Reducing the Computational Complexity of Massive MIMO using Pre-coding Techniques under Some Lower Orders

Teerapat Sanguankotchakorn and Ganti V. Sowmya
Asian Institute of Technology, Pathumthani 12120, Thailand
Email: teerapat@ait.ac.th; sowmyaganti96@gmail.com

Abstract**—**Massive Multiple-input and Multiple-output (MIMO) is considered as a solution to the next generation cellular systems. It is visualized to provide extensive upgrade in capacity, along with the computational complexity as well as hardware. The main drawback of massive MIMO is the computational complexity in pre-coding, particularly when the "relative antenna-efficient Regularized Zero-Forcing (RZF)" is chosen to simplify Maximum Ratio Transmission (MRT). In this work, we propose to use the beam-forming methods, especially a hybrid pre-coding to reduce the system complexity in Massive MIMO. However, not only the system complexity, but also the computational complexity in pre-coding is the significant issue. In this regard, we propose another technique called Truncated Polynomial Expansion (TPE) pre-coding. It can emulate the same advantages of RZF, while offering the lower and extensible computational complexity that is achievable in an efficient pipelined fashion. By using random matrix theory, we can derive a closed-form expression of the SINR under TPE pre-coding. The proposed scheme is executed in an ideal Rayleigh fading channels, so that it produces highly desirable performance. Finally, we compare the results achieved from our proposed TPE pre-coding using three lower orders with RZF under various Channel State Information (CSI). It is obvious that our proposed method can provide the closest match to RZF, while the computational complexity is lower.

Index Terms**—**Massive MIMO, Zero-Forcing (ZF) pre-coding, RZF

I. INTRODUCTION

Wireless communication is defined as "the transmission of information over a distance without the help of wires, cables or any other forms of electrical conductors." The sender sends the signals through a device that has a capacity to generate electro-magnetic signals to the receiver. There are various communication platforms such as Satellite, Mobile, Wireless network, Infrared and Bluetooth Communications. The main objective of all these platforms is to provide the high data rates. This desire for higher data rates leads the telecommunication engineers to implement new technology, such as the 5th Generation Mobile communication Technology.

The invention of 5G has enhanced the coverage, signaling & spectral efficiency [1], [2]. There are numerous advantages of 5G technology, namely: (a) High resolution along with bi-directional broad bandwidth assembling (b) Assembly of all networks on a single platform (c) Highly effective and efficient. It is expected that technology for 5G cellular communications is "Massive MIMO (Multiple-Input Multiple-Output) systems [1], [3], which contains a huge number of antennas. The number of the antennas at the base station is much larger than the number of the user equipment. The performance of massive MIMO systems heavily relies on the availability of Channel State Information (CSI) at the Base Station (BS). When the base station knows the CSI of the users, it can apply beam-forming techniques to improve the spectral and energy efficiencies of the system. The pre-coding and combining can be performed at the baseband by digital beam-forming techniques where it has a full control over the phase and amplitude of the signals at/from each antenna element. However, the beam-forming algorithms implemented at the digital baseband might become very complex [3]. Further, if all beam-forming is implemented at baseband and RF feed is necessary for each antenna, then, the system complexity, cost and loss increase with the increment of the number of antennas at high frequency.

To solve these problems, the hybrid beam-forming/hybrid pre-coding is proposed where a part of beam-forming is executed using only analog instead of digital components. There also exists an issue not only the system complexity but also the computational complexity in the pre-coding. To remedy this issue, a different class of "Truncated Polynomial Expansion (TPE)" pre-coder that restores the matrix inversion in place of matrix polynomial along with some orders of truncated polynomial is proposed. We investigate the value of TPE order that provides the close match to RZF under various Channel State Information (CSI), while providing the lower computational complexity. The performance is measured in terms of achievable data rate by simulation using MATLAB.

According to the simulations results, it is obvious that our proposed TPE with some lower orders can provide the close match to RZF. However, the computational complexity of our proposed method at high order of TPE is quite high comparing to RZF [4].

Manuscript received October 5, 2018; revised April 30, 2019.
Corresponding author email: teerapat@ait.ac.th
doi:10.12720/jcm.14.6.498-503

The rest of paper is structured as follows: In section 2, we describe the relevant literature review. Section 3 presents the system model as well as the analysis of RZF and TPE. Section 4 presents the results and discussion. Finally, we conclude our work in Conclusions section.

II. Related Works

The ZF pre-coding is simple to implement, but the performance is poor [5]. In [5], the RZF is proposed to address the drawback of ZF pre-coding and improve its performance significantly. Another version of ZF, proposed in [6], is a hybrid ZF called PZF (Phased Zero Forcing) pre-coding. It is shown that the PZF can reduce the system complexity comparing to the ZF pre-coding [6]. In [3], the MMSE (Minimum Mean Square Error), which is one type of hybrid pre-coding method, is considered. It is shown that the computational complexity drastically increases with the increment of number of antennas in MMSE. The RZF in massive MIMO is considered in [7]. Its complexity per coherence time is analyzed and provided concretely. The RZF provides a good balance between the complexity and the spectral efficiency among all linear pre-coding methods. However, its computational complexity is still high. The method, which can reduce the computation complexity of RZF is known as Truncated Polynomial Expansions (TPE) [8]. This method is reported to reduce the computational complexity and also be easy to implement.

III. System Model

A. System Model of Hybrid Precoding

Here, we analyse a downlink communication channel which consists of a BS equipped with N_t transmitting antennas, driven by a small amount of RF chains (K), as shown in Fig. 1. These chains restrict the amount of transmitting antennas to be K, so that we can estimate the amount of K single-antennas users that should be scheduled. The baseband $\mathbf{W}$ (of dimension $K\times K$) and the RF processing $\mathbf{F}$ (of dimension $N_t\times K$) are derived from the downlink pre-codings.

It can be observed that the amplitude and phase modifications in the baseband pre-coder $\mathbf{W}$ are possible, while only the phase changes are possible by the help of the phase shifters and the combiners in case of RF precoder ($\mathbf{F}$). Therefore, initially the entry of each $\mathbf{F}$ is normalized, so that it satisfies $\left|\mathbf{F}_{i,j}\right|=1/\sqrt{Nt}$ where, $|\mathbf{F}_{i,j}|$ denotes the magnitude of the $(i, j)^{\text{th}}$ element of $\mathbf{F}$.

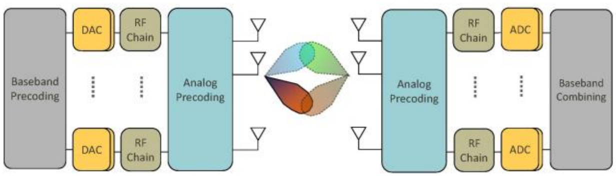

Fig. 1. Hybrid millimetre wave precoding [9]

Since, we chose a narrowband flat fading channel, then a baseband signal at the k^{th} user is

$$y_k = \mathbf{h}_k^H \mathbf{F}\mathbf{W}\mathbf{s} + n_k \tag{1}$$

Here, $\mathbf{h}_k^H$ represents the downlink channel and $\mathbf{s} \in \mathbb{C}^{K\times 1}$ represents the signal vector satisfying $E\left[\mathbf{s}\mathbf{s}^H\right]=\left(P/K\right)\mathbf{I}_k$ where $E[.]$ is the expectation operator and P is the transmitting power at the BS. Baseband is normalized so that we can accomplish $\left\|\mathbf{F}\mathbf{W}\right\|_F^2 = K$ such that it obtains the total power constraints. Finally, n_k denotes the additive noise.

Therefore, the SINR of the received signal at the k^{th} user is

$$SINR_k = \frac{\frac{P}{K}\left|\mathbf{h}_k^H\mathbf{F}\mathbf{w}_k\right|^2}{1+\sum_{j\neq k}\frac{P}{K}\left|\mathbf{h}_k^H\mathbf{F}\mathbf{w}_j\right|^2} \tag{2}$$

where $\mathbf{w}_j$ denotes the j^{th} column of $\mathbf{W}$. We use Gaussian inputs to achieve the long term spectral efficiency as follows:

$$R = \sum_{k=1}^{K} E\left[\log_2\left(1+SINR_k\right)\right] \tag{3}$$

B. Block Diagram of the Proposed Method

Fig. 2 illustrates the block diagram of our investigation in massive MIMO.

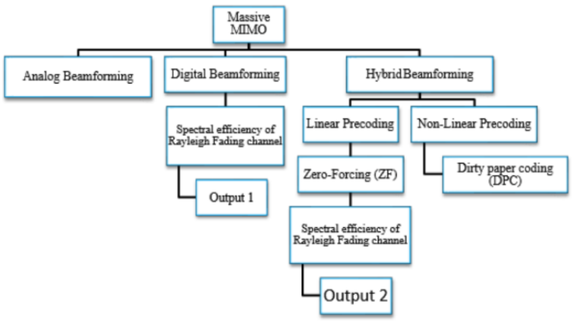

Fig. 2. Block diagram of our proposed methodology

Firstly, we investigate the spectral efficiency by creating the spectral efficiency of Rayleigh Fading channel of both the Digital beam-forming as well as the proposed Zero-Forcing of Hybrid beam-forming. Then, the spectral efficiencies of both schemes are compared. The results are as follows:

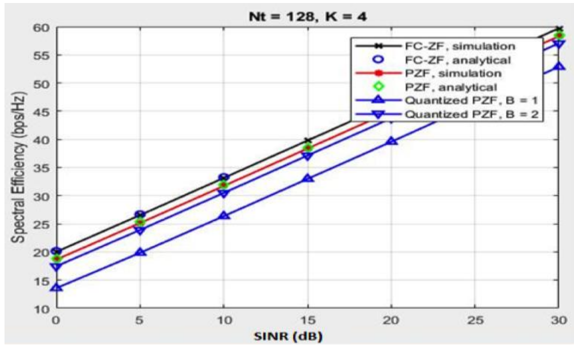

Fig. 3. Spectral efficiency for N_t = 128 and K = 4

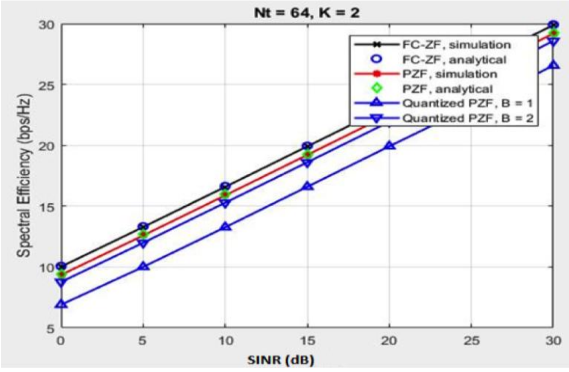

Fig. 4. Spectral efficiency for $N_t = 64$ and $K = 2$

According to the results shown in Fig. 3 and 4, we can note that the derived mathematical expressions provide exactly the same results as the simulations for the whole SINR values. In addition, it is obvious that the PZF [10] and Full-Complexity ZF (FC-ZF) pre-coding method provide almost the same spectral efficiency for all SINR range, thus contributing helpful guidance in designing the real time system.

Next, the achieved average bit rate of various beam-forming methods, namely Optimal Beam-Forming, Transmitted MMSE/Regularized Zero Forcing (RZF) Beam-Forming, Zero Forcing Beam-Forming (ZFBF) and MRT are compared, as shown in Fig. 5.

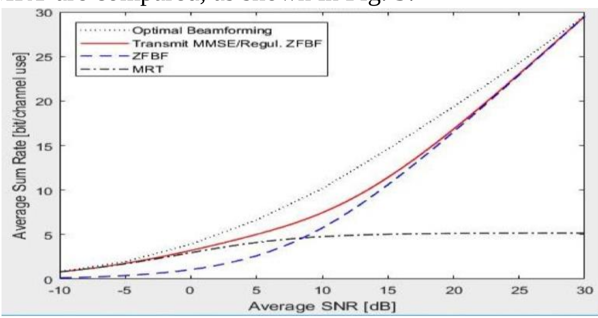

Fig. 5. Comparison of linear pre-coding methods [7]

Here, it can be observed that Transmit MMSE/RZF pre-coding provides an approximated result to that of an Optimal Beam-Forming, while the RZF pre-coding provides the closest approximation of the Optimal Beam-Forming at both low and high SNRs. However, the MRT (Maximum Ratio Transmission) provides a close approximation only at low SNR and ZF pre-coding provides a close approximation only at high SNR. Therefore, the Hybrid beam-forming is the best and most balanced beam-forming among all type of beam-forming methods.

Our main objective is to reduce the complexity, which is possible by considering the linear pre-coding method whose sub-methods are available [3]. One sub-method called Truncated Polynomial Expansions (TPE) is able to reduce the computational complexity and also be easy to implement. Therefore, we focus mainly on TPE.

C. System Models of RZF and TPE Precoding

The system model of linear pre-coding, which is a sub-method of hybrid pre-coding methods is already illustrated. So, let us go through the system model of the RZF and TPE, which are the sub-methods of the linear precodings.6. In this regard, the analysis of a single-cell downlink technique with a flat-fading channel and its errors is considered. A system under consideration here consists of a Base Station (BS) with M antennas and K single-antenna User Terminals (UTs). Its received signal can be shown as

$$y_k = \mathbf{h}_k^H \mathbf{x} + n_k \tag{4}$$

where, $\mathbf{x} \in \mathrm{C}^{\mathrm{M\,x\,1}}$, $\mathbf{h}_k \in \mathrm{C}^{\mathrm{M\,x\,1}}$, and $n_k \approx CN\left(0,\sigma^2\right)$ denote the transmit signal, random channel vector and additive complex Gaussian noise with zero mean and variance σ^2, respectively. Each and every channel vector $\mathbf{h}_\mathbf{k}$, for k = 1,…, K is modeled as

$$\mathbf{h}_\mathbf{k} \approx CN\left(\mathbf{0}_{Mx1}, \frac{1}{K}\Phi\right) \tag{5}$$

where, $\Phi \in C^{MxM}$ denotes the channel covariance matrix and is run with leaping strides of spectral norm, as $M \to \infty$. $\mathbf{0}_{MxM}$ denotes a vector with M zeros. Here, the covariance scaling $1/K$ is only for understanding purpose.

Now, let us analyze a Rayleigh fading channel in which the $\mathbf{h}_\mathbf{k}$ has a definite discernment for a coherence period and then derive a new uncommitted awareness. Having roots on this speculation, the transmitted signal can be given as

$$\mathbf{x} = \sum g_n s_n = \mathbf{GS} \tag{6}$$

The matrix syllabary is acquired by letting $\mathbf{G} = \left[g_1,...,g_k\right] \in \mathbf{C}^{MxK}$ be the pre-coding matrix and $\mathbf{S} = \left[s_1,...,s_K\right]^T \approx CN\left(\mathbf{0}_{Kx1}, \mathbf{I}_K\right)$ be the vector comprising all UT data symbols. Therefore, the received signal in (4) can be illustrated as

$$y_k = \mathbf{h}_k^H g_k s_k + \sum_{n=1, n\neq k}^{K} \mathbf{h}_k^H g_n s_n + n_k \tag{7}$$

Let $\mathbf{G}_k$ be represented as a matrix $\mathbf{G}$ while column g_k is isolated. Then, the SINR at the k^{th} user [7] is

$$SINR_k = \frac{\mathbf{h}_k^H g_k \mathbf{g}_k^H h_k}{\mathbf{h}_k^H \mathbf{G}_k \mathbf{G}_k^H h_k + \sigma^2} \tag{8}$$

By figuring that every user has faultless CSI (Channel State Information), the ergodic attaining the data rate at the UTs [7] is

$$r_k = E\left[\log_2\left(1 + SINR_k\right)\right] \tag{9}$$

The transmitter is undertaken to have unsound comprehension of the lighting channel awareness of each UT, for $k = 1,\ldots,K$.

By applying TPE to the RZF pre-coder, we can approximate the RZF with a low complexity by using TPE. It is observed in [11] that to approximate the RZF, the calculation of inverse of the Hermitian matrix is needed. Additionally, the inverse of any Hermitian matrix can be expressed as a matrix polynomial. That is, we can state that the lower terms in the expression are more important. Therefore, we can use the same concept in TPE and consider only the first *J* terms.

IV. Results and Discussion

In this work, we apply the previous mentioned TPE with some orders to approximate the RZF, so that it reduces the complexity caused by the RZF algorithm. After a few iterations, we have observed that the TPE orders vary with the change of *M* and *K*. Therefore, we can select the appropriate TPE order that matches the RZF, so that we can reduce the computational delay caused by RZF for the pre-computing process.

All the graphs in this section are plotted by the mean values of a number of iterations executed.

A. Achievable Average Data Rate at (M,K)=(512,128)

Let us consider the case of *M*=512 and *K*=128. By varying the quality of CSI (τ = 0.1, 0.4 and 0.7), we then compare the results of RZF and TPE with some lower orders (*J*=2,3,4).

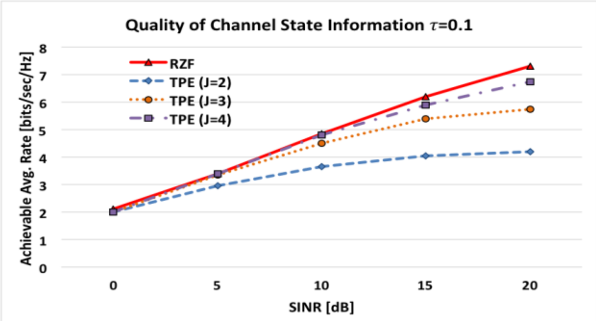

Fig. 6. Average Data Rate vs SINR for at (M,K) =(512, 128) and quality of CSI τ=0.1

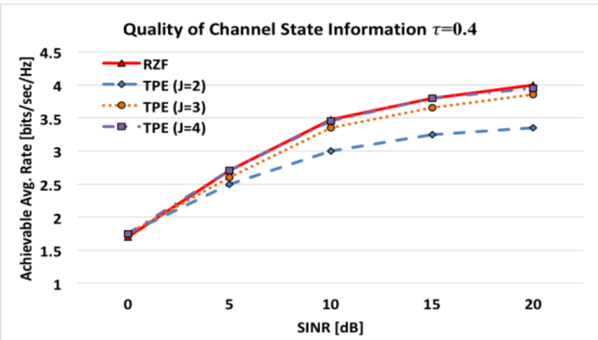

Fig. 7. Average data rate vs SINR at (*M, K*)=(512, 128) and quality of CSI τ =0.4

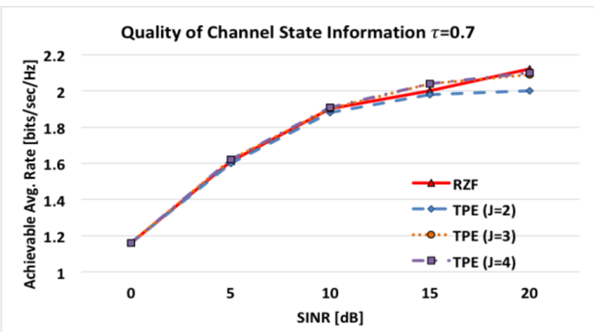

Fig. 8. Average data rate vs SINR at (*M, K*) =(512, 128) and quality of CSI τ=0.7

From Fig. 6 to 8, it is obvious that the proposed method can approximate closely to the RZF depending on TPE orders as well as channel quality.

For example, in Fig. 6, the channel quality is poor, so TPE order *J*=2 is not appropriate. However, at *J*=4, TPE provides the closest approximation of RZF and its loss is negligible. Similarly, if we consider Fig. 8, where the channel quality is quite good, we can see that the TPE order *J*=2, 3, 4 are appropriate values when compared to RZF. Therefore, in this case, we select *J*=2, which takes lower executing time compared to RZF, as shown in the Fig. 12.

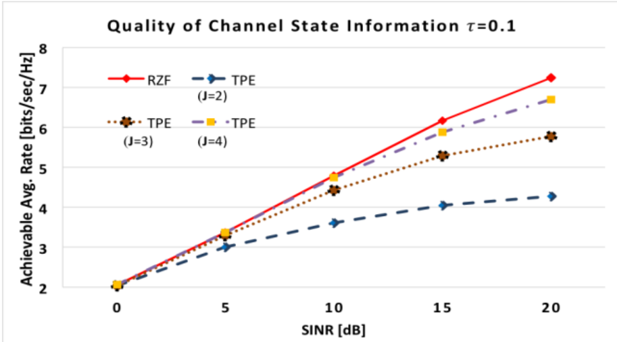

Fig. 9. Average data rate vs SINR at (*M, K*)=(256, 64) and Quality of CSI τ =0.1

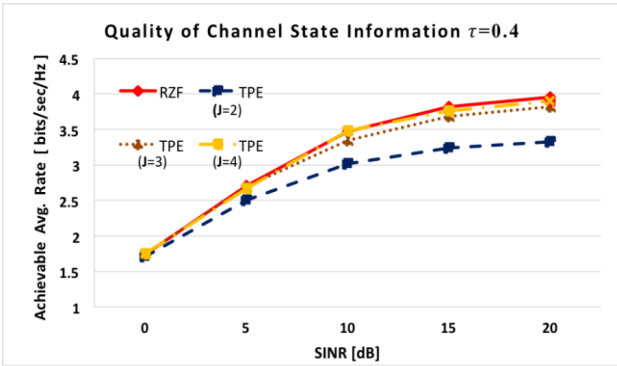

Fig. 10. Average data rate vs SINR at (*M, K*)=(256, 64) and quality of CSI τ=0.4

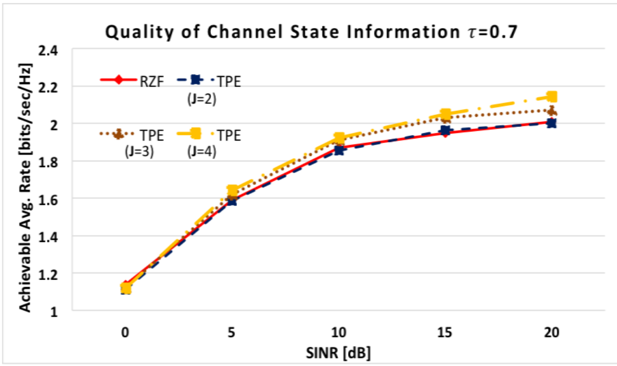

Fig. 11. Average data rate vs SINR at (*M, K*)=(256, 64) and quality of CSI τ=0.7

Finally, it can be concluded that if the channel quality is good (CSI value is high), TPE can provide the close approximation of RZF regardless of the order of TPE. However, when the channel quality is poor (CSI value is low), the higher the order of TPE, the closer the approximation of RZF is. That is, we need to select the nearest TPE order such that it provides the closest approximation of RZF, given that the execution time should not so high.

B. Achievable Average Data Rate at (M,K)=(256,64)

Now, we consider the case of (M, K)=(256, 64). Again, by varying the channel quality (τ = 0.1, 0.4 and 0.7), we can compare the outputs for RZF and TPE orders.

Similarly, in Fig. 9, where the channel quality is poor, so, TPE order J=2 is not appropriate. However at J=4, TPE provides the closest approximation of RZF and its loss is negligible. In Fig. 11, where the channel quality is good (CSI is high), it is apparent that the TPE provides the close approximation to RZF regardless of the order. Therefore, in this case, we select the order J=2, which takes the lower executing time comparing to RZF, as shown in the Fig. 13.

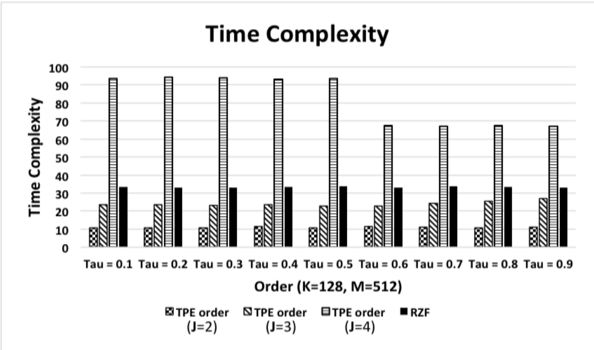

Fig. 12. Time complexity at (M, K)=(512, 128)

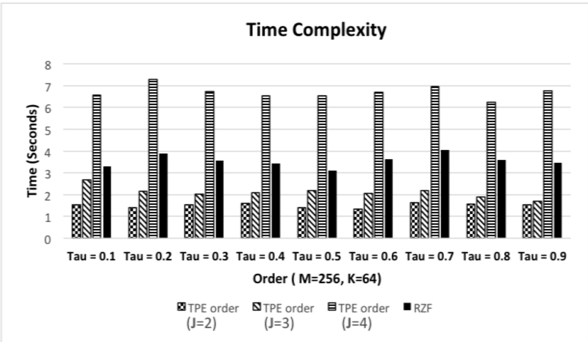

Fig. 13. Time complexity for (M, K)=(256, 64)

C. Time Complexity of RZF and TPE Pre-Coding

TABLE I: DETAILS OF IMPLEMENTED SYSTEM

Processor	Intel (R) Core™ i5-3437 CPU @ 1.90-2.40GHZ
RAM	8.00 GB
Operating System	Windows 10 Home, 64-bit
MATLAB version	R2017a

In this section, the time complexity is measured as an indirect evaluation metric of the computational complexity, since it is directly proportional to the computational complexity. Now, let us compare the pre-coding time for executing the proposed TPE for all CSI (τ) with the original RZF for 2 scenarios: (M, K)=(512,128) and (256, 64). Table I provides the details of platform used to execute the proposed method and measure the time complexity.

From Fig. 12 and 13, it is apparent that the time consumed by the RZF is between the TPE order J=3 and 4. As RZF requires time for pre-computing before the execution starts, we can select the nearby TPE order which provides closest approximation to RZF as it does not require any pre-computing.

V. CONCLUSIONS

Hybrid Beam-forming is considered as an attractive method adopted in massive MIMO, since it reduces the hardware complexity. However, to reduce the computational complexity, the linear pre-coding methods are used.

One of the main attractive linear pre-coding methods is the conventional RZF, since it provides the best throughput comparing to all other linear pre-coding methods. However, its disadvantage is the implementation, which is quite difficult since it involves channel matrix inversion. So, in order to reduce the computational complexity, we propose to adopt the TPE pre-coding method, in which we replace the matrix inversion with the truncated polynomials. Our proposed method is evaluated by simulation using MATLAB. In our simulations, the TPE with some orders and channel quality are varied, then the achievable average bit rate as well as the time complexity are measured. The time complexity is measured to illustrate the computational complexity of the system.

According to our simulation results, it can be concluded that if the channel quality is good (CSI value is high), TPE can provide the close approximation of RZF regardless of the order of TPE. However, when the channel quality is poor (CSI value is low), the higher the order of TPE, the closer the approximation of RZF is. That is, we need to select the TPE order such that it provides the closest approximation of RZF, with the time complexity directly reflecting the computational complexity should be taken into account. That is the execution time should not so high.

Finally, the main advantages of this method can be summarized as follows:

1. Pre-computing of the matrix is not necessary, so channel usage for data transmission is reduced.
2. The computational delay reduces.
3. The multistage structure is possible, since this proposed methods use a pipeline process

Therefore, TPE is considered as an effective option due to its wide range of applicability. It can be used even though the number of antenna increases, unlike RZF, where it needs not go through pre-computing process iteratively. Additionally, the previous work is not needed to put into account.

Recommendation for the future works:

1. To reduce the time complexity which reflect the computational complexity of the proposed TPE.

Because, on one hand, the higher the TPE order is, the closer the result matches to the optimal value. However, on the other hands, the higher the TPE order, the higher the time complexity becomes.

2. The optimal balance between the closest approximation of RZF and the computational complexity should be considered.

ACKNOWLEDGMENT

The authors wish to thank Asian Institute of Technology (AIT) for providing grant and some necessary tools to carry out this work.

REFERENCES

[1] S. Dutta and T. Sanguankotchakorn, "4-Element 28 GHz PIFA MIMO antenna design for future 5G applications," in *Proc. 2nd IEEE International Conference on Recent Trends in Electronics, Information and Communication Technology*, India, May 19-20, 2017.

[2] Y. A. M. K. Hashem, O. M. Haraz, and E. M. El-Sayed, "6-element 28/38 GHz dual-band MIMO PIFA for future 5G cellular systems," in *Proc. IEEE International Symposium on Antennas and Propagation*, June 26-July 1, 2016, pp. 393-394.

[3] E. Bjornson, J. Hoydis, and L. Sanguinetti, "Massive MIMO networks: Spectral, energy and hardware efficiency," *Foundations and Trends: Signal Processing*, vol. 11, 2017.

[4] S. Wagner, R. Couillet, M. Debbah, and D. T. M. Slock, "Large system analysis of linear precoding in correlated MISO broadcast channels under limited feedback," *IEEE Transactions on Information Theory*, pp. 4509-4537, 2012.

[5] C. M. Peel, B. M. Hochwald, and A. L. Swindlehurst, "A Vector perturbation technique for near-capacity multiantenna multiuser communications-part I: Channel inversion and regularization," *IEEE Trans. of Communications*, vol. 53, pp. 195-202, January 2005.

[6] L. Liang, W. Xu, and X. Dong, "Low-complexity hybrid precoding in massive multiuser MIMO systems," *IEEE Wireless Communications Letters*, pp. 653-656, 2014.

[7] M. Debbah, A. Müller, E. Björnson, and A. Kammoun, "Linear precoding based on polynomial expansion: Large-scale multi-cell MIMO systems," *IEEE Journal of Selected Topics in Signal Processing*, pp. 861-875, 2014.

[8] S. Payami, *Hybrid Beamforming for Massive MIMO Systems*, Guildford: University of Surrey, 2017.

[9] [Online]. Available: www.profheath.org/wp-content/uploads/2016/02/HybridArch.jpg

[10] M. S. Alouini and A. J. Goldsmith, "Capacity of rayleigh fading channels under different adaptive transmission and diversity-combining techniques," *IEEE Transaction on Vehicular Technology*, 1999, pp. 1165–1181.

[11] G. M. A. Sessles and F. K. Jondral, "Low complexity polynomial expansion multiuser detector for CDMA systems," *IEEE Transactions on Vehicular Technology*, 2005, pp. 1379-1391.

Teerapat Sanguankotchakorn was born in Bangkok, Thailand on December 8, 1965. He received the B. Eng. in electrical engineering from Chulalongkorn University, Thailand in 1987, M. Eng. and D. Eng. in information processing from Tokyo Institute of Technology, Japan in 1990 and 1993, respectively.

In 1993, he joined Telecommunication and Information Systems Research Laboratory at Sony Corporation, Japan where he holds two patents on Signal Compression. Since October 1998, he has been with Asian Institute of Technology where he is currently an Associate Professor at Telecommunications Field of Study, School of Engineering and Technology.

Dr. Sanguankotchakorn is a Senior member of IEEE and member of IEICE, Japan.

Ganti V. Sowmya was born in India. She received her Bachelor of Technology in electronics and communications engineering from Jawaharlal Nehru Technological University, Hyderabad, Telangana, India in 2016 and M. Sc. in Information and Communications Technologies from Asian Institute of Technology, Thailand in 2018. Her current research interests are in the area of Wireless Communication, and Network Protocols.