Multiple Antenna Transmitter Optimization Schemes for Multiuser Systems

Dragan Samardzija
Wireless Research Laboratory,
Bell Labs, Lucent Technologies,
791 Holmdel-Keyport Road,
Holmdel, NJ 07733, USA
Email: dragan@bell-labs.com

Narayan Mandayam
WINLAB, Rutgers University,
73 Brett Road,
Piscataway, NJ 08854, USA
Email: narayan@winlab.rutgers.edu

***Abstract*— In this paper we present multiple antenna transmitter optimization (i.e., spatial pre-filtering) schemes that are based on linear transformations and transmit power optimization (keeping average transmit power conserved). We consider the downlink of a wireless system with multiple transmit antennas at the base station and a number of mobile terminals (i.e., users) each with a single receive antenna. We consider maximum achievable data rates in the case of the zero-forcing and triangularization spatial pre-filtering coupled with dirty paper coding transmission scheme. We also present the effects of channel mismatch.**

I. Introduction

There is ever increasing need for higher data rates and greater number of users to be supported by wireless networks. To address this need, the latest 3G cellular wireless standards are increasing the spectral efficiency of the already deployed 2G wireless networks. For example by employing more advanced modulations, channel coding, variable orthogonal spreading, fast power control with fast feedback, and soft hand-offs. These approaches are some of important improvements introduced in the 3G wireless standards primarily to support wireless telephony (e.g., UMTS [1] and cdma2000 [2]). Solutions optimized for data-centric wireless communications are resulting in even higher spectral efficiency. For example, the HSDPA (high speed downlink packet access, part of the 3GPP group) and its equivalent EV-DO/DV [3], [4](evolution data only / data and voice, part of the 3GPP2 group) are applying a slotted scheduled downlink packet based access, adaptive modulation and rate matching that are exploiting time variations in channel quality among multiple users. In addition, fast retransmissions, Chase combining and incremental redundancy schemes are further improving the efficiency of those systems. The above solutions have improved the spectral efficiency of the 3G networks two to three times versus the existing 2G networks (i.e., approximately, 1.5bps/Hz versus 0.5 bps/Hz).

Assuming spectrum as a limited resource, application of multiple antenna systems appear to be one of the most promising solutions leading to even higher data rates and/or the ability to support greater number of users. Multiple-transmit multiple-receive antenna systems represent an implementation of the MIMO (multiple input multiple output) concept in wireless communications [5]. This particular multiple antenna architecture provides high capacity (i.e., spectral efficiency) wireless communications in rich scattering environments. It has been shown that the theoretical capacity (approximately) increases linearly as the number of antennas is increased [5], [6].

Recent studies are focusing on multiple antenna systems with multiple users (see [7] and references therein). In this paper we study multiple antenna transmitter optimization (i.e, spatial pre-filtering) schemes that are based on linear transformations and transmit power optimization (keeping average transmit power conserved). Perfect knowledge of the channel is assumed at the transmitter. We consider the downlink of a wireless system with multiple transmit antennas at the base station and a number of mobile terminals (i.e., users) each with a single receive antenna. The downlink corresponds to information theoretical definition of a *broadcast channel* [8]. We consider maximum achievable data rates in the case of the zero-forcing, modified zero-forcing and triangularization spatial pre-filtering coupled with dirty paper coding transmission scheme. Furthermore, we present the effects of channel mismatch.

II. System Model and Transmitter Optimization Schemes

In the following we introduce the system model. We use a MIMO model [5] that corresponds to a system presented in Figure 1. It consists of M transmit antennas and N mobile terminals.

In Figure 1, x_n is the information bearing signal intended for mobile terminal n and y_n is the received signal at the corresponding terminal (for $n = 1, \cdots, N$). x_n are assumed to be circularly symmetric complex random variables having Gaussin distribution $\mathcal{N}(0, P_{av})$. Further, the received vector $\mathbf{y} = [y_1, \cdots, y_N]^{\mathrm{T}}$ is

$$\begin{aligned} \mathbf{y} &= \mathbf{HSx} + \mathbf{n}, \\ \mathbf{y} &\in \mathcal{C}^N, \mathbf{x} \in \mathcal{C}^N, \mathbf{n} \in \mathcal{C}^N, \mathbf{S} \in \mathcal{C}^{M\times N}, \mathbf{H} \in \mathcal{C}^{N\times M} \end{aligned} \quad (1)$$

where $\mathbf{x} = [x_1, \cdots, x_N]^{\mathrm{T}}$ is the transmitted vector ($\mathrm{E}[\mathbf{xx}^{\mathrm{H}}] = P_{av}\ \mathbf{I}_{N\times N}$), $\mathbf{n}$ is AWGN ($\mathrm{E}[\mathbf{nn}^{\mathrm{H}}] = N_0\ \mathbf{I}_{N\times N}$), $\mathbf{H}$ is the MIMO channel response matrix, and $\mathbf{S}$ is a transformation

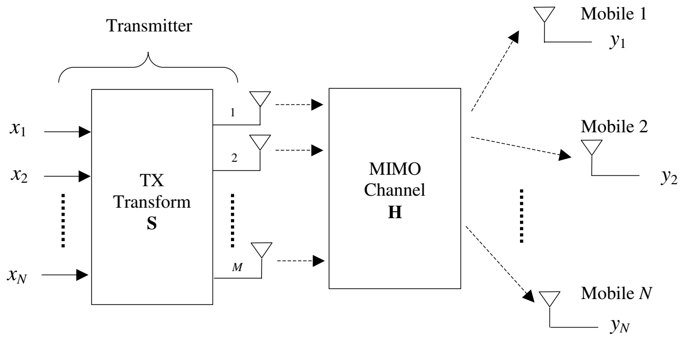

Fig. 1. System model consisting of M transmit antennas and N mobile terminals.

(spatial pre-filtering) performed at the transmitter. Note that the vectors $\mathbf{x}$ and $\mathbf{y}$ have the same dimensionality. Further, h_{nm} is the n-th row and m-th column element of the matrix $\mathbf{H}$ corresponding to a channel between mobile terminal n and transmit antenna m.

Application of the spatial pre-filtering results in the composite MIMO channel (true MIMO channel $\mathbf{H}$ multiplied by the spatial pre-filter $\mathbf{S}$)

$$\mathbf{G} = \mathbf{HS}, \quad \mathbf{G} \in \mathcal{C}^{N\times N} \tag{2}$$

where g_{nm} is the n-th row and m-th column element of the composite MIMO channel response matrix $\mathbf{G}$. The signal received at the n-th mobile terminal is

$$y_n = \underbrace{g_{nn}x_n}_{\text{Desired signal for user } n} + \underbrace{\sum_{i=1, i\neq n}^{N} g_{ni}x_i}_{\text{Interference}} + n_n. \tag{3}$$

In the above representation, the interference is the signal that is intended for other mobile terminals than terminal n. As said earlier, the matrix $\mathbf{S}$ is a spatial pre-filter at the transmitter. It is determined based on optimization criteria that we address later in the text and has to satisfy the following constraint

$$\text{trace}\left(\mathbf{SS}^{\text{H}}\right) \leq N \tag{4}$$

keeping the average transmit power conserved. As a further simplification, we restrict the matrix $\mathbf{S}$ to be

$$\mathbf{S} = \mathbf{AP}, \quad \mathbf{A} \in \mathcal{C}^{M\times N}, \mathbf{P} \in \mathcal{C}^{N\times N} \tag{5}$$

where $\mathbf{A}$ is a linear transformation and $\mathbf{P}$ is a diagonal matrix. $\mathbf{P}$ is determined such that the transmit power remains conserved. Considering different forms of the matrix $\mathbf{A}$ we study the following solutions.

1) Zero-forcing (ZF) scheme where $\mathbf{A}$ is represented by

$$\mathbf{A} = \mathbf{H}^{\text{H}}(\mathbf{HH}^{\text{H}})^{-1}. \tag{6}$$

As can be seen, the above linear transformation is zeroing the interference between the signals dedicated to different mobile terminals since after the transforming the channel $\mathbf{H}$, $\mathbf{HA} = \mathbf{I}_{N\times N}$. Consequently, the maximum achievable data rate (capacity) for mobile terminal n is

$$R_n^{\text{ZF}} = \log_2\left(1 + \frac{P_{av}|p_{nn}|^2}{N_0}\right) \tag{7}$$

where p_{nn} is the n-th diagonal element of the matrix $\mathbf{P}$ defined in (5). In (6) it is assumed that $\mathbf{HH}^{\text{H}}$ is invertible, this is when the rows of $\mathbf{H}$ are linearly independent.

2) Modified zero-forcing (MZF) scheme that assumes

$$\mathbf{A} = \mathbf{H}^{\text{H}}(\mathbf{HH}^{\text{H}} + \frac{N_0}{P_{av}}\mathbf{I})^{-1}. \tag{8}$$

The above transformation appears to be of a similar form like a MMSE linear receiver. The important difference is that the transformation is performed at the transmitter. In the case of the above transformation, in addition to the knowledge of the channel $\mathbf{H}$ the transmitter has to know the noise variance N_0. The maximum achievable data rate (capacity) for mobile terminal n now becomes

$$R_n^{\text{MZF}} = \log_2\left(1 + \frac{P_{av}|g_{nn}|^2}{P_{av}\sum_{i=1,i\neq n}^{N}|g_{ni}|^2 + N_0}\right). \tag{9}$$

3) Triangularization that permits dirty paper coding where the matrix $\mathbf{A}$ assumes the form

$$\mathbf{A} = \mathbf{H}^{\text{H}}\mathbf{R}^{-1} \tag{10}$$

where $\mathbf{H} = (\mathbf{QR})^{\text{H}}$ and $\mathbf{Q}$ is unitary and $\mathbf{R}$ is upper triangular (see [9] for details on *QR factorization*). In general, $\mathbf{R}^{-1}$ is a pseudo inverse. The composite MIMO channel $\mathbf{G}$ in (2) becomes $\mathbf{G} = \mathbf{L} = \mathbf{HS}$ a lower triangular matrix. It permits application of dirty paper coding (DPC) designed for single input single output systems. Check [10], [11] for details on the DPC schemes.

By applying the transformation represented in (10), the signal intended for terminal 1 is received without interference. The signal at terminal 2 suffers from the interference arising from the signal dedicated to terminal 1. In general, the signal at terminal n suffers from the interference arising from the signals dedicated to terminals 1 to $n-1$. In other words,

$$\begin{aligned} y_1 &= g_{11}x_1 + n_1, \\ y_2 &= g_{22}x_2 + g_{21}x_1 + n_2, \\ &\vdots \\ y_n &= g_{nn}x_n + \sum_{i=1}^{n-1} g_{ni}x_i + n_n, \\ &\vdots \\ y_N &= g_{NN}x_N + \sum_{i=1}^{N-1} g_{Ni}x_i + n_N. \end{aligned} \tag{11}$$

The interference is known at the transmitter, therefore DPC is applied to mitigate the interference. Based on the results in [10], the achievable rate for mobile terminal n is

$$\begin{aligned} R_n^{\mathrm{DPC}} &= \log_2\left(1+\frac{P_{av}|g_{nn}|^2}{N_0}\right) = \\ &= \log_2\left(1+\frac{P_{av}|r_{nn}p_{nn}|^2}{N_0}\right) \end{aligned} \tag{12}$$

where r_{nn} is the n-th diagonal element of the matrix $\mathbf{R}$ defined in (10). Note that DPC is applied just in the case of the linear transformation in (10), with corresponding rate given in (12).

Note that $\mathrm{trace}(\mathbf{A}\mathbf{A}^{\mathrm{H}}) = N$ satisfying the constraint in (4). Consequently, we can select $\mathbf{P} = \mathbf{I}_{N\times N}$ and present the following proposition.

Proposition 1: For high SNR ($P_{av} \gg N_0$) the achievable sum rate of the triangularization and DPC scheme is equal to the rate of the equivalent (open loop) MIMO system. In other words, for $P_{av} \gg N_0$

$$\sum_{n=1}^{N} R_n^{\mathrm{DPC}} = \log_2\left(\det\left(\mathbf{I}_{N\times N} + \frac{P_{av}}{N_0}\mathbf{H}\mathbf{H}^{\mathrm{H}}\right)\right). \tag{13}$$

Proof: Starting from right side term in (13) and with $\mathbf{H}\mathbf{H}^{\mathrm{H}} = \mathbf{R}^{\mathrm{H}}\mathbf{R}$, for $P_{av} \gg N_0$

$$\begin{aligned} & \log_2\left(\det\left(\mathbf{I}_{N\times N} + \frac{P_{av}}{N_0}\mathbf{H}\mathbf{H}^{\mathrm{H}}\right)\right) \approx \\ \approx\ & \log_2\left(\det\left(\frac{P_{av}}{N_0}\mathbf{H}\mathbf{H}^{\mathrm{H}}\right)\right) = \\ =\ & \log_2\left(\frac{P_{av}}{N_0}|r_{11}|^2 \cdots \frac{P_{av}}{N_0}|r_{NN}|^2\right) = \\ =\ & \sum_{i=1}^{N}\log_2\left(\frac{P_{av}}{N_0}|r_{ii}|^2\right) \approx \\ \approx\ & \sum_{i=1}^{N}\log_2\left(1+\frac{P_{av}}{N_0}|r_{ii}|^2\right) = \\ =\ & \sum_{n=1}^{N} R_n^{\mathrm{DPC}} \end{aligned} \tag{14}$$

which concludes the proof.

Based on [12] we describe the DPC scheme. The transmitted signal intended for terminal n is

$$x_n = f_k\left(\hat{x}_n - \alpha_n\, I_n + t_n\right) \tag{15}$$

where $f_k(.)$ is a modulo operation over a k-dimensional region. Further, $\hat{x}_n$ is the information bearing signal for terminal n, and α_n is a parameter to be optimized ($0 < \alpha_n \le 1$). I_n is the interference $I_n = \sum_{i=1}^{n-1} g_{ni}x_i/g_{nn}$ and t_n is a dither (uniformly distributed pseudo noise over the k-dimensional region). At terminal n the following operation is performed,

$$f_k\left(y_n/g_{nn}\right) = \hat{x}_n + (1-\alpha_n)u_n + \alpha_n n_n^* \tag{16}$$

where n_n^* is a wrapped-around AWGN (due to the nonlinear operation $f_k(.)$) and u_n is uniformly distributed over the k-dimensional region. For $k \to \infty$ and $\hat{x}_n$ being uniformly distributed over the k-dimensional region, the rate in (12) can be achieved [12].

One practical, but suboptimal single-dimensional DPC solution is described in [12], [13]. The transmitted signal intended for terminal n is

$$x_n = f_{mod}\left(\hat{x}_n - I_n\right) \tag{17}$$

where $f_{mod}(.)$ is a modulo operation (i.e., a uniform scalar quantizer). At terminal n the following operation is performed

$$f_{mod}\left(y_n/g_{nn}\right) = \hat{x}_n + n_n^* \tag{18}$$

where n_n^* is a wrapped-around AWGN (due to the nonlinear operation $f_{mod}(.)$). For high SNR and with $\hat{x}_n$ being uniformly distributed over the single-dimensional region, the achievable rate is approximately 1.53dB away from the rate in (12) [12], [13].

Once the matrix $\mathbf{A}$ is selected, elements of the diagonal matrix $\mathbf{P}$ are determined such that the transmit power remains conserved and the sum rate is maximized. Constraint on the transmit power is

$$s_{m1}^2 + \cdots + s_{mN}^2 \le \frac{N}{M}, \quad m = 1, \cdots, M, \tag{19}$$

where s_{mn} ($m = 1, \cdots, M$ and $n = 1, \cdots, N$) is the element of the matrix $\mathbf{S}$ (defined in (1)). The above condition satisfies the constraint in (4). Actually it is even a stronger constraint because it limits the average transmit power per each antenna, individually. The elements of the matrix $\mathbf{P}$ are selected such that

$$\mathrm{diag}(\mathbf{P}) = [p_{11}, \cdots, p_{NN}]^{\mathrm{T}} = \arg \max_{\text{constraint in (19)}} \sum_{i=1}^{N} R_n. \tag{20}$$

III. Numerical Results

To evaluate the performance of the schemes described in Section II, we consider the following base line solutions.

1) No pre-filtering solution where each mobile terminal is served by one transmit antenna dedicated to that mobile. This is equivalent to $\mathbf{S} = \mathbf{I}$. A transmit antenna is assigned to a particular terminal corresponding to the best channel (maximum channel magnitude) among all available transmit antennas and that terminal.
2) Equal resource TDMA and coherent beam-forming (denoted as TDMA-CBF) is a solution where signals for

different terminals are sent in different (isolated) time slots. In this case, there is no interference, and each terminal is using $1/N$ of the over all resources. When serving a particular mobile ideal coherent beam forming is applied using all M transmit antennas.

3) Closed loop MIMO (using the water pouring optimization on eigen modes) is a solution that is used as an upper bound on achievable sum rates. In the following it is denoted as CL-MIMO. This solution would require that multiple terminals act as a joint multiple antenna receiver. This solution is not practical because the terminals are assumed to be individual entities in the network and they do not cooperate when receiving signals on the downlink.

In Figure 2 we present average rates per user for a system consisting of $M = 3$ transmit antennas and $N = 3$ terminals. The channel is Rayleigh, i.e., the elements of the matrix $\mathbf{H}$ are independent identically distribute Gaussian random variables with distribution $\mathcal{N}(0, 1)$. From the figure we observe the following. The triangularization and DPC scheme is approaching the closed loop MIMO rates for higher SNR. The MZF solution is performing very well for lower SNRs (approaching CL-MIMO and DPC rates), while for higer SNRs the rates for the ZF scheme are converging to the MZF rates. The TDMA-CBF rates are increasing with SNR, but still significantly lower than the rates of the proposed optimization schemes. The solution where no pre-filtering is applied clearly exhibits properties of an interference limited system (i.e., after a certain SNR, the rates are not increasing). Corresponding cumulative distribution functions (cdf) of the rates are given in Figure 3, for $SNR = 10$dB, per user (see more on the "capacity versus outage" approach in [14]).

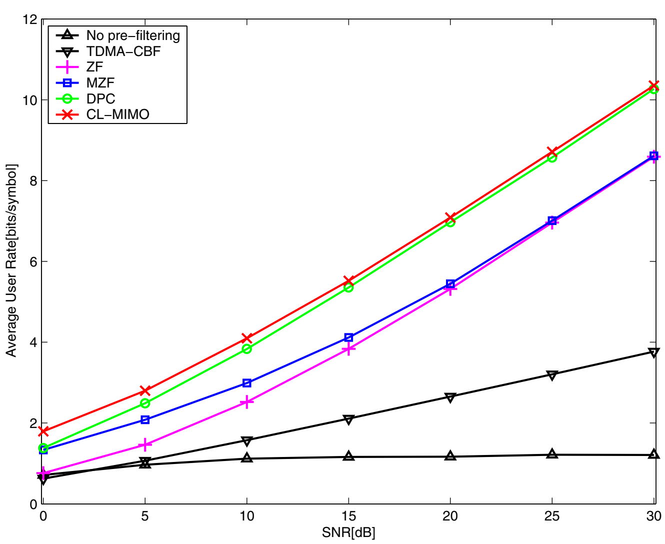

Fig. 2. Average rate per user vs. SNR, $M = 3$, $N = 3$, Rayleigh 3×3 channel.

In Figure 4 we present the behaviour of the average rates per user vs. number of transmit antennas. The average rates are observed for $SNR = 10$dB, $N = 3$, and variable number

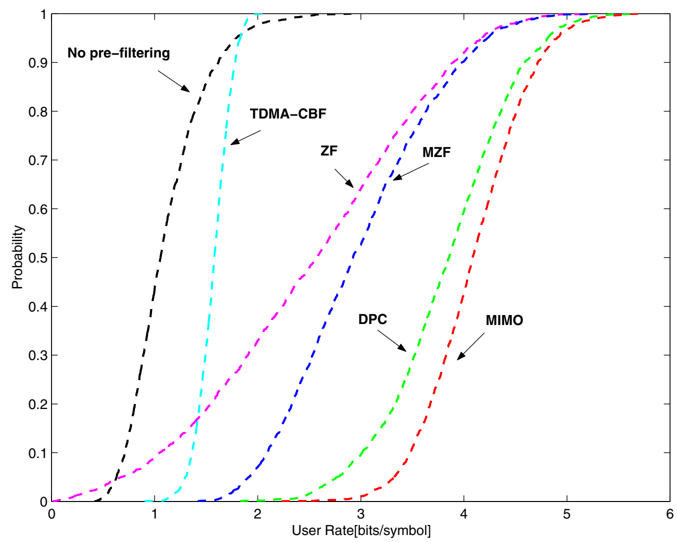

Fig. 3. CDF of rates, $SNR = 10$dB, per user, $M = 3$, $N = 3$, Rayleigh 3×3 channel.

of transmit antenna $M = 3, 6, 12, 24$. The Rayleigh channel is again assumed. With the number of transmit antennas increasing the rates are increasing, and the difference between the rates for different schemes is becoming smaller. This particular case speaks in favor of the simplest ZF scheme if higher number of transmit antennas can be applied.

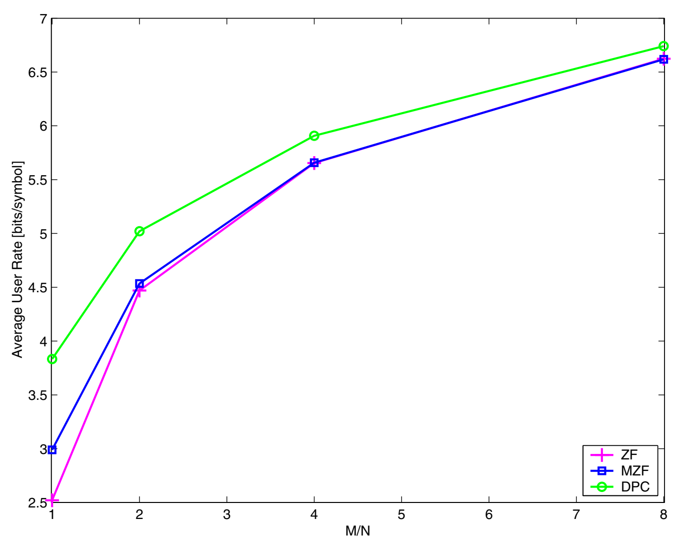

Fig. 4. Average rate per user vs. M/N, $SNR = 10$dB, $N = 3$, variable number of transmit antenna $M = 3, 6, 12, 24$, Rayleigh $M \times 3$ channel.

Let us now analyze the effects of imperfect channel state knowledge. In practical communication systems, the channel state $\mathbf{H}$ has to be estimated at the receiver, and then fed to the transmitter. In the case of a time varying channel, this practical procedure results in noisy and delayed (temporally mismatched) estimates being available to the transmitter to perform the optimization.

As said earlier, the MIMO channel is time varying. Let $\mathbf{H}_{i-1}$ and $\mathbf{H}_i$ correspond to consecutive block faded channel responses. In the following, the subscripts i and $i-1$ on different variables will indicate values corresponding to the block channel periods i and $i-1$, respectively. The temporal characteristic of the channel is described using the correlation

$$k = \mathrm{E}\left[h_{(i-1)nm}\, h^*_{inm}\right]/\Gamma \quad (21)$$

where $\Gamma = \mathrm{E}[h_{inm}h^*_{inm}]$, and h_{inm} is a stationary random process (for $m = 1, \cdots, M$ and $n = 1, \cdots, N$, denoting transmit and receive antenna indices, respectively). We assume that the knowledge of the correlation k is not known at the receiver and the transmitter. Low values of the correlation k correspond to higher mismatch between $\mathbf{H}_{i-1}$ and $\mathbf{H}_i$. Note that the above channel is modeled as a first order discrete Markov process. In the case of Jake's model, $k = J_0(2\pi f_d \tau)$, where f_d is the maximum Doppler frequency and τ is the time difference between $h_{(i-1)nm}$ and h_{inm}.

We assume that the receiver feeds back $\mathbf{H}_{i-1}$. Because the ideal channel state $\mathbf{H}_i$ is not available at the transmitter, we assume that $\mathbf{H}_{i-1}$ is used instead to perform the transmitter optimization for the i-th block. In other words the transmitter is ignoring the fact that $\mathbf{H}_i \neq \mathbf{H}_{i-1}$. In Figure 5 we present cdf of rates for the ZF scheme for different correlation $k = 0.5, 0.8, 0.99, 1$. From these results we note very high sensitivity of the scheme to the channel mismatch. Similar results can be shown for other two schemes presented in Section II. See [15] on a related study of channel mismatch and achievable data rates for single user MIMO systems.

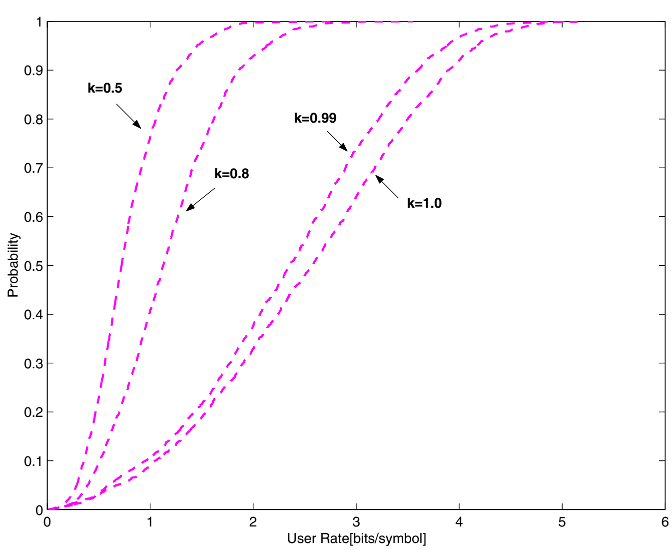

Fig. 5. CDF of rates for different correlation $k = 0.5, 0.8, 0.99, 1$, $SNR =$ 10dB, $M = 3$, $N = 3$, Rayleigh 3×3 channel.

IV. Conclusion

We have presented a study on multiple antenna transmitter optimization (i.e, spatial pre-filtering) schemes that are based on linear transformations and transmit power optimization. We have considered maximum achievable data rates in the case of the zero-forcing, modified zero-forcing and triangularization spatial pre-filtering coupled with the dirty paper coding transmission scheme. We have shown that the triangularization and DPC scheme is approaching the closed loop MIMO rates (upper bound) for higher SNR. Further, the MZF solution is performing very well for lower SNRs (approaching CL-MIMO and DPC rates), while for higer SNRs the rates for the ZF scheme are converging to the MZF rates. In addition, we have presented how the average rates depend on number of transmit antennas, while keeping the number of terminals constant. With the number of transmit antennas increasing the rates are increasing, and the difference between the rates for different schemes is getting smaller. Furthermore, we have presented the effects of the channel mismatch and have shown very high sensitivity of the schemes to the channel mismatch. Channel state information and transmitter optimization schemes is the subject of our future studies.

Acknowledgment

The authors would like to thank Dr. Jack Salz for his valuable ideas, guidance and discussions.

References

[1] H. Holma and A. Toskala, eds., *WCDMA for UMTS*. Wiley, first ed., 2000.

[2] TIA/EIA/IS-2002.2, "Physical Layer Standard for cdma2000 Spread Spectrum System,"

[3] TIA/EIA/IS-856, "CDMA200, High Rate Packet Data Air Interface Specification,"

[4] Q. Bi and S. Vitebsky, "Performance Analysis of 3G-1X EVDO High Data Rate System," *Wireless Communications and Networking Conference*, vol. 1, pp. 389 –395, March 2002.

[5] G. J. Foschini, "Layered Space-Time Architecture for Wireless Communication in a Fading Environment When Using Multiple Antennas," *Bell Labs Technical Journal*, vol. 1, no. 2, pp. 41–59, 1996.

[6] G. J. Foschini and M. J. Gans, "On Limits of Wireless Communications in a Fading Environment when Using Multiple Antennas," *Wireless Personal Communications*, no. 6, pp. 315–335, 1998.

[7] S. Vishwanath, G. Kramer, S. Shamai, S. Jafar, and A. Goldsmith, "Capacity Bounds for Gaussian Vector Broadcast Channel," *Technical Memorandum, Bell Laboratories, Lucent Technologies*, May 2003.

[8] T. M. Cover and J. A. Thomas, *Elements of Information Theory*. Wiley Publications, 1st ed., 1991.

[9] G. Strang, *Linear Algebra and its Applications*. Harcourt Brace Jovanovich, third ed., 1988.

[10] M. H. M. Costa, "Writing on Dirty Paper," *IEEE Transactions on Information Theory*, vol. 29, pp. 439–441, May 1983.

[11] A. S. Cohen and A. Lapidoth, "Generalized Writing on Dirty Paper," *IEEE International Symposium on Information Theory*, p. 227, 2002.

[12] G. J. Foschini and A. H. Diaz, "Dirty Paper Coding: Perturbing Off the Infinite Dimensional Lattice Limit," *The DIMACS Workshop on Signal Processing for Wireless Transmission*, October 2002. Rutgers University.

[13] U. Erez, R. Zamir, and S. Shamai, "Additive Noise Channels with Side Information at the Transmitter ," *The 21st IEEE Convention of the Electrical and ELectronic Engineers in Israel*, pp. 373–376, April 2000.

[14] E. Biglieri, J. Proakis, and S. Shamai, "Fading channels: Information-Theoretic and Communications Aspects," *IEEE Transactions on Information Theory*, vol. 44, pp. 2619 –2692, October 1998.

[15] D. Samardzija and N. Mandayam, "Pilot Assisted Estimation of MIMO Fading Channel Response and Achievable Data Rates," *To Appear in IEEE Transactions on Signal Processing, Special Issue on MIMO*, 2003.