MIMO IBC Beamforming with Combined Channel Estimate and Covariance CSIT

Wassim Tabikh[¶‡] Dirk Slock[¶] Yi Yuan-Wu[‡]
[¶]EURECOM, 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 MIMO Interfering Broadcast Channel (IBC), i.e. the Multi-Input Multi-Output (MIMO) 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 (Posterior) partial CSIT can optimally combine channel estimate and channel covariance information. We introduce the first large system analysis for optimized beamformers with partial CSIT, here for the Massive MISO (MaMISO) case. 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:

- 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. Furthermore, three types of channel estimate are considered (deterministic, quantized, LMMSE). The analysis also applies to suboptimal EWSMSE and Naive EWSR designs.

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

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

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 $\widehat{x}_k = f_k^H y_k$

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

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)

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

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:

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

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,

$$f_k = R_k^{-1} H_{k,b_k} g_k\;. \tag{5}$$

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

$$\sum_{k:b_k=j} \text{tr}\{Q_k\} \leq P_j\;. \tag{6}$$

In a classical difference of convex functions (DC programming) approach, Kim and Giannakis [9] 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

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

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

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

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 $WSR(g, \hat{g}, \lambda)$

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

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

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

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

and hence we get the following MaMIMO limit matrices

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

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

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

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: 3 Channel Models

In this case $C_r = 1$ and we shall denote the matrices R, H^H as the scalar r and the vector h. The channel h, its estimate $\widehat{h}$ and estimation error $\widetilde{h}$ have covariance matrices $\Theta = C_{hh}$, $\widehat{\Theta} = C_{\widehat{h}\widehat{h}}$ and $C_{\widetilde{h}\widetilde{h}}$.

(i) In the case of a *deterministic channel estimate*, we have $\widehat{h} = h + \widetilde{h}$ where h and $\widetilde{h}$ are independent.

(ii) The *channel quantization error model* considered in [5] is a variation on the deterministic model in which we have a quantized channel of the form $\widehat{h}' = \eta h + \widetilde{h}$ with again $\widetilde{h}$ independent from h. Let $\widehat{h} = \eta^{-1} \widehat{h}'$.

(iii) In the case of a *LMMSE channel estimate*, we have $h = \widehat{h} + \widetilde{h}$ in which $\widehat{h}$ and $\widetilde{h}$ are decorrelated and hence independent in the Gaussian case. In the partial CSIT case, the term hh^H of the perfect CSIT case gets replaced by its estimate
$S = \mathrm{E}_{h|\widehat{h}} hh^H = \widehat{h}\widehat{h}^H + \widetilde{\Theta}$ where in the deterministic case, $\widetilde{\Theta} = -C_{\widetilde{h}\widetilde{h}} = -\sigma_{\widetilde{h}}^2 I$, in the quantized case $\widetilde{\Theta} = -\eta^{-2} C_{\widetilde{h}\widetilde{h}} = -\frac{\tau^2}{1-\tau^2} C_{hh}$ and actually $\widehat{\Theta} = \frac{1+\tau^2}{1-\tau^2} C_{hh}$, and in the LMMSE case $\widetilde{\Theta} = C_{\widetilde{h}\widetilde{h}}$ is the posterior covariance
$\widetilde{\Theta} = \Theta - \Theta(\Theta + \sigma_{\widetilde{h}}^2 I)^{-1} \Theta$.

VII. Max EWSR BF in the MaMISO Limit

The EWSR represents two rounds of averaging over the partial CSIT

$$\begin{array}{l} EWSR = \mathrm{E}_{\widehat{h}} \max_g EWSR(g) \\ EWSR(g) = \mathrm{E}_{h|\widehat{h}} WSR(g) = \sum_{k=1}^{K} u_k\; \mathrm{E}_{h|\widehat{h}} \ln(r_k / r_{\overline{k}}) \\ \stackrel{(a)}{=} \sum_{k=1}^{K} u_k\; [\ln(\overline{r}_k) - \ln(\overline{r}_{\overline{k}})] \end{array} \tag{20}$$

where transition (a) represents the MaMISO limit and

$$\begin{array}{l} \overline{r}_{\overline{k}} = 1 + \sum_{i\neq k} \mathrm{E}_{h|\widehat{h}} |h_{k,b_i}^H g_i|^2 = 1 + \sum_{i\neq k} g_i^H S_{k,b_i} g_i \\ \overline{r}_k = \overline{r}_{\overline{k}} + g_k^H S_{k,b_k} g_k \;. \end{array} \tag{21}$$

By adding the Lagrange terms for the BS power constraints, $\sum_{c=1}^{C} \lambda_c (P_c - \sum_{k:b_k=c} ||g_k||^2)$, to the EWSR in (20), we get the gradient

$$\begin{array}{l} \dfrac{\partial EWSR}{\partial g_k^*} = \alpha_k S_{k,b_k} g_k - \left[\sum_{i\neq k} \beta_i S_{i,b_k} + \lambda_{b_k} I \right] g_k = 0 \\ \alpha_k = \frac{u_k}{\overline{r}_k} \;,\; \beta_k = u_k (\frac{1}{\overline{r}_{\overline{k}}} - \frac{1}{\overline{r}_k}) \;. \end{array} \tag{22}$$

This leads to the iterative (power method like) solution

$$\begin{array}{l} g_k^{'}(\lambda_{b_k}) = \left[\sum_{i\neq k} \beta_i \widehat{h}_{i,b_k} \widehat{h}_{i,b_k}^H + \widetilde{\Theta}_k + \lambda_{b_k} I \right]^{-1} \widehat{h}_{k,b_k} \alpha_k \widehat{h}_{k,b_k}^H g_k \\ g_k = \xi_c\, g_k^{'} \;,\; \xi_c = \sqrt{P_c / \sum_{k:b_k=c} ||g_k^{'}(\lambda_{b_k})||^2} \\ \widetilde{\Theta}_k = \sum_{i\neq k} \beta_i \widetilde{\Theta}_{i,b_k} - \alpha_k \widetilde{\Theta}_{k,b_k} \;. \end{array} \tag{23}$$

The BF scale factors ξ_c are introduced because instead of a bisection method to force satisfaction of the power constraints, the Lagrange multipliers (if non-zero) can be adapted analytically as in [1] by exploiting $\sum_{k:b_k=c} g_k^H \dfrac{\partial EWSR}{\partial g_k^*} = 0$ and

$\sum_{k:b_k=c} ||g_k||^2 = P_c$. Then from (22) we get

$$\lambda_c = \begin{cases} \lambda_c^{'} & , \zeta_c > P_c \\ 0 & , \zeta_c \leq P_c \end{cases} \quad (24)$$

where

$$\begin{array}{l} \lambda_c^{'} = \frac{1}{P_c} \sum_{k:b_k=c} [\alpha_k g_k^H S_{k,c} g_k - \sum_{i \neq k} \beta_i g_k^H S_{i,c} g_k] \\ \zeta_c = \sum_{k:b_k=c} ||g_k^{'}(0)||^2 \end{array} \quad (25)$$

Indeed, in the case of multiple power constraints, not all constraints are necessarily satisfied with equality.

A. Min EWSMSE and Naive Max EWSR

The minimum EWSMSE BF design of [1] can be obtained from the EWSR design above by setting $\widetilde{\Theta}_{k,b_k} = 0$ in (21)-(25) and hence $S_{k,b_k} = \widehat{h}_{k,b_k} \widehat{h}_{k,b_k}^H$. On the other hand, the naive EWSR approach, which ignores the covariance of $\widetilde{h}$ in any occurrence, is obtained from the EWSR design above by setting $\widetilde{\Theta}_k = 0$ in (21)-(25) and setting $S_{i,c} = \widehat{h}_{i,c} \widehat{h}_{i,c}^H$ for any (i, c). Of course, these simplifications should be carried out in the BF design from (21)-(25), but not in the EWSR evaluation in (20)-(21).

VIII. LARGE SYSTEM APPROXIMATION OF THE EWSR

Due to the law of large numbers, the scalars r_k, $r_{\overline{k}}$, $a_k = \widehat{h}_{k,b_k}^H g_k^{'}$, $||g_k^{'}(\lambda_{b_k})||^2$, $g_k^H S_{k,c} g_k$, $g_k^H S_{i,c} g_k$ and hence α_k, β_k, ξ_c, $\lambda_c^{'}$ and ζ_c converge to deterministic limits as $M, K \to \infty$ at fixed ratio $\beta = M/K$. We shall perform a large system analysis to determine these deterministic limits, which will also provide the limiting value for

$$EWSR = \sum_{k=1}^{K} u_k \ \ln(\overline{r}_k / \overline{r}_{\overline{k}}) = \sum_{k=1}^{K} u_k \ \ln(1 + \overline{\gamma}_k) \quad (26)$$

where the $\overline{\gamma}_k$ are the limiting SINRs. The deterministic limits will follow the same iterations as the BF design algorithm for which we can rewrite iteration j as

$$\begin{array}{l}
\mu_{k,i}^{(j)} = \breve{g}_k^{(j-1)H} S_{i,b_k} \breve{g}_k^{(j-1)} \ , \ \forall i, k \\
\overline{r}_{\overline{k}}^{(j)} = 1 + \sum_{i \neq k} \xi_{b_i}^{2,(j-1)} \mu_{i,k}^{(j)} \ , \ \overline{r}_k^{(j)} = \overline{r}_{\overline{k}}^{(j)} + \xi_{b_k}^{2,(j-1)} \mu_{k,k}^{(j)} \\
\alpha_k^{(j)} = \frac{u_k}{\overline{r}_k^{(j)}} \ , \ \beta_k^{(j)} = u_k (\frac{1}{\overline{r}_{\overline{k}}^{(j)}} - \frac{1}{\overline{r}_k^{(j)}}) \\
a_k^{(j-1)} = \widehat{h}_{k,b_k}^H \breve{g}_k^{(j-1)} \\
\widetilde{\Theta}_k^{(j)} = \sum_{i \neq k} \beta_i^{(j)} \widetilde{\Theta}_{i,b_k} - \alpha_k^{(j)} \widetilde{\Theta}_{k,b_k} \\
\psi_k^{(j)}(0) = \widehat{h}_{k,b_k}^H \left[\sum_{i \neq k} \beta_i^{(j)} \widehat{h}_{i,b_k} \widehat{h}_{i,b_k}^H + \widetilde{\Theta}_k^{(j)} \right]^{-2} \widehat{h}_{k,b_k} \\
\nu_k^{(j)} = \alpha_k^{(j)} \xi_{b_k}^{(j-1)} a_k^{(j-1)} \\
\zeta_c^{(j)} = \xi_c^{2,(j-1)} \sum_{k:b_k=c} \psi_k^{(j)}(0) \, \nu_k^{2,(j)} \\
\breve{\lambda}_c^{(j)} = \frac{1}{P_c} \sum_{k:b_k=c} [\alpha_k^{(j)} \xi_{b_k}^{2,(j-1)} \mu_{k,k}^{(j)} - \sum_{i \neq k} \beta_i^{(j)} \xi_{b_k}^{2,(j-1)} \mu_{k,i}^{(j)}] \\
\lambda_c^{(j)} = \begin{cases} \breve{\lambda}_c^{(j)} & , \zeta_c^{(j)} > P_c \\ 0 & , \zeta_c^{(j)} \leq P_c \end{cases}
\end{array} \quad (27)$$

$$\begin{array}{l}
\breve{g}_k^{(j)} = \left[\sum_{i \neq k} \beta_i^{(j)} \widehat{h}_{i,b_k} \widehat{h}_{i,b_k}^H + \widetilde{\Theta}_k^{(j)} + \lambda_{b_k}^{(j)} I \right]^{-1} \widehat{h}_{k,b_k} \nu_k^{(j)} \\
\psi_k^{(j)}(\lambda_{b_k}^{(j)}) = \widehat{h}_{k,b_k}^H \left[\sum_{i \neq k} \beta_i^{(j)} \widehat{h}_{i,b_k} \widehat{h}_{i,b_k}^H + \widetilde{\Theta}_k^{(j)} + \lambda_{b_k}^{(j)} I \right]^{-2} \widehat{h}_{k,b_k} \\
\xi_c^{(j)} = \sqrt{P_c / \sum_{k:b_k=c} \psi_k^{(j)}(\lambda_{b_k}^{(j)}) \, \nu_k^{2,(j)}} \\
g_k^{(j)} = \xi_c^{(j)} \, \breve{g}_k^{(j)}
\end{array} \quad (28)$$

where we used the short-hand notation for e.g. $\nu_k^{2,(j)} = (\nu_k^{(j)})^2$. Note that $a_k^{(j)}$ can be computed recursively by introducing

$$\begin{array}{l}
\phi_k^{(j)} = \widehat{h}_{k,b_k}^H \left[\sum_{i \neq k} \beta_i^{(j)} \widehat{h}_{i,b_k} \widehat{h}_{i,b_k}^H + \widetilde{\Theta}_k^{(j)} + \lambda_{b_k}^{(j)} I \right]^{-1} \widehat{h}_{k,b_k} \ , \\
\Rightarrow \ a_k^{(j)} = \phi_k^{(j)} \nu_k^{(j)} = \phi_k^{(j)} \alpha_k^{(j)} \xi_{b_k}^{(j-1)} a_k^{(j-1)} \ .
\end{array} \quad (29)$$

A. Large System Analysis

In the large system analysis, we do not compute the BFs g_k. Instead, deterministic limits are determined for the following quantities: $\phi_k^{(j)}$, $\psi_k^{(j)}(0)$, $\psi_k^{(j)}(\lambda_{b_k}^{(j)})$, $\mu_{k,i}^{(j)}$.

The following results can now be obtained by applying the principles developed in [5]. For (with asymptotic equalities)

$$\begin{array}{l}
\phi_k = \widehat{h}_{k,b_k}^H [\sum_{i \neq k} \beta_i \widehat{h}_{i,b_k} \widehat{h}_{i,b_k}^H + \widetilde{\Theta}_k + \lambda_{b_k} I]^{-1} \widehat{h}_{k,b_k} \\
= \text{tr}\{\widehat{\Theta}_{k,b_k} [\sum_{i \neq k} \beta_i \widehat{h}_{i,b_k} \widehat{h}_{i,b_k}^H + \widetilde{\Theta}_k + \lambda_{b_k} I]^{-1}\} \ ,
\end{array} \quad (30)$$

the deterministic limit $\phi_k = e_{k,b_k}(\lambda_{b_k})$ can be computed from the C sets of implicit equations

$$e_{k,c}(\lambda_c) = \text{tr}\{\widehat{\Theta}_{k,c} (\underbrace{\sum_{i \neq k} \frac{\beta_i}{1 + \beta_i e_{i,c}(\lambda_c)} \widehat{\Theta}_{i,c} + \widetilde{\widetilde{\Theta}}_{k,c} + \lambda_c I}_{T_{k,c}(\lambda_c)})^{-1}\} \quad (31)$$

where $\widetilde{\widetilde{\Theta}}_{k,c} = \sum_{i \neq k} \beta_i \widetilde{\Theta}_{i,c} - \alpha_k \widetilde{\Theta}_{k,c}$, hence $\widetilde{\Theta}_k = \widetilde{\widetilde{\Theta}}_{k,b_k}$. For $\psi_k(\lambda) = \psi_{k,b_k}(\lambda)$, we get

$$\begin{array}{l}
\psi_{k,b_k}(\lambda) = \widehat{h}_{k,b_k}^H [\sum_{i \neq k} \beta_i \widehat{h}_{i,b_k} \widehat{h}_{i,b_k}^H + \widetilde{\Theta}_k + \lambda I]^{-2} \widehat{h}_{k,b_k} \\
= \text{tr}\{\widehat{\Theta}_{k,b_k} [\sum_{i \neq k} \beta_i \widehat{h}_{i,b_k} \widehat{h}_{i,b_k}^H + \widetilde{\Theta}_k + \lambda I]^{-2}\} \\
= -\frac{d}{d\lambda} \text{tr}\{\widehat{\Theta}_{k,b_k} T_{k,b_k}^{-1}(\lambda)\} = -e'_{k,b_k}(\lambda)
\end{array} \quad (32)$$

The $\psi_{k,b_k}(\lambda_{b_k})$ can be solved from the linear equations

$$\begin{array}{l}
\psi_{k,c}(\lambda_c) = \text{tr}\{\widehat{\Theta}_{k,c} T_{k,c}^{-1}(\lambda_c) T'_{k,c}(\lambda_c) T_{k,c}^{-1}(\lambda_c)\} \\
\text{with } T'_{k,c}(\lambda_c) = \sum_{i \neq k} \frac{\beta_i \psi_{i,c}(\lambda_c)}{(1 + \beta_i e_{i,c}(\lambda_c))^2} \widehat{\Theta}_{i,c} + I \ .
\end{array} \quad (33)$$

The C sets of equations (30) for the C values λ_c have to be augmented with a set $C+1$ for the $e_k(0)$. And also (33) has to be considered for $\psi_k(0)$.

For $\mu_{k,i}$ we introduce $\mu_{k,i} = \nu_k^2 \widehat{\mu}_{k,i} + \nu_k^2 \widetilde{\mu}_{k,i}$ with $\widehat{\mu}_{k,i} = \breve{g}_k^H \widehat{h}_{i,b_k} \widehat{h}_{i,b_k}^H \breve{g}_k / \nu_k^2$ and $\widetilde{\mu}_{k,i} = \breve{g}_k^H \widetilde{\Theta}_{i,b_k} \breve{g}_k / \nu_k^2$. We shall obtain $\widetilde{\mu}_{k,i}$ as $\widetilde{\mu}_{k,i} = e'_{k,b_k,i,b_k} = \frac{\partial}{\partial z} e_{k,b_k,i,b_k}(\lambda_{b_k}, 0)$ where

$$\begin{array}{l}
e_{k,c,i,d}(\lambda_c, z) = \text{tr}\{\widehat{\Theta}_{k,c} T_{k,c,i,d}^{-1}(\lambda_c, z)\} \text{ with} \\
T_{k,c,i,d}(\lambda_c, z) = \sum_{j \neq k} \frac{\beta_j \widehat{\Theta}_{j,c}}{1 + \beta_j e_{j,c,i,d}(\lambda_c, z)} + \widetilde{\widetilde{\Theta}}_{k,c} + \lambda_c I - z \, \widetilde{\Theta}_{i,d}
\end{array} \quad (34)$$

We can then obtain the $e'_{k,c,i,d}$ as the solution of the linear equations

$$\begin{aligned} e'_{k,c,i,d} &= \mathrm{tr}\{\widehat{\Theta}_{k,c} T^{-1}_{k,c,i,d}(\lambda_c,0)\, T'_{k,c,i,d}(\lambda_c,0)\, T^{-1}_{k,c,i,d}(\lambda_c,0)\} \\ T'_{k,c,i,d}(\lambda_c,0) &= \sum_{j\neq k} \frac{\beta_j e'_{j,c,i,d}(\lambda_c)}{(1+\beta_j e_{j,c}(\lambda_c))^2}\widehat{\Theta}_{j,c} + \widetilde{\Theta}_{i,d} \end{aligned} \quad (35)$$

On the other hand, we get $\widehat{\mu}_{k,k} = a_k^2 = \nu_k^2\, e^2_{k,b_k}(\lambda_{b_k})$. For $i \neq k$,

$$\begin{aligned} \widehat{\mu}_{k,i} &= |\widehat{h}^H_{i,b_k}[\textstyle\sum_{j\neq k}\beta_j \widehat{h}_{j,b_k}\widehat{h}^H_{j,b_k} + \widetilde{\Theta}_k + \lambda_{b_k} I]^{-1}\widehat{h}_{k,b_k}|^2 \\ &= \frac{|\widehat{h}^H_{i,b_k}[\sum_{j\neq k,i}\beta_j \widehat{h}_{j,b_k}\widehat{h}^H_{j,b_k}+\widetilde{\Theta}_k+\lambda_{b_k} I]^{-1}\widehat{h}_{k,b_k}|^2}{(1+\beta_i\widehat{h}^H_{i,b_k}[\sum_{j\neq k,i}\beta_j\widehat{h}_{j,b_k}\widehat{h}^H_{j,b_k}+\widetilde{\Theta}_k+\lambda_{b_k} I]^{-1}\widehat{h}_{i,b_k})^2} \end{aligned} \quad (36)$$

Hence $\widehat{\mu}_{k,i} = \widehat{\widehat{\mu}}_{k,i}/(1+\beta_i e_{i,b_k}(\lambda_{b_k}))^2$ where the $\widehat{\widehat{\mu}}_{k,i}$ are obtained from a system of equations similar to that for $\widetilde{\mu}_{k,i}$ (as in (34), (35)), by replacing the $\widetilde{\Theta}_{i,b_k}$ by $\widehat{\Theta}_{i,b_k}$.

IX. Numerical Results

We plot the performance of the proposed EWSR BF with MF initialization and compare it to the proposed large system approximation. The channel correlation matrix is modeled as in [5]. We consider the quantization error channel model. Fig. 1 shows the performance of the precoder and its large system approximation for i.i.d. channels for $C = 2$ cells. Monte Carlo simulations are averaged over 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 large system analysis for the sum rate seems complex, we need to calculate it only once per given SNR (independent of channel realizations). In Fig. 2 we consider

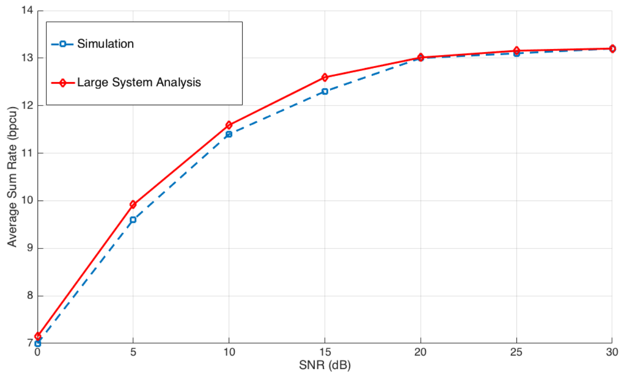

Fig. 1. Expected sum rate comparison for $C = 2, K = 3, M = 10, N = 1$, $\Theta_{k,b_i} = I_M\ \forall i, \forall k$ and $\tau^2 = \frac{1}{10}$.

the actual EWSR (averaged over 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. The EWSR curve does not saturate at high SNR because the $rank(\Theta_{k,b_i}) = 2$ information allows covariance CSIT based ZF for 4 users. We perceive an unlimited gain for the proposed approach with respect to EWSMSE of [1] in this case of low rank correlation matrices, and of course even more w.r.t. the naive approach of perfect CSIT in which the true channels simply get replaced by their estimates.

X. Conclusion

In this work, we derived and presented an asymptotically optimal beamforming algorithm for the case of partial

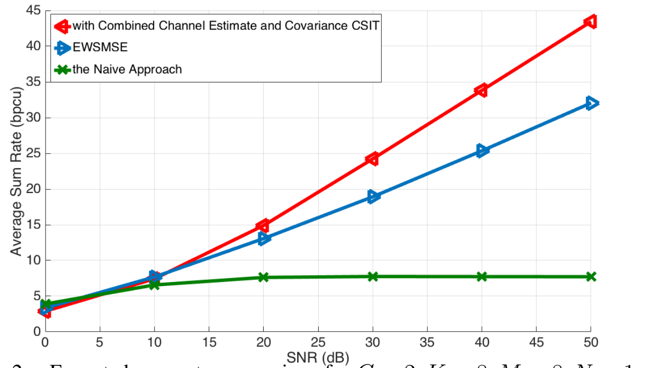

Fig. 2. Expected sum rate comparison for $C = 2, K = 8, M = 8, N = 1$, $rank(\Theta_{k,b_i}) = 2,\ \forall i, \forall k$ and $\tau^2 = \frac{1}{2}$.

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 HIGHTS (EU H2020), DUPLEX (French ANR) and MASS-START (French FUI). The research of Orange Labs is partially supported by the EU H2020 5GPPP project One5G.

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 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 Tr. 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.