Minimum Mean-Square Error and Maximum Likelihood Multiuser Detection: Statistical Properties and Applications

Lie-Liang Yang
School of ECS, University of Southampton, SO17 1BJ, United Kingdom
Tel: 0044-(0)23-8059 3364, Email: lly@ecs.soton.ac.uk, http://www-mobile.ecs.soton.ac.uk

Abstract—Although multiuser detection (MUD) has received intensive research since its invention by Verdu in 1983 [1], however, the statistical properties in MUD have not received enough attention in research and application. In this contribution, we first study the statistical properties of minimum mean-square error (MMSE)-MUD, maximum likelihood (ML)-MUD and hybrid MMSE/ML-MUD from different aspects in the context of direct-sequence code-division multiple-access (DS-CDMA) communication and over additive white Gaussian noise (AWGN) channels. The extensions to the other communications scenarios with our emphasis on finding their possible applications, in particular for designing good spreading sequences to achieve enhanced system performance, are also addressed. The statistical properties of MUD are discussed. As shown in Section V, the linear MMSE-MUD and nonlinear ML-MUD are studied. The statistical properties related to both random variables.

II. PRELIMINARIES

Consider a synchronous DS-CDMA system supporting \( K \) number of users, which transmit binary phase-shift keying (BPSK) modulated signals over AWGN channels, the received discrete signals at the base-station (BS) can be written as [2]

\[
y = Cb + n = \sum_{k=1}^{K} c_k b_k + n
\]

where \( y \) denotes a \( N \)-length observation vector, \( C = [c_1, c_2, \ldots, c_K] \) is a \((N \times K)\) spreading matrix, which is constructed by the \( N \)-length spreading sequences assigned to the \( K \) users, \( c_k = [c_{k0}, c_{k1}, \ldots, c_{k(N-1)}]^{T}/\sqrt{N} \) is the \( k \)th user’s spreading sequence, where \( c_{kn} \in \{+1, -1\} \) and \( \|c_k\|^2 = 1 \). In (1) \( b = [b_1, b_2, \ldots, b_K]^T \) contains the data bits transmitted by the \( K \) users, \( b_k \) is assumed to be an independent binary random variable taking a value of +1 or −1 with equal probability. Finally, in (1) \( n \) is a \( N \)-length Gaussian noise vector obeying the multivariate normal distribution associated with mean zero and a covariance matrix of \( \sigma^2 R \), where \( \sigma^2 = (2\gamma_0)^{-1} \) and \( \gamma_0 \) denotes the signal-to-noise ratio (SNR) per bit. It has been shown in [2] that the sufficient statistics for detecting \( b \) can be obtained as

\[
y = C^T y = Rb + n
\]

where \( R = C^T C \) denotes the auto-correlation matrix of \( C \), while \( n = C^T n \) is now a \( K \)-length Gaussian noise vector distributed with mean zero and a covariance matrix \( \sigma^2 R \). In the context of the optimum ML-MUD, the estimate to \( b \) can be obtained by solving any one of the following equivalent optimization problems [2, 4]

\[
\hat{b} = \arg\min_{b \in \{+1, -1\}^K} \left\{ \lambda_0(\hat{b}) = \|y - C\hat{b}\|^2 \right\} \quad (3)
\]

\[
\hat{b} = \arg\min_{b \in \{+1, -1\}^K} \left\{ \lambda_\delta(\hat{b}) = \|y - R\hat{b}\|^2 + \|R^{-1}(y - R\hat{b})\|^2 \right\} \quad (4)
\]

\[
\hat{b} = \arg\min_{b \in \{+1, -1\}^K} \left\{ \lambda_\lambda(\hat{b}) = \|\bar{b}^* Rb - 2\gamma \bar{b}\|^2 \right\} \quad (5)
\]

where \( \hat{b} \) is a \( K \)-length binary test vector. Note that, the condition for using (4) is that the auto-correlation matrix \( R \) is invertible, which often implies that \( K \leq N \). By contrast, there is no such constraint for using (3) and (5). Explicitly, the optimum ML-MUD is a nonlinear MUD and requires an exhaustive search over all the \( 2^K \) possible combinations of the components in the binary data vector \( b \in \{+1, -1\}^K \), in order to find the optimum solution to \( b \). Hence, the computational complexity of the ML-MUD is exponentially dependent on the number of users of \( K \), making the optimum ML-MUD impractical to implement, when the number of users supported is relatively high. For this sake, efficient suboptimum algorithms, such as those summarized in [4], have usually been proposed for finding near-optimum solutions for the ML-MUD, yielding the suboptimum ML-MUDs.

Furthermore, in order to improve the efficiency of the suboptimum ML-MUDs, the detection process may be started with a low-complexity linear single-user or linear MUD procedure. By doing this, the detector has obtained certain amount of information about the transmitted bits before starting the suboptimum ML-MUD procedure, which may hence assist to find the near-optimum solution for the
suboptimum ML-MUD. In this contribution, as an example, the linear MMSE-MUD is invoked, which carries out the detection by forming a decision variable vector \( \mathbf{z} = [z_1, z_2, \ldots, z_K]^T \) as

\[
\mathbf{z} = \mathbf{W}^T \mathbf{y}
\]  

(6)

where \( \mathbf{W} \) is the optimum weight matrix in MMSE sense, which can be expressed as \([2, 4]\)

\[
\mathbf{W} = \left( \mathbf{C} \mathbf{C}^T + \sigma^2 \mathbf{I}_N \right)^{-1} \mathbf{C}
\]  

(7)

where \( \mathbf{R}_y = (\mathbf{C} \mathbf{C}^T + \sigma^2 \mathbf{I}_N) \) is the auto-correlation matrix of the observation vector \( \mathbf{y} \). It has been shown \([5]\) that, after the MMSE-MUD, the decision variables \( z_k, k = 1, \ldots, K \) can be closely approximated as independent Gaussian random variables expressed as

\[
z_k = b_k + \bar{n}_k, \quad k = 1, 2, \ldots, K
\]  

(8)

where \( \bar{n}_k \) is Gaussian distributed with mean zero and a variance given by \([4, 5]\)

\[
\bar{\sigma}^2 = \left[ 1 - c_k^2 \mathbf{R}_y^{-1} c_k \right]^{-1} - 1
\]  

(9)

Let us now study the statistics, when various detection scenarios are considered.

III. AMPLITUDE STATISTICS

In this section we study the amplitude statistics of the ML-metric \( \lambda(\cdot) \) in (3) and the decision variable \( z_k, k = 1, \ldots, K \), of (6). The PDFs depicted are either evaluated from the formulas provided or obtained by simulations. Let us first consider the amplitude statistics related to the ML-MUD.

A. Maximum Likelihood Multiuser Detection

1) Probability Distribution of \( \lambda(\cdot) \): Letting \( \mathbf{b} = \mathbf{b} \), from (1) and (3) we know that

\[
\lambda(\mathbf{b}) = ||\mathbf{y} - \mathbf{Cb}||^2 = ||\mathbf{n}||^2
\]  

(10)

Since \( \mathbf{n} \) is a multivariate Gaussian random vector distributed with mean zero and a covariance matrix \( \sigma^2 \mathbf{I}_N \), it can be shown that the normalized ML-metric \( \lambda(\mathbf{b}) / \sigma^2 \) obeys the central \( \chi^2 \)-distribution of \( N \) degrees-of-freedom with the probability density function (PDF) given by \([7, 8]\)

\[
f_{x_j}(x)(y) = \frac{1}{2^{(N/2)} \Gamma(N/2)} \left( \frac{y}{2} \right)^{N/2 - 1} \exp \left( -\frac{y}{2} \right), \quad y \geq 0
\]  

(11)

where \( \Gamma(\cdot) \) is the gamma function \([7]\).

2) Probability Distribution of \( \lambda(\mathbf{b}^{(U)}) \): Let \( \mathbf{b}^{(U)} \) denote a \( K \)-length test vector having \( U \leq K \) differences from the transmitted data vector \( \mathbf{b} \) as seen in (1). Then, the ML-metric \( \lambda(\mathbf{b}^{(U)}) \) is given by

\[
\lambda(\mathbf{b}^{(U)}) = ||\mathbf{y} - \mathbf{Cb}^{(U)}||^2 = ||\mathbf{C} \left( \mathbf{b} - \mathbf{b}^{(U)} \right) + \mathbf{n}||^2
\]  

(12)

It can be shown that, when assuming that \( \mathbf{C} \) is a random binary spreading matrix, the PDF of \( \lambda(\mathbf{b}^{(U)}) \) can be obtained. However, this PDF has a complicated structure, and furthermore, it is very hard to evaluate, when the value of \( U \) is high. For this sake, Monte-Carlo based simulation approaches are applied to obtain the PDF of \( \lambda(\mathbf{b}^{(U)}) \) or of its normalized version \( f_{x_j}(x^{(U)})(y) \). Fig. 1 illustrates the probability distributions of the ML-metrics in the DS-CDMA systems using random spreading sequences, when the test vector \( \mathbf{b} \) has \( U = 0, 1, 2 \) or 3 differences from the transmitted data vector \( \mathbf{b} \). Note that, the reason for us to consider only the cases of \( U = 0, 1, 2 \) and 3 is that their PDFs can provide us important information about the convergence behaviour of the ML-MUD. The PDFs in Fig. 1 show that the distribution curve moves toward the righthand-side as the number of differences between the test vector \( \mathbf{b} \) and the transmitted vector \( \mathbf{b} \) increases. Hence, when the test vector \( \mathbf{b} \) has more differences from the transmitted vector \( \mathbf{b} \), it generally results in a higher ML-metric value.

Additionally, according to the statistics of the ML-MUD, we know that the number of \( K \)-length binary vectors having low differences, such as \( U = 0, 1, 2 \), etc., from the transmitted vector is small. Hence, when a sufficiently low ML-metric appears during a random search process, it is then highly confident that the corresponding test vector \( \mathbf{b} \) is the actual vector \( \mathbf{b} \) transmitted or close to the vector \( \mathbf{b} \) transmitted. Therefore, the distribution information of the ML-metrics may be exploited for improving the efficiency of search in ML-MUD.

B. Minimum Mean-Square Error Multiuser Detection

Given the decision variables as shown in (8), if \( \{z_k\} \) can be approximated as independent identically distributed (iid) Gaussian random variables, the reliabilities of the detected bits in \( \mathbf{b} \) based on \( \{z_k\} \) can be measured by the amplitudes of \( \{z_k\} \) or the normalized amplitudes \( \{z_k / \bar{\sigma}\} \) \([9]\). Let us define the following variables:

\[
A_i = \min \{ |z_1|, |z_2|, \ldots, |z_K| \}, \quad i = 1, 2, \ldots, K
\]  

(13)

where \( A_i \) represents the reliability of the \( i \)-th most unreliable bit in \( \mathbf{b} \). We have \( A_1 \leq A_2 \leq \ldots \leq A_K \); \( A_1 \) and \( A_K \) are the reliabilities of the most unreliable and most reliable bits in \( \mathbf{b} \), respectively.

1) Probability Distribution of \( A_i \): The PDF of \( A_i, i = 1, \ldots, K \) can be derived, which can be expressed as

\[
f_{A_i}(y) = \sum_{u=0}^{i-1} \sum_{v=0}^{K-i} (-1)^{u-1} u! (u+1)! |y|^{u-1} K! Q^n \left( \sqrt{2y^2} - y \right) Q^n \left( 2\sqrt{2y^2} \right) f_{|\mathbf{b}_i|}(y)
\]  

(14)

where \( \gamma = 1/2\bar{\sigma}^2 \), \( Q(x) \) is the Gaussian Q-function defined as \( Q(x) = (2\pi)^{-1/2} \int_x^\infty e^{-t^2/2} dt \), and \( f_{|\mathbf{b}_i|}(y) \) is given by

\[
f_{|\mathbf{b}_i|}(y) = \frac{1}{\sqrt{2\pi}} \left[ e^{-y^2/2} + e^{-\left( y + \sqrt{2\gamma} \right)^2/2} \right], \quad y \geq 0
\]  

(15)

Fig. 2 shows the statistics of the ordered reliabilities of the MMSE-MUD in a DS-CDMA system employing a spreading factor \( N = 31 \) and supporting \( K = 31 \) users, when communicating over AWGN channels. It can be observed that, after the MMSE-MUD, the reliabilities of the detected data bits are highly sparse, some of them are much more reliable than the others. As shown in our forthcoming discourse, the reliability information extracted from the MMSE-MUD may be
Fig. 2. Probability distribution of the ordered reliabilities of the MMSE-MUD for a DS-CDMA system using random spreading sequences.

Fig. 3. PDF of reliability of the most unreliable bit on condition that the MMSE-MUD is correct ($H_1$) or incorrect ($H_0$), for a DS-CDMA system using random spreading sequences.

Fig. 4. PDF of the minimum ML-metric $B_0$ when $L = 0$ or $B_1$ when $L > 0$, on condition that the MMSE-MUD is correct ($H_1$) or incorrect ($H_0$).

MMSE-MUD, while with a complexity significantly lower than that of the ML-MUD. For convenience, we refer to this MUD scheme as the hybrid MMSE/ML-MUD. In this subsection we study the related statistics of the hybrid MMSE/ML-MUD. Its BER performance will be provided in Section V.

Let after the MMSE-MUD $\hat{b}_i = [b_{1i}, b_{2i}, \ldots, b_{Li}]^T$ contains the $L (0 \leq L \leq K)$ number of most unreliable bits, while the other $(K - L)$ bits are assumed to have been detected reliably by the MMSE-MUD. Explicitly, $L = 0$ corresponds to the case that the detector accepts the MMSE-MUD’s estimation as the final estimation, while $L = K$ corresponds the scenario that the detector ignores the MMSE-MUD’s estimation and carries out the complete ML-MUD. Letting $b_{li}$, $l = 1, \ldots, L$ takes values randomly in $\{+1, -1\}^L$, we can generate $M = 2^{KL}$ ML-metrics, which can be expressed in the form as

$$\{Z_m\} = \{\|y -Cb(\hat{b}_i)\|^2 : \hat{b}_i \in \{+1, -1\}^L\}$$

where $b(\hat{b}_i)$ denotes a $K$-length binary test vector constructed by the bits assumed to have been reliably detected and the bits from one realization of $\hat{b}_i \in \{+1, -1\}^L$. Let the n-th-minimum of $\{Z_1, Z_2, \ldots, Z_M\}$ be expressed as

$$B_{n} = \min \{Z_1, Z_2, \ldots, Z_M\} , n = 1, 2, \ldots, M$$

Additionally, we can see that, when $L = 0$, there is only one ML-metric, which is defined as $B_0 = Z_1$, where $Z_1$ is obtained by letting the test vector in ML-MUD equal to the MMSE-MUD’s estimated vector. Below we consider the statistics of $B_0$ or $B_1$, when the MMSE-MUD is correct (event $H_1$) or incorrect (event $H_0$). For both figures the DS-CDMA systems using random spreading sequences were assumed, and the DS-CDMA systems were full-load corresponding to $K = N = 31$. As shown in Figs. 4 and 5, if the detection of the MMSE-MUD or the ML-MUD is correct, the distributions of $B_0$ and $B_1$ for different values of $L$ retain nearly the same. This observation implies that, when the detection of the MMSE-MUD or ML-MUD is correct, the minimum of $\{Z_m\}$ generally follows the distribution of the ML-metric generated directly using the transmitted data vector. However, when the detection of the MMSE-MUD or ML-MUD is incorrect, as shown in Figs. 4 and 5, the PDF of $B_1$ of the minimum ML-metric in $\{Z_m\}$ is different for different $L$ values. Furthermore, as shown in Fig. 4, when the value of $L$ is lower than $8$ dB, the probability distribution of $B_1$ is higher than that of $B_0$.

C. Hybrid Minimum Mean-Square Error Maximum-Likelihood Multiuser Detection

From Sections III-A and III-B, we are implied that the reliability information provided by the MMSE-MUD may be exploited by the ML-MUD, in order to achieve a better BER performance than the
increases, the distribution of $B_1$ on condition that the MMSE-MUD is incorrect becomes close to and tends to resemble the distribution of $B_1$ on condition that the MMSE-MUD is correct. This is because, even although the MMSE-MUD is incorrect, the minimum ML-metric $B_1$ of $\{Z_m\}$ might still be the one generated by the transmitted data vector. In this case, the ML-MUD produces correct detection.

### IV. RATIO STATISTICS

In this section the ratio statistics is studied in the context of the hybrid MMSE/ML-MUD. The ratio concerned is defined as

$$ R = \frac{B_1}{B_2} \quad (18) $$

where $B_1$ and $B_2$ are defined by (17). $B_1$ represents the minimum of $\{Z_m\}$, while $B_2$ represents the second minimum of $\{Z_m\}$.

Straightforwardly, if the detection of the MMSE-MUD or ML-MUD is correct, then, the minimum $B_1$ is generated by a test vector that is the same as the transmitted vector $\mathbf{b}$, but the second minimum $B_2$ is generated by a test vector different from the transmitted vector $\mathbf{b}$. In this case, according to Fig. 1, the ratio $R = B_1/B_2$ will be significantly lower than one. By contrast, if the detection of the MMSE-MUD or ML-MUD is incorrect, then, the minimum $B_1$, the second minimum $B_2$, and even $B_3$, $B_4$, etc. may be similar. In this case, the ratio $R = B_1/B_2$ will be close to one.

Therefore, in Fig. 6 we illustrate the PDF of $R = B_1/B_2$ on condition that the detection of the MMSE-MUD is correct ($H_1$) or incorrect ($H_0$), while in Fig. 7 the PDF of $R = B_1/B_2$ on condition that the detection of the ML-MUD is correct ($H_1$) or incorrect ($H_0$). It is observed that the ratio $R = B_1/B_2$ is mainly distributed in the region close to one, when the detection of the MMSE-MUD or the ML-MUD is incorrect. Otherwise, when the detection of the MMSE-MUD or the ML-MUD is correct, the ratio $R = B_1/B_2$ is then distributed within a region having the ratio values explicitly lower than one.

Let us below discuss some possible applications inspired by the statistical properties of the MMSE- and ML-MUD.

### V. APPLICATION EXAMPLES

When the statistical properties related to the MMSE-MUD or/and ML-MUD are available, they may be exploited for enhancing the performance of the corresponding DS-CDMA system in various ways. Below are some examples showing the possible applications.

#### A. Design of Efficient Search Algorithms

For example, when the statistics of $\tilde{\lambda}(\mathbf{b}^{(U)})$ or $\bar{\lambda}(\mathbf{b}^{(U)})$ for $u = 0, 1, \ldots, K$ are available, i.e., when the receiver knows the PDFs of $f_{\tilde{\lambda}(\mathbf{b}^{(U)})}(y)$ or $f_{\bar{\lambda}(\mathbf{b}^{(U)})}(y)$ for $u = 0, 1, \ldots, K$, then, given a ML-metric $\lambda$ or $\bar{\lambda}$, the receiver may obtain certain information about how many differences between the corresponding test vector and the transmitted vector $\mathbf{b}$. Consequently, the ML-MUD may improve its search efficiency with the aid of this information.

#### B. Design of Spreading Sequences

Assuming a ML test vector $\mathbf{b}^{(U)}$, which has $U$ differences from the transmitted data vector $\mathbf{b}$. Then, according to (12) and when ignoring the background noise, the ML-metric is given by

$$ \lambda(\mathbf{b}^{(U)}) = \|y - C\mathbf{b}^{(U)}\|^2 = \|C(\mathbf{b} - \mathbf{b}^{(U)})\|^2 \quad (19) $$

This equation implies that $\lambda(\mathbf{b}^{(U)})$ for $U \neq 0$ should be as large as possible in order to minimize the error probability after the ML-MUD. This observation can also be reflected by Fig. 1: in order for the ML-MUD to achieve the minimum error probability, the overlapping area between the PDF of $f_{\lambda(\mathbf{b}^{(U)})}(y)$ and the other PDFs of $f_{\lambda(\mathbf{b}^{(U')})}(y)$ for $U \neq 0$ should be as low as possible. Therefore, an optimum set of spreading sequences for the DS-CDMA system using ML-MUD should be designed according to

$$ C_{\text{opt}} = \arg \max_{C} \left\{ \lambda(\mathbf{b}^{(U)}) = \|C(\mathbf{b} - \mathbf{b}^{(U)})\|^2 \right\} \quad (20) $$
for any $U \neq 0$.

Note that, the optimization problem of (20) may be extremely hard to solve. For this sake, sub-optimum solutions may be obtained by considering the following facts. First, the PDFs $f_{\hat{Z}(b^{(i)})}(y)$ and $f_{Z(b^{(i)})}(y)$ maintain the same for any set of spreading sequences. Second, from Fig. 1 we can be implied that only those test vectors having a small number of differences from the transmitted vector may have noticeable effect on the achievable error performance. Hence, the spreading sequences may be designed by considering only some low values of $U$, such as $U = 2$ and 3.

C. Enhancing Post-MUD Processing

The statistics obtained after the MMSE-MUD or ML-MUD can be exploited to enhance the post-MUD processing. For example, with the aid of the reliability information provided by the MMSE-MUD or ML-MUD, in an error-control coding assisted DS-CDMA system soft-input soft-output (SISO) decoding or iterative decoding [10] may be carried out in order to improve the system performance. Additionally, based on the reliability information, some unreliably detected data bits may be erased. Then, errors-and-erasures decoding can be carried out by the DS-CDMA receiver. It has been illustrated that the errors-and-erasures decoding usually outperforms the errors-only decoding [11, 12].

D. Design of Novel Multiuser Detection Schemes

With the aid of the statistical properties of the multiuser detection, such as that of the MMSE-MUD and ML-MUD, novel MUD schemes may be designed. For example, a hybrid MMSE/ML-MUD scheme may be designed, which first carries out the MMSE-MUD. Then, as shown in Fig. 1, we can be implied that only those test vectors having a small number of differences from the transmitted vector may have noticeable effect on the achievable error performance. Hence, the spreading sequences may be designed by considering only some low values of $U$, such as $U = 2$ and 3.

Fig. 8. BER versus SNR per bit $\gamma_0$ performance of the DS-CDMA systems using conventional MMSE-MUD ($L = 0$) and hybrid MMSE/ML-MUD, when random spreading sequences are employed.

Fig. 9. BER versus SNR per bit $\gamma_0$ performance of the DS-CDMA systems using conventional MMSE-MUD ($L = 0$) and hybrid MMSE/ML-MUD, when Gold-sequences are employed.

In Summary, in this paper we have studied a range of statistics in the MMSE-MUD, ML-MUD and the hybrid MMSE/ML-MUD. It can be shown that these statistics may be found many applications for improving the performance of DS-CDMA systems. Finally, note that, the work considered in this paper may be extended to various communications scenarios in wireless communications, including a variety of MIMO systems requiring MUD or equalization for suppressing ISI, CCI, MUI, etc.

REFERENCES