Indefinite Hermitian Quadratic Forms in Gaussian Random Variables: Distribution, Scaling, and Some Applications

Tareq Y. Al-Naffouri

Electrical Engineering Department
King Fahd University of Petroleum and Minerals
www.kfupm.edu.sa/faculty/ee/naffouri

Part of this work was done jointly with Masoud Sharif and Babak Hassibi

A Brief about KFUPM

- Established in 1963
- Located in Dhahran, in the heart of oil fields and industrial cities
- Has a large world class campus but small enough to build close relationships
- Disciplines: Engineering, Sciences, Environmental Design, & Industrial Management
- Around 450 faculty members and 8,000 Students (10% Graduate students)
- The QS Times Ranking (2008): 338 out of 500 Top World Universities

Quadratic Forms in Gaussian Variables

- Gaussian variables play a very important role in statistics, signal processing, and communications
- Quadratic forms in Gaussian random variables are of particular importance
- We will characterize the distribution of quadratic forms and apply our findings to
  - Multiuser information theory
  - Mean-square analysis of the normalized LMS adaptive filter

Simplest Case of Quadratic Forms

- $Y = \sum_{i=1}^{M} |H(i)|^2$ is a sum of iid Gaussian random variables
- Y is Chi-square with M degrees of freedom
- We find the pdf of Y using the characteristic function approach

$$
\begin{aligned}
f_y(y) &= f_{|H(1)|^2}(y) * f_{|H(2)|^2}(y) * \cdots * f_{|H(M)|^2}(y) \\
\phi_y(w) &= \phi^M(w) \\
&\rightarrow f_Y(y) \rightarrow F_Y(y)
\end{aligned}
$$

- This gives the pdf of the quadratic form $Y = \|H\|^2$

More Complicated Forms

- Follow same approach to find pdf of a weighted sum

$$Y = \sum_{i=1}^{M} \lambda_i |H(i)|^2 \quad \lambda_i \geq 0$$

- This corresponds to the weighted Euclidean norm

$$\|H\|_\Lambda^2 \triangleq H^* \Lambda H$$

- This is equivalent to finding the CDF of

$$\|H\|_A^2 \triangleq H^* A H$$

where A is positive semi-definite.

- Approach fails when the sum is mixed (not all λ_i's are postive)

What about Indefinite Quadratic Forms

- We tackle the most general problem

$$Y = \|H\|_A^2 \triangleq H^* A H \tag{1}$$

  - A is a general Hermitian matrix
  - H is circularly symmetric Gaussian random variable, i.e. $H \sim C\mathcal{N}(0, \mathbf{R})$.

- We abandon the conventional approach of

$$\text{Charcateristic function} \rightarrow \text{pdf} \rightarrow \text{CDF}$$

CDF is a more Convenient Tool than the pdf

- The CDF is a direct expression of probability
- The pdf can be obtained from the CDF by differentiation
- Moments can be obtained from the CDF using integration by parts
- The CDF expression can be obtained directly without going through the characteristic function.
- The approach is unified: it applies to real/complex, white/correlated, cental/non-central Gaussian r. v.'s
- Can even be extended to non-Gaussian r. v.'s

Evaluating the CDF the Usual Way is Difficult

- Consider the random Hermitian quadratic form

$$Y = \|H\|_A^2 \triangleq H^* A H$$

- H has the pdf $p(H) = \frac{1}{(2\pi)^N} e^{-\|H\|^2}$
- The CDF of Y is defined by

$$\begin{aligned} F_Y(y) &= P\{Y \leq y\} \\ &= \int \cdots \int_{\mathcal{A}} p(H) dH \end{aligned}$$

$\mathcal{A}$ is an area in M multidimensional plane defined by $\|H\|_A^2 \leq y$

- Integral very difficult to evaluate/manipulate

An Alternative Way...

- Express inequality that defines $\mathcal{A}$ in terms of unit step function
- CDF takes the form

$$
\begin{aligned}
F_Y(y) &= \int \cdots \int_{\mathcal{A}} p(H)dH \\
&= \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} p(H)u(y - \|H\|_A^2)dH \\
&= \frac{1}{2\pi^M} \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} e^{-H^*H} u(y - \|H\|_A^2)dH
\end{aligned}
$$

The constraint appears in the integrand and not in the integration limits

- Difficult to deal with the unit step as is
- Replace unit step with its Fourier transform

$$u(x) = \frac{1}{2\pi} \int_{-\infty}^{\infty} \frac{e^{x(j\omega+\beta)}}{j\omega + \beta} d\omega$$

  which is valid for any $\beta > 0$
- So

$$u(y - \|H\|_A^2) = \frac{1}{2\pi} \int_{-\infty}^{\infty} \frac{e^{(y-\|H\|_A^2)(j\omega+\beta)}}{j\omega + \beta} d\omega$$

The CDF as a 1-D Integral

- CDF can be written as $M+1$ integral

$$F_Y(y) = \frac{1}{2\pi^{M+1}} \int_{-\infty}^{\infty} d\omega \frac{e^{y(j\omega+\beta)}}{j\omega+\beta} \int \cdots \int dH e^{-H^*(I+A(j\omega+\beta))H} dH$$

- Inner integral looks like integral of a Gaussian pdf

$$\frac{1}{\pi^M} \int \cdots \int e^{-H^*(I+A(j\omega+\beta))H} dH = \frac{1}{\det(I+A(j\omega+\beta))}$$

- We are left with

$$F_Y(y) = \frac{1}{2\pi} \int_{-\infty}^{\infty} \frac{e^{y(j\omega+\beta)}}{j\omega+\beta} \prod_{i=1}^{M} \frac{1}{1+\lambda_i(j\omega+\beta)} d\omega$$

The CDF in Closed Form

- Use partial fraction expansion and contour integration to write CDF in closed form

$$F_Y(y) = u(y) + \sum_{l=1}^{M} \frac{\lambda_i^{M+1}}{\prod_{l \neq i}(\lambda_i - \lambda_l)} \frac{1}{|\lambda_i|} e^{-\frac{y}{\lambda_l}} u(\frac{y}{\lambda_l})$$

- Note that this result applies irrespective of the correlation of H and irrespective of the weight matrix A

λ_i's are the eigenvalues of A

The Nonzero Mean Problem

- In the nonzero mean problem, we have the quadratic form

$$Y = \|H - a\|_A^2$$

- We can evaluate the CDF up to a 1-D integral

$$Pr\{Y \leq y\} = \frac{1}{2\pi}\int_{-\infty}^{\infty} \frac{e^{y(j\omega+\beta)}}{j\omega+\beta} e^{-c} \frac{1}{det(I + (j\omega+\beta)\Lambda)} d\omega \quad (2)$$

where

$$c = a^*(I + \frac{1}{j\omega+\beta}\Lambda^{-1})^{-1} a$$

- The integral can be written as a real integral.

Distribution of a Ratio of Weighted Norms

- We can use this approach to find the distribution of a ratio of Gaussian norms

$$\Pr\left\{\frac{\epsilon_1 + \|H\|_{B_1}^2}{\epsilon_2 + \|H\|_{B_2}^2} \leq x\right\} = \Pr\left\{\|H\|_{B_1 - xB_2} \leq \epsilon_2 x - \epsilon_1\right\}$$

- Recall the CDF $F_Y(y) = Pr\{\|H\|_A^2 \leq y\}$

$$F_Y(y) = u(y) + \sum_{l=1}^{M} \frac{\lambda_i^{M+1}}{\prod_{l \neq i}(\lambda_i - \lambda_l)} \frac{1}{|\lambda_i|} e^{-\frac{y}{\lambda_l}} u(\frac{y}{\lambda_l})$$

- To get the CDF of the ratio, perform the substitution

$$\begin{aligned} A &\rightarrow B_1 - xB_2 \\ y &\rightarrow \epsilon_2 x - \epsilon_1 \end{aligned}$$

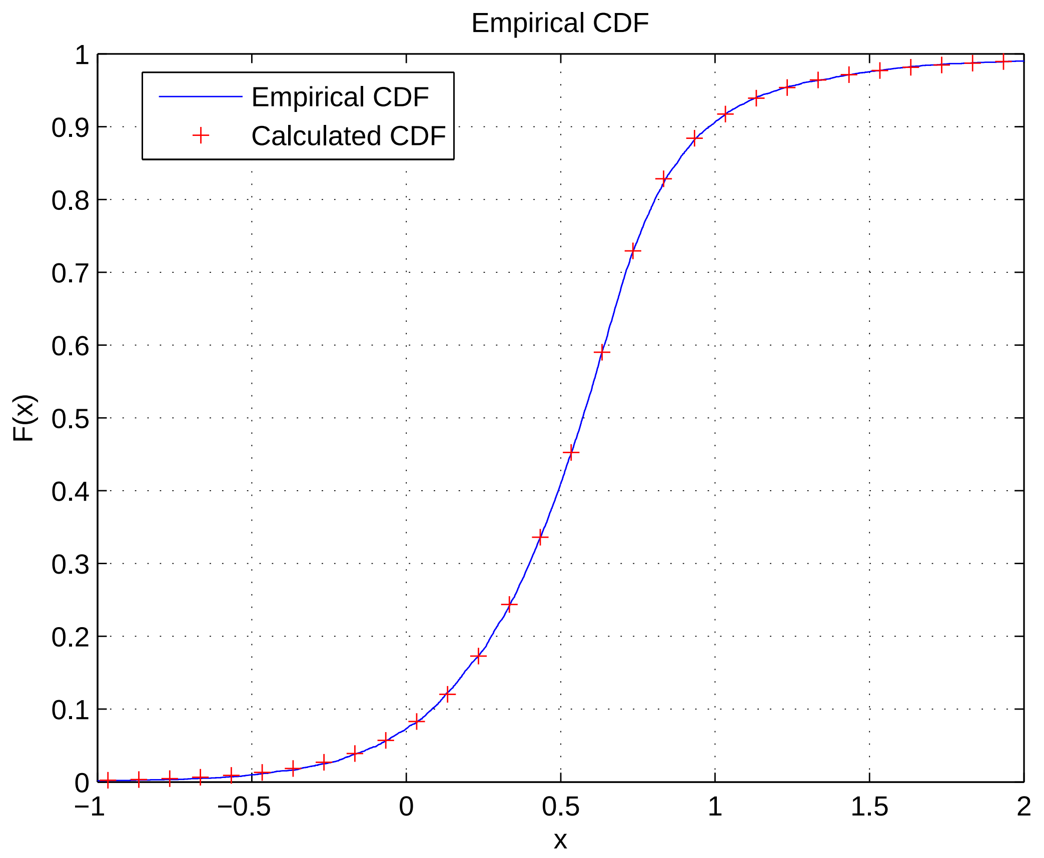


Figure 1: Empirical and calculated CDF's of a ratio of norms

Real Quadratic Forms

- The Hermitian quadratic form (1) is a special case of the real quadratic form

$$Y_r = \|H_r\|^2_{A_r} \triangleq H'_r A_r H_r \tag{3}$$

- We can show that the CDF is given by

$$Pr\{Y \leq y\} = \frac{1}{2\pi} \int_{-\infty}^{\infty} \frac{e^{y(j\omega+\beta)}}{j\omega + \beta} \frac{1}{\sqrt{\det(I + A(j\omega + \beta))}}$$

- We can express this as a real integral
- Can use the same approach to obtain closed form Chernof upper bounds on the CDF

Joint Distribution of Quadratic Forms

- We can use this approach to find the joint distribution of two or more weighted norms

$$F_{X_a,X_b}(x_a, x_b) = \Pr\left\{\|H\|_A^2 \le x_a, \|H\|_B^2 \le x_b\right\}$$

- We replace the two step functions with their F.T.'s

$$F_{X_a,X_b}(x_a, x_b) = \frac{1}{2^2\pi^{M+2}} \int d\omega_1 \frac{e^{x_a(j\omega_1+\beta_1)}}{j\omega_1+\beta_1} \int d\omega_2 \frac{e^{x_b(j\omega_2+\beta_2)}}{j\omega_2+\beta_2} \int dH e^{-\|H\|^2 - \|H\|^2_{(j\omega_1+\beta_1)A} - \|H\|^2_{(j\omega_2+\beta_2)B}}$$

- Boils down to a 2-D integral

$$F_{X_a,X_b}(x_a,x_b) = \frac{1}{2^2\pi^2}\int d\omega_1 \frac{e^{x_a(j\omega_1+\beta_1)}}{j\omega_1+\beta_1}$$
$$\int d\omega_2 \frac{e^{x_b(j\omega_2+\beta_2)}}{j\omega_2+\beta_2}\frac{1}{\det(I+(j\omega_1+\beta_1)A+(j\omega_2+\beta_2)B)}$$

- If A and B are jointly diagonalizable by an orthonormal transform, this can be written in closed form

$$F_{X_a,X_b} = \frac{1}{2^2\pi^2}\int d\omega_1 \frac{e^{x_a(j\omega_1+\beta_1)}}{j\omega_1+\beta_1}$$
$$\int d\omega_2 \frac{e^{x_b(j\omega_2+\beta_2)}}{j\omega_2+\beta_2}\frac{1}{\prod_{i=1}^{M}(I+(j\omega_1+\beta_1)a_i+(j\omega_2+\beta_2)b_i)}$$

Non-Quadratic Forms: Craig's Formula

- Approach can be used to derive Craig's representation of the Q function in a straight-forward manner (instead of resorting to polar coordinates)

$$Q(x) = \frac{1}{\pi}\int_0^{\frac{\pi}{2}} exp(-\frac{x^2}{\sin^2\theta})d\theta \tag{4}$$

$$\begin{aligned} Q(x) &= P\{y < x\} \\ &= \int_{-\infty}^{\infty} e^{-\frac{y^2}{2}} u(x-y)dy \\ &= \frac{1}{2\pi}\int_{-\infty}^{\infty} d\omega \frac{e^{(x-y)(j\omega+\beta)}}{j\omega+\beta} \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\infty} dye^{-\frac{y^2}{2}} \end{aligned}$$

- By choosing β judicially, we arrive at Craig's formula with proper change of variables.

Non-Gaussian Variables

- Can use same approach to evaluate the CDF of quadratic forms in isotropic random variables (or ratios of such forms)

$$R = \|\phi\|_A^2$$

- Isotropic random vectors have the pdf

$$p(\phi) = \frac{\Gamma(M)}{\pi^M}\delta(\|\phi\|^2 - 1) \qquad (5)$$

- The CDF is given by

$$P\{R \leq x\} \quad = \quad \int p(\phi)u(x - \|\phi\|_A^2)d\phi \qquad (6)$$

- To evaluate the delta and unit step functions with their Fourier transforms

$$
\begin{aligned}
p(\phi) &= \frac{\Gamma(M)}{\pi^M}\delta(\|\phi\|^2 - 1)e^{-\alpha\|\phi\|^2-1} \\
&= \frac{\Gamma(M)}{\pi^M}\frac{1}{2\pi}\int_{-\infty}^{\infty} e^{j\omega_2(\|\phi\|^2-1)}e^{-\alpha(\|\phi\|^2-1)}d\omega_2
\end{aligned}
$$

This is valid for all $\alpha > 0$.

- We arrive at the CDF ($\lambda_i' s$ are the eigenvalues of A)

$$
P\{R \leq x\} = \sum_{i=1}^{M} \eta_i (x - \lambda_i)^{M-1} u(x - \lambda_i)
$$

For more details ...

1. T. Y. Al-Naffouri and B. Hassibi, "On the distribution of indefinite Hermitian quadratic forms in Gaussian random variables," *Submitted to International Symposium on Information Theory (ISIT)*

2. T. Y. Al-Naffouri and B. Hassibi, "On the distribution of indefinite Hermitian quadratic forms in Gaussian random variables," *under preparation for submission to IEEE Transactions on Information Theory*

Application I:

Effect of Correlation on the Sum-Rate of Broadcast Channels

An Application to Broadcast Channels

- Multiple antennas add tremendous value to point to point systems
- Research shifted recently to the role of multiple antennas in multiuser systems
  - (Uplink) Multiple Access (MAC)
  - (Downlink) Broadcast (BC)
- Broadcast scenarios (point to multi-point) are especially important because downlink scheduling is the major bottleneck for broadband wireless networks

Three Main Questions in a Broadcast Scenario (1)

Q1) Quantify the maximum sum rate possible to all users

A1) Sum-rate is achieved using dirty paper coding (DPC) (Caire and Shamai '02, Viswanath and Tse '02, Vishwanath et al. '02, Yu and Cioffi '02)

- **(-)** DPC is computationally complex at both Tx and Rx
- **(-)** Requires a great deal of Feedback (CSI for all users at Tx)

Three Main Questions in a Broadcast Scenario (2)

The second question is motivated by the drawbacks of DPC

Q2) Devise computationally efficient algorithms for capturing capacity

A2) Utilize multi-user diversity to achieve performance close to capacity

(+) Opportunist multiple random beamforming coincides asymptotically with DPC (Sharif and Hassibi '06)

$$R = M \log \log n + M \log \frac{P}{M}$$

(+) Requires simply SINR feedback to Tx

Three Main Questions in a Broadcast Scenario (3)

Q3) With this promising performance, how does opportunist beam-forming perform under various non-idealities

A3) **(i)** Time correlation (Kountouris and Gesbert '05)

(ii) Frequency correlation (Fakhereddin, Sharif, and Hassibi'06)

(iii) Channel estimation error (Vikali, Sharif, and Hassibi '06)

(iv) Spatial correlation (D. Park and S Y. Park '05)

Main Problem to be Addressed:

- For a Gaussian broadcast channel, we would like to quantify the hit that transmit correlation causes to scaling laws of the sum-rate capacity. We consider DPC and various beamforming schemes.

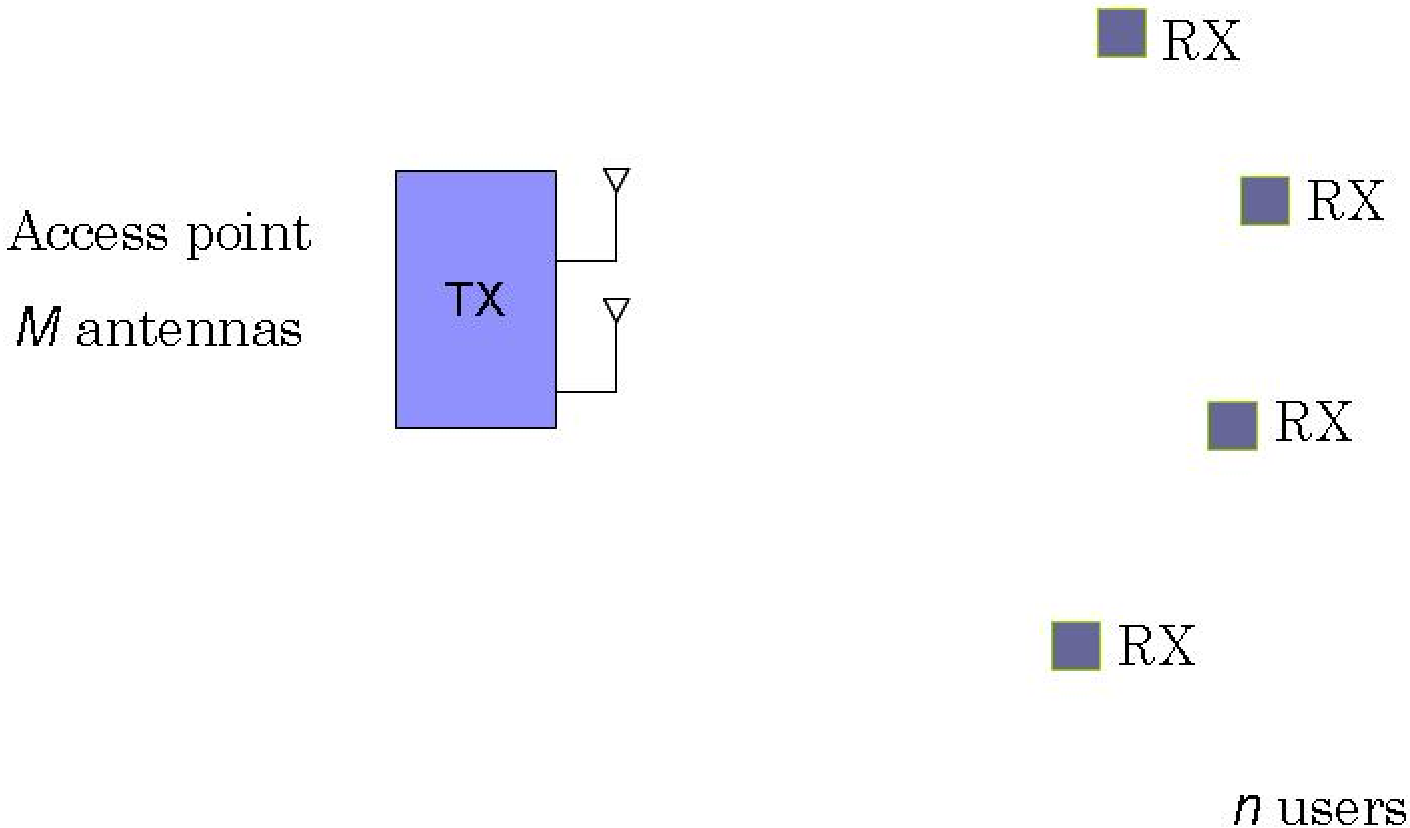
System Model

- Base station with M antennas broadcasting to n single-antenna users
- Received signal at each antenna

$$Y_i = \sqrt{P} H_i S + W_i, \qquad i = 1, \ldots, n$$

  with $E[S^* S] = 1$ and Gaussian noise $W_i \sim CN(0, I)$

- Channel H_i of i-th user is $1 \times M$ vector
  - Distributed as $CN(0, \mathbf{R})$; $\mathbf{R}$ is nonsingular with $\text{tr}(\mathbf{R}) = M$
  - Known perfectly at receiver
  - Follows a bock fading model (with coherence interval T)
  - H_i is independent from one user to another


Access point
M antennas
TX
RX
RX
RX
RX
n users

Scaling of DPC under Correlation

- Rate for spatially uncorrelated channel

$$\boxed{R_{DPC} = M \log \log n + M \log \left(\frac{P}{M}\right)}$$

- Rate in the presence of correction

$$\boxed{R_{DPC} = M \log \log n + M \log \left(\frac{P}{M}\right) + M \log \sqrt[M]{\det \mathbf{R}}}$$

- Since $\mathrm{tr}(R) = M$, the geometric mean satisfies

$$\sqrt[M]{\det(R)} \leq \frac{\mathrm{tr}(R)}{M} \leq 1$$

What is Random Beam Forming?

- Choose M random orthonormal vectors ϕ_m, $m = 1, \ldots, M$ (according to an isotropic distribution)
- Construct the signal

$$S(t) = \sum_{m=1}^{M} \phi_m s_m, \qquad t = 1, \ldots, T$$

where T is less than the coherence interval of the channel.

- After T channel uses we independently choose another isotropic set of orthonormal vectors $\{\phi_m\}$, and so on. So we are transmitting M random beams.

Exploit Multi-User Diversity

- Signal received by the ith user

$$Y_i = \sqrt{P} H_i \phi_1 s_1 + \sqrt{P} H_i \phi_2 s_2 + \cdots + \sqrt{P} H_i \phi_M s_M + W_i$$

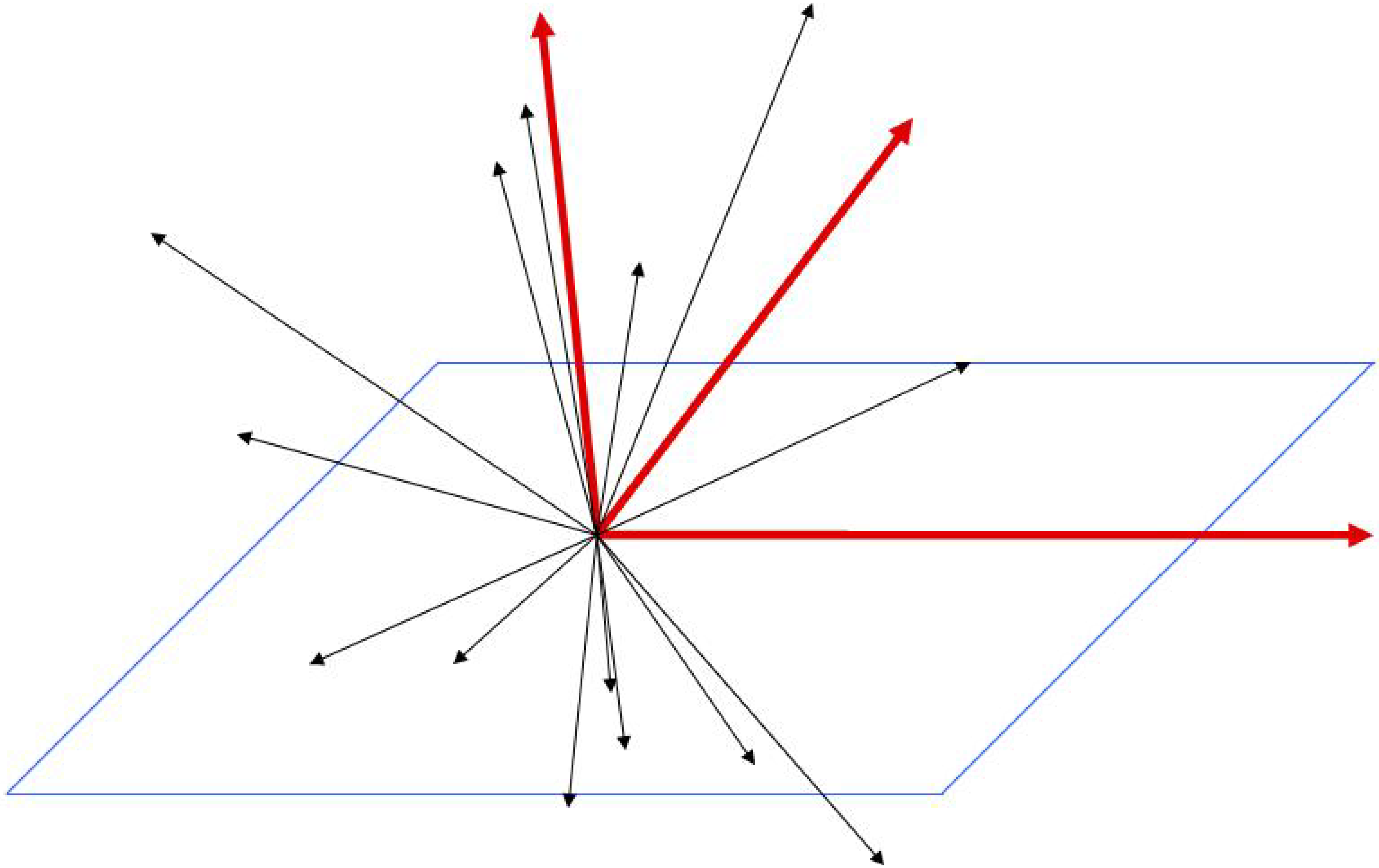
- Each receiver $i = 1, \ldots, n$ computes the following M SINRs

$$\text{SINR}_{i,m} = \frac{|H_i \phi_m|^2}{1/\rho + \sum_{k \neq m} |H_i \phi_k|^2}, \qquad m = 1, \ldots, M$$

  and feeds back the best SINR

- Transmitter assigns signal s_m to the user with the best SINR

$$\begin{aligned} C &= E \sum_{m=1}^{M} \log \left(1 + \max_{i=1,\ldots,n} \text{SINR}_{i,m} \right) \\ &= ME \log \left(1 + \max_{i=1,\ldots,n} \text{SINR}_{i,1} \right) \end{aligned}$$



How to Characterize the Scaling

- Need to find the scaling of $\max_{i=1,\ldots,n} \mathrm{SINR}_{i,1}$ for large n
- An order statistics problem; need to find pdf and CDF of SINR
- SINR is a ratio of two weighted norms

$$SINR_{i,1} = \frac{\|H_i\|^2_{\phi_1 \phi_1^*}}{\frac{1}{\rho} + \|H_i\|^2_{I-\phi_1 \phi_1^*}}$$

Statistics of $\mathrm{SINR}_{i,m}$ (White Channel)

- Easy to find distribution of $\mathrm{SINR}_{i,m}|\Phi$ when H_i is white

$$
\begin{aligned}
f(x) &= \frac{e^{-\frac{x}{\rho}}}{(1+x)^M}\left(\frac{1}{\rho}(1+x)+M-1\right) \\
F(x) &= 1-\frac{e^{-\frac{x}{\rho}}}{(1+x)^M}
\end{aligned}
$$

- Finding these statistics in the correlated case is challenging

Statistics of $\mathrm{SINR}_{i,m}$ Given Φ (Correlated Case)

- We can show that the CDF of SINR in the correlated case

$$\boxed{F(x) = 1 - \frac{1}{2\pi^M \det(R)} \lambda_M \prod_{i=1}^{M-1} \frac{\lambda_i \lambda_M}{x(\lambda_i - \lambda_M)} e^{-\frac{1}{\rho}\frac{x}{\lambda_M}}}$$

where $\lambda_1 \le \lambda_2 \le \cdots \le \lambda_M$ are the eigenvalues of the matrix

$$A = (1+x)\Lambda^{1/2}\phi_m\phi_m^*\Lambda^{1/2} - x\Lambda \qquad \rho = \frac{P}{M}$$

Note that eigenvalues are a function of x.

- pdf is given by

$$f(x) = \frac{1}{2\pi^M \det(R)} e^{-\frac{1}{\rho}\frac{x}{\lambda_M}} \prod_{i=1}^{M-1} \frac{\lambda_i \lambda_M}{x(\lambda_i - \lambda_M)} \times$$
$$\left\{ \frac{1}{\rho} \frac{\|q_M\|_C^2}{\lambda_M} - \|q_M\|_B^2 - \sum_{i=1}^{M} \frac{1}{\lambda_i} \frac{\lambda_M^2 \|q_i\|_C^2 - \lambda_i^2 \|q_M\|_C^2}{x(\lambda_i - \lambda_M)} \right\}$$

where $B = \Lambda^{1/2}(\phi_m\phi_m^* - I)\Lambda^{1/2} \qquad C = \Lambda^{1/2}\phi_m\phi_m^*\Lambda^{1/2}$

Scaling of Maximum SINR

- Can now show

$$\lim_{x\to\infty} \frac{1-F(x)}{f(x)} = \frac{P}{M} \frac{1}{\|\phi_m\|^2_{\Lambda^{-1}}}$$

- Using extreme value theory, we can show that for large n

$$\max_{i=1,\ldots,n} \mathrm{SINR}_{i,m} = \frac{P}{M} \frac{1}{\|\phi_m\|^2_{\Lambda^{-1}}} \log n$$

- Conditional sum-rate capacity scales as

$$R_{\mathrm{BF}|\Phi} = M \log\log n + M \log \frac{P}{M} + M \log \left(\frac{1}{\|\phi_m\|^2_{\Lambda^{-1}}} \right)$$

- Sum-rate capacity of random beam-forming

$$\boxed{R_{\mathrm{RBF}} = M \log\log n + M \log \tfrac{P}{M} + M E_{\Phi} \log \left(\tfrac{1}{\|\phi_m\|^2_{\Lambda^{-1}}} \right)}$$

Sum rate of RBF with Channel Whitening

- For random beam forming with channel whitening,

$$S(t) = \sum_{m=1}^{M} \sqrt{\alpha}\mathbf{R}^{-1/2}\phi_m s_m(t)$$

- Set $\alpha = \frac{\mathrm{tr}(\mathbf{R}^{-1})}{M}$ to guarantee $E[S^*S] \leq 1$
- Scaling becomes the same as for white channel case with reduced signal power

$$\boxed{R_{BF-W} = M \log\log n + M \log \frac{P}{M} + M \log \frac{M}{\mathrm{tr}(\mathbf{R}^{-1})}}$$

Summary: Scaling of Sum-Rate

- White channel

$$R = M \log\log n + M \log \frac{P}{M}$$

- DPC with corelation

$$R_{DPC} = M \log\log n + M \log \frac{P}{M} + M \log \sqrt[M]{\det \mathbf{R}}$$

- Random Beamforming with correlation

$$R_{\mathrm{RBF}} = M \log\log n + M \log \frac{P}{M} + M E_{\Phi} \log \left(\frac{1}{\|\phi_m\|^2_{\Lambda^{-1}}} \right)$$

- Random beamforming with channel whitening

$$R_{BF-W} = M \log\log n + M \log \frac{P}{M} + M \log \frac{M}{\mathrm{tr}(\mathbf{R}^{-1})}$$

Simulations

- Consider a base station with $M = 2$ and $M = 3$ antennas
- The corresponding correlation matrix is

$$\mathbf{R} = \begin{bmatrix} 1 & \alpha \\ \alpha & 1 \end{bmatrix}$$

$$\mathbf{R} = \begin{bmatrix} 1 & \alpha & \alpha^2 \\ \alpha & 1 & \alpha \\ \alpha^2 & \alpha & 1 \end{bmatrix}$$

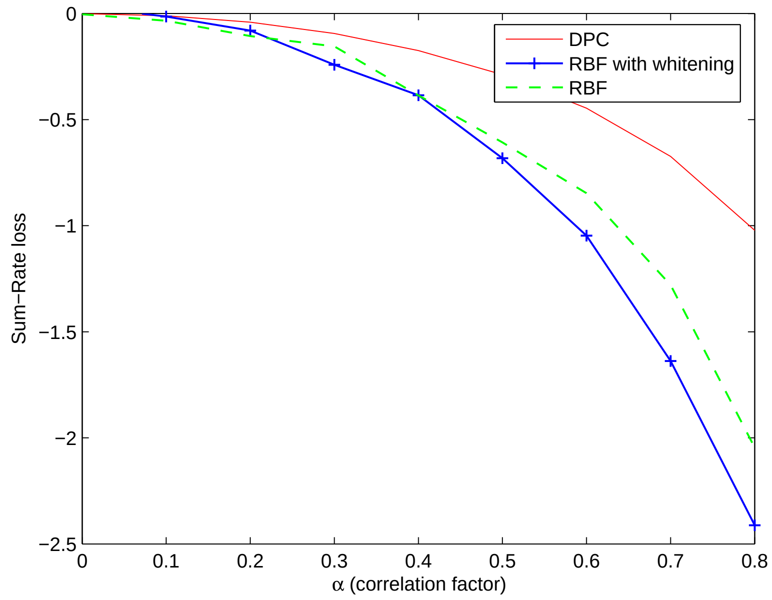


Figure 2: Sum-rate loss versus the correlation factor α for a system with $M = 2$ and $n = 100$.

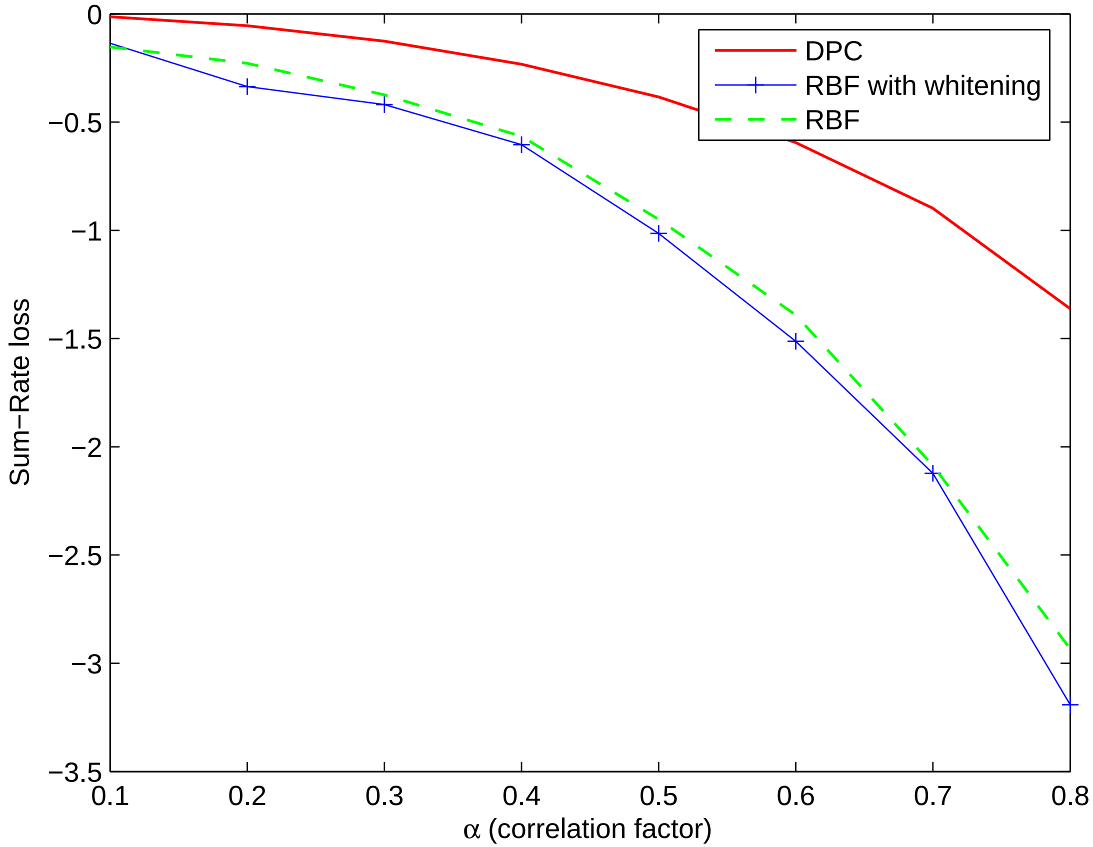


Figure 3: Sum-rate loss versus the correlation factor α for a system with $M = 3$ and $n = 100$.

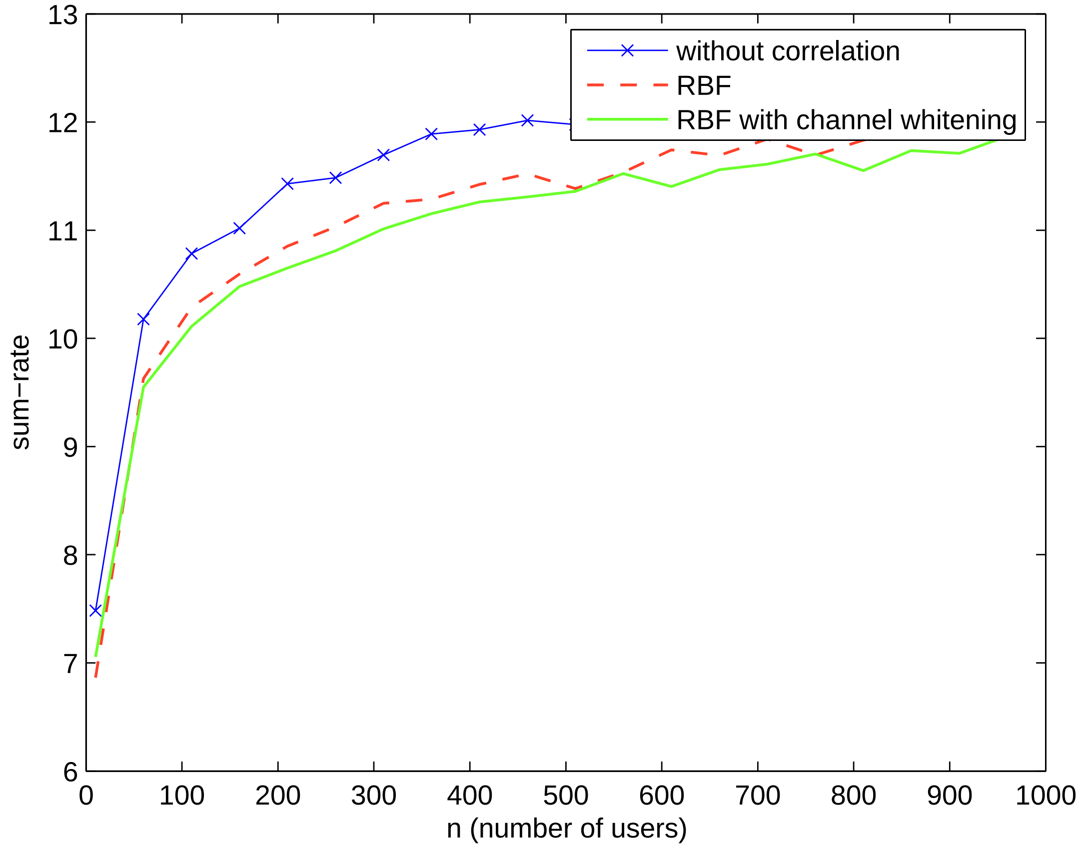


Figure 4: Sum-rate versus the number of users in a system with $M = 2$ and $\alpha = 0.5$

For more Details ...

1. T. Y. Al-Naffouri "Opportunistic beamforming with MMSE precoding for spatially correlated channels," *Accepted in IEEE Communication Letters.*

2. T. Y. Al-Naffouri, M. Sharif, and B. Hassibi " How much does transmit correlation affect the sum-rate of MIMO downlink channels?" *IEEE Transactions on Communications*, no. 2, Feb. 2009.

Application II:

Mean-Square Analysis of Normalized LMS

The Normalized LMS Algorithm

- LMS algorithm has found wide-spread application in control, signal processing, and communication
- It suffers from slow convergence in the presence of input correlation
- NLMS reduces the effect of correlation by normalizing the input by its energy

$$\mathbf{w}_i = \mathbf{w}_{i-1} + \mu \frac{\mathbf{u}_i^*}{\epsilon + ||\mathbf{u}_i||^2} e(i), \qquad i \geq 0$$

$$e(i) = d(i) - \mathbf{u}_i \mathbf{w}_{i-1} = \mathbf{u}_i \mathbf{w}^o - \mathbf{u}_i \mathbf{w}_i + v(i)$$

- Performance determined by the behavior of weight error vector

$$\|\tilde{\mathbf{w}}_i\|^2 = \|\mathbf{w}_i - \mathbf{w}_0\|^2$$

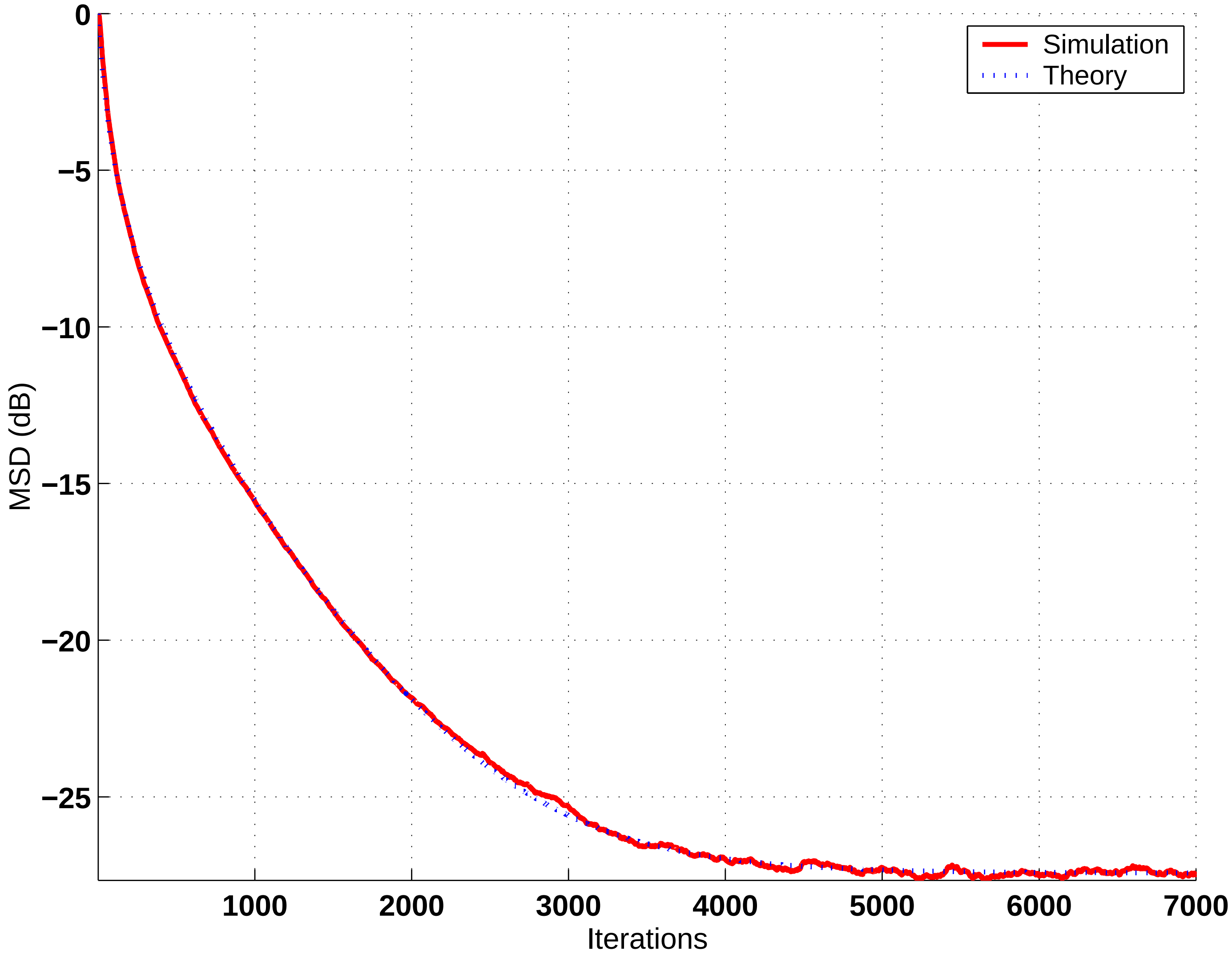


Figure 5: Learning curves of the LMS algorithm

The Normalized LMS Algorithm

- Stability and transient and steady-state behavior are completely determined by input moments
- The matrix $\mathbf{F}$ determines the transient and steady-state preformance

$$\mathbf{F} = \mathbf{I} - \mu\mathbf{A} + \mu^2\mathbf{B}$$

$$\mathbf{A} \triangleq 2E\left[\frac{\mathbf{u}^*\mathbf{u}}{\epsilon + \|\mathbf{u}\|^2}\right] \quad \mathbf{B} = E\left[\frac{(\mathbf{u} \odot \mathbf{u})^* (\mathbf{u} \odot \mathbf{u})}{(\epsilon + \|\mathbf{u}\|^2)^2}\right]$$

- $\mathbf{A}$ and $\mathbf{B}$ are determined by finding 1st and 2nd moments of the r.v.'s

$$r_k = \frac{|u(k)|^2}{\epsilon + ||\mathbf{u}||^2} \quad \text{and} \quad s_{kl} = \frac{|u(k)|^2 + |u(l)|^2}{\epsilon + ||\mathbf{u}||^2}$$

- Moments can found in closed form by finding the CDF of these variables.

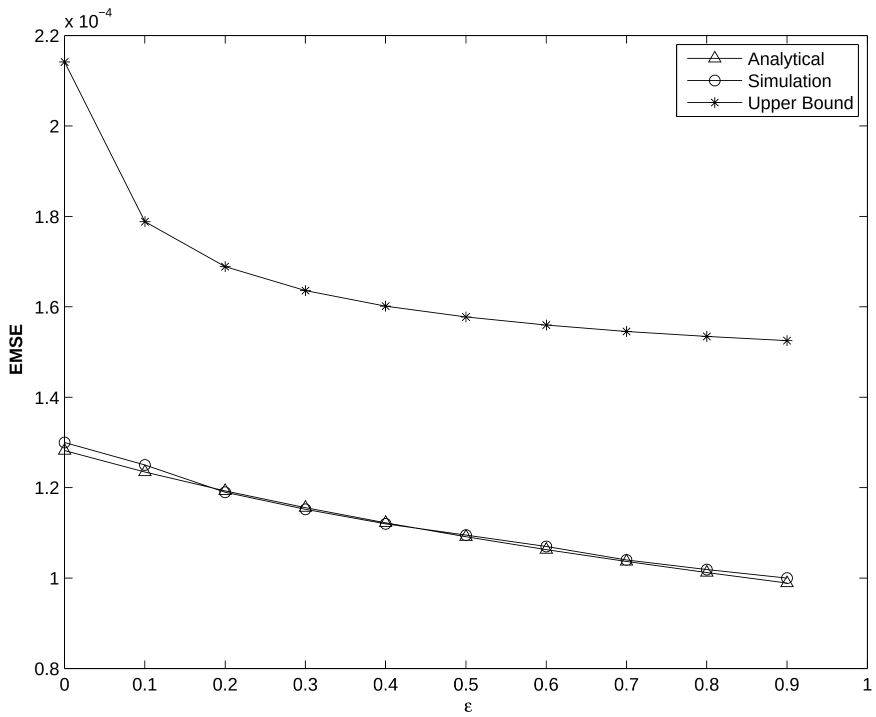


Figure 6: Steady-state EMSE of the ϵ-NLMS w.r.t. ϵ.

For more Details ...

T. Y. Al-Naffouri and M. Moinuddin "Exact mean-square analysis of the (ϵ-) normalized LMS," *submitted to IEEE Transactions on Signal Processing.*

Conclusion

- Presented a new approach for calculating the (joint) distribution of indefinite quadratic forms in (Gaussian) random variables.
- Used these results to study the effect of spatial correlation on various multiuser schemes for MIMO broadcast channels.

$$\boxed{R_{\text{sum}-\text{rate}} = M \log\log n + M \log \tfrac{P}{M} + M \log c}$$

- The constant c depends 1) the scheduling technique used and 2) the eigenvalues of the correlation matrix $\mathbf{R}$
- Used the weighted norms ratio result to evaluate the performance of NLMS in closed form.

Other Research Interests

- Compressive sensing for impulsive noise estimation/cancellation in OFDM (Giuseppe Caire, USC)
- Application of compressive sensing in multiuser information theory (KFUPM students)
- Group Broadcast Channels (Amir Dana and Babak Hassibi, Cal Tech)
- Adaptive filtering analysis and design (Vitor Nascimento; previously with Ali Sayed, UCLA)
- Receiver design for (MIMO) OFDM in (block) time-variant channels (Naofal Al-Dhahir U. T. Dallas) (previously with A. Paulraj, Stanford)
- Blind channel estimation (KFUPM Students)

Can We Do Better?

- Apply a general precoding matrix

$$\alpha AS(t) = \alpha A \sum_{m=1}^{M} \phi_m(t) s_m(t), \qquad\qquad t = 1, \ldots, T$$

- The factor α ensures that we have a fixed power constraint

$$\alpha \leq \sqrt{\frac{M}{\operatorname{tr}(A^* A)}}$$

- This produces the effective channel

$$\tilde{H}_i = \alpha H_i A$$

with correlation $\alpha^2 \tilde{R} = \alpha^2 A^* R A$.

What is the Sum-Rate with a General Precoding?

- Sum-rate is given by

$$R_{\mathrm{PC}} = M \log \log n + M \log \frac{P}{M} - h_{\mathrm{PC}} \tag{7}$$

where h_{PC} is the hit incurred by using a general precoding matrix A

$$\boxed{h_{\mathrm{PC}} = M \log \frac{\mathrm{tr}(A^* A)}{M} + ME \log \|\phi_m\|^2_{\tilde{R}^{-1}}} \tag{8}$$

- Finding the optimum A that minimizes hit is difficult. But we can show that optimum precoding matrix A_{opt} can be written as

$$A_{opt} = Q_{Aopt} D_{Aopt}$$

where Q_{Aopt} is an orthornormal matrix and D_{opt} is diag with positive entries.

Special choices of A

- Difficult to optimize Q_{opt} and D_{opt} jointly.
- Set $Q_{opt} = Q_R$ as this will diagonalize R and optimize over D_{opt}.
- Zero forcing

$$A_{\mathrm{ZF}} = Q_R \Lambda_R^{-\frac{1}{2}}$$

and resulting hit

$$h_{ZF} = M \log \frac{tr(R^{-1})}{M}$$

Special Choices of A

- MMSE precoding

$$A = Q_R(\Lambda + \beta I)^{-\frac{1}{2}}$$

with β obtained as a solution to a fixed-pt problem

$$\frac{Tr(\Lambda + \beta^* I)^{-2}}{Tr(\Lambda + \beta^* I)^{-1}} = E\left(\frac{1}{\beta^* + \frac{1}{\|\Phi_m\|^2_{\Lambda^{-1}}}}\right)$$

- More generally, we can set $Q_{opt} = Q_R$ and find the optimum D_{opt}. Need to solve a set of M implicit equations

$$\frac{1}{d_i} E\left[\frac{\frac{1}{d_i \lambda_i}|\phi(i)|^2}{\|\phi\|^2_{D^{-1}\Lambda^{-1}}}\right] = \frac{1}{tr(D)}$$

Minimize an Upper Bound Instead

- Difficulty in minimizing h_{PC} due to the *phi* term

$$h_{\mathrm{PC}} = M \log \frac{\mathrm{tr}(A^* A)}{M} + M E \log \|\phi_m\|^2_{\tilde{R}^{-1}}$$

- Minimize an upper bound

$$h \leq M \log \mathrm{tr}(A^* A) + M \log \mathrm{tr}((A^* R A)^{-1})$$

- Can show that optimum A in this case is

$$\boxed{A = Q_R \Lambda_R^{-1/4}}$$

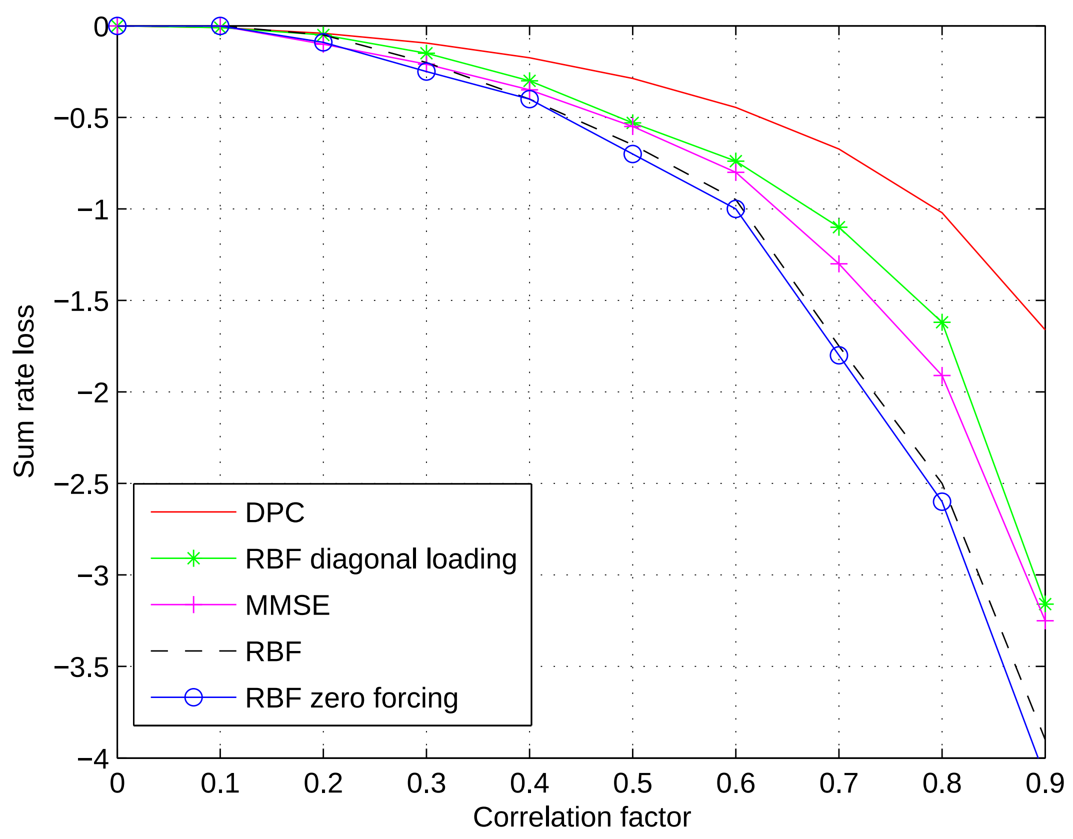


Figure 7: Sum-rate loss versus correlation factor α for a system with M = 3, P=10 and n = 200.