Theoretical performance evaluation of MU-MIMO THP with user scheduling

Karen Taguchi, Ryota Mizutani, Yukiko Shimbo, Hirofumi Suganuma, and Fumiaki Maehara

Abstract: This paper presents the theoretical system-level performance of multi-user multiple-input and multiple-output (MU-MIMO) Tomlinson-Harashima precoding (THP) with user scheduling. In our performance evaluation, proportional fairness (PF), which makes a reasonable compromise between fairness among users and the benefit of multi-user diversity, is implemented as a user scheduling technique, and the effect of modulo loss resulting from THP modulo operation at the receiver is taken into account using mod-Λ channel-based analysis, which provides accurate theoretical performance. Moreover, considering the application to the PF metric, the performance of the mod-Λ channel-based PF metric is compared with that of the traditional Shannon-Hartley theorem-based metric.

Keywords: MU-MIMO, Tomlinson-Harashima precoding (THP), system capacity, mod-Λ channel, proportional fairness (PF)

Classification: Wireless Communication Technologies

References


1 Introduction

With the rapid growth in the use of smart devices, the demand for mobile wireless services has increased exponentially, leading to expectations of higher speed, larger capacity, and lower latency. In 2022, the amount of wireless traffic is estimated to reach 71% of all IP traffic [1], and the fifth-generation mobile communication system (5G) is soon to be commercialized in terms of enhanced mobile broadband (eMBB), ahead of ultra-reliable and low-latency communications (URLLC) and massive machine type communications (mMTC). Multi-user multiple-input and multiple-output (MU-MIMO) is an essential technique for 5G systems because larger capacity can be realized via a single antenna mounted on a mobile station (MS) [2].

Precoding techniques are essential for implementation of MU-MIMO, and are categorized into two approaches: linear precoding (LP) and non-linear precoding (NLP). NLP provides better system capacity than LP because it reduces noise enhancement, and has thus emerged as a candidate technique for 5G systems [3, 4, 5]. Of the various NLP schemes, Tomlinson-Harashima precoding (THP) is considered a practical approach because the perturbation vector can be generated by a simple modulo operation [4, 6, 7].

From a system-level perspective, the combination of MU-MIMO with user scheduling must be taken into account [8, 9, 10]. This is because the capacity of
the entire system strongly depends on how simultaneous users are selected from existing users within a certain cell. Therefore, the performance evaluation is expected to consider the impact of user scheduling as well as MU-MIMO. Computer simulations are considered the surest ways to investigate system performance, but require large computational cost because wireless signal processing, such as modulation and demodulation, needs to be conducted. In this sense, theoretical analysis is considered a powerful tool because mathematical expressions enable us to comprehensively investigate the influence of system parameters on system performance without any time-consuming computer simulations [11].

Considering the above background, we investigate the theoretical system capacity of MU-MIMO THP in terms of user scheduling. The focus of our work is to account for the impact of the modulo loss peculiar to THP, which provides more accurate theoretical performance than our previous work [9] based on the Shannon-Hartley theorem-based approach [6, 7]. Moreover, although the authors in [10] considered the effect of modulo loss for performance evaluation, its effect is given as a constant margin degradation of 0.5 dB, which has left further room for improvement. Thus, in this paper, proportional fairness (PF) [8], which makes a reasonable compromise between fairness among users and the benefit of multi-user diversity, is considered as a user scheduling technique, and the theoretical system performance of MU-MIMO THP with PF is analyzed based on the mod-A channel [12]. Moreover, to clarify the required accuracy of the PF metric, the performance of the mod-A channel-based PF metric is compared with that of the traditional Shannon-Hartley theorem-based metric.

2 System-level performance evaluation of MU-MIMO THP

2.1 Operating principle of MU-MIMO THP

In this section, we briefly introduce the operating principle of MU-MIMO THP with user scheduling. Fig. 1 shows the system configuration, where $N_t$ and $N_r$ denote the number of base station (BS) antennas and MSs with one received antenna element, respectively. In Fig. 1, user scheduling is performed prior to MU-MIMO THP to select the suitable MSs and then, the feedforward (FF) and feedback (FB) filters are determined to retain spatial orthogonality among selected MSs.

In general, THP can be implemented by an LQ decomposition [6, 7], and the channel matrix $H \in \mathbb{C}^{N_r \times N_t}$ can be decomposed as

$$H = LQ,$$

where $L \in \mathbb{C}^{N_t \times N_t}$ and $Q \in \mathbb{C}^{N_r \times N_r}$ are lower triangular and unitary matrices, respectively. Assuming that the precoding weight is determined by the zero-forcing (ZF) criterion, both FF filter $F$ and FB filter $B$ for the THP algorithm can be obtained as

$$G = \text{diag}(L_{11}^{-1}, \cdots, L_{N_rN_t}^{-1}),$$

$$F = Q^H G,$$

$$B = HF - I = LG - I,$$

where $L_{ii}$ is the $i$-th diagonal element of $L$.

In THP, the modulo operation is performed to limit the transmit power increased by the addition of an interference subtraction vector generated by the FB filter $B$. 
Moreover, because the transmit power is changed by the FF filter $F$, a power normalization factor $g$ is required, which is given by

$$g = \sqrt{\frac{\text{tr}(F C_v F^H)}{E_{tx}}},$$

where $E_{tx}$ denotes the total transmit power and $C_v \in \mathbb{C}^{N, N}$ is the covariance matrix of the transmit signal after the modulo operation $v \in \mathbb{C}^N$.

2.2 Mod-$\Lambda$ channel-based analysis for MU-MIMO THP

The system capacity of MU-MIMO can be generally analyzed using the power normalization factor $g$ shown in Eq. (5). This is because this normalization factor indicates the SNR. Therefore, the sum-rate based on the Shannon-Hartley theorem is given by [6, 7]

$$C_{\text{sum}} = \sum_{i=1}^{N} \log_2 \left( 1 + \frac{\sigma_s^2}{g^2 \sigma_n^2} \right) \text{[bps/Hz]},$$

where $\sigma_s^2$ and $\sigma_n^2$ are the transmit signal power and noise power, respectively.

The approach shown in Eq. (6) has the problem that the impact of the modulo loss resulting from the THP modulo operation at the receiver is not taken into account. Therefore, in this paper, we investigate the system-level performance of THP in consideration of the modulo loss. In detail, the impact of the modulo loss is considered using the mod-$\Lambda$ channel-based analysis [12], and it is clarified by comparing it with the traditional Shannon-Hartley-based approach.

The achievable rate of the mod-$\Lambda$ channel is given by

$$C = 2(\log_2 \tau - H(Z_{\text{mod}})) \text{[bps/Hz]},$$

where $\tau$ and $H(Z_{\text{mod}})$ denote the modulo width and differential entropy of the white Gaussian noise (WGN) after the modulo operation. Thus, in order to obtain the differential entropy $H(Z_{\text{mod}})$, it is necessary to derive the probability density function (PDF) of the WGN after the modulo operation $p_{Z_{\text{mod}}}(z_{\text{mod}})$. The PDF of the WGN $p_Z(z)$ is represented by

$$p_Z(z) = \frac{1}{\sqrt{2\pi g^2 \sigma_n^2}} e^{-\frac{z^2}{2g^2 \sigma_n^2}}.$$

The actual impact of the WGN after the modulo operation is represented as the sum of shifted versions of the PDF $p_Z(z)$ in the fundamental region $[-\tau/2, \tau/2]$. The shifts are integral multiples of the modulo width $\tau$. Thus, the PDF of the WGN after the modulo operation $p_{Z_{\text{mod}}}(z_{\text{mod}}) (-\tau/2 < z_{\text{mod}} < \tau/2)$ is given by

![Fig. 1. System configuration of MU-MIMO THP.](image)
\[ p_{Z_{\text{mod}}}(z_{\text{mod}}) = \sum_{k=-\infty}^{\infty} p_z(z_{\text{mod}} + k\tau). \]  

(9)

In consequence, the sum-rate of MU-MIMO THP is represented as

\[ C_{\text{sum}} = \sum_{i=1}^{N_r} 2(\log_2 \tau - H(Z_{\text{mod}})) \]

\[ = \sum_{i=1}^{N_r} 2 \left( \log_2 \tau + \int_{-\tau/2}^{\tau/2} p_{Z_{\text{mod}}}(z_{\text{mod}}) \log_2 p_{Z_{\text{mod}}}(z_{\text{mod}}) dz_{\text{mod}} \right) \text{[bps/Hz].} \]  

(10)

### 2.3 Application of user scheduling to MU-MIMO THP

User scheduling is generally performed before precoding because the number of MSs in a radio zone is more than the number of BS antennas. In this paper, we consider PF [8] as user scheduling and analyze the system-level performance using the above-mentioned mod-$\Lambda$ channel-based approach.

In application of PF to MU-MIMO THP, the system capacity of all possible combinations of MSs has to be calculated because it is used as a criterion in PF-based user selection. The PF metric in the $k$-th combination $M_k$ is given by

\[ M_k = \sum_{i=1}^{N_r} \frac{R_{k,i}(t)}{T_{k,i}(t)} \quad (k = 1, 2, \cdots, k C_N), \]  

(11)

where $K$ is the number of the existing users, and $R_{k,i}(t)$ is the instantaneous system capacity of the $i$-th MS, in the case that the $k$-th combination is admitted to the transmission at time $t$. $T_{k,i}(t)$ is average system capacity of the $i$-th MS in the $k$-th combination until time $t$, which is represented as

\[ T_{k,i}(t + 1) = \left( 1 - \frac{1}{t_c} \right) T_{k,i}(t) + \frac{1}{t_c} R'_{k,i}(t), \]  

(12)

where $t_c$ is the average time range of the system capacity and $R'_{k,i}(t)$ is the instantaneous system capacity of the $i$-th MS at time $t$. Here, $R'_{k,i}(t)$ is zero if the $i$-th MS is not scheduled at time $t$. Eqs. (11) and (12) proves that the PF metric requires the instantaneous system capacity, which is obtained from Eq. (6) or (10).

### 3 Numerical results

In this section, we evaluate system-level performance of MU-MIMO THP with PF user scheduling based on the mod-$\Lambda$ channel-based approach, and compare its performance to that of the traditional Shannon-Hartley theorem-based approach with or without considering modulo loss. Fig. 2 shows the evaluation model and its system parameters. In our performance evaluation, MSs are randomly distributed and the ordering process [7, 12] is adopted to enhance the transmission performance of THP. In addition, the MIMO channel is assumed to follow spatially uncorrelated Rayleigh fading. Moreover, perfect channel state information (CSI) feedback is assumed, and its feedback error and delay are negligible.

Fig. 3 shows the sum-rate versus the number of existing users $K$, where the MIMO antenna configurations are set to be $4 \times 4$ and $6 \times 6$. Fig. 3(a) demonstrates that the performance of the mod-$\Lambda$ channel-based analysis is lower than that of the traditional Shannon-Hartley theorem-based approach, regardless of MIMO antenna
configuration, which indicates that the traditional approach overestimates its performance. To clarify the required accuracy of the PF metric, in Fig. 3(b), the sum-rate is obtained from the mod-Λ channel-based and traditional Shannon-Hartley theorem-based PF metrics. From Fig. 3(b), the sum-rates of both approaches are the same regardless of the MIMO antenna configuration, which implies that the traditional Shannon-Hartley theorem-based approach is only useful for PF metric calculation.

4 Conclusion

In this paper, we presented the exact system-level performance of MU-MIMO THP with PF user scheduling by means of the mod-Λ channel-based analysis. Moreover, we clarified the required accuracy of PF metric by comparing the performance of the mod-Λ channel-based PF metric and the traditional Shannon-Hartley theorem-based metric. Numerical results showed that the system-level performance of mod-Λ channel-based analysis is slightly lower than that of the traditional Shannon-Hartley theorem-based approach, which implies that the traditional approach overestimates its performance. However, in use of the PF metric, the traditional Shannon-Hartley theorem-based approach can be adopted because there is no performance difference between these two approaches.