Postprint

This is the accepted version of a paper presented at *International Conference of Communications*.

N.B. When citing this work, cite the original published paper.

Low-Overhead Coordination in Sub-28 Millimeter-Wave Networks

Hadi Ghauch
School of Electrical Engrng
Royal Inst. of Techn., KTH
ghauch@kth.se

Taejoon Kim
Dept. of Electrical Engrng
and Computer Science
Univ of Kansas, Lawrence
taejoonkim@ku.edu

Mikael Skoglund
School of Electrical Engineering
Royal Inst. of Techn., KTH
skoglund@kth.se

Carlo Fischione
School of Electrical Engrng
Royal Inst. of Techn., KTH
carlofi@kth.se

Abstract—In this paper, we present some contributions from our recent investigation. We address the open issue of interference coordination for sub-28 GHz millimeter-wave communication, by proposing fast-converging coordination algorithms, for dense multi-user multi-cell networks. We propose to optimize a lower bound on the network sum-rate, after investigating its tightness. The bound in question results in distributed optimization, requiring local information at each base station and user. We derive the optimal solution to the transmit and receive filter updates, that we dub non-homogeneous waterfilling, and show its convergence to a stationary point of the bound. We also underline a built-in mechanism to turn-off data streams with low SINR, and allocate power to high-SNR streams. This "stream control" is a at the root of the fast-converging nature of the algorithm. Our numerical result conclude that low-overhead coordination offers large gains, for dense sub-28 GHz systems. These findings bear direct relevance to the ongoing discussions around 5G New Radio.

Index Terms—Sub-28 GHz Millimeter-wave, low-overhead coordination, Difference of Log and Trace (DLT), Non-homogeneous Waterfilling, max-DLT.

I. INTRODUCTION

To address the exponentially increasing demand in 5G systems, communications in the millimeter-wave (mmWave) band are among the most promising candidates [2], due to the large mmWave spectrum. While most investigations of mmWave communcation have been focused on systems above 28 GHz, in the current work, we study multi-user multi-cell coordination, in *sub-28 GHz systems*, e.g., X-band (8-12) GHz, Ku-band (12-18) GHz, and 28 GHz in the Ka band. These systems are characterized by a relatively large antenna spacing (compared to systems beyond 60 GHz), thereby implying that tens (rather than hundreds) of antennas can be fitted on transmitters/receivers. Thus, the urge for hybrid analog-digital precoding is not stringent and *fully digital* precoding/combining is preferred. Moreover, propagation channels are still dominated by Rayleigh/Rician components in non line-of-sight environments [3]. As a result, conventional pilot-based channel estimation techniques are more efficient than beam alignment/sounding [4].

The implication of highly directional wireless links at 60 GHz (and above) is that they are almost interference-free. However, in sub-28 GHz systems, channels are *less sparse* (in terms of eigenmodes), and beamforming has relatively lower directivity than systems in the higher bands. This is attributed to the presence of significant multi-path components, in urban propagation (confirmed by narrowband/wideband channel measurements in the 9.6 GHz, 11.4 GHz, and 28.8 GHz bands [5]).

Consequently, interference may still be a limiting factor in these systems, especially when considering *dense* multi-cell scenarios, where interference management and coordination are still beneficial. Coordination in multi-user multi-cell networks generally refers to the exchanges of information among base stations, to increase the network sum-rate. While these aspects have been investigated at the MAC layer [6], they are still essentially unaddressed at the physical layer. Indeed, the benefits/costs of coordination in sub-28 GHz systems is still an open problem, especially in the case of ultra dense networks - believed to be pervasive in future networks [7]: In these scenarios, ignoring interference for cell-edge users may be a limiting factor on the sum-rate.

In the context of multi-user multi-cell networks, coordination is done using the framework of *Forward-Backward (F-B) iterations*: this over-the-air training leverages local Channel State Information (CSI) at each Base Station (BS) and user, to iteratively optimize the filter at each BS/user, in a fully distributed manner. This framework has been at the heart of most distributed coordination algorithms, such as interference leakage minimization [8], max-SINR [8], minimum mean-squared error [9], and (weighted) sum-rate maximization [10].

Unfortunately, these conventional schemes suffer from extremely elevated overhead, as they require hundreds/thousands of F-B iterations before convergence [12], where the latter increases with the number of BSs, users and transmit/receive antennas [12]. Moreover, mmWave systems will have a larger number of BSs/cells per unit-area (due to their inherent short range), and require a much larger number of BS/user antennas (to mitigate pathloss through array gain), compared to sub-6 GHz systems. Consequently, this shortcoming severely *impairs* the applicability of conventional coordination, to the systems in question. This limitation is reinforced by the *lower coherence time* of mmWave channels. Despite bearing direct relevance to conventional sub-6 GHz, this major limitation has only been addressed in a few recent works [13]–[16], and remains essentially unexplored.
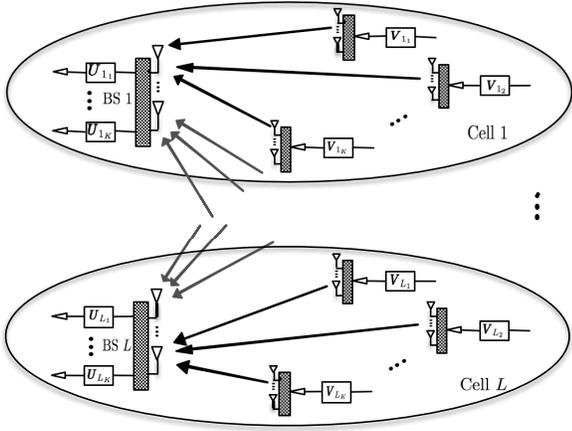
Fig. 1: L-cell MIMO Interfering Multiple-Access Channel

In this work, we design low-overhead distributed coordination algorithms, constrained to operate in a just a few F-B iterations for *increasing dimensions* of the networks. In the algorithm design, we further aim at a *tenfold* reduction in the communication overhead of conventional algorithms. We derive and optimize a lower bound the sum-rate maximization problem in MIMO Interfering Multiple-Access Channels (MIMO IMAC), that we dub *Difference of Log and trace (DLT)*. Unlike the sum-rate, when combined with alternating optimization methods, the DLT expression results in subproblems that are distributed (only requiring local CSI at each BS and user) [1]. Despite their non-convexity, we derive the optimal solution to each of these subproblems, that we dub *non-homogeneous waterfilling* (a variation on the classical waterfilling). This solution turns-off data streams with low-SINR, and allocates power to streams with high SNR. The built-in *"stream-control"* is key to achieving the tenfold increase in convergence speed. Moreover, we show that the devised algorithm converges to a locally optimal solution of the DLT bound. It is revealed that the proposed fast-converging algorithm offers large sum-rate gains, compared to many standard and fast-converging benchmarks. Coordination is still a vital aspect of these systems, and offers major performance gains over uncoordinated transmission While the approach is developed for the MIMO IMAC (i.e., uplink communication), the methods/results are applicable to the downlink, and all its special cases.

II. System Model

The proposed algorithms operate within the framework of F-B iterations/training by exploiting the uplink (UL) and downlink (DL) channel reciprocity. The proposed scheme is designed to operate in the low-overhead regime, by restricting the number of F-B iterations, T, to $T \leq 5$ (a tenfold reduction in communication overhead, over conventional coordination algorithms). We assume that each BS and user posses local CSI only, that is assumed to be perfect (i.e., no CSI errors).

We consider a multi-user multi-cell setting, with L cells/BSs, serving K users each. In the MIMO IMAC case, transmitters at users and receivers are BSs. Each transmitter and receiver have M and N antennas, respectively, and

communicate d data streams. Let $\mathcal{L}$ be the set of BSs, $\mathcal{K}_l$ the set of users served by BS $l \in \mathcal{L}$, and $\mathcal{I}$ the total set of users. Moreover, we denote by user $l_j \in \mathcal{I}$ the jth user ($j \in \mathcal{K}_l$), in the lth cell ($l \in \mathcal{L}$). The recovered signal for user $j \in \mathcal{K}_l$, in cell $l \in \mathcal{L}$ (denoted as user $l_j \in \mathcal{I}$ hereafter),

$$\tilde{s}_{l_j} = U_{l_j}^{\dagger} H_{l,l_j} V_{l_j} s_{l_j}$$
$$+ \sum_{\substack{i \in \mathcal{L} \\ i \neq l}} \sum_{k \in \mathcal{K}} U_{l_j}^{\dagger} H_{l,i_k} V_{i_k} s_{i_k} + U_{l_j}^{\dagger} n_l, \ \forall \ l_j \in \mathcal{I} \quad (1)$$

where $^1 V_{i_k} \in \mathbb{C}^{M \times d}$ is the linear transmit filter for user $i_k \in \mathcal{I}$, $U_{l_j} \in \mathbb{C}^{N \times d}$ is the linear receive filter for user $l_j \in \mathcal{I}$, and H_{l,i_k} the $N \times M$ MIMO channel from user $i_k \in \mathcal{I}$, to BS l assumed to be block-fading.[2] $s_{i_k} \in \mathbb{C}^d$ is the transmit signal of user $i_k \in \mathcal{I}$ with unit power symbols, and n_l is the AWGM noise at receiver l, with $\mathbb{E}[n_l n_l^{\dagger}] = \sigma_l^2 I_N$.

We assume simple decoding, i.e., treating interference as noise , without successive interference cancellation. Then, the achievable rate of user $l_j \in \mathcal{I}$ is given by,

$$r_{l_j} = \log_2 |I_d + (U_{l_j}^{\dagger} R_{l_j} U_{l_j})(U_{l_j}^{\dagger} Q_{l_j} U_{l_j})^{-1}|, \ l_j \in \mathcal{I} \quad (2)$$

where R_{l_j} and Q_{l_j} are the desired signal and interference-plus-noise (I+N) covariance matrices for user j, at BS l, respectively, and are given by,

$$R_{l_j} = H_{l,l_j} V_{l_j} V_{l_j}^{\dagger} H_{l,l_j}^{\dagger}, l_j \in \mathcal{I}$$

$$Q_{l_j} = \sum_{i=1}^{L} \sum_{k=1}^{K} H_{l,i_k} V_{i_k} V_{i_k}^{\dagger} H_{l,i_k}^{\dagger} + \sigma_l^2 I_N - R_{l_j}, l_j \in \mathcal{I}.$$

We let

$$\bar{R}_{i_k} = H_{i,i_k}^{\dagger} U_{i_k} U_{i_k}^{\dagger} H_{i,i_k}, \ i_k \in \mathcal{I}$$

$$\bar{Q}_{i_k} = \sum_{l=1}^{L} \sum_{j=1}^{K} H_{l,i_k}^{\dagger} U_{l_j} U_{l_j}^{\dagger} H_{l,i_k} + \bar{\sigma}_{i_k}^2 I_M - \bar{R}_{i_k}, i_k \in \mathcal{I}$$

denote the desired and I+N covariance matrices of user i_k, in the backward DL network (where $\bar{\sigma}_{i_k}^2$ is the noise power at receiver $i_k \in \mathcal{I}$). Moreover, we let $L_{l_j} L_{l_j}^{\dagger}$ be the Cholesky Decomposition of Q_{l_j}, and $K_{i_k} K_{i_k}^{\dagger}$ as that of $\bar{Q}_{i_k}$.

We aim at maximizing the sum-rate, i.e.,

$$(P) \begin{cases} \max_{\{U_{l_j}, V_{l_j}\}} \sum_{l_j \in \mathcal{I}} r_{l_j} \\ \text{s. t. } \|U_{l_j}\|_F^2 = P_r, \|V_{l_j}\|_F^2 = P_t, \forall l_j \in \mathcal{I} \end{cases} \quad (3)$$

While distributed multi-user multi-cell optimization generally entails a sum-power constraint on the users of a cell (e.g. W-MMSE [10]), we adopt an equal power allocation among the all users within a cell: The BS power is equally split among all its UL (or DL) users, to simplify the presentation. Note

---

[1]*Notation:* we use bold upper-case letters to denote matrices, and bold lower-case denote vectors. For a given matrix A, we define tr(A) as its trace, $\|A\|_F^2$ as its Frobenius norm, $|A|$ as its determinant, $A^{\dagger}$ as its conjugate transpose, and $A^{-\dagger}$ as $(A^{\dagger})^{-1}$. In addition, $A_{(i)}$ denotes its ith column, $A_{i:j}$ columns i to j, $A_{(i,j)}$ element (i,j) in A, $\lambda_i[A]$ the ith eigenvalue of a Hermitian matrix A (assuming the eigenvalues are sorted in decreasing order), and $v_{1:d}[A]$ denotes the d dominant eigenvectors of A. Furthermore, $A \succ 0$ (resp. $A \succeq 0$) implies that A is positive definite (resp. positive semi-definite). Finally, I_n denotes the $n \times n$ identity matrix, $\{n\} = \{1, \cdots, n\}$, and $x^+ = \max\{0, x\}$.

[2]The model/results assume M, N and d for simplicity, and can be easily extended to differ across users and BSs.

that this does not affect the generality of the results.

III. PROPOSED APPROACH

Sum-rate maximization problems, such as (P), are known to be *NP-hard* [17]. Our proposed approach is based on a tractable lower bound formulation that transfers the sum-rate maximization, which is an originally coupled problem, into separable subproblems.

A. Problem Formulation

We focus on the interference-limited regime (in a dense deployment for instance), where we assume

$$\lambda_i[\boldsymbol{U}_{l_j}^\dagger \boldsymbol{Q}_{l_j} \boldsymbol{U}_{l_j}] \to \infty, \; \forall i \in \{d\} \qquad (4)$$

Proposition 1. *Under the conditions in* (4), *the* r_{l_j} *in* (2) *satisfies*

$$r_{l_j} \geq \log_2 |\boldsymbol{I}_d + \boldsymbol{U}_{l_j}^\dagger \boldsymbol{R}_{l_j} \boldsymbol{U}_{l_j}| - \operatorname{tr}(\boldsymbol{U}_{l_j}^\dagger \boldsymbol{Q}_{l_j} \boldsymbol{U}_{l_j}) \triangleq r_{l_j}^{(LB)}, \quad (5)$$

where the gap $\Delta_{l_j} \triangleq r_{l_j} - r_{l_j}^{(LB)}$ *is characterized by*

$$\Delta_{l_j} = \operatorname{tr}(\boldsymbol{U}_{l_j}^\dagger \boldsymbol{Q}_{l_j} \boldsymbol{U}_{l_j}) - \log_2 |\boldsymbol{U}_{l_j}^\dagger \boldsymbol{Q}_{l_j} \boldsymbol{U}_{l_j}|$$
$$+ \mathcal{O}(\operatorname{tr}[(\boldsymbol{U}_{l_j}^\dagger \boldsymbol{Q}_{l_j} \boldsymbol{U}_{l_j})(\boldsymbol{U}_{l_j}^\dagger \boldsymbol{R}_{l_j} \boldsymbol{U}_{l_j})^{-1}]), \; \forall l_j \in \mathcal{I}. \quad (6)$$

Proof: Refer to [1][Appendix B] ∎

In what follows, we shall dub the quantity $r_{l_j}^{(LB)}$ Difference of Log-Trace (DLT). The DLT becomes significant when it is used as an alternative objective of the sum-rate objective in (P). Note that DLT is a lower bound on the sum-rate, R_Σ, and can be written in the following ways:

$$R_\Sigma^{(LB)} = \sum_{l_j \in \mathcal{I}} \log_2 |\boldsymbol{I}_d + \boldsymbol{U}_{l_j}^\dagger \boldsymbol{R}_{l_j} \boldsymbol{U}_{l_j}| - \operatorname{tr}(\boldsymbol{U}_{l_j}^\dagger \boldsymbol{Q}_{l_j} \boldsymbol{U}_{l_j}) \quad (7)$$

$$= \sum_{i_k \in \mathcal{I}} \log_2 |\boldsymbol{I}_d + \boldsymbol{V}_{i_k}^\dagger \bar{\boldsymbol{R}}_{i_k} \boldsymbol{V}_{i_k}| - \operatorname{tr}(\boldsymbol{V}_{i_k}^\dagger \bar{\boldsymbol{Q}}_{i_k} \boldsymbol{V}_{i_k}) \quad (8)$$

The above expressions reveal that DLT makes both the receive filters in (7) and the transmit filters in (8) decoupled, and facilitates the aimed distributed F-B implementation. We formulate the maximal DLT (max-DLT) criterion as a surrogate objective to the sum-rate maximization in (P),

$$(P_{LB}) \begin{cases} \max\limits_{\{\boldsymbol{V}_{l_j}, \boldsymbol{U}_{l_j}\}} R_\Sigma^{(LB)} \\ \text{s. t. } \|\boldsymbol{U}_{l_j}\|_F^2 = P_r, \; \|\boldsymbol{V}_{l_j}\|_F^2 = P_t, \; \forall l_j \in \mathcal{I}. \end{cases} \quad (9)$$

It should be noted that although DLT allows distributed F-B implementation, the problem in (P_{LB}) is not directly solvable since it is non-convex due to the coupling between the transmit and receive filters, and the quadratic equality constraints.

B. Proposed Algorithm

We underline that a coupled optimization like (P_{LB}) can ideally be handled by a Block Coordinate Descent (BCD) approach. This means if the superscript $^{(n)}$ is adopted to denote the iteration number, problem (P_{LB}) is decomposed into a sequence of subproblems that are solved via F-B iterations as below

$$(J1) \; \{\boldsymbol{U}_{l_j}^{(n+1)}\} \triangleq \operatorname*{argmax}_{\{\boldsymbol{U}_{l_j}\}} R_\Sigma^{(LB)} \left(\{\boldsymbol{U}_{l_j}\}, \{\boldsymbol{V}_{l_j}^{(n)}\} \right)$$

$$(J2) \; \{\boldsymbol{V}_{l_j}^{(n+1)}\} \triangleq \operatorname*{argmax}_{\{\boldsymbol{V}_{l_j}\}} R_\Sigma^{(LB)} \left(\{\boldsymbol{U}_{l_j}^{(n+1)}\}, \{\boldsymbol{V}_{l_j}\} \right), \quad (10)$$

Algorithm 1 Maximal DLT (max-DLT)
___
for $t = 1, 2, ..., T$ **do**
  // forward network optimization: receive filter update
    Estimate $\boldsymbol{R}_{l_j}, \boldsymbol{Q}_{l_j}$, and compute $\boldsymbol{L}_{l_j}, \forall l_j$
    $\boldsymbol{U}_{l_j} \leftarrow \boldsymbol{L}_{l_j}^{-\dagger} v_{1:d}[\boldsymbol{L}_{l_j}^{-1} \boldsymbol{R}_{l_j} \boldsymbol{L}_{l_j}^{-\dagger}] \boldsymbol{\Sigma}_{l_j}, \quad \forall l_j$
  // backwark network optimization: transmit filter update
    Estimate $\bar{\boldsymbol{R}}_{i_k}, \bar{\boldsymbol{Q}}_{i_k}$, and compute $\boldsymbol{K}_{i_k}, \forall i_k$
    $\boldsymbol{V}_{i_k} \leftarrow \boldsymbol{K}_{i_k}^{-\dagger} v_{1:d}[\boldsymbol{K}_{i_k}^{-1} \bar{\boldsymbol{R}}_{i_k} \boldsymbol{K}_{i_k}^{-\dagger}] \boldsymbol{\Lambda}_{i_k}, \quad \forall i_k$
end for
___

for $n = 1, 2, 3 \dots$. As seen from (7), at each iteration n, given the fixed $\{\boldsymbol{V}_{l_j}^{(n)}\}$, problem (J1) is decouples in the receive filters $\{\boldsymbol{U}_{l_j}\}$, yielding

$$(J1) \begin{cases} \min\limits_{\boldsymbol{U}_{l_j}} \operatorname{tr}(\boldsymbol{U}_{l_j}^\dagger \boldsymbol{Q}_{l_j} \boldsymbol{U}_{l_j}) - \log_2 |\boldsymbol{I}_d + \boldsymbol{U}_{l_j}^\dagger \boldsymbol{R}_{l_j} \boldsymbol{U}_{l_j}| \\ \text{s. t. } \|\boldsymbol{U}_{l_j}\|_F^2 = P_r. \end{cases} \quad (11)$$

Likewise, if the receive filters are fixed, problem (J2) decouples, as seen from (8), in the transmit filters, resulting in

$$(J2) \begin{cases} \min\limits_{\boldsymbol{V}_{i_k}} \operatorname{tr}(\boldsymbol{V}_{i_k}^\dagger \bar{\boldsymbol{Q}}_{i_k} \boldsymbol{V}_{i_k}) - \log_2 |\boldsymbol{I}_d + \boldsymbol{V}_{i_k}^\dagger \bar{\boldsymbol{R}}_{i_k} \boldsymbol{V}_{i_k}| \\ \text{s. t. } \|\boldsymbol{V}_{i_k}\|_F^2 = P_t. \end{cases} \quad (12)$$

As aforementioned, the feasible sets of (11) and (12) are non-convex. Nevertheless, we show in the result below, that their globally optimal solutions can still be found.

Lemma 1. *Non-homogeneous Waterfilling.*
Consider the following problem,

$$\begin{cases} \min\limits_{\boldsymbol{X} \in \mathbb{C}^{n \times r}} f(\boldsymbol{X}) \triangleq \operatorname{tr}(\boldsymbol{X}^\dagger \boldsymbol{Q} \boldsymbol{X}) - \log_2 |\boldsymbol{I}_d + \boldsymbol{X}^\dagger \boldsymbol{R} \boldsymbol{X}| \\ \text{s. t. } \|\boldsymbol{X}\|_F^2 = \zeta. \end{cases} \quad (13)$$

where $\boldsymbol{Q} \succ \boldsymbol{0}$ *and* $\boldsymbol{R} \succeq \boldsymbol{0}$, $r < n$. *Let* $\boldsymbol{Q} \triangleq \boldsymbol{L}\boldsymbol{L}^\dagger$ *be the Cholesky factorization of* $\boldsymbol{Q}$, *and* $\boldsymbol{M} \triangleq \boldsymbol{L}^{-1} \boldsymbol{R} \boldsymbol{L}^{-\dagger}$, $\boldsymbol{M} \succeq \boldsymbol{0}$, *and define the following,* $\{\alpha_i \triangleq \lambda_i[\boldsymbol{M}]\}_{i=1}^r$, $\boldsymbol{\Psi} \triangleq v_{1:r}[\boldsymbol{M}]$, $\{\beta_i \triangleq \boldsymbol{\Psi}_{(i)}^\dagger (\boldsymbol{L}^\dagger \boldsymbol{L})^{-1} \boldsymbol{\Psi}_{(i)}\}_{i=1}^r$. *Then the optimal solution for* (13) *is,*

$$\boldsymbol{X}^\star = \boldsymbol{L}^{-\dagger} \boldsymbol{\Psi} \boldsymbol{\Sigma}^\star, \quad (14)$$

where $\boldsymbol{\Sigma}^\star$ *(diagonal) is the optimal power allocation. Moreover, optimal power allocation in* $\boldsymbol{\Sigma}^\star$ *is,*

$$\boldsymbol{\Sigma}_{(i,i)}^\star = \sqrt{\left(1/(1 + \mu^\star \beta_i) - 1/\alpha_i\right)^+}, \forall i, \quad (15)$$

where $\mu^\star$ *is the unique root to* $g(\mu) \triangleq \sum_{i=1}^r \beta_i \left(1/(1 + \mu\beta_i) - 1/\alpha_i\right)^+ - \zeta$, *on the interval* $[-1/(\max_i \beta_i), \infty]$, *and* $g(\mu)$ *is monotonically decreasing on that interval.*

Proof: Refer to [1][Appendix C] ∎

With Lemma 1, the optimal transmit and receive filter updates are formulated as below

$$\boldsymbol{U}_{l_j}^\star = \boldsymbol{L}_{l_j}^{-\dagger} \boldsymbol{\Psi}_{l_j} \boldsymbol{\Sigma}_{l_j}^\star, \quad \boldsymbol{\Psi}_{l_j} \triangleq v_{1:d}[\boldsymbol{L}_{l_j}^{-1} \boldsymbol{R}_{l_j} \boldsymbol{L}_{l_j}^{-\dagger}], \forall l_j,$$

$$\boldsymbol{V}_{i_k}^\star = \boldsymbol{K}_{i_k}^{-\dagger} \boldsymbol{\Theta}_{i_k} \boldsymbol{\Lambda}_{i_k}^\star, \quad \boldsymbol{\Theta}_{i_k} \triangleq v_{1:d}[\boldsymbol{K}_{i_k}^{-1} \bar{\boldsymbol{R}}_{i_k} \boldsymbol{K}_{i_k}^{-\dagger}], \forall i_k, \quad (16)$$

where $\boldsymbol{\Sigma}_{l_j}^\star$ and $\boldsymbol{\Lambda}_{i_k}^\star$ are the optimal power allocation, given in Lemma 1. Denoting by T the predefined number of F-B iterations, the max-DLT algorithm is in Algorithm 1.

Discussions

Based on the generalized eigenvalue analysis, we have that $\{\alpha_i \triangleq \lambda_i[\boldsymbol{L}^{-1}\boldsymbol{R}\boldsymbol{L}^{-\dagger}]\}_{i=1}^r$ are also the eigenvalues of $\boldsymbol{Q}^{-1}\boldsymbol{R}$. This means $\{\alpha_i\}$ can be viewed as a (quasi)-SINR measure of each data stream. The proposed method in (15) allocates no power to streams that have low-SINR, since $\boldsymbol{\Sigma}^\star_{(i,i)}$ tends to zero as $\alpha_i \to 0$. Moreover, as seen from (15), $\{\beta_i\}$ models the price of activating each of the streams, mimicking the original waterfilling principle. The difference however is that (14) fills the power level based on the SINR and cost for the stream activation, namely the non-homogeneous waterfilling solution. This readily enables the algorithm to not allocate power to some low SINR streams. Finally, since the global optimizer is found at each iteration in Algorithm 1, we can conclude that $R_\Sigma^{(LB)}(\{\boldsymbol{U}_{l_j}^{(n)}\}, \{\boldsymbol{V}_{l_j}^{(n)}\})$ in (P_{LB}) is monotonically decreasing with n, and converges to a stationary point of the DLT bound. While the 'stream-control' greatly speeds up the convergence, it evidently raises fairness issues, as some users/streams with low-SINR, may not get served. This can be remedied by introducing user weights in (P), with minor modifications in the problem/solutions.

IV. PRACTICAL ASPECTS

A. Comparisons

Our approach is applicable to other communication scenarios such as the MIMO Interfering Broadcast Channel (MIMO IBC), the MIMO Interference Channel (MIMO IFC). We benchmark our algorithms against widely adopted ones,

- o *max-SINR* [8] in the MIMO IMAC / MIMO IFC / MIMO IBC
- o *MMSE and Weighted-MMSE* [10], [18] in the MIMO IFC / MIMO IBC
- o *Uncoordinated* (Eigen-beamforming): each transmit (resp. receive) filter uses right (resp. let) singular eigenvectors of the desired channel

We also include relevant fast-converging algorithms,

- o *CCP-WMMSE* [14]: an accelerated version of WMMSE algorithm for the MIMO IMAC
- o *IWU* [16]: a fast-convergent leakage minimization algorithm for the MIMO IFC
- o *AIMS*: our previously proposed generalization of max-SINR [16], for MIMO IMAC / MIMO IFC / MIMO IBC

Algorithms such as IWU and CCP-WMMSE use so-called *turbo iterations*, where I inner-loop iterations are performed within each F-B iteration. Unlike IWU where the turbo iterations are done at the BS/user (i.e., offline), these iterations are carried over-the-air for CCP-WMMSE.

B. Communication Overhead

The operation of the proposed scheme hinges on each transmitter and receiver's having knowledge of effective channels, for the desired and interfering links. We note that investigating different mechanisms for the distributed acquisition of CSI is outside the scope of the current work (we refer the reader to [19]). However, we have outlined a simple mechanism that goes hand-in-hand with F-B iterations, in Fig. 2. We recall that T F-B iterations are carried out.
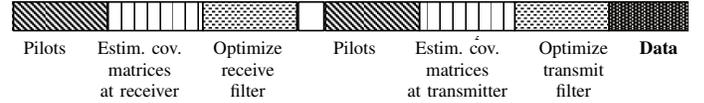

Fig. 2: Basic structure of Forward-Backward Iteration

It becomes clear that each F-B iteration has an associated *communication overhead*. While total overhead comprises of bidirectional transmission of pilots, synchronization, frequency offset calibration, etc, it is dominated by the pilot overhead, if the case of cellular coordination [20]. Thus, we can safely approximate the communication overhead by the total number of pilot symbols, for channel estimation, after T F-B iterations. In conventional coordination, it is typical to assume $T = 100 \sim 1000$ until convergence, even for small systems [12]. Moreover, this number increases with more BSs, cells and transmit/receive antennas, all of which are prevalent in sub-28 GHz systems. This limitation is compounded by the naturally lower coherence time of mmWave channels, thus further restricting the possible number of F-B iteration (before the channel changes). Indeed, simple calculations reveal that conventional algorithms would fail in these systems, as the overhead would destroy the sum-rate gains from coordination. Thus, we aggressively limit the number of F-B iteration to $T \leq 5$, thereby resulting in a drastic tenfold reduction in the communication overhead.

For simplicity, we additionally assume that the minimal number of orthogonal pilots is used, i.e. d pilot symbols for each UL/DL effective channel, resulting in a total of KLd orthogonal pilots for each UL/DL phase. The total overhead for max-DLT, in the number of channel uses (c.u.), is given by,

$$\Omega_{\text{prop}} = T(\underbrace{KLd}_{UL} + \underbrace{KLd}_{DL}) = 2TKLd$$

The overhead is the same for schemes such as max-SINR, IWU and MMSE. Similar calculations can be made to estimate the overhead of CCP-WMMSE and WMMSE (in c.u.),

$$\Omega_{\text{ccp-wmmse}} = T[(\underbrace{KLM}_{\substack{UL\ chann.\\estim}}) \times (L-1) + \underbrace{I}_{turbo} \times (\underbrace{KLN}_{\substack{cov.\ mat\\upd.}})]$$

$$\Omega_{\text{w-mmse}} = T(\underbrace{KLd}_{UL} + \underbrace{KLM}_{weights} + \underbrace{KLd}_{DL})$$

where I denotes the number of turbo iterations. These simple calculation reveal that the overhead for W-MMSE and CCP-WMMSE is significantly higher than that of max-DLT. Furthermore, the turbo iteration in CCP-WMMSE (outlined in Sec IV-A) is carried over-the-air, and thus induces a massively higher overhead, compared to other schemes. We include the overhead of these algorithms in the simulation results.

C. Complexity

We can approximate the computational complexity of max-DLT, by noticing that it is dominated by the complexity of the Cholesky Decomposition of the I+N covariance matrix, $\mathcal{O}(N^3)$, and that of Eigenvalue Decomposition of $\boldsymbol{M}$, $\mathcal{O}(M^3)$,
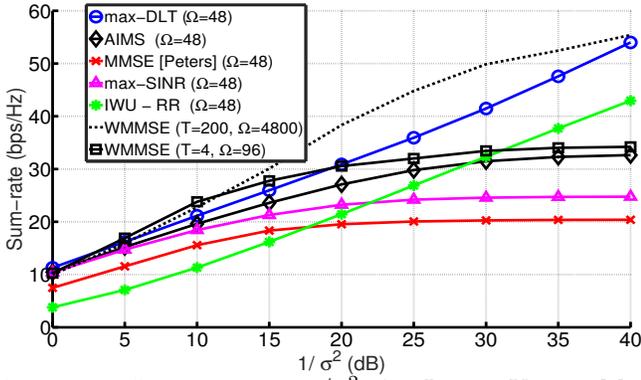
$$C_{\text{prop}} = \mathcal{O}((M+N)^3) \ .$$

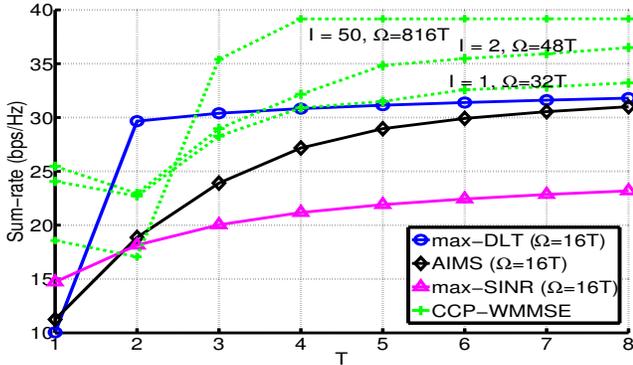Fig. 3: Ergodic sum-rate vs $1/\sigma^2$, for $L = 3, K = 1, M = N = 4, d = 2, T = 4$ (MIMO IFC)

Fig. 4: Ergodic sum-rate vs T, for $L = 2, K = 2, M = 4, N = 4, d = 2$

One can verify that the above also holds for max-SINR, IWU, MMSE, and WMMSE. Unlike other methods, the acceleration does not require gradient/Hessian, and thus comes at a negligible added computational cost, compared to conventional algorithms. However, each turbo iteration for CCP-WMMSE involves running a series of semidefinite programs (using interior point solvers), which render the algorithm very costly.

V. NUMERICAL RESULTS

A. Performance in Sub-6 GHz systems

We start with presenting results for conventional multi-cell multi-user MIMO, to illustrate desired features of max-DLT. We refer the reader to [1] for a detailed discussion of the simulation setup.

1) Single-user Multi-cell MIMO Uplink

We start with a widely used coordination test case, a MIMO IFC with $L = 3, K = 1, M = N = 4, d = 2$ where the set $T = 4$ for all algorithms. We include W-MMSE results for $T = 4$, and $T = 200$ F-B iterations (as an upper bound). Fig 3 reveals that while max-DLT and W-MMSE (with $T = 4$) have similar performance in the low- and medium-SNR range, this gap increases sharply as the SNR increases. This is in spite of two-fold increase in communication overhead for WMMSE. Moreover, the proposed scheme yields better sum-rate performance than all benchmarks, with this gap becoming significant in the high-SNR: as the following results will show, the gap increases further with more users, antennas, and BS/cells, under a low number of F-B iterations.
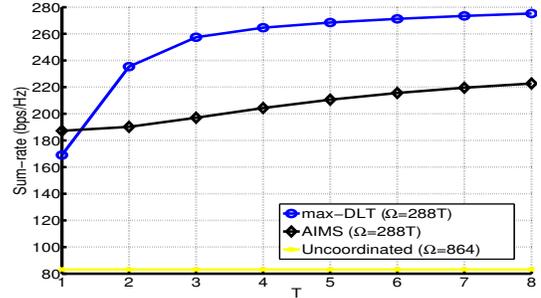

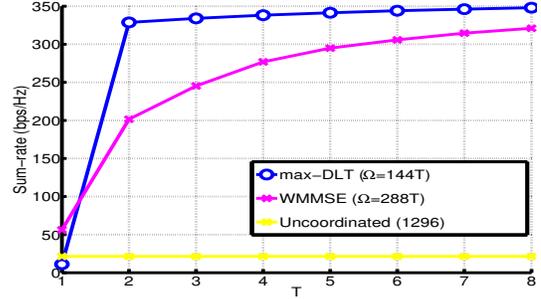Fig. 5: Ergodic sum-rate vs T, for dense uplink ($N = 8, M = 4, d = 2, L = 9, K = 8$).

Fig. 6: Ergodic sum-rate vs T for dense downlink ($M = 16, N = 4, d = 1, L = 9, K = 8$)

2) Multi-user Multi-cell MIMO uplink

Moving on to a larger setup with $L = 2, K = 2, M = 4, N = 4, d = 2$ (MIMO IMAC), we benchmark max-DLT against the fast-converging CCP-WMMSE (Sec. IV-A), by varying the number of turbo iterations I, for CCP-WMMSE. Fig. 4 clearly exhibits the fast converging nature of max-DLT, that achieves 95% of its final performance, after just 2 iterations. In the low overhead regime (for $T \le 5$), max-DLT outperforms CCP-WMMSE (for $I = 1$), although the overhead of the latter is twice that of former. While additional turbo iterations improve slightly the CCP-WMMSE performance, the overhead increases linearly with I, e.g., the overhead of CCP-WMMSE with $I = 2$ is threefold that of max-DLT. Achieving the nominal performance of CCP-WMMSE relies on convergence of the turbo iteration, which implies a (possibly arbitrary) large number of turbo iteration. This results in (potentially) orders-of-magnitude higher overhead/complexity (e.g. CCP-WMMSE with $I = 50$). Despite its fast-converging nature, CCP-WMMSE is clearly ill-suited for the systems studied here.

B. Performance in Dense mmWave Deployments

Following recent measurements in the 28 GHz band [21], we consider a dense urban mmWave setting. The full parametrization is detailed in [1][Sec. VI-C].

1) Dense Multi-user Multi-cell uplink

We consider a dense UL system with $L = 9, K = 8, N = 8, M = 4, d = 2$, where the average SNR (across users) is set to 19 dB. Fig. 5 reveals that max-DLT offers significantly better sum-rate, than all benchmarks. Interestingly, max-DLT (with $T = 3$) provides a threefold increase in sum-rate with respect to the uncoordinated scheme, however, with a similar overhead.

2) Dense Multi-user Multi-cell downlink

We next consider DL scenario with $L = 9, K = 8, M = 16, N = 4, d = 1$, while setting the average SNR to 21dB, and following the above simulation method. The fast-converging nature of max-DLT is embodied in Fig. 6, where most the performance is delivered in just 2 F-B iterations: this is due to inherent stream control feature, that allows poor quality streams to be shut down, thus converging quickly to a good sum-rate. Note that max-DLT assumes equal power allocation for users in each cell. In contrast, WMMSE performs power allocation for users in each cell, as part of the algorithm. Despite this unfavorable setup for max-DLT, we observe a large sum-rate gain compared to WMMSE, while resulting in a 50% decrease in overhead. Evidently, the sum-rate for WMMSE will exceed that of max-DLT, as T increases (with a huge overhead).

C. Discussions

We note the significant gap between the proposed scheme and the benchmarks, may be attributed to the fast-converging nature of the max-DLT, which is in turn due to the inherent stream-control mechanism of the non homogeneous waterfilling solution. Moreover, the drastically limited number of F-B iterations limits the performance of conventional algorithms, due to significant levels of residual interference. As seen in Figs. 5 and 6, that uncoordinated transmission performs extremely poorly: max-DLT provides a threefold sum-rate improvement over uncoordinated scheme, with a similar communication overhead. This provides a clear answer that low-overhead coordination is a crucial, to achieving huge sum-rate improvements in a dense multi-cell 28 GHz mmWave system. This also implies that the same conclusions hold for sub-28 GHz systems, which are naturally more sensitive to inference.

VI. CONCLUSIONS

We have proposed a low-overhead algorithm for coordination, in dense multi-cell sub-28 GHz systems. The DLT bound - a lower bound on the sum-rate, was derived and its tightness was investigated. Moreover, we have proposed a distributed optimization algorithm (max-DLT), and showed its convergence to a stationary point of the DLT bound. The non-homogeneous waterfilling was derived as a solution to the optimal BS/user filter update, and its ability to turn-off low-SINR streams was underlined. We have tied this to the fast-convergence of the algorithm, thus enabling a tenfold reduction in communication overhead (over conventional coordination). Our numerical results have showed that low-overhead coordination offers huge gains, in dense sub-28 GHz systems.

REFERENCES

[1] H. Ghauch, T. Kim, M. Bengtsson, and M. Skoglund, "Sum-rate maximization in sub-28-ghz millimeter-wave mimo interfering networks," *IEEE Journal on Selected Areas in Communications*, vol. 35, pp. 1649–1662, July 2017.

[2] J. G. Andrews, S. Buzzi, W. Choi, S. V. Hanly, A. Lozano, A. C. K. Soong, and J. C. Zhang, "What will 5g be?," *IEEE Journal on Selected Areas in Communications*, vol. 32, pp. 1065–1082, June 2014.

[3] M. K. Samimi and T. S. Rappaport, "Characterization of the 28 GHz millimeter-wave dense urban channel for future 5G mobile cellular," March 2014.

[4] S. Hur, T. Kim, D. Love, J. Krogmeier, T. Thomas, and A. Ghosh, "Millimeter wave beamforming for wireless backhaul and access in small cell networks," *IEEE Transactions on Communications,*, vol. 61, pp. 4391–4403, October 2013.

[5] E. J. Violette, R. H. Espeland, R. O. DeBolt, and F. K. Schwering, "Millimeter-wave propagation at street level in an urban environment," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 26, pp. 368–380, May 1988.

[6] H. Shokri-Ghadikolaei, C. Fischione, G. Fodor, P. Popovski, and M. Zorzi, "Millimeter wave cellular networks: A MAC layer perspective," *IEEE Transactions on Communications*, vol. 63, pp. 3437–3458, Oct 2015.

[7] METIS D6.2, "Initial report on horizontal topics, first results and 5G system concept," March 2014.

[8] K. Gomadam, V. R. Cadambe, and S. A. Jafar, "A distributed numerical approach to interference alignment and applications to wireless interference networks," *IEEE Transactions on Information Theory*, vol. 57, pp. 3309–3322, June 2011.

[9] D. Schmidt, C. Shi, R. Berry, M. Honig, and W. Utschick, "Minimum mean squared error interference alignment," in *2009 Conference Record of the Forty-Third Asilomar Conference on Signals, Systems and Computers*, pp. 1106 –1110, Nov. 2009.

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

[11] J. Kaleva, A. Tlli, and M. Juntti, "Weighted sum rate maximization for interfering broadcast channel via successive convex approximation," in *IEEE Global Communications Conference (GLOBECOM)*, pp. 3838–3843, Dec 2012.

[12] D. Schmidt, C. Shi, R. Berry, M. Honig, and W. Utschick, "Comparison of distributed beamforming algorithms for MIMO interference networks," *IEEE Transactions on Signal Processing*, vol. 61, pp. 3476–3489, July 2013.

[13] P. Komulainen, A. Tölli, and M. Juntti, "Effective CSI signaling and decentralized beam coordination in TDD multi-cell MIMO systems," *IEEE Transactions on Signal Processing*, vol. 61, pp. 2204–2218, May 2013.

[14] D. H. N. Nguyen and T. Le-Ngoc, "Sum-rate maximization in the multicell MIMO multiple-access channel with interference coordination," *IEEE Transactions on Wireless Communications*, vol. 13, pp. 36–48, January 2014.

[15] R. Brandt and M. Bengtsson, "Fast-convergent distributed coordinated precoding for TDD multicell MIMO systems," in *IEEE 6th International Workshop on Computational Advances in Multi-Sensor Adaptive Processing (CAMSAP)*, pp. 457–460, Dec 2015.

[16] H. Ghauch, T. Kim, M. Bengtsson, and M. Skoglund, "Distributed low-overhead schemes for multi-stream MIMO interference channels," *IEEE Transactions on Signal Processing*, vol. 63, pp. 1737–1749, April 2015.

[17] M. Razaviyayn, G. Lyubeznik, and Z.-Q. Luo, "On the degrees of freedom achievable through interference alignment in a MIMO interference channel," in *Signal Processing Advances in Wireless Communications (SPAWC), 2011 IEEE 12th International Workshop on*, pp. 511–515, June 2011.

[18] S. W. Peters and R. W. Heath, "Cooperative algorithms for MIMO interference channels," *IEEE Transactions on Vehicular Technology*, vol. 60, pp. 206–218, Jan. 2011.

[19] R. Brandt and M. Bengtsson, "Distributed CSI acquisition and coordinated precoding for TDD multicell MIMO systems," *IEEE Transactions on Vehicular Technology*, vol. PP, no. 99, pp. 1–1, 2015.

[20] O. El Ayach, A. Lozano, and R. Heath, "On the overhead of interference alignment: Training, feedback, and cooperation," *IEEE Transactions on Wireless Communications*, vol. 11, no. 11, pp. 4192–4203, 2012.

[21] G. R. MacCartney, M. K. Samimi, and T. S. Rappaport, "Omnidirectional path loss models in new york city at 28 GHz and 73 GHz," in *2014 IEEE 25th Annual International Symposium on Personal, Indoor, and Mobile Radio Communication (PIMRC)*, pp. 227–231, Sept 2014.