Importance Sampling in Wireless Communication Systems with Ultra-Low Error Rates

BHARAT MAHESHWARI AND YEWEN ZHOU
MASTER´S THESIS
DEPARTMENT OF ELECTRICAL AND INFORMATION TECHNOLOGY
FACULTY OF ENGINEERING | LTH | LUND UNIVERSITY

Importance Sampling in Wireless Communication Systems with Ultra-Low Error Rates

Bharat Maheshwari and Yewen Zhou
bh2128ma-s@student.lu.se, ye4324zh-s@student.lu.se

Department of Electrical and Information Technology
Lund University

Supervisor: Dzevdan Kapetanovic (Huawei Lund Research Center)
and Michael Lentmaier (EIT)

Examiner: Ove Edfors (EIT)

June 13, 2022

Printed in Sweden
Tryckeriet i E-huset, Lund

Abstract

Simulation mimics the behaviour of real world processes or the system over time. It helps us to understand the impact of modification and the effect of introducing various interventions to a system. One such simulation method is known as Monte Carlo (MC) simulation, which has been utilized to evaluate the performance of digital communication systems over the last 70 years.

MC has been the most exploited simulation method to assess modern communication systems due to its ability to cope with arbitrary complex system. This method utilizes the concept of repeated sampling, i.e., it blindly samples from a pseudo-random number generator without any knowledge of rare (error) events, to obtain the statistical properties of the system. Hence, to estimate the performance metric down to very low probabilities with high accuracy, long MC simulations are needed and require significant computational power.

Therefore, in this thesis we will explore a modified MC simulation technique called importance sampling (IS), which reduces the variance of the estimator by sampling from the error (rare) events of the input space and thus achieves a given accuracy with shorter simulation time. A detailed evaluation and implementation of current state-of-the-art IS techniques is presented across the additive white Gaussian noise (AWGN) and the Rayleigh fading channel.

The limitation of IS is the requirement of the input probability density function (pdf) which helps in identifying the error region. Obtaining a pdf for 3^{rd} generation partnership project (3GPP) channel models is often not possible and therefore researchers and standardization engineers still resort to MC for system evaluations. In this thesis, we derive an optimal channel pdf for a multiple importance sampling (MIS) technique called ALOE ("At Least One rare Event") in an orthogonal frequency-division multiplexing (OFDM) system. It is further observed that channel samples from the optimal pdf are obtainable via rejection sampling (RS). Significant gain over MC, and better or satisfactory performance compared to the current state-of-the-art IS technique for the Rayleigh fading channel is obtained.

Also, a significant improvement over the current state-of-the-art IS technique for the Rayleigh fading channel has been achieved. This is accomplished by using the Kullback-Leibler divergence (KLD) to estimate an optimal pdf for ALOE using another Rayleigh channel pdf. The system and methods are implemented using MATLAB, and to obtain 3GPP channel models we have utilized QuaDRiGa

version 2.6.1.

Keywords: error region, Monte Carlo, importance sampling, rejection sampling, Kullback-Leibler divergence.

Acknowledgement

"If people do not believe that mathematics is simple, it is only because they do not realize how complicated life is".

-John von Neumann

First, we would like to dedicate this master thesis to our parents. Their support and encouragement have helped us throughout our education and especially during the master thesis.

Then, we would like to express our deep gratitude to Huawei, especially to Dr. Dzevdan Kapetanovic, our supervisor, for his guidance and patience throughout the master thesis as well as for giving critical remarks to improve our report structure.

Further, we would like to express our sincere thanks to Associate Professor Michael Lentmaier, our supervisor at university for his constant support throughout our master's education and his valuable comments on the thesis work. Also, we would like to thank all the professors at Lund university for assisting us with their professional teaching in different aspects of wireless communication and showing us the importance of understanding the concepts rather than memorizing them.

Lastly, we would also like to thank our family, friends and colleagues for their support throughout the master's education.

Bharat and Yewen

Popular Science Summary

Performing simulations using the statistical sampling techniques was made possible by the miraculous development of the first digital computer called electronic numerical integrator and computer (ENIAC) in 1945. Since then, utilization of simulation to asses communication systems has increased significantly, as it provides an insight about the system behaviour under different conditions before the actual implementation with hardware and is much more cost effective than a hardware prototype. One of the vastly utilized method to evaluate communication systems is MC simulation.

MC simulates the system without any knowledge of the systems error region, using the repeated sampling technique in order to obtain the performance of a system under certain conditions, which results in very long simulations time to estimate the low error probabilities with good accuracy. Therefore, in this thesis we will investigate a variance reduction (VR) technique known as IS, to obtain a superior performance compared to MC, which utilizes the knowledge of the input pdf to locate the rare event or error region. The main idea of this thesis work can be understood by the following example:

Consider a supermarket where there are different sections and each section has a worker assigned to it. MC corresponds to you blindly going around in the market looking for the taco (error region) without asking anyone. IS without considering channel pdf corresponds to you blindly going to different sections (bathroom, gardening, food, etc.) and asking the corresponding worker whether the tortilla is in his section. If it is, the worker shows you and you are done. If it is not, you move randomly to the next section. Now, adding signs into the supermarket (or each worker recommending which worker to ask next) corresponds to changing the channel pdf, so that you move only to the relevant sections and reject others.

Table of Contents

List of Figures

List of Tables

List of Acronyms

3GPP	3^{rd} generation partnership project
5G	5^{th} generation cellular network technology
ALOE	At least one rare event
ALOE-MVP	At least one rare event with minimum variance prior
AWGN	Additive white Gaussian noise
BEP	Bit error probability
BLER	Block error rate
BPSK	Binary phase-shift keying
CIS	Conventional importance sampling
CP	Cyclic prefix
CSCG	Circular symmetric complex Gaussian
E2E	End to end
ENIAC	Electronic numerical integrator and computer
EPA	Extended pedestrian A model
EVA	Extended vehicular A model
i.i.d	independent, identically-distributed
IIS	Improved importance sampling
IS	Importance sampling
ISI	Inter-symbol interference
KLD	Kullback-Leibler divergence
MC	Monte Carlo
MIS	Multiple importance sampling
ML	Maximum likelihood
MT	Mean translation
NR	New radio
OFDM	Orthogonal frequency-division multiplexing
OVER	Overbiasing
pdf	Probability density function
QAM	Quadrature amplitude modulation
SEP	Symbol error probability
RS	Rejection sampling
RVs	Random variables

SISO	Single-input single-output
SNR	Signal-to-noise ratio
SCS	Subcarrier spacing
URLLC	Ultra-reliable low latency communications
VR	Variance reduction
VS	Variance scaling
ZF	Zero forcing

Chapter 1

Introduction

1.1 Background and Motivation

Ultra-reliable low latency communications (URLLC) is one of the key technologies of 5^{th} generation cellular network technology (5G) new radio (NR) and has gained significant interest in the wireless community since it was introduced by 3^{rd} generation partnership project (3GPP) in release 15 [1]. URLLC is recognized as the enabler for tactile internet consisting of use cases such as autonomous driving vehicles, augmented reality, robotics in healthcare, smart gird, and are incorporated within two categories: human-to-machine and machine-to-machine communication. In order to fully achieve the potential of these applications, low latency and high reliability is required. The 3GPP has characterized the basic URLLC reliability essentials for a single data frame of 32 byte as 99.9%, and an end to end (E2E) communication latency of <1 ms [1]. Due to stringent requirements for URLLC introduced in release 16, a block error rate (BLER) of 10^{-6} is needed for motion control use cases in industrial automation [2].

Digital communication systems performance (such as BLER) is estimated by averaging a performance metric across the random variables in the system [3]. The most straightforward and common method for averaging is Monte Carlo (MC) simulation, which has the ability to estimate arbitrarily complex communication systems. This technique works on the principle repeatedly sampling the random variables to calculate the mean of the random process [4]. In the case of BLER, to obtain a high degree of accuracy with MC simulation at high signal-to-noise ratio (SNR), large number of samples are required to cause an error event. As a rule of thumb, to estimate the error probability of order P_e, MC simulation requires a sample size of order $100/P_e$ to obtain relative precision of 10% [3]. Applied to the requirement for URLLC stated above, MC simulation will require around 10^8 samples, which is approaching the infeasible simulation complexity.

In this thesis, we will investigate a variance reduction (VR) technique called importance sampling (IS), which obtains high accuracy with much lower number of samples.

1.2 Methodology

The selection of a coherent algorithm is necessary to accelerate the simulations and minimize the storage overhead for estimating a performance metric. However, for advanced systems, huge speed-up cannot be achieved only by efficient algorithms unless the statistical efficiency of the simulation is enhanced. The statistical efficiency of the system is measured by the variance of the estimator [5]. VR in e.g. bit error probability (BEP) or symbol error probability (SEP) estimation provides speed-up as better accuracy can be achieved with smaller sample size, or equivalently, fewer samples are required for the same accuracy. Most VR techniques require more complex computations than MC simulation, but the improvement in statistical efficiency compensates for the decrement in computational efficiency [5].

VR is achieved by modifying MC simulation such that it encourages important events. This modified MC simulation is called IS. IS works on the fundamental that certain input random variables in simulation have more impact on the estimator output than others. Therefore, in order to decrease the variance, these "important" values are sampled more frequently.

The resulting biased sampling is accomplished by modifying the statistical distribution to encourage the occurrence of these important events. In order to obtain an unbiased result (same mean value), the output of the modified simulation must be "weighted" in a particular way [3].

1.3 Literature Review

The idea of utilizing a different simulation technique than MC to reduce estimator variance was first introduced in the mid-fifties by Herman Kahn [6], where he introduced six different sampling methods to combat the statistical problem faced by MC.

IS was formally introduced for a digital communication system by K. S. Shanmugam and P. Balaban in the early eighties [7]. They showed that significant speed-up can be obtained over MC simulation by sampling in the tail region of Gaussian noise with zero mean and unit variance. Although these concepts were applied to the additive white Gaussian noise (AWGN) channel model, they paved the way to explore the application of IS to other channel models.

IS for the flat Rayleigh fading channel was explored in [8], where IS provided significant speed-up over MC for BEP estimation, and the concept of overbiasing was also explained. The main takeaway from this paper was that in order to obtain gain over MC, the statistics of both noise and channel needed to be altered, as only changing the Gaussian noise did not provide any gain.

Another IS concept is multiple importance sampling (MIS) which was introduced in the late 1980s [9]. It proposes to sample from multiple distributions instead of one as seen in conventional IS. A general way of implementing IS with a mixture proposal was explained in [10].

IS implementation for single-input single-output (SISO) orthogonal frequency-division multiplexing (OFDM) systems across frequency selective Rayleigh fading channels was presented in [11]. The same authors have also optimized the conventional IS (CIS) and improved importance sampling (IIS) over frequency selective

fading channels to minimize the variance of the estimator by modifying the variance and mean of noise.

1.4 Contributions

As discussed in [8], biasing the noise probability density function (pdf) is not efficient to obtain speed-up across fading channels and to obtain shorter simulation run time channel pdf knowledge is required. Therefore, this master's thesis introduces new and efficient IS simulation techniques for scalar and OFDM system across fading channels which provides massive speed-up compared to MC and satisfactory performance compared to the current state-of-the-are IS technique *without* utilizing the channel pdf knowledge. Also, an another IS technique is introduced which obtained significant improvement over the current state-of-the-are IS technique by utilizing the channel pdf knowledge.

In order to achieve this speed-up across fading channels an optimal channel pdf is derived for the MIS technique discussed in [10], i.e., we have extended [10] to fading. Further, two techniques are discussed to obtain samples from the optimal channel pdf; rejection sampling (RS) and Kullback-Leibler divergence (KLD).

1.5 Thesis Structure

The layout of this report is as follows: Chapter 2 introduces the basic concept, mathematical background of simulation methods and OFDM system model for IS analysis. In Chapter 3, we introduce the state-of-the-art IS methods for modifying the statistical properties of the distribution and their simulation results are presented across AWGN and Rayleigh fading channels. Chapter 4 derives a close to optimal IS pdf for a scalar fading channels, along with a method to sample from it. Simulation results of its performance are also given in Chapter 4. Chapter 5 generalizes the method in Chapter 4 to OFDM channels and illustrates its performance across 3GPP channels. Finally, the report is concluded and future work is summarized in Chapter 6.

Chapter 2

Simulation Methods and System Model

This chapter introduces the different simulation techniques used to evaluate digital communication systems. Here we discuss the traditional MC simulation technique followed by its drawback in advanced communication systems. Further, we explain the IS methodology and its advantages over MC. Also, the evaluation parameters are introduced to compare the performance of IS and MC simulation. Lastly, we introduce the OFDM system model on which IS techniques are applied.

2.1 Monte Carlo Simulation

MC simulation has been widely utilized to evaluate the performance of digital communication over the past seventy years. The MC method was developed during the time of World War II at the University of Pennsylvania in Philadelphia, where a team of scientists, engineers, and technicians were trying to build the electronic numerical integrator and computer (ENIAC), the first electronic computer [12].

The spark which led to the evolution of the MC method happened in 1946 during the review of ENIAC at Los Alamos, where Stanislaw Ulam was astonished by the feasibility and speed of ENIAC. Ulam, himself was already aware of the tedious calculation required for statistical sampling required for electromechanical computers, but after seeing the speed and versatility of the ENIAC he thought that the sampling technique should be resuscitated and discussed the idea with John von Neumann. Subsequently, von Neumann wrote the proposal to solve the problem of neutron diffusion in fissionable materials. As both of them wanted to keep the work on this method a secret, they went for the code name Monte Carlo which was suggested by their colleague Nicholas Metropolis and refers to the Monte Carlo Casino in Monaco where Ulam's uncle used to borrow money from relatives and gamble [12].

MC simulation suggests to draw N independent, identically-distributed (i.i.d) samples $\{x_i\}_{i=1}^{N}$ from a density $f(x)$ defined on a high dimensional space Γ. These N samples are utilized to approximate the distribution $f(x)$ by the empirical point-mass function [13]

$$f_N(x) = \frac{1}{N}\sum_{i=1}^{N}\delta_{x_i}(x), \tag{2.1}$$

where $\delta_{x_i}(x)$ denotes the delta-Dirac mass located at x_i. The above empirical approach of approximating a density can be applied to evaluate integrals $I(p)$ with a tractable sum $I_N(p)$ as

$$I_N(p) = \frac{1}{N}\sum_{i=1}^{N} p(x_i) \xrightarrow[N\to\infty]{\text{a.s}} I(p) = \int_{\Gamma} p(x) f(x)\,\mathrm{d}x, \tag{2.2}$$

where $p(x)$ is an integrable function and a.s. denotes almost sure convergence. Equation (2.2) implies that the estimator $I_N(p)$ converges to the mean value of $p(x)$ under $f(x)$ and is therefore unbiased [13]. Consider a communication system with a discrete random variable x as an input. Generally, this can be a passband, binary or M-ary complex-valued input symbol with the probability density function $f_X(x)$. Let the receiver output be υ , which is composed of the input, the channel and noise. The systems probability of error P_e can be expressed as [7]

$$P_e = \sum_{x}\int_{\upsilon} \mathbb{1}_{\epsilon}(\upsilon|x) f_{V|X}(\upsilon|x) f_X(x)\,\mathrm{d}\upsilon, \tag{2.3}$$

where $f_{V|X}(\upsilon|x)$ is a conditional pdf of channel and noise for a given x and $\mathbb{1}_{\epsilon}(\upsilon|x)$ is the error indicator function

$$\mathbb{1}_{\epsilon}(\upsilon|x) = \begin{cases} 1 & \upsilon \in \Gamma_x \\ 0 & \upsilon \notin \Gamma_x, \end{cases} \tag{2.4}$$

where Γ_x is the error region of the receiver for system input x. Using (2.4) we can re-write (2.3) as

$$P_e = \sum_{x}\int_{\Gamma_x} f_{V|X}(\upsilon|x) f_X(x)\,\mathrm{d}\upsilon. \tag{2.5}$$

The MC estimator for P_e is denoted by $\hat{P}_e$ and is given by

$$\hat{P}_e = \frac{1}{N_{MC}}\sum_{i=1}^{N_{MC}} \mathbb{1}_{\epsilon}(\upsilon_i|x_i) = \frac{N_e}{N_{MC}},\ x_i \sim f_X(x),\ \upsilon_i \sim f_{V|X}(\upsilon|x), \tag{2.6}$$

where $\{x_i, \upsilon_i\}_{i=1}^{N_{MC}}$ are the input and output sequences, respectively, with x_i and υ_i being distributed according to $f_X(x)$ and $f_{V|X}(\upsilon|x)$, respectively. N_e and N_{MC} are number of error events and number of samples, respectively. By definition, $\mathbb{1}_{\epsilon}(\upsilon_i|x_i)$ is a Bernoulli distributed random variable with probability of success P_e. Hence, N_e is binomially distributed as $B(N_{MC}, P_e)$. Thus, the mean and variance of the MC estimator can be deduced using the mean and variance of N_e as

$$E\left[\hat{P}_e\right] = P_e, \tag{2.7}$$

and

$$\sigma^2_{MC} = \frac{P_e(1-P_e)}{N_{MC}} \approx \frac{P_e}{N_{MC}}. \tag{2.8}$$

From (2.7), we can confirm that the MC estimator in (2.6) is unbiased and (2.8) shows that the variance of estimator will be reduced as the number of samples increases [11].

The relative precision ϵ_r of an estimator with variance σ^2 is defined as [7]

$$\epsilon_r \triangleq \frac{\sigma}{P_e}. \tag{2.9}$$

Using (2.9) and the approximation in (2.8), the number of samples required by MC to estimate P_e with a given ϵ_r can be approximated as

$$N_{MC} \approx \frac{1}{\epsilon_r^2 P_e}. \tag{2.10}$$

Equation (2.10) indicates that the simulation run time increases exponentially to obtain reliable low error probability. Due to this, MC simulation is infeasible to evaluate the algorithms for URLLC.

2.2 Importance Sampling

Importance sampling is a VR technique that has the potential to offer substantial run time savings for communication systems. This technique provides speed-up by simulating the system with a modified (biased) distribution, i.e., varying the statistical distribution which encourages important events (errors) to occur more often. Knowledge of the biased distribution is exploited at the output to make the estimator unbiased.

IS introduces the biased conditional pdf $f^*_{V|X}(\upsilon|x)$ which is used to simulate the communication system for efficient performance. One can also modify the input pdf $f_X(x)$, which is known as probabilistic constellation shaping, however, this technique is not investigated in this thesis. Using the biased pdf $f^*_{V|X}(\upsilon|x)$ (2.5) can be written as

$$P_e = \sum_x \int_{\Gamma_x} f_{V|X}(\upsilon|x) f_X(x)\,\mathrm{d}\upsilon = \sum_x \int_{\Gamma_x} \frac{f_{V|X}(\upsilon|x)}{f^*_{V|X}(\upsilon|x)} f^*_{V|X}(\upsilon|x) f_X(x)\,\mathrm{d}\upsilon, \tag{2.11}$$

where $W(\upsilon|x) \triangleq \frac{f_{V|X}(\upsilon|x)}{f^*_{V|X}(\upsilon|x)}$ is the weight function of IS. The estimator for P_e becomes

$$\hat{P}^*_e = \frac{1}{N_{IS}} \sum_{i=1}^{N_{IS}} \mathbb{1}_\epsilon(\upsilon_i|x_i) W(\upsilon_i|x_i),\ x_i \sim f_X(x),\ \upsilon_i \sim f^*_{V|X}(\upsilon|x), \tag{2.12}$$

where $\hat{P}^*_e$ is the IS estimator with N_{IS} samples. This estimator is unbiased, which can be shown by re-writing (2.12) as

$$\begin{aligned} E\left[\hat{P}_e{}^*\right] &= \frac{1}{N_{IS}} \sum_{i=1}^{N_{IS}} E\left[\mathbb{1}_\epsilon(\upsilon_i|x_i) \frac{f_{V|X}(\upsilon_i|x_i)}{f^*_{V|X}(\upsilon_i|x_i)}\right] \\ &= \frac{1}{N_{IS}} \times N_{IS} \sum_x \int_{\Gamma_x} \frac{f_{V|X}(\upsilon|x)}{f^*_{V|X}(\upsilon|x)} f^*_{V|X}(\upsilon|x) f_X(x)\,\mathrm{d}\upsilon \\ &= \sum_x \int_{\Gamma_x} f_{V|X}(\upsilon|x) f_X(x)\,\mathrm{d}\upsilon = P_e, \end{aligned}$$

The variance of $\hat{P}_e^*$ is given as [14]

$$\sigma_{IS}^2 = \frac{1}{N_{IS}} \sum_x \int_{\Gamma_x} f_{V|X}(\upsilon|x)\, W(\upsilon|x)\, f_X(x)\, \mathrm{d}\upsilon - \frac{P_e^2}{N_{IS}}. \tag{2.13}$$

The idea of IS is to choose a biased conditional density that reduces the variance. The reward of selecting a good biased density are substantial time savings compared to MC simulation. On the other hand, the penalty of selecting a bad density is longer run time and worse performance than MC. Such biased densities underestimate or overestimate the SEP which results in bad performance of IS [3]. This phenomenon is called overbiasing (OVER).

In general, good system knowledge is required in order to select a good biased density $f_{V|X}^*(\upsilon|x)$, which means that IS is a system specific simulation technique. The optimal density which achieves $\sigma_{IS}^2 = 0$ can be derived from (2.13) as [3]

$$f_{opt}^*(\upsilon|x) = \frac{\mathbb{1}_\epsilon(\upsilon|x)\, f_{V|X}(\upsilon|x)}{P_e(x)}. \tag{2.14}$$

The expression (2.14) is not realizable as it contains the true mean P_e that we want to calculate. However, it provides us with an intuition that a good IS density should have all of its mass in the error region proportional to the original density.

To identify the run time saving and speed-up obtained by IS compared to MC, evaluation of its biased density is required. This is generally measured by the variance reduction factor $\zeta_{MC/IS}$ for a sample size $N = N_{MC} = N_{IS}$, given as

$$\zeta_{MC/IS} \triangleq \frac{\sigma_{MC}^2}{\sigma_{IS}^2}\bigg|_N, \tag{2.15}$$

which can also be interpreted as a speed-up factor by which the IS estimator achieves the same precision as MC. The variance reduction factor has to be calculated empirically by running many error probability simulations and computing the variance of the resulting error estimates for MC and IS. It is noteworthy that this evaluation parameter does not involve the computation overhead incurred by weight calculation $W(\upsilon|x)$. Therefore some authors define the net run time, which besides variance reduction also counts mathematical operations as a part of computation overhead, to be the main performance metric [15].

2.3 System Model

The simulation of MC and IS techniques with different channel models is executed over an OFDM system shown in Fig. 2.1. The OFDM signal consists of N subcarriers, where each subcarrier is modulated by rectangular M-ary quadrature amplitude modulation (QAM), M being the order of QAM. The baseband modulated symbols have a symbol time of T_s and are grouped into blocks of N symbols. A block can be denoted as

$$\mathbf{x} = [x_0, x_1, x_2, ..., x_{N-1}]^T, \tag{2.16}$$

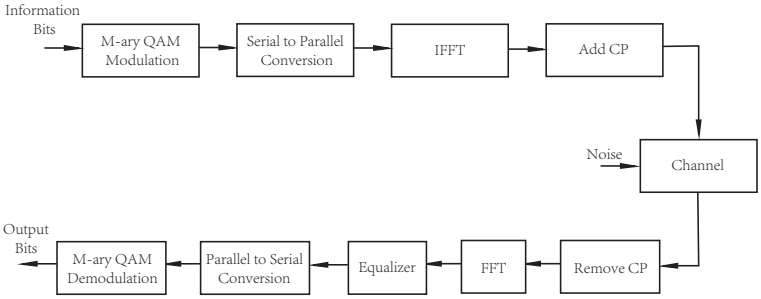


Figure 2.1: OFDM transmitter and receiver structure

where $[.]^T$ denotes the transpose of a matrix and x_i is the modulated symbol selected from the QAM constellation S. Each symbol x_i carries n information bits where $n = \log_2 M$. The M-ary QAM constellation is the set [11]

$$S = \left\{ \left[(\pm 2d_1 + 1) \pm j\,(2d_2 + 1) \right] / \sqrt{E_s}; d_1, d_2 \in \{0, 1, ..., \log_2 M - 1\} \right\}, \quad (2.17)$$

where $E_s = 2(M-1)/3$. The QAM constellation symbol power $E\left[|x_i|^2\right]$ and bit power $E\left[|x_i|^2\right]/n$ are equal to 1 and $1/n$, respectively. Symbols on each subcarrier are generated randomly and have equal distribution of power. This is an appropriate choice, in cases when there is no channel knowledge at the transmitter [16].

The baseband modulated symbols are OFDM modulated and transmitted through the channel. The OFDM transmitted signal can be written as

$$r(t) = \sum_{i=0}^{N-1} x_i e^{j2\pi i \frac{t}{T}}, -T_{cp} \leq t \leq T, \quad (2.18)$$

where $T = NT_s$ is the OFDM symbol duration and T_{cp} is the cyclic prefix (CP) duration which is inserted between consecutive OFDM symbols to avoid intersymbol interference (ISI) and maintain the orthogonality between OFDM subcarriers.

The OFDM signal is transmitted through a communication channel with impulse response $h(t)$. Therefore, the received OFDM signal $y(t)$ is given as

$$y(t) = h(t) * r(t) + z(t), -T_{cp} \leq t \leq T, \quad (2.19)$$

where $z(t)$ is the complex Gaussian distribution with zero mean and variance σ_z^2 per dimension and '$*$' represents convolution.

If we consider a communication channel that is a multipath, frequency-selective fading channel that is constant during an OFDM symbol, the channel response $h(t)$ can be written as [17]

$$h(t) = \sum_{l=0}^{L-1} h_l \delta(t - \tau_l), \quad (2.20)$$

where h_l is the attenuation factor for the signal received at l_{th} path, τ_l is the propagation delay for l_{th} path and L is the number of paths. Inserting (2.20) in

(2.19) we obtain that (2.19) can be written as

$$y(t) = \sum_{l=0}^{L-1} h_l r(t - \tau_l) + z(t), -T_{cp} \leq t \leq T. \tag{2.21}$$

The AWGN channel is a special case of (2.21) with $L = 1$, $\tau_0 = 0$ and $h_0 = 1$.

Considering T_{cp} larger than the channel maximum delay spread, the output symbol for i_{th} subcarrier after OFDM demodulation is given as [11]

$$Y_i = \frac{1}{T} \int_0^T y(t) e^{-j2\pi \frac{i}{T} t} \mathrm{d}t, i = 0, 1, ..., N-1. \tag{2.22}$$

Using (2.18), (2.20) and (2.21), equation (2.22) can be expressed as [18]

$$Y_i = H_i x_i + Z_i, \tag{2.23}$$

where Z_i is a circular symmetric complex Gaussian (CSCG) random variable with zero mean and variance σ_z^2 per dimension, $Z_i \sim \mathscr{CN}\left(0, 2\sigma_z^2\right)$. H_i is the channel frequency response at i_{th} subcarrier

$$H_i = \sum_{l=0}^{L-1} h_l e^{-j2\pi i \frac{\tau_l}{T}}, i = 0, 1, ..., N-1. \tag{2.24}$$

A coherent detection is assumed at the receiver, i.e., channel state information (CSI) is assumed to be known at the receiver. In this thesis we have utilized the zero-forcing (ZF) equalizer, whose weighting factor can be described as [19]

$$G_i = \frac{1}{H_i}, i = 0, 1, ..., N-1. \tag{2.25}$$

ZF is chosen here as it equals the maximum likelihood (ML) equalizer for SISO channels and coherent detection. Therefore, the output sample of channel equalization at i_{th} subcarrier equals

$$V_i = G_i Y_i = x_i + Z_i', i = 0, 1, ..., N-1, \tag{2.26}$$

where Z_i' is a CSCG variable $\mathscr{CN}\left(0, 2\sigma_i^2\right)$ with $\sigma_i^2 = \frac{1}{|H_i|^2}\sigma_z^2$. The detector will make an independent decision on each subcarrier i to estimate the SEP.

Chapter 3

State-of-the-Art IS

In this chapter, different state-of-the-art IS techniques are explained and their performance investigated.

3.1 Variance Scaling and Mean Translation of Noise

As described in Chapter 2, the main idea of IS is to as good as possible approximate the optimal biased distribution. This is accomplished by approximating the original distribution in the error region, i.e., the biased distribution should be selected such that it places as much of its mass as possible inside the error region. For the communication model in (2.26), $f_{V|X}$ from Chapter 2 is a Gaussian distribution. The error region of (2.26) lies in the tail region of the Gaussian distribution. Therefore a good approximation of f^*_{opt} for the model in (2.26) is a scaled/translated version of $f_Z(z)$ [3]. The scaling of noise is achieved by modifying the variance of the $f_Z(z)$, and this technique is called variance scaling (VS) or CIS. The other way to modify the original distribution is by translating/shifting $f_Z(z)$, and this method is known as mean translation (MT) or IIS.

For the communication model presented in (2.26), we can represent the SEP P_{e_i} for the i_{th} subcarrier from (2.11) as

$$P_{e_i} = \sum_x \int_{\Gamma_x} W(z_i) f_X(x) f^*_{Z'_i}(z_i) \, \mathrm{d}\upsilon_i, \tag{3.1}$$

where $z_i = \upsilon_i - x_i$, $W(z_i) = \frac{f_{Z'_i}(z_i)}{f^*_{Z'_i}(z_i)}$ is the weight function, $f_{Z'_i}(z_i)$ and $f^*_{Z'_i}(z_i)$ are the input and biased noise density functions, respectively. The estimator for P_{e_i} is given as

$$\hat{P}_{e_i} = \mathbb{1}_\epsilon(\upsilon_i|x_i) W(z_i), \; x_i \sim f_X(x), \tag{3.2}$$

and the variance of $\hat{P}_{e_i}$ is

$$\sigma^2_{IS_i} = \frac{1}{N} \sum_x \int_{\Gamma_x} W(z_i) f_X(x) f^*_{Z'_i}(z_i) \, \mathrm{d}\upsilon_i - \frac{P^2_{e_i}}{N}. \tag{3.3}$$

3.1.1 Variance Scaling

As described above, variance scaling provides a biased density that is scaled in the tail region of the input density as shown in Fig. 3.1. This is achieved by scaling the input density by VS factor β. Let $\sigma^*_{z_i}$ be the standard deviation of the biased noise and β_i the VS factor on the i_{th} subcarrier, which are related as

$$\sigma^*_{z_i} = \beta_i \sigma_z. \tag{3.4}$$

The IS weighting function for VS equals [11]

$$W_{VS}\left(z'_i\right) = \beta_i^2 \exp\left(-\frac{|z'_i|^2}{2\sigma_i^2}\left(1 - \frac{1}{\beta_i^2}\right)\right), \tag{3.5}$$

with biased noise pdf being

$$f^*_{Z'_i}\left(z'_i\right) = \frac{1}{\sqrt{2\pi\beta_i^2\sigma_i^2}} \exp\left(-\frac{|z'_i|^2}{2\beta_i^2\sigma_i^2}\right). \tag{3.6}$$

The VS estimator variance σ^2_{VS} for a binary phase-shift keying (BPSK) system with amplitude A is given by [11]

$$\sigma^2_{VS} = \frac{\beta_i^2}{2 - \frac{1}{\beta_i^2}} Q\left(\frac{\left(\sqrt{\left(2 - \frac{1}{\beta_i^2}\right)}\right)A}{\sigma_i}\right), \tag{3.7}$$

where

$$Q\left(t\right) = \frac{1}{\sqrt{2\pi}} \int_t^\infty e^{\frac{-t^2}{2}} \mathrm{d}t. \tag{3.8}$$

The value of β that minimizes the variance in (3.7) is [11]

$$\beta^2 = \frac{5\sigma^2 + 4D^2 + \left(25\sigma^4 + 8\sigma^2 D^2 + 16D^4\right)^{0.5}}{8\sigma^2}, \tag{3.9}$$

where $\sigma = \max_i \left\{\frac{1}{|H_i|}\sigma_z\right\}$ and $D = A$ for BPSK. For higher order QAM modulation, $D = A = \frac{d_{min}}{2}$ where d_{min} is the minimum Euclidean distance between two symbols in the QAM constellation.

3.1.2 Mean Translation

Mean translation provides a biased density by shifting the mean of the input density towards the error region as shown in Fig. 3.2. Let Z'_i denote the AWGN noise on subcarrier i and let $m_{z'_i} = m_{x'} + jm_{y'}$ denote its mean value. The weighting function for the i_{th} subcarrier equals [11]

$$W_{MT}\left(z'_i\right) = \exp\left(-\frac{|z'_i|^2 - |z'_i - m_{z'_i}|^2}{2\sigma_i^2}\right), \tag{3.10}$$

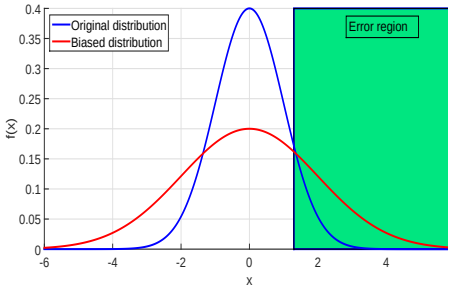


Figure 3.1: Variance scaling of noise pdf.

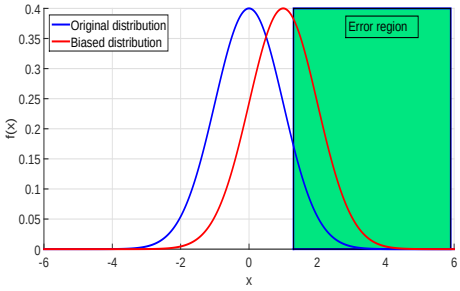


Figure 3.2: Mean translation of noise pdf.

with the biased noise pdf as

$$f^*_{Z'_i}(z'_i) = \frac{1}{\sqrt{2\pi\sigma_i^2}} \exp\left(-\frac{|z'_i - m_{z'_i}|^2}{2\sigma_i^2}\right). \tag{3.11}$$

Assuming a BPSK system with $-A$ as the transmitted symbol, the MT estimator variance σ^2_{MT} can be approximated as [11]

$$\sigma^2_{MT} \approx \frac{\sigma_i}{\sqrt{2\pi}\,(A + m_{x'})} \exp\left(-\frac{3m^2_{x'} + 2m^2_{y'} + 2Am_{x'} + A^2}{2\sigma_i^2}\right). \tag{3.12}$$

Minimizing (3.12) with respect to $m_{x'}$ and $m_{y'}$ yields the mean

$$m_{z'_i} = \frac{2A + \sqrt{(A^2 - 3\sigma_i^2)}}{3}. \tag{3.13}$$

For high SNR scenario we can assume $A^2 \gg \sigma_i^2$, therefore the expression for optimum mean of MT can be approximated as

$$m_{z'_i} \approx A. \tag{3.14}$$

The expression (3.14) shows that the mean that minimizes the IS estimator variance would shift V_i to the midpoint between the transmitted symbol and the error symbol, thereby generating an error with a probability of $\frac{1}{2}$.

This technique can be applied to higher order constellations by targeting specific error centres of the signal constellation in each simulation run. Assuming that the all zero bits sequence corresponding to symbol s_0 is transmitted, the SEP for s_0 can be computed by

$$\hat{P}_e(s_0) = \sum_{j=1}^{M-1} \hat{P}(s_j|s_0), \tag{3.15}$$

where s_j is the target error center and $\hat{P}(s_j|s_0)$ is the estimated probability of decoding s_j given that s_0 is transmitted. The decision variable in (2.26) is now

biased so that $E^*[V_i] = (s_j + s_0)/2$, where $E^*[.]$ represents the expectation with respect to the biased noise pdf [11].

Finally, the estimator for MT is given as

$$\hat{P}_e = \frac{1}{M} \sum_{k=0}^{M-1} \sum_{\substack{j=1 \\ j \neq k}}^{M-1} \hat{P}(s_j|s_k). \tag{3.16}$$

3.2 Variance Scaling of the Rayleigh Fading Channels

As discussed in [8], to achieve speed-up for fading channels, biasing of the channel fading coefficient is required. Therefore, here we discuss the VS technique to bias the fading coefficient. Let the time domain multipath coefficients h_l be mutually independent complex random variables (RVs) represented as

$$h_l = |h_l| e^{j\phi_l}, \tag{3.17}$$

where $|h_l|$ and ϕ_l are the amplitude and phase of h_l, respectively. Assuming a Rayleigh pdf of $|h_l|$, we get

$$f_{|h_l|}(R) = \frac{R}{\sigma_l^2} \exp\left(-\frac{R^2}{2\sigma_l^2}\right) \tag{3.18}$$

where $2\sigma_l^2 = E\left[|h_l|^2\right]$ is the power of h_l. The pdf of ϕ_l is uniformly distributed within $[0, 2\pi]$.

The frequency domain multipath coefficient H_i in (2.24) is now the summation of Gaussian variables and is thus Gaussian itself; therefore the pdf of $|H|$ is a Rayleigh distribution with $E\left[|H|^2\right] = \sum_{l=0}^{L-1} 2\sigma_l^2 = 2\sigma_H^2$

$$f_{|H|}^*(R) = \frac{R}{\sigma_H^2} \exp\left(-\frac{R^2}{2\sigma_H^2}\right). \tag{3.19}$$

The channel coefficients $|H|$ can be biased by using the VS factor β_R^2; this biasing is referred to as VS-Rayleigh [20]. The biased variance of $|H|$ is $\beta_R^2 \sigma_H^2$ and of h_l it is $\beta_R^2 \sigma_l^2$. The IS weighting function of VS-Rayleigh equals

$$W_{VS-R}(R) = \beta_R^2 \exp\left(-\frac{R^2}{2\sigma_H^2}\left(1 - \frac{1}{\beta_R^2}\right)\right). \tag{3.20}$$

In the case of a BPSK system with $-A$ as the transmitted symbol and the output symbol of subcarrier i given by (2.23), the VS-Rayleigh estimator variance equals [20]

$$\sigma_{VS-R}^2 = \frac{\beta_R^3 \sigma_z^2}{2A\sigma_H \left(A^2 \beta_R^2 \sigma_H^2 + 2\sigma_z^2 \beta_R^2 - \sigma_z^2\right)^{0.5}}. \tag{3.21}$$

Minimizing σ_{VS-R}^2 in (3.21) with respect to β_R yields the optimum β_R for BPSK as

$$\beta_R^2 = \frac{3\sigma_z^2}{2A^2\sigma_H^2 + 4\sigma_z^2}. \tag{3.22}$$

The optimum β_R [20] for higher order QAM modulation can be derived similarly by substituting $A = \frac{d_{min}}{2}$ in (3.22).

3.3 ALOE

We know that the optimal bias density in (2.14) is proportional to the original density in the error region. In Section 3.1, we have discussed the methods to bias the noise pdf using variance and mean, which minimizes the variance of the estimator but the biased pdf doesn't follow the tails of the original density perfectly as depicted in Fig. 3.1 and Fig. 3.2. In order to achieve significant gains we would like to have a method which can approximate the original distribution in the error region as good as possible. For this purpose, we will utilize an MIS technique called ALOE ("At Least One rare Event").

The SEP of a point x_i, given a channel realization H_i, equals the integral of a Gaussian over the union of half-spaces formed by $K(x_i)$ hyperplanes [10]. For a rectangular QAM constellation the inner, edge and corner points have 4, 3 and 2 hyperplanes, respectively, as shown in Fig. 3.3. The goal of ALOE is to sample the received point within this region.

Assume a symbol x_i is being transmitted from the pdf $f_X(x)$ and the output of the receiver is υ_i. The SEP of x_i is the integral of the Gaussian distribution centered at x_i outside $\mathscr{R}_i$ (shaded region) in Fig. 3.3: $p_i \triangleq P(e|x_i) = \mathbb{P}(\upsilon_i \notin \mathscr{R}_i|x_i)$. Thus, the SEP is

$$P_e = \frac{1}{M}\sum_{i=1}^{M} p_i. \tag{3.23}$$

The idea of ALOE is to efficiently estimate $p_i, i = 1, 2, .., M$ in (3.23).

The SEP of x_i can be expressed as

$$p_i = \int_{\mathbb{R}^2} \mathbb{1}_{\mathscr{R}_i^c}(\upsilon_i|x_i) f_{V|X}(\upsilon_i|x_i)\mathrm{d}\upsilon_i, \tag{3.24}$$

where $\mathbb{1}_{\mathscr{R}_i^c}$ is the indicator function taking value 1 for all $\upsilon_i \notin \mathscr{R}_i$. As described above, the error region $\mathscr{R}_i^c$ can be written as $\bigcup_{k=1}^{K(x_i)} S_k^i$, where $S_k^i = \left\{\upsilon \in \mathbb{C} | \Re\left\{\upsilon\bar{\gamma}_k^i\right\} \geq \beta_k^i\right\}$ is the half-space parametrized by γ_k^i and β_k^i which are defined by the transmitted symbol x_i, where γ_k^i is the direction of hyperplane represented in complex domain, $\bar{L}$ represents the complex conjugate of a complex number L and $\Re\{C\}$ denotes the real part of complex number C.

The ALOE mixture proposal with $K(x_n)$ hyperplanes is given as [10]

$$\upsilon_n \sim t_{n,\alpha} = \sum_{k=1}^{K(x_n)} \alpha_{n,k} q_{n,k}\left(\upsilon_n|x_n\right), n = 1, 2, .., N, \tag{3.25}$$

where $\alpha_{n,k}$ is the non-negative weight of the k-th hyperplane such that $\sum_{k=1}^{K(x_n)} \alpha_{n,k} = 1$. The number of proposals, $K(x_n)$, is equal to the number of hyperplanes, with each proposal being a truncated version of the target distribution (Gaussian centered at x_n) beyond each hyperplane

$$q_k\left(\upsilon_n|x_n\right) = \frac{\mathbb{1}_{S_k}\left(\upsilon_n|x_n\right) f_{V|X}\left(\upsilon_n|x_n\right)}{P_{n,k}}, \tag{3.26}$$

where $P_{n,k} = \int_{\mathscr{R}^c} \mathbb{1}_{S_k(\upsilon_n|x_n)} f_{V|X}\left(\upsilon_n|x_n\right)\mathrm{d}\upsilon_n$ is the integral of the target distribution beyond a hyperplane. The procedure to generate a sample from the

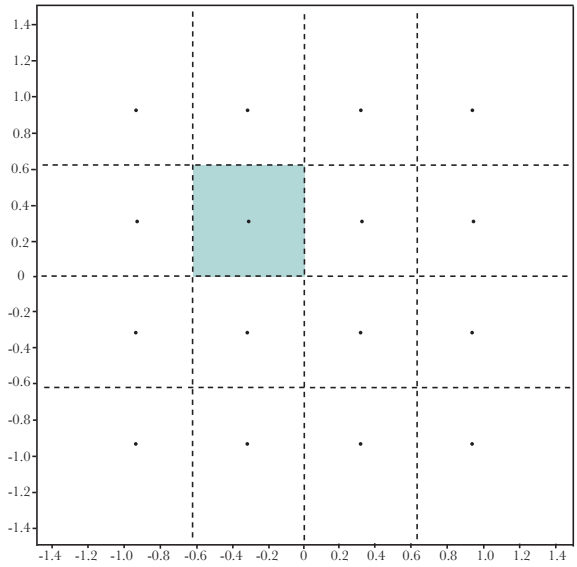


Figure 3.3: 16 QAM constellation with hyperplanes and $\mathscr{R}_i$

truncated Gaussian distribution is described in Appendix A. It should be noted that the union upper bound equals $\tilde{p}_n \triangleq \sum_{k=1}^{K(x_n)} P_{n,k}$.

Equation (2.12) of the IS estimator can now be extended as

$$\begin{aligned}\hat{P}_e^{(ALOE)} &= \frac{1}{N}\sum_{n=1}^{N} \mathbb{1}_{\mathscr{R}_n^c}(\upsilon_n|x_n)\, W(\upsilon_n|x_n) \\ &= \frac{1}{N}\sum_{n=1}^{N} \mathbb{1}_{\mathscr{R}_n^c}(\upsilon_n|x_n)\, \frac{f_{V|X}(\upsilon_n|x_n)}{t_{n,\alpha}(\upsilon_n|x_n)}. \end{aligned} \tag{3.27}$$

Inserting (3.26) in (3.25) it follows that (3.27) equals

$$\hat{P}_e{}^{(ALOE)} = \frac{1}{N}\sum_{n=1}^{N} \frac{\mathbb{1}_{\mathscr{R}_n^c}(\upsilon_n|x_n)}{\sum_{k=1}^{K(x_n)} \alpha_{n,k}\mathbb{1}_{S_k}(\upsilon_n|x_n)\, P_{n,k}^{-1}}. \tag{3.28}$$

The weight of each proposal in (3.25) is chosen as $\alpha_{n,k} = \frac{P_{n,k}}{\tilde{p}_n}$, for $k = 1, 2, .., K(x_n)$. Then, (3.28) becomes

$$\begin{aligned}\hat{P}_e{}^{(ALOE)} &= \frac{1}{N}\sum_{n=1}^{N} \frac{\tilde{p}_n}{\sum_{k=1}^{K(x_n)} \mathbb{1}_{S_k}(\upsilon_n|x_n)}, \\ &= \frac{1}{N}\sum_{n=1}^{N} \frac{\tilde{p}_n}{C(\upsilon_n|x_n)}, \end{aligned} \tag{3.29}$$

where $C(\upsilon_n|x_n) = \sum_{k=1}^{K(x_n)} \mathbb{1}_{S_k}(\upsilon_n|x_n)$ is the number of half-spaces S_k where υ_n is observed. In (3.28), $\mathbb{1}_{\mathscr{R}_n^c}(\upsilon_n|x_n)$ is always equal to 1 as all the samples are generated in $\mathscr{R}_n^c$.

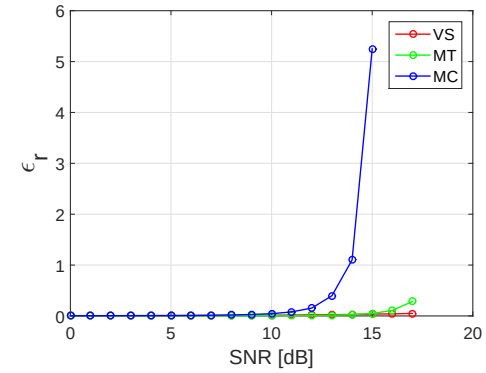
3.4 Performance Across the AWGN Channel

In this section, we will demonstrate the performance of the different sampling techniques from this chapter across the AWGN channel. To compute the variance of the SEP estimates, we repeat the SEP simulation N_s number of times and compute the variance of the resulting N_s SEP estimates. The techniques are simulated with $N = 7.168 \times 10^4$ and $N_s = 200$.

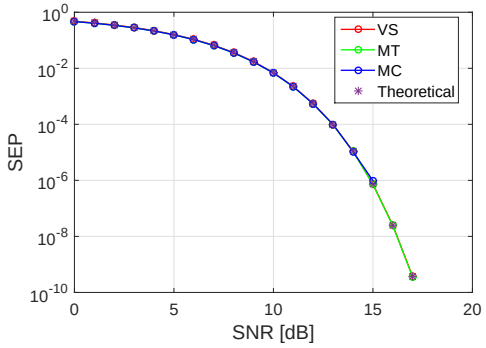
Variance Scaling and Mean Translation of Noise

Here, we present the simulation result of biasing the noise pdf using VS and MT of noise.

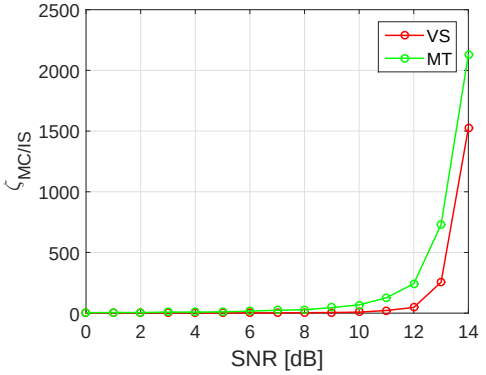
According to Fig. 3.4 for 16 QAM, VS and MT achieves significant gain over MC. Fig. 3.4 (a) shows that MC achieves similar precision as VS and MT at low SNRs. At SNR = 11 dB, the estimated SEP in Fig. 3.4 (b) is 10^{-3}, for which MC obtains a good approximation of the true mean, since approximately 10^5 samples are required (see Section 1.1), which is close to the number of samples we use in the simulation. However, at high SNR the precision of MC decreases significantly compared to VS and MT, which is expected as both VS and MT bias the noise pdf with an optimal constant that minimizes the estimator variance, i.e., in each simulation they obtain a good approximate of the true SEP. This also concurs the accuracy of SEPs as low as 10^{-9} achieved by VS and MT. The speed-up in simulation run time is measured by the variance reduction factor. As depicted in Fig. 3.4 (c), at SNR = 14 dB, MT requires 2100 times less samples than MC to obtain an estimate of the SEP with the same precision. The performance with 64 QAM is shown in Fig. 3.5, where a similar improvement in relative precision and SEP is seen as for 16 QAM.



(a) Relative precision vs. SNR

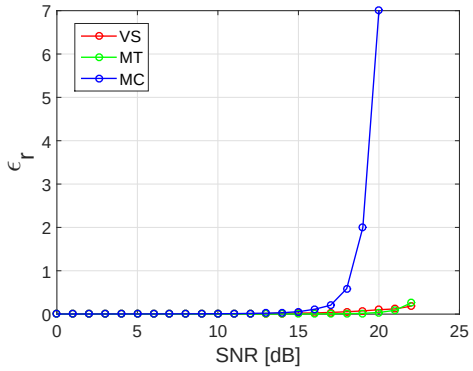


(b) SEP vs. SNR curve for VS, MT, MC and theoretical values

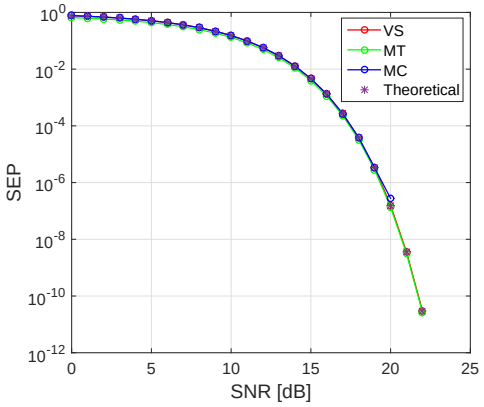


(c) Reduction in estimator variance across SNR

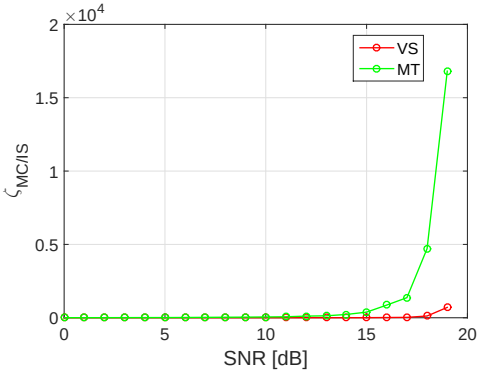
Figure 3.4: Performance of VS and MT across AWGN channel for 16 QAM



(a) Relative precision vs. SNR



(b) SEP vs. SNR curve for VS, MT, MC and theoretical values



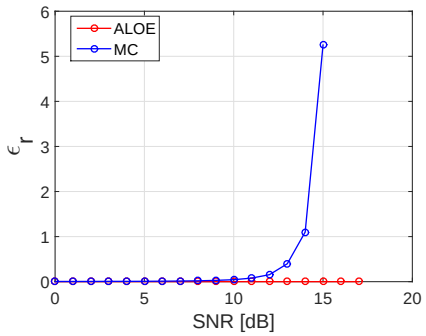
(c) Reduction in estimator variance across SNR

Figure 3.5: Performance of VS and MT across AWGN channel for 64 QAM

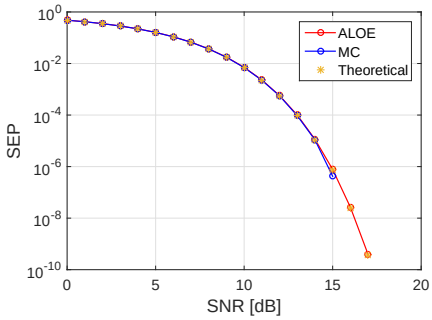
ALOE

Here, we present the simulation result of biasing the noise pdf using ALOE.

As per Fig. 3.6 for 16 QAM, ALOE obtains a significant gain over MC. Fig. 3.6 (a) depicts the gain achieved in precision by ALOE over MC and substantiate the precision of estimated SEP obtained at high SNR shown in Fig. 3.6 (b). The performance with 64 QAM shown in Fig. 3.7 indicates similar enhancements in relative precision and SEP as for 16 QAM. The superiority of ALOE over VS and MT can be confirmed via the variance reduction factor shown in Table 3.1. As the table demonstrates, enormous improvements on the order of at least 10^9, are achieved at moderate and high SNRs.

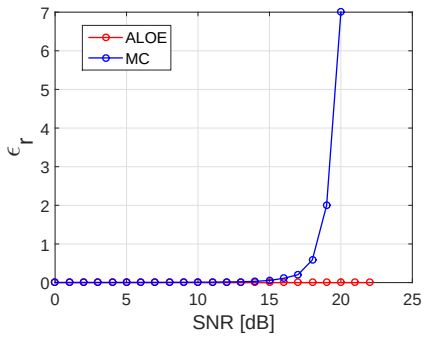


(a) Relative precision vs. SNR

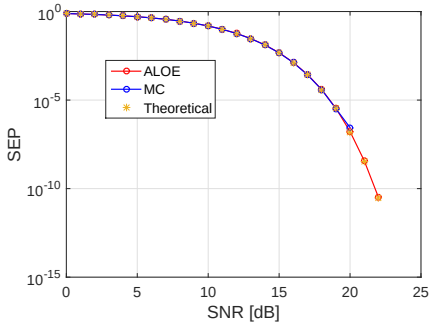


(b) SEP vs. SNR curve for ALOE, MC and theoretical values

Figure 3.6: Performance of ALOE across AWGN channel for 16 QAM



(a) Relative precision vs. SNR



(b) SEP vs. SNR curve for ALOE, MC and theoretical values

Figure 3.7: Performance of ALOE across AWGN channel for 64 QAM

SNR (dB)	16 QAM		64 QAM	
	$\zeta_{VS/ALOE}$	$\zeta_{MT/ALOE}$	$\zeta_{VS/ALOE}$	$\zeta_{MT/ALOE}$
10	1.63×10^{4}	2.00×10^{3}	1.80×10^{2}	5.72
15	9.46×10^{8}	1.64×10^{9}	6.02×10^{4}	9.54×10^{2}
16	4.14×10^{9}	2.79×10^{10}	2.44×10^{5}	4.68×10^{3}
17	7.64×10^{8}	2.29×10^{10}	2.61×10^{6}	5.25×10^{4}
18	-	-	3.63×10^{7}	9.19×10^{5}
19	-	-	1.06×10^{9}	4.47×10^{7}
20	-	-	7.55×10^{26}	7.60×10^{25}
21	-	-	1.15×10^{27}	4.79×10^{26}
22	-	-	3.27×10^{8}	5.19×10^{8}

Table 3.1: Gain obtained by ALOE over VS and MT

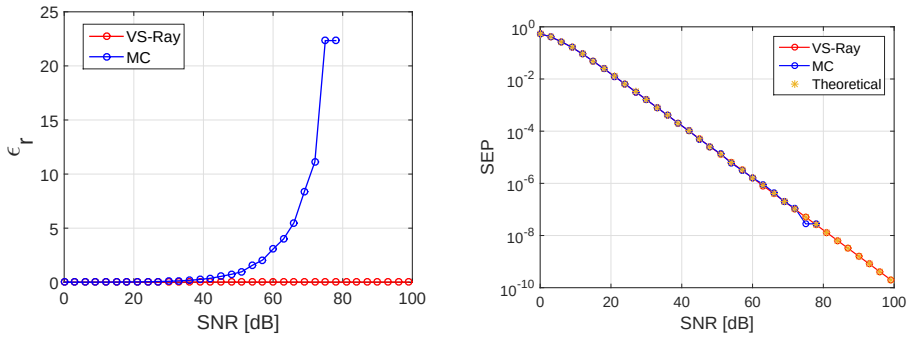
3.5 Performance Across the Rayleigh Distribution Channels

Here, we present the simulation result of biasing the Rayleigh pdf using VS-Rayleigh, where we now use $N_s = 500$.

Fig. 3.8 shows the performance with 16 QAM. From Fig. 3.8 (a) we can observe that MC has comparable precision with VS-Rayleigh at low SNR; at 30 dB the estimated SEP is 10^{-3} for which approximately 10^5 samples are required. After 30 dB, MC precision will decrease significantly and will require large number of samples to obtain the same precision as VS-Rayleigh, which can also be verified from Table 3.2. The reliability of SEP estimate can be validated by comparing the theoretical Rayleigh SEP curve, with the estimated SEP as shown in Fig. 3.8 (b). The performance with 64 QAM, shown in Fig. 3.9 and Table 3.2, obtains is similar as for 16 QAM.

SNR (dB)	16 QAM $\zeta_{MC/VS-Ray}$	64 QAM $\zeta_{MC/VS-Ray}$
30	43.48	5.56
40	354.29	56.75
50	3.58×10^{3}	524.24
60	4.19×10^{4}	6.31×10^{3}
70	3.99×10^{5}	6.71×10^{4}

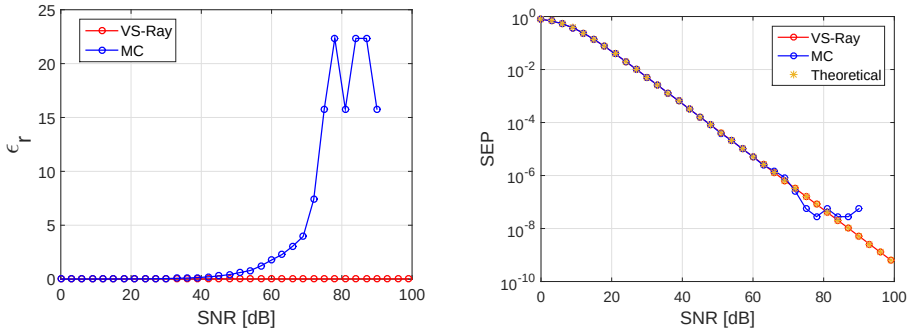
Table 3.2: Reduction in estimator variance across SNR for 16, 64 QAM and Rayleigh fading



(a) Relative precision vs. SNR

(b) SEP vs. SNR curve for VS, MC and theoretical values

Figure 3.8: Performance of VS-Rayleigh across Rayleigh fading for 16 QAM

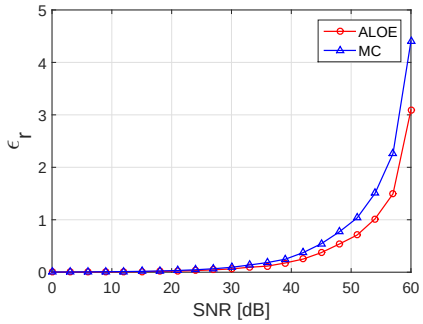


(a) Relative precision vs. SNR

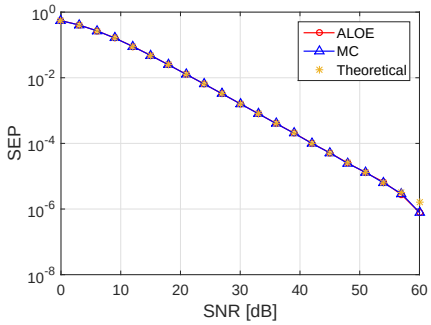
(b) SEP vs. SNR curve for VS, MC and theoretical values

Figure 3.9: Performance of VS-Rayleigh across Rayleigh fading for 64 QAM

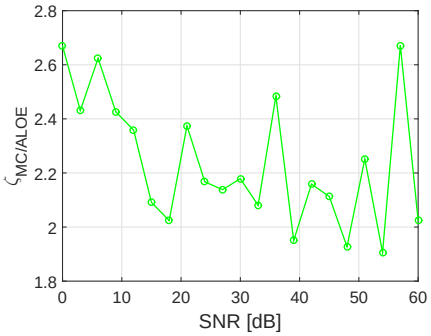
Fig. 3.10 shows the implementation of ALOE with 16 QAM across the Rayleigh fading channel. As we can observe from Fig. 3.10 (a) and Fig. 3.10 (c), the enormous gain obtained with ALOE across an AWGN channel has completely vanished for a Rayleigh fading channel. This behaviour is consistent with [8], where it was showed that only biasing the noise pdf is not efficient to obtain significant speed-up across the Rayleigh fading channel.



(a) Relative precision vs. SNR



(b) SEP vs. SNR curve for ALOE, MC and theoretical values



(c) Reduction in estimator variance across SNR

Figure 3.10: Performance of ALOE across Rayleigh fading for 16 QAM

Chapter 4

MIS for Flat Fading Channels

As observed in Chapter 3, ALOE achieved significant gain across the AWGN channel compared to MC, whereas, across the Rayleigh fading channel its performance was equivalent to MC. On the other hand, VS-Rayleigh obtained significant gain over MC. This speed-up in simulation run time of VS-Rayleigh was made possible by utilizing the knowledge of the channel pdf to create a biased channel pdf that causes more fading and thus more error events for the MC estimator. Observing the performance of ALOE across the AWGN channel, we realized that it has the potential to achieve speed-up, if we could obtain a biased channel pdf optimal for ALOE. Therefore, in this chapter we have extended ALOE across fading channels by deriving an optimal biased channel pdf for ALOE in a scalar channel system. We refer to the new method as "At Least One rare Event with Minimum Variance Prior" (ALOE-MVP). Further, the procedure to generate samples from ALOE-MVP is introduced. Lastly, the simulation results are presented to demonstrate the gain obtained by ALOE-MVP.

4.1 ALOE-MVP for Scalar Channel System

Assuming a scalar fading channel and a ZF equalizer, the model we assume is $v = x + z/h$. Since h is independent of x, the optimal biased pdf in (2.14) is now equal to

$$f^*_{opt}(v|x) = \frac{\mathbb{1}_\epsilon(v|x)\, f_{V|H,X}(v|h,x)\, f_H(h)}{P_e(x)}. \tag{4.1}$$

Note that $f_{V|H,X}$ is a Gaussian distribution while $f_H(h)$ is the prior distribution of the channel. Further manipulation of (4.1) allows us to express it as

$$f^*_{opt}(v|x) = \frac{\mathbb{1}_\epsilon(v|x)\, f_{V|H,X}(v|h,x)}{P(h,x)} \frac{P(h,x)\, f_H(h)}{P_e(x)}. \tag{4.2}$$

In (4.2), $P(h,x) = \int_{\Gamma_x} f_{V|H,X}(v|h,x)\,\mathrm{d}v$, which makes the first ratio $\frac{\mathbb{1}_\epsilon(v|x) f_{V|H,X}(v|h,x)}{P(h,x)}$ a conditional pdf. Furthermore, the second ratio $\frac{P(h,x) f_H(h)}{P_e(x)}$ is also a pdf and can be interpreted as the conditional pdf of the channel given a transmit symbol x. Hence, sampling from the optimal pdf can be seen as first sampling the channel from its new conditional pdf and then sampling the received

signal in the detector error region for the given channel. Sampling exactly in this way is unfortunately still intractable (this would result in zero estimator variance). However, sampling in the detector error region is exactly what ALOE is suitable for. Thus, we can approximate the first ratio in (4.2) with the ALOE pdf. The second ratio is the channel pdf, and as seen it depends on the transmitted symbol x. This would result in generating the channel with a new pdf for every transmitted symbol. In a real communication system simulation, the channel is generated independently of the transmitted signal and we have no possibility of changing the channel pdf given a transmitted signal. Due to this reason, we replace the second ratio in (4.2) with a channel pdf $g_H(h)$ that is independent of x (even though, as just mentioned, this is a loss). Hence, with these changes, the pdf we propose is

$$f^*(\upsilon|x) = \left[\sum_{k=1}^{K(x)} \frac{\mathbb{1}_{S_k}(\upsilon|x) f_{V|H,X}(\upsilon|h,x)}{P_k(h,x)} \alpha_k\right] g_H(h), \tag{4.3}$$

where $P_k(h,x) = \int_{\mathscr{R}^c} \mathbb{1}_{S_k}(\upsilon|x) f_{V|H,X}(\upsilon|h,x) \mathrm{d}\upsilon$, $\alpha_k = \frac{P_k(h,x)}{\tilde{p}}$ and $K(x)$ is the number of neighboring hyperplanes for transmit point x. Due to the utilization of a ZF equalizer the distance of x from any hyperplane k is equal; therefore, the upper union bound becomes $\tilde{p} = \sum_{k=1}^{K(x)} P_k(h,x) = K(x)T(h)$, where $T(h)$ is the probability of crossing one of the $K(x)$ hyperplanes and is the same for all x in a QAM constellation, expressed as

$$T(h) = \mathrm{Q}\left(\frac{A}{\frac{\sigma_z}{|h|}}\right),$$

where Q(.) represents the Q-function, $A = \frac{d_{min}}{2}$ and σ_z is standard deviation of the AWGN. With the pdf in (4.3), equation (3.29) is modified as

$$\hat{P}_e^{(ALOE-MVP)} = \frac{1}{N}\left(\sum_{n=1}^{N} \frac{f_H(h_n)}{g_H(h_n)} \frac{\tilde{p}_n}{C(\upsilon_n|x_n,h_n)}\right), \tag{4.4}$$

where $\frac{f_H(h_n)}{g_H(h_n)}$ is the weight function for channel compensation. A summation term in (4.4), i.e., the weight function of ALOE-MVP can be expressed as

$$w(h,x) = \frac{f_H(h)}{g_H(h)} \frac{K(x)T(h)}{C(\upsilon|x,h)}. \tag{4.5}$$

In (4.5), $C(\upsilon|x,h)$ is a random variable that denotes the number of half-spaces in which υ is present for a given channel realization h and transmit symbol x (see Section 3.3). The randomness of $C(\upsilon|h,x)$ is caused by the residual AWGN after conditioning on x and h. For a QAM constellation, $C(\upsilon|x,h)$ equals 1 or 2, which makes $w(h,x)$ binomially distributed. $C(\upsilon|x,h)$ equals 2 when υ is diagonally beyond the hyperplane with respect to $\mathscr{R}$ (see Fig. 3.3). The optimal $g_H(h)$ is derived by minimizing the variance of $w(h,x)$ in (4.5) averaged across x and can be shown to equal (see Appendix B)

$$g_H(h) = \frac{f_H(h) T(h) \sqrt{(1-\eta T(h))\rho}}{c}, \tag{4.6}$$

where $\eta = \frac{6a^2+9a+3}{4a^2+9a+4}$, $\rho = \frac{16a^2+36a+16}{M}$, $a = \sqrt{M}-2$ and $c = \int_h f_H(h)\, T(h)$ $\sqrt{(1-\eta T(h))\,\rho}\mathrm{d}h$. We can observe that (4.6) resembles the optimal channel pdf in (4.2) as both of them can be expressed as a product of the original channel pdf and a term that depends on the error probability of the channel.

4.2 Sampling from ALOE-MVP

From (4.6), we see that $g_H(h)$ depends on the original channel pdf $f_H(h)$. Apparently, this seems to imply that knowledge of $f_H(h)$ is necessary in order to sample from $g_H(h)$. Remarkably, we can overcome this limitation by using the rejection sampling technique. The reason for why this can be done is due to the specific expression the optimal $g_H(h)$ exhibits - as a product of the original channel pdf $f_H(h)$ and a known function of h that we can easily compute. However, if $f_H(h)$ is known in advance, then it is possible to approximate $g_H(h)$ with another tractable distribution using Kullback-Leibler divergence. Hence, in this section we will present the methodology to obtain samples from $g_H(h)$ using RS and KLD. We refer to ALOE-MVP combined with these two sampling techniques as ALOE-MVP-RS and ALOE-MVP-KLD, respectively.

4.2.1 Rejection Sampling

Rejection sampling is a Monte Carlo algorithm to generate samples from a target distribution by utilizing the samples from a proposal distribution [24]. For our scenario, $g_H(h)$ and $f_H(h)$ corresponds to the target distribution and proposal distribution, respectively. RS requires the restriction

$$B > \frac{g_H(h)}{f_H(h)}, \tag{4.7}$$

where B is a constant, finite bound on the likelihood ratio $\frac{g_H(h)}{f_H(h)}$. Inserting (4.6) in (4.7), we obtain $B = \max_h \left(\frac{T(h)\sqrt{(1-\eta T(h))\rho}}{c} \right)$ as the smallest such upper bound.

The procedure to obtain a sample from $g_H(h)$ with RS is as follows [24] :

Step 1: Generate $u \sim \mathscr{U}(0,1)$.

Step 2: Obtain a sample $\tilde{h}$ from $f_H(h)$.

Step 3: Check if $u < \frac{g_H(\tilde{h})}{B f_H(\tilde{h})}$. If true, accept $\tilde{h}$, otherwise go to step 1.

The sampled values provided by the above algorithm will be distributed according to $g_H(h)$.

It is evident from the RS procedure that B dictates the number of samples which should be "rejected" in order to obtain one sample from $g_H(h)$. B will be large when the proposal distribution is far from the target distribution, resulting in longer simulation time to obtain samples for $g_H(h)$, which can be considered a drawback of this method. In principle, one can run the above RS algorithm offline to obtain a set of channels on which simulation is performed. Moreover, we will see from the simulation results that the drawback is also compensated by the speed-up gain obtained when simulating across channels sampled from $g_H(h)$.

4.2.2 ALOE-MVP Approximation Using Kullback-Leibler Divergence

As described in Section 4.2.1, RS enables sampling from the optimal channel pdf $g_H(h)$ even though $g_H(h)$ depends on the unknown, original channel pdf $f_H(h)$. Instead of continuously generating and rejecting channel samples, another alternative is to approximate $g_H(h)$ with a tractable pdf that is easy to sample from. As an example, if it is known in advance that h is Rayleigh distributed, then it turns out that $g_H(h)$ can be well approximated by another Rayleigh distribution $Q(h;\sigma)$ with some variance σ^2. This implies that we can obtain channel realizations from $g_H(h)$ as $\sigma \times h$, where h is a sample from $f_H(h)$. Hence, this provides a significant speed-up compared to RS, as there is no need to wait for RS to provide optimally distributed channels. The approximation of $g_H(h)$ can be achieved by applying a statistical distance measure called Kullback-Leibler divergence. KLD ($D_{KL}(R||J)$) is a non-symmetric measure of the difference between two pdfs R and J, which is defined as [25]

$$D_{KL}(R||J) \triangleq \int_{-\infty}^{\infty} R(b)\, log\left(\frac{R(b)}{J(b)}\right) \mathrm{d}b. \tag{4.8}$$

For our scenario, the pdf $R(h)$ corresponds to $g_H(h)$ while $J(b)$ corresponds to a Rayleigh pdf $Q(h;\sigma)$ with a certain variance σ^2. The goal is to find a σ^2 that minimizes (4.8), which produces a pdf $Q(h;\sigma)$ that approximates $g_H(h)$. The minimization is accomplished by a numerical search of σ^2 around a starting value, where we choose the VS-Rayleigh variance $\beta_R^2\sigma_H^2$ as a starting value. The reason for this choice of starting value is that the VS-Rayleigh pdf is close to optimal for Rayleigh distributed channels. Then, $D_{KL}(R||J)$ is evaluated for different variances and the σ^2 which minimized (4.8) is chosen for approximating $g_H(h)$ with $Q(h;\sigma)$.

4.3 Performance Evaluation of ALOE-MVP Across Scalar Channels

This section demonstrates the performance of ALOE-MVP-RS and ALOE-MVP-KLD. To compute the variance of the SEP estimates, we repeat the SEP simulation N_s number of times and compute the variance of the resulting N_s SEP estimates. We choose $N_s = 300$ and each simulation consists of 50 transmissions. These simulation parameters are used for comparing ALOE-MVP-RS, ALOE-MVP-KLD and VS-Rayleigh across Rayleigh fading whereas, for simulating ALOE-MVP-RS and MC across Rician distribution we use 500 transmissions in each simulation.

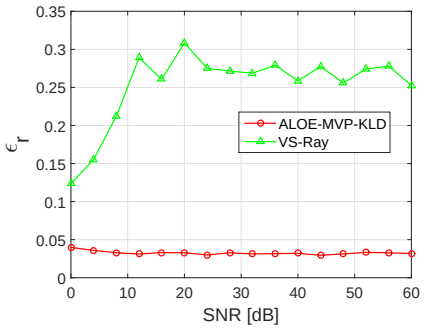
ALOE-MVP-RS

Here, we evaluate the performance of ALOE-MVP-RS across Rayleigh and Rician fading channels.

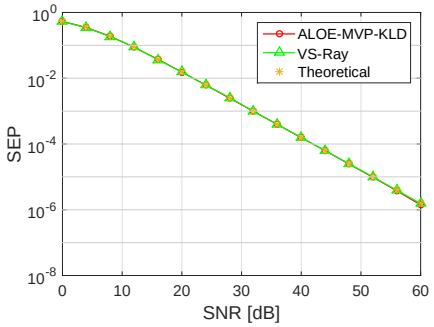
Fig. 4.1 shows the performance of ALOE-MVP-RS across Rayleigh fading with 16 QAM. Fig. 4.1 (a) depicts the gain achieved in precision by ALOE-MVP-RS

over VS-Rayleigh and substantiate the precision of estimated SEP obtained at high SNR shown in Fig. 4.1 (b). Also, significant speed-up is achieved by ALOE-MVP-RS compared to VS-Rayleigh as shown in Fig. 4.1 (c), where at SNR = 45 dB, ALOE-MVP-RS requires approximately 90 times less samples than VS-Rayleigh to obtain an estimate of the SEP with the same precision. The performance with 64 QAM is shown in Fig. 4.2, where a similar improvement in relative precision and variance reduction factor is seen as for 16 QAM.

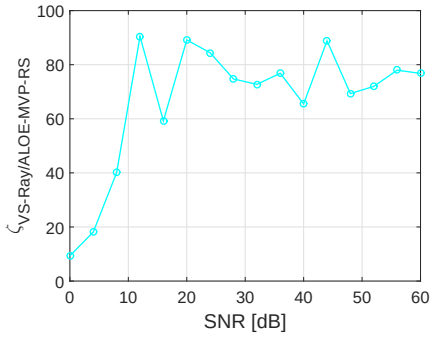
These results confirm the optimality of $g_H(h)$ for ALOE by providing significant gains to VS-Rayleigh, where the latter can be seen as the optimal biased channel pdf for the *MC estimator* in Rayleigh channels. On the other hand, $g_H(h)$ is the optimal biased channel pdf for the *ALOE estimator*, which is inherently a more efficient estimator than MC. Furthermore, what is most important, is that these significant gains were obtained *without* knowing $f_H(h)$, whereas VS-Rayleigh only works for a perfectly known Rayleigh pdf $f_H(h)$.



(a) Relative precision vs. SNR

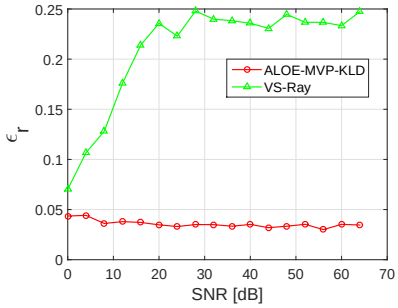


(b) SEP vs. SNR curves for ALOE-MVP-RS, VS-Rayleigh and theoretical values

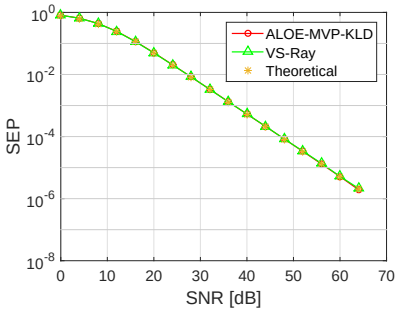


(c) Reduction in estimator variance across SNR

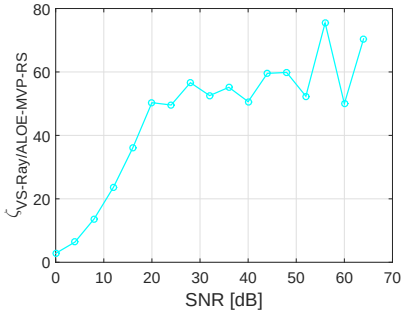
Figure 4.1: Comparison of ALOE-MVP-RS and VS-Rayleigh across Rayleigh fading for 16 QAM



(a) Relative precision vs. SNR



(b) SEP vs. SNR curves for ALOE-MVP-RS, VS-Rayleigh and theoretical values



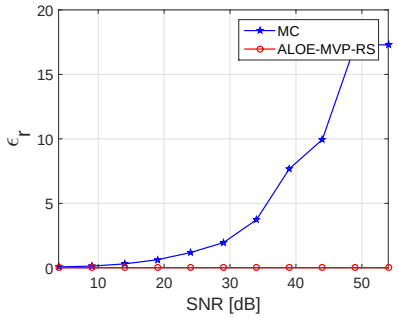
(c) Reduction in estimator variance across SNR

Figure 4.2: Comparison of ALOE-MVP-RS and VS-Rayleigh across Rayleigh fading for 64 QAM

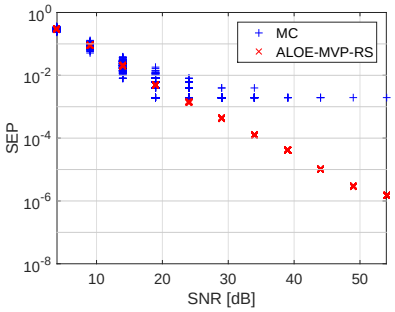
To show that this sampling technique can be utilized for arbitrarily distributed fading channels, we present the performance of ALOE-MVP-RS across the Rician distribution with Rician K-factor equal to 3 and 16 QAM in Fig. 4.3. Fig. 4.3 (a) depicts the enormous gain obtained by ALOE-MVP-RS in precision compared to MC. Comparing Figures 4.1 (a), 4.2 (a) and 4.3 (a) with Fig. 3.6 (a), it is striking that ALOE-MVP-RS achieves the same constant performance behaviour across a fading channel as across an AWGN channel. It is as if ALOE-MVP-RS completely removes the detrimental impact of fading on simulation time. Fig. 4.3 (b) plots the SEP obtained in each simulation so to visualize the variation of the N_s different SEP estimates for each SNR. As seen, ALOE-MVP-RS provides remarkably accurate estimates across the whole SNR range, whereas MC estimation becomes very unreliable at relatively low SNRs.

Also, we can observe from Figures 4.1 (b) and 4.3 (b) that ALOE-MVP-RS achieves good approximation of the SEP across all SNRs with only 50 samples per simulation. In contrast, MC could only obtain good SEP estimate until SNR = 20 dB with 500 samples, which again depicts the massive gain obtained in simulation

run time with ALOE-MVP-RS.



(a) Relative precision vs. SNR



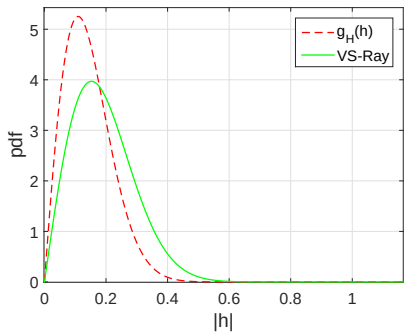
(b) SEP vs. SNR curves for ALOE-MVP-RS and MC

Figure 4.3: Comparison of ALOE-MVP-RS and MC across Rician distributed fading channel with K-factor = 3 for 16 QAM

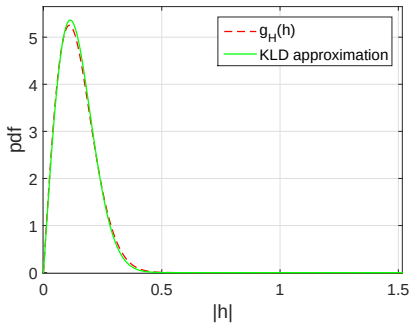
ALOE-MVP-KLD

Here, we evaluate the performance of ALOE-MVP-KLD across Rayleigh fading channels.

As described in Section 4.2.2, the approximation of $g_H(h)$ by a Rayleigh pdf can be corroborated from Fig. 4.4. Fig. 4.4 (a) indicates that $g_H(h)$ resembles the Rayleigh distribution. In Fig. 4.4 (b), we can observe that KL divergence provides a satisfactory approximation of $g_H(h)$ with a Rayleigh pdf.



(a) $g_H(h)$ and VS-Rayleigh pdf

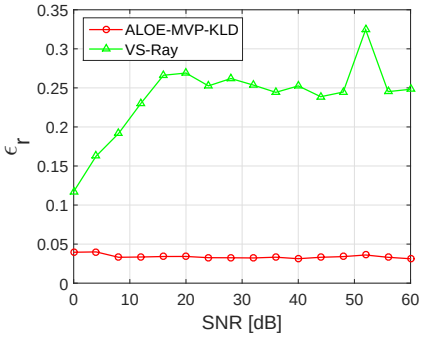


(b) $g_H(h)$ approximated by KLD

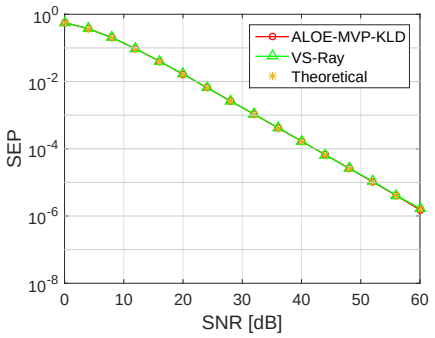
Figure 4.4: $g_H(h)$ approximation using KLD

According to Fig. 4.5 for 16 QAM, ALOE-MVP-KLD achieves a significant gain over VS-Rayleigh. Fig. 4.5 (a) shows that ALOE-MVP-KLD has obtained better and stable precision than VS-Rayleigh. The reliability of SEP estimate

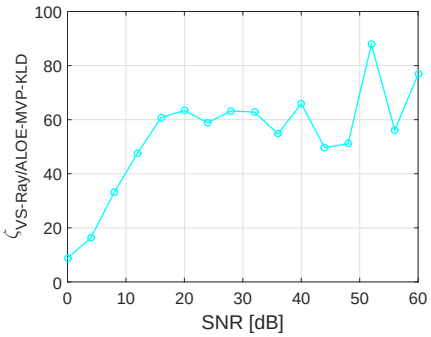
obtained by ALOE-MVP-KLD can be validated by comparing it with theoretical Rayleigh SEP curve as shown in Fig. 4.5 (b). Also, significant reduction in estimator variance is achieved by ALOE-MVP-KLD as shown in Fig. 4.5 (c). The performance with 64 QAM is shown in Fig. 4.6, where a similar improvement in relative precision and variance reduction factor is seen as for 16 QAM.



(a) Relative precision vs. SNR

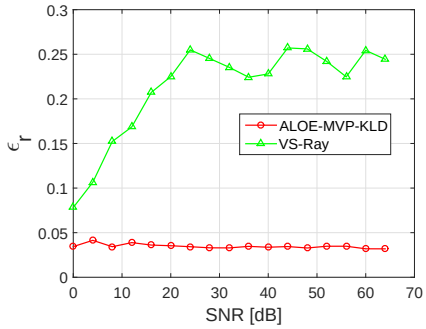


(b) SEP vs. SNR curves for ALOE-MVP-KLD, VS-Rayleigh and theoretical value

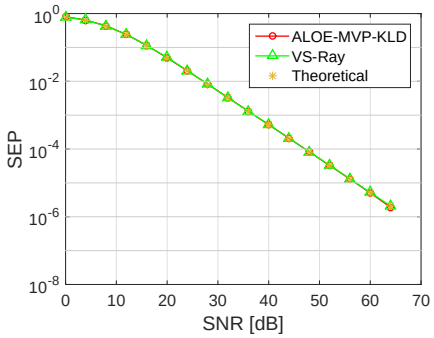


(c) Reduction in estimator variance across SNR

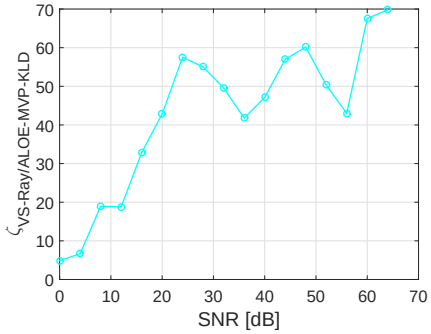
Figure 4.5: Comparison of ALOE-MVP-KLD and VS-Rayleigh across Rayleigh fading for 16 QAM



(a) Relative precision vs. SNR



(b) SEP vs. SNR curves for ALOE-MVP-KLD, VS-Rayleigh and theoretical values



(c) Reduction in estimator variance across SNR

Figure 4.6: Comparison of ALOE-MVP-KLD and VS-Rayleigh across Rayleigh fading for 64 QAM

Chapter 5

MIS for Frequency Selective Channels

In Chapter 4, ALOE-MVP in a scalar channel system obtained significant gain over MC and VS-Rayleigh, as its variance was minimized with respect to the pdf of a scalar channel. However, to apply ALOE-MVP across frequency selective channels in an OFDM system we need an optimal channel pdf for a vector channel (representing the tapped delay model). In this chapter we will derive ALOE-MVP for an OFDM system across a frequency selective channel that minimizes the estimator variance for an OFDM symbol consisting of N subcarriers. We refer to ALOE-MVP in an OFDM system as ALOE-MVP-OFDM. Further, the simulation results are presented to demonstrate the gain obtained by ALOE-MVP-OFDM.

5.1 ALOE-MVP-OFDM for General Distribution of Channel Delays

Simulation across a frequency selective channel introduces correlation between subcarriers. Due to this reason, we will derive ALOE-MVP-OFDM to minimize the variance of the sum of weights across the N subcarriers in an OFDM symbol. Let $\mathbf{h} = [h_0, h_1 ..., h_{L-1}]$ denote the vector of L channel taps, x_i the QAM symbol on subcarrier i and $w_i(\mathbf{h}, x_i)$ the weight that multiplies the detector output on subcarrier i. The considered variance can be expressed with Bienaymé's identity [26] as

$$\begin{aligned}\mathrm{Var}\left(\sum_{i=1}^{N} w_i\left(\mathbf{h}, x_i\right)\right) &= \sum_{i=1}^{N}\sum_{j=1}^{N} \mathrm{Cov}\left(w_i(\mathbf{h}, x_i), w_j(\mathbf{h}, x_j)\right) \\ &= \sum_{i=1}^{N} \mathrm{Var}(w_i(\mathbf{h}, x_i)) + 2 \sum_{1\le i<j\le N} \mathrm{Cov}(w_i(\mathbf{h}, x_i), w_j(\mathbf{h}, x_j)),\end{aligned} \tag{5.1}$$

where Cov(.,.) is covariance. The sum of the variances in (5.1) equals

$$\sum_{i=1}^{N} \mathrm{Var}(w_i(\mathbf{h}, x_i)) = \sum_{i=1}^{N} E_{g_{\mathbf{H}}(\mathbf{h}),x_i} \left[(w_i(\mathbf{h}, x_i) - E_{g_{\mathbf{H}}(\mathbf{h}),x_i}[w_i(\mathbf{h}, x_i)])^2 \right]$$
$$= \sum_{i=1}^{N} E_{g_{\mathbf{H}}(\mathbf{h}),x_i} \left[w_i(\mathbf{h}, x_i)^2 \right] - \sum_{i=1}^{N} P_{e_i}^2, \tag{5.2}$$

where P_{e_i} is the SEP of the i_{th} subcarrier and $g_{\mathbf{H}}(\mathbf{h})$ is the new channel pdf. Using (B.9), (5.2) becomes

$$\sum_{i=1}^{N} \mathrm{Var}(w_i(\mathbf{h}, x_i)) = \sum_{i=1}^{N} \int_{\mathbf{h}} \left[\frac{\left(f_{\mathbf{H}}(\mathbf{h}) T_i(\mathbf{h})\right)^2}{g_{\mathbf{H}}(\mathbf{h})} \left(1 - \eta T_i(\mathbf{h})\right) \rho \right] \mathrm{d}\mathbf{h} - \sum_{i=1}^{N} P_{e_i}^2, \tag{5.3}$$

where $\eta = \frac{6a^2+9a+3}{4a^2+9a+4}$, $\rho = \frac{16a^2+36a+16}{M}$, $a = \sqrt{M} - 2$, $f_{\mathbf{H}}(\mathbf{h})$ is the original channel pdf and $T_i(\mathbf{h})$ can be written as

$$T_i(\mathbf{h}) = \mathrm{Q}\left(\frac{A}{\left(\frac{\sigma_z}{\mathrm{DFT}(\mathbf{h})_i} \right)} \right),$$

where $\mathrm{DFT}(\mathbf{h})_i$ represents the channel frequency response on subcarrier i, which can be obtained from (2.24), $A = \frac{d_{min}}{2}$ and σ_z is the standard deviation of the AWGN. The covariance term in (5.1) can be written as

$$\begin{aligned}
& 2 \sum_{1 \le i < j \le N} \mathrm{Cov}\Big(w_i(\mathbf{h}, x_i), w_j(\mathbf{h}, x_j) \Big) \\
=& 2 \sum_{1 \le i < j \le N} E_{g_{\mathbf{H}}(\mathbf{h}),x_i,x_j} \Big[\left(w_i(\mathbf{h}, x_i) - E_{g_{\mathbf{H}}(\mathbf{h}),x_i}[w_i(\mathbf{h}, x_i)] \right) \\
& \times \left(w_j(\mathbf{h}, x_j) - E_{g_{\mathbf{H}}(\mathbf{h}),x_j}[w_j(\mathbf{h}, x_j)] \right) \Big] \\
=& 2 \sum_{1 \le i < j \le N} E_{g_{\mathbf{H}}(\mathbf{h}),x_i,x_j} \left[w_i(\mathbf{h}, x_i) w_j(\mathbf{h}, x_j) \right] - 2 \sum_{1 \le i < j \le N} P_{e_i} P_{e_j}.
\end{aligned} \tag{5.4}$$

For a QAM constellation there are potentially 9 different values of $E_{g_{\mathbf{H}}(\mathbf{h}),x_i,x_j} \left[w_i(\mathbf{h}, x_i) w_j(\mathbf{h}, x_j) \right]$ corresponding to the 9 different combinations of corner, edge or inner QAM points for the two subcarriers i and j. Taking into account all the possible combinations we get (see Appendix C)

$$2 \sum_{1 \le i < j \le N} \mathrm{Cov}\Big(w_i(\mathbf{h}, x_i), w_j(\mathbf{h}, x_j) \Big) = \int_{\mathbf{h}} \frac{f_{\mathbf{H}}^2(\mathbf{h})}{g_{\mathbf{H}}(\mathbf{h})} \left[\left(\sum_{i=1}^{N} t_i \right)^2 - \sum_{i=1}^{N} t_i^2 \right] \mathrm{d}\mathbf{h}$$
$$- 2 \sum_{1 \le i < j \le N} P_{e_i} P_{e_j}. \tag{5.5}$$

where $t_i = T_i(\mathbf{h}) \left(\frac{4a^2+12a+8}{M} - \frac{4a^2+8a+4}{M} T_i(\mathbf{h}) \right)$.

Inserting (5.3) and (5.5) into (5.1), it follows that

$$\mathrm{Var}\left(\sum_{i=1}^{N} w_i\left(\mathbf{h}, x_i\right)\right) = \int_{\mathbf{h}} \frac{f_{\mathbf{H}}^2\left(\mathbf{h}\right)}{g_{\mathbf{H}}\left(\mathbf{h}\right)} \Bigg\{ \sum_{i=1}^{N} T_i^2\left(\mathbf{h}\right)\left(\xi_1 - \xi_2 T_i\left(\mathbf{h}\right) - \xi_3^2 T_i^2\left(\mathbf{h}\right)\right) + \left[\sum_{i=1}^{N} T_i\left(\mathbf{h}\right)\left(\xi_4 - \xi_3 T_i\left(\mathbf{h}\right)\right)\right]^2 \Bigg\} \mathrm{d}\mathbf{h} - \left(\sum_{i=1}^{N} P_{e_i}\right)^2, \quad (5.6)$$

where $\xi_1 = \frac{16a^2\left(M-a^2-6a-13\right)+12a(3M-16)+16(M-4)}{M^2}$, $\xi_2 = \frac{8a^2\left(3M-4a^2-20a-36\right)}{M^2} + \frac{4a(9M-56)+4(3M-16)}{M^2}$, $\xi_3 = \frac{4a^2+8a+4}{M}$ and $\xi_4 = \frac{4a^2+12a+8}{M}$.

As in Chapter 4, the integral in (5.6) can be minimized by obtaining a Lagrangian function with the constraint $\int_{\mathbf{h}} g_{\mathbf{H}}(\mathbf{h})\mathrm{d}\mathbf{h} = 1$, and then applying the Euler-Lagrange equation to obtain the optimal $g_{\mathbf{H}}(\mathbf{h})$ for ALOE-MVP-OFDM as

$$g_{\mathbf{H}}\left(\mathbf{h}\right) = \frac{f_{\mathbf{H}}\left(\mathbf{h}\right)\sqrt{\sum_{i=1}^{N} {T_i}^2\left(\mathbf{h}\right)\left(\xi_1 - \xi_2 T_i\left(\mathbf{h}\right) - \xi_3^2 T_i^2\left(\mathbf{h}\right)\right) + \left[\sum_{i=1}^{N} T_i\left(\mathbf{h}\right)\left(\xi_4 - \xi_3 T_i\left(\mathbf{h}\right)\right)\right]^2}}{c}, \quad (5.7)$$

where c equals the integral of the numerator across $\mathbf{h}$. We can observe that (5.7) resembles the optimal channel pdf in (4.2) as both of them can be expressed as a product of the original channel pdf $f_{\mathbf{H}}(\mathbf{h})$ and a known function of the channel realization. Hence, as in Chapter 4, we can apply RS to obtain channel vector samples from $g_{\mathbf{H}}(\mathbf{h})$. This technique is named ALOE-MVP-OFDM-RS. Moreover, since we know that for independent Rayleigh channel taps the subcarrier channels are also Rayleigh distributed (albeit correlated) with the same Rayleigh distribution, we can apply the KLD method from Chapter 4 to approximate the *optimal* subcarrier channel distribution with another Rayleigh distribution. We denote this method as ALOE-MVP-OFDM-KLD. In this specific case, as will be explained below, the KLD approximation has even better performance than sampling according to $g_{\mathbf{H}}\left(\mathbf{h}\right)$ in (5.7). The following section will shed more light on these techniques.

5.2 Performance Evaluation of ALOE-MVP-OFDM

In this section, we will demonstrate the performance of ALOE-MVP-OFDM-RS and ALOE-MVP-OFDM-KLD across 3GPP multipath fading channel models in our OFDM system as introduced in Section 2.3. To compute the variance of the SEP estimates, we repeat the SEP simulation N_s number of times and compute the variance of the resulting N_s SEP estimates. Table 5.1 summarizes the OFDM system parameters utilized for simulation. The parameters of the 3GPP channel models extended pedestrian A model (EPA) and extended vehicular A model (EVA) are presented in Table 5.2 [28].

Parameter	Value
Total Bandwidth (W)	3.072 MHz
Subcarrier frequency spacing (SCS)	30 KHz
Total number of OFDM symbols (N)	140
Number of subcarrier (N_{sub})	1024
Length of cyclic prefix (N_{cp})	72
OFDM symbol duration (T)	$33.33(\mu s)$
Cyclic prefix duration (T_{cp})	$2.34(\mu s)$
Modulation order (M)	$16, 64$ QAM
Number of simulations (N_s)	30

Table 5.1: OFDM system parameters

Extended Pedestrian A model		Extended Vehicular A Model	
Excess tap delay (ns)	Relative power (dB)	Excess tap delay (ns)	Relative power (dB)
0	0.0	0	0.0
30	-1.0	30	-1.5
70	-2.0	150	-1.4
90	-3.0	310	-3.6
110	-8.0	370	-0.6
190	-17.2	710	-9.1
410	-20.8	1090	-7.0
-	-	1730	-12.0
-	-	2510	-16.9

Table 5.2: Parameters of 3GPP EPA and EVA models

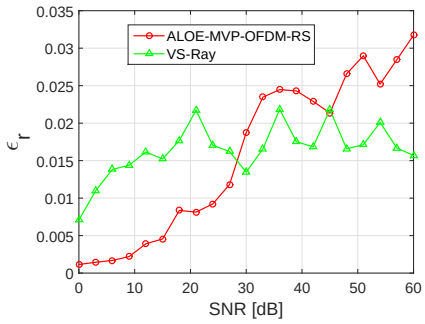
ALOE-MVP-OFDM-RS

Here, we evaluate the performance of ALOE-MVP-OFDM-RS by sampling from $g_{\mathbf{H}}(\mathbf{h})$ with RS across EPA and EVA models.

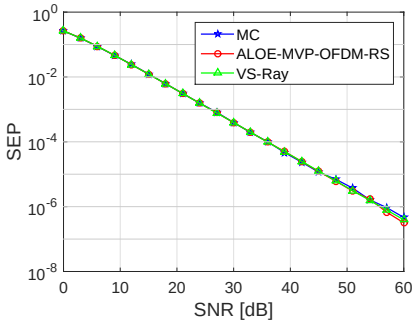
Fig. 5.1 shows the performance of ALOE-MVP-OFDM-RS with 16 QAM across EVA model. Fig. 5.1 (a) illustrates the gain obtained in precision by ALOE-MVP-OFDM-RS compared to VS-Rayleigh at low SNRs, while from SNR = 30 to 50 dB its performance is equivalent to VS-Rayleigh and at SNR greater than 50 dB it has precision slightly worse than VS-Rayleigh. Same behaviour can be seen in estimator reduction factor in Fig. 5.1 (c). The reason for slightly worse performance of ALOE-MVP-OFDM-RS at high SNR, compared to VS-Rayleigh, is because ALOE-MVP-OFDM-RS minimizes the variance of the sum of weights which is suboptimal compared to minimizing the variance of the weight on each subcarrier. Since EPA and EVA models have i.i.d. Rayleigh fading channel tap,

each subcarrier channel in (2.24) has the same Rayleigh distribution. Thus, VS-Rayleigh uses this known Rayleigh channel pdf on each subcarrier and minimizes the variance of the MC estimator on each subcarrier.

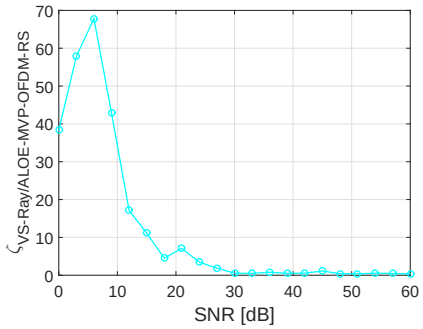
On the other hand, Table 5.3 and Fig. 5.1 (d) concurs that ALOE-MVP-OFDM-RS has achieved significant gain in precision and estimator variance compared to MC. ALOE-MVP-OFDM-RS at SNR = 50 dB, requires approximately 1300 times less samples than MC to achieve the same precision. Also, according to the rule of thumb (see Section 1.1), MC would require 10^8 samples to achieve good precision until SEP 10^{-6}, which corresponds to simulating an OFDM system with approximately 97650 OFDM symbols, which is infeasible to simulate. ALOE-MVP-OFDM-RS gives unbiased SEP estimates which can be seen in Fig. 5.1 (b). The performance with 64 QAM is shown in Fig. 5.2 and Table 5.4, where a similar behaviour across EPA model is seen as for 16 QAM across EVA model.



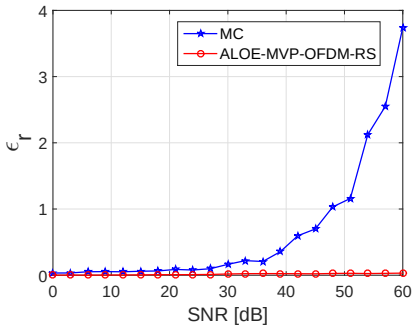
(a) Relative precision vs. SNR



(b) SEP vs. SNR curves for ALOE-MVP-OFDM-RS, VS-Rayleigh and MC



(c) Reduction in estimator variance across SNR

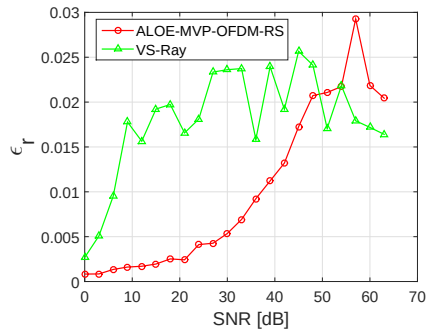


(d) Comparison of ϵ_r across SNR for ALOE-MVP-OFDM-RS and MC

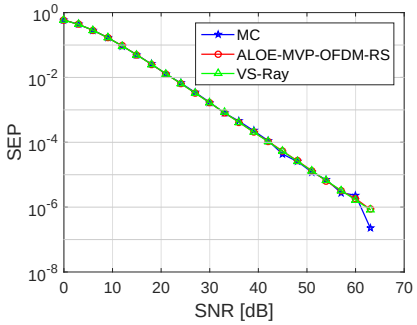
Figure 5.1: Comparison of ALOE-MVP-OFDM-RS, VS-Rayleigh and MC across EVA model for 16 QAM

SNR(dB)	$\zeta_{MC/ALOE-MVP-OFDM-RS}$
40	183.27
45	1.49×10^3
50	1.29×10^3
55	6.51×10^3
60	2.70×10^4

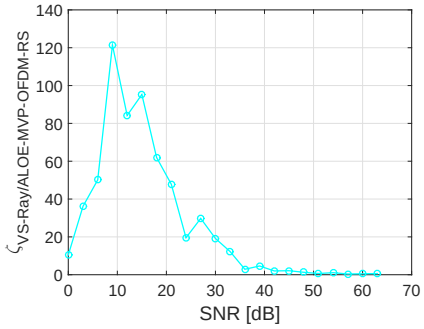
Table 5.3: Reduction in estimator variance across SNR for 16 QAM and EVA model



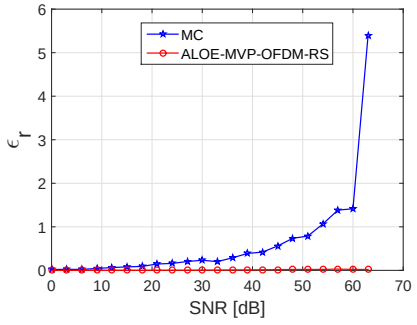
(a) Relative precision vs. SNR



(b) SEP vs. SNR curves for ALOE-MVP-OFDM-RS, VS-Rayleigh and MC



(c) Reduction in estimator variance across SNR



(d) Comparison of ϵ_r across SNR for ALOE-MVP-OFDM-RS and MC

Figure 5.2: Comparison of ALOE-MVP-OFDM-RS, VS-Rayleigh and MC across EPA model for 64 QAM

SNR(dB)	$\zeta_{MC/ALOE-MVP-OFDM-RS}$
40	1.46×10^3
45	2.04×10^3
50	2.53×10^3
55	2.89×10^3
60	1.27×10^4

Table 5.4: Reduction in estimator variance across SNR for 64 QAM and EPA model

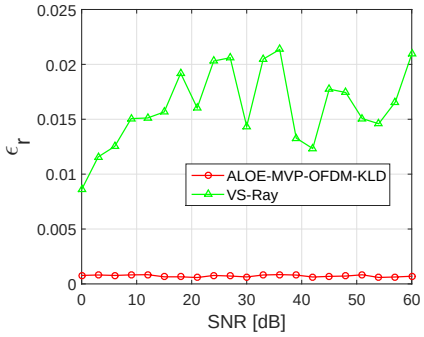
These results illustrate that ALOE-MVP-OFDM-RS, which requires no knowledge on the distribution of the channel taps, obtains satisfactory performance compared to VS-Rayleigh at high SNR. Moreover, compared to MC, which currently is the only method to compute SEP in case where we have no knowledge of channel pdf, ALOE-MVP-OFDM-RS achieves significant gain in precision compared to MC. This results into significant reduction in simulation run time.

The $g_{\mathbf{H}}(\mathbf{h})$ obtained in (5.7) can be utilized to simulate an OFDM system across channels whose distribution is different for each subcarrier, e.g. non-standard, correlated Gaussian distributions, which is not possible with VS-Rayleigh as it requires that each subcarrier must have the same Rayleigh distribution. Although ALOE-MVP-OFDM-RS requires additional computations to obtain samples from $g_{\mathbf{H}}(\mathbf{h})$, this reduction in computational efficiency is compensated by the enormous speed-up gain obtained by ALOE-MVP-OFDM-RS when simulating across the chosen channels.

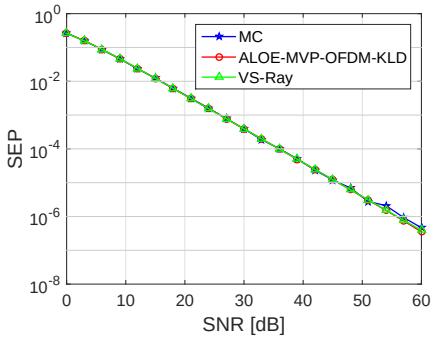
ALOE-MVP-OFDM-KLD

Here, we evaluate the performance of ALOE-MVP-OFDM-KLD across EPA and EVA models. Just as VS-Rayleigh, we utilize the fact that we know the distribution of the Rayleigh fading channel taps, which produce the same known Rayleigh distribution on each subcarrier, and obtain an approximation to $g_H(h)$ from (4.6) on each subcarrier using KLD. Hence, this method results in obtaining a variance scaling factor σ that is used to multiply each channel tap in the EPA and EVA models. As VS-Rayleigh, this method is minimizing the weight variance on each subcarrier.

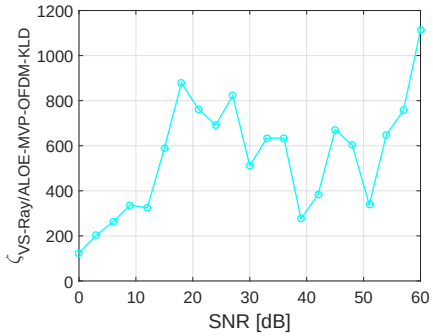
According to Fig. 5.3 for 16 QAM, ALOE-MVP-OFDM-KLD has achieved significant gain over VS-Rayleigh. This should not come as a surprise since ALOE-MVP-KLD, with $g_H(h)$ in (4.6), has already been shown to be a much better sampling method than VS-Rayleigh across scalar channels. Fig. 5.3 (a) shows that ALOE-MVP-OFDM-KLD has obtained a better and more stable precision than VS-Rayleigh across all SNRs. The reliability of the SEP estimate obtained by ALOE-MVP-OFDM-KLD can be validated by comparing it with MC SEP curve as shown in Fig. 5.3 (b). Also, significant reduction in estimator variance is achieved by ALOE-MVP-OFDM-KLD as shown in Fig. 5.3 (c). The performance with 64 QAM is shown in Fig. 5.4, where a similar improvement in relative precision and SEP is seen as for 16 QAM.



(a) Relative precision vs. SNR

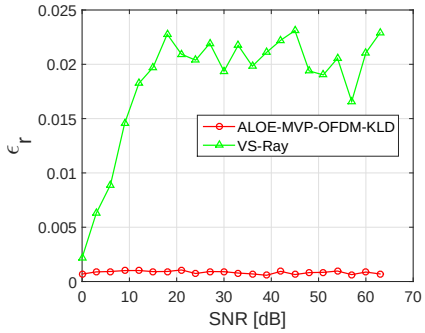


(b) SEP vs. SNR curves for ALOE-MVP-OFDM-KLD, VS-Rayleigh and MC

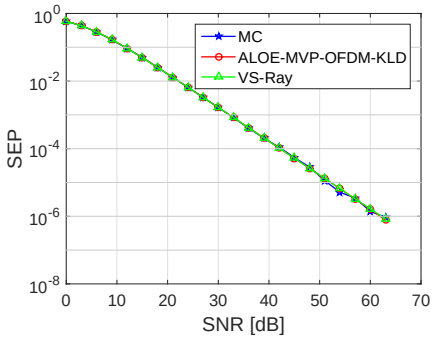


(c) Reduction in estimator variance across SNR

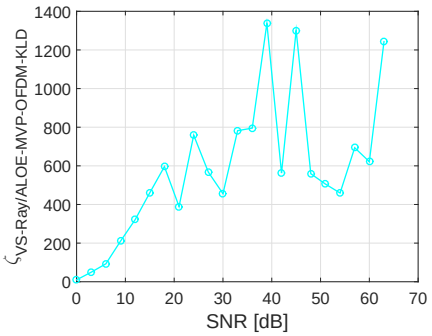
Figure 5.3: Comparison of ALOE-MVP-OFDM-KLD and VS-Rayleigh across EVA model for 16 QAM



(a) Relative precision vs. SNR



(b) SEP vs. SNR curves for ALOE-MVP-OFDM-KLD, VS-Rayleigh and MC



(c) Reduction in estimator variance across SNR

Figure 5.4: Comparison of ALOE-MVP-OFDM-KLD and VS-Rayleigh across EPA model for 64 QAM

ALOE-MVP-OFDM-RS failed to achieve similar gains as ALOE-MVP-OFDM-KLD which uses the fact that each subcarrier has the same distribution to minimize the weight variance on each subcarrier. However, KLD for an OFDM systems will only work if the channel pdf is perfectly known and generates the same, known, pdf on each subcarrier. Despite the loss in gains, ALOE-MVP-OFDM-RS has this advantage over ALOE-MVP-OFDM-KLD, as it can be utilized to simulate any channel distribution with enormous gains compared to naive MC.

Chapter 6
Conclusion and Future Work

6.1 Conclusion

URLLC, a key technology of 5G, has imposed stringent requirements on latency and reliability to the tactile internet applications. In order to meet these requirement, simulations with BLER as low as of 10^{-6} is needed for industrial automation systems. Hence, in this thesis we have discussed, evaluated and compared MC and IS simulation techniques across different channels. A new and efficient IS simulation techniques for scalar and OFDM system across fading channels is introduced and its performance evaluated with simulations.

IS techniques work by modifying the statistical properties of the input distribution to reduce estimator variance, resulting in shorter simulation run time. For a communication system this corresponds to altering either channel, noise, transmit symbol distribution or a combination of them. We started with the evaluation of altering noise pdf via VS and MT in an AWGN channel and obtained significant speed-up and gains in estimator variance compared to MC. Further, we evaluated the VS-Rayleigh technique which alters the variance of the Rayleigh distribution and as expected huge gains were achieved compared to MC in both precision and estimator variance. However, optimal IS density shows that the biased distribution should be proportional to the original (input) distribution in the error region. VS and MT of noise nor VS-Rayleigh techniques satisfy this property.

To fully utilize the power of IS by biasing the distribution according to the optimal IS density, we evaluated a current state-of-the-art MIS technique called ALOE. As expected, ALOE achieved massive gains over VS and MT of noise across AWGN channel both in precision and estimator variance. However, across fading channels, the AWGN gains completely vanished and ALOE obtained equivalent performance as MC. This behaviour was in-line with the results obtained in [8]. This result left us astonished (even though we expected it could happen), however at the same time the significant gain of ALOE across the AWGN channel inspired us to dig deeper in the methodology of ALOE and try to understand the reason for its failure across the fading channel.

As it followed, we found that the culprit for the loss in gains of ALOE across fading channels is that instead of generating channels with the original channel pdf, we need a channel pdf that is optimized for ALOE. First we derived an optimal channel pdf for the scalar channel system and we named ALOE with the improved

channel pdf as "ALOE-MVP". To obtain samples from ALOE-MVP, we introduced two different techniques based on RS and KLD minimization; the two methods were named ALOE-MVP-RS and ALOE-MVP-KLD, respectively. ALOE-MVP-RS can be used for any original channel pdf, while ALOE-MVP-KLD assumes knowledge of the original channel pdf in order to approximate it with another pdf from which we easily can obtain samples. ALOE-MVP-RS was evaluated across a fading channel and the results left us awestruck. Simulations showed us that ALOE-MVP-RS completely removed the detrimental effect of fading. For Rayleigh fading, it performed much better than the current state-of-the-art VS-Rayleigh which needs knowledge of the Rayleigh fading pdf. ALOE-MVP-KLD also obtained similar gains over VS-Rayleigh, which enables even simpler simulation than ALOE-MVP-RS across Rayleigh fading since it results in a mere scaling of the original Rayleigh fading samples.

Next, we derived an optimal channel pdf for ALOE in an OFDM system, by minimizing the variance of the sum of weights across the subcarriers and referred to this method as "ALOE-MVP-OFDM". We showed how to obtain samples from ALOE-MVP-OFDM with RS and named this method ALOE-MVP-OFDM-RS. ALOE-MVP-OFDM-RS gave a satisfactory performance compared to VS-Rayleigh, while obtaining massive gain in both precision and estimator variance compared to MC. For independent Rayleigh distributed channel taps, each subcarrier channel has the same Rayleigh distribution, which enables the use of the KLD approximation method from Chapter 4 to minimize the weight variance on each subcarrier. This method is superior than minimizing the variance of the total weight sum (something that ALOE-MVP-OFDM-RS does). The KLD approximation method in OFDM systems is denoted as ALOE-MVP-OFDM-KLD and provides a large speed-up compared to ALOE-MVP-OFDM-RS and VS-Rayleigh for independent Rayleigh distributed channel taps.

To summarize, the main contribution of this thesis work are the IS methods ALOE-MVP-RS, ALOE-MVP-KLD for scalar channels and ALOE-MVP-OFDM-RS, ALOE-MVP-OFDM-KLD for OFDM channels. As described above, the RS methods are suitable for simulation when no knowledge of the channel pdf is at hand. On the other hand, if the channel pdf is known, then the KLD methods have the potential to simplify the simulation process significantly as well as provide additional speed-up in OFDM systems compared to RS methods.

6.2 Future Work

In order to extend our new IS methods further, following points can be considered:

- Implementation in different types of detectors, for example minimum mean-square error.
- How will the methods behave for an imperfect channel estimation? This question should be further investigated.
- Further examination is required to implement the methods in a SISO coded system.

- Further research is needed to find all the possible channel tap distributions on which the KLD approximation can be applied.

References

[1] *Study on Scenarios and Requirements for Next Generation Access Technologies*, 3GPP TR 38.913 v15.0.0, 2018.

[2] *Study on physical layer enhancements for NR ultra-reliable and low latency case (URLLC)*, 3GPP TR 38.824 v16.0.0, 2019.

[3] P. Smith, M. Shafi, and H. Gao, "Quick simulation: a review of importance sampling techniques in communications systems," *IEEE Journal on Selected Areas in Communications*, vol. 15, no. 4, pp. 597–613, 1997. doi: 10.1109/49.585771.

[4] S. S. Sawilowsky, "You think youve got trivials?," *Journal of Modern Applied Statistical Methods*, vol. 2, no. 1, p. 21, 2003.

[5] A. M. Law, W. D. Kelton, and W. D. Kelton, *Simulation modeling and analysis*, vol. 3. Mcgraw-hill New York, 2007.

[6] H. Kahn, *Use of Different Monte Carlo Sampling Techniques*. Santa Monica, CA: RAND Corporation, 1955.

[7] K. Shanmugam and P. Balaban, "A modified Monte-Carlo simulation technique for the evaluation of error rate in digital communication systems," *IEEE Transactions on Communications*, vol. 28, no. 11, pp. 1916–1924, 1980. doi: 10.1109/TCOM.1980.1094613.

[8] A. Dekorsy and V. Kuhn, "Applicability of importance sampling for Rayleigh fading mobile radio channels," in *Proceedings of ISSSTA'95 International Symposium on Spread Spectrum Techniques and Applications*, vol. 1, pp. 224–228 vol.1, 1996. doi: 10.1109/ISSSTA.1996.563775.

[9] T. Hesterberg, "Weighted average importance sampling and defensive mixture distributions," *Technometrics*, vol. 37, no. 2, pp. 185–194, 1995.

[10] V. Elvira and I. Santamaría, "Multiple importance sampling for efficient symbol error rate estimation," *IEEE Signal Processing Letters*, vol. 26, no. 3, pp. 420–424, 2019. doi: 10.1109/LSP.2019.2892835.

[11] P. Wijesinghe, "Biased Monte Carlo methods for efficient simulation of communication systems," 2011. [Online]. Available: `https:`

//researchdirect.westernsydney.edu.au/islandora/object/uws:11188/datastream/PDF/view

[12] N. Metropolis, "The beginning of the Monte Carlo method," Los Alamos Sci., no. 15, pp. 125130, 1987.

[13] C. Andrieu, N. De Freitas, A. Doucet, and M. I. Jordan, "An introduction to MCMC for machine learning," *Machine learning*, vol. 50, no. 1, pp. 5–43, 2003. doi: 10.1023/A:1020281327116.

[14] P. Hahn and M. Jeruchim, "Developments in the theory and application of importance sampling," *IEEE Transactions on Communications*, vol. 35, no. 7, pp. 706–714, 1987. doi: 10.1109/TCOM.1987.1096838.

[15] S. X. Chen, P. J. Smith, M. Shafi, and D. Vere-Jones, "Some improvements to conventional importance sampling techniques for coded systems using Viterbi decoding," *Electron. Lett.*, vol. 26, no. 12, 1990.

[16] A. Goldsmith, *Wireless communications*. Cambridge university press, 2005.

[17] J. G. Proakis and M. Salehi, *Digital communications*, vol. 4. McGraw-hill New York, 2001.

[18] R. Prasad, *OFDM for wireless communications systems*. Artech House, 2004.

[19] H. Sari, G. Karam, and I. Jeanclaude, "Transmission techniques for digital terrestrial TV broadcasting," *IEEE Communications Magazine*, vol. 33, no. 2, pp. 100–109, 1995. doi: 10.1109/35.350382.

[20] P. Wijesinghe, U. Gunawardana, and R. Liyanapathirana, "Efficient simulation of OFDM and OSTBC-OFDM systems over multipath rayleigh fading channels," in *2011 Australian Communications Theory Workshop*, pp. 136–141, 2011. doi: 10.1109/AUSCTW.2011.5728751.

[21] A. B. Owen, Y. Maximov, and M. Chertkov, "Importance sampling the union of rare events with an application to power systems analysis," *Electronic Journal of Statistics*, vol. 13, no. 1, pp. 231–254, 2019. doi: 10.48550/arXiv.1710.06965.

[22] L. D. Hoffmann, G. L. Bradley, and K. H. Rosen, *Calculus for business, economics, and the social and life sciences*. McGraw-Hill, 1989.

[23] C. Fox, *An introduction to the calculus of variations*. Courier Corporation, 1987.

[24] G. Casella, C. P. Robert, and M. T. Wells, "Generalized accept-reject sampling schemes," *Lecture Notes-Monograph Series*, vol. 45, pp. 342–347, 2004.

[25] Bishop, Christopher M. Pattern Recognition and Machine Learning. New York :Springer, 2006.

[26] M. Loeve, *Probability Theory: Third Edition*. Dover Books on Mathematics, Dover Publications, 2017.

[27] D. W. Lozier, "Nist digital library of mathematical functions," *Annals of Mathematics and Artificial Intelligence*, vol. 38, no. 1, pp. 105–119, 2003.

[28] *Evolved Universal Terrestrial Radio Access (E-UTRA); Base Station (BS) radio transmission and reception*, 3GPP TS 36.104 v16.5.0, 2020.

Appendix A

Algorithm to Generate Samples from a Truncated Gaussian Distribution

Here we will represent a complex number c by its real-valued vector $\boldsymbol{c} = [\Re\{c\}, \mathrm{Im}\{c\}]^T$ of real and imaginary part, respectively. Let $\boldsymbol{\Sigma}$ denote a non-singular covariance matrix, $\boldsymbol{I}$ the identity matrix and $\mathbf{0}$ the all-zero vector. Furthermore, $\boldsymbol{A}^T$ denotes the transpose of matrix $\boldsymbol{A}$.

A. Algorithm of normalizing $\mathcal{N}(\boldsymbol{\mu}, \boldsymbol{\Sigma})$ to a standard Gaussian $\mathcal{N}(\mathbf{0}, \mathbf{I})$.

The half-space where $\boldsymbol{v}$ belongs can be described as $\boldsymbol{\gamma}_k^T \boldsymbol{v} \geq \beta_k$. Since $\boldsymbol{\Sigma}$ is a non-singular covariance matrix, its square root $\sqrt{\boldsymbol{\Sigma}}$ exists. With $\boldsymbol{v}_{norm} \sim \mathcal{N}(\mathbf{0}, \mathbf{I})$, the equation for the half-space $\boldsymbol{\gamma}_k^T \boldsymbol{v} \geq \beta_k$ becomes [21]

$$\boldsymbol{\omega}_k^T \boldsymbol{v}_{norm} \geq \tau_k, \quad \boldsymbol{\omega}_k = \frac{\boldsymbol{\gamma}_k^T \boldsymbol{\Sigma}^{\frac{1}{2}}}{\sqrt{\boldsymbol{\gamma}_k^T \boldsymbol{\Sigma} \boldsymbol{\gamma}_k}}, \quad \tau_k = \frac{\beta_k - \boldsymbol{\gamma}_k^T \boldsymbol{\mu}}{\sqrt{\boldsymbol{\gamma}_k^T \boldsymbol{\Sigma} \boldsymbol{\gamma}_k}}. \tag{A.1}$$

B. Algorithm to sample from a truncated Gaussian $\mathcal{N}(\mathbf{0}, \mathbf{I})$

Assume a truncated Gaussian $\boldsymbol{v}_{norm} \in \mathcal{N}(\mathbf{0}, \mathbf{I})$ in the half space described by $\boldsymbol{v}_{norm}^T \boldsymbol{\omega} \geq \tau_k$. The sample $\boldsymbol{v}_{norm}$ is first obtained from the complementary half-space (i.e., $\boldsymbol{v}_{norm}^T \boldsymbol{\omega} < \tau_k$), then transformed to $-\boldsymbol{v}_{norm}$ for numerical stability. The procedure for doing this is described in [21]

1). Simulate a standard Gaussian distribution $\mathbf{z} \sim \mathcal{N}(\mathbf{0}, \mathbf{I})$.

2). Simulate a uniform distribution $u \sim \mathcal{U}(0, 1)$.

3). Let $y = \Phi^{-1}(u\Phi(-\tau))$, where $\Phi(\tau)$, is the cumulative distribution function for the standard Gaussian distribution.

4). Let $\boldsymbol{v}_{norm} = \boldsymbol{\omega} y + \left(\mathbf{I} - \boldsymbol{\omega}\boldsymbol{\omega}^{\boldsymbol{T}}\right)\mathbf{z}$.

5). Output $\boldsymbol{v} = \boldsymbol{\mu} - \boldsymbol{\Sigma}^{\frac{1}{2}} \boldsymbol{v}_{norm}$.

Appendix B

Derivation of Optimal Scalar Fading Channel PDF

In order to denote different QAM constellation points, we will break the QAM constellation set S into 3 sub-sets as $S = S_{in} \cup S_{ed} \cup S_{co}$, where S_{in}, S_{ed} and S_{co} are the sets of inner, edge and corner points, respectively. The value of $w(h,x)$ in (4.5) for these points equals

$$w\left(h, x \in S_{co}\right) = \begin{cases} \frac{2f_H(h)T(h)}{g_H(h)}; & 1 - T(h) \\ \frac{2f_H(h)T(h)}{2g_H(h)}; & T(h) \end{cases} \tag{B.1}$$

$$w\left(h, x \in S_{ed}\right) = \begin{cases} \frac{3f_H(h)T(h)}{g_H(h)}; & 1 - \frac{4}{3}T(h) \\ \frac{3f_H(h)T(h)}{2g_H(h)}; & \frac{4}{3}T(h) \end{cases} \tag{B.2}$$

$$w\left(h, x \in S_{in}\right) = \begin{cases} \frac{4f_H(h)T(h)}{g_H(h)}; & 1 - 2T(h) \\ \frac{4f_H(h)T(h)}{2g_H(h)}; & 2T(h) \end{cases} \tag{B.3}$$

where $T(h)$ is given as

$$T(h) = \mathrm{Q}\left(\frac{A}{\frac{\sigma_z}{|h|}}\right),$$

where Q(.) represents the Q-function, $A = \frac{d_{min}}{2}$ with d_{min} being the minimum Euclidean distance between two symbols in the QAM constellation and σ_z is the standard deviation of the AWGN. The variance of $w(h,x)$ equals

$$\begin{aligned} \mathrm{Var}(w(h,x)) &= E_{g_H(h),x}\left[\left(w(h,x) - E_{g_H(h),x}[w(h,x)]\right)^2\right] \\ &= E_{g_H(h),x}\left[w(h,x)^2\right] - E^2_{g_H(h),x}\left[w(h,x)\right] \\ &= E_{g_H(h),x}\left[w(h,x)^2\right] - P_e^2, \end{aligned} \tag{B.4}$$

where P_e is the sought probability of error and Var(.) denotes variance. $E_{g_H(h),x}\left[w(h,x)^2\right]$ in (B.4) can be further expressed with (B.1) - (B.3) as

$$\begin{aligned} E_{g_H(h),x}\left[w(h,x)^2\right] =& P(x \in S_{co}) E_{g_H(h)}\left[w(h, x \in S_{co})^2\right] \\ &+ P(x \in S_{ed}) E_{g_H(h)}\left[w(h, x \in S_{ed})^2\right] \\ &+ P(x \in S_{in}) E_{g_H(h)}\left[w(h, x \in S_{in})^2\right]. \end{aligned} \tag{B.5}$$

It holds that

$$\begin{aligned} E_{g_H(h)}\left[w(h, x \in S_{co})^2\right] =& 4 \int_h \left(\frac{f_H(h)\, T(h)}{g_H(h)}\right)^2 g_H(h)\,(1 - T(h))\,\mathrm{d}h \\ &+ 4 \int_h \frac{1}{4} \left(\frac{f_H(h)\, T(h)}{g_H(h)}\right)^2 g_H(h)\, T(h)\ \mathrm{d}h \\ =& 4 \int_h \frac{\left(f_H(h)\, T(h)\right)^2}{g_H(h)} \left(1 - \frac{3}{4} T(h)\right) \mathrm{d}h. \end{aligned} \tag{B.6}$$

Similarly,

$$E_{g_H(h)}\left[w(h, x \in S_{ed})^2\right] = 9 \int_h \frac{\left(f_H(h)\, T(h)\right)^2}{g_H(h)} (1 - T(h)) \mathrm{d}h. \tag{B.7}$$

$$E_{g_H(h)}\left[w(h, x \in S_{in})^2\right] = 16 \int_h \frac{\left(f_H(h)\, T(h)\right)^2}{g_H(h)} (1 - 1.5T(h)) \mathrm{d}h. \tag{B.8}$$

Using the fact that $P(x \in S_{co}) = \frac{4}{M}$, $P(x \in S_{ed}) = \frac{4\left(\sqrt{M}-2\right)}{M}$, $P(x \in S_{in}) = (\sqrt{M} - 2)^2/M$ we get that (B.4) becomes

$$\mathrm{Var}(w(h, x)) = \int_h \frac{\left(f_H(h)\, T(h)\right)^2}{g_H(h)} (1 - \eta T(h)) \rho \mathrm{d}h - {P_e}^2, \tag{B.9}$$

where $\eta = \frac{6a^2+9a+3}{4a^2+9a+4}$, $\rho = \frac{16a^2+36a+16}{M}$ and $a = \sqrt{M} - 2$. The integral in (B.9) can be represented as a functional [23]

$$R[g_H(h)] = \int_h L\left(h, g_H(h)\right) \mathrm{d}h, \tag{B.10}$$

where $L\left(h, g_H(h)\right)$ is the integrand in (B.9). To minimize the variance in (B.9) with respect to $g_H(h)$, we will utilize the Lagrange multiplier method [22] with the constraint $I(h, g_H(h)) = \int_h g_H(h) \mathrm{d}h = 1$, which expresses the fact that $g_H(h)$ should be a pdf. Since P_e is a constant, the Lagrangian function for (B.9) can be expressed as

$$\mathscr{L}(g_H(h), \lambda) = R[g_H(h)] + \lambda I(h, g_H(h)), \tag{B.11}$$

where $\mathscr{L}$ is the Lagrangian function of $g_H(h)$ and a scalar Lagrange multiplier λ. Applying the Euler-Lagrange equation [23] to (B.11), it follows that the following equality must be satisfied

$$\frac{\partial[L\left(h, g_H(h)\right) + \lambda g_H(h)]}{\partial g_H(h)} = 0. \tag{B.12}$$

Further manipulation of equation (B.12) leads to

$$\begin{aligned} \frac{\partial L\left(h, g_H(h)\right)}{\partial g_H(h)} &= -\lambda \frac{\partial g_H(h)}{\partial g_H(h)} \\ \frac{\partial L\left(h, g_H(h)\right)}{\partial g_H(h)} &= \frac{\left(f_H(h)\, T(h)\right)^2}{g_H^2(h)} (1 - \eta T(h))\, \rho = \lambda. \end{aligned} \tag{B.13}$$

From (B.13) we can express $g_H(h)$ as

$$\begin{aligned} g_H(h) &= \sqrt{\frac{\left(f_H(h)\,T(h)\right)^2}{\lambda}\left(1-\eta T(h)\right)\rho} \\ &= \frac{f_H(h)\,T(h)\sqrt{\left(1-\eta T(h)\right)\rho}}{c}, \end{aligned} \tag{B.14}$$

where $\sqrt{\lambda}$ is replaced with c for ease of representation. Using the constraint $I(h, g_H(h))$, we obtain that

$$c = \int_h f_H(h)\,T(h)\sqrt{\left(1-\eta T(h)\right)\rho}\mathrm{d}h. \tag{B.15}$$

Inserting (B.15) in (B.14), we finally obtain that the optimal $g_H(h)$ equals

$$g_H(h) = \frac{f_H(h)\,T(h)\sqrt{\left(1-\eta T(h)\right)\rho}}{\int_h f_H(h)\,T(h)\sqrt{\left(1-\eta T(h)\right)\rho}\mathrm{d}h}. \tag{B.16}$$

Appendix C

Derivation of the Covariance

In order to denote different QAM constellation points, we will break the QAM constellation set S into three sub-sets as $S = S_{in} \cup S_{ed} \cup S_{co}$, where S_{in}, S_{ed} and S_{co} are the sets of inner, edge and corner points, respectively. $\mathrm{Cov}(w_i(\mathbf{h}, x_i), w_j(\mathbf{h}, x_j))$ in (5.4) is obtained by evaluating $E_{g_{\mathbf{H}}(\mathbf{h}),x_i,x_j}\left[w_i\left(\mathbf{h}, x_i\right) w_j\left(\mathbf{h}, x_j\right)\right]$ for each of the 9 combinations of S across subcarriers i and j by utilizing the functions in (B.1) - (B.3) as

$$\begin{aligned} & E_{g_{\mathbf{H}}(\mathbf{h}),x_i,x_j}\left[w_i\left(\mathbf{h}, x_i\right) w_j\left(\mathbf{h}, x_j\right)\right] \\ = & \sum_{a=in,ed,co} \sum_{b=in,ed,co} P(x_i \in S_a) P(x_j \in S_b) E_{g_{\mathbf{H}}(\mathbf{h})}[w_i(\mathbf{h}, x_i \in S_a) w_j(\mathbf{h}, x_j \in S_b)], \end{aligned} \tag{C.1}$$

Using $P(x_{i,j} \in S_{in}) = (\sqrt{M}-2)^2/M$, we get

$$\begin{aligned} & P(x_i \in S_{in}) P(x_j \in S_{in}) E_{g_{\mathbf{H}}(\mathbf{h})}[w_i(\mathbf{h}, x_i \in S_{in}) w_j(\mathbf{h}, x_j \in S_{in})] \\ = & \frac{(\sqrt{M}-2)^4}{M^2} \int_{\mathbf{h}} \frac{16 f_{\mathbf{H}}^2\left(\mathbf{h}\right) T_i\left(\mathbf{h}\right) T_j\left(\mathbf{h}\right)}{g_{\mathbf{H}}\left(\mathbf{h}\right)} \Big[(1-2T_i\left(\mathbf{h}\right))(1-2T_j\left(\mathbf{h}\right)) \\ & + (1-2T_i\left(\mathbf{h}\right)) T_j\left(\mathbf{h}\right) + T_i\left(\mathbf{h}\right) T_j\left(\mathbf{h}\right) + T_i\left(\mathbf{h}\right)(1-2T_j\left(\mathbf{h}\right))\Big] \mathrm{d}\mathbf{h} \\ = & \frac{(\sqrt{M}-2)^4}{M^2} \int_{\mathbf{h}} \frac{16 f_{\mathbf{H}}^2\left(\mathbf{h}\right) T_i\left(\mathbf{h}\right) T_j\left(\mathbf{h}\right)}{g_{\mathbf{H}}\left(\mathbf{h}\right)} \Big[(1-T_i\left(\mathbf{h}\right))(1-T_j\left(\mathbf{h}\right))\Big] \mathrm{d}\mathbf{h}, \end{aligned} \tag{C.2}$$

where $T_{i,j}\left(\mathbf{h}\right)$ is given as

$$T_{i,j}(\mathbf{h}) = \mathrm{Q}\left(\frac{A}{\left(\frac{\sigma_z}{\mathrm{DFT}(\mathbf{h})_i}\right)}\right),$$

where $\mathrm{DFT}\left(\mathbf{h}\right)_{i,j}$ represents the channel frequency response on subcarrier i, j which is obtained from (2.24), Q(.) denotes the Q-function, $A = \frac{d_{min}}{2}$ with d_{min} being the minimum Euclidean distance between two symbols in the signal constellation and σ_z is the standard deviation of the AWGN.

Similarly,

$$\begin{aligned}&P(x_i \in S_{in})P(x_j \in S_{co})E_{g_{\mathbf{H}}(\mathbf{h})}[w_i(\mathbf{h}, x_i \in S_{in})w_j(\mathbf{h}, x_j \in S_{co})]\\ =&4\frac{(\sqrt{M}-2)^2}{M^2}\int_{\mathbf{h}}\frac{8f_{\mathbf{H}}^2(\mathbf{h})T_i(\mathbf{h})T_j(\mathbf{h})}{g_{\mathbf{H}}(\mathbf{h})}\left[(1-T_i(\mathbf{h}))(1-\frac{1}{2}T_j(\mathbf{h}))\right]\mathrm{d}\mathbf{h}. \quad \text{(C.3)}\end{aligned}$$

$$\begin{aligned}&P(x_i \in S_{in})P(x_j \in S_{ed})E_{g_{\mathbf{H}}(\mathbf{h})}[w_i(\mathbf{h}, x_i \in S_{in})w_j(\mathbf{h}, x_j \in S_{ed})]\\ =&\frac{4(\sqrt{M}-2)^3}{M^2}\int_{\mathbf{h}}\frac{12f_{\mathbf{H}}^2(\mathbf{h})T_i(\mathbf{h})T_j(\mathbf{h})}{g_{\mathbf{H}}(\mathbf{h})}\left[(1-T_i(\mathbf{h}))(1-\frac{2}{3}T_j(\mathbf{h}))\right]\mathrm{d}\mathbf{h}. \quad \text{(C.4)}\end{aligned}$$

$$\begin{aligned}&P(x_i \in S_{ed})P(x_j \in S_{in})E_{g_{\mathbf{H}}(\mathbf{h})}[w_i(\mathbf{h}, x_i \in S_{ed})w_j(\mathbf{h}, x_j \in S_{in})]\\ =&\frac{4(\sqrt{M}-2)^3}{M^2}\int_{\mathbf{h}}\frac{12f_{\mathbf{H}}^2(\mathbf{h})T_i(\mathbf{h})T_j(\mathbf{h})}{g_{\mathbf{H}}(\mathbf{h})}\left[(1-\frac{2}{3}T_i(\mathbf{h}))(1-T_j(\mathbf{h}))\right]\mathrm{d}\mathbf{h}, \quad \text{(C.5)}\end{aligned}$$

$$\begin{aligned}&P(x_i \in S_{ed})P(x_j \in S_{co})E_{g_{\mathbf{H}}(\mathbf{h})}[w_i(\mathbf{h}, x_i \in S_{ed})w_j(\mathbf{h}, x_j \in S_{co})]\\ =&\frac{16(\sqrt{M}-2)}{M^2}\int_{\mathbf{h}}\frac{6f_{\mathbf{H}}^2(\mathbf{h})T_i(\mathbf{h})T_j(\mathbf{h})}{g_{\mathbf{H}}(\mathbf{h})}\left[(1-\frac{2}{3}T_i(\mathbf{h}))(1-\frac{1}{2}T_j(\mathbf{h}))\right]\mathrm{d}\mathbf{h}, \quad \text{(C.6)}\end{aligned}$$

$$\begin{aligned}&P(x_i \in S_{ed})P(x_j \in S_{ed})E_{g_{\mathbf{H}}(\mathbf{h})}[w_i(\mathbf{h}, x_i \in S_{ed})w_j(\mathbf{h}, x_j \in S_{ed})]\\ =&\frac{16(\sqrt{M}-2)^2}{M^2}\int_{\mathbf{h}}\frac{9f_{\mathbf{H}}^2(\mathbf{h})T_i(\mathbf{h})T_j(\mathbf{h})}{g_{\mathbf{H}}(\mathbf{h})}\left[(1-\frac{2}{3}T_i(\mathbf{h}))(1-\frac{2}{3}T_j(\mathbf{h}))\right]\mathrm{d}\mathbf{h}. \quad \text{(C.7)}\end{aligned}$$

$$\begin{aligned}&P(x_i \in S_{co})P(x_j \in S_{in})E_{g_{\mathbf{H}}(\mathbf{h})}[w_i(\mathbf{h}, x_i \in S_{co})w_j(\mathbf{h}, x_j \in S_{in})]\\ =&\frac{4(\sqrt{M}-2)^2}{M^2}\int_{\mathbf{h}}\frac{8f_{\mathbf{H}}^2(\mathbf{h})T_i(\mathbf{h})T_j(\mathbf{h})}{g_{\mathbf{H}}(\mathbf{h})}\left[(1-\frac{1}{2}T_i(\mathbf{h}))(1-T_j(\mathbf{h}))\right]\mathrm{d}\mathbf{h}. \quad \text{(C.8)}\end{aligned}$$

$$\begin{aligned}&P(x_i \in S_{co})P(x_j \in S_{ed})E_{g_{\mathbf{H}}(\mathbf{h})}[w_i(\mathbf{h}, x_i \in S_{co})w_j(\mathbf{h}, x_j \in S_{ed})\\ =&\frac{16(\sqrt{M}-2)}{M^2}\int_{\mathbf{h}}\frac{6f_{\mathbf{H}}^2(\mathbf{h})T_i(\mathbf{h})T_j(\mathbf{h})}{g_{\mathbf{H}}(\mathbf{h})}\left[(1-\frac{1}{2}T_i(\mathbf{h}))(1-\frac{2}{3}T_j(\mathbf{h}))\right]\mathrm{d}\mathbf{h}, \quad \text{(C.9)}\end{aligned}$$

$$\begin{aligned}&P(x_i \in S_{co})P(x_j \in S_{co})E_{g_{\mathbf{H}}(\mathbf{h})}[w_i(\mathbf{h}, x_i \in S_{co})w_j(\mathbf{h}, x_j \in S_{co})]\\ =&\frac{16}{M^2}\int_{\mathbf{h}}\frac{4f_H^2(\mathbf{h})T_i(\mathbf{h})T_j(\mathbf{h})}{g_{\mathbf{H}}(\mathbf{h})}\left[(1-\frac{1}{2}T_i(\mathbf{h}))(1-\frac{1}{2}T_j(\mathbf{h}))\right]\mathrm{d}\mathbf{h}, \quad \text{(C.10)}\end{aligned}$$

Inserting (C.2)-(C.10) in (C.1) it follows that

$$\begin{aligned}&E_{g_{\mathbf{H}}(\mathbf{h}),x_i,x_j}\left[w_i(\mathbf{h},x_i)\,w_j(\mathbf{h},x_j)\right]\\ =&\int_{\mathbf{h}}\frac{f_{\mathbf{H}}^2(\mathbf{h})T_i(\mathbf{h})T_j(\mathbf{h})}{g_{\mathbf{H}}(\mathbf{h})}\left[\left(\frac{4a^2+12a+8}{M}-\frac{4a^2+8a+4}{M}T_i(\mathbf{h})\right)\right.\\ &\left.\times\left(\frac{4a^2+12a+8}{M}-\frac{4a^2+8a+4}{M}T_j(\mathbf{h})\right)\right]\mathrm{d}\mathbf{h}\\ =&\int_{\mathbf{h}}\frac{f_{\mathbf{H}}^2(\mathbf{h})}{g_{\mathbf{H}}(\mathbf{h})}t_it_j\mathrm{d}\mathbf{h}, \quad \text{(C.11)}\end{aligned}$$

where $t_i = T_i(\mathbf{h})\left(\frac{4a^2+12a+8}{M} - \frac{4a^2+8a+4}{M}T_i(\mathbf{h})\right)$,
$t_j = T_j(\mathbf{h})\left(\frac{4a^2+12a+8}{M} - \frac{4a^2+8a+4}{M}T_j(\mathbf{h})\right)$ and $a = \sqrt{M} - 2$.

Inserting (C.11) into (5.4) we get

$$2\sum_{1\le i<j\le N} Cov\left(w_i(\mathbf{h}, x_i)\, w_j(\mathbf{h}, x_j)\right) = 2\sum_{1\le i<j\le N}\int_{\mathbf{h}} \frac{f^2_{\mathbf{H}}(\mathbf{h})}{g_{\mathbf{H}}(\mathbf{h})} t_i t_j \mathrm{d}\mathbf{h} - 2\sum_{1\le i<j\le N} P_{e_i} P_{e_j}, \tag{C.12}$$

Using the multinomial theorem [27], (C.12) becomes

$$2\sum_{1\le i<j\le N} Cov\left(w_i(\mathbf{h}, x_i)\, w_j(\mathbf{h}, x_j)\right) = \int_{\mathbf{h}} \frac{f^2_{\mathbf{H}}(\mathbf{h})}{g_{\mathbf{H}}(\mathbf{h})} \left[\left(\sum_{i=1}^{N} t_i\right)^2 - \sum_{i=1}^{N} t_i^2\right] \mathrm{d}\mathbf{h} - 2\sum_{1\le i<j\le N} P_{e_i} P_{e_j}. \tag{C.13}$$

Printed by Tryckeriet i E-huset, Lund 2022

LUND
UNIVERSITY

Series of Master's theses
Department of Electrical and Information Technology
LU/LTH-EIT 2022-870
http://www.eit.lth.se