CELL-FREE MASSIVE MIMO: EXPLOITING THE WAX DECOMPOSITION

Juan Vidal Alegría[*], Jinliang Huang[†], Fredrik Rusek[*]

[*]Department of Electrical and Information Technology, Lund University, Lund, Sweden
[†]Huawei Technologies Sweden AB, R&D Center, Kista, Sweden
juan.vidal_alegria@eit.lth.se

ABSTRACT

Cell-free massive multiple-input multiple-output (MIMO) consists of a large set of distributed access points (APs) serving a number of users. The APs can be far from each other, and they can also have a big number of antennas. Thus, decentralized architectures have to be considered so as to reduce the interconnection bandwidth to a central processing unit (CPU) and make the system scalable. On the other hand, the APs in a heterogeneous network might have limited processing capabilities and fully-decentralized processing may not be available. In a recent paper, a trade-off between level of decentralization and decentralized processing complexity has been identified. Furthermore, a novel matrix decomposition – the WAX decomposition – which, if applicable to the channel matrix, allows for exploitation of said trade-off without loss of information. The results on WAX decomposition are only available for random channel matrices without specific structures, while in a cell-free massive MIMO scenario the channel can have sparse structures. In this work, we study the applicability of WAX decomposition to cell-free massive MIMO with its implications to the above-mentioned trade-off.

***Index Terms*—** WAX decomposition, cell-free massive MIMO, decentralized processing, linear equalization, distributed LIS.

1. INTRODUCTION

Cell-free massive multiple-input multiple-output (MIMO) [1–4] is currently gaining importance as an alternative to cell-based massive MIMO systems. The idea is to have a large number of distributed access points (APs) which jointly serve a number of user equipments (UEs) within the same time-frequency resource. This idea directly relates to the concept of distributed large intelligent surface (LIS), where the APs would correspond to LIS panels [5–7].

Cell-free massive MIMO scenarios depend on a great number of distributed APs. Thus, the availability of decentralized processing capabilities at the APs is crucial to reduce the interconnection bandwidth between the APs and a central processing unit (CPU). There is extensive research on decentralized approaches for large multi-antenna systems [6, 8–12]. In these approaches, part of the processing is performed at the antenna-end (or panel-end), so that a CPU does not need to gather all the channel information and perform all the processing tasks. However, in future heterogeneous networks, the processing capabilities of the cell-free APs might be limited due to the use of cheap equipment easily deployable and scalable.

In [13] a trade-off between level of decentralization (inputs to a CPU) and decentralized processing capabilities (multiplications per antenna) for information-lossless processing is identified, together with a framework that allows the exploitation of said trade-off. [13] also introduces the WAX decomposition, which is a novel matrix decomposition that, if applicable to the channel matrix, allows for information-lossless processing within the framework studied. However, the conditions for the applicability of WAX decomposition, which define the previously mentioned trade-off, are only given for a randomly chosen channel matrix with probability 1, leaving some gaps in the case of channel matrices with some specific structure.

The current work aims at closing some gaps from [13] by considering the application of WAX decomposition to cell-free massive MIMO channels, where the sparsity of the channel can be problematic. We have adapted the framework from [13] to a cell-free massive MIMO scenario, which consists of a set of APs with an arbitrary number of antennas serving several users. Our findings show that the sparsity of the cell-free massive MIMO channel can degrade the trade-off presented in [13] under some parameter configurations, but WAX decomposition can still offer a viable solution for applying information-lossless processing in a cell-free massive MIMO scenario with limited processing capabilities at the APs.

The rest of the paper is organized as follows. The system model is presented in Section 2 together with some background on the previous results on WAX decomposition. Section 3 presents new results on the applicability of WAX decomposition for cell-free massive MIMO scenarios. In Section 4, numerical results are given showing the implications of the new results on the WAX decomposition trade-off in cell-free massive MIMO. Section 5 concludes the paper with some final remarks.

Notation: In this paper, lowercase, bold lowercase and bold uppercase letters stand for scalars, column vectors and matrices, respectively. When using the mutual information operator, $I(.;.)$, bold uppercase in the sub-scripts refers to random vectors instead of their realizations. The operations $(\cdot)^{\mathrm{T}}$, $(\cdot)^*$ and $(\cdot)^{\mathrm{H}}$ denote transpose, conjugate, and conjugate transpose, respectively. The operation $\mathrm{diag}(\cdot)$ outputs a block diagonal matrix with the input matrices as the diagonal blocks. The operation $\|\cdot\|_1$ denotes L^1-norm. $\mathbf{0}_{i\times j}$ denotes the $i \times j$ all-zeros matrix. In this paper, a randomly chosen matrix corresponds to a realization of a random matrix whose elements are independently drawn from a continuous probability distribution function.

2. SYSTEM MODEL

Let us consider the uplink of a cell-free network of P APs serving K single-antenna users. Each AP is equipped with N antennas, such that the total number of antennas is $M = PN$. The aggregated $M \times 1$ received complex vector, $\boldsymbol{y}$, can be expressed as

$$\boldsymbol{y} = \boldsymbol{H}\boldsymbol{s} + \boldsymbol{n}, \tag{1}$$

where $\boldsymbol{H}$ is an $M \times K$ channel matrix, $\boldsymbol{s}$ is the $K \times 1$ vector of symbols transmitted by the users, and $\boldsymbol{n}$ is a zero-mean complex white Gaussian noise vector with sample variance N_0.

The current research has been carried out as part of an internship at Huawei Technologies AB, R&D Center, Kista, Sweden.

In cell-free massive MIMO scenarios the APs can be physically situated far from each other. Therefore, each AP might only see a small subset of the users, while each user can still be seen by several APs. The implication of such a scenario is thus that the matrix $\boldsymbol{H}$ would become a sparse matrix having 0s at the positions associated to hidden users. Furthermore, in a rich scattering environment the entries associated to non-hidden users can be assumed to be IID complex Gaussian (Rayleigh fading).

Let us consider the framework presented in Fig. 1 so that each AP only needs to apply $L \times L$ filters to each of the N_L groups of L antennas, i.e., AP i applies $\boldsymbol{W}_{ij}^{\mathrm{H}}$ to the jth L-group. Note that $N_{\mathrm{L}} = N/L$ should evaluate to an integer. We can thus identify L as the number of multiplications per antenna, which corresponds to the trade-off parameter associated to decentralized processing capabilities. The outputs from the APs are then combined through a $T \times M$ fixed combining module, $\boldsymbol{A}^H$, which reduces the dimensions of the data to T. The resulting T entries would be the inputs to the CPU, which corresponds to the trade-off parameter associated to level of decentralization. An equivalent framework would be to consider that each antenna in an AP multiplies its received signals by L numbers and the corresponding sums would be applied in the combining module; however, this framework would scale by L the number of outputs per AP.

$$\boldsymbol{z} = \boldsymbol{X}^{\mathrm{H}}\boldsymbol{A}^{\mathrm{H}}\boldsymbol{W}^{\mathrm{H}}\boldsymbol{y}, \tag{2}$$

where $\boldsymbol{W}$ is an $M \times M$ block diagonal matrix of the form

$$\boldsymbol{W} = \mathrm{diag}\left(\boldsymbol{W}_{11}, \ldots, \boldsymbol{W}_{1N_{\mathrm{L}}}, \boldsymbol{W}_{21}, \ldots, \boldsymbol{W}_{PN_{\mathrm{L}}}\right). \tag{3}$$

The matrices $\boldsymbol{W}$ and $\boldsymbol{X}$ can be recalculated for every channel realization, while the matrix $\boldsymbol{A}$ remains unchanged once the system is deployed (think of it as a fixed hardware combining module). The framework under study is represented in Fig. 1.

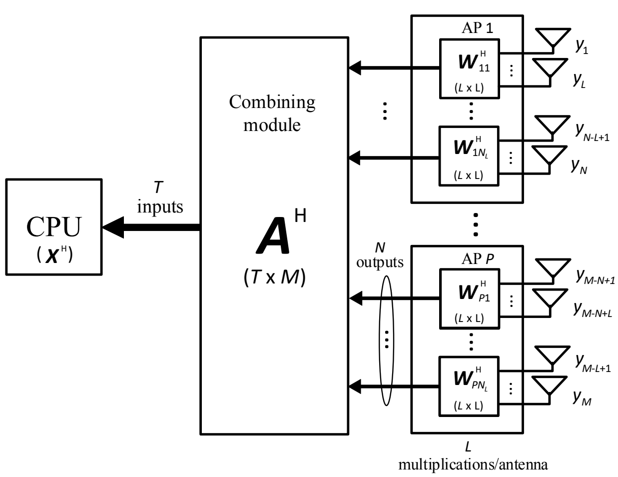

Fig. 1: Framework used in this paper during the data phase.

The framework under study allows for the exploitation of the trade-off between the number of multiplications per antenna, associated to L, and the number of inputs to the CPU, associated to T. Said trade-off, which can be seen as a trade-off between decentralized processing complexity and level of decentralization, was identified in [13], which considers an equivalent version of the framework under study.

We constrain the information rate at which the users can transmit to the BS, i.e., $I_{\boldsymbol{Z},\boldsymbol{S}}(\boldsymbol{z}, \boldsymbol{s})$, or, correspondingly[1], $I_{\boldsymbol{Y},\boldsymbol{S}}(\boldsymbol{A}^{\mathrm{H}}\boldsymbol{W}^{\mathrm{H}}\boldsymbol{y}, \boldsymbol{s})$ to be lossless. Thus, $I_{\boldsymbol{Y},\boldsymbol{S}}(\boldsymbol{A}^{\mathrm{H}}\boldsymbol{W}^{\mathrm{H}}\boldsymbol{y}, \boldsymbol{s}) = I_{\boldsymbol{Y},\boldsymbol{S}}(\boldsymbol{y}, \boldsymbol{s})$.

2.1. Background

As we mentioned earlier, the architecture considered in this paper is based on the framework presented in [13], where important results are shown that will be the required for our analysis. From [13, Lemma 1], the framework under study can achieve information lossless processing if and only if we can decompose the channel matrix $\boldsymbol{H}$ into the so called WAX decomposition

$$\boldsymbol{H} = \boldsymbol{W}\boldsymbol{A}\boldsymbol{X}, \tag{4}$$

where $\boldsymbol{W}$, $\boldsymbol{A}$ and $\boldsymbol{X}$ correspond to the matrices from (2), i.e., $\boldsymbol{A}$ is fixed by design while $\boldsymbol{W}$ and $\boldsymbol{X}$ can be tuned to $\boldsymbol{H}$. Note that selecting $\boldsymbol{W}$ and $\boldsymbol{X}$ in (2) such that (4) is fulfilled leads to information-lossless processing within our framework. The main result of the applicability of WAX decomposition is given in [13, Theorem 1], which states that, for a randomly chosen $\boldsymbol{A} \in \mathbb{C}^{M\times T}$, a randomly chosen $\boldsymbol{H} \in \mathbb{C}^{M\times K}$ admits WAX decomposition with probability 1 if

$$T > \max\left(M\frac{K-L}{K}, K-1\right). \tag{5}$$

3. WAX DECOMPOSITION FOR SPARSE CHANNELS

Much of the WAX theory, including methods for practical computation of the constituent matrices $\boldsymbol{W}$ and $\boldsymbol{X}$, is underpinned by [13, Lemma 3]. However, for sparse channels, an important assumption of [13, Lemma 3] is not fulfilled. Let $\boldsymbol{H} = [\boldsymbol{H}_{11}^{\mathrm{T}}\ \boldsymbol{H}_{12}^{\mathrm{T}}\ \ldots\ \boldsymbol{H}_{PN_{\mathrm{L}}}^{\mathrm{T}}]^{\mathrm{T}}$, where we used notation from (3), and $\boldsymbol{H}_{pn}$ is an $L \times K$ submatrix of $\boldsymbol{H}$. For [13, Lemma 3] to hold it should be assumed that $\mathrm{rank}(\boldsymbol{H}_{pn}) = L$, $\forall p, n$. However, in the current scenario, it could very well happen that a certain panel sees far less UEs than L, so $\mathrm{rank}(\boldsymbol{H}_{pn}) = L$ cannot be guaranteed. Wherefore, [13, Lemma 3] needs to be generalized, which we do next.

Lemma 1 ([13, Lemma 3] for sparse channels) *For all matrices $\boldsymbol{H}$ satisfying $\mathrm{rank}(\boldsymbol{H}_{pn}) = \min(k_{pn}, L)$, where k_{pn} can be seen as the number of non-zero columns in $\boldsymbol{H}_{pn}$, there exists a block diagonal matrix $\boldsymbol{W}$ and a matrix $\boldsymbol{X}$ such that $\boldsymbol{W}\boldsymbol{A}\boldsymbol{X} = \boldsymbol{H}$, if and only if, there exists a block diagonal invertible matrix $\hat{\boldsymbol{W}}$ with the same form as $\boldsymbol{W}$ such that $\boldsymbol{A}\boldsymbol{X} - \hat{\boldsymbol{W}}\boldsymbol{H} = \boldsymbol{0}_{M\times K}$.*

Proof The *if* part is trivial. If an invertible matrix $\hat{\boldsymbol{W}}$ exists, then we can set $\boldsymbol{W} = \hat{\boldsymbol{W}}^{-1}$. To prove the *only if* part, assume that $\boldsymbol{W}\boldsymbol{A}\boldsymbol{X} = \boldsymbol{H}$. Specifically,

$$\boldsymbol{W}_{pn}\boldsymbol{A}_{pn}\boldsymbol{X} = \boldsymbol{H}_{pn},\ \forall p, n.$$

Let us now focus on indices p, n where $\mathrm{rank}(\boldsymbol{H}_{pn}) = k_{pn} < L$, as rank L blocks are treated in [13, Lemma 3]. Let $\boldsymbol{X}_{pn}$ be the columns of $\boldsymbol{X}$ corresponding to the k_{pn} nonzero columns of $\boldsymbol{H}_{pn}$. Then, $\boldsymbol{A}_{pn}\boldsymbol{X}_{pn}$ is an $L \times k_{pn}$ matrix whose rank must be k_{pn}. Moreover, $\mathrm{rank}(\boldsymbol{W}_{pn}) \geq k_{pn}$. Being a full rank matrix, the left nullspace of $\boldsymbol{A}_{pn}\boldsymbol{X}_{pn}$ has dimensionality of precisely $L - k_{pn}$, i.e., there exists a full rank $(L - k_{pn}) \times L$ matrix $\boldsymbol{N}_{pn}$ such that $\boldsymbol{N}_{pn}\boldsymbol{A}_{pn}\boldsymbol{X}_{pn} = \boldsymbol{0}_{(L-k_{pn})\times k_{pn}}$. Decompose $\boldsymbol{W}_{pn}$ as $\boldsymbol{W}_{pn} = \tilde{\boldsymbol{W}}_{pn} + \boldsymbol{Z}_{pn}\boldsymbol{N}_{pn}$,

[1]Note that $\boldsymbol{X}$ cannot possibly increase the maximum information rate at which the users can transmit (recall data-processing inequality [14]).

where $\tilde{\boldsymbol{W}}_{pn}\boldsymbol{N}_{pn} = \boldsymbol{0}_{L\times k_{pn}}$ and $\boldsymbol{Z}_{pn}$ is $L \times (L - k_{pn})$. We have,

$$\begin{aligned}\boldsymbol{W}_{pn}\boldsymbol{A}_{pn}\boldsymbol{X} &= (\tilde{\boldsymbol{W}}_{pn} + \boldsymbol{Z}_{pn}\boldsymbol{N}_{pn})\boldsymbol{A}_{pn}\boldsymbol{X} \\ &= \tilde{\boldsymbol{W}}_{pn}\boldsymbol{A}_{pn}\boldsymbol{X} \\ &= \boldsymbol{H}_{pn}.\end{aligned}$$

Therefore, it follows that $\operatorname{rank}(\tilde{\boldsymbol{W}}_{pn}) = k_{pn}$, but $\boldsymbol{W}_{pn}$ is constructed from $\tilde{\boldsymbol{W}}_{pn}$ by adding a rank $L - k_{pn}$ matrix that is orthogonal to $\tilde{\boldsymbol{W}}_{pn}$, which directly implies that there exists a full rank solution $\boldsymbol{W}_{pn}$ whenever a solution exists. Specifically, this implies that there exists a full rank $\boldsymbol{W}$ that solves $\boldsymbol{W}\boldsymbol{A}\boldsymbol{X} = \boldsymbol{H}$ whenever a solution exists. But this implies that we can set $\hat{\boldsymbol{W}} = \boldsymbol{W}^{-1}$, which concludes the proof. □

3.1. 2 APs scenario

In order to get an understanding of how to deal with WAX decomposition in the case of sparse channel matrices, let us first take a look at a simplified scenario where there are only 2 APs. Assume that K_1 users are only seen by AP 1, K_2 users are only seen by AP 2s, and K_3 users are seen by both APs. Without any loss of generality, we can express the channel matrix for this scenario as

$$\boldsymbol{H} = \begin{pmatrix} \boldsymbol{H}_{11} & \boldsymbol{0}_{N\times K_2} & \boldsymbol{H}_{13} \\ \boldsymbol{0}_{N\times K_1} & \boldsymbol{H}_{22} & \boldsymbol{H}_{23} \end{pmatrix}, \tag{6}$$

where $\boldsymbol{H}_{13}$ and $\boldsymbol{H}_{23}$ correspond to the $N \times K_3$ channel matrices from the shared users to AP 1 and AP 2, respectively, while $\boldsymbol{H}_{11}$ and $\boldsymbol{H}_{22}$ correspond to the $N \times K_1$ and $N \times K_2$ channel matrices from the non-shared users to AP 1 and AP 2, respectively. Note that any permutation of the columns of $\boldsymbol{H}$ only corresponds to a re-indexing of the user IDs[2]. Incidentally, the structure of $\boldsymbol{H}$ corresponds to the structure of the matrices being decomposed in the WAX modules from the binary tree architecture presented in [7], which only considered (4) as a first limit to the dimension reduction. Thus, the results we obtain in this subsection have direct implications in the dimension reduction that can be achieved by such architecture.

Let us express the WAX decomposition (4) of (6) as

$$\boldsymbol{H} = \begin{pmatrix} \boldsymbol{W}_1 & \boldsymbol{0}_{N\times N} \\ \boldsymbol{0}_{N\times N} & \boldsymbol{W}_2 \end{pmatrix} \begin{pmatrix} \boldsymbol{A}_1 \\ \boldsymbol{A}_2 \end{pmatrix} \begin{pmatrix} \boldsymbol{X}_1 & \boldsymbol{X}_2 & \boldsymbol{X}_3 \end{pmatrix}, \tag{7}$$

where $\boldsymbol{W}_1$ and $\boldsymbol{W}_2$ are the two $N \times N$ blocks of $\boldsymbol{W}$, which are also formed by $L \times L$ diagonal blocks, $\boldsymbol{A}_1$ and $\boldsymbol{A}_2$ are the two $N \times T$ row blocks of $\boldsymbol{A}$, and $\boldsymbol{X}_1$, $\boldsymbol{X}_2$, and $\boldsymbol{X}_3$ are the $T \times K_1$, $T \times K_2$, and K_3 column blocks of $\boldsymbol{X}$. Assume that the non-zero blocks of $\boldsymbol{H}$ are randomly chosen matrices (e.g., IID complex Gaussian), and so is the $\boldsymbol{A}$ matrix from the WAX decomposition.

From Lemma 1 we assure that the existence of solution to $\boldsymbol{W}\boldsymbol{A}\boldsymbol{X} = \boldsymbol{H}$ is equivalent to the existence of solution to $\boldsymbol{A}\boldsymbol{X} = \boldsymbol{W}^{-1}\boldsymbol{H}$ with full-rank $\boldsymbol{W}^{-1}$. Let us thus focus on the equivalent problem

$$\boldsymbol{A}\boldsymbol{X} = \boldsymbol{W}^{-1}\boldsymbol{H}, \tag{8}$$

for full-rank $\boldsymbol{W}^{-1}$. If (5) is fulfilled, (8) will have a non-zero solution for $\boldsymbol{X}$ and $\boldsymbol{W}^{-1}$, but it remains to assure full-rank $\boldsymbol{W}^{-1}$. Considering the structure of $\boldsymbol{H}$, we can translate (7) into the following set of equations

$$\boldsymbol{A}_1\boldsymbol{X}_1 = \boldsymbol{W}_1^{-1}\boldsymbol{H}_{11}, \tag{9a}$$

$$\boldsymbol{A}_1\boldsymbol{X}_2 = \boldsymbol{0}_{N\times K_2}, \tag{9b}$$

$$\boldsymbol{A}_1\boldsymbol{X}_3 = \boldsymbol{W}_1^{-1}\boldsymbol{H}_{13}, \tag{9c}$$

$$\boldsymbol{A}_2\boldsymbol{X}_1 = \boldsymbol{0}_{N\times K_1}, \tag{9d}$$

$$\boldsymbol{A}_2\boldsymbol{X}_2 = \boldsymbol{W}_2^{-1}\boldsymbol{H}_{22}, \tag{9e}$$

$$\boldsymbol{A}_2\boldsymbol{X}_3 = \boldsymbol{W}_2^{-1}\boldsymbol{H}_{23}, \tag{9f}$$

where $\boldsymbol{W}_1^{-1}$ and $\boldsymbol{W}_2^{-1}$ are now restricted to be full-rank matrices. In order for (9b) and (9d) to hold $\boldsymbol{X}_1$ and $\boldsymbol{X}_2$ must be in the null-space of $\boldsymbol{A}_2$ and $\boldsymbol{A}_1$, respectively. Since $\boldsymbol{A}$ is randomly chosen, $\boldsymbol{A}_1$ and $\boldsymbol{A}_2$ will be full rank with probability 1. We thus have

$$\boldsymbol{X}_1 = \boldsymbol{N}_2\widetilde{\boldsymbol{X}}_1, \quad \boldsymbol{N}_2 = \mathcal{N}\{\boldsymbol{A}_2\}, \tag{10a}$$

$$\boldsymbol{X}_2 = \boldsymbol{N}_1\widetilde{\boldsymbol{X}}_2, \quad \boldsymbol{N}_1 = \mathcal{N}\{\boldsymbol{A}_1\}, \tag{10b}$$

where $\boldsymbol{N}_1$ and $\boldsymbol{N}_2$ are the $T \times (T - N)$ matrices conforming the null-spaces of $\boldsymbol{A}_1$ and $\boldsymbol{A}_2$, respectively. We can now rewrite the set of equations (9) as

$$\boldsymbol{A}_1\boldsymbol{N}_2\widetilde{\boldsymbol{X}}_1 = \boldsymbol{W}_1^{-1}\boldsymbol{H}_{11}, \tag{11a}$$

$$\boldsymbol{A}_1\boldsymbol{X}_3 = \boldsymbol{W}_1^{-1}\boldsymbol{H}_{13}, \tag{11b}$$

$$\boldsymbol{A}_2\boldsymbol{N}_1\widetilde{\boldsymbol{X}}_2 = \boldsymbol{W}_2^{-1}\boldsymbol{H}_{22}. \tag{11c}$$

$$\boldsymbol{A}_2\boldsymbol{X}_3 = \boldsymbol{W}_2^{-1}\boldsymbol{H}_{23}, \tag{11d}$$

Let us rewrite (11a) and (11c) as

$$\widetilde{\boldsymbol{A}}_1\widetilde{\boldsymbol{X}}_1 = \boldsymbol{W}_1^{-1}\boldsymbol{H}_{11}, \tag{12a}$$

$$\widetilde{\boldsymbol{A}}_2\widetilde{\boldsymbol{X}}_2 = \boldsymbol{W}_2^{-1}\boldsymbol{H}_{22}, \tag{12b}$$

where $\widetilde{\boldsymbol{A}}_1$ and $\widetilde{\boldsymbol{A}}_2$ can be seen as two $N \times (T - N)$-sized randomly chosen matrices, since they are computed as a product of a randomly chosen matrix with the null-space of another randomly chosen matrix. Equations (12a) and (12b) can be transformed into linear equations by vectorizing. Thus, in order have non-zero solution the following conditions should be fulfilled

$$(T - N) > \max\left(N\frac{K_1 - L}{K_1}, K_1 - 1\right), \tag{13a}$$

$$(T - N) > \max\left(N\frac{K_2 - L}{K_2}, K_2 - 1\right). \tag{13b}$$

Furthermore, the set of $\boldsymbol{W}_1^{-1}$ and $\boldsymbol{W}_2^{-1}$ matrices solving (12a) and (12b) will be full-rank except for a subset of measure 0.[3]On the other hand, if (5) is fulfilled, (11b) and (11d) will have at least one solution randomly situated in the continuous set of $\boldsymbol{W}_1^{-1}$ and $\boldsymbol{W}_2^{-1}$ that solve (12a) and (12b).[3] Therefore, a matrix $\boldsymbol{H}$ selected as in (6) admits WAX decomposition with probability 1 for randomly chosen $\boldsymbol{A}$ if and only if the conditions (5), (13a), and (13b) are jointly fulfilled.

3.2. Extension to any number of APs

Let us consider now the general case with P APs. The channel matrix for this scenario can be now expressed as

$$\boldsymbol{H} = \begin{bmatrix} b_{11}\boldsymbol{H}_{11} & b_{12}\boldsymbol{H}_{12} & \cdots & b_{1C}\boldsymbol{H}_{1C} \\ b_{21}\boldsymbol{H}_{21} & b_{22}\boldsymbol{H}_{22} & \cdots & b_{2C}\boldsymbol{H}_{1C} \\ \vdots & \vdots & \ddots & \vdots \\ b_{P1}\boldsymbol{H}_{P1} & b_{P2}\boldsymbol{H}_{P2} & \cdots & b_{PC}\boldsymbol{H}_{PC} \end{bmatrix}, \tag{14}$$

where $C = 2^P - 1$, $\boldsymbol{b}_j = (b_{1j}b_{2j}\cdots b_{Pj})_2$ is the binary expansion of j, and $\boldsymbol{H}_{ij}$ are randomly chosen matrices (e.g., IID complex Gaussian entries) of dimensions $N \times K_j$.[4] It can thus be noted that selecting $\boldsymbol{H}$ as (14) allows us to consider any possible channel matrix within the presented cell-free scenario ($P = 2$ leads to (6)).

[2] Such a permutation can be applied by multiplying $\boldsymbol{H}$ from the right with an invertible matrix, which, in fact, could be absorbed by $\boldsymbol{X}$ in its WAX decomposition

[3] For further understanding see [13, Proof of Theorem 1]

[4] Note that the K_j values are also allowed to be 0.

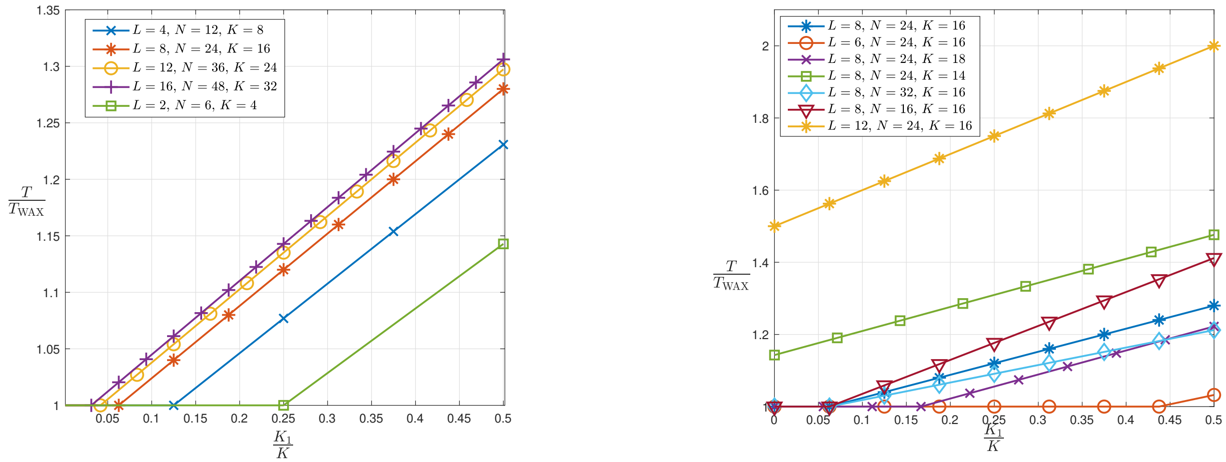

Fig. 2: APs scenario with $K_2 = K_1$. Plot of T/T_{WAX} with respect to K/K_1.

Proposition 1 *A matrix* $\boldsymbol{H}$ *selected as* (14) *admits WAX decomposition with probability 1 for randomly chosen* $\boldsymbol{A}$ *if and only if the set of conditions*

$$T - (P - \|\boldsymbol{b}_j\|_1)\, N > \max\left(\|\boldsymbol{b}_j\|_1 N \frac{K_j - L}{K_j}, K_j - 1 \right), \forall j, \tag{15}$$

together with (5) *are jointly fulfilled.*

Proof *The proof follows from direct extension of the arguments used in the 2 APs scenario. We should note that the* j*th column block of* $\boldsymbol{H}$ *has* $\|\boldsymbol{b}_j\|_1$ *randomly chosen blocks of size* $N \times K_j$*, and* $P - \|\boldsymbol{b}_j\|_1$ *blocks with* $\mathbf{0}_{N \times K_j}$. □

The previous proposition states the conditions of existence of WAX decomposition for cell-free massive MIMO channels within the considered framework. This poses an important advancement in the definition and exploitation of the trade-off between processing complexity at the APs and level of decentralization of the system.

4. NUMERICAL RESULTS

We next present some results on the impact of cell-free massive MIMO sparse channels on the WAX decomposition trade-off. We define

$$T_{\text{WAX}} = \max\left(\left\lfloor M \frac{K - L}{K} + 1 \right\rfloor, K \right),$$

which corresponds to the minimum T that can be used to fulfill (5), i.e., the trade-off for non-sparse channels.

Fig. 2 shows how T is degraded with respect to T_{WAX} in the 2 APs case as we increase $K_1 = K_2$, i.e. the number users seen by only one of the APs. From Fig. 2 (left) we can see that increasing L, N and K at the same rate degrades the trade-off by shifting the curves towards the left (faster sensitivity to unseen users), but it seems to converge. Fig. 2 (right) indicates that decreasing the total number of users K also degrades the trade-off by shifting the curves to the left, decreasing the antennas per panel N degrades the trade-off only by increasing the slope, while increasing the multiplications per antenna L degrades trade-off by both shifting to the left and increasing the slope.

In Fig. 3 we can see how the sparsity degrades the trade-off in a more general scenario. For the simulation we have drawn the number of panels seen by each user, n, from a discrete distribution with $p_n = a/n^q$, where a is a constant for normalization purposes, and the n APs are randomly selected with equal probabilities (i.e., select uniformly $\boldsymbol{b}_j$ such that $\|\boldsymbol{b}_j\| = n$). Thus, variable q can be related to the sparsity level since a bigger value of it means that each user will see less panels with more probability. From the plot, we can also see that increasing the total number of users K decreases the impact of sparsity on the trade-off. Increasing the number of panels further, degrades the trade-off since it increases the inherent sparsity. We should note that having $q = 0$ still incurs a degradation of the trade-off since in this case a row of $\boldsymbol{H}$ would have any number of zero blocks (between 1 and P) with equal probability, i.e., $q = 0$ is not equivalent to having no sparsity.

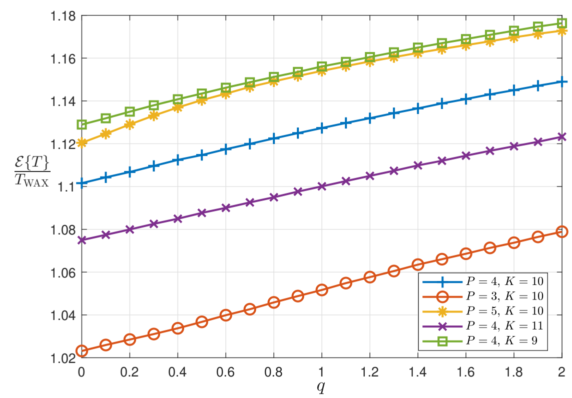

Fig. 3: Average T/T_{WAX} with respect to sparsity factor q. $L = 2$ and $N = 16$.

5. CONCLUSIONS

We have extended the application of the WAX decomposition to a cell-free massive MIMO scenario with sparse channel matrix. We have found a new set of conditions that limit the information-lossless trade-off between multiplications per antenna and number of input to a CPU in the case of channel sparsity. Our numerical results confirm that the trade-off presented in [13] is degraded as the sparsity of the channel increases.

6. REFERENCES

[1] H. Q. Ngo, A. Ashikhmin, H. Yang, E. G. Larsson, and T. L. Marzetta, "Cell-free massive MIMO versus small cells," *IEEE Transactions on Wireless Communications*, vol. 16, no. 3, pp. 1834–1850, 2017.

[2] J. Zhang, S. Chen, Y. Lin, J. Zheng, B. Ai, and L. Hanzo, "Cell-free massive MIMO: A new next-generation paradigm," *IEEE Access*, vol. 7, pp. 99 878–99 888, 2019.

[3] S. Buzzi and C. D'Andrea, "Cell-free massive MIMO: User-centric approach," *IEEE Wireless Communications Letters*, vol. 6, no. 6, pp. 706–709, 2017.

[4] E. Björnson and L. Sanguinetti, "Scalable cell-free massive MIMO systems," *IEEE Transactions on Communications*, vol. 68, no. 7, pp. 4247–4261, 2020.

[5] S. Hu, K. Chitti, F. Rusek, and O. Edfors, "User assignment with distributed large intelligent surface (lis) systems," in *2018 IEEE 29th Annual International Symposium on Personal, Indoor and Mobile Radio Communications (PIMRC)*, 2018, pp. 1–6.

[6] J. Vidal Alegría, J. R. Sánchez, F. Rusek, L. Liu, and O. Edfors, "Decentralized equalizer construction for large intelligent surfaces," in *2019 IEEE 90th Vehicular Technology Conference (VTC2019-Fall)*, Sep. 2019, pp. 1–6.

[7] J. V. Alegría, F. Rusek, J. R. Sánchez, and O. Edfors, "Modular binary tree architecture for distributed large intelligent surface," in *ICASSP 2021 - 2021 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, 2021, pp. 4565–4569.

[8] K. Li, R. R. Sharan, Y. Chen, T. Goldstein, J. R. Cavallaro, and C. Studer, "Decentralized baseband processing for massive MU-MIMO systems," *IEEE Journal on Emerging and Selected Topics in Circuits and Systems*, vol. 7, no. 4, pp. 491–507, Dec 2017.

[9] E. Bertilsson, O. Gustafsson, and E. G. Larsson, "A scalable architecture for massive MIMO base stations using distributed processing," in *2016 50th Asilomar Conference on Signals, Systems and Computers*, Nov 2016, pp. 864–868.

[10] J. R. Sánchez, J. Vidal Alegría, and F. Rusek, "Decentralized massive MIMO systems: Is there anything to be discussed?" in *2019 IEEE International Symposium on Information Theory (ISIT)*, July 2019, pp. 787–791.

[11] J. R. Sanchez, F. Rusek, M. Sarajlic, O. Edfors, and L. Liu, "Fully decentralized massive MIMO detection based on recursive methods," in *2018 IEEE International Workshop on Signal Processing Systems (SiPS)*, Oct 2018, pp. 53–58.

[12] A. Amiri, S. Rezaie, C. N. Manchon, and E. de Carvalho, "Distributed receivers for extra-large scale MIMO arrays: A message passing approach," 2020.

[13] J. V. Alegría, F. Rusek, and O. Edfors, "Trade-offs in decentralized multi-antenna architectures: The wax decomposition," *IEEE Transactions on Signal Processing*, vol. 69, pp. 3627–3641, 2021.

[14] T. M. Cover and J. A. Thomas, *Elements of Information Theory (Wiley Series in Telecommunications and Signal Processing)*. USA: Wiley-Interscience, 2006.