Partial Perturbation to Alleviate the Performance Degradation of Vector Perturbation With Inaccurate Power Scaling Factors

Yafei Hou, Senior Member, IEEE, and Satoshi Denno, Member, IEEE

Abstract—For multi-user multiple-input-multiple-output (MU-MIMO) system, transmitter utilizes a pre-equalizer based precoding to cancel inter-stream interference for parallel transmission with the aid of accurate feedback of channel state information (CSI) from all receivers. The correct power scaling factor, which normalizes the symbols power of precoding to be a constant one, is required at each receiver. Due to CSI error and limitation of feed-forward link, the received inaccurate power scaling factor will shrink or expand the received constellation points which severely degrades the performance of MIMO system. In this paper, using nonlinear vector perturbation (VP) and linear zero-forcing (ZF) precoding, we analyze the impact of inaccurate power scaling factor on the performance of a MIMO precoding system. The analyzed results show that the mismatched modulo size from inaccurate power scaling factor severely degrades the performance of the VP system, especially for high order M-QAM modulation. In addition, the performance degradation is strongly related with the distribution of representation points for each M-QAM symbol. For linear ZF precoding, the performance loss from the shrinked or expanded constellation points is severe for system using high order M-QAM. In addition, to alleviate the performance degradation, we propose a VP system using partial perturbation points (PPP). By limiting the region of redundant points with vector perturbation for partial M-QAM symbols, the performance degradation due to the inaccurate power scaling factor can be alleviated.

Index Terms—MIMO, precoding, dirty paper coding, performance analysis, inaccurate power scaling factor, vector perturbation.

I. INTRODUCTION

MULTIPLE-INPUT-MULTIPLE-OUTPUT (MIMO) technique is key part of evolving wireless access standards such as WiMAX, WLAN, LTE-Advanced systems [1]. Massive MIMO is an emerging technology that scales up MIMO by possibly orders of magnitude to further dramatically improve spectrum efficiency of systems [2]. Combined with OFDM technique, MIMO and massive MIMO techniques become one of major techniques for next generation wireless communication system, next WLAN system and broadcasting systems [3]. For fifth generation (5G) wireless communication system, the high wireless traffic is assumed as major occurred in indoor environment. Therefore, multi-user MIMO (MU-MIMO) technique is largely applied in the picocell or femtocell system with a small number of user equipments (UEs) [4]. On the other hand, MIMO technology has been largely used for broadcasting system [5] and combined with scalable video coding for TV broadcasting [6].

To achieve a large capacity, MU-MIMO algorithms usually utilize a pre-equalizer to cancel inter-stream interference at the base station (BS) side to support parallel transmission with the aid of accurate feedback of channel state information (CSI) from all UEs. Major linear methods with low computational complexity are zero-forcing (ZF) precoding using channel inversion and minimum-mean square error (MMSE) precoding using regularized channel inversion [7]. However, for the environments such as indoor high correlated channel, the MIMO channel becomes an ill-conditioned one and largely reduces the power efficiency of parallel transmission and the system capacity. To overcome this problem, nonlinear MU-MIMO algorithms using dirty paper coding (DPC) [8] have been proposed. DPC has been utilized in many wireless systems to improve their capacities, especially for the technology of layered division multiplexing (LDM) which has been broadly used for broadcasting system [9], [10].

DPC utilizes a relaxed symbol mapping to represent the original M-quadrature amplitude modulation (QAM) symbols. This relaxed constellation copies the original M-QAM constellation into I-phase and/or Q-phase with a real constant spacing named as “modulo size”. The M-QAM symbol representation in the relaxed constellation is therefore redundant which means that each M-QAM symbol is represented using several constellation points. These redundant points with constant spacing in I-phase and/or Q-phase are perfectly remapped to the original M-QAM symbol using a modulo operator at the receiver if the modulo size is known at the UEs. Therefore, DPC can provide more candidates for each M-QAM symbol for optimizing power efficiency or other targets. Two major nonlinear MU-MIMO precodings named as Tomlinson-Harashima precoding (THP) [11] and vector perturbation (VP) precoding [12] have been proposed. THP utilizes DPC to limit the power range of each transmitted symbol during the pre-equalization process. VP precoding aims to reduce the total power scaling factor to increase power efficiency of system. Compared with THP

Manuscript received November 13, 2019; revised January 22, 2020; accepted February 3, 2020. Date of publication March 12, 2020; date of current version March 4, 2021. (Corresponding author: Yafei Hou.)

The authors are with the Graduate School of Natural Science and Technology, Okayama University, Okayama 700-8530, Japan (e-mail: yfhou@okayama-u.ac.jp; denno@okayama-u.ac.jp).

Digital Object Identifier 10.1109/TBC.2020.2977514

method, VP algorithm is known to be able to largely improve the system capacity of linear MU-MIMO to the capacity bound [13].

To further deploy VP precoding in a more practical manner, several issues must be solved. The first issue is about computational complexity (CC) for finding optimal perturbed vector to minimize the power scaling factor which is an integer-lattice least-square (ILLS) problem [14]. There exists a large number of approximate algorithms [15] which have trade-off the CC and precoding gain. To further reduce the complexity, instead of searching all candidate vectors, a constrained region can be formed to search for the optimal vector [16]–[20]. Another issue is how to obtain the exact power scaling factor to remove the perturbation quantity before the correct symbol demodulation at each UE. In fast fading environment or under a constraint of short-term average transmit power, to get the correct power scaling factor, it causes significant transmission overhead to feed-forward the instantaneous data- and channel-related parameters to the UEs. In addition, it is impractical to get the correct channel status information (CSI) for VP precoding before its transmission due to CSI estimation error and limited feed-forward resources. Therefore, the power scaling factor used at each UE side always includes error. This error shrinks or expands the symbol constellation points [21]. Due to the inaccurate power scaling factor, the modulo operator tends to remap the received symbols into a wrong region owing to noise, especially near the boundary of a constellation. We name this impact as “modulo loss”. The modulo loss largely reduces the system capacity. For example, THP has been shown to suffer from a 4–5 [dB] modulo loss [22].

Several approaches have been proposed to cope with the CSI error and error of power scaling factors. Using MMSE criterion, the authors in [23] showed that the statistical property of CSI error can be treated as one component for precoding calculation, which improves the performance of VP algorithm. Using the rate-distortion theory, References [24], [25] investigated the MMSE-VP precoder design for CSI error with CSI quantization error and showed how the CSI quantization error would affect VP performance. Similar to [25], [26] has proposed a robust VP precoder design which considered the CSI error and power scaling factor jointly with MMSE criterion for improving the VP algorithm. Some proposals aimed to circumvent the need of power scaling factors, References [27] and [28] proposed a VP based precoding scheme that the detection can be perfectly performed without knowing the power scaling factor at the UEs. However, the scheme is only applicable for a QPSK based modulation and cannot support higher order QAM schemes. In [16], a modified VP precoding method named as transmit outage precoding (TOP) has been proposed, where a known power scaling factor is agreed before transmission. However, the predefined power scaling factor lead to a violation of the power constraint referred to as transmit outage. TOP intentionally discards some data which perhaps causes retransmission, and the transmission will drop into outage for a long time when the channel varies slowly. However, obtaining the predefined power scaling factor is impractical when in a fast fading scenario.

In this paper, we analyze the impact of mismatched modulo size from inaccurate power scaling factor on performance of a MU-MIMO system. The major contributions of this paper are summarized as follows.

1. We theoretically analyze the impact of modulo loss on performance of VP based MU-MIMO system. The analyzed results are also applicable for the ZF and MMSE precoding. The analysis shows that the inaccurate power scaling factor severely reduces the performance of the MIMO system, especially for high order M-QAM modulation.

2. The analyzed results show that the performance degradation of precoding system is strongly related with the distribution of representation points for each M-QAM symbol. In addition, the mismatched modulo size has different impact on performance degradation according to positive or negative error of power scaling factor. It also shows that, different with ZF precoding, MMSE precoding can largely shrink the range of power scaling factor which can be used to reduce the impact of inaccurate power scaling factor.

3. To alleviate the performance loss, we propose a VP system using partial perturbation points (PPP). Using a limited region of redundant points, the performance degradation of VP precoding due to the inaccurate power scaling factor can be alleviated.

The paper is organized as followings. The ZF precoding, MMSE precoding and VP precoding, impact of inaccurate power scaling factor and a low-complexity QR-decomposition M-algorithm (QRD-M) are described in Section II. The impact of mismatched modulo size on the MIMO precoding due to the inaccurate power scaling factor is analyzed and simulated in Section III. The scheme of the proposed partial perturbation points (PPP) for the VP system is explained in Section IV and Section V gives the simulated results. The paper ends with conclusions in Section VI.

Notation: Upper (lower) boldface letter denotes matrix (vector), $(.)^H$, $(.)^T$ and $E(.)$ denote conjugate transpose, transpose and expectation operators. We use superscripts $I$ and $Q$ to represent the I- and Q-components of symbol and use $x^2 + \sqrt{\frac{1}{Q}}1^Q$ to represent complex number $x$. The $v_k$ or $v(k)$ is representing the $k$th element of vector $v$. We also use $\text{Prob}(event)$ to represent the probability of event.

II. ALGORITHM OF VECTOR PERTURBATION AND ITS PERTURBATION VECTOR SEARCH ALGORITHM

A. Vector Perturbation

A general model of a downlink MU-MIMO system includes a BS with $N_t$ transmitting antennas and $K$ users, each with $N_r$ receiving antennas. Suppose the MIMO channel is a $KN_r \times N_t$ matrix $H$ with i.i.d. complex element $h_{lk}$. We firstly assume $N_l = KN_r$ and will explain that the similar results will be obtained when considering the case of $N_l \neq KN_r$. We also assume that $s_l = 1$. The vector $n$ is an $N_t \times 1$ AWGN noise vector as $[n_1, \ldots, n_{N_t}]^T$. Then $E[\mathbf{mm}^H]$ equals $\sigma^2 \mathbf{I}$ and $\sigma^2$ is the noise power at each receiving antenna, here $\mathbf{I}$ is the $(N_l \times N_l)$ identity matrix.
Fig. 1. Relaxed constellation points and VP algorithm for MU-MIMO system.

We suppose $x$ is transmitted M-QAM symbols and $W$ is a precoding weight matrix utilized to eliminate the interference among the transmitting streams. The SNR at receiving antenna (SNR per stream) is defined as $P/(N_0\sigma^2)$ with $i$-th M-QAM symbol power as $E[x_i^2] = P$. Supposing the BS has the perfect CSI matrix, the matrix $W$ becomes

$$W = H^H \left( HH^H + \frac{1}{SNR} I \right)^{-1}$$

for MMSE precoding, and

$$W = H^H (HH^H)^{-1}$$

for ZF precoding [7].

The received signal $y$ is given by

$$y = Hu + n + \sqrt{P}HWx + n.$$  

where $y = [y_1, \ldots, y_{KN}]^T$. The transmitted signal $u$ ($u = [u_1, \ldots, u_{KN}]^T$) is represented as $u = \sqrt{P}/\gamma Wx$. The normalization parameter $\gamma = ||Wx||^2$ is used to make the average transmit power be $P$. It should be noted that a small value of $\gamma$ can provide a large SNR.

To reduce $\gamma$, the VP algorithm perturbs the transmitted symbol $x$ with an offset vector called perturbation vector (PV). We use Fig. 1 to show the major idea of VP algorithm for MU-MIMO system with perturbation vector. As shown in Fig. 1(a), let us set $\hat{x} = x + \ell \tau$ where $\tau$ is $2(|c_{\max}| + d_{min}/2)$ [12]. Here $|c_{\max}|$ is the value of constellation point with the largest amplitude and $d_{min}$ is the spacing between two adjacent constellation points. The M-QAM symbol $x_k$ can be represented as

$$x_k = x_k^I + \sqrt{-1}x_k^Q,$$

$$\{x_k^I, x_k^Q\} \in \{s_ι | s_ι = -\tau + (2i - 1)\tau, 2\mu\}$$

$$i = [1, \ldots, \mu].$$

where parameter $\mu$ is 2 for QPSK, 4 for 16-QAM and 8 for 64-QAM, respectively. $\ell \tau$ is a perturbation vector with complex element as $(\ell^I(k) + \sqrt{-1}\ell^Q(k))\tau$, where $\ell^I(k)$ and $\ell^Q(k)$, $(k = 1, \ldots, N_t)$ are integers. This also means that each transmitted M-QAM symbol $x_k$ is represented using many constellation points with a constant spacing $\tau$ in I-phase and/or Q-phase.

As shown in Fig. 1(b), the scalar $\gamma$ is computed as

$$\gamma = ||W(x + \ell \tau)||^2.$$  

Due to this redundant representation, the BS can now choose the best representation of the transmitted M-QAM symbols with the minimum $\gamma$. Therefore, the core design of the VP algorithm is to find the optimal $\ell_{\text{opt}}$ which minimizes $\gamma$, that is

$$\ell_{\text{opt}} = \arg\min_{\ell} ||W(x + \ell \tau)||^2,$$

and then $\gamma$ is computed as

$$\gamma = ||W(x + \ell_{\text{opt}} \tau)||^2.$$  

To make the transmitted signal be a constant transmit power $P$ after the PV search, the transmit signal $u$ is represented as

$$u = \frac{1}{\beta} W(x + \ell_{\text{opt}} \tau).$$

where $\beta$ is power scaling factor and represented as $\sqrt{\gamma}/P$. Then $u$ is transmitted over channel $H$, and the received signal $y$ is

$$y = \frac{1}{\beta} HWx + \frac{1}{\beta} HW\ell_{\text{opt}} \tau + n.$$  

To perfectly remove the effect of perturbation vector, the receiver needs the correct value $\beta$ to multiply the received symbol $y$ as

$$\hat{y} = \beta y = HWx + HW\ell_{\text{opt}} \tau + \beta n.$$  

After that, the receiver utilizes the modulo operator (Mod) on $\hat{y}_k$ for the $k$th stream as

$$f_r(\hat{y}_k) = \hat{y}_k - \left[\frac{\hat{y}_k + \frac{1}{\tau}}{\tau}\right],$$

where the function $[\cdot]$ gives the largest integer which does not exceed its argument. The receiver demodulates its own data from $f_r(\hat{y}_k)$.

From Fig. 1(b) and above explanations, we can find that, if $\ell_{\text{opt}}$ is a zero vector, ZF-VP precoding and MMSE-VP precoding are totally identical to ZF precoding and MMSE precoding. Therefore, VP precoding can be regarded as a general model for analysis.
B. Impact of Inaccurate Power Scaling Factor

As shown in Eq. (8), the power scaling factor $\beta$ is related to the constant transmit power $P$ and $\gamma$. The value $\gamma$ is decided by the symbol $(x + \ell_{\text{opt}}\tau)$ and precoding matrix $W$ related the instantaneous CSI. Therefore, as previous explanation, due to CSI error and limitation of feed-forward resource, UEs cannot get the accurate value of $\beta$.

Let us suppose the received power scaling factor is $\hat{\beta}$ as $(1 + \Delta_1)\beta$ with an error item $\Delta_1\beta$. Therefore, the received symbol $\hat{y}$ in Eq. (10) is represented as

$$\hat{y} = \hat{\beta} y = \frac{(1 + \Delta_1)Wx + HW\ell_{\text{opt}}(1 + \Delta_1)\tau + (1 + \Delta_1)\beta n}{\text{symbol signal}} + \frac{HWx + HW\ell_{\text{opt}}(1 + \Delta_1)\tau + (1 + \Delta_1)\beta n}{\text{PV part}} + \frac{1}{\text{noise}}$$

As shown in Eq. (12), the inaccurate power scaling factor cannot change the ratio of signal power and noise power, but shrinks or expands the symbol constellation and PV before removing perturbation quantity. Due to this impact, the modulo operator tends to remap the received symbols into a wrong region, especially for a high order M-QAM constellation points. We name this impact as “modulo loss”. The modulo loss largely reduces the system capacity.

In addition, for linear ZF and MMSE precoding when $\ell_{\text{opt}}$ is a zero vector, the symbol constellation point will be shrunked or expanded. The linear system has no performance degradation from modulo loss. But, similar to VP system, the shrinked or expanded received constellation points are wrongly demodulated near constellation points from impact of inaccurate power scaling factor.

To reduce the modulo loss, one efficient way is to using more resources of feed-forward link to transmit the correct value $\beta$. The value of $\beta$ is data- and channel- dependent which varies with data in a large dynamic manner with one example as shown in Fig. 3 of this paper, the $\beta$ of 8-by-8 ZF-VP based precoding varies from small value to near 400 for both 16-QAM and 64-QAM modulation. This requires more bits for encoding and transmitting $\beta$ to UE sides for sufficient accuracy. The similar explanation on the limitation of quantization of $\beta$ is also given in [26].

On the other hand, the number of users $K$ may not be a fixed value, it could be a time-dependent value which makes $N_t \neq KN_r$. When $KN_r > N_t$, usually BS selects a set of UEs with good channel condition to meet $N_t = KN_r$. When $N_t > KN_r$, the channel $H$ in Eq. (3) will be a fat matrix where the row number $KN_r$ is smaller than column number $N_t$. Both linear precoding and nonlinear VP algorithms only further reduce the value of $\gamma$ as transmit diversity gain [30] when utilize this fat matrix for precoding calculation. Although it reduces the dynamic range of $\gamma$, the impact of inaccurate power scaling factor still exists which is similar to that of $N_t = KN_r$. Therefore, we use the case of $N_t = KN_r$ to analyze the impact in the whole paper.

Algorithm 1 QRD-M Algorithm

1: input: $R; A; x; \tau; M$
2: init: $X_{set} \leftarrow 0_{N_t \times 1}; cnt = 1; X_{setbk} \leftarrow \{\}$
3: loop 1: for idx $= N_t; 1: 1$
4: $X_{setbk} \leftarrow X_{set}; X_{setbk} \leftarrow \{\}$
5: loop 2: for $i = 1 : 1 : cnt$
6: $V = X_{setbk}[i];$
7: loop 3: for $ij = 1 : 1 : 2T - 1$
8: $V(idx) = x(idx) + A(ij)\tau;$
9: $X_{set} \leftarrow V; cnt++$
end Loop3
end Loop2
10: if $(cnt >= M)$
11: $(k = 1, \ldots, cnt)$
12: $\Gamma^k = RX_{set}(k), p^k = \sum_{j=N_t}^{idx}||J^k(j)||^2;$
13: if $(idx > 1)$
14: Find $M$ smallest $p^k$ with the indexes as $\{k = \{k_1, \ldots, k_M\}\}$
15: $X_{setbk} \leftarrow X_{set}; X_{set} \leftarrow \{\}$
16: $X_{set} \leftarrow X_{setbk}[k]; cnt = M$
else
19: $k = \arg\min_{i \in \{1, \ldots, cnt\}} p^i$
20: end
end Loop1
21: output: $X_{set}(k)$

C. QR Decomposition M Algorithm

The perturbation vector search of Eq. (6) is a $2N_t$-dimensional integer-lattice least-squares problem which can be implemented using a sphere encoder [15]. However, the computational complexity of sphere decoder is large. The QR decomposition M (QRD-M) algorithm [31] is known as an efficient algorithm, which can balance the trade-off between the computational complexity and the performance of the VP system. The QRD-M limits the candidates of $\ell_{opt}(k)$ and $\ell_{opt}(k)$ to $A$ as $A = \{-T, -T + 1, \ldots, 0, \ldots, T - 1, T\}$ in order to reduce the search complexity. In addition, the QRD-M algorithm factorizes $W$ as

$$W = QR.$$

where $Q$ is a unitary matrix and $R$ is an upper triangular matrix. Thus, the search problem in Eq. (6) is simplified to

$$\ell_{opt} = \arg\min_{\ell \in A^{N_t}} ||QR(x + \ell\tau)||^2$$

and $\gamma$ is represented as

$$\gamma = ||R(x + \ell_{opt}\tau)||^2.$$
M-QAM symbols are distributed into a small set $A$ as $T = 1$. Therefore, the QRD-M algorithm can find a near-optimal $\ell_{opt}$ and reduce the computational complexity of sphere decoder if $T$ is set as an appropriate value ($T > 0$) with a large $M$.

III. ANALYSIS OF PRECODING USING VECTOR PERTURBATION AND CHANNEL INVERSION WITH INACCURATE POWER SCALING FACTOR

As shown in the previous section, due to the inaccurate power scaling factor, the modulo operator tends to remap the received symbols into a wrong region, especially for a high order M-QAM constellation points. In this section, we give a theoretical analysis on the symbol error ratio (SER) performance of vector perturbation with the mismatched modulo size to show how the modulo operator with a mismatched modulo size degrades the VP system performance.

A. Analysis of Vector Perturbation With Inaccurate Power Scaling Factor

To make the analysis tractable, we use a simple analysis model. We assume the matrix $W$ is the ZF weight which makes $HW$ equal $I$. If the value of $\sqrt{\gamma/P}$ is known at the receivers, the correct modulo size $\tau$ can be obtained at the receiver. However we suppose that the mismatched modulo size is $\tau'$ as $\tau' = \tau + \Delta \tau$ and $\Delta$ is an offset from Eq. (12) as $(1 + x_i/\ell_{opt}(i))$. The item $\{x_i/\ell_{opt}(i)\}$ is a small value when $x_i$ is an interior point from high-order M-QAM constellation. We also assume the average transmit power $P$ is 1 without loss of generality of analysis.

After multiplication by $\sqrt{\gamma}$, the received signal $\hat{y}$ of Eq. (10) is represented as

$$\hat{y} = t + \sqrt{\gamma}n = (x + \ell_{opt}\tau) + \sqrt{\gamma}n.$$  

Here, a perturbed symbol $t_k$ as $x_k + \ell_{opt}(k)\tau$ can be represented as

$$t'_k = x'_k + \ell'_{opt}(k)\tau,$$

$$\ell'_{opt} = x'_k + \ell'_{opt}(k)\tau.$$  

We use $\text{CASE}(i,j,p,q)$ to represent the event that $\{x'_k = s_i, \ell'_{opt}(k) = p, \ell'_{opt}(k) = q\}$ for the $k$th stream. The integer values $\{p, q\}$ are independently limited into $\{-T, \ldots, -1, 0, 1, \ldots, T\}$ according to the QRD-M algorithm. Therefore, the received symbol is independent in I-phase and/or Q-phase. The probability distribution function (PDF) of $\gamma$ is denoted by $f(\gamma)$. For each specific value $\gamma$, there exists a distribution as $\text{Prob}[\text{CASE}(i,j,p,q)|\gamma]$ for each transmitted constellation point $\text{CASE}(i,j,p,q)$. The received signal $\hat{y}_k$ for the $k$th stream can be computed using Eq. (16) and its PDF can be represented as

$$g(\hat{y}_k|\gamma) = \frac{1}{\pi \gamma \sigma^2} e^{-\frac{(\lambda_1 - (\gamma + \sigma^2))}{\gamma \sigma^2}}.$$  

The SER is computed as $(1 - \text{Prob}(\hat{x}_k = x_k))$ or

$$\text{SER} = 1 - \int_{\gamma} f(\gamma) \text{Prob}[\text{CASE}(i,j,p,q)|\gamma] \times \lambda f(\gamma) d\gamma.$$  

Here we use $\text{Prob}$ to represent $\text{Prob}\{\hat{x}_k = s_i, \ell'_{opt}(k) = s_j|\text{CASE}(i,j,p,q)\}$.

$$\text{Prob} = \int_{\gamma} f(\gamma) g(\hat{y}_k|\gamma) d\gamma.$$  

The SER is computed as $(1 - \text{Prob}(\hat{x}_k = x_k))$ or

$$\text{SER} = 1 - \int_{\gamma} f(\gamma) \text{Prob}[\text{CASE}(i,j,p,q)|\gamma] \times \lambda f(\gamma) d\gamma.$$  

Here we use $\text{Prob}$ to represent $\text{Prob}\{\hat{x}_k = s_i, \ell'_{opt}(k) = s_j|\text{CASE}(i,j,p,q)\}$. It can be computed as

$$\text{Prob} = \int_{\gamma} f(\gamma) g(\hat{y}_k|\gamma) d\gamma.$$  

The SER is computed as $(1 - \text{Prob}(\hat{x}_k = x_k))$ or

$$\text{SER} = 1 - \int_{\gamma} f(\gamma) \text{Prob}[\text{CASE}(i,j,p,q)|\gamma] \times \lambda f(\gamma) d\gamma.$$  

Here we use $\text{Prob}$ to represent $\text{Prob}\{\hat{x}_k = s_i, \ell'_{opt}(k) = s_j|\text{CASE}(i,j,p,q)\}$. It can be computed as

$$\text{Prob} = \int_{\gamma} f(\gamma) g(\hat{y}_k|\gamma) d\gamma.$$  

The mismatched $\tau'$ changes regions for correct decision which largely changes the SER performance. As an example, Fig. 2 shows the region of correct decision for one QPSK constellation point $(s_1, s_2)$ at the receiver and $(c^I_1, c^Q_1)$ with $(\text{integer } c^I_1, c^Q_1 \in [-\infty, \infty])$ is the spacing of region of correct decision in I-phase and/or Q-phase for modulo operator with the mismatched modulo size $\tau'$. Therefore, for each specific $\text{CASE}(i,j,p,q)$ and $\gamma$, Eq. (20) provides a relationship between the correct modulo size $\tau$ and mismatched modulo size $\tau'$.

The mismatched $\tau'$ changes regions for correct decision which largely changes the SER performance. As an example, Fig. 2 shows the region of correct decision for one QPSK constellation point $(s_1, s_2)$ at the receiver and $(c^I_1, c^Q_1)$ with $(\text{integer } c^I_1, c^Q_1 \in [-\infty, \infty])$ is the spacing of region of correct decision in I-phase and/or Q-phase for modulo operator with the mismatched modulo size $\tau'$. Therefore, for each specific $\text{CASE}(i,j,p,q)$ and $\gamma$, Eq. (20) provides a relationship between the correct modulo size $\tau$ and mismatched modulo size $\tau'$.

To further analyze the relationship between the value of $\text{Prob}$ and the region of $R_{ij}$, especially if constellation points are at the boundary, and makes the copying region in I-phase and/or Q-phase using a small spacing $\tau'$. As shown in Fig. 2 (b), $R_{1,2}$ is reduced and the $\text{Prob}$ of $R_{1,2}$ is computed as

$$\text{Prob} = \int_{\gamma} f(\gamma) g(\hat{y}_k|\gamma) d\gamma.$$  

which is only related with $f(\gamma)$.

However, if $\tau' \neq \tau$ as shown in Fig. 2 (b) and (c), the transmitted redundant points, which locate in a different location of each region of correct decision, contribute with a different value of SER. In other words, the value of $\text{Prob}$ is highly influenced by $\text{Prob}[\text{CASE}(i,j,p,q)|\gamma]$.

To further analyze the relationship between the value of $\text{Prob}$ and the region of $R_{ij}$, especially if constellation points are at the boundary, and makes the copying region in I-phase and/or Q-phase using a small spacing $\tau'$. As shown in Fig. 2 (b), $R_{1,2}$ is reduced and the $\text{Prob}$ of $R_{1,2}$ is computed as

$$\text{Prob} = \int_{\gamma} f(\gamma) g(\hat{y}_k|\gamma) d\gamma.$$  

which is only related with $f(\gamma)$.

However, if $\tau' \neq \tau$ as shown in Fig. 2 (b) and (c), the transmitted redundant points, which locate in a different location of each region of correct decision, contribute with a different value of SER. In other words, the value of $\text{Prob}$ is highly influenced by $\text{Prob}[\text{CASE}(i,j,p,q)|\gamma]$.

To further analyze the relationship between the value of $\text{Prob}$ and the region of $R_{ij}$, especially if constellation points are at the boundary, and makes the copying region in I-phase and/or Q-phase using a small spacing $\tau'$. As shown in Fig. 2 (b), $R_{1,2}$ is reduced and the $\text{Prob}$ of $R_{1,2}$ is computed as

$$\text{Prob} = \int_{\gamma} f(\gamma) g(\hat{y}_k|\gamma) d\gamma.$$  

which is only related with $f(\gamma)$.
ble especially for external constellation points generated by QAM. Therefore, for the same level of mismatched modulo
high order QAM is smaller than that of using low order
rability points than that of low order QAM. When the transmit
\begin{align}
\gamma(\gamma) & = (1 + \Delta) N \beta n \tag{26} \\
\text{symbol signal} & + (1 + \Delta) \beta n \tag{26}
\end{align}
noise
The inaccurate power scaling factor as \( \Delta \) will also shrink or expand the received constellation points. If we make \( \tau' = (1 + \Delta) \tau \), the analysis process is similar to previous subsection with \((p, q) = (0, 0)\). Therefore, the analysis of SER performance of the ZF based precoding with inaccurate power scaling factor can be calculated from Eqs. (18)-(25).

In [7], for M-QAM modulation, the distribution of \( \gamma \) for ZF based precoding \( f_{ZF}(\gamma) \) and then \( \text{Prob} \{ \text{CASE}(i, j, p, q)|\gamma \} \) are given as
\begin{align}
f_{ZF}(\gamma) &= \frac{N \gamma^{(N-1)}}{(1 + \gamma)^N (N_r + 1)} \tag{27} \\
\text{Prob} \{ \text{CASE}(i, j, p = 0, q = 0)|\gamma \} &= \frac{1}{\mu_2} \tag{28}
\end{align}
respectively.
Therefore, the SER of ZF based precoding with inaccurate power scaling factor is computed as

$$\text{SER} = 1 - \frac{1}{\mu^2} \int_{\gamma} \sum_{(i,j)} \text{Prob}1 \times f_{ZF}(\gamma) \, d\gamma,$$

where \(\text{Prob}1\) is calculated using the same formulas as Eqs. (22)-(25) but with the \(g(\hat{y}_k, \phi_k|\gamma)\) revised as

$$g(\hat{y}_k, \phi_k|\gamma) = \frac{1}{\pi \gamma^2} \exp \left[ \frac{-(|\hat{y}_k|_2^2 + |\phi_k|_2^2)}{\gamma^2} \right].$$

(C. Simulated Results for Vector Perturbation and ZF Based Precoding With Inaccurate Power Scaling Factor)

To verify these analysis, we simulate an \((8 \times 8)\) ZF-VP based MU-MIMO system. The PV search is realized using QRD-M algorithm with \((T = 3, M = 7)\). We also obtain the \(f(\gamma)\) and \(\text{Prob}[\text{CASE}(i, j, p, q)|\gamma]\) by preliminary simulation. The channel model is independent and identically distributed (i.i.d.) channel and the BS has the perfect CSI. It is noted that there are many wireless channel models for the performance evaluation of OFDM-based system. The simple i.i.d. channel model assumes that the distance between antennas is large enough and no any correlation for each frequency subcarrier. Many precoding and MIMO related research papers have utilized this simple channel model to evaluate the performance of algorithms. Therefore, we also select this channel model for the performance evaluation in this paper.

Fig. 3 shows the simulated results of \(f(\gamma)\) for \((8 \times 8)\) i.i.d. MIMO channel using 16-QAM and 64-QAM modulations with a unit power for each M-QAM symbol. For the case of \((\tau' = \tau)\), the SER performance is computed only using the information of \(f(\gamma)\) as Eq. (21). Fig. 4 (a) and (b) give the simulated results of \(\text{Prob}[\text{CASE}(i, j, p, q)|\gamma]\) for 16-QAM and 64-QAM when \(\gamma\) are 2.5 and 25, respectively. From Fig. 3 and Fig. 4, we can find that, when \(\gamma\) is 2.5, the \(\ell_{\text{opt}}\) is mostly distributed among the \([-1, 0, 1]\). However, some \(\ell_{\text{opt}}\) will appear out of \([-1, 0, 1]\) when \(\gamma\) is 25. In addition, for large \(\gamma\), the quantization error will be enlarged if number of quantization bits cannot be increased. Therefore, more performance degradation due to mismatched modulo size will occur on the VP system with larger \(\gamma\). To further compare MMSE-VP precoding and ZF-VP precoding, we use Fig. 5 to show the \(f(\gamma)\) for \((8 \times 8)\) MMSE-VP precoding over i.i.d. MIMO channel using 16-QAM and 64-QAM modulations. Here we fix the SNR values as 5dB, 15dB and 25dB, respectively. From Fig. 3 and Fig. 5, we can find that MMSE-VP precoding has smaller range of \(\gamma\) than that of ZF-VP precoding because of it’s regularization item using SNR value as Eq. (1). For the feed-forward link, the small range of power scaling factors can simplify the design of quantization method and reduce the quantization error. Therefore, MMSE-VP precoding usually
has better performance than that of ZF-VP precoding when considering the impact of inaccurate power scaling factors.

On the other hand, both Fig. 4 (a) and (b) show that it is difficult to find a clear relationship between the \( \text{Prob} \{ \text{CASE}(i, j, p, q) \} \) and \( \gamma \) which increase the difficulty of finding a specific equation to represent the SER performance if \( (\tau' \neq \tau) \). Therefore, for the case of \( (\tau' \neq \tau) \), the computation of SER requires both simulated \( f(\gamma) \) and simulated \( \text{Prob} \{ \text{CASE}(i, j, p, q) | \gamma \} \).

Fig. 6 (a) and (b) show the analyzed and simulated SER performance using 16-QAM and 64-QAM modulations for the case of \( \Delta \tau \leq 0 \) and \( \Delta \tau > 0 \), respectively. From the results, we can confirm that the analyzed performance matches well with the simulated results for both cases. Fig. 6 (a) and (b) also show that the mismatched modulo offset largely degrades the system performances. For 16-QAM modulation, to achieve a SER of \( 5 \times 10^{-3} \), \( \pm 10\% \) mismatched modulo offset results in over 10 [dB] performance gap compared with the system using the perfect modulo size. For 64-QAM, \( \pm 5\% \) mismatched modulo offset causes the level of the SER floor to be higher than \( 1 \times 10^{-3} \). On the other hand, both figures show that 16QAM modulation is more tolerant of mismatched modulo size than 64-QAM. This is because the spacing between the constellation points of 64-QAM is smaller than that of 16QAM for the identical value \( \tau \).

In addition, as shown in Fig. 6(a) and (b), negative modulo offset has greater impact on the SER performance than a positive one with the identical absolute value \( |\Delta \tau| \). The reason has been explained in the previous section. To further confirm the analysis, we compute the probability that the transmitted constellation point has not been perturbated as \( \sum_{(i,j,p,q)} \text{Prob} \{ \text{CASE}(i, j, p, q) | \gamma \} \) with \( (p, q) = (0, 0) \). Table I showed the results where \( \gamma \) equals 2.5 and 25 for 16-QAM and 64-QAM. As shown in Table I, more than 54\% of the original constellation points in \( x \) are not perturbed. For these transmitted constellation points, the SER performance loss due to \( E_{\text{mis}} \) of Eq. (22) is severer than that of \( E_{\text{exp}} \) of Eq. (24) which supports that negative modulo offset has greater impact on the SER performance than a positive one.

For linear ZF precoding, we showed the analyzed results using Eq. (29) and simulated SER performance of \( 8 \times 8 \) ZF precoding system using 16-QAM and 64-QAM modulations over i.i.d. channel in Fig. 7. The inaccurate power scaling factor is set as \( \pm 10\% \) for both modulations. The simulated SER value are nearly well-fitted to the analyzed results. It is should be noted that the simulated results are slightly better than that of analyzed results. The reason is that the decision ranges of most external constellation points are larger than that of interior constellation points for the simulation process. For theoretical analysis, the decision range is assumed to be the same for every constellation point. This makes the performance slightly difference. On the other hand, similar to the non-linear VP precoding, high-order QAM modulation has larger

<table>
<thead>
<tr>
<th>Modulation</th>
<th>( \gamma )</th>
<th>( (p, q) = (0, 0) )</th>
<th>( (p, q) \in (-1, 0, 1) )</th>
</tr>
</thead>
<tbody>
<tr>
<td>16QAM</td>
<td>2.5</td>
<td>0.8300</td>
<td>0.9997</td>
</tr>
<tr>
<td></td>
<td>25</td>
<td>0.7983</td>
<td>0.9996</td>
</tr>
<tr>
<td>64QAM</td>
<td>2.5</td>
<td>0.6119</td>
<td>0.9996</td>
</tr>
<tr>
<td></td>
<td>25</td>
<td>0.5976</td>
<td>0.9785</td>
</tr>
</tbody>
</table>
performance loss than that of low-order QAM modulation as shown in Fig. 7.

Based on these analyzed and simulated results, we can get following major corollaries.

* Corollary 1: The negative modulo offset has greater impact on performance degradation MIMO system than that of a positive one.

* Corollary 2: The performance degradation is severer for linear and nonlinear precoding using high order QAM than that of using low order QAM modulation, especially when $\ell_{\text{opt}}$ has a large absolute value to increase redundant constellation points for nonlinear precoding.

* Corollary 3: The MMSE based linear precoding and MMSE-VP precoding have more capable advantage to alleviate the performance degradation from inaccurate power scaling factor than that of ZF based linear precoding and ZF-VP precoding.

IV. THE PROPOSED VP WITH PARTIAL PERTURBATION POINTS

From previous analysis, the more redundant constellation points the VP system have, the severer performance degradation due to mismatched modulo size will occur. One efficient way to alleviate the performance degradation is reducing the unnecessary redundant constellation points.

A. Main Idea of Proposed VP With PPP

For each M-QAM modulation, we define the number of value of constellation points in each I-phase or Q-phase as $\mu$ where $\mu$ is 2 for QPSK, 4 for 16-QAM and 8 for 64-QAM, respectively. The perturbed symbol $\ell_{\text{opt}}\tau$ will generates external tiers of constellation points around the core constellation points of M-QAM. We assume there are $L$ external tiers of constellation points with $L = \mu T$. Therefore, the constellation points of perturbed symbol $x + \ell_{\text{opt}}\tau$ can be represented as the original constellation points $x$ and external $L$ tiers of constellation points generated by perturbation process when the $\ell_{\text{opt}}$ are limited into $\{-T, -T + 1, \ldots, 0, \ldots, T - 1, T\}$ and $T$ is a value defined in Section II-C. By increasing the value of external tiers, $L$, from 0 to $\mu T$, the transmitted constellation points are transformed from the original constellation points, which can be regarded as linear precoding system, to $x + \ell_{\text{opt}}\tau$ as a nonlinear VP system. We named this method as a VP system with PPP.

Fig. 8 shows the constellations of a VP system with PPP for QPSK modulation. If the value of $L$ equals $T\mu$, the PPP constellation points are identical to that of the original VP system where $\ell_{\text{opt}}$ are limited into $\{-T, -T + 1, \ldots, 0, \ldots, T - 1, T\}$. Therefore, PPP model can finely control the number of perturbation points by adjusting the value of $L$. On the other hand, Table I provides the probability that the transmitted constellation point is within $\ell_{\text{opt}} \in \{0, 1, -1\}$ in I-phase and/or Q-phase as $\sum_{(i,j,p,q)} \text{Prob}\{\text{CASE}(i,j,p,q)|\gamma_{\text{opt}}\}$ with $(p,q) \in \{-1, 0, 1\}$. From Table I, we can find that more than 97% perturbation vectors $\ell_{\text{opt}}$ are limited into $\{-1, 0, 1\}$. In other words, the value of $L$ can be limited as $2\mu T$.

The pseudocode of proposed VP system with PPP is shown in Algorithm 2. Different with Algorithm 1, the proposed method has different set of perturbation points for different modulated M-QAM constellation points. We use $\mathcal{A}(x)\mathcal{N}(\mathcal{A}(x))$ to represent the perturbation values and their numbers, respectively. For an example, when $x(i) = 1 + j$ is a constellation point of QPSK modulation used for proposed method with $L = 1$, then, from Fig. 8, the $\mathcal{A}(x(i))$ and $\mathcal{N}(\mathcal{A}(x_i))$ will be $\{0; j; -1 + 0j; 1 - j\}$ and 3, respectively. Therefore, the proposed method just changes the Steps 7 and 8 with the varied perturbation points during each iterative process.

As shown in Fig. 2, the constellation points have different SER performance if the received modulo size has an offset. Fig. 9 shows the analyzed SER performance of the VP system.
with PPP with different $L$. The modulo size offset is $-0.1\tau$ for 16-QAM and $-0.05\tau$ for 64-QAM, respectively. We assume the $\gamma_{\text{opt}}$ is fixed as 1 and every constellation point has an identical probability to be transmitted. As shown in Fig. 9, compared with the system with $\Delta\tau = 0$ and $L = 0$, the SER performance is deteriorated by increasing $L$ when $\tau' \neq \tau$.

For both 16-QAM modulation with $-0.1\tau$ offset and 64-QAM modulation with $-0.05\tau$ offset, when $L$ equals $\mu - 1$, the SER performance can still be largely reduced if the SNR value is increased as shown in Fig. 9. However, error floors occur when $L$ is $\mu$.

We use Fig. 10 to further explain the reason that the error floors occur when $L = \mu$ in Fig. 9. As shown in Fig. 10 for 16-QAM where $\mu = 4$, if the position of constellation point $P_4$ is larger than $3\tau'/2$, the received symbol cannot be demodulated correctly and the error floor occurs. Therefore, the condition for that there is no error floor can be obtained as

$$\frac{3\tau'}{2} > \frac{\tau}{2} - \frac{\tau}{2\mu} + \frac{L\tau}{\mu}$$

(31)

for the negative modulo offset in Fig. 10. It can be computed that $L$ should be smaller than 4 for 16-QAM modulation with $-0.1\tau$ offset and be smaller than 8 for 64-QAM modulation with $-0.05\tau$ offset.

In other words, in such case, the VP system still causes error floors even if the original VP limits $\ell_{\text{opt}}$ as $\{-1, 0, 1\}$. The results mean that a small $L$ can alleviate the performance degradation due to the mismatched modulo size. On the other hand, a small $L$ usually increases $\gamma$ because the range of the PV becomes small. However, for the same computational complexity (CC), with a small $L$, the proposed VP system with PPP can increase the number of branches, $M$, of M-algorithm enabling it to find a better perturbation vector which have the least accumulative metrics to reduce $\gamma$. Therefore, the performance of the VP system with PPP needs to be evaluated with different $L$ and $M$.

### B. Computational Complexity of Proposed VP With PPP

The computational complexity (CC) of VP algorithm using QRD-M algorithm is analyzed in [31]. However, it is difficult to get the correct value of CC because there are many methods to realize channel inversion and QR decomposition with different CC. Therefore, we just utilize the order of total multiplication and addition operations to show the CC of VP algorithm. The orders of CC for different matrix operations are shown in Table II.

The CC order of ZF based VP algorithm $C_{\text{ZF-VP}}$ are approximately computed as

$$C_{\text{ZF-VP}} = C_{\text{ZF}} + 8N_3^3 + 32/3N_3^3 + C_{\text{QRD-M}}.$$  (32)

Here $C_{\text{ZF}}$ equals $8N_3^3$ which includes one multiplication operation between two $(N_i \times N_i)$ complex matrices and one operation of $(N_i \times N_i)$ complex matrix inversion operation as shown in Eq. (2). The CC order of PV search based on the QRD-M algorithm using $(N_i \times N_i)$ complex matrix $\mathbf{R}$ is approximately given as

$$C_{\text{QRD-M}} = 2M(2T + 1)(N_i^2 + N_i).$$  (33)

Here $(2T + 1)$ is the number of constellation points per each M-QAM symbol during the iterative process.

Here we provide some simple explanations on the CC of proposed VP with PPP. The PV search for the VP with PPP is also realized using the QRD-M algorithm but the average number of constellation points for each M-QAM symbol is $(1 + 2L/\mu)$. The PV search for the VP with PPP with $(L \leq \mu, M = 7)$ has about 25% computational complexity of the original QRD-M algorithm with $(T = 3, M = 7)$. Therefore the proposed VP system with PPP can achieve almost comparable performance to that of VP system using the QRD-M algorithm with low CC. On the other hand, considering the initial stage for building $M$ survivors, the PV search for the VP system with PPP with $(L \leq \mu, M = 28)$ has almost identical CC to that of original QRD-M algorithm with $(T = 3, M = 7)$.

### V. THE PERFORMANCE OF PROPOSED VP WITH PPP

We simulate an $(8 \times 8)$ ZF-VP and MMSE-VP based MIMO system for evaluation of the proposed VP system with PPP.
The channel model is i.i.d. model and the BS has the perfect CSI. We also simulate the VP system using QRD-M algorithm with \((T = 3, M = 7)\) which is marked as “VP + Original” for comparison.

To compare the performance of the proposed ZF-VP with PPP and original one, we simulated both distribution of \(\gamma\). The \(\gamma\) of ZF-VP with PPP using \((L = 2, M = 28)\) for 16-QAM and \((L = 4, M = 28)\) for 64-QAM are shown in Fig. 11 and compared with the original VP system using the QRD-M algorithm with \((T = 3, M = 7)\). The comparison confirms that the proposed ZF-VP with PPP can realize the smaller \(\gamma\) than that of original VP system using the QRD-M algorithm. The reason is that QRD-M algorithm sometimes just finds a suboptimal \(\ell_{opt}\) with small value of \(M\) during the iterative process. By limiting \(L\) and increasing \(M\), the number of PV search path can be increased and the optimal \(\ell_{opt}\) can be found as explained in [18], [20].

A. The Performance of Proposed ZF-VP and MMSE-VP Using PPP

Fig. 12 (a) and (b) show the simulated BER performance of the proposed ZF-VP with PPP for 16-QAM modulation and 64-QAM, respectively. The modulo size has no offset. For 16-QAM with \((M = 7)\), the number of tiers \(L\) is increased from 1 to 4. The simulated BER performance is improved from \((L = 1)\) to \((L = 4)\). When \((L = 1)\), the VP system with PPP only utilizes the boundary constellation points of 16-QAM to reduce \(\gamma\). As shown in Fig. 11 (a), the performance of the VP with PPP is worse than that of the original QRD-M algorithm with \((T = 3, M = 7)\). On the other hand, the performance improvement of the proposed system from \((L = 2)\) to \((L = 4)\) is limited. When \((L = 4)\), the simulated system is almost identical to that of the VP system using QRD-M algorithm with \((T = 3, M = 7)\). The similar results also appear in Fig. 11(b). When \((L = 8)\), the performance of the proposed VP with PPP is almost identical to that of the original one. On the other hand, the performance degradation of the proposed system by reducing \(L\) from \((L = 8)\) to \((L = 4)\) is slight.

Fig. 12 (a) and (b) also show the simulated BER performance of the proposed VP with PPP for 16-QAM modulation and 64-QAM with \((M = 28)\). The proposed VP with PPP with \((M = 28)\) provides a better performance than that of with \((M = 7)\) for each identical \(L\). In addition, the proposed VP system with PPP with \((L \geq \mu/2, M = 28)\) can achieve better BER performance than that of original VP system using the QRD-M algorithm with \((T = 3, M = 7)\).

To further evaluate the proposed VP system with PPP, we simulate the BER performance of an MMSE-based VP system with PPP. The simulated results are shown in Fig. 13 (a) for 16-QAM modulation and in Fig. 13 (b) for 64-QAM modulation, respectively. The simulated results show that the MMSE-based VP system using PPP has a similar performance characteristic to that of the ZF-based VP system using PPP. On the other hand, it can achieve better BER performance than that of the ZF-based VP system using PPP with identical parameters \(L\) and \(M_1\).

B. The Performance of MMSE Based VP Using PPP With Random Modulo Size Offset

Fig. 14 (a) and (b) show the simulated BER performance of the proposed MMSE based VP with PPP for 16-QAM
modulation and 64-QAM, respectively. The modulo size offset $\Delta \tau$ is Gaussian distributed as $\Delta \tau \sim \mathcal{N}(0, \sigma^2)$. In our simulation, the value of $\sigma^2$ is set as 0.05 and 0.1.

For 16-QAM, the number of tiers $L$ is set as 1, 2 and 4. As shown in Fig. 13 (a), the BER performance of the VP system is dramatically reduced even if the $\sigma^2$ is small as 0.05. Compared with the VP system using original QRD-M algorithm with $(T = 3, M = 7)$, the proposed VP algorithm with PPP with $(L = 1)$ can achieve a high performance improvement. In addition, the proposed system with $(L = 1)$ also achieves better performance than that of $(L = 2)$. For 64-QAM, the number of tiers $L$ is set as 1, 2, 4 and 8. As shown in Fig. 13 (a) and (b), the impact of mismatched modulo size on the performance of 64-QAM is more significant than that of 16-QAM. These simulated results of both figures confirm that the proposed method can alleviate the performance degradation due to the inaccurate power scaling factor by perturbing the partial M-QAM symbols for the VP algorithm with small value $L$.

VI. CONCLUSION

This paper analyzed the impact of inaccurate power scaling factor on the performance degradation of MU-MIMO precoding system. It showed that the mismatched modulo size from inaccurate power scaling factor severely degrades the performance of the VP system, especially for high order M-QAM modulation. The similar impact on performance degradation of linear ZF and MMSE precoding was discussed and show the advantages of MMSE and MMSE-VP precodings. To alleviate the performance degradation, we proposed a VP system with partial perturbation points. Through computer simulation, it was confirmed that the performance degradation due to the inaccurate power scaling factor can be alleviated by reducing the number of redundant points.

REFERENCES


