Progress Report

Tianpei Chen
Department of Electrical and Computer Engineering
McGill University
Montreal, Quebec, Canada

Abstract

The abstract goes here.

Index Terms

IEEEtran, journal, LaTeX, paper, template.

I. System Model

Consider a uncoded complex Large-Scale MIMO (LS-MIMO) uplink spatial multiplexing (SM) system with N_t users, where each is equipped with one transmit antenna. The number of receive antennas at the Base Station (BS) is N_r, $N_r \geq N_t$. Typically LS-MIMO systems have hundreds of antennas at the BS,

Bit sequences, which are modulated to complex symbols, are transmitted by the users over a flat fading channel. The discrete time model of the system is given by:

$$\mathbf{y} = \mathbf{H}\mathbf{s} + \mathbf{n}, \tag{1}$$

where $\mathbf{y} \in \mathbb{C}^{N_r \times 1}$ is the received symbol vector, $\mathbf{s} \in \mathbb{C}^{N_t}$ is the transmitted symbol vector, with components that are mutually independent and taken from a finite signal constellation alphabet $\mathbb{O}$ (e.g. BPSK, 4-QAM, 16-QAM, 64-QAM), $|\mathbb{O}| = M$. The transmitted symbol vectors $\mathbf{s} \in \mathbb{O}^{N_t}$, satisfy $\mathbb{E}[\mathbf{s}\mathbf{s}^H] = \mathbf{I}_{N_t} E_s$, where E_s denotes the symbol average energy, $\mathbb{E}[\cdot]$ denotes the expectation operation, $\mathbf{I}_{N_t}$ denotes identity matrix of size $N_r \times N_t$. Furthermore $\mathbf{H} \in \mathbb{C}^{N_r \times N_t}$

denotes the Rayleigh fading channel propagation matrix, $\mathbf{H}_{ij}$ denotes the component of $\mathbf{H}$ at ith row and jth column, representing the channel response from ith receive antenna to the jth transmit antenna. Each component is independent identically distributed (i.i.d) circularly symmetric complex Gaussian (CSCG) random variable with unit variance. Finally, $\mathbf{n} \in \mathbb{C}^{N_r}$ is the additive white Gaussian noise (AWGN) vector with zero mean components and $\mathbb{E}[\mathbf{n}\mathbf{n}^H] = \mathbf{I}_{N_r} N_0$, where N_0 denotes the noise power spectrum density, and hence $\frac{E_s}{N_0}$ is the signal to noise ratio (SNR).

The task of a MIMO detector is to estimate the transmit symbol vector $\mathbf{s}$, based on the knowledge of receive symbol vector $\mathbf{y}$ and channel matrix $\mathbf{H}$.

The Optimal (in a sense of lowest average error probability) Maximum Likelihood Detector (MLD) for MIMO system is given by

$$\hat{\mathbf{s}}^{ML} = \min_{\mathbf{s} \in \mathbb{O}^{N_t}} ||\mathbf{y} - \mathbf{H}\mathbf{s}||^2, \tag{2}$$

where $||\cdot||$ denotes the 2-norm operation, therefore from (2), the solution of MLD is the $\hat{\mathbf{s}}$ that can generate the minimum Euclidean distance between vector $\mathbf{y}$ and $\mathbf{H}\mathbf{s}$.

Consider an alternative representation of MLD principle, let $\mathbf{H}_1 \in \mathbb{C}^{N_r \times N}$ denotes a sub matrix composed of N columns from $\mathbf{H}$, let $\mathbf{s}_1 \in \mathbb{C}^N$ denote the sub symbol vector whose components are transmitted from the users corresponding to $\mathbf{H}_1$. Similarly, let $\mathbf{H}_2 \in \mathbb{C}^{N_r \times (N_t - N)}$ denotes the sub matrix composed of the remaining columns from $\mathbf{H}$ and $\mathbf{s}_2 \in \mathbb{C}^{(N_t - N)}$ is the sub symbol vector whose components are transmitted by the users corresponding to $\mathbf{H}_2$. Thus (1) can be rewritten as

$$\mathbf{y} = \mathbf{H}_1\mathbf{s}_1 + \mathbf{H}_2\mathbf{s}_2 + \mathbf{n}. \tag{3}$$

Let $\hat{\mathbf{s}}_1, \hat{\mathbf{s}}_2$ denote the estimations of $\mathbf{s}_1, \mathbf{s}_2$, we have $\hat{\mathbf{s}}_1 \in [\hat{\mathbf{s}}_1^1, \hat{\mathbf{s}}_1^2, \cdots \hat{\mathbf{s}}_1^K], , K = M^N$ and $\hat{\mathbf{s}}_2 \in [\hat{\mathbf{s}}_2^1, \hat{\mathbf{s}}_2^2, \cdots \hat{\mathbf{s}}_2^Q], Q = M^{N_t - N}$, MLD in (2) can be rewritten as

$$[\tilde{k}, \tilde{q}] = \min_{k \in [1,2\ldots K]} \min_{q \in [1,2,\ldots,Q]} ||\mathbf{y} - \mathbf{H}_1\hat{\mathbf{s}}_1^k - \mathbf{H}_2\hat{\mathbf{s}}_2^q||^2, \mathbf{s}_1^{ML} = \hat{\mathbf{s}}_1^{\tilde{k}}, \mathbf{s}_2^{ML} = \hat{\mathbf{s}}_2^{\tilde{q}}, \tag{4}$$

(4) can be divided into three steps, first define

$$\mathbf{y}^k = \mathbf{y} - \mathbf{H}_1\hat{\mathbf{s}}_1^k \quad k \in [1, 2, \ldots, K], \tag{5}$$

Then solve

$$\hat{\mathbf{x}}_2^k = \min_{\hat{\mathbf{s}}_2^q \in [\hat{\mathbf{s}}_2^1, \hat{\mathbf{s}}_2^3 \ldots, \hat{\mathbf{s}}_2^Q]} ||\mathbf{y}^k - \mathbf{H}_2 \hat{\mathbf{s}}_2^q||^2, \tag{6}$$

$$\tilde{k} = \min_{k \in [1,2,\ldots,K]} ||\mathbf{y}^k - \mathbf{H}_2 \hat{\mathbf{x}}_2^k||^2, \tag{7}$$

Finally we have $\mathbf{s}_1^{ML} = \hat{\mathbf{s}}_1^{\tilde{k}}, \mathbf{s}_2^{ML} = \hat{\mathbf{x}}_2^{\tilde{k}} = \hat{\mathbf{s}}_2^{\tilde{q}}$.

II. Minimum Achievable Diversity based Channel Partition

A. Diversity Maximization Selection

Based on the alternative form of MLD in (4)-(7), General Parallel Interference Cancellation (GPIC) algorithm first generate a list of symbol vector candidates and then choose the candidate with the minimum Euclidean distance as the solution. To elaborate further, GPIC first generate a list of symbol vector candidate by exhaustive search for all the possible $\hat{\mathrm{s}}_1$, then rather than performing exhaustive search in (6), GPIC exploit low complexity sub optimal linear detectors (i.e., zeros forcing (ZF) and minimum mean square error (MMSE)) and their successive interference cancellation (SIC) counterparts to get the estimation of s_2, for sake of simplicity, hereinafter we use LD/SIC. Then the best candidate in the list with the minimum Euclidean distance is chosen as the solution.

Performance analysis of GPIC algorithm is provided in [1], here we briefly conclude the main results in [1], at the receiving side, the frame error is defined as there is at least one erroneous symbol in the estimation of transmitted symbol vector. Let P_e denotes the average frame error probability. The diversity order, denoted by d is defined as the slope of the P_e in log-scale in high SNR region, which is given by [2]

$$d = - \lim_{SNR \to \infty} \frac{\log P_e}{\log(SNR)}, \tag{8}$$

Let P_e^{ML} denotes the average frame error probability of MLD, and $P_{e2} = \mathbb{E}_{\mathbf{H}_2, \mathbf{s}_2}(Pr(\hat{\mathrm{s}}_2^{k_1} \neq \mathrm{s}_2 | \mathbf{H}_2, \mathrm{s}_2))$, where $P_r(A)$ denotes the probability that a event A occurs, $\hat{\mathrm{s}}_1^{k_1} = \mathrm{s}_1$, therefore P_{e2} is the average frame error probability of the LD/SICs for the sub system in which the interference from the users that transmitted $\mathbf{s}_1$ are perfectly cancelled from the observation $\mathbf{y}$. given by

$$\mathbf{y}^{k_1} = \mathbf{y} - \mathbf{H}_1 \hat{\mathbf{s}}_1^{k_1}, \tag{9}$$

because $\hat{\mathbf{s}}_1^{k1} = \mathbf{s}_1$, we have

$$\mathbf{y}^{k1} = \mathbf{H}_2 \mathbf{s}_2 + \mathbf{n}, \tag{10}$$

The average frame error probability of GPIC algorithm P_{et} is bounded by

$$\max(P_e^{ML}, P_{e2}) \leq P_{et} \leq P_e^{ML} + P_{e2}, \tag{11}$$

in (11), the diversity order of MLD is N_r, P_{e2} is the average frame error probability of sub optimal LD/SICs, let d_2 denotes the diversity order of LD/SICs, based on (8), we have

$$\lim_{SNR \to \infty} P_e^{ML} \propto SNR^{-N_r}, \tag{12}$$

$$\lim_{SNR \to \infty} P_{e2} \propto SNR^{-d_2}, \tag{13}$$

Let d_t denotes the diversity order of GPIC algorithm, now consider the following conditions

- $d_2 > N_r$,

  $$\lim_{SNR \to \infty} \frac{P_{e2}}{P_e^{ML}} = SNR^{N_r - d_2} = 0, \tag{14}$$

  based on (11), we have

  $$\lim_{SNR \to \infty} \frac{P_{et}}{P_e^{ML}} = 1, \tag{15}$$

  thus $d_t = N_r$.

- $d_2 = N_r$, we have

  $$\lim_{SNR \to \infty} P_{et} \propto SNR^{-N_r}, \tag{16}$$

  thus $d_t = N_r$.

- $d_2 < N_r$,

  $$\lim_{SNR \to \infty} \frac{P_e^{ML}}{P_{e2}} = SNR^{d_2 - N_r} = 0, \tag{17}$$

based on (11), we have

$$\lim_{SNR\to\infty} P_{et} = P_{e2} \propto SNR^{-d_2}, \tag{18}$$

thus $d_t = d_2$.

Therefore when $d_2 \geq N_r$, GPIC algorithm can achieve ML performance asymptotically, when $d_2 < N_r$, the diversity order that GPIC can achieve at high SNR region is equal to that of LD/SICs, in [1], the authors employ diversity maximization selection (DMS) scheme for channel partition, which is proposed in [3]. The optimum diversity order can be guaranteed with low complexity.

To elaborate further, for a given number of antennas chosen at the channel partition stage N, there is $N_u = \begin{bmatrix} N_t \\ N \end{bmatrix}$ possible combinations of $[\mathbf{H}_1, \mathbf{H}_2]$, according to DMS principle, the maximum diversity order of LD/SIC is achieved by choosing the subset $\mathbf{H}_2$ that has the strongest weakest substream, in a sense of post processing SNR. Therefore, when using MMSE/SIC as the sub optimal detector, the subset $\mathbf{H}_2^{opt}$ selected based on DMS principle should be [3]

$$\mathbf{H}_2^{opt} = \mathbf{H}_2^{p}, \tag{19}$$

$$p = \min_{j=1,2,\ldots,N_u} \theta_j, \tag{20}$$

$$\theta_j = \max_{k=1,2,\ldots,N_t-N} ((\mathbf{H}_2^j)^H \mathbf{H}_2^j + SNR^{-1}\mathbf{I})^{-1}_{kk}, \quad \mathbf{H}_2^j \in [\mathbf{H}_2^1, \mathbf{H}_2^2, \ldots, \mathbf{H}_2^{N_u}]; \tag{21}$$

Where $\mathbf{A}_{kk}$ denotes the kth diagonal component of matrix $\mathbf{A}$. The diversity order of LD/SIC detector after DMS channel partition process is given by $d_{2DMS} = (N+1)(N_r - N_t + N + 1)$. In [1], the authors derive the minimum number of antennas N_{min} at channel partition stage that can guarantee $d_{2DMS} \geq N_r$, which is given by

$$N_{min} = \lceil \sqrt{\frac{(N_r - N_t)^2}{4} + N_r} - \frac{N_r - N_t}{2} - 1 \rceil, \tag{22}$$

where $\lceil \alpha \rceil$ denotes the minimum integer no less than α.

B. Minimum Achievable Diversity: How many antennas do we need for LS-MIMO?

In LS-MIMO V-BLAS systems, the maximum diversity order is N_r, which is extremely large under the condition that there is tens to hundreds of receive antennas. On the one hand, diversity order is achieved at high SNR region, however, in practical, such a high diversity order is not necessary, because the SNR region of interest is the range in which the bit error rate (BER) or symbol error rate (SER) can be 10^{-5} to 10^{-7}. On the other hand, the computational complexity to guarantee ML performance by DMS channel partition is excessive in LS-MIMO, for example, when $N_r = N_t$, based on (22), the $N_{min} = \lceil \sqrt{N_r} - 1 \rceil$, the number of the symbol vector candidates in the list generated by exhaustive search for $\mathbf{s}_1$ is $M^{N_{min}}$.

Therefore, here we consider minimum achievable diversity (MAD) principle, which sacrifices redundant diversity gain in LS-MIMO, in order to reduce the complexity. Let d_{2MAD} denotes the diversity order of LD/SIC detectors, $\tilde{N}_{min}$ denotes the minimum number of antennas chosen by MAD principle. Our goal is to guarantee the overall diversity is no less than a given minimum diversity while keep $\tilde{N}_{min}$ small. Let $1 \leq g < N_r$ denote the given minimum achievable diversity order, based on (18), we have

$$\begin{aligned} &if \quad d_{2MAD} < N_r \quad and \quad d_{2MAD} \geq g \\ &d_t = d_{2MAD} \geq g. \end{aligned} \tag{23}$$

Therefore we can guarantee that the overall diversity order of GPIC algorithm is no less than g, in order to derive the minimum number of antennas required based on MAD principle, we define function

$$f(N) = d_{2MAD} - g = (N+1)(N_r - N_t + N + 1) - g, \tag{24}$$

where $N \in [1, 2, \ldots, N_t]$, our goal is to find minimum N, denoted by $\tilde{N}_{min}$, that satisfy $f(N) \geq 0$. The two zeros points of quadratic function $f(N)$ are

$$N_1 = -\sqrt{\frac{(N_r - N_t)^2}{4} + g} - \frac{N_r - N_t}{2} - 1, \tag{25}$$

$$N_2 = \sqrt{\frac{(N_r - N_t)^2}{4} + g} - \frac{N_r - N_t}{2} - 1 \tag{26}$$

when $N < N_1$ or $N > N_2$, $f(N) \geq 0$, obviously $N_1 < 0$, there is no effective N exits that satisfy the former condition, thus, we have

$$\tilde{N}_{min} = \lceil \sqrt{\frac{(N_r - N_t)^2}{4} + g} - \frac{N_r - N_t}{2} - 1 \rceil \tag{27}$$

III. Theoretical Analysis of Channel Hardening

A. Preliminary

Orthogonality deficiency ϕ_{od} measures how orthogonal a matrix is [4], which is defined by

$$\phi_{od} = 1 - \frac{det(\mathbf{W})}{\prod_{i=1}^{N_t} ||\mathbf{h}_i||^2}, \tag{28}$$

where $\mathbf{W} = \mathbf{H}^H\mathbf{H}$ denotes Wishart matrix, $\mathbf{h}_i$ denotes the i th column of $\mathbf{H}$, $det(\cdot)$ denotes determinant operation, $||\cdot||$ denotes 2-norm operation. Based on Hadamard's inequality $\prod_{i=1}^{N_t} ||\mathbf{h}_i|| \geq det(\mathbf{H})$, we have $0 \leq \phi_{od} \leq 1$, if $\mathbf{H}$ is singular, $\phi_{od} = 1$, if $\mathbf{H}$ is orthogonal, then $\phi_{od} = 0$.

$||\mathbf{h}_i||^2 = \sum_{j=1}^{N_r} |\mathbf{H}_{ji}|^2$, where $|\cdot|$ denotes magnitude operation. $\mathbf{H}_{ij} \sim \mathbb{C}N(0,1)$, where $\mathbb{C}N(0,\sigma^2)$ denotes CSCG distribution with variance σ^2, thus $|\mathbf{H}_{ji}| \sim Rayleigh(1/\sqrt{2})$, where $Rayleigh(\sigma)$ denotes the Rayleigh distribution with shape parameter σ, therefore $||\mathbf{h}_i||^2 \sim Gamma(N_r, 1)$ [5]. $Gamma(k,\theta)$ denotes Gamma distribution, with k degrees of freedom and scale parameter θ.

For sake of simplicity, we define orthogonality measure ϕ_{om}, which is given by

$$\phi_{om} = \frac{det(\mathbf{W})}{\prod_{i=1}^{N_t} ||\mathbf{h}_i||^2}, \tag{29}$$

$0 \leq \phi_{om} \leq 1$, if ϕ_{om} is closer to 1, $\mathbf{H}$ is more closer to an orthogonal matrix.

B. Introduction to Channel Hardening Phenomenon

C. Logarithmic Expectation of Orthogonality Measure

Based on (29), The logarithm of ϕ_{om} can be written as

$$\ln(\phi_{om}) = \ln(det(\mathbf{W})) - \sum_{i=1}^{N_t} \ln(||\mathbf{h}_i||^2), \tag{30}$$

Taking expectation of (30), we have

$$\mathbb{E}[\ln(\phi_{om})] = \mathbb{E}[\ln(det(\mathbf{W}))] - \sum_{i=1}^{N_t} \mathbb{E}[\ln(||\mathbf{h}_i||^2)]. \tag{31}$$

First we consider the first summand on the right hand of (31) $\mathbb{E}[\ln(det(\mathbf{W}))]$. Let $\mathbb{C}W_m(n, \Sigma)$ denotes complex Wishart distribution, which is the joint distribution of sample covariance matrix from multivariate complex Gaussian random variable [6]. n is the degree of freedom and $\Sigma \in \mathbb{C}^{m\times m}$ is the covariance matrix. Let Π_i denotes the ith row of $\mathbf{H}$. Π_i is a N_t-variate complex Gaussian random variable so does Π_i^H. Therefore $\mathbf{W} = \mathbf{H}^H\mathbf{H} = \sum_{i=1}^{N_r} \Pi_i^H \Pi_i \sim \mathbb{C}W_{N_t}(N_r, \mathbf{I})$. The logarithmic expectation of $det(\mathbf{W})$ is

$$\mathbb{E}[\ln(det(\mathbf{W}))] = \sum_{i=1}^{N_t} \psi(N_r - i + 1), \tag{32}$$

where $\psi(n)$ denotes Digamma function, which is given by [5]

$$\psi(n) = \frac{\Gamma'(n)}{\Gamma(n)}, \tag{33}$$

$\Gamma(n)$ denotes Gamma function. Proof: see Appendix A.

Now we consider the second summand on the right hand of (31) $\sum_{i=1}^{N_t} \mathbb{E}[\ln(||\mathbf{h}_i||^2)]$. $||\mathbf{h}_i||^2 \sim Gamma(N_r, 1)$, the logarithmic expectation of a Gamma random variable $\gamma \sim Gamma(n, \theta)$ can be written as:

$$\mathbb{E}[\ln(\gamma)] = \psi(n) + \ln(\theta), \tag{34}$$

Thus (34), we have

$$\sum_{i=1}^{N_t} \mathbb{E}[\ln(||\mathbf{h}_i||^2)] = \sum_{i=1}^{N_t} \psi(N_r), \tag{35}$$

proof: see Appendix B.

Based on (32) and (35), the logarithmic expectation of orthogonality measure ϕ_{om} is

$$\mathbb{E}[\ln(\phi_{om})] = \sum_{i=1}^{N_t} [\psi(N_r - i + 1) - \psi(N_r)], \tag{36}$$

D. Probability Density Function of Orthogonality Measure

IV. Simplified Channel Partition based on Orthogonality Measure

V. Computer Simulations

A. Full Loaded Systems

B. Systems with Different Loading Factors

VI. Conclusion

The conclusion goes here.

Appendix A

Let $\mathbf{A} \in \mathbb{C}^{m\times m}$, $A \sim \mathbb{C}W_m(n, \mathbf{\Sigma})$, According to the definition of complex Wishart matrix, it is obvious $\mathbf{A}$ is Hermition positive definite matrix (i.e., $\mathbf{A} = \mathbf{A}^H > 0$).

The pdf of $\mathbf{A}$ can be written as [7]:

$$f(\mathbf{A}) = \{\tilde{\Gamma}_m(n) det(\mathbf{\Sigma})^n\}^{-1} det(\mathbf{A})^{n-m} etr(-\mathbf{\Sigma}^{-1}\mathbf{A}), \tag{37}$$

where $\tilde{\Gamma}_m(\beta)$ denotes multivariate complex Gamma function defined by [7]:

$$\tilde{\Gamma}_m(\beta) = \pi^{\frac{m(m-1)}{2}} \prod_{i=1}^{m} \Gamma(\beta - i + 1) \quad Re(\beta) > m - 1. \tag{38}$$

Furthermore, from [7], we have

$$\tilde{\Gamma}_m(\beta) = \int_{\mathbf{X}=\mathbf{X}^H>0} etr(-\mathbf{X}) det(\mathbf{X})^{\beta-m} d\mathbf{X} \quad Re(\beta) > m - 1. \tag{39}$$

We derive logarithmic expectation of $det(\mathbf{A})$

$$\begin{aligned}
\mathbb{E}[\ln(det(\mathbf{A}))] &= \int_{\mathbf{A}=\mathbf{A}^H>0} \ln(det(\mathbf{A})) f(\mathbf{A}) d\mathbf{A} \\
&= \int_{\mathbf{A}=\mathbf{A}^H>0} \ln(det(\mathbf{A})) \{\tilde{\Gamma}_m(n) det(\mathbf{\Sigma})^n\}^{-1} det(\mathbf{A})^{n-m} etr(-\mathbf{\Sigma}^{-1}\mathbf{A}) d\mathbf{A} \\
&= \frac{det(\mathbf{\Sigma})^{-n}}{\tilde{\Gamma}_m(n)} \int_{\mathbf{A}=\mathbf{A}^H>0} \ln(det(\mathbf{A})) det(\mathbf{A})^{n-m} etr(-\mathbf{\Sigma}^{-1}\mathbf{A}) d\mathbf{A},
\end{aligned} \tag{40}$$

if $\mathbf{\Sigma} = \mathbf{I}$, (40) can be written as

$$\mathbb{E}[\ln(det(\mathbf{A}))] = \frac{1}{\tilde{\Gamma}_m(n)} \int_{\mathbf{A}=\mathbf{A}^H>0} \ln(det(\mathbf{A}))det(\mathbf{A})^{n-m} etr(-\mathbf{A}) d\mathbf{A}. \tag{41}$$

Because $\frac{d}{dn}[det(\mathbf{A})]^{n-m} = \ln(det(\mathbf{A}))det(\mathbf{A})^{n-m}$, (41) can be rewritten as

$$\mathbf{E}[\ln(det(\mathbf{A}))] = \frac{1}{\tilde{\Gamma}_m(n)} \frac{d}{dn} \int_{\mathbf{A}=\mathbf{A}^H>0} etr(-\mathbf{A}) det(\mathbf{A})^{n-m} d\mathbf{A}, \tag{42}$$

Based on (39), in (42), we have

$$\tilde{\Gamma}_m^{'}(n) = \frac{d}{dn} \int_{\mathbf{A}=\mathbf{A}^H>0} etr(-\mathbf{A}) det(\mathbf{A})^{n-m} d\mathbf{A}, \tag{43}$$

Therefore (42) can be rewritten as

$$\mathbf{E}[\ln(\mathbf{A})] = \frac{\tilde{\Gamma}_m^{'}(n)}{\tilde{\Gamma}_m(n)}. \tag{44}$$

Based on (38), we have

$$\tilde{\Gamma}_m^{'}(n) = \pi^{\frac{m(m-1)}{2}} \sum_{i=1}^{m} [\Gamma^{'}(n-i+1) \prod_{j\neq i}^{m} \Gamma(n-j+1)], \tag{45}$$

Using (38) and (45), we have

$$\begin{aligned} \frac{\tilde{\Gamma}_m^{'}(n)}{\tilde{\Gamma}_m(n)} &= \frac{\sum_{i=1}^{m} [\Gamma^{'}(n-i+1) \prod_{j\neq i}^{m} \Gamma(n-j+1)]}{\prod_{k=1}^{m} \Gamma(n-k+1)} \\ &= \sum_{i=1}^{m} [\frac{\Gamma^{'}(n-i+1) \prod_{j\neq i}^{m} \Gamma(n-j+1)}{\prod_{k=1}^{m} \Gamma(n-k+1)}] = \sum_{i=1}^{m} \frac{\Gamma^{'}(n-i+1)}{\Gamma(n-i+1)}, \end{aligned} \tag{46}$$

Therefore (44) can be rewritten as

$$\mathbf{E}[\ln(det(\mathbf{A}))] = \sum_{i=1}^{m} \psi(n-i+1), \tag{47}$$

where $\psi(n) = \frac{\Gamma^{'}(n)}{\Gamma(n)}$ denotes Digamma function.

Appendix B

If $x \sim Gamma(n, \theta)$, with shape parameter k and scale parameter θ, $x > 0$, $\Gamma(k)$ denotes Gamma function, the density function of Gamma distribution is

$$f(x, k, \theta) = \frac{x^{k-1} e^{-x/\theta}}{\Gamma(k)\theta^k}. \tag{48}$$

where $\Gamma(n)$ satisfies [5]

$$\Gamma(n) = \int_0^\infty x^{n-1} e^{-x} dx, \tag{49}$$

Thus the logarithmic expectation of x can be written as

$$\mathbf{E}[\ln(x)] = \frac{1}{\Gamma(k)} \int_0^\infty \ln(x) x^{k-1} e^{-x/\theta} \theta^{-k} dx, \tag{50}$$

define $z = x/\theta$ and based on (49), (50) can be rewritten as

$$\mathbf{E}[\ln(x)] = \ln(\theta) + \frac{1}{\Gamma(k)} \int_0^\infty \ln(z) z^{k-1} e^{-z} dz. \tag{51}$$

Because $\frac{d(z^{k-1})}{dk} = \ln(z) z^{k-1}$, (51) can be rewritten as

$$\mathbf{E}[\ln(x)] = \ln(\theta) + \frac{1}{\Gamma(k)} \frac{d}{dk} \int_0^\infty z^{k-1} e^{-z} dz, \tag{52}$$

Based on (49), we have

$$\Gamma'(k) = \frac{d}{dk} \int_0^\infty z^{k-1} e^{-z} dz, \tag{53}$$

Thus (52) can be rewritten as

$$\mathbf{E}(\ln(x)) = \ln(\theta) + \frac{\Gamma'(k)}{\Gamma(k)} = \ln(\theta) + \psi(k), \tag{54}$$

where $\psi(k)$ denotes Digamma function.

References

[1] D. Radji and H. Leib, "Interference cancellation based detection for v-blast with diversity maximizing channel partition," *Selected Topics in Signal Processing, IEEE Journal of*, vol. 3, no. 6, pp. 1000–1015, 2009.

[2] L. Zheng and D. N. Tse, "Diversity and multiplexing: a fundamental tradeoff in multiple-antenna channels," *Information Theory, IEEE Transactions on*, vol. 49, no. 5, pp. 1073–1096, 2003.

[3] H. Zhang, H. Dai, Q. Zhou, and B. L. Hughes, "On the diversity order of spatial multiplexing systems with transmit antenna selection: A geometrical approach," *Information Theory, IEEE Transactions on*, vol. 52, no. 12, pp. 5297–5311, 2006.

[4] X. Ma and W. Zhang, "Performance analysis for MIMO systems with lattice-reduction aided linear equalization," *Communications, IEEE Transactions on*, vol. 56, no. 2, pp. 309–318, 2008.

[5] A. Papoulis and S. U. Pillai, *Probability, random variables, and stochastic processes*. Tata McGraw-Hill Education, 2002.

[6] N. Goodman, "Statistical analysis based on a certain multivariate complex gaussian distribution (an introduction)," *Annals of mathematical statistics*, pp. 152–177, 1963.

[7] D. K. Nagar and A. K. Gupta, "Expectations of functions of complex Wishart matrix," *Acta applicandae mathematicae*, vol. 113, no. 3, pp. 265–288, 2011.