MaMISO IBC Beamforming Design with Combined Channel Estimate and Covariance CSIT: a Large System Analysis

Wassim Tabikh[¶‡] Dirk Slock[¶] Yi Yuan-Wu[‡]
[¶]EURECOM, BIOT Sophia-Antipolis, France, Email:{tabikh,slock}@eurecom.fr
[‡]Orange Labs, Châtillon, France, Email:{wassim.tabikh,yi.yuan}@orange.com

***Abstract*—This work deals with beamforming for the MaMISO Interfering Broadcast Channel (IBC), i.e. the Massive Multi-Input Single-Output (MaMISO) Multi-User Multi-Cell downlink (DL). The novel beamformers are here optimized for the Expected Weighted Sum Rate (EWSR) for the case of Partial Channel State Information at the Transmitters (CSIT). Gaussian partial CSIT can optimally combine channel estimate and channel covariance information. We introduce the first large system analysis for optimized beamformers with partial CSIT. In the case of Gaussian partial CSIT, the beamformers only depend on the means and covariances of the channels. The large system analysis furthermore allows to predict the EWSR performance on the basis of the channel statistics only.**

***Index Terms*—Massive MIMO; multi-user; multi-cell; sum rate; beamforming; partial CSIT; large system analysis**

I. Introduction

In this work, Tx may denote transmit/transmitter/ transmission and Rx may denote receive/receiver/reception. Interference is the main limiting factor in wireless transmission. Base stations (BSs) disposing of multiple antennas are able to serve multiple User Equipments (UEs) simultaneously. Two popular approaches for maximum Weighted Sum Rate (WSR) beamformer (BF) optimization are the exploitation of the WSR-WSMSE (Weighted Sum MSE) correspondence as in [10] or the concave WSR minorization approach as in [9]. In the MISO case, the BF are proportional in both cases, but the stream powers are optimized with an interference aware waterfilling algorithm in [9] whereas [10] alternates between DL and dual UL MMSE updates for Rx/Tx. In the MIMO case, the BF in [9] are found as generalized eigenvectors resulting from optimal Signal-to-Leakage-plus-Noise (SLNR) considerations, whereas the DL/UL ping-pong MMSE updates in [10] correspond to power method iterations for these generalized eigenvectors.

However, Multi-User (MU) systems have precise requirements for CSIT which is more difficult to acquire than CSI at the Rx (CSIR). Hence we focus here on the more challenging downlink (DL) and we consider maximizing the Expected WSR (EWSR) with partial CSIT. Earlier works have attempted optimal partial CSIT designs for MU MIMO, e.g. the Expected WSMSE (EWSMSE) approach applied in [1] for MIMO IBC beamformer (BF) design, based on an extension of [10] to the partial CSIT case. However, the EWSMSE approach is suboptimal and cannot even be used in the zero channel mean case (case of covariance CSIT only). In spite of that, it has been mistakenly considered as optimal as recently as in [2]. It is also desirable to have deterministic alternatives to the cumbersome (though optimal) stochastic approximation solution of [3]. We treat the Gaussian CSIT case, optimally combining mean (channel estimate) and (channel) covariance information. The Gaussian model allows to exploit both mean (channel estimate) and covariance information, which are actually the only statistics that matter in the large system limit, regardless of actual channel (estimate) didstribution. The goal here is to go beyond the extreme of Zero-Forcing (ZF) and to introduce a meaningful BF design at finite SNR and with partial CSIT. Over the last year or so, a number of research works have proposed to exploit the channel hardening in Massive MIMO (MaMIMO) to reduce global instantaneous CSIT requirements to local instantaneous CSIT plus global statistical CSIT. This only works for MISO systems [4] in which all the work needs to be done by the transmitters. We may remark that in the case of MIMO, in which UEs possess a limited number of antennas which contribute actively (e.g. to Zero-Forcing (ZF)/Interference Alignment (IA) at high SNR), the interference subspace and hence the receiver at the UEs does not harden.

A significant push for large system analysis in MaMIMO systems appeared in [5]. It allows to obtain deterministic (instead of fast fading channel dependent) expressions for various scalar quantities, facilitating the analysis and design of wireless systems. E.g. it may allow to evaluate beamforming performance without computing explicit beamformers. The analysis in [5] allowed e.g. the determination of the optimal regularization factor in Regularized ZF (R-ZF) BF, both with perfect and partial CSIT. A little known extension appeared in [6] for optimal beamformers, but only for the perfect CSIT MISO BC case. Some other extensions appeared recently in [7] or [8] where MISO IBC is considered with perfect CSIT and weighted R-ZF BF, with two optimized weight levels, for intracell or intercell interference. The contributions of this paper are (see also the Simulations and the Conclusions):

- A novel optimal (max EWSR) beamforming design for MISO IBC with partial CSIT, an extension of the concave

WSR minorization approach in [9] to the partial CSIT case. Whereas this approach finds BF's as generalized eigenvectors of higher rank matrices, also in the case of MISO with partial CSIT, we propose Jacobi iterations for reduced computational complexity and simplified large system analysis.
- A novel large system analysis of the proposed BF design, that constitutes the first large system analysis of optimized BF design with partial CSIT.

II. The IBC Signal Model

In the rest of this paper we shall consider a per stream approach (which in the perfect CSI case would be equivalent to per user). In an IBC formulation, one stream per user can be expected to be the usual scenario. In the development below, in the case of more than one stream per user, we shall treat each stream as an individual user. So, consider again an IBC with C cells with a total of K users. We shall consider a system-wide numbering of the users. User k is served by BS b_k. We shall initially consider users equipped with N_k antennas but the partial CSIT developments and associated large system analysis will be valid only for the MISO case ($N_k = 1$). The $N_k \times 1$ received signal at user k in cell b_k is

$$y_k = \underbrace{H_{k,b_k}\, g_k\, x_k}_{\text{signal}} + \underbrace{\sum_{\substack{i\neq k\\ b_i=b_k}} H_{k,b_k}\, g_i\, x_i}_{\text{intracell interf.}} + \underbrace{\sum_{j\neq b_k}\sum_{i:b_i=j} H_{k,j}\, g_i\, x_i}_{\text{intercell interf.}} + v_k \quad (1)$$

where x_k is the intended (white, unit variance) scalar signal stream, H_{k,b_k} is the $N_k \times M_{b_k}$ channel from BS b_k to user k. BS b_k serves $K_{b_k} = \sum_{i:b_i=b_k} 1$ users. We considering a noise whitened signal representation so that we get for the noise $v_k \sim \mathcal{CN}(0, I_{N_k})$. The $M_{b_k} \times 1$ spatial Tx filter or beamformer (BF) is g_k. Treating interference as noise, user k will apply a linear Rx filter f_k to maximize the signal power (diversity) while reducing any residual interference that would not have been (sufficiently) suppressed by the BS Tx. The Rx filter output is $\widehat{x}_k = f_k^H y_k$

$$\begin{aligned}\widehat{x}_k &= f_k^H H_{k,b_k}\, g_k\, x_k + \sum_{i=1,\neq k}^{K} f_k^H H_{k,b_i}\, g_i\, x_i + f_k^H v_k \\ &= f_k^H h_{k,k}\, x_k + \sum_{i\neq k} f_k^H h_{k,i}\, x_i + f_k^H v_k\end{aligned} \quad (2)$$

where $h_{k,i} = H_{k,b_i}\, g_i$ is the channel-Tx cascade vector.

III. Max WSR with Perfect CSIT

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

$$WSR = WSR(g) = \sum_{k=1}^{K} u_k \ \ln\frac{1}{e_k} \quad (3)$$

where g represents the collection of BFs g_k, the $e_k = e_k(g)$ are the Minimum Mean Squared Errors (MMSEs) for estimating the x_k:

$$\begin{aligned}&\frac{1}{e_k} = 1 + g_k^H H_{k,b_k}^H R_{\overline{k}}^{-1} H_{k,b_k} g_k = (1 - g_k^H H_{k,b_k}^H R_k^{-1} H_{k,b_k} g_k)^{-1}\\ &R_k = H_{k,b_k} Q_k\, H_{k,b_k}^H + R_{\overline{k}}\,,\ Q_i = g_i g_i^H\,,\\ &R_{\overline{k}} = \sum_{i\neq k} H_{k,b_i} Q_i\, H_{k,b_i}^H + I_{N_k}\,.\end{aligned} \quad (4)$$

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

$$f_k = R_k^{-1} H_{k,b_k} g_k\ . \quad (5)$$

The WSR cost function needs to be augmented with the power constraints

$$\sum_{k:b_k=j} \mathrm{tr}\{Q_k\} \leq P_j\ . \quad (6)$$

In a classical difference of convex functions (DC programming) approach, Kim and Giannakis [10] propose to keep the concave signal terms and to replace the convex interference terms by the linear (and hence concave) tangent approximation. Note that it is more appropriate to consider this DC programming as an instance of a minorization approach because whereas the linearization is carried out w.r.t. the Tx covariance matrices Q_k, the resulting problem then gets reparameterized in terms of the BF's g_k, and the minorization is insensitive to the parameterization. More specifically, consider the dependence of WSR on Q_k alone. Then

$$\begin{aligned}WSR &= u_k \ln\det(R_{\overline{k}}^{-1} R_k) + WSR_{\overline{k}}\,,\\ WSR_{\overline{k}} &= \textstyle\sum_{i=1,i\neq k}^{K} u_i \ln\det(R_{\overline{i}}^{-1} R_i)\end{aligned} \quad (7)$$

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

$$\begin{aligned}&WSR_{\overline{k}}(Q_k, \hat{Q}) \approx WSR_{\overline{k}}(\hat{Q}_k, \hat{Q}) - \mathrm{tr}\{(Q_k - \hat{Q}_k)\hat{A}_k\}\\ &\hat{A}_k = -\left.\frac{\partial WSR_{\overline{k}}(Q_k,\hat{Q})}{\partial Q_k}\right|_{\hat{Q}_k,\hat{Q}} = \sum_{i\neq k}^{K} u_i H_{i,b_k}^H (\hat{R}_{\overline{i}}^{-1} - \hat{R}_i^{-1}) H_{i,b_k}\end{aligned} \quad (8)$$

Note that the linearized (tangent) expression for $WSR_{\overline{k}}$ constitutes a lower bound for it. Now, dropping constant terms, reparameterizing the $Q_k = g_k g_k^H$, performing this linearization for all users, and augmenting the WSR cost function with the constraints, we get the Lagrangian $WSR(g, \hat{g}, \lambda)$

$$= \sum_{j=1}^{C} \lambda_j P_j + \sum_{k=1}^{K} u_k \ln(1 + g_k^H \hat{B}_k g_k) - g_k^H (\hat{A}_k + \lambda_{b_k} I) g_k \quad (9)$$

$$\text{where}\ \ \hat{B}_k = H_{k,b_k}^H \hat{R}_{\overline{k}}^{-1} H_{k,b_k}\ . \quad (10)$$

The gradient (w.r.t. g_k) 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

$$\hat{B}_k\, g_k = \frac{1+g_k^H \hat{B}_k g_k}{u_k}\,(\hat{A}_k + \lambda_{b_k} I) g_k \tag{11}$$

or hence $g_k^{'} = V_{max}(\hat{B}_k, \hat{A}_k + \lambda_{b_k} I)$ is the (normalized) "max" generalized eigenvector of the two indicated matrices, with max eigenvalue $\sigma_k = \sigma_{max}(\hat{B}_k, \hat{A}_k + \lambda_{b_k} I)$. Let $\sigma_k^{(1)} = g_k^{'H} \hat{B}_k g_k^{'}$, $\sigma_k^{(2)} = g_k^{'H} \hat{A}_k g_k^{'}$. The advantage of formulation (9) is that it allows straightforward power adaptation: introducing stream powers $p_k \geq 0$ and substituting $g_k = \sqrt{p_k}\, g_k^{'}$ in (9) yields

$$WSR = \sum_{j=1}^{C} \lambda_j P_j + \sum_{k=1}^{K} \{u_k \ln(1+p_k \sigma_k^{(1)}) - p_k(\sigma_k^{(2)} + \lambda_{b_k})\}$$

which leads to the following interference leakage ($\sigma_k^{(2)}$) aware water filling

$$p_k = \left(\frac{u_k}{\sigma_k^{(2)} + \lambda_{b_k}} - \frac{1}{\sigma_k^{(1)}} \right)^{+} \tag{12}$$

where the Lagrange multipliers are adjusted to satisfy the power constraints $\sum_{k:b_k=j} p_k = P_j$. With $\sigma_k^{(2)} = 0$ this would be standard waterfilling. The minorization approach is crucial for this waterfilling power optimization.

IV. Joint Mean and Covariance Gaussian CSIT

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

$$H = \overline{H} + C_r^{1/2}\, \widetilde{H}\, C_t^{1/2} \tag{13}$$

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

$$\begin{array}{l} \mathrm{E}(H-\overline{H})(H-\overline{H})^H = \mathrm{tr}\{C_t\}\, C_r \\ \mathrm{E}(H-\overline{H})^H(H-\overline{H}) = \mathrm{tr}\{C_r\}\, C_t \end{array} \tag{14}$$

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

$$S_t = \mathrm{E}\, H^H H = \overline{H}^H \overline{H} + \mathrm{tr}\{C_r\}\, C_t\ . \tag{15}$$

V. MaMIMO Limit

If the number of Tx antennas M becomes very large, we get a convergence for any quadratic term of the form

$$HQH^H \stackrel{M\rightarrow\infty}{\longrightarrow} \mathrm{E}_H\, HQH^H = \overline{H} Q \overline{H}^H + \mathrm{tr}\{QC_t\}\, C_r\ . \tag{16}$$

and hence we get the following MaMIMO limit matrices

$$\begin{array}{l} \breve{R}_k = I_{N_k} + \sum_{i=1}^{K} \left\{ \overline{H}_{k,b_i} Q_i \overline{H}_{k,b_i}^H + \mathrm{tr}\{Q_i C_{t,k,b_i}\}\, C_{r,k} \right\} \\ \breve{R}_{\overline{k}} = I_{N_k} + \sum_{i=1,\neq k}^{K} \left\{ \overline{H}_{k,b_i} Q_i \overline{H}_{k,b_i}^H + \mathrm{tr}\{Q_i C_{t,k,b_i}\}\, C_{r,k} \right\} \end{array} \tag{17}$$

$$\breve{B}_k = \overline{H}_{k,b_k}^H \breve{R}_{\overline{k}}^{-1} \overline{H}_{k,b_k} + \mathrm{tr}\{C_{r,k} \breve{R}_{\overline{k}}^{-1}\} C_{t,k,b_k} \tag{18}$$

$$\begin{array}{l} \breve{A}_k = \sum_{i\neq k}^{K} u_i \Big[\overline{H}_{i,b_k}^H \left(\breve{R}_{\overline{i}}^{-1} - \breve{R}_i^{-1} \right) \overline{H}_{i,b_k} \\ \qquad + \mathrm{tr}\{ \left(\breve{R}_{\overline{i}}^{-1} - \breve{R}_i^{-1} \right) C_{r,i} \} C_{t,i,b_k} \Big]\ . \end{array} \tag{19}$$

It suffices now to replace the matrices A_k, B_k in the DC programming approach of Section III by the matrices $\breve{A}_k$, $\breve{B}_k$ above to get a maximum EWSR design.

VI. The MISO Case

In this case $C_r = 1$ and we shall denote the matrices $\breve{R}$, H^H as the scalar $\breve{r}$ and the vector h. In the partial CSIT case, the term hh^H of the perfect CSIT case gets replaced by the posterior correlation matrix $S = \overline{h}\,\overline{h}^H + C_t$. For the optimization of g_k, we get

$$\begin{array}{l} WSR = u_k \ln(1 + g_k^H \breve{B}_k g_k) - g_k^H (\breve{A}_k + \lambda_{b_k} I) g_k \\ \text{with} \quad \breve{B}_k = \breve{r}_{\overline{k}}^{-1} S_{k,b_k}\ ,\ \breve{A}_k = \sum_{i\neq k}^{K} u_i d_i S_{i,b_k} \end{array} \tag{20}$$

and $d_i = \breve{r}_{\overline{i}}^{-1} - \breve{r}_i^{-1} = \frac{g_i^H S_{i,b_k} g_i}{\breve{r}_i \breve{r}_{\overline{i}}}$. The gradient of (20) w.r.t. g_k yields

$$a_k\, S_{k,b_k}\, g_k = (\breve{A}_k + \lambda_{b_k} I)\, g_k\ ,\ a_k = \frac{u_k / \breve{r}_{\overline{k}}}{1 + g_k^H \breve{B}_k g_k} \tag{21}$$

which leads to the generalized eigenvector solution $g_k^{'} = V_{max}(S_{k,b_k}, \breve{A}_k + \lambda_{b_k} I)$ and associated generalized eigenvalue $1/a_k = \lambda_{max}(S_{k,b_k}, \breve{A}_k + \lambda_{b_k} I)$. Note that in the MISO case with full CSIT, this $g_k^{'}$ would just be proportional to $(\breve{A}_k + \lambda_{b_k} I)^{-1} h_{k,b_k}$. However, in the MIMO case, or in the MISO case with partial CSIT, a genuine eigenvector computation is required (which in the WSMSE method gets done iteratively using the power method). The following (Jacobi style) iterative method avoids explicit (large dimension) eigenvector computation and is amenable to large system analysis. Assuming that $\overline{h}_{k,b_k}$ is non-zero (otherwise a random vector can be taken), (21) can be rewritten as

$$a_k\, \overline{h}_{k,b_k} \overline{h}_{k,b_k}^H\, g_k = (\breve{A}_k + \lambda_{b_k} I - a_k\, C_{t,k,b_k})\, g_k \tag{22}$$

which leads to the following explicit solution

$$g_k^{'} = (\breve{A}_k + \lambda_{b_k} I - a_k\, C_{t,k,b_k})^{-1}\, \overline{h}_{k,b_k}\ \ g_k = \sqrt{p_k} g_k^{'}\ . \tag{23}$$

We will perform a large system analysis on this solution in the next section. At this point, we should note that the expectations in Section IV are in fact conditional expectations $\mathrm{E}_{.|\overline{H}}\{.\}$. Whereas the beamformers are determined for given $\overline{H}$, the resulting EWSR ultimately depends also on the distribution of $\overline{H}$. We shall assume, as in [5], that the covariance matrix of $\overline{H}$ is also proportional to C_t so that we only have to consider one covariance matrix for both $\overline{H}$ and $\widetilde{H}$ (and even H). Hence (13) can be rewritten as follows

$$\overline{h}_{k,b_k}^H = (\sqrt{1-\tau^2} h_{k,b_k}^{'H} + \tau \widetilde{h}_{k,b_k}^H) C_{t,k,b_k}^{\frac{1}{2}} \tag{24}$$

where $h_{k,b_k}^{'}$ and $\widetilde{h}_{k,b_k}^H$ are independent with i.i.d. elements and $h_{k,b_k} = \sqrt{1-\tau^2} C_{t,k,b_k}^{\frac{1}{2}} h_{k,b_k}^{'}$ represents the true channel.

VII. The Determinstic equivalent of the SINR

The deterministic equivalents for scalar quantities that arise here in MaMISO are basically a consequence of the law of large numbers. We shall restrict the analysis to $M_i \equiv M$. The precoder for user k is expressed as the following:

$$g_k^{'} = (\breve{A}_k + \lambda_{b_k} I - a_k C_{t,k,b_k})^{-1} \overline{h}_{k,b_k} \tag{25}$$

where $\breve{A}_k$ is given in (20) is suitable for large system analysis. However, in order to ease the procedure the Lagrangian term λ_{b_k} can be fixed optimally to $\lambda_{b_k} = tr\frac{\overline{D}_{b_k}}{\rho_{b_k}}$; $\overline{D}_{b_k} = diag(vec(\mathcal{D}_{b_k}))$ with $\mathcal{D}_{b_k} = \{d_i; \forall i : b_i = b_k\}$ as in [10], $vec(\mathcal{D}_{b_k})$ is the vector composed of all (diagonal) elements of $\mathcal{D}_{b_k}$, $d_i = (\breve{r}_{\bar{i}}^{-1} – \breve{r}_i^{-1}) = -\breve{r}_{\bar{i}}^{-1}(\overline{h}_{i,b_i} Q_i \overline{h}_{i,b_i}^H)\breve{r}_i^{-1} = |f_i|^2\, w_i$ [Lemma 2; [5]]; where f_i and w_i are the Rx filter and the precoding weight respectively as in [6]; and $\rho_{b_k} = \frac{P_{b_k}}{1}$ is the signal-to-noise ratio.

In the following, a performance analysis is conducted for the proposed precoder. The large-system limit is considered, in which M and K go to infinity while keeping the ratio K/M finite such that $\limsup_M K/M < \infty$ and $\liminf_M K/M > 0$. The procedure should be understood in the way that, for each set of system dimension parameters M and K we provide an approximate expression for the SINR γ_k, which becomes more accurate as the system dimensions increase. The large system analysis follows the DC programming iterations. The precoders are initialized using Matched Filter (MF) precoders. Let γ_k^{MF} be the SINR of user k under MF precoding then, $\gamma_k^{MF} - \overline{\gamma}_k^{MF} \xrightarrow{M\to\infty} 0$ [7], almost surely, with

$$\overline{\gamma}_k^{MF} = \frac{1}{\frac{1}{\beta_{b_k}\rho_{b_k}} + \frac{1}{M^2}\sum_{i=1, i\neq k}^{K} tr\Theta_{k,b_i}\Theta_{i,b_i}} \tag{26}$$

where $\Theta_{k,b_i} = C_{t,k,b_i} \forall i, \forall k$ and ξ_{b_k} is the BF normalization parameter. For our precoder, a deterministic equivalent of the SINR is provided in the following theorem:

Theorem 1: Let γ_k be the SINR of the kth user with the precoder defined in (25). Then, a deterministic equivalent $\overline{\gamma}_k^{(j)}$ at iteration $j > 0$ and under MF initialization, is given by $\overline{\gamma}_k^{(j)}$

$$\overline{\gamma}_k^{(j)} = \frac{(1-\tau^2)\overline{m}_k^{(j)\,2}(1+\overline{m}_k^{(j)})^2}{\overline{\Upsilon}_k^{(j)} + \overline{\hat{\Upsilon}}_k^{(j)} + \overline{d}_k^{(j)}\frac{\overline{\Psi}_{b_k}^{(j)}}{\rho_{b_k}}(1+\overline{m}_k^{(j)})^2} \tag{27}$$

where

$$\overline{m}_k^{(j)} = \frac{1}{M} tr\{\overline{\Theta}_k^{(j)} V_{k,b_k}\}\,, \; \overline{\Psi}_{b_k}^{(j)} = \frac{1}{M}\sum_{i=1}^{K} e'_{i,b_k} \tag{28}$$

$$\overline{\Upsilon}_k^{(j)} = \frac{1}{M}\sum_{l=1, l\neq k}^{K} \overline{d}_k\, e'_{k,l,b_k} \tag{29}$$

$$\overline{\hat{\Upsilon}}_k^{(j)} = \overline{d}_k \frac{1}{M}(1+\overline{m}_k^{(j)})^2 \sum_{l=1}^{K}\frac{1}{(1+\overline{m}_l^{(j)})^2} e'_{k,l,b_l} \tag{30}$$

with $\overline{\Theta}_{k,b_i} = \overline{d}_k \Theta_{k,b_i}$, $a_k = (1/\overline{d}_k)$, $\overline{m}_{k,b_i}^{(j)} = \frac{1}{M} tr\{\overline{\Theta}_{k,b_i}^{(j)} V_{k,b_i}\}$ and $\overline{f}_k^{(j)}, \overline{w}_k^{(j)}$ and $\overline{d}_k^{(j)}$ are given by

$$\overline{f}_k^{(j)} = \frac{1}{\sqrt{\overline{P}_k^{(j-1)}}} \frac{\overline{\gamma}_k^{(j-1)}}{1+\overline{\gamma}_k^{(j-1)}} \tag{31}$$

$$\sqrt{\overline{P}_k^{(j-1)}} = \frac{1}{\overline{d}_k^{(j-1)}} \sqrt{\frac{P_{b_k}}{\overline{\Psi}_{b_k}^{(j-1)}}}\, \overline{m}_k^{(j-1)} \tag{32}$$

$$\overline{w}_k^{(j)} = (1+\overline{\gamma}_k^{(j-1)})\,;\; \overline{d}_k^{(j)} = \overline{w}_k^{(j)} \overline{f}_k^{2,(j)}. \tag{33}$$

$$\text{Let } V_{k,b_i} = (F_{b_i} + \overline{\lambda}_{b_i} I_M - a_k C_{t,k,b_i} + \sum_{j\neq k}^{K} d_j C_{t,j,b_i})^{-1} \tag{34}$$

with $\overline{\lambda}_{b_k}^{(j)} = \frac{tr\overline{D}_{b_k}^{(j)}}{M\rho_{b_k}}$, three systems of coupled equations have to be solved. First, we need to introduce $e_{k,b_i} \forall\{b_i, k\} \in \{\mathcal{C}, \mathcal{K}\}$ ($\mathcal{K}$ is the set of all users) which form the unique positive solutions of

$$e_{k,b_i} = \frac{1}{M} tr\{\overline{\Theta}_{k,b_i} V_{k,b_i}\}, \tag{35}$$

$$F_{b_i} = \frac{1}{M}\sum_{j=1}^{K}\frac{\overline{\Theta}_{j,b_i}}{1+e_{j,b_i}}. \tag{36}$$

e_{k,b_k} and m_k denote e_k and $m_{b_k,k}$ respectively. Secondly, we need $e'_1, ..., e'_{1,K}$ which form the unique positive solutions of

$$e'_k = \frac{1}{M} tr\overline{\Theta}_k V_{k,b_k}(F'_{b_k} + I_M)V_{k,b_k}, \tag{37}$$

$$F'_{b_k} = \frac{1}{M}\sum_{j=1}^{K}\frac{\overline{\Theta}_{j,b_k} e'_j}{(1+e_{j,b_k})^2}. \tag{38}$$

And finally, we provide $e'_{k,i,b_i} \forall\{b_i, k, i\} \in \{\mathcal{C}, \mathcal{K}, \mathcal{K}\}$

$$e'_{k,i,b_i} = \frac{1}{M} tr\{\overline{\Theta}_{i,b_i} V_{i,b_i}(F'_{i,b_i} + \overline{\Theta}_i)V_{i,b_i}\} \tag{39}$$

$$F'_{i,b_i} = \frac{1}{M}\sum_{j=1}^{K}\frac{\overline{\Theta}_{j,b_i} e'_{i,j,b_i}}{(1+e_{j,b_i})^2}. \tag{40}$$

For $j \geq 1$, define $\Gamma_{b_k}^{(j)} = \frac{1}{M} H_{b_k}^H \overline{D}^{(j)} H_{b_k} + \overline{\lambda}_{b_k}^{(j)} I_M$, with $H_{b_k} = [h_{1,b_k}, ..., h_{k-1,b_k}, h_{k+1,b_k}, ..., h_{K,b_k}]^H$ and $\overline{D} = diag(\overline{D}_1, ...\overline{D}_C)$. The precoder at the end of iteration j is given by

$$\overline{g}_k^{(j)} = \frac{\xi_{b_k}^{(j)}}{M}(\Gamma_{b_k}^{(j)})^{-1} h_{k,b_k} \tag{41}$$

for each user k , where $\xi_{b_k}^{(j)}$ is

$$\xi_{b_k}^{(j)} = \sqrt{\frac{P_{b_k}}{\frac{1}{M^2} tr(\Gamma_{b_k}^{(j)})^{-2} H_{\hat{b_k}}^H \overline{A}_{b_k}^{H,(j)} \overline{W}_{b_k}^{2,(j)} \overline{A}_{b_k}^{(j)} H_{\hat{b_k}}}} = \sqrt{\frac{P_{b_k}}{\Psi_{b_k}^{(j)}}}.$$

$H_{\hat{b_k}} = [H_{b_k,i} \;\; for \;\; all \;\; i \;\; s.t. \;\; b_i = b_k]^H$. We derive the deterministic equivalents of the normalization term $\xi_{b_k}^{(j)}$, the signal power $|\overline{g}_k^{H,(j)} h_{k,b_k}|^2$ and the interference power $\sum_{i=1}^{K} h_{k,b_i}^H \overline{g}_i^{(j)} \overline{g}_i^{H,(j)} h_{k,b_i}$ similarly to [5] and [6]. The proof is omitted due to lack of space.

VIII. NUMERICAL RESULTS

We plot the performance of the precoder in (25) with MF initialization and compare it to the large system approximation in Theorem 1. The channel correlation matrix is modeled as in [9]. Figure 1 shows the performance of the precoder and its approximation for i.i.d. channels for $C = 2$ cells. For the simulations of (25), we have used 1000 channel realizations. It can be observed that for i.i.d channels the approximation is very accurate which validates our asymptotic approach. Although the sum rate expression for the approximation approach (27) seems to be complex, we need to calculate it only once per a given SNR (independent of channel realization). In Figure 2 we consider the actual EWSR (averaged over

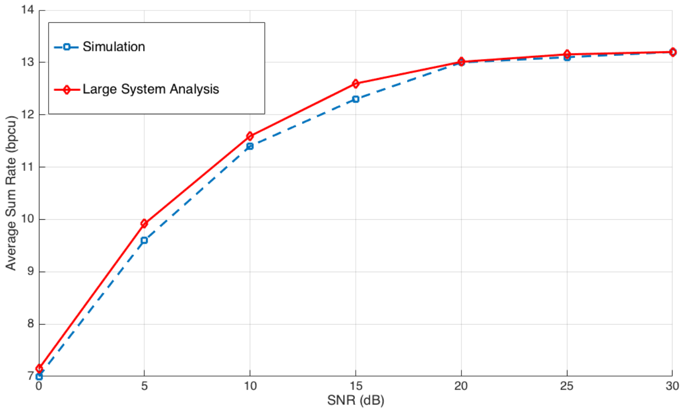

Fig. 1. Expected sum rate comparison for $C = 2, K = 3, M = 10, N = 1$, $C_{t,k,b_i} = I_M \, \forall i, \forall k$ and $\tau^2 = \frac{1}{10}$.

1000 channel realizations) for another scenario, in which the channels exhibit low channel covariance rank (few dominating multipath), in which the covariance information adds significant contributions to the channel estimate based CSIT. We perceive an unlimited gain for the proposed MaMISO asymptotic EWSR approach with respect to EWSMSE of [1] in the case of low rank correlation matrices, as SNR increases, though at medium SNR the order is slightly reversed for reasons unexplained at the time of writing.

IX. CONCLUSION

In this work, we derived and presented an asymptotically optimal beamforming algorithm for the case of partial CSIT and its large system performance analysis. Important EWSR gains have been illustrated over the naive approach in which the true channels are replaced by their estimates in a perfect CSIT approach, and also over the more sophisticated approach which relates EWSR to EWSME. The gain over EWSMSE comes from exploiting the channel covariance information also in the signal power (and not only in the interference terms).

Acknowledgments

EURECOM's research is partially supported by its industrial members: ORANGE, BMW, SFR, ST Microelectronics, Symantec, SAP, Monaco Telecom, iABG, by the projects ADEL

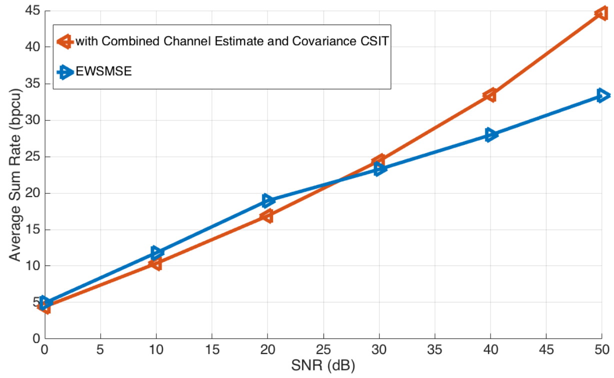

Fig. 2. Expected sum rate comparison for $C = 2, K = 4, M = 8, N = 1$, $rank(C_{t,k,b_i}) = 2, \; \forall i, \forall k$ and $\tau^2 = \frac{1}{10}$.

(EU FP7) and DIONISOS (French ANR). The research of Orange Labs is partially supported by the EU H2020 project Fantastic5G.

REFERENCES

[1] 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.

[2] X. Li, E. Björnson, E.G. Larsson, S. Zhou, and J. Wang, "Multi-cell MMSE Precoder for Massive MIMO Systems and New Large System Analysis," in IEEE Global Comm's Conf. (Globecom), San Diego, CA, USA, 2015.

[3] M. Razaviyayn, M.S. Boroujeni, and Z.-Q. Luo, "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.

[4] H. Asgharimoghaddam, A. Tolli, and N. Rajatheva, "Decentralizing the Optimal Multi-cell Beamforming via Large System Analysis," in IEEE Int'l Conf. on Communications (ICC), 2014.

[5] S. Wagner, R. Couillet, M. Debbah, and D. Slock, "Large System Analysis of Linear Precoding in Correlated MISO Broadcast Channels Under Limited Feedback," IEEE Trans. Info. Theory, Jul. 2012.

[6] S. Wagner and D.T.M. Slock, "Weighted Sum Rate Maximization of Correlated MISO Broadcast Channels under Linear Precoding : a Large System Analysis," in IEEE Int'l Workshop on Sig. Proc. Advances in Wireless Comm's (SPAWC), San Fransisco, CA, USA, 2011.

[7] S. Lakshminarayana, M. Assaad, and M. Debbah, "Coordinated Multi-cell Beamforming for Massive MIMO: A Random Matrix Approach," IEEE Trans. Information Theory, Jun. 2015.

[8] A. Muller, R. Couillet, E. Bjornson, S. Wagner, and M. Debbah, "Interference-Aware RZF Precoding for Multi Cell Downlink Systems," IEEE Trans. Signal Proc., Aug. 2015.

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

[10] S.S. Christensen, R. Agarwal, E. de Carvalho, and J. Cioffi, "Weighted sum-rate maximization using weighted MMSE for MIMO-BC beamforming design," IEEE Trans. on Wireless Comm's, Dec. 2008.