Low Complexity Moore-Penrose Inverse for Large CoMP Areas with Sparse Massive MIMO Channel Matrices

Amir M. Ahmadian, Wolfgang Zirwas, Rakash Sivasiva Ganesan and Berthold Panzner

Nokia Bell Labs, Radio Systems Research, Munich, Germany

E-mails: {amir.ahmadian_tehrani.ext; wolfgang.zirwas; rakash.sivasivaganesan; berthold.panzner}@nokia-bell-labs.com

***Abstract*—Joint transmission coordinated multipoint (JT CoMP) has been identified as a potential differentiator for future 5G radio systems due to its superior interference mitigation capabilities. Further, the combination with massive multiple-input-multiple-output (mMIMO) when a set of fixed grid of beams is used at eNodeB results in sparse overall channel matrices with a relatively low number of relevant channel components, which reduces the feedback overhead for reporting of channel state information (CSI). Although JT CoMP faces several challenges from synchronization to CSI outdating, in this paper the focus will be on complexity reduction of the precoding over large clustered cells utilizing massive MIMO. It will be derived how the typical sparse number of relevant channel components per user equipment can be exploited to reduce the number of floating point operations (FLOPs) by a factor of ten compared to state-of-the-art solutions for the calculation of the Moore Penrose pseudo inverse of the channel matrix.**

***Index Terms*—Moore-Penrose inverse, FLOPs, massive MIMO**

I. Introduction

Currently EU funded 5GPPP projects like Fantastic5G [1] lay the foundation for a future 5G mobile radio access system. One focus is on the frequency range below 6GHz and the options to maximize its spectral efficiency, capacity and coverage. From previous projects like METIS [2] it is known that joint transmission coordinated multipoint (JT CoMP) over adjacent sites in combination with massive multiple-input-multiple-output (mMIMO) have the potential to combat interference and support large multi-user multiple-input-multiple-output (MU-MIMO) gains. Furthermore, network clustering which turns a potentially interference free system in case of network wide cooperation into an interference limited one due to the inter cluster interference, plays a vital rule as an essential part of each CoMP scheme. This clustered network in CoMP literatures is referred to as cooporation area (CA). Since coordination in the spatial domain relies on the use of suitable multiple antenna combining, there has been few techniques for advanced beamforming. Reducing the complexity of massive MIMO antenna array in such architecture, the massive number of antennas can be restricted to a set of essential ones through generating a narrowed and fixed grid of beams (GoBs) [3] at the evolved node B (eNodeB). Assuming a fixed GoB concept with eight beams in azimuth and two elevations in direction as well as two polarizations, there will be 32 beams per cell. Thus, allocating a group of nine clustered small cells or three adjacent sites, the total number of beams will equal to 288 . Further, the current understanding is that massive MIMO gains for below 6GHZ will rely mainly on MU-MIMO like spatial multiplexing of ten or more users per cell [4], [2]. As such, there will be in total 90 user equipment (UEs) distributed in a three adjacent site cooperation area. Out of these, the overall channel matrix will be in order of size 90×288.

Zero forcing which has been identified as common method used for precoding process, requires accurate calculation of the Moore Penrose pseudo inverse. The processing complexity in number of FLOPs can be expressed as $2np^2 + 2n^3$ for a matrix size of $n \times p$ applying the common Singular Value Decomposition (SVD) method. Thus, the number of FLOPs for inverse of a 90×288 matrix will be in order of 10^7. Although the processing complexity for such matrix size can be reduced to the order of 10^6 in number of FLOPs utilizing the state-of-the-art algorithms [5], in case the transmission time interval (TTI) length is 1ms and the frequency band is sub divided into 100 physical resource blocks (PRBs), this will end up to $10^3 \times 10^2 \times 10^6 = 10^{11}$ FLOPs per second only for the precoding. Although many aspects of JT CoMP - ranging from user grouping, clustering, back-haul requirements up to channel prediction - have been evaluated and leading to reasonable solutions, the existing literature have paid very limited attention to the processing complexity of large cooperation areas over several sites or small cells, where each of the cells is equipped with a massive MIMO antenna array. In other words, JT CoMP which requires adjacent sites to be synchronized exchanging user data constantly, will be facing with such high processing power which can not be manageable with todays processor cards. Therefore, computational complexity will be a challenging issue bringing a significant burden for future eNodeB.

The above complexity has been achieved neglecting any special matrix structure. But, for a typical JT CoMP channel matrix where the fixed beams are gridded directly to the users, high percentage of the channel coefficients are below a certain

power threshold of e.g. 25dB and might therefore being set to zero with minor impact to the precoding performance.

In this paper, we bridge the gap by proposing a novel algorithm that essentially calculates the Moore-Penrose inverse of sparse channel components matrix used in precoding generation process at the base station. More specifically, the main contribution of this paper is reducing the number of FLOPs for Moore-Penrose inverse calculation of channel components matrix taking into account the sparsity as the main feature of such matrices.

The rest of the paper is formed as follows. First, we detail the general system model features in Section II. The main analytical algorithm is presented in Section III. Numerical results including system parameters and performance examples are discussed in Section IV. Conclusion accomplishes this paper in Section V.

II. System model

We consider a 5G system concept explicitly described in [3]. This integrated system is assumed to support massive MIMO along with JT CoMP (which provides synchronization of distant sites, multi-cell channel estimation, feedback of the channel state information (CSI), synchronous exchange of user data and joint processing) to increase the spectral efficiency. Forming cell clusters in the network denoted as cooperation areas (CA), the served users are expected to gain through cooperation. In this paper we limit our analysis to a single cooperation area comprising three adjacent sites with three small cells per site, i.e. overall 9 cells (See Figure 1).

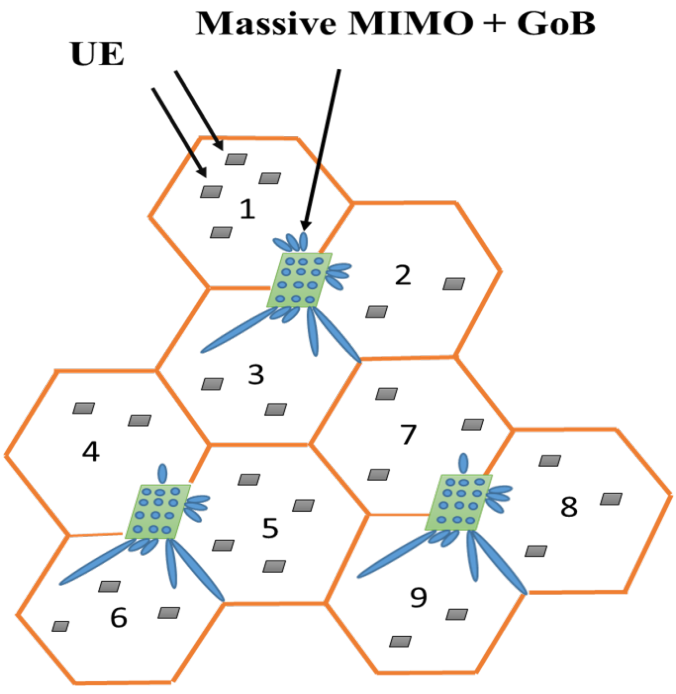

Fig. 1: Cooperation area architecture.

In order to reduce the number of antenna elements since each cell is equipped with a massive MIMO antenna array, a limited set of effective beams or antenna ports can be used. these effective beams can be generated through a fixed GoBs. Moreover, the interference can be more localized when the radial subsectors created via fixed GoBs is combined with massive MIMO. If we assume that GoB precoding matrix at the eNodeB denoted as **V** generates eight beams in azimuth and two in elevation direction combines with two polarizations per beam, it will result in 32 beams per cell (See Figure 2).

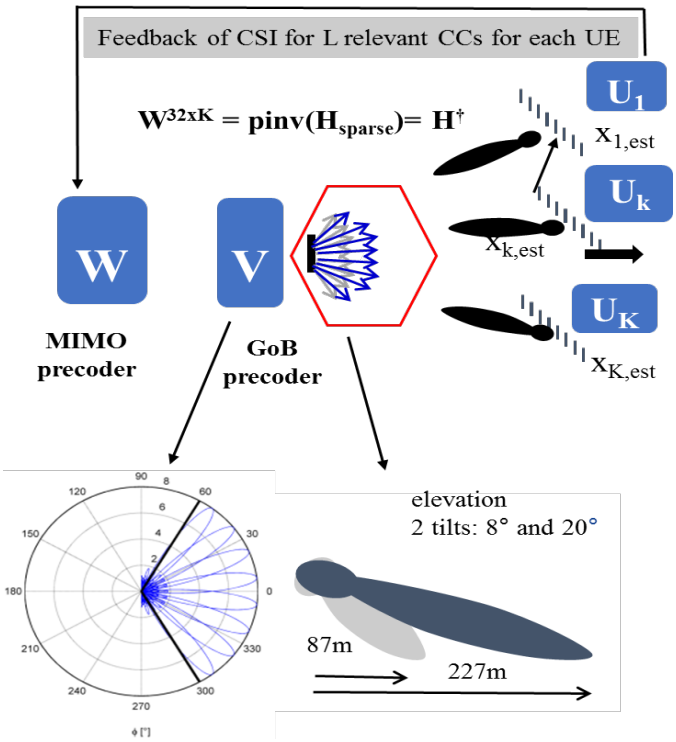

Fig. 2: System model precoding process in each cell.

Furthermore, it is assumed that the UEs are in rank one reception mode and apply in case of multiple UE antennas a suitable Rx beamformer. In case of FDD, all the UEs report the so-called relevant channel components (RCCs) which defines the number of channel coefficients above a certain power threshold for each user with respect to the strongest channel component to the eNodeB. In case of single cell MU-MIMO, the MIMO channel precoding matrix **W** will be calculated through the pseudo inverse of the matrix **H** which becomes sparse after applying the power threshold through the UEs CSI feedback process. As such, the form of the matrix **W** depends on the certain power threshold applied. The lower the power threshold relative to the strongest channel component, the higher is the sparsity of **W**. Obviously, in case of JT CoMP up to 9 cells, a single precoder spans all cells with an accordingly high matrix dimension of e.g. K=90 UEs times N=288 antenna ports, which is significantly larger compared to todays 3GPP LTE matrix dimensions of e.g. 4×4 or maximal 8×8.

III. Performance Analysis

In this section, we provide our detailed analysis allowing us to calculate the pseudo inverse of the channel component matrix efficiently.

A. State-of-the-art Moore-Penrose Inverse computation algorithm

Our below analysis extends the known results for so-called $geninv$ method of Moore-Penrose inverse computation published in [5]. More specifically, we target to investigate the number of FLOPs using such method with respect to sparsity of the matrix. If we denote $\mathbf{H}$ as complex-valued channel components matrix of size $m \times n$ where $m < n$, we could consider the symmetric positive matrix $\mathbf{HH'}$ of size $m \times m$ and rank of $r \leq n$ where $\mathbf{H'}$ corresponds to transpose of the channel matrix $\mathbf{H}$. By using the Cholesky factorization of matrix $\mathbf{HH'}$, the matrix $\mathbf{L}$ of size $n \times r$ is obtained:

$$\mathbf{Chol}(\mathbf{HH'}) = \mathbf{L}. \tag{1}$$

Using the general relation concerning the product of two matrices $\mathbf{A}$ and $\mathbf{B}$, we have:

$$(\mathbf{AB})^{+} = \mathbf{A'B'}\left(\mathbf{A'ABB'}\right)^{+}, \tag{2}$$

where $(\mathbf{AB})^{+}$ represents the Moore-Penrose inverse of the product of two matrices $\mathbf{A}$ and $\mathbf{B}$. If $\mathbf{B}=\mathbf{A'}$ and $\mathbf{A}$ is $n \times r$ matrix of rank r, then from (2) we have:

$$\left(\mathbf{AA'}\right)^{+} = \mathbf{A'A}\left(\mathbf{AA'}\right)^{-1}\left(\mathbf{AA'}\right)^{-1}. \tag{3}$$

Proposition 1. *The pseudo inverse of matrix **H** defined above can be obtained by the following expression [5]:*

$$\mathbf{H}^{+} = \mathbf{H'L}\left(\mathbf{L'L}\right)^{-1}\left(\mathbf{L'L}\right)^{-1}\mathbf{L'}$$

Proof. Considering eq.(2) and (3), we may write down the following:

$$\mathbf{H}^{+} = \mathbf{H'}\left(\mathbf{HH'}\right)^{+}. \tag{4}$$

If we denote $\mathbf{L'}$ as the transpose matrix of $\mathbf{L}$, from eq.(3) and eq.(4) we obtain:

$$\left(\mathbf{HH'}\right)^{+} = \mathbf{L}\left(\mathbf{L'L}\right)^{-1}\left(\mathbf{L'L}\right)^{-1}\mathbf{L'}. \tag{5}$$

Therefore, the Moore-Penrose inverse of matrix $\mathbf{H}$ using the discussed approach can be expressed as follow:

$$\mathbf{H}^{+} = \mathbf{H'L}\left(\mathbf{L'L}\right)^{-1}\left(\mathbf{L'L}\right)^{-1}\mathbf{L'}. \tag{6}$$

□

B. Efficient sparse Moore-Penrose inverse algorithm

After above analysis of the state-of-the-art Moore-Penrose inverse algorithm, we continue by proposing an efficient approach comprising a set of techniques in order to reduce the computational complexity specifically when the sparsity in the channel components matrix is considered. For every $\mathbf{H}^{+}$ matrix obtained through the algorithm detailed above, the two main operations include the full rank Cholesky factorization of $\mathbf{HH'}$ and inverse of $\mathbf{L'L}$. Considering the eq.(6), we may divide this process into three steps as follows:

$$\begin{aligned}
\mathbf{P_1} &= \mathbf{L} \times \left(\mathbf{L'L}\right)^{-1}, \\
\mathbf{P_2} &= \mathbf{P1} \times \ \mathbf{P'_1}, \\
\mathbf{P_3} &= \mathbf{H'} \times \mathbf{P_2}.
\end{aligned} \tag{7}$$

In the following, we mainly focus on reducing the number of operations for each matrix multiplication exploiting the sparsity of matrix $\mathbf{H}$. The $Sgeninv$ algorithm for efficient inverse of sparse matrices comprises four steps, being applied to the three multiplications as defined in (7):

1) Store the location of non-zero elements in the matrix before reordering,
2) Reorder the sparse matrix applying Reserve Cuthill-Mckee (RCM) algorithm [6],
3) Store the location of non-zero elements after reordering,
4) Multiply two new reordered matrices together with avoiding unnecessary multiplication and additions.

1) Store the location of non-zero elements in the matrix before reordering: First, we store the location of non-zero elements which could be done by obtaining the adjacent graph of the matrix.

2) Reorder the sparse matrix applying Reserve Cuthill-Mckee algorithm [6]: The adjacent graph of the sparse matrix is used as input to produce a matrix with much smaller bandwidth by employing the Reserve Cuthill-Mckee algorithm. As such, we will obtain the sparse matrix elements where the non-zero of initial sparse matrix are relabeled. Figure 3 represents an effect of RCM reordering process on the order of adjacent graph nodes related to sparse matrix $\mathbf{A}$. Matrix $\mathbf{A}^{*}$ denotes the reordered sparse matrix of $\mathbf{A}$.

3) Store the location of non-zero elements after reordering: The relabeled nodes (non-zero elements) are stored in this step.

4) Multiply new reordered matrices together with avoiding unnecessary multiplications and additions: Obtaining the reordered sparse matrix where all the non-zero elements of each row become closed to the diagonal, we will be able to specify the position of non-zero elements in the matrix. As such, unnecessary multiplications (zero by non-zero elements multiplications) and inessential additions (like zero and non-zero elements additions) can be avoided which leads to reducing the processing complexity.

However, the additional computational cost in this method includes RCM reordering and identifying non zero positions. RCM algorithm cost is defined in order of $O(\log(f)|E|)$ where f and E denote the maximum degree in the adjacent graph and the total number of edges in the considered sparse matrix, respectively. Thus, the proposed $Sgeninv$ is applied as part of existing $geninv$ to improve the efficiency in terms of FLOPs specifically when the sparsity is considered. The method which counts number of operations on top of discussed algorithms will be detailed in Section IV (C).

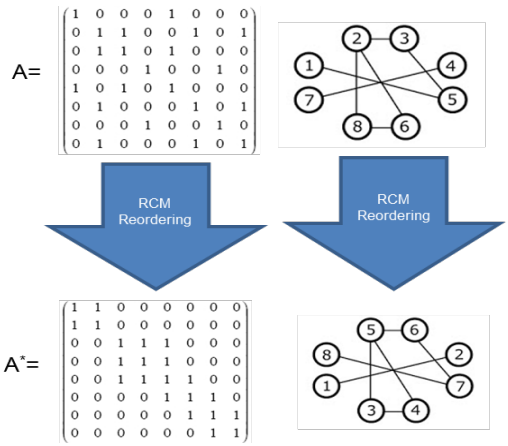

Fig. 3: The effect of RCM algorithm on matrix of A and its adjacent graph.

IV. Numerical Results and Summary

A. System Parameters

In what follows, we detail all the parameters and assumptions regarding the numerical data presentation based on the system model discussed above.

For evaluation we have concentrated on two different channel matrices representing inter-site JT CoMP over a 3 site cooperation area. Channel matrix I of size 40×288 has been obtained through a channel model based on Quasi Deterministic Radio Channel Generator (QuaDRiGa) [7]. This channel matrix belongs to the case where we distribute 40 users in a 3 site coopearion area featuring 288 fixed grid of beams. As the second examined channel, we have employed a 3D urban macro channel matrix of size 90×288 based on 3GPP 3D channel model for LTE [8] where 10 users have been placed in each cell. These two matrices have different channel characteristics which have been demonstrated in Table I. Additionally, we have been focused on intra-site cooperation scenario taken exclusively from channel matrix I where a matrix of size 30×96 assuming 10 users per site and 32×3 = 96 beams (three cells per site) are considered. To exemplify the characteristics of both channel matrices in terms of sparsity level, we have used an approach by employing three different thresholds referring to as the minimum power level with respect to the strongest channel component of the matrix. As such, we set those matrix elements of each UE to zero for which the Rx power is below the power of the strongest channel component minus the threshold value.

B. Signal-to-Interference-Plus-Noise Ratio and Spectral Efficiency

In order to examine the effect of channel matrix sparsity level (caused by employing power thresholds) on the system-level performance, we have taken the channel matrices described in Section IV (A) and compare the spectral efficiency in the noisy cooperation area. If we use a bandwidth of 20 MHz around a carrier frequency of 2.1 GHz, we will obtain

Parameter	Channel Matrix I	Channel Matrix II
Number of Cells	9	9
Number of Users	40	90
Number of Sites	3	3
Inter-Site Distance (m)	500	500
Channel Type	QuaDRiGa [6]	3GPP LTE 3D [8]
Number of Channel Taps	24	24
Number of Beams	288	288

TABLE I: Channel Matrices Properties

100 PRBs over LTE. We assume that the transmit power is equally distributed over 1200 subcarriers. The simulation parameters are summarized in Table II. The spectral efficiency (SE) per cell can be defined as the following:

$$SE = \frac{1/N_{Cell} \times N_{UE} \times C_{LTE} \times N_{sub} \times N_{OFDM}}{TTI \times B}, \quad (8)$$

where N_{Cell} is number of cells, N_{UE} is the number of UEs, C_{LTE} is the achievable capacity of a memoryless channel in LTE, N_{sub} denotes the number of subcarriers over 100 PRBs, N_{OFDM} displays the number of OFDM blocks per subframe, TTI denotes the duration of one subframe and B is used bandwidth. Achievable number of bits per second C_{LTE} is based on the 3GPP physical layer procedures described in [9] where the measured signal-to-interference-plus-noise ratio (SINR) is mapped to CQI value and eventually the spectral efficiency in bit/sec/Hz/cell can be obtained. The derived SINR of the channel model per user in each iteration can be obtained as the following:

$$SINR[\text{dB}] = \text{y}_{UE(i)} \times P_{tx}[\text{dBm}] - \text{P}_n[\text{dBm}] - \text{I}_{UE(i)}, \quad (9)$$

where

$$P_{tx} = \text{Total Tx Power}_{dBm}/N_{UE} - 10\log_{10} N_{sub} \times N_{PRB}, \quad (10)$$

$$P_n = -173.83 + 10\log_{10} N_{SubSp} + ReceiverNF[\text{dB}]. \quad (11)$$

$\text{y}_{UE(i)}$ in (9) denotes the received signal at the UE with index i. As discussed above, if we use zero forcing beamforming (ZFBF) where the linear precoders are determined according to an interference zero forcing objective, then the precoding matrix can be obtained:

$$\mathbf{W} = \mathbf{H^H}(\mathbf{H}\mathbf{H^H})^{-1}, \quad (12)$$

where $\mathbf{H}$ is complex-valued channel matrix and $\mathbf{H^H}$ denotes the complex conjugate transpose of matrix $\mathbf{H}$. With this choice of precoding matrix, the received signal is assumed as follows:

$$\mathbf{y} = \mathbf{H}\mathbf{W}. \quad (13)$$

If the proposed pseudo inverse method is applied, the received signal in (13) will not be diagonal due to applying the manipulated channel matrix in precoding process (See eq. (12)). P_{tx} in (9) specifies the transmit power distributed equally among all the users which is obtained via (10). $\text{I}_{UE(i)}$ denotes the

PHY Layer Parameters	Value
Carrier Frequency (GHz)	2.1
Total Bandwidth (MHz)	20
Used Bandwidth (MHz)	18
Subcarrier Spacing (KHz)	15
Number of Subcarriers	1200
Number of PRBs	100
Noise level (dBm)	-125.0691
Total Tx Power per cell (dBm)	46
Receiver NF (dB)	7
Number of UEs (Channel Matrix I)	40
Number of UEs (Channel Matrix iI)	90
Number of Cells	9
Number of OFDM blocks per subframe	14
Duration of one subframe (ms)	1

TABLE II: Simulation Parameters

received interference at the UE with the index i in dBm. N_{UE} is number of UEs in one cell, N_{sub} is the number of subcarriers and N_{PRB} is the number of PRBs expressed in (10). In order to observe the effect of noise to the channel, the noise formulation referred to as P_n has been used. N_{SubSp} denotes the subcarrier spacing and $ReceiverNF$ refers to the receiver noise figure in dB. Obviously, channel interference shown in eq.10 is equal to zero when the perfect zero forcing utilizing the Moore-Penrose pseudo inverse is applied. However, when the channel components are set to zero based on the explained power threshold approach we will receive interferences from neighboring cells after ZFBF process at the receiver. Thus, we essentially deal with not only the noise level but also the imposed interferences when the spectral efficiency discussed in the proposed approach.

C. Performance examples

In what follows, we discuss the obtained numerical results. In order to calculate the number of FLOPs, we have used the Lightspeed toolbox version 2.7 [10] since built in MATLAB function FLOPs retired right after MATLAB release 6 due to its infeasibility to keep running count of FLOPs. Lightspeed provides count routines for various arithmetic operations specifically required to calculate the discussed Moore-Penrose inverse methods with higher accuracy compared with FLOPs function included in MATLAB version 6. As such, we are able to count the FLOPs through this toolbox by collecting and combination of essential arithmetic operations required in the presented Moore-Penrose inverse methods. Performance analysis described in Section III (B) can be applied to any method using for Moore-Penrose inverse. In order to observe the pseudo inverse computational complexity using different power thresholds applied on channel matrix I, two traditional methods include QR decomposition and singular value decomposition (SVD) as well the proposed $SGeninv$ have been compared (See Figure4). Sparsity level of the channel matrix of size 40×288 is gained through changing the power threshold with respect to the strongest

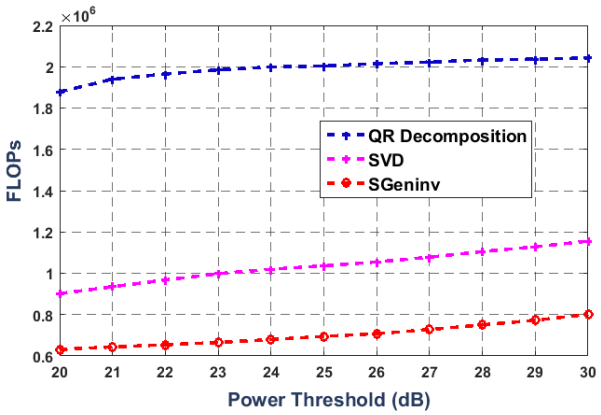

Fig. 4: FLOPs comparison between different pseudo inverse methods .

Method	Intra-Site Cooperation	Inter-Site Cooperation	Receiving Power Threshold (dB)
Pinv	2416162	55344466	15,20,30
geninv	522349	4986003	15,20,30
Sgeninv	3819	278013	15
Sgeninv	10920	512925	20
Sgeninv	42229	617588	30

TABLE III: Average Number of FLOPs (Channel Matrix I)

Method	Intra-Site Cooperation	Inter-Site Cooperation	Receiving Power Threshold (dB)
Pinv	2416162	97397366	15,20,30
geninv	522349	27670499	15,20,30
Sgeninv	6738	330062	15
Sgeninv	16783	684546	20
Sgeninv	93712	135351	30

TABLE IV: Average Number of FLOPs (Channel Matrix II)

channel component per user. As it can be seen, both QR decomposition and SVD impose higher number of FLOPs to calculate the pseudo inverse even if we use the proposed method before the precoding process. By contrast, if we apply the proposed methods (See Figure4 the red curve) specifically on top of (6), we could obtain 30 gain comparing to SVD and nearly 70 percent gain comparing to QR decomposition. Thus, the properties of sparse matrix can be applied to any pseudo inversion method. However, the processing cost will be increased dramatically.Table III compares the average number of FLOPs for calculation of 100 precoders for 100 physical resource blocks using channel matrix I applying three methods including the MATLAB Pinv function which is mainly based on SVD, state-of-the-art $geninv$ algorithm as discussed in Section III and the proposed $Sgeninv$.

Interestingly, number of FLOPs calculated by $geninv$ and MATLAB Pinv in both inter-site (one cooperation area) and

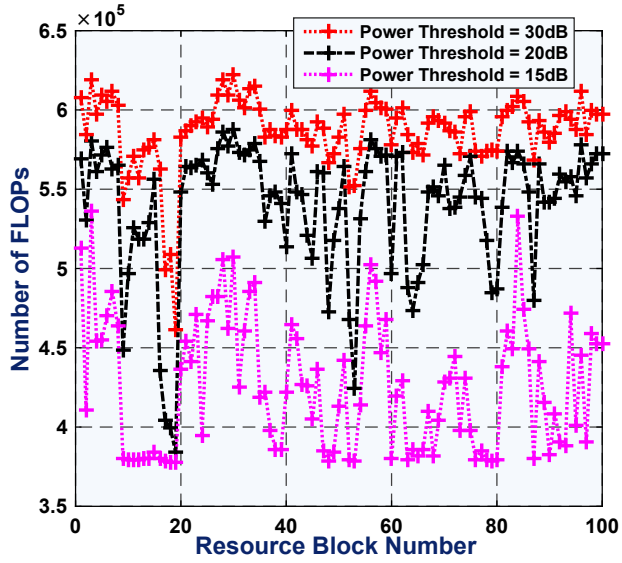

Fig. 5: FLOPs for inter-site cooperation (Channel I).

intra-site (one site) cooperation is independent from sparsity of the matrix. In other words, changing the receiving power threshold has no effect on the number of FLOPs when these two methods are applied. By contrast, the proposed $Sgeninv$ gains from sparsity. As it can be seen, this approach clearly reduces the computational complexity approximately by factor of 10 compared with $geninv$ and by factor of 100 comparing with MATLAB Pinv function. Average number of FLOPs counted for channel matrix II are provided in Table IV. it

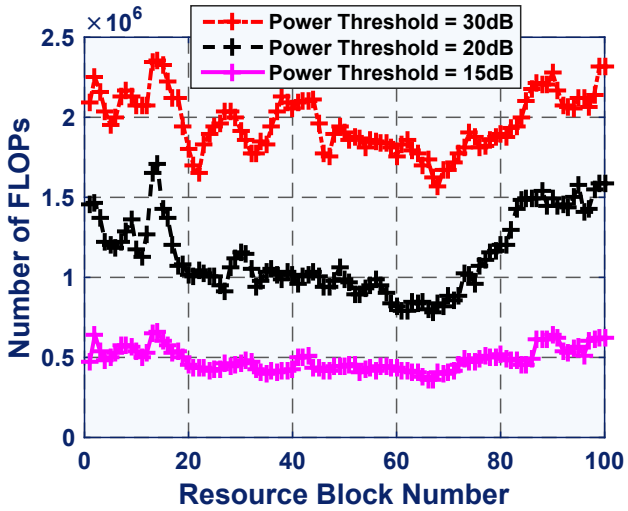

Fig. 6: FLOPs for inter-site cooperation (Channel II).

is shown that the precoding process complexity in FLOPs can be also reduced up to order of 10 comparing with the state-of-the-art algorithm in both inter-site and intra-site scenarios. As such, the proposed approach can be used for larger sparse channel matrices to combat complexity efficiently. Figure 5 illustrates the FLOPs comparison between different receiving

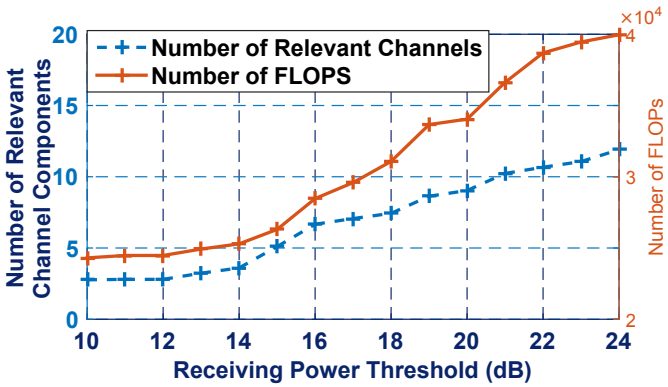

Fig. 7: FLOPs for Intra-site cooperation using Sgeninv method (Channel I).

power thresholds in inter-site cooperation when the proposed $Seninv$ is applied on channel matrix I for 100 resource blocks. Figure 6 depicts the effect of setting components of channel matrix II to zero based on three different power thresholds applied in one cooperation area. Similarly, complexity in terms of FLOPs increases as the sparsity level of the channel matrix is reduced. However, the complexity difference in channel matrix II between different thresholds is higher compared to the experienced channel matrix of size 40×288.

Figure 8 confirms that for JT CoMP channel matrix I, total number of relevant channel components per user will not be greater than 70 when the power threshold approaches 30dB. Thus, we will obtain a channel matrix with approximately 80 percent sparsity which can be efficiently used for precoding matrix generation with lowest complexity.

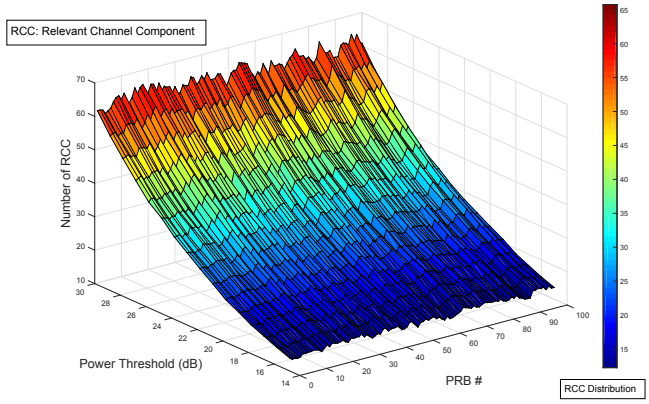

Fig. 8: Relevant Channel Components per User (Channel I).

Figure 7 contrasts FLOPs and number of intra-site cooperation relevant channel components per user in the channel matrix I when different receiving power thresholds are being applied. As it can be seen, $Sgeninv$ method confirms that both FLOPs and relevant channel components per user increase linearly with respect to the sparsity level of the matrix. Therefore, using higher thresholds (lower sparsity level) will result

in higher relevant channel components and higher number of FLOPs. In order to conceive the effect of proposed approach on the distribution of relevant channel components, Figure 8 and 9 are provided.

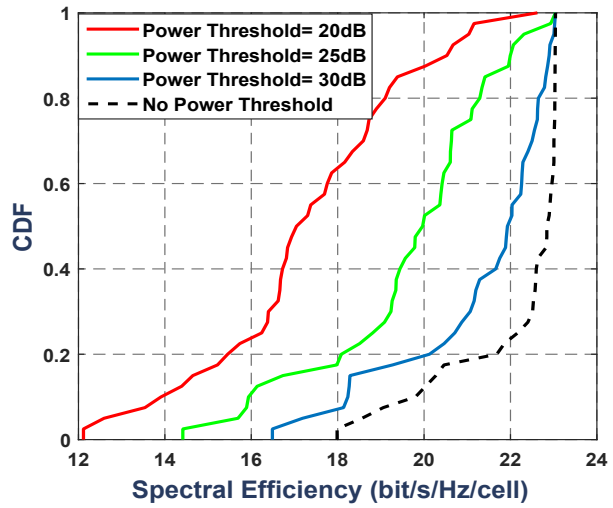

Fig. 9: CDF plot for Spectral Efficiency per Cell (Channel I).

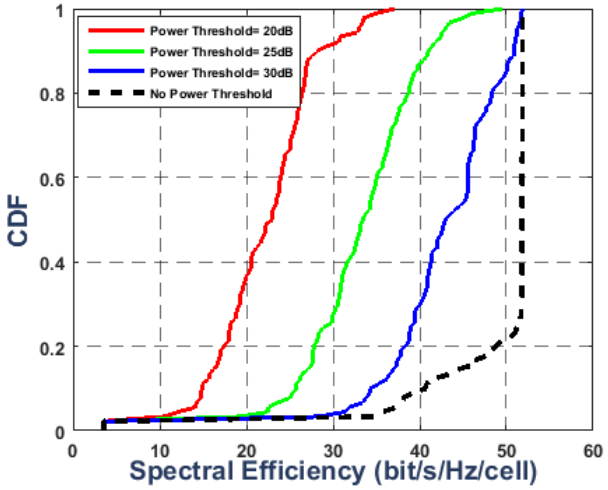

Fig. 10: CDF plot for Spectral Efficiency per Cell (Channel II).

To exemplify the spectral efficiency per cell using the proposed approach, CDF plots in Figure 9 and Figure 10 are provided. As it can be observed, a reasonable spectral efficiency comparing with the perfect channel case (See the dashed curve in Figure 9) with approximately 16 percent loss can be obtained in channel matrix I applying power threshold of 25 dB. Nevertheless, performance can be degraded slightly with the same threshold of 25dB applied on the larger channel matrix (See Figure 10). Choosing appropriate threshold, the effect of interferences received at the eNodeB due to cutting off channel components can result in not only with reasonable efficiency but also it provides a very low complex precoding process based on the proposed approach.

V. Conclusion

The processing complexity for a central JT CoMP precoder can be kept under control for future 5G radio systems even in case of very large cooperation areas comprising 3 sites or nine small cells and high number of beams. Simulation results for such large cooperation areas demonstrate that computational complexity in channel matrix pseudo inverse process used in precoding process can be reduced by factor 100 compared to SVD and by 10 compared to state-of-the-art algorithm. Exploiting the sparsity of typical massive MIMO channel matrices allows to replace many arithmetic operations by low cost operations. System level simulation results strongly confirm that we could achieve reasonable spectral efficiency with moderate degradation of the precoding performance by finding appropriate power threshold leading to setting 70 to 80 percentage of the channel components to zero for such purpose.

References

[1] F. Schaich, B. Sayrac, M. Schubert, H. Lin, K. Pedersen, M. Shaat, G. Wunder, and A. Georgakopoulos, "Fantastic - 5G: 5G-PPP project on 5G air interface below 6 GHz," in *Proc. of the IEEE European Conference on Networks and Communications*, July 2015.

[2] A. Osseiran, F. Boccardi, V. Braun, K. Kusume, P. Marsch, M. Maternia, O. Queseth, M. Schellmann, H. Schotten, H. Taoka, H. Tullberg, M. Uusitalo, B. Timus, and M. Fallgren, "Scenarios for 5G mobile and wireless communications: The vision of the metis project," *IEEE Wireless Communications Magazine*, vol. 52, no. 5, pp. 26–35, May 2014.

[3] V. Jungnickel, K. Manokolis, W. Zirwas, B. Panzner, V. Braun, M. Lossow, M. Sternad, R. Apelfrojd, and T. Svensson, "The roll of small cells, coordinated multipoint, and massive MIMO in 5G," *IEEE Wireless Communications Magazine*, vol. 52, no. 5, pp. 44–51, May 2014.

[4] D. Gesbert, T. Svensson, and Ed. (2012) Deliverable d1. 4 in artist4g, interference avoidance techniques and system design. [Online]. Available: http://ict-artist4g.eu/project/deliverables/

[5] P. Courrieu, "Fast computation of moore-penrose inverse matrices," *Neural Information Processing*, pp. 25–29, 2005.

[6] E. Cuthill and J. Mckee, "Reducing bandwidth of sparse symmetric matrices," in *Proc. of the ACM 24th Nat. Conference*, 1969, pp. 157–172.

[7] S. Jaeckel, L. Raschkowski, K. Börner, and L. Thiele, "Quadriga: A 3-d multi-channel model with time evolution for enabling virtual field trials in 5G," *IEEE Trans. Antennas Propag.*, vol. 62, no. 5, pp. 3242–3256, 2014.

[8] "Study on 3D channel model for LTE (Release 12)," *3GPP, Technical Report (TP) 36.873: Technical Specification Group Radio Access Network, 3GPP, Tech. Rep. V12.2.0*, 2015.

[9] "Study on Physical Layer Procesures," *3GPP Technical Report (TR) 136.213,3GPP, Tech. Rep. V8.8.0*, 2010.

[10] T. Minka. (2013) The lightspeed toolbox. [Online]. Available: http://research.microsoft.com/en-us/um/people/minka/software/lightspeed/