# Massive MISO IBC Beamforming - a Multi-Antenna Stochastic Geometry Perspective

Christo Kurisummoottil Thomas, Dirk Slock,
EURECOM, Sophia-Antipolis, France, Email: {kurisumm,slock}@eurecom.fr

***Abstract*—This work deals with coordinated beamforming (BF) for the Multi-Input Single-Output (MISO) Interfering Broadcast Channel (IBC), i.e. the MISO 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. With Gaussian partial CSIT, the beamformers only depend on the means (estimates) and (error) covariances of the channels. We extend a recently introduced large system analysis for optimized beamformers with partial CSIT, by a stochastic geometry inspired randomization of the channel covariance eigen spaces, leading to much simpler analytical results, which depend only on some essential channel characteristics. In the Massive MISO (MaMISO) limit, we obtain deterministic approximations of the signal and interference plus noise powers at the receivers, which are tight as the number of antennas and number of users $M, K \to \infty$ at fixed ratio. Simulation results exhibit the correctness of the large system results and the performance superiority of optimal BF designs based on both the MaMISO limit of the EWSR and using Linear Minimum Mean Squared Error (LMMSE) channel estimates.**

***Index terms*—** Massive MIMO, stochastic geometry, partial CSIT, ergodic weighted sum rate, optimal beamforming

## I. INTRODUCTION

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. Base stations (BSs) disposing of multiple antennas are able to serve multiple Mobile Terminals (MTs) simultaneously, which is called Spatial Division Multiple Access (SDMA) or Multi-User (MU) Multi-Input Multiple-Output (MIMO). The recent development of Massive MIMO (MaMIMO) [1] opens new possibilities for increased system capacity while at the same time simplifying system design. However, MU systems have precise requirements for Channel State Information at the Tx (CSIT) which is more difficult to acquire than CSI at the Rx (CSIR). [2] proposes optimal beamformers (BFs) in the case of partial CSIT in the massive MISO limit.

Gaussian (posterior) partial CSIT can optimally combine channel estimate and channel covariance information. Recently 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. Indeed, in Massive MISO systems, the received signal and interference powers converge to their expected value due to the law of large numbers. We may remark that in the case of MIMO (not considered here), in which UEs possess a limited number of Rx 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 major work on large system analysis for MaMISO systems is [3]. The authors obtain deterministic (instead of channel realization 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 [3] allowed e.g. the determination of the optimal regularization factor in Regularized ZF (R-ZF) BF, both with perfect and partial CSIT. In [4], the authors investigate the deterministic limits for optimal beamformers, but only for the perfect CSIT MISO BC (broadcast channel) case. Some other extensions appeared recently in [5] or [6] where MISO IBC is considered with perfect CSIT and weighted Regularized Zero-Forcing (R-ZF) BF, with two optimized weight levels, for intracell or intercell interference. [7] studied the energy consumption dynamics in a MISO BC with users moving around according to a random walk model. Recently large system analysis was applied to investigate the BF designs for the power minimization problem with quality-of-service (QoS) targets at the users [8], [9]. [10] applies large system analysis to reduced order ZF beamforming for the simplified case of differently attenuated channel covariance matrices of the users. [11], [12] proposes the Expected Weighted Sum MSE (EWSMSE) based approach for BF design under partial CSIT. However EWSMSE is suboptimal and cannot even be used in the case of zero channel mean (covariance CSIT only information). [2] proposes a large system analysis for optimized BF with partial CSIT as considered here. Furthermore, the channel, channel estimate and channel error covariances can all be arbitrary and different for all users. However, the resulting deterministic analysis is quite cumbersome and does not allow much analytical insight. In stochastic geometry based methods [13], the location of the users is assumed to be random, their geographic distribution then inducing a certain probability distribution for the channel attenuations. Whereas most stochastic geometry work focuses on the distribution of the attenuations, here we consider an extension to multi-antenna systems. The multipath propagation for the various users leads to randomized angles of arrival at the BS which can be translated into spatial channel response contributions that depend on the antenna array response. In the massive MIMO regime in which the number of BS antennas gets very large, it has been observed and exploited that despite complex multipath propagation, the channel covariance matrix tends to be low rank. Exploiting the randomized nature of the user and scatterer positions and making abstraction of the antenna array response, we proposed to model the user channel subspaces as isotropically randomly oriented. This allows us to assume the eigen vectors of the channel covariance matrix to be Haar distributed, and this identically and independently for all users.

### A. *Contributions of this paper*

In this paper:

- We first review optimal BFs for the expected weighted sum rate (EWSR) criterion in the MaMISO limit.
- We evaluate the ergodic sum rate performance for Least-Squares (LS), LMMSE and subspace projection channel estimators. Numerical results suggest that there is substantial gain by exploiting the channel covariance information compared to just using the LS estimates.
- New large system analysis for various cases of BF with partial CSIT is proposed, with a randomized analysis of the covariance subspaces, leading to much simpler results. This constitutes a marriage between large system analysis and multi-antenna stochastic geometry.
- Simulation results indicate that the large system approximations are very accurate even for small system dimensions and reveal the deterministic dependence of the system performance on several important scalar parameters, such as the channel multipath attenuation profile, signal powers and SNR (whereas [2] doesn't lead to any such tractable analytical solutions).

Notation: In the following, boldface lower-case and upper-case characters denote vectors and matrices respectively. The operators $\mathrm{E}(\cdot)$, $tr(\cdot)$ , $(\cdot)^H$, $(\cdot)^T$ represents expectation, trace , conjugate transpose and transpose respectively. diag$(\cdot)$ represents the diagonal matrix formed by the elements $(\cdot)$. A circularly complex Gaussian random vector with mean $\boldsymbol{\mu}$ and covariance matrix $\boldsymbol{\Theta}$ is distributed as $\boldsymbol{x} \sim \mathcal{CN}(\boldsymbol{\mu}, \boldsymbol{\Theta})$. $\mathbf{V}_{max}(\mathbf{A}, \mathbf{B})$ or $\mathbf{V}_{max}(\mathbf{A})$ represents (normalized) dominant generalized eigen vector of $\mathbf{A}$ and $\mathbf{B}$ or (normalized) dominant eigen vector of $\mathbf{A}$ respectively and $\lambda_{max}(\mathbf{A})$ is the corresponding max eigen value.

## II. MISO IBC Signal Model

We consider here an IBC with $C$ cells with a total of $K$ single antenna users. We shall consider a system-wide numbering of the users. User $k$ is served by BS $b_k$. The received signal at user $k$ in cell $b_k$ is

$$\boldsymbol{y}_k = \underbrace{\mathbf{h}_{k,b_k}^H \mathbf{g}_k x_k}_{\text{signal}} + \underbrace{\sum_{\substack{i \neq k \\ b_i = b_k}} \mathbf{h}_{k,b_k}^H \mathbf{g}_i x_i}_{\text{intracell interf.}} + \underbrace{\sum_{j \neq b_k} \sum_{i: b_i = j} \mathbf{h}_{k,j}^H \mathbf{g}_i x_i}_{\text{intercell interf.}} + \boldsymbol{v}_k \tag{1}$$

where $x_k$ is the intended (white, unit variance) scalar signal stream, $\mathbf{h}_{k,b_k}$ is the $M_{b_k} \times 1$ channel from BS $b_k$ to user $k$. The Rx signal (and hence the channel) is assumed to be scaled so that we get for the noise $v_k \sim \mathcal{CN}(0,1)$. BS $b_k$ serves $K_{b_k} = \sum_{i:b_i=b_k} 1$ users. The $M_{b_k} \times 1$ spatial Tx filter or beamformer (BF) is $\mathbf{g}_k$.

## III. Channel and CSIT Model

For simplicity, we omit all the user indices $k$. We start from a deterministic Least-Squares (LS) channel estimate

$$\widehat{\mathbf{h}}_{LS} = \mathbf{h} + \widetilde{\mathbf{h}}, \tag{2}$$

where $\mathbf{h}$ is the true MISO channel, and the error is modeled as circularly symmetric white Gaussian noise $\widetilde{\mathbf{h}} \sim \mathcal{CN}(0, \widetilde{\sigma}^2\mathbf{I})$. Now each MISO channel is modeled according to a correlation structure as follows,

$$\mathbf{h} = \mathbf{C}\,\mathbf{c}, \;\; \mathbf{c} = \mathbf{D}^{1/2}\mathbf{c}', \tag{3}$$

where $\mathbf{c}' \sim \mathcal{CN}(0, \mathbf{I}_L)$ and $\mathbf{D}$ is diagonal. Here $\mathbf{C}$ is the $M \times L$ eigen vector matrix of the reduced rank channel covariance $R_{\mathbf{hh}} = \mathbf{CDC}^H$. The total sum rank across all users $N_p = \sum_{k=1}^{K} L_{k,c}$ is assumed to be less than $M_c$, where $L_{k,c}$ is the channel rank between user $k$ and BS $c$. Assuming the channel covariance subspace is known, the LMMSE channel estimate can be written as $\widehat{\mathbf{h}} = \mathbf{CDC}^H \left(\mathbf{CDC}^H + \widetilde{\sigma}^2\mathbf{I}\right)^{-1} \widehat{\mathbf{h}}_{LS}$. Applying the matrix inversion lemma and exploiting $\mathbf{C}^H\mathbf{C} = \mathbf{I}_L$, this simplifies to

$$\widehat{\mathbf{h}} = \mathbf{C}\left(\widetilde{\sigma}^2\mathbf{D}^{-1} + \mathbf{I}\right)^{-1} \mathbf{C}^H \widehat{\mathbf{h}}_{LS} = \mathbf{C}\widehat{\mathbf{D}}^{1/2}\widehat{\mathbf{c}}, \tag{4}$$

where $\widehat{\mathbf{D}} = \left(\widetilde{\sigma}^2\mathbf{D}^{-1} + \mathbf{I}\right)^{-1}\mathbf{D}$ and $\widehat{\mathbf{c}} = \mathbf{D}^{-1/2}(\sigma^2\mathbf{D}^{-1} + \mathbf{I})^{-1/2}\mathbf{C}^H\widehat{\mathbf{h}}_{LS}$. The posterior error covariance becomes

$$R_{\widetilde{\mathbf{h}}\widetilde{\mathbf{h}}} = \mathbf{CDC}^H - \mathbf{CDC}^H\left(\mathbf{CDC}^H + \widetilde{\sigma}^2 I\right)^{-1}\mathbf{CDC}^H, \tag{5}$$

which the matrix inversion lemma allows to simplify to,

$$R_{\widetilde{\mathbf{h}}\widetilde{\mathbf{h}}} = \mathbf{C}\left[\mathbf{D} - \left(\widetilde{\sigma}^2\mathbf{D}^{-1} + \mathbf{I}\right)^{-1}\mathbf{D}\right]\mathbf{C}^H = \mathbf{C}\widetilde{\mathbf{D}}\mathbf{C}^H. \tag{6}$$

So we can write for $\mathbf{S} = \mathrm{E}_{\mathbf{h}|\widehat{\mathbf{h}}}\left(\mathbf{hh}^H\right) = \widehat{\mathbf{h}}\,\widehat{\mathbf{h}}^H + R_{\widetilde{\mathbf{h}}\widetilde{\mathbf{h}}} = \mathbf{CWC}^H$, where $\mathbf{W} = \widehat{\mathbf{D}}^{1/2}\widehat{\mathbf{c}}\widehat{\mathbf{c}}^H\widehat{\mathbf{D}}^{1/2} + \widetilde{\mathbf{D}}$.

## IV. Partial CSIT BF based on Different Channel Estimates

In the MaMIMO limit, BF design with partial CSIT will depend on the quantities $\mathbf{S} = \mathrm{E}_{\mathbf{h}|\widehat{\mathbf{h}}}\left(\mathbf{hh}^H\right) = \widehat{\mathbf{h}}\,\widehat{\mathbf{h}}^H + \widetilde{\boldsymbol{\Theta}}$. We shall consider three possible channel estimates.

(i) *LS Channel Estimate*

We have $\widehat{\mathbf{h}}_{LS} = \mathbf{h} + \widetilde{\mathbf{h}}$ where $\mathbf{h}$ and $\widetilde{\mathbf{h}}$ are independent. In the LS case, $\widetilde{\boldsymbol{\Theta}} = \widetilde{\sigma}^2\mathbf{I}$. If we want however $\mathbf{S}$ to be an unbiased estimate for $\mathbf{hh}^H$, then we shall take $\widetilde{\boldsymbol{\Theta}} = -\widetilde{\sigma}^2\mathbf{I}$.

(ii) *LMMSE Channel Estimate*

We have $\mathbf{h} = \widehat{\mathbf{h}} + \widetilde{\mathbf{h}}$ in which $\widehat{\mathbf{h}}$ and $\widetilde{\mathbf{h}}$ are decorrelated and hence independent in the Gaussian case. In the LMMSE case, $\widetilde{\boldsymbol{\Theta}} = C_{\widetilde{\mathbf{h}}\widetilde{\mathbf{h}}}$ is the posterior covariance. The resulting $\mathbf{S} = \widehat{\mathbf{h}}\widehat{\mathbf{h}}^H + \widetilde{\boldsymbol{\Theta}}$ now forms an unbiased estimate of $\mathbf{hh}^H$ : $\mathrm{E}_{\widehat{\mathbf{h}}}\mathbf{S} = R_{\mathbf{hh}}$.

(iii) *Subspace Projection based Channel Estimate*

We also investigate the effect of limiting channel estimation error to the covariance subspace (without the LMMSE weighting, this is a simplification of the LMMSE estimate). The subspace channel estimate is given as,

$$\widehat{\mathbf{h}}_S = \boldsymbol{P}_C \widehat{\mathbf{h}}_{LS} = \mathbf{h} + \boldsymbol{P}_C \widetilde{\mathbf{h}}_{LS} \, , \; C_{\widetilde{\mathbf{h}}_S \widetilde{\mathbf{h}}_S} = \widetilde{\sigma}^2 \boldsymbol{P}_C, \quad (7)$$

where $\boldsymbol{P}_C = \mathbf{C}(\mathbf{C}^H\mathbf{C})^{-1}\mathbf{C}^H$ represents the projection onto the covariance subspace. Estimates for $\mathbf{h}\mathbf{h}^H$ :
(a) Naive Subspace Channel Estimator, $\mathbf{S} = \widehat{\mathbf{h}}_S\widehat{\mathbf{h}}_S^H = \mathbf{C}\widehat{\mathbf{c}}\widehat{\mathbf{c}}^H\mathbf{C}^H$,
(b) Subspace Channel Estimator, $\mathbf{S} = \widehat{\mathbf{h}}_S\widehat{\mathbf{h}}_S^H + \mathbf{C}_{\widetilde{h}_S\widetilde{h}_S} = \mathbf{C}(\widehat{\mathbf{c}}\widehat{\mathbf{c}}^H + \widetilde{\sigma}^2\mathbf{I})\mathbf{C}^H$,
(c) Unbiased $(\mathbf{S})$ Subspace Channel Estimator, $\mathbf{S} = \widehat{\mathbf{h}}_S\widehat{\mathbf{h}}_S^H - \mathbf{C}_{\widetilde{h}_S\widetilde{h}_S} = \mathbf{C}(\widehat{\mathbf{c}}\widehat{\mathbf{c}}^H - \widetilde{\sigma}^2\mathbf{I})\mathbf{C}^H$.

### *A. BF with Partial CSIT*

Three types of BF design with partial CSIT can be analyzed. In the case of partial CSIT we get for the Rx signal,

$$\begin{aligned} y_k = & \; \widehat{\mathbf{h}}_{k,b_k}^H \, \mathbf{g}_k \, x_k + \underbrace{\widetilde{\mathbf{h}}_{k,b_k}^H \, \mathbf{g}_k \, x_k}_{\text{sig. ch. error}} \\ & + \sum_{i=1,\neq k}^{K} (\widehat{\mathbf{h}}_{k,b_i}^H \, \mathbf{g}_i \, x_i + \underbrace{\widetilde{\mathbf{h}}_{k,b_i}^H \, \mathbf{g}_i \, x_i}_{\text{interf. ch. error}} ) + v_k \, . \end{aligned} \quad (8)$$

1) Naive BF EWSR: just replace $\mathbf{h}$ by $\widehat{\mathbf{h}}$ in a perfect CSIT approach. Ignore $\widetilde{\mathbf{h}}$ everywhere. 2) Optimal BF EWSR: accounts for covariance CSIT in the signal and interference terms. 3) EWSMSE BF (Expected Weighted Sum MSE) [11]: accounts for covariance CSIT in the interference terms, but also associates the signal $\widetilde{\mathbf{h}}$ term with the interference !

### *B. Max EWSR BF in the MaMISO limit (ESEI-WSR)*

The scenario of interest here is to design optimal beamformers when there is only partial CSIT. Once the CSIT is imperfect, various optimization criteria such as outage capacity can be considered. Here the design is based on expected weighted sum rate (EWSR) (and in a first instance with LMMSE channel estimates). The actual EWSR represents two rounds of averaging. In a first stage, the WSR is averaged over the channels given the channel estimates and covariance information (i.e. the partial CSIT), leading to a cost function that can be optimized by the Tx. The optimized result then needs to be averaged over the channel estimates to obtain the final ergodic WSR. In the MaMISO limit, due to the law of large numbers, a number of scalars converge to their expected value, facilitating averaging the WSR. From the law of total expectation and motivated from the ergodic capacity formulations [14] (point to point MIMO systems), [15] (multi user MISO systems),

$$\begin{aligned} & EWSR = \mathrm{E}_{\widehat{\mathbf{h}}} \max_{\mathbf{g}} EWSR(\mathbf{g}), \\ & EWSR(\mathbf{g}) = \mathrm{E}_{\mathbf{h}|\widehat{\mathbf{h}}} WSR(g) = \sum_{k=1}^{K} u_k \, \mathrm{E}_{\mathbf{h}|\widehat{\mathbf{h}}} \ln(s_k/s_{\overline{k}}) \\ & \stackrel{(a)}{=} \sum_{k=1}^{K} u_k \, \ln((\mathrm{E}_{\mathbf{h}|\widehat{\mathbf{h}}} s_k)/(\mathrm{E}_{\mathbf{h}|\widehat{\mathbf{h}}} s_{\overline{k}})) = \sum_{k=1}^{K} u_k \, \ln(r_{\overline{k}}^{-1} r_k), \end{aligned} \quad (9)$$

where transition $(a)$ represents the MaMISO limit leading to ESEI-WSR (Expected Signal Expected Interference WSR), $u_k$ are the rate weights, $\mathbf{g}$ represents the collection of BFs $\mathbf{g}_k$. $s_{\overline{k}}$ is the (channel dependent) interference plus noise power and $s_k$ is the signal plus interference plus noise power. Their conditional expectations are

$$\begin{aligned} & r_{\overline{k}} = 1 + \sum_{i \neq k} \mathrm{E}_{\mathbf{h}|\widehat{\mathbf{h}}} |\mathbf{h}_{k,b_i}^H \mathbf{g}_i|^2 = 1 + \sum_{i \neq k} \mathbf{g}_i^H \mathbf{S}_{k,b_i} \mathbf{g}_i, \\ & r_k = r_{\overline{k}} + \mathbf{g}_k^H \mathbf{S}_{k,b_k} \mathbf{g}_k, \; \mathbf{S}_{k,b_k} = \mathbf{C}_{k,b_k} \mathbf{W}_{k,b_k} \mathbf{C}_{k,b_k}. \end{aligned} \quad (10)$$

Further we split $\mathbf{g}_k = \mathbf{g}'_k p_k^{1/2}$, where $p_k$ is the power allocated to user $k$, and $||\mathbf{g}'_k|| = 1$. By adding the Lagrange terms for the BS power constraints, $\sum_{c=1}^{C} \mu_c (P_c - \sum_{k:b_k=c} ||\mathbf{g}_k||^2)$, to the EWSR in (9), we get the gradient (with $\alpha_k = \frac{u_k}{r_k}$, $\beta_k = u_k (\frac{1}{r_{\overline{k}}} - \frac{1}{r_k})$)

$$\frac{\partial EWSR}{\partial \mathbf{g}_k^*} = \alpha_k \mathbf{S}_{k,b_k} \, \mathbf{g}_k - (\sum_{i \neq k} \beta_i \mathbf{S}_{i,b_k} + \mu_{b_k} \mathbf{I}) \mathbf{g}_k = 0, \quad (11)$$

with $\mathbf{S}_{i,b_k} = \mathbf{C}_{i,b_k} \mathbf{W}_{i,b_k} \mathbf{C}_{i,b_k}^H$. This leads to the generalized eigen vector,

$$\mathbf{g}'_k = \mathbf{V}_{max}(\mathbf{S}_{k,b_k}, \sum_{i \neq k} \beta_i \mathbf{S}_{i,b_k} + \mu_{b_k} \mathbf{I}) \, . \quad (12)$$

While (11) can be interpreted many ways, (12) comes from the following DC programming. Introducing the Tx covariance matrices $\mathbf{Q}_i = \mathbf{g}_i \mathbf{g}_i^H$, the power constraints can be written as $\sum_{k:b_k=c} tr\{\mathbf{Q}_k\} \leq P_c$. The EWSR problem is non-concave in the $\mathbf{Q}_k$ due to the interference terms. Therefore finding the global optimum is challenging. In order to find at least a local optimum, we consider the difference of convex functions programming (DCP) approach as in [16]. Whereas [16] however solves the Lagrange multipliers by Lagrangean duality, here we solve them together with the powers (as in standard water filling) in an alternating optimization approach (alternating with optimizing the $\mathbf{g}'_k$). In DCP one keeps the concave signal term and linearizes the convex term, leading to a concave cost function in the $\mathbf{Q}_i$ (or a minorizer actually in the $\mathbf{g}'_i$ and $p_i$), which can be optimized iteratively.

$$\begin{aligned} EWSR \; & = \; u_k \ln\det(r_{\overline{k}}^{-1} r_k) \; + \; EWSR_{\overline{k}}, \\ EWSR_{\overline{k}} \; & = \; \sum_{i=1,\neq k}^{K} u_i \ln(r_{\overline{i}}^{-1} r_i), \end{aligned} \quad (13)$$

where $\ln(r_{\overline{k}}^{-1} r_k)$ 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 of $WSR_{\overline{k}}$ in $\mathbf{Q}_k$ around $\widehat{\mathbf{Q}}$ (i.e. all $\widehat{\mathbf{Q}}_i$).

$$\begin{aligned} & EWSR_{\overline{k}}(\mathbf{Q}_k, \widehat{\mathbf{Q}}) \approx EWSR_{\overline{k}}(\widehat{\mathbf{Q}}_k, \widehat{\mathbf{Q}}) - \\ & tr\left\{ (\mathbf{Q}_k - \widehat{\mathbf{Q}}_k) \widehat{\mathbf{A}}_k \right\}, \end{aligned} \quad (14)$$

$$\text{where,} \;\; \widehat{\mathbf{A}}_k = - \left. \frac{\partial EWSR_{\overline{k}}(\mathbf{Q}_k, \widehat{\mathbf{Q}})}{\partial \mathbf{Q}_k} \right|_{\widehat{\mathbf{Q}}_k, \widehat{\mathbf{Q}}} = \sum_{i=1,\neq k}^{K} \widehat{\beta}_i \mathbf{S}_{i,b_k} \, . \quad (15)$$

Note that the linearized tangent expression for $EWSR_{\overline{k}}$ constitutes a lower bound for it and hence the DC approach is also a minorization approach. Now dropping the constant terms and reparameterizing the $\mathbf{Q}_k$ in terms of the $\mathbf{g}_k$, we can write the original WSR as the Lagrangian,

$$EWSR(\mathbf{g}) = \sum_{k=1}^{K} \Big[ u_k \ln \Big( 1 + \mathbf{g}_k^H \widehat{\mathbf{B}}_k \mathbf{g}_k \Big) \\ -\text{tr} \Big\{ \mathbf{g}_k^H \left( \widehat{\mathbf{A}}_k + \mu_{b_k} \mathbf{I} \right) \mathbf{g}_k \Big\} \Big] + \sum_{j=1}^{C} \mu_j P_j, \ \widehat{\mathbf{B}}_k = \widehat{r}_{\overline{k}}^{-1} \mathbf{S}_{k,b_k}. \tag{16}$$

(16) leads again to (11) and esp. (12). The advantage of formulation (16) is that it allows straightforward power adaptation: substituting $\mathbf{g}_k = \sqrt{p_k}\,\mathbf{g}'_k$ in (16) and optimizing leads to the following interference leakage ($\sigma_k^{(2)}$) aware water filling

$$p_k = \left( \frac{u_k}{\sigma_k^{(2)} + \mu_{b_k}} - \frac{1}{\sigma_k^{(1)}} \right)^+, \tag{17}$$

where $(x)^+ = \max\{0, x\}$ and the Lagrange multipliers $\mu_c$ are adjusted (e.g. by bisection) to satisfy the power constraints. Also, $\sigma_k^{(1)} = \mathbf{g}_k'^H \widehat{\mathbf{B}}_k \mathbf{g}'_k$, $\sigma_k^{(2)} = \mathbf{g}_k'^H \widehat{\mathbf{A}}_k \mathbf{g}'_k$. With $\sigma_k^{(2)} = 0$ this would be standard waterfilling.

## V. Stochastic Geometry MaMIMO Regime

As argued earlier, an appropriate model for the covariance subspaces $\mathbf{C}$ is a Haar distribution (randomly oriented semi-unitary matrices). However, as we shall consider that the rank $L$ remains finite, whereas $M$ grows unboundedly, for the large system analysis we may equivalently consider the elements of $\mathbf{C}$ as i.i.d. with zero mean and variance $1/M$ so that the expected squared norm (and asymptotically the actual squared norm) of the columns of $\mathbf{C}$ is normalized to 1. The subspaces $\mathbf{C}$ of different channels are considered independent. For the large system analysis, we use Theorem 1, Lemma $1, 4, 6$ from [3]. From (12), we can see that $\mathbf{g}'_k$ will be of the form $[\sum_{i\neq k} \beta_i \mathbf{S}_{i,b_k} + \mu_{b_k} \mathbf{I}]^{-1} \mathbf{C}_{k,b_k} \mathbf{b}_k$, where $\mathbf{b}_k = \alpha_k \mathbf{W}_{k,b_k} \mathbf{C}_{k,b_k}^H \mathbf{g}'_k$ is of size $L_{k,b_k} \times 1$. $\mathbf{b}_k$ can be written as,

$$\mathbf{b}_k = \alpha_k \mathbf{W}_{k,b_k} \mathbf{C}_{k,b_k}^H [\sum_{i\neq k} \beta_i \mathbf{S}_{i,b_k} + \mu_{b_k} \mathbf{I}]^{-1} \mathbf{C}_{k,b_k} \mathbf{b}_k. \tag{18}$$

Hence, $\mathbf{b}_k$ is the max eigen vector of $\mathbf{W}_{k,b_k} \mathbf{C}_{k,b_k}^H (\sum_{i\neq k} \beta_i \mathbf{S}_{i,b_k} + \mu_{b_k} \mathbf{I})^{-1} \mathbf{C}_{k,b_k}$. Asymptotically $\mathbf{C}_{k,b_k}^H (\sum_{i\neq k} \beta_i \mathbf{S}_{i,b_k} + \mu_{b_k} \mathbf{I})^{-1} \mathbf{C}_{k,b_k}$ converges to a deterministic limit which is a multiple of identity, $e_{b_k} \mathbf{I}$, where $e_{b_k}$ is defined below. This leads to $\mathbf{b_k} = \mathbf{V}_{max}(\mathbf{W}_{k,b_k})$. Now we derive the deterministic equivalents for terms of the form $\mathbf{C}_{k,b_k}^H [\sum_{i\neq k} \beta_i \mathbf{S}_{i,b_k} + \mu_{b_k} \mathbf{I}]^{-1} \mathbf{C}_{k,b_k}$. Considering the eigen decomposition of $\mathbf{W}_{k,b_k} = \mathbf{V}_{k,b_k} \mathbf{\Lambda}_{k,b_k} \mathbf{V}_{k,b_k}^H$, where $\mathbf{\Lambda}_{k,b_k} = \text{diag}(\zeta_{k,b_k}^{(1)}, ..., \zeta_{k,b_k}^{(L_{k,b_k})})$ and let $\mathbf{C}_{k,b_k} \mathbf{V}_{k,b_k} = \mathbf{C}'_{k,b_k}$, with $\mathbf{C}'_{k,b_k}$ still being Haar matrix. Considering the term $\mathbf{C}_{k,b_k}^H [\sum_{i\neq k} \beta_i \mathbf{C}_{i,b_k} \mathbf{W}_{i,b_k} \mathbf{C}_{i,b_k}^H + \mu_{b_k} \mathbf{I}]^{-1} \mathbf{C}_{k,b_k}$, we can use Lemma 4 in Appendix VI of [3], that $\boldsymbol{x}_N^H \mathbf{A}_N \boldsymbol{x}_N \xrightarrow{N\to\infty} (1/N) tr \mathbf{A}_N$ when the elements of $\boldsymbol{x}_N$ are iid with variance $1/N$ and independent of $\mathbf{A}_N$, and similarly when $\boldsymbol{y}_N$ is independent of $\boldsymbol{x}_N$, that $\boldsymbol{x}_N^H \mathbf{A}_N \boldsymbol{y}_N \xrightarrow{N\to\infty} 0$. Hence the expression above goes to,

$$\mathbf{C}_{k,b_k}^H [\sum_{i\neq k} \beta_i \mathbf{C}_{i,b_k} \mathbf{W}_{i,b_k} \mathbf{C}_{i,b_k}^H + \mu_{b_k} \mathbf{I}]^{-1} \mathbf{C}_{k,b_k} = \\ \frac{1}{M_{b_k}} \text{tr}\{[\sum_{i\neq k} \beta_i \mathbf{C}_{i,b_k} \mathbf{W}_{i,b_k} \mathbf{C}_{i,b_k}^H + \mu_{b_k} \mathbf{I}]^{-1}\} \mathbf{I}_{L_{k,b_k}}. \tag{19}$$

Further we apply Lemma 6 from [3] which states that $\frac{1}{N}\text{tr}\{\mathbf{A}_N^{-1}\} - \frac{1}{N}\text{tr}\{(\mathbf{A}_N + \boldsymbol{v}\boldsymbol{v}^H)^{-1}\} \xrightarrow{N\to\infty} 0$ to approximate $[\sum_{i\neq k} \beta_i \mathbf{C}_{i,b_k} \mathbf{W}_{i,b_k} \mathbf{C}_{i,b_k}^H + \mu_{b_k} \mathbf{I}]^{-1} = [\sum_{i=1} \beta_i \mathbf{C}_{i,b_k} \mathbf{W}_{i,b_k} \mathbf{C}_{i,b_k}^H + \mu_{b_k} \mathbf{I}]^{-1}$. Now using Theorem 1 from [3] that any terms of the form $\frac{1}{N}\text{tr}\{(\mathbf{A}_N - z\mathbf{I}_N)^{-1}\}$, where $\mathbf{A}_N$ is the summation of independent rank one matrices with covariance matrix $\mathbf{\Theta}_i$ is equal to the unique positive solution of $e_j = \frac{1}{N}\text{tr}\{(\sum_{i=1}^{K} \frac{\mathbf{\Theta}_i}{1+e_i} - z\mathbf{I}_N)^{-1}\}$. Thus $\text{tr}\{[\sum_{i\neq k} \beta_i \mathbf{C}_{i,b_k} \mathbf{W}_{i,b_k} \mathbf{C}_{i,b_k}^H + \mu_{b_k} \mathbf{I}]^{-1}\}$ can be simplified as $e_{b_k}$, which is defined as the solution to the following fixed point equation,

$$e_c = \left( \frac{1}{M_c} \sum_{i=1}^{K} \sum_{r=1}^{L_{i,c}} \frac{\beta_i \zeta_{i,c}^{(r)}}{1 + \beta_i \zeta_{i,c}^{(r)} e_c} + \mu_c \right)^{-1}. \tag{20}$$

Now (18) gets simplified as, $\mathbf{b}_k = \alpha_k e_{b_k} \mathbf{W}_{k,b_k} \mathbf{b}_k$ and thus $\mathbf{b}_k$ will be the max eigen vector of $\mathbf{W}_{k,b_k}$. Finally we write the optimized BF w.r.t. partial CSIT as,

$$\mathbf{g}_k = [\sum_{i\neq k} \beta_i \mathbf{S}_{i,b_k} + \mu_{b_k} \mathbf{I}]^{-1} \mathbf{C}_{k,b_k} \boldsymbol{v}_{k,b_k}, \ \text{where,} \\ \boldsymbol{v}_{k,b_k} = \mathbf{V}_{max}(\mathbf{W}_{k,b_k}). \tag{21}$$

Computation of eigen values $\zeta_{k,b_i}^{(r)}$ of $\mathbf{W}_{k,b_i}$: from Section III,

$$\mathbf{W}_{k,b_i} = \check{\mathbf{c}}_{k,b_i} \check{\mathbf{c}}_{k,b_i}^H + \widetilde{\mathbf{D}}_{k,b_i}, \ \check{\mathbf{c}}_{k,b_i} = \widehat{\mathbf{D}}_{k,b_i}^{1/2} \widehat{\mathbf{c}}_{k,b_i}, \forall i, k \tag{22}$$

In (3), we assume that all the eigen values are equal and positive, i.e $\mathbf{D}_{k,b_i} = \eta_{k,b_i} \mathbf{I}$, $\widetilde{\mathbf{D}}_{k,b_i} = \widetilde{\eta}_{k,b_i} \mathbf{I}$. Thus the eigen values of $\mathbf{W}_{k,b_i}$ can be shown to be $\zeta_{k,b_i}^{(1)} = \lambda_{max}(\mathbf{W}_{k,b_i}) = \|\check{\mathbf{c}}_{k,b_i}\|^2 + \widetilde{\eta}_{k,b_i}$ and $\zeta_{k,b_i}^{(2)} = .... = \zeta_{k,b_i}^{(L_{k,b_i})} = \widetilde{\eta}_{k,b_i}$, where $\widetilde{\eta}_{k,b_i} = \frac{\widetilde{\sigma}_{k,b_i}^2 \eta_{k,b_i}}{\widetilde{\sigma}_{k,b_i}^2 + \eta_{k,b_i}}$, using the definition of $\widetilde{\mathbf{D}}_{k,b_i}$ from (6). $\lambda_{max}(\mathbf{W}_{k,b_i})$ is random since $\check{\mathbf{c}}_{k,b_i}$ is random. By the law of large numbers (assuming $L_{k,b_i}$ is large but finite) we replace it by the expectation which can be computed as follows. $\text{E}(\lambda_{max}(\mathbf{W}_{k,b_i})) = \text{E}(\widehat{\mathbf{c}}_{k,b_i}^H \widehat{\mathbf{D}}_{k,b_i} \widehat{\mathbf{c}}_{k,b_i}) + \widetilde{\eta}_{k,b_i}$. This gets simplified as, $\text{E}\left(\lambda_{max}(\mathbf{W}_{k,b_i})\right) = L_{k,b_i} \widehat{d}_{k,b_i} + \widetilde{\eta}_{k,b_i}$, where $\widehat{d}_{k,b_i} = \frac{\eta_{k,b_i}^2}{\eta_{k,b_i} + \widetilde{\sigma}_{k,b_i}^2}$ from (4) ($\widehat{\mathbf{D}}_{k,b_i} = \widehat{d}_{k,b_i} \mathbf{I}$) and $\text{E}(\widehat{\mathbf{c}}_{k,b_i}^H \widehat{\mathbf{c}}_{k,b_i}) = L_{k,b_i}$ from (4).

## VI. Large System Analysis of SINR and Power

In this section, we derive under large system limit, with $M_c, K_c \to \infty$ at a fixed ratio $\frac{K_c}{M_c} < 1, \forall c$, approximations to the scalar quantities involved in the rate expression, which we denote as the deterministic equivalent.

*Theorem 1:* In the large system limit, the quantities $\sigma_k^{(1)} - \overline{\sigma}_k^{(1)} \xrightarrow[a.s]{M_{b_k}\to\infty} 0$, $\sigma_k^{(2)} - \overline{\sigma}_k^{(2)} \xrightarrow[a.s]{M_{b_k}\to\infty} 0$, $r_{\overline{k}} - \overline{r}_{\overline{k}} \xrightarrow[a.s]{M_{b_k}\to\infty} 0$ and $r_k - \overline{r}_k \xrightarrow[a.s]{M_{b_k}\to\infty} 0$, where $\overline{\sigma}_k^{(1)}, \overline{\sigma}_k^{(2)}, \overline{r}_{\overline{k}}, \overline{r}_k$ are the deterministic equivalents. Here $\xrightarrow[a.s]{M_{b_k}\to\infty}$ denotes almost sure convergence. Further we can show that, since the logarithm is a continuous function, by applying the continuous mapping theorem [17], it follows from the almost sure convergence of $r_k$ and $r_{\overline{k}}$ that, $R_k - \overline{R}_k \xrightarrow[a.s]{M_{b_k}\to\infty} 0$, where $R_k$ is the rate of user $k$, with $\overline{R}_k = \ln(\frac{\overline{r}_k}{\overline{r}_{\overline{k}}})$. By using similar argument, we state that $\beta_k - \overline{\beta}_k \xrightarrow[a.s]{M_{b_k}\to\infty} 0$ and $\alpha_k - \overline{\alpha}_k \xrightarrow[a.s]{M_{b_k}\to\infty} 0$. The deterministic limits are obtained as,

$$\begin{aligned}
&\overline{\sigma}_k^{(1)} = \frac{e_{b_k}^2 \lambda_{max}(\mathbf{W}_{k,b_k})}{e'_{b_k}\left(1+\Upsilon_{\overline{k}}\right)},\\
&\overline{\sigma}_k^{(2)} = \frac{1}{M_{b_k}} \sum_{i=1, i\neq k}^{K} \overline{\beta}_i [\sum_{r=1}^{L_{i,b_k}} \frac{\zeta_{i,b_k}^{(r)}}{(1+\overline{\beta}_i \zeta_{i,b_k}^{(r)} e_{b_k})^2}],\\
&\overline{r}_{\overline{k}} = 1 + \Upsilon_{\overline{k}},\\
&\overline{r}_k = 1 + \Upsilon_{\overline{k}} + p_k \frac{e_{b_k}^2 \lambda_{max}(\mathbf{W}_{k,b_k})}{e'_{b_k}},
\end{aligned} \tag{23}$$

where,

$$\begin{aligned}
&\Upsilon_{\overline{k}} = \sum_{\substack{i=1,\\ i\neq k}}^{K} p_i \frac{1}{M_{b_i}} [\sum_{r=1}^{L_{k,b_i}} \frac{\zeta_{k,b_i}^{(r)}}{(1+\beta_k \zeta_{k,b_i}^{(r)} e_{b_i})^2}],\\
&\overline{\beta}_k = u_k \left(\frac{1}{\overline{r}_{\overline{k}}} - \frac{1}{\overline{r}_k}\right),\ \overline{\alpha}_k = \frac{u_k}{\overline{r}_k},
\end{aligned} \tag{24}$$

*Proof:* Main steps leading to this using standard results from random matrix theory [3] are outlined in Appendix A.
All the deterministic equivalents described above depend just on the scalar parameters such as eigen values of the channel covariance matrices, transmit powers and channel estimation error variances. Note that the BF computation algorithm based on EWSR still remains iterative in $p_i, \beta_i$ and $\mu_{b_i}$.

## VII. Various BF Sum Rate Expressions for Multi Cell with LMMSE/Subspace Channel Estimator

In this section, we consider the simplified sum rate expressions for naive, EWSMSE and EWSR ZF BFs for LMMSE/Subspace channel estimators under multi cell (C cells), with identical parameters, $\widetilde{\sigma}_{k,c}^2 = \widetilde{\sigma}^2$, $L_{k,c} = L$, $\mathbf{D}_{k,c} = \eta\mathbf{I}$ and $M_c = M, \forall k, c$. Number of users in cell $c$ is denoted as $K_c = K/C, \forall c$. We denote,

$$\zeta_{k,b_k}^{(1)} = L\frac{\eta^2}{\widetilde{\sigma}^2+\eta} + \frac{\widetilde{\sigma}^2\eta}{\widetilde{\sigma}^2+\eta}, \tag{25}$$

and rest of the eigen values $\zeta_{k,b_k}^{(r)} = \frac{\widetilde{\sigma}^2\eta}{\widetilde{\sigma}^2+\eta}, \forall r = 2, ..., L$. For convenience, we define the terms $\lambda_1 = \frac{\eta^2}{\eta+\widetilde{\sigma}^2}$ and $\lambda_2 = \frac{\widetilde{\sigma}^2\eta}{\widetilde{\sigma}^2+\eta}$. Further at high SNR, BS power becomes equally distributed among the users. Substituting these values, leads to the following simplified expressions for sum rate.

**Optimal ZF BF with LMMSE:**

$$\overline{R}_{LMMSE} = K\ln\left(1 + (1 - \tfrac{(K-1)L}{M_{b_k}})(L\lambda_1 + \lambda_2)\tfrac{P}{K}\right). \tag{26}$$

**Naive BF with LMMSE:** $\widetilde{\mathbf{D}}_{k,b_k} = 0$ leading to $\zeta_{k,b_k}^{(1)} = L\lambda_1$ and rest of the eigen values $\zeta_{k,b_k}^{(r)} = 0, \forall r = 2, ..., L$.

$$\overline{R}_{naiveLMMSE} = K\ln\left(1 + (1 - \tfrac{(K-1)L}{M_{b_k}})L\lambda_1\tfrac{P}{K}\right). \tag{27}$$

For naive LMMSE, maximum eigen value for $\mathbf{W}_{k,b_k}$ ($\zeta_{k,b_k}^{(1)} = L\lambda_1$) is lesser compared to the optima ZF BF case and hence the performance degrades compared to the optimal BF case.

**EWSMSE BF with LMMSE:** In this case, the error covariance for the intended user get moved to the interference part, thus the sum rate expression will be,

$$\overline{R}_{EWSMSE} = K\ln\left(1 + \frac{(1-\frac{(K-1)L}{M_{b_k}})L\lambda_1}{(1-\frac{(K-1)L}{M_{b_k}})L\lambda_2+1}\frac{P}{K}\right). \tag{28}$$

Compared to naive LMMSE, there is scaling factor in the denominator factor accounting for the channel estimation error power being moved to the noise part (8) and this explains the performance degradation for EWSMSE.

**Optimal BF with Subspace Estimator:** For subspace estimator, the sum rate expression $\overline{R}_{Subspace}$ is similar to the optimal BF with LMMSE case, (29), but with different eigen values. The eigen values can be derived as $\zeta_{k,b_k}^{(1)} = L + \widetilde{\sigma}^2$ and $\zeta_{k,b_k}^{(r)} = \widetilde{\sigma}^2, \forall r = 2, ..., L$.

$$\overline{R}_{Subspace} = K\ln\left(1 + (1 - \tfrac{(K-1)L}{M_{b_k}})(L + \widetilde{\sigma}^2)\tfrac{P}{K}\right). \tag{29}$$

## VIII. Simulation Results

In this section, we present the Ergodic Sum Rate Evaluations for BF design for the various channel estimates. Monte Carlo evaluations of ergodic sum rates are done with the following parameters: $C$, number of cells. $K_c$, number of (single-antenna) users in cell $c$ and $K = \sum_c K_c$. $M$, number of transmit antennas in each cell. We consider a path-wise or low rank channel model as in section III, with $L$ = number of paths = channel covariance rank. $d$ : scale factor in the LS channel estimation error variance $\widetilde{\sigma}^2 = d/SNR$. Notations: in the figures, iCSIT refers to the optimal BF design for the instantaneous CSIT case [18].

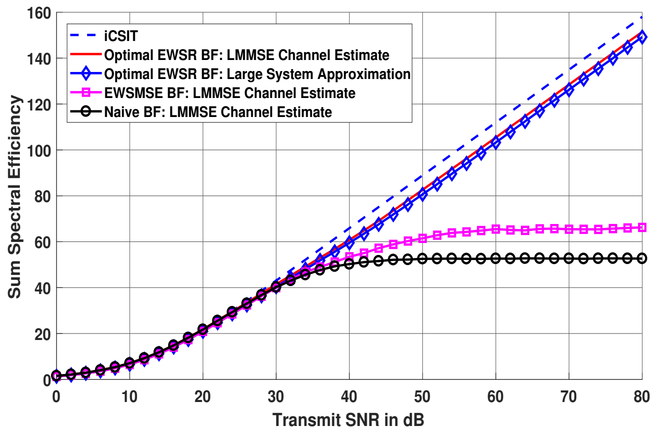


Fig. 1. EWSR for $C = 1$ cell, $K = 10$ users, $M = 64$, $L = 4$, $\widetilde{\sigma}^2 = 0.1$.

In Figure 1, we plot the optimal BF performance with LMMSE channel estimator comparing to the optimal BF performance for the case of large system approximation. It is evident that the deterministic approximations are accurate even for finite $M, K$. Further, we have fixed the channel estimator error variance $\widetilde{\sigma}^2$ to be 0.1 and thus it is evident from the figure that exploiting the channel estimation error covariance information has significant performance gain compared to the sub-optimal methods such as EWSMSE and naive BFs.

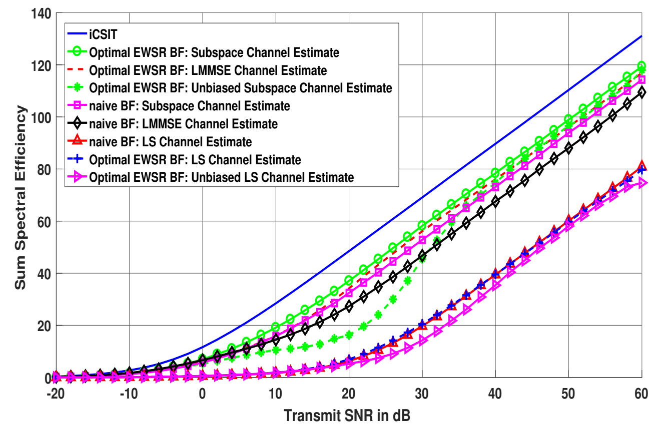


Fig. 2. EWSR for $C = 1$ cell, $K_1 = K = 10$ users, $M = 64$, $L = 2$, $\widetilde{\sigma}^2 = 1/SNR$.

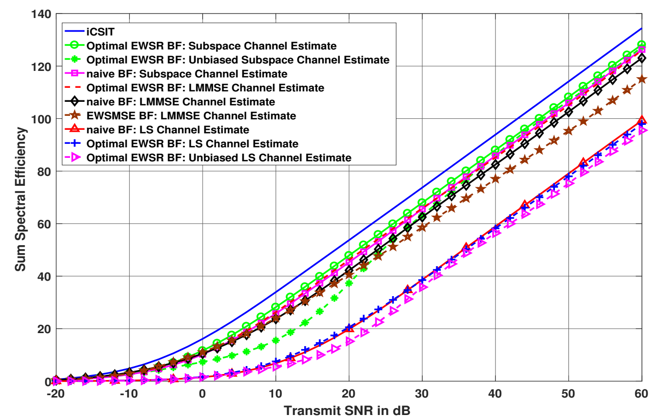


Fig. 3. EWSR for $C = 2$ cells, $K_1 = K_2 = 5$ users, $M = 32$, $L = 2$, $\widetilde{\sigma}^2 = 1/SNR$.

In Figure 3, we plot the EWSMSE beamforming performance also and it is evident from the figure that ESEI-WSR based beamformers (i.e. MaMISO limit based) perform better EWSR approximations than a EWSMSE design. From the numerical simulations in both Figures 2,3, it is quite evident that just using LS channel estimates may lead to substantial EWSR loss. In Massive MIMO, the exploitation of channel subspaces (reduced rank covariances) in channel estimates may lead to substantial reductions in SNR loss due to partial CSIT. Moreover, there is significant gain from exploiting (error) channel covariances in addition to (LMMSE) channel estimates and proper handling of channel error covariance in the direct link in the BF design.

## IX. Conclusion

This paper investigated the optimal linear precoder based on partial CSIT in the multi-cell MU-MISO downlink. We introduced a stochastic geometry inspired randomization of the channel covariance eigen spaces and analyzed the large system behavior. This leads to simpler analytical results with SINR or the user rates depending only on some scalar quantities such as eigen value profile, channel rank, the number of antennas $M$ or users $K$ and the channel estimation error variance. Moreover, we show the improvement in performance by using an LMMSE channel estimate compared to just having LS estimates, and by furthermore properly exploiting all covariance information. Numerical simulations suggest that the large system approximations are accurate even for finite values of $M, K$.

## Appendix A
## Proof of Theorem 1

First we compute the deterministic equivalent for $\sigma_k^{(1)}$,

$$\sigma_k^{(1)} = \widetilde{r}_{\overline{k}}^{-1}\mathbf{g}_k'^H\widehat{\mathbf{S}}_{k,b_k}\mathbf{g}_k' = \widetilde{r}_{\overline{k}}^{-1}\mathbf{g}_k'^H\mathbf{C}_{k,b_k}\mathbf{W}_{k,b_k}\mathbf{C}_{k,b_k}^H\mathbf{g}_k'. \tag{30}$$

Using the eigen decomposition of $\mathbf{W}_{k,b_k} = \mathbf{V}_{k,b_k}\mathbf{\Lambda}_{k,b_k}\mathbf{V}_{k,b_k}^H$,

$$\begin{aligned} &\mathbf{g}_k'^H\mathbf{S}_{k,b_k}\mathbf{g}_k' = \mathbf{g}_k'^H\mathbf{C}_{k,b_k}\mathbf{W}_{k,b_k}\mathbf{C}_{k,b_k}^H\mathbf{g}_k', \\ &\mathbf{g}_k' = \mathbf{g}_k''/\left\|\mathbf{g}_k''\right\|, \mathbf{g}_k'' = \mathbf{\Gamma}_k^{-1}\mathbf{C}_{k,b_k}\boldsymbol{v}_{k,b_k}, \\ &\mathbf{C}_{k,b_k}^H\mathbf{g}_k' = \mathbf{C}_{k,b_k}^H\mathbf{\Gamma}_k^{-1}\mathbf{C}_{k,b_k}\boldsymbol{v}_{k,b_k}/||\mathbf{g}_k''||, \end{aligned} \tag{31}$$

where $\mathbf{\Gamma}_k = \sum_{i \neq k}\beta_i\mathbf{S}_{i,b_k} + \mu_{b_k}\mathbf{I}$. Using large system analysis simplifications shown in (20), $\mathbf{C}_{k,b_k}^H\mathbf{\Gamma}_k^{-1}\mathbf{C}_{k,b_k} = e_{b_k}\mathbf{I}$,

$$\mathbf{g}_k'^H\mathbf{S}_{k,b_k}\mathbf{g}_k' = \frac{e_{b_k}^2\boldsymbol{v}_{k,b_k}^H\mathbf{W}_{k,b_k}\boldsymbol{v}_{k,b_k}}{\|\mathbf{g}_k''\|^2} = \frac{e_{b_k}^2\lambda_{max}(\mathbf{W}_{k,b_k})}{\|\mathbf{g}_k''\|^2}, \tag{32}$$

We define $\mathbf{\Gamma}_{b_k} = \mathbf{\Gamma}_k + \beta_k\mathbf{S}_{k,b_k}$. Further we consider simplifying $\|\mathbf{g}_k''\|^2 = \boldsymbol{v}_{k,b_k}^H\mathbf{C}_{k,b_k}\mathbf{\Gamma}_k^{-2}\mathbf{C}_{k,b_k}\boldsymbol{v}_{k,b_k}$. By using Lemma 4 from [3] leads to $\|\mathbf{g}_k''\|^2 = \frac{1}{M_{b_k}}\text{tr}\{\mathbf{\Gamma}_k^{-2}\}\left\|\boldsymbol{v}_{k,b_k}\right\|^2 = \frac{1}{M_{b_k}}\text{tr}\{\mathbf{\Gamma}_k^{-2}\}$. Further, by using Lemma 6 we approximate $\mathbf{\Gamma}_k^{-1} \approx (\mathbf{\Gamma}_k + \beta_k\mathbf{S}_{k,b_k})^{-1} = \mathbf{\Gamma}_{b_k}^{-1}$. From [3], in the large system limit, for $(1/M_{b_k})\text{tr}\{\mathbf{\Gamma}_{b_k}^{-2}\}$, we have an almost sure convergence value as $e'_{b_k}$, where $e'_{b_k}$ is the derivative of $e_{b_k}$ w.r.t. $\mu_{b_k}$, and thus $\|\mathbf{g}''_k\|^2 = e'_{b_k}$,

$$\begin{aligned} e'_{b_k} &= e^2_{b_k}\big(\tfrac{1}{M_{b_k}}\sum_{i=1}^{K}\sum_{r=1}^{L_{i,b_k}}\frac{\beta_i^2\zeta_{i,b_k}^{(r),2}e'_{b_k}}{(1+\beta_i\zeta_{i,b_k}^{(r)}e_{b_k})^2}+1\big) \\ \implies e'_{b_k} &= \frac{e^2_{b_k}}{1-\frac{e^2_{b_k}}{M_{b_k}}\sum_{i=1}^{K}\sum_{r=1}^{L_{i,b_k}}\frac{\beta_i^2\zeta_{i,b_k}^{(r),2}}{(1+\beta_i\zeta_{i,b_k}^{(r)}e_{b_k})^2}}. \end{aligned} \quad (33)$$

Deterministic limit for $r_{\overline{k}}, r_k$: Each term in $r_{\overline{k}}$ is of the form $p_i\mathbf{g}''_i\mathbf{S}_{k,b_i}\mathbf{g}''_i/\|\mathbf{g}''_i\|^2$, where $\mathbf{g}''^H_i\mathbf{S}_{k,b_i}\mathbf{g}''_i = \boldsymbol{v}'^H_{i,b_i}\boldsymbol{\Gamma}_i^{-1}\mathbf{C}_{k,b_i}\mathbf{W}_{k,b_i}\mathbf{C}^H_{k,b_i}\boldsymbol{\Gamma}_i^{-1}\boldsymbol{v}'_{i,b_i}$ and we defined $\boldsymbol{v}'_{i,b_i} = \mathbf{C}_{i,b_i}\boldsymbol{v}_{i,b_i}$. Since $\boldsymbol{v}'_{i,b_i}$ is independent of all other random quantities in this expression, we apply Lemma 4 and then Lemma 6 to get, $\boldsymbol{v}'^H_{i,b_i}\boldsymbol{\Gamma}_i^{-1}\mathbf{C}_{k,b_i}\mathbf{W}_{k,b_i}\mathbf{C}^H_{k,b_i}\boldsymbol{\Gamma}_i^{-1}\boldsymbol{v}'_{i,b_i} = \frac{1}{M_{b_i}}\text{tr}\{\boldsymbol{\Gamma}_{b_i}^{-1}\mathbf{C}_{k,b_i}\mathbf{W}_{k,b_i}\mathbf{C}^H_{k,b_i}\boldsymbol{\Gamma}_{b_i}^{-1}\}$. Applying Lemma 1 to each of the rows of $\mathbf{V}^H_{k,b_i}\mathbf{C}^H_{k,b_i}\boldsymbol{\Gamma}_i^{-1}$, then Lemma 4 and 6, we obtain the following simplified expression,

$$\begin{aligned} &\tfrac{1}{M_{b_i}}\text{tr}\{\boldsymbol{\Gamma}_{b_i}^{-1}\mathbf{S}_{k,b_i}\boldsymbol{\Gamma}_{b_i}^{-1}\} = \tfrac{1}{M^2_{b_i}}\text{tr}\{\boldsymbol{\Gamma}_{b_i}^{-2}\}\text{tr}\{\boldsymbol{\Lambda}_{k,b_i}\mathbf{B}^{-2}_{k,b_i}\}, \\ &\text{where, } \mathbf{B}_{k,b_i} = \text{diag}(1+\beta_k\zeta_{k,b_i}^{(1)}e_{b_i},...,1+\beta_k\zeta_{k,b_i}^{(L)}e_{b_i}). \end{aligned} \quad (34)$$

Finally we obtain,

$$\begin{aligned} &\sum_{\substack{i=1,\\ i\neq k,}}^{K} p_i\mathbf{g}'^H_i\mathbf{S}_{k,b_k}\mathbf{g}'_i = \\ &\sum_{\substack{i=1,\\ i\neq k}}^{K} p_i\frac{1}{M_{b_i}}\Big[\sum_{r=1}^{L_{k,b_i}}\frac{\zeta_{k,b_i}^{(r)}}{(1+\beta_k\zeta_{k,b_i}^{(r)}e_{b_i})^2}\Big] = \Upsilon_{\overline{k}}, \end{aligned} \quad (35)$$

Thus we can write $\overline{r}_k$ and $\overline{r}_{\overline{k}}$,

$$\begin{aligned} &\overline{r}_{\overline{k}} = 1+\Upsilon_{\overline{k}},\ \overline{r}_k = 1+\Upsilon_{\overline{k}}+p_k\frac{e^2_{b_k}\lambda_{max}(\mathbf{W}_{k,b_k})}{e'_{b_k}}, \\ &\text{Also, } \overline{\beta}_k = u_k\left(\tfrac{1}{\overline{r}_{\overline{k}}}-\tfrac{1}{\overline{r}_k}\right),\ \overline{\alpha}_k = \tfrac{u_k}{\overline{r}_k}. \end{aligned} \quad (36)$$

Finally, combining (32), (33), (36), we can write the deterministic equivalent for $\sigma_k^{(1)}$ as, $\overline{\sigma}_k^{(1)} = \frac{e^2_{b_k}\lambda_{max}(\mathbf{W}_{k,b_k})}{e'_{b_k}(1+\Upsilon_{\overline{k}})}$. Each term in $\sigma_k^{(2)}$ is of the form $\widehat{\beta}_i\mathbf{g}''^H_k\mathbf{S}_{i,b_k}\mathbf{g}''^H_k/\|\mathbf{g}''_k\|^2$, which gets simplified as follows:

$$\begin{aligned} &\mathbf{g}''^H_k\mathbf{S}_{i,b_k}\mathbf{g}''^H_k = \boldsymbol{v}'^H_{k,b_k}\boldsymbol{\Gamma}_k^{-1}\mathbf{C}_{i,b_k}\mathbf{W}_{i,b_k}\mathbf{C}^H_{i,b_k}\boldsymbol{\Gamma}_k^{-1}\boldsymbol{v}'_{k,b_k} \\ &\overset{(a)}{=} \tfrac{1}{M_{b_k}}\text{tr}\{\boldsymbol{\Gamma}_k^{-1}\mathbf{C}_{i,b_k}\mathbf{W}_{i,b_k}\mathbf{C}^H_{i,b_k}\boldsymbol{\Gamma}_k^{-1}\}, \end{aligned} \quad (37)$$

where (a) follows from Lemma 4 (since $\boldsymbol{v}'_{k,b_k}$ is independent of all other matrices involved). By following the same steps as in (34)-(35), this gets simplified and we write $\overline{\sigma}_k^{(2)}$ as,

$$\begin{aligned} &\overline{\sigma}_k^{(2)} = \sum_{i=1,i\neq k}^{K}\overline{\beta}_i\mathbf{g}''^H_k\mathbf{S}_{i,b_k}\mathbf{g}''^H_k/\|\mathbf{g}''_k\|^2 = \\ &\tfrac{1}{M_{b_k}}\sum_{i=1,i\neq k}^{K}\overline{\beta}_i\Big[\sum_{r=1}^{L_{i,b_k}}\frac{\zeta_{i,b_k}^{(r)}}{(1+\overline{\beta}_i\zeta_{i,b_k}^{(r)}e_{b_k})^2}\Big]. \end{aligned} \quad (38)$$


## ACKNOWLEDGMENTS

EURECOM's research is partially supported by its industrial members: ORANGE, BMW, SFR, ST Microelectronics, Symantec, SAP, Monaco Telecom, iABG, and by the projects MASS-START (French FUI) and DUPLEX (French ANR).


## REFERENCES


[1] E.G. Larsson, O. Edfors, F. Tufvesson, and T.L. Marzetta, "Massive MIMO for Next Generation Wireless Systems," *IEEE Comm's Mag.*, Feb. 2014.

[2] Wassim Tabikh, Dirk Slock, and Yi Yuan-Wu, "MIMO IBC beamforming with combined channel estimate and covariance CSIT," in *IEEE International Symposium on Information Theory*, 2017.

[3] S. Wagner, R. Couillet, M. Debbah, and D. T. M. Slock, "Large system analysis of linear precoding in MISO broadcast channels with limited feedback," *IEEE Trans. Inf. Theory*, vol. 58, no. 7, pp. 4509–4538, July 2012.

[4] S. Wagner and Dirk Slock, "Weighted sum rate maximization of correlated MISO broadcast channels under linear precoding : a large system analysis," in *Proc. IEEE Int'l Workshop on Sig. Proc. Advances in Wireless Comm's (SPAWC), San Fransisco, USA*, 2011.

[5] S. Lakshminarayana, M. Assaad, and M. Debbah, "Coordinated multicell beamforming for massive MIMO: A random matrix approach," *IEEE Trans. Information Theory*, vol. 61, no. 6, pp. 3387 – 3412, June 2015.

[6] A. Muller, R. Couillet, E. Bjornson, S. Wagner, and M. Debbah, "Interference-Aware RZF Precoding for Multi Cell Downlink Systems," *IEEE Trans. Signal Proc.*, vol. 63, no. 15, pp. 3959 – 3973, August 2015.

[7] Luca Sanguinetti, Aris L. Moustakas, Emil Bjornson, and Merouane Debbah, "Large system analysis of the energy consumption distribution in multi-user MIMO systems with mobility," *IEEE Trans. on Wirel. Commun.*, vol. 14, no. 3, pp. 1730–1745, March 2015.

[8] R. Zakhour and S. V. Hanly, "Base station cooperation on the downlink: Large system analysis," *IEEE Trans. on Info. Theo.*, vol. 58, no. 4, pp. 2079–2106, April 2012.

[9] Y. Huang, C. W. Tan, and B. D. Rao, "Large system analysis of power minimization in multiuser MISO downlink with transmit-side channel correlation," in *Intl. Sympo. on Info. Theo. and its Applic. (ISIT)*, October 2012.

[10] C. K. Thomas and Dirk Slock, "Reduced-order zero-forcing beamforming vs optimal beamforming and dirty Paper coding and massive MIMO Analysis," in *10th IEEE Sens. Arr. and Mul.chnl. Sig. Process. Wkshp.*, Sheffield, UK, 2018.

[11] F. Negro, I. Ghauri, and Dirk T. M. Slock, "Sum rate maximization in the noisy MIMO interfering broadcast channel with partial CSIT via the expected weighted MSE," in *Proc. Int'l Symposium on Wireless Communication Systems, ISWCS, Paris, France*, 2012.

[12] X. Li, E. Bjornson, 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.

[13] G. George, R. K. Mungara, A. Lozano, and M. Haenggi, "Ergodic Spectral Efficiency in MIMO Cellular Networks," *IEEE Trans. on Wireless Communications*, May 2017.

[14] G. Caire and K. R. Kumar, "Information theoretic foundations of adaptive coded modulation," *Proc. IEEE*, vol. 95, no. 12, pp. 2274–2298, December 2007.

[15] Hamdi Joudeh and Bruno Clerckx, "Sum-rate maximization for linearly precoded downlink multiuser MISO systems with partial CSIT: A rate-splitting approach," *IEEE Transactions on Communications*, vol. 64, no. 11, pp. 4847–4861, November 2016.

[16] S. J. Kim and G. B. Giannakis, "Optimal resource allocation for MIMO ad-hoc cognitive radio networks," in *IEEE Trans. on Inf. Theory*, May 2011, vol. 57, pp. 3117–3131.

[17] P. Billingsley, "Probability and measure.," in *Hoboken, NJ: Wiley*, 1995.

[18] 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 Commun.*, pp. 4792–4799, December 2008.