Distributed Detection for Handoff Macrodiversity in Cellular Communication Systems

Frederic Martin

Department of Electrical & Computer Engineering
McGill University
Montreal, Canada

January 2005

A thesis submitted to McGill University in partial fulfillment of the requirements for the degree of Master of Engineering.

Your file Votre référence
ISBN: 0-494-12631-0
Our file Notre référence
ISBN: 0-494-12631-0

Abstract

This thesis considers the application of the principles of distributed signal detection to the uplink of a mobile communication unit engaged in soft handoff, when all base stations involved are equipped with multiple receiving antennas. The system consists of a local detector at each base station and a fusion center at the Mobile Switching Center (MSC). Optimum decision rules are derived for systems without channel coding, as well as for systems using channel coding, over a quasi-static spatially uncorrelated Rayleigh fading channel. Two different cases are considered. In the first case, accurate estimates of the base station channel states are available at the MSC, while in the second case only the statistics of the channels are known. For both cases, when the system is not using channel coding the optimum local decision rules are likelihood ratio quantizers for which the defining thresholds are optimized numerically with respect to the probability of bit error at the output of the MSC. With channel coding it is shown that the complexity of either the implementation or the optimization of the optimum decision rules increases exponentially with the frame size. Hence, for coded systems, sub-optimum alternatives are proposed where the local decision rules are likelihood ratio quantizers. The performances of these systems are investigated. For the uncoded systems the probability of bit error is evaluated numerically, and for coded systems the probability of bit error and frame error are estimated through computer simulations. Finally, it is demonstrated that by carefully selecting the thresholds defining the local decision rules, 8 quantization levels are sufficient to make the performances almost identical to the performances of an optimum centralized system, implementing at the MSC a maximum likelihood test using the actual signals received at the involved base stations.

Sommaire

Le sujet de cette thèse concerne l'application des principes de détection décentralisée au lien montant d'une unité de communication mobile engagée dans un transfert intercellulaire doux ("soft handoff"), lorsque les stations de base sont équipées de plusieurs antennes de réception. Le système est formé d'un détecteur local à chaque station de base et d'un centre de fusion au Centre de Commutation Mobile (CCM). Les règles optimales de décision sont dérivées pour les systèmes sans codage de canal de même que pour les systèmes qui utilisent le codage de canal pour des canaux à évanouissement quasi-statiques de type Rayleigh, non corrélés spatialement. Deux cas différents sont considérès. Dans le premier cas, il est présumé que l'atténuation sur les différents canaux est connue au CCM tandis que, dans le deuxième cas, seulement les statistiques de ces canaux sont connues. Dans ces deux cas, lorsque le système n'utilise pas de codage de canal, les règles locales optimales de décision sont basées sur une quantification des rapports de vraisemblance. Les différents seuils correspondant à cette quantification sont ajustés de façon à minimiser la probabilité d'erreur à la sortie du CCM. Lorsque le système utilise le codage de canal, il est démontré que, dans les deux cas, la complexité des règles de décision augmente exponentiellement avec la longueur des trames d'information. Par conséquent, des alternatives sous optimales qui effectuent, à chaque station de base, une quantification des rapports de vraisemblance sont donc proposées pour les systèmes codés. Les performances de ces systèmes sont étudiées. La probabilité d'erreur par bit des systèmes sans aucun codage de canal est évaluée numériquement tandis que la probabilité d'erreur par bit et par trame des systèmes qui utilisent le codage de canal est estimée en utilisant des simulations par ordinateur. Finalement, il est démontré que lorsque les seuils sont choisis avec attention, 8 niveaux de quantification sont suffisants pour obtenir des performances presque identiques aux performances obtenues avec un système optimal de détection centralisée qui applique au CCM une méthode de maximum de vraisemblance pour les signaux reçus aux stations de base impliquées.

Acknowledgement

I would like to thank a number of people who have made significant contributions to the development of this thesis. First of all, I will like to thank my supervisor, Professor Harry Leib, to whom I am greatly indebted for his guidance and encouragement. In addition, I would also like to thank Professor Leib for providing me financial support throughout the course of my graduate study. I would like to thank all the previous and current members in our research group for providing fruitful cooperations and a pleasant atmosphere. Finally, I would like to thank my family and friends whose support and encouragements have contributed to make this work possible.

Contents

List of Figures

List of Tables

List of abbreviations

ASA	Adaptive Simulated Annealing
BER	Bit Error Rate
BPSK	Binary Phase Shift Keying
CDMA	Code Division Multiple Access
CHM	Conventional Handoff Macrodiversity
CSHDD	Coded Soft Handoff Distributed Detection
CSI	Channel State Information
DPSK	Differential Phase Shift Keying
FA	Fast Annealing
FER	Frame Error Rate
FSK	Frequency Shift Keying
LLR	Log-Likelihood Ratio
MMSE	Minimum Mean Square Error
MPEP	Minimum Pairwise Error Probability
MSC	Mobile Switching Center
MSE	Mean Square Error
OC	Optimum Centralized
PEP	Pairwise Error Probability
SA	Simulated Annealing
SHDD	Soft Handoff Distributed Detection
SNR	Signal-to-Noise Ratio

Notational convention

$(\cdot)^T$	Transpose of the argument
$(\cdot)^*$	Conjugate of the argument
$\Re\{\cdot\}$	Real part of the argument
$\Im\{\cdot\}$	Imaginary part of the argument
$\lvert\cdot\rvert$	Magnitude of the argument
$\lfloor\cdot\rfloor$	Largest integer $\leq$ than the argument
$sgn(\cdot)$	Sign of the argument
$\mathbf{a}$	Vector $\mathbf{a}$
X	Random variable X
x	Realization of the random variable X
$P(X = x)$	Probability that $X = x$
$P(X = x \mid Y)$	Probability that $X = x$ given the random variable Y
$P(X = x \mid Y = y)$	Probability that $X = x$ given $Y = y$
$P(X = x \mid y)$	Simpler notation for $P(X = x \mid Y = y)$
$f_X(x)$	PDF of the random variable X
$f_X(x \mid Y)$	PDF of the random variable X given the random variable Y
$f_X(x \mid Y = y)$	PDF of the random variable X given $Y = y$
$f_X(x \mid y)$	Simpler notation for $f_X(x \mid Y = y)$
$F_X(x)$	CDF of the random variable X
$F_X(x \mid Y)$	CDF of the random variable X given the random variable Y
$F_X(x \mid Y = y)$	CDF of the random variable X given $Y = y$
$F_X(x \mid y)$	Simpler notation for $F_X(x \mid Y = y)$
$E_{X,Y}[g(X, Y)]$	$= \int_x \int_y g(x, y) f_{X,Y}(x, y) dx dy$, where $f_{X,Y}(x, y)$ is the joint PDF of the random variables X and Y
$E_{X\mid Y}[g(X, Y)]$	$= \int_x g(x, Y) f_X(x \mid Y) dx$, where $f_X(x \mid Y)$ is the PDF of the random variable X given the random variable Y

Chapter 1

Introduction

Third generation wireless systems are based on Code Division Multiple Access (CDMA) techniques allowing the implementation of soft handoff at the cell boundaries. During soft handoff, the uplink consists of a mobile unit communicating simultaneously with multiple base stations that are often separated by a few kilometers. In current systems, all base stations transmit through wireline their respective decoded data frame to the mobile switching center (MSC) where the best frame is selected based on a reliability criterion[1]. This selection diversity technique reduces the impairments due to shadowing near the cell boundaries. On the other hand, selection diversity does not take advantage of the information contained in the signals received by base stations that are not selected. A better technique would be to perform maximum ratio combining [2] of the actual signals received by the different base stations and then to channel decode the output of the combiner, providing protection against multipath fading as well as shadowing. The feasibility of such a system, however, is questionable due to the required bandwidth between the base stations and the MSC where combining is performed. This work proposes a bandwidth efficient solution which is a compromise between these two alternatives and makes use of the principles of distributed signal detection.

Distributed signal detection is a generalization of classical detection theory to systems where the observations are first processed at distributed sensors before sending the outcome of the processing to a fusion center where a final decision is made. A good introduction to distributed detection can be found in [3], where an analysis employing a Bayesian framework is included. For the past several years, distributed detection has received increasing interest in many application areas, including diversity combining for wireless communication. The first of such studies is reported in [4] where the optimum hard decision combiner is derived and the performances of the combiner are evaluated for systems using binary non-coherent Frequency Shift Keying (FSK) modulation over a Rayleigh fading channel. The performances of the optimum hard decision combiner have been also investigated for other system configurations and channel models. For example in [5] and [6], the performances of the hard decision combiner are analyzed for DS-CDMA in a shadowed Rician fading land mobile satellite channel using Binary Phase Shift Keying (BPSK) and Differential Phase Shift Keying (DPSK) modulation respectively. In [7] [8], [9] and [10], an adaptive implementation of the optimum hard decision combiner is derived and the performances of the combiner are evaluated for a three base stations macroscopic diversity scheme. These studies show that the application of distributed detection with hard decisions at the local detectors can be used to reduce the probability of bit error although the performance is still far from the performance of the optimum centralized detection scheme.

In order to close the gap between the performance of the optimum centralized detection scheme and the performance of the distributed detection scheme, the local detectors must provide soft decisions to the fusion center. The first work that studied distributed detection with soft decisions at the local detectors in the context of wireless communication is reported in [11]. In this paper, the optimum design of a soft decision distributed detection system is studied for diversity reception with non-coherent FSK over a Rayleigh fading channel. It is shown that soft decision distributed detection with a few bits of resolution provides performances close to the optimum centralized detection scheme. Then, in [12] and [13], the application of distributed detection with soft decisions at the local detectors is considered for the detection of data from multiple users using BPSK modulation in the presence of interference and additive noise. The optimum fusion rule and optimum local detector decision rules are derived for a jointly optimum decision criterion and an individually optimum decision criterion. In [14] and [15], the subjects of distributed detection and distributed multiuser detection are revisited and more practical sub-optimum schemes are presented where Minimum Mean Square Error (MMSE) Log-Likelihood Ratio (LLR) quantizers are used by the local detectors.

The problem of distributed detection with channel coding is addressed in [16] for the uplink when mobiles using BPSK modulation are in soft handoff with three base stations. The paper proposes different approaches based on hard decision combining prior to de-interleaving/channel decoding as well as a hybrid combining technique. In the hybrid combining technique, the combining unit is located at the base station with the maximum average carrier power allowing direct observations from this base station, as well as hard decisions from the remote base stations, to be used in the combining procedure.

Motivated by these results, we study in this work the application of soft decision distributed detection to handoff macrodiversity in cellular communication systems when the channel fading is spatially uncorrelated, quasi-static and Rayleigh distributed. We first consider uncoded communication systems using BPSK modulation and generalize results from [11] for handoff macrodiversity where each base station (local detector) is equipped with multiple antennas. A global optimization algorithm is proposed to optimize the thresholds defining the likelihood ratio quantizers used at the local detectors, as opposed to the local optimization algorithms proposed in [11][12][13][14][15]. Furthermore, we consider coded communication systems using BPSK modulation, where as opposed to [16] we assume soft decisions are made at the local detectors and the channel fading may or may not be known at the MSC. In addition, we generalize the MMSE-LLR quantizer proposed in [14] and [15] to coded communication systems using multiple receiving antennas at each local detector.

This thesis is structured as follow. In chapter II, the optimum distributed detection scheme for handoff macrodiversity is derived assuming an uncoded communication system using BPSK modulation, where each base station is equipped with multiple receiving antennas. We refer to this scheme as the Soft Handoff Distributed Detection (SHDD) scheme. Two cases are considered. In the first case, the channel state is assumed to change slowly enough such that estimates[1] with infinite precision can be transmitted by the base stations to the MSC while, in the second case, the channel state information (CSI) is not available at the MSC. For almost all system configurations that are considered, the results show that the optimum local detector decision rule is a likelihood ratio quantizer, which is defined by a set of thresholds. These thresholds must be adjusted in order to minimize the probability of error at the output of the MSC, which is a nonlinear non-convex function of these thresh-

[1]The CSI estimates are assumed to be accurate

olds. Since local optimization techniques only provide locally optimum solutions, we propose to use a global optimization technique called Adaptive Simulated Annealing (ASA) [17] to perform the required optimization. The performances of the SHDD schemes are investigated in term of Bit Error Rate (BER) and compared to the performances of the Optimum Centralized (OC) detector and Conventional Handoff Macrodiversity (CHM). Analytical expressions are derived in Appendix B for the BER of the SHDD schemes allowing the BER to be evaluated numerically. In chapter III, channel coding is included in the design of the SHDD scheme. The optimum distributed detection scheme for handoff macrodiversity is derived assuming a coded communication system using BPSK modulation, where each base station is equipped with multiple receiving antennas. Since the complexity of the optimum decision rules grows exponentially with the frame size, in this case sub-optimum alternatives are proposed where the local decision rules are LLR quantizers. The performances of the proposed sub-optimum Coded Soft Handoff Distributed Detection (CSHDD) schemes are investigated in term of BER and Frame Error Rate (FER), and compared to the performances of the OC decoder and CHM. As opposed to the uncoded case, a computer simulator was constructed in order to estimate the BER and FER. Finally, chapter IV presents concluding remarks.

It is important to mention that a compact disk is included with this thesis and contains all software necessary to reproduce the results presented in Chapter 2 and Chapter 3. More precisely, the compact disk contains the software used to evaluate numerically the BER of the SHDD schemes as well as the software used to optimize the LLR quantizer thresholds of the CSHDD schemes, where both softwares are written in Matlab programming language. Furthermore, the compact disk also contains the software simulator, written in C programming language, used to estimate the performances of the CSHDD and reference schemes. In addition, the compact disk also includes a software manual discussing the implementation of the different softwares and providing instructions for the utilization of these softwares.

In the development and performance evaluation of the designed handoff macrodiversity schemes based on distributed detection, the following novel and original contributions were made

1. Application of the principles of soft decision distributed detection to handoff macrodiversity for uncoded communication systems using BPSK modulation,

where each base station is equipped with multiple receiving antennas and the channel is a spatially uncorrelated quasi-static Rayleigh fading channel.

2. Extension of the principles of soft decision distributed detection in order to take channel coding into account in the design of the handoff macrodiversity schemes.

3. The utilization of a global optimization technique called Adaptive Simulated Annealing in the optimization of the designed distributed detection schemes, as opposed to local optimization techniques usually used.

4. Evaluation of the performances of the designed handoff macrodiversity schemes for various system configurations.

Chapter 2

SHDD scheme for uncoded communication systems

In this chapter we study the application of distributed detection, with soft decisions at the local detectors, to the uplink when a mobile unit is in soft handoff. In section 2.1, the SHDD scheme is presented. In section 2.2, the optimum SHDD scheme is derived for BPSK modulation for the case when the channel state is known at the MSC and also when it is not known. In section 2.3, the optimization of the local detector decision rules is considered using the ASA global optimization technique. Finally, in section 2.4, the performances of the designed SHDD schemes are evaluated numerically for a quasi-static spatially-uncorrelated Rayleigh fading channel.

2.1 System model

We consider the uplink of a mobile unit in soft handoff with N_{BS} base stations, each equipped with N_R antennas, as illustrated in Fig. 2.1. At the mobile unit, prior to transmission the information bit B is sent to a symbol mapper to generate the BPSK symbol $S \in \{-1, 1\}$. The symbol is then transmitted to the N_{BS} base stations involved in the handoff process. At the receiving end, all base stations make individually a soft decision on the transmitted bit. For instance, the kth base station makes a soft decision $U_k \in \{0, \dots, L-1\}$ on the transmitted bit B based on the received signal vector $\mathbf{R}_k = [R_{k,1}, \dots, R_{k,N_R}]^T$. The decisions contained in the local decision vector $\mathbf{U} = [U_1, \dots, U_{N_{BS}}]^T$ are sent from the involved base stations to the MSC where a

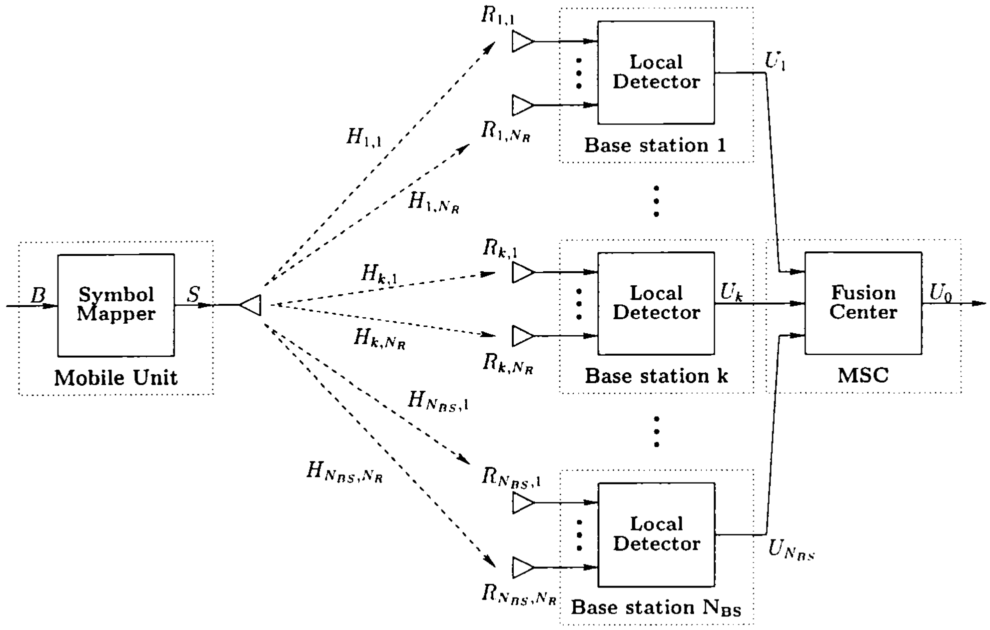


Fig. 2.1 Uncoded Soft Handoff Distributed Detection system model

final decision U_0 is made on the transmitted information bit B.

The signal received at the nth antenna of the kth base station is modeled as follows

$$R_{k,n} = H_{k,n}\sqrt{E_k}S + N_{k,n}. \tag{2.1}$$

The parameters $N_{k,n}$ model white Gaussian noise as independent zero mean circular complex Gaussian random variables with variance $N_0/2$ per real and imaginary component. The parameters $H_{k,n}$ model spatially-uncorrelated Rayleigh fading as independent zero mean circular complex Gaussian random variables with variance 0.5 per real and imaginary component. The parameters E_k model the average received energy per antenna at the different base stations and are dependent on the position of the mobile unit in the cellular network as well as power control. It is assumed that each base station provides to the MSC an accurate estimate of the average signal-to-noise ratio (SNR) received at each individual antenna, which is defined for the kth base station as $SNR_k = \frac{E_k}{N_0}$. On the other hand, the channel state vector $\mathbf{H} = [\mathbf{H}_1, \ldots, \mathbf{H}_{N_{BS}}]^T$, where $\mathbf{H}_k = [H_{k,1}, \ldots, H_{k,N_R}]$, may or may not be available at the MSC although $\mathbf{H}_k$ is perfectly known at the kth base station.

Since almost all third generation cellular systems employ CDMA, it is important to mention that the SHDD scheme can be applied to CDMA. In fact, considering the received signal in a CDMA system is modeled as follows

$$\mathbf{R}_{k,n} = H_{k,n}\sqrt{E_k}S\mathbf{a} + \mathbf{N}_{k,n} \tag{2.2}$$

where $\mathbf{a}$ is the spreading code, the SHDD scheme can be applied as presented in this thesis by assuming the kth base station local detector observes the output of a correlator to $\mathbf{R}_{k,n}$, which can be modeled by expression (2.1).

2.2 Optimum Distributed Detection

In this section, we consider the optimum SHDD scheme generalizing the results from [11] for the system presented in the previous section. As presented in the previous section, the SHDD scheme consists of N_{BS} local detectors (base stations) and one fusion center (MSC) where the final decision is made. Hence, the objective in optimizing this type of scheme is to obtain the set of local decision rules used at the different base stations, denoted by γ_k $k = 1, \ldots, N_{BS}$, and the fusion rule used at the MSC, denoted by γ_0, that jointly minimize the optimality criterion. The considered optimality criterion is the average probability of bit error at the output of the MSC which can be defined as follows

$$P_b = \int_{\mathbf{h}} P_{b|\mathbf{h}} f_{\mathbf{H}}(\mathbf{h}) d\mathbf{h} \tag{2.3}$$

where $P_{b|\mathbf{h}}$ is the probability of bit error given the channel state vector $\mathbf{H} = \mathbf{h}$ and $f_{\mathbf{H}}(\mathbf{h})$ is the Probability Density Function (PDF) of $\mathbf{H}$. Let $P(U_0 = u_0 \mid \mathbf{h}, \mathbf{u})$ denote the probability that the final decision U_0 equals u_0 given the local decision vector $\mathbf{U} = \mathbf{u}$ and the channel state vector $\mathbf{H} = \mathbf{h}$, $P(\mathbf{U} = \mathbf{u} \mid \mathbf{h}, B = b)$ denote the probability that the local decision vector $\mathbf{U}$ equals $\mathbf{u}$ given the bit $B = b$ and the channel state vector $\mathbf{H} = \mathbf{h}$. The conditional probability of bit error $P_{b|\mathbf{h}}$ equals

$$\begin{aligned} P_{b|\mathbf{h}} &= \frac{1}{2}\sum_{\mathbf{u}} P(U_0 = 1 \mid \mathbf{h}, \mathbf{u}) P(\mathbf{U} = \mathbf{u} \mid \mathbf{h}, B = 0) \\ &\quad + \frac{1}{2}\sum_{\mathbf{u}} P(U_0 = 0 \mid \mathbf{h}, \mathbf{u}) P(\mathbf{U} = \mathbf{u} \mid \mathbf{h}, B = 1). \end{aligned} \tag{2.4}$$

Furthermore, since $P(U_0 = 0 \mid \mathbf{h}, \mathbf{u}) = 1 - P(U_0 = 1 \mid \mathbf{h}, \mathbf{u})$, the conditional probability of bit error (2.4) can be reformulated as follows

$$\begin{aligned} P_{b|\mathbf{h}} &= \frac{1}{2}\sum_{\mathbf{u}} P(U_0 = 1 \mid \mathbf{h}, \mathbf{u}) P(\mathbf{U} = \mathbf{u} \mid \mathbf{h}, B = 0) \\ &\qquad + \frac{1}{2} - \frac{1}{2}\sum_{\mathbf{u}} P(U_0 = 1 \mid \mathbf{h}, \mathbf{u}) P(\mathbf{U} = \mathbf{u} \mid \mathbf{h}, B = 1) \\ &= \frac{1}{2} + \frac{1}{2}\sum_{\mathbf{u}} P(U_0 = 1 \mid \mathbf{h}, \mathbf{u}) \Big[P(\mathbf{U} = \mathbf{u} \mid \mathbf{h}, B = 0) - P(\mathbf{U} = \mathbf{u} \mid \mathbf{h}, B = 1) \Big] \end{aligned} \tag{2.5}$$

which is a more appropriate form for the optimization of the decision rules.

It is important to mention that, since the decision rules $\gamma = \{\gamma_0, \gamma_1, \ldots, \gamma_{N_{BS}}\}$ have a common optimality criterion, they are interdependent on each others and cannot be selected individually. Hence, we consider that the decision rules are all selected at the MSC and that the local decision rules are updated at the base stations when the average SNR or the channel state varies, depending on the information available at the MSC and other assumptions as will be discussed next. In fact, it will be shown that the optimum decision rules are likelihood ratio quantizers such that the MSC only needs to transmit through the fixed network new threshold values to the base stations in order to update the local decision rules. As mentioned in the previous section, the channel state vector $\mathbf{H} = [\mathbf{H}_1, \ldots, \mathbf{H}_{N_{BS}}]^T$ may or may not be available at the MSC although the channel state vector $\mathbf{H}_k$ is perfectly known at the kth base station. Since both cases provide different decision rules, the derivation is separated in two parts treating separately both cases.

2.2.1 Known channel state information at the fusion center

In this section, it is assumed that the channel state is varying slowly enough at each base station such that accurate estimates can be transmitted to the MSC, where the decision rules are optimized and the final decision U_0 is made. Therefore, since the channel state vector $\mathbf{H}$ is known at the MSC, the channel state information is available to the fusion center such that the optimum fusion rule should take advantage of this information and be a function of the channel state vector $\mathbf{H}$. In addition, since the decision rules are optimized at the MSC, it is also possible for the local decision rules to be functions of the channel state vector $\mathbf{H}$. This requires the MSC to update the local decision rules used at the base stations every time the channel state varies. The optimum decision rules are therefore functions of the channel state vector $\mathbf{H}$ and minimize, given $\mathbf{H} = \mathbf{h}$, the conditional probability of error (2.5).

It can be argued that, depending on the rate at which the channel state varies, such a scheme may require more bandwidth from the fixed network than the OC scheme (see appendix D) contradicting our original goal of designing bandwidth efficient handoff macrodiversity schemes. However, it is still important to consider such a scheme since its probability of bit error represents a lower bound to the probability of bit error of any possible SHDD scheme. A bandwidth efficient alternative to the optimum scheme will be to limit the MSC to update the local decision rules only when the average SNR varies at any base station. Hence, as opposed to the optimum scheme, the kth base station local decision rule is not a function of the channel state vector $\mathbf{H}$ anymore. However, since $\mathbf{H}_k$ is perfectly known at the kth base station, it is possible for the kth base station local decision rule to be a function of the channel state vector $\mathbf{H}_k$.

In this section, we are considering the optimum scheme and the bandwidth efficient scheme. We first derive a fusion rule which is optimum in the sense that for fixed local decision rules at the base stations, it provides the minimum average probability of bit error at the output of the fusion center. Then, we derive, for the optimum scheme and bandwidth efficient scheme, the kth base station decision rule which is optimum in the sense that for a fixed fusion rule and fixed local decision rules at the remaining base stations, it provides the minimum average probability of bit error at the output of the fusion center.

A. Optimum fusion rule

At the MSC, the only information available to the fusion rule to make a final decision U_0 on the transmitted bit B is the local decision vector $\mathbf{U} = [U_1, \ldots, U_{N_{BS}}]^T$ and the channel state vector $\mathbf{H}$. Furthermore, since the optimality criterion is the probability of bit error which is a Bayesian criterion, it can be assumed that the fusion rule is deterministic. In fact, it is shown in [19] that even under a Neyman-Pearson criterion, a randomized fusion rule is never optimum, when the local detector likelihood ratios contain no point mass. Hence, when the channel state vector $\mathbf{H}$ is equal to $\mathbf{h}$, the fusion rule should partition the observation set $\mathcal{Z}$ containing all possible realizations of $\mathbf{U}$ into the mutually exclusive sets $\mathcal{Z}_0(\mathbf{h})$ and $\mathcal{Z}_1(\mathbf{h})$. The conditional probability $P(U_0 = 1 \mid \mathbf{h}, \mathbf{u})$ in (2.5) can thus be expressed as follows

$$P(U_0 = 1 \mid \mathbf{h}, \mathbf{u}) = 1 - P(U_0 = 0 \mid \mathbf{h}, \mathbf{u}) = \begin{cases} 1 \text{ if } \mathbf{u} \in \mathcal{Z}_1(\mathbf{h}) \\ 0 \text{ if } \mathbf{u} \in \mathcal{Z}_0(\mathbf{h}) \end{cases} \tag{2.6}$$

and the conditional probability of bit error (2.5) can be rewritten as follows

$$P_{b|\mathbf{h}} = \frac{1}{2} + \frac{1}{2} \sum_{\mathbf{u} \in \mathcal{Z}_1(\mathbf{h})} \left[P(\mathbf{U} = \mathbf{u} \mid \mathbf{h}, B = 0) - P(\mathbf{U} = \mathbf{u} \mid \mathbf{h}, B = 1) \right]. \tag{2.7}$$

From expression (2.7), it is clear that the realizations of $\mathbf{U}$ that make the summand negative must be included in $\mathcal{Z}_1(\mathbf{h})$ and the realizations of $\mathbf{U}$ that make the summand positive must be included in $\mathcal{Z}_0(\mathbf{h})$ in order to minimize the probability of bit error. However, the realizations of $\mathbf{U}$ that make the summand equal to 0 can be included in either set without affecting the performances of the system. Using these facts, the optimum fusion rule can be formulated as follows

$$\Lambda_{1,0}^{(0)}(\mathbf{u}, \mathbf{h}) = \frac{P(\mathbf{U} = \mathbf{u} \mid \mathbf{h}, B = 1)}{P(\mathbf{U} = \mathbf{u} \mid \mathbf{h}, B = 0)} \quad \underset{U_0 = 0}{\overset{U_0 = 1}{\gtrless}} \quad 1 \tag{2.8}$$

or equivalently in the maximum a-posterior form

$$P(B = 1 \mid \mathbf{h}, \mathbf{U} = \mathbf{u}) \quad \underset{U_0 = 0}{\overset{U_0 = 1}{\gtrless}} \quad P(B = 0 \mid \mathbf{h}, \mathbf{U} = \mathbf{u}) \tag{2.9}$$

since $P(B = 1) = P(B = 0)$. Then, considering the local decisions contained in $\mathbf{U} = [U_1, \ldots, U_{N_{BS}}]^T$ are conditionally independent since no communication is assumed between the base stations, the likelihood ratio $\Lambda_{1,0}^{(0)}(\mathbf{u}, \mathbf{h})$ simplifies to

$$\Lambda_{1,0}^{(0)}(\mathbf{u}, \mathbf{h}) = \prod_{k=1}^{N_{BS}} \frac{P(U_k = u_k \mid \mathbf{h}, B = 1)}{P(U_k = u_k \mid \mathbf{h}, B = 0)}. \tag{2.10}$$

It is interesting to note that even if the channel fading is not spatially uncorrelated, the decision rule stays valid. In addition, the derived fusion rule is dependent on the local decision rules and is optimum, regardless of the rate at which the local decision rules are updated at the base stations.

B. Optimum local decision rules assuming the MSC updates the local decision rules when the channel state varies at any base station

Considering that, in the optimum scheme, the MSC updates the decision rules used by the local detectors when the channel state varies, the information available to the local detector of the kth base station to make the decision U_k is the received signal vector $\mathbf{R}_k$ and the channel state vector $\mathbf{H}$. Assuming first all base stations are making a hard decision, the local decision rule $\gamma_k(\mathbf{r}_k, \mathbf{h})$, which determines the value of the local decision U_k given $\mathbf{R}_k = \mathbf{r}_k$ and $\mathbf{H} = \mathbf{h}$, should therefore partition the observation set $\mathcal{R}^k$ containing all possible realizations of $\mathbf{R}_k$ into the mutually exclusive sets $\mathcal{R}_1^k(\mathbf{h})$ and $\mathcal{R}_0^k(\mathbf{h})$. In order to determine which realizations of $\mathbf{R}_k$ should be included in these sets, it is necessary to expand the conditional probability of bit error (2.5) as a function of $\mathbf{r}_k$.

Before expanding the conditional probability of bit error (2.5) as a function of $\mathbf{r}_k$, it is advantageous to first expand expression (2.5) as follows

$$
\begin{aligned}
P_{b|\mathbf{h}} & \\
= \; & \frac{1}{2} + \frac{1}{2}\sum_{\mathbf{u}^k} P(U_0 = 1 \mid \mathbf{h}, \mathbf{U} = \mathbf{u}^{k0}) \left[P(\mathbf{U} = \mathbf{u}^{k0} \mid \mathbf{h}, B = 0) - P(\mathbf{U} = \mathbf{u}^{k0} \mid \mathbf{h}, B = 1)\right] \\
& + \frac{1}{2}\sum_{\mathbf{u}^k} P(U_0 = 1 \mid \mathbf{h}, \mathbf{U} = \mathbf{u}^{k1}) \left[P(\mathbf{U} = \mathbf{u}^{k1} \mid \mathbf{h}, B = 0) - P(\mathbf{U} = \mathbf{u}^{k1} \mid \mathbf{h}, B = 1)\right], \quad (2.11)
\end{aligned}
$$

where $\mathbf{u}^k = [u_1, \ldots, u_{k-1}, u_{k+1}, \ldots, u_{N_{BS}}]^T$, $\mathbf{u}^{k1} = [u_1, \ldots, u_{k-1}, 1, u_{k+1}, \ldots, u_{N_{BS}}]^T$ and $\mathbf{u}^{k0} = [u_1, \ldots, u_{k-1}, 0, u_{k+1}, \ldots, u_{N_{BS}}]^T$. Then, given that $P(\mathbf{U} = \mathbf{u}^{k0} \mid \mathbf{h}, B = b) = P(\mathbf{U}^k = \mathbf{u}^k \mid \mathbf{h}, B = b) - P(\mathbf{U} = \mathbf{u}^{k1} \mid \mathbf{h}, B = b)$, the conditional probability of bit error (2.11) can be written as

$$\begin{aligned}
P_{b|\mathbf{h}} &= \frac{1}{2}+\frac{1}{2}\sum_{\mathbf{u}^k} P(U_0=1 \mid \mathbf{h}, \mathbf{U}=\mathbf{u}^{k0}) \left[P(\mathbf{U}^k=\mathbf{u}^k \mid \mathbf{h}, B=0) - P(\mathbf{U}^k=\mathbf{u}^k \mid \mathbf{h}, B=1)\right] \\
&\quad -\frac{1}{2}\sum_{\mathbf{u}^k} P(U_0=1 \mid \mathbf{h}, \mathbf{U}=\mathbf{u}^{k0}) \left[P(\mathbf{U}=\mathbf{u}^{k1} \mid \mathbf{h}, B=0) - P(\mathbf{U}=\mathbf{u}^{k1} \mid \mathbf{h}, B=1)\right] \\
&\quad +\frac{1}{2}\sum_{\mathbf{u}^k} P(U_0=1 \mid \mathbf{h}, \mathbf{U}=\mathbf{u}^{k1}) \left[P(\mathbf{U}=\mathbf{u}^{k1} \mid \mathbf{h}, B=0) - P(\mathbf{U}=\mathbf{u}^{k1} \mid \mathbf{h}, B=1)\right] \\
&= A_k(\mathbf{h}) + \sum_{\mathbf{u}^k} B_k(\mathbf{u}^k, \mathbf{h}) \left[P(\mathbf{U}=\mathbf{u}^{k1} \mid \mathbf{h}, B=0) - P(\mathbf{U}=\mathbf{u}^{k1} \mid \mathbf{h}, B=1)\right], \qquad (2.12)
\end{aligned}$$

where $\mathbf{U}^k = [U_1, \ldots, U_{k-1}, U_{k+1}, \ldots, U_{N_{BS}}]^T$,

$$A_k(\mathbf{h}) = \frac{1}{2}+\frac{1}{2}\sum_{\mathbf{u}^k} P(U_0=1 \mid \mathbf{h}, \mathbf{U}=\mathbf{u}^{k0}) \left[P(\mathbf{U}^k=\mathbf{u}^k \mid \mathbf{h}, B=0) - P(\mathbf{U}^k=\mathbf{u}^k \mid \mathbf{h}, B=1)\right] \qquad (2.13)$$

and

$$B_k(\mathbf{u}^k, \mathbf{h}) = \frac{1}{2}P(U_0=1 \mid \mathbf{h}, \mathbf{U}=\mathbf{u}^{k1}) - \frac{1}{2}P(U_0=1 \mid \mathbf{h}, \mathbf{U}=\mathbf{u}^{k0}). \qquad (2.14)$$

Let $f_{\mathbf{R}_k}(\mathbf{r}_k \mid \mathbf{h}_k, B=b)$ denote the joint PDF of the received signals $R_{k,1}, \ldots, R_{k,N_R}$ given the bit $B=b$ and the channel state vector $\mathbf{H}_k = \mathbf{h}_k$. Since no communication is assumed between the base stations, the local decisions are conditionally independent and the conditional probability $P(\mathbf{U}=\mathbf{u}^{k1} \mid \mathbf{h}, B=b)$ in equation (2.12) can be expanded as a function of $\mathbf{r}_k$ as follows

$$\begin{aligned}
&P(\mathbf{U}=\mathbf{u}^{k1} \mid \mathbf{h}, B=b) \\
&= P(\mathbf{U}^k=\mathbf{u}^k \mid \mathbf{h}, B=b) P(U_k=1 \mid \mathbf{h}, B=b) \\
&= P(\mathbf{U}^k=\mathbf{u}^k \mid \mathbf{h}, B=b) \int_{\mathbf{r}_k} P(U_k=1 \mid \mathbf{h}, \mathbf{r}_k, B=b) f_{\mathbf{R}_k}(\mathbf{r}_k \mid \mathbf{h}, B=b) d\mathbf{r}_k \\
&= P(\mathbf{U}^k=\mathbf{u}^k \mid \mathbf{h}, B=b) \int_{\mathbf{r}_k} P(U_k=1 \mid \mathbf{h}, \mathbf{r}_k) f_{\mathbf{R}_k}(\mathbf{r}_k \mid \mathbf{h}_k, B=b) d\mathbf{r}_k, \qquad (2.15)
\end{aligned}$$

where $P(U_k=1 \mid \mathbf{h}, \mathbf{r}_k, B=b) = P(U_k=1 \mid \mathbf{h}, \mathbf{r}_k)$, since the local decision U_k is specified by the deterministic decision rule $\gamma_k(\mathbf{r}_k, \mathbf{h})$ given $\mathbf{R}_k = \mathbf{r}_k$ and $\mathbf{H} = \mathbf{h}$, and $f_{\mathbf{R}_k}(\mathbf{r}_k \mid \mathbf{h}, B=b) = f_{\mathbf{R}_k}(\mathbf{r}_k \mid \mathbf{h}_k, B=b)$, since $\mathbf{R}_k$ is independent of $\mathbf{H}^k =$

$[\mathbf{H}_1, \ldots, \mathbf{H}_{k-1}, \mathbf{H}_{k+1}, \ldots, \mathbf{H}_{N_{BS}}]^T$. The conditional probability of bit error (2.12) can therefore be expanded as a function of $\mathbf{r}_k$ using (2.15) as follows

$$\begin{aligned}
P_{b|\mathbf{h}} &= A_k(\mathbf{h}) + \sum_{\mathbf{u}^k} B_k(\mathbf{u}^k, \mathbf{h}) \times \\
&\quad \Big[P(\mathbf{U}^k = \mathbf{u}^k \mid \mathbf{h}, B = 0) \int_{\mathbf{r}_k} P(U_k = 1 \mid \mathbf{h}, \mathbf{r}_k) f_{\mathbf{R}_k}(\mathbf{r}_k \mid \mathbf{h}_k, B = 0) d\mathbf{r}_k \\
&\quad -P(\mathbf{U}^k = \mathbf{u}^k \mid \mathbf{h}, B = 1) \int_{\mathbf{r}_k} P(U_k = 1 \mid \mathbf{h}, \mathbf{r}_k) f_{\mathbf{R}_k}(\mathbf{r}_k \mid \mathbf{h}_k, B = 1) d\mathbf{r}_k\Big] \\
&= A_k(\mathbf{h}) + \int_{\mathbf{r}_k} P(U_k = 1 \mid \mathbf{h}, \mathbf{r}_k) \times \\
&\quad \sum_{\mathbf{u}^k} B_k(\mathbf{u}^k, \mathbf{h}) \Big[P(\mathbf{U}^k = \mathbf{u}^k \mid \mathbf{h}, B = 0) f_{\mathbf{R}_k}(\mathbf{r}_k \mid \mathbf{h}_k, B = 0) \\
&\quad -P(\mathbf{U}^k = \mathbf{u}^k \mid \mathbf{h}, B = 1) f_{\mathbf{R}_k}(\mathbf{r}_k \mid \mathbf{h}_k, B = 1)\Big] d\mathbf{r}_k
\end{aligned} \tag{2.16}$$

Since the local decision U_k is specified by the deterministic decision rule $\gamma_k(\mathbf{r}_k, \mathbf{h})$ given $\mathbf{R}_k = \mathbf{r}_k$ and $\mathbf{H} = \mathbf{h}$, the conditional probability $P(U_k = 1 \mid \mathbf{h}, \mathbf{r}_k)$ in (2.16) can thus be expressed as follows

$$P(U_k = 1 \mid \mathbf{h}, \mathbf{r}_k) = 1 - P(U_k = 0 \mid \mathbf{h}, \mathbf{r}_k) = \begin{cases} 1 \text{ if } \mathbf{r}_k \in \mathcal{R}_1^k(\mathbf{h}) \\ 0 \text{ if } \mathbf{r}_k \in \mathcal{R}_0^k(\mathbf{h}) \end{cases} \tag{2.17}$$

and the conditional probability of bit error (2.16) can be rewritten as follows

$$\begin{aligned}
P_{b|\mathbf{h}} &= A_k(\mathbf{h}) + \int_{\mathbf{r}_k \in \mathcal{R}_1^k(\mathbf{h})} \Big[f_{\mathbf{R}_k}(\mathbf{r}_k \mid \mathbf{h}_k, B = 0) \sum_{\mathbf{u}^k} B_k(\mathbf{u}^k, \mathbf{h}) P(\mathbf{U}^k = \mathbf{u}^k \mid \mathbf{h}, B = 0) \\
&\quad -f_{\mathbf{R}_k}(\mathbf{r}_k \mid \mathbf{h}_k, B = 1) \sum_{\mathbf{u}^k} B_k(\mathbf{u}^k, \mathbf{h}) P(\mathbf{U}^k = \mathbf{u}^k \mid \mathbf{h}, B = 1)\Big] d\mathbf{r}_k.
\end{aligned} \tag{2.18}$$

From expression (2.18), it can be concluded that, in order to minimize the probability of bit error, the realizations of $\mathbf{R}_k$ that make the integrand negative must be included in $\mathcal{R}_1^k(\mathbf{h})$. The optimum local decision rule for the kth base station can therefore be formulated as follows

$$m_k \Lambda_{1,0}^{(k)}(\mathbf{r}_k, \mathbf{h}_k) - b_k \underset{U_k = 0}{\overset{U_k = 1}{\gtrless}} 0 \tag{2.19}$$

where

$$\Lambda_{1,0}^{(k)}(\mathbf{r}_k, \mathbf{h}_k) = \frac{f_{\mathbf{R}_k}(\mathbf{r}_k \mid \mathbf{h}_k, B = 1)}{f_{\mathbf{R}_k}(\mathbf{r}_k \mid \mathbf{h}_k, B = 0)}, \tag{2.20}$$

$$m_k = \sum_{\mathbf{u}^k} B_k(\mathbf{u}^k, \mathbf{h}) P(\mathbf{U}^k = \mathbf{u}^k \mid \mathbf{h}, B = 1) \tag{2.21}$$

and

$$b_k = \sum_{\mathbf{u}^k} B_k(\mathbf{u}^k, \mathbf{h}) P(\mathbf{U}^k = \mathbf{u}^k \mid \mathbf{h}, B = 0) \tag{2.22}$$

Since the coefficients m_k and b_k are independent of $\mathbf{r}_k$ but vary with $\mathbf{h}$, it can be seen on Fig. 2.2 that, for a given $\mathbf{h}$, the optimum decision rule at the kth base station is a likelihood ratio threshold test, with the exception of the case when the local decision U_k is discarded by the fusion rule.

It is important to mention that the value assigned to U_k only identifies the interval, delimited by the threshold, in which the likelihood ratio $\Lambda_{1,0}^{(k)}(\mathbf{r}_k, \mathbf{h}_k)$ appears. The mapping can therefore be permuted as long as the same change is made in the fusion rule. If the local decision rules are made nondecreasing functions of the likelihood ratios by permuting the mapping when necessary, it implies that $\prod_{k=1}^{N_{BS}} P(U_k = u_k \mid \mathbf{h}, B = 1)/P(U_k = u_k \mid \mathbf{h}, B = 0)$ is a nondecreasing function of $\mathbf{u}$ [20]. Consequently, the fusion rule is a nondecreasing function of $\mathbf{u}$. The converse is also true, when the fusion rule is a nondecreasing function of $\mathbf{u}$, all local decision rules are nondecreasing functions of the likelihood ratios since $B_k(\mathbf{u}^k, \mathbf{h}) \geq 0$ such that m_k and b_k are always larger than 0 (see Fig. 2.2). Therefore, assuming that the fusion rule is a nondecreasing function, the optimum local decision rule for the kth base station can be formulated as follows

$$\Lambda_{1,0}^{(k)}(\mathbf{r}_k, \mathbf{h}_k) \underset{U_k = 0}{\overset{U_k = 1}{\gtrless}} \frac{\sum_{\mathbf{u}^k} B_k(\mathbf{u}^k, \mathbf{h}) P(\mathbf{U}^k = \mathbf{u}^k \mid \mathbf{h}, B = 0)}{\sum_{\mathbf{u}^k} B_k(\mathbf{u}^k, \mathbf{h}) P(\mathbf{U}^k = \mathbf{u}^k \mid \mathbf{h}, B = 1)} = t_k(\mathbf{h}) \tag{2.23}$$

where the threshold $t_k(\mathbf{h})$ is a function of $\mathbf{h}$ and must therefore be updated as the channel state varies. Furthermore, the threshold $t_k(\mathbf{h})$ is dependent on the fusion rule and the other local decision rules such that expression (2.23) is optimal only if the other decision rules are optimal.

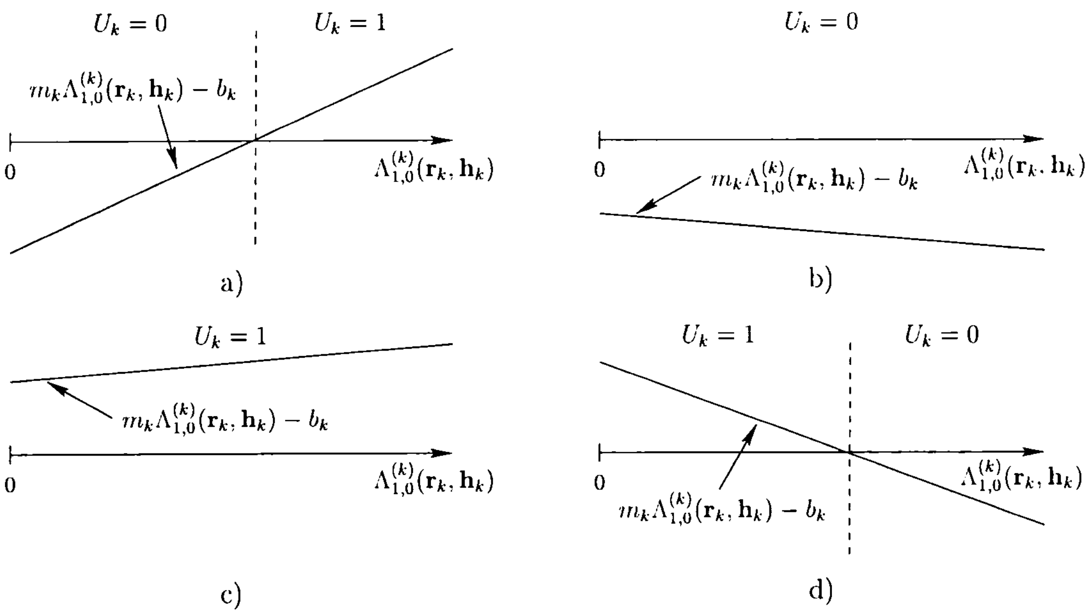


Fig. 2.2 Graphical representation of the kth base station local decision rule: a) $m_k > 0$ and $b_k > 0$, b) $m_k \leq 0$ and $b_k > 0$, c) $m_k \geq 0$ and $b_k \leq 0$, d) $m_k < 0$ and $b_k < 0$

These results can be extended to the case when local detectors are making soft decisions. In fact, in [20] it is proved that the local decision rules, for a distributed detection scheme using soft decisions, are deterministic monotone likelihood ratio threshold tests, as long as the observations made at the distributed sensors are conditionally independent. Using this result, the optimum local decision rule at the kth base station can be generalized, in a logarithmic form, as follows

$$U_k = u_k \text{ if } t_{k,u_k}(\mathbf{h}) \le \Psi_{1,0}^{(k)}(\mathbf{r}_k, \mathbf{h}_k) < t_{k,u_k+1}(\mathbf{h}) \tag{2.24}$$

where

$$\Psi_{1,0}^{(k)}(\mathbf{r}_k, \mathbf{h}_k) = \ln\left(\Lambda_{1,0}^{(k)}(\mathbf{r}_k, \mathbf{h}_k)\right) \tag{2.25}$$

given $u_k = 0, \ldots, L-1$, $k = 1, \ldots, N_{BS}$, $t_{k,0}(\mathbf{h}) = -\infty$ and $t_{k,L}(\mathbf{h}) = \infty$. The local decision rule at the kth base station is therefore defined by the $L-1$ thresholds contained in the vector $\mathbf{t}_k(\mathbf{h}) = [t_{k,1}(\mathbf{h}), \ldots, t_{k,L-1}(\mathbf{h})]$ and partitioning the Log-Likelihood Ratio (LLR) $\Psi_{1,0}^{(k)}(\mathbf{r}_k, \mathbf{h}_k)$.

It is important to mention that the values of these thresholds cannot be determined analytically due to the interdependence of these thresholds and other local detector thresholds caused by the common optimality criterion. The set of thresholds contained in the vector $\mathbf{t}(\mathbf{h}) = [\mathbf{t}_1(\mathbf{h}), \ldots, \mathbf{t}_{N_{BS}}(\mathbf{h})]^T$ must therefore be optimized simultaneously using a numerical optimization algorithm in order to determine their optimum values. Since the channel state is known at the MSC where the numerical optimization takes place and new thresholds are transmitted to the base stations every-time the channel state varies, it is obvious that, in order to minimize the probability of bit error, the cost function $J(\mathbf{t})$ used for the threshold optimization, given $\mathbf{H} = \mathbf{h}$, is $P_{b|\mathbf{h}}(\mathbf{t}(\mathbf{h}))$ representing the conditional probability of bit error (2.5) as a function of the thresholds in $\mathbf{t}(\mathbf{h})$. The optimization of the local detector thresholds is treated in more details in section 2.3 and an analytical expression is derived for $P_{b|\mathbf{h}}(\mathbf{t}(\mathbf{h}))$ in Appendix B.

C. Optimum local decision rules assuming the MSC updates the local decision rules when the average SNR varies at any base station

Considering that, in the bandwidth efficient scheme, the MSC only updates the decision rules used by the local detectors when the average SNR varies at any base station, the information available to the local detector of the kth base station to make its decision is the received signal vector $\mathbf{R}_k$ and the channel state vector $\mathbf{H}_k$. Assuming first all base stations are making a hard decision, the local decision rule $\gamma_k(\mathbf{r}_k, \mathbf{h}_k)$, which determine the value of the local decision U_k given $\mathbf{R}_k = \mathbf{r}_k$ and $\mathbf{H}_k = \mathbf{h}_k$, should therefore partition the observation set $\mathcal{R}^k$ containing all possible realizations of $\mathbf{R}_k$ into the mutually exclusive sets $\mathcal{R}_1^k(\mathbf{h}_k)$ and $\mathcal{R}_0^k(\mathbf{h}_k)$. In order to determine which realizations of $\mathbf{R}_k$ should be included in these sets, it is necessary to expand as a function of $\mathbf{r}_k$ the probability of bit error conditioned on the channel state vector $\mathbf{H}_k = \mathbf{h}_k$, which can be defined as follows

$$P_{b|\mathbf{h}_k} = \int_{\mathbf{h}^k} P_{b|\mathbf{h}} f_{\mathbf{H}^k}(\mathbf{h}^k|\mathbf{h}_k) d\mathbf{h}^k = \int_{\mathbf{h}^k} P_{b|\mathbf{h}} f_{\mathbf{H}^k}(\mathbf{h}^k) d\mathbf{h}^k, \tag{2.26}$$

where $P_{b|\mathbf{h}}$ is defined in expression (2.5) and $f_{\mathbf{H}^k}(\mathbf{h}^k|\mathbf{h}_k) = f_{\mathbf{H}^k}(\mathbf{h}^k)$ since $\mathbf{H}^k = [\mathbf{H}_1, \ldots, \mathbf{H}_{k-1}, \mathbf{H}_{k+1}, \ldots, \mathbf{H}_{N_{BS}}]^T$ is independent of $\mathbf{H}_k$. In the derivation of the kth base station local decision rule of the optimum scheme, we already expanded the conditional probability of bit error (2.5) as a function of $\mathbf{r}_k$ in (2.16). Hence, by substituting (2.16) in (2.26), the conditional probability of bit error $P_{b|\mathbf{h}_k}$ can therefore be expanded as a function of $\mathbf{r}_k$ as follows

$$\begin{aligned} P_{b|\mathbf{h}_k} = & \int_{\mathbf{h}^k} f_{\mathbf{H}^k}(\mathbf{h}^k) A_k(\mathbf{h}) d\mathbf{h}^k + \int_{\mathbf{h}^k} f_{\mathbf{H}^k}(\mathbf{h}^k) \Big(\int_{\mathbf{r}_k} P(U_k = 1 \mid \mathbf{h}, \mathbf{r}_k) \times \\ & \sum_{\mathbf{u}^k} B_k(\mathbf{u}^k, \mathbf{h}) \Big[P(\mathbf{U}^k = \mathbf{u}^k \mid \mathbf{h}, B = 0) f_{\mathbf{R}_k}(\mathbf{r}_k \mid \mathbf{h}_k, B = 0) \\ & - P(\mathbf{U}^k = \mathbf{u}^k \mid \mathbf{h}, B = 1) f_{\mathbf{R}_k}(\mathbf{r}_k \mid \mathbf{h}_k, B = 1) \Big] d\mathbf{r}_k \Big) d\mathbf{h}^k, \end{aligned} \tag{2.27}$$

where $A_k(\mathbf{h})$ and $B_k(\mathbf{u}^k, \mathbf{h})$ are defined in (2.13) and (2.14) respectively. However, as opposed to the optimum scheme, the kth base station local decision of the bandwidth efficient scheme is independent of $\mathbf{H}^k$ such that $P(U_k = u_k \mid \mathbf{h}, \mathbf{r}_k) = P(U_k = u_k \mid \mathbf{h}_k, \mathbf{r}_k)$ and $P(\mathbf{U}^k = \mathbf{u}^k \mid \mathbf{h}, B = b) = P(\mathbf{U}^k = \mathbf{u}^k \mid \mathbf{h}^k, B = b)$. The conditional probability of bit error (2.27) can therefore be simplified as follows

$$\begin{aligned} & P_{b|\mathbf{h}_k} \\ & = \hat{A}_k(\mathbf{h}_k) + \int_{\mathbf{r}_k} P(U_k = 1 \mid \mathbf{h}_k, \mathbf{r}_k) \times \\ & \int_{\mathbf{h}^k} f_{\mathbf{H}^k}(\mathbf{h}^k) \Big(\sum_{\mathbf{u}^k} B_k(\mathbf{u}^k, \mathbf{h}) \Big[P(\mathbf{U}^k = \mathbf{u}^k \mid \mathbf{h}^k, B = 0) f_{\mathbf{R}_k}(\mathbf{r}_k \mid \mathbf{h}_k, B = 0) \\ & - P(\mathbf{U}^k = \mathbf{u}^k \mid \mathbf{h}^k, B = 1) f_{\mathbf{R}_k}(\mathbf{r}_k \mid \mathbf{h}_k, B = 1) \Big] d\mathbf{r}_k \Big) d\mathbf{h}^k, \end{aligned} \tag{2.28}$$

where

$$\hat{A}_k(\mathbf{h}_k) = \frac{1}{2} + \frac{1}{2}\int_{\mathbf{h}^k} f_{\mathbf{H}^k}(\mathbf{h}^k) \sum_{\mathbf{u}^k} P(U_0 = 1 \mid \mathbf{h}, \mathbf{U} = \mathbf{u}^{k0}) \times \Big[P(\mathbf{U}^k = \mathbf{u}^k \mid \mathbf{h}^k, B = 0) - P(\mathbf{U}^k = \mathbf{u}^k \mid \mathbf{h}^k, B = 1)\Big] d\mathbf{h}^k. \quad (2.29)$$

Since the local decision U_k is specified by the deterministic decision rule $\gamma_k(\mathbf{r}_k, \mathbf{h}_k)$ given $\mathbf{R}_k = \mathbf{r}_k$ and $\mathbf{H}_k = \mathbf{h}_k$, the conditional probability $P(U_k = u_k \mid \mathbf{h}_k, \mathbf{r}_k)$ in (2.28) can thus be expressed as follows

$$P(U_k = 1 \mid \mathbf{h}_k, \mathbf{r}_k) = 1 - P(U_k = 0 \mid \mathbf{h}_k, \mathbf{r}_k) = \begin{cases} 1 & \text{if } \mathbf{r}_k \in \mathcal{R}_1^k(\mathbf{h}_k) \\ 0 & \text{if } \mathbf{r}_k \in \mathcal{R}_0^k(\mathbf{h}_k) \end{cases} \quad (2.30)$$

and the conditional probability of bit error (2.28) can be rewritten as follows

$$\begin{aligned} &P_{b|\mathbf{h}_k} \\ &= \hat{A}_k(\mathbf{h}_k) + \int_{\mathbf{r}_k \in \mathcal{R}_1^k(\mathbf{h}_k)} \\ &\quad \Bigg[f_{\mathbf{R}_k}(\mathbf{r}_k \mid \mathbf{h}_k, B = 0) \sum_{\mathbf{u}^k} \int_{\mathbf{h}^k} f_{\mathbf{H}^k}(\mathbf{h}^k) B_k(\mathbf{u}^k, \mathbf{h}) P(\mathbf{U}^k = \mathbf{u}^k \mid \mathbf{h}^k, B = 0) d\mathbf{h}^k \\ &\quad - f_{\mathbf{R}_k}(\mathbf{r}_k \mid \mathbf{h}_k, B = 1) \sum_{\mathbf{u}^k} \int_{\mathbf{h}^k} f_{\mathbf{H}^k}(\mathbf{h}^k) B_k(\mathbf{u}^k, \mathbf{h}) P(\mathbf{U}^k = \mathbf{u}^k \mid \mathbf{h}^k, B = 1) d\mathbf{h}^k \Bigg] d\mathbf{r}_k. \quad (2.31) \end{aligned}$$

From expression (2.31), it can be concluded that, in order to minimize the conditional probability of bit error, the realizations of $\mathbf{R}_k$ that make the integrand negative must be included in $\mathcal{R}_1^k(\mathbf{h}_k)$. Similarly to the optimum scheme, the optimum local decision rule of the bandwidth efficient scheme for the kth base station can therefore be formulated as follows

$$m_k \Lambda_{1,0}^{(k)}(\mathbf{r}_k, \mathbf{h}_k) - b_k \underset{U_k = 0}{\overset{U_k = 1}{\gtrless}} 0 \quad (2.32)$$

where $\Lambda_{1,0}^{(k)}(\mathbf{r}_k, \mathbf{h}_k)$ is defined in (2.20). However, as opposed to the optimum scheme,

$$m_k = \sum_{\mathbf{u}^k} \int_{\mathbf{h}^k} f_{\mathbf{H}^k}(\mathbf{h}^k) B_k(\mathbf{u}^k, \mathbf{h}) P(\mathbf{U}^k = \mathbf{u}^k \mid \mathbf{h}^k, B = 1) d\mathbf{h}^k \tag{2.33}$$

and

$$b_k = \sum_{\mathbf{u}^k} \int_{\mathbf{h}^k} f_{\mathbf{H}^k}(\mathbf{h}^k) B_k(\mathbf{u}^k, \mathbf{h}) P(\mathbf{U}^k = \mathbf{u}^k \mid \mathbf{h}^k, B = 0) d\mathbf{h}^k. \tag{2.34}$$

The coefficients m_k and b_k are now independent of $\mathbf{h}^k$ and $\mathbf{r}_k$, although they are still functions of $\mathbf{h}_k$. Hence, as shown in Fig. 2.2, for a given $\mathbf{h}_k$ the optimum decision rule at the kth base station is a likelihood ratio threshold test, where the threshold is a function of $\mathbf{h}_k$ and equals

$$t_k(\mathbf{h}_k) = \frac{\sum_{\mathbf{u}^k} \int_{\mathbf{h}^k} f_{\mathbf{H}^k}(\mathbf{h}^k) B_k(\mathbf{u}^k, \mathbf{h}) P(\mathbf{U}^k = \mathbf{u}^k \mid \mathbf{h}^k, B = 0) d\mathbf{h}^k}{\sum_{\mathbf{u}^k} \int_{\mathbf{h}^k} f_{\mathbf{H}^k}(\mathbf{h}^k) B_k(\mathbf{u}^k, \mathbf{h}) P(\mathbf{U}^k = \mathbf{u}^k \mid \mathbf{h}^k, B = 1) d\mathbf{h}^k}. \tag{2.35}$$

The threshold $t_k(\mathbf{h}_k)$ is also dependent on the fusion rule and the other local decision rules such that expression (2.35) is optimal only if the other decision rules are optimal. However, since the other decision rules are not known a priori, expression (2.35) cannot be used directly to determine the optimum threshold value. All base stations thresholds must be optimized simultaneously using a numerical optimization algorithm in order to determine their optimum values. Unfortunately, since the MSC updates the local decision rules only when the average SNR varies, it makes the optimization very difficult, since the threshold used by kth base station decision rule does not appear as a scalar anymore but as a function of $\mathbf{h}_k$. It is important to mention that for the optimum scheme we do not have this problem since the thresholds are optimized for a given $\mathbf{h}$, every-time the channel state varies.

Therefore, as a sub-optimum alternative, we propose that the thresholds be independent of the channel state vector $\mathbf{H}_k$. Assuming soft decisions are made at the local detectors, the local decision rule at the kth base station can be expressed, in a logarithmic form, as follows

$$U_k = u_k \text{ if } t_{k,u_k} \leq \Psi_{1,0}^{(k)}(\mathbf{r}_k, \mathbf{h}_k) < t_{k,u_k+1} \tag{2.36}$$

where $\Psi_{1,0}^{(k)}(\mathbf{r}_k, \mathbf{h}_k)$ is defined in (2.25), $u_k = 0, \ldots, L-1$, $k = 1, \ldots, N_{BS}$, $t_{k,0} = -\infty$ and $t_{k,L} = \infty$. The local decision rule at the kth base station is therefore defined by the

$L-1$ thresholds contained in $\mathbf{t}_k = [t_{k,1}, \ldots, t_{k,L-1}]$ partitioning the LLR $\Psi_{1,0}^{(k)}(\mathbf{r}_k, \mathbf{h}_k)$.

Since in this case the channel state information cannot be used in the optimization of the thresholds contained in the vector $\mathbf{t} = [\mathbf{t}_1, \ldots, \mathbf{t}_{N_{BS}}]^T$, the cost function $J(\mathbf{t})$ used for the threshold optimization is $P_b(\mathbf{t})$, representing the average probability of bit error (2.3) as a function of the thresholds contained in the vector $\mathbf{t}$, rather than the conditional probability of bit error (2.5). The optimization of the local detector thresholds is treated in more details in section 2.3 and an analytical expression is derived for $P_b(\mathbf{t})$ in Appendix B considering the assumptions made in this section. It is important to mention that the evaluation of $P_b(\mathbf{t})$ for a given $\mathbf{t}$ requires a N_{BS}-fold integral to be performed numerically, making the optimization process time consuming.

2.2.2 Unknown channel state at the fusion center

In this section, it is assumed that the channel state vector $\mathbf{H}$ is not available at the MSC. However, the statistical properties of the channel state vector $\mathbf{H}$ are known. Therefore, as opposed to the previous case, the fusion rule is not a function of the channel state vector $\mathbf{H}$ and the MSC updates the local decision rules at the base stations every time the average SNR varies at any base station. However, since $\mathbf{H}_k$ is perfectly known at the kth base station, it is possible for the kth base station local decision rule to be a function of the channel state vector $\mathbf{H}_k$.

In this section, we first derive a fusion rule which is optimum in the sense that for fixed local decision rules at the base stations, it provides the minimum average probability of bit error at the output of the fusion center. Then, we derive the kth base station local decision rule which is optimum in the sense that for a fixed fusion rule and fixed local decision rules at the remaining base stations, it provides the minimum average probability of bit error at the output of the MSC.

A. Optimum fusion rule

At the MSC, the only information available to make a final decision U_0 on the transmitted bit B is the local decision vector $\mathbf{U} = [U_1, \ldots, U_{N_{BS}}]^T$. As mentioned in the previous section, since the optimality criterion is the probability of bit error which is a Bayesian criterion, it can be assumed that the fusion rule is deterministic. The fusion rule should therefore partition the observation set $\mathcal{Z}$ containing all possible realizations of $\mathbf{U}$ into the mutually exclusive sets $\mathcal{Z}_0$ and $\mathcal{Z}_1$. Hence, considering the fusion rule is independent of the channel state vector $\mathbf{H}$, the conditional probability $P(U_0 = u_0 \mid \mathbf{h}, \mathbf{u})$ reduces to $P(U_0 = u_0 \mid \mathbf{u})$ and the average probability of bit error (2.3) can be reformulated as follows

$$\begin{aligned} P_b &= \frac{1}{2} + \frac{1}{2} \sum_{\mathbf{u}} P(U_0 = 1 \mid \mathbf{u}) \times \\ &\quad \Big[\int_{\mathbf{h}} f_{\mathbf{H}}(\mathbf{h}) P(\mathbf{U} = \mathbf{u} \mid \mathbf{h}, B = 0) d\mathbf{h} - \int_{\mathbf{h}} f_{\mathbf{H}}(\mathbf{h}) P(\mathbf{U} = \mathbf{u} \mid \mathbf{h}, B = 1) d\mathbf{h} \Big] \\ &= \frac{1}{2} + \frac{1}{2} \sum_{\mathbf{u}} P(U_0 = 1 \mid \mathbf{u}) \Big[P(\mathbf{U} = \mathbf{u} \mid B = 0) - P(\mathbf{U} = \mathbf{u} \mid B = 1) \Big] \end{aligned} \tag{2.37}$$

since $P(\mathbf{U} = \mathbf{u} \mid B = b) = \int_{\mathbf{h}} f_{\mathbf{H}}(\mathbf{h}) P(\mathbf{U} = \mathbf{u} \mid \mathbf{h}, B = b) d\mathbf{h}$. Furthermore, the conditional probability $P(U_0 = 1 \mid \mathbf{u})$ in (2.37) can thus be expressed as follows

$$P(U_0 = 1 \mid \mathbf{u}) = 1 - P(U_0 = 0 \mid \mathbf{u}) = \begin{cases} 1 & \text{if } \mathbf{u} \in \mathcal{Z}_1 \\ 0 & \text{if } \mathbf{u} \in \mathcal{Z}_0 \end{cases} \tag{2.38}$$

such that the average probability of bit error (2.37) can be rewritten as follows

$$P_b = \frac{1}{2} + \frac{1}{2} \sum_{\mathbf{u} \in \mathcal{Z}_1} \Big[P(\mathbf{U} = \mathbf{u} \mid B = 0) - P(\mathbf{U} = \mathbf{u} \mid B = 1) \Big]. \tag{2.39}$$

From expression (2.39), it is clear that the realizations of $\mathbf{U}$ that make the summand negative must be included in $\mathcal{Z}_1$ and the realizations of $\mathbf{U}$ that make the summand positive must be included in $\mathcal{Z}_0$ in order to minimize the probability of bit error. However, the realizations of $\mathbf{U}$ that make the summand equal to 0 can be included in either set without affecting the performances of the system. Therefore, it is unnecessary to consider randomized test to take care of the equality case since it only make the test more complicated. Using these facts, the optimum fusion rule can be formulated as follows

$$\Lambda_{1,0}^{(0)}(\mathbf{u}) = \frac{P(\mathbf{U} = \mathbf{u} \mid B = 1)}{P(\mathbf{U} = \mathbf{u} \mid B = 0)} \underset{U_0 = 0}{\overset{U_0 = 1}{\substack{\geq \\ <}}} 1 \tag{2.40}$$

or equivalently in the maximum a-posterior form

$$P(B = 1 \mid \mathbf{U} = \mathbf{u}) \quad \underset{U_0 = 0}{\overset{U_0 = 1}{\geq \atop <}} \quad P(B = 0 \mid \mathbf{U} = \mathbf{u}). \tag{2.41}$$

Then, considering that the local decisions contained in the local decision vector $\mathbf{U} = [U_1, \ldots, U_{N_{BS}}]^T$ are conditionally independent since no communication is assumed between the base stations and the channel fading is spatially uncorrelated meaning that $f_{\mathbf{H}}(\mathbf{h}) = f_{\mathbf{H}_1}(\mathbf{h}_1) \times \cdots \times f_{\mathbf{H}_{N_{BS}}}(\mathbf{h}_{N_{BS}})$, the likelihood ratio $\Lambda_{1,0}^{(0)}(\mathbf{u})$ simplifies as follows

$$\Lambda_{1,0}^{(0)}(\mathbf{u}) = \frac{\prod_{k=1}^{N_{BS}} P(U_k = u_k \mid B = 1)}{\prod_{k=1}^{N_{BS}} P(U_k = u_k \mid B = 0)]}. \tag{2.42}$$

B. Optimum local decision rules

As the bandwidth efficient scheme presented in section 2.2.1, in this case the information available to the local detector of the kth base station to make the decision U_k is the received signal vector $\mathbf{R}_k$ and the channel state vector $\mathbf{H}_k$. Hence, assuming all base stations are making a hard decision, the optimum kth base station decision rule we derived in section 2.2.1 for the bandwidth efficient scheme is also valid in this case. The optimum local decision rule can therefore be formulated as follows

$$m_k \Lambda_{1,0}^{(k)}(\mathbf{r}_k, \mathbf{h}_k) - b_k \quad \underset{U_k = 0}{\overset{U_k = 1}{\geq \atop <}} \quad 0, \tag{2.43}$$

where $\Lambda_{1,0}^{(k)}(\mathbf{r}_k, \mathbf{h}_k)$ is defined in (2.20),

$$m_k = \sum_{\mathbf{u}^k} \int_{\mathbf{h}^k} f_{\mathbf{H}^k}(\mathbf{h}^k) B_k(\mathbf{u}^k, \mathbf{h}) P(\mathbf{U}^k = \mathbf{u}^k \mid \mathbf{h}^k, B = 1) d\mathbf{h}^k \tag{2.44}$$

and

$$b_k = \sum_{\mathbf{u}^k} \int_{\mathbf{h}^k} f_{\mathbf{H}^k}(\mathbf{h}^k) B_k(\mathbf{u}^k, \mathbf{h}) P(\mathbf{U}^k = \mathbf{u}^k \mid \mathbf{h}^k, B = 0) d\mathbf{h}^k. \tag{2.45}$$

However, since in this case the fusion rule is independent of $\mathbf{H}$, $P(U_0 = u_0 \mid \mathbf{h}, \mathbf{u}) = P(U_0 = u_0 \mid \mathbf{u})$ such that, in (2.44) and (2.45), $B_k(\mathbf{u}^k, \mathbf{h})$, which is defined in (2.14), can be replaced by

$$\hat{B}_k(\mathbf{u}^k) = \frac{1}{2} P(U_0 = 1 \mid \mathbf{U} = \mathbf{u}^{k1}) - \frac{1}{2} P(U_0 = 1 \mid \mathbf{U} = \mathbf{u}^{k0}). \tag{2.46}$$

The coefficient m_k and b_k can therefore be simplified to

$$m_k = \sum_{\mathbf{u}^k} \hat{B}_k(\mathbf{u}^k) \int_{\mathbf{h}^k} f_{\mathbf{H}^k}(\mathbf{h}^k) P(\mathbf{U}^k = \mathbf{u}^k \mid \mathbf{h}^k, B = 1) d\mathbf{h}^k = \sum_{\mathbf{u}^k} \hat{B}_k(\mathbf{u}^k) P(\mathbf{U}^k = \mathbf{u}^k \mid B = 1) \tag{2.47}$$

and

$$b_k = \sum_{\mathbf{u}^k} \hat{B}_k(\mathbf{u}^k) \int_{\mathbf{h}^k} f_{\mathbf{H}^k}(\mathbf{h}^k) P(\mathbf{U}^k = \mathbf{u}^k \mid \mathbf{h}^k, B = 0) d\mathbf{h}^k = \sum_{\mathbf{u}^k} \hat{B}_k(\mathbf{u}^k) P(\mathbf{U}^k = \mathbf{u}^k \mid B = 0), \tag{2.48}$$

since $P(\mathbf{U}^k = \mathbf{u}^k \mid B = b) = \int_{\mathbf{h}^k} f_{\mathbf{H}^k}(\mathbf{h}^k) P(\mathbf{U}^k = \mathbf{u}^k \mid \mathbf{h}^k, B = b) d\mathbf{h}^k$.

Thus, the coefficients m_k and b_k are in this case independent of both $\mathbf{r}_k$ and $\mathbf{h}_k$. Hence, as shown in Fig.2.2, the optimum decision rule at the kth base station is a likelihood ratio threshold test, where the threshold is independent of $\mathbf{h}_k$. It is important to mention that, as the bandwidth efficient scheme presented in section 2.2.1, it is assumed that the MSC only updates the local decision rules when the average SNR varies at any base station. However, since in this case the threshold is not a function of $\mathbf{h}_k$, it is therefore possible to derive its optimum value at the MSC without imposing any constraint on the threshold.

As mentioned previously, the mapping between the interval in which the likelihood ratio $\Lambda_{1,0}^{(k)}(\mathbf{r}_k, \mathbf{h}_k)$ appears and the local decision U_k can be permuted as long as the same change is made in the fusion rule. Hence, it can be assumed without loss of generality the local decision rules are nondecreasing functions of the likelihood ratios. It implies that $\prod_{k=1}^{N_{BS}} P(U_k = u_k \mid B = 1)/P(U_k = u_k \mid B = 0)$ is a nondecreasing function of $\mathbf{u}$ [20]. Consequently, the fusion rule is a nondecreasing function of $\mathbf{u}$. The converse is also true, when the fusion rule is a nondecreasing function of $\mathbf{u}$, all local decision rules are nondecreasing functions of the likelihood ratios since $\hat{B}_k(\mathbf{u}^k) \geq 0$ such that m_k and b_k are always larger than 0 (see Fig. 2.2). Therefore, assuming that the fusion rule is a nondecreasing function, the optimum local decision rule for the kth base station can be formulated as follows

$$\Lambda_{1,0}^{(k)}(\mathbf{r}_k, \mathbf{h}_k) \underset{U_k=0}{\overset{U_k=1}{\gtrless}} \frac{\sum_{\mathbf{u}^k} \hat{B}_k(\mathbf{u}^k) P(\mathbf{U}^k = \mathbf{u}^k \mid B = 0)}{\sum_{\mathbf{u}^k} \hat{B}_k(\mathbf{u}^k) P(\mathbf{U}^k = \mathbf{u}^k \mid B = 1)} = t_k \tag{2.49}$$

where the threshold t_k is independent of $\mathbf{h}_k$.

These results can be extended to the case when the local detectors are making soft decisions. As mentioned previously, in [20] it is proved that the local decision rules, for distributed detection schemes using soft decisions, are deterministic monotone likelihood ratio threshold tests, as long as the observations made at the distributed sensors are conditionally independent. Using this result, the optimum local decision rule at the kth base station can be generalized, in a logarithmic form, as follows

$$U_k = u_k \ \text{ if } \ t_{k,u_k} \leq \Psi_{1,0}^{(k)}(\mathbf{r}_k, \mathbf{h}_k) < t_{k,u_k+1} \tag{2.50}$$

where $\Psi_{1,0}^{(k)}(\mathbf{r}_k, \mathbf{h}_k)$ is defined in (2.25), $u_k = 0, \ldots, L-1$, $k = 1, \ldots, N_{BS}$, $t_{k,0} = -\infty$ and $t_{k,L} = \infty$. Hence, as the sub-optimum local decision rule proposed for the bandwidth efficient scheme presented in 2.2.1, the optimum local decision rule at the kth base station is defined by the $L-1$ thresholds contained in $\mathbf{t}_k = [t_{k,1}, \ldots, t_{k,L-1}]$ partitioning the LLR $\Psi_{1,0}^{(k)}(\mathbf{r}_k, \mathbf{h}_k)$. Furthermore, since only the statistics of the channel state are known at the MSC where the optimization takes place, as the bandwidth efficient scheme the cost function $J(\mathbf{t})$ used for the optimization of the thresholds contained in the vector $\mathbf{t} = [\mathbf{t}_1, \ldots, \mathbf{t}_{N_{BS}}]^T$ is $P_b(\mathbf{t})$ representing the average probability of bit error (2.3) as a function of the thresholds contained in the vector $\mathbf{t}$. In fact, what differentiates both schemes is that, in this case, the fusion rule is independent of the channel state. The optimization of the local detector thresholds is treated in more details in section 2.3 and an analytical expression is derived for $P_b(\mathbf{t})$ in Appendix B. It important to mention that, in this case, a closed form expression can be derived for $P_b(\mathbf{t})$, accelerating the optimization process.

2.3 Optimization of local detector thresholds

In the previous section, we designed SHDD schemes employing a fusion rule γ_0 and a set of local decision rules $\{\gamma_1, \ldots, \gamma_{N_{BS}}\}$, defined by the thresholds contained in the vector $\mathbf{t} = [\mathbf{t}_1, \ldots, \mathbf{t}_{N_{BS}}]^T$ where $\mathbf{t}_k = [t_{k,1}, \ldots, t_{k,L-1}]$. These thresholds must be adjusted in order to minimize the cost function $J(\mathbf{t})$ defined by the probability of bit error (2.3) or the conditional probability of bit error (2.5), depending on how often the thresholds are updated. Therefore, the SHDD schemes pose the following multivariate optimization problem

$$\mathbf{t}^* = \arg\min_{\mathbf{t} \in \mathcal{T}} J(\mathbf{t}) \tag{2.51}$$

where the set $\mathcal{T}$ contains all possible values of $\mathbf{t}$. The major difficulty in the implementation of the SHDD schemes is caused by this numerical optimization problem and more precisely by the fact that the cost function $J(\mathbf{t})$ is a nonlinear non-convex function of $\mathbf{t}$. Since this cost function may have many local minima, local optimization techniques do not guarantee to provide a global optimum solution. Until now, the optimization algorithms used to optimize distributed detection schemes related to communication applications were local optimization algorithm. In [11] for example, this optimization problem was solved by a numerical gradient descent based algorithm. In [13], a Gauss-Seidel[1] procedure was used to optimize the proposed distributed multiuser detection scheme.

In this work, we propose to solve the optimization problem using an improved Simulated Annealing (SA) algorithm called Adaptive Simulated Annealing (ASA). SA was originally developed in 1983 [21] as a technique to solve combinatorial optimization problems. This approach has been later extended to solve continuous global optimization problems with multivariate nonlinear non-convex cost functions. SA algorithms try to mimic the principles of thermodynamic that make metal freeze in a minimum energy crystalline structure, when it is cooled slowly enough (annealing process). The major advantage of SA over other local optimization algorithms is that it has the ability to avoid staying trapped in local minima. This ability is due in part to the fact that the algorithm employs a random search which not only accepts changes that decrease the cost function but also some changes that increase it, where

[1]The Gauss-Seidel procedure converges to a locally optimum solution for which the performance cannot be improved by changing only one of the decision rules.

the probability of taking such changes decreases with an artificial temperature parameter. It has been proved in [22] that by cooling the temperature according to an inverse-log law the SA algorithm converges in probability to a global optimum solution. However, the convergence of the algorithm can be too slow for many applications. In order to accelerate the convergence of the algorithm, a method known as Fast Annealing (FA) [23] was proposed which permits lowering the temperature exponentially faster, guaranteeing that the algorithm converges to a value close to the global minimum in a reasonable amount of time. However, FA still requires quite a lot of time to converge. An algorithm called Very Fast Simulated Re-annealing (VFSR)[17][24] [25] or Adaptive Simulated Annealing (ASA), which is exponentially faster than FA, was developed by A.L.Ingber, making SA a viable solution to global optimization. It is important to mention that a C language code implementation of the ASA algorithm has been made publicly available by Ingber since 1993[2]. In addition, a Matlab gateway routine to Ingber's ASA C code, called ASAMIN[3], is also publicly available via the World Wide Web and allows the ASA algorithm to be used directly in a Matlab environment as any Matlab function.

In this section, the ASA algorithm is first presented in details. Then, the convergence of the algorithm is discussed. The tuning of the ASA algorithm is then studied for the optimization of the SHDD schemes. Finally, some simplification assumptions are proposed to simplify the optimization process.

2.3.1 ASA algorithm

ASA is a global optimization algorithm designed to solve continuous optimization problems of the form

$$\mathbf{x}^* = \arg\min_{\mathbf{x}\in\mathcal{X}} J(\mathbf{x}), \tag{2.52}$$

where $\mathbf{x} = [x_1, \ldots, x_D]^T$ and $\mathcal{X} \subseteq \Re^D$ is a continuous domain defined by

$$Lo^{(i)} \le x_i \le Up^{(i)}, \quad 1 \le i \le D \tag{2.53}$$

[2]ASA C language code is available at http://www.ingber.com.

[3]ASAMIN mex-file is available at http://www.econ.ubc.ca/ssakata/public_html/software

and the inequality constraints

$$f_j(\mathbf{x}) \leq 0, \quad 1 \leq j \leq N. \tag{2.54}$$

Similarly to any SA algorithm designed to solve continuous optimization problems, the ASA algorithm performs a random walk through the D-dimensional domain $\mathcal{X}$, searching for the optimum solution $\mathbf{x}^*$. Also, as for any SA algorithm, the random walk is controlled by three main functions which are the generating function, the acceptance function and the cooling function. In general, a SA algorithm can be described as follows:

Step 0 A starting point $\mathbf{x}(0)$ is selected randomly from the domain $\mathcal{X}$

Step 1 Assuming $\mathbf{x}(t)$ is the tth visited point by the random walk, the generating function generates randomly a new candidate point $\mathbf{y}(t+1)$ from $\mathbf{x}(t)$.

Step 3 The acceptance function accepts or rejects the new candidate point $\mathbf{y}(t+1)$ by comparing the cost at the new candidate point $\mathbf{y}(t+1)$ and the last visited point $\mathbf{x}(t)$. It determines if the random walk stays still or moves to the candidate point $\mathbf{y}(t+1)$, such that $\mathbf{x}(t+1) = \mathbf{y}(t+1)$, when the candidate point is accepted, and $\mathbf{x}(t+1) = \mathbf{x}(t)$, when the candidate point is rejected.

Step 4 Assuming the point $\mathbf{x}^*(t)$ is the lowest observed cost value after t iterations, it is verified if $\mathbf{x}(t+1)$ is a new minimum such that $\mathbf{x}^*(t+1) = \mathbf{x}(t+1)$, if $J(\mathbf{x}(t+1)) \leq J(\mathbf{x}^*(t))$, and $\mathbf{x}^*(t+1) = \mathbf{x}^*(t)$ otherwise.

Step 5 The cooling function adjusts the temperatures controlling the random behavior of the walk insuring the convergence of the algorithm.

Step 6 It is verified if predefined criterion of convergence are satisfied. If the convergence criterion are not met, the time index t is incremented by 1 and the algorithm goes back to Step 1.

However, what differentiate the different implementations of the SA algorithm is how the generating function, the acceptance function and the cooling function are defined. A good overview of the different implementations of the SA algorithm can be found in [26].

One of the particularities of the ASA algorithm is that an artificial temperature is not only employed in the acceptance function, but also in the generating function. In order to distinguish the different temperature parameters, the temperature associated with the acceptance function is denoted as $T_{accept}(k_a)$ and the temperature associated with the ith dimension in the generating function is denoted as $T_{i,gen}(k_i)$, where k_a and k_i are the annealing time index associated with these artificial temperatures. It is important to mention that each temperature parameter is varied independently of the others allowing the ASA algorithm to adapt to the different sensitivity in each parameter dimension and to adapt to the current status of the cost function, through a process called re-annealing. In the next three sections, the generating function, the acceptance function and the cooling function used by the ASA algorithm are defined.

A. Generating function

The main task of the generating function in the random walk is to generate a new candidate point $\mathbf{y}(t+1)$ from the current point $\mathbf{x}(t)$. In the ASA algorithm, the ith component of the new candidate point $\mathbf{y}(t+1)$ is determined as follows

$$y_i(t+1) = x_i(t) + \Delta x_i(Up^{(i)} - Lo^{(i)}) \tag{2.55}$$

where $\Delta x_i \in [-1, 1]$ is a sample of a random variable with the following PDF,

$$g_i(\Delta x_i) = \frac{1}{2(|\Delta x_i| + T_{i,gen}(k_i)) \ln(1 + \frac{1}{T_{i,gen}(k_i)})}. \tag{2.56}$$

It is shown in [17] that the sample Δx_i can be generated as follows

$$\Delta x_i = sgn\left(\nu_i - \frac{1}{2}\right) T_{i,gen}(k_i) \left(\left(1 + \frac{1}{T_{i,gen}(k_i)}\right)^{|2\nu_i - 1|} - 1 \right), \tag{2.57}$$

where ν_i is a sample of a random variable uniformly distributed over the interval $[0, 1]$. The temperature $T_{i,gen}(k_i)$ controls the width and scale of the PDF (2.56). In fact, as can be seen from Fig. 2.3, at high temperature the random variable Δx_i is almost uniformly distributed over the interval $[-1, 1]$, favoring a global exploration of the domain $\mathcal{X}$. Then, as the temperature is gradually decreased by the cooling function, the PDF favorizes more and more a local exploration of the domain $\mathcal{X}$ by generating with higher probability new candidate points in the vicinity of the current point.

Finally, it can be noticed that equation (2.55) does not take in consideration the conditions set by the lower and upper bounds (2.53) and by the inequality constraints (2.54) when it generates the new candidate point $\mathbf{y}(t+1)$. Hence, new candidate points are generated by the generating function, using expression (2.55), until a candidate point that satisfies the conditions is generated.

B. Acceptance function

The main task of the acceptance function in the random walk is to determine if a new candidate point $\mathbf{y}(t+1)$ of poorer quality than the current point $\mathbf{x}(t)$ is accepted or rejected. In the ASA algorithm as in most SA algorithms, the acceptance function is the Metropolis acceptance function. The Metropolis acceptance function generates a sample $p \in [0, 1]$ of a uniformly distributed random variable and set

$$\mathbf{x}(t+1) = \begin{cases} \mathbf{y}(t+1) & \text{if } p \le A(\mathbf{y}(t+1), \mathbf{x}(t), T_{accept}(k_a)) \\ \mathbf{x}(t) & \text{if } p > A(\mathbf{y}(t+1), \mathbf{x}(t), T_{accept}(k_a)) \end{cases} \tag{2.58}$$

where

$$A(\mathbf{y}(t+1), \mathbf{x}(t), T_{accept}(k_a)) = \min\left\{1, \exp\left(\frac{-(J(\mathbf{y}(t+1)) - J(\mathbf{x}(t)))}{T_{accept}(k_a)}\right)\right\}. \tag{2.59}$$

Therefore, independently of the temperature $T_{accept}(k_a)$, the acceptance function always accepts a new candidate point $\mathbf{y}(t+1)$ that improves the cost value with respect to the current point $\mathbf{x}(t)$. However, new candidate points of poorer quality are only accepted with probability

$$P_{accept} = \exp\left(\frac{-(J(\mathbf{y}(t+1)) - J(\mathbf{x}(t)))}{T_{accept}(k_a)}\right), \tag{2.60}$$

which decreases with the temperature parameter $T_{accept}(k_a)$. In the final stage of the algorithm, $T_{accept}(k_a) \to 0$ and the probability of accepting a point of poorer quality is almost null such that the algorithm acts much more like a local optimization algorithm.

It is important to mention that the choice of acceptance function is much less important in the ASA algorithm then the choice of generating function. First, as

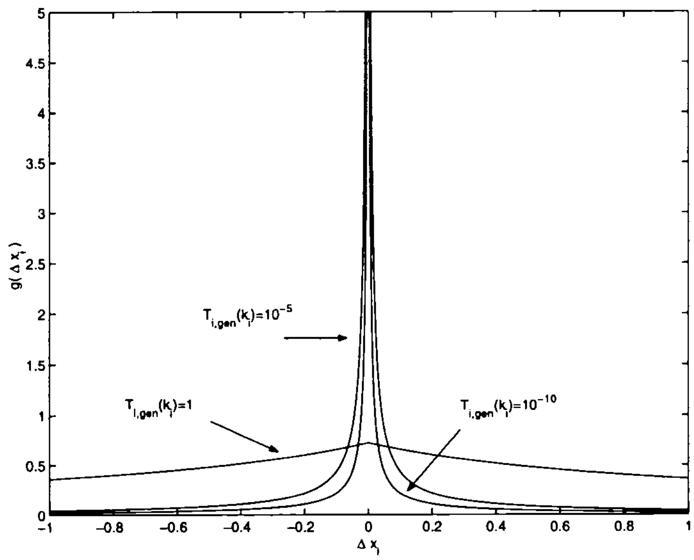


Fig. 2.3 PDF $g_i(\Delta x_i)$ for $T_{i,gen}(k_i) = 1$, $T_{i,gen}(k_i) = 10^{-5}$ and $T_{i,gen}(k_i) = 10^{-10}$

discussed in [26], most acceptance functions, which do not depend directly on $\mathbf{y}(t+1)$ and $\mathbf{x}(t)$ but depend on the difference of their cost values, can be substituted by the Metropolis acceptance function after a simple modification of the cooling function. Furthermore, the convergence of the ASA algorithm, as opposed to more conventional SA algorithms, is much less dependent on the acceptance function. This is due to the fact that the PDF used by the generating function allows for much wider displacements in the domain such that the algorithm does not depend as much on the acceptance function to escape from local minima.

C. Cooling function

The main task of the cooling function in the random walk is to gradually decrease (Annealing function) as well as periodically rescale (Re-annealing function) the temperature parameters associated with the acceptance function and the generating function. It is important to mention that the cooling schedules employed by the Annealing function must be chosen carefully in order to guarantee that the algorithm converges in probability to a globally optimum solution [17] . More precisely, in the ASA algorithm, the cooling schedules of the temperatures $T_{i,gen}(k_i)$ $i = 1, \ldots, D$, which control the width and scale of the distribution associated with the new candidate point, are crucial to guarantee the convergence of the ASA algorithm, as will be discussed in the next section treating the convergence of the ASA algorithm. Hence, in the ASA algorithm, the cooling schedule associated with the ith dimension of the generating function can be defined as follows

$$T_{i,gen}(k_i) = T_{i,gen}(0) \exp\left(-c_{gen} k_i^{\frac{1}{D}}\right), \tag{2.61}$$

where the temperatures $T_{i,gen}(0)$ is usually initially set to 1, c_{gen} must be adjusted by the user and k_i is incremented every time a new candidate point is generated. It is important to mention that, even if the algorithm converges statistically for any appropriate value of c_{gen}, practically the convergence of the algorithm is influenced by this parameter such that adjustments must be made by the user to tune the ASA algorithm to a specific problem. On the other hand, the cooling schedule of the temperature $T_{accept}(k_a)$ is less crucial to the convergence of the ASA algorithm. In fact, it is only required that [26]

$$\lim_{k_a \to \infty} T_{accept}(k_a) = 0. \tag{2.62}$$

Practically, the cooling schedule used by the ASA algorithm can be defined as follows,

$$T_{accept}(k_a) = T_{accept}(0) \exp\left(-c_a k_a^{\frac{1}{D}}\right) \tag{2.63}$$

where the temperature $T_{accept}(0)$ is usually initially set to $J(\mathbf{x}(0))$, c_a must be adjusted by the user and k_a is incremented every time a new candidate point is accepted.

As mentioned previously, one of the particularities of the ASA algorithm as opposed to other SA algorithms is that it allows for re-annealing. However, it is important to mention that the re-annealing process is not essential to the statistical convergence of the algorithm. It is in fact possible to use the ASA algorithm without re-annealing and it is sometime advantageous for some applications. However, it has been shown in [25] that the re-annealing process can in practice accelerate the convergence of the algorithm. The re-annealing process improves the convergence speed in two ways. First, the re-annealing process allows the ASA algorithm to adapt itself to the difference in sensitivity of the cost function in the different dimensions by periodically rescaling the generating function temperatures $T_{i,gen}(k_i)$ $i = 1, \ldots, D$. Furthermore, since the acceptance function depends on the difference of the cost values at $\mathbf{y}(t+1)$ and $\mathbf{x}(t)$ and the relative importance of a difference in cost value varies depending on the order of magnitude of the cost function in the region explored by the ASA algorithm, the re-annealing process allows the ASA algorithm to better suit the status of the cost function throughout the random walk by periodically rescaling the temperature $T_{accept}(k_a)$.

In the ASA algorithm the re-annealing process takes place every time r_{gen} points are generated or r_{accept} points are accepted, where recommended values are in the order of 10000 for r_{gen} and in the order of 100 for r_{accept}. When re-annealing is performed, the temperature $T_{i,gen}(k_i)$ is rescaled as follows

$$T_{i,gen}(k_i) = \frac{s_{i_{max}}}{s_i} T_{i,gen}(k_i), \tag{2.64}$$

where

$$s_i = \left| \left. \frac{d(J(\mathbf{x}))}{dx_i} \right|_{\mathbf{x}^*(t)} \right|, \tag{2.65}$$

and $i_{max} = \arg\max_{i=1,\ldots,D}\{s_i\}$. Then, the annealing index k_i is reset as follows

$$k_i = \left(-\frac{1}{c_{gen}} \ln \left(\frac{T_{i,gen}(k_i)}{T_{i,gen}(0)} \right) \right)^D. \tag{2.66}$$

In addition, the temperature $T_{accept}(0)$ is reset to the cost value at the last accepted point $\mathbf{x}(t)$, $T_{accept}(k_a)$ is reset to the cost value at the current minimum $\mathbf{x}^*(t)$ and then the annealing index k_a is reset as follows

$$k_a = \left(-\frac{1}{c_a} \ln \left(\frac{T_{accept}(k_a)}{T_{accept}(0)} \right) \right)^D. \tag{2.67}$$

2.3.2 Convergence of the ASA algorithm

In [17], it is proved that the ASA algorithm converges in probability to a global optimum solution as the number of iterations goes to infinity, meaning that,

$$\lim_{t\to\infty} Pr[\mathbf{x}^*(t) \in \mathcal{B}_\epsilon] = 1 \tag{2.68}$$

where $\mathcal{B}_\epsilon = \{\mathbf{x} \in \mathcal{X} : J(\mathbf{x}) \leq J(\mathbf{x}^*) + \epsilon\}$ for all $\epsilon > 0$, given $\mathbf{x}^*$ is the global minimum. The proof is based on the fact that by carefully controlling the rate of cooling of the temperatures associated with the new candidate point distribution, any subset $\mathcal{B}_\epsilon$ of the domain $\mathcal{X}$ with positive Lebesgue measure is visited infinitely often.

Practically it is sufficient to find a solution that approaches closely the global minimum $\mathbf{x}^*$ such that the search can be halted when it ceases to make sufficient progress. Lack of progress can be defined in a number of ways, where an overview of the different stopping rule found in the literature can be found in [26]. For example, in Ingber ASA C code, the search is stopped when

- $T_{i,gen}(k_i) \leq T_{gen}^{min}$ or $T_{accept}(k_a) \leq T_{accept}^{min}$
- the cost $J(\mathbf{x}(t))$ repeats at $N_{repetition}$ successive re-annealing time instants within a predefined precision Δ_{min}
- the ratio of accepted points and generated points is smaller than $t_{a/g}^{min}$
- the number of accepted new candidate points $\mathbf{y}(t)$ is larger than N_{accept}^{max}
- the number of generated new candidate points $\mathbf{y}(t)$ is larger than N_{gen}^{max}

where N_{accept}^{max}, N_{gen}^{max}, Δ_{min}, $t_{a/g}^{min}$, T_{gen}^{min}, T_{accept}^{min} and $N_{repetition}$ are parameters defined by the users. It is important to mention that it is almost impossible to select stopping rules that guarantee with a given probability that the global minimum has been detected within a certain accuracy. However, it is obvious that, longer the algorithm will run, better are the chance that a minimum in the vicinity of the global minimum $\mathbf{x}^*$ has been found.

2.3.3 Tuning of the ASA algorithm for the optimization of the SHDD schemes

In this section, we explain how to tune the ASA algorithm to optimize the thresholds defining the SHDD schemes. More precisely, three main points are discussed, which are the selection of the domain $\mathcal{X}$, the tuning of the parameter c_{gen}, and the tuning of the stopping rule parameters. We discuss in details these three points since they influence the convergence of the ASA algorithm.

The domain $\mathcal{T}$ contains all possible values of $\mathbf{t}$ that satisfy the inequalities $-\infty < t_{k,1} < t_{k,2} < \cdots < t_{k,L-2} < t_{k,L-1} < \infty$, where $k = 1, \ldots, N_{BS}$. However, the ASA algorithm is designed to work with a domain having a finite range. It is therefore necessary to select a domain $\mathcal{X}$ as a subset of $\mathcal{T}$ that still includes the global optimum solution but only contains the values of $\mathbf{t}$ satisfying the inequalities $-L_k < t_{k,1} < t_{k,2} < \cdots < t_{k,L-2} < t_{k,L-1} < L_k$, where $k = 1, \ldots, N_{BS}$. Obviously, the more the domain $\mathcal{X}$ can be constrained to a smaller region of the domain $\mathcal{T}$, the easier is the convergence of the ASA algorithm. The limit L_k must therefore be selected such that only plausible values of $t_{k,1}$ and $t_{k,L-1}$ are included in the domain $\mathcal{X}$. To this end, we use the fact that the probability of bit error of SHDD schemes using L quantization levels is upper bounded by the probability of bit error of SHDD schemes using $L-1$ quantization levels.

For SHDD schemes for which the cost function is the conditional probability of bit error (2.5), the value of L_k, for $\mathbf{H}_k = \mathbf{h}_k$, is selected to eliminate the region of the domain $\mathcal{T}$, where $P(U_k = 0|\mathbf{h}_k)$ and $P(U_k = L-1|\mathbf{h}_k)$ approach 0. However, it is also necessary to make sure that the domain $\mathcal{X}$ still includes the optimum solution. Hence, the value of L_k is selected such that the thresholds $t_{k,1}$ and $t_{k,L-1}$ be able to take values for which $P(U_k = 0|\mathbf{h}_k)$ and $P(U_k = L-1|\mathbf{h}_k)$ are at least as low as $\alpha = 10^{-10}$. To determine the values of L_k that satisfy this condition, we first derive an upper bound to $P(U_k = L-1|\mathbf{h}_k)$ as follows

$$\begin{aligned} P(U_k = L-1|\mathbf{h}_k) &= \frac{1}{2}P(U_k = L-1|\mathbf{h}_k, B=1) + \frac{1}{2}P(U_k = L-1|\mathbf{h}_k, B=0) \\ &< P(U_k = L-1|\mathbf{h}_k, B=1) \end{aligned} \tag{2.69}$$

where

$$P(U_k = L-1|\mathbf{h}_k, B) = \int_{t_{k,L-1}}^{\infty} f_{\Psi_{1,0}^{(k)}(\mathbf{R}_k,\mathbf{H}_k)}(x \mid \mathbf{h}_k, B)dx = 1 - F_{\Psi_{1,0}^{(k)}(\mathbf{R}_k,\mathbf{H}_k)}(t_{k,L-1} \mid \mathbf{h}_k, B), \tag{2.70}$$

given $f_{\Psi_{1,0}^{(k)}(\mathbf{R}_k,\mathbf{H}_k)}(x \mid \mathbf{h}_k, B)$ and $F_{\Psi_{1,0}^{(k)}(\mathbf{R}_k,\mathbf{H}_k)}(x \mid \mathbf{h}_k, B)$ are derived in Appendix A and represent, respectively, the PDF and cumulative distribution function (CDF) of the LLR $\Psi_{1,0}^{(k)}(\mathbf{R}_k, \mathbf{H}_k)$ given the transmitted bit B and $\mathbf{H}_k = \mathbf{h}_k$. Then, since

$$P(U_k = L-1|\mathbf{h}_k, B=1) = Q\left[\frac{t_{k,L-1}}{2\sqrt{2\frac{E_k}{N_0}\omega_k^2}} - \frac{\sqrt{2\frac{E_k}{N_0}\omega_k^2}}{1}\right], \tag{2.71}$$

it can be concluded that $P(U_k = L-1|\mathbf{h}_k, B=1) = \alpha$ and $P(U_k = L-1|\mathbf{h}_k) < \alpha$ if

$$t_{k,L-1} = 2\sqrt{2\frac{E_k}{N_0}\omega_k^2}\left(\sqrt{2}erfinv\left(-(2\alpha-1)\right) + \sqrt{2\frac{E_k}{N_0}\omega_k^2}\right), \tag{2.72}$$

where $\omega_k = \sqrt{\sum_{n=1}^{N_R}|h_{k,n}|^2}$. Hence, if for a given α and ω_k L_k equals to $t_{k,L-1}$ in (2.72), it is guaranteed that the limit L_k satisfies the condition set above. Fig.2.4 presents L_k as a function of $\bar{\omega}_k = \frac{E_k}{N_0}\omega_k^2$.

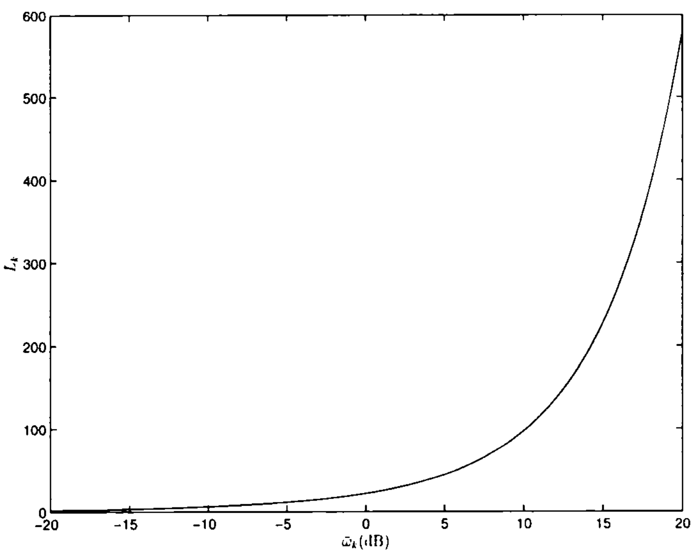


Fig. 2.4 Limit L_k as a function of $\bar{\omega}_k$, where $\alpha = 10^{-10}$

For SHDD schemes for which the cost function is the probability of bit error (2.3), it is possible to use the same strategy as for the previous case but replacing $P(U_k = 0|\mathbf{h}_k)$ and $P(U_k = L-1|\mathbf{h}_k)$ with $P(U_k = 0)$ and $P(U_k = L-1)$. However, the domain that results from this strategy is relatively large. A more efficient alternative uses the fact that the probability of bit error after the threshold optimization can be expressed as follows

$$P_b^* = \int_\omega P_{b|\omega}^* f_{\mathbf{\Omega}}(\omega)d\omega = \int_\omega I(\omega)d\omega \tag{2.73}$$

where $P_{b|\omega}^*$ represents the probability of bit error after the threshold optimization conditioned on $\mathbf{\Omega} = \omega$, $\mathbf{\Omega} = [\Omega_1, \ldots, \Omega_{N_{BS}}]^T$ given $\Omega_k = \sqrt{\sum_{n=1}^{N_R}|H_{k,n}|^2}$ and $f_{\mathbf{\Omega}}(\omega) = f_{\Omega_1}(\omega_1)\cdots f_{\Omega_{N_{BS}}}(\omega_{N_{BS}})$ given $f_{\Omega_k}(\omega_k)$ is the PDF of Ω_k defined in (B.8). Hence, the value of L_k can be selected such that the thresholds $t_{k,1}$ and $t_{k,L-1}$ be able to take values for which $P(U_k = 0|\mathbf{h}_k)$ and $P(U_k = L-1|\mathbf{h}_k)$ are at least as low as $\alpha = 10^{-10}$, for all values of ω_k for which the integrand $I(\omega)$ of (2.73) can possibly take a value that has an impact on the result of the integral within the desired precision. Since the probability $P(U_k = 0|\mathbf{h}_k)$ and $P(U_k = L-1|\mathbf{h}_k)$ are monotonically increasing functions of ω_k, it is first necessary to determine a value of ω_k over which the integrand $I(\omega)$ is lower than the desired precision independently of $\omega_1 \dots \omega_{k-1}, \omega_{k+1} \dots, \omega_{N_{BS}}$. Then, we use this value of ω_k in expression (2.72) to determine the limit L_k as in the previous case. It is obviously necessary to use an upper bound $U(\omega_k)$ to the integrand $I(\omega)$, independent of $\omega_1 \dots \omega_{k-1}, \omega_{k+1} \dots, \omega_{N_{BS}}$, since the optimum thresholds are not known a priori.

For instance, for a SHDD scheme for which the channel state vector $\mathbf{H}$ is known at the MSC and the thresholds are forced to be even symmetric, the conditional probability of bit error $P^*_{b|\omega}$ can be upper bounded by the conditional probability of bit error of a single base station scheme, which can be defined as follows

$$P_{s|\omega_k} = Q\left[\sqrt{2\frac{E_k}{N_0}\omega_k^2}\right]. \tag{2.74}$$

Hence, in this case, we use the upper bound

$$U(\omega_k) = P_{s|\omega_k} f_{\Omega_k}(\omega_k) \prod_{k' \neq k} \max_{\omega_{k'}} \left\{ f_{\Omega_{k'}}(\omega_{k'}) \right\}, \tag{2.75}$$

where $k' = 1, \dots, N_{BS}$. However if either the assumption that the thresholds are symmetric or the assumption that the channel state vector $\mathbf{H}$ is known at the MSC is removed, this upper bound is not valid and it becomes more difficult to find a relatively tight upper bound $U(\omega_k)$. Hence, in these cases, we use the fact that $P^*_{b|\omega} < 0.5$ and set

$$U(\omega_k) = 0.5 f_{\Omega_k}(\omega_k) \prod_{k' \neq k} \max_{\omega_{k'}} \left\{ f_{\Omega_{k'}}(\omega_{k'}) \right\}, \tag{2.76}$$

where $k' = 1, \dots, N_{BS}$.

The ASA algorithm also requires that the cooling schedule parameter c_{gen} be chosen carefully. In order to help its adjustment, the parameter c_{gen} is given by

$$c_{gen} = m_{gen} \exp\left(-\frac{n_{gen}}{D}\right), \tag{2.77}$$

where the parameters m_{gen} and n_{gen} can be adjusted according to the following relations

$$m_{gen} = -\ln\left(\frac{T(k_f)}{T(0)}\right) \tag{2.78}$$

and

$$n_{gen} = \ln(k_f). \tag{2.79}$$

The parameter $T(0)$ is the initial temperature, $T(k_f)$ is the desired final temperature and k_f is the desired number of annealing steps, assuming no re-annealing is performed. It is suggested that the initial temperature $T(0)$ be set to 1 and final temperature $T(k_f)$ be set to 10^{-10} although it can be necessary to adjust these parameters differently for certain problems. However, the value of k_f need to be selected experimentally. Table 2.1a) presents the values of k_f for which we obtain a good trade-off between efficiency and accuracy.

Finally, it is necessary to determine when to stop the ASA algorithm. In our case, the only stopping rule parameter that we consider is N_{gen}^{max} representing the maximum number of candidate points generated, since it is probably the more secure criterion. Given we know that the ASA algorithm is in the final stage of the search when the number of iterations is larger than k_f, we can expect relatively good results by setting N_{gen}^{max} equal to k_f but the accuracy can be improved by increasing the value of N_{gen}^{max}, with respect to k_f.

2.3.4 Simplification assumptions

The main disadvantage of the ASA algorithm as well as any global optimization algorithm is that it is computationally expensive, as can be seen from Table 2.1a). However, since the optimization can be performed off-line, the computational complexity has little influence on the actual scheme complexity. On the other hand, the complexity of the algorithm increases rapidly with the number of thresholds to be optimized, where the number of thresholds increases exponentially with the number of bits of resolution at each base station and linearly with the number of base stations. It may therefore be advantageous to add constraints on the thresholds in order

N_{BS}	L	D	k_f
2	2	2	$2.0 \cdot 10^3$
	4	6	$3.2 \cdot 10^4$
	8	14	$1.0 \cdot 10^6$
3	2	3	$4.0 \cdot 10^3$
	4	9	$2.6 \cdot 10^5$
	8	21	$2.0 \cdot 10^6$

a)

N_{BS}	L	D	k_f
2	2	0	
	4	2	$2.0 \cdot 10^3$
	8	6	$3.2 \cdot 10^4$
3	2	0	
	4	3	$4.0 \cdot 10^3$
	8	9	$2.6 \cdot 10^5$

b)

N_{BS}	L	D	k_f
2	2	0	
	4	1	$1.0 \cdot 10^3$
	8	3	$4.0 \cdot 10^3$
3	2	0	
	4	1	$1.0 \cdot 10^3$
	8	3	$4.0 \cdot 10^3$

c)

Table 2.1 Values of k_f used for the optimization of the different SHDD schemes a) when there is no constraint applied to the thresholds, b) when the thresholds are constrained to be even symmetric, c) when the thresholds are constrained to be even symmetric as well as identical at all base stations

to keep the optimization complexity to a reasonable level. More precisely, the local decision rules can be for example constrained to be even symmetric meaning that $P(U_k = l \mid B = j) = P(U_k = L - 1 - l \mid B = -j + 1)$, for $0 \leq l \leq L - 1$. When the local detector likelihood ratio is expressed in its logarithmic form, as shown in Fig. 2.5 the number of dimensions of the optimization problem reduces from $N_{BS}(L-1)$ to $N_{BS}\left(\frac{L-2}{2}\right)$. Hence, when $L = 2$, the number of dimensions is equal to 0 since the only threshold $t_{k,1}$ defining the kth local decision rule is forced to be equal to 0 in order for the decision rule to be even symmetric, avoiding any optimization. However, for systems using more than two quantization levels at each base station, the optimization complexity can be further reduced by forcing the thresholds to be equal at all base stations such that the number of dimensions of the optimization problem reduces to $\frac{L-2}{2}$. Since the number of cost function evaluations required by the ASA algorithm increases almost exponentially with the number of dimensions, the proposed constraints greatly simplify the optimization as can be seen in Table 2.1b) and Table 2.1c). On the other hand, it is important to mention that when any of the constraints is applied to the thresholds, the solution cannot be considered optimum anymore. The influence of these constraints on the performances of the system is evaluated for different system configurations in section 2.4.

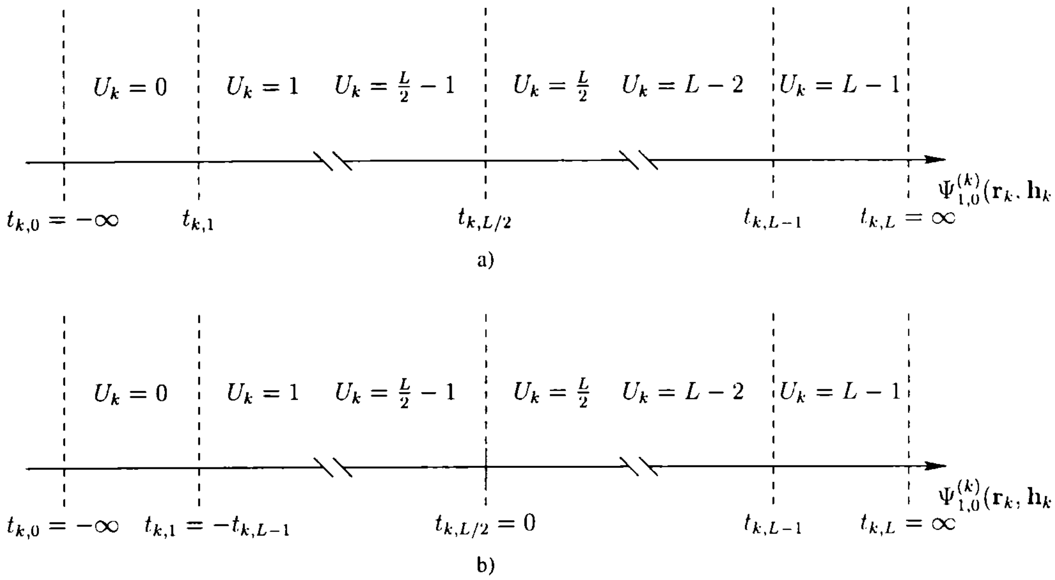


Fig. 2.5 Graphical representation of the kth base station local decision rule a) when there is no constraint applied to the thresholds, b) when the thresholds are constrained to be even symmetric

2.4 Numerical results

In this section, we study the performances of the designed SHDD schemes, in term of BER, for different numbers of base stations, different numbers of receiving antennas, different numbers of quantization levels and different types of constraints on the local detector thresholds. In addition, since the average SNR received at each base station is dependent on the mobile unit location in the cellular network, as well as on power control, we study the performances of the designed SHDD schemes for the case when the average received SNR is equal at all base stations as well as for the case when the average received SNR is different at each base station.

Before presenting the performances of the designed SHDD schemes, it is interesting to evaluate the potential gain that can be achieved by the SHDD schemes with respect to the CHM scheme. The potential gain is obtained by evaluating the difference between the performances of the CHM and OC schemes (see Appendix D). Fig 2.6 presents the potential gain in SNR achievable, when operating at $BER = 10^{-5}$, by 2 base stations SHDD schemes as a function of the difference in average SNR between the two base stations. Obviously, the potential gain decreases with an increase of the difference in average SNR between the two base stations and the number of receiving antennas at each base station. However, it is clear from this figure that important gains can be made by using more sophisticated techniques such as the proposed SHDD schemes.

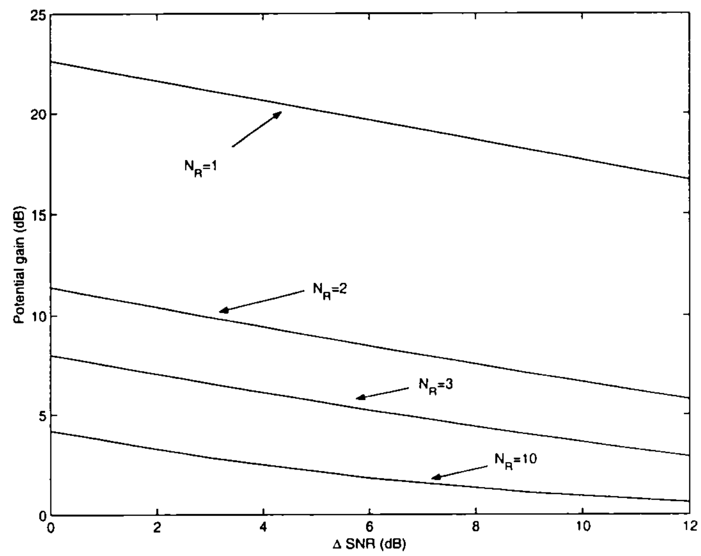


Fig. 2.6 Potential gain at $BER = 10^{-5}$ for 2 base stations SHDD schemes with respect to the CHM scheme as a function of $\Delta SNR = \frac{SNR_1}{SNR_2}$.

2.4.1 Known channel state information at the fusion center

In section 2.2.1, we have presented two alternative SHDD schemes for the case when the channel state is known at the MSC. The first is the optimum SHDD scheme when the channel state is known at the MSC referred as the $\text{SHDD}_{1,opt}$ scheme. As mentioned previously, such a scheme requires that the MSC transmits to each base station new threshold values every time the channel state varies. The second alternative is a bandwidth efficient sub-optimum SHDD scheme referred as the $\text{SHDD}_{1,sub}$ scheme. The only difference between the $\text{SHDD}_{1,opt}$ and $\text{SHDD}_{1,sub}$ schemes is that, in the $\text{SHDD}_{1,sub}$ scheme, the thresholds defining the local decision rules are independent of the channel state vector $\mathbf{H}$.

We evaluated numerically the BER of both the $\text{SHDD}_{1,opt}$ and $\text{SHDD}_{1,sub}$ schemes assuming the mobile unit is communicating simultaneously with 2 base stations. The evaluation of the probability of bit error for both schemes is discussed in Appendix B. In addition, the software used to obtain these results is included in the compact disk provided with this thesis. Results for BER as a function of the first base station average SNR, defined as follows $SNR_1 = \frac{E_1}{N_0}$, are respectively illustrated for the case when each base station is equipped with 1, 2 and 3 receiving antennas in Fig.2.7, Fig.2.8 and Fig.2.9. Each figure is made of two sub-figures where in part a) the average SNR is equal at both base stations while in part b) there is a difference of 6dB between the average SNR at the first and the second base station. It is important to mention that, in these figures, when the difference between two curves is less than $0.05dB$ at fixed BER, the two curves are plotted as one single curve on the figures.

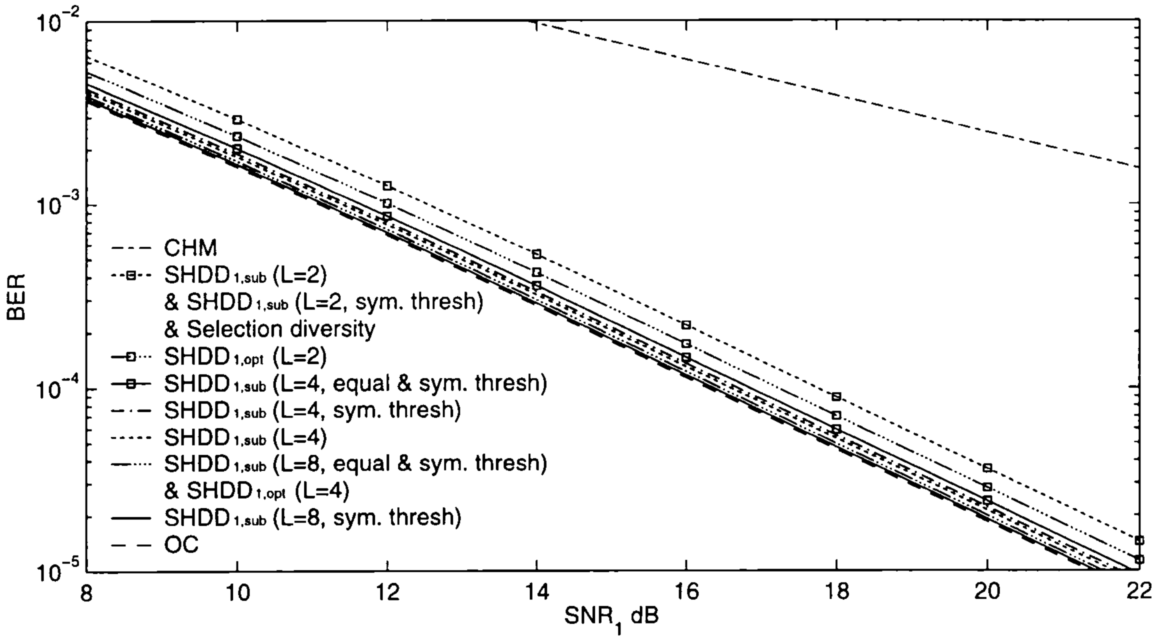
All figures present BER curves for $\text{SHDD}_{1,opt}$ and $\text{SHDD}_{1,sub}$ schemes using 2 and 4 quantization levels when there is no constraint applied on the thresholds. In addition, all figures also present BER curves for $\text{SHDD}_{1,sub}$ schemes using 2, 4 and 8 quantization levels for the case when the thresholds are constrained to be even symmetric and for the case when the thresholds are constrained to be even symmetric as well as identical at all base stations. For comparison purposes, all figures also include the BER curves of the OC scheme, the CHM scheme and a selection diversity scheme which assumes the channel state is known at the MSC. Appendix D presents these 3 reference schemes and discusses the evaluation of the probability of bit error for each scheme. It is important to remember that the BER of the OC scheme is obviously a lower bound to the BER of the $\text{SHDD}_{1,opt}$ and $\text{SHDD}_{1,sub}$ schemes while, as shown in Appendix D, the BER of the selection diversity and CHM schemes are upper bounds.

A. Effect of the number of receiving antennas

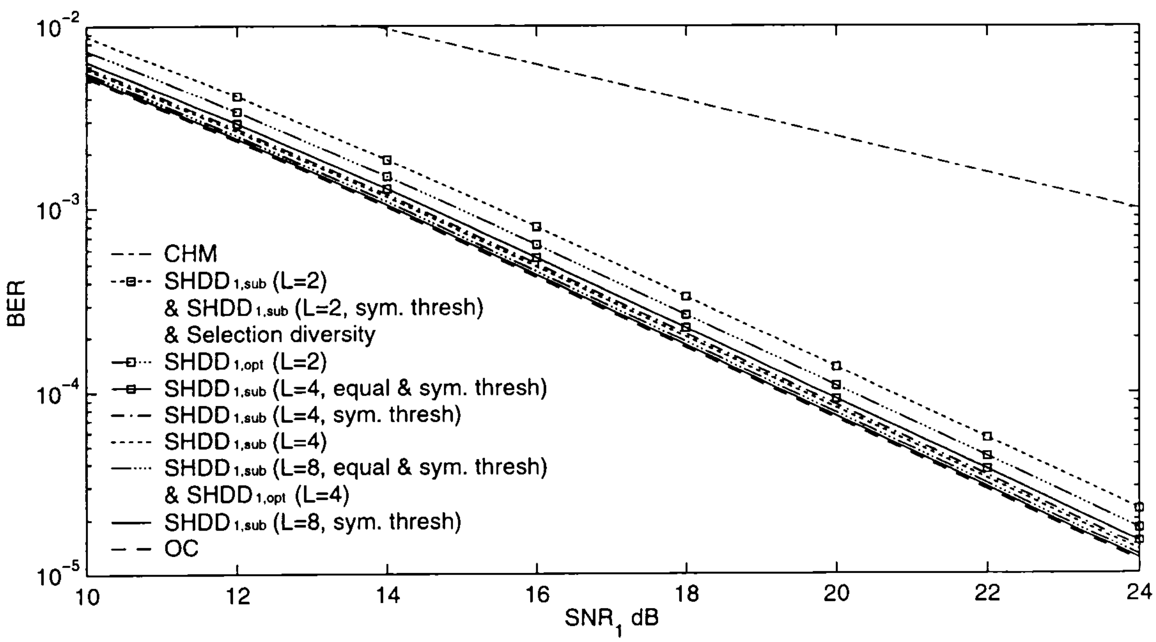
By comparing Fig.2.7, Fig.2.8 and Fig.2.9, we see that, as expected, increasing the number of receiving antennas increases the slope of the BER curves at large SNR and consequently the diversity order provided by the $\text{SHDD}_{1,opt}$ and $\text{SHDD}_{1,sub}$ schemes, which can be defined as

$$d = \frac{-10\log_{10}\left(\frac{BER^{(2)}}{BER^{(1)}}\right)}{\left(10\log_{10}\left(SNR_k^{(2)}\right) - 10\log_{10}\left(SNR_k^{(1)}\right)\right)}. \tag{2.80}$$

In fact, it appears that all considered $\text{SHDD}_{1,opt}$ and $\text{SHDD}_{1,sub}$ schemes provide the same asymptotic diversity order as the OC scheme, which equals approximately $N_{BS}N_R$. Hence, by increasing the number of receiving antennas per base station from

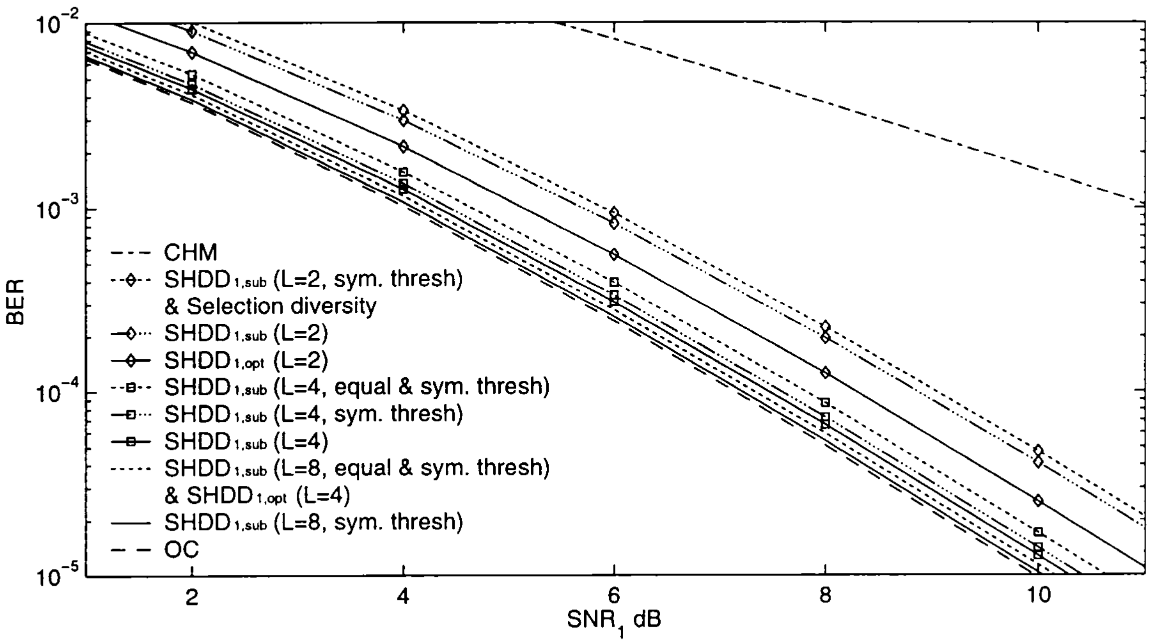


a)

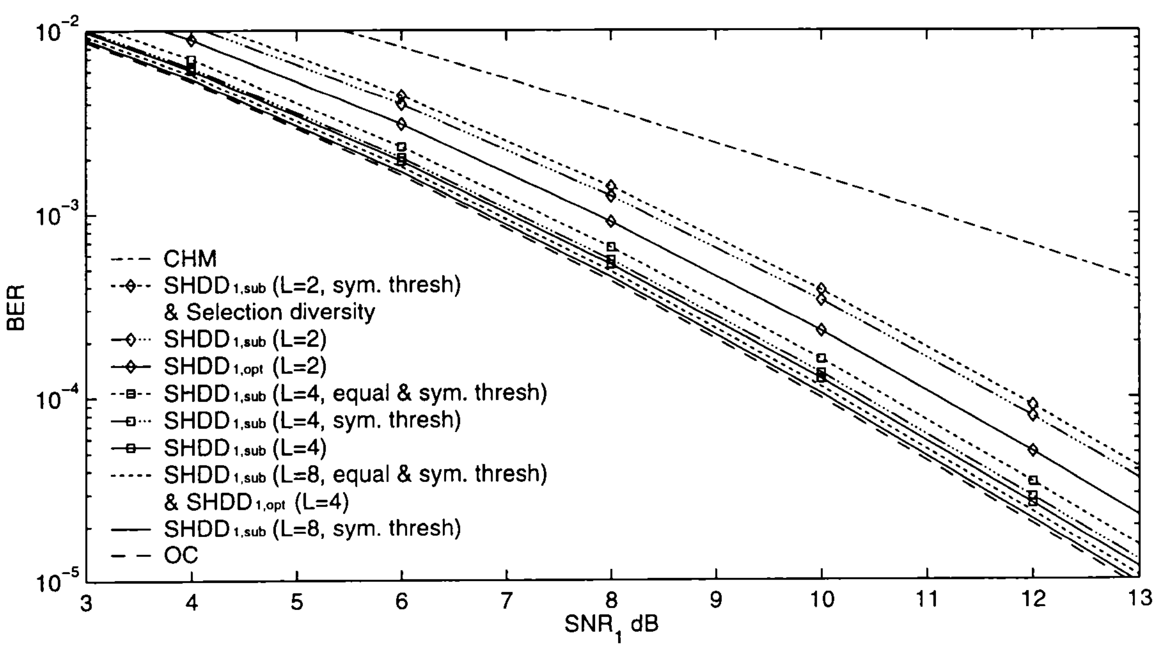


b)

Fig. 2.7 BER vs SNR_1: $\mathrm{SHDD}_{1,opt}$ and $\mathrm{SHDD}_{1,sub}$, $N_{BS}=2$, $N_R=1$, a) $SNR_1 = SNR_2$ b) $SNR_1 = 4SNR_2$

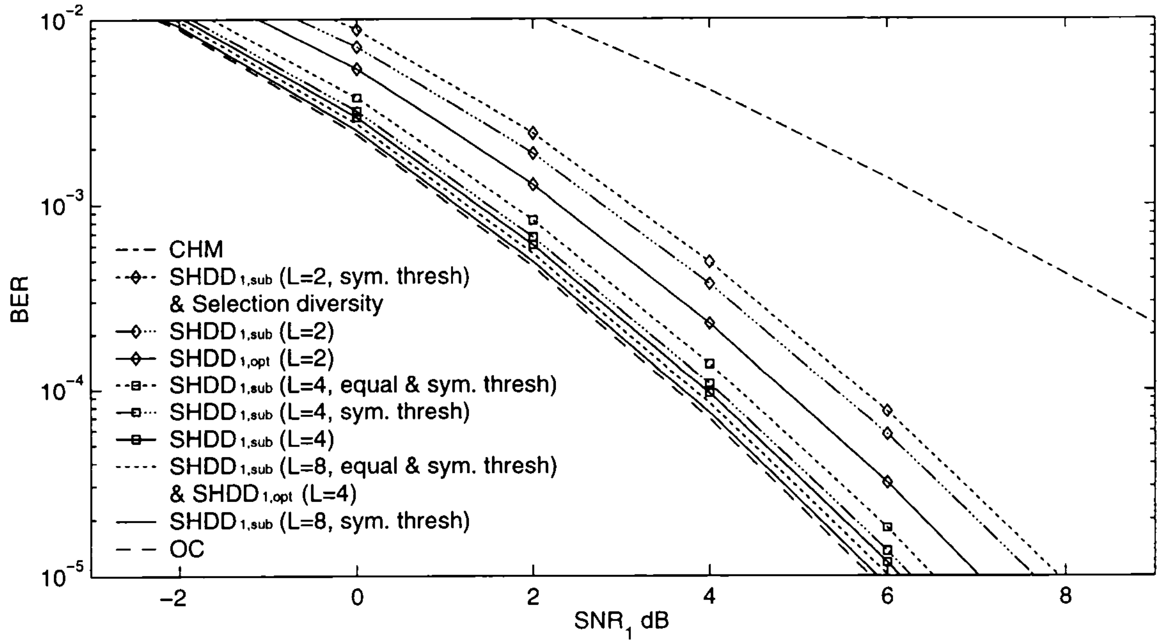


a)

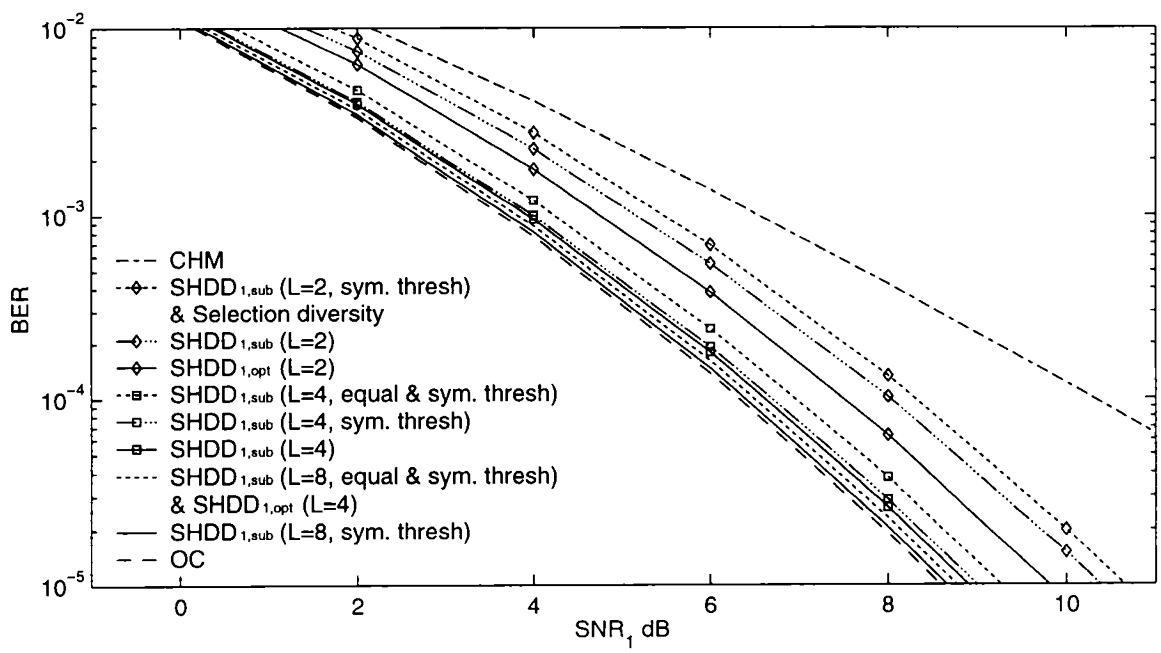


b)

Fig. 2.8 BER vs SNR_1: SHDD$_{1,opt}$ and SHDD$_{1,sub}$, $N_{BS} = 2$, $N_R = 2$, a) $SNR_1 = SNR_2$ b) $SNR_1 = 4SNR_2$



a)



b)

Fig. 2.9 BER vs SNR_1: $\mathrm{SHDD}_{1,opt}$ and $\mathrm{SHDD}_{1,sub}$, $N_{BS} = 2$, $N_R = 3$, a) $SNR_1 = SNR_2$ b) $SNR_1 = 4SNR_2$

1 to 2, a gain in SNR of approximately 11.0-11.5dB can be observed at $BER = 10^{-5}$ while the gain, obtained by increasing the number of receiving antennas per base station from 2 to 3, reduces to approximately 4.0-4.3dB. The gain reduction is due to the fact that as the diversity order increases the gain obtained by further increasing the diversity order decreases.

It is important to mention that the diversity order of the CHM scheme equals only to N_R, since this scheme selects a base station based on the average SNR and is therefore not taking advantage of the diversity provided by the remaining base stations. The diversity order of the $\text{SHDD}_{1,opt}$ and $\text{SHDD}_{1,sub}$ schemes is therefore N_{BS} times larger than the diversity order of the CHM scheme, independently of the number of receiving antennas. For this reason, the $\text{SHDD}_{1,opt}$ and $\text{SHDD}_{1,sub}$ schemes provide important gains with respect to CHM scheme. For instance, from part a) of Fig. 2.7, it appears that, when the SNR is equal at both base stations and each base station is equipped with a single receiving antenna, the $\text{SHDD}_{1,opt}$ and $\text{SHDD}_{1,sub}$ schemes provide respectively gains in SNR with respect to the CHM scheme of 21.7dB and 21.2dB at $BER = 10^{-5}$, with only 2 levels of quantization, which increase to 22.5dB and 22.3dB, with 4 levels of quantization. However, as mentioned previously, as the diversity order increases, the gain obtained by further increasing the diversity order decreases, such that the gains obtained by the $\text{SHDD}_{1,opt}$ and $\text{SHDD}_{1,sub}$ schemes with respect to the CHM scheme decrease with the number of receiving antennas per base station. From part a) of Fig. 2.8, it appears that, when each base station is equipped with 2 receiving antennas, the gains in SNR with respect to the CHM scheme reduce respectively to 10.2dB and 9.6dB at $BER = 10^{-5}$, with 2 levels of quantization, and to 11.2dB and 11.1dB, with 4 levels of quantization. Furthermore, from part a) of Fig. 2.9, it appears that, when each base station is equipped with 3 receiving antennas, the gains in SNR with respect to the CHM scheme reduce respectively to 6.8dB and 6.2dB at $BER = 10^{-5}$, with 2 levels of quantization, and to 7.8dB and 7.7dB, with 4 levels of quantization. Hence, from these results it can be concluded that, even if each base station is equipped with 3 receiving antennas, important gains with respect to the CHM scheme can still be obtained by using handoff macrodiversity schemes that further increase the diversity order such as the $\text{SHDD}_{1,opt}$ and $\text{SHDD}_{1,sub}$ schemes.

B. Effect of the difference in average SNR between two base stations

By comparing part a) to part b) of Fig.2.7, Fig.2.8 and Fig.2.9, we see that, when the average SNR at the second base station is lower by 6dB from the SNR at the first base station, it has the effect of shifting horizontally the BER curves of the $SHDD_{1,opt}$ and $SHDD_{1,sub}$ schemes, obtained when the average SNR is equal at both base stations, by approximately 2.7-3.0dB toward the BER curves of the CHM scheme, independently of the number of receiving antennas. However, it is important to mention that, even if a difference of 6dB in the average SNR is not sufficient to affect the diversity order of the scheme, it is expected that, as the difference in average SNR increases, the diversity order of the $SHDD_{1,opt}$ and $SHDD_{1,sub}$ schemes will eventually tend toward the diversity order of the CHM scheme. This is due to the fact that the performances of the $SHDD_{1,opt}$ and $SHDD_{1,sub}$ schemes, as the OC scheme, have to converge to the performances of the CHM scheme when the asymmetry between the two links is very large.

C. Effect of the number of quantization levels

From Fig.2.7 - Fig.2.9, we see that in all considered cases most of the potential gain with respect to the CHM scheme is reached by $SHDD_{1,opt}$ and $SHDD_{1,sub}$ schemes using only 2 quantization levels. However, additional gains can still be obtained by increasing the number of quantization levels. In fact, if the number of quantization levels is increased from 2 to 4, the BER curves are shifted horizontally toward the OC scheme by approximately 0.8-1.0dB, for the $SHDD_{1,opt}$ scheme, and 1.2-1.5dB, for the $SHDD_{1,sub}$ scheme. Hence, with only 4 levels of quantization at each base station, the potential additional gain that can be obtained by further increasing the number of quantization levels equals approximately 0.3dB, for the $SHDD_{1,sub}$ scheme, and approximately 0.2dB, for the $SHDD_{1,opt}$ scheme. Furthermore, using the BER of the $SHDD_{1,sub}$ scheme when the thresholds are symmetric as an upper bound to the BER of the $SHDD_{1,opt}$ and $SHDD_{1,sub}$ schemes, the potential additional gain that can be obtained by further increasing the number of quantization levels of $SHDD_{1,opt}$ and $SHDD_{1,sub}$ schemes, using 8 quantization levels, can be upper bounded to less than 0.1dB for all considered cases. It can be concluded that no practical gains can be obtained by increasing the number of quantization levels over 8.

D. Comparison between the $\text{SHDD}_{1,opt}$ and $\text{SHDD}_{1,sub}$ schemes

From Fig.2.7 - Fig.2.9, we see that, when operating at a fixed BER, in all considered cases the difference in SNR between the $\text{SHDD}_{1,opt}$ and $\text{SHDD}_{1,sub}$ schemes decreases from approximately 0.5-0.6dB, when $L = 2$, to approximately 0.1-0.2dB, when $L = 4$. Furthermore, when $L = 8$ the SNR difference at fixed BER between the $\text{SHDD}_{1,opt}$ and $\text{SHDD}_{1,sub}$ schemes can be upper bounded to approximately 0.1dB, the SNR difference between the OC scheme and the $\text{SHDD}_{1,sub}$ scheme when the local detector thresholds are forced to be even symmetric which are respectively a lower and upper bound to the BER of both schemes. Hence, when the number of quantization levels is larger or equal to 4, the performances of the $\text{SHDD}_{1,opt}$ and $\text{SHDD}_{1,sub}$ schemes are almost identical, proving that the $\text{SHDD}_{1,sub}$ scheme is a viable alternative to the $\text{SHDD}_{1,opt}$ scheme.

E. Effect of the threshold constraints on the $\text{SHDD}_{1,sub}$ scheme

From Fig.2.7 - Fig.2.9, we see that, in all considered cases, the $\text{SHDD}_{1,sub}$ scheme is not severely affected by the constraints applied on the local detector thresholds, although they significantly reduce the complexity of the optimization process. In fact, when $L = 2$ the losses caused by forcing the thresholds to be symmetric vary from within the precision of the results when $N_R = 1$, to approximately 0.2dB, when $N_R = 2$, and to approximately 0.3dB, when $N_R = 3$. For $L = 4$, the losses, caused by forcing the thresholds to be symmetric, equal approximately 0.1dB while the losses, caused by forcing the thresholds to be symmetry and equal at both base stations, equal approximately 0.3-0.4dB. For $L = 8$, the losses caused by forcing the thresholds to be symmetric can be upper bounded to 0.1dB, while the losses caused by forcing the thresholds to be symmetric and equal at both base stations can be upper bounded to 0.2dB, using the OC scheme as a lower bound to the BER of the $\text{SHDD}_{1,sub}$ scheme with no constraints on the thresholds.

It is interesting to mention that, when $L = 2$ and the thresholds are even symmetric, the local decisions are locally optimum decision rules meaning that they minimize the probability of bit error on the local decisions as opposed to globally optimum decision rules which minimize the probability of bit error at the output of the MSC. Furthermore, if the $\text{SHDD}_{1,sub}$ scheme involves only 2 base stations and is using locally optimum decision rules, the scheme is equivalent to the selection diversity scheme since, as shown in Appendix D, the fusion rule selects the local decision of the base station for which the probability of bit error on the local decision is the lowest. Finally, for any number of quantization levels, the losses caused by forcing the thresholds to be symmetric and equal at both base stations do not seem to increase when there is a difference of 6dB between the average SNR of the two base stations.

2.4.2 Unknown channel state information at the fusion center

In section 2.2.2, we have presented the optimum SHDD scheme for the case when the channel state is unknown at the MSC referred as the $\text{SHDD}_{2,opt}$ scheme. Similarly to the $\text{SHDD}_{1,sub}$ scheme, the local decision rules of the $\text{SHDD}_{2,opt}$ scheme are likelihood ratio quantizers for which the thresholds are independent of the channel state vector $\mathbf{H}$. However, as opposed to the $\text{SHDD}_{1,sub}$ scheme, the fusion rule of the $\text{SHDD}_{2,opt}$ is independent of the channel state vector $\mathbf{H}$.

We evaluated numerically the BER of the $\text{SHDD}_{2,opt}$ scheme for the case when the mobile unit is communicating with 2 base stations as well as for the case when the mobile unit is communicating with 3 base stations. The evaluation of the probability of bit error is discussed in Appendix B. In addition, the software used to obtain these results is included in the compact disk provided with this thesis. Results for BER as a function of the first base station average SNR are illustrated in Fig.2.10, Fig.2.11 and Fig.2.12 for 2 base stations equipped with 1, 2 and 3 receiving antennas respectively. Similarly Fig.2.13 and Fig.2.14 present results for 3 base stations equipped with 1 and 2 receiving antennas respectively. Each figure is made of two sub-figures where in part a) the average SNR is equal at both base stations while in part b) there is a difference of 6dB between the average SNR at the first and the second base station. It is important to mention that, in these figures, when the difference between two curves is less than $0.05dB$ at fixed BER, the two curves are plotted as one single curve on the figures.

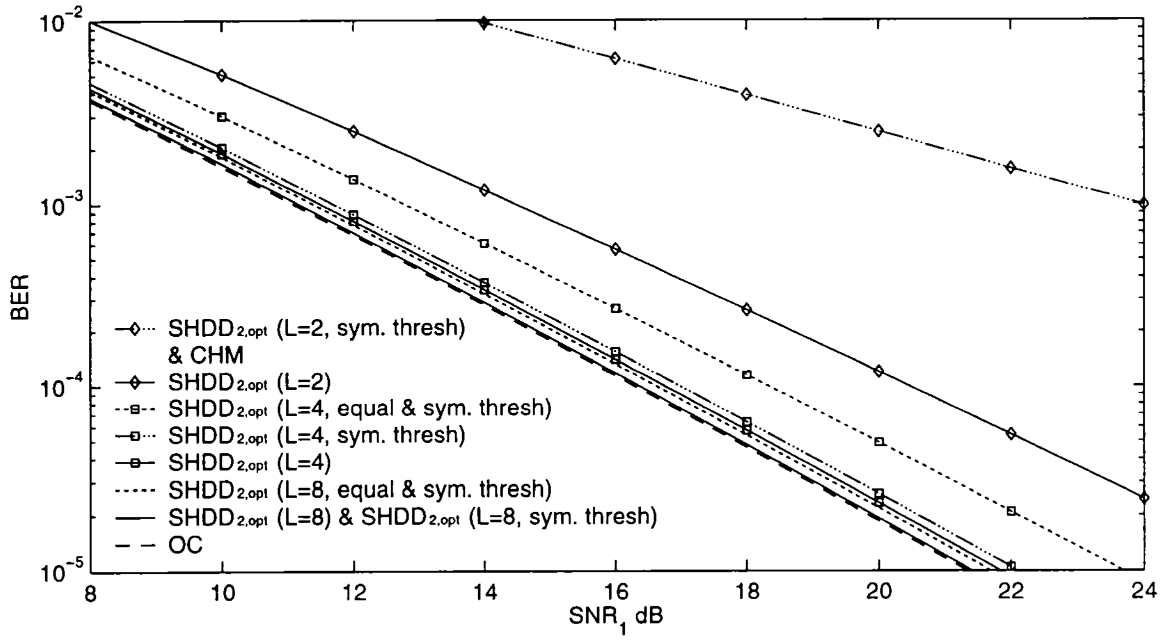
All figures present BER curves for $\text{SHDD}_{2,opt}$ schemes using 2, 4 and 8 quantization levels for the case when there is no constraint applied on the thresholds, for the case when the thresholds are constrained to be even symmetric and for the case when the thresholds are constrained to be even symmetric as well as identical at all base stations. All figures also include for comparison purposes the BER curves of the OC and the CHM schemes. Appendix D presents these 2 reference schemes and discusses the evaluation of the probability of bit error for each scheme. It is important to remember that the BER of the OC scheme is obviously a lower bound to the BER of the $\mathrm{SHDD}_{2,opt}$ schemes while the BER of the CHM scheme is an upper bound.

A. Effect of the number of receiving antennas and the number of base stations

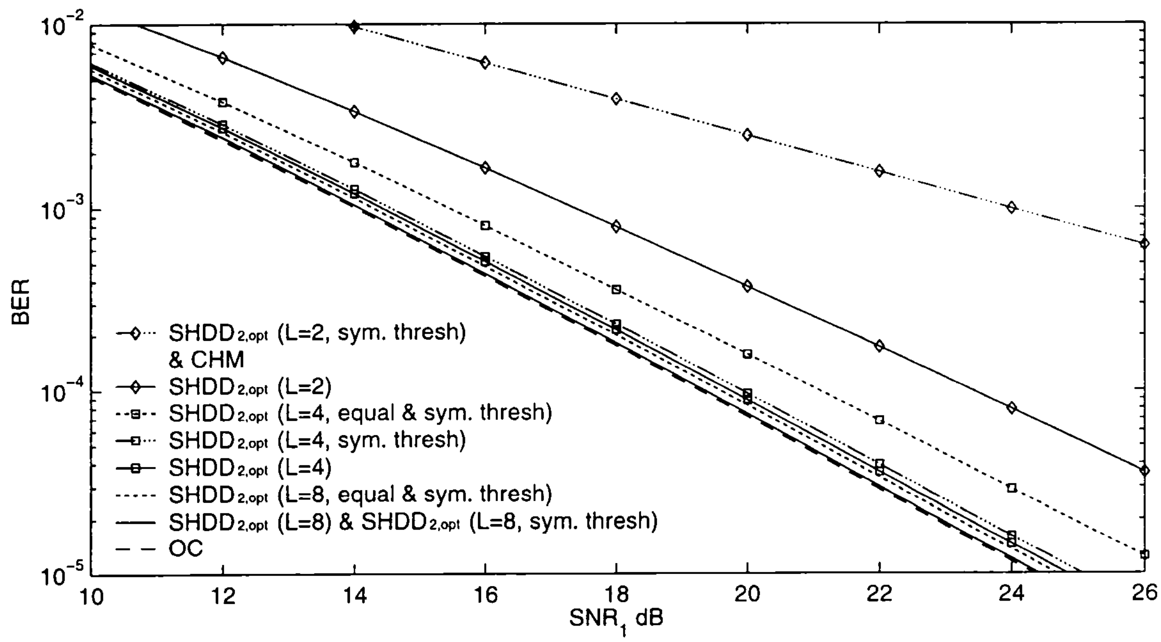
From Fig.2.10 - Fig.2.14, we see that, as expected, increasing the number of base stations or the number of receiving antennas per base station increases the diversity order of the $\mathrm{SHDD}_{2,opt}$ scheme, as defined in (2.80). However, as opposed to the $\mathrm{SHDD}_{1,opt}$ and $\mathrm{SHDD}_{1,sub}$ schemes, it appears that the diversity order provided by the $\mathrm{SHDD}_{2,opt}$ scheme is not equal to the diversity order provided by the OC scheme, which equals $N_{BS}N_R$, although the difference is not dramatic. Furthermore, the diversity order of the $\mathrm{SHDD}_{2,opt}$ scheme seems to increase as the number of quantization levels increases such that, for schemes using 4 or 8 quantization levels, the difference with respect to the diversity order of the OC scheme is only minor.

Hence, by comparing Fig.2.10, Fig.2.11 and Fig.2.12, it appears that, when 2 base station are involved in the handoff macrodiversity scheme, the gains obtained by increasing the number of receiving antennas of the $\mathrm{SHDD}_{2,opt}$ scheme are similar to the gains observed for the $\mathrm{SHDD}_{1,opt}$ and $\mathrm{SHDD}_{1,sub}$ schemes, especially for schemes using 4 and 8 quantization levels. Furthermore, by comparing Fig.2.13 and Fig.2.14, a gain of approximately 8.0-9.6dB, at $BER = 10^{-5}$, can be observed when the number of receiving antennas per base station is increased from 1 to 2 in a 3 base station $\mathrm{SHDD}_{2,opt}$ scheme.

Similarly to the $\mathrm{SHDD}_{1,opt}$ and $\mathrm{SHDD}_{1,sub}$ schemes, the diversity order provided by the $\mathrm{SHDD}_{2,opt}$ scheme is approximately N_{BS} time larger than the diversity order of the CHM scheme, independently of the number of receiving antennas. For this reason, the $\mathrm{SHDD}_{2,opt}$ scheme also provides important gains with respect to the CHM scheme. For instance, from part a) of Fig.2.10, it appears that, when the SNR is equal at both base stations and each base station is equipped with a single receiving antenna, the $\mathrm{SHDD}_{2,opt}$ scheme provides a gain in SNR with respect to the CHM scheme of 17.8dB at $BER = 10^{-5}$, with only 2 levels of quantization, which increases to $22.1dB$, with 4 levels of quantization, and $22.6dB$, with 8 levels of quantization. However, as for the $\mathrm{SHDD}_{1,opt}$ and $\mathrm{SHDD}_{1,sub}$ schemes, the gains obtained by the $\mathrm{SHDD}_{2,opt}$ scheme with

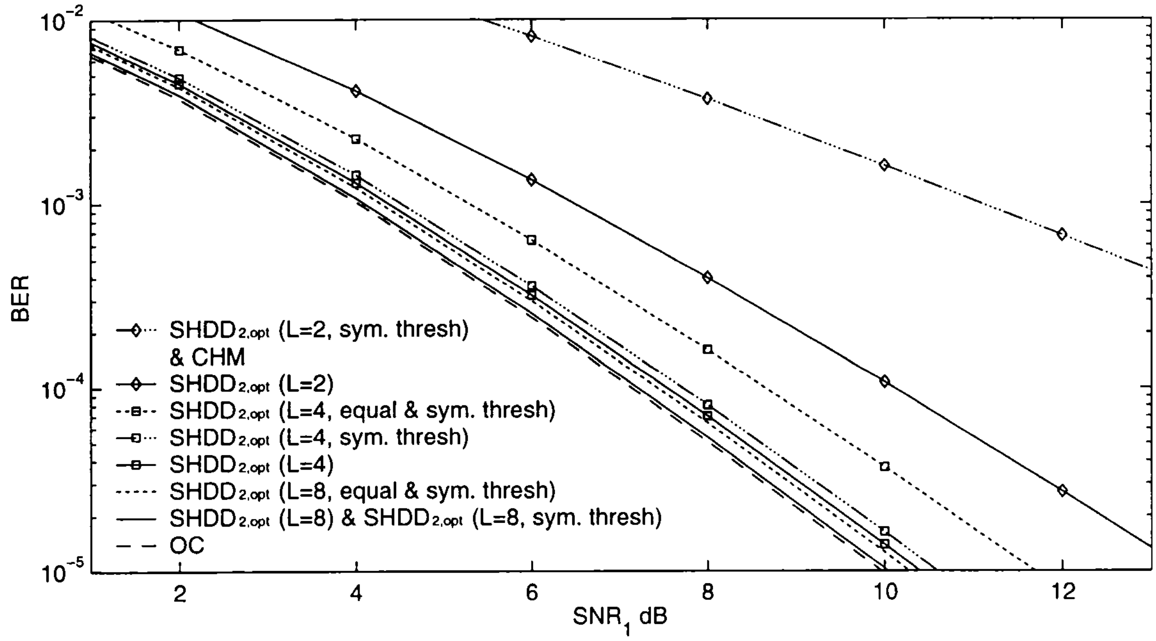


a)

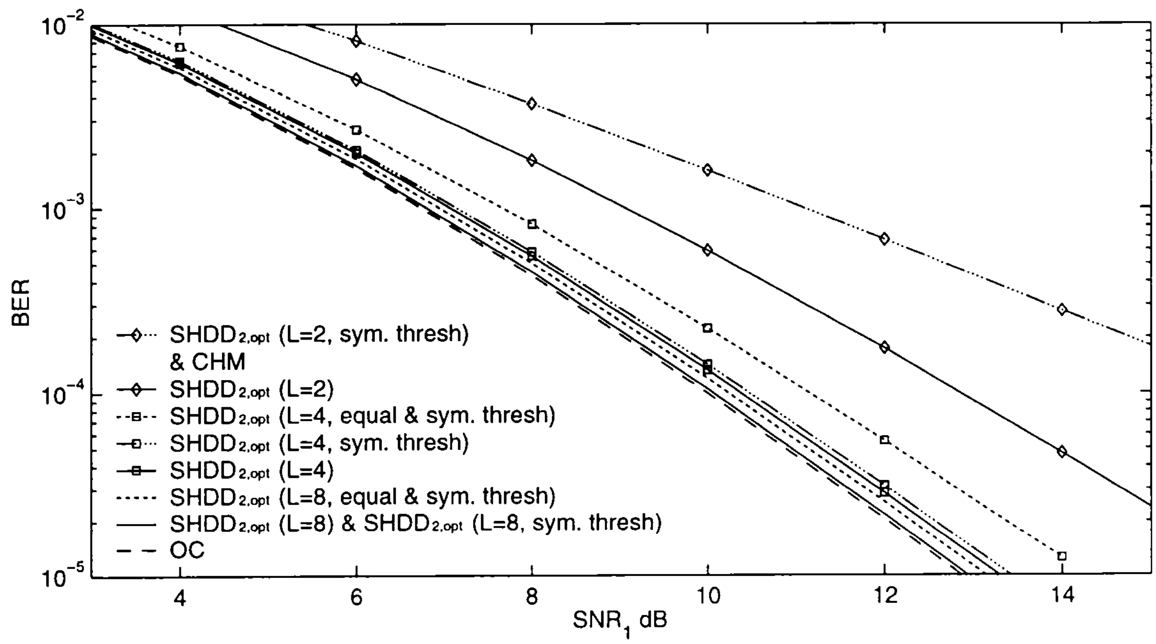


b)

Fig. 2.10 BER vs SNR_1: SHDD$_{2,opt}$, $N_{BS} = 2$, $N_R = 1$, a) $SNR_1 = SNR_2$ b) $SNR_1 = 4SNR_2$

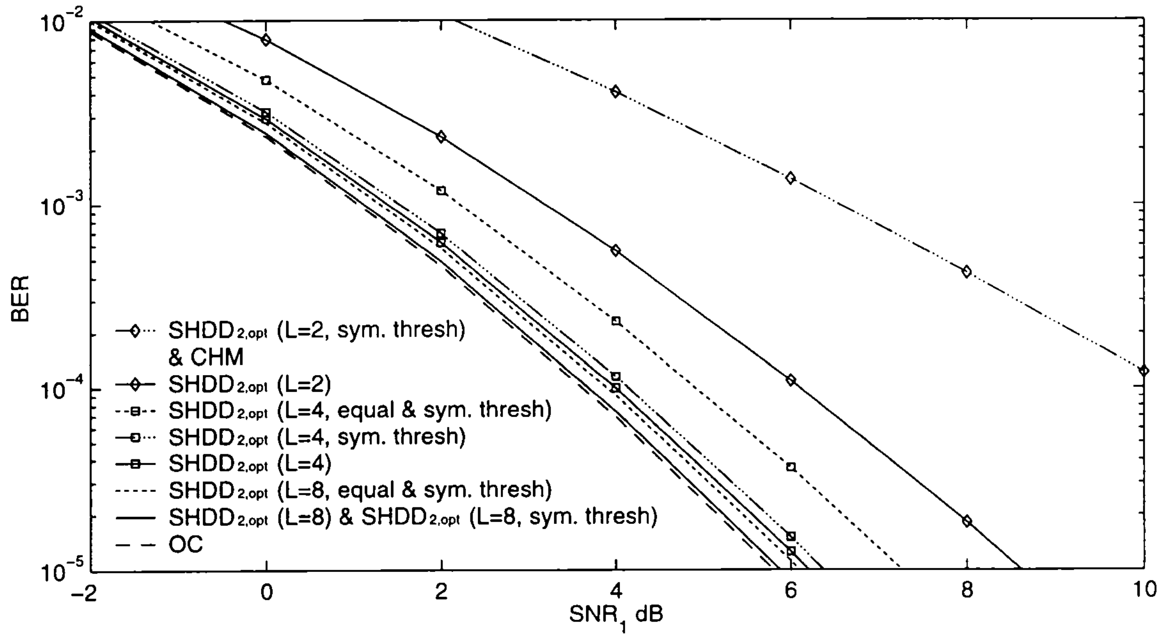


a)

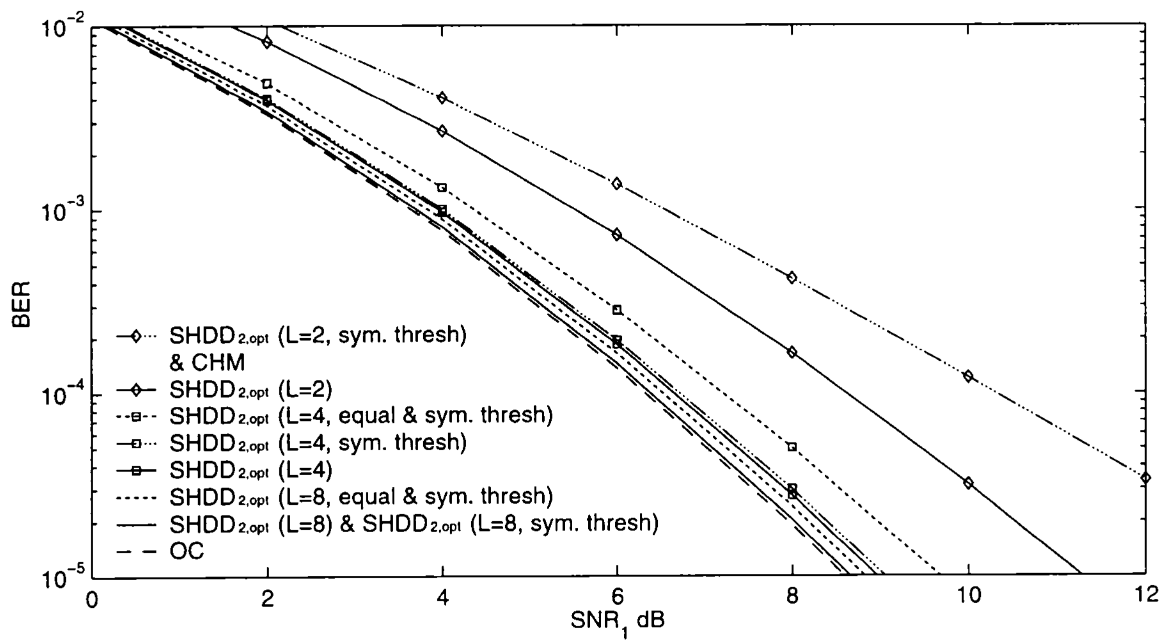


b)

Fig. 2.11 BER vs SNR_1: SHDD$_{2,opt}$, $N_{BS} = 2$, $N_R = 2$, a) $SNR_1 = SNR_2$ b) $SNR_1 = 4SNR_2$



a)



b)

Fig. 2.12 BER vs SNR_1: SHDD$_{2,opt}$, $N_{BS} = 2$, $N_R = 3$, a) $SNR_1 = SNR_2$ b) $SNR_1 = 4SNR_2$

respect to the CHM scheme decrease with the number of receiving antennas per base station. From Fig.2.11a), it appears that, when each base station is equipped with 2 receiving antennas, the gain in SNR, at $BER = 10^{-5}$, reduces to 7.9dB, with 2 levels of quantization, to 11.0dB, with 4 levels of quantization, and to 11.3dB, with 8 levels of quantization. Furthermore, from part a) of Fig.2.12, it appears that, when each base station is equipped with 3 receiving antennas, the gain in SNR with respect to the CHM scheme reduces to 5.2dB at $BER = 10^{-5}$, with 2 levels of quantization, to 7.8dB, with 4 levels of quantization, and 7.9dB, with 8 levels of quantization.

Furthermore, since the diversity order of the CHM scheme is independent of the number of base stations, additional gains with respect to the CHM scheme can be obtained by increasing the number of base stations involved in the scheme. In fact, by comparing Fig.2.10 and Fig.2.11 with Fig.2.13 and Fig.2.14, we see that the additional gain obtained by increasing the number of base stations from 2 to 3 varies from approximately 7.0-7.4dB, when $N_R = 1$, to approximately 3.7-4.1dB, when $N_R = 2$.

Finally, it is interesting to notice that more gain is obtained by increasing the number of receiving antennas of a two base stations $\text{SHDD}_{2,opt}$ scheme from 2 to 3 than increasing the number of base stations of the same scheme from 2 to 3, although both schemes provide the same diversity order. In fact, the difference decreases with the number of quantization levels and varies from up to 1.1dB, when $L = 2$, to 0.4dB, when $L = 4$, and 0.1dB, when $L = 8$. Hence, it seems that, in this case, it is more advantageous to provide to the fusion center fewer local decisions of better quality than more of worst quality.

B. Effect of the difference in average SNR between base stations

By comparing part a) to part b) of Fig.2.10, Fig.2.11 and Fig.2.12, it appears that, similarly to the $\text{SHDD}_{1,opt}$ and the $\text{SHDD}_{1,sub}$ schemes, the BER curves of a $\text{SHDD}_{2,opt}$ scheme using 2 base stations are shifted horizontally by approximately 2.7-3.1dB toward the BER curves of the CHM scheme, when the average SNR at the second base station is lowered by 6dB from the SNR at the first base station. It is important to mention that the BER curves of the OC scheme are also affected similarly by the asymmetry in the two links, such that the losses with respect to the OC scheme appear independent of this factor.

C. Effect of the number of quantization

From Fig.2.10 - Fig.2.14, we see that in all considered cases a large part of the potential gain that can be obtained by the SHDD$_{2,opt}$ scheme with respect to the CHM scheme is reached with only 2 quantization levels although less than the SHDD$_{1,opt}$ and SHDD$_{1,sub}$ schemes. However, additional gains can still be obtained by increasing the number of quantization levels.

More precisely, from Fig.2.10 - Fig.2.12, it appears that for the 2 base station SHDD$_{2,opt}$ scheme, if the number of quantization levels is increase from 2 to 4, the performance is improved at $BER = 10^{-5}$ by 4.4dB, when $N_R = 1$, by 3.0dB, when $N_R = 2$, and by 2.4dB, when $N_R = 3$. Hence, with only 4 levels of quantization at each base station, the difference between the BER curves of the OC scheme and the 2 base station SHDD$_{2,opt}$ scheme equals approximately 0.4-0.5dB, for all considered cases. The gain obtained by further increasing the number of quantization levels from 4 to 8 equals approximately 0.3-0.4dB. Hence, when $L = 8$ the difference in SNR, at fixed BER, between the OC scheme and the 2 base station SHDD$_{2,opt}$ scheme is lower than 0.1dB for all considered cases.

From Fig.2.13 and Fig.2.14, it appears that for the 3 base station SHDD$_{2,opt}$ scheme, if the number of quantization levels is increased from 2 to 4, the performance is improved at $BER = 10^{-5}$ by 4.6dB, when $N_R = 1$, and 3.1dB, when $N_R = 2$. Hence, with only 4 levels of quantization at each base station, the difference between the BER curves of the OC scheme and the 3 base station SHDD$_{2,opt}$ scheme equals approximately 0.8dB, for all considered cases. The gain obtained by increasing the number of quantization from 4 to 8 equals approximately 0.5-0.7dB. In fact, when $L = 8$ the difference in SNR, at fixed BER, between the OC scheme and the 3 base station SHDD$_{2,opt}$ scheme is lower than 0.2dB for all considered cases.

Hence, since the difference between the SHDD$_{2,opt}$ scheme using 8 quantization levels and the OC scheme is less than 0.2dB for all considered cases, it can be concluded that the additional gains obtained by increasing the number of quantization levels over 8 can only be marginal.

D. Effect of the threshold constraints

From Fig.2.10 - Fig.2.14, we see that the SHDD$_{2,opt}$ scheme is much more affected by the constraints than the SHDD$_{1,sub}$ scheme. As opposed to the SHDD$_{1,sub}$ scheme, the performance losses are not only caused by an horizontal shift of the BER curves but the constraints have also an impact on the diversity order of the $SHDD_{2,opt}$ scheme. The more striking example can be observed when the thresholds of a $SHDD_{2,opt}$ scheme, using 2 quantization levels, are forced to be even symmetric when only two base stations are involved in the handoff macrodiversity scheme. It implies that the local decision rules are locally optimum and, as shown in Appendix D, the scheme is now equivalent to the CHM scheme, reducing the diversity order of the scheme by one half. From Fig.2.10, it appears that, when the scheme is using only one receiving antenna, the losses reach up to 17.8dB when operating at a fixed BER of 10^{-5}. However, it can also be seen from these figures that the performance losses diminish with the number of quantization levels, the number of base stations and the number of receiving antennas. In fact, in Fig.2.14 it can be observed that, independently of the number of quantization levels, the constraints produce no noticeable loss of performance for a $SHDD_{2,opt}$ scheme involving 3 base stations, when each base station is equipped with 2 receiving antennas. Similarly, when the number of quantization levels at each base station is equal to 8, in all considered cases the losses caused by forcing the thresholds to be even symmetric are as low as 0.01dB, although the reduction in optimization complexity is considerable.

2.4.3 Comparison of the two cases

In this section, we use the results presented in the two previous sections and compare the BER of the SHDD schemes for the case when the channel state is known at the MSC and the case when the channel state is not known at the MSC. First, by comparing the performances of the $SHDD_{1,sub}$ and $SHDD_{2,opt}$ schemes, we see that, when both schemes are using 2 quantization levels, the SNR difference when operating at $BER = 10^{-5}$ reaches up to 3.4dB, when $N_R = 1$, but reduces to 1.7dB, when $N_R = 2$, and 1.0dB, when $N_R = 3$. Similar, the gain in SNR at $BER = 10^{-5}$ obtained by the $SHDD_{1,opt}$ scheme with respect to the $SHDD_{2,opt}$ scheme equals 3.9dB, when $N_R = 1$, 2.3dB, when $N_R = 2$, and 1.6dB, when $N_R = 3$. Hence, when the SHDD schemes are using 2 quantization levels at each base station, the additional complexity and additional fixed network bandwidth required by the $SHDD_{1,opt}$ scheme and $SHDD_{1,sub}$ scheme can be justified by the relatively important performance gains that provide these schemes with respect to the $SHDD_{2,opt}$ scheme, especially when

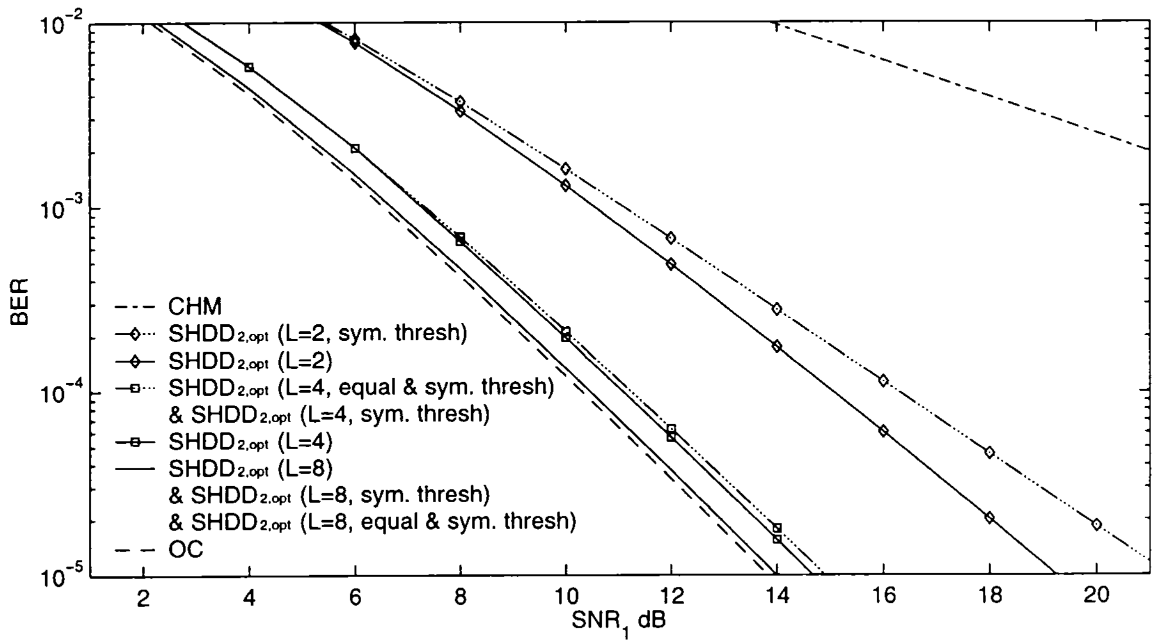


Fig. 2.13 BER vs SNR_1: SHDD$_{2,opt}$, $N_{BS} = 3$, $N_R = 1$, $SNR_1 = SNR_2 = SNR_3$

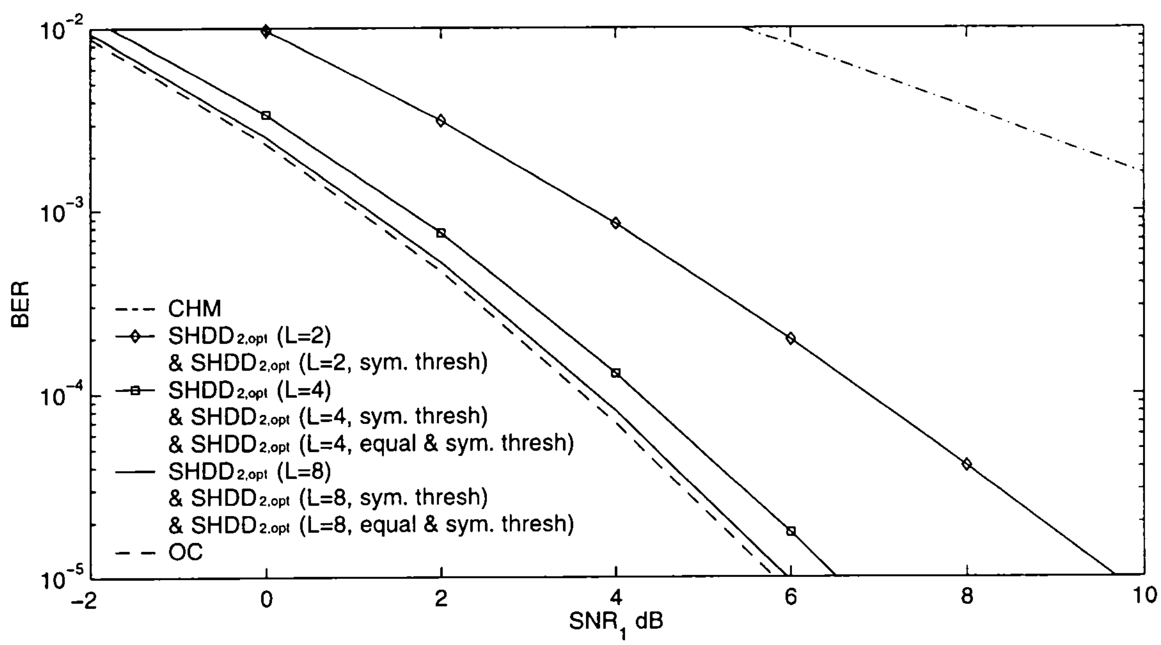


Fig. 2.14 BER vs SNR_1: SHDD$_{2,opt}$, $N_{BS} = 3$, $N_R = 2$, $SNR_1 = SNR_2 = SNR_3$

each base station is using a single receiving antenna.

However, it appears that the performance difference diminishes as the number of quantization levels increases. More precisely, the performance difference between either the $\mathrm{SHDD}_{1,opt}$ or $\mathrm{SHDD}_{1,sub}$ scheme and the $\mathrm{SHDD}_{2,opt}$ scheme using 4 quantization levels reduces to less than 0.3dB at $BER = 10^{-5}$, independently of the number of receiving antennas. Furthermore, the performance difference for the case when $L = 8$ can be upper bounded to approximately $0.07dB$ since the difference between the OC scheme and the $\mathrm{SHDD}_{2,sub}$ scheme, using even symmetric thresholds, is approximately $0.07dB$ for any number of receiving antennas. Hence, when $L > 2$, the performance difference between the $\mathrm{SHDD}_{1,opt}$, $\mathrm{SHDD}_{1,sub}$ and $\mathrm{SHDD}_{2,opt}$ schemes is much less significant. Consequently, when $L > 2$, the $\mathrm{SHDD}_{2,opt}$ scheme becomes a much more attractive alternative since it does not require the channel state to be known at the MSC, only requires that the thresholds be updated when the average SNR varies and a closed form expression can be derived for the threshold optimization cost function.

Chapter 3

SHDD scheme for coded communication systems

In this chapter we study the application of distributed detection, with soft decisions at the local detectors, to the uplink when a mobile unit is in soft handoff and error control coding is used. In section 3.1, the CSHDD scheme is presented. In section 3.2, the optimum CSHDD scheme is derived for BPSK modulation for the case when the channel state is known at the MSC and also when it is not known. Since the complexity of the optimum CSHDD scheme grows exponentially with the frame size, section 3.3 considers sub-optimum alternatives using likelihood ratio quantizers at the base stations. Finally, in section 3.4, the performances of the designed sub-optimum CSHDD schemes are evaluated by computer simulations for a quasi-static spatially-uncorrelated Rayleigh fading channel.

3.1 System model

We consider the uplink of a mobile unit in soft handoff with N_{BS} base stations, each equipped with N_R antennas, as illustrated in Fig. 3.1. At the mobile unit, prior to transmission, the information bits in the frame $\mathbf{B} = [B_1, B_2, \ldots, B_N]^T$, where N is the frame size, are channel encoded, by a rate k_c/n_c convolutional code with constraint length μ, yielding the coded frame $\mathbf{C} = [C_1, C_2, \ldots, C_{N_c}]^T$, where N_c is the coded frame size. Before encoding, N_{tail} tail bits are added to the information frame $\mathbf{B}$ to ensure that the code trellis terminates in the zero state, such that $N_c = (N + N_{tail})\frac{n_c}{k_c}$.

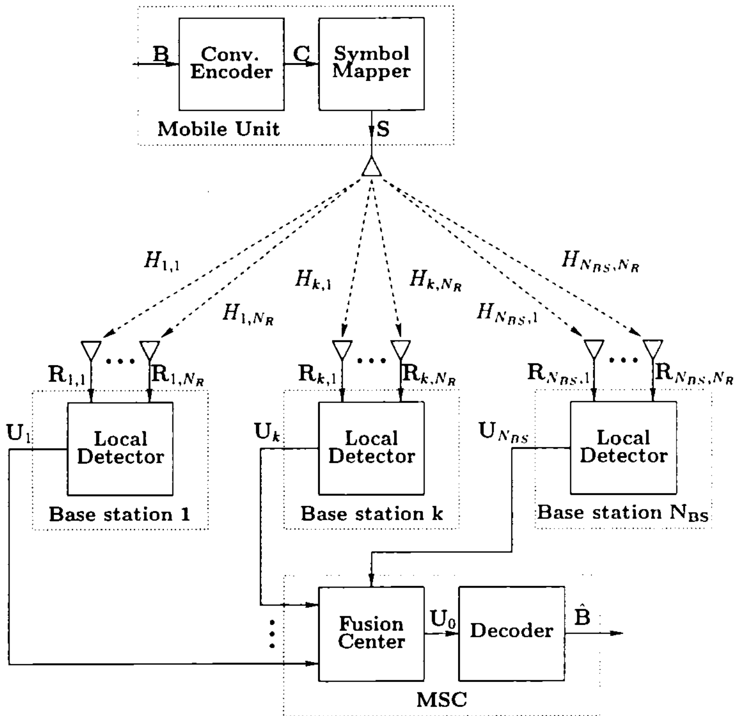


Fig. 3.1 Coded Soft Handoff Distributed detection system model

The coded bits $C_1, C_2, \ldots, C_{N_c}$ are then sent individually to a symbol mapper to generate the BPSK symbols in the frame $\mathbf{S} = [S_1, S_2, \ldots, S_{N_c}]^T$, where $S_t \in \{-1, 1\}$. The symbols are then transmitted sequentially to the N_{BS} base stations taking part in the handoff process.

At the receiving end, all base stations make soft decisions on the transmitted coded bits. For instance, the local detector at the kth base station makes a soft decision $U_{k,t} \in \{0, \ldots, L-1\}$ on the transmitted coded bit C_t based on the received signal vector $\mathbf{R}_{k,t} = [R_{k,t,1}, \ldots, R_{k,t,N_R}]^T$. The decisions contained in the local decision vector $\mathbf{U} = [\mathbf{U}_1, \ldots, \mathbf{U}_{N_{BS}}]^T$, where $\mathbf{U}_k = [U_{k,1}, \ldots, U_{k,N_c}]$, are sent from the different base stations to the MSC where a final decision $\mathbf{U}_0 = [U_{0,1}, \ldots, U_{0,N_c}]^T$ is made on the transmitted coded frame $\mathbf{C}$. More precisely, the final decision is selected from the codebook $\mathcal{C}$ containing all possible codewords such that the decoding process only requires to determine the information frame $\hat{\mathbf{B}} = [\hat{B}_1, \hat{B}_2, \ldots, \hat{B}_N]^T$ associated with the final decision $\mathbf{U}_0$.

The signal received at time-index t at the nth antenna of the kth base station can be represented as follows

$$R_{k,t,n} = H_{k,n}\sqrt{E_k}S_t + N_{k,t,n}. \tag{3.1}$$

The parameters $N_{k,t,n}$ model white Gaussian noise as independent zero mean circular complex Gaussian random variables with variance $N_0/2$ per real and imaginary component. The parameters $H_{k,n}$ model quasi-static spatially-uncorrelated Rayleigh fading as independent zero mean circular complex Gaussian random variables with variance 0.5 per real and imaginary component. The parameters E_k model the average received energy per antenna at the different base stations and are dependent on the position of the mobile unit in the cellular network as well as power control. It is assumed that each base station provides to the MSC an accurate estimate of the average signal-to-noise ratio (SNR) received at each individual antenna, which can be defined as $SNR_k = \frac{E_k}{N_0}$. On the other hand, the channel state vector $\mathbf{H} = [\mathbf{H}_1, \ldots, \mathbf{H}_{N_{BS}}]^T$, where $\mathbf{H}_k = [H_{k,1}, \ldots, H_{k,N_R}]$, may or may not be available at the MSC although $\mathbf{H}_k$ is perfectly known at the kth base station.

3.2 Optimum Distributed Detection

In this section, we derive the optimum CSHDD scheme for which the optimality criterion is the probability of frame error at the output of the MSC. Similarly to the uncoded case considered in the previous chapter, the objective in optimizing such a system is to obtain the set of local decision rules used at different base stations and time instants, denoted by $\gamma_{k,t}$ where $k = 1, \ldots, N_{BS}$ and $t = 1, \ldots, N_c$, and the fusion rule used at the MSC, denoted by γ_0, that jointly minimize the optimality criterion. However, the major difference with respect to the uncoded case is that correlation, introduced by the error control coding, exists between the different transmitted bits of a frame. Consequently, the distributed detection scheme must take in consideration the information contained in the whole frame when making a decision on any of the transmitted bits.

The average probability of frame error at the output of the MSC can be defined as follows

$$P_f = \int_{\mathbf{h}} P_{f|\mathbf{h}} f_{\mathbf{H}}(\mathbf{h}) d\mathbf{h}, \tag{3.2}$$

where $P_{f|\mathbf{h}}$ is the probability of frame error given the channel state vector $\mathbf{H} = \mathbf{h}$ and

$f_{\mathbf{H}}(\mathbf{h})$ is the PDF of $\mathbf{H}$. Let $P(\mathbf{U}_0 = u_0 \mid \mathbf{h}, \mathbf{u})$ denote the probability that the final decision $\mathbf{U}_0$ equals $\mathbf{u}_0$ given the channel state vector $\mathbf{H} = \mathbf{h}$ and the local decision vector $\mathbf{U} = \mathbf{u}$, $P(\mathbf{U} = \mathbf{u} \mid \mathbf{h}, \mathbf{c})$ denote the probability that the local decision vector $\mathbf{U}$ equals $\mathbf{u}$ given the channel state vector $\mathbf{H} = \mathbf{h}$ and the transmitted coded frame $\mathbf{C} = \mathbf{c}$ and $P(\mathbf{C} = \mathbf{c})$ denote the probability that the mobile unit transmitted the codeword $\mathbf{c}$. The conditional probability of frame error $P_{f|\mathbf{h}}$ equals

$$P_{f|\mathbf{h}} = \sum_{\mathbf{c}\in\mathcal{C}} \sum_{\mathbf{u}_0\neq\mathbf{c}} \sum_{\mathbf{u}} P(\mathbf{U}_0 = \mathbf{u}_0 \mid \mathbf{h}, \mathbf{u}) P(\mathbf{U} = \mathbf{u} \mid \mathbf{h}, \mathbf{c}) P(\mathbf{C} = \mathbf{c}). \tag{3.3}$$

where $P(\mathbf{C} = \mathbf{c}) = \frac{1}{2^N}$. Furthermore, the conditional probability of frame error (3.3) can be reformulated as follows

$$\begin{aligned} P_{f|\mathbf{h}} &= \frac{1}{2^N} \sum_{\mathbf{c}\in\mathcal{C}} \sum_{\mathbf{u}} P(\mathbf{U} = \mathbf{u} \mid \mathbf{h}, \mathbf{c}) \sum_{\mathbf{u}_0\neq\mathbf{c}} P(\mathbf{U}_0 = \mathbf{u}_0 \mid \mathbf{h}, \mathbf{u}) \\ &= \frac{1}{2^N} \sum_{\mathbf{c}\in\mathcal{C}} \sum_{\mathbf{u}} P(\mathbf{U} = \mathbf{u} \mid \mathbf{h}, \mathbf{c}) \left(1 - P(\mathbf{U}_0 = \mathbf{c} \mid \mathbf{h}, \mathbf{u})\right) \\ &= 1 - \frac{1}{2^N} \sum_{\mathbf{c}\in\mathcal{C}} \sum_{\mathbf{u}} P(\mathbf{U}_0 = \mathbf{c} \mid \mathbf{h}, \mathbf{u}) P(\mathbf{U} = \mathbf{u} \mid \mathbf{h}, \mathbf{c}), \end{aligned} \tag{3.4}$$

which is a more appropriate form for the optimization of the decision rules.

It is important to mention that, as the uncoded case, the decision rules have a common optimality criterion and are therefore dependent on each others. Since the decision rules cannot be selected individually, we consider that they are selected at the MSC and that the MSC has a mean to update the local decision rules at the base stations when the average SNR or the channel state varies, depending on the information available at the MSC. As mentioned in the previous section, the channel state vector $\mathbf{H} = [\mathbf{H}_1, \ldots, \mathbf{H}_{N_{BS}}]^T$ may or may not be available at the MSC although the channel state vector $\mathbf{H}_k$ is perfectly known at the kth base station. Since both cases provide different decision rules, the derivation is separated in two parts treating separately both cases.

3.2.1 Known channel state information at the fusion center

In this section, it is assumed that the channel state at each base station is varying slowly enough such that accurate estimates can be transmitted to the MSC, where the decision rules are optimized and the final decision $\mathbf{U}_0$ is made. Hence, as for the uncoded case, the fusion rule is obviously a function of the channel state vector $\mathbf{H}$. In addition, since the decision rules are optimized at the MSC, the local decision rules are also functions of the channel state vector $\mathbf{H}$ as long as the MSC updates the local decision rules used at each base station every time the channel state varies. The optimum decision rules are therefore functions of the channel state vector $\mathbf{H}$, and minimize, for $\mathbf{H} = \mathbf{h}$, the conditional probability of frame error (3.4).

In this section, we first derive a fusion rule which is optimum in the sense that, for fixed local decision rules, it provides the minimum average probability of frame error at the output of the fusion center. Then, we derive the kth base station local decision rule which is optimum in the sense that, for a fixed fusion rule and fixed local decision rules at the remaining base stations, it provides the minimum average probability of frame error at the output of the fusion center.

A. Optimum fusion rule

At the MSC, the information available to the fusion rule to make a final decision on the transmitted frame is the local decision vector $\mathbf{U} = [\mathbf{U}_1, \ldots, \mathbf{U}_{N_{BS}}]^T$ and the channel state vector $\mathbf{H}$. Furthermore, since the optimality criterion is the probability of frame error which is a Bayesian criterion, it can be assumed that the fusion rule is deterministic. Hence, when $\mathbf{H} = \mathbf{h}$, the fusion rule should therefore partition the observation set $\mathcal{Z}$ containing all possible realizations of $\mathbf{U}$ into the 2^N mutually exclusive sets $\mathcal{Z}_{\mathbf{c}}(\mathbf{h})$. The conditional probability $P(\mathbf{U}_0 = \mathbf{c} \mid \mathbf{h}, \mathbf{u})$ in (3.4) can thus be expressed as follows

$$P(\mathbf{U}_0 = \mathbf{c} \mid \mathbf{h}, \mathbf{u}) = \begin{cases} 1 \text{ if } \mathbf{u} \in \mathcal{Z}_{\mathbf{c}}(\mathbf{h}) \\ 0 \text{ if } \mathbf{u} \notin \mathcal{Z}_{\mathbf{c}}(\mathbf{h}) \end{cases} \tag{3.5}$$

and the conditional probability of frame error (3.4) can be rewritten as follows

$$P_{f|\mathbf{h}} = 1 - \frac{1}{2^N} \sum_{\mathbf{c} \in C} \sum_{\mathbf{u} \in \mathcal{Z}_{\mathbf{c}}(\mathbf{h})} P(\mathbf{U} = \mathbf{u} \mid \mathbf{h}, \mathbf{c}). \tag{3.6}$$

From expression (3.6), it can be concluded that, in order to minimize P_f, each realization of $\mathbf{U}$ should be included in the set $\mathcal{Z}_{\mathbf{c}}(\mathbf{h})$ associated with the codeword $\mathbf{c}$ that maximizes $P(\mathbf{U} = \mathbf{u} \mid \mathbf{h}, \mathbf{c})$. Therefore, the optimal fusion rule, given $\mathbf{U} = \mathbf{u}$ and

$\mathbf{H} = \mathbf{h}$, is

$$\mathbf{U}_0 = \arg\max_{\mathbf{c}\in\mathcal{C}} \Big\{ P(\mathbf{U} = \mathbf{u} \mid \mathbf{h}, \mathbf{c}) \Big\} \tag{3.7}$$

or equivalently in the maximum a-posterior form

$$\mathbf{U}_0 = \arg\max_{\mathbf{c}\in\mathcal{C}} \Big\{ P(\mathbf{C} = \mathbf{c} \mid \mathbf{h}, \mathbf{u}) \Big\}, \tag{3.8}$$

since all codewords contained in $\mathcal{C}$ are equiprobable. In case of equality, the final decision can be made randomly among the codewords in $\mathcal{C}$ providing the same maximum since it does not affect the performances of the CSHDD scheme. Considering that the local decision vectors $\mathbf{U}_1, \ldots, \mathbf{U}_{N_{BS}}$ are conditionally independent since no communication is assumed between the base stations, the optimum fusion rule simplifies to

$$\mathbf{U}_0 = \arg\max_{\mathbf{c}\in\mathcal{C}} \left\{ \prod_{k=1}^{N_{BS}} P(\mathbf{U}_k = \mathbf{u}_k \mid \mathbf{h}, \mathbf{c}) \right\}. \tag{3.9}$$

Furthermore, since it is assumed that each local decision $U_{k,t}$ is made only based on the information contained in $\mathbf{R}_{k,t}$ ignoring the information contained in the received signal vectors $\mathbf{R}_{k,1}, \ldots, \mathbf{R}_{k,t-1}, \mathbf{R}_{k,t+1}, \ldots, \mathbf{R}_{k,N_c}$, the local decisions $U_{k,1}, \ldots, U_{k,N_c}$ are also conditionally independent such that the fusion rule simplifies to

$$\mathbf{U}_0 = \arg\max_{\mathbf{c}\in\mathcal{C}} \left\{ \prod_{t=0}^{N_c} \prod_{k=1}^{N_{BS}} P(U_{k,t} = u_{k,t} \mid \mathbf{h}, c_t) \right\}. \tag{3.10}$$

The Viterbi algorithm can therefore be used to implement the optimum fusion rule where the branch metric $m_i(s, s')$ associated with the branch starting in state s and ending in state s' in the trellis step i, where $i = 1, \ldots, \frac{N+N_{tail}}{k_c}$, can be defined as follows

$$m_i(s, s') = \prod_{j=1}^{n_c} \prod_{k=1}^{N_{BS}} P(U_{k,(i-1)n_c+j} = u_{k,(i-1)n_c+j} \mid \mathbf{h}, C_{(i-1)n_c+j} = c_j(s, s')) \tag{3.11}$$

given $c_j(s, s')$ is the jth coded bit associated with the set of branches starting in state s and ending in state s'.

B. Optimum local decision rules

Considering the MSC updates the local decision rules every-time the channel state varies, the information available to the local detector of the kth base station, to make the decision $U_{k,t}$ on the transmitted coded bit C_t, is the received signal vectors $\mathbf{R}_{k,1}, \ldots, \mathbf{R}_{k,N_c}$ and the channel state vector $\mathbf{H}$. However, it is assumed that the local decision rule that determines the local decision $U_{k,t}$ ignores the information contained in the received signal vectors $\mathbf{R}_{k,1}, \ldots, \mathbf{R}_{k,t-1}, \mathbf{R}_{k,t+1}, \ldots, \mathbf{R}_{k,N_c}$ in order for the kth base station local decisions $U_{k,1}, \ldots, U_{k,N_c}$ to be conditionally independent and $P(\mathbf{U}_k = \mathbf{u}_k \mid \mathbf{h}, \mathbf{c}) = \prod_{k=1}^{N_c} P(U_{k,t} = u_{k,t} \mid \mathbf{h}, c_t)$. The local decision $U_{k,t}$ is therefore conditionally independent from all the other decisions contained in $\mathbf{U}$ such that $P(\mathbf{U} = \mathbf{u} \mid \mathbf{h}, \mathbf{c}) = P(\mathbf{U}^{k,t} = \mathbf{u}^{k,t} \mid \mathbf{h}, \mathbf{c}) P(U_{k,t} = u_{k,t} \mid \mathbf{h}, c_t)$ and expression (3.4) can be expressed as follows

$$\begin{aligned} P_{f|\mathbf{h}} &= 1 - \frac{1}{2^N} \sum_{\mathbf{c} \in \mathcal{C}} \sum_{u_{k,t}} \sum_{\mathbf{u}^{k,t}} P(\mathbf{U}_0 = \mathbf{c} \mid \mathbf{h}, \mathbf{u}^{k,t}, u_{k,t}) \times \\ & \qquad P(\mathbf{U}^{k,t} = \mathbf{u}^{k,t} \mid \mathbf{h}, \mathbf{c}) P(U_{k,t} = u_{k,t} \mid \mathbf{h}, c_t), \quad (3.12) \end{aligned}$$

where $\mathbf{U}^{k,t} = [\mathbf{U}_1, \ldots, \mathbf{U}_k^t, \ldots, \mathbf{U}_{N_{BS}}]^T$ and $\mathbf{U}_k^t = [U_{k,1}, \ldots, U_{k,t-1}, U_{k,t+1}, \ldots, U_{k,N_c}]^T$.

The local decision rule $\gamma_{k,t}(\mathbf{r}_{k,t}, \mathbf{h})$, which determines the value of the local decision $U_{k,t}$ given $\mathbf{R}_{k,t} = \mathbf{r}_{k,t}$ and $\mathbf{H} = \mathbf{h}$, should partition the observation set $\mathcal{R}^{k,t}$ containing all possible realizations of $\mathbf{R}_{k,t}$ into the mutually exclusive sets $\mathcal{R}_l^{k,t}(\mathbf{h})$, where $l = 0, \ldots, L-1$. In order to determine which realizations of $\mathbf{R}_{k,t}$ should be included in these sets, it is necessary to expand the conditional probability of frame error (3.12) as a function of $\mathbf{r}_{k,t}$. Let $f_{\mathbf{R}_{k,t}}(\mathbf{r}_{k,t} \mid \mathbf{h}_k, c_t)$ denote the joint PDF of the received signals $R_{k,t,1}, \ldots, R_{k,t,N_R}$ given the coded bit $C_t = c_t$ and the channel state vector $\mathbf{H}_k = \mathbf{h}_k$. The conditional probability $P(U_{k,t} = u_{k,t} \mid \mathbf{h}, c_t)$ can be expanded as a function of $\mathbf{r}_{k,t}$ as follows

$$\begin{aligned} P(U_{k,t} = u_{k,t} \mid \mathbf{h}, c_t) &= \int_{\mathbf{r}_{k,t}} P(U_{k,t} = u_{k,t} \mid \mathbf{h}, \mathbf{r}_{k,t}, c_t) f_{\mathbf{R}_{k,t}}(\mathbf{r}_{k,t} \mid \mathbf{h}, c_t) d\mathbf{r}_{k,t} \\ &= \int_{\mathbf{r}_{k,t}} P(U_{k,t} = u_{k,t} \mid \mathbf{h}, \mathbf{r}_{k,t}) f_{\mathbf{R}_{k,t}}(\mathbf{r}_{k,t} \mid \mathbf{h}_k, c_t) d\mathbf{r}_{k,t}, \quad (3.13) \end{aligned}$$

where $P(U_{k,t} = u_{k,t} \mid \mathbf{h}, \mathbf{r}_{k,t}, c_t) = P(U_{k,t} = u_{k,t} \mid \mathbf{h}, \mathbf{r}_{k,t})$, since the knowledge of $\mathbf{H}$ and $\mathbf{R}_{k,t}$ is sufficient to determine the local decision $U_{k,t}$, and $f_{\mathbf{R}_{k,t}}(\mathbf{r}_{k,t} \mid \mathbf{h}, c_t) =$

$f_{\mathbf{R}_{k,t}}(\mathbf{r}_{k,t} \mid \mathbf{h}_k, c_t)$, since $\mathbf{R}_{k,t}$ is independent of $\mathbf{H}^k = [\mathbf{H}_1, \ldots, \mathbf{H}_{k-1}, \mathbf{H}_{k+1}, \ldots, \mathbf{H}_{N_{BS}}]^T$. Substituting (3.13) in (3.12), the conditional probability of frame error can therefore be expanded as a function of $\mathbf{r}_{k,t}$, such that

$$\begin{aligned} P_{f|\mathbf{h}} &= 1 - \frac{1}{2^N} \sum_{\mathbf{c}\in\mathcal{C}} \sum_{u_{k,t}} \sum_{\mathbf{u}^{k,t}} P(\mathbf{U}_0 = \mathbf{c} \mid \mathbf{h}, \mathbf{u}^{k,t}, u_{k,t}) \times \\ &\qquad P(\mathbf{U}^{k,t} = \mathbf{u}^{k,t} \mid \mathbf{h}, \mathbf{c}) \int_{\mathbf{r}_{k,t}} P(U_{k,t} = u_{k,t} \mid \mathbf{h}, \mathbf{r}_{k,t}) f_{\mathbf{R}_{k,t}}(\mathbf{r}_{k,t} \mid \mathbf{h}_k, c_t) d\mathbf{r}_{k,t} \\ &= 1 - \frac{1}{2^N} \sum_{u_{k,t}} \int_{\mathbf{r}_{k,t}} P(U_{k,t} = u_{k,t} \mid \mathbf{h}, \mathbf{r}_{k,t}) \sum_{\mathbf{c}\in\mathcal{C}} \sum_{\mathbf{u}^{k,t}} P(\mathbf{U}_0 = \mathbf{c} \mid \mathbf{h}, \mathbf{u}^{k,t}, u_{k,t}) \times \\ &\qquad P(\mathbf{U}^{k,t} = \mathbf{u}^{k,t} \mid \mathbf{h}, \mathbf{c}) f_{\mathbf{R}_{k,t}}(\mathbf{r}_{k,t} \mid \mathbf{h}_k, c_t) d\mathbf{r}_{k,t} \\ &= 1 - \frac{1}{2^N} \sum_{u_{k,t}} \int_{\mathbf{r}_{k,t}} P(U_{k,t} = u_{k,t} \mid \mathbf{h}, \mathbf{r}_{k,t}) L_{u_{k,t}}(\mathbf{r}_{k,t}, \mathbf{h}) d\mathbf{r}_{k,t}, \end{aligned} \tag{3.14}$$

where

$$L_{u_{k,t}}(\mathbf{r}_{k,t}, \mathbf{h}) = \sum_{\mathbf{c}\in\mathcal{C}} \sum_{\mathbf{u}^{k,t}} P(\mathbf{U}_0 = \mathbf{c} \mid \mathbf{h}, \mathbf{u}^{k,t}, u_{k,t}) P(\mathbf{U}^{k,t} = \mathbf{u}^{k,t} \mid \mathbf{h}, \mathbf{c}) f_{\mathbf{R}_{k,t}}(\mathbf{r}_{k,t} \mid \mathbf{h}_k, c_t). \tag{3.15}$$

Since the local decision $U_{k,t}$ is specified by the deterministic decision rule $\gamma_{k,t}(\mathbf{r}_{k,t}, \mathbf{h})$ given $\mathbf{R}_{k,t} = \mathbf{r}_{k,t}$ and $\mathbf{H} = \mathbf{h}$, the conditional probability $P(U_{k,t} = u_{k,t} \mid \mathbf{h}, \mathbf{r}_{k,t})$ in (3.14) can thus be expressed as follows

$$P(U_{k,t} = u_{k,t} \mid \mathbf{h}, \mathbf{r}_{k,t}) = \begin{cases} 1 & \text{if } \mathbf{r}_{k,t} \in \mathcal{R}^{k,t}_{u_{k,t}}(\mathbf{h}) \\ 0 & \text{if } \mathbf{r}_{k,t} \notin \mathcal{R}^{k,t}_{u_{k,t}}(\mathbf{h}) \end{cases} \tag{3.16}$$

and the conditional probability of frame error (3.14) can be rewritten as follows

$$P_{f|\mathbf{h}} = 1 - \frac{1}{2^N} \sum_{u_{k,t}} \int_{\mathbf{r}_{k,t} \in \mathcal{R}^{k,t}_{u_{k,t}}(\mathbf{h})} L_{u_{k,t}}(\mathbf{r}_{k,t}, \mathbf{h}) d\mathbf{r}_{k,t}. \tag{3.17}$$

From expression (3.17), it can be concluded that, in order to minimize P_f, each realization of $\mathbf{R}_{k,t}$ should be included in the set $\mathcal{R}^{k,t}_{u_{k,t}}(\mathbf{h})$ associated with the index $u_{k,t}$ that maximizes $L_{u_{k,t}}(\mathbf{r}_{k,t}, \mathbf{h})$. Therefore, the optimum local decision rule, given

$\mathbf{R}_{k,t} = \mathbf{r}_{k,t}$ and $\mathbf{H} = \mathbf{h}$, equals

$$U_{k,t} = \arg \max_{u_{k,t}=0,\ldots,L-1} \left\{ L_{u_{k,t}}(\mathbf{r}_{k,t}, \mathbf{h}) \right\}. \tag{3.18}$$

Then, given the set $\mathcal{C}_t^1$ of all codewords for which $C_t = 1$ and the set $\mathcal{C}_t^0$ of all codewords for which $C_t = 0$, expression (3.15) can be expressed as follows

$$\begin{aligned}
&L_{u_{k,t}}(\mathbf{r}_{k,t}, \mathbf{h}) \\
&= \sum_{\mathbf{c}\in\mathcal{C}_t^1} \sum_{\mathbf{u}^{k,t}} P(\mathbf{U}_0 = \mathbf{c} \mid \mathbf{h}, \mathbf{u}^{k,t}, u_{k,t}) P(\mathbf{U}^{k,t} = \mathbf{u}^{k,t} \mid \mathbf{h}, \mathbf{c}) f_{\mathbf{R}_{k,t}}(\mathbf{r}_{k,t} \mid \mathbf{h}_k, C_t = 1) \\
&\quad + \sum_{\mathbf{c}\in\mathcal{C}_t^0} \sum_{\mathbf{u}^{k,t}} P(\mathbf{U}_0 = \mathbf{c} \mid \mathbf{h}, \mathbf{u}^{k,t}, u_{k,t}) P(\mathbf{U}^{k,t} = \mathbf{u}^{k,t} \mid \mathbf{h}, \mathbf{c}) f_{\mathbf{R}_{k,t}}(\mathbf{r}_{k,t} \mid \mathbf{h}_k, C_t = 0) \\
&= f_{\mathbf{R}_{k,t}}(\mathbf{r}_{k,t} \mid \mathbf{h}_k, C_t = 1) \sum_{\mathbf{c}\in\mathcal{C}_t^1} \sum_{\mathbf{u}^{k,t}} P(\mathbf{U}_0 = \mathbf{c} \mid \mathbf{h}, \mathbf{u}^{k,t}, u_{k,t}) P(\mathbf{U}^{k,t} = \mathbf{u}^{k,t} \mid \mathbf{h}, \mathbf{c}) \\
&\quad + f_{\mathbf{R}_{k,t}}(\mathbf{r}_{k,t} \mid \mathbf{h}_k, C_t = 0) \sum_{\mathbf{c}\in\mathcal{C}_t^0} \sum_{\mathbf{u}^{k,t}} P(\mathbf{U}_0 = \mathbf{c} \mid \mathbf{h}, \mathbf{u}^{k,t}, u_{k,t}) P(\mathbf{U}^{k,t} = \mathbf{u}^{k,t} \mid \mathbf{h}, \mathbf{c}).
\end{aligned} \tag{3.19}$$

The local decision rule (3.18) can therefore be reformulated as follows

$$U_{k,t} = \arg \max_{u_{k,t}=0,\ldots,L-1} \left\{ f_{\mathbf{R}_{k,t}}(\mathbf{r}_{k,t} \mid \mathbf{h}_k, C_t = 0) \left(m_{u_{k,t}} \Lambda_{1,0}^{(k,t)}(\mathbf{r}_{k,t}, \mathbf{h}_k) + b_{u_{k,t}} \right) \right\} \tag{3.20}$$

or equivalently, since $f_{\mathbf{R}_{k,t}}(\mathbf{r}_{k,t} \mid \mathbf{h}_k, C_t = 0)$ is independent of $u_{k,t}$, as follows

$$U_{k,t} = \arg \max_{u_{k,t}=0,\ldots,L-1} \left\{ m_{u_{k,t}} \Lambda_{1,0}^{(k,t)}(\mathbf{r}_{k,t}, \mathbf{h}_k) + b_{u_{k,t}} \right\}, \tag{3.21}$$

where

$$\Lambda_{1,0}^{(k,t)}(\mathbf{r}_{k,t}, \mathbf{h}_k) = \frac{f_{\mathbf{R}_{k,t}}(\mathbf{r}_{k,t} \mid \mathbf{h}_k, C_t = 1)}{f_{\mathbf{R}_{k,t}}(\mathbf{r}_{k,t} \mid \mathbf{h}_k, C_t = 0)}, \tag{3.22}$$

$$m_{u_{k,t}} = \sum_{\mathbf{c}\in\mathcal{C}_t^1} \sum_{\mathbf{u}^{k,t}} P(\mathbf{U}_0 = \mathbf{c} \mid \mathbf{h}, \mathbf{u}^{k,t}, u_{k,t}) P(\mathbf{U}^{k,t} = \mathbf{u}^{k,t} \mid \mathbf{h}, \mathbf{c}) \tag{3.23}$$

and

$$b_{u_{k,t}} = \sum_{\mathbf{c}\in\mathcal{C}_t^0} \sum_{\mathbf{u}^{k,t}} P(\mathbf{U}_0 = \mathbf{c} \mid \mathbf{h}, \mathbf{u}^{k,t}, u_{k,t}) P(\mathbf{U}^{k,t} = \mathbf{u}^{k,t} \mid \mathbf{h}, \mathbf{c}). \tag{3.24}$$

Since the coefficients $m_{u_{k,t}}$ and $b_{u_{k,t}}$ are independent of $\mathbf{r}_{k,t}$ but varying with $\mathbf{h}$, as can be seen on Fig. 3.2 the optimum decision rule $\gamma_{k,t}(\mathbf{r}_{k,t}, \mathbf{h})$ is, as for the uncoded

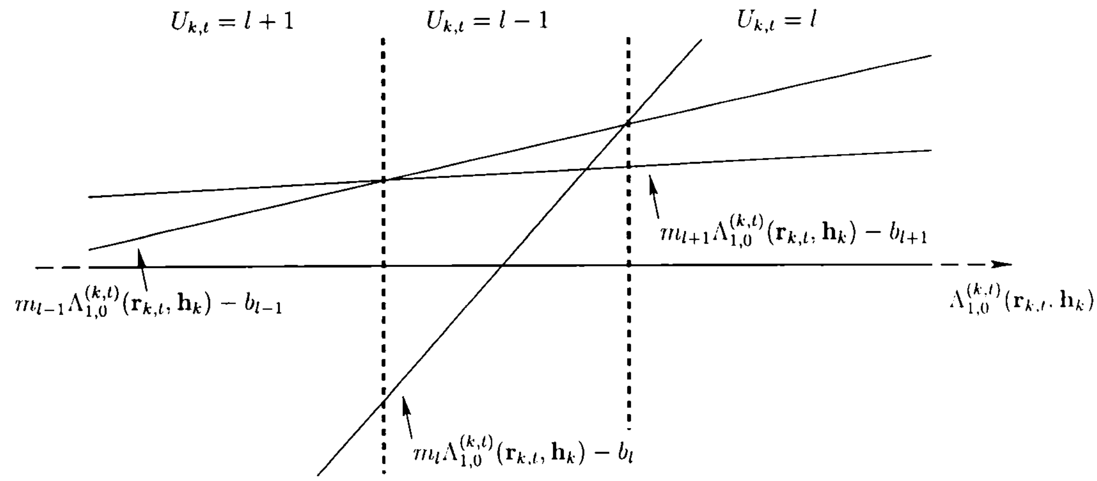


Fig. 3.2 Graphical representation of the local decision rule $\gamma_{k,t}(\mathbf{r}_{k,t}, \mathbf{h})$ and $\gamma_{k,t}(\mathbf{r}_{k,t}, \mathbf{h}_k)$

case, a likelihood ratio threshold test, where the thresholds partitioning the likelihood ratio $\Lambda_{1,0}^{(k,t)}(\mathbf{r}_{k,t}, \mathbf{h}_k)$ are functions of $\mathbf{h}$ and must therefore be updated as the channel state varies. Furthermore, as for the uncoded case, the values of these thresholds cannot be determined analytically due to the interdependence of these thresholds and other local detector thresholds with respect to the common optimality criterion. The thresholds must therefore be optimized simultaneously using a numerical optimization algorithm in order to determine their optimum values. Since the channel state is known at the MSC where the numerical optimization takes place, in this case the cost function minimized by the thresholds, for $\mathbf{H} = \mathbf{h}$, is the conditional probability of frame error (3.4). However, as opposed to the uncoded case, the number of summations required to evaluate this cost function grows exponentially with the frame size. Therefore, in section 3.3 sub-optimum schemes are proposed.

3.2.2 Unknown channel state information at the fusion center

In this section, it is assumed that the channel state vector $\mathbf{H}$ is not known at the MSC, where the decision rules are optimized and the final decision $\mathbf{U}_0$ is made. However, the statistics necessary to describe the random behavior of the channel state vector $\mathbf{H}$ are known. Therefore, as opposed to the previous case, the fusion rule is not a function of the channel state vector $\mathbf{H}$ and the MSC only updates the local decision rules every time the average SNR varies at any base station. However, since $\mathbf{H}_k$ is perfectly known at the kth base station, it is possible for the kth base station local decision rules to be functions of the channel state vector $\mathbf{H}_k$.

In this section, we first derive a fusion rule which is optimum in the sense that, for fixed local decision rules at the base stations, it provides the minimum average probability of frame error at the output of the fusion center. Then, we derive the kth base station local decision rules which are optimum in the sense that, for a fixed fusion rule and fixed local decision rules at the remaining base stations, it provides the minimum average probability of frame error at the output of the fusion center.

A. Optimum fusion rule

At the MSC, the only information available to the fusion rule to make a final decision on the transmitted frame $\mathbf{C}$ is the local decision vector $\mathbf{U} = [\mathbf{U}_1, \ldots, \mathbf{U}_{N_{BS}}]^T$. Since the optimality criterion is the probability of frame error which is a Bayesian criterion, it can be assumed that the fusion rule is deterministic. The fusion rule should therefore partition the observation set $\mathcal{Z}$ containing all possible realizations of $\mathbf{U}$ into the 2^N mutually exclusive sets $\mathcal{Z}_{\mathbf{c}}$. As opposed to the previous case, the fusion rule is not a function of the channel state vector $\mathbf{H}$ such that the conditional probability $P(\mathbf{U}_0 = \mathbf{u}_0 \mid \mathbf{h}, \mathbf{u})$ simplifies as follows $P(\mathbf{U}_0 = \mathbf{u}_0 \mid \mathbf{u})$. The average probability of frame error (3.2) can thus be reformulated as

$$\begin{aligned} P_f &= 1 - \frac{1}{2^N} \sum_{\mathbf{c} \in \mathcal{C}} \sum_{\mathbf{u}} P(\mathbf{U}_0 = \mathbf{c} \mid \mathbf{u}) \int_{\mathbf{h}} f_{\mathbf{H}}(\mathbf{h}) P(\mathbf{U} = \mathbf{u} \mid \mathbf{h}, \mathbf{c}) d\mathbf{h} \\ &= 1 - \frac{1}{2^N} \sum_{\mathbf{c} \in \mathcal{C}} \sum_{\mathbf{u}} P(\mathbf{U}_0 = \mathbf{c} \mid \mathbf{u}) P(\mathbf{U} = \mathbf{u} \mid \mathbf{c}) \end{aligned} \tag{3.25}$$

since

$$P(\mathbf{U} = \mathbf{u} \mid \mathbf{c}) = \int_{\mathbf{h}} f_{\mathbf{H}}(\mathbf{h}) P(\mathbf{U} = \mathbf{u} \mid \mathbf{h}, \mathbf{c}) d\mathbf{h}. \tag{3.26}$$

Furthermore, since the final decision $\mathbf{U}_0$ is specified by the fusion rule $\gamma_0(\mathbf{u})$ given $\mathbf{U} = \mathbf{u}$, the conditional probability $P(\mathbf{U}_0 = \mathbf{c} \mid \mathbf{u})$ in (3.25) can thus be expressed as follows

$$P(\mathbf{U}_0 = \mathbf{c} \mid \mathbf{u}) = \begin{cases} 1 \text{ if } \mathbf{u} \in \mathcal{Z}_{\mathbf{c}} \\ 0 \text{ if } \mathbf{u} \notin \mathcal{Z}_{\mathbf{c}} \end{cases} \tag{3.27}$$

and the probability of frame error can be rewritten as follows

$$P_f = 1 - \frac{1}{2^N} \sum_{\mathbf{c} \in \mathcal{C}} \sum_{\mathbf{u} \in \mathcal{Z}_\mathbf{c}} P(\mathbf{U} = \mathbf{u} \mid \mathbf{c}). \tag{3.28}$$

From expression (3.28), it can be concluded that, in order to minimize P_f, each realization of $\mathbf{U}$ should be included in the set $\mathcal{Z}_\mathbf{c}$ associated with the codeword $\mathbf{c}$ that maximizes $P(\mathbf{U} = \mathbf{u} \mid \mathbf{c})$. Therefore, the optimal fusion rule, given $\mathbf{U} = \mathbf{u}$, is

$$\mathbf{U}_0 = \arg\max_{\mathbf{c} \in \mathcal{C}} \left\{ P(\mathbf{U} = \mathbf{u} \mid \mathbf{c}) \right\} \tag{3.29}$$

or equivalently in the maximum a-posterior form

$$\mathbf{U}_0 = \arg\max_{\mathbf{c} \in \mathcal{C}} \left\{ P(\mathbf{C} = \mathbf{c} \mid \mathbf{u}) \right\}, \tag{3.30}$$

since all codewords contained in $\mathcal{C}$ are equiprobable. Then, considering that the local decision vectors $\mathbf{U}_1, \ldots, \mathbf{U}_{N_{BS}}$ are conditionally independent since no communication is assumed between the base stations and the channel fading is spatially uncorrelated, the optimum fusion rule simplifies as follows

$$\mathbf{U}_0 = \arg\max_{\mathbf{c} \in \mathcal{C}} \left\{ \prod_{k=1}^{N_{BS}} P(\mathbf{U}_k = \mathbf{u}_k \mid \mathbf{c}) \right\}. \tag{3.31}$$

However, even if each local decision $U_{k,t}$ is made only based on the information contained in $\mathbf{R}_{k,t}$, the kth base station local decisions $U_{k,1}, \ldots, U_{k,N_c}$ are still dependent on the unknown channel state vector $\mathbf{H}_k$ such that $P(\mathbf{U}_k = \mathbf{u}_k \mid \mathbf{c}) \neq \prod_{t=1}^{N_c} P(U_{k,t} = u_{k,t} \mid c_t)$. Therefore, as opposed to the previous case, the Viterbi algorithm cannot be used to progressively eliminate possible candidates using the convolution code structure. The implementation of such a fusion rule requires a number of comparisons that increases exponentially with the frame size.

B. Optimum local decision rules

At the kth base station, the information available to the local detector, to make the local decision $U_{k,t}$ on the transmitted coded bit C_t, is the received signal vectors $\mathbf{R}_{k,1}, \ldots, \mathbf{R}_{k,N_c}$ and the channel state vector $\mathbf{H}_k$. However, as for the previous case, it is assumed that the local decision rule that determine the local decision $U_{k,t}$ ignores the information contained in the received signal vectors $\mathbf{R}_{k,1}, \ldots, \mathbf{R}_{k,t-1}, \mathbf{R}_{k,t+1}, \ldots, \mathbf{R}_{k,N_c}$ in order for the kth base station local decisions $U_{k,1}, \ldots, U_{k,N_c}$ to be conditionally independent.

The local decision rule $\gamma_{k,t}(\mathbf{r}_{k,t}, \mathbf{h}_k)$, which determines the value of the local decision $U_{k,t}$ given $\mathbf{R}_{k,t} = \mathbf{r}_{k,t}$ and $\mathbf{H}_k = \mathbf{h}_k$, should therefore partition the observation set $\mathcal{R}^{k,t}$ containing all possible realizations of $\mathbf{R}_{k,t}$ into the mutually exclusive sets $\mathcal{R}_l^{k,t}(\mathbf{h}_k)$, where $l = 0, \ldots, L-1$. In order to determine which realizations of $\mathbf{R}_{k,t}$ should be included in these sets, it is necessary to expand as a function of $\mathbf{r}_{k,t}$ the probability of frame error conditioned on the channel state vector $\mathbf{H}_k = \mathbf{h}_k$ which can be defined as

$$P_{f|\mathbf{h}_k} = \int_{\mathbf{h}^k} P_{f|\mathbf{h}} f_{\mathbf{H}^k}(\mathbf{h}^k|\mathbf{h}_k) d\mathbf{h}^k = \int_{\mathbf{h}^k} P_{f|\mathbf{h}} f_{\mathbf{H}^k}(\mathbf{h}^k) d\mathbf{h}^k, \tag{3.32}$$

where $P_{f|\mathbf{h}}$ is defined in expression (3.4) and $f_{\mathbf{H}^k}(\mathbf{h}^k|\mathbf{h}_k) = f_{\mathbf{H}^k}(\mathbf{h}^k)$ since $\mathbf{H}^k = [\mathbf{H}_1, \ldots, \mathbf{H}_{k-1}, \mathbf{H}_{k+1}, \ldots, \mathbf{H}_{N_{BS}}]^T$ is independent of $\mathbf{H}_k$. Hence, by substituting (3.4) in (3.32), the conditional probability of frame error can be expressed as follows

$$P_{f|\mathbf{h}_k} = 1 - \frac{1}{2^N} \sum_{\mathbf{c} \in \mathcal{C}} \sum_{\mathbf{u}} P(\mathbf{U}_0 = \mathbf{c} \mid \mathbf{u}) \int_{\mathbf{h}^k} f_{\mathbf{H}^k}(\mathbf{h}^k) P(\mathbf{U} = \mathbf{u} \mid \mathbf{h}, \mathbf{c}) d\mathbf{h}^k, \tag{3.33}$$

where we used the fact that $P(U_0 = \mathbf{c} \mid \mathbf{h}, \mathbf{u}) = P(U_0 = \mathbf{c} \mid \mathbf{u})$ since the fusion rule is independent of $\mathbf{H}$. Furthermore, since the local decisions are conditionally independent, it implies that, in (3.33), $P(\mathbf{U} = \mathbf{u} \mid \mathbf{h}, \mathbf{c}) = P(\mathbf{U}^k = \mathbf{u}^k \mid \mathbf{h}^k, \mathbf{c}) P(\mathbf{U}_k^t = \mathbf{u}_k^t \mid \mathbf{h}_k, \mathbf{c}) P(U_{k,t} = u_{k,t} \mid \mathbf{h}_k, c_t)$ and the conditional probability of frame error can be expressed as follows

$$\begin{aligned}
P_{f|\mathbf{h}_k} &= 1 - \frac{1}{2^N} \sum_{\mathbf{c} \in \mathcal{C}} \sum_{u_{k,t}} \sum_{\mathbf{u}_k^t} \sum_{\mathbf{u}^k} P(\mathbf{U}_0 = \mathbf{c} \mid \mathbf{u}^k, \mathbf{u}_k^t, u_{k,t}) \times \\
&\quad \int_{\mathbf{h}^k} f_{\mathbf{H}^k}(\mathbf{h}^k) P(\mathbf{U}^k = \mathbf{u}^k \mid \mathbf{h}^k, \mathbf{c}) d\mathbf{h}^k P(\mathbf{U}_k^t = \mathbf{u}_k^t \mid \mathbf{h}_k, \mathbf{c}) P(U_{k,t} = u_{k,t} \mid \mathbf{h}_k, c_t) \\
&= 1 - \frac{1}{2^N} \sum_{\mathbf{c} \in \mathcal{C}} \sum_{u_{k,t}} \sum_{\mathbf{u}_k^t} \sum_{\mathbf{u}^k} P(\mathbf{U}_0 = \mathbf{c} \mid \mathbf{u}^k, \mathbf{u}_k^t, u_{k,t}) \times \\
&\quad P(\mathbf{U}^k = \mathbf{u}^k \mid \mathbf{c}) P(\mathbf{U}_k^t = \mathbf{u}_k^t \mid \mathbf{h}_k, \mathbf{c}) P(U_{k,t} = u_{k,t} \mid \mathbf{h}_k, c_t),
\end{aligned} \tag{3.34}$$

where $\mathbf{U}^k = [\mathbf{U}_1, \ldots, \mathbf{U}_{k-1}, \mathbf{U}_{k+1}, \ldots, \mathbf{U}_{N_{BS}}]^T$, $\mathbf{U}_k^t = [U_{k,1}, \ldots, U_{k,t-1}, U_{k,t+1}, \ldots, U_{k,N_c}]^T$

and

$$P(\mathbf{U}^k = \mathbf{u}^k \mid \mathbf{c}) = \int_{\mathbf{h}^k} f_{\mathbf{H}^k}(\mathbf{h}^k) P(\mathbf{U}^k = \mathbf{u}^k \mid \mathbf{h}^k, \mathbf{c}) d\mathbf{h}^k. \tag{3.35}$$

The conditional probability $P(U_{k,t} = u_{k,t} \mid \mathbf{h}_k, c_t)$ in (3.34) can be expanded as a function of $\mathbf{r}_{k,t}$ as follows

$$\begin{aligned} P(U_{k,t} = u_{k,t} \mid \mathbf{h}_k, c_t) &= \int_{\mathbf{r}_{k,t}} P(U_{k,t} = u_{k,t} \mid \mathbf{h}_k, \mathbf{r}_{k,t}, c_t) f_{\mathbf{R}_{k,t}}(\mathbf{r}_{k,t} \mid \mathbf{h}_k, c_t) d\mathbf{r}_{k,t} \\ &= \int_{\mathbf{r}_{k,t}} P(U_{k,t} = u_{k,t} \mid \mathbf{h}_k, \mathbf{r}_{k,t}) f_{\mathbf{R}_{k,t}}(\mathbf{r}_{k,t} \mid \mathbf{h}_k, c_t) d\mathbf{r}_{k,t}, \end{aligned} \tag{3.36}$$

where $P(U_{k,t} = u_{k,t} \mid \mathbf{h}_k, \mathbf{r}_{k,t}, c_t) = P(U_{k,t} = u_{k,t} \mid \mathbf{h}_k, \mathbf{r}_{k,t})$, since the knowledge of $\mathbf{H}_k$ and $\mathbf{R}_{k,t}$ is sufficient to determine the local decision $U_{k,t}$. Hence, by substituting (3.36) in (3.34), the conditional probability of frame error can therefore be expanded as a function of $\mathbf{r}_{k,t}$, such that

$$\begin{aligned} & P_{f|\mathbf{h}_k} \\ &= 1 - \frac{1}{2^N} \sum_{\mathbf{c} \in \mathcal{C}} \sum_{u_{k,t}} \sum_{\mathbf{u}_k^t} \sum_{\mathbf{u}^k} P(\mathbf{U}_0 = \mathbf{c} \mid \mathbf{u}^k, \mathbf{u}_k^t, u_{k,t}) \times \\ & \quad P(\mathbf{U}^k = \mathbf{u}^k \mid \mathbf{c}) P(\mathbf{U}_k^t = \mathbf{u}_k^t \mid \mathbf{h}_k, \mathbf{c}) \int_{\mathbf{r}_{k,t}} P(U_{k,t} = u_{k,t} \mid \mathbf{h}_k, \mathbf{r}_{k,t}) f_{\mathbf{R}_{k,t}}(\mathbf{r}_{k,t} \mid \mathbf{h}_k, c_t) d\mathbf{r}_{k,t} \\ &= 1 - \frac{1}{2^N} \sum_{u_{k,t}} \int_{\mathbf{r}_{k,t}} P(U_{k,t} = u_{k,t} \mid \mathbf{h}_k, \mathbf{r}_{k,t}) \sum_{\mathbf{c} \in \mathcal{C}} \sum_{\mathbf{u}_k^t} \sum_{\mathbf{u}^k} P(\mathbf{U}_0 = \mathbf{c} \mid \mathbf{u}^k, \mathbf{u}_k^t, u_{k,t}) \times \\ & \qquad P(\mathbf{U}^k = \mathbf{u}^k \mid \mathbf{c}) P(\mathbf{U}_k^t = \mathbf{u}_k^t \mid \mathbf{h}_k, \mathbf{c}) f_{\mathbf{R}_{k,t}}(\mathbf{r}_{k,t} \mid \mathbf{h}_k, c_t) d\mathbf{r}_{k,t} \\ &= 1 - \frac{1}{2^N} \sum_{u_{k,t}} \int_{\mathbf{r}_{k,t}} P(U_{k,t} = u_{k,t} \mid \mathbf{h}_k, \mathbf{r}_{k,t}) L_{u_{k,t}}(\mathbf{r}_{k,t}, \mathbf{h}_k) d\mathbf{r}_{k,t}, \end{aligned} \tag{3.37}$$

where

$$\begin{aligned} L_{u_{k,t}}(\mathbf{r}_{k,t}, \mathbf{h}_k) &= \sum_{\mathbf{c} \in \mathcal{C}} \sum_{\mathbf{u}_k^t} \sum_{\mathbf{u}^k} P(\mathbf{U}_0 = \mathbf{c} \mid \mathbf{u}^k, \mathbf{u}_k^t, u_{k,t}) \times \\ & \qquad P(\mathbf{U}^k = \mathbf{u}^k \mid \mathbf{c}) P(\mathbf{U}_k^t = \mathbf{u}_k^t \mid \mathbf{h}_k, \mathbf{c}) f_{\mathbf{R}_{k,t}}(\mathbf{r}_{k,t} \mid \mathbf{h}_k, c_t). \end{aligned} \tag{3.38}$$

Since the local decision $U_{k,t}$ is specified by the deterministic decision rule $\gamma_{k,t}(\mathbf{r}_{k,t}, \mathbf{h}_k)$ given $\mathbf{R}_{k,t} = \mathbf{r}_{k,t}$ and $\mathbf{H}_k = \mathbf{h}_k$, the conditional probability $P(U_{k,t} = u_{k,t} \mid \mathbf{h}_k, \mathbf{r}_{k,t})$ in

(3.37) can thus be expressed as follows

$$P(U_{k,t} = u_{k,t} \mid \mathbf{h}_k, \mathbf{r}_{k,t}) = \begin{cases} 1 & \text{if } \mathbf{r}_{k,t} \in \mathcal{R}^{k,t}_{u_{k,t}}(\mathbf{h}_k) \\ 0 & \text{if } \mathbf{r}_{k,t} \notin \mathcal{R}^{k,t}_{u_{k,t}}(\mathbf{h}_k) \end{cases} \tag{3.39}$$

and the conditional probability of frame error (3.37) can be rewritten as follows

$$P_{f|\mathbf{h}_k} = 1 - \frac{1}{2^N} \sum_{u_{k,t}} \int_{\mathbf{r}_{k,t} \in \mathcal{R}^{k,t}_{u_{k,t}}(\mathbf{h}_k)} L_{u_{k,t}}(\mathbf{r}_{k,t}, \mathbf{h}_k) d\mathbf{r}_{k,t} \tag{3.40}$$

From expression (3.40), it can be concluded that, in order to minimize P_f, each realization of $\mathbf{R}_{k,t}$ should be included in the set $\mathcal{R}^{k,t}_{u_{k,t}}(\mathbf{h}_k)$ associated with the index $u_{k,t}$ that maximizes $L_{u_{k,t}}(\mathbf{r}_{k,t}, \mathbf{h}_k)$. Therefore, the optimum local decision rule, given $\mathbf{R}_{k,t} = \mathbf{r}_{k,t}$ and $\mathbf{H}_k = \mathbf{h}_k$, equals

$$U_{k,t} = \arg \max_{u_{k,t}=0,\dots,L-1} \left\{ L_{u_{k,t}}(\mathbf{r}_{k,t}, \mathbf{h}_k) \right\} \tag{3.41}$$

Then, given the set $\mathcal{C}^1_t$ of all codewords for which $C_t = 1$ and the set $\mathcal{C}^0_t$ of all codewords for which $C_t = 0$, expression (3.38) can be expressed as follows

$$\begin{aligned} & L_{u_{k,t}}(\mathbf{r}_{k,t}, \mathbf{h}_k) \\ & = f_{\mathbf{R}_{k,t}}(\mathbf{r}_{k,t} \mid \mathbf{h}_k, C_t = 1) \times \\ & \quad \sum_{\mathbf{c} \in \mathcal{C}^1_t} \sum_{\mathbf{u}^t_k} \sum_{\mathbf{u}^k} P(\mathbf{U}_0 = \mathbf{c} \mid \mathbf{u}^k, \mathbf{u}^t_k, u_{k,t}) P(\mathbf{U}^k = \mathbf{u}^k \mid \mathbf{c}) P(\mathbf{U}^t_k = \mathbf{u}^t_k \mid \mathbf{h}_k, \mathbf{c}) \\ & \quad + f_{\mathbf{R}_{k,t}}(\mathbf{r}_{k,t} \mid \mathbf{h}_k, C_t = 0) \times \\ & \quad \sum_{\mathbf{c} \in \mathcal{C}^0_t} \sum_{\mathbf{u}^t_k} \sum_{\mathbf{u}^k} P(\mathbf{U}_0 = \mathbf{c} \mid \mathbf{u}^k, \mathbf{u}^t_k, u_{k,t}) P(\mathbf{U}^k = \mathbf{u}^k \mid \mathbf{c}) P(\mathbf{U}^t_k = \mathbf{u}^t_k \mid \mathbf{h}_k, \mathbf{c}) \end{aligned} \tag{3.42}$$

As the previous case, the local decision rule (3.41) can therefore be reformulated as follows

$$U_{k,t} = \arg \max_{u_{k,t}=0,\dots,L-1} \left\{ f_{\mathbf{R}_{k,t}}(\mathbf{r}_{k,t} \mid \mathbf{h}_k, C_t = 0) \left(m_{u_{k,t}} \Lambda^{(k,t)}_{1,0}(\mathbf{r}_{k,t}, \mathbf{h}_k) + b_{u_{k,t}} \right) \right\} \tag{3.43}$$

or equivalently, since $p(\mathbf{r}_{k,t} \mid \mathbf{h}_k, C_t = 0)$ is independent of $u_{k,t}$, as follows

$$U_{k,t} = \arg \max_{u_{k,t}=0,\ldots,L-1} \left\{ m_{u_{k,t}} \Lambda_{1,0}^{(k,t)}(\mathbf{r}_{k,t}, \mathbf{h}_k) + b_{u_{k,t}} \right\}, \tag{3.44}$$

where $\Lambda_{1,0}^{(k,t)}(\mathbf{r}_{k,t}, \mathbf{h}_k)$ is defined in (3.22). However, as opposed to the previous case,

$$m_{u_{k,t}} = \sum_{\mathbf{c}\in C_t^1} \sum_{\mathbf{u}_k^t} \sum_{\mathbf{u}^k} P(\mathbf{U}_0 = \mathbf{c} \mid \mathbf{u}^k, \mathbf{u}_k^t, u_{k,t}) P(\mathbf{U}^k = \mathbf{u}^k \mid \mathbf{c}) P(\mathbf{U}_k^t = \mathbf{u}_k^t \mid \mathbf{h}_k, \mathbf{c}) \tag{3.45}$$

and

$$b_{u_{k,t}} = \sum_{\mathbf{c}\in C_t^0} \sum_{\mathbf{u}_k^t} \sum_{\mathbf{u}^k} P(\mathbf{U}_0 = \mathbf{c} \mid \mathbf{u}^k, \mathbf{u}_k^t, u_{k,t}) P(\mathbf{U}^k = \mathbf{u}^k \mid \mathbf{c}) P(\mathbf{U}_k^t = \mathbf{u}_k^t \mid \mathbf{h}_k, \mathbf{c}) \tag{3.46}$$

Since the coefficients $m_{u_{k,t}}$ and $b_{u_{k,t}}$ are independent of $\mathbf{r}_{k,t}$ but varying with $\mathbf{h}_k$, as can be seen on Fig. 3.2 the optimum decision rule $\gamma_{k,t}(\mathbf{r}_{k,t}, \mathbf{h}_k)$ is a likelihood ratio threshold test, where as opposed to the uncoded case the thresholds are functions of $\mathbf{h}_k$. The values of these thresholds cannot be determined analytically due to the interdependence of these thresholds and other local detector thresholds with respect to the common optimality criterion. The thresholds must therefore be optimized simultaneously using a numerical optimization algorithm in order to determine their optimum values. However, since new local decision rules are only transmitted to the base stations when the average SNR varies at any base station, it is very difficult to optimize such a decision rule since the thresholds do not appear as scalars anymore but as functions of $\mathbf{h}_k$. In fact, we encounter the same difficulty in chapter 2 when we design a bandwidth efficient scheme for the case when the channel state is known at the MSC. Hence, as a sub-optimum alternative, it can be proposed that the thresholds be independent of $\mathbf{h}_k$. However, the number of summations required to evaluate the cost function, which is in this case the average probability of frame error, grows exponentially with the frame size, the number of base stations and the number of bits of resolution per decision made at each base station. Therefore, in section 3.3 sub-optimum schemes are proposed.

3.3 Sub-optimum alternatives

In the previous section, we derived decision rules that minimize the probability of frame error at the output of the MSC. We showed that the complexity associated with either the implementation or the numerical optimization of these decision rules grows exponentially with the frame size. Therefore, in this section, we propose for both considered cases sub-optimum decision rules based on the results from the previous section with some simplifications to keep the complexity to a reasonable level.

It is first assumed that the local decision rules used at the kth base station are identical for all values of t, such that $P(U_{k,1} = l \mid C_1 = j) = \cdots = P(U_{k,N_c} = l \mid C_{N_c} = j)$, and even symmetric, such that $P(U_{k,t} = l \mid C_t = j) = P(U_{k,t} = L - 1 - l \mid C_t = -j + 1)$, for $0 \le l \le L - 1$ and $j = 0, 1$. In fact, for both considered cases we propose to use LLR quantizers as local decision rules which can be defined as follows,

$$U_{k,t} = u_{k,t} \text{ if } t_{k,u_{k,t}} \le \Psi_{1,0}^{(k,t)}(\mathbf{r}_{k,t}, \mathbf{h}_k) < t_{k,u_{k,t}+1}, \tag{3.47}$$

where

$$\Psi_{1,0}^{(k,t)}(\mathbf{r}_{k,t}, \mathbf{h}_k) = \ln\left(\Lambda_{1,0}^{(k,t)}(\mathbf{r}_{k,t}, \mathbf{h}_k)\right) \tag{3.48}$$

and $-\infty = t_{k,0} < \cdots < t_{k,L/2} = 0 < \cdots < t_{k,L} = \infty$ and $t_{k,l} = -t_{k,L-l}$, for $l = 0, \ldots, L-1$. Furthermore, it is assumed that the thresholds contained in the vector $\mathbf{t}_k = [t_{k,1}, \ldots, t_{k,L-1}]^T$ partitioning the LLRs at the kth base station are independent of $\mathbf{h}$, such that new threshold values need only to be transmitted to the base stations when the average SNR varies at any base station.

Then, for both considered cases we propose to use fusion rules that can be implemented using the Viterbi algorithm. As derived in section 3.2.1, when the channel state is known at the MSC and each local decision $U_{k,t}$ is made only based on the information contained in the received signal vector $\mathbf{R}_{k,t}$, the optimum fusion rule presented in (3.10) can be implemented using the Viterbi algorithm. On the other hand, as derived in section 3.2.2, when the channel state is not available at the fusion center, the optimum fusion rule cannot be implemented using the Viterbi algorithm since the local decisions $U_{k,1}, \ldots, U_{k,N_c}$ are dependent on the unknown channel state vector $\mathbf{H}$. Therefore, we propose a sub-optimum fusion rule that ignores the dependence between the local decisions $U_{k,1}, \ldots, U_{k,N_c}$ made at the kth base station. The sub-optimum fusion rule given $\mathbf{U} = \mathbf{u}$ can therefore be expressed as follows

$$\mathbf{U}_0 = \arg\max_{\mathbf{c}\in\mathcal{C}} \left\{ \prod_{t=1}^{N_c} \prod_{k=1}^{N_{BS}} P(U_{k,t} = u_{k,t} \mid C_t = c_t) \right\}. \tag{3.49}$$

The Viterbi algorithm can therefore be used to implement the optimum fusion rule where the branch metric $m_i(s, s')$ associated with the branch starting in state s and ending in state s' in the trellis step i, where $i = 1, \ldots, \frac{N+N_{tail}}{k_c}$, can be defined as follows

$$m_i(s, s') = \prod_{j=1}^{n_c} \prod_{k=1}^{N_{BS}} P(U_{k,(i-1)n_c+j} = u_{k,(i-1)n_c+j} \mid C_{(i-1)n_c+j} = c_j(s, s')) \tag{3.50}$$

given $c_j(s, s')$ is the jth coded bit associated with the set of branches starting in state s and ending in state s'.

Finally, in order to completely define these decision rules, the thresholds contained in the vector $\mathbf{t} = [\mathbf{t}_1, \ldots, \mathbf{t}_{N_{BS}}]^T$ defining the LLR quantizers must be specified. Obviously, it is a computationally intensive problem to select local detector thresholds that minimize directly the average probability of frame error. To reduce the computation difficulty, we propose two optimality criteria that are expected to reduce the probability of frame error at the output of the MSC, which are the pairwise error probability (PEP) and the Mean Square Error (MSE) between the LLRs used by the CSHDD scheme and the OC scheme at the fusion center. We refer respectively to the quantizers optimized with respect to these criteria as the Minimum PEP-LLR (MPEP-LLR) quantizer and the Minimum MSE-LLR (MMSE-LLR) quantizer. It is important to mention that, in both cases, the cost function used to optimize the thresholds is a nonlinear non-convex function of these thresholds. Since the cost functions may have many local minima, we propose to use the ASA algorithm presented in the the previous chapter to perform the optimization. Hence, the two proposed LLR quantizers are first presented and then the tuning of the ASA algorithm for the optimization of the quantizer thresholds is discussed. The numerical evaluation of the proposed cost functions is discussed in Appendix C.

3.3.1 MPEP-LLR quantizer

In this section, we consider the design of LLR quantizers that minimize the PEP at the output of the fusion center. The PEP is the basic component of the probability of frame error union bound [27] and can be defined as the average probability of frame error when $\mathbf{c}^{(1)}$ and $\mathbf{c}^{(2)}$ are the only two codewords in the codebook $\mathcal{C}$. Hence, the PEP of the CSHDD scheme can be defined as follows

$$P_2(\mathbf{c}^{(1)}, \mathbf{c}^{(2)}) = \int_{\mathbf{h}} P_{2|\mathbf{h}}(\mathbf{c}^{(1)}, \mathbf{c}^{(2)}) f_{\mathbf{H}}(\mathbf{h}) d\mathbf{h}, \tag{3.51}$$

where $P_{2|\mathbf{h}}(\mathbf{c}^{(1)}, \mathbf{c}^{(2)})$ represents the PEP associated with the codewords $\mathbf{c}^{(1)}$ and $\mathbf{c}^{(2)}$, given the channel state vector $\mathbf{H} = \mathbf{h}$. Using expression (3.3) defining the conditional probability of frame error, the conditional PEP $P_{2|\mathbf{h}}(\mathbf{c}^{(1)}, \mathbf{c}^{(2)})$ can be formulated as follows

$$\begin{aligned}
&P_{2|\mathbf{h}}(\mathbf{c}^{(1)}, \mathbf{c}^{(2)}) \\
&= \frac{1}{2}\sum_{\mathbf{u}} P(\mathbf{U}_0 = \mathbf{c}^{(1)} \mid \mathbf{h}, \mathbf{u}) P(\mathbf{U} = \mathbf{u} \mid \mathbf{h}, \mathbf{C} = \mathbf{c}^{(2)}) \\
&\qquad + \frac{1}{2}\sum_{\mathbf{u}} P(\mathbf{U}_0 = \mathbf{c}^{(2)} \mid \mathbf{h}, \mathbf{u}) P(\mathbf{U} = \mathbf{u} \mid \mathbf{h}, \mathbf{C} = \mathbf{c}^{(1)}) \\
&= \frac{1}{2}\sum_{\mathbf{u}} P(\mathbf{U}_0 = \mathbf{c}^{(1)} \mid \mathbf{h}, \mathbf{u}) P(\mathbf{U} = \mathbf{u} \mid \mathbf{h}, \mathbf{C} = \mathbf{c}^{(2)}) \\
&\qquad + \frac{1}{2} - \frac{1}{2}\sum_{\mathbf{u}} P(\mathbf{U}_0 = \mathbf{c}^{(1)} \mid \mathbf{h}, \mathbf{u}) P(\mathbf{U} = \mathbf{u} \mid \mathbf{h}, \mathbf{C} = \mathbf{c}^{(1)}) \\
&= \frac{1}{2} + \frac{1}{2}\sum_{\mathbf{u}} P(\mathbf{U}_0 = \mathbf{c}^{(1)} \mid \mathbf{h}, \mathbf{u}) \left[P(\mathbf{U} = \mathbf{u} \mid \mathbf{C} = \mathbf{h}, \mathbf{c}^{(2)}) - P(\mathbf{U} = \mathbf{u} \mid \mathbf{h}, \mathbf{C} = \mathbf{c}^{(1)}) \right].
\end{aligned} \tag{3.52}$$

since $P(\mathbf{U}_0 = \mathbf{c}^{(2)} \mid \mathbf{h}, \mathbf{u}) = 1 - P(\mathbf{U}_0 = \mathbf{c}^{(1)} \mid \mathbf{h}, \mathbf{u})$. Considering the structure of the proposed sub-optimum CSHDD schemes, we prove next that the PEP can be expressed as a function of the Hamming distance d between the two considered codewords.

First, considering $\mathbf{c}^{(1)}$ and $\mathbf{c}^{(2)}$ are the only two codewords in the codebook $\mathcal{C}$, the fusion rule (3.10), for the case when the channel state is known at the MSC, can be reformulated as follows

$$\Lambda^{(0)}_{\mathbf{c}^{(1)},\mathbf{c}^{(2)}}(\mathbf{u},\mathbf{h}) = \prod_{t=1}^{N_c}\prod_{k=1}^{N_{BS}} \frac{P(U_{k,t}=u_{k,t}\mid \mathbf{h}_k, C_t=c_t^{(1)})}{P(U_{k,t}=u_{k,t}\mid \mathbf{h}_k, C_t=c_t^{(2)})} \underset{\mathbf{U}_0=\mathbf{c}^{(2)}}{\overset{\mathbf{U}_0=\mathbf{c}^{(1)}}{\gtrless}} 1, \tag{3.53}$$

while the fusion rule (3.49), for the case when the channel state is unknown at the MSC, can be reformulated as follows

$$\Lambda^{(0)}_{\mathbf{c}^{(1)},\mathbf{c}^{(2)}}(\mathbf{u}) = \prod_{t=1}^{N_c}\prod_{k=1}^{N_{BS}} \frac{P(U_{k,t}=u_{k,t}\mid C_t=c_t^{(1)})}{P(U_{k,t}=u_{k,t}\mid C_t=c_t^{(2)})} \underset{\mathbf{U}_0=\mathbf{c}^{(2)}}{\overset{\mathbf{U}_0=\mathbf{c}^{(1)}}{\gtrless}} 1. \tag{3.54}$$

Considering the set $\mathcal{T}_{uq}$ contains the indexes t for which $c_t^{(1)} \neq c_t^{(2)}$ and the set $\mathcal{T}_{eq}$ contains the indexes t for which $c_t^{(1)} = c_t^{(2)}$, the likelihood ratio $\Lambda^{(0)}_{\mathbf{c}^{(1)},\mathbf{c}^{(2)}}(\mathbf{u},\mathbf{h})$ simplifies as follows

$$\begin{aligned}
&\Lambda^{(0)}_{\mathbf{c}^{(1)},\mathbf{c}^{(2)}}(\mathbf{u},\mathbf{h})\\
&= \prod_{t\in\mathcal{T}_{uq}}\prod_{k=1}^{N_{BS}} \frac{P(U_{k,t}=u_{k,t}\mid \mathbf{h}_k, C_t=c_t^{(1)})}{P(U_{k,t}=u_{k,t}\mid \mathbf{h}_k, C_t=c_t^{(2)})} \prod_{t\in\mathcal{T}_{eq}}\prod_{k=1}^{N_{BS}} \frac{P(U_{k,t}=u_{k,t}\mid \mathbf{h}_k, C_t=c_t^{(1)})}{P(U_{k,t}=u_{k,t}\mid \mathbf{h}_k, C_t=c_t^{(2)})}\\
&= \prod_{t\in\mathcal{T}_{uq}}\prod_{k=1}^{N_{BS}} \frac{P(U_{k,t}=u_{k,t}\mid \mathbf{h}_k, C_t=c_t^{(1)})}{P(U_{k,t}=u_{k,t}\mid \mathbf{h}_k, C_t=c_t^{(2)})},
\end{aligned} \tag{3.55}$$

and similarly the likelihood ratio $\Lambda^{(0)}_{\mathbf{c}^{(1)},\mathbf{c}^{(2)}}(\mathbf{u})$ simplifies as follows

$$\begin{aligned}
&\Lambda^{(0)}_{\mathbf{c}^{(1)},\mathbf{c}^{(2)}}(\mathbf{u})\\
&= \prod_{t\in\mathcal{T}_{uq}}\prod_{k=1}^{N_{BS}} \frac{P(U_{k,t}=u_{k,t}\mid C_t=c_t^{(1)})}{P(U_{k,t}=u_{k,t}\mid C_t=c_t^{(2)})} \prod_{t\in\mathcal{T}_{eq}}\prod_{k=1}^{N_{BS}} \frac{P(U_{k,t}=u_{k,t}\mid C_t=c_t^{(1)})}{P(U_{k,t}=u_{k,t}\mid C_t=c_t^{(2)})}\\
&= \prod_{t\in\mathcal{T}_{uq}}\prod_{k=1}^{N_{BS}} \frac{P(U_{k,t}=u_{k,t}\mid C_t=c_t^{(1)})}{P(U_{k,t}=u_{k,t}\mid C_t=c_t^{(2)})},
\end{aligned} \tag{3.56}$$

proving that both fusion rules are independent of the local decisions associated with the indexes $t \in \mathcal{T}_{eq}$. Let $\mathbf{U}^{(uq)}$ contain all local decisions $U_{k,t}$ for which $t \in \mathcal{T}_{uq}$ and $\mathbf{U}^{(eq)}$ contain all local decisions $U_{k,t}$ for which $t \in \mathcal{T}_{eq}$. Since the fusion rules are independent of the local decisions in the vector $\mathbf{U}^{(eq)}$, the conditional probability $P(\mathbf{U}_0 = \mathbf{u}_0|\mathbf{h}, \mathbf{u})$ reduces to $P(\mathbf{U}_0 = \mathbf{u}_0|\mathbf{h}, \mathbf{U}^{(uq)} = \mathbf{u}^{(uq)})$ and the conditional PEP (3.52) can be simplified as follows

$$\begin{aligned} P_{2|\mathbf{h}}(\mathbf{c}^{(1)}, \mathbf{c}^{(2)}) &= \frac{1}{2} + \frac{1}{2} \sum_{\mathbf{u}^{(uq)}} P(\mathbf{U}_0 = \mathbf{c}^{(1)}|\mathbf{h}, \mathbf{U}^{(uq)} = \mathbf{u}^{(uq)}) \times \\ & \left[P(\mathbf{U}^{(uq)} = \mathbf{u}^{(uq)} \mid \mathbf{h}, \mathbf{C} = \mathbf{c}^{(2)}) - P(\mathbf{U}^{(uq)} = \mathbf{u}^{(uq)} \mid \mathbf{h}, \mathbf{C} = \mathbf{c}^{(1)})\right], \end{aligned} \tag{3.57}$$

since the local decisions are conditionally independent.

In addition, since the local decision rules are assumed to be even symmetric, $P(\mathbf{U}^{(uq)} = \mathbf{u}^{(uq)}|\mathbf{h}, \mathbf{C} = \mathbf{c}^{(m)}) = P(\mathbf{U}^{(uq)} = \hat{\mathbf{u}}^{(uq)}|\mathbf{h}, \mathbf{C} = \hat{\mathbf{c}}^{(m)})$ and $P(\mathbf{U}_0 = \mathbf{c}^{(m)}|\mathbf{h}, \mathbf{U}^{(uq)} = \mathbf{u}^{(uq)}) = P(\mathbf{U}_0 = \hat{\mathbf{c}}^{(m)}|\mathbf{h}, \mathbf{U}^{(uq)} = \hat{\mathbf{u}}^{(uq)})$ as long as $\hat{u}_{k,t} = L - 1 - u_{k,t}$ if $\hat{c}_t^{(m)} = 1 - c_t^{(m)}$, where $m = 1, 2$. Hence, the conditional PEP can be reformulated as a function of $\hat{\mathbf{u}}^{(uq)}$ and $\hat{\mathbf{c}}^{(m)}$ as follows

$$\begin{aligned} P_{2|\mathbf{h}}(\mathbf{c}^{(1)}, \mathbf{c}^{(2)}) &= \frac{1}{2} + \frac{1}{2} \sum_{\hat{\mathbf{u}}^{(uq)}} P(\mathbf{U}_0 = \hat{\mathbf{c}}^{(1)}|\mathbf{h}, \mathbf{U}^{(uq)} = \hat{\mathbf{u}}^{(uq)}) \times \\ & \left[P(\mathbf{U}^{(uq)} = \hat{\mathbf{u}}^{(uq)} \mid \mathbf{h}, \mathbf{C} = \hat{\mathbf{c}}^{(2)}) - P(\mathbf{U}^{(uq)} = \hat{\mathbf{u}}^{(uq)} \mid \mathbf{h}, \mathbf{C} = \hat{\mathbf{c}}^{(1)})\right]. \end{aligned} \tag{3.58}$$

Since the sum is performed with respect to all possible values of $\hat{\mathbf{u}}^{(uq)}$, $\hat{\mathbf{u}}^{(uq)}$ can be replaced by $\mathbf{u}^{(uq)}$ in (3.58) without affecting the result. It shows that the PEP is independent of the specific choice of codewords $\mathbf{c}^{(1)}$ and $\mathbf{c}^{(2)}$ as long as $c_t^{(1)} \neq c_t^{(2)}$ for $t \in \mathcal{T}_{uq}$. It can therefore be assumed without loss of generality that $\mathbf{c}^{(2)}$ is the all zero codeword and $\mathbf{c}^{(1)}$ differs from the all zero codeword at $t \in \mathcal{T}_{uq}$.

Finally, since the local decision rules at the kth base station are assumed identical for all time indexes t, it can be assumed that $P_k^{lj} = P(U_{k,1} = l \mid \mathbf{h}_k, C_1 = j) = \cdots = P(U_{k,N_c} = l \mid \mathbf{h}_k, C_{N_c} = j)$, such that

$$P(\mathbf{U}^{(uq)} = \mathbf{u}^{(uq)} \mid \mathbf{h}, \mathbf{C} = \mathbf{c}^{(1)}) = \prod_{k=1}^{N_{BS}} \prod_{l=0}^{L-1} (P_k^{l1}(\mathbf{h}_k))^{n_{k,l}} \tag{3.59}$$

and

$$P(\mathbf{U}^{(uq)} = \mathbf{u}^{(uq)} \mid \mathbf{h}, \mathbf{C} = \mathbf{c}^{(2)}) = \prod_{k=1}^{N_{BS}} \prod_{l=0}^{L-1} (P_k^{l0}(\mathbf{h}_k))^{n_{k,l}} \tag{3.60}$$

where $\mathbf{n}_k = [n_{k,0}, \ldots, n_{k,L-1}]^T$ and $n_{k,0} + \cdots + n_{k,L-1} = d$, since $n_{k,l}$ represents the number of local decisions equal to l at the kth base station for $t \in \mathcal{T}_{uq}$ and the set $\mathcal{T}_{uq}$ contains d elements. The coefficients $n_{k,0}, \ldots, n_{k,L-1}$ can be interpreted as the number of elements in sets, partitioning the set $\mathcal{T}_{uq}$, which contain the time indexes t for which the kth base station local decisions equal respectively $0, \ldots, L-1$. From [28], it is known that the number of distinct partitions of $\mathcal{T}_{uq}$ into L sets of $n_{k,0}, \ldots, n_{k,L-1}$ elements equals $M(\mathbf{n}_k)$ which is defined as follows

$$M(\mathbf{n}_k) = \frac{d!}{n_{k,0}! \cdots n_{k,L-1}!} \tag{3.61}$$

and is called the multinominal coefficient. Hence, $M(\mathbf{n}_1) \cdots M(\mathbf{n}_{N_{BS}})$ realizations of $\mathbf{U}^{(uq)}$ provide the same $\mathbf{n}_1, \ldots, \mathbf{n}_{N_{BS}}$ and consequently the same probability $P(\mathbf{U}^{(uq)} = \mathbf{u}^{(uq)} \mid \mathbf{h}, \mathbf{C} = \mathbf{c}^{(m)})$. The conditional PEP (3.52) can therefore be reformulated as follows

$$\begin{aligned} P_{2|\mathbf{h}}(d) \;=\; & \frac{1}{2} + \frac{1}{2} \sum_{\mathbf{n}_1} \cdots \sum_{\mathbf{n}_{N_{BS}}} M(\mathbf{n}_1) \cdots M(\mathbf{n}_{N_{BS}}) 1(\mathbf{n}_1, \ldots, \mathbf{n}_{N_{BS}}) \times \\ & \left[\prod_{k=1}^{N_{BS}} \prod_{l=0}^{L-1} (P_k^{l0}(\mathbf{h}_k))^{n_{k,l}} - \prod_{k=1}^{N_{BS}} \prod_{l=0}^{L-1} (P_k^{l1}(\mathbf{h}_k))^{n_{k,l}} \right], \end{aligned} \tag{3.62}$$

where $1(\mathbf{n}_1, \ldots, \mathbf{n}_{N_{BS}})$ is an indicator function which, for the case when the channel state is known at the MSC, equals

$$1(\mathbf{n}_1, \ldots, \mathbf{n}_{N_{BS}}) = \begin{cases} 1 & \text{if } \prod_{k=1}^{N_{BS}} \prod_{l=0}^{L-1} \left(\frac{P_k^{l1}(\mathbf{h}_k)}{P_k^{l0}(\mathbf{h}_k)} \right)^{n_{k,l}} \geq 1 \\ 0 & \text{if } \prod_{k=1}^{N_{BS}} \prod_{l=0}^{L-1} \left(\frac{P_k^{l1}(\mathbf{h}_k)}{P_k^{l0}(\mathbf{h}_k)} \right)^{n_{k,l}} < 1 \end{cases} \tag{3.63}$$

while, for the case when the channel state is unknown at the MSC, equals

$$1(\mathbf{n}_1, \ldots, \mathbf{n}_{N_{BS}}) = \begin{cases} 1 & \text{if } \prod_{k=1}^{N_{BS}} \prod_{l=0}^{L-1} \left(\frac{E_{\mathbf{H}_k}[P_k^{l1}(\mathbf{H}_k)]}{E_{\mathbf{H}_k}[P_k^{l0}(\mathbf{H}_k)]} \right)^{n_{k,l}} \geq 1 \\ 0 & \text{if } \prod_{k=1}^{N_{BS}} \prod_{l=0}^{L-1} \left(\frac{E_{\mathbf{H}_k}[P_k^{l1}(\mathbf{H}_k)]}{E_{\mathbf{H}_k}[P_k^{l0}(\mathbf{H}_k)]} \right)^{n_{k,l}} < 1 \end{cases} . \tag{3.64}$$

Expression (3.62) proves that the PEP can be expressed as a function of the Hamming distance d between the two codewords $\mathbf{c}^{(1)}$ and $\mathbf{c}^{(2)}$.

It is assumed that the PEP with respect to the two closer codewords in $\mathcal{C}$ represents the more dominant error event. Therefore, by optimizing the decision rules with respect to the PEP associated with the free Hamming distance d_f, it can be expected that the probability of frame error will be reduced. The numerical evaluation of the PEP as a function of the thresholds contained in $\mathbf{t}$ is discussed in Appendix C.

3.3.2 MMSE-LLR quantizer

In this section, we consider the design of LLR quantizers that minimize a MSE criterion, generalizing results from [15] to distributed detection system using channel coding.

As mentioned in Appendix D, the decision rule of the OC scheme is a maximum likelihood rule which can be reformulated as a function of the LLR $\Psi_{1,0}^{(k,t)}(\mathbf{r}_{k,t}, \mathbf{h}_k)$, defined in (3.48), as follows

$$\mathbf{U}_0 = \arg\max_{\mathbf{c}\in\mathcal{C}} \left\{ \sum_{t=1}^{N_c} \sum_{k=1}^{N_{BS}} (-1)^{c_t+1} \Psi_{1,0}^{(k,t)}(\mathbf{r}_{k,t}, \mathbf{h}_k) \right\}. \tag{3.65}$$

Hence, the OC scheme decision rule can therefore be presented as a function of the LLR $\Psi_{1,0}^{(k,t)}(\mathbf{r}_{k,t}, \mathbf{h}_k)$, which is the input to the kth base station LLR quantizer used by the proposed CSHDD schemes. Similarly, the fusion rule (3.10) used by the proposed sub-optimum CSHDD scheme, for the case when the channel state is known at the MSC, is equivalent to [29]

$$\mathbf{U}_0 = \arg\max_{\mathbf{c}\in\mathcal{C}} \left\{ \sum_{t=1}^{N_c} \sum_{k=1}^{N_{BS}} (-1)^{c_t+1} \Psi_{1,0}^{(k,t)}(u_{k,t}, \mathbf{h}_k) \right\}, \tag{3.66}$$

where

$$\Psi_{1,0}^{(k,t)}(u_{k,t}, \mathbf{h}_k) = \ln \left(\frac{P(U_{k,t} = u_{k,t} \mid \mathbf{h}_k, C_t = 1)}{P(U_{k,t} = u_{k,t} \mid \mathbf{h}_k, C_t = 0)} \right). \tag{3.67}$$

In addition, the fusion rule (3.49) used by the sub-optimum CSHDD scheme, for the case when the channel state is unknown at the MSC, is equivalent to [29]

$$\mathbf{U}_0 = \arg\max_{\mathbf{c} \in \mathcal{C}} \left\{ \sum_{t=1}^{N_c} \sum_{k=1}^{N_{BS}} (-1)^{c_t+1} \Psi_{1,0}^{(k,t)}(u_{k,t}) \right\}, \tag{3.68}$$

where

$$\Psi_{1,0}^{(k,t)}(u_{k,t}) = \ln \left(\frac{P(U_{k,t} = u_{k,t} \mid C_t = 1)}{P(U_{k,t} = u_{k,t} \mid C_t = 0)]} \right). \tag{3.69}$$

It is important to mention that the LLRs $\Psi_{1,0}^{(k,t)}(u_{k,t}, \mathbf{h}_k)$ and $\Psi_{1,0}^{(k,t)}(u_{k,t})$, used by the CSHDD schemes, are functions of the LLR $\Psi_{1,0}^{(k,t)}(\mathbf{r}_{k,t}, \mathbf{h}_k)$, since the value of the local decision $U_{k,t}$ is determined, given $\mathbf{R}_{k,t} = \mathbf{r}_{k,t}$ and $\mathbf{H}_k = \mathbf{h}_k$, by a local decision rule that quantizes the LLR $\Psi_{1,0}^{(k,t)}(\mathbf{r}_{k,t}, \mathbf{h}_k)$. Hence, by comparing expression (3.65) with either (3.66) or (3.68), the LLRs $\Psi_{1,0}^{(k,t)}(u_{k,t}, \mathbf{h}_k)$ and $\Psi_{1,0}^{(k,t)}(u_{k,t})$ can be interpreted as discretized representations of the LLR $\Psi_{1,0}^{(k,t)}(\mathbf{r}_{k,t}, \mathbf{h}_k)$, used by the OC scheme. Furthermore, the thresholds contained in the vector $\mathbf{t}$ defining the LLR quantizers used by the CSHDD schemes determine the discrete values that take the LLRs $\Psi_{1,0}^{(k,t)}(u_{k,t}, \mathbf{h}_k)$ and $\Psi_{1,0}^{(k,t)}(u_{k,t})$ as well as the mapping between $\Psi_{1,0}^{(k,t)}(\mathbf{r}_{k,t}, \mathbf{h}_k)$ and these discrete values.

Hence, given $\Psi_{1,0}^{(k,t)}(\mathbf{r}_{k,t}, \mathbf{h}_k)$, $\Psi_{1,0}^{(k,t)}(u_{k,t}, \mathbf{h}_k)$ and $\Psi_{1,0}^{(k,t)}(u_{k,t})$ are respectively realization of the random variables $\Psi_{1,0}^{(k,t)}(\mathbf{R}_{k,t}, \mathbf{H}_k)$, $\Psi_{1,0}^{(k,t)}(U_{k,t}, \mathbf{H}_k)$ and $\Psi_{1,0}^{(k,t)}(U_{k,t})$, we propose that the thresholds defining the LLR quantizers, used by the CSHDD schemes, be adjusted in order for the LLRs $\Psi_{1,0}^{(k,t)}(U_{k,t}, \mathbf{H}_k)$ and $\Psi_{1,0}^{(k,t)}(U_{k,t})$ to approximate under a MMSE criterion the LLR $\Psi_{1,0}^{(k,t)}(\mathbf{R}_{k,t}, \mathbf{H}_k)$. The MSE between the LLR used by the OC scheme and the LLRs used by the CSHDD schemes can be defined as follows

$$\varepsilon_{k,t} = E_{\mathbf{R}_{k,t}, \mathbf{H}_k, U_{k,t}} \left[\left(\Psi_{1,0}^{(k,t)}(\mathbf{R}_{k,t}, \mathbf{H}_k) - \Psi_{1,0}^{(k,t)}(U_{k,t}, \mathbf{H}_k) \right)^2 \right], \tag{3.70}$$

for the case when the channel state is known at the MSC, and as follows

$$\varepsilon_{k,t} = E_{\mathbf{R}_{k,t}, \mathbf{H}_k, U_{k,t}} \left[\left(\Psi_{1,0}^{(k,t)}(\mathbf{R}_{k,t}, \mathbf{H}_k) - \Psi_{1,0}^{(k,t)}(U_{k,t}) \right)^2 \right], \tag{3.71}$$

for the case when the channel state is unknown at the MSC. Hence, this quantizer is expected to minimize indirectly the probability of frame error by minimizing the MSE between the LLRs used by the OC and CSHDD schemes. The numerical evaluation of the MMSE as a function of the thresholds $\mathbf{t}$ is discussed in Appendix C.

3.3.3 Tuning of the ASA algorithm for the optimization of the CSHDD schemes

In this section, we explain how to tune the ASA algorithm to optimize the thresholds defining the LLR quantizers used by the proposed sub-optimum CSHDD schemes. The reader is referred to chapter 2 for a detailed description of the ASA algorithm. Hence, as discussed in chapter 2, the convergence of the ASA algorithm is influenced principally by the selection of the domain $\mathcal{X}$, the tuning of the cooling schedule parameter c_{gen} and the tuning of the stopping rule parameters.

As for the uncoded case, it is necessary to select a domain $\mathcal{X}$ having a finite range but still includes the global minimum. Considering the fact that the PEP and MSE of CSHDD schemes using L quantization levels are respectively upper bounded by the PEP and MSE of CSHDD schemes using $L-1$ quantization levels, the domain $\mathcal{X}$ can be selected such that it contains all possible values of $\mathbf{t}$ that satisfy the inequalities $-L_k < t_{k,1} < t_{k,2} < \cdots < t_{k,L-2} < t_{k,L-1} < L_k$ $k = 1, \ldots, N_{BS}$.

Hence, for CSHDD schemes using MMSE-LLR quantizers, the value of L_k is selected to eliminate the region of the domain $\mathcal{T}$, where $P(U_k = 0)$ and $P(U_k = L-1)$ approach 0. However, it is also necessary to make sure that the domain still includes the optimum solution. The value of L_k is thus selected such that the thresholds $t_{k,1}$ and $t_{k,L-1}$ be able to take values for which $P(U_{k,t} = 0)$ and $P(U_{k,t} = L-1)$ are at least as low as $\alpha = 10^{-10}$.

However, for CSHDD schemes using MPEP-LLR quantizers, it is possible to limit $\mathcal{X}$ to a smaller region of $\mathcal{T}$ by using the fact that the PEP after the threshold optimization can be expressed as follows

$$P_2^*(d) = \int_\omega P_{2|\omega}^*(d) f_{\mathbf{\Omega}}(\omega) d\omega = \int_\omega I(\omega) d\omega, \tag{3.72}$$

where $P_{2|\omega}^*(d)$ represents the PEP after the threshold optimization conditioned on $\mathbf{\Omega} = \omega$, $\mathbf{\Omega} = [\Omega_1, \ldots, \Omega_{N_{BS}}]^T$ given $\Omega_k = \sqrt{\sum_{n=1}^{N_R} |H_{k,n}|^2}$ and $f_{\mathbf{\Omega}}(\omega) = f_{\Omega_1}(\omega_1) \cdots f_{\Omega_{N_{BS}}}(\omega_{N_{BS}})$ given $f_{\Omega_k}(\omega_k)$ is the PDF of Ω_k defined in (B.8). Hence, the value of L_k is selected such that the thresholds $t_{k,1}$ and $t_{k,L-1}$ be able to take values for which $P(U_{k,t} = 0|\mathbf{h}_k)$ and $P(U_{k,t} = L-1|\mathbf{h}_k)$ are at least as low as $\alpha = 10^{-10}$, for all values of ω_k for which the integrand of (3.72) can possibly take a value that has an impact on the results of the integral within the desired precision. Since the probability $P(U_{k,t} = 0|\mathbf{h}_k)$ and $P(U_{k,t} = L-1|\mathbf{h}_k)$ are monotonically increasing functions of ω_k, it is first necessary to determine a value of ω_k over which the integrand $I(\omega)$ is lower than the desired precision independently of $\omega_1 \ldots \omega_{k-1}, \omega_{k+1} \ldots, \omega_{N_{BS}}$. Then, as for the uncoded case, we use this value of ω_k in expression (2.72) to determine the limit L_k. It is obviously necessary to use an upper bound $U(\omega_k)$ to the integrand $I(\omega)$, independent of $\omega_1 \ldots \omega_{k-1}, \omega_{k+1} \ldots, \omega_{N_{BS}}$, since the optimum thresholds are not known a priori.

For instance, for a CSHDD scheme for which the channel state vector $\mathbf{H}$ is known at the MSC, the conditional PEP $P^*_{2|\omega}(d)$ can be upper bounded by the conditional PEP of a CSHDD scheme for which only the local decisions of the kth base station are considered and hard decisions are made at the local detectors. This represents the conditional PEP of a single base station scheme using a maximum likelihood rule based on hard decisions and equals [30]

$$P_{h|\omega_k}(d) = \left(\left\lfloor\frac{d}{2}\right\rfloor - \frac{d-1}{2}\right)\binom{d}{\lfloor\frac{d}{2}\rfloor} P_{s|\omega_k}^{\frac{d}{2}}(1-P_{s|\omega_k})^{\frac{d}{2}} + \sum_{i=\lfloor\frac{d}{2}\rfloor+1}^{d}\binom{d}{k} P_{s|\omega_k}^{i}(1-P_{s|\omega_k})^{d-i}, \tag{3.73}$$

where $\lfloor x \rfloor$ represents the greatest integer $\leq x$ and $P_{s|\omega_k}$ represents the conditional probability of error on the local decisions which equals either $P(U_{k,t} = 0|\mathbf{h}_k C_t = 1)$ or $P(U_{k,t} = 1|\mathbf{h}_k C_t = 0)$. Hence, in this case, we use the upper bound

$$U(\omega_k) = P_{h|\omega_k}(d) f_{\Omega_k}(\omega_k) \prod_{k' \neq k} \max_{\omega_{k'}} \left\{ f_{\Omega_{k'}}(\omega_{k'}) \right\} \tag{3.74}$$

where $k' = 1, \ldots, N_{BS}$. However, if the channel state vector $\mathbf{H}$ is unknown at the MSC, this upper bound is not valid and it becomes more difficult to find a relatively tight upper bound $U(\omega_k)$. Hence, in these cases, we use the fact that $P^*_{2|\omega} < 0.5$ and set

$$U(\omega_k) = 0.5 f_{\Omega_k}(\omega_k) \prod_{k' \neq k} \max_{\omega_{k'}} \left\{ f_{\Omega_{k'}}(\omega_{k'}) \right\} \tag{3.75}$$

where $k' = 1, \ldots, N_{BS}$.

The ASA algorithm also requires that the cooling schedule parameter c_{gen} be adjusted. As mentioned when we discussed the tuning of the ASA options for the uncoded case, c_{gen} must be adjusted as shown in (2.77) which is a function of the initial temperature $T(0)$, the desired final temperature $T(k_f)$ and the desired number of annealing steps k_f. It is suggested that the initial temperature $T(0)$ be set to 1 and the final temperature $T(k_f)$ be set to 10^{-10} although it can be necessary to adjust these parameters differently for certain problems. However, the value of k_f need to be selected experimentally. Table 3.1 presents for both the MPEP and MMSE quantizers the values of k_f for which we obtain a good trade-off between efficiency and accuracy.

Finally, it is necessary to determine when to stop the ASA algorithm. As for the uncoded case, the only stopping rule parameter that we consider is N_{gen}^{max}, which represents the maximum number of sampling point generated, since it is probably the more secure criterion. Given we know that the ASA algorithm is in the final stage of the search when the number of iterations is larger than k_f, we can expect relatively good results by setting N_{gen}^{max} equal to k_f but the accuracy can be improved by increasing the value of N_{gen}^{max}, with respect to k_f.

N_{BS}	L	D	k_f
2	2	0	0
	4	2	$2.0 \cdot 10^3$
	8	6	—

a)

N_{BS}	L	D	k_f
-	2	0	0
-	4	1	$1.0 \cdot 10^3$
-	8	3	$4.0 \cdot 10^3$

b)

Table 3.1 Values of k_f used for the optimization of the different LLR quantizers used by the proposed sub-optimum CSHDD schemes: a) MPEP-LLR quantizer, b) MMSE-LLR quantizer

3.4 Computer Simulation Results

In this section, we study the performances of the proposed sub-optimum CSHDD schemes, in term of BER and FER, for a 16-states, rate 1/2 convolution code with generator polynomials $G_1(D) = 1 + D^3 + D^4$ and $G_2(D) = 1 + D + D^2 + D^4$ [30]. It is important to mention that the considered convolution code is designed for an AWGN channel and may not necessarily be optimum for a quasi-static fading channel. We consider only 2 base station CSHDD schemes but study the performances for different numbers of receiving antennas and different numbers of quantization levels. In addition, since the average SNR received at each base station is dependent on the mobile unit location in the cellular network, as well as on power control, we study the performances of the CSHDD schemes for the case when the average received SNR is equal both base stations as well as for the case when the average received SNR differs at both base stations.

Before evaluating the performances of the designed CSHDD schemes, it is interesting to estimate the potential gain that can be achieved by the CSHDD schemes with respect to CHM. The potential gain is obtained by evaluating the difference between the performances of the CHM and OC schemes (see Appendix D). Table 3.2 presents the potential gain in SNR achievable, when operating at $FER = 10^{-3}$ and $BER = 10^{-4}$, by 2 base stations CSHDD schemes. Similarly to the uncoded case, the potential gain decreases when the difference in average SNR between the two base stations and the number of receiving antennas at each base station increases. However, it is clear from these results that important gains can be made by using more sophisticated techniques such as the proposed CSHDD schemes.

Unfortunately, the evaluation of the probability of frame error and bit error is a computationally expensive problem that grows exponentially with the frame size. In addition, bounding techniques have shown to provide, for the centralized case, very loose bounds for the probability of frame and bit error over quasi-static channels [27]. Therefore, we constructed a software simulator in order to estimate the probability of frame and bit error of the CSHDD schemes and reference schemes through Monte Carlo simulations. The considered reference schemes are the OC scheme as well as CHM scheme which are presented in Appendix D.

N_R	ΔSNR (dB)	Pot. gain (dB)
1	0	16.5
	6	13.4
2	0	9.6
	6	6.6
3	0	7.3
	6	4.5

a)

N_R	ΔSNR (dB)	Pot. gain (dB)
1	0	19.7
	6	16.6
2	0	10.9
	6	8.0
3	0	8.1
	6	5.3

b)

Table 3.2 Potential gain for 2 base stations CSHDD schemes with respect to the CHM scheme as a function of $\Delta SNR = \frac{SNR_1}{SNR_2}$: a) at $FER = 10^{-3}$ b) at $BER = 10^{-4}$

3.4.1 Computer simulation implementation

In order to estimate the performances of the proposed sub-optimum CSHDD schemes and reference schemes, a software simulator was implemented and is included on the compact disk provided with this thesis. However, it is important to mention that, before using the software simulator to estimate the performances of the CSHDD schemes, the local detector thresholds were optimized using the ASA algorithm presented in the previous chapter. More precisely, the ASAMIN Matlab gateway routine to Ingber's ASA C code is used to determine the local detector thresholds minimizing the MSE and PEP criterion proposed in section 3.3. The evaluation of the MSE and PEP is discussed in Appendix C. In addition, the optimization software is also included on the compact disk provided with this thesis.

The software simulator is made of two modules which are the communication system module and the performance estimation module. In the next two sub-sections, we describe in details the role of both modules and their interactions.

A. Communication system module

The communication system module reproduces the transmission of a frame in the different considered handoff macrodiversity scheme which can be either the CSHDD scheme, the OC scheme or the CHM scheme. In addition, the communication system module compares the detected information frame with the transmitted information frame to determine the number of bit errors obtained. This number of bit errors is transmitted at the performance estimation module to form the basic experiment used to estimate the performances of the tested scheme.

For the CSHDD scheme, the communication system module implements the system presented in section 3.1 for $N = 130$ and $N_{tail} = 4$, considering the convolution code investigated. The local detector LLRs $\Psi_{1,0}^{(k,t)}(\mathbf{r}_{k,t}, \mathbf{h}_k)$ $k = 1, \ldots, N_{BS}$ $t = 1, \ldots, N_c$ are evaluated, using the expression derived in Appendix A, and quantized using the optimized thresholds to determine the local decisions. The fusion rule is implemented using a Viterbi algorithm where no decision is made on any of the transmitted bits before the end of the trellis is reached. In the Viterbi algorithm, the branch metric $\hat{m}_i(s, s')$ associated with the branch starting in state s and ending in state s' in the trellis step i, where $i = 1, \ldots, \frac{N+N_{tail}}{k_c}$, equals

$$\hat{m}_i(s, s') = \sum_{j=1}^{n_c} (-1)^{c_j(s,s')+1} \sum_{k=1}^{N_{BS}} \ln \left(\frac{P(U_{k,(i-1)n_c+j} = u_{k,(i-1)n_c+j} \mid \mathbf{h}_k, C_{(i-1)n_c+j} = 1) + \epsilon}{P(U_{k,(i-1)n_c+j} = u_{k,(i-1)n_c+j} \mid \mathbf{h}_k, C_{(i-1)n_c+j} = 0) + \epsilon} \right), \tag{3.76}$$

when the channel state is known at the MSC, and

$$\hat{m}_i(s, s') = \sum_{j=1}^{n_c} (-1)^{c_j(s,s')+1} \sum_{k=1}^{N_{BS}} \ln \left(\frac{P(U_{k,(i-1)n_c+j} = u_{k,(i-1)n_c+j} \mid C_{(i-1)n_c+j} = 1) + \epsilon}{P(U_{k,(i-1)n_c+j} = u_{k,(i-1)n_c+j} \mid C_{(i-1)n_c+j} = 0) + \epsilon} \right), \tag{3.77}$$

when the channel state is unknown at the MSC, where the constant $\epsilon = 10^{-15}$ ensures there is no division per 0 and $c_j(s, s')$ is the jth coded bit associated with the set of branches starting in state s and ending in state s'. It is important to mention that

$$P(U_{k,t} = u_{k,t} \mid \mathbf{H}_k, C_t) = F_{\Psi_{1,0}^{(k,t)}(\mathbf{R}_{k,t},\mathbf{H}_k)}(t_{k,u_{k,t}+1} \mid \mathbf{H}_k, C_t) - F_{\Psi_{1,0}^{(k,t)}(\mathbf{R}_{k,t},\mathbf{H}_k)}(t_{k,u_{k,t}} \mid \mathbf{H}_k, C_t) \tag{3.78}$$

and

$$P(U_{k,t} = u_{k,t} \mid C_t) = F_{\Psi_{1,0}^{(k,t)}(\mathbf{R}_{k,t},\mathbf{H}_k)}(t_{k,u_{k,t}+1} \mid C_t) - F_{\Psi_{1,0}^{(k,t)}(\mathbf{R}_{k,t},\mathbf{H}_k)}(t_{k,u_{k,t}} \mid C_t), \tag{3.79}$$

where the CDF $F_{\Psi_{1,0}^{(k,t)}(\mathbf{R}_{k,t},\mathbf{H}_k)}(x \mid \mathbf{H}_k, C_t)$ and $F_{\Psi_{1,0}^{(k,t)}(\mathbf{R}_{k,t},\mathbf{H}_k)}(x \mid C_t)$ are defined in Appendix A.

B. Performance estimation module

The performance estimation module uses iteratively the communication system module to estimate the FER and BER of the considered communication system as well as the degree of accuracy of these estimates. In addition, the performance estimation module initializes the simulation and periodically saves partial results to avoid large data losses in case of system crash.

In the first part of the performance estimation, new frames are transmitted until $N_e = 100$ frame errors are detected. Given N_f frames were necessary to obtain the 100 frame errors, a new experiment is formed where N_f frames are transmitted and the number of detected bit errors X_i^b and frame errors X_i^f are recorded. This experiment is performed $N_{rep} = 10$ times such that estimates of the mean and variance of X^b and

X^f can be estimated, where X^b and X^f are random variables representing respectively the number of frame errors and bit errors when N_f frames are transmitted. In fact, assuming X denotes either X^b or X^f and X_i denotes either X_i^b or X_i^f, an estimate of the mean μ_X can be evaluated as follows

$$\hat{\mu}_X(X_1, \ldots, X_{N_{rep}}) = \frac{1}{N_{rep}} \sum_{i=1}^{N_{rep}} X_i \tag{3.80}$$

and an estimate of the variance σ_X^2 can be evaluated as follows

$$\hat{\sigma}_X^2(X_1, \ldots, X_{N_{rep}}) = \frac{1}{N_{rep}} \sum_{i=1}^{N_{rep}} \left(X_i - \hat{\mu}_X(X_1, \ldots, X_{N_{rep}})\right)^2. \tag{3.81}$$

Furthermore, the mean and variance estimates can be used to estimate the standard deviation of the error $e(X_1, \ldots, X_{N_{rep}}) = \mu_X - \hat{\mu}_X(X_1, \ldots, X_{N_{rep}})$. Since $E[e(X_1, \ldots, X_{N_{rep}})] = 0$, the variance of $e(X_1, \ldots, X_{N_{rep}})$ simplifies as follows

$$\begin{aligned} \sigma_e^2(X_1, \ldots, X_{N_{rep}}) &= Var\left\{e(X_1, \ldots, X_{N_{rep}})\right\} = Var\left\{\hat{\mu}_X(X_1, \ldots, X_{N_{rep}})\right\} \\ &= Var\left\{\frac{1}{N_{rep}} \sum_{i=1}^{N_{rep}} X_i\right\} = \frac{1}{N_{rep}^2} Var\left\{\sum_{i=1}^{N_{rep}} X_i\right\}. \end{aligned} \tag{3.82}$$

Since all experiments are i.i.d.,

$$Var\left\{\sum_{i=1}^{N_{rep}} X_i\right\} = \sum_{i=1}^{N_{rep}} Var\left\{X_i\right\} = N_{rep}\sigma_X^2, \tag{3.83}$$

$$\sigma_e^2(X_1, \ldots, X_{N_{rep}}) = \frac{\sigma_X^2}{N_{rep}}, \tag{3.84}$$

such that the error standard deviation $\sigma_e(X_1, \ldots, X_{N_{rep}})$ can be estimated as follows

$$\hat{\sigma}_e(X_1, \ldots, X_{N_{rep}}) = \frac{\hat{\sigma}_X(X_1, \ldots, X_{N_{rep}})}{\sqrt{N_{rep}}} = \frac{\sqrt{\sum_{i=1}^{N_{rep}} \left(X_i - \hat{\mu}_X(X_1, \ldots, X_{N_{rep}})\right)^2}}{N_{rep}} \tag{3.85}$$

Hence, given $\hat{\mu}_{X^f}(X_1^f, \ldots, X_{N_{rep}}^f)$ represents an estimate of the mean number of frame errors when N_f frames are transmitted, the FER can be estimated by the performance estimation module as follows

$$FER = \frac{\hat{\mu}_{X^f}(X_1^f, \ldots, X_{N_{rep}}^f)}{N_f} = \frac{1}{N_f N_{rep}} \sum_{i=1}^{N_{rep}} X_i^f \tag{3.86}$$

and the standard deviation on the FER estimate error can be estimated as follows

$$\hat{\sigma}_{FER} = \frac{\hat{\sigma}_e(X_1^f, \ldots, X_{N_{rep}}^f)}{N_f} = \frac{\sqrt{\sum_{i=1}^{N_{rep}} \left(X_i^f - \hat{\mu}_{X^f}(X_1^f, \ldots, X_{N_{rep}}^f)\right)^2}}{N_f N_{rep}}. \tag{3.87}$$

Similarly, given $\hat{\mu}_{X^b}(X_1^b, \ldots, X_{N_{rep}}^b)$ represents an estimate of the mean number of bit errors when N_f frames of N information bits are transmitted, the BER can be estimated by the performance estimation module as follows

$$BER = \frac{\hat{\mu}_{X^b}(X_1^b, \ldots, X_{N_{rep}}^b)}{N_f N} = \frac{1}{N_f N N_{rep}} \sum_{i=1}^{N_{rep}} X_i^b \tag{3.88}$$

and the standard deviation on the BER estimate error can be estimated as follows

$$\hat{\sigma}_{BER} = \frac{\hat{\sigma}_e(X_1^b, \ldots, X_{N_{rep}}^b)}{N_f N} = \frac{\sqrt{\sum_{i=1}^{N_{rep}} \left(X_i^b - \hat{\mu}_{X^b}(X_1^b, \ldots, X_{N_{rep}}^b)\right)^2}}{N_f N N_{rep}}. \tag{3.89}$$

3.4.2 Known channel state information at the fusion center

In section 3.3, we have presented a sub-optimum CSHDD scheme for the case when the channel state is known at the MSC referred as the $\text{CSHDD}_{1,sub}$ scheme. The local decision rules of the $\text{CSHDD}_{1,sub}$ scheme are LLR quantizers. Two different types of LLR quantizers were considered in section 3.3: the MPEP-LLR quantizer and the MMSE-LLR quantizer. In this section, we further propose a third LLR quantizer that uses the thresholds optimized for the uncoded $\text{SHDD}_{1,sub}$ scheme, presented in chapter 2, when the thresholds are constrained to be even symmetric and we refer to this LLR quantizer as the U-LLR quantizer.

We estimated using our software simulator the BER and FER of the $\text{CSHDD}_{1,sub}$ scheme for the three choices of LLR quantizers assuming the mobile unit is communicating simultaneously with 2 base stations. Results for the BER and FER as a function of the first base station average SNR, defined as follows $SNR_1 = \frac{E_1}{N_0}$, are illustrated in Fig.3.3, Fig.3.4 and Fig.3.5 for 1, 2 and 3 receiving antennas respectively, for the case when the average SNR is equal at both base stations. Similarly, Fig.3.6, Fig.3.7 and Fig.3.8 present results for 1, 2 and 3 receiving antennas respectively, for the case when there is a difference of 6dB between the average SNR at the first and the second base station. Each figure is made of two sub-figures where in part a) the FER curves are presented and in part b) the BER curves are presented.

For comparison purposes, all figures also include the BER and FER curves of the OC scheme and the CHM scheme presented in Appendix D. It is important to mention that the BER and FER of the OC scheme are obviously lower bounds to the BER and FER of the CSHDD$_{1,sub}$ scheme. On the other hand, as opposed to the uncoded case, the BER and FER of the CSHDD$_{1,sub}$ scheme are not upper bounded by the BER and FER of the CHM scheme. It is due to the fact that, even if the CHM scheme does not take advantage of the information received at the non-selected base stations, the selected base station has access to direct observations to perform the decoding as opposed to the CSHDD$_{1,sub}$ scheme. Hence, when the difference in average SNR between the two base stations is significant, the CHM scheme can provide better performances than the CSHDD$_{1,sub}$ scheme.

A. Effect of the number of receiving antennas

From Fig.3.3 - Fig.3.8, we see that, as expected, increasing the number of receiving antennas increases the slope of the BER and FER curves at large SNR and consequently the diversity order provided by the CSHDD$_{1,sub}$ scheme, which can be defined as

$$d = \frac{-10\log_{10}\left(\frac{ER^{(2)}}{ER^{(1)}}\right)}{\left(10\log_{10}\left(SNR_k^{(2)}\right) - 10\log_{10}\left(SNR_k^{(1)}\right)\right)} \tag{3.90}$$

where $ER^{(i)}$ represents either the BER or FER at the ith point. In fact, it appears that in all considered cases the CSHDD$_{1,sub}$ scheme provides the same asymptotic diversity order as the OC scheme, which equals approximately $N_{BS}N_R$. Hence, by increasing the number of receiving antennas per base station from 1 to 2, a gain in SNR of approximately 9.3-9.6dB at $FER = 10^{-3}$ and 10.7-11.1dB at $BER = 10^{-4}$ can be observed while, by increasing the number of receiving antennas per base station from 2 to 3, the gain reduces to approximately 3.9-4.2dB at $FER = 10^{-3}$ and 4.1-4.5dB at $BER = 10^{-4}$. However, it can be observed that, when the CSHDD$_{1,sub}$

scheme is using either MPEP-LLR or U-LLR quantizers, the asymptotic diversity is attained more quickly as opposed to the case when MMSE-LLR quantizers are used, especially when $N_R = 2$ and $N_R = 3$.

It is important to mention that, as in the uncoded case, the diversity order of the CHM scheme is only equal to N_R, since this scheme selects a base station based on the average SNR. The diversity order of the $\text{CSHDD}_{1,sub}$ scheme is therefore N_{BS} times larger than the diversity order of the CHM scheme, independently of the number of receiving antennas. For this reason, the $\text{CSHDD}_{1,sub}$ scheme provides in the considered cases important gains with respect to the CHM scheme, especially when the average SNR is equal at both base station. In fact, the gains obtained when the $\text{CSHDD}_{1,sub}$ scheme is using U-LLR quantizers with 8 levels of quantization are within 0.3dB of the potential gain presented in Table 3.2. Hence, as the potential gain, the gain obtained by the $\text{CSHDD}_{1,sub}$ scheme with respect to the CHM scheme decreases with the number of receiving antennas per base station. However, from Fig.3.5 and Fig.3.8, it can be concluded that, even if each base station is equipped with 3 receiving antennas, important gains with respect to the CHM scheme can still be obtained by using handoff macrodiversity schemes that further increase the diversity order such as the $\text{CSHDD}_{1,opt}$ scheme.

B. Effect of the difference in average SNR between two base stations

By comparing Fig.3.3 - Fig.3.5 with Fig.3.6 - Fig.3.8, we see that, similarly to the uncoded case, when the average SNR at the second base station is lower by 6dB from the SNR at the first base station, it has for effect to shift horizontally the BER and FER curves of the $\text{CSHDD}_{1,sub}$ scheme, obtained when the average SNR is equal at both base stations, by approximately 2.6-3.1dB toward the BER and FER curves of the CHM scheme, independently of the number of receiving antennas and the choice of LLR quantizers. Hence, the diversity order of the $\text{CSHDD}_{1,sub}$ scheme is not affected by average SNR difference of 6dB between the two base stations.

C. Comparison between the MPEP-LLR, MMSE-LLR and U-LLR quantizer

First, it is important to mention that, when $L = 2$, the three LLR quantizers are equivalent because the even symmetry constraint implies that the only threshold defining each quantizer is equal to 0. Furthermore, from Fig.3.3 - Fig.3.8, we see that in all considered cases most of the potential gain that can be obtained by the $\mathrm{CSHDD}_{1,sub}$ scheme with respect to the CHM scheme is reached when $L = 2$. However, additional gains can still be obtained by increasing the number of quantization levels. In fact, the potential additional gain that can be obtained by increasing the number of quantization levels of the $\mathrm{CSHDD}_{1,opt}$ scheme to more than 2 equals approximately 1.6-2.0dB, when operating at fixed BER or FER.

On the other hand, it is interesting to see that, when $L > 2$, the three designed quantizers perform quite differently, especially the MMSE-LLR quantizer. In fact, the worst performances are obtained by the MMSE-LLR quantizer, especially at high SNR where almost no gain (<0.3dB) can be observed by increasing the number of quantization levels from 2 to 8. This can be explained by the fact that this criterion does not focus on minimizing the probability of error associated with the more probable error event.

As expected, the best performances are obtained with the MPEP-LLR quantizer which minimizes the criterion that is the more correlated with the probability of frame error and bit error. In fact, if the number of quantization levels is increased from 2 to 4, the FER and BER curves are shifted horizontally toward the BER and FER curves of the OC scheme by approximately 1.1-1.5dB. Hence, with only 4 levels of quantization at each base station, the potential additional gain that can be obtained by further increasing the number of quantization levels equals approximately 0.4-0.8dB, when operating at fixed BER or FER. However, it is important to mention that the optimization of the thresholds with respect to the PEP is a computationally intensive problem due to the numerical integrations and the large number of operations involved. This is why the performances of the MPEP-LLR quantizer were not estimate for $L = 8$.

The quantizer that offered the best trade-off between performance and computation complexity is the U-LLR quantizer. In fact, the performance difference between the MPEP-LLR quantizer and the U-LLR quantizer, when using 4 quantization levels, is inferior to 0.3dB for almost all considered cases. However, in Fig.3.4 and Fig.3.7, it can be observed that, when $N_R = 2$, the performance difference reaches 0.6dB at $BER = 10^{-4}$, when the average SNR is equal at the two base stations, and 0.4dB, when there is a difference of 6dB between the average SNR at the two base stations. Finally, with 8 levels of quantization, the U-LLR quantizer provide performances within 0.3dB of the OC scheme, for all considered cases. It demonstrates that by carefully selecting the thresholds defining the local decision rules, 8 quantization levels are sufficient to make the performances almost identical to the performances of the OC scheme.

3.4.3 Unknown channel state information at the fusion center

In section 3.3, we have presented a sub-optimum CSHDD scheme for the case when the channel state is unknown at the MSC and, in this section, we refer to this scheme as the $\text{CSHDD}_{2,sub}$ scheme. As for the $\text{CSHDD}_{1,sub}$ scheme, the $\text{CSHDD}_{2,sub}$ scheme uses LLR quantizers at the local detectors, where two different types of LLR quantizers are proposed in section 3.3: the MPEP-LLR quantizer and the MMSE-LLR quantizer. In this section, we further propose a third LLR quantizer that uses the thresholds optimized for the uncoded $\text{SHDD}_{2,opt}$ scheme, presented in chapter 2, when the thresholds are constrained to be even symmetric and we refer to this LLR quantizer as the U-LLR quantizer.

We estimated, using our software simulator, the BER and FER of the $\text{CSHDD}_{2,sub}$ scheme for the three choices of LLR quantizers assuming the mobile unit is communicating simultaneously with 2 base stations. As for the previous case, results for the BER and FER as a function of the first base station average SNR, defined as follows $SNR_1 = \frac{E_1}{N_0}$, are illustrated in Fig.3.9, Fig.3.10 and Fig.3.11 for 1, 2 and 3 receiving antennas respectively, for the case when the average SNR is equal at both base stations. Similarly, Fig.3.12, Fig.3.13 and Fig.3.14 present results for 1, 2 and 3 receiving antennas respectively, for the case when there is a difference of 6dB between the average SNR at the first and the second base station. Each figure is made of two sub-figures where in part a) the FER curves are presented and in part b) the BER curves are presented.

For comparison purposes, all figures also include the BER and FER curves of the OC scheme and CHM scheme presented in Appendix D. It is important to mention that the BER and FER of the OC scheme are also lower bounds to the BER and FER of the $\text{CSHDD}_{2,sub}$ scheme while, as opposed to the uncoded case, the BER and FER of the CHM scheme are not upper bounds.

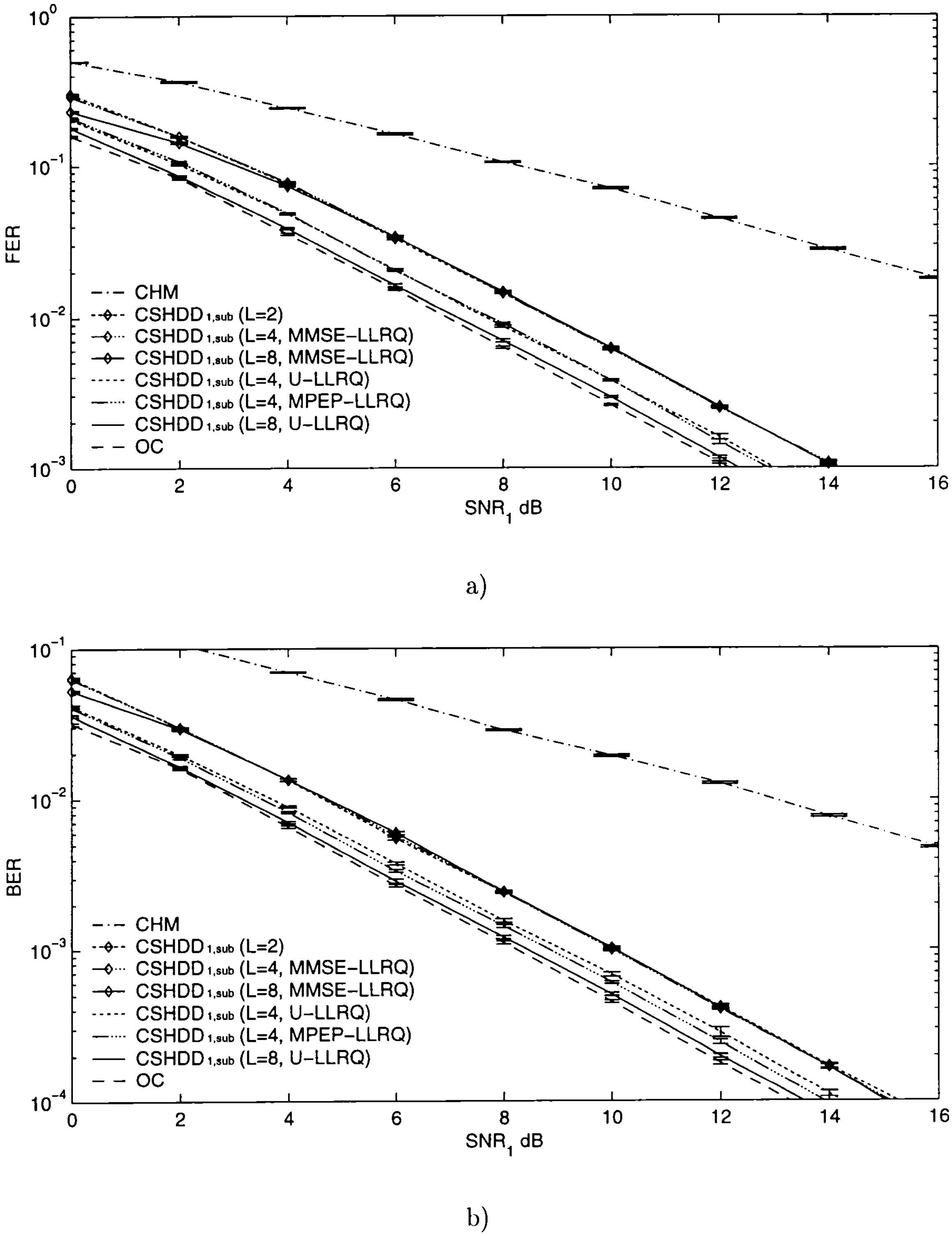


Fig. 3.3 FER and BER vs SNR_1: $\mathrm{CSHDD}_{1,sub}$, $N_{BS}=2$, $N_R=1$, $SNR_1 = SNR_2$

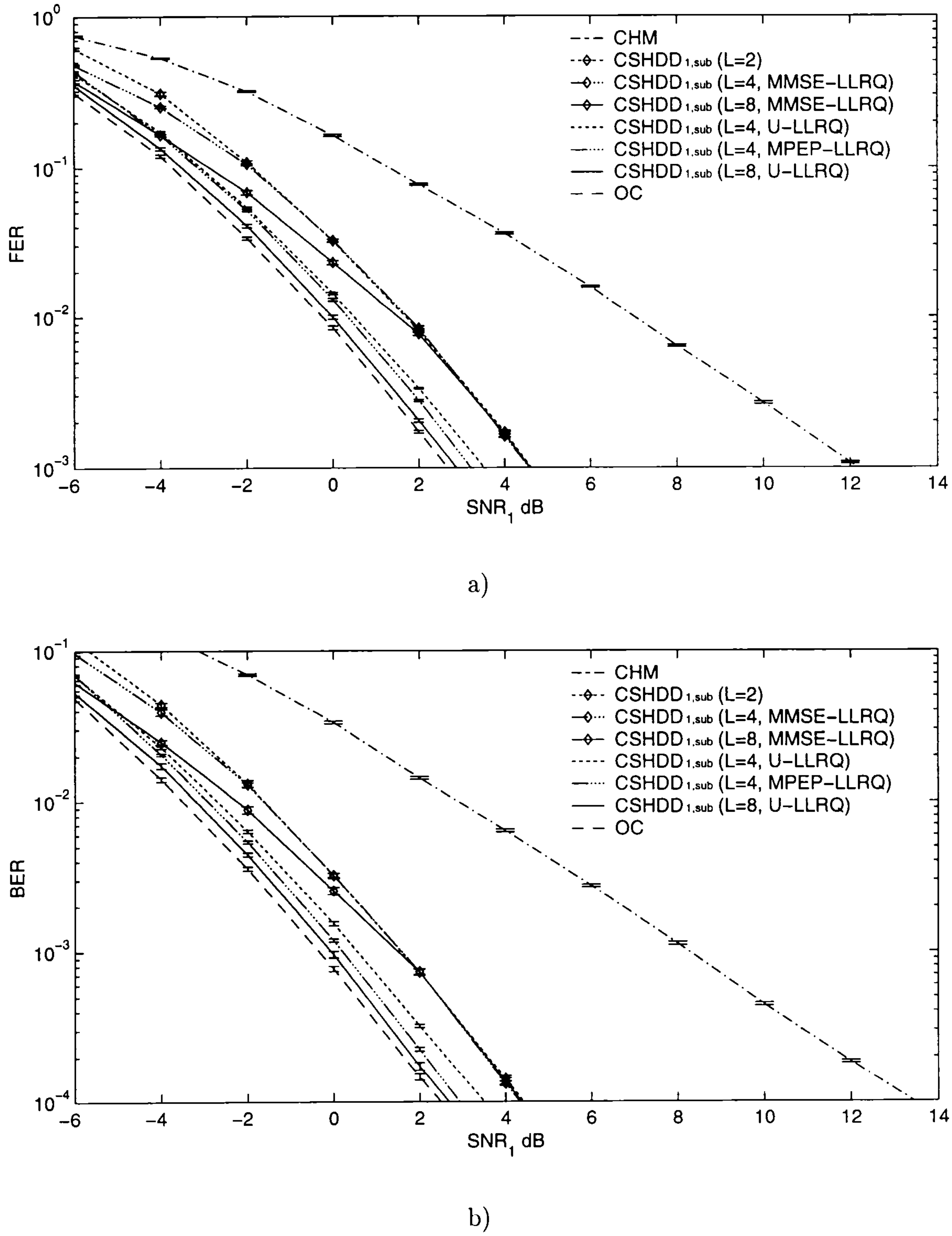
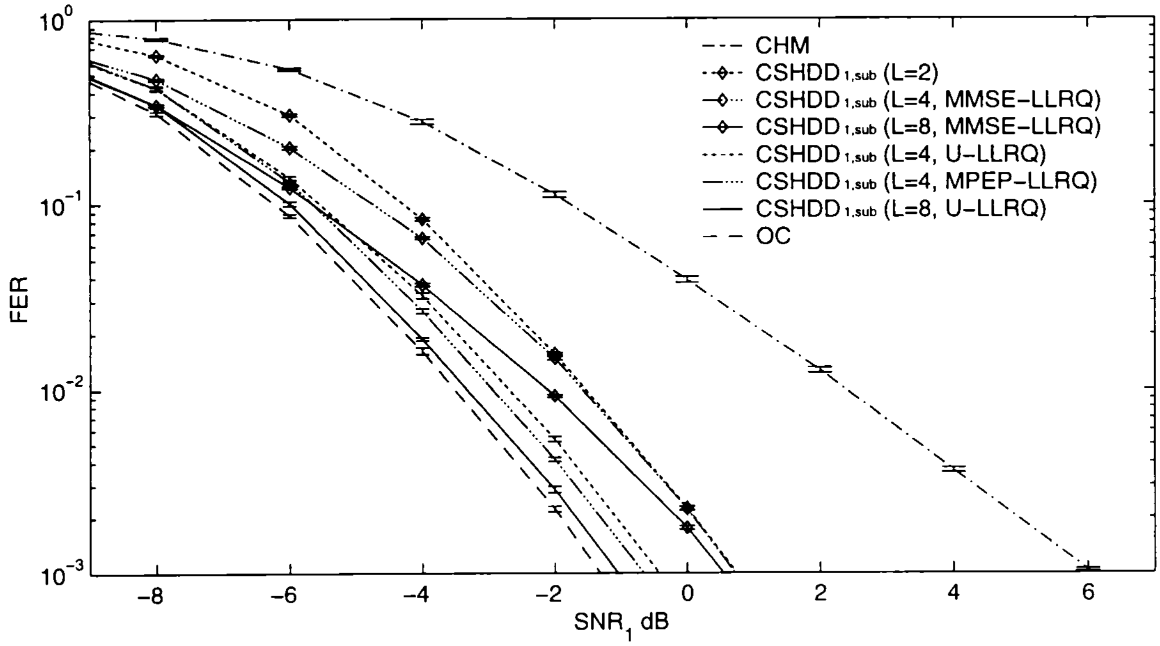
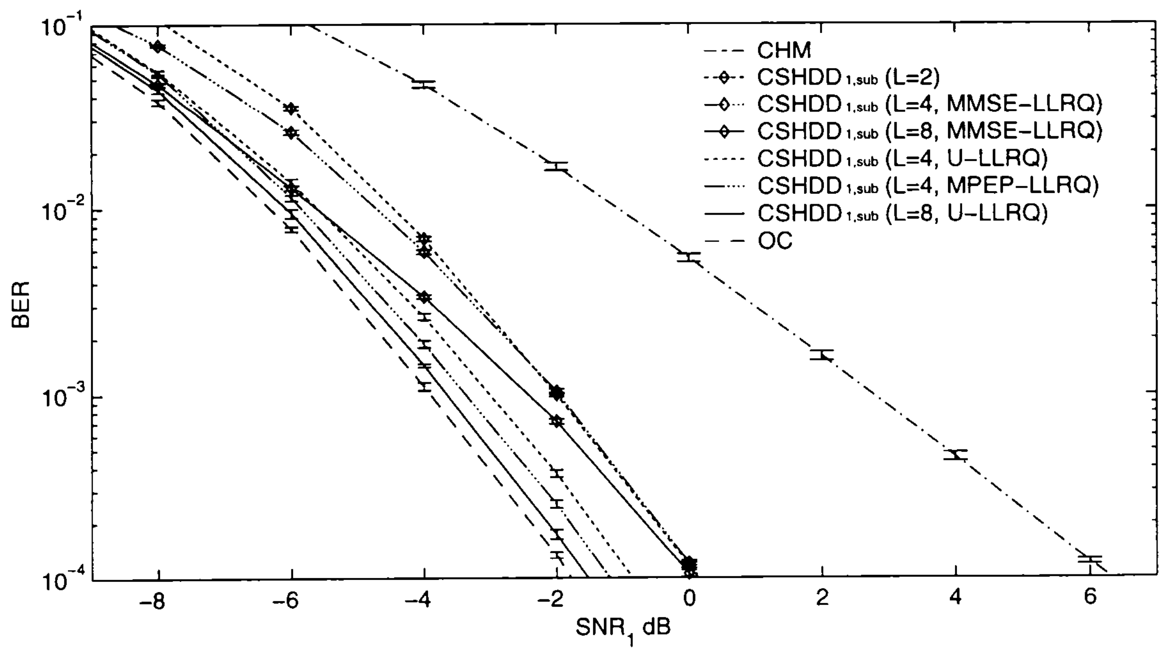


Fig. 3.4 FER and BER vs SNR_1: CSHDD$_{1,sub}$, $N_{BS} = 2$, $N_R = 2$, $SNR_1 = SNR_2$

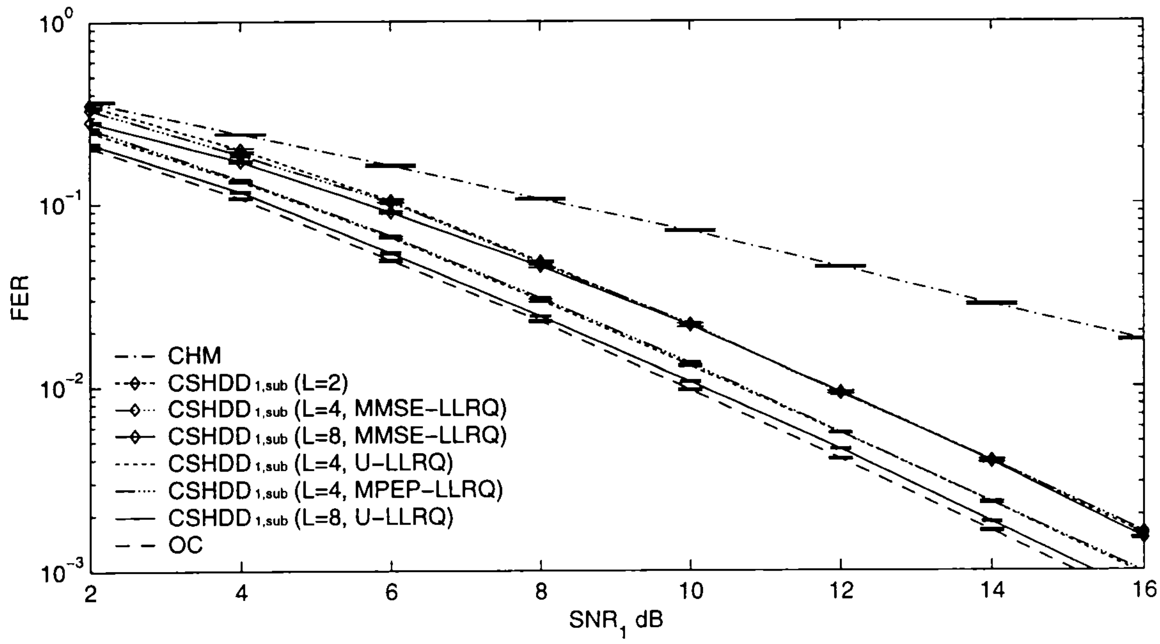


a)

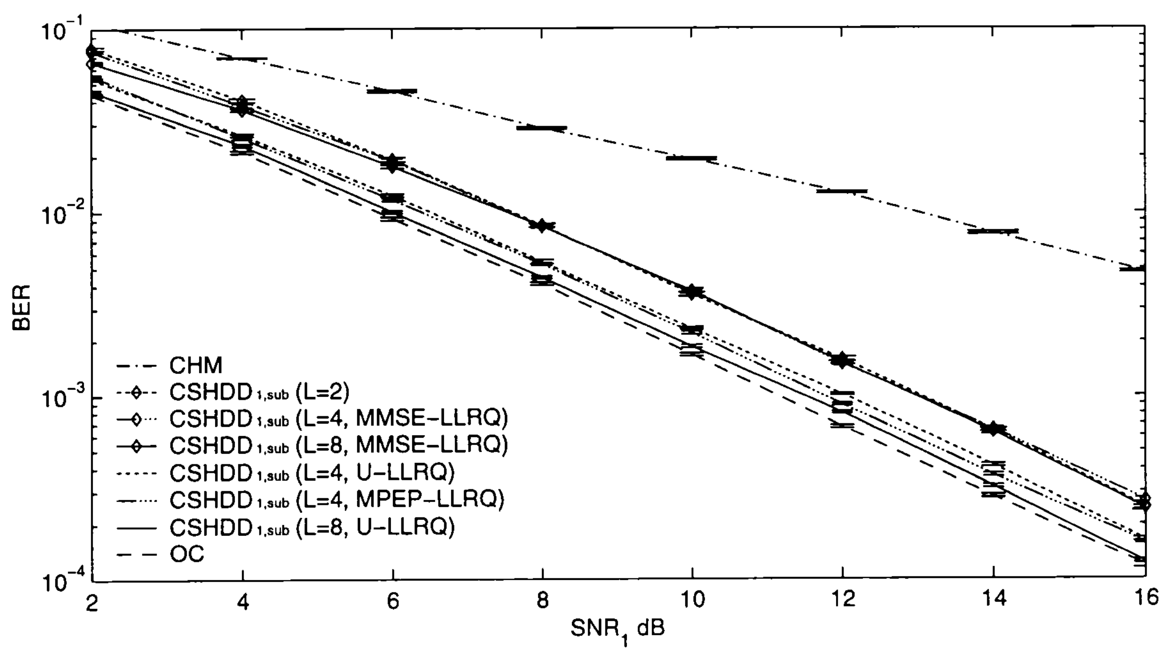


b)

Fig. 3.5 FER and BER vs SNR_1: CSHDD$_{1,sub}$, $N_{BS} = 2$, $N_R = 3$, $SNR_1 = SNR_2$

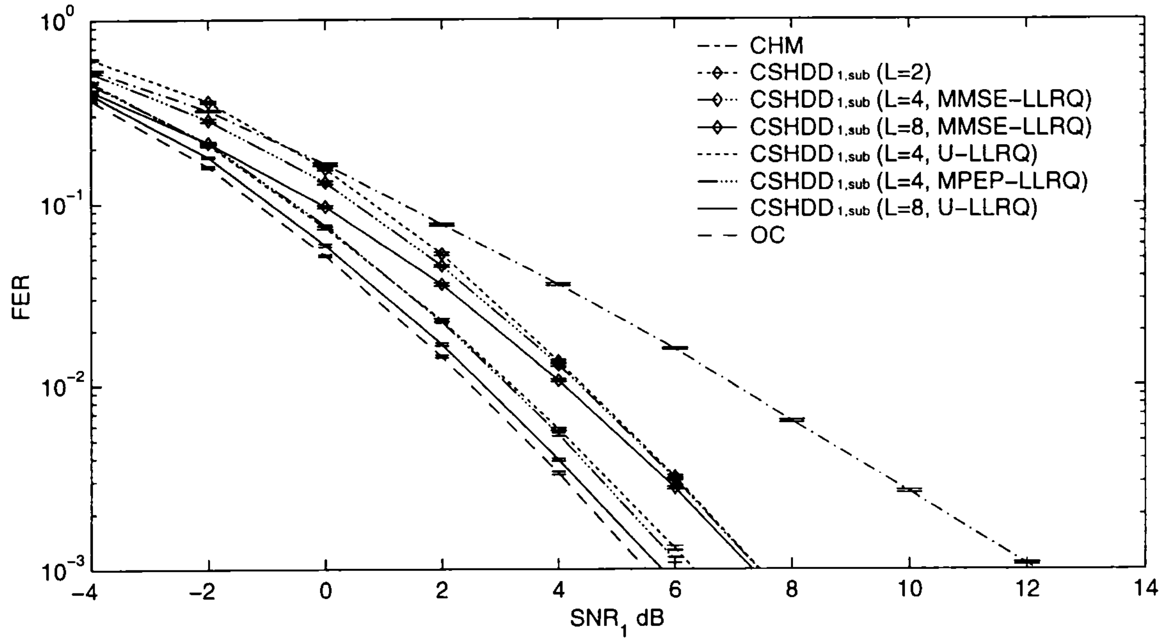


a)

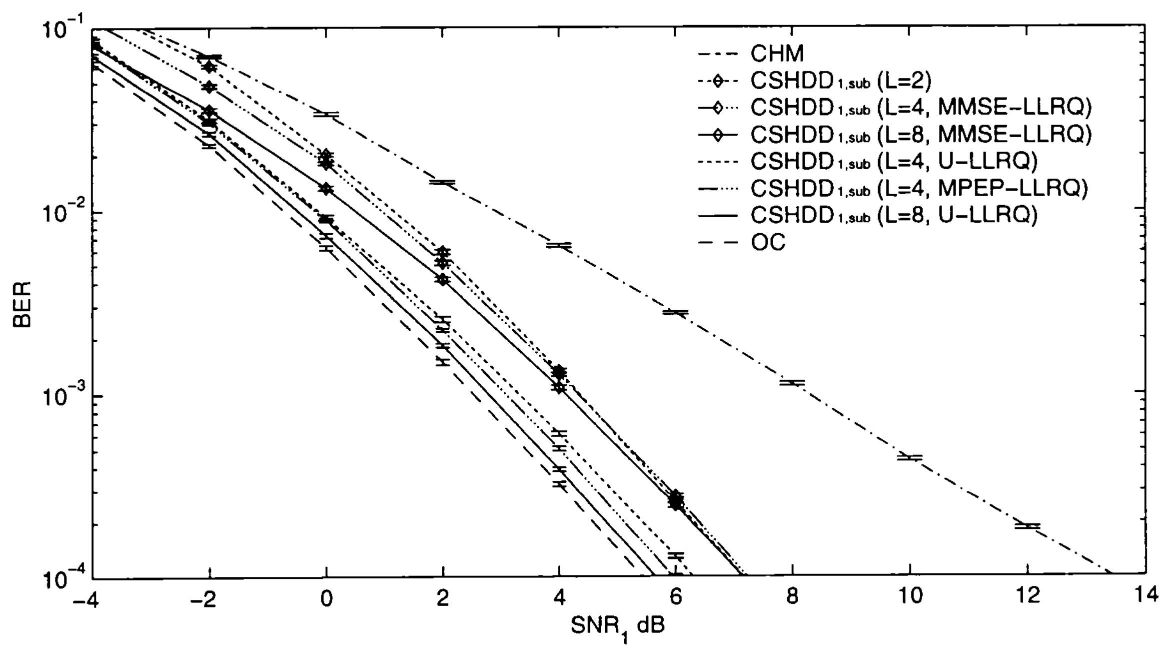


b)

Fig. 3.6 FER and BER vs SNR_1: CSHDD$_{1,sub}$, $N_{BS} = 2$, $N_R = 1$, $SNR_1 = 4SNR_2$

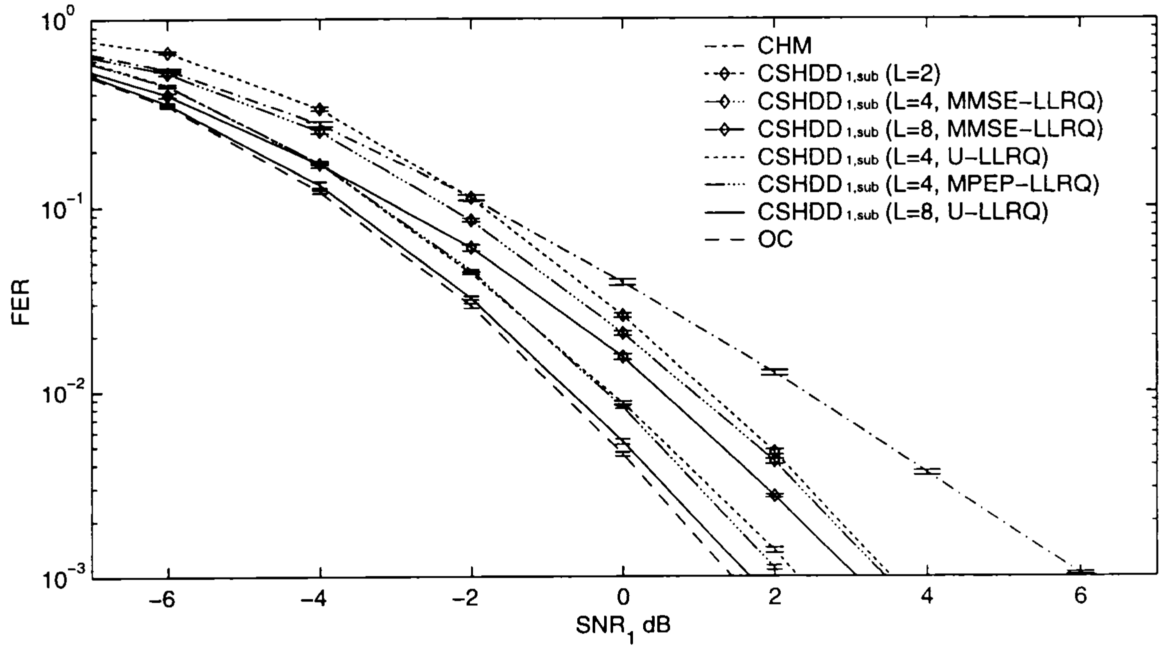


a)

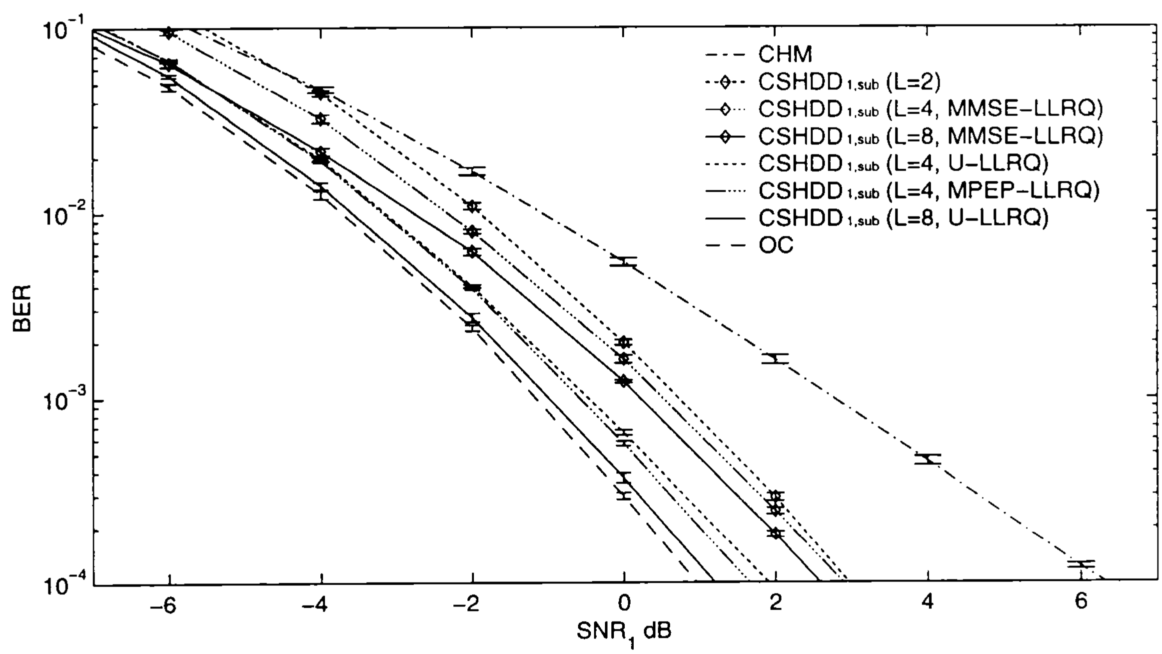


b)

Fig. 3.7 FER and BER vs SNR_1: CSHDD$_{1,sub}$, $N_{BS} = 2$, $N_R = 2$, $SNR_1 = 4SNR_2$



a)



b)

Fig. 3.8 FER and BER vs SNR_1: CSHDD$_{1,sub}$, $N_{BS} = 2$, $N_R = 3$, $SNR_1 = 4SNR_2$

A. Effect of the number of receiving antennas

From Fig.3.9 - Fig.3.14, we see that, as expected, increasing the number of receiving antennas increases the slope of the BER and FER curves at large SNR and consequently the diversity order provided by the CSHDD$_{2,sub}$ scheme, as defined in (3.90). However, as opposed to the CSHDD$_{1,sub}$ scheme, it seems that, when the CSHDD$_{2,sub}$ scheme is using only 2 quantization levels or MMSE-LLR quantizers, the CSHDD$_{2,sub}$ scheme does not reach the same asymptotic diversity order as the OC scheme. It is mostly apparent in Fig.3.9 and Fig.3.12 presenting the results for the case when $N_R = 1$. However, the gain obtained by increasing the number of receiving antennas for these schemes seems to be superior as for the other cases that reach the full diversity. In fact, by increasing the number of receiving antennas per base station from 1 to 2, it can be observed that the performances of the CSHDD$_{2,sub}$ scheme using 2 quantization levels are improved by approximately 13.0-14.8dB at $FER = 10^{-3}$ and by 13.9-14.0dB at $BER = 10^{-4}$ while, by increasing the number of receiving antennas per base station from 2 to 3, the performance improvement reduces to approximately 4.8dB at $FER = 10^{-3}$ and 4.6-5.2dB at $BER = 10^{-4}$. On the other hand, for the other cases where the CSHDD$_{2,sub}$ scheme reaches the same asymptotic diversity order as the OC scheme, the gains obtained by increasing the number of receiving antennas are similar to the gains obtained with the CSHDD$_{1,sub}$ scheme.

B. Effect of the difference in average SNR between two base stations

By comparing Fig.3.9 - Fig.3.11 with Fig.3.12 to Fig.3.14, it can be observed that in most of the considered cases, when the average SNR at the second base station is lower by 6dB from the SNR at the first base station, the BER and FER curves of the CSHDD$_{2,sub}$ scheme, obtained when the average SNR is equal at both base stations, are shifted horizontally by approximately 2.6-3.1dB toward the BER and FER curves of the CHM scheme. However, when the CSHDD$_{2,sub}$ scheme is using only 2 quantization levels or MMSE quantizers, it seems that the performance losses are not only caused by an horizontal shift of the performance curves but also by a change in the slope of these curves at high SNR.

Furthermore, when examining the curves associated with $L = 2$ in Fig.3.12 - Fig.3.14, it can be observed that the CSHDD$_{2,sub}$ scheme provides, at low SNR, worst performances than the CHM scheme, proving that the performances of the

$\text{CSHDD}_{2,sub}$ scheme are not lower bounded by the performances of the CHM scheme. Consequently, in this case when the asymmetry between the quality of the local decisions becomes too significant, it may be advantageous to implement the CHM scheme.

C. Comparison between the MPEP-LLR, MMSE-LLR and U-LLR quantizer

As for the $\text{CSHDD}_{1,sub}$ scheme, when $L = 2$ the three LLR quantizers are equivalent because the even symmetry constraint implies that the only threshold defining each quantizer is equal to 0. From Fig.3.9 - Fig.3.14, we see that, as opposed to the $\text{CSHDD}_{1,sub}$ scheme, when $L = 2$ the performances obtained are far from reaching the potential gain that CSHDD schemes can obtained with respect to the CHM scheme. In this case, it is thus much more advantageous to use soft decisions at the local detectors. In fact, the potential additional gain at $FER = 10^{-3}$ that can be obtained, by increasing the number of quantization levels of the $\text{CSHDD}_{2,opt}$ scheme to more than 2, goes up to 9.1dB, when $N_R = 1$ and the average SNR is equal at both base stations.

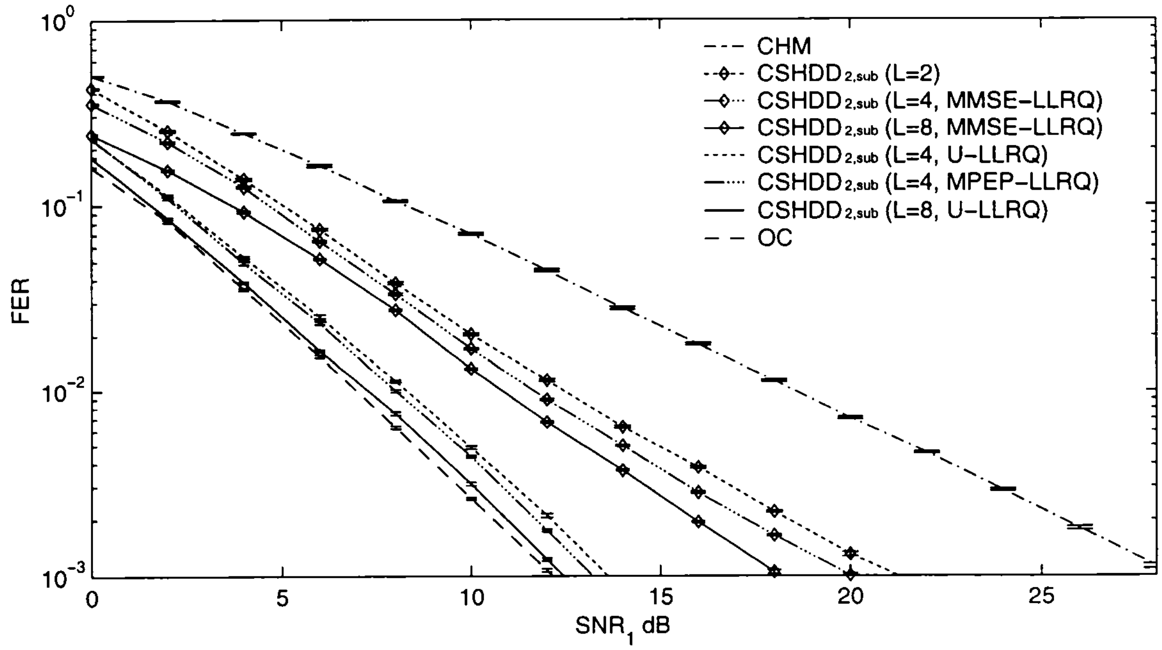
Hopefully, by increasing the number of quantization levels, the performance curves of the $\text{CSHDD}_{2,sub}$ scheme move gradually toward the performance curves of the OC scheme. The worst performances are still obtained with the MMSE-LLR quantizer. On the other hand, as opposed to the $\text{CSHDD}_{1,sub}$ scheme, the performances obtained with the MMSE-LLR quantizer are in most cases not converging at high SNR to the performances of the 2 quantization level $\text{CSHDD}_{2,sub}$ scheme. However, even with 8 quantization levels, the performances of the $\text{CSHDD}_{2,sub}$ scheme using MMSE-LLR quantizers are far from the performances of the OC scheme. From Fig.3.9, it can be observed that the difference is still as high as 6.1dB at $FER = 10^{-3}$, when $N_R = 1$ and the average SNR is equal at both base stations.

Again, as expected the best performances are obtained with the MPEP-LLR quantizer which minimize the criterion the more correlated with the probability of frame error. In fact, from Fig.3.9, it can be observed that the gain in SNR obtained at $FER = 10^{-3}$ by increasing the number of quantization levels from 2 to 4 goes up to 7.8dB, when $N_R = 1$ and the average SNR is equal at both base stations. Hence, with only 4 levels of quantization at each base station, the potential additional SNR gain that can be obtained at $FER = 10^{-3}$, by further increasing the number of quantization levels, varies from approximately $1.3dB$, when $N_R = 1$, to approximately $0.8dB$, when $N_R = 2$ or $N_R = 3$. However, similarly to the CSHDD$_{1,sub}$ scheme, the optimization of the thresholds with respect to the PEP is still a computationally intensive problem due to the numerical integrations and the large number of operations involved.

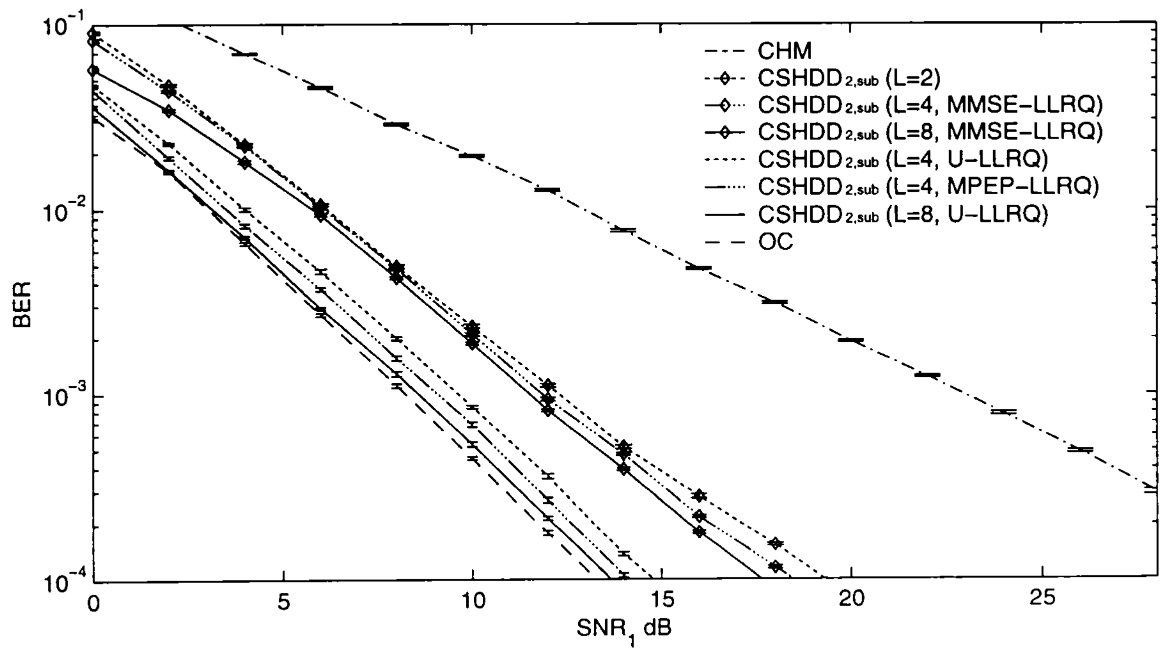
The quantizer that offered the best trade-off between performance and computation complexity is the U-LLR quantizer. In fact, from Fig.3.12 - Fig.3.14, it appears that the performance difference between the MPEP-LLR quantizer and the U-LLR quantizer, using 4 quantization levels, is negligible for the case when the average SNR is unequal at the two base stations. On the other hand, from Fig.3.9 - Fig.3.11, it appears that, when the average SNR is equal at the two base stations, the difference between the performances obtained with the two quantizers is more considerable and varies from $0.3dB$ to $0.6dB$ at $FER = 10^{-3}$, depending on the number of receiving antennas. Furthermore, with only 8 levels of quantization, the CSHDD$_{2,sub}$ scheme using U-LLR quantizers provides performances within $0.3dB$ of the OC scheme, for all considered cases. It demonstrates that by carefully selecting the thresholds defining the local decision rules, 8 quantization levels are sufficient to make the performances almost identical to the OC scheme without even knowing the channel state at the MSC.

3.4.4 Comparison of the CSHDD$_{1,sub}$ and CSHDD$_{2,sub}$ schemes

In this section, we use the results presented in the two previous sections and compare the performances of the CSHDD$_{1,sub}$ and CSHDD$_{2,sub}$ schemes. Similarly to the uncoded case, the more significant performance difference between the CSHDD$_{1,sub}$ and CSHDD$_{2,sub}$ schemes can be observed when $L = 2$. In fact, by comparing the results presented in Fig.3.3 - Fig.3.8 with the results presented in Fig.3.9 - Fig.3.14, we see that, when both schemes are using 2 quantization levels, the SNR difference when operating at $FER = 10^{-3}$ reaches up to 7.2dB, when $N_R = 1$. However, the difference diminishes as the number of receiving antennas increases and equals to 1.8dB, when $N_R = 2$, and 1.0dB, when $N_R = 3$. Furthermore, as the uncoded case, the performance difference diminishes also as the number of quantization levels increases. For instance, for CSHDD schemes using MPEP-LLR quantizers and 4 quantization levels, the performance difference reduces to less than 0.5dB at $FER = 10^{-3}$, independently of the

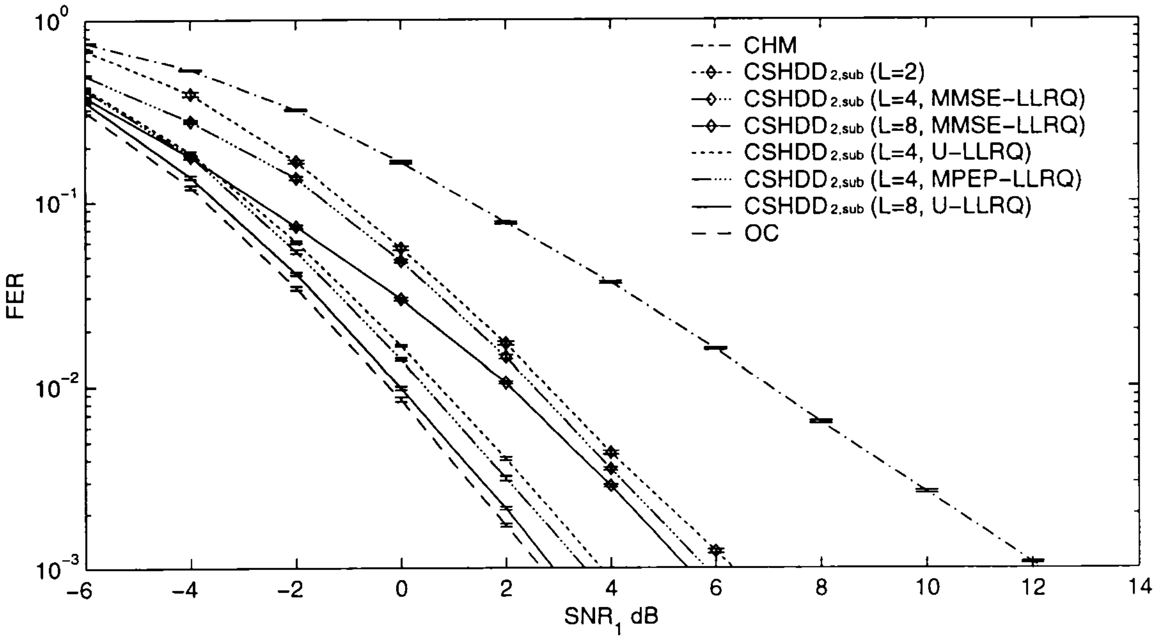


a)

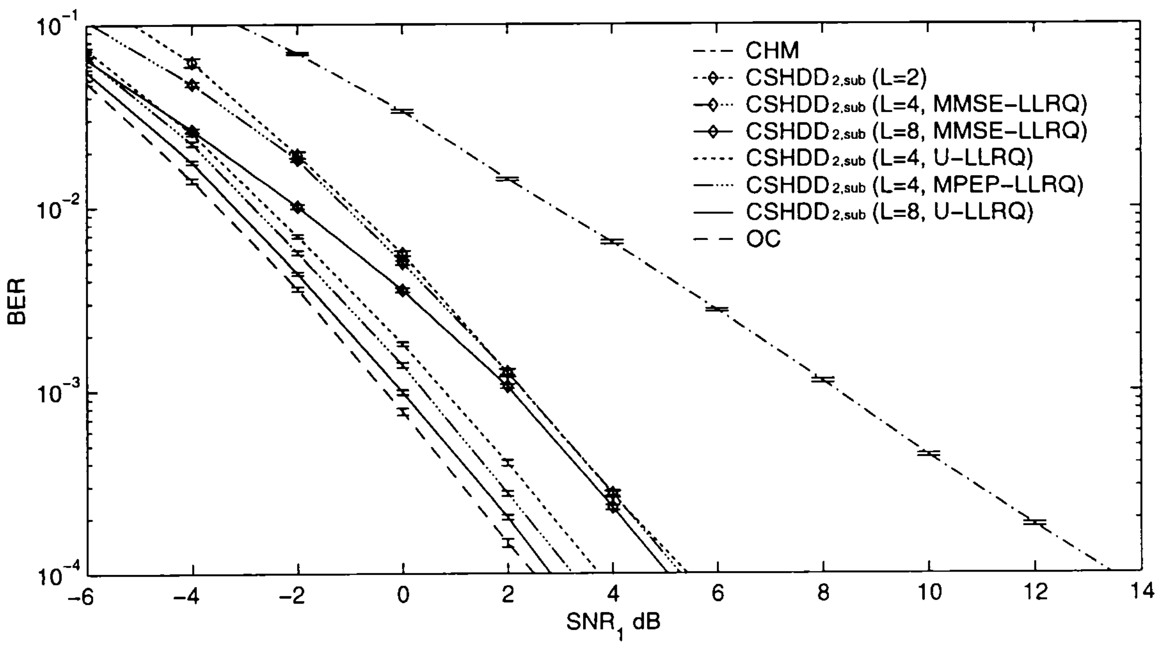


b)

Fig. 3.9 FER and BER vs SNR_1: CSHDD$_{2,sub}$, $N_{BS} = 2$, $N_R = 1$, $SNR_1 = SNR_2$

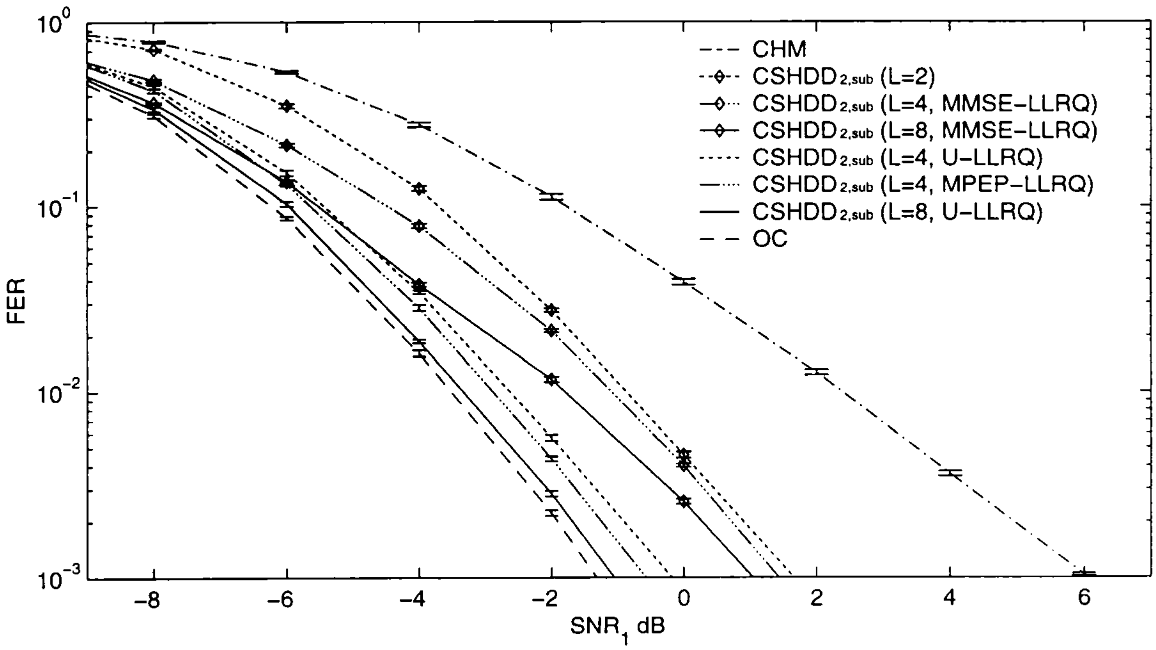


a)

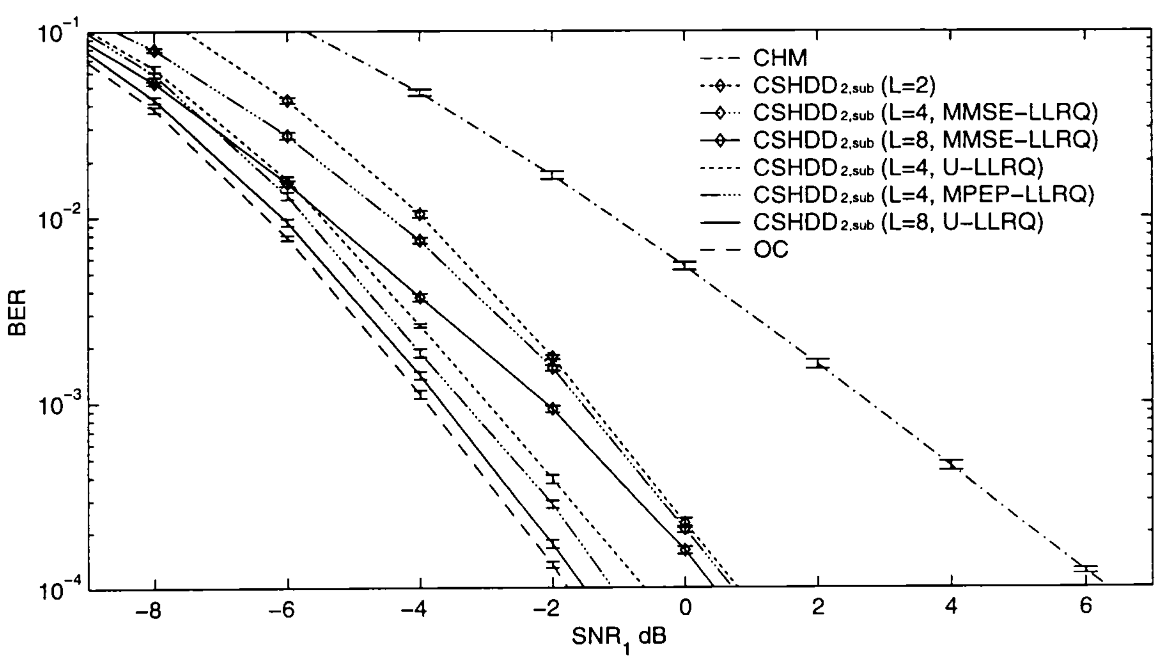


b)

Fig. 3.10 FER and BER vs SNR_1: CSHDD$_{2,sub}$, $N_{BS} = 2$, $N_R = 2$, $SNR_1 = SNR_2$

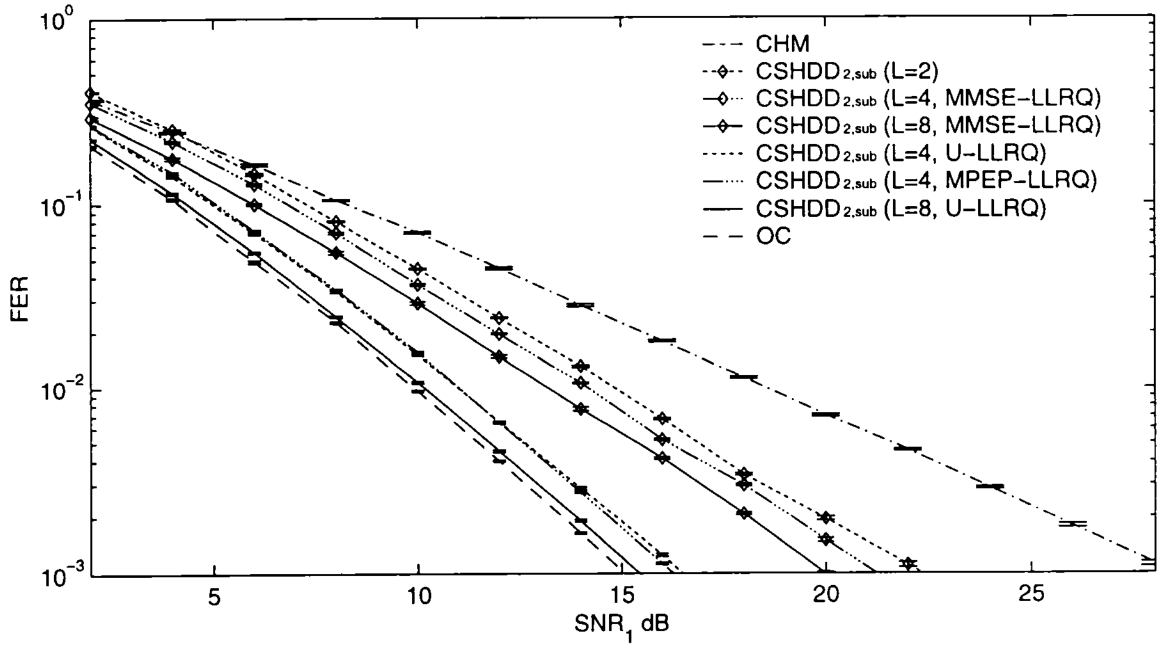


a)

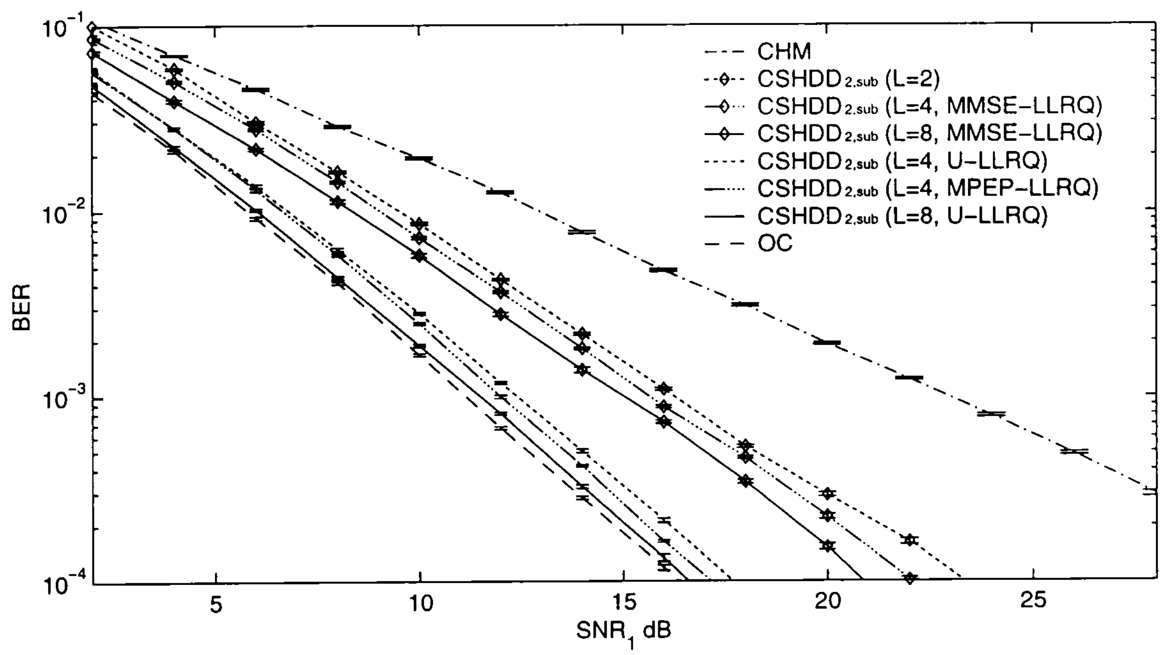


b)

Fig. 3.11 FER and BER vs SNR_1: CSHDD$_{2,sub}$, $N_{BS} = 2$, $N_R = 3$, $SNR_1 = SNR_2$

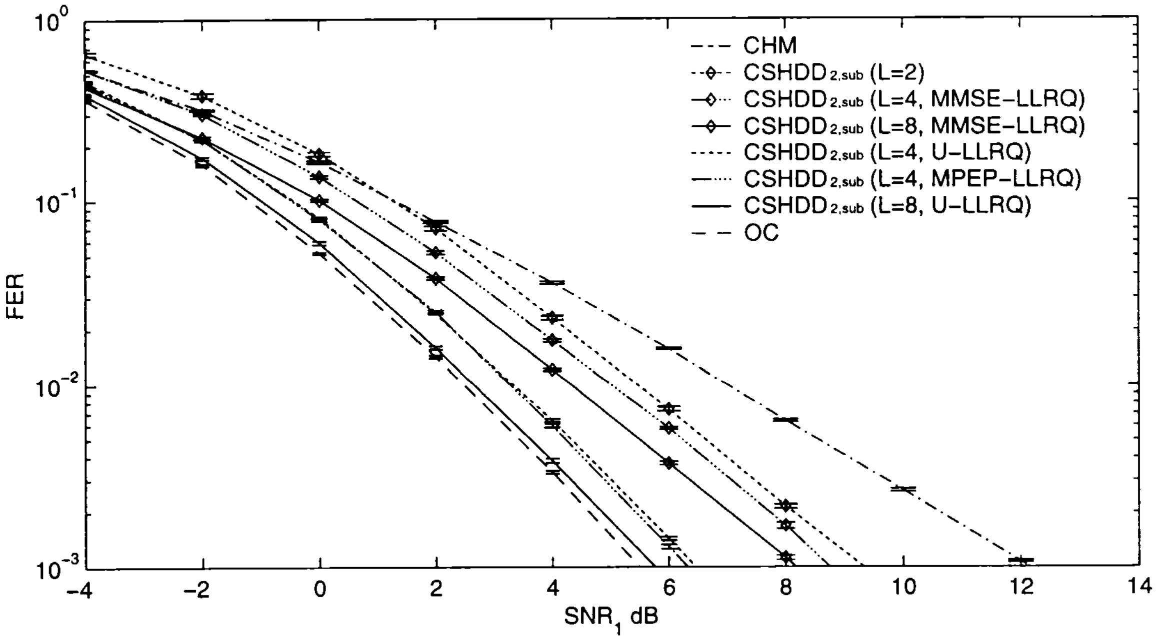


a)

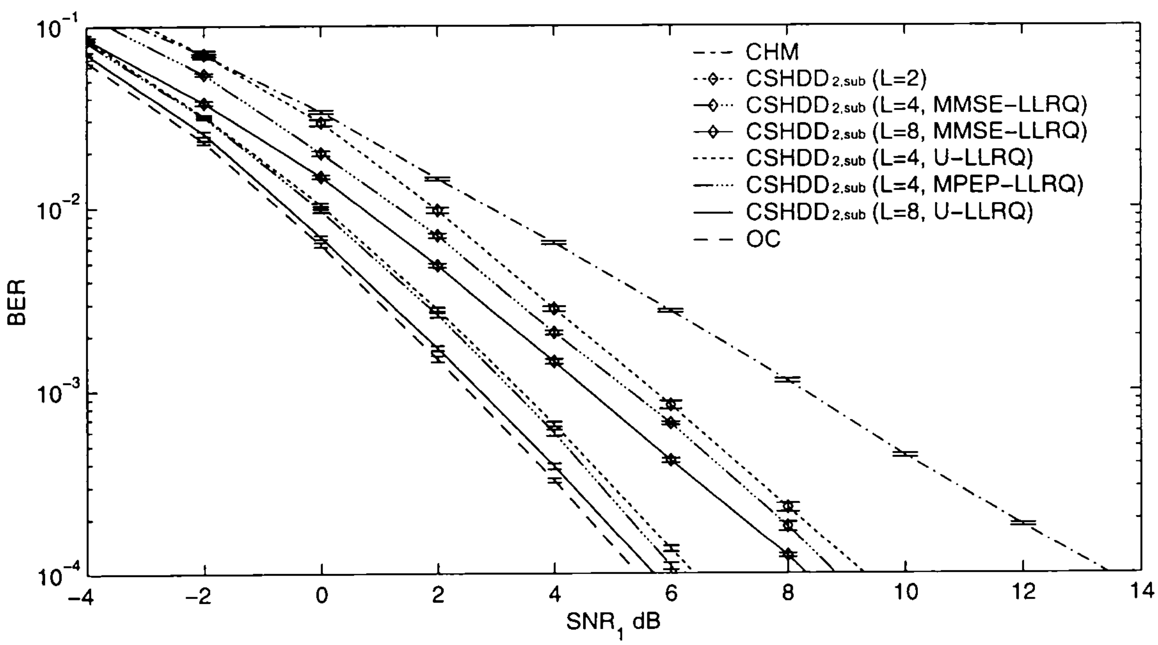


b)

Fig. 3.12 FER and BER vs SNR_1: CSHDD$_{2,sub}$, $N_{BS} = 2$, $N_R = 1$, $SNR_1 = 4SNR_2$

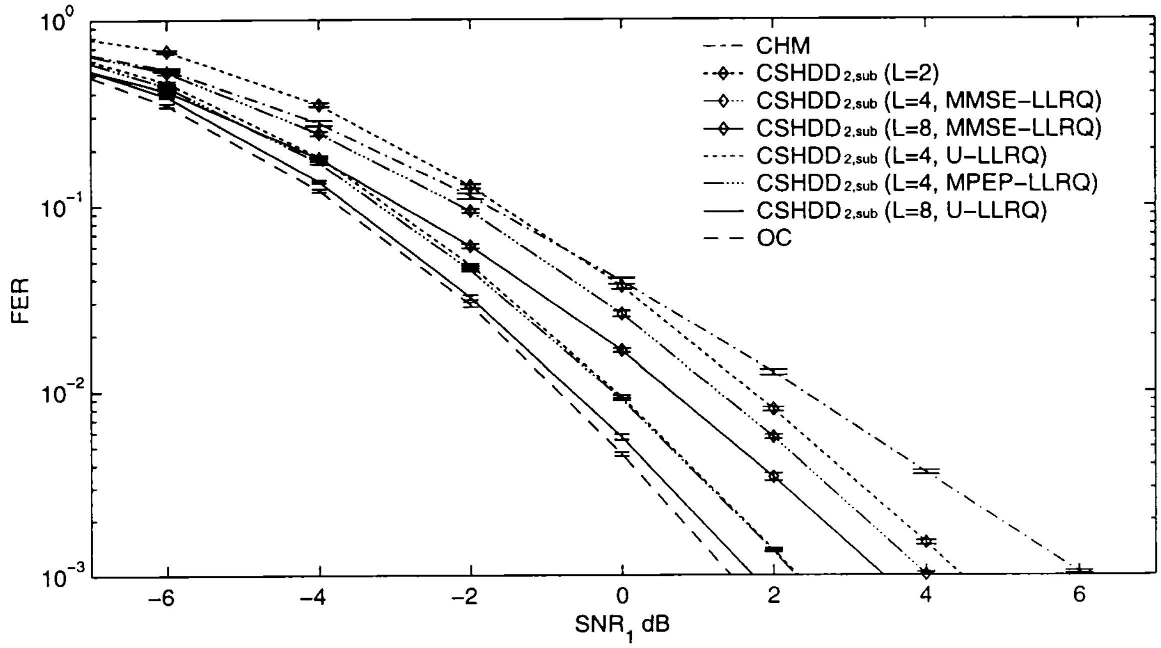


a)

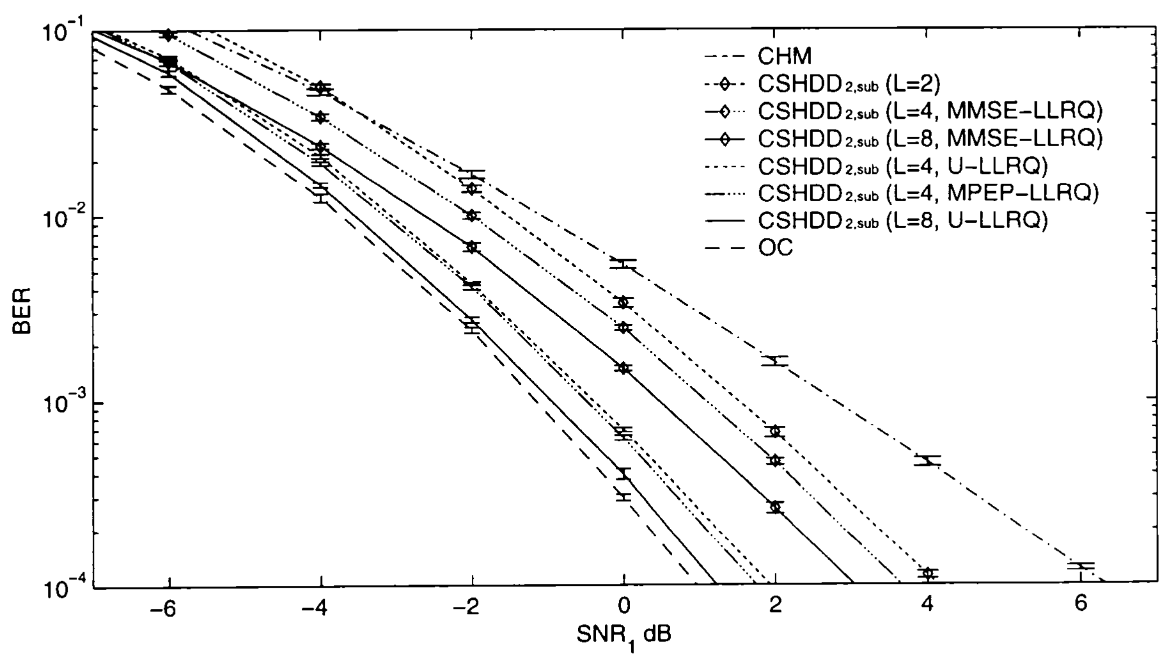


b)

Fig. 3.13 FER and BER vs SNR_1: CSHDD$_{2,sub}$, $N_{BS} = 2$, $N_R = 2$, $SNR_1 = 4SNR_2$



a)



b)

Fig. 3.14 FER and BER vs SNR_1: CSHDD$_{2,sub}$, $N_{BS} = 2$, $N_R = 3$, $SNR_1 = 4SNR_2$

number of receiving antennas. Similarly, for CSHDD schemes using U-LLR quantizers, the performance difference reduces to less than 0.7dB, when $L = 4$, and 0.1dB, when $L = 8$. Hence, when $L = 2$, the significant gains obtained by the $\text{CSHDD}_{1,sub}$ scheme with respect to the $\text{CSHDD}_{2,sub}$ scheme can justify the additional complexity and additional fixed network bandwidth required by the $\text{CSHDD}_{1,sub}$ scheme, especially when each base station is using a single receiving antenna. On the other hand, when $L > 2$, the $\text{CSHDD}_{2,sub}$ scheme becomes, as the number of quantization levels increases, more attractive since it provides almost the same performances as the $\text{CSHDD}_{1,sub}$ scheme without the added complexity and fixed network bandwidth.

3.4.5 Comparison with the uncoded case

In this section, we compare the BER performances of the $\text{SHDD}_{1,sub}$ and $\text{SHDD}_{2,opt}$ schemes presented in chapter 2 with the BER performances of the $\text{CSHDD}_{1,sub}$ and $\text{CSHDD}_{2,sub}$ schemes. It is important to mention that in order to fairly compare the performances of these schemes, it is necessary to express the results as a function of $\frac{E_b}{N_0}$, where E_b represents the received energy per information bit at the first base station and, for the coded case, $E_b = \frac{N_c}{N}E_1$. In addition, the $\text{SHDD}_{1,sub}$ and $\text{SHDD}_{2,opt}$ schemes considered are using even symmetric thresholds as the $\text{CSHDD}_{1,sub}$ and $\text{CSHDD}_{2,sub}$ schemes.

Before discussing the results obtained for the CSHDD schemes, it is important to first look at the coding gain obtained with the OC scheme. Hence, by comparing the BER of the OC scheme for coded and uncoded communication systems, it can be remarked that no coding gain is obtained at $BER = 10^{-4}$ when $N_R = 1$. This can be explained in part by the fact that the channel coding does not increase the diversity order of the system since the fading is quasi-static and is therefore constant over the whole frame. Furthermore, since all error events have the same diversity order and the probability associated with these error events is not decreasing, as for AWGN channel, exponentially with the hamming distance between the transmitted codeword and the erroneously decoded codeword, there is no dominant error event and all error events contribute to the probability of bit error [27][31]. However, the coding gain increases with the number of receiving antennas since by increasing the diversity order the channel tends to behave more like a AWGN channel. In fact, the coding gain reaches 1.6dB, when $N_R = 2$, and 2.4dB, when $N_R = 3$.

For the CSHDD schemes, the more important gain is observed when comparing the performances of the SHDD$_{2,opt}$ and CSHDD$_{2,sub}$ schemes when $L = 2$. In fact, the gain at $BER = 10^{-4}$ reaches up to 11.7dB, when $N_R = 1$, up to 7.9dB, when $N_R = 2$, and up to 6.5dB, when $N_R = 3$. It is due to the fact that a 2 base station SHDD$_{2,opt}$ scheme, when $L = 2$ and the thresholds are even symmetric, is equivalent to the CHM scheme and only provides half the diversity of the CSHDD$_{2,sub}$ scheme. However, this is an exception and, in all other cases, the diversity order provided by both the coded and uncoded schemes approaches the asymptotic diversity of the OC scheme. Furthermore, for all the other cases, the gain provided by the CSHDD schemes with respect to the SHDD schemes is inferior to the coding gain provided by the OC scheme. It is not surprising since the CSHDD$_{1,sub}$ and CSHDD$_{2,sub}$ schemes are sub-optimum schemes designed to reduce the probability of frame error as opposed to the SHDD$_{1,sub}$ and SHDD$_{2,opt}$ schemes which are optimum scheme designed to minimize the probability of bit error.

This is obviously the CSHDD schemes using MMSE-LLR quantizers that provide the less gain. In fact, for most of the considered cases, CSHDD schemes using MMSE-LLR quantizers provides worst performance than the SHDD schemes, where the losses can reach up to 4.5dB when $N_R = 1$. This actually proves the inefficiency of the MMSE-LLR quantizers. On the other hand, it can be observed that CSHDD schemes using MPEP-LLR quantizers with 4 quantization levels provide gains within 0.3dB of the coding gain obtained with the OC scheme. Similarly, the CSHDD schemes using U-LLR quantizers provide gains within 0.3dB of the coding gain obtained with the OC scheme, when $L = 8$, and within 0.8dB when $L = 4$.. Consequently, the performances of the SHDD schemes can be improved significantly by adding channel coding to the system as long as the diversity order is sufficient and the local detector quantizers are well designed.

Chapter 4

Conclusions and future work

This work studied the application of soft decision distributed detection to the uplink of a mobile in soft handoff, in an effort to design improved alternatives to conventional handoff macrodiversity techniques based on selection diversity. We first considered uncoded communication system over a quasi-static spatially uncorrelated Rayleigh fading channel. Two different cases were considered. In the first case, the average SNR as well as the channel state at each base station is assumed known at the MSC. In the second case, only the average SNR at each base station is known at the MSC. For both cases, it is shown that the optimum local decision rules can be expressed as likelihood ratio quantizers and the optimum fusion rule is a maximum likelihood decision rule based on the local decisions. When the channel state is known at the MSC, new threshold values need to be transmitted to the base stations every time the channel state varies at any base station. However, it was observed that, when the thresholds are only updated when the average SNR varies at any base station, the performance degradation is neglectable as long as the number of quantization levels is larger than 2. The performances of the distributed detection schemes were evaluated numerically assuming that each base station is equipped with multiple receiving antennas. It appears that with only 3 bits of resolution at each base station, performances less than $0.1dB$ from the optimum centralized scheme are obtained for all considered cases.

The proposed distributed detection schemes provide large performance gains with respect to the conventional handoff macrodiversity scheme, based on selection diversity, at the expense of a small increase in the required bandwidth from the fixed network and an increase in the offline computational requirements, where the offline computation is caused by the optimization of the local detector thresholds. The optimization of the thresholds is complicated by the fact that the probability of bit error is a non-convex nonlinear function of the thresholds. An improved Simulated Annealing (SA) algorithm, called Adaptive Simulated Annealing (ASA), was thus used to perform the numerical optimization. The influence of assumptions simplifying the optimization process were investigated. It was observed that as the number of quantization levels and the number of base stations increase, these assumptions have less and less impact on the performances.

Extension of the principles of distributed detection were derived for communication system using channel coding over a quasi-static spatially uncorrelated Rayleigh fading channel. The two same cases as for uncoded communication systems were considered, and optimum decision rules were derived. It was shown that the complexity, that implies either the implementation or the numerical optimization of these decision rules, grows exponentially with the frame size. Sub-optimum alternatives using likelihood ratio quantizers were proposed. The selection of local detector thresholds to minimize directly the average probability of frame error is a computationally intensive problem. We therefore investigated different optimality criteria of a lower computational complexity that reduce the probability of frame error. The optimality criterion that provided the best performances is the pairwise error probability. However, it is still a computationally intensive problem, because it requires the evaluation of numerical integrals. The best trade-off between computation complexity and performance was obtained by simply using the thresholds of the uncoded system. It appears that with only 3 bits of resolution at each base station, performances less than $0.3dB$ from the optimum centralized scheme are obtained at $FER = 10^{-3}$. It should be mentioned that the mean square error is the criterion that provided the worst performances, especially at high SNR.

We also investigated the performance gains of the coded schemes over the corresponding uncoded schemes, with respect to the BER. It was observed that the added complexity and bandwidth required by the coded systems only improve the performances as long as the diversity order is sufficient and the local detector quantizers are well designed.

In closing, the designed handoff macrodiversity schemes based on distributed detection have proved to be an effective alternatives to the more conventional handoff macrodiversity scheme. However, much more still need to be investigated, commencing by the extension of these results to higher order modulation schemes. In fact, preliminary results show that the results presented in this thesis can easily be extended to communication systems using QPSK modulation. Furthermore, additional future research can consider other type of channel fading (spatial correlation, frequency selective fading,...), the impact of error on the estimates of the channel coefficients used at the base stations and the MSC and much more.

Appendix A

Local detector LLR

In this appendix, the local detector likelihood ratio is first formulated in its logarithmic form. The PDF and CDF of the LLR are then derived for the case when the channel state vector $\mathbf{H}$ is known and for the case when the channel state vector $\mathbf{H}$ is unknown. These results are necessary for the evaluation of the cost functions and for the implementation of the decision rules of both the SHDD and CSHDD schemes. It is important to mention that, in this appendix, we consider the LLR used by the SHDD scheme. However, these results can be extended to the LLR used by the CSHDD scheme by replacing in all expressions the bit B by the coded bit C_t as well as the received signal vector $\mathbf{R}_k$ and its realization $\mathbf{r}_k$ by the received signal vector $\mathbf{R}_{k,t}$ and its realization $\mathbf{r}_{k,t}$.

A.1 Local detector likelihood ratio

Considering the system model presented in section 2.1, the local detector likelihood ratio is defined as follows

$$\Lambda_{1,0}^{(k)}(\mathbf{r}_k, \mathbf{h}_k) = \frac{f_{\mathbf{R}_k}(\mathbf{r}_k \mid \mathbf{h}_k, B = 1)}{f_{\mathbf{R}_k}(\mathbf{r}_k \mid \mathbf{h}_k, B = 0)} \tag{A.1}$$

where the joint conditional PDF of the N_R received signals at the kth base station equals

$$f_{\mathbf{R}_k}(\mathbf{r}_k \mid \mathbf{h}_k, B) = \frac{1}{(\pi N_0)^{N_R}} \exp\left\{ \frac{-\sum_{n=1}^{N_R} \left| r_{k,n} - h_{k,n}\sqrt{E_k}(-1)^{B+1} \right|^2}{N_0} \right\}. \tag{A.2}$$

The likelihood ratio at the kth base station can therefore be simplified as follows

$$\begin{aligned}
\Lambda_{1,0}^{(k)}(\mathbf{r}_k, \mathbf{h}_k) &= \exp\left\{\sum_{n=1}^{N_R} \frac{-\left|r_{k,n} - h_{k,n}\sqrt{E_k}\right|^2}{N_0}\right\} / \exp\left\{\sum_{n=1}^{N_R} \frac{-\left|r_{k,n} + h_{k,n}\sqrt{E_k}\right|^2}{N_0}\right\} \\
&= \exp\left\{\sum_{n=1}^{N_R} \frac{-\left|r_{k,n} - h_{k,n}\sqrt{E_k}\right|^2 + \left|r_{k,n} + h_{k,n}\sqrt{E_k}\right|^2}{N_0}\right\} \\
&= \exp\left\{\sum_{n=1}^{N_R} \frac{-\left|r_{k,n}\right|^2 - \left|h_{k,n}\sqrt{E_k}\right|^2 + 2\Re\{h_{k,n}^*\sqrt{E_k} r_{k,n}\}}{N_0}\right. \\
&\qquad \left. + \sum_{n=1}^{N_R} \frac{\left|r_{k,n}\right|^2 + \left|h_{k,n}\sqrt{E_k}\right|^2 + 2\Re\{h_{k,n}^*\sqrt{E_k} r_{k,n}\}}{N_0}\right\} \\
&= \exp\left\{\sum_{n=1}^{N_R} \frac{4\Re\{h_{k,n}^*\sqrt{E_k} r_{k,n}\}}{N_0}\right\} \qquad \text{(A.3)}
\end{aligned}$$

such that the LLR equals

$$\Psi_{1,0}^{(k)}(\mathbf{r}_k, \mathbf{h}_k) = \sum_{n=1}^{N_R} \frac{4\Re\{h_{k,n}^*\sqrt{E_k} r_{k,n}\}}{N_0} = \sum_{n=1}^{N_R} \Psi_{1,0}^{(k,n)}(r_{k,n}, h_{k,n}), \qquad \text{(A.4)}$$

where $\Psi_{1,0}^{(k,n)}(r_{k,n}, h_{k,n}) = \frac{4\Re\{h_{k,n}^*\sqrt{E_k} r_{k,n}\}}{N_0}$ is the nth antenna LLR.

A.2 LLR PDF given $\mathbf{H}_k$ is known

Let $\Psi_{1,0}^{(k,n)}(r_{k,n}, h_{k,n})$ and $\Psi_{1,0}^{(k)}(\mathbf{r}_k, \mathbf{h}_k)$ be respectively the realizations of the random variables $\Psi_{1,0}^{(k,n)}(R_{k,n}, H_{k,n})$ and $\Psi_{1,0}^{(k)}(\mathbf{R}_k, \mathbf{H}_k)$. Assuming the transmitted bit B and the channel state vector $\mathbf{H}_k$ are known, the LLRs $\Psi_{1,0}^{(k,n)}(R_{k,n}, H_{k,n})$ $n = 1, \ldots, N_R$ are normally distributed random variables with mean $\mu(B, H_{k,n})$ and variance $\sigma^2(H_{k,n})$:

$$\begin{aligned}
\mu(B, H_{k,n}) &= E_{R_{k,n}|H_{k,n}}\left[\Psi_{1,0}^{(k,n)}(R_{k,n}, H_{k,n})\right] = E_{R_{k,n}|H_{k,n}}\left[\frac{4\sqrt{E_k}}{N_0}\Re\left\{H_{k,n}^* R_{k,n}\right\}\right] \\
&= E_{N_{k,n}}\left[\frac{4\sqrt{E_k}}{N_0}\Re\left\{H_{k,n}^*(H_{k,n}\sqrt{E_k}(-1)^{B+1} + N_{k,n})\right\}\right] \\
&= E_{N_{k,n}}\left[\frac{4E_k}{N_0}|H_{k,n}|^2(-1)^{B+1} + \frac{4\sqrt{E_k}}{N_0}\Re\left\{H_{k,n}^* N_{k,n}\right\}\right] \\
&= \frac{4E_k}{N_0}|H_{k,n}|^2(-1)^{B+1} + \frac{4\sqrt{E_k}}{N_0}\Re\left\{H_{k,n}^* E_{N_{k,n}}\left[N_{k,}\right]\right\} = \frac{4E_k}{N_0}|H_{k,n}|^2(-1)^{B+1} \qquad \text{(A.5)}
\end{aligned}$$

and

$$\begin{aligned}
&\sigma^2(H_{k,n}) \\
&= E_{R_{k,n}|H_{k,n}}\left[\left(\Psi_{1,0}^{(k,n)}(R_{k,n}, H_{k,n}) - \mu(B, H_{k,n})\right)^2\right] \\
&= E_{R_{k,n}|H_{k,n}}\left[\left(\frac{4\sqrt{E_k}}{N_0}\Re\left\{H_{k,n}^* R_{k,n}\right\} - \frac{4E_k}{N_0}|H_{k,n}|^2(-1)^{B+1}\right)^2\right] \\
&= E_{N_{k,n}}\left[\left(\frac{4\sqrt{E_k}}{N_0}\Re\left\{H_{k,n}^*\left(H_{k,n}\sqrt{E_k}(-1)^{B+1} + N_{k,n}\right)\right\} - \frac{4E_k}{N_0}|H_{k,n}|^2(-1)^{B+1}\right)^2\right] \\
&= E_{N_{k,n}}\left[\left(\frac{4\sqrt{E_k}}{N_0}\Re\left\{H_{k,n}^* N_{k,n}\right\}\right)^2\right] \\
&= \frac{16E_k}{(N_0)^2}E_{N_{k,n}}\Bigg[\Bigg(\Re\Big\{\left(\Re\{H_{k,n}\}\Re\{N_{k,n}\} + \Im\{H_{k,n}\}\Im\{N_{k,n}\}\right) \\
&\qquad\qquad + j\left(\Re\{H_{k,n}\}\Im\{N_{k,n}\} - \Im\{H_{k,n}\}\Re\{N_{k,n}\}\right)\Big\}\Bigg)^2\Bigg] \\
&= \frac{16E_k}{(N_0)^2}E_{N_{k,n}}\left[\left(\Re\{H_{k,n}\}\Re\{N_{k,n}\} + \Im\{H_{k,n}\}\Im\{N_{k,n}\}\right)^2\right] \\
&= \frac{16E_k}{(N_0)^2}E_{N_{k,n}}\left[\left(\Re\{H_{k,n}\}\Re\{N_{k,n}\}\right)^2\right] + E_{N_{k,n}}\left[\left(\Im\{H_{k,n}\}\Im\{N_{k,n}\}\right)^2\right] \\
&= \frac{16E_k}{(N_0)^2}\left(\Re\{H_{k,n}\}\right)^2 E_{N_{k,n}}\left[\left(\Re\{N_{k,n}\}\right)^2\right] + \left(\Im\{H_{k,n}\}\right)^2 E_{N_{k,n}}\left[\left(\Im\{N_{k,n}\}\right)^2\right] \\
&= \frac{8E_k}{N_0}\left(\Re\{H_{k,n}\}\right)^2 + \left(\Im\{H_{k,n}\}\right)^2 = \frac{8E_k}{N_0}|H_{k,n}|^2 \qquad \text{(A.6)}
\end{aligned}$$

Hence, given the transmitted bit B and the channel state vector $\mathbf{H}_k$ are known, the LLR $\Psi_{1,0}^{(k)}(\mathbf{R}_k, \mathbf{H}_k)$ is thus the sum of N_R independent normally distributed random variables and is itself normally distributed as follow

$$f_{\Psi_{1,0}^{(k)}(\mathbf{R}_k,\mathbf{H}_k)}(y \mid \mathbf{H}_k, B) = \frac{1}{\sqrt{2\pi\sigma^2(\mathbf{H}_k)}}\exp\left\{-\frac{(y - \mu(B,\mathbf{H}_k))^2}{2\sigma^2(\mathbf{H}_k)}\right\} \qquad \text{(A.7)}$$

where $\mu(B, \mathbf{H}_k) = \sum_{n=1}^{N_R}\mu(B, H_{k,n})$ and $\sigma^2(\mathbf{H}_k) = \sum_{n=1}^{N_R}\sigma^2(H_{k,n})$. Since

$$\mu(B, \mathbf{H}_k) = \sum_{n=1}^{N_R}\frac{4E_k}{N_0}|H_{k,n}|^2(-1)^{B+1} = \frac{4E_k}{N_0}\Omega_k^2(-1)^{B+1} = \mu(B, \Omega_k) \qquad \text{(A.8)}$$

and

$$\sigma^2(\mathbf{H}_k) = \sum_{n=1}^{N_R} \frac{8E_k}{N_0}|H_{k,n}|^2 = \frac{8E_k}{N_0}\Omega_k^2 = \sigma^2(\Omega_k), \tag{A.9}$$

the PDF $f_{\Psi_{1,0}^{(k)}(\mathbf{R}_k,\mathbf{H}_k)}(y \mid \mathbf{H}_k, B)$ reduces to $f_{\Psi_{1,0}^{(k)}(\mathbf{R}_k,\mathbf{H}_k)}(y \mid \Omega_k, B)$, where $\Omega_k = \sqrt{\sum_{n=1}^{N_R} |H_{k,n}|^2}$.

A.3 LLR CDF given $\mathbf{H}_k$ is known

Using (A.7), the LLR CDF $F_{\Psi_{1,0}^{(k)}(\mathbf{R}_k,\mathbf{H}_k)}(x \mid \mathbf{H}_k, B)$ can be evaluated as follows

$$\begin{aligned}
&F_{\Psi_{1,0}^{(k)}(\mathbf{R}_k,\mathbf{H}_k)}(x \mid \mathbf{H}_k, B) \\
&= \int_{-\infty}^{x} f_{\Psi_{1,0}^{(k)}(\mathbf{R}_k,\mathbf{H}_k)}(y \mid \mathbf{H}_k, B)dy = \int_{-\infty}^{x} f_{\Psi_{1,0}^{(k)}(\mathbf{R}_k,\mathbf{H}_k)}(y \mid \Omega_k, B)dy \\
&= \int_{-\infty}^{x} \frac{1}{\sqrt{2\pi\sigma^2(\Omega_k)}} \exp\left\{-\frac{(y-\mu(B,\Omega_k))^2}{2\sigma^2(\Omega_k)}\right\} dy = 1 - Q\left[\frac{x-\mu(B,\Omega_k)}{\sigma(\Omega_k)}\right] \\
&= 1 - Q\left[\frac{x}{2\sqrt{2\frac{E_k}{N_0}\Omega_k^2}} - \frac{(-1)^{B+1}\sqrt{2\frac{E_k}{N_0}\Omega_k^2}}{1}\right] = F_{\Psi_{1,0}^{(k)}(\mathbf{R}_k,\mathbf{H}_k)}(x \mid \Omega_k, B). \qquad \text{(A.10)}
\end{aligned}$$

A.4 LLR PDF given $\mathbf{H}_k$ is unknown

Assuming now the transmitted bit B is known but the channel fading $\mathbf{H}$ is unknown, the LLRs $\Psi_{1,0}^{(k,n)}(R_{k,n}, H_{k,n})$ $n = 1, \ldots, N_R$ are independent random variables since the channel fading is spatially uncorrelated. Thus, the kth base station LLR $\Psi_{1,0}^{(k)}(\mathbf{R}_k, \mathbf{H}_k)$ is the sum of the N_R conditionally independent antenna LLRs $\Psi_{1,0}^{(k,n)}(R_{k,n}, H_{k,n})$ $n = 1, \ldots, N_R$ and its PDF equals to

$$f_{\Psi_{1,0}^{(k)}(\mathbf{R}_k,\mathbf{H}_k)}(y \mid B) = f_{\Psi_{1,0}^{(k,1)}(R_{k,1},H_{k,1})}(y \mid B) \star \cdots \star f_{\Psi_{1,0}^{(k,N_R)}(R_{k,N_R},H_{k,N_R})}(y \mid B), \tag{A.11}$$

where $\star$ represents the convolution operator and

$$\begin{aligned}
f_{\Psi_{1,0}^{(k,n)}(R_{k,n},H_{k,n})}(y \mid B) &= \int_{h_{k,n}} f_{\Psi_{1,0}^{(k,n)}(R_{k,n},H_{k,n})}(y \mid h_{k,n}, B) f_{H_{k,n}}(h_{k,n}) dh_{k,n} \\
&= \int_{0}^{\infty} f_{\Psi_{1,0}^{(k,n)}(R_{k,n},H_{k,n})}(y \mid |h_{k,n}|, B) f_{|H_{k,n}|}(|h_{k,n}|) d|h_{k,n}| \qquad \text{(A.12)}
\end{aligned}$$

since the mean $\mu(B, h_{k,n})$ and variance $\sigma^2(h_{k,n})$ defining $f_{\Psi_{1,0}^{(k,n)}(R_{k,n},H_{k,n})}(x \mid h_{k,n}, B)$ are independent of the phase of $h_{k,n}$. Hence, since the fading magnitude $|H_{k,n}|$ is Rayleigh distributed as follows

$$f_{|H_{k,n}|}(|h_{k,n}|) = 2|h_{k,n}| \exp\left\{-|h_{k,n}|^2\right\}, \tag{A.13}$$

the antenna LLR PDF $f_{\Psi_{1,0}^{(k,n)}(R_{k,n},H_{k,n})}(y \mid B)$ can be evaluated using (A.12) as follows

$$\begin{aligned}
&f_{\Psi_{1,0}^{(k,n)}(R_{k,n},H_{k,n})}(y \mid B) \\
&= \int_0^\infty \frac{1}{\sqrt{2\pi\sigma^2(h_{k,n})}} \exp\left\{-\frac{(y-\mu(B,h_{k,n}))^2}{2\sigma^2(h_{k,n})}\right\} 2|h_{k,n}| \exp\left\{-|h_{k,n}|^2\right\} d|h_{k,n}| \\
&= \int_0^\infty \frac{1}{2\sqrt{\pi}} \frac{1}{\sqrt{\frac{E_k}{N_0}}} \exp\left\{-|h_{k,n}|^2 - \frac{(y - 4\frac{E_k}{N_0}|h_{k,n}|^2(-1)^{B+1})^2}{16\frac{E_k}{N_0}|h_{k,n}|^2}\right\} d|h_{k,n}| \\
&= \int_0^\infty \frac{1}{2\sqrt{\pi}} \frac{1}{\sqrt{\frac{E_k}{N_0}}} \exp\left\{-|h_{k,n}|^2 - \left(\frac{y^2 + 16(\frac{E_k}{N_0})^2|h_{k,n}|^4 - 8y\frac{E_k}{N_0}|h_{k,n}|^2(-1)^{B+1}}{16\frac{E_k}{N_0}|h_{k,n}|^2}\right)\right\} d|h_{k,n}| \\
&= \int_0^\infty \frac{1}{2\sqrt{\pi}} \frac{1}{\sqrt{\frac{E_k}{N_0}}} \exp\left\{-|h_{k,n}|^2 - \frac{y^2}{16}\frac{1}{|h_{k,n}|^2\frac{E_k}{N_0}} - \frac{|h_{k,n}|^2\frac{E_k}{N_0}}{1} + \frac{(-1)^{B+1}y}{2}\right\} d|h_{k,n}| \\
&= \frac{1}{2\frac{E_k}{N_0}} \exp\left\{\frac{(-1)^{B+1}y}{2}\right\} \frac{1}{\sqrt{\pi}} \int_0^\infty \exp\left\{-u^2\left(\frac{1}{\frac{E_k}{N_0}}+1\right) - \frac{y^2}{16u^2}\right\} du \qquad \text{(A.14)} \\
&= \frac{(\beta_k^2-1)}{2} \exp\left\{\frac{(-1)^{B+1}y}{2}\right\} \frac{1}{\sqrt{\pi}} \int_0^\infty \exp\left\{-u^2\beta_k^2 - \frac{y^2}{16u^2}\right\} du \qquad \text{(A.15)}
\end{aligned}$$

where the change of variable $u = \sqrt{\frac{E_k}{N_0}}|h_{k,n}|$ is made in expression (A.14),

$$\beta_k = \sqrt{\frac{\frac{E_k}{N_0}+1}{\frac{E_k}{N_0}}} \quad \text{and} \quad (\beta_k^2 - 1) = \frac{\frac{E_k}{N_0}+1}{\frac{E_k}{N_0}} - 1 = \frac{1}{\frac{E_k}{N_0}}.$$

Since as shown in [32]

$$\int_0^\infty \exp\left\{-ax^2 - \frac{b}{x^2}\right\} dx = \frac{1}{2}\sqrt{\frac{\pi}{a}} \exp\left\{-2\sqrt{ab}\right\} \qquad a > 0 \;\; b > 0, \tag{A.16}$$

the antenna LLR PDF $f_{\Psi_{1,0}^{(k,n)}(R_{k,n},H_{k,n})}(y \mid B)$ simplifies to the following closed form expression

$$f_{\Psi_{1,0}^{(k,n)}(R_{k,n},H_{k,n})}(y \mid B) = A_k \begin{cases} \exp\left\{B_k^+ y\right\} & \text{if } y \geq 0 \\ \exp\left\{B_k^- y\right\} & \text{if } y < 0 \end{cases} \tag{A.17}$$

where $A_k = \frac{(\beta_k^2-1)}{4\beta_k}$, $B_k^+ = \frac{(-1)^{B+1}-\beta_k}{2} < 0$ and $B_k^- = \frac{(-1)^{B+1}+\beta_k}{2} > 0$. Since $f_{\Psi_{1,0}^{(k,n)}(R_{k,n},H_{k,n})}(y \mid B)$ is a PDF,

$$\begin{aligned}
&\int_{-\infty}^{\infty} f_{\Psi_{1,0}^{(k,n)}(R_{k,n},H_{k,n})}(y \mid B)dy \\
&= \left[\frac{A_k}{B_k^-}\exp\left\{B_k^- y\right\}\right]_{-\infty}^{0} + \left[\frac{A_k}{B_k^+}\exp\left\{B_k^+ y\right\}\right]_{0}^{\infty} = \frac{A_k}{B_k^-} - \frac{A_k}{B_k^+} \\
&= \frac{(\beta_k^2-1)}{4\beta_k}\left(\frac{2}{(-1)^{B+1}+\beta_k} - \frac{2}{(-1)^{B+1}-\beta_k}\right) \\
&= \frac{(\beta_k^2-1)}{4\beta_k}\left(\frac{2((-1)^{B+1}-\beta_k) - 2((-1)^{B+1}+\beta_k)}{1-\beta_k^2}\right) \\
&= \frac{(\beta_k^2-1)}{4\beta_k}\left(\frac{-4\beta_k}{1-\beta_k^2}\right) = 1.
\end{aligned} \tag{A.18}$$

Using the derived antenna LLR PDF $f_{\Psi_{1,0}^{(k,n)}(R_{k,n},H_{k,n})}(y \mid B)$ and expression (A.11), it is possible to derive the kth base station LLR PDF $f_{\Psi_{1,0}^{(k)}(\mathbf{R}_k,\mathbf{H}_k)}(y \mid B)$ for any number of receiving antennas. In the next two sub-sections, we evaluate the convolution integral (A.11) for the cases where the kth base station is equipped with 2 and 3 receiving antennas.

A.4.1 Two receiving antennas

When the kth base station is equipped with two receiving antennas, the local detector LLR PDF $f_{\Psi_{1,0}^{(k)}(\mathbf{R}_k,\mathbf{H}_k)}(y \mid B)$ equals to

$$\begin{aligned}
f_{\Psi_{1,0}^{(k)}(\mathbf{R}_k,\mathbf{H}_k)}(y \mid B) &= f_{\Psi_{1,0}^{(k,1)}(R_{k,1},H_{k,1})}(y \mid B) \star f_{\Psi_{1,0}^{(k,2)}(R_{k,2},H_{k,2})}(y \mid B) \\
&= \int_{-\infty}^{\infty} f_{\Psi_{1,0}^{(k,1)}(R_{k,1},H_{k,1})}(x \mid B) f_{\Psi_{1,0}^{(k,2)}(R_{k,2},H_{k,2})}(y - x \mid B)dx.
\end{aligned} \tag{A.19}$$

Hence, since $B_k^- - B_k^+ = \beta_k$, if y is larger than 0

$$
\begin{aligned}
& f_{\Psi_{1,0}^{(k)}(\mathbf{R}_k,\mathbf{H}_k)}(y \mid B) \\
= & \int_{-\infty}^{0} A_k \exp\left\{B_k^- x\right\} A_k \exp\left\{B_k^+(y-x)\right\} dx + \int_{0}^{y} A_k \exp\left\{B_k^+ x\right\} A_k \exp\left\{B_k^+(y-x)\right\} dx \\
& + \int_{y}^{\infty} A_k \exp\left\{B_k^+ x\right\} A_k \exp\left\{B_k^-(y-x)\right\} dx \\
= & A_k^2 \left\{ \exp\left\{B_k^+ y\right\} \int_{-\infty}^{0} \exp\left\{\beta_k x\right\} dx + \exp\left\{B_k^+ y\right\} \int_{0}^{y} dx + \exp\left\{B_k^- y\right\} \int_{y}^{\infty} \exp\left\{-\beta_k x\right\} dx \right\} \\
= & A_k^2 \left\{ \exp\left\{B_k^+ y\right\} \left[\frac{1}{\beta_k} \exp\left\{x\beta_k\right\}\right]_{-\infty}^{0} + \exp\left\{B_k^+ y\right\} \left[x\right]_{0}^{y} + \exp\left\{B_k^- y\right\} \left[\frac{-1}{\beta_k} \exp\left\{-x\beta_k\right\}\right]_{y}^{\infty} \right\} \\
= & A_k^2 \left\{ \frac{2}{\beta_k} + y \right\} \exp\left\{B_k^+ y\right\}
\end{aligned} \tag{A.20}
$$

and, if y is smaller than 0,

$$
\begin{aligned}
& f_{\Psi_{1,0}^{(k)}(\mathbf{R}_k,\mathbf{H}_k)}(y \mid B) \\
= & \int_{-\infty}^{y} A_k \exp\left\{B_k^- x\right\} A_k \exp\left\{B_k^+(y-x)\right\} dx + \int_{y}^{0} A_k \exp\left\{B_k^- x\right\} A_k \exp\left\{B_k^-(y-x)\right\} dx \\
& + \int_{0}^{\infty} A_k \exp\left\{B_k^+ x\right\} A_k \exp\left\{B_k^-(y-x)\right\} dx \\
= & A_k^2 \left\{ \exp\left\{B_k^+ y\right\} \int_{-\infty}^{y} \exp\left\{x\beta_k\right\} dx + \exp\left\{B_k^- y\right\} \int_{y}^{0} dx + \exp\left\{B_k^- y\right\} \int_{0}^{\infty} \exp\left\{-x\beta_k\right\} dx \right\} \\
= & A_k^2 \left\{ \exp\left\{B_k^+ y\right\} \left[\frac{1}{\beta_k} \exp\left\{x\beta_k\right\}\right]_{-\infty}^{y} + \exp\left\{B_k^- y\right\} \left[x\right]_{y}^{0} + \exp\left\{B_k^- y\right\} \left[\frac{-1}{\beta_k} \exp\left\{-x\beta_k\right\}\right]_{0}^{\infty} \right\} \\
= & A_k^2 \left\{ \frac{2}{\beta_k} - y \right\} \exp\left\{B_k^- y\right\}.
\end{aligned} \tag{A.21}
$$

In summary, when the kth base station is equipped with two antennas, the local detector LLR PDF $f_{\Psi_{1,0}^{(k)}(\mathbf{R}_k,\mathbf{H}_k)}(y \mid B)$ can be expressed as follows

$$
f_{\Psi_{1,0}^{(k)}(\mathbf{R}_k,\mathbf{H}_k)}(y \mid B) = A_k^2 \begin{cases} \left(\frac{2}{\beta_k} + y\right) \exp\left\{B_k^+ y\right\} & \text{if } y \geq 0 \\ \left(\frac{2}{\beta_k} - y\right) \exp\left\{B_k^- y\right\} & \text{if } y < 0 \end{cases}. \tag{A.22}
$$

A.4.2 Three receiving antennas

When the kth base station is equipped with three receiving antennas, the local detector LLR PDF $f_{\Psi_{1,0}^{(k)}(\mathbf{R}_k,\mathbf{H}_k)}(y \mid B)$ equals to

$$\begin{aligned}
f_{\Psi_{1,0}^{(k)}(\mathbf{R}_k,\mathbf{H}_k)}(y \mid B) &= f_{\Psi_{1,0}^{(k,1)}(R_{k,1},H_{k,1})}(y \mid B) \star f_{\Psi_{1,0}^{(k,2)}(R_{k,2},H_{k,2})}(y \mid B) \star f_{\Psi_{1,0}^{(k,3)}(R_{k,3},H_{k,3})}(y \mid B) \\
&= g(y) \star f_{\Psi_{1,0}^{(k,3)}(R_{k,3},H_{k,3})}(y \mid B) \\
&= \int_{-\infty}^{\infty} g(x) f_{\Psi_{1,0}^{(k,3)}(R_{k,3},H_{k,3})}(y - x \mid B) dx, \qquad \text{(A.23)}
\end{aligned}$$

where $g(y) = f_{\Psi_{1,0}^{(k,1)}(R_{k,1},H_{k,1})}(y \mid B) \star f_{\Psi_{1,0}^{(k,2)}(R_{k,2},H_{k,2})}(y \mid B)$ is derived in the previous section. Therefore, using (A.23) the local detector LLR PDF $f_{\Psi_{1,0}^{(k)}(\mathbf{R}_k,\mathbf{H}_k)}(y \mid B)$ can be evaluated such that, if y is larger than 0,

$$\begin{aligned}
& f_{\Psi_{1,0}^{(k)}(\mathbf{R}_k,\mathbf{H}_k)}(y \mid B) \\
&= \int_{-\infty}^{0} A_k^2 \left(\frac{2}{\beta_k} - x\right) \exp\left\{B_k^- x\right\} A_k \exp\left\{B_k^+ (y-x)\right\} dx \\
&\quad + \int_{0}^{y} A_k^2 \left(\frac{2}{\beta_k} + x\right) \exp\left\{B_k^+ x\right\} A_k \exp\left\{B_k^+ (y-x)\right\} dx \\
&\quad + \int_{y}^{\infty} A_k^2 \left(\frac{2}{\beta_k} + x\right) \exp\left\{B_k^+ x\right\} A_k \exp\left\{B_k^- (y-x)\right\} dx \\
&= A_k^3 \exp\left\{B_k^+ y\right\} \left(\int_{-\infty}^{0} \frac{2}{\beta_k} \exp\{x\beta_k\} dx - \int_{-\infty}^{0} x \exp\{x\beta_k\} dx + \int_0^y \left(\frac{2}{\beta_k} + x\right) dx\right) \\
&\quad + A_k^3 \exp\left\{B_k^- y\right\} \left(\int_y^{\infty} \frac{2}{\beta_k} \exp\{-x\beta_k\} dx + \int_y^{\infty} x \exp\{-x\beta_k\} dx\right) \\
&= A_k^3 \exp\left\{B_k^+ y\right\} \left(\left[\frac{2}{\beta_k^2} \exp\{x\beta_k\} - \frac{x \exp\{x\beta_k\}}{\beta_k} + \frac{\exp\{x\beta_k\}}{\beta_k^2}\right]_{-\infty}^{0} + \left[\frac{2x}{\beta_k} + \frac{x^2}{2}\right]_0^y\right) \\
&\quad + A_k^3 \exp\left\{B_k^- y)\right\} \left(\left[-\frac{2}{\beta_k^2} \exp\{-x\beta_k\} - \frac{x \exp\{-x\beta_k\}}{\beta_k} - \frac{\exp\{-x\beta_k\}}{\beta_k^2}\right]_y^{\infty}\right) \\
&= A_k^3 \left[\frac{6}{\beta_k^2} + \frac{3y}{\beta_k} + \frac{y^2}{2}\right] \exp\left\{B_k^+ y\right\} \qquad \text{(A.24)}
\end{aligned}$$

and, if y is smaller than 0,

$$\begin{aligned}
&f_{\Psi_{1,0}^{(k)}(\mathbf{R}_k,\mathbf{H}_k)}(y \mid B) \\
&= \int_{-\infty}^{y} A_k^2\left(\frac{2}{\beta_k}-x\right)\exp\left\{B_k^- x\right\} A_k \exp\left\{B_k^+(y-x)\right\} dx \\
&\quad + \int_{y}^{0} A_k^2\left(\frac{2}{\beta_k}-x\right)\exp\left\{B_k^- x\right\} A_k \exp\left\{B_k^-(y-x)\right\} dx \\
&\quad + \int_{0}^{\infty} A_k^2\left(\frac{2}{\beta_k}+x\right)\exp\left\{B_k^+ x\right\} A_k \exp\left\{B_k^-(y-x)\right\} dx \\
&= A_k^3 \exp\left\{B_k^+ y\right\}\left(\int_{-\infty}^{y}\frac{2}{\beta_k}\exp\{x\beta_k\}\,dx - \int_{-\infty}^{y} x\exp\{x\beta_k\}\,dx\right) \\
&\quad + A_k^3 \exp\left\{B_k^- y\right\}\left(\int_{y}^{0}\left(\frac{2}{\beta_k}-x\right)dx + \int_{0}^{\infty}\frac{2}{\beta_k}\exp\{-x\beta_k\}\,dx + \int_{0}^{\infty} x\exp\{-x\beta_k\}\,dx\right) \\
&= A_k^3 \exp\left\{B_k^+ y\right\}\left[\frac{2}{\beta_k^2}\exp\{x\beta_k\} - \frac{x\exp\{x\beta_k\}}{\beta_k} + \frac{\exp\{x\beta_k\}}{\beta_k^2}\right]_{-\infty}^{y} \\
&\quad + A_k^3 \exp\left\{B_k^- y\right\}\left(\left[\frac{2x}{\beta_k}-\frac{x^2}{2}\right]_{y}^{0} + \left[-\frac{2}{\beta_k^2}\exp\{-x\beta_k\} - \frac{x\exp\{-x\beta_k\}}{\beta_k} - \frac{\exp\{-x\beta_k\}}{\beta_k^2}\right]_{0}^{\infty}\right) \\
&= A_k^3\left[\frac{6}{\beta_k^2}-\frac{3y}{\beta_k}+\frac{y^2}{2}\right]\exp\left\{B_k^- y\right\}, \qquad \text{(A.25)}
\end{aligned}$$

where we used the following integral

$$H_a^{(n)}(x) = \int x^n \exp\{ax\}\,dx = \sum_{k=0}^{n}(-1)^k \frac{n!}{(n-k)!}\frac{x^{n-k}}{a^{k+1}}\exp\{ax\} \qquad \text{(A.26)}$$

to obtain expression (A.24) and (A.25).

In summary, when the kth base station is equipped with three antennas, the local detector LLR PDF $f_{\Psi_{1,0}^{(k)}(\mathbf{R}_k,\mathbf{H}_k)}(y \mid B)$ can be expressed as follows

$$f_{\Psi_{1,0}^{(k)}(\mathbf{R}_k,\mathbf{H}_k)}(y \mid B) = A_k^3\begin{cases}\left(\frac{6}{\beta_k^2}+\frac{3y}{\beta_k}+\frac{y^2}{2}\right)\exp\left\{B_k^+ y\right\} & \text{if } y \geq 0 \\ \left(\frac{6}{\beta_k^2}-\frac{3y}{\beta_k}+\frac{y^2}{2}\right)\exp\left\{B_k^- y\right\} & \text{if } y < 0\end{cases}. \qquad \text{(A.27)}$$

A.5 LLR CDF given $\mathbf{H}_k$ is unknown

Using (A.17), (A.22) or (A.27), the kth base station LLR CDF $F_{\Psi_{1,0}^{(k)}(\mathbf{R}_k,\mathbf{H}_k)}(x \mid B)$ can be evaluated using integral (A.26) for 1, 2 and 3 receiving antennas.

A.5.1 Single receiving antenna

If x is larger than 0,

$$\begin{aligned} F_{\Psi_{1,0}^{(k)}(\mathbf{R}_k,\mathbf{H}_k)}(x \mid B) &= \int_{-\infty}^{x} f_{\Psi_{1,0}^{(k)}(\mathbf{R}_k,\mathbf{H}_k)}(y \mid B)dy = 1 - \int_{x}^{\infty} f_{\Psi_{1,0}^{(k)}(\mathbf{R}_k,\mathbf{H}_k)}(y \mid B)dy \\ &= 1 - \int_{x}^{\infty} A_k \exp\left\{B_k^+ y\right\} dy = 1 + \frac{A_k}{B_k^+} \exp\left\{B_k^+ x\right\}, \end{aligned} \quad (A.28)$$

and, if x is smaller than 0,

$$F_{\Psi_{1,0}^{(k)}(\mathbf{R}_k,\mathbf{H}_k)}(x \mid B) = \int_{-\infty}^{x} f_{\Psi_{1,0}^{(k)}(\mathbf{R}_k,\mathbf{H}_k)}(y \mid B)dy = \int_{-\infty}^{x} A_k \exp\left\{B_k^- y\right\} dy = \frac{A_k}{B_k^-} \exp\left\{B_k^- x\right\}. \quad (A.29)$$

A.5.2 Two receiving antennas

If x is larger than 0,

$$\begin{aligned} F_{\Psi_{1,0}^{(k)}(\mathbf{R}_k,\mathbf{H}_k)}(x \mid B) &= \int_{-\infty}^{x} f_{\Psi_{1,0}^{(k)}(\mathbf{R}_k,\mathbf{H}_k)}(y \mid B)dy = 1 - \int_{x}^{\infty} f_{\Psi_{1,0}^{(k)}(\mathbf{R}_k,\mathbf{H}_k)}(y \mid B)dy \\ &= 1 - \int_{x}^{\infty} A_k^2 \left(\frac{2}{\beta_k} + y\right) \exp\left\{B_k^+ y\right\} dy \\ &= 1 - A_k^2 \left(\frac{2}{\beta_k}\right) \left[\frac{1}{B_k^+} \exp\left\{B_k^+ y\right\}\right]_x^{\infty} \\ &\quad - A_k^2 \left[\frac{y}{B_k^+} \exp\left\{B_k^+ y\right\} - \frac{1}{(B_k^+)^2} \exp\left\{B_k^+ y\right\}\right]_x^{\infty} \\ &= 1 + A_k^2 \left(\left(\frac{2}{\beta_k} + x\right) \frac{1}{B_k^+} - \frac{1}{(B_k^+)^2}\right) \exp\left\{B_k^+ x\right\} \end{aligned} \quad (A.30)$$

and, if x is smaller than 0,

$$
\begin{aligned}
& F_{\Psi_{1,0}^{(k)}(\mathbf{R}_k,\mathbf{H}_k)}(x \mid B) \\
= & \int_{-\infty}^{x} f_{\Psi_{1,0}^{(k)}(\mathbf{R}_k,\mathbf{H}_k)}(y \mid B)dy = \int_{-\infty}^{x} A_k^2 \left(\frac{2}{\beta_k} - y\right) \exp\left\{B_k^- y\right\} dy \\
= & A_k^2 \left(\frac{2}{\beta_k}\right) \left[\frac{1}{B_k^-} \exp\left\{B_k^- y\right\}\right]_{-\infty}^{x} - A_k^2 \left[\frac{y}{B_k^-} \exp\left\{B_k^- y\right\} - \frac{1}{(B_k^-)^2} \exp\left\{B_k^- y\right\}\right]_{-\infty}^{x} \\
= & A_k^2 \left(\left(\frac{2}{\beta_k} - x\right) \frac{1}{B_k^-} + \frac{1}{(B_k^-)^2}\right) \exp\left\{B_k^- x\right\} \qquad (A.31)
\end{aligned}
$$

A.5.3 Three receiving antennas

If x is larger than 0,

$$
\begin{aligned}
& F_{\Psi_{1,0}^{(k)}(\mathbf{R}_k,\mathbf{H}_k)}(x \mid B) \\
= & \int_{-\infty}^{x} f_{\Psi_{1,0}^{(k)}(\mathbf{R}_k,\mathbf{H}_k)}(y \mid B)dy = 1 - \int_{x}^{\infty} f_{\Psi_{1,0}^{(k)}(\mathbf{R}_k,\mathbf{H}_k)}(y \mid B)dy \\
= & 1 - \int_{x}^{\infty} A_k^3 \left[\frac{6}{\beta_k^2} + \frac{3y}{\beta_k} + \frac{y^2}{2}\right] \exp\left\{B_k^+ y\right\} dy \\
= & 1 - A_k^3 \left(\frac{6}{\beta_k^2}\right) \left[\frac{1}{B_k^+} \exp\left\{B_k^+ y\right\}\right]_{x}^{\infty} - A_k^3 \left(\frac{3}{\beta_k}\right) \left[\left(\frac{y}{B_k^+} - \frac{1}{(B_k^+)^2}\right) \exp\left\{B_k^+ y\right\}\right]_{x}^{\infty} \\
& - A_k^3 \left(\frac{1}{2}\right) \left[\left(\frac{y^2}{B_k^+} - \frac{2y}{(B_k^+)^2} + \frac{2}{(B_k^+)^3}\right) \exp\left\{B_k^+ y\right\}\right]_{x}^{\infty} \\
= & 1 + A_k^3 \left[\left(\frac{6}{\beta_k^2} + \frac{3x}{\beta_k} + \frac{x^2}{2}\right) \frac{1}{B_k^+} - \left(\frac{3}{\beta_k} + x\right) \frac{1}{(B_k^+)^2} + \frac{1}{(B_k^+)^3}\right] \exp\left\{B_k^+ x\right\} \qquad (A.32)
\end{aligned}
$$

and, if x is smaller than 0,

$$
\begin{aligned}
& F_{\Psi_{1,0}^{(k)}(\mathbf{R}_k,\mathbf{H}_k)}(x \mid B) \\
= & \int_{-\infty}^{x} f_{\Psi_{1,0}^{(k)}(\mathbf{R}_k,\mathbf{H}_k)}(y \mid B)dy = \int_{-\infty}^{x} A_k^3 \left[\frac{6}{\beta_k^2} - \frac{3y}{\beta_k} + \frac{y^2}{2}\right] \exp\left\{B_k^- y\right\} dy \\
= & A_k^3 \left(\frac{6}{\beta_k^2}\right) \left[\frac{1}{B_k^-} \exp\left\{B_k^- y\right\}\right]_{-\infty}^{x} - A_k^3 \left(\frac{3}{\beta_k}\right) \left[\left(\frac{y}{B_k^-} - \frac{1}{(B_k^-)^2}\right) \exp\left\{B_k^- y\right\}\right]_{-\infty}^{x} \\
& + A_k^3 \left(\frac{1}{2}\right) \left[\left(\frac{y^2}{B_k^-} - \frac{2y}{(B_k^-)^2} + \frac{2}{(B_k^-)^3}\right) \exp\left\{B_k^- y\right\}\right]_{-\infty}^{x} \\
= & A_k^3 \left[\left(\frac{6}{\beta_k^2} - \frac{3x}{\beta_k} + \frac{x^2}{2}\right) \frac{1}{B_k^-} + \left(\frac{3}{\beta_k} - x\right) \frac{1}{(B_k^-)^2} + \frac{1}{(B_k^-)^3}\right] \exp\left\{B_k^- x\right\} \qquad \text{(A.33)}
\end{aligned}
$$

Appendix B

Evaluation of the BER of the SHDD schemes

In chapter 2, three different SHDD schemes are proposed which are the $\mathrm{SHDD}_{1,opt}$, $\mathrm{SHDD}_{1,sub}$ and $\mathrm{SHDD}_{2,opt}$ schemes. In this appendix, we discuss the evaluation of the probability of bit error for each scheme.

B.1 $\mathrm{SHDD}_{1,opt}$ scheme

The $\mathrm{SHDD}_{1,opt}$ scheme is a SHDD scheme for which the channel state vector $\mathbf{H}$ is known at the MSC and the local detector thresholds $\mathbf{t}(\mathbf{h})$ are optimized with respect to the conditional probability of bit error $P_{b|\mathbf{h}}(\mathbf{t}(\mathbf{h}))$. For such a scheme, the average probability of bit error after the threshold optimization can be evaluated as follows

$$P_b^* = \int_{\mathbf{h}} P_{b|\mathbf{h}}^* f_{\mathbf{H}}(\mathbf{h}) d\mathbf{h}, \tag{B.1}$$

where $P_{b|\mathbf{h}}^*$ represents the conditional probability of bit error after the threshold optimization and

$$P_{b|\mathbf{h}}^* = \min_{\mathbf{t}(\mathbf{h})} \left\{ P_{b|\mathbf{h}}(\mathbf{t}(\mathbf{h})) \right\}. \tag{B.2}$$

The conditional probability of bit error $P_{b|\mathbf{h}}(\mathbf{t}(\mathbf{h}))$ is defined by expression (2.5). However, using results from [33], it is possible to reformulate the conditional probability of bit error as

$$P_{b|\mathbf{h}}(\mathbf{t}(\mathbf{h})) = \frac{1}{2} - \frac{1}{4}\sum_{\mathbf{u}}\Big|P(\mathbf{U}=\mathbf{u} \mid \mathbf{h}, B=0) - P(\mathbf{U}=\mathbf{u} \mid \mathbf{h}, B=1)\Big|, \quad \text{(B.3)}$$

where, since the local decisions are conditionally independent,

$$P(\mathbf{U}=\mathbf{u} \mid \mathbf{h}, B) = \prod_{k=1}^{N_{BS}} P(U_k = u_k \mid \mathbf{h}, B). \quad \text{(B.4)}$$

Using results from Appendix A,

$$\begin{aligned} P(U_k = u_k \mid \mathbf{h}, B) &= \int_{t_{k,u_k}(\mathbf{h})}^{t_{k,u_k+1}(\mathbf{h})} f_{\Psi_{1,0}^{(k)}(\mathbf{R}_k,\mathbf{H}_k)}(x \mid \mathbf{h}_k, B)dx \\ &= F_{\Psi_{1,0}^{(k)}(\mathbf{R}_k,\mathbf{H}_k)}\left(t_{k,u_k+1}(\mathbf{h}) \mid \mathbf{h}_k, B\right) - F_{\Psi_{1,0}^{(k)}(\mathbf{R}_k,\mathbf{H}_k)}\left(t_{k,u_k}(\mathbf{h}) \mid \mathbf{h}_k, B\right) \quad \text{(B.5)} \end{aligned}$$

such that the conditional probability of bit error $P_{b|\mathbf{h}}(\mathbf{t}(\mathbf{h}))$ has a closed form expression, simplifying the optimization process.

Due to the optimization in the integrand of expression (B.1), it is obvious that the integration required to evaluate the average probability of bit error needs to be performed numerically. As shown in Appendix A, the CDF $F_{\Psi_{1,0}^{(k)}(\mathbf{R}_k,\mathbf{H}_k)}(x \mid \mathbf{H}_k, B)$ can be expressed as a function of $\Omega_k = \sum_{n=1}^{N_R} |H_{k,n}|^2$ such that the conditional probability of bit error can be expressed as a function of $\mathbf{\Omega} = [\Omega_1, \ldots, \Omega_{N_{BS}}]^T$. Consequently, the average probability of bit error (B.1) can be reformulated as follows

$$P_b^* = \int_{\omega} P_{b|\omega}^* f_{\mathbf{\Omega}}(\omega)d\omega, \quad \text{(B.6)}$$

where

$$f_{\mathbf{\Omega}}(\omega) = f_{\Omega_1}(\omega_1)\cdots f_{\Omega_{N_{BS}}}(\omega_{N_{BS}}) \quad \text{(B.7)}$$

and, as shown in [30],

$$f_{\Omega_k}(\omega_k) = \frac{2\omega_k^{2N_R-1}}{(N_R-1)!}\exp\left\{-\omega_k^2\right\} \quad \text{(B.8)}$$

which is a generalized Rayleigh distribution. This has for effect to simplify the evaluation of the probability of bit error since it reduces the order of the numerical integration from $2N_{BS}N_R$ to N_{BS}.

B.2 $\text{SHDD}_{1,sub}$ scheme

The $\text{SHDD}_{1,sub}$ scheme is a SHDD scheme for which the channel state vector $\mathbf{H}$ is also known at the MSC but, as opposed to the $\text{SHDD}_{1,opt}$ scheme, the local detector thresholds $\mathbf{t}$ are independent of $\mathbf{H}$ and are optimized with respect to the average probability of bit error $P_b(\mathbf{t})$. For such a scheme, the average probability of bit error after the threshold optimization can be evaluated as follows

$$P_b^* = \min_{\mathbf{t}} \left\{ P_b(\mathbf{t}) \right\}, \tag{B.9}$$

where

$$P_b(\mathbf{t}) = \int_{\mathbf{h}} P_{b|\mathbf{h}}(\mathbf{t}) f_{\mathbf{H}}(\mathbf{h}) d\mathbf{h} \tag{B.10}$$

given $P_{b|\mathbf{h}}(\mathbf{t})$ is defined as for the $\text{SHDD}_{1,opt}$ by expression (B.3). As opposed to the previous case, the minimization is outside the integral but, unfortunately, due to the absolute value operation in the conditional probability of bit error $P_{b|\mathbf{h}}(\mathbf{t})$ (B.3), the integration still needs to be performed numerically. Furthermore, since $P_b(\mathbf{t})$ is the threshold optimization cost function, the integral has to be evaluated multiple times during the optimization process making the optimization time consuming. It is therefore advantageous, as for the $\text{SHDD}_{1,opt}$ scheme, to perform the averaging with respect to $\boldsymbol{\Omega}$ such that

$$P_b(\mathbf{t}) = \int_{\omega} P_{b|\omega}(\mathbf{t}) f_{\boldsymbol{\Omega}}(\omega) d\omega, \tag{B.11}$$

where $f_{\boldsymbol{\Omega}}(\omega)$ is defined in (B.7). However, the evaluation of the average probability of bit error $P_b(\mathbf{t})$ still requires a relatively large amount of time, making the optimization of the thresholds time consuming, especial for systems using more than two base stations since the order of the numerical integration increases with the number of base stations.

B.3 $\text{SHDD}_{2,opt}$ scheme

The $\text{SHDD}_{2,opt}$ scheme is a SHDD scheme for which the channel state vector $\mathbf{H}$ is unknown at the MSC and, as the $\text{SHDD}_{1,sub}$ scheme, the local detector thresholds $\mathbf{t}$ are optimized with respect to the average probability of bit error $P_b(\mathbf{t})$. Similarly to the $\text{SHDD}_{1,sub}$ scheme, the average probability of bit error after the threshold optimization can be evaluated as follows

$$P_b^* = \min_{\mathbf{t}} \Big\{ P_b(\mathbf{t}) \Big\}. \tag{B.12}$$

However, as opposed to the SHDD$_{1,sub}$ scheme, as shown in (2.37) the probability of bit error simplifies as follows

$$P_b(\mathbf{t}) = \frac{1}{2} + \frac{1}{2} \sum_{\mathbf{u}} P(U_0 = 1 \mid \mathbf{U} = \mathbf{u}) \Big[P(\mathbf{U} = \mathbf{u} \mid B = 0) \big] - P(\mathbf{U} = \mathbf{u} \mid B = 1) \Big]. \tag{B.13}$$

since the fusion rule is independent of the channel state vector $\mathbf{H}$. Using results from [33], it is possible to reformulate the probability of bit error as follows

$$P_b(\mathbf{t}) = \frac{1}{2} - \frac{1}{4} \sum_{\mathbf{u}} \Big| P(\mathbf{U} = \mathbf{u} \mid B = 0) - P(\mathbf{U} = \mathbf{u} \mid B = 1) \Big|, \tag{B.14}$$

where, since the local decisions are conditionally independent and the fading is spatially uncorrelated,

$$P(\mathbf{U} = \mathbf{u} \mid B) = \prod_{k=1}^{N_{BS}} P(U_k = u_k \mid B). \tag{B.15}$$

Using results from Appendix A,

$$\begin{aligned} P(U_k = u_k \mid B) &= \int_{t_{k,u_k}}^{t_{k,u_k+1}} f_{\Psi_{1,0}^{(k)}(\mathbf{R}_k, \mathbf{H}_k)}(x \mid B) dx \\ &= F_{\Psi_{1,0}^{(k)}(\mathbf{R}_k, \mathbf{H}_k)}(t_{k,u_k+1} \mid B) - F_{\Psi_{1,0}^{(k)}(\mathbf{R}_k, \mathbf{H}_k)}(t_{k,u_k} \mid B) \end{aligned} \tag{B.16}$$

such that, as opposed of the SHDD$_{1,sub}$ scheme, the average probability of bit error $P_b(\mathbf{t})$, which is the threshold optimization cost function, has a closed form expression. It has for effect to greatly accelerate the optimization process.

Appendix C

Evaluation of the cost functions of the CSHDD schemes

In Chapter 3, we proposed sub-optimum CSHDD schemes for which the local detectors are LLR quantizers. The thresholds defining these LLR quantizers are optimized to minimize either the PEP associated with the free Hamming distance d_f or the MSE between the LLRs used at the fusion center by the OC and CSHDD schemes. In this appendix, we discuss the evaluation of the PEP and MSE.

C.1 MPEP-LLR quantizer

In section 3.3.1, we showed that the conditional PEP and consequently the average PEP can be expressed as a function of the Hamming distance d between two codewords. The average PEP can thus be defined as

$$P_2(d, \mathbf{t}) = \int_{\mathbf{h}} P_{2|\mathbf{h}}(d, \mathbf{t}) f_{\mathbf{H}}(\mathbf{h}) d\mathbf{h}, \tag{C.1}$$

where $P_{2|\mathbf{h}}(d, \mathbf{t})$ represents the conditional PEP associated with the thresholds $\mathbf{t}$ and the Hamming distance d between two codewords. The conditional PEP $P_{2|\mathbf{h}}(d, \mathbf{t})$ is defined in expression (3.62) as follows

$$\begin{aligned} P_{2|\mathbf{h}}(d,\mathbf{t}) &= \frac{1}{2}+\frac{1}{2}\sum_{\mathbf{n}_1}\cdots\sum_{\mathbf{n}_{N_{BS}}} M(\mathbf{n}_1)\cdots M(\mathbf{n}_{N_{BS}})1(\mathbf{n}_1,\ldots,\mathbf{n}_{N_{BS}})\times \\ & \left[\prod_{k=1}^{N_{BS}}\prod_{l=0}^{L-1}(P_k^{l0}(\mathbf{h}_k))^{n_{k,l}} - \prod_{k=1}^{N_{BS}}\prod_{l=0}^{L-1}(P_k^{l1}(\mathbf{h}_k))^{n_{k,l}}\right], \end{aligned} \tag{C.2}$$

given the indicator function $1(\mathbf{n}_1,\ldots,\mathbf{n}_{N_{BS}})$ is defined in (3.63), for the case when the channel state is known at the MSC, and in (3.64), for the case when the channel state is unknown at the MSC. In order to evaluate (C.2), it is necessary to specify $P_k^{lj}(\mathbf{h}_k)$ which equals

$$\begin{aligned} P_k^{lj}(\mathbf{h}_k) &= \int_{t_{k,l}}^{t_{k,l+1}} f_{\Psi_{1,0}^{(k,t)}(\mathbf{R}_{k,t},\mathbf{H}_k)}(x \mid \mathbf{h}_k, C_t=j)\,dx \\ &= F_{\Psi_{1,0}^{(k,t)}(\mathbf{R}_{k,t},\mathbf{H}_k)}(t_{k,l+1} \mid \mathbf{h}_k, C_t=j) - F_{\Psi_{1,0}^{(k,t)}(\mathbf{R}_{k,t},\mathbf{H}_k)}(t_{k,l} \mid \mathbf{h}_k, C_t=j), \end{aligned} \tag{C.3}$$

where the CDF $F_{\Psi_{1,0}^{(k,t)}(\mathbf{R}_{k,t},\mathbf{H}_k)}(x \mid \mathbf{H}_k, C_t)$ is derived in Appendix A.

C.1.1 Known channel state information at the MSC

For the case when the channel state is known at the MSC, the indicator function $1(\mathbf{n}_1,\ldots,\mathbf{n}_{N_{BS}})$ is a function of $\mathbf{h}$ such that the integral in (C.1) needs to be performed numerically. As shown in Appendix A, the CDF $F_{\Psi_{1,0}^{(k,t)}(\mathbf{R}_{k,t},\mathbf{H}_k)}(x \mid \mathbf{H}_k, C_t)$ can be expressed as a function of $\Omega_k = \sqrt{\sum_{n=1}^{N_R}|H_{k,n}|^2}$, such that the average PEP can be reformulated as follows

$$P_2(d,\mathbf{t}) = \int_{\omega} P_{2|\omega}(d,\mathbf{t}) f_{\mathbf{\Omega}}(\omega)d\omega, \tag{C.4}$$

where $\mathbf{\Omega} = [\Omega_1,\ldots,\Omega_{N_{BS}}]^T$, $f_{\mathbf{\Omega}}(\omega) = f_{\Omega_1}(\omega_1)\cdots f_{\Omega_{N_{BS}}}(\omega_{N_{BS}})$ and the PDF $f_{\Omega_k}(\omega_k)$ is defined in (B.8). This simplifies the evaluation of the average PEP since it reduces the order of the numerical integration from $2N_{BS}N_R$ to N_{BS}. However, the evaluation of the average PEP still requires a relatively large amount of time, making the optimization of the thresholds time consuming. In fact, the time required to perform the numerical integration increases exponentially with the number of resolution bits, Hamming distance d and even more dramatically with the number of base stations since it determines the order of integration.

C.1.2 Unknown channel state information at the MSC

For the case when the channel state is unknown at the MSC, the average PEP can be evaluated as for the previous case although, in this case, in order to determine the value of the indicator function $1(\mathbf{n}_1, \ldots, \mathbf{n}_{N_{BS}})$ it is necessary to evaluate $E_{\mathbf{H}_k}[P_k^{lj}(\mathbf{H}_k)]$ which equals

$$\begin{aligned} E_{\mathbf{H}_k}[P_k^{lj}(\mathbf{H}_k)] &= \int_{\mathbf{h}_k} \int_{t_{k,l}}^{t_{k,l+1}} f_{\Psi_{1,0}^{(k,t)}(\mathbf{R}_{k,t},\mathbf{H}_k)}(x \mid \mathbf{h}_k, C_t = j)\, f_{\mathbf{H}_k}(\mathbf{h}_k) dx d\mathbf{h}_k \\ &= \int_{t_{k,l}}^{t_{k,l+1}} f_{\Psi_{1,0}^{(k,t)}(\mathbf{R}_{k,t},\mathbf{H}_k)}(x \mid C_t = j)\, dx \\ &= F_{\Psi_{1,0}^{(k,t)}(\mathbf{R}_{k,t},\mathbf{H}_k)}(t_{k,l+1} \mid C_t = j) - F_{\Psi_{1,0}^{(k,t)}(\mathbf{R}_{k,t},\mathbf{H}_k)}(t_{k,l} \mid C_t = j)\,, \end{aligned} \tag{C.5}$$

where the CDF $F_{\Psi_{1,0}^{(k,t)}(\mathbf{R}_{k,t},\mathbf{H}_k)}(x \mid C_t)$ is also derived in Appendix A. On the other hand, since the indicator function $1(\mathbf{n}_1, \ldots, \mathbf{n}_{N_{BS}})$ is in this case independent of $\mathbf{H}$, the average PEP (C.1) can also be simplified as follows

$$\begin{aligned} P_2(d, \mathbf{t}) &= \frac{1}{2} + \frac{1}{2} \sum_{\mathbf{n}_1} \cdots \sum_{\mathbf{n}_{N_{BS}}} M(\mathbf{n}_1) \cdots M(\mathbf{n}_{N_{BS}}) 1(\mathbf{n}_1, \ldots, \mathbf{n}_{N_{BS}}) \times \\ & \left[\prod_{k=1}^{N_{BS}} E_{\mathbf{H}_k} \left[\prod_{l=0}^{L-1} (P_k^{l0}(\mathbf{H}_k))^{n_{k,l}} \right] - \prod_{k=1}^{N_{BS}} E_{\mathbf{H}_k} \left[\prod_{l=0}^{L-1} (P_k^{l1}(\mathbf{H}_k))^{n_{k,l}} \right] \right], \end{aligned} \tag{C.6}$$

where

$$E_{\mathbf{H}_k} \left[\prod_{l=0}^{L-1} (P_k^{lj}(\mathbf{H}_k))^{n_{k,l}} \right] = \int_{\mathbf{h}_k} \prod_{l=0}^{L-1} (P_k^{lj}(\mathbf{h}_k))^{n_{k,l}} f_{\mathbf{H}_k}(\mathbf{h}_k) d\mathbf{h}_k = \int_{\omega_k} \prod_{l=0}^{L-1} (P_k^{lj}(\omega_k))^{n_{k,l}} f_{\Omega_k}(\omega_k) d\omega_k \tag{C.7}$$

and the integration needs to be performed numerically. Hence, as opposed to the other alternative, it is necessary to evaluate numerically multiple one-dimensional integrals as opposed to a single N_{BS}-dimensional integral.

C.2 MMSE-LLR quantizer

C.2.1 Known channel state information at the MSC

The MSE between the LLR $\Psi_{1,0}^{(k,t)}(\mathbf{R}_{k,t}, \mathbf{H}_k)$ used at the fusion center by the OC scheme and the LLR $\Psi_{1,0}^{(k,t)}(U_{k,t}, \mathbf{H}_k)$ used at the fusion center by the CSHDD scheme, when the channel state is known at the MSC, has been defined in expression (3.70). Since the local decision $U_{k,t}$ is determined by the MMSE-LLR quantizer that quantizes the LLR $\Psi_{1,0}^{(k,t)}(\mathbf{R}_{k,t}, \mathbf{H}_k)$, $U_{k,t}$ can be expressed as a function of $\Psi_{1,0}^{(k,t)}(\mathbf{R}_{k,t}, \mathbf{H}_k)$ such that $U_{k,t} = f_k\left(\Psi_{1,0}^{(k,t)}(\mathbf{R}_{k,t}, \mathbf{H}_k)\right)$. The MSE can thus be reformulated as follows

$$\varepsilon_{k,t} = E_{\mathbf{R}_{k,t},\mathbf{H}_k}\left[\left(\Psi_{1,0}^{(k,t)}(\mathbf{R}_{k,t}, \mathbf{H}_k) - \Psi_{1,0}^{(k,t)}\left(f_k\left(\Psi_{1,0}^{(k,t)}(\mathbf{R}_{k,t}, \mathbf{H}_k)\right), \mathbf{H}_k\right)\right)^2\right] \tag{C.8}$$

In order to evaluate the MSE $\varepsilon_{k,t}$, it is advantageous to decompose the expectation as follows

$$\begin{aligned}\varepsilon_{k,t} = E_{\mathbf{H}_k,C_t}\Bigg[& \underbrace{E_{\Psi_{1,0}^{(k,t)}(\mathbf{R}_{k,t},\mathbf{H}_k)|\mathbf{H}_k,C_t}\left[\left(\Psi_{1,0}^{(k,t)}(\mathbf{R}_{k,t}, \mathbf{H}_k)\right)^2\right]}_{I_1 C_t,\mathbf{H}_k)} \\ & + \underbrace{E_{\Psi_{1,0}^{(k,t)}(\mathbf{R}_{k,t},\mathbf{H}_k)|\mathbf{H}_k,C_t}\left[\left(\Psi_{1,0}^{(k,t)}\left(f_k\left(\Psi_{1,0}^{(k,t)}(\mathbf{R}_{k,t}, \mathbf{H}_k)\right), \mathbf{H}_k\right)\right)^2\right]}_{I_2 C_t,\mathbf{H}_k)} \\ & -2\underbrace{E_{\Psi_{1,0}^{(k,t)}(\mathbf{R}_{k,t},\mathbf{H}_k)|\mathbf{H}_k,C_t}\left[\Psi_{1,0}^{(k,t)}(\mathbf{R}_{k,t}, \mathbf{H}_k)\Psi_{1,0}^{(k,t)}\left(f_k\left(\Psi_{1,0}^{(k,t)}(\mathbf{R}_{k,t}, \mathbf{H}_k)\right), \mathbf{H}_k\right)\right]}_{I_3 C_t,\mathbf{H}_k)}\Bigg],\end{aligned} \tag{C.9}$$

where closed form expressions can be derived for $I_1(C_t, \mathbf{H}_k)$, $I_2(C_t, \mathbf{H}_k)$ and $I_3(C_t, \mathbf{H}_k)$.

First, $I_1(C_t, \mathbf{H}_k)$ represents the second moment of $\Psi_{1,0}^{(k,t)}(\mathbf{R}_{k,t}, \mathbf{H}_k)$, given the transmitted coded bit C_t and the channel state vector $\mathbf{H}_k$ are known, and can be expressed as a function of the mean $\mu(\mathbf{H}_k, C_t)$ and variance $\sigma^2(\mathbf{H}_k)$, which are derived in Appendix A, as follows

$$I_1(C_t, \mathbf{H}_k) = \sigma^2(\mathbf{H}_k) + (\mu(\mathbf{H}_k, C_t))^2 = \frac{8E_k}{N_0}\sum_{n=1}^{N_R}|H_{k,n}|^2 + \left(\frac{4E_k}{N_0}\sum_{n=1}^{N_R}|H_{k,n}|^2 10^{C_t+1}\right)^2. \tag{C.10}$$

Then, since $\Psi_{1,0}^{(k,t)}\left(f_k\left(\Psi_{1,0}^{(k,t)}(\mathbf{r}_{k,t}, \mathbf{h}_k)\right), \mathbf{h}_k\right) = \Psi_{1,0}^{(k,t)}(u_{k,t}, \mathbf{h}_k)$ if $t_{k,u_{k,t}} < \Psi_{1,0}^{(k,t)}(\mathbf{r}_{k,t}, \mathbf{h}_k) <$

$t_{k,u_{k,t}+1}$,

$$\begin{aligned}
I_2(C_t,\mathbf{H}_k) &= \int_{-\infty}^{\infty}\left(\Psi_{1,0}^{(k,t)}\left(f_k\left(x\right),\mathbf{H}_k\right)\right)^2 f_{\Psi_{1,0}^{(k,t)}(\mathbf{R}_{k,t},\mathbf{H}_k)}(x|\mathbf{H}_k,C_t)dx \\
&= \sum_{u_{k,t}}\left(\Psi_{1,0}^{(k,t)}(u_{k,t},\mathbf{H}_k)\right)^2\int_{t_{k,u_{k,t}}}^{t_{k,u_{k,t}+1}} f_{\Psi_{1,0}^{(k,t)}(\mathbf{R}_{k,t},\mathbf{H}_k)}(x|\mathbf{H}_k,C_t)dx \\
&= \sum_{u_{k,t}}\left(\Psi_{1,0}^{(k,t)}(u_{k,t},\mathbf{H}_k)\right)^2\times \\
&\quad \left(F_{\Psi_{1,0}^{(k,t)}(\mathbf{R}_{k,t},\mathbf{H}_k)}(t_{k,u_{k,t}+1}|\mathbf{H}_k,C_t)-F_{\Psi_{1,0}^{(k,t)}(\mathbf{R}_{k,t},\mathbf{H}_k)}(t_{k,u_{k,t}}|\mathbf{H}_k,C_t)\right),
\end{aligned} \tag{C.11}$$

where the CDF $F_{\Psi_{1,0}^{(k,t)}(\mathbf{R}_{k,t},\mathbf{H}_k)}(t_{k,l+1}|\mathbf{H}_k,C_t)$ is derived in Appendix A. Finally,

$$\begin{aligned}
&I_3(C_t,\mathbf{H}_k) \\
&= \int_{-\infty}^{\infty}\Psi_{1,0}^{(k,t)}\left(f_k\left(x\right),\mathbf{H}_k\right)xf_{\Psi_{1,0}^{(k,t)}(\mathbf{R}_{k,t},\mathbf{H}_k)}(x|\mathbf{H}_k,C_t)dx \\
&= \sum_{u_{k,t}}\Psi_{1,0}^{(k,t)}(u_{k,t},\mathbf{H}_k)\int_{t_{k,u_{k,t}}}^{t_{k,u_{k,t}+1}} xf_{\Psi_{1,0}^{(k,t)}(\mathbf{R}_{k,t},\mathbf{H}_k)}(x|\mathbf{H}_k,C_t)dx \\
&= \sum_{u_{k,t}}\Psi_{1,0}^{(k,t)}(u_{k,t},\mathbf{H}_k)\int_{t_{k,u_{k,t}}}^{t_{k,u_{k,t}+1}}\frac{x}{\sqrt{2\pi\sigma^2(\mathbf{H}_k)}}\exp\left\{-\frac{(x-\mu(\mathbf{H}_k,C_t))^2}{2\sigma^2(\mathbf{H}_k)}\right\}dx \\
&= \sum_{u_{k,t}}\Psi_{1,0}^{(k,t)}(u_{k,t},\mathbf{H}_k)\int_{\frac{t_{k,u_{k,t}}-\mu(\mathbf{H}_k,C_t)}{\sigma(\mathbf{H}_k)}}^{\frac{t_{k,u_{k,t}+1}-\mu(\mathbf{H}_k,C_t)}{\sigma(\mathbf{H}_k)}}\frac{v\sigma(\mathbf{H}_k)+\mu(\mathbf{H}_k,C_t)}{\sqrt{2\pi}}\exp\left\{-\frac{v^2}{2}\right\}dv \\
&= \sum_{u_{k,t}}\Psi_{1,0}^{(k,t)}(u_{k,t},\mathbf{H}_k)\Bigg(-\frac{\sigma(\mathbf{H}_k)}{\sqrt{2\pi}}\Bigg[\exp\left\{-\frac{(t_{k,u_{k,t}+1}-\mu(\mathbf{H}_k,C_t))^2}{2\sigma^2(\mathbf{H}_k)}\right\} \\
&\qquad -\exp\left\{-\frac{(t_{k,u_{k,t}}-\mu(\mathbf{H}_k,C_t))^2}{2\sigma^2(\mathbf{H}_k)}\right\}\Bigg] \\
&\quad +\mu(\mathbf{H}_k,C_t)\left(F_{\Psi_{1,0}^{(k,t)}(\mathbf{R}_{k,t},\mathbf{H}_k)}(t_{k,u_{k,t}+1}|\mathbf{H}_k,C_t)-F_{\Psi_{1,0}^{(k,t)}(\mathbf{R}_{k,t},\mathbf{H}_k)}(t_{k,u_{k,t}}|\mathbf{H}_k,C_t)\right)\Bigg).
\end{aligned} \tag{C.12}$$

However, even if we found closed form expressions for $I_1(C_t,\mathbf{H}_k)$, $I_2(C_t,\mathbf{H}_k)$ and $I_3(C_t,\mathbf{H}_k)$, it is not possible to derive a closed form expression for $\varepsilon_{k,t}$. which can be formulated as a function of $I_1(C_t, \mathbf{H}_k)$, $I_2(C_t, \mathbf{H}_k)$ and $I_3(C_t, \mathbf{H}_k)$ as follows

$$\varepsilon_{k,t} = \sum_{c_t} \int_{\mathbf{h}_k} (I_1(c_t, \mathbf{h}_k) + I_2(c_t, \mathbf{h}_k) - 2I_3(c_t, \mathbf{h}_k)) f_{\mathbf{H}_k}(\mathbf{h}_k) d\mathbf{h}_k P(C_t = c_t). \quad \text{(C.13)}$$

Hence, since $I_1(C_t, \mathbf{H}_k)$, $I_2(C_t, \mathbf{H}_k)$ and $I_3(C_t, \mathbf{H}_k)$ can all be expressed as function of $\Omega_k = \sqrt{\sum_{n=1}^{N_R} |H_{k,n}|^2}$, it is advantageous to reformulate $\varepsilon_{k,t}$ as follows

$$\varepsilon_{k,t} = \sum_{c_t} \int_{\omega_k} (I_1(c_t, \omega_k) + I_2(c_t, \omega_k) - 2I_3(c_t, \omega_k)) f_{\Omega_k}(\omega_k) d\omega_k P(C_t = c_t), \quad \text{(C.14)}$$

reducing the number of dimensions of the required numerical integral from $2N_R$ to 1. The PDF $f_{\Omega_k}(\omega_k)$ is presented in (B.8).

C.2.2 Unknown channel state information at the MSC

The MSE between the LLR $\Psi_{1,0}^{(k,t)}(\mathbf{R}_{k,t}, \mathbf{H}_k)$ used at the fusion center by the OC scheme and the LLR $\Psi_{1,0}^{(k,t)}(U_{k,t})$ used at the fusion center by the CSHDD scheme, when the channel state is unknown at the MSC, has been defined in expression (3.71). As in the previous case, the local decision $U_{k,t}$ is determined by the MSE-LLR quantizer that quantizes the LLR $\Psi_{1,0}^{(k,t)}(\mathbf{R}_{k,t}, \mathbf{H}_k)$. The local decision $U_{k,t}$ can thus be expressed as a function of $\Psi_{1,0}^{(k,t)}(\mathbf{R}_{k,t}, \mathbf{H}_k)$ such that $U_{k,t} = f_k\left(\Psi_{1,0}^{(k,t)}(\mathbf{R}_{k,t}, \mathbf{H}_k)\right)$ and the MSE can be reformulated as follows

$$\begin{aligned}
\varepsilon_{k,t} &= E_{\Psi_{1,0}^{(k,t)}(\mathbf{R}_{k,t},\mathbf{H}_k)}\left[\left(\Psi_{1,0}^{(k,t)}(\mathbf{R}_{k,t}, \mathbf{H}_k) - \Psi_{1,0}^{(k,t)}\left(f_k(\Psi_{1,0}^{(k,t)}(\mathbf{R}_{k,t}, \mathbf{H}_k))\right)\right)^2\right] \\
&= \underbrace{E_{\Psi_{1,0}^{(k,t)}(\mathbf{R}_{k,t},\mathbf{H}_k)}\left[\left(\Psi_{1,0}^{(k,t)}(\mathbf{R}_{k,t}, \mathbf{H}_k)\right)^2\right]}_{I_1} \\
&\quad + \underbrace{E_{\Psi_{1,0}^{(k,t)}(\mathbf{R}_{k,t},\mathbf{H}_k)}\left[\left(\Psi_{1,0}^{(k,t)}\left(f_k(\Psi_{1,0}^{(k,t)}(\mathbf{R}_{k,t}, \mathbf{H}_k))\right)\right)^2\right]}_{I_2} \\
&\quad \underbrace{-2\,E_{\Psi_{1,0}^{(k,t)}(\mathbf{R}_{k,t},\mathbf{H}_k)}\left[\Psi_{1,0}^{(k,t)}(\mathbf{R}_{k,t}, \mathbf{H}_k)\Psi_{1,0}^{(k,t)}\left(f_k(\Psi_{1,0}^{(k,t)}(\mathbf{R}_{k,t}, \mathbf{H}_k))\right)\right]}_{I_3}. \quad \text{(C.15)}
\end{aligned}$$

It is important to mention that closed form expressions can be derived for I_1, I_2 and I_3 such that, as opposed to the previous case, a closed form expression can be obtained for $\varepsilon_{k,t}$. First,

$$\begin{aligned} I_1 &= \sum_{c_t} \int_{-\infty}^{\infty} x^2 f_{\Psi_{1,0}^{(k,t)}(\mathbf{R}_{k,t},\mathbf{H}_k)}(x|C_t = c_t) P(C_t = c_t) dx \\ &= \sum_{c_t} \int_{-\infty}^{\infty} \frac{x^2}{2} f_{\Psi_{1,0}^{(k,t)}(\mathbf{R}_{k,t},\mathbf{H}_k)}(x|C_t = c_t) dx, \end{aligned} \tag{C.16}$$

where $f_{\Psi_{1,0}^{(k,t)}(\mathbf{R}_{k,t},\mathbf{H}_k)}(x|C_t)$ has been derived in Appendix A. Hence, given the function $H_a^{(n)}(x)$ defined in (A.26), if $N_R = 1$

$$\begin{aligned} I_1 &= \sum_{c_t} \left(\int_0^{\infty} \frac{x^2}{2} A_k \exp\left\{B_k^+(c_t)x\right\} dx + \int_{-\infty}^0 \frac{x^2}{2} A_k \exp\left\{B_k^-(c_t)x\right\} \right) dx \\ &= \sum_{c_t} \frac{A_k}{2} \left(-H_{B_k^+(c_t)}^{(2)}(0) + H_{B_k^-(c_t)}^{(2)}(0) \right), \end{aligned} \tag{C.17}$$

if $N_R = 2$,

$$\begin{aligned} I_1 &= \sum_{c_t} \left(\int_0^{\infty} \frac{x^2}{2} A_k^2 \left(\frac{2}{\beta_k} + x \right) \exp\left\{B_k^+(c_t)x\right\} dx \right. \\ &\qquad \left. + \int_{-\infty}^0 \frac{x^2}{2} A_k^2 \left(\frac{2}{\beta_k} - x \right) \exp\left\{B_k^-(c_t)x\right\} dx \right) \\ &= \sum_{c_t} \frac{A_k^2}{2} \left(\frac{-2}{\beta_k} H_{B_k^+(c_t)}^{(2)}(0) - H_{B_k^+(c_t)}^{(3)}(0) + \frac{2}{\beta_k} H_{B_k^-(c_t)}^{(2)}(0) - H_{B_k^-(c_t)}^{(3)}(0) \right) \end{aligned} \tag{C.18}$$

and, if $N_R = 3$,

$$\begin{aligned} I_1 &= \sum_{c_t} \left(\int_0^{\infty} \frac{x^2}{2} A_k^3 \left(\frac{6}{\beta_k^2} + \frac{3x}{\beta_k} + \frac{x^2}{2} \right) \exp\left\{B_k^+(c_t)x\right\} dx \right. \\ &\qquad \left. \int_{-\infty}^0 \frac{x^2}{2} A_k^3 \left(\frac{6}{\beta_k^2} - \frac{3x}{\beta_k} + \frac{x^2}{2} \right) \exp\left\{B_k^-(c_t)x\right\} dx \right) \\ &= \sum_{c_t} \frac{A_k^3}{2} \left(\frac{-6}{\beta_k^2} H_{B_k^+(c_t)}^{(2)}(0) - \frac{3}{\beta_k} H_{B_k^+(c_t)}^{(3)}(0) - \frac{1}{2} H_{B_k^+(c_t)}^{(4)}(0) \right. \\ &\qquad \left. + \frac{6}{\beta_k^2} H_{B_k^-(c_t)}^{(2)}(0) - \frac{3}{\beta_k} H_{B_k^-(c_t)}^{(3)}(0) + \frac{1}{2} H_{B_k^-(c_t)}^{(4)}(0) \right), \end{aligned} \tag{C.19}$$

where in these expressions it is emphasized that the parameters B_k^+ and B_k^- defined in

Appendix A are functions of C_t. Then, since $\Psi_{1,0}^{(k,t)}\left(f_k\left(\Psi_{1,0}^{(k,t)}(\mathbf{r}_{k,t},\mathbf{h}_k)\right)\right) = \Psi_{1,0}^{(k,t)}(u_{k,t})$ when $t_{k,u_{k,t}} < \Psi_{1,0}^{(k,t)}(\mathbf{r}_{k,t},\mathbf{h}_k) < t_{k,u_{k,t}+1}$,

$$
\begin{aligned}
I_2 &= \sum_{c_t}\int_{-\infty}^{\infty}\left(\Psi_{1,0}^{(k,t)}(f_k(x))\right)^2 f_{\Psi_{1,0}^{(k,t)}(\mathbf{R}_{k,t},\mathbf{H}_k)}(x|C_t=c_t)P(C_t=c_t)dx \\
&= \sum_{c_t}\sum_{u_{k,t}}\frac{\left(\Psi_{1,0}^{(k,t)}(u_{k,t})\right)^2}{2}\int_{t_{k,u_{k,t}}}^{t_{k,u_{k,t}+1}} f_{\Psi_{1,0}^{(k,t)}(\mathbf{R}_{k,t},\mathbf{H}_k)}(x|C_t=c_t)dx \\
&= \sum_{c_t}\sum_{u_{k,t}}\frac{\left(\Psi_{1,0}^{(k,t)}(u_{k,t})\right)^2}{2}\times \\
&\quad \left(F_{\Psi_{1,0}^{(k,t)}(\mathbf{R}_{k,t},\mathbf{H}_k)}(t_{k,u_{k,t}+1}|C_t=c_t) - F_{\Psi_{1,0}^{(k,t)}(\mathbf{R}_{k,t},\mathbf{H}_k)}(t_{k,u_{k,t}}|C_t=c_t)\right),
\end{aligned}
\tag{C.20}
$$

where the CDF $F_{\Psi_{1,0}^{(k,t)}(\mathbf{R}_{k,t},\mathbf{H}_k)}(x|C_t)$ is derived in Appendix A. Finally,

$$
\begin{aligned}
I_3 &= \sum_{c_t}\int_{-\infty}^{\infty}\Psi_{1,0}^{(k,t)}(f_k(x))x f_{\Psi_{1,0}^{(k,t)}(\mathbf{R}_{k,t},\mathbf{H}_k)}(x|C_t=c_t)P(C_t=c_t)dx \\
&= \sum_{c_t}\sum_{u_{k,t}}\frac{\Psi_{1,0}^{(k,t)}(u_{k,t})}{2}\int_{t_{k,u_{k,t}}}^{t_{k,u_{k,t}+1}} x f_{\Psi_{1,0}^{(k,t)}(\mathbf{R}_{k,t},\mathbf{H}_k)}(x|C_t=c_t)dx
\end{aligned}
\tag{C.21}
$$

such that, if $N_R = 1$,

$$
\begin{aligned}
I_3 &= \sum_{c_t}\left(\sum_{u_{k,t}=L/2}^{L-1}\frac{\Psi_{1,0}^{(k,t)}(u_{k,t})}{2}\int_{t_{k,u_{k,t}}}^{t_{k,u_{k,t}+1}} xA_k\exp\left\{B_k^+(c_t)x\right\}dx\right. \\
&\qquad \left.+\sum_{u_{k,t}=0}^{(L/2)-1}\frac{\Psi_{1,0}^{(k,t)}(u_{k,t})}{2}\int_{t_{k,u_{k,t}}}^{t_{k,u_{k,t}+1}} xA_k\exp\left\{B_k^-(c_t)x\right\}dx\right) \\
&= \sum_{c_t}\frac{A_k}{2}\left(\sum_{u_{k,t}=L/2}^{L-1}\Psi_{1,0}^{(k,t)}(u_{k,t})\left(H^{(1)}_{B_k^+(c_t)}(t_{k,u_{k,t}+1}) - H^{(1)}_{B_k^+(c_t)}(t_{k,u_{k,t}})\right)\right. \\
&\qquad \left.+\sum_{u_{k,t}=0}^{(L/2)-1}\Psi_{1,0}^{(k,t)}(u_{k,t})\left(H^{(1)}_{B_k^-(c_t)}(t_{k,u_{k,t}+1}) - H^{(1)}_{B_k^-(c_t)}(t_{k,u_{k,t}})\right)\right),
\end{aligned}
\tag{C.22}
$$

if $N_R = 2$,

$$\begin{aligned}
I_3 &= \sum_{c_t} \left(\sum_{u_{k,t}=L/2}^{L-1} \frac{\Psi_{1,0}^{(k,t)}(u_{k,t})}{2} \int_{t_{k,u_{k,t}}}^{t_{k,u_{k,t}+1}} x A_k^2 \left(\frac{2}{\beta_k} + x \right) \exp\left\{ B_k^+(c_t) x \right\} dx \right. \\
&\qquad \left. + \sum_{u_{k,t}=0}^{(L/2)-1} \frac{\Psi_{1,0}^{(k,t)}(u_{k,t})}{2} \int_{t_{k,u_{k,t}}}^{t_{k,u_{k,t}+1}} x A_k^2 \left(\frac{2}{\beta_k} - x \right) \exp\left\{ B_k^-(c_t) x \right\} dx \right) \\
&= \sum_{c_t} \frac{A_k^2}{2} \Bigg(\sum_{u_{k,t}=L/2}^{L-1} \Psi_{1,0}^{(k,t)}(u_{k,t}) \left[\frac{2}{\beta_k} \left(H^{(1)}_{B_k^+(c_t)}(t_{k,u_{k,t}+1}) - H^{(1)}_{B_k^+(c_t)}(t_{k,u_{k,t}}) \right) \right. \\
&\qquad \left. + \left(H^{(2)}_{B_k^+(c_t)}(t_{k,u_{k,t}+1}) - H^{(2)}_{B_k^+(c_t)}(t_{k,u_{k,t}}) \right) \right] \\
&\qquad + \sum_{u_{k,t}=0}^{(L/2)-1} \Psi_{1,0}^{(k,t)}(u_{k,t}) \left[\frac{2}{\beta_k} \left(H^{(1)}_{B_k^-(c_t)}(t_{k,u_{k,t}+1}) - H^{(1)}_{B_k^-(c_t)}(t_{k,u_{k,t}}) \right) \right. \\
&\qquad \left. - \left(H^{(2)}_{B_k^-(c_t)}(t_{k,u_{k,t}+1}) - H^{(2)}_{B_k^-(c_t)}(t_{k,u_{k,t}}) \right) \right] \Bigg)
\end{aligned} \tag{C.23}$$

and, if $N_R = 3$,

$$\begin{aligned}
I_3 & \\
&= \sum_{c_t} \left(\sum_{u_{k,t}=L/2}^{L-1} \frac{\Psi_{1,0}^{(k,t)}(u_{k,t})}{2} \int_{t_{k,u_{k,t}}}^{t_{k,u_{k,t}+1}} x A_k^3 \left(\frac{6}{\beta_k^2} + \frac{3x}{\beta_k} + \frac{x^2}{2} \right) \exp\left\{ B_k^+(c_t) x \right\} dx \right. \\
&\qquad \left. + \sum_{u_{k,t}=0}^{(L/2)-1} \frac{\Psi_{1,0}^{(k,t)}(u_{k,t})}{2} \int_{t_{k,u_{k,t}}}^{t_{k,u_{k,t}+1}} x A_k^3 \left(\frac{6}{\beta_k^2} - \frac{3x}{\beta_k} + \frac{x^2}{2} \right) \exp\left\{ B_k^-(c_t) x \right\} dx \right) \\
&= \sum_{c_t} \frac{A_k^3}{2} \Bigg(\sum_{u_{k,t}=L/2}^{L-1} \Psi_{1,0}^{(k,t)}(u_{k,t}) \left[\frac{6}{\beta_k^2} \left(H^{(1)}_{B_k^+(c_t)}(t_{k,u_{k,t}+1}) - H^{(1)}_{B_k^+(c_t)}(t_{k,u_{k,t}}) \right) \right. \\
&\qquad \left. + \frac{3}{\beta_k} \left(H^{(2)}_{B_k^+(c_t)}(t_{k,u_{k,t}+1}) - H^{(2)}_{B_k^+(c_t)}(t_{k,u_{k,t}}) \right) + \frac{1}{2} \left(H^{(3)}_{B_k^+(c_t)}(t_{k,u_{k,t}+1}) - H^{(3)}_{B_k^+(c_t)}(t_{k,u_{k,t}}) \right) \right] \\
&\qquad + \sum_{u_{k,t}=0}^{(L/2)-1} \Psi_{1,0}^{(k,t)}(u_{k,t}) \left[\frac{6}{\beta_k^2} \left(H^{(1)}_{B_k^-(c_t)}(t_{k,u_{k,t}+1}) - H^{(1)}_{B_k^-(c_t)}(t_{k,u_{k,t}}) \right) \right. \\
&\qquad \left. - \frac{3}{\beta_k} \left(H^{(2)}_{B_k^-(c_t)}(t_{k,u_{k,t}+1}) - H^{(2)}_{B_k^-(c_t)}(t_{k,u_{k,t}}) \right) + \frac{1}{2} \left(H^{(3)}_{B_k^-(c_t)}(t_{k,u_{k,t}+1}) - H^{(3)}_{B_k^-(c_t)}(t_{k,u_{k,t}}) \right) \right] \Bigg)
\end{aligned} \tag{C.24}$$

Appendix D

Reference schemes

In this appendix, we present three different handoff macrodiversity reference schemes. We consider the OC scheme, the CHM scheme and a selection diversity scheme that assumes the channel state is known at the MSC. We consider these three schemes for both uncoded and coded cellular communication systems.

D.1 OC scheme

In the OC scheme, the local detectors transmit to the fusion center the received signals as well as all the channel state information. Hence, classical detection theory can be applied to determine the final decision.

For uncoded communication systems, the optimum decision rule at the fusion center is therefore a maximum-likelihood rule which is, considering the assumptions made in this thesis, equivalent to a maximum-ratio-combining receiver. The probability of bit error of such systems is derived in [34] and these results were used to produce the BER curves of the OC scheme presented in section 2.4.

Similarly, for coded communication systems, the decision rule of the OC scheme is a maximum likelihood rule which can be formulated as follows

$$\mathrm{U}_0 = \arg\max_{\mathbf{c}\in\mathcal{C}} \left\{ \prod_{k=1}^{N_{BS}} \prod_{t=1}^{N_c} f_{\mathbf{R}_{k,t}}(\mathbf{r}_{k,t} \mid \mathbf{h}_k, C_t = c_t) \right\} \tag{D.1}$$

given $\mathbf{R}_{k,t} = \mathbf{r}_{k,t}$ and $\mathbf{H}_k = \mathbf{h}_k$, where the PDF $p_{\mathbf{R}_{k,t}}(\mathbf{r}_{k,t} \mid \mathbf{h}_k, C_t = c_t)$ is defined in (A.2). However, it can be proved that the decision rule (D.1), when expressed in a logarithmic form, can be simplified as follows

$$\mathbf{U}_0 = \arg\max_{\mathbf{c}\in\mathcal{C}} \left\{ \sum_{k=1}^{N_{BS}} \sum_{t=1}^{N_c} (-1)^{c_t+1} \Psi_{1,0}^{(k,t)}(\mathbf{r}_{k,t}, \mathbf{h}_k) \right\}, \tag{D.2}$$

where the LLR $\Psi_{1,0}^{(k,t)}(\mathbf{r}_{k,t}, \mathbf{h}_k)$ is defined in (A.4). In fact, this is the decision rule used by the computer simulator used to estimate the performances of the OC scheme.

D.2 CHM and selection diversity schemes

In the CHM scheme as in the selection diversity scheme, each base station makes hard decisions on the transmitted information bits, using locally optimum decision rules that minimize the probability of error on the local decisions. Considering the assumptions made in this thesis, the optimum local decision rules are maximum likelihood rules. Then, at the MSC, the decisions of the base station for which the probability of error on the local decisions is the lowest are selected.

What differentiate both schemes is the amount of information on the channel available at the MSC. For the CHM scheme, it is assumed that only the average SNR received at each base station is known at the MSC. The link with the highest SNR is thus selected. Hence, for fixed values of average SNR, the probability of error for the CHM scheme is equal to the probability of error at the base station with the maximum average SNR. For the selection diversity scheme, it is assumed that the channel state as well as the average SNR at each base station is known at the MSC. Since the probability of error at the kth base station given $\mathbf{H}_k = \mathbf{h}_k$ is a monotonically decreasing function of $\bar{\omega}_k = \frac{E_k}{N_0} \sum_{n=1}^{N_R} |h_{k,n}|^2$, the base station for which $\bar{\omega}_k$ is the highest is selected.

It is important to mention that, for uncoded communication systems, the local decision rules of a SHDD scheme, making hard decisions on the transmitted bits, are locally optimum when the local decisions are forced to be even symmetric, or equivalently when the only threshold defining each local decision rule equals to 0. Furthermore, when the handoff macrodiversity scheme involves two base stations and the two base stations are making hard decisions on the transmitted bits using locally optimum decision rules, the optimum fusion rule selects the local decisions of the base station for which the probability of bit error on the local decisions is the lowest, as the

CHM scheme and the selection diversity scheme. For instance, assuming the channel state is known at the MSC, the optimum decision rule for a 2 base station SHDD scheme equals

$$\frac{P(U_1 = u_1|\mathbf{h}_1, B = 1)P(U_2 = u_2|\mathbf{h}_2, B = 1)}{P(U_1 = u_1|\mathbf{h}_1, B = 0)P(U_2 = u_2|\mathbf{h}_2, B = 0)} \underset{U_0 = 0}{\overset{U_0 = 1}{\gtrless}} 1 \tag{D.3}$$

which is equivalent to

$$\frac{P(U_1 = u_1|\mathbf{h}_1, B = 1)}{P(U_1 = u_1|\mathbf{h}_1, B = 0)} \underset{U_0 = 0}{\overset{U_0 = 1}{\gtrless}} \frac{P(U_2 = u_2|\mathbf{h}_2, B = 0)}{P(U_2 = u_2|\mathbf{h}_2, B = 1)}. \tag{D.4}$$

Then, since locally optimum decision rules are even symmetric, $P(U_k = u_k|\mathbf{h}_k, B = 0) = 1 - P(U_k = u_k|\mathbf{h}_k, B = 1)$ and we have

$$\frac{P(U_1 = u_1|\mathbf{h}_1, B = 1)}{1 - P(U_1 = u_1|\mathbf{h}_1, B = 1)} \underset{U_0 = 0}{\overset{U_0 = 1}{\gtrless}} \frac{1 - P(U_2 = u_2|\mathbf{h}_2, B = 1)}{P(U_2 = u_2|\mathbf{h}_2, B = 1)}$$

$$0 \underset{U_0 = 0}{\overset{U_0 = 1}{\gtrless}} 1 - P(U_1 = u_1|\mathbf{h}_1, B = 1) - P(U_2 = u_2|\mathbf{h}_2, B = 1). \tag{D.5}$$

Using (D.5), the local decision rule can be reformulated as follows

$$P(U_1 = u_1|\mathbf{h}_1, B = 1) \underset{U_0 = 0}{\overset{U_0 = 1}{\gtrless}} P(U_2 = u_2|\mathbf{h}_2, B = 0) \tag{D.6}$$

or as follows

$$P(U_2 = u_2|\mathbf{h}_2, B = 1) \underset{U_0=0}{\overset{U_0=1}{\gtrless}} P(U_1 = u_1|\mathbf{h}_1, B = 0). \tag{D.7}$$

Hence, assuming that $\bar{\omega}_1 > \bar{\omega}_2$, it is clear from (D.6) that if $U_1 = 1$ the final decision U_0 equals 1 independently of U_2 while it is clear from (D.7) that if $U_1 = 0$ the final decision U_0 equals 0 independently of U_2. Similar results can also be obtained for the case when the channel state is unknown at the MSC, proving that the CHM scheme and the selection diversity scheme involving two base stations are equivalent to SHDD schemes for which $L = 2$ and the only threshold defining each decision rule is equal to 0. Furthermore, it can be concluded that the selection diversity scheme probability of bit error is an upper bound to the probability of bit error of the SHDD scheme for the case when the channel state is known at the MSC while the probability of bit error of the CHM scheme is an upper bound to both cases.

References

[1] V. K. Garg, *Wireless Network Evolution: 2G to 3G.* Prentice Hall PTR, 2001.

[2] T. S. Rappaport, *Wireless Communications: Principles and Practice.* Prentice Hall PTR, 1999.

[3] P. K. Varshney, *Distributed Detection and Data Fusion.* Springer-Verlag, 1997.

[4] A. D. Kot and C. Leung, "Optimal partial decision combining in diversity systems," *IEEE Transactions on Communications*, vol. 38, pp. 981–991, July 1990.

[5] Y. A. Chau and J.-T. Sun, "Diversity with distributed decisions combining for direct-sequence CDMA in a shadowed Rician-fading land-mobile satellite channel," *IEEE Transactions on Vehicular Technology*, vol. 45, pp. 237–247, May 1996.

[6] Y. A. Chau and A.-S. Laih, "Diversity for noncoherent DPSK direct-sequence CDMA over a shadowed Rician-fading satellite channel," in *IEE proceedings in Communications*, vol. 142, pp. 386–392, December 1995.

[7] J. G. Chen, *Adaptive distributed detection with applications to cellular CDMA.* PhD thesis, New Jersey Institute of Technology, January 1997.

[8] J. G. Chen, N. Ansari, and Z. Siveski, "Improving multiuser detection performance by data fusion," in *Proceedings of Globecom'96*, vol. 3, pp. 1794–1798, November 1996.

[9] J. G. Chen, N. Ansari, and Z. Siveski, "Distributed detection for cellular CDMA," *Electronics Letters*, vol. 32, pp. 169–171, February 1996.

[10] J. G. Chen, N. Ansari, and Z. Siveski, "Performance analysis and realization of decision fusion for macroscopic diversity in cellular wireless systems," in *Proceeding of Globecom'97*, vol. 1, pp. 424–428, November 1997.

[11] R. S. Blum, "Distributed detection for diversity reception of fading signals in noise," *IEEE Transactions on Information Theory*, vol. 45, pp. 158–164, January 1999.

[12] R. Blum and J. Hu, "Distributed multiuser detection," in *Proceedings of the thirty-third Asilomar Conference on signals, systems and computers*, vol. 2, pp. 1567–1571, October 1999.

[13] J. Hu and R. S. Blum, "Application of distributed signal detection to multiuser communication systems," *IEEE Transactions on Aerospace and Electronic Systems*, vol. 38, pp. 1220–1229, October 2002.

[14] Q. Zhang and P. K. Varshney, "Diversity signal reception via soft decision combining," in *Proceedings of IEEE International Conference on Acoustics, Speech and Signal processing*, vol. 5, pp. 2937–2940, May 2001.

[15] Q. Zhang, W. Ye, and P. K. Varshney, "Multiuser detection with cell diversity for DS/CDMA systems," in *Proceedings of IEEE WCNC 2000*, vol. 2, pp. 670–675, September 2000.

[16] W. Papen, "Improved soft handoff and macro-diversity for mobile radio," in *Proceedings of the International Conference on Communications*, vol. 3, pp. 1828–1833, June 1995.

[17] L. Ingber, "Very fast simulated re-annealing," *Mathematical and Computer modelling*, vol. 12, no. 8, pp. 967–973, 1989.

[18] Z. J. Haas and C.-P. Li, "The multiply-detected macrodiversity scheme for wireless cellular systems," *IEEE Transactions on Vehicular Technology*, vol. 47, pp. 506–530, May 1998.

[19] P. Willett and D. Warren, "The suboptimality of randomized tests in distributed and quantized detection systems," *IEEE Transactions on Information Theory*, vol. 38, pp. 355–361, March 1992.

[20] J. N. Tsitsiklis, "Decentralized detection," in *Vol. 2: Signal Detection*, Advances in Statistical Signal Processing, pp. 297–344, Jai Press, H.V. Poor and J.B. Thomas ed., 1991.

[21] S. Kirkpatrick, C. D. Gelatt, and M. P. Vecchi, "Optimization by simulated annealing," *Science*, vol. 220, no. 4598, pp. 671–680, 1983.

[22] S. Geman and D. Geman, "Stochastic relaxation, Gibbs distribution and the Bayesian restoration in images," *IEEE Transaction on Pattern Analysis and Machine Intelligence*, vol. 6, no. 6, pp. 721–741, 1984.

[23] H. Szu and R. Hartley, "Fast simulated annealing," *Physics Letter A*, vol. 122, no. 3-4, pp. 157–162, 1987.

[24] L. Ingber, "Adaptive simulated annealing(ASA): lessons learned," *Control and Cybernetics*, vol. 25, no. 1, pp. 33–54, 1996.

[25] L. Ingber, "Simulated annealing: Practice versus theory," *Mathematical and Computer modelling*, vol. 18, no. 11, pp. 29–57, 1993.

[26] M. Locatelli, *Simulated Annealing Algorithms for Continuous Global Optimization*, chapter 6, pp. 179–230. Handbook of Global Optimization II, Kluwer Academic Publishers, 2002.

[27] E. Malkamäki and H. Leib, "Evaluating the performance of convolution codes over block fading channels," *IEEE Transactions on Information Theory*, vol. 45, pp. 1643–1646, July 1999.

[28] A. Leon-Garcia, *Probability and Random Processes for Electrical Engineering Second Edition.* Addison-Wesley Publishing Company Inc, 1994.

[29] J. Hagenauer, "Viterbi decoding of convolutional codes for fading and burst channels," in *Proceedings of the 1980 Zurich Seminar on Digital Communications*, IEEE Catalog No. 80CH1521-4 COM, pp. G2.1–G2.7, 1980.

[30] J. G. Proakis, *Digital Communications Fourth Edition.* McGraw-Hill, 2000.

[31] R. Knopp and P. A. Humblet, "On coding for block fading channels," *IEEE Transactions on Information Theory*, vol. 46, pp. 189–205, January 2000.

[32] G. Ryzhik, *Tables of Integrals, Series and Products.* Academic Press Inc., 1980.

[33] W. A. Hashlamoun and P. K. Varshney, "An approach to the design of distributed Bayesian detection structures," *IEEE Transactions on Systems, Man and Cybernetics*, vol. 21, pp. 1206–1211, September 1991.

[34] M. K. Simon and M.-S. Alouini, *Digital Communication over Fading Channels: A Unified Approach to Performance Analysis.* Wiley-Interscience, 2000.