpha \boldsymbol{h}_1 - \boldsymbol{d}) + (\boldsymbol{h}_1 - \boldsymbol{d}_1)^H \boldsymbol{D}_1^{-1} (\boldsymbol{h}_1 - \boldsymbol{d}_1) \\
= \; & ((\boldsymbol{h}_2 - \boldsymbol{d}) - \alpha \boldsymbol{h}_1)^H \boldsymbol{C}^{-1} ((\boldsymbol{h}_2 - \boldsymbol{d}) - \alpha \boldsymbol{h}_1) + (\boldsymbol{h}_1 - \boldsymbol{d}_1)^H \boldsymbol{D}_1^{-1} (\boldsymbol{h}_1 - \boldsymbol{d}_1) \\
= \; & |\alpha|^2 \boldsymbol{h}_1^H \boldsymbol{C}^{-1} \boldsymbol{h}_1 - \alpha (\boldsymbol{h}_2 - \boldsymbol{d})^H \boldsymbol{C}^{-1} \boldsymbol{h}_1 - \alpha^H \boldsymbol{h}_1^H \boldsymbol{C}^{-1} (\boldsymbol{h}_2 - \boldsymbol{d}) + (\boldsymbol{h}_2 - \boldsymbol{d})^H \boldsymbol{C}^{-1} (\boldsymbol{h}_2 - \boldsymbol{d}) \\
& + \boldsymbol{h}_1^H \boldsymbol{D}_1^{-1} \boldsymbol{h}_1 - \boldsymbol{d}_1^H \boldsymbol{D}_1^{-1} \boldsymbol{h}_1 - \boldsymbol{h}_1^H \boldsymbol{D}_1^{-1} \boldsymbol{d}_1 + \boldsymbol{d}_1^H \boldsymbol{D}_1^{-1} \boldsymbol{d}_1 \\
= \; & \boldsymbol{h}_1^H (|\alpha|^2 \boldsymbol{C}^{-1} + \boldsymbol{D}_1^{-1}) \boldsymbol{h}_1 + \boldsymbol{d}_1^H \boldsymbol{D}_1^{-1} \boldsymbol{d}_1 + (\boldsymbol{h}_2 - \boldsymbol{d})^H \boldsymbol{C}^{-1} (\boldsymbol{h}_2 - \boldsymbol{d}) \\
& - \Big[\alpha (\boldsymbol{h}_2 - \boldsymbol{d})^H \boldsymbol{C}^{-1} + \boldsymbol{d}_1^H \boldsymbol{D}_1^{-1}\Big] \boldsymbol{h}_1 - \boldsymbol{h}_1^H \big[\alpha^H \boldsymbol{C}^{-1} (\boldsymbol{h}_2 - \boldsymbol{d}) + \boldsymbol{D}_1^{-1} \boldsymbol{d}_1\big] \\
= \; & \big[\boldsymbol{h}_1^H \boldsymbol{\Omega}_1^{-1} \boldsymbol{h}_1 - \boldsymbol{\mu}_1^H \boldsymbol{\Omega}_1^{-1} \boldsymbol{h}_1 - \boldsymbol{h}_1^H \boldsymbol{\Omega}_1^{-1} \boldsymbol{\mu}_1\big] + \boldsymbol{d}_1^H \boldsymbol{D}_1^{-1} \boldsymbol{d}_1 + (\boldsymbol{h}_2 - \boldsymbol{d})^H \boldsymbol{C}^{-1} (\boldsymbol{h}_2 - \boldsymbol{d}) \\
= \; & (\boldsymbol{h}_1 - \boldsymbol{\mu}_1)^H \boldsymbol{\Omega}_1^{-1} (\boldsymbol{h}_1 - \boldsymbol{\mu}_1) - \boldsymbol{\mu}_1^H \boldsymbol{\Omega}_1^{-1} \boldsymbol{\mu}_1 + \boldsymbol{d}_1^H \boldsymbol{D}_1^{-1} \boldsymbol{d}_1 + (\boldsymbol{h}_2 - \boldsymbol{d})^H \boldsymbol{C}^{-1} (\boldsymbol{h}_2 - \boldsymbol{d}),
\end{aligned} \tag{4.12}
$$

where $\boldsymbol{\mu}_1 \triangleq \boldsymbol{\Omega}_1 \left[\alpha^H \boldsymbol{C}^{-1} (\boldsymbol{h}_2 - \boldsymbol{d}) + \boldsymbol{D}_1^{-1} \boldsymbol{d}_1\right]$ and $\boldsymbol{\Omega}_1^{-1} \triangleq |\alpha|^2 \boldsymbol{C}^{-1} + \boldsymbol{D}_1^{-1}$. In (4.12), only the first term is relevance to the integrater $\boldsymbol{h}_1$. Hence, taking the first term into (4.11), and integrating with respect to $\boldsymbol{h}_1$ yields:

$$
\int_{\mathcal{C}^M} \exp\left[-(\boldsymbol{h}_1 - \boldsymbol{\mu}_1)^H \boldsymbol{\Omega}_1^{-1} (\boldsymbol{h}_1 - \boldsymbol{\mu}_1)\right] d\boldsymbol{h}_1 = |\pi \boldsymbol{\Omega}_1|\,. \tag{4.13}
$$

Let $\beta_k \triangleq 1 + |\alpha|^2 + |\alpha|^4 + \cdots + |\alpha|^{2k}$. By observing that $\boldsymbol{C}^{-1} = \beta_0 \boldsymbol{D}_1^{-1}$ and $\boldsymbol{\Omega}_1^{-1} = \beta_1 \boldsymbol{D}_1^{-1}$

and $\boldsymbol{\mu}_1 = \frac{1}{\beta_1}\left[\alpha^H \beta_0 (\boldsymbol{h}_2 - \boldsymbol{d}) + \boldsymbol{d}_1\right]$, the remaining terms in (4.12) put:

$$\begin{aligned}
& \boldsymbol{d}_1^H \boldsymbol{D}_1^{-1} \boldsymbol{d}_1 + (\boldsymbol{h}_2 - \boldsymbol{d})^H \boldsymbol{C}^{-1} (\boldsymbol{h}_2 - \boldsymbol{d}) - \boldsymbol{\mu}_1^H \boldsymbol{\Omega}_1^{-1} \boldsymbol{\mu}_1 \\
= & \ \boldsymbol{d}_1^H \boldsymbol{D}_1^{-1} \boldsymbol{d}_1 + \beta_0 (\boldsymbol{h}_2 - \boldsymbol{d})^H \boldsymbol{D}_1^{-1} (\boldsymbol{h}_2 - \boldsymbol{d}) \\
& -\frac{1}{\beta_1}\left[\alpha^H \beta_0 (\boldsymbol{h}_2 - \boldsymbol{d}) + \boldsymbol{d}_1\right]^H \boldsymbol{D}_1^{-1}\left[\alpha^H \beta_0 (\boldsymbol{h}_2 - \boldsymbol{d}) + \boldsymbol{d}_1\right] \\
= & \ \boldsymbol{d}_1^H \boldsymbol{D}_1^{-1} \boldsymbol{d}_1 + \beta_0 (\boldsymbol{h}_2 - \boldsymbol{d})^H \boldsymbol{D}_1^{-1} (\boldsymbol{h}_2 - \boldsymbol{d}) \\
& -\frac{|\alpha|^2 \beta_0^2}{\beta_1} (\boldsymbol{h}_2 - \boldsymbol{d})^H \boldsymbol{D}_1^{-1} (\boldsymbol{h}_2 - \boldsymbol{d}) - \frac{\alpha \beta_0}{\beta_1} (\boldsymbol{h}_2 - \boldsymbol{d})^H \boldsymbol{D}_1^{-1} \boldsymbol{d}_1 \\
& -\frac{\alpha^H \beta_0}{\beta_1} \boldsymbol{d}_1^H \boldsymbol{D}_1^{-1} (\boldsymbol{h}_2 - \boldsymbol{d}) - \frac{1}{\beta_1} \boldsymbol{d}_1^H \boldsymbol{D}_1^{-1} \boldsymbol{d}_1 \\
= & \ \frac{\beta_0}{\beta_1} (\boldsymbol{h}_2 - \boldsymbol{d})^H \boldsymbol{D}_1^{-1} (\boldsymbol{h}_2 - \boldsymbol{d}) - \frac{\alpha \beta_0}{\beta_1} (\boldsymbol{h}_2 - \boldsymbol{d})^H \boldsymbol{D}_1^{-1} \boldsymbol{d}_1 - \frac{\alpha^H \beta_0}{\beta_1} \boldsymbol{d}_1^H \boldsymbol{D}_1^{-1} (\boldsymbol{h}_2 - \boldsymbol{d}) \\
& +\frac{|\alpha|^2 \beta_0}{\beta_1} \boldsymbol{d}_1^H \boldsymbol{D}_1^{-1} \boldsymbol{d}_1 \\
= & \ (\boldsymbol{h}_2 - \boldsymbol{d} - \alpha \boldsymbol{d}_1)^H \boldsymbol{D}_2^{-1} (\boldsymbol{h}_2 - \boldsymbol{d} - \alpha \boldsymbol{d}_1),
\end{aligned}$$

where $\boldsymbol{D}_2 \triangleq (\beta_1/\beta_0)\boldsymbol{D}_1$. Summarizing the above derivation, we obtain:

$$f(\boldsymbol{h}_2) = \frac{1}{|\pi \boldsymbol{D}_2|} \exp\left[-(\boldsymbol{h}_2 - \alpha \boldsymbol{d}_1 - \boldsymbol{d})^H \boldsymbol{D}_2^{-1} (\boldsymbol{h}_2 - \alpha \boldsymbol{d}_1 - \boldsymbol{d})\right], \tag{4.14}$$

and hence, $\boldsymbol{h}_2$ is complex Gaussian distributed with mean $\boldsymbol{d}_2 \triangleq \alpha \boldsymbol{d}_1 + \boldsymbol{d}$ and covariance matrix $\boldsymbol{D}_2$.

We now turn to the derivation of $f(\boldsymbol{h}_3)$. Observe that

$$\begin{aligned}
f(\boldsymbol{h}_3) &= \int_{\mathcal{C}^M} f(\boldsymbol{h}_3|\boldsymbol{h}_2) f(\boldsymbol{h}_2) d\boldsymbol{h}_2 \\
&= \frac{1}{|\pi \boldsymbol{C}| |\pi \boldsymbol{D}_2|} \int_{\mathcal{C}^M} \exp\left(-(\boldsymbol{h}_3 - \alpha \boldsymbol{h}_2 - \boldsymbol{d})^H \boldsymbol{C}^{-1} (\boldsymbol{h}_3 - \alpha \boldsymbol{h}_2 - \boldsymbol{d})\right) \\
& \qquad \exp\left[-(\boldsymbol{h}_2 - \boldsymbol{d}_2)^H \boldsymbol{D}_2^{-1} (\boldsymbol{h}_2 - \boldsymbol{d}_2)\right] d\boldsymbol{h}_2,
\end{aligned} \tag{4.15}$$

which has the same form as in (4.11). So, the negative exponent inside the integral in (4.15) is likewisely equal to:

$$(\boldsymbol{h}_3 - \alpha \boldsymbol{h}_2 - \boldsymbol{d})^H \boldsymbol{C}^{-1} (\boldsymbol{h}_3 - \alpha \boldsymbol{h}_2 - \boldsymbol{d}) + (\boldsymbol{h}_2 - \boldsymbol{d}_2)^H \boldsymbol{D}_2^{-1} (\boldsymbol{h}_2 - \boldsymbol{d}_2)$$

$$= (\boldsymbol{h}_2 - \boldsymbol{\mu}_2)^H \boldsymbol{\Omega}_2^{-1} (\boldsymbol{h}_2 - \boldsymbol{\mu}_2) - \boldsymbol{\mu}_2^H \boldsymbol{\Omega}_2^{-1} \boldsymbol{\mu}_2 + \boldsymbol{d}_2^H \boldsymbol{D}_2^{-1} \boldsymbol{d}_2 + (\boldsymbol{h}_3 - \boldsymbol{d})^H \boldsymbol{C}^{-1} (\boldsymbol{h}_3 - \boldsymbol{d}), \quad (4.16)$$

where $\boldsymbol{\mu}_2 \triangleq \boldsymbol{\Omega}_2 \left[\alpha^H \boldsymbol{C}^{-1} (\boldsymbol{h}_3 - \boldsymbol{d}) + \boldsymbol{D}_2^{-1} \boldsymbol{d}_2\right]$ and $\boldsymbol{\Omega}_2^{-1} \triangleq |\alpha|^2 \boldsymbol{C}^{-1} + \boldsymbol{D}_2^{-1}$. By observing that $\boldsymbol{C}^{-1} = \beta_1 \boldsymbol{D}_2^{-1}$ and $\boldsymbol{\Omega}_2^{-1} = \beta_2 \boldsymbol{D}_2^{-1}$ and $\boldsymbol{\mu}_2 = \frac{1}{\beta_2} \left[\alpha^H \beta_1 (\boldsymbol{h}_3 - \boldsymbol{d}) + \boldsymbol{d}_2\right]$, the last three terms in (4.16), as the first term is removed due to the integration with respect to $\boldsymbol{h}_2$, put:

$$\boldsymbol{d}_2^H \boldsymbol{D}_2^{-1} \boldsymbol{d}_2 + (\boldsymbol{h}_3 - \boldsymbol{d})^H \boldsymbol{C}^{-1} (\boldsymbol{h}_3 - \boldsymbol{d}) - \boldsymbol{\mu}_2^H \boldsymbol{\Omega}_2^{-1} \boldsymbol{\mu}_2 = (\boldsymbol{h}_3 - \boldsymbol{d} - \alpha \boldsymbol{d}_2)^H \boldsymbol{D}_3^{-1} (\boldsymbol{h}_3 - \boldsymbol{d} - \alpha \boldsymbol{d}_2),$$

where $\boldsymbol{D}_3 = (\beta_2/\beta_1) \boldsymbol{D}_2$. Hence, $\boldsymbol{h}_3$ is complex Gaussian distributed with mean $\boldsymbol{d}_3 \triangleq \alpha \boldsymbol{d}_2 + \boldsymbol{d}$ and covariance matrix $\boldsymbol{D}_3$.

By repeating the above procedure, we conclude that $\boldsymbol{h}_k$ is complex Gaussian distributed with mean $\boldsymbol{d}_k = \alpha \boldsymbol{d}_{k-1} + \boldsymbol{d}$ and covariance matrix $\boldsymbol{D}_k = (\beta_{k-1}/\beta_{k-2}) \boldsymbol{D}_{k-1}$ with the initial values $\boldsymbol{d}_0 = \boldsymbol{h}_0$, $\boldsymbol{D}_0 = \boldsymbol{C}$ and $\beta_k = 1 + |\alpha|^2 + \cdots + |\alpha|^{2k}$ with $\beta_{-1} \triangleq 1$. As a result, $\boldsymbol{d}_k = \alpha^k \boldsymbol{h}_0 + \frac{1-\alpha^k}{1-\alpha} \boldsymbol{d}$ and $\boldsymbol{D}_k = \beta_{k-1} \boldsymbol{C} = \frac{1-|\alpha|^{2k}}{1-|\alpha|^2} \boldsymbol{C}$.

After obtaining the probability distribution of the fading coefficient $\boldsymbol{h}_k$, the probability distribution of $f(y_k|\boldsymbol{x})$ can be calculated by integrating the product of $f(y_k|\boldsymbol{x}, \boldsymbol{h}_k)$ and $f(\boldsymbol{h}_k)$ as follows:

$$f(y_k|\boldsymbol{x}) = \int_{\mathcal{C}^M} f(y_k|\boldsymbol{x}, \boldsymbol{h}_k) f(\boldsymbol{h}_k) \, d\boldsymbol{h}_k$$
$$= \frac{1}{\pi \sigma_z^2} \frac{1}{|\pi \boldsymbol{D}_k|} \int_{\mathcal{C}^M} \exp\left(-\frac{\left|y_k - \boldsymbol{x}_k^T \boldsymbol{h}_k\right|^2}{\sigma_z^2}\right) \exp\left[-(\boldsymbol{h}_k - \boldsymbol{d}_k)^H \boldsymbol{D}_k^{-1} (\boldsymbol{h}_k - \boldsymbol{d}_k)\right] d\boldsymbol{h}_k. \quad (4.17)$$

The negative exponent in (4.17) is equal to:

$$\begin{aligned}
& \frac{\left(y_k - \boldsymbol{x}_k^T \boldsymbol{h}_k\right)^H \left(y_k - \boldsymbol{x}_k^T \boldsymbol{h}_k\right)}{\sigma_z^2} + (\boldsymbol{h}_k - \boldsymbol{d}_k)^H \boldsymbol{D}_k^{-1} \left(\boldsymbol{h}_k^H - \boldsymbol{d}_k\right) \\
= & \frac{|y_k|^2}{\sigma_z^2} - \frac{y_k^H \boldsymbol{x}_k^T}{\sigma_z^2} \boldsymbol{h}_k - \boldsymbol{h}_k^H \frac{\boldsymbol{x}_k^* y_k}{\sigma_z^2} + \boldsymbol{h}_k^H \frac{\boldsymbol{x}_k^* \boldsymbol{x}_k^T}{\sigma_z^2} \boldsymbol{h}_k + \boldsymbol{h}_k^H \boldsymbol{D}_k^{-1} \boldsymbol{h}_k - \boldsymbol{h}_k^H \boldsymbol{D}_k^{-1} \boldsymbol{d}_k \\
& - \boldsymbol{d}_k^H \boldsymbol{D}_k^{-1} \boldsymbol{h}_k + \boldsymbol{d}_k^H \boldsymbol{D}_k^{-1} \boldsymbol{d}_k \\
= & \boldsymbol{h}_k^H \left(\frac{\boldsymbol{x}_k^* \boldsymbol{x}_k^T}{\sigma_z^2} + \boldsymbol{D}_k^{-1} \right) \boldsymbol{h}_k - \boldsymbol{h}_k^H \left(\frac{\boldsymbol{x}_k^* y_k}{\sigma_z^2} + \boldsymbol{D}_k^{-1} \boldsymbol{d}_k \right) - \left(\frac{\boldsymbol{x}_k^* y_k}{\sigma_z^2} + \boldsymbol{D}_k^{-1} \boldsymbol{d}_k \right)^H \boldsymbol{h}_k \\
& + \boldsymbol{d}_k^H \boldsymbol{D}_k^{-1} \boldsymbol{d}_k + \frac{|y_k|^2}{\sigma_z^2} \\
= & \boldsymbol{h}_k^H \boldsymbol{E}_k^{-1} \boldsymbol{h}_k - \boldsymbol{h}_k^H \boldsymbol{E}_k^{-1} \boldsymbol{e}_k - \boldsymbol{e}_k^H \boldsymbol{E}_k^{-1} \boldsymbol{h}_k + \boldsymbol{d}_k^H \boldsymbol{D}_k^{-1} \boldsymbol{d}_k + \frac{|y_k|^2}{\sigma_z^2} \\
= & (\boldsymbol{h}_k - \boldsymbol{e}_k)^H \boldsymbol{E}_k^{-1} (\boldsymbol{h}_k - \boldsymbol{e}_k) - \boldsymbol{e}_k^H \boldsymbol{E}_k^{-1} \boldsymbol{e}_k + \boldsymbol{d}_k^H \boldsymbol{D}_k^{-1} \boldsymbol{d}_k + \frac{|y_k|^2}{\sigma_z^2}
\end{aligned} \tag{4.18}$$

where $\boldsymbol{E}_k^{-1} \triangleq \left(\frac{\boldsymbol{x}_k^* \boldsymbol{x}_k^T}{\sigma_z^2} + \boldsymbol{D}_k^{-1} \right)$ and $\boldsymbol{e}_k \triangleq \boldsymbol{E}_k \left(\frac{\boldsymbol{x}_k^* y_k}{\sigma_z^2} + \boldsymbol{D}_k^{-1} \boldsymbol{d}_k \right)$. Therefore,

$$f(y_k|\boldsymbol{x}_k) = \frac{|\pi \boldsymbol{E}_k|}{\pi \sigma_z^2 |\pi \boldsymbol{D}_k|} \exp \left\{ \boldsymbol{e}_k^H \boldsymbol{E}_k^{-1} \boldsymbol{e}_k - \boldsymbol{d}_k^H \boldsymbol{D}_k^{-1} \boldsymbol{d}_k - \frac{|y_k|^2}{\sigma_z^2} \right\}. \tag{4.19}$$

It remains to simplify the exponent of the above expression.

Referring to [11],

$$\left(\boldsymbol{A} + \boldsymbol{a}\boldsymbol{b}^H\right)^{-1} = \boldsymbol{A}^{-1} - \frac{\boldsymbol{A}^{-1} \boldsymbol{a} \boldsymbol{b}^H \boldsymbol{A}^{-1}}{1 + \boldsymbol{b}^H \boldsymbol{A}^{-1} \boldsymbol{a}} \tag{4.20}$$

for any k-by-k matrix $\boldsymbol{A}$, and $k \times 1$ vectors $\boldsymbol{a}$ and $\boldsymbol{b}$. Hence,

$$\boldsymbol{E}_k = \left(\boldsymbol{D}_k^{-1} + \frac{\boldsymbol{x}_k^* \boldsymbol{x}_k^T}{\sigma_z^2} \right)^{-1} = \boldsymbol{D}_k - \frac{\boldsymbol{D}_k \boldsymbol{x}_k^* \boldsymbol{x}_k^T \boldsymbol{D}_k}{\sigma_z^2 + \boldsymbol{x}_k^T \boldsymbol{D}_k \boldsymbol{x}_k^*}, \tag{4.21}$$

and

$$\boldsymbol{e}_k = \boldsymbol{E}_k \left(\frac{\boldsymbol{x}_k^* y_k}{\sigma_z^2} + \boldsymbol{D}_k^{-1} \boldsymbol{d}_k \right) = \left(\boldsymbol{D}_k - \frac{\boldsymbol{D}_k \boldsymbol{x}_k^* \boldsymbol{x}_k^T \boldsymbol{D}_k}{\sigma_z^2 + \boldsymbol{x}_k^T \boldsymbol{D}_k \boldsymbol{x}_k^*} \right) \left(\frac{\boldsymbol{x}_k^* y_k}{\sigma_z^2} + \boldsymbol{D}_k^{-1} \boldsymbol{d}_k \right) \tag{4.22}$$

Also, for any k-by-k' matrix $\boldsymbol{A}$ and k'-by-k matrix $\boldsymbol{B}$ [8],

$$|\boldsymbol{I}_k + \boldsymbol{A}\boldsymbol{B}| = |\boldsymbol{I}_{k'} + \boldsymbol{B}\boldsymbol{A}|, \tag{4.23}$$

where $\boldsymbol{I}_k$ represents the k-by-k identity matrix. This gives that:

$$\begin{aligned} |\pi \boldsymbol{E}_k| &= \left|\pi \boldsymbol{D}_k \left(\boldsymbol{I} - \frac{\boldsymbol{x}_k^* \boldsymbol{x}_k^T \boldsymbol{D}_k}{\sigma_z^2 + \boldsymbol{x}_k^T \boldsymbol{D}_k \boldsymbol{x}_k^*}\right)\right| \\ &= |\pi \boldsymbol{D}_k| \left|\boldsymbol{I} - \boldsymbol{x}_k^* \frac{\boldsymbol{x}_k^T \boldsymbol{D}_k}{\sigma_z^2 + \boldsymbol{x}_k^T \boldsymbol{D}_k \boldsymbol{x}_k^*}\right| \\ &= |\pi \boldsymbol{D}_k| \left|1 - \frac{\boldsymbol{x}_k^T \boldsymbol{D}_k \boldsymbol{x}_k^*}{\sigma_z^2 + \boldsymbol{x}_k^T \boldsymbol{D}_k \boldsymbol{x}_k^*}\right| \\ &= |\pi \boldsymbol{D}_k| \left(\frac{\sigma_z^2}{\sigma_z^2 + \boldsymbol{x}_k^T \boldsymbol{D}_k \boldsymbol{x}_k^*}\right) \end{aligned} \tag{4.24}$$

By letting $\sigma_X^2 \triangleq \sigma_z^2 + \boldsymbol{x}_k^T \boldsymbol{D}_k \boldsymbol{x}_k^*$ for notational convenience, the negative exponent in (4.19) can be reduced to:

$$
\begin{aligned}
& -\boldsymbol{e}_k^H \boldsymbol{E}_k^{-1} \boldsymbol{e}_k + \boldsymbol{d}_k^H \boldsymbol{D}_k^{-1} \boldsymbol{d}_k + \frac{|y_k|^2}{\sigma_z^2} \\
= & -\left(\boldsymbol{E}_k^{-1} \boldsymbol{e}_k\right)^H \boldsymbol{E}_k \left(\boldsymbol{E}_k^{-1} \boldsymbol{e}_k\right) + \boldsymbol{d}_k^H \boldsymbol{D}_k^{-1} \boldsymbol{d}_k + \frac{|y_k|^2}{\sigma_z^2} \\
= & -\left(\frac{\boldsymbol{x}_k^* y_k}{\sigma_z^2} + \boldsymbol{D}_k^{-1} \boldsymbol{d}_k\right)^H \left(\boldsymbol{D}_k - \frac{\boldsymbol{D}_k \boldsymbol{x}_k^* \boldsymbol{x}_k^T \boldsymbol{D}_k}{\sigma_X^2}\right) \left(\frac{\boldsymbol{x}_k^* y_k}{\sigma_z^2} + \boldsymbol{D}_k^{-1} \boldsymbol{d}_k\right) + \boldsymbol{d}_k^H \boldsymbol{D}_k^{-1} \boldsymbol{d}_k + \frac{|y_k|^2}{\sigma_z^2} \\
= & \; y_k^H \left(\frac{\boldsymbol{x}_k^T \boldsymbol{D}_k \boldsymbol{x}_k^* \boldsymbol{x}_k^T \boldsymbol{D}_k \boldsymbol{x}_k^*}{\left(\sigma_z^2\right)^2 \sigma_X^2} - \frac{\boldsymbol{x}_k^T \boldsymbol{D}_k \boldsymbol{x}_k^*}{\left(\sigma_z^2\right)^2} + \frac{1}{\sigma_z^2}\right) y_k + y_k^H \left(\boldsymbol{d}_k^H \boldsymbol{x}_k^* \left(\frac{\boldsymbol{x}_k^T \boldsymbol{D}_k \boldsymbol{x}_k^*}{\sigma_z^2 \sigma_X^2} - \frac{1}{\sigma_z^2}\right)\right)^H \\
& + \boldsymbol{d}_k^H \boldsymbol{x}_k^* \left(\frac{\boldsymbol{x}_k^T \boldsymbol{D}_k \boldsymbol{x}_k^*}{\sigma_z^2 \sigma_X^2} - \frac{1}{\sigma_z^2}\right) y_k + \frac{\left(\boldsymbol{x}_k^T \boldsymbol{d}_k\right)^H \left(\boldsymbol{x}_k^T \boldsymbol{d}_k\right)}{\sigma_X^2} \\
= & \; y_k^H \left(\frac{(\boldsymbol{x}_k^T \boldsymbol{D}_k \boldsymbol{x}_k^*)^2}{\left(\sigma_z^2\right)^2 \sigma_X^2} - \frac{\boldsymbol{x}_k^T \boldsymbol{D}_k \boldsymbol{x}_k^*}{\left(\sigma_z^2\right)^2} + \frac{\sigma_X^2 - \boldsymbol{x}_k^T \boldsymbol{D}_k \boldsymbol{x}_k^*}{\left(\sigma_z^2\right)^2}\right) y_k + y_k^H \left(\frac{\boldsymbol{x}_k^T \boldsymbol{D}_k \boldsymbol{x}_k^*}{\sigma_z^2 \sqrt{\sigma_X^2}} - \frac{\sqrt{\sigma_X^2}}{\sigma_z^2}\right)^H \\
& \frac{\left(\boldsymbol{x}_k^T \boldsymbol{d}_k\right)}{\sqrt{\sigma_X^2}} + \frac{\left(\boldsymbol{x}_k^T \boldsymbol{d}_k\right)^H}{\sqrt{\sigma_X^2}} \left(\frac{\boldsymbol{x}_k^T \boldsymbol{D}_k \boldsymbol{x}_k^*}{\sigma_z^2 \sqrt{\sigma_X^2}} - \frac{\sqrt{\sigma_X^2}}{\sigma_z^2}\right) y_k + \frac{\left(\boldsymbol{x}_k^T \boldsymbol{d}_k\right)^H}{\sqrt{\sigma_X^2}} \frac{\left(\boldsymbol{x}_k^T \boldsymbol{d}_k\right)}{\sqrt{\sigma_X^2}} \\
= & \left[\left(\frac{\boldsymbol{x}_k^T \boldsymbol{D}_k \boldsymbol{x}_k^*}{\sigma_z^2 \sqrt{\sigma_X^2}} - \frac{\sqrt{\sigma_X^2}}{\sigma_z^2}\right) y_k + \frac{\left(\boldsymbol{x}_k^T \boldsymbol{d}_k\right)}{\sqrt{\sigma_X^2}}\right]^H \left[\left(\frac{\boldsymbol{x}_k^T \boldsymbol{D}_k \boldsymbol{x}_k^*}{\sigma_z^2 \sqrt{\sigma_X^2}} - \frac{\sqrt{\sigma_X^2}}{\sigma_z^2}\right) y_k + \frac{\left(\boldsymbol{x}_k^T \boldsymbol{d}_k\right)}{\sqrt{\sigma_X^2}}\right] \\
= & \left|\left(\frac{\boldsymbol{x}_k^T \boldsymbol{D}_k \boldsymbol{x}_k^*}{\sigma_z^2 \sqrt{\sigma_X^2}} - \frac{\sqrt{\sigma_X^2}}{\sigma_z^2}\right) y_k + \frac{\left(\boldsymbol{x}_k^T \boldsymbol{d}_k\right)}{\sqrt{\sigma_X^2}}\right|^2 \\
= & \frac{1}{\sigma_X^2} \left|\left(\frac{\boldsymbol{x}_k^T \boldsymbol{D}_k \boldsymbol{x}_k^*}{\sigma_z^2} - \frac{\sigma_X^2}{\sigma_z^2}\right) y_k + \boldsymbol{x}_k^T \boldsymbol{d}_k\right|^2 \\
= & \frac{1}{\sigma_X^2} \left|\left(\frac{\boldsymbol{x}_k^T \boldsymbol{D}_k \boldsymbol{x}_k^*}{\sigma_z^2} - \frac{\sigma_z^2 + \boldsymbol{x}_k^T \boldsymbol{D}_k \boldsymbol{x}_k^*}{\sigma_z^2}\right) y_k + \left(\boldsymbol{x}_k^T \boldsymbol{d}_k\right)\right|^2 \\
= & \frac{\left|y_k - \boldsymbol{x}_k^T \boldsymbol{d}_k\right|^2}{\sigma_z^2 + \boldsymbol{x}_k^T \boldsymbol{D}_k \boldsymbol{x}_k^*}.
\end{aligned} \tag{4.25}
$$

Finally, the above derivations summarize to:

$$
\begin{aligned}
f\left(y_k | \boldsymbol{x}_k\right) & = \frac{|\pi \boldsymbol{E}_k|}{\pi \sigma_z^2 |\pi \boldsymbol{D}_k|} \exp\left(-\frac{\left|y_k - \left(\boldsymbol{x}_k^T \boldsymbol{d}_k\right)\right|^2}{\sigma_z^2 + \boldsymbol{x}_k^T \boldsymbol{D}_k \boldsymbol{x}_k^*}\right) \\
& = \frac{|\pi \boldsymbol{D}_k|}{\pi \sigma_z^2 |\pi \boldsymbol{D}_k|} \left|\frac{\sigma_z^2}{\sigma_z^2 + \boldsymbol{x}_k^T \boldsymbol{D}_k \boldsymbol{x}_k^*}\right| \exp\left(-\frac{\left|y_k - \left(\boldsymbol{x}_k^T \boldsymbol{d}_k\right)\right|^2}{\sigma_z^2 + \boldsymbol{x}_k^T \boldsymbol{D}_k \boldsymbol{x}_k^*}\right) \\
& = \frac{1}{\pi\left(\sigma_z^2 + \boldsymbol{x}_k^T \boldsymbol{D}_k \boldsymbol{x}_k^*\right)} \exp\left(-\frac{\left|y_k - \left(\boldsymbol{x}_k^T \boldsymbol{d}_k\right)\right|^2}{\sigma_z^2 + \boldsymbol{x}_k^T \boldsymbol{D}_k \boldsymbol{x}_k^*}\right),
\end{aligned} \tag{4.26}
$$

where $\boldsymbol{D}_k \triangleq (1-|\alpha|^{2k})/(1-|\alpha|^2)\boldsymbol{C}$ and $\boldsymbol{d}_k \triangleq \alpha^k \boldsymbol{h}_0 + (1-\alpha^k)\boldsymbol{d}/(1-\alpha)$. Accordingly, the probability distribution of $f(y_k)$ is

$$f(y_k) = \sum_{\boldsymbol{x}_k \in \mathcal{X}^M} P_{\boldsymbol{X}_k}(\boldsymbol{x}_k) \frac{1}{\pi\left(\sigma_z^2 + \boldsymbol{x}_k^T \boldsymbol{D}_k \boldsymbol{x}_k^*\right)} \exp\left(-\frac{\left|y_k - \left(\boldsymbol{x}_k^T \boldsymbol{d}_k\right)\right|^2}{\sigma_z^2 + \boldsymbol{x}_k^T \boldsymbol{D}_k \boldsymbol{x}_k^*}\right). \tag{4.27}$$

4.3 Independent bound for $M = 1$

Recall that:

$$C(S) \le \frac{1}{n}\left[C_1(S) + \cdots + C_n(S)\right]. \tag{4.28}$$

where $C_k(S) = \max_{\{P_{\boldsymbol{X}}\,:\,(\forall\ i)E[X_i^2]\le S\}} I(X_k;Y_k)$. Now we will derive $C_k(S)$ for the case of $M = 1$.

For $M = 1$,

$$\begin{aligned} f(y_k|\boldsymbol{x}_k) = f(y_k|x_k) &= \frac{1}{\pi\left(\sigma_z^2 + |x_k|^2 D_k\right)} \exp\left(-\frac{|y_k - x_k d_k|^2}{\sigma_z^2 + |x_k|^2 D_k}\right) \\ &= \frac{1}{\pi\left(\sigma_z^2 + |x_k|^2 D_k\right)} \exp\left(-\frac{|d_k|^2\,|y_k/d_k - x_k|^2}{\sigma_z^2 + |x_k|^2 D_k}\right). \end{aligned} \tag{4.29}$$

Since $I(X_k;Y_k) = I(X_k;\tilde{Y}_k)$ for invertible transformation $\tilde{Y}_k \triangleq Y_k/d_k$, we can transform $f(y_k|x_k)$ to obtain $f(\tilde{y}_k|x_k)$ as:

$$f(\tilde{y}_k|x_k) = \frac{1}{\pi\left(\sigma_z^2 + |x_k|^2 D_k\right)/|d_k|^2} \exp\left(-\frac{|\tilde{y}_k - x_k|^2}{(\sigma_z^2 + |x_k|^2 D_k)/|d_k|^2}\right). \tag{4.30}$$

In our system, the transmitted symbol x_k is assumed antipodally modulated. In other words, $x_k \in \{-s, +s\}$ for some real s. In such case, the imaginary part of $\tilde{y}_k$ is irrelevant to the channel input; hence, we can further reduce the complex channel to the real channel without affecting the mutual information as:

$$f(\tilde{y}_{k,r}|x_k) = \frac{1}{\sqrt{\pi(\sigma_z^2 + s^2 D_k)/|d_k|^2}} \exp\left(-\frac{(\tilde{y}_{k,r} - x_k)^2}{(\sigma_z^2 + s^2 D_k)/|d_k|^2}\right), \tag{4.31}$$

where $\tilde{y}_{k,r}$ is the real part of $\tilde{y}_k$. For this real-valued symmetric additive Gaussian channel, its capacity-cost function is achieved by uniform input with $s^2 = S$; hence, by denoting $\sigma_N^2 \triangleq (\sigma_z^2 + s^2 D_k)/(2|d_k|^2)$,

$$\begin{aligned} C_k(S) &= I(X_k; \tilde{Y}_{k,r}) \\ &= h(\tilde{Y}_{k,r}) - \sum_{x_k \in \mathcal{X}} \frac{1}{2} \int_{\Re} f\left(\tilde{y}_{k,r}|x_k\right) \log\left[\frac{1}{f\left(\tilde{y}_{k,r}|x_k\right)}\right] d\tilde{y}_{k,r} \\ &= h(\tilde{Y}_{k,r}) - \frac{1}{2}\log\left[2\pi e \sigma_N^2\right]. \end{aligned} \tag{4.32}$$

Now, for uniform channel input,

$$\begin{aligned} h(\tilde{Y}_{k,r}) &= -\int_{\Re} f(\tilde{y}_{k,r}) \log(f(\tilde{y}_{k,r})) d\tilde{y}_{k,r} \\ &= \int_{\Re} f(\tilde{y}_{k,r}) \left[\frac{1}{2}\log(2\pi\sigma_N^2) + \frac{\tilde{y}_{k,r}^2 + s^2}{2\sigma_N^2} - \log\left(\frac{e^{\tilde{y}_{k,r}s/\sigma_N^2} + e^{-\tilde{y}_{k,r}s/\sigma_N^2}}{2}\right)\right] d\tilde{y}_{k,r} \\ &= \frac{1}{2}\log(2\pi e\sigma_N^2) + \frac{s^2}{\sigma_N^2} - \int_{\Re} \frac{1}{\sqrt{2\pi\sigma_N^2}} e^{-(\tilde{y}_{k,r}-s)^2/(2\sigma_N^2)} \log\left(\cosh(\tilde{y}_{k,r}s/\sigma_N^2)\right) d\tilde{y}_{k,r} \\ &= \frac{1}{2}\log(2\pi e\sigma_N^2) + \frac{s^2}{\sigma_N^2} - \int_{\Re} \frac{1}{\sqrt{2\pi}} e^{-t^2/2} \log\left(\cosh\left(\frac{s^2}{\sigma_N^2} + t\sqrt{\frac{s^2}{\sigma_N^2}}\right)\right) dt \\ &= \frac{1}{2}\log(2\pi e\sigma_N^2) + \frac{S}{\sigma_N^2} - \int_{\Re} \frac{1}{\sqrt{2\pi}} e^{-t^2/2} \log\left(\cosh\left(\frac{S}{\sigma_N^2} + t\sqrt{\frac{S}{\sigma_N^2}}\right)\right) dt, \end{aligned} \tag{4.33}$$

which immediately gives that:

$$C_k(S) = \frac{2|d_k|^2 S}{\sigma_z^2 + SD_k} - \int_{\Re} \frac{1}{\sqrt{2\pi}} e^{-t^2/2} \log\left(\cosh\left(\frac{2|d_k|^2 S}{\sigma_z^2 + SD_k} + t\sqrt{\frac{2|d_k|^2 S}{\sigma_z^2 + SD_k}}\right)\right) dt. \tag{4.34}$$

By Cesáro-mean theorem [3], when taking n to infinity, we have:

$$C(S) \le \lim_{n\to\infty} \frac{1}{n} \sum_{k=1}^{n} C_k(S) = \lim_{n\to\infty} C_n(S) = C_\infty(S), \tag{4.35}$$

where

$$C_\infty(S) = \frac{2|d_\infty|^2 S}{\sigma_z^2 + SD_\infty} - \int_{\Re} \frac{1}{\sqrt{2\pi}} e^{-t^2/2} \log\left(\cosh\left(\frac{2|d_\infty|^2 S}{\sigma_z^2 + SD_\infty} + t\sqrt{\frac{2|d_\infty|^2 S}{\sigma_z^2 + SD_\infty}}\right)\right) dt \tag{4.36}$$

and $d_\infty = d/(1-\alpha)$ and $D_\infty = C/(1-|\alpha|^2)$, provided that $|\alpha| < 1$.

We will demonstrate in the next three figures how channel model parameters affect the independent bound of channel capacity.

In Fig.4.1, $C_k(S)$ is plotted for different initial channel coefficients $\boldsymbol{h}_0$. Obviously, $\boldsymbol{h}_0$ affects $C_k(S)$ only at small k. As k grows, $\boldsymbol{d}_k$ will be dominated by α and $\boldsymbol{d}_0$, and $C_k(S)$ will converge to the same limit $C_\infty(S)$.

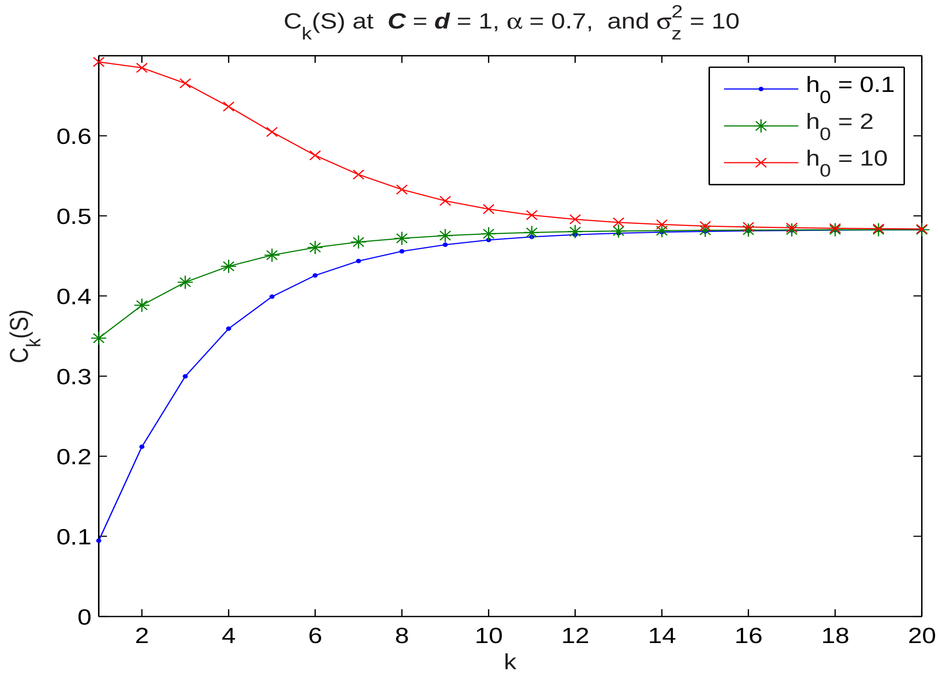

Figure 4.1: Illustration of $C_k(S)$ for different initial fading coefficients $\boldsymbol{h}_0 = 0.1, 2$ and 10. Other parameters for Gauss-Markov channels are $S = 1$, $\boldsymbol{C} = 1$, $\boldsymbol{d} = 1$, $\alpha = 0.7$ and $\sigma_z^2 = 10$.

In Fig. 4.2, $C_k(S)$ is plotted for different initial Gauss-Markov fading mean $\boldsymbol{d}$. In principle, the value of $\boldsymbol{d}$ determines the strength of line-of-sight (LOS) propagated signal. Figure 4.2 indicates that a stronger LOS signal can result in better capacity bound $C_\infty(S)$.

In Fig.4.3, $C_k(S)$ is plotted for different initial fading covariance matrix $\boldsymbol{C}$. The figure indicates that a larger fading variance $\boldsymbol{C}$ makes a lower capacity bound.

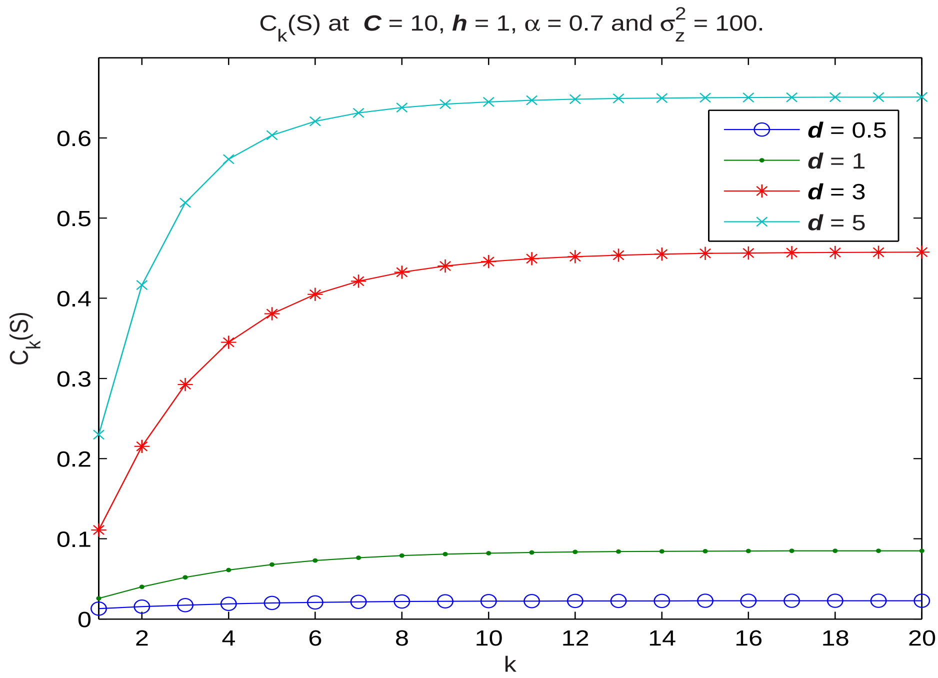

Figure 4.2: Illustration of $C_k(S)$ for different initial Gauss-Markov fading mean $\boldsymbol{d} = 0.5, 1, 3$ and 5. Other parameters for Gauss-Markov channels are $S = 1$, $\boldsymbol{C} = 10$, $\boldsymbol{h}_0 = 1$, $\alpha = 0.7$ and $\sigma_z^2 = 100$.

An interesting observation that can be made on $C_\infty(S)$ is that it is equal to zero once $\boldsymbol{d} = \boldsymbol{0}$. This means that the channel capacity $C(S) = 0$ if there exists no LOS signals in the communications via Gauss-Markov channels.

The Gauss-Markov channel model can be reduced to the additive white Gaussian noise (AWGN) channel model by letting $\alpha = 0$, $\boldsymbol{d} = \boldsymbol{1}$ and $\boldsymbol{C} = \boldsymbol{0}$. As a result,

$$C_\infty^{\text{AWGN}}(S) = C_k^{\text{AWGN}}(S) = \frac{S}{\sigma_z^2/2} - \int_{\Re} \frac{1}{\sqrt{2\pi}} e^{-t^2/2} \log\left(\cosh\left(\frac{S}{\sigma_z^2/2} + t\sqrt{\frac{S}{\sigma_z^2/2}}\right)\right) dt. \quad (4.37)$$

Notably, this is no longer an upper bound, but the exact channel capacity formula for binary-input AWGN channels [3].

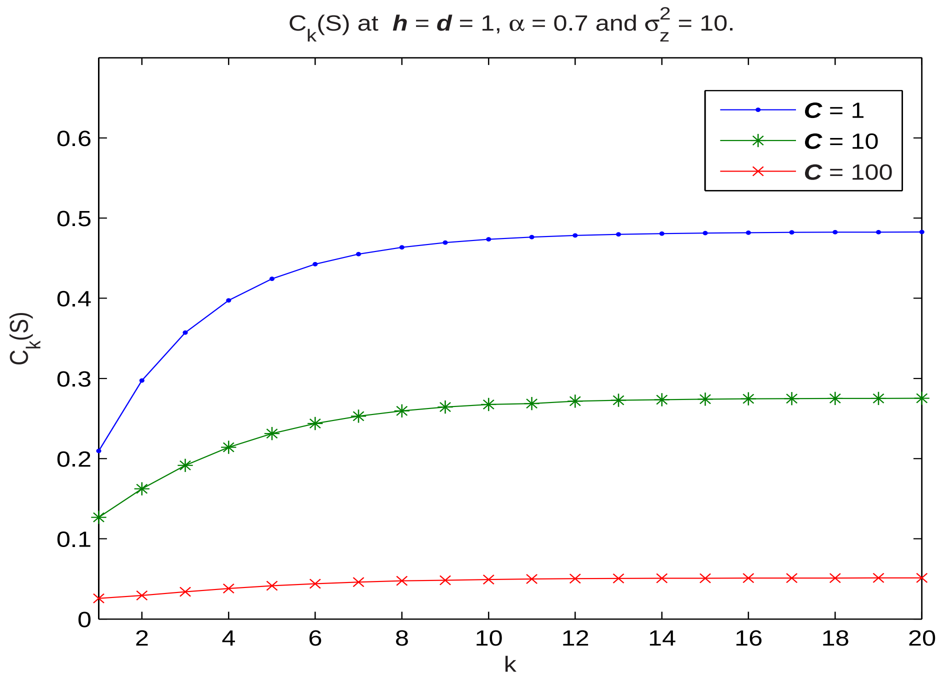

Figure 4.3: Illustration of $C_k(S)$ for different initial fading covariance matrix $\boldsymbol{C} = 1, 10$ and 100. Other parameters for Gauss-Markov channels are $S = 1$, $\boldsymbol{h}_0 = 1$, $\boldsymbol{d} = 1$, $\alpha = 0.7$ and $\sigma_z^2 = 10$.

4.4 Independent bound for $M = 2$

At $M = 2$,

$$\begin{aligned} C_k(S) &= \max_{\{P_{\mathbf{X}} \,:\, (\forall\, i) E[X_i^2] \leq S\}} I(X_k; Y_k) \\ &= \max_{\{P_{X_{k-1}^k} \,:\, E[X_k]^2 \leq S \text{ and } E[X_{k-1}^2] \leq S\}} I(X_k; Y_k). \end{aligned} \tag{4.38}$$

It is in general hard to find $C_k(S)$ for the case of $M > 1$. Hence, we made the assumption on the channel statistics below.

Assumption 4.1 *$d_{k,1} = \rho_1 d_k$ and $d_{k,2} = \rho_2 d_k$ for some real numbers ρ_1 and ρ_2, where $\boldsymbol{d}_k = \begin{bmatrix} d_{k,1} & d_{k,2} \end{bmatrix}$. Also, $\boldsymbol{C}$ is diagonal; hence, there exists $D_{k,1}$ and $D_{k,2}$ such that*

$$\boldsymbol{D}_k = \begin{bmatrix} D_{k,1} & 0 \\ 0 & D_{k,2} \end{bmatrix} = (1 - |\alpha|^{2k})/(1 - |\alpha|^2) \begin{bmatrix} C_{11} & 0 \\ 0 & C_{22} \end{bmatrix}.$$

Recall that

$$
\begin{aligned}
f\left(y_k|\boldsymbol{x}_k\right) &\triangleq \frac{1}{\pi\left(\sigma_z^2+\boldsymbol{x}_k^T\boldsymbol{D}_k\boldsymbol{x}_k^*\right)}\exp\left(-\frac{\left|y_k-\left(\boldsymbol{x}_k^T\boldsymbol{d}_k\right)\right|^2}{\sigma_z^2+\boldsymbol{x}_k^T\boldsymbol{D}_k\boldsymbol{x}_k^*}\right)\\
&= f\left(y_k|x_k,x_{k-1}\right)\\
&= \frac{1}{\pi\left(\sigma_z^2+|x_k|^2D_{k,1}+|x_{k-1}|^2D_{k,2}\right)}\exp\left(-\frac{|y_k-x_kd_{k,1}-x_{k-1}d_{k,2}|^2}{\sigma_z^2+|x_k|^2D_{k,1}+|x_{k-1}|^2D_{k,2}}\right)\\
&= \frac{1}{\pi\left(\sigma_z^2+s^2D_{k,1}+s^2D_{k,2}\right)}\exp\left(-\frac{|y_k/d_k-(\rho_1x_k+\rho_2x_{k-1})|^2}{(\sigma_z^2+s^2D_{k,1}+s^2D_{k,2})/|d_k|^2}\right). \qquad (4.39)
\end{aligned}
$$

Then, by letting $\tilde{y}_k = y_k/d_k$, we obtain:

$$
f\left(\tilde{y}_k|x_k,x_{k-1}\right) = \frac{1}{\pi\sigma^2}\exp\left(-\frac{|\tilde{y}_k-(\rho_1x_k+\rho_2x_{k-1})|^2}{\sigma^2}\right), \qquad (4.40)
$$

where $\sigma^2 \triangleq (\sigma_z^2+s^2D_{k,1}+s^2D_{k,2})/|d_k|^2$. By following the same reasoning as in the previous section,

$$
I(X_k;Y_k) = I(X_k;\tilde{Y}_k) = I(X_k;\tilde{Y}_{k,r}),
$$

where $\tilde{Y}_k = \tilde{Y}_{k,r}+j\tilde{Y}_{k,i}$. Then, we derive:

$$
I(X_k;\tilde{Y}_{k,r}) = \sum_{x_{k-1}\in\mathcal{X}}\sum_{x_k\in\mathcal{X}}\int_{\Re}P_{X_k,X_{k-1}}(x_k,x_{k-1})f(\tilde{y}_{k,r}|x_k,x_{k-1})
$$

$$
\log\frac{\displaystyle\sum_{\bar{x}_{k-1}\in\mathcal{X}}P_{X_k,X_{k-1}}(x_k,\bar{x}_{k-1})f(\tilde{y}_{k,r}|x_k,\bar{x}_{k-1})}{\left(\displaystyle\sum_{\hat{x}_{k-1}\in\mathcal{X}}P_{X_k,X_{k-1}}(x_k,\hat{x}_{k-1})\right)\left(\displaystyle\sum_{x'_{k-1}\in\mathcal{X}}\sum_{x'_k\in\mathcal{X}}P_{X_k,X_{k-1}}(x'_k,x'_{k-1})f(\tilde{y}_{k,r}|x'_k,x'_{k-1})\right)}d\tilde{y}_{k,r};
$$

hence, by taking the derivative of $I(X_k;\tilde{Y}_{k,r})+\lambda\left(\sum_{x_{k-1}\in\mathcal{X}}\sum_{x_k\in\mathcal{X}}P_{X_k,X_{k-1}}(x_k,x_{k-1})-1\right)$

with respect to $P_{X_k,X_{k-1}}(x''_k, x''_{k-1})$, we obtain:

$$
\begin{aligned}
& \frac{\partial\left[I(X_k;\tilde{Y}_{k,r})+\lambda\left(\sum_{x_{k-1}\in\mathcal{X}}\sum_{x_k\in\mathcal{X}}P_{X_k,X_{k-1}}(x_k,x_{k-1})-1\right)\right]}{\partial P_{X_k,X_{k-1}}(x''_k,x''_{k-1})} \\
= & \int_{\Re} f(\tilde{y}_{k,r}|x''_k,x''_{k-1})\log\left(\sum_{\bar{x}_{k-1}\in\mathcal{X}}P_{X_k,X_{k-1}}(x''_k,\bar{x}_{k-1})f(\tilde{y}_{k,r}|x''_k,\bar{x}_{k-1})\right)d\tilde{y}_{k,r}+1 \\
& -\int_{\Re} f(\tilde{y}_{k,r}|x''_k,x''_{k-1})\log\left(\sum_{\hat{x}_{k-1}\in\mathcal{X}}P_{X_k,X_{k-1}}(x''_k,\hat{x}_{k-1})\right)d\tilde{y}_{k,r}-1 \\
& -\int_{\Re} f(\tilde{y}_{k,r}|x''_k,x''_{k-1})\log\left(\sum_{x'_{k-1}\in\mathcal{X}}\sum_{x'_k\in\mathcal{X}}P_{X_k,X_{k-1}}(x'_k,x'_{k-1})f(\tilde{y}_{k,r}|x'_k,x'_{k-1})\right)-1+\lambda \\
= & I(x''_k,x''_{k-1};\tilde{Y}_{k,r})-1+\lambda,
\end{aligned}
$$

where

$$
\begin{aligned}
& I(x''_k,x''_{k-1};\tilde{Y}_{k,r}) \triangleq \int_{\Re} f(\tilde{y}_{k,r}|x''_k,x''_{k-1}) \\
& \times\log\frac{\sum_{\bar{x}_{k-1}\in\mathcal{X}}P_{X_k,X_{k-1}}(x''_k,\bar{x}_{k-1})f(\tilde{y}_{k,r}|x''_k,\bar{x}_{k-1})}{\left(\sum_{\hat{x}_{k-1}\in\mathcal{X}}P_{X_k,X_{k-1}}(x''_k,\hat{x}_{k-1})\right)\left(\sum_{x'_{k-1}\in\mathcal{X}}\sum_{x'_k\in\mathcal{X}}P_{X_k,X_{k-1}}(x'_k,x'_{k-1})f(\tilde{y}_{k,r}|x'_k,x'_{k-1})\right)}d\tilde{y}_{k,r}
\end{aligned}
$$

By similar reasoning in [3],

$$
I(x_k,x_{k-1};\tilde{Y}_{k,r})\begin{cases} = C_k(S), & \text{if } P_{X_k,X_{k-1}}(x_k,x_{k-1})>0 \\ \leq C_k(S), & \text{if } P_{X_k,X_{k-1}}(x_k,x_{k-1})=0. \end{cases} \tag{4.41}
$$

For notational convenience, let $I_k(a,b)$ denote $I(x_k = as, x_{k-1} = bs; \tilde{Y}_{k,r})$, where $a, b = \pm 1$. Also, brief $P_{X_k,X_{k-1}}(x_k, x_{k-1})$ and $f(\tilde{y}_{k,r}|x_k, x_{k-1})$ by $p_{a,b}$ and $f(\tilde{y}_{k,r}|a, b)$, respectively. Then,

$$
\begin{aligned}
I_k(a,b) & \triangleq \int_{\Re} f(\tilde{y}_{k,r}|a,b)\log\frac{\sum_{\bar{b}=\pm1}p_{a,\bar{b}}f(\tilde{y}_{k,r}|a,\bar{b})}{\left(\sum_{\hat{b}=\pm1}p_{a,\hat{b}}\right)\left(\sum_{a'=\pm1}\sum_{b'=\pm1}p_{a',b'}f(\tilde{y}_{k,r}|a',b')\right)}d\tilde{y}_{k,r} \\
& = h_k(a,b)-h_k(a,b|a),
\end{aligned}
$$

where

$$
\begin{aligned}
h_k(a,b) &\triangleq -\int_{\Re} f(\tilde{y}_{k,r}|a,b)\log\left(\sum_{a'=\pm1}\sum_{b'=\pm1} p_{a',b'} f(\tilde{y}_{k,r}|a',b')\right) d\tilde{y}_{k,r} \\
&= \int_{\Re} f(\tilde{y}_{k,r}|a,b)\left[\frac{1}{2}\log(\pi\sigma^2) + \frac{\tilde{y}_{k,r}^2 + s^2\rho_1^2 + s^2\rho_2^2}{\sigma^2}\right. \\
&\quad - \log\left(p_{1,1}e^{\frac{2s(\rho_1+\rho_2)\tilde{y}_{k,r}-2s^2\rho_1\rho_2}{\sigma^2}} + p_{1,-1}e^{\frac{2s(\rho_1-\rho_2)\tilde{y}_{k,r}+2s^2\rho_1\rho_2}{\sigma^2}}\right. \\
&\quad \left.\left. + p_{-1,1}e^{\frac{-2s(\rho_1-\rho_2)\tilde{y}_{k,r}+2s^2\rho_1\rho_2}{\sigma^2}} + p_{-1,-1}e^{\frac{-2s(\rho_1+\rho_2)\tilde{y}_{k,r}-2s^2\rho_1\rho_2}{\sigma^2}}\right)\right] d\tilde{y}_{k,r} \\
&= \frac{1}{2}\log(\pi e\sigma^2) + \frac{s^2\rho_1^2 + s^2\rho_2^2 + s^2(a\rho_1+b\rho_2)^2}{\sigma^2} \\
&\quad - \int_{\Re} f(\tilde{y}_{k,r}|a,b)\left[\log\left(p_{1,1}e^{\frac{2s(\rho_1+\rho_2)\tilde{y}_{k,r}-2s^2\rho_1\rho_2}{\sigma^2}} + p_{1,-1}e^{\frac{2s(\rho_1-\rho_2)\tilde{y}_{k,r}+2s^2\rho_1\rho_2}{\sigma^2}}\right.\right. \\
&\quad \left.\left. + p_{-1,1}e^{\frac{-2s(\rho_1-\rho_2)\tilde{y}_{k,r}+2s^2\rho_1\rho_2}{\sigma^2}} + p_{-1,-1}e^{\frac{-2s(\rho_1+\rho_2)\tilde{y}_{k,r}-2s^2\rho_1\rho_2}{\sigma^2}}\right)\right] d\tilde{y}_{k,r},
\end{aligned}
$$

and

$$
\begin{aligned}
h_k(a,b|a) &\triangleq -\int_{\Re} f(\tilde{y}_{k,r}|a,b)\log\frac{\sum_{\bar{b}=\pm1} p_{a,\bar{b}} f(\tilde{y}_{k,r}|a,\bar{b})}{\left(\sum_{\hat{b}=\pm1} p_{a,\hat{b}}\right)} d\tilde{y}_{k,r} \\
&= \int_{\Re} f(\tilde{y}_{k,r}|a,b)\left[\frac{1}{2}\log(\pi\sigma^2) + \frac{\tilde{y}_{k,r}^2 + s^2\rho_1^2 + s^2\rho_2^2}{\sigma^2}\right. \\
&\quad \left. - \log\left(p_{a,1}e^{\frac{2s(a\rho_1+\rho_2)\tilde{y}_{k,r}-2s^2\rho_1\rho_2 a}{\sigma^2}} + p_{a,-1}e^{\frac{2s(a\rho_1-\rho_2)\tilde{y}_{k,r}+2s^2\rho_1\rho_2 a}{\sigma^2}}\right)\right] d\tilde{y}_{k,r} + \log\left(p_{a,1}+p_{a,-1}\right) \\
&= \log\left(p_{a,1}+p_{a,-1}\right) + \frac{1}{2}\log(\pi e\sigma^2) + \frac{s^2\rho_1^2 + s^2\rho_2^2 + s^2(a\rho_1+b\rho_2)^2}{\sigma^2} \\
&\quad - \int_{\Re} f(\tilde{y}_{k,r}|a,b)\log\left(p_{a,1}e^{\frac{2s(a\rho_1+\rho_2)\tilde{y}_{k,r}-2s^2\rho_1\rho_2 a}{\sigma^2}} + p_{a,-1}e^{\frac{2s(a\rho_1-\rho_2)\tilde{y}_{k,r}+2s^2\rho_1\rho_2 a}{\sigma^2}}\right) d\tilde{y}_{k,r}.
\end{aligned}
$$

Accordingly,

$$\begin{aligned}
I_k(a,b) &= \log\left(p_{a,1}+p_{a,-1}\right) \\
&+ \int_{\Re} f(\tilde{y}_{k,r}|a,b)\log\left(p_{a,1}e^{\frac{2s(a\rho_1+\rho_2)\tilde{y}_{k,r}-2s^2\rho_1\rho_2 a}{\sigma^2}} + p_{a,-1}e^{\frac{2s(a\rho_1-\rho_2)\tilde{y}_{k,r}+2s^2\rho_1\rho_2 a}{\sigma^2}}\right)d\tilde{y}_{k,r} \\
&- \int_{\Re} f(\tilde{y}_{k,r}|a,b)\left[\log\left(p_{1,1}e^{\frac{2s(\rho_1+\rho_2)\tilde{y}_{k,r}-2s^2\rho_1\rho_2}{\sigma^2}} + p_{1,-1}e^{\frac{2s(\rho_1-\rho_2)\tilde{y}_{k,r}+2s^2\rho_1\rho_2}{\sigma^2}}\right.\right. \\
&\left.\left. + p_{-1,1}e^{\frac{-2s(\rho_1-\rho_2)\tilde{y}_{k,r}+2s^2\rho_1\rho_2}{\sigma^2}} + p_{-1,-1}e^{\frac{-2s(\rho_1+\rho_2)\tilde{y}_{k,r}-2s^2\rho_1\rho_2}{\sigma^2}}\right)\right]d\tilde{y}_{k,r} \\
&= \log\left(p_{a,1}+p_{a,-1}\right) \\
&+ \int_{\Re}\frac{1}{\sqrt{2\pi}}e^{-u^2/2}\log\left(p_{a,1}e^{\frac{\sqrt{2}s\sigma(a\rho_1+\rho_2)u+2s^2(\rho_1^2+b\rho_2^2+ab\rho_1\rho_2)}{\sigma^2}}\right. \\
&\left. + p_{a,-1}e^{\frac{\sqrt{2}s\sigma(a\rho_1-\rho_2)u+2s^2(\rho_1^2-b\rho_2^2+ab\rho_1\rho_2)}{\sigma^2}}\right)du \\
&- \int_{\Re}\frac{1}{\sqrt{2\pi}}e^{-u^2/2}\left[\log\left(p_{1,1}e^{\frac{\sqrt{2}s\sigma(\rho_1+\rho_2)u+2s^2[a\rho_1^2+b\rho_2^2+(a+b-1)\rho_1\rho_2]}{\sigma^2}}\right.\right. \\
&+ p_{1,-1}e^{\frac{\sqrt{2}s\sigma(\rho_1-\rho_2)u+2s^2[a\rho_1^2-b\rho_2^2-(a-b-1)\rho_1\rho_2]}{\sigma^2}} + p_{-1,1}e^{\frac{-\sqrt{2}s\sigma(\rho_1-\rho_2)u-2s^2[a\rho_1^2-b\rho_2^2-(a-b+1)\rho_1\rho_2]}{\sigma^2}} \\
&\left.\left. + p_{-1,-1}e^{\frac{-\sqrt{2}s\sigma(\rho_1+\rho_2)u-2s^2[a\rho_1^2+b\rho_2^2+(a+b+1)\rho_1\rho_2]}{\sigma^2}}\right)\right]du,
\end{aligned}$$

where $u \triangleq \frac{\sqrt{2}}{\sigma}(\tilde{y}_{k,r}-s(a\rho_1+b\rho_2))$. As a result,

$$\begin{aligned}
I_k(1,1) &= \log\left(p_{1,1}+p_{1,-1}\right) \\
&+ \int_{\Re}\frac{1}{\sqrt{2\pi}}e^{-u^2/2}\log\left(p_{1,1}e^{\frac{\sqrt{2}s\sigma(\rho_1+\rho_2)u+2s^2(\rho_1^2+\rho_2^2+\rho_1\rho_2)}{\sigma^2}}\right. \\
&\left. + p_{1,-1}e^{\frac{\sqrt{2}s\sigma(\rho_1-\rho_2)u+2s^2(\rho_1^2-\rho_2^2+\rho_1\rho_2)}{\sigma^2}}\right)du \\
&- \int_{\Re}\frac{1}{\sqrt{2\pi}}e^{-u^2/2}\left[\log\left(p_{1,1}e^{\frac{\sqrt{2}s\sigma(\rho_1+\rho_2)u+2s^2[\rho_1^2+\rho_2^2+\rho_1\rho_2]}{\sigma^2}}\right.\right. \\
&+ p_{1,-1}e^{\frac{\sqrt{2}s\sigma(\rho_1-\rho_2)u+2s^2[\rho_1^2-\rho_2^2+\rho_1\rho_2]}{\sigma^2}} + p_{-1,1}e^{\frac{-\sqrt{2}s\sigma(\rho_1-\rho_2)u-2s^2[\rho_1^2-\rho_2^2-\rho_1\rho_2]}{\sigma^2}} \\
&\left.\left. + p_{-1,-1}e^{\frac{-\sqrt{2}s\sigma(\rho_1+\rho_2)u-2s^2[\rho_1^2+\rho_2^2+3\rho_1\rho_2]}{\sigma^2}}\right)\right]du
\end{aligned}$$

$$
\begin{aligned}
I_k(1,-1) \;=\;& \log\left(p_{1,1}+p_{1,-1}\right)\\
&+\int_{\Re}\frac{1}{\sqrt{2\pi}}e^{-u^2/2}\log\Bigg(p_{1,1}e^{\frac{\sqrt{2}s\sigma(\rho_1+\rho_2)u+2s^2(\rho_1^2-\rho_2^2-\rho_1\rho_2)}{\sigma^2}}\\
&+p_{1,-1}e^{\frac{\sqrt{2}s\sigma(\rho_1-\rho_2)u+2s^2(\rho_1^2+\rho_2^2-\rho_1\rho_2)}{\sigma^2}}\Bigg)\,du\\
&-\int_{\Re}\frac{1}{\sqrt{2\pi}}e^{-u^2/2}\Bigg[\log\Bigg(p_{1,1}e^{\frac{\sqrt{2}s\sigma(\rho_1+\rho_2)u+2s^2[\rho_1^2-\rho_2^2-\rho_1\rho_2]}{\sigma^2}}\\
&+p_{1,-1}e^{\frac{\sqrt{2}s\sigma(\rho_1-\rho_2)u+2s^2[\rho_1^2+\rho_2^2-\rho_1\rho_2]}{\sigma^2}}+p_{-1,1}e^{\frac{-\sqrt{2}s\sigma(\rho_1-\rho_2)u-2s^2[\rho_1^2+\rho_2^2-3\rho_1\rho_2]}{\sigma^2}}\\
&+p_{-1,-1}e^{\frac{-\sqrt{2}s\sigma(\rho_1+\rho_2)u-2s^2[\rho_1^2-\rho_2^2+\rho_1\rho_2]}{\sigma^2}}\Bigg)\Bigg]\,du
\end{aligned}
$$

$$
\begin{aligned}
I_k(-1,1) \;=\;& \log\left(p_{a,1}+p_{a,-1}\right)\\
&+\int_{\Re}\frac{1}{\sqrt{2\pi}}e^{-v^2/2}\log\Bigg(p_{-1,1}e^{\frac{\sqrt{2}s\sigma(\rho_1-\rho_2)v+2s^2(\rho_1^2+\rho_2^2-\rho_1\rho_2)}{\sigma^2}}\\
&+p_{-1,-1}e^{\frac{\sqrt{2}s\sigma(\rho_1+\rho_2)v+2s^2(\rho_1^2-\rho_2^2-\rho_1\rho_2)}{\sigma^2}}\Bigg)\,dv\\
&-\int_{\Re}\frac{1}{\sqrt{2\pi}}e^{-v^2/2}\Bigg[\log\Bigg(p_{1,1}e^{\frac{-\sqrt{2}s\sigma(\rho_1+\rho_2)v-2s^2[\rho_1^2-\rho_2^2+\rho_1\rho_2]}{\sigma^2}}\\
&+p_{1,-1}e^{\frac{-\sqrt{2}s\sigma(\rho_1-\rho_2)v-2s^2[\rho_1^2+\rho_2^2-3\rho_1\rho_2]}{\sigma^2}}+p_{-1,1}e^{\frac{\sqrt{2}s\sigma(\rho_1-\rho_2)v+2s^2[\rho_1^2+\rho_2^2-\rho_1\rho_2]}{\sigma^2}}\\
&+p_{-1,-1}e^{\frac{\sqrt{2}s\sigma(\rho_1+\rho_2)v+2s^2[\rho_1^2-\rho_2^2-\rho_1\rho_2]}{\sigma^2}}\Bigg)\Bigg]\,dv
\end{aligned}
$$

$$
\begin{aligned}
I_k(-1,-1) \;=\;& \log\left(p_{-1,1}+p_{-1,-1}\right)\\
&+\int_{\Re}\frac{1}{\sqrt{2\pi}}e^{-v^2/2}\log\Bigg(p_{-1,1}e^{\frac{\sqrt{2}s\sigma(\rho_1-\rho_2)v+2s^2(\rho_1^2-\rho_2^2+\rho_1\rho_2)}{\sigma^2}}\\
&+p_{-1,-1}e^{\frac{\sqrt{2}s\sigma(\rho_1+\rho_2)v+2s^2(\rho_1^2+\rho_2^2+\rho_1\rho_2)}{\sigma^2}}\Bigg)\,dv\\
&-\int_{\Re}\frac{1}{\sqrt{2\pi}}e^{-v^2/2}\Bigg[\log\Bigg(p_{1,1}e^{\frac{-\sqrt{2}s\sigma(\rho_1+\rho_2)v-2s^2[\rho_1^2+\rho_2^2+3\rho_1\rho_2]}{\sigma^2}}\\
&+p_{1,-1}e^{\frac{-\sqrt{2}s\sigma(\rho_1-\rho_2)v-2s^2[\rho_1^2-\rho_2^2-\rho_1\rho_2]}{\sigma^2}}+p_{-1,1}e^{\frac{\sqrt{2}s\sigma(\rho_1-\rho_2)v+2s^2[\rho_1^2-\rho_2^2+\rho_1\rho_2]}{\sigma^2}}\\
&+p_{-1,-1}e^{\frac{\sqrt{2}s\sigma(\rho_1+\rho_2)v+2s^2[\rho_1^2+\rho_2^2+\rho_1\rho_2]}{\sigma^2}}\Bigg)\Bigg]\,dv
\end{aligned}
$$

where for $I_k(-1,-1)$ and $I_k(-1,1)$, we take $v=-u$.

Numerically evaluation of the above four terms shows that for positive ρ_1 and ρ_2, the largest $C_k(S)$ occurs at $p_{1,1} = p_{-1,-1} = 1/2$ and $p_{1,-1} = p_{-1,1} = 0$, in which case

$$C_k(S) \begin{cases} = I_k(1,1) = I_k(-1,-1) \\ \geq I_k(1,-1) \\ \geq I_k(-1,1). \end{cases}$$

An interpretation of the result is that since all of the four possible inputs, i.e., $(+s,+s)$, $(+s,-s)$, $(-s,+s)$ and $(-s,-s)$, for (x_k, x_{k-1}) share the same power, and since $f(y_k|x_k, x_{k-1})$ for different (x_k, x_{k-1}) has common variance but is with aligned means (cf. Fig. 4.4), it is advantageous to use the two inputs that are farthest to each other. When $\rho_1 > 0$ and $\rho_2 > 0$, these two inputs should be $(x_k, x_{k-1}) = (s, s)$ and $(x_k, x_{k-1}) = (-s, -s)$. For general ρ_1 and ρ_2, the two inputs become $(s \cdot \operatorname{sgn}(\rho_1), s \cdot \operatorname{sgn}(\rho_2))$ and $(-s \cdot \operatorname{sgn}(\rho_1), -s \cdot \operatorname{sgn}(\rho_2))$, where $\operatorname{sgn}(\cdot)$ represents the sign function.

In summary, for $d_{k,1} = \rho_1 d_k$ and $d_{k,2} = \rho_2 d_k$, we can transform the original complex channel to its equivalent real channel as $f(\tilde{y}_{k,r}|x_k, x_{k-1})$ is Gaussian distributed with mean $\sum_{i=1}^{M} \rho_i x_{k-i+1}$ and variance $(1/2)\sigma^2 = (\sigma_z^2 + S\sum_{i=1}^{M} D_{k,i})/(2|d_k|^2)$. The input statistics that places equal probability mass on $(x_k, x_{k-1}) = (s \cdot \operatorname{sgn}(\rho_1), s \cdot \operatorname{sgn}(\rho_2))$ and $(x_k, x_{k-1}) = (-s \cdot \operatorname{sgn}(\rho_1), -s \cdot \operatorname{sgn}(\rho_2))$ then maximizes $I(X_k, \tilde{Y}_{k,r})$. Hence,

$$\begin{aligned} C_k(S) &= I(X_k; \tilde{Y}_{k,r}) \\ &= \sum_{x_{k-1}\in\mathcal{X}} \sum_{x_k\in\mathcal{X}} \int_{\Re} P_{X_k,X_{k-1}}(x_k, x_{k-1}) f(\tilde{y}_{k,r}|x_k, x_{k-1}) \\ &\quad \log \frac{\sum_{\bar{x}_{k-1}\in\mathcal{X}} P_{X_k,X_{k-1}}(x_k, \bar{x}_{k-1}) f(\tilde{y}_{k,r}|x_k, \bar{x}_{k-1})}{\left(\sum_{\hat{x}_{k-1}\in\mathcal{X}} P_{X_k,X_{k-1}}(x_k, \hat{x}_{k-1})\right)\left(\sum_{x'_{k-1}\in\mathcal{X}} \sum_{x'_k\in\mathcal{X}} P_{X_k,X_{k-1}}(x'_k, x'_{k-1}) f(\tilde{y}_{k,r}|x'_k, x'_{k-1})\right)} d\tilde{y}_{k,r} \end{aligned}$$

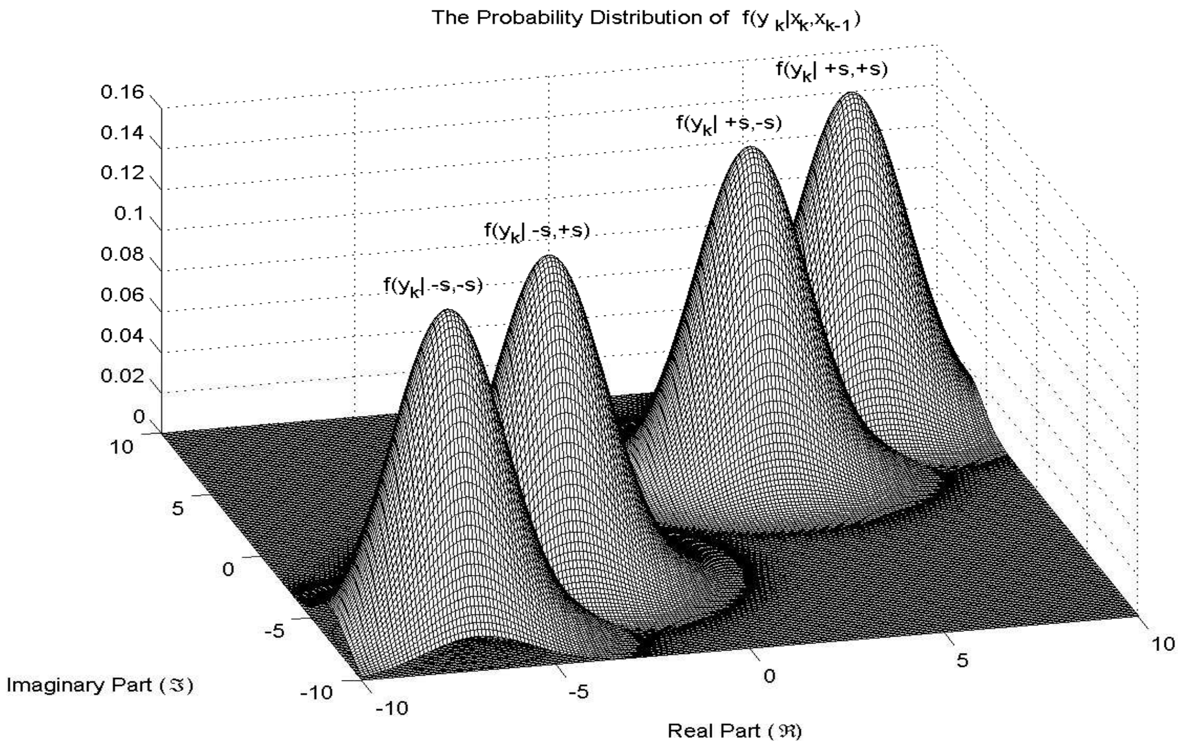

Figure 4.4: $f(y_k|x_k, x_{k-1})$ for four different (x_k, x_{k-1}). The parameters used in this figure are $s = 1$, $d_{k,1} = 5.25 + j5.25$, $d_{k,2} = 1.75 + j1.75$ and $D_{k,1} + D_{k,2} + \sigma_z^2 = 2$.

$$
\begin{aligned}
&= \int_{\Re} \frac{1}{2} \cdot \frac{1}{\sqrt{\pi\sigma^2}} e^{-\frac{(\tilde{y}_{k,r}-s\rho)^2}{\sigma^2}} \log \frac{\frac{1}{\sqrt{\pi\sigma^2}} e^{-\frac{(\tilde{y}_{k,r}-s\rho)^2}{\sigma^2}}}{\frac{1}{2} \cdot \frac{1}{\sqrt{\pi\sigma^2}} e^{-\frac{(\tilde{y}_{k,r}-s\rho)^2}{\sigma^2}} + \frac{1}{2} \cdot \frac{1}{\sqrt{\pi\sigma^2}} e^{-\frac{(\tilde{y}_{k,r}+s\rho)^2}{\sigma^2}}} d\tilde{y}_{k,r} \\
&+ \int_{\Re} \frac{1}{2} \cdot \frac{1}{\sqrt{\pi\sigma^2}} e^{-\frac{(\tilde{y}_{k,r}+s\rho)^2}{\sigma^2}} \log \frac{\frac{1}{\sqrt{\pi\sigma^2}} e^{-\frac{(\tilde{y}_{k,r}+s\rho)^2}{\sigma^2}}}{\frac{1}{2} \cdot \frac{1}{\sqrt{\pi\sigma^2}} e^{-\frac{(\tilde{y}_{k,r}-s\rho)^2}{\sigma^2}} + \frac{1}{2} \cdot \frac{1}{\sqrt{\pi\sigma^2}} e^{-\frac{(\tilde{y}_{k,r}+s\rho)^2}{\sigma^2}}} d\tilde{y}_{k,r} \\
&= \frac{2S}{\sigma^2/\rho^2} - \int_{\Re} \frac{1}{\sqrt{2\pi}} e^{-t^2/2} \left[\log \left(\cosh \left(\frac{\sqrt{2S}}{\sigma/\rho} t + \frac{2S}{\sigma^2/\rho^2} \right) \right) \right] dt, \qquad (4.42)
\end{aligned}
$$

where $\rho = \sum_{i=1}^{M} |\rho_i|$; hence,

$$
\frac{\sigma^2}{\rho^2} = \frac{\sigma_z^2 + S \sum_{i=1}^{M} D_{k,i}}{\left(\sum_{i=1}^{M} |d_{k,i}| \right)^2}. \qquad (4.43)
$$

Finally, by replacing

$$\frac{\sigma^2}{\rho^2} = \frac{\sigma_z^2 + \frac{S}{1-|\alpha|^2}\sum_{i=1}^{M} C_{i,i}}{\left(\frac{1}{|1-\alpha|}\sum_{i=1}^{M}|d_i|\right)^2}, \tag{4.44}$$

where $\boldsymbol{C}$ is diagonal with diagonal elements $\{C_{i,i}\}_{i=1}^{M}$ and $\boldsymbol{d} = \begin{bmatrix} d_1 & d_2 & \cdots & d_M \end{bmatrix}^T$, we yield $C_\infty(S)$ for the case of $M = 2$.

4.5 Independent bound for general M

Following similar argument to the previous section, we can generalize $C_k(S)$ for general M below.

Theorem 4.1 *Let M be the true memory order of the channel. Assume that there exists a complex number d_k such that $d_{k,i} = \rho_i d_k$ for some real number ρ_i for every $1 \leq i \leq M$, where $\boldsymbol{d}_k = \begin{bmatrix} d_{k,1} & d_{k,2} & \cdots & d_{k,M} \end{bmatrix}$. Also, $\boldsymbol{C}$ is diagonal. Then, the component independent bound $C_k(S)$ is given by:*

$$C_k(S) = \frac{2S}{\delta^2} - \int_{\Re} \frac{1}{\sqrt{2\pi}} e^{-t^2/2} \left[\log\left(\cosh\left(\frac{\sqrt{2S}}{|\delta|} t + \frac{2S}{\delta^2} \right) \right) \right] dt, \tag{4.45}$$

where

$$\delta^2 = \frac{\sigma_z^2 + S\sum_{i=1}^{M} D_{k,i}}{\left(\sum_{i=1}^{M}|d_{k,i}|\right)^2}. \tag{4.46}$$

Furthermore, the ultimate independent bound $C_\infty(S)$ has the same form as (4.45) *with*

$$\delta^2 = \frac{\sigma_z^2 + \frac{S}{1-|\alpha|^2}\sum_{i=1}^{M} C_{i,i}}{\left(\frac{1}{|1-\alpha|}\sum_{i=1}^{M}|d_i|\right)^2}. \tag{4.47}$$

Figure 4.5 shows the independent bounds for Gauss-Markov channels of different memory orders. By intuition, for fixed $C_{i,i}$ and d_i, the higher the channel memory order, the more involved in received vector $\boldsymbol{y}$ at the receiver end. Thus, it is reasonable to expect a lower capacity for larger M. However, the independent bound shows that $C_\infty(S)$ grows as M

increases. This indicates that in the case we considered, the independent bound could be looser for higher M.

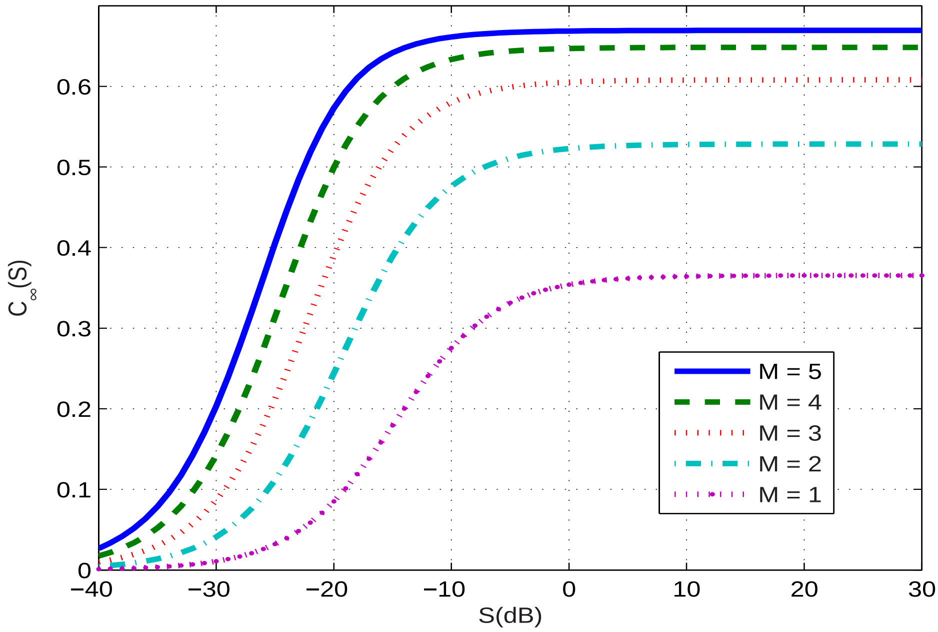

Figure 4.5: Illustration of $C_\infty(S)$. Other parameters for Gauss-Markov channels are $C_{1,1} = C_{2,2} = C_{3,3} = C_{4,4} = C_{5,5} = 10$, $d_1 = d_2 = d_3 = d_4 = d_5 = 1$, $\alpha = 0.7$ and $\sigma_z^2 = 1$.

Figure 4.6 shows the independent bounds for Gauss-Markov channels for another set of parameters. Since $C_{i,i}$ is getting larger as i increases, the influence to the current output y_k by the distant input x_{k-M+1} grows (from the "power" standpoint). In such case, the channel output should be much more involved than the case considered in Fig. 4.5, if M increases. Following the intuition, the independent bounds decrease as M increases for SNR beyond -10 dB. Nevertheless, when SNR is below -10 dB, the independent bounds become messy in channel memory order M, which indicates that they could be loose in this SNR range.

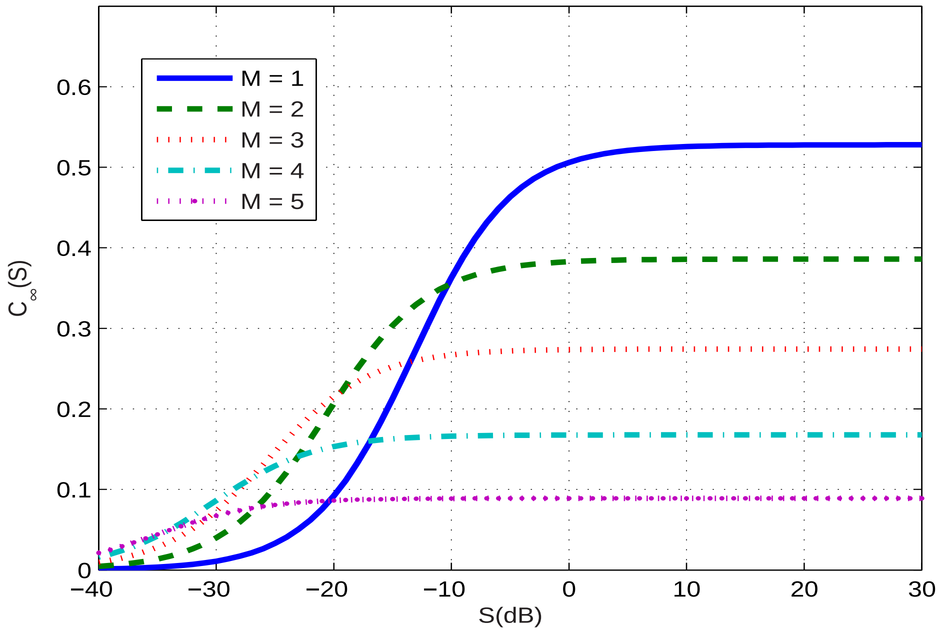

Figure 4.6: Illustration of $C_\infty(S)$. Other parameters for Gauss-Markov channels are $C_{1,1} = 10^{0.7}$, $C_{2,2} = 10^{1.5}$, $C_{3,3} = 10^{2}$, $C_{4,4} = 10^{2.5}$, $C_{5,5} = 10^{3}$, $d_1 = d_2 = d_3 = d_4 = d_5 = 1$, $\alpha = 0.7$ and $\sigma_z^2 = 1$.

4.6 The lower bound of bit error probability

In general, the bit error rate (P_b) (for information bits) is used as the typical performance measure in practical communication system. Referring to [3], we can obtain a lower bound of this typical performance measure by means of the rate-distortion function and the capacity bound just derived.

First, we need to derive the average E_b/N_0 for the Gauss-Markov fading channel. The average SNR is equal to:

$$\begin{aligned}
\overline{SNR} &= \lim_{n\to\infty} \frac{\sum_{i=1}^{n} E[|\boldsymbol{x}_k^T \boldsymbol{h}_k|^2]}{\sum_{i=1}^{n} E[z_k^2]} \\
&= \lim_{n\to\infty} \frac{\sum_{i=1}^{n} \sum_{i=1}^{M} E[|x_{k-i+1} h_{k,i}|^2]}{n\sigma_z^2} \\
&= \lim_{n\to\infty} \frac{S \sum_{i=1}^{n} \sum_{i=1}^{M} E[|h_{k,i}|^2]}{n\sigma_z^2} \\
&= \lim_{n\to\infty} \frac{S \sum_{i=1}^{n} (\boldsymbol{d}_k^H \boldsymbol{d}_k + \mathrm{tr}(\boldsymbol{D}_k))}{n\sigma_z^2} \\
&= \frac{S}{\sigma_z^2} \lim_{n\to\infty} \frac{1}{n} \sum_{i=1}^{n} (\boldsymbol{d}_k^H \boldsymbol{d}_k + \mathrm{tr}(\boldsymbol{D}_k)) \\
&= \frac{S}{\sigma_z^2} (\boldsymbol{d}_\infty^H \boldsymbol{d}_\infty + \mathrm{tr}(\boldsymbol{D}_\infty)) \\
&= \frac{S}{\sigma_z^2} \left(\frac{1}{|1-\alpha|^2} \boldsymbol{d}^H \boldsymbol{d} + \frac{1}{1-|\alpha|^2} \mathrm{tr}(\boldsymbol{C}) \right). \qquad (4.48)
\end{aligned}$$

Hence,

$$\gamma_b \triangleq \frac{E_b}{N_0} = \frac{\overline{SNR}}{R} = \frac{1}{R} \frac{S}{\sigma_z^2} \left(\frac{1}{|1-\alpha|^2} \boldsymbol{d}^H \boldsymbol{d} + \frac{1}{1-|\alpha|^2} \mathrm{tr}(\boldsymbol{C}) \right), \qquad (4.49)$$

where R is the channel code rate.

Referring to [3], the rate distortion for binary input and Hamming additive distortion measure is

$$R(D) = \begin{cases} \log(2) - H_b(D), & \text{for } 0 \le D \le 0.5 \\ 0, & \text{for } D > 0.5 \end{cases} \qquad (4.50)$$

According to the joint source-channel coding theorem, good codes exists when

$$R(P_b) < \frac{C(S)}{R}. \qquad (4.51)$$

Therefore, we obtain a lower bound for P_b as:

$$H_b(P_b) > \log(2) - \frac{1}{R} C_\infty(S). \qquad (4.52)$$

The lower bounds of bit error rate P_b corresponding to those in Figs. 4.5 and 4.6 are summarized in Figs.4.7 and 4.8.

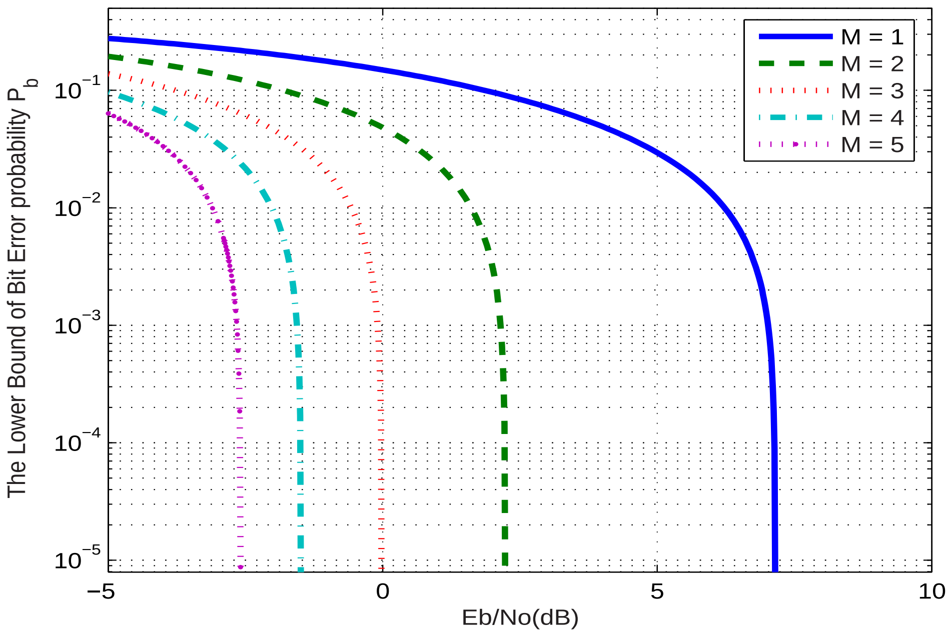

Figure 4.7: Illustration of lower bounds for P_b. The code rate adopted is $R = 1/3$. Other parameters used in this figure are the same as those in Fig. 4.5.

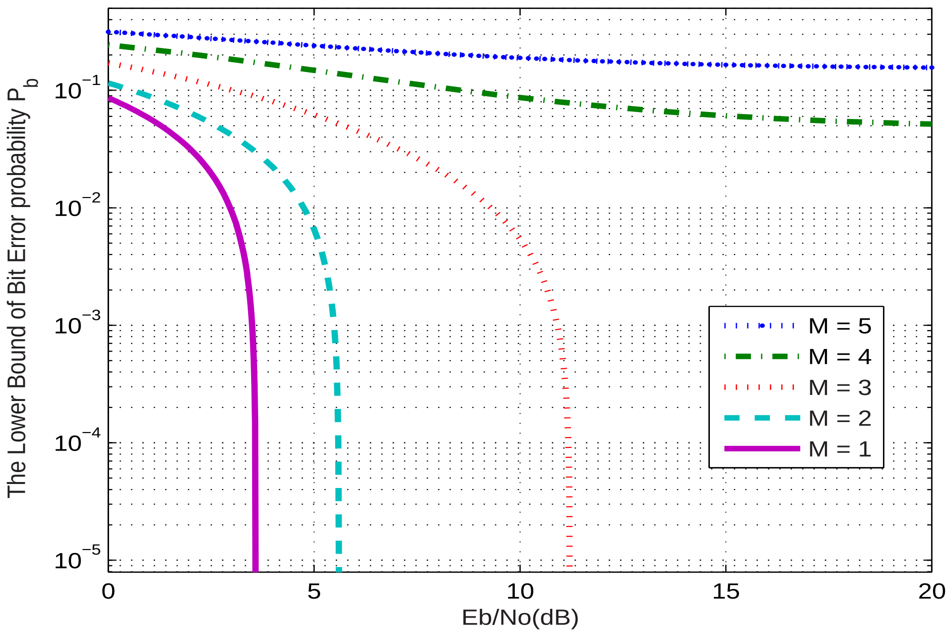

Figure 4.8: Illustration of lower bounds for P_b. The code rate adopted is $R = 1/3$. Other parameters used in this figure are the same as those in Fig. 4.6.

Chapter 5

Conclusions

In this thesis, we have remarked on four different definitions of channel capacities according to the transmiter/receiver with/without channel state information. We then turn to the derivation of the independent bounds for the channel capacity without CSI in both transmitter and receiver. We then found that if there is no LOS signal existing, the capacity of the blind-CSI system will be reduced to zero.

Appendix A

The derivation of probability distribution $f(\boldsymbol{y}|\boldsymbol{x})$

Recall that:

$$f\left(\boldsymbol{y}|\boldsymbol{x},\boldsymbol{H}\right)=\frac{1}{\left(\pi\sigma_z^2\right)^n}\prod_{k=1}^{n}\exp\left(-\frac{\left|y_k-\boldsymbol{x}_k^T\boldsymbol{h}_k\right|^2}{\sigma_z^2}\right), \tag{A.1}$$

and

$$\begin{aligned} f(\boldsymbol{H}) &= f(\boldsymbol{h}_1)\prod_{k=2}^{n} f(\boldsymbol{h}_k|\boldsymbol{h}_{k-1}) \\ &= \frac{1}{|\pi\boldsymbol{C}|^n}\prod_{k=1}^{n}\exp\left\{-\left(\boldsymbol{h}_k-\alpha\boldsymbol{h}_{k-1}-\boldsymbol{d}\right)^H\boldsymbol{C}^{-1}\left(\boldsymbol{h}_k-\alpha\boldsymbol{h}_{k-1}-\boldsymbol{d}\right)\right\}. \end{aligned} \tag{A.2}$$

Hence,

$$\begin{aligned}
f(\boldsymbol{y}|\boldsymbol{x}) &= \int_{\boldsymbol{\mathcal{H}}} f(\boldsymbol{y}|\boldsymbol{x},\boldsymbol{H})\, f(\boldsymbol{H})\, d\boldsymbol{H} \\
&= \int_{\mathcal{C}^M} \cdots \int_{\mathcal{C}^M} \frac{1}{\pi^n \sigma_z^{2n}} \prod_{k=1}^{n} \exp\left(-\frac{\left|y_k - \boldsymbol{x}_k^T \boldsymbol{h}_k\right|^2}{\sigma_z^2}\right) \\
&\quad \frac{1}{\pi^{nM} |\boldsymbol{C}|^n} \prod_{k=1}^{n} \exp\left\{-\left(\boldsymbol{h}_k - \alpha \boldsymbol{h}_{k-1} - \boldsymbol{d}\right)^H \boldsymbol{C}^{-1} \left(\boldsymbol{h}_k - \alpha \boldsymbol{h}_{k-1} - \boldsymbol{d}\right)\right\} d\boldsymbol{h}_1 \cdots d\boldsymbol{h}_n \\
&= \frac{1}{\pi^{n(M+1)} \sigma_z^{2n} |\boldsymbol{C}|^n} \int_{\mathcal{C}^M} \cdots \int_{\mathcal{C}^M} \prod_{k=2}^{n} e^{-\frac{\left|y_k - \boldsymbol{x}_k^T \boldsymbol{h}_k\right|^2}{\sigma_z^2}} \\
&\quad \prod_{k=3}^{n} e^{-(\boldsymbol{h}_k - \alpha \boldsymbol{h}_{k-1} - \boldsymbol{d})^H \boldsymbol{C}^{-1} (\boldsymbol{h}_k - \alpha \boldsymbol{h}_{k-1} - \boldsymbol{d})} \\
&\quad \left(\int_{\mathcal{C}^M} e^{-\frac{\left|y_1 - \boldsymbol{x}_1^T \boldsymbol{h}_1\right|^2}{\sigma_z^2}} e^{-(\boldsymbol{h}_1 - \alpha \boldsymbol{h}_0 - \boldsymbol{d})^H \boldsymbol{C}^{-1} (\boldsymbol{h}_1 - \alpha \boldsymbol{h}_0 - \boldsymbol{d})} e^{-(\boldsymbol{h}_2 - \alpha \boldsymbol{h}_1 - \boldsymbol{d})^H \boldsymbol{C}^{-1} (\boldsymbol{h}_2 - \alpha \boldsymbol{h}_1 - \boldsymbol{d})} d\boldsymbol{h}_1\right) \\
&\quad d\boldsymbol{h}_2 \cdots d\boldsymbol{h}_n.
\end{aligned}$$

The exponent in the inner integral can be re-written as:

$$
\begin{aligned}
& \frac{1}{\sigma_z^2}|y_1 - \boldsymbol{x}_1^T \boldsymbol{h}_1|^2 + (\boldsymbol{h}_1 - \alpha \boldsymbol{h}_0 - \boldsymbol{d})^H \boldsymbol{C}^{-1} (\boldsymbol{h}_1 - \alpha \boldsymbol{h}_0 - \boldsymbol{d}) \\
& + (\boldsymbol{h}_2 - \alpha \boldsymbol{h}_1 - \boldsymbol{d})^H \boldsymbol{C}^{-1} (\boldsymbol{h}_2 - \alpha \boldsymbol{h}_1 - \boldsymbol{d}) \\
= & \frac{1}{\sigma_z^2}(y_1^* - \boldsymbol{h}_1^H \boldsymbol{x}_1^*)(y_1 - \boldsymbol{x}_1^T \boldsymbol{h}_1) + (\boldsymbol{h}_1^H - (\alpha \boldsymbol{h}_0 + \boldsymbol{d})^H)\boldsymbol{C}^{-1}(\boldsymbol{h}_1 - (\alpha \boldsymbol{h}_0 + \boldsymbol{d})) \\
& +((\boldsymbol{h}_2 - \boldsymbol{d})^H - \alpha^* \boldsymbol{h}_1^H)\boldsymbol{C}^{-1}((\boldsymbol{h}_2 - \boldsymbol{d}) - \alpha \boldsymbol{h}_1) \\
= & \frac{1}{\sigma_z^2}\left[|y_1|^2 - y_1^* \boldsymbol{x}_1^T \boldsymbol{h}_1 - y_1 \boldsymbol{h}_1^H \boldsymbol{x}_1 + \boldsymbol{h}_1^H \boldsymbol{x}_1 \boldsymbol{x}_1^T \boldsymbol{h}_1\right] \\
& + \left[\boldsymbol{h}_1^H \boldsymbol{C}^{-1} \boldsymbol{h}_1 - (\alpha \boldsymbol{h}_0 + \boldsymbol{d})^H \boldsymbol{C}^{-1} \boldsymbol{h}_1 - \boldsymbol{h}_1^H \boldsymbol{C}^{-1}(\alpha \boldsymbol{h}_0 + \boldsymbol{d}) + (\alpha \boldsymbol{h}_0 + \boldsymbol{d})^H \boldsymbol{C}^{-1}(\alpha \boldsymbol{h}_0 + \boldsymbol{d})\right] \\
& + \left[(\boldsymbol{h}_2 - \boldsymbol{d})^H \boldsymbol{C}^{-1}(\boldsymbol{h}_2 - \boldsymbol{d}) - \alpha(\boldsymbol{h}_2 - \boldsymbol{d})^H \boldsymbol{C}^{-1} \boldsymbol{h}_1 - \alpha^* \boldsymbol{h}_1^H \boldsymbol{C}^{-1}(\boldsymbol{h}_2 - \boldsymbol{d}) + |\alpha|^2 \boldsymbol{h}_1^H \boldsymbol{C}^{-1} \boldsymbol{h}_1\right] \\
= & \boldsymbol{h}_1^H \left(\frac{\boldsymbol{x}_1 \boldsymbol{x}_1^T}{\sigma_z^2} + \boldsymbol{C}^{-1} + |\alpha|^2 \boldsymbol{C}^{-1}\right) \boldsymbol{h}_1 - \left(\frac{y_1^* \boldsymbol{x}_1^T}{\sigma_z^2} + (\alpha \boldsymbol{h}_0 + \boldsymbol{d})^H \boldsymbol{C}^{-1} + \alpha(\boldsymbol{h}_2 - \boldsymbol{d})^H \boldsymbol{C}^{-1}\right) \boldsymbol{h}_1 \\
& - \boldsymbol{h}_1^H \left(\frac{y_1 \boldsymbol{x}_1}{\sigma_z^2} + \boldsymbol{C}^{-1}(\alpha \boldsymbol{h}_0 + \boldsymbol{d}) + \alpha^* \boldsymbol{C}^{-1}(\boldsymbol{h}_2 - \boldsymbol{d})\right) \\
& + \left(\frac{|y_1|^2}{\sigma_z^2} + (\alpha \boldsymbol{h}_0 + \boldsymbol{d})^H \boldsymbol{C}^{-1}(\alpha \boldsymbol{h}_0 + \boldsymbol{d}) + (\boldsymbol{h}_2 - \boldsymbol{d})^H \boldsymbol{C}^{-1}(\boldsymbol{h}_2 - \boldsymbol{d})\right) \\
= & \boldsymbol{h}_1^H \boldsymbol{G}_1^{-1} \boldsymbol{h}_1 - \boldsymbol{h}_1^H \boldsymbol{G}_1^{-1} \boldsymbol{g}_1 - \boldsymbol{g}_1^H \boldsymbol{G}_1^{-1} \boldsymbol{h}_1 \\
& + \left(\frac{|y_1|^2}{\sigma_z^2} + (\alpha \boldsymbol{h}_0 + \boldsymbol{d})^H \boldsymbol{C}^{-1}(\alpha \boldsymbol{h}_0 + \boldsymbol{d}) + (\boldsymbol{h}_2 - \boldsymbol{d})^H \boldsymbol{C}^{-1}(\boldsymbol{h}_2 - \boldsymbol{d})\right) \\
= & (\boldsymbol{h}_1 - \boldsymbol{g}_1)^H \boldsymbol{G}_1^{-1}(\boldsymbol{h}_1 - \boldsymbol{g}_1) - \boldsymbol{g}_1^H \boldsymbol{G}_1^{-1} \boldsymbol{g}_1 \\
& + \left(\frac{|y_1|^2}{\sigma_z^2} + (\alpha \boldsymbol{h}_0 + \boldsymbol{d})^H \boldsymbol{C}^{-1}(\alpha \boldsymbol{h}_0 + \boldsymbol{d}) + (\boldsymbol{h}_2 - \boldsymbol{d})^H \boldsymbol{C}^{-1}(\boldsymbol{h}_2 - \boldsymbol{d})\right),
\end{aligned}
$$

where

$$
\boldsymbol{g}_1 \triangleq \boldsymbol{G}_1 \left(\boldsymbol{q}_1 + \alpha^* \boldsymbol{C}^{-1}(\boldsymbol{h}_2 - \boldsymbol{d})\right), \quad \boldsymbol{G}_1^{-1} \triangleq \frac{\boldsymbol{x}_1^* \boldsymbol{x}_1^T}{\sigma_z^2} + (1 + |\alpha|^2)\boldsymbol{C}^{-1},
$$

and

$$
\boldsymbol{q}_1 \triangleq \frac{y_1 \boldsymbol{x}_1^*}{\sigma_z^2} + \boldsymbol{C}^{-1}\boldsymbol{d} + \alpha \boldsymbol{C}^{-1}\boldsymbol{h}_0.
$$

Since

$$
\int_{\mathcal{C}^M} e^{-(\boldsymbol{h}_1 - \boldsymbol{g}_1)^H \boldsymbol{G}_1^{-1}(\boldsymbol{h}_1 - \boldsymbol{g}_1)} d\boldsymbol{h}_1 = |\pi \boldsymbol{G}_1|,
$$

the exponent terms remained after the integration of the inner integral are given by:

$$
\begin{aligned}
& -\boldsymbol{g}_1^H \boldsymbol{G}_1^{-1} \boldsymbol{g}_1 + \left(\frac{|y_1|^2}{\sigma_z^2} + (\alpha \boldsymbol{h}_0 + \boldsymbol{d})^H \boldsymbol{C}^{-1} (\alpha \boldsymbol{h}_0 + \boldsymbol{d}) + (\boldsymbol{h}_2 - \boldsymbol{d})^H \boldsymbol{C}^{-1} (\boldsymbol{h}_2 - \boldsymbol{d}) \right) \\
= & - \left(\boldsymbol{q}_1 + \alpha^* \boldsymbol{C}^{-1} (\boldsymbol{h}_2 - \boldsymbol{d}) \right)^H \boldsymbol{G}_1 \left(\boldsymbol{q}_1 + \alpha^* \boldsymbol{C}^{-1} (\boldsymbol{h}_2 - \boldsymbol{d}) \right) \\
& + \left(\frac{|y_1|^2}{\sigma_z^2} + (\alpha \boldsymbol{h}_0 + \boldsymbol{d})^H \boldsymbol{C}^{-1} (\alpha \boldsymbol{h}_0 + \boldsymbol{d}) + (\boldsymbol{h}_2 - \boldsymbol{d})^H \boldsymbol{C}^{-1} (\boldsymbol{h}_2 - \boldsymbol{d}) \right) \\
= & -\boldsymbol{q}_1^H \boldsymbol{G}_1 \boldsymbol{q}_1 - \alpha^* \boldsymbol{q}_1^H \boldsymbol{G}_1 \boldsymbol{C}^{-1} (\boldsymbol{h}_2 - \boldsymbol{d}) - \alpha (\boldsymbol{h}_2 - \boldsymbol{d})^H \boldsymbol{C}^{-1} \boldsymbol{G}_1 \boldsymbol{q}_1 \\
& -|\alpha|^2 (\boldsymbol{h}_2 - \boldsymbol{d})^H \boldsymbol{C}^{-1} \boldsymbol{G}_1 \boldsymbol{C}^{-1} (\boldsymbol{h}_2 - \boldsymbol{d}) \\
& + \left(\frac{|y_1|^2}{\sigma_z^2} + (\alpha \boldsymbol{h}_0 + \boldsymbol{d})^H \boldsymbol{C}^{-1} (\alpha \boldsymbol{h}_0 + \boldsymbol{d}) + (\boldsymbol{h}_2 - \boldsymbol{d})^H \boldsymbol{C}^{-1} (\boldsymbol{h}_2 - \boldsymbol{d}) \right) \\
= & (\boldsymbol{h}_2 - \boldsymbol{d})^H \left(\boldsymbol{C}^{-1} - |\alpha|^2 \boldsymbol{C}^{-1} \boldsymbol{G}_1 \boldsymbol{C}^{-1} \right) (\boldsymbol{h}_2 - \boldsymbol{d}) - \alpha^* \boldsymbol{q}_1^H \boldsymbol{G}_1 \boldsymbol{C}^{-1} (\boldsymbol{h}_2 - \boldsymbol{d}) - \alpha (\boldsymbol{h}_2 - \boldsymbol{d})^H \boldsymbol{C}^{-1} \boldsymbol{G}_1 \boldsymbol{q}_1 \\
& -\boldsymbol{q}_1^H \boldsymbol{G}_1 \boldsymbol{q}_1 + \frac{|y_1|^2}{\sigma_z^2} + (\alpha \boldsymbol{h}_0 + \boldsymbol{d})^H \boldsymbol{C}^{-1} (\alpha \boldsymbol{h}_0 + \boldsymbol{d}) \\
= & (\boldsymbol{h}_2 - \alpha \bar{\boldsymbol{h}}_1 - \boldsymbol{d})^H \boldsymbol{\Sigma}_1^{-1} (\boldsymbol{h}_2 - \alpha \bar{\boldsymbol{h}}_1 - \boldsymbol{d}) - \boldsymbol{q}_1^H \boldsymbol{G}_1 \boldsymbol{q}_1 \\
& -|\alpha|^2 \bar{\boldsymbol{h}}_1^H \boldsymbol{\Sigma}_1^{-1} \bar{\boldsymbol{h}}_1 + \frac{|y_1|^2}{\sigma_z^2} + (\alpha \boldsymbol{h}_0 + \boldsymbol{d})^H \boldsymbol{C}^{-1} (\alpha \boldsymbol{h}_0 + \boldsymbol{d}),
\end{aligned}
$$

where

$$
\boldsymbol{\Sigma}_1^{-1} \triangleq \boldsymbol{C}^{-1} - |\alpha|^2 \boldsymbol{C}^{-1} \boldsymbol{G}_1 \boldsymbol{C}^{-1} \quad \text{and} \quad \bar{\boldsymbol{h}}_1 \triangleq \boldsymbol{\Sigma}_1 \boldsymbol{C}^{-1} \boldsymbol{G}_1 \boldsymbol{q}_1.
$$

Consequently,

$$
\begin{aligned}
f(\boldsymbol{y}|\boldsymbol{x}) = & \frac{e^{|\alpha|^2 \bar{\boldsymbol{h}}_1^H \boldsymbol{\Sigma}_1^{-1} \bar{\boldsymbol{h}}_1 - (\alpha \boldsymbol{h}_0 + \boldsymbol{d})^H \boldsymbol{C}^{-1} (\alpha \boldsymbol{h}_0 + \boldsymbol{d})}}{\pi^{M(n-1)+n} \sigma_z^{2n} |\boldsymbol{C}|^n} e^{-|y_1|^2/\sigma_z^2} |\boldsymbol{G}_1| e^{\boldsymbol{q}_1^H \boldsymbol{G}_1 \boldsymbol{q}_1} \int_{\mathcal{C}^M} \cdots \int_{\mathcal{C}^M} \prod_{k=3}^{n} e^{-\frac{\left| y_k - \boldsymbol{x}_k^T \boldsymbol{h}_k \right|^2}{\sigma_z^2}} \\
& \prod_{k=4}^{n} e^{-(\boldsymbol{h}_k - \alpha \boldsymbol{h}_{k-1} - \boldsymbol{d})^H \boldsymbol{C}^{-1} (\boldsymbol{h}_k - \alpha \boldsymbol{h}_{k-1} - \boldsymbol{d})} \\
& \left(\int_{\mathcal{C}^M} e^{-\frac{\left| y_2 - \boldsymbol{x}_2^T \boldsymbol{h}_2 \right|^2}{\sigma_z^2}} e^{-\left(\boldsymbol{h}_2 - \alpha \bar{\boldsymbol{h}}_1 - \boldsymbol{d} \right)^H \boldsymbol{\Sigma}_1^{-1} \left(\boldsymbol{h}_2 - \alpha \bar{\boldsymbol{h}}_1 - \boldsymbol{d} \right)} e^{-(\boldsymbol{h}_3 - \alpha \boldsymbol{h}_2 - \boldsymbol{d})^H \boldsymbol{C}^{-1} (\boldsymbol{h}_3 - \alpha \boldsymbol{h}_2 - \boldsymbol{d})} d\boldsymbol{h}_2 \right) \\
& d\boldsymbol{h}_3 \cdots d\boldsymbol{h}_n. \qquad \text{(A.3)}
\end{aligned}
$$

Similarly, the exponent in the inner integral in (A.3) can be re-written as:

$$\begin{aligned}
& \frac{1}{\sigma_z^2}|y_2 - \boldsymbol{x}_2^T \boldsymbol{h}_2|^2 + \left(\boldsymbol{h}_2 - \alpha\bar{\boldsymbol{h}}_1 - \boldsymbol{d}\right)^H \boldsymbol{\Sigma}_1^{-1} \left(\boldsymbol{h}_2 - \alpha\bar{\boldsymbol{h}}_1 - \boldsymbol{d}\right) \\
& + (\boldsymbol{h}_3 - \alpha\boldsymbol{h}_2 - \boldsymbol{d})^H \boldsymbol{C}^{-1} (\boldsymbol{h}_3 - \alpha\boldsymbol{h}_2 - \boldsymbol{d}) \\
= \; & \frac{1}{\sigma_z^2}(y_2^* - \boldsymbol{h}_2^H \boldsymbol{x}_2^*)(y_2 - \boldsymbol{x}_2^T \boldsymbol{h}_2) + (\boldsymbol{h}_2^H - (\alpha\bar{\boldsymbol{h}}_1 + \boldsymbol{d})^H)\boldsymbol{\Sigma}_1^{-1}(\boldsymbol{h}_2 - (\alpha\bar{\boldsymbol{h}}_1 + \boldsymbol{d})) \\
& +((\boldsymbol{h}_3 - \boldsymbol{d})^H - \alpha^* \boldsymbol{h}_2^H)\boldsymbol{C}^{-1}((\boldsymbol{h}_3 - \boldsymbol{d}) - \alpha\boldsymbol{h}_2) \\
= \; & \frac{1}{\sigma_z^2}\left[|y_2|^2 - y_2^* \boldsymbol{x}_2^T \boldsymbol{h}_2 - y_2 \boldsymbol{h}_2^H \boldsymbol{x}_2 + \boldsymbol{h}_2^H \boldsymbol{x}_2 \boldsymbol{x}_2^T \boldsymbol{h}_2\right] \\
& + \left[\boldsymbol{h}_2^H \boldsymbol{\Sigma}_1^{-1} \boldsymbol{h}_2 - (\alpha\bar{\boldsymbol{h}}_1 + \boldsymbol{d})^H \boldsymbol{\Sigma}_1^{-1} \boldsymbol{h}_2 - \boldsymbol{h}_2^H \boldsymbol{\Sigma}_1^{-1}(\alpha\bar{\boldsymbol{h}}_1 + \boldsymbol{d}) + (\alpha\bar{\boldsymbol{h}}_1 + \boldsymbol{d})^H \boldsymbol{\Sigma}_1^{-1}(\alpha\bar{\boldsymbol{h}}_1 + \boldsymbol{d})\right] \\
& + \left[(\boldsymbol{h}_3 - \boldsymbol{d})^H \boldsymbol{C}^{-1}(\boldsymbol{h}_3 - \boldsymbol{d}) - \alpha(\boldsymbol{h}_3 - \boldsymbol{d})^H \boldsymbol{C}^{-1}\boldsymbol{h}_2 - \alpha^* \boldsymbol{h}_2^H \boldsymbol{C}^{-1}(\boldsymbol{h}_3 - \boldsymbol{d}) + |\alpha|^2 \boldsymbol{h}_2^H \boldsymbol{C}^{-1}\boldsymbol{h}_2\right] \\
= \; & \boldsymbol{h}_2^H \left(\frac{\boldsymbol{x}_2 \boldsymbol{x}_2^T}{\sigma_z^2} + \boldsymbol{\Sigma}_1^{-1} + |\alpha|^2 \boldsymbol{C}^{-1}\right)\boldsymbol{h}_2 - \left(\frac{y_2^* \boldsymbol{x}_2^T}{\sigma_z^2} + (\alpha\bar{\boldsymbol{h}}_1 + \boldsymbol{d})^H \boldsymbol{\Sigma}_1^{-1} + \alpha(\boldsymbol{h}_3 - \boldsymbol{d})^H \boldsymbol{C}^{-1}\right)\boldsymbol{h}_2 \\
& -\boldsymbol{h}_2^H \left(\frac{y_2 \boldsymbol{x}_2}{\sigma_z^2} + \boldsymbol{\Sigma}_1^{-1}(\alpha\bar{\boldsymbol{h}}_1 + \boldsymbol{d}) + \alpha^* \boldsymbol{C}^{-1}(\boldsymbol{h}_3 - \boldsymbol{d})\right) \\
& + \left(\frac{|y_2|^2}{\sigma_z^2} + (\alpha\bar{\boldsymbol{h}}_1 + \boldsymbol{d})^H \boldsymbol{\Sigma}_1^{-1}(\alpha\bar{\boldsymbol{h}}_1 + \boldsymbol{d}) + (\boldsymbol{h}_3 - \boldsymbol{d})^H \boldsymbol{C}^{-1}(\boldsymbol{h}_3 - \boldsymbol{d})\right) \\
= \; & \boldsymbol{h}_2^H \boldsymbol{G}_2^{-1} \boldsymbol{h}_2 - \boldsymbol{h}_2^H \boldsymbol{G}_2^{-1} \boldsymbol{g}_2 - \boldsymbol{g}_2^H \boldsymbol{G}_2^{-1} \boldsymbol{h}_2 \\
& + \left(\frac{|y_2|^2}{\sigma_z^2} + (\alpha\bar{\boldsymbol{h}}_1 + \boldsymbol{d})^H \boldsymbol{\Sigma}_1^{-1}(\alpha\bar{\boldsymbol{h}}_1 + \boldsymbol{d}) + (\boldsymbol{h}_3 - \boldsymbol{d})^H \boldsymbol{C}^{-1}(\boldsymbol{h}_3 - \boldsymbol{d})\right) \\
= \; & (\boldsymbol{h}_2 - \boldsymbol{g}_2)^H \boldsymbol{G}_2^{-1}(\boldsymbol{h}_2 - \boldsymbol{g}_2) - \boldsymbol{g}_2^H \boldsymbol{G}_2^{-1} \boldsymbol{g}_2 \\
& + \left(\frac{|y_2|^2}{\sigma_z^2} + (\alpha\bar{\boldsymbol{h}}_1 + \boldsymbol{d})^H \boldsymbol{\Sigma}_1^{-1}(\alpha\bar{\boldsymbol{h}}_1 + \boldsymbol{d}) + (\boldsymbol{h}_3 - \boldsymbol{d})^H \boldsymbol{C}^{-1}(\boldsymbol{h}_3 - \boldsymbol{d})\right),
\end{aligned}$$

where

$$\boldsymbol{g}_2 \triangleq \boldsymbol{G}_2 \left(\boldsymbol{q}_2 + \alpha^* \boldsymbol{C}^{-1}(\boldsymbol{h}_3 - \boldsymbol{d})\right),$$

$$\boldsymbol{G}_2^{-1} \triangleq \frac{\boldsymbol{x}_2^* \boldsymbol{x}_2^T}{\sigma_z^2} + \boldsymbol{\Sigma}_1^{-1} + |\alpha|^2 \boldsymbol{C}^{-1} = \frac{\boldsymbol{x}_k^* \boldsymbol{x}_k^T}{\sigma_z^2} + (1 + |\alpha|^2)\boldsymbol{C}^{-1} - |\alpha|^2 \boldsymbol{C}^{-1} \boldsymbol{G}_{k-1} \boldsymbol{C}^{-1},$$

and

$$\boldsymbol{q}_2 \triangleq \frac{y_2 \boldsymbol{x}_2^*}{\sigma_z^2} + \boldsymbol{\Sigma}_1^{-1}(\alpha\bar{\boldsymbol{h}}_1 + \boldsymbol{d}) = \frac{y_k \boldsymbol{x}_k^*}{\sigma_z^2} + \boldsymbol{C}^{-1}\boldsymbol{d} + \alpha \boldsymbol{C}^{-1} \boldsymbol{G}_1 \boldsymbol{q}_1 - |\alpha|^2 \boldsymbol{C}^{-1} \boldsymbol{G}_1 \boldsymbol{C}^{-1} \boldsymbol{d}.$$

Since

$$\int_{\mathcal{C}^M} e^{-(\boldsymbol{h}_2 - \boldsymbol{g}_2)^H \boldsymbol{G}_2^{-1}(\boldsymbol{h}_2 - \boldsymbol{g}_2)} d\boldsymbol{h}_2 = |\pi \boldsymbol{G}_2|,$$

the exponent terms remained after the integration of the inner integral are given by:

$$\begin{aligned}
& -\boldsymbol{g}_2^H \boldsymbol{G}_2^{-1} \boldsymbol{g}_2 + \left(\frac{|y_2|^2}{\sigma_z^2} + (\alpha \bar{\boldsymbol{h}}_1 + \boldsymbol{d})^H \boldsymbol{\Sigma}_1^{-1} (\alpha \bar{\boldsymbol{h}}_1 + \boldsymbol{d}) + (\boldsymbol{h}_3 - \boldsymbol{d})^H \boldsymbol{C}^{-1} (\boldsymbol{h}_3 - \boldsymbol{d}) \right) \\
= & -\left(\boldsymbol{q}_2 + \alpha^* \boldsymbol{C}^{-1} (\boldsymbol{h}_3 - \boldsymbol{d}) \right)^H \boldsymbol{G}_2 \left(\boldsymbol{q}_2 + \alpha^* \boldsymbol{C}^{-1} (\boldsymbol{h}_3 - \boldsymbol{d}) \right) \\
& + \left(\frac{|y_2|^2}{\sigma_z^2} + (\alpha \bar{\boldsymbol{h}}_1 + \boldsymbol{d})^H \boldsymbol{\Sigma}_1^{-1} (\alpha \bar{\boldsymbol{h}}_1 + \boldsymbol{d}) + (\boldsymbol{h}_3 - \boldsymbol{d})^H \boldsymbol{C}^{-1} (\boldsymbol{h}_3 - \boldsymbol{d}) \right) \\
= & -\boldsymbol{q}_2^H \boldsymbol{G}_2 \boldsymbol{q}_2 - \alpha^* \boldsymbol{q}_2^H \boldsymbol{G}_2 \boldsymbol{C}^{-1} (\boldsymbol{h}_3 - \boldsymbol{d}) - \alpha (\boldsymbol{h}_3 - \boldsymbol{d})^H \boldsymbol{C}^{-1} \boldsymbol{G}_2 \boldsymbol{q}_2 \\
& -|\alpha|^2 (\boldsymbol{h}_3 - \boldsymbol{d})^H \boldsymbol{C}^{-1} \boldsymbol{G}_2 \boldsymbol{C}^{-1} (\boldsymbol{h}_3 - \boldsymbol{d}) \\
& + \left(\frac{|y_2|^2}{\sigma_z^2} + (\alpha \bar{\boldsymbol{h}}_1 + \boldsymbol{d})^H \boldsymbol{\Sigma}_1^{-1} (\alpha \bar{\boldsymbol{h}}_1 + \boldsymbol{d}) + (\boldsymbol{h}_3 - \boldsymbol{d})^H \boldsymbol{C}^{-1} (\boldsymbol{h}_3 - \boldsymbol{d}) \right) \\
= & (\boldsymbol{h}_3 - \boldsymbol{d})^H \left(\boldsymbol{C}^{-1} - |\alpha|^2 \boldsymbol{C}^{-1} \boldsymbol{G}_2 \boldsymbol{C}^{-1} \right) (\boldsymbol{h}_3 - \boldsymbol{d}) - \alpha^* \boldsymbol{q}_2^H \boldsymbol{G}_2 \boldsymbol{C}^{-1} (\boldsymbol{h}_3 - \boldsymbol{d}) - \alpha (\boldsymbol{h}_3 - \boldsymbol{d})^H \boldsymbol{C}^{-1} \boldsymbol{G}_2 \boldsymbol{q}_2 \\
& -\boldsymbol{q}_2^H \boldsymbol{G}_2 \boldsymbol{q}_2 + \frac{|y_2|^2}{\sigma_z^2} + (\alpha \bar{\boldsymbol{h}}_1 + \boldsymbol{d})^H \boldsymbol{\Sigma}_1^{-1} (\alpha \bar{\boldsymbol{h}}_1 + \boldsymbol{d}) \\
= & (\boldsymbol{h}_3 - \alpha \bar{\boldsymbol{h}}_2 - \boldsymbol{d})^H \boldsymbol{\Sigma}_2^{-1} (\boldsymbol{h}_3 - \alpha \bar{\boldsymbol{h}}_2 - \boldsymbol{d}) - \boldsymbol{q}_2^H \boldsymbol{G}_2 \boldsymbol{q}_2 \\
& -|\alpha|^2 \bar{\boldsymbol{h}}_2^H \boldsymbol{\Sigma}_2^{-1} \bar{\boldsymbol{h}}_2 + \frac{|y_2|^2}{\sigma_z^2} + (\alpha \bar{\boldsymbol{h}}_1 + \boldsymbol{d})^H \boldsymbol{\Sigma}_1^{-1} (\alpha \bar{\boldsymbol{h}}_1 + \boldsymbol{d}),
\end{aligned}$$

where

$$\boldsymbol{\Sigma}_2^{-1} \triangleq \boldsymbol{C}^{-1} - |\alpha|^2 \boldsymbol{C}^{-1} \boldsymbol{G}_2 \boldsymbol{C}^{-1} \quad \text{and} \quad \bar{\boldsymbol{h}}_2 \triangleq \boldsymbol{\Sigma}_2 \boldsymbol{C}^{-1} \boldsymbol{G}_2 \boldsymbol{q}_2.$$

Consequently,

$$\begin{aligned}
f(\boldsymbol{y}|\boldsymbol{x}) = & \frac{\prod_{v=1}^2 e^{|\alpha|^2 \bar{\boldsymbol{h}}_v^H \boldsymbol{\Sigma}_v^{-1} \bar{\boldsymbol{h}}_v - (\alpha \bar{\boldsymbol{h}}_{v-1} + \boldsymbol{d})^H \boldsymbol{\Sigma}_{v-1}^{-1} (\alpha \bar{\boldsymbol{h}}_{v-1} + \boldsymbol{d})}}{\pi^{M(n-2)+n} \sigma_z^{2n} |\boldsymbol{C}|^n} \prod_{u=1}^2 e^{-|y_u|^2/\sigma_z^2} |\boldsymbol{G}_u| e^{\boldsymbol{q}_u^H \boldsymbol{G}_u \boldsymbol{q}_u} \\
& \int_{\mathcal{C}^M} \cdots \int_{\mathcal{C}^M} \prod_{k=4}^n e^{-\frac{\left| y_k - \boldsymbol{x}_k^T \boldsymbol{h}_k \right|^2}{\sigma_z^2}} \prod_{k=5}^n e^{-(\boldsymbol{h}_k - \alpha \boldsymbol{h}_{k-1} - \boldsymbol{d})^H \boldsymbol{C}^{-1} (\boldsymbol{h}_k - \alpha \boldsymbol{h}_{k-1} - \boldsymbol{d})} \\
& \left(\int_{\mathcal{C}^M} e^{-\frac{\left| y_3 - \boldsymbol{x}_3^T \boldsymbol{h}_3 \right|^2}{\sigma_z^2}} e^{-\left(\boldsymbol{h}_3 - \alpha \bar{\boldsymbol{h}}_2 - \boldsymbol{d} \right)^H \boldsymbol{\Sigma}_2^{-1} \left(\boldsymbol{h}_3 - \alpha \bar{\boldsymbol{h}}_2 - \boldsymbol{d} \right)} e^{-(\boldsymbol{h}_4 - \alpha \boldsymbol{h}_3 - \boldsymbol{d})^H \boldsymbol{C}^{-1} (\boldsymbol{h}_4 - \alpha \boldsymbol{h}_3 - \boldsymbol{d})} d\boldsymbol{h}_3 \right) \\
& d\boldsymbol{h}_4 \cdots d\boldsymbol{h}_n, \qquad \text{(A.4)}
\end{aligned}$$

where $\mathbf{\Sigma}_v^{-1} = \boldsymbol{C}^{-1} - |\alpha|^2 \boldsymbol{C}^{-1}\boldsymbol{G}_v\boldsymbol{C}^{-1}$ and $\bar{\boldsymbol{h}}_v = \mathbf{\Sigma}_v\boldsymbol{C}^{-1}\boldsymbol{G}_v\boldsymbol{q}_v$ for $1 \leq v \leq 2$, $\bar{\boldsymbol{h}}_0 = \boldsymbol{h}_0$ and $\mathbf{\Sigma}_0 = \boldsymbol{C}$. Continue the above procedure until we obtain:

$$\begin{aligned} f(\boldsymbol{y}|\boldsymbol{x}) &= \frac{\prod_{v=1}^{n-1} e^{|\alpha|^2 \bar{\boldsymbol{h}}_v^H \mathbf{\Sigma}_v^{-1}\bar{\boldsymbol{h}}_v - (\alpha\bar{\boldsymbol{h}}_{v-1}+\boldsymbol{d})^H \mathbf{\Sigma}_{v-1}^{-1}(\alpha\bar{\boldsymbol{h}}_{v-1}+\boldsymbol{d})}}{\pi^{M+n}\sigma_z^{2n}|\boldsymbol{C}|^n} \prod_{u=1}^{n-1} e^{-|y_u|^2/\sigma_z^2}|\boldsymbol{G}_u| e^{\boldsymbol{q}_u^H\boldsymbol{G}_u\boldsymbol{q}_u} \\ & \left(\int_{\mathcal{C}^M} e^{-\frac{\left|y_n - \boldsymbol{x}_n^T\boldsymbol{h}_n\right|^2}{\sigma_z^2}} e^{-\left(\boldsymbol{h}_n - \alpha\bar{\boldsymbol{h}}_{n-1} - \boldsymbol{d}\right)^H \mathbf{\Sigma}_{n-1}^{-1}\left(\boldsymbol{h}_n - \alpha\bar{\boldsymbol{h}}_{n-1} - \boldsymbol{d}\right)} d\boldsymbol{h}_n\right). \end{aligned} \tag{A.5}$$

The exponent in the integral in (A.5) equals:

$$\begin{aligned} & \frac{1}{\sigma_z^2}|y_n - \boldsymbol{x}_n^T\boldsymbol{h}_n|^2 + \left(\boldsymbol{h}_n - \alpha\bar{\boldsymbol{h}}_{n-1} - \boldsymbol{d}\right)^H \mathbf{\Sigma}_{n-1}^{-1}\left(\boldsymbol{h}_n - \alpha\bar{\boldsymbol{h}}_{n-1} - \boldsymbol{d}\right) \\ = & \frac{1}{\sigma_z^2}(y_n^* - \boldsymbol{h}_n^H\boldsymbol{x}_n^*)(y_n - \boldsymbol{x}_n^T\boldsymbol{h}_n) + (\boldsymbol{h}_n^H - (\alpha\bar{\boldsymbol{h}}_{n-1} + \boldsymbol{d})^H)\mathbf{\Sigma}_{n-1}^{-1}(\boldsymbol{h}_n - (\alpha\bar{\boldsymbol{h}}_{n-1} + \boldsymbol{d})) \\ = & \frac{1}{\sigma_z^2}\left[|y_n|^2 - y_n^*\boldsymbol{x}_n^T\boldsymbol{h}_n - y_n\boldsymbol{h}_n^H\boldsymbol{x}_n + \boldsymbol{h}_n^H\boldsymbol{x}_n\boldsymbol{x}_n^T\boldsymbol{h}_n\right] + \left[\boldsymbol{h}_n^H\mathbf{\Sigma}_{n-1}^{-1}\boldsymbol{h}_n\right. \\ & \left. -(\alpha\bar{\boldsymbol{h}}_{n-1} + \boldsymbol{d})^H\mathbf{\Sigma}_{n-1}^{-1}\boldsymbol{h}_n - \boldsymbol{h}_n^H\mathbf{\Sigma}_{n-1}^{-1}(\alpha\bar{\boldsymbol{h}}_{n-1} + \boldsymbol{d}) + (\alpha\bar{\boldsymbol{h}}_{n-1} + \boldsymbol{d})^H\mathbf{\Sigma}_{n-1}^{-1}(\alpha\bar{\boldsymbol{h}}_{n-1} + \boldsymbol{d})\right] \\ = & \boldsymbol{h}_n^H\left(\frac{\boldsymbol{x}_n\boldsymbol{x}_n^T}{\sigma_z^2} + \mathbf{\Sigma}_{n-1}^{-1}\right)\boldsymbol{h}_n - \left(\frac{y_n^*\boldsymbol{x}_n^T}{\sigma_z^2} + (\alpha\bar{\boldsymbol{h}}_{n-1} + \boldsymbol{d})^H\mathbf{\Sigma}_{n-1}^{-1}\right)\boldsymbol{h}_n \\ & -\boldsymbol{h}_n^H\left(\frac{y_n\boldsymbol{x}_n}{\sigma_z^2} + \mathbf{\Sigma}_{n-1}^{-1}(\alpha\bar{\boldsymbol{h}}_{n-1} + \boldsymbol{d})\right) + \left(\frac{|y_n|^2}{\sigma_z^2} + (\alpha\bar{\boldsymbol{h}}_{n-1} + \boldsymbol{d})^H\mathbf{\Sigma}_{n-1}^{-1}(\alpha\bar{\boldsymbol{h}}_{n-1} + \boldsymbol{d})\right) \\ = & \boldsymbol{h}_n^H\boldsymbol{G}_n^{-1}\boldsymbol{h}_n - \boldsymbol{h}_n^H\boldsymbol{G}_n^{-1}\boldsymbol{g}_n - \boldsymbol{g}_n^H\boldsymbol{G}_n^{-1}\boldsymbol{h}_n + \left(\frac{|y_n|^2}{\sigma_z^2} + (\alpha\bar{\boldsymbol{h}}_{n-1} + \boldsymbol{d})^H\mathbf{\Sigma}_{n-1}^{-1}(\alpha\bar{\boldsymbol{h}}_{n-1} + \boldsymbol{d})\right) \\ = & (\boldsymbol{h}_n - \boldsymbol{g}_n)^H\boldsymbol{G}_n^{-1}(\boldsymbol{h}_n - \boldsymbol{g}_n) - \boldsymbol{g}_n^H\boldsymbol{G}_n^{-1}\boldsymbol{g}_n + \left(\frac{|y_n|^2}{\sigma_z^2} + (\alpha\bar{\boldsymbol{h}}_{n-1} + \boldsymbol{d})^H\mathbf{\Sigma}_{n-1}^{-1}(\alpha\bar{\boldsymbol{h}}_{n-1} + \boldsymbol{d})\right), \end{aligned}$$

where

$$\boldsymbol{g}_n \triangleq \boldsymbol{G}_n\boldsymbol{q}_n, \quad \boldsymbol{G}_n^{-1} \triangleq \frac{\boldsymbol{x}_n^*\boldsymbol{x}_n^T}{\sigma_z^2} + \mathbf{\Sigma}_{n-1}^{-1} = \frac{\boldsymbol{x}_n^*\boldsymbol{x}_n^T}{\sigma_z^2} + \boldsymbol{C}^{-1} - |\alpha|^2\boldsymbol{C}^{-1}\boldsymbol{G}_{n-1}\boldsymbol{C}^{-1},$$

and

$$\boldsymbol{q}_n \triangleq \frac{y_n\boldsymbol{x}_n^*}{\sigma_z^2} + \mathbf{\Sigma}_{n-1}^{-1}(\alpha\bar{\boldsymbol{h}}_{n-1} + \boldsymbol{d}) = \frac{y_n\boldsymbol{x}_n^*}{\sigma_z^2} + \boldsymbol{C}^{-1}\boldsymbol{d} + \alpha\boldsymbol{C}^{-1}\boldsymbol{G}_{n-1}\boldsymbol{q}_{n-1} - |\alpha|^2\boldsymbol{C}^{-1}\boldsymbol{G}_{n-1}\boldsymbol{C}^{-1}\boldsymbol{d}.$$

Since

$$\begin{aligned} & -\boldsymbol{g}_n^H\boldsymbol{G}_n^{-1}\boldsymbol{g}_n + \left(\frac{|y_n|^2}{\sigma_z^2} + (\alpha\bar{\boldsymbol{h}}_{n-1} + \boldsymbol{d})^H\mathbf{\Sigma}_{n-1}^{-1}(\alpha\bar{\boldsymbol{h}}_{n-1} + \boldsymbol{d})\right) \\ = & -\boldsymbol{q}_n^H\boldsymbol{G}_n\boldsymbol{q}_n + \frac{|y_n|^2}{\sigma_z^2} + (\alpha\bar{\boldsymbol{h}}_{n-1} + \boldsymbol{d})^H\mathbf{\Sigma}_{n-1}^{-1}(\alpha\bar{\boldsymbol{h}}_{n-1} + \boldsymbol{d}), \end{aligned}$$

the final expression for $f(\boldsymbol{y}|\boldsymbol{x})$ is established as:

$$\begin{aligned}
f(\boldsymbol{y}|\boldsymbol{x}) &= \frac{1}{\pi^n\sigma_z^{2n}|\boldsymbol{C}|^n}\prod_{k=1}^{n} e^{|\alpha|^2\bar{\boldsymbol{h}}_k^H\boldsymbol{\Sigma}_k^{-1}\bar{\boldsymbol{h}}_k-(\alpha\bar{\boldsymbol{h}}_{k-1}+\boldsymbol{d})^H\boldsymbol{\Sigma}_{k-1}^{-1}(\alpha\bar{\boldsymbol{h}}_{k-1}+\boldsymbol{d})} e^{-|y_k|^2/\sigma_z^2}|\boldsymbol{G}_k|e^{\boldsymbol{q}_k^H\boldsymbol{G}_k\boldsymbol{q}_k} \\
&= \frac{1}{(\pi\sigma_z^2)^n|\pi\boldsymbol{C}|^n}\left(\prod_{k=1}^{n}|\pi\boldsymbol{G}_k|e^{-|y_k|^2/\sigma_z^2}\right) \\
&\quad \left(\prod_{k=1}^{n-1} e^{|\alpha|^2\bar{\boldsymbol{h}}_k^H\boldsymbol{\Sigma}_k^{-1}\bar{\boldsymbol{h}}_k-(\alpha\bar{\boldsymbol{h}}_k+\boldsymbol{d})^H\boldsymbol{\Sigma}_k^{-1}(\alpha\bar{\boldsymbol{h}}_k+\boldsymbol{d})} e^{\boldsymbol{q}_k^H\boldsymbol{G}_k\boldsymbol{q}_k}\right) e^{\boldsymbol{q}_n^H\boldsymbol{G}_n\boldsymbol{q}_n-(\alpha\boldsymbol{h}_0+\boldsymbol{d})^H\boldsymbol{C}^{-1}(\alpha\boldsymbol{h}_0+\boldsymbol{d})} \\
&= \frac{1}{(\pi\sigma_z^2)^n|\pi\boldsymbol{C}|^n}\left(\prod_{k=1}^{n}|\pi\boldsymbol{G}_k|e^{-|y_k|^2/\sigma_z^2}\right) \\
&\quad \left(\prod_{k=1}^{n-1} e^{\boldsymbol{q}_k^H\boldsymbol{G}_k\boldsymbol{q}_k-\alpha^*\bar{\boldsymbol{h}}_k^H\boldsymbol{\Sigma}_k^{-1}\boldsymbol{d}-\alpha\boldsymbol{d}^H\boldsymbol{\Sigma}_k^{-1}\bar{\boldsymbol{h}}_k-\boldsymbol{d}^H\boldsymbol{\Sigma}_k^{-1}\boldsymbol{d}}\right) e^{\boldsymbol{q}_n^H\boldsymbol{G}_n\boldsymbol{q}_n-(\alpha\boldsymbol{h}_0+\boldsymbol{d})^H\boldsymbol{C}^{-1}(\alpha\boldsymbol{h}_0+\boldsymbol{d})} \\
&= \frac{1}{(\pi\sigma_z^2)^n|\pi\boldsymbol{C}|^n} e^{\boldsymbol{q}_n^H\boldsymbol{G}_n\boldsymbol{q}_n-(\alpha\boldsymbol{h}_0+\boldsymbol{d})^H\boldsymbol{C}^{-1}(\alpha\boldsymbol{h}_0+\boldsymbol{d})}\left(\prod_{k=1}^{n}|\pi\boldsymbol{G}_k|e^{-|y_k|^2/\sigma_z^2}\right) \\
&\quad \left(\prod_{k=1}^{n-1} e^{\boldsymbol{q}_k^H\boldsymbol{G}_k\boldsymbol{q}_k-\alpha^*\boldsymbol{q}_k^H\boldsymbol{G}_k\boldsymbol{C}^{-1}\boldsymbol{d}-\alpha\boldsymbol{d}^H\boldsymbol{C}^{-1}\boldsymbol{G}_k\boldsymbol{q}_k+|\alpha|^2\boldsymbol{d}^H\boldsymbol{C}^{-1}\boldsymbol{G}_k\boldsymbol{C}^{-1}\boldsymbol{d}-\boldsymbol{d}^H\boldsymbol{C}^{-1}\boldsymbol{d}}\right) \\
&= \frac{1}{(\pi\sigma_z^2)^n|\pi\boldsymbol{C}|^n} e^{\boldsymbol{q}_n^H\boldsymbol{G}_n\boldsymbol{q}_n-(\alpha\boldsymbol{h}_0+\boldsymbol{d})^H\boldsymbol{C}^{-1}(\alpha\boldsymbol{h}_0+\boldsymbol{d})}\left(\prod_{k=1}^{n}|\pi\boldsymbol{G}_k|e^{-|y_k|^2/\sigma_z^2}\right) \\
&\quad \left(\prod_{k=1}^{n-1} e^{(\boldsymbol{q}_k-\alpha^*\boldsymbol{C}^{-1}\boldsymbol{d})^H\boldsymbol{G}_k(\boldsymbol{q}_k-\alpha^*\boldsymbol{C}^{-1}\boldsymbol{d})-\boldsymbol{d}^H\boldsymbol{C}^{-1}\boldsymbol{d}}\right),
\end{aligned}$$

where

$$\boldsymbol{G}_k = \begin{cases} \left(\dfrac{\boldsymbol{x}_1^*\boldsymbol{x}_1^T}{\sigma_z^2}+(1+|\alpha|^2)\boldsymbol{C}^{-1}\right)^{-1}, & \text{if } k=1 \\ \left(\dfrac{\boldsymbol{x}_k^*\boldsymbol{x}_k^T}{\sigma_z^2}+(1+|\alpha|^2)\boldsymbol{C}^{-1}-|\alpha|^2\boldsymbol{C}^{-1}\boldsymbol{G}_{k-1}\boldsymbol{C}^{-1}\right)^{-1}, & \text{if } 1<k<n \\ \left(\dfrac{\boldsymbol{x}_n^*\boldsymbol{x}_n^T}{\sigma_z^2}+\boldsymbol{C}^{-1}-|\alpha|^2\boldsymbol{C}^{-1}\boldsymbol{G}_{n-1}\boldsymbol{C}^{-1}\right)^{-1}, & \text{if } k=n \end{cases}$$

and

$$\boldsymbol{q}_k = \begin{cases} \dfrac{y_1\boldsymbol{x}_1^*}{\sigma_z^2}+\boldsymbol{C}^{-1}\boldsymbol{d}+\alpha\boldsymbol{C}^{-1}\boldsymbol{h}_0, & \text{if } k=1 \\ \dfrac{y_k\boldsymbol{x}_k^*}{\sigma_z^2}+\boldsymbol{C}^{-1}\boldsymbol{d}+\alpha\boldsymbol{C}^{-1}\boldsymbol{G}_{k-1}\boldsymbol{q}_{k-1}-|\alpha|^2\boldsymbol{C}^{-1}\boldsymbol{G}_{k-1}\boldsymbol{C}^{-1}\boldsymbol{d}, & \text{if } 1<k\le n. \end{cases}$$

Bibliography

[1] I. C. Abou-Faycal, M. D. Trottand and S. Shamai, “The capacity of discrete-time memoryless Rayleigh-fading channels,” *IEEE Transactions on Information Theory*, vol. 47 , no. 4 , pp. 1290-1301, May 2001.

[2] S. A. Butman and R. J. McEliece, “The ultimate limits of binary coding for a wideband Gaussian channel,” *DSN Progress Report 42-22*, Jet Propulsion Lab., Pasadena, Ca, pp. 78–80, August 1974.

[3] Po-Ning Chen and Fady Alajaji, *Lecture Note in Infomraiton Theory Volume I*, Chap. 6, pp. 152-156, 2005. Available: http://shannon.cm.nctu.edu.tw.

[4] H. Chen, K. Buckleyand R. Perry, “Time-recursive maximum-likelihood- based sequence estimation for unknown ISI channels,” *in Proc. 34th Asilomar Conf. Circuits, Systems, Computers*, Nov. 2000, pp. TA8a14: 1V5.

[5] H. Chen, R. Perry and K. Buckley, “On MLSE algorithms for unknown fast time-varying channels,” *IEEE Transactions on Communications*, vol. 51 , no. 5 , pp. 730-734, May 2003.

[6] T. M. Cover and Joy A. Thomas, *Elements of Information Theory* John Wiley & Sons, 1991.

[7] W. C. Y. Lee, "Estimate of channel capacity in Rayleigh fading environment," *IEEE Trans. Veh. Technol.*, vol. 39, no. 3 , pp. 187-189, Aug. 1990.

[8] Jan R. Magnus and Heinz Neudecker , *Matrix Differential Calculus with Applications in Statistics and Econometrics*, John Wiley, pp. 15 , 1999.

[9] M. Medard, "The effect upon channel capacity in wireless communications of perfect and imperfect knowledge of the channel," *IEEE Transactions on Information Theory*, vol. 46 , no. 3, pp. 933-946, May 2000.

[10] Y.-D. Yaoand A. Sheikh, "Evaluation of channel capacity in a generalized fading channel," *in 1993 IEEE Vehicular Technology Conference*, pp.134 - 137 May 1993.

[11] Fuzhen Zhang, *Matrix theory: Basic Results and Techniques*," Springer-Verlag, May 1999.