MIMO Broadcast Channels with Gaussian CSIT and Application to Location based CSIT

Manijeh Bashar‡, Yohan Lejosne‡¶, Dirk Slock‡, Yi Yuan-Wu¶
‡EURECOM, Sophia-Antipolis, France, Email: {bashar,lejosne,slock}@eurecom.fr
¶Orange Labs, Issy-les-Moulineaux, France, Email: {yohan.lejosne,yi.yuan}@orange.com

***Abstract*—Channel State Information at the Transmitter (CSIT), which is crucial in multi-user systems, is always imperfect in practice. In this paper we focus on the optimization of beamformers for the expected weighted sum rate (EWSR) in the MIMO Broadcast Channel (BC) (multi-user MIMO downlink). We first review some beamformer (BF) designs for the perfect CSIT case, such as Weighted Sum MSE (WSMSE) and we introduce the Weighted Sum SINR (WSSINR) point of view, an optimal form of the Signal to Leakage plus Noise Ratio (SLNR) or Signal to Jamming plus Noise Ratio (SJNR) approaches. The discussion then turns to mean and covariance Gaussian CSIT. We review an exact Monte Carlo based approach and a variety of approximate techniques and bounds that all reduce to problems of the (deterministic) form of perfect CSIT. Other simplified exact solutions can be obtained through massive MIMO asymptotics, or the more precise large MIMO asymptotics. Whereas in the perfect CSI case, all reviewed approaches are equivalent, they differ in the partial CSIT case. In particular the expected WSSINR approach is significantly better than expected WSMSE, with large MIMO asymptotics introducing some further tweaking weights that yield a deterministic approach that becomes exact when the number of antennas increases. The complexity and relative performance of the in the end many possible approaches and approximations are then compared.**

I. Introduction

Interference is the main limiting factor in wireless transmission, due to its open nature. In cellular systems, one can distinguish between the cell interior where a single cell design is appropriate and the cell edge where a multi-cell approach is mandatory. Even if interference would be treated as noise in a simplified approach for both cases, in the single cell design it is only the receiver that handles the interference noise whereas in the multicell approach, transmitter coordination is required (as in Interference Alignment (IA) for the Interference Channel (IC) formulation). Since Channel State Information at the Transmitter (Tx) (CSIT) is more difficult to obtain than the CSIR at the Receiver (Rx) (except perhaps in the TDD case), we focus here on the single cell downlink which in the multi user (MU) case becomes the Broadcast Channel (BC).

Partial CSIT formulations can typically be categorized as either bounded error / worst case (relevant for quantization error in digital feedback) or Gaussian error (relevant for analog feedback, prediction error, second-order statistics information). The Gaussian CSIT formulation with mean and covariance information was first introduced for SDMA (a Direction of Arrival (DoA) based historical precedent of MU MIMO), in which the channel outer product was typically replaced by the transmit side channel correlation matrix, and worked out in more detail for single user (SU) MIMO, e.g. [1], [2]. The use of covariance CSIT has recently reappeared in the context of Massive MIMO, [3], [4] where a not so rich propagation environment leads to subspaces (slow CSIT) for the channel vectors so that the fast CSIT can be reduced to the smaller dimension of the subspace, which is especially crucial for Massive MIMO.

With partial CSIT, outage is possible. Hence both Sum Rate (SR) and outage probability P_{out} have been considered as optimization criteria. As a result, there is some utility for adding space-time coding. This can be added independently of the beamformer design [1]. Also, in wireless Tx there is always fading, although with full CSIT outage can be avoided. In any case, we shall consider the ergodic sum rate as optimization criterion.

The contributions here are significantly better partial CSIT approaches compared to the EWSMSE approach in [5], and present deterministic alternatives to the stochastic approximation solution of [6]. We first treat the general Gaussian CSIT case. Then we focus on a location aided CSIT case with zero mean and identity plus rank one Tx side covariance matrix and no Rx side correlations. The goal here is to introduce a meaningful beamforming design at finite SNR and finite Ricean factor when not much more than the (location based) LoS information is available at the Tx. In this paper, Tx may denote transmit/transmitter/transmission and Rx may denote receive/receiver/reception.

II. Max WSR Techniques with Perfect CSI

We shall focus on MU MIMO designs in which each user gets one stream since some user selection can make this typically preferable over multiple streams/user. The $N_k \times 1$ received signal at user k is

$$\mathbf{y}_k = \mathbf{H}_k\, \mathbf{g}_k\, x_k + \sum_{i=1,\neq k}^{K} \mathbf{H}_k\, \mathbf{g}_i\, x_i + \mathbf{v}_k \qquad (1)$$

where x_i is the signal intended for user i, channel $\mathbf{H}_k$ has size $N_k \times M$, and $\mathbf{v}_k$ is additive noise. We shall assume that the $K \leq M$ signal streams x_i have unit variance and that the noise is white with $\mathbf{v}_k \sim \mathcal{CN}(0, \sigma_{v,k}^2 I_{N_k})$ (for N_k Rx antennas). We shall assume that the received signal in (1) is rescaled[1] so that

[1]For the case of spatially correlated noise (interference) at the receiver, one may need to equivalently consider Rx side channel correlation $\mathbf{C}_{r,k}$ after noise whitening.

the noise variance becomes $\sigma_{v,k}^2 = 1$. The spatial Tx filter or beamformer (BF) is $\mathbf{g}_k$. Consider as a starting point for the optimization the weighted sum rate (WSR)

$$WSR = WSR(\mathbf{g}) = \sum_{k=1}^{K} u_k \ \ln \frac{1}{e_k} \qquad (2)$$

where $\mathbf{g}$ represents the collection of BFs $\mathbf{g}_k$, the u_k are rate weights, the $e_k = e_k(\mathbf{g})$ are the Minimum Mean Squared Errors (MMSEs)

$$\begin{aligned} &\frac{1}{e_k} = 1 + \mathbf{g}_k^H \mathbf{H}_k^H \mathbf{R}_{\overline{k}}^{-1} \mathbf{H}_k \mathbf{g}_k = (1 - \mathbf{g}_k^H \mathbf{H}_k^H \mathbf{R}_k^{-1} \mathbf{H}_k \mathbf{g}_k)^{-1} \\ &\mathbf{R}_k = \mathbf{R}_{\overline{k}} + \mathbf{H}_k \mathbf{g}_k \mathbf{g}_k^H \mathbf{H}_k^H \\ &\mathbf{R}_{\overline{k}} = \textstyle\sum_{i \neq k} \mathbf{H}_k \mathbf{g}_i \mathbf{g}_i^H \mathbf{H}_k^H + I_{N_k} \ , \end{aligned} \qquad (3)$$

$\mathbf{R}_k$, $\mathbf{R}_{\overline{k}}$ are the total and interference plus noise Rx covariance matrices resp. and e_k is the MMSE obtained at the output $\widehat{x}_k = \mathbf{f}_k^H \mathbf{y}_k$ of the optimal (MMSE) linear Rx $\mathbf{f}_k$,

$$\mathbf{f}_k = \mathbf{R}_k^{-1} \mathbf{H}_k \mathbf{g}_k \ . \qquad (4)$$

A. Minimum Weighted Sum MSE (WSMSE)

For a general Rx filter $\mathbf{f}_k$ we have the MSE

$$\begin{aligned} e_k(\mathbf{f}_k, \mathbf{g}) &= (1 - \mathbf{f}_k^H \mathbf{H}_k \mathbf{g}_k)(1 - \mathbf{g}_k^H \mathbf{H}_k^H \mathbf{f}_k) \\ &\quad + \textstyle\sum_{i \neq k} \mathbf{f}_k^H \mathbf{H}_k \mathbf{g}_i \mathbf{g}_i^H \mathbf{H}_k^H \mathbf{f}_k + ||\mathbf{f}_k||^2 = \\ 1 - \mathbf{f}_k^H &\mathbf{H}_k \mathbf{g}_k - \mathbf{g}_k^H \mathbf{H}_k^H \mathbf{f}_k + \textstyle\sum_i \mathbf{f}_k^H \mathbf{H}_k \mathbf{g}_i \mathbf{g}_i^H \mathbf{H}_k^H \mathbf{f}_k + ||\mathbf{f}_k||^2. \end{aligned} \qquad (5)$$

The $WSR(\mathbf{g})$ is a non-convex and complicated function of $\mathbf{g}$. Inspired by [7], we introduced in [8], [9] an augmented cost function, the Weighted Sum MSE, $WSMSE(\mathbf{g}, \mathbf{f}, w)$

$$= \sum_{k=1}^{K} u_k (w_k \ e_k(\mathbf{f}_k, \mathbf{g}) - \ln w_k) + \lambda (\sum_{k=1}^{K} ||\mathbf{g}_k||^2 - P) \qquad (6)$$

where λ is a Lagrange multiplier and P is the Tx power constraint. After optimizing over the aggregate auxiliary Rx filters $\mathbf{f}$ and weights w, we get the WSR back:

$$\min_{\mathbf{f}, w} WSMSE(\mathbf{g}, \mathbf{f}, w) = -WSR(\mathbf{g}) + \sum_{k=1}^{K} u_k \qquad (7)$$

where we shall typically ignore the last constant term. The advantage of the augmented cost function is however that alternating optimization for one of the three sets of quantities, $\mathbf{g}, \mathbf{f}, w$, keeping the other two fixed, leads to solving simple quadratic or convex functions

$$\begin{aligned} &\min_{w_k} WSMSE \ \Rightarrow \ w_k = 1/e_k \\ &\min_{\mathbf{f}_k} WSMSE \ \Rightarrow \ \mathbf{f}_k = (\sum_i \mathbf{H}_k \mathbf{g}_i \mathbf{g}_i^H \mathbf{H}_k^H + I_{N_k})^{-1} \mathbf{H}_k \mathbf{g}_k \\ &\min_{\mathbf{g}_k} WSMSE \ \Rightarrow \\ &\qquad \mathbf{g}_k = (\textstyle\sum_i u_i w_i \mathbf{H}_i^H \mathbf{f}_i \mathbf{f}_i^H \mathbf{H}_i + \lambda I_M)^{-1} \mathbf{H}_k^H \mathbf{f}_k u_k w_k \end{aligned} \qquad (8)$$

Indeed, after substituting (5) into (6), one can notice the UL/DL duality, leading to a duality between Tx and Rx filters and the optimal Tx filter g_k in (8) is indeed of the form of a MMSE linear Rx for the dual UL in which λ plays the role of Rx noise variance and $u_k w_k$ plays the role of stream variance.

B. SINRs

The WSR can be rewritten as

$$WSR = WSR(\mathbf{g}) = \sum_{k=1}^{K} u_k \ \ln(1 + \text{SINR}_k) \qquad (9)$$

where $1 + \text{SINR}_k = 1/e_k$ or for general $\mathbf{f}_k$:

$$\text{SINR}_k = \frac{|\mathbf{f}_k \mathbf{H}_k \mathbf{g}_k|^2}{\sum_{i=1, \neq k}^{K} |\mathbf{f}_k \mathbf{H}_k \mathbf{g}_i|^2 + ||\mathbf{f}_k||^2} \ . \qquad (10)$$

From (9), one can obtain the WSR variation as

$$\partial WSR = \sum_{k=1}^{K} \frac{u_k}{1 + \text{SINR}_k} \partial \text{SINR}_k \qquad (11)$$

which can be interpreted as the variation of a weighted sum SINR (WSSINR) criterion. The BFs obtained from this criterion interpretation are of course the same as those of the WSR or WSMSE criteria, but this interpretation shows that the WSR approach is an optimal approach to the SLNR or SJNR heuristics. The details of this WSSINR approach correspond to the Kim-Giannakis discussion below.

C. Optimal Lagrange Multiplier

The optimal λ for the update of $\mathbf{g}_k$ in (8) can be found using a (bisection) line search on $\sum_{k=1}^{K} ||\mathbf{g}_k||^2 - P = 0$ as in [10]. Alternatively, it can be updated analytically as in [8], [9] by exploiting $\sum_k \mathbf{g}_k^H \frac{\partial WSMSE}{\partial \mathbf{g}_k^*} = 0$. This leads to the same result as in [11] where the problem of introducing and finding a Lagrange multiplier was avoided by reparameterizing the stream powers to satisfy the power constraint. If we reparameterize using normalized BF vectors $\mathbf{g}_k^{'}$ ($||\mathbf{g}_k^{'}|| = 1$) and stream powers p_k: $\mathbf{g}_k = \sqrt{p_k} \, \mathbf{g}_k^{'}$, then the power constraint becomes $\sum_{k=1}^{K} p_k - P = 0$. The reparameterized powers become $p_k = p_k^{'} / \sum_{i=1}^{K} p_i^{'}$ in which the new power parameters $p_k^{'}$ are now unconstrained and lead to

$$\text{SINR}_k = \frac{|\mathbf{f}_k \mathbf{H}_k \mathbf{g}_k^{'}|^2 p_k^{'}}{\sum_{i=1, \neq k}^{K} |\mathbf{f}_k \mathbf{H}_k \mathbf{g}_i^{'}|^2 p_i^{'} + ||\mathbf{f}_k||^2 \sum_{m=1}^{K} p_m^{'}} \ . \qquad (12)$$

This leads to the same Lagrange multiplier expression obtained in [7] on the basis of a heuristic that was introduced in [12] as was pointed out in [13].

D. Kim-Giannakis [14]

Let $\mathbf{Q}_k = \mathbf{g}_k \mathbf{g}_k^H$ be the transmit covariance for stream k. The WSR can be rewritten as

$$WSR = \sum_{k=1}^{K} u_k [\ln \det(\mathbf{R}_k) - \ln \det(\mathbf{R}_{\overline{k}})] \qquad (13)$$

where $\mathbf{R}_k = \mathbf{H}_k (\sum_i \mathbf{Q}_i) \mathbf{H}_k^H + I_{N_k}$, and $\mathbf{R}_{\overline{k}} = \mathbf{H}_k (\sum_{i \neq k} \mathbf{Q}_i) \mathbf{H}_k^H + I_{N_k}$. Kim and Giannakis propose to keep the concave signal terms and to replace the convex interference terms by the linear (and hence concave) tangent approximation. More specifically, consider the dependence of WSR on $\mathbf{Q}_k$ alone. Then

$$\begin{aligned} WSR &= u_k \ln\det(\mathbf{R}_{\overline{k}}^{-1}\mathbf{R}_k) + WSR_{\overline{k}}\,, \\ WSR_{\overline{k}} &= \textstyle\sum_{i=1,\neq k}^{K} u_i \ln\det(\mathbf{R}_{\overline{i}}^{-1}\mathbf{R}_i) \end{aligned} \quad (14)$$

where $\ln\det(\mathbf{R}_{\overline{k}}^{-1}\mathbf{R}_k)$ is concave in $\mathbf{Q}_k$ and $WSR_{\overline{k}}$ is convex in $\mathbf{Q}_k$. Since a linear function is simultaneously convex and concave, consider the first order Taylor series expansion in $\mathbf{Q}_k$ around $\widehat{\mathbf{Q}}$ (i.e. all $\widehat{\mathbf{Q}}_i$) with e.g. $\widehat{\mathbf{R}}_i = \mathbf{R}_i(\widehat{\mathbf{Q}})$, then

$$\begin{aligned} & WSR_{\overline{k}}(\mathbf{Q}_k,\widehat{\mathbf{Q}}) \approx WSR_{\overline{k}}(\widehat{\mathbf{Q}}_k,\widehat{\mathbf{Q}}) - \mathrm{tr}\{(\mathbf{Q}_k - \widehat{\mathbf{Q}}_k)\widehat{\mathbf{A}}_k\} \\ & \widehat{\mathbf{A}}_k = -\left.\frac{\partial WSR_{\overline{k}}(\mathbf{Q}_k,\widehat{\mathbf{Q}})}{\partial \mathbf{Q}_k}\right|_{\widehat{\mathbf{Q}}_k,\widehat{\mathbf{Q}}} = \sum_{i=1,\neq k}^{K} u_i \mathbf{H}_i^H(\widehat{\mathbf{R}}_{\overline{i}}^{-1} - \widehat{\mathbf{R}}_i^{-1})\mathbf{H}_i \end{aligned} \quad (15)$$

Note that the linearized (tangent) expression for $WSR_{\overline{k}}$ constitutes a lower bound for it. Now, dropping constant terms, reparameterizing the $Q_k = \mathbf{g}_k\mathbf{g}_k^H$ and performing this linearization for all users, we get

$$\begin{aligned} & WSR(\mathbf{g},\widehat{\mathbf{g}}) = \lambda P + \\ & \sum_{k=1}^{K} u_k \ln(1 + \mathbf{g}_k^H\mathbf{H}_k^H\widehat{\mathbf{R}}_{\overline{k}}^{-1}\mathbf{H}_k\mathbf{g}_k) - \mathbf{g}_k^H(\widehat{\mathbf{A}}_k + \lambda I)\mathbf{g}_k\,. \end{aligned} \quad (16)$$

The gradient of this concave WSR lower bound is actually still the same as that of the original WSR or of the WSMSE criteria! And it allows an interpretation as a generalized eigenvector condition

$$\mathbf{H}_k^H\widehat{\mathbf{R}}_{\overline{k}}^{-1}\mathbf{H}_k\mathbf{g}_k = \frac{1 + \mathbf{g}_k^H\mathbf{H}_k^H\widehat{\mathbf{R}}_{\overline{k}}^{-1}\mathbf{H}_k\mathbf{g}_k}{u_k}(\widehat{\mathbf{A}}_k + \lambda I)\mathbf{g}_k \quad (17)$$

or hence $\mathbf{g}_k^{'} = V_{max}(\mathbf{H}_k^H\widehat{\mathbf{R}}_{\overline{k}}^{-1}\mathbf{H}_k, \widehat{\mathbf{A}}_k + \lambda I)$ is the "max" generalized eigenvector of the two indicated matrices and is proportional to the "LMMSE" $\mathbf{g}_k$ in (8), with max eigenvalue $\sigma_k = \sigma_{max}(\mathbf{H}_k^H\widehat{\mathbf{R}}_{\overline{k}}^{-1}\mathbf{H}_k, \widehat{\mathbf{A}}_k + \lambda I)$. This can be viewed as an optimally weighted version of the SLNR solution [15] which takes as Tx filter $\mathbf{g}_k^{'} = V_{max}(\mathbf{H}_k^H\mathbf{H}_k, \sum_{i\neq k}\mathbf{H}_i^H\mathbf{H}_i + I)$.

Let $\sigma_k^{(1)} = \mathbf{g}_k^{'H}\mathbf{H}_k^H\widehat{\mathbf{R}}_{\overline{k}}^{-1}\mathbf{H}_k\mathbf{g}_k^{'}$ and $\sigma_k^{(2)} = \mathbf{g}_k^{'H}\widehat{\mathbf{A}}_k\mathbf{g}_k^{'}$. The advantage of formulation (16) is that it allows straightforward power adaptation: substituting $\mathbf{g}_k = \sqrt{p_k}\,\mathbf{g}_k^{'}$ in (16) yields

$$WSR = \lambda P + \sum_{k=1}^{K}\{u_k \ln(1 + p_k\sigma_k^{(1)}) - p_k(\sigma_k^{(2)} + \lambda)\} \quad (18)$$

which leads to the following interference leakage aware water filling

$$p_k = \left(\frac{u_k}{\sigma_k^{(2)} + \lambda} - \frac{1}{\sigma_k^{(1)}}\right)^{+}. \quad (19)$$

For a given λ, $\mathbf{g}$ needs to be iterated till convergence. And λ can be found by duality (line search):

$$\min_{\lambda\geq 0}\max_{\mathbf{g}} \lambda P + \sum_k\{u_k \ln\det(\mathbf{R}_{\overline{k}}^{-1}\mathbf{R}_k) - \lambda p_k\} = \min_{\lambda\geq 0} WSR(\lambda). \quad (20)$$

E. High/Low SNR Behavior

At high SNR, the max WSR BF converge to the ZF solutions with uniform power distribution. For ZF BF, the BS shall use for user k a spatial filter $\mathbf{g}_k$ such that

$$\mathbf{g}_k^H = \mathbf{f}_k\mathbf{H}_k P^{\perp}_{(\mathbf{fH})^H_{\overline{k}}} / ||\mathbf{f}_k\mathbf{H}_k P^{\perp}_{(\mathbf{fH})^H_{\overline{k}}}|| \quad (21)$$

where $P^{\perp}_{\mathbf{X}} = I - P_{\mathbf{X}}$ and $P_{\mathbf{X}} = \mathbf{X}(\mathbf{X}^H\mathbf{X})^{-1}\mathbf{X}^H$ are projection matrices (onto the column space of $\mathbf{X}$ or its orthogonal complement, so here the row space of $(\mathbf{fH})_{\overline{k}}$) and $(\mathbf{fH})_{\overline{k}}$ denotes the (up-down) stacking of $\mathbf{f}_i\mathbf{H}_i$ for users $i = 1,\ldots,K,\, i\neq k$.

At low SNR, we get the matched filter (MF) solution for the user with largest $||\mathbf{H}_k||_2$ (max singular value).

F. Number of Local Optima

In the MIMO case, more Rx antennas does not help to increase the total number of streams. However, they may lead to different distributions of ZF (high SNR) between Tx and Rx, yielding different ZF channel gains ($\mathbf{f}_k\mathbf{H}_k\mathbf{g}_k$)! If the Rx ZF's n streams, then the Tx only has to ZF $M-1-n$ streams! So, the number of possible solutions (for one stream/user) becomes [16]:

$$\prod_{k=1}^{K}\left(\sum_{n=0}^{N_k-1}\frac{(M-1)!}{n!(M-1-n)!}\right) \quad (22)$$

i.e., for each user, the Rx can ZF n between 0 and $N-1$ streams, to choose among $M-1$. These different ZF solutions are the possible local optima for max WSR at infinite SNR. By homotopy [9] this remains the number of max WSR local optima as the SNR decreases from infinity. Of course, as the SNR decreases further, a stream for some user may get turned off until only a single stream remains at low SNR. Hence, the number of local optima (ZF possibilities) reduces as streams disappear at finite SNR .

As a corollary, in the MISO case, the max WSR optimum is unique, since there is only one way to perform ZF BF.

G. Finding the Global MU-MIMO Optimum: Deterministic Annealing

At low SNR, the BF solution can be written down analytically. The usual water filling at low SNR leads to a Matched Filter (MF) solution for the user with the largest $||\mathbf{H}_k||_2$ (max singular value). Deterministic annealing can be used as in [9] to track this global optimum from SNR $= 0$ to the desired SNR. Along this SNR trajectory, a homotopy method is used to track the evolution of the BF filters: as the SNR gets incremented, one iteration of the iterative methods above is sufficient to update the $\mathbf{g}_k$ to the higher SNR. In deterministic annealing, the homotopy evolution gets complemented by phase transitions which correspond to a stream for one more user getting switched on. As this stream then barely gets switched on (the SNR for this stream is near zero), its optimal BF initialization corresponds to a colored noise (existing streams) MF, whereas the existing streams get barely affected.

III. MEAN AND COVARIANCE GAUSSIAN CSIT

In this section we drop the user index k for simplicity. Mean information about the channel can come from channel feedback or reciprocity, and prediction, or it may correspond to the non fading (e.g. LoS) part of the channel (note that an unknown phase factor $e^{j\phi}$ in the overall channel mean does not affect the BF design). Covariance information may correspond to channel estimation (feedback, prediction) errors and/or to information about spatial correlations. The separable (or Kronecker) correlation model (for the channel itself, as opposed to its estimation error or knowledge) below is acceptable when the number of propagation paths N_p becomes large ($N_p \gg MN$) as possibly in indoor propagation. Given only mean and covariance information, the fitting maximum entropy distribution is Gaussian. Hence consider

$$\text{vec}(\mathbf{H}) \sim \mathcal{CN}(\text{vec}(\overline{\mathbf{H}}), \mathbf{C}_t^T \otimes \mathbf{C}_r) \tag{23}$$

which can be rewritten as

$$\mathbf{H} = \overline{\mathbf{H}} + \mathbf{C}_r^{1/2}\, \widetilde{\mathbf{H}}\, \mathbf{C}_t^{1/2} \tag{24}$$

where $\mathbf{C}_r^{1/2}$, $\mathbf{C}_t^{1/2}$ are Hermitian square-roots of the Rx and Tx side covariance matrices

$$\begin{array}{l} \text{E}(\mathbf{H}-\overline{\mathbf{H}})(\mathbf{H}-\overline{\mathbf{H}})^H = \text{tr}\{\mathbf{C}_t\}\ \mathbf{C}_r \\ \text{E}(\mathbf{H}-\overline{\mathbf{H}})^H(\mathbf{H}-\overline{\mathbf{H}}) = \text{tr}\{\mathbf{C}_r\}\ \mathbf{C}_t \end{array} \tag{25}$$

and the elements of $\widetilde{\mathbf{H}}$ are i.i.d. $\sim \mathcal{CN}(0,1)$. Obviously, a scale factor needs to be fixed in the product $\text{tr}\{\mathbf{C}_r\}\text{tr}\{\mathbf{C}_t\}$ for unicity. In what follows, it will also be of interest to consider the total Tx side correlation matrix

$$\mathbf{R}_t = \text{E}\,\mathbf{H}^H\mathbf{H} = \overline{\mathbf{H}}^H\overline{\mathbf{H}} + \text{tr}\{\mathbf{C}_r\}\,\mathbf{C}_t\ . \tag{26}$$

Note that the Gaussian CSIT model could be considered an instance of Ricean fading in which the ratio $\text{tr}\{\overline{\mathbf{H}}^H\overline{\mathbf{H}}\}/(\text{tr}\{\mathbf{C}_r\}\text{tr}\{\mathbf{C}_t\})$ could be considered the Ricean factor.

IV. EXPECTED WSR (EWSR)

Now, so far we have assumed that the channel $\mathbf{H}$ is known. The scenario of interest however is that of perfect CSIR but partial (LoS) CSIT. Once the CSIT is imperfect, various optimization criteria could be considered, such as outage capacity. Here we shall consider the expected weighted sum rate $\text{E}_{\mathbf{H}} WSR(\mathbf{g},\mathbf{H}) =$

$$EWSR(\mathbf{g}) = \text{E}_{\mathbf{H}} \sum_k u_k \ln(1 + \mathbf{g}_k^H \mathbf{H}_k^H \mathbf{R}_{\overline{k}}^{-1} \mathbf{H}_k \mathbf{g}_k) \tag{27}$$

where we now underlign the dependence of various quantities on $\mathbf{H}$. The EWSR in (27) corresponds to perfect CSIR since the optimal Rx filters $\mathbf{f}_k$ as a function of the aggregate $\mathbf{H}$ have been substituted, namely $WSR(\mathbf{g},\mathbf{H}) = \max_{\mathbf{f}} \sum_k u_k(-\ln(e_k(\mathbf{f}_k,\mathbf{g})))$. At high SNR we get:

Theorem 1: **Sufficiency of Incomplete CSIT for Full DoF in MIMO BC** In the MIMO BC with perfect CSIR, it is sufficient that for each of the K users rank$(\mathbf{R}_{t,k}) \leq N_k$ and that the BS knows any vector $\mathbf{h}_k \in$ Range$(\mathbf{R}_{t,k})$ (as long as the resulting vectors $\mathbf{h}_k$ are linearly independent) in order for ZF BF to produce $\min(M,K)$ interference free streams (degrees of freedom (DoF)).

V. MAX EWSR BY STOCHASTIC APPROXIMATION

In [6] a stochastic approximation approach for maximizing the EWSR was introduced. In this approach the statistical average gets replaced by a sample average (samples of $\mathbf{H}$ get generated according to its Gaussian CSIT distribution in a Monte Carlo fashion), and one iteration of the min WSMSE approach gets executed per term added in the sample average.

Some issues with this approach are that in this case the number of iterations may get dictated by a sufficient size for the sample average rather than by a convergence requirement for the iterative approach. Another issue is that this approach converges to a local maximum of the EWSR. It is not immediately clear how to combine this stochastic approximation approach with deterministic annealing.

In the rest of this paper we discuss various deterministic approximations and bounds for the EWSR, which can then be optimized as in the full CSI case.

VI. EXPECTED WSR (EWSR) BOUNDS

1) EWSR Lower Bound: EWSMSE: $EWSR(\mathbf{g})$ is first of all difficult to compute and also again difficult to maximize directly. As observed in [5], it appears much more attractive to consider $\text{E}_{\mathbf{H}} e_k(\mathbf{f}_k,\mathbf{g},\mathbf{H})$ since $e_k(\mathbf{f}_k,\mathbf{g},\mathbf{H})$ is quadratic in $\mathbf{H}$. Hence in [5], the cost function optimized is $\text{E}_{\mathbf{H}} WSMSE(\mathbf{g},\mathbf{f},w,\mathbf{H})$ where $WSMSE(\mathbf{g},\mathbf{f},w,\mathbf{H})$ appears in (6). However,

$$\begin{array}{l} \min_{\mathbf{f},w}\ \text{E}_{\mathbf{H}} WSMSE(\mathbf{g},\mathbf{f},w,\mathbf{H}) \\ \geq \text{E}_{\mathbf{H}} \min_{\mathbf{f},w} WSMSE(\mathbf{g},\mathbf{f},w,\mathbf{H}) = -EWSR(\mathbf{g}) \end{array} \tag{28}$$

or hence $EWSR(\mathbf{g}) \geq -\min_{\mathbf{f},w}\ \text{E}_{\mathbf{H}} WSMSE(\mathbf{g},\mathbf{f},w,\mathbf{H})$. So now only a lower bound to the EWSR gets maximized, which corresponds in fact to the CSIR being equally partial as the CSIT. The EWSR gap can be reduced by following the optimization over the Tx filters $\mathbf{g}_k$ with an optimization over the Rx filters $\mathbf{f}_k$ for full CSIR, namely by taking the $\mathbf{f}_k$ as in (4).

From (5), we get

$$\begin{array}{l} \text{E}_{\mathbf{H}} e_k = 1 - 2\Re\{\mathbf{f}_k^H \overline{\mathbf{H}}_k \mathbf{g}_k\} + \sum_{i=1}^K \mathbf{f}_k^H \overline{\mathbf{H}}_k \mathbf{g}_i \mathbf{g}_i^H \overline{\mathbf{H}}_k^H \mathbf{f}_k \\ + \mathbf{f}_k^H \mathbf{R}_{r,k} \mathbf{f}_k \sum_{i=1}^K \mathbf{g}_i^H \mathbf{R}_{t,k} \mathbf{g}_i + ||\mathbf{f}_k||^2. \end{array} \tag{29}$$

From the first line we see that the signal term disappears if $\overline{\mathbf{H}}_k = 0$! Hence the EWSMSE lower bound is (very) loose unless the Rice factor is high, and is useless in the absence of mean CSIT.

2) EWSR Upper Bound: Using the concavity of $\ln(.)$, we get

$$EWSR(\mathbf{g}) \leq \sum_{k=1}^K u_k\ \ln(1 + \text{E}_{\mathbf{H}_k}\, \text{SINR}_k(\mathbf{g},\mathbf{H}_k))\ . \tag{30}$$

VII. Max ES-EI-NR Approach

Consider the approximation

$$\mathrm{E}_{\mathbf{H}} \ln(1+\mathrm{SINR}_k) \approx \ln(1+\frac{\mathrm{E}S}{\mathrm{E}I+N}) \quad . \tag{31}$$

This can be solved as easily as min (E)WSMSE. However, here the $\widetilde{\mathbf{H}}_k$ part in the signal gets also counted in the signal power, unlike in the EWSMSE criterion where it gets ignored. The approximation (31) becomes exact in Massive MIMO, as $M \to \infty$. The WSR can still be rewritten as

$$WSR = \sum_{k=1}^{K} u_k [\ln\det(\widetilde{\mathbf{R}}_k) - \ln\det(\widetilde{\mathbf{R}}_{\overline{k}})] \tag{32}$$

where $\widetilde{\mathbf{R}}_k = (\sum_i \mathbf{Q}_i)\mathbf{H}_k^H\mathbf{H}_k + I_M$, and $\widetilde{\mathbf{R}}_{\overline{k}} = (\sum_{i\neq k} \mathbf{Q}_i)\mathbf{H}_k^H\mathbf{H}_k + I_M$. We now apply the algorithm in section II.D, replacing $\widetilde{\mathbf{R}}_k$, $\widetilde{\mathbf{R}}_{\overline{k}}$ by $\mathrm{E}\,\widetilde{\mathbf{R}}_k$, $\mathrm{E}\,\widetilde{\mathbf{R}}_{\overline{k}}$, and hence $\mathbf{H}_k^H\mathbf{H}_k$ by $\mathbf{R}_{t,k}$ and expressions of the form $\mathbf{H}_k^H\mathbf{R}^{-1}\mathbf{H}_k$ by $\mathbf{R}_{t,k}\mathbf{R}^{-1}$.

A. Large MIMO Asymptotics Refinement

The SU MIMO asymptotics from [17], [18] (in which both $M, N \to \infty$, which tends to give more precise approximations when M is not so large) for a term of the form $\ln\det(\mathbf{Q}\mathbf{H}^H\mathbf{H}+I)$ (as in (32)) correspond to replacing $\mathbf{H}_k^H\mathbf{H}_k$ in the $\widetilde{\mathbf{R}}_k$ and $\widetilde{\mathbf{R}}_{\overline{k}}$ in (32) with a kind of $\mathbf{R}_{t,k}$ with a different weighting of the $\overline{\mathbf{H}}_k^H\overline{\mathbf{H}}_k$ and $\mathbf{C}_{t,k}$ portions, of the form $\mathbf{R}_{t,k}^{'} = a_k\mathbf{C}_{t,k} + \overline{\mathbf{H}}_k^H\mathbf{B}_k\overline{\mathbf{H}}_k$ for some scalar a_k and matrix $\mathbf{B}_k$ that depends on $\mathbf{C}_{r,k}$.

For the general case of Gaussian CSIT with separable (Kronecker) covariance structure, [17], [18] lead to asymptotic expressions of the form

$$\begin{aligned} &\mathrm{E}_{\mathbf{H}} \ln\det(I+\mathbf{H}\mathbf{Q}\mathbf{H}^H) \\ &= \max_{z,w}\left\{ \ln\det\begin{bmatrix} I+w\mathbf{C}_r & \overline{\mathbf{H}} \\ -\mathbf{Q}\overline{\mathbf{H}}^H & I+z\mathbf{Q}\mathbf{C}_t \end{bmatrix} - zw \right\} . \end{aligned} \tag{33}$$

For the simpler case of zero channel means $\overline{\mathbf{H}}_k = 0$ and no Rx side correlations $\mathbf{C}_r = I$, and with per user Tx side correlations $\mathbf{C}_t \leftarrow \mathbf{C}_k$, the EWSR can be rewritten with large MIMO asymptotics as

$$\begin{aligned} &EWSR = \\ &\sum_{k=1}^{K}\left\{ u_k \max_{z_k,w_k} \left[\ln\det(I+z_k\mathbf{G}\mathbf{G}^H\mathbf{C}_k)+N_k\ln(1+w_k)-z_kw_k\right]\right. \\ &\left. -u_k \max_{z_{\overline{k}},w_{\overline{k}}} \left[\ln\det(I+z_{\overline{k}}\mathbf{G}_{\overline{k}}\mathbf{G}_{\overline{k}}^H\mathbf{C}_k)+N_k\ln(1+w_{\overline{k}})-z_{\overline{k}}w_{\overline{k}}\right]\right\} \end{aligned} \tag{34}$$

where $\mathbf{G} = [\mathbf{g}_1 \cdots \mathbf{g}_K]$ and $\mathbf{G}_{\overline{k}}$ is the same as $\mathbf{G}$ except for column $\mathbf{g}_k$. The expression in (34) can be maximized by alternating optimization.

B. Other possible WSR Approximations

1) Absorbing the Mean in the Covariance:: Replacing $\overline{\mathbf{H}}_k$ by 0 and $\mathbf{C}_{t,k}$ by $\mathbf{R}_{t,k}$ as suggested in [1] for SU MIMO leads to one simplification. Other simplifications can be obtained by either absorbing the noise term in the "Rayleigh" channel part of the interference or vice versa.

2) Improvements upon ESEINR: One simple such improvement can be obtained by acknowledging the quadratic explicit appearance of the channels in the gradient of the WSR w.r.t. the Tx filters, and then compute the corresponding second-order moments, similar to EWSMSE. Of course, the actual dependence of the gradient of the WSR on the channels is highly nonlinear but we replace terms like the MSE e_k and $\mathbf{R}_{\overline{k}}$ by their mean. This approach acknowledges that the Rx contains the channel matched filter as factor and applies the second order statistics to the resulting quadratic appearances of the channel.

3) Higher-Order Taylor Series Expansions: One possibility is to go to the next (second) order term in the Taylor series expansion of the log as in (15) in [19].

VIII. Propagation Channel Model

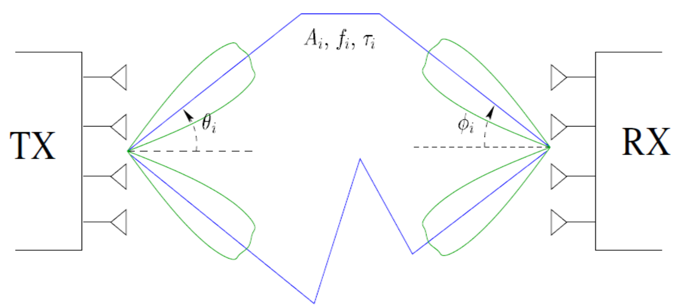

Fig. 1. MIMO transmission with M transmit and N receive antennas.

A. Specular Wireless MIMO Channel Model

Consider a MIMO transmission configuration as depicted in Fig. 1. We get for the matrix impulse response of the time-varying channel $\mathbf{h}(t,\tau)$ [20]

$$\mathbf{h}(t,\tau) = \sum_{i=1}^{N_p} A_i(t)\, e^{j2\pi\, f_i\, t}\, \mathbf{h}_r(\phi_i)\, \mathbf{h}_t^H(\theta_i)\, p(\tau-\tau_i) \quad . \tag{35}$$

The channel impulse response $\mathbf{h}$ has per path a rank 1 contribution in 4 dimensions (Tx and Rx spatial multi-antenna dimensions, delay spread and Doppler spread); there are N_p pathwise contributions where

- A_i: complex attenuation
- f_i: Doppler shift
- θ_i: angle of departure (AoD)
- ϕ_i: direction of arrival (DoA)
- τ_i: path delay (ToA)
- $\mathbf{h}_t^*(.)/\mathbf{h}_r(.)$: $M/N \times 1$ Tx/Rx antenna array response
- $p(.)$: pulse shape (Tx filter)

The fast variation of the phase in $e^{j2\pi\, f_i\, t}$ and possibly the variation of the A_i (when the nominal path represents in fact a superposition of paths with similar parameters) correspond to the fast fading. All the other parameters (including the Doppler frequency) vary on a slower time scale and correspond to slow fading. We shall assume here OFDM transmission, as is typical for 4G systems, with the Doppler variation over the OFDM symbol duration being negligible. We then get for the channel transfer matrix at any particular subcarrier of a given OFDM symbol

$$\mathbf{H} = \sum_{i=1}^{N_p} A_i \, \mathbf{h}_r(\phi_i) \, \mathbf{h}_t^H(\theta_i) \tag{36}$$

where with some abuse of notation we use the same complex amplitude A_i in which we ignored the dependence on time (particular OFDM symbol), through at least the Doppler shift, and on frequency (subcarrier), through the Tx (and Rx) filter(s).

B. Narrow AoD Aperture (NADA) case

The idea here is to focus on the category of mobiles for which the angular spread seen from the BS is limited [21]. This is a small generalization of the LoS case. In the NADA case, the MIMO channel $\mathbf{H}$ is of the form

$$\mathbf{H} = \sum_i A_i \, \mathbf{h}_r(\phi_i)\mathbf{h}_t^H(\theta_i) \approx \mathbf{B} \, \mathbf{A}^H \; , \;\; \mathbf{A} = \begin{bmatrix} \mathbf{h}_t(\theta) \; \dot{\mathbf{h}}_t(\theta) \end{bmatrix} \; . \tag{37}$$

In the case of narrow AoD spread, we have

$$\theta_i = \theta + \Delta\theta_i \tag{38}$$

where θ is the nominal (LoS) AoD and $\Delta\theta_i$ is small. Hence

$$\mathbf{h}_t(\theta_i) \approx \mathbf{h}_t(\theta) + \Delta\theta_i \; \dot{\mathbf{h}}_t(\theta) \; . \tag{39}$$

This leads to the second equality in (37). Hence $\mathbf{H}$ is of rank 2 (regardless of the DoA spread). The LOS case is a limiting case in which the power of the $\dot{\mathbf{h}}_t(\theta)$ term becomes negligible and the channel rank becomes 1. The factor $\mathbf{A}$ in $\mathbf{H}$ depends straightforwardly on position (which translates into LOS AoD), only $\mathbf{B}$ (which depends on the $A_i \, \mathbf{h}_r(\phi_i)$ and the $\Delta\theta_i$) remains random.

C. Location Aided Partial CSIT LoS Channel Model

Assuming the Tx disposes of not much more than the LoS component information, we shall consider the following MIMO channel model

$$\mathbf{H} = \mathbf{h}_r \, \mathbf{h}_t^H(\theta) + \widetilde{\mathbf{H}}^{'} \tag{40}$$

where θ is the LoS AoD and the Tx side array response is normalized: $||\mathbf{h}_t(\theta)||^2 = 1$. Since the orientation of the MT is random, and the LoS case can be considered as a limiting NADA case in which a multitude of DoAs could appear, we shall model the Rx side LoS array response $\mathbf{h}_r$ as a vector of i.i.d. complex Gaussian variables

$$\begin{array}{l} \mathbf{h}_r \;\; \text{i.i.d.} \; \sim \mathcal{CN}(0, \frac{\mu}{\mu+1}) \quad \text{and} \\ \widetilde{\mathbf{H}}^{'} \;\; \text{i.i.d.} \; \sim \mathcal{CN}(0, \frac{1}{\mu+1}\frac{1}{M}) \; , \text{independent of } \mathbf{h}_r, \end{array} \tag{41}$$

where the matrix $\widetilde{\mathbf{H}}$ represents the aggregate NLoS components. Note that

$$\begin{array}{l} \mathrm{E}||\mathbf{H}||_F^2 = \, \mathrm{E}\,\mathrm{tr}\{\mathbf{H}^H\mathbf{H}\} = \\ ||\mathbf{h}_t(\theta)||^2 \;\; \mathrm{E}||\mathbf{h}_r||^2 + \mathrm{E}||\widetilde{\mathbf{H}}^{'}||_F^2 = \frac{\mu N}{\mu+1} + \frac{N}{\mu+1} = N \, , \end{array} \tag{42}$$

reflecting that Rx power augments proportionally with N, and $(\,\mathrm{E}||\mathbf{h}_r \, \mathbf{h}_t^T(\theta)||_F^2)/(\,\mathrm{E}||\widetilde{\mathbf{H}}^{'}||_F^2) = \mu$ which can be considered as a Rice factor. In fact the only parameter additional to the LoS AoD θ assumed in (40) is μ.

So, this is a case of zero mean CSIT and Tx side covariance CSIT

$$\mathbf{R}_t = \, \mathrm{E}\,\mathbf{H}^H\mathbf{H} = \frac{\mu N}{\mu+1}\mathbf{h}_t(\theta)\mathbf{h}_t^H(\theta) + \frac{N}{\mu+1}\frac{1}{M}I_M \; . \tag{43}$$

IX. The Location Aided Case

A. LoS ZF BF

For ZF BF, the BS shall use for user k a spatial filter $\mathbf{g}_k = \sqrt{p_k}\,\mathbf{g}_k^{'}$ such that $\mathbf{g}_k^{'} = \mathbf{g}_k^{''}/||\mathbf{g}_k^{''}||$

$$\mathbf{g}_k^{''} = P_{\mathbf{h}_{t,\overline{k}}}^{\perp}\mathbf{h}_{t,k} \tag{44}$$

where $\mathbf{h}_{t,\overline{k}} = [\mathbf{h}_{t,1} \cdots \mathbf{h}_{t,k-1} \, \mathbf{h}_{t,k+1} \cdots \mathbf{h}_{t,K}]$. And uniform power distribution $p_k = P/K, \; k = 1,\ldots,K$. The $\mathbf{g}_k^{''}$ can also be computed from

$$\mathbf{g}^{''} = [\mathbf{g}_1^{''} \cdots \mathbf{g}_K^{''}] = \mathbf{h}_t(\mathbf{h}_t^H\mathbf{h}_t)^{-1} \; , \;\; \mathbf{h}_t = [\mathbf{h}_{t,1} \cdots \mathbf{h}_{t,K}] \, . \tag{45}$$

B. Beyond ZF

In the previous subsection we considered the attainable DoF with LoS CSIT, attained by ZF Tx BF design on the LoS components. Note in passing that in practical multipath scenarios, even if only the interference passing through the LoS paths would be handled, this would already lead to a substantial SINR increase. Here we shall explore how to go beyond the asymptotics of high SNR and high Ricean factor: even if the Tx ignores the multipath and the Rx can handle it, it would be better to have a multipath aware Tx design. To this end, various intermediate forms of CSIT could be considered beyond the LoS knowledge only. Here we shall consider the perhaps simplest model, the partial CSIT LoS model of (40). Note that the Ricean factor μ satisfies uplink/downlink (UL/DL) reciprocity, even in a FDD system, and hence can easily be estimated.

C. Ricean Model Specific Approximations

1) Absorbing the Rayleigh Component in the Noise: Consider again the received signal model

$$\begin{array}{rcl} \mathbf{y}_k & = & \mathbf{H}_k \sum_{i=1}^K \mathbf{g}_i \, x_i + \mathbf{v}_k \\ & = & \mathbf{h}_{r,k} \, \mathbf{h}_{t,k}^H \sum_{i=1}^K \mathbf{g}_i \, x_i + \widetilde{\mathbf{H}}_k^{'} \sum_{i=1}^K \mathbf{g}_i \, x_i + \mathbf{v}_k \, . \end{array} \tag{46}$$

Going to this MIMO model to an equivalent SIMO model with the same SINR (or ESINR), we get

$$y_k = \sqrt{\frac{\mu_k}{\mu_k+1}}\mathbf{h}_{t,k}^H \sum_{i=1}^K \mathbf{g}_i \, x_i + \frac{1}{\sqrt{(\mu_k+1)M}}\widetilde{\mathbf{h}}_k^H \sum_{i=1}^K \mathbf{g}_i \, x_i + v_k^{'} \, . \tag{47}$$

Considering that $\widetilde{\mathbf{h}}_k$ is a vector of i.i.d. variables, the last two terms in (47) represent a spatially white noise with variance

$\sigma_{v,k}^2 + \dfrac{P}{(\mu_k+1)M} = \sigma_{v,k}^2(1+\dfrac{\mathrm{SNR}_k}{(\mu_k+1)M})$. Hence (47) leads to an equivalent MISO model

$$y_k = \mathbf{h}_{t,k}^H \sum_{i=1}^{K} \mathbf{g}_i \, x_i + v_k \tag{48}$$

with effective SNR

$$\mathrm{SNR}_{eff,k} = \frac{\mu_k \mathrm{SNR}_k}{\mu_k + 1 + \mathrm{SNR}_k/M} \tag{49}$$

which is now a deterministic MISO BC model.

2) Absorbing the Noise in the Rayleigh Component: Alternatively, the previous trick of assimilating the Rayleigh component with the noise can be used to associate the noise with the Rayleigh component and leads to a modified $\mathbf{C}_{t,k}$ (reduced Rice factor). The absence of noise and hence of the $I+$ term in the argument of $\ln\det(.)$ allows us to have a $\ln\det(.)$ of a product of factors. Assuming for a moment that all factors would be square allows us to put the i.i.d. random factors of the channel in a separate term, thus effectively replacing $\mathbf{H}_k^H\mathbf{H}_k$ by $\mathbf{C}_{t,k}$, again resulting in a determinstic WSR scenario.

X. Some First Simulations

A simulation result is presented in Fig. 2 where the optimal stochastic approximation result for the EWSR is compared to naive ZF BF on the LoS component only and to an optimized deterministic BF design in which the Rayleigh channel component is simply associated with the noise. This very simple approximation turns out to be as good as optimal in this scenario for an SNR up to 30dB.

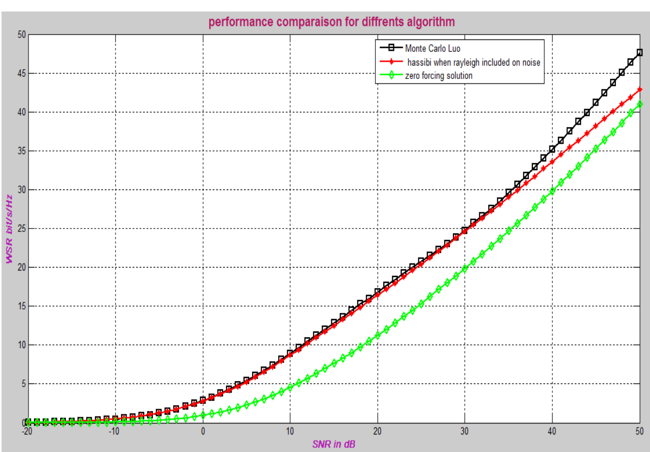

Fig. 2. EWSR vs SNR for $K = M = N_k = 4$ with Rice factor $\mu = 10$. The black curve corresponds to stochastic approximation, the green curve to ZF on the LoS component, and the red curve to an optimized deterministic BF design when the Rayleigh part is absorbed in the noise.

Acknowledgments

EURECOM's research is partially supported by its industrial members: ORANGE, BMW Group, Swisscom, SFR, ST Microelectronics, Symantec, SAP, Monaco Telecom, iABG, and also by the EU FP7 projects ADEL and NEWCOM#. The research of Orange Labs is partially supported by the EU FP7 project METIS.

References

[1] R. de Francisco and D. Slock, "Spatial Transmit Prefiltering for Frequency-Flat MIMO Transmission with Mean and Covariance Information," in *IEEE Asilomar Conf, on Signals, Systems, and Computers*, Pacific Grove, US, Oct. 2005.

[2] M. Vu and A. Paulraj, "MIMO Wireless Linear Precoding," *IEEE Signal Processing Magazine*, Sept. 2007.

[3] H. Yin, D. Gesbert, M. Filippou, and Y. Liu, "A coordinated approach to channel estimation in large-scale multiple-antenna systems," *IEEE J. on Selected Areas in Comm's (JSAC), Special Issue on: Large Scale Antenna Systems*, Jan. 2013.

[4] A. Adhikary, J. Nam, J.-Y. Ahn, and G. Caire, "Joint Spatial Division and Multiplexing," 2012, arxiv1209.1402.

[5] F. Negro, I. Ghauri, and D. Slock, "Sum Rate Maximization in the Noisy MIMO Interfering Broadcast Channel with Partial CSIT via the Expected Weighted MSE," in *IEEE Int'l Symp. on Wireless Commun. Systems (ISWCS)*, Paris, France, Aug. 2012.

[6] M. Razaviyayn, M. Boroujeni, and Z.-Q. Luo, "A Stochastic Weighted MMSE Approach to Sum Rate Maximization for a MIMO Interference Channel," in *IEEE workshop on Signal Processing Advances in Wireless Communications (SPAWC)*, Darmstadt, Germany, June 2013.

[7] S. Christensen, R. Agarwal, E. Carvalho, and J. Cioffi, "Weighted sum-rate maximization using weighted MMSE for MIMO-BC beamforming design," *IEEE Trans. on Wireless Communications*, vol. 7, no. 12, pp. 4792–4799, December 2008.

[8] F. Negro, S. Shenoy, I. Ghauri, and D. Slock, "On the MIMO Interference Channel," in *Information Theory and Applications Workshop (ITA)*, Feb. 2010.

[9] F. Negro, I. Ghauri, and D. T. M. Slock, "Deterministic Annealing Design and Analysis of the Noisy MIMO Interference Channel," in *Proc. IEEE Information Theory and Applications workshop (ITA)*, San Diego, CA, USA, Feb. 2011.

[10] Q. Shi, M. Razaviyayn, Z.-Q. Luo, and C. He, "An iteratively weighted MMSE approach to distributed sum-utility maximization for a MIMO interfering broadcast channel," *Signal Processing, IEEE Transactions on*, vol. 59, no. 9, pp. 4331 –4340, sept. 2011.

[11] M. Stojnic, H. Vikalo, and B. Hassibi, "Rate maximization in multi-antenna broadcast channels with linear preprocessing," *IEEE Trans. on Wireless Communications*, Sept. 2006.

[12] M. Joham, K. Kusume, M. Gzara, W. Utschick, and J. Nossek, "Transmit Wiener Filter for the Downlink of TDD DS-CDMA Systems," in *Proc. IEEE Int. Symp. on Spread Spectrum Techniques and Applications (ISSSTA)*, 2002.

[13] F. Negro, S. Prasad Shenoy, I. Ghauri, and D. Slock, "On the MIMO Interference Channel," in *Proc. IEEE Information Theory and Applications workshop (ITA)*, San Diego, CA, USA, Feb. 2010.

[14] S.-J. Kim and G. Giannakis, "Optimal Resource Allocation for MIMO Ad Hoc Cognitive Radio Networks," *IEEE Trans. Info. Theory*, May 2011.

[15] M. Sadek, A. Tarighat, and A. Sayed, "A Leakage-Based Precoding Scheme for Downlink Multi-User MIMO Channels," *IEEE Trans. on Signal Proc.*, May 2007.

[16] Y. Lejosne, D. Slock, and Y. Yuan Wu, "User Selection in the MIMO BC," in *European Signal Proc. Conf. (EUSIPCO)*, Bucharest, Romania, Aug. 2012.

[17] J. Dumont, W. Hachem, S. Lasaulce, P. Loubaton, and J. Najim, "On the Capacity Achieving Covariance Matrix for Rician MIMO Channels: An Asymptotic Approach," *IEEE Trans. Info. Theory*, Mar. 2010.

[18] G. Taricco, "Asymptotic mutual information statistics of separately correlated Rician fading MIMO channels," *IEEE Trans. Info. Theory*, Aug. 2008.

[19] C. Martin and B. Ottersten, "Asymptotic Eigenvalue Distributions and Capacity for MIMO Channels Under Correlated Fading," *IEEE Trans. on Signal Proc.*, July 2004.

[20] D. Shutin and B. Fleury, "Sparse Variational Bayesian SAGE Algorithm With Application to the Estimation of Multipath Wireless Channels," *IEEE Trans. on Signal Proc.*, Aug. 2011.

[21] D. Astely and B. Ottersten, "The Effects of Local Scattering on Direction of Arrival Estimation with MUSIC," *IEEE Trans. on Signal Processing*, Dec. 1999.