Cognitive Multi-User MIMO Downlink with Mixed Feedback/Location based Gaussian CSIT

Manijeh Bashar, Dirk Slock
EURECOM, Sophia-Antipolis, France, Email: {bashar,slock}@eurecom.fr

***Abstract*—The Multi-User MIMO downlink or MIMO Broadcast Channel (BC) formulation is relevant for cell center users. Whereas multiple receive antennas do not allow to increase the total number of streams (or degrees of freedom (DoF)) in the BC, they allow the sharing of zero-forcing (ZF) between transmitter and receivers so that a secondary base station (SBS) can serve its secondary users (SU) while ZF beamforming (BF) to primary users (PU). Channel State Information at the Transmitter (CSIT), which is crucial in multi-user systems, is always imperfect in practice, especially for the SBS-PU link. We consider mean and covariance Gaussian partial CSIT, and the special case of a (possibly location based) MIMO Ricean channel model. In this paper we focus on the optimization of beamformers for the secondary expected weighted sum rate (EWSR) under expected PU interference power constraints. We apply a perfect CSI technique, based on a difference of convex functions approach, to a number of deterministic approximations of the EWSR, involving the Massive MIMO limit (large number of transmit antennas), Massive MIMO with a second-order refinement, and the large MIMO limit (both large transmit and receive antenna numbers).**

I. Introduction

Define CBC

In this paper, Tx may denote transmit/transmitter/transmission and Rx may denote receive/receiver/reception. 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. Since Channel State Information at the Tx (CSIT) is more difficult to obtain than the CSIR at the 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).

In [1] we studied a cognitive MISO Interference Channel (IC) with K MISO secondary base station (SBS) - secondary user (SU) pairs and an additional set of L single-antenna Primary Users (PUs). In this paper the objective is to find the set of beamforming (BF) vectors that maximize the Weighted Sum Rate (WSR) of the Cognitive BC (CBC) network, under Tx power constraints for the secondary BS (SBS), and interference level constraints at the primary Rxs. Unfortunately, this problem is non-convex. The proposed solution, which is an iterative algorithm based on alternating optimization of subsets of variables, converges to a local optimum. Deterministic Annealing (DA) could be added as in [2] to find the global optimum. In [1] the alternative problem formulation of SINR balancing is considered.

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 etc.). 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. [3]. The use of covariance CSIT has recently reappeared in the context of Massive MIMO [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. Such CSIT (feedback) reduction is especially crucial for Massive MIMO.

The contributions here are significantly better partial CSIT approaches compared to the EWSMSE approach in [5] (which cannot even be used in the zero channel mean case), 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 go beyond zreo-forcing (ZF) and to introduce a meaningful beamforming design at finite SNR and partial CSIT, for e.g. a finite Ricean factor when not much more than the (location based) LoS information of the PUs is available at the SBS Tx.

II. System Model

We shall focus on MIMO CBC designs in which each user gets one stream since some user selection can make this typically preferable over multiple streams/user. Fig. 1 illustrates the system model, where K SUs with N (or N_k) antennas access the spectrum used by L multiple antenna PUs. Consider downlink transmission with M antennas at the SBS. The SBS designs its beamformers such that it limits the interference power that it causes to the PUs. The $N \times 1$ 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 $N \times M$. We shall assume that the $K \leq M$ signal streams x_i have unit variance and that the noise is white with $\mathbf{v}_k \sim$

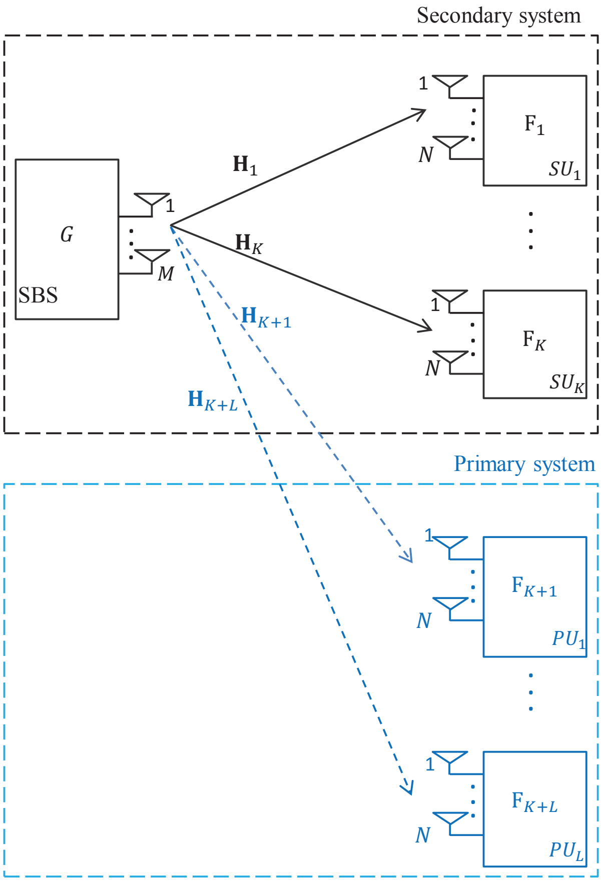

Fig. 1. The Cognitive Broadcast Channel (CBC) system model.

$\mathcal{CN}(0, \sigma_{v,k}^2 I_N)$. We shall assume that the received signal in (1) is rescaled[1] so that the noise variance becomes $\sigma_{v,k}^2 = 1$. The spatial Tx filter or beamformer (BF) is $\mathbf{g}_k$.

III. Cognitive BC (CBC) ZF Feasibility with Multi-Antenna SUs

One key observation we wish to make here is the advantage of the MIMO BC over MISO BC for cognitive radio purposes. Whereas multiple Rx antennas at the SUs do not allow to increase the total number of streams that the SBS can send, in the MIMO case the zero forcing (ZF) of streams can be shared between the SBS and the SUs. In this way, if the SUs are equipped with N antennas, they can cancel $N-1$ of the $K-1$ interfering streams. As a result, the SBS only needs to ZF $K-N$ of the secondary streams, leaving the possibility to ZF to $M-K+N-1$ PU antennas. In the case of Line of Sight (LoS) propagation from the SBS to the PUs, the ZF to PUs introduces only a single constraint per PU, regardless of the number of PU antennas. However, here we shall consider designs beyond ZF also.

[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.

IV. Max WSR with Perfect CSIT

Consider as a starting point for the optimization the weighted sum rate (WSR)

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

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)

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

$\mathbf{R}_k$, $\mathbf{R}_{\overline{k}}$ 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 $\mathbf{f}_k$,

$$\mathbf{f}_k = \mathbf{H}_k \mathbf{R}_{\overline{k}}^{-1} \mathbf{g}_k \,. \tag{4}$$

Note that the use of the non-Hermitian matrices $\mathbf{R}_k$, $\mathbf{R}_{\overline{k}}$ will be of interest in the partial CSIT case, and is made possible due to the properties $\det(I+\mathbf{XY}) = \det(I+\mathbf{YX})$ and $(I+\mathbf{XY})^{-1}\mathbf{X} = \mathbf{X}(I+\mathbf{YX})^{-1}$. Also note that a matrix of the form $\mathbf{H}^H\mathbf{H}(\mathbf{Q}\mathbf{H}^H\mathbf{H}+I_M)^{-1}$ is Hermitian though. The WSR cost function needs to be augmented with the secondary power and primary interference constraints

$$\text{tr}\{\mathbf{Q}\} \leq P \,,\ \ \text{tr}\{\mathbf{Q}\, \mathbf{H}_{K+l}^H \mathbf{H}_{K+l}\} \leq P_l,\, l = 1, \ldots, L \,. \tag{5}$$

In a classical difference of convex functions (DC programming) approach, Kim and Giannakis [7] 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

$$\begin{aligned}
&WSR = u_k \ln\det(\mathbf{R}_{\overline{k}}^{-1} \mathbf{R}_k) + WSR_{\overline{k}} \,, \\
&WSR_{\overline{k}} = \textstyle\sum_{i=1,\neq k}^{K} u_i \ln\det(\mathbf{R}_{\overline{i}}^{-1} \mathbf{R}_i)
\end{aligned} \tag{6}$$

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

$$\begin{aligned}
&WSR_{\overline{k}}(\mathbf{Q}_k, \widehat{\mathbf{Q}}) \approx WSR_{\overline{k}}(\widehat{\mathbf{Q}}_k, \widehat{\mathbf{Q}}) - \text{tr}\{(\mathbf{Q}_k - \widehat{\mathbf{Q}}_k)\widehat{\mathbf{A}}_k\} \\
&\widehat{\mathbf{A}}_k = - \left.\frac{\partial WSR_{\overline{k}}(\mathbf{Q}_k, \widehat{\mathbf{Q}})}{\partial \mathbf{Q}_k}\right|_{\widehat{\mathbf{Q}}_k, \widehat{\mathbf{Q}}} = \sum_{i=1,\neq k}^{K} u_i \mathbf{H}_i^H \mathbf{H}_i (\widehat{\mathbf{R}}_{\overline{i}}^{-1} - \widehat{\mathbf{R}}_i^{-1})
\end{aligned} \tag{7}$$

Note that the linearized (tangent) expression for $WSR_{\overline{k}}$ constitutes a lower bound for it. Now, dropping constant terms, reparameterizing the $Q_k = \mathbf{g}_k \mathbf{g}_k^H$, introducing $P_{L+1} = P$, $\mathbf{D}_{K+L+1} = I$, performing this linearization for all users, and augmenting the WSR cost function with the constraints, we get the Lagrangian

$$
\begin{aligned}
&WSR(\mathbf{g}, \widehat{\mathbf{g}}, \lambda) = \sum_{l=1}^{L+1} \lambda_l P_l + \\
&\sum_{k=1}^{K} u_k \ln(1 + \mathbf{g}_k^H \widehat{\mathbf{B}}_k \mathbf{g}_k) - \mathbf{g}_k^H (\widehat{\mathbf{A}}_k + \sum_{l=1}^{L+1} \lambda_l \mathbf{D}_l) \mathbf{g}_k
\end{aligned} \tag{8}
$$

where

$$
\widehat{\mathbf{B}}_k = \mathbf{H}_k^H \mathbf{H}_k \widehat{\mathbf{R}}_{\overline{k}}^{-1}\ , \quad \mathbf{D}_l = \mathbf{H}_{K+l}^H \mathbf{H}_{K+l} \tag{9}
$$

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

$$
\widehat{\mathbf{B}}_k \, \mathbf{g}_k = \frac{1 + \mathbf{g}_k^H \widehat{\mathbf{B}}_k \mathbf{g}_k}{u_k} \, (\widehat{\mathbf{A}}_k + \sum_{l=1}^{L+1} \lambda_l \mathbf{D}_l) \mathbf{g}_k \tag{10}
$$

or hence $\mathbf{g}_k^{'} = V_{max}(\widehat{\mathbf{B}}_k, \widehat{\mathbf{A}}_k + \sum_l \lambda_l \mathbf{D}_l)$ is the (normalized) "max" generalized eigenvector of the two indicated matrices, with max eigenvalue $\sigma_k = \sigma_{max}(\widehat{\mathbf{B}}_k, \widehat{\mathbf{A}}_k + \sum_l \lambda_l \mathbf{D}_l)$. Let $\sigma_k^{(1)} = \mathbf{g}_k^{'H} \widehat{\mathbf{B}}_k \mathbf{g}_k^{'}$, $\sigma_k^{(2)} = \mathbf{g}_k^{'H} \widehat{\mathbf{A}}_k \mathbf{g}_k^{'}$ and $\sigma_{k,l} = \mathbf{g}_k^{'H} \mathbf{D}_l \mathbf{g}_k^{'}$. The advantage of formulation (8) is that it allows straightforward power adaptation: introducing stream powers $p_k \geq 0$ and substituting $\mathbf{g}_k = \sqrt{p_k}\, \mathbf{g}_k^{'}$ in (8) yields

$$
WSR = \sum_l \lambda_l P_l + \sum_{k=1}^{K} \{u_k \ln(1 + p_k \sigma_k^{(1)}) - p_k (\sigma_k^{(2)} + \sum_l \lambda_l \sigma_{k,l})\} \tag{11}
$$

which leads to the following interference leakage aware water filling

$$
p_k = \left(\frac{u_k}{\sigma_k^{(2)} + \sum_l \lambda_l \sigma_{k,l}} - \frac{1}{\sigma_k^{(1)}} \right)^+ . \tag{12}
$$

For a given vector $\lambda = [\lambda_1 \cdots \lambda_{L+1}]^T$, $\mathbf{g}$ needs to be iterated till convergence. And λ can be found by duality (line search):

$$
\min_{\lambda \geq 0} \max_{\mathbf{g}} \sum_l \lambda_l P_l + \sum_k \{u_k \ln\det(\mathbf{R}_{\overline{k}}^{-1} \mathbf{R}_k) - p_k \sum_l \lambda_l \sigma_{k,l}\}. \tag{13}
$$

Note that for each value of λ, a maximization over $\mathbf{g}^{'}$ is required, which means a maximization over both the $\mathbf{g}_k^{'}$ and the powers p_k. Since the optimization over $\mathbf{g}^{'}$ is iterative anyway, the iterations can be simplified by freezing the $\mathbf{g}_k^{'}$ and just performing $\min_{\lambda \geq 0} \max_{\mathbf{p}}$ in (13). This corresponds to taking the solution for the powers p_k in (12) and increasing the consecutive λ_l beyond zero if so required to satisfy the power/interference constraints (5). This can be done by the ellipsoid algorithm [7] or with a greedy approach that focuses on the constraint violation terms in order of decreasing sensitivity. In the non-cognitive case ($L = 0$, only one power constraint) this becomes the usual water filling procedure. Note that as with any alternating optimization procedure, there are many updating schedules possible, with different impact on convergence speed. The quantities to be updated are the $\mathbf{g}_k^{'}$, the p_k and the λ_l.

V. Mean and Covariance Gaussian CSIT

In this section we drop the user index k for simplicity. The separable correlation model is

$$
\mathbf{H} = \overline{\mathbf{H}} + \mathbf{C}_r^{1/2} \, \widetilde{\mathbf{H}} \, \mathbf{C}_t^{1/2} \tag{14}
$$

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

$$
\begin{aligned}
\mathrm{E}(\mathbf{H} - \overline{\mathbf{H}})(\mathbf{H} - \overline{\mathbf{H}})^H &= \mathrm{tr}\{\mathbf{C}_t\}\ \mathbf{C}_r \\
\mathrm{E}(\mathbf{H} - \overline{\mathbf{H}})^H(\mathbf{H} - \overline{\mathbf{H}}) &= \mathrm{tr}\{\mathbf{C}_r\}\ \mathbf{C}_t
\end{aligned} \tag{15}
$$

and the elements of $\widetilde{\mathbf{H}}$ are i.i.d. $\sim \mathcal{CN}(0, 1)$. It is also of interest to consider the total Tx side correlation matrix

$$
\mathbf{R}_t = \mathrm{E}\,\mathbf{H}^H \mathbf{H} = \overline{\mathbf{H}}^H \overline{\mathbf{H}} + \mathrm{tr}\{\mathbf{C}_r\}\ \mathbf{C}_t\ . \tag{16}
$$

A. Location Aided Partial CSIT LoS Channel Model

Assuming the SBS disposes of not much more than the LoS component information of PUs (and possibly SUs), consider the following MIMO channel model

$$
\mathbf{H} = \mathbf{h}_r \, \mathbf{h}_t^H(\theta) + \widetilde{\mathbf{H}}^{'} \tag{17}
$$

where θ is the LoS AoD and the SBS side array response is normalized: $||\mathbf{h}_t(\theta)||^2 = 1$. We shall model the unknown Rx side LoS array response $\mathbf{h}_r$ as a vector of i.i.d. complex Gaussian variables

$$
\begin{aligned}
&\mathbf{h}_r \ \text{ i.i.d. } \sim \mathcal{CN}(0, \tfrac{\mu}{\mu+1}) \quad \text{and} \\
&\widetilde{\mathbf{H}}^{'} \ \text{ i.i.d. } \sim \mathcal{CN}(0, \tfrac{1}{\mu+1} \tfrac{1}{M})\ , \text{ independent of } \mathbf{h}_r,
\end{aligned} \tag{18}
$$

where the matrix $\widetilde{\mathbf{H}}$ represents the aggregate NLoS components. Note that $(\mathrm{E}||\mathbf{h}_r \, \mathbf{h}_t^T(\theta)||_F^2)/(\mathrm{E}||\widetilde{\mathbf{H}}^{'}||_F^2) = \mu$ can be considered as a Rice factor. In fact the only parameter additional to the LoS AoD θ assumed in (17) is μ. So, this is a case of zero mean CSIT and Tx side covariance CSIT

$$
\mathbf{R}_t = \mathrm{E}\,\mathbf{H}^H \mathbf{H} = \frac{\mu N}{\mu + 1} \mathbf{h}_t(\theta) \mathbf{h}_t^H(\theta) + \frac{N}{\mu + 1} \frac{1}{M} I_M\ . \tag{19}
$$

VI. Expected WSR (EWSR)

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

$$
EWSR(\mathbf{g}) = \mathrm{E}_{\mathbf{H}} \sum_k u_k \ln(1 + \mathbf{g}_k^H \mathbf{H}_k^H \mathbf{H}_k \mathbf{R}_{\overline{k}}^{-1} \mathbf{g}_k) \tag{20}
$$

where we now underlign the dependence of various quantities on $\mathbf{H}$. The EWSR in (20) corresponds to perfect CSIR since the optimal Rx filters $\mathbf{f}_k$ as a function of the aggregate $\mathbf{H}$ have been substituted, namely $WSR(\mathbf{g}, \mathbf{H}) = \max_{\mathbf{f}} \sum_k u_k(-\ln(e_k(\mathbf{f}_k, \mathbf{g})))$. For the PU interference constraints, we shall consider also the expected interference. Hence we shall now have $\mathbf{D}_l = \mathrm{E}\,\mathbf{H}_{K+l}^H \mathbf{H}_{K+l} = \mathbf{R}_{t,K+l}$.

At high SNR, max EWSR attempts ZF and we get:

Theorem 1: **Sufficiency of Incomplete CSIT for Full DoF in the MIMO CBC** In the MIMO CBC with perfect CSIR, it is sufficient that for each of K users $\text{rank}\,(\mathbf{R}_{t,k}) \leq N_k$ and that the BS knows any vector $\mathbf{h}_k \in \text{Range}\,(\mathbf{R}_{t,k})$ (as long as the K resulting vectors $\mathbf{h}_k$ are linearly independent) and the column space of $[\mathbf{R}_{t,K+1} \cdots \mathbf{R}_{t,K+L}]$ in order for ZF BF to produce $K = \max(0, M - \text{rank}([\mathbf{R}_{t,K+1} \cdots \mathbf{R}_{t,K+L}]))$ interference free streams (degrees of freedom (DoF)). $\square$

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 (Weighted Sum MSE) 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. Deterministic annealing can be used as in [2] for a deterministic algorithm as in Section IV to track the global optimum from SNR ≈ 0 (where the solution is clear analytically) to the desired SNR. This is essentially a homotopy method in which the problem gets resolved for an SNR that increases in small steps. At each higher SNR, the global optimum will be in the region of attraction of the global optimum at the lower SNR.

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

VII. Massive MIMO Limit

If the number of Tx antennas M becomes very large, then quantities of the form $\mathbf{H}_k^H \mathbf{H}_k$ converge to their mean (LLN). Hence in the Massive MIMO limit, the EWSR gets maximized by the algorithm in Section IV by replacing $\mathbf{H}_k^H \mathbf{H}_k$ by $\mathbf{R}_{t,k}$ everywhere.

VIII. Massive MIMO 2nd Order Refinement

This second-order refinement was first suggested in [8, eq. (15)]. Consider the second-order Taylor series expansion [9, p. 644]

$$\ln\det(\mathbf{X}+\mathbf{Y}) \approx \ln\det(\mathbf{X}) + \text{tr}\{\mathbf{X}^{-1}\mathbf{Y}\} - \frac{1}{2}\text{tr}\{\mathbf{X}^{-1}\mathbf{Y}\mathbf{X}^{-1}\mathbf{Y}\} \tag{21}$$

assuming small $\mathbf{X}^{-1}\mathbf{Y}$. We shall take $\mathbf{X}+\mathbf{Y} = I + \mathbf{Q}\mathbf{H}^H\mathbf{H}$, $\mathbf{X} = I + \mathbf{Q}\mathbf{R}_t$, $\mathbf{Y} = \mathbf{Q}(\mathbf{H}^H\mathbf{H} - \mathbf{R}_t)$ with $\mathbf{H}$ as in (14), and assuming $\text{tr}\{\mathbf{C}_r\} = 1$. Hence we get

$$\begin{aligned} &\mathrm{E}_{\mathbf{H}} \ln\det(I + \mathbf{Q}\mathbf{H}^H\mathbf{H}) \approx \ln\det(I + \mathbf{Q}\mathbf{R}_t) \\ &-\tfrac{1}{2}\,\mathrm{E}_{\mathbf{H}}\text{tr}\{\mathbf{A}(\mathbf{H}^H\mathbf{H} - \mathbf{R}_t)\mathbf{A}(\mathbf{H}^H\mathbf{H} - \mathbf{R}_t)\} \end{aligned} \tag{22}$$

where $\mathbf{A} = (I + \mathbf{Q}\mathbf{R}_t)^{-1}\mathbf{Q}$. Using 4th order Gaussian moments [10] in (22) and using also becomes

$$\begin{aligned} &\mathrm{E}_{\mathbf{H}} \ln\det(I + \mathbf{Q}\mathbf{H}^H\mathbf{H}) \approx \ln\det(I + \mathbf{Q}\mathbf{R}_t) \\ &-\tfrac{1}{2}\text{tr}\{(\overline{\mathbf{H}}\mathbf{A}\overline{\mathbf{H}}^H + \text{tr}\{\mathbf{A}\mathbf{C}_t\}\mathbf{C}_r)^2\} + \tfrac{1}{2}\text{tr}\{\mathbf{A}\overline{\mathbf{H}}^H\overline{\mathbf{H}}\mathbf{A}\overline{\mathbf{H}}^H\overline{\mathbf{H}}\}\,. \end{aligned} \tag{23}$$

For the case $\mathbf{C}_r = \frac{1}{N} I_N$ and $\overline{\mathbf{H}} = 0$ (hence $\mathbf{R}_t = \mathbf{C}_t$), this becomes

$$\begin{aligned} &\mathrm{E}_{\mathbf{H}} \ln\det(I + \mathbf{Q}\mathbf{H}^H\mathbf{H}) \approx \ln\det(I + \mathbf{Q}\mathbf{C}_t) \\ &-\tfrac{1}{2N}\left(\text{tr}\{(I + \mathbf{Q}\mathbf{C}_t)^{-1}\mathbf{Q}\mathbf{C}_t\}\right)^2 \end{aligned} \tag{24}$$

with differential

$$\begin{aligned} &d(\,\mathrm{E}_{\mathbf{H}} \ln\det(I + \mathbf{Q}\mathbf{H}^H\mathbf{H})) \approx \text{tr}\{\mathbf{C}_t(I + \mathbf{Q}\mathbf{C}_t)^{-1}\,d\mathbf{Q}\} \\ &-\tfrac{1}{N}\text{tr}\{(I + \mathbf{Q}\mathbf{C}_t)^{-1}\mathbf{Q}\mathbf{C}_t\}\text{tr}\{\mathbf{C}_t(I + \mathbf{Q}\mathbf{C}_t)^{-2}\,d\mathbf{Q}\} \\ &= \text{tr}\{\mathbf{C}_t\overline{\mathbf{R}}^{-1}\,d\mathbf{Q}\} - \tfrac{1}{N}\text{tr}\{\overline{\mathbf{R}}^{-1}\mathbf{Q}\mathbf{C}_t\}\text{tr}\{\mathbf{C}_t\overline{\mathbf{R}}^{-2}\,d\mathbf{Q}\} \end{aligned} \tag{25}$$

with $\mathbf{R} = I + \mathbf{Q}\mathbf{H}^H\mathbf{H}$ and $\overline{\mathbf{R}} = \mathrm{E}\,\mathbf{R} = I + \mathbf{Q}\mathbf{C}$. When applying the RHS of (24) to the signal part in (6), we shall count the $\ln\det()$ term in the signal and the correction term in the interference plus noise. We can now apply the algorithm in Section IV with the following conventions:

$$\begin{aligned} &\widehat{\mathbf{B}}_k = \mathbf{C}_k\,\widehat{\overline{\mathbf{R}}}_{\overline{k}}^{-1} \text{ with } \widehat{\overline{\mathbf{R}}}_k = I + \widehat{\mathbf{Q}}\mathbf{C}_k,\; \widehat{\overline{\mathbf{R}}}_{\overline{k}} = I + \widehat{\mathbf{Q}}_{\overline{k}}\mathbf{C}_k \\ &\widehat{\mathbf{A}}_k = u_k\alpha_k\mathbf{C}_k\widehat{\overline{\mathbf{R}}}_k^{-2} + \sum_{i=1,\neq k}^{K} u_i\mathbf{C}_i(\widehat{\overline{\mathbf{R}}}_{\overline{i}}^{-1} - \widehat{\overline{\mathbf{R}}}_i^{-1} + \alpha_i\widehat{\overline{\mathbf{R}}}_i^{-2} - \alpha_{\overline{i}}\widehat{\overline{\mathbf{R}}}_{\overline{i}}^{-2}) \end{aligned} \tag{26}$$

and $\alpha_i = \frac{\text{tr}\{\widehat{\overline{\mathbf{R}}}_i\widehat{\mathbf{Q}}\mathbf{C}_i\}}{N}$, $\alpha_{\overline{i}} = \frac{\text{tr}\{\widehat{\overline{\mathbf{R}}}_{\overline{i}}\widehat{\mathbf{Q}}_{\overline{i}}\mathbf{C}_i\}}{N}$ and $\mathbf{C}_k = \mathbf{C}_{t,k}$. Note that the Massive MIMO limit is obtained by putting all $\alpha = 0$.

IX. Large MIMO Asymptotics

The large MIMO asymptotics from [11], [12], in which both $M, N \to \infty$ at constant ratio, tend to give more precise approximations when M is not so large. For the general case of Gaussian CSIT with separable (Kronecker) covariance structure, [11], [12] lead to asymptotic expressions of the form

$$\begin{aligned} &\mathrm{E}_{\mathbf{H}} \ln\det(I + \mathbf{H}\mathbf{Q}\mathbf{H}^H) \\ &= \max_{z\geq 0,\, w\geq 0} \left\{ \ln\det \begin{bmatrix} I + w\mathbf{C}_r & \overline{\mathbf{H}} \\ -\mathbf{Q}\overline{\mathbf{H}}^H & I + z\mathbf{Q}\mathbf{C}_t \end{bmatrix} - zw \right\}. \end{aligned} \tag{27}$$

where the maximization over z and w should be carried out alternatingly (and not jointly: the joint optimization may correspond to a global maximum or a saddle point. The cost function is concave however in z or w separately). 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 $\mathbf{C}_k$, the EWSR can be rewritten with large MIMO asymptotics as

$$\begin{aligned} &EWSR = \\ &\sum_{k=1}^{K} \Big\{ u_k \max_{z_k, w_k} [\ln\det(I + z_k\mathbf{Q}\mathbf{C}_k) + N\ln(1+w_k) - z_k w_k] \\ &-u_k \max_{z_{\overline{k}}, w_{\overline{k}}} \left[\ln\det(I + z_{\overline{k}}\mathbf{Q}_{\overline{k}}\mathbf{C}_k) + N\ln(1+w_{\overline{k}}) - z_{\overline{k}}w_{\overline{k}}\right] \Big\}. \end{aligned} \tag{28}$$

This criterion can be used to evaluate the EWSR for given $\mathbf{Q}$. It can also be used to optimize $\mathbf{Q}$, in which case we can apply again the algorithm of Section IV, with the following conventions:

$$\begin{aligned} &\widehat{\widehat{\mathbf{R}}}_k(z) = I + z\,\widehat{\mathbf{Q}}\,\mathbf{C}_k\,,\ \widehat{\widehat{\mathbf{R}}}_{\overline{k}}(z) = I + z\,\widehat{\mathbf{Q}}_{\overline{k}}\,\mathbf{C}_k \\ &\widehat{\mathbf{B}}_k = z_k \mathbf{C}_k \widehat{\widehat{\mathbf{R}}}_{\overline{k}}^{-1}(z_k)\,,\ \widehat{\mathbf{A}}_k = \sum_{i\neq k} u_i \mathbf{C}_i \big(z_{\overline{i}} \widehat{\widehat{\mathbf{R}}}_{\overline{i}}^{-1}(z_{\overline{i}}) - z_i \widehat{\widehat{\mathbf{R}}}_i^{-1}(z_i)\big) \end{aligned} \tag{29}$$

where z_k, $z_{\overline{k}}$ are obtained from

$$\begin{aligned} &\max_{z_k, w_k} g(z_k, w_k, \mathbf{Q}, \mathbf{C}_k)\,,\ \max_{z_{\overline{k}}, w_{\overline{k}}} g(z_{\overline{k}}, w_{\overline{k}}, \mathbf{Q}_{\overline{k}}, \mathbf{C}) \quad \text{where} \\ &g(z, w, \mathbf{Q}, \mathbf{C}) = \ln\det(I + z\mathbf{Q}\mathbf{C}) + N\ln(1+w) - z\,w\ . \end{aligned} \tag{30}$$

For these optimizations, we get from $\partial g/\partial w = 0$ that $z = N/(1+w)$. From this and $\partial g/\partial z = 0$ we get

$$w = f(w) = \frac{1+w}{N}\mathrm{tr}\{(\frac{1+w}{N}I_M + \mathbf{Q}\mathbf{C})^{-1}\mathbf{Q}\mathbf{C}\} \tag{31}$$

The curves $y = w$ and $y = f(w)$ have a unique intersection in the first quadrant, with $y = f(w)$ lying initially above $y = w$. Hence the optimal w can be found by iterating $w^{(i)} = f(w^{(i-1)})$. The first time, one can initialize with $w^{(0)} = 0$. In the iterative algorithm from Section IV, w can be initialized with the value obtained in the previous iteration for $\mathbf{g}$. The corresponding optimal z is then $z = N/(1+w)$.

X. Numerical Results

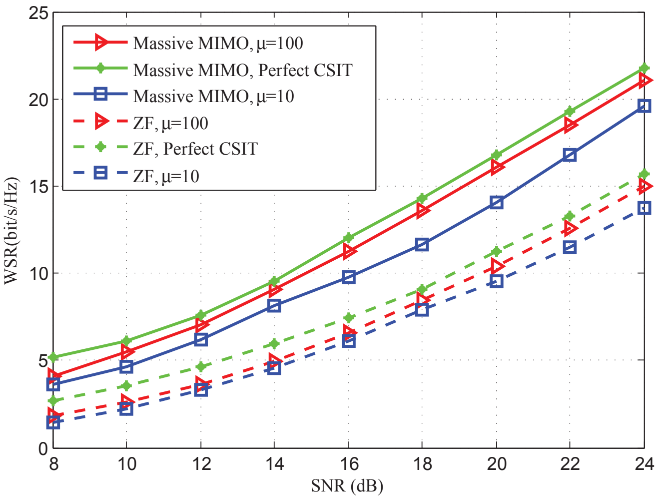

Fig. 2. EWSR vs SNR for $M = N = 4$, $K = 2$, $L = 0$.

In these initial simulations, we consider just the regular MIMO BC, with CSIT based on the MIMO Ricean channel model (hence the CSIT comprises the (downlink) Tx side LoS antenna array response $\mathbf{h}_t$ and the Rice factor μ, both of which can be estimated from the uplink channel). In Fig. 2 the expected sum rate (SR) is plotted versus SNR for $M = N = 4$ with $K = 2$, $L = 0$. For the Tx design, we consider either ZF on the LoS component, with uniform power loading, or an optimized design based on the Massive MIMO limit. For each design, three cases of Rice factor are considered: $\mu = 10, 100, \infty$ (this last case is labeled "Perfect CSIT" in the figure). The expected SR is obtained by averaging over channel realizations, according to the Ricean distribution with respectively one of the three possible values for μ.

XI. Conclusions

We have studied Expected WSR maximization in the MU-MIMO cognitive radio Broadcast Channel, for both perfect and partial CSIT. The paper introduced a beamforming design at finite SNR and finite Ricean factor when location based LoS CSIT of PUs is available at the SBS. A simple solution can be obtained through Massive MIMO asymptotics. Refinements based on a second-order logdet expansion or on large MIMO asymptotics were also introduced. All these approaches lead to variations of a basic difference of convex functions approach to which deterministic annealing can straightforwardly be applied to find the global optimum. Furture work will include a more extensive evaluation of the proposed approaches.

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#.

References

[1] F. Negro, M. Cardone, I. Ghauri, and D.T.M. Slock, "SINR Balancing and Beamforming for the MISO Interference Channel," in *Proc. IEEE Int'l Symp. Personal Indoor and Mobile Radio Communications (PIMRC)*, Sept. 2011.

[2] Francesco Negro, Irfan Ghauri, and Dirk 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.

[3] R. de Francisco and D.T.M. 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.

[4] 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.

[5] F. Negro, I. Ghauri, and D.T.M. 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.S. 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.-J. Kim and G.B. Giannakis, "Optimal Resource Allocation for MIMO Ad Hoc Cognitive Radio Networks," *IEEE Trans. Info. Theory*, May 2011.

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

[9] S. Boyd and L. Vandenberghe, *Convex Optimization*, Cambridge Univ. Press, Cambridge, U.K., 2003.

[10] P.H.M. Janssen and P. Stoica, "On the Expectation of the Product of Four Matrix-Valued Gaussian Random Variables," *IEEE Trans. Automatic Control*, Sept. 1988.

[11] 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.

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