Article

Hybrid Precoding Design for Wireless Fading Channels

Bangwon Seo [ORCID]

Division of Electrical, Electronics and Control Engineering, The Institute of IT Convergence Technology, Kongju National University, Cheonan 31080, Republic of Korea; seobw@kongju.ac.kr

Abstract: Precoding techniques are widely used to eliminate interference between multiple user signals in a wireless fading channel environment. The linear MMSE precoding technique, one of the most widely used techniques so far, has low computational complexity but has the disadvantage of relatively poor symbol error rate (SER) performance. The symbol-level precoding (SLP) technique, on which much research has been conducted recently, has excellent SER performance but has the disadvantage of being too computationally complex. In this paper, we propose a hybrid precoding technique that simultaneously applies SLP and MMSE precoding to appropriately adjust SER performance and computational complexity performance. If two different types of precoding techniques are applied simultaneously, interference may occur between signals to which different types of precoding are applied, which can significantly deteriorate SER performance. Therefore, in this paper, precoding is designed to prevent interference between the two signals using the null space of the channel matrices. Through computer simulation, the proposed scheme showed that its SER performance was superior to that of the linear MMSE scheme, and the computational complexity was much lower than that of the SLP scheme.

Keywords: symbol-level precoding (SLP); linear MMSE precoding; computation complexity; symbol error rate (SER)

1. Introduction

One of the most important features aimed at by next-generation wireless communication systems is the ability to enable high-speed data transmission with very low latency. The most straightforward and easiest way to increase data transmission rates is to expand the frequency bandwidth used for data transmission. However, until the fourth generation of mobile communication systems, frequencies below 3 GHz were used for mobile communication. Globally, most frequencies below 3 GHz are already allocated for other types of communications, making it almost impossible to use additional bandwidth for next-generation wireless communications. Therefore, starting with fifth-generation mobile communication, techniques using frequency bands near 6 GHz or millimeter-wave (mmWave) frequency bands above 24 GHz have been proposed [1,2]. However, as the frequency of the transmitted signal increases, there is the disadvantage of limited usability, as the transmission range of the radio waves becomes shorter, making it suitable only for cases where the distance between the transmitter and receiver is short.

Among methods to increase data transmission rates without increasing bandwidth, a representative method is to apply multiple-input, multiple-output (MIMO) technology using multiple antennas at the transmitter and receiver. If the transmitting end and the receiving end have multiple antennas, the data transmission speed can be increased by transmitting multiple data streams simultaneously with the same frequency resources using single-user MIMO (SU-MIMO) technology. Meanwhile, even when the receiving end has only one receiving antenna, the data transmission rate of the entire system can be increased by simultaneously transmitting data to multiple users using the same frequency through multi-user MIMO (MU-MIMO) technology [3–5].

In MU-MIMO transmission cases, however, inter-user interference may occur between multiple user signals simultaneously transmitted using the same frequency, and if such interference is not properly removed, the probability of correct symbol detection at the receiving end may severely decrease, and overall communication performance may deteriorate. If the receiving end has multiple antennas, inter-user interference can be reduced to some extent by using well-known multi-user detection techniques [6,7], but if the receiving end has only one receiving antenna, it is impossible to reduce interference at the receiving end. Therefore, in this case, signal processing must be performed at the transmitting end to prevent inter-user interference from occurring, and the most common method among these technologies is to apply precoding to the transmitter [8].

Precoding technologies can be divided into linear precoding and non-linear precoding. Among linear precoding techniques, the most representative methods are zero-forcing (ZF) precoding and minimum mean-squared error (MMSE) precoding techniques [9–11]. ZF precoding technology uses only channel state information and completely eliminates interference signals between users but has the disadvantage of increasing background noise. On the other hand, MMSE precoding technology is a technology that utilizes both channel state information and the statistical characteristics of noise and shows better performance than ZF precoding techniques by appropriately reducing inter-user interference and background noise. Since these linear precoding techniques only use channel state information regardless of the current value of each user's information data, the precoding matrix only needs to be obtained once during the time when the channel state information does not change. Therefore, the longer the coherence time of the channel, the less computational complexity required to obtain the precoding matrix.

Meanwhile, among non-linear precoding techniques, the most frequently mentioned techniques include dirty paper coding (DPC), vector perturbation precoding, and symbol-level precoding (SLP) [12–15]. These methods have excellent SER performance because they use both channel state information and the current information data values of each user to obtain a precoding vector for each symbol interval, but they have the disadvantage of greatly increasing computational complexity. Recently, among non-linear precoding techniques, much research has been conducted on the SLP technique, which has relatively less computational complexity than the DPC technique [16–20]. This technique defines a constructive interference region (CIR) in which inter-user interference helps improve receiver performance and SER performance by designing the interference signal so that the received signal exists in the CIR and is as far away from the decision boundary as possible. However, this technique still has the disadvantage of requiring significantly more calculations than linear precoding techniques because the precoding result must be obtained for each symbol duration. In particular, the SLP technique has the disadvantage that when the number of users increases or the number of transmitter antennas increases, its computational complexity increases significantly, making it difficult to actually implement [21–25].

In this paper, we propose a hybrid precoding technique that simultaneously applies linear precoding and non-linear precoding to appropriately control the SER performance and computational complexity of the entire system. When the number of users for which the base station needs to transmit data is high and the hardware performance of the base station is not satisfactory, it becomes impossible for the base station to apply non-linear precoding to all user data due to computational complexity. In such cases, the proposed technique divides users into two groups, namely User Group 1 and User Group 2. Therefore, to reduce the computational complexity required for applying precoding to the transmitter for the data of users belonging to Group 1, linear precoding is applied, while non-linear precoding is applied to the data of users belonging to Group 2 to enhance SER performance. When applying the existing linear precoding technique and non-linear precoding technique as they are, interference occurs between received signals with different precoding schemes, leading to significantly degraded SER performance.

To prevent interference between received signals when applying two different types of precoding, we propose the orthogonal MMSE (OMMSE) precoding scheme and the orthogonal SLP (OSLP) scheme by utilizing the null space of the channel matrix for each user group. The OMMSE precoding matrix is designed to be orthogonal to the channels of User Group 2, while the OSLP precoding vector is designed to be orthogonal to the channels of User Group 1. This ensures that the received signals for User Group 1 do not contain signals from User Group 2, and vice versa. Considering the drawback of decreased SER performance as the number of users in User Group 1 increases, we can adjust the number of users in User Group 1 while taking into account the performance of the transmitter hardware. Through computer simulation, the proposed hybrid precoding scheme shows better performance than the scheme that applies linear MMSE precoding to all users and has less computational complexity than the scheme that applies SLP to all users. We also show that SER performance and computational complexity can be controlled by adjusting the number of users applying the SLP scheme.

Overall, the contributions of this paper can be summarized as follows:

- For the first time, we addressed the problem of simultaneously applying two types of distinct precoding schemes to transmit data to multiple users. To the best of the authors' knowledge, there is no existing research that simultaneously applies linear precoding and non-linear precoding among the papers published to date.
- We proposed a hybrid precoding scheme utilizing null space information about the channel to prevent interference between the received signals for users applying linear precoding and those applying non-linear precoding. The proposed hybrid precoding consists of OMMSE precoding and the OSLP scheme.
- Using the proposed hybrid precoding scheme, it is possible to adjust the number of users in each user group based on the hardware performance of the base station transmitter. This allows for the adjustment of the overall system's SER performance and computational complexity.

The structure of this paper is as follows. Section 2 describes the system model considered in this paper. Section 3 explains the proposed hybrid precoding technique. Section 4 presents the simulation results, and Section 5 concludes.

Notations: Matrices and vectors are denoted by boldface capital- and lower-case letters, respectively. The notations $(\cdot)^*$, $(\cdot)^T$, and $(\cdot)^H$ represent the complex conjugate, transpose, and Hermitian of a matrix or vector. Also, $\text{tr}(\cdot)$ and $||\cdot||_F$ represent the trace and the Frobenius norm of a matrix, respectively. The notation $E[\cdot]$ denotes the statistical expectation operation.

2. System Model

Since the emergence of orthogonal frequency division multiplexing (OFDM) technology, research on precoding techniques in transmitters with multiple antennas has assumed an OFDM system, even without specific mention. In an OFDM system, independent transmission, reception, and signal processing operations are performed for each subcarrier through fast Fourier transform (FFT) and inverse FFT (IFFT) operations. The time-domain channel between the transmitter and receiver is generally a multipath fading channel in OFDM systems; however, the appropriate selection of OFDM parameters results in a frequency flat-fading channel for each subcarrier. Therefore, the received signal considered in this paper refers to the received signal at each subcarrier, and the channel model applied to the received signal is assumed to be a frequency flat-fading channel.

Figure 1 shows the system model considered in this paper. Assume that the base station is transmitting data to (K + J) users, the number of transmitting antennas at the base station is M, and the number of receiving antennas for each user is one. In cases where there are multiple receiving antennas, the proposed technique can be easily applied by modifying it, so, for simplicity, this paper only considers the case where there is one receiving antenna.
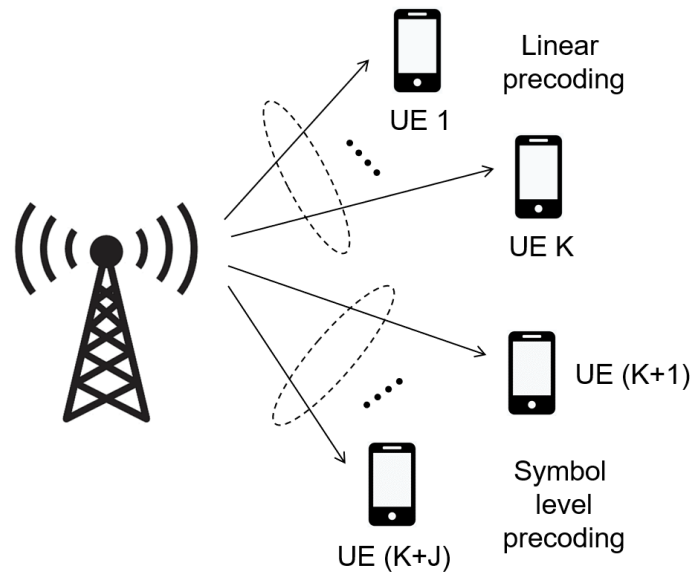
Figure 1. System model considered in this paper, where the base station with multiple transmitting antennas is transmitting data to (K + J) users with one receiving antenna each.

Let the data symbol for the m-th user be s_m, and its mean and variance are $E[s_m] = 0$ and $E[|s_m|^2] = 1$, respectively. If symbol-level-based precoding is applied to all user data at the transmitter, the SER performance of each user will be excellent, but the computational complexity at the transmitter will become too high. Therefore, average-based linear MMSE precoding is applied to the data $\mathbf{s}_1 = [s_1, \ldots, s_K]^T$ for K users, and symbol-based precoding is applied to the data $\mathbf{s}_2 = [s_{K+1}, \ldots, s_{K+J}]^T$ for the remaining J users to implement a trade-off between the SER performance and the computation complexity.

Let us assume that the precoding matrix for linear precoding is $\mathbf{W}_1$ of size M × K and that the data vector obtained through symbol-level precoding is $\mathbf{x}_2 = [x_1, \ldots, x_M]^T \in \mathbb{C}^{M \times 1}$. In this case, the transmission signal, $\mathbf{c}$, can be expressed as follows:

$$\mathbf{c} = \beta_1 \mathbf{W}_1 \mathbf{s}_1 + \beta_2 \mathbf{x}_2, \tag{1}$$

where β_1 represents a parameter for adjusting the transmission power for linear precoding to be P_1, and β_2 represents a parameter for adjusting the transmission power for symbol level precoding to be P_2. Therefore, β_1 and β_2 are given as follows:

$$E[||\beta_1 \mathbf{W}_1 \mathbf{s}_1||^2] = \beta_1^2 \mathrm{tr}(\mathbf{W}_1 \mathbf{W}_1^H) = P_1 \Rightarrow \beta_1 = \sqrt{\frac{P_1}{\mathrm{tr}(\mathbf{W}_1 \mathbf{W}_1^H)}}, \tag{2}$$

$$||\beta_2 \mathbf{x}_2||^2 = P_2 \Rightarrow \beta_2 = \sqrt{\frac{P_2}{||\mathbf{x}_2||^2}}, \tag{3}$$

where the properties $E[\mathbf{s}_1 \mathbf{s}_1^H] = \mathbf{I}_K$ and $\mathrm{tr}[\mathbf{W}_1^H \mathbf{W}_1] = \mathrm{tr}[\mathbf{W}_1 \mathbf{W}_1^H]$ are used. As can be seen in Equation (1), since the base station applies both average-based linear precoding and symbol-level precoding, this technique is called hybrid precoding in this paper.

Assuming that the channel between the base station and the k-th user is $\mathbf{h}_k \in \mathbb{C}^{M \times 1}$, the received signal, y_k, by the k-th user is given as follows:

$$y_k = \mathbf{h}_k^H (\beta_1 \mathbf{W}_1 \mathbf{s}_1 + \beta_2 \mathbf{x}_2) + z_k, \tag{4}$$

where z_k represents additive white Gaussian noise (AWGN) with a mean of 0 and a variance of σ_z^2.

Let us define the group of all users applying average-based linear precoding as group 1 and the group of users applying symbol-based precoding as group 2. Also, if we express the received signals for group 1 and group 2 as $\mathbf{y}_1 = [y_1, \ldots, y_K]^T$ and $\mathbf{y}_2 = [y_{K+1}, \ldots, y_{K+J}]^T$, respectively, $\mathbf{y}_1$ and $\mathbf{y}_2$ can be written as follows:

$$\mathbf{y}_1 = \beta_1 \mathbf{H}_1 \mathbf{W}_1 \mathbf{s}_1 + \beta_2 \mathbf{H}_1 \mathbf{x}_2 + \mathbf{z}_1, \tag{5}$$

$$\mathbf{y}_2 = \beta_1 \mathbf{H}_2 \mathbf{W}_1 \mathbf{s}_1 + \beta_2 \mathbf{H}_2 \mathbf{x}_2 + \mathbf{z}_2, \tag{6}$$

where the noise vectors are $\mathbf{z}_1 = [z_1, \ldots, z_K]^T$ and $\mathbf{z}_2 = [z_{K+1}, \ldots, z_{K+J}]^T$ with sizes of K × 1 and J × 1, respectively, and the channel matrices, $\mathbf{H}_1$ and $\mathbf{H}_2$, of user groups 1 and 2, respectively, are given as follows:

$$\mathbf{H}_1 = [\mathbf{h}_1, \ldots, \mathbf{h}_K]^H \in \mathbb{C}^{K \times M}, \tag{7}$$

$$\mathbf{H}_2 = [\mathbf{h}_{K+1}, \ldots, \mathbf{h}_{K+J}]^H \in \mathbb{C}^{J \times M}. \tag{8}$$

3. Proposed Hybrid Precoding

Precoding for the users in group 1 was designed based on the average of the data symbols during the coherence time of the channel, and precoding for the users in group 2 was designed based on the actual symbol value of each symbol interval. Since these two types of precoding have very different characteristics, it is very difficult to remove the interference between groups at the receiving end if there are two types of signals together at the received signal. Therefore, it is desirable to implement precoding at the base station to prevent interference between the two user groups. In order to satisfy this condition, in this paper, precoding is implemented so that no signal related to $\mathbf{x}_2$ exists in $\mathbf{y}_1$ and no signal related to $\mathbf{s}_1$ exists in $\mathbf{y}_2$. Therefore, we design the linear precoding matrix $\mathbf{W}_1$ and the SLP vector $\mathbf{x}_2$ to satisfy the following conditions:

$$\mathbf{H}_1 \mathbf{x}_2 = \mathbf{0}, \tag{9}$$

$$\mathbf{H}_2 \mathbf{W}_1 = \mathbf{0}, \tag{10}$$

To design $\mathbf{W}_1$ and $\mathbf{x}_2$ such that the conditions in Equations (9) and (10) are satisfied, the base station needs to have knowledge of the channel. Research on channel estimation has been extensively conducted [26], and in this paper, it is assumed that the channel estimation is perfect, with no channel estimation errors.

In order to design $\mathbf{x}_2$ to satisfy the condition in (9), $\mathbf{x}_2$ is designed to exist in the null space of $\mathbf{H}_1$. To design like this, let us assume that the eigenvalue decomposition for the channel matrix $\mathbf{H}_1$ of group 1 is $\mathbf{H}_1 = \mathbf{U}_1 \mathbf{\Sigma}_1 \mathbf{V}_1^H$, where the matrices $\mathbf{U}_1$ of size K × K and $\mathbf{V}_1$ of size M × M are unitary matrices, and the eigenvalue matrix $\mathbf{\Sigma}_1$ is a diagonal matrix of size K × M. Additionally, $\mathbf{U}_1$ and $\mathbf{V}_1$ have the characteristics of $\mathbf{U}_1 \mathbf{U}_1^H = \mathbf{U}_1^H \mathbf{U}_1 = \mathbf{I}_K$ and $\mathbf{V}_1 \mathbf{V}_1^H = \mathbf{V}_1^H \mathbf{V}_1 = \mathbf{I}_M$, where $\mathbf{I}_M$ is an identity matrix of size M × M. If we define $\mathbf{\Sigma}_1 = \begin{bmatrix} \mathbf{\Sigma}_{11} & \mathbf{0}_{K \times (M-K)} \end{bmatrix}$ and $\mathbf{V}_1 = \begin{bmatrix} \mathbf{V}_{11} & \mathbf{V}_{12} \end{bmatrix}$, the channel matrix can be expressed as $\mathbf{H}_1 = \mathbf{U}_1 \mathbf{\Sigma}_{11} \mathbf{V}_{11}^H$, where the sizes of matrices $\mathbf{\Sigma}_{11}$, $\mathbf{V}_{11}$, and $\mathbf{V}_{12}$ are K × K, M × (M-K), and M × K, respectively. Now that we have $\mathbf{H}_1 \mathbf{x}_2 = \mathbf{U}_1 \mathbf{\Sigma}_{11} \mathbf{V}_{11}^H \mathbf{x}_2 = \mathbf{0}$, to satisfy Equation (9), we can write $\mathbf{x}_2 = \mathbf{V}_{12} \mathbf{f}_2$, $\mathbf{f}_2 \in \mathbb{C}^{(M-K) \times 1}$. Thus, designing $\mathbf{x}_2$ is ultimately the same as designing $\mathbf{f}_2$, and how to design the vector $\mathbf{f}_2$ will be explained in detail later.

Also, in order to design $\mathbf{W}_1$ to satisfy Equation (10), let us assume that the eigen-value decomposition for the channel matrix, $\mathbf{H}_2$, of group 2 is $\mathbf{H}_2 = \mathbf{U}_2 \mathbf{\Sigma}_2 \mathbf{V}_2^H$, where the sizes of matrices $\mathbf{U}_2$, $\mathbf{V}_2$, and $\mathbf{\Sigma}_2$ are J × J, M × M, and J × M, respectively. Similar to the method used for $\mathbf{H}_1$, if we define $\mathbf{\Sigma}_2 = \begin{bmatrix} \mathbf{\Sigma}_{21} & \mathbf{0}_{J \times (M-J)} \end{bmatrix}$, $\mathbf{V}_2 = \begin{bmatrix} \mathbf{V}_{21} & \mathbf{V}_{22} \end{bmatrix}$, $\mathbf{\Sigma}_{21} \in \mathbb{C}^{J \times J}$, $\mathbf{V}_{21} \in \mathbb{C}^{M \times J}$, and $\mathbf{V}_{22} \in \mathbb{C}^{M \times (M-J)}$, $\mathbf{H}_2$ can be expressed as $\mathbf{H}_2 = \mathbf{U}_2 \mathbf{\Sigma}_{21} \mathbf{V}_{21}^H$. Since $\mathbf{H}_2 \mathbf{W}_1 = \mathbf{U}_2 \mathbf{\Sigma}_{21} \mathbf{V}_{21}^H \mathbf{W}_1 = \mathbf{0}$, $\mathbf{W}_1$ can be designed as $\mathbf{W}_1 = \mathbf{V}_{22} \mathbf{G}_1$, $\mathbf{G}_1 \in \mathbb{C}^{(M-J) \times K}$ to satisfy

Equation (10). Therefore, designing $\mathbf{W}_1$ is the same as designing $\mathbf{G}_1$, and the method for designing the matrix $\mathbf{G}_1$ will be explained in detail later.

In summary, designing the linear precoding matrix, $\mathbf{W}_1$, and the symbol-level precoding vector, $\mathbf{x}_2$, is the same as designing $\mathbf{G}_1$ and $\mathbf{f}_2$, respectively. In this paper, $\mathbf{G}_1$ is referred to as the linear orthogonal MMSE (OMMSE) matrix, and $\mathbf{f}_2$ is referred to as the non-linear orthogonal SLP (OSLP) vector.

3.1. Linear OMMSE Precoder Design

Here, we explain how to design the average-based linear precoding matrix, $\mathbf{G}_1$. Since $\mathbf{H}_1\mathbf{x}_2 = \mathbf{0}$, $\mathbf{y}_1$ can be simply written as follows:

$$\mathbf{y}_1 = \beta_1 \mathbf{H}_1 \mathbf{W}_1 \mathbf{s}_1 + \mathbf{z}_1 = \beta_1 \mathbf{H}_1 \mathbf{V}_{22} \mathbf{G}_1 \mathbf{s}_1 + \mathbf{z}_1. \tag{11}$$

If we define the equalizer for the received signal in group 1 as $\mathbf{A}_1 = \mathrm{diag}(a_{1,1}, \dots, a_{1,K})$, the estimate $\hat{\mathbf{s}}_1$ for $\mathbf{s}_1$ can be obtained as follows:

$$\hat{\mathbf{s}}_1 = \frac{\mathbf{A}_1}{\beta_1} \mathbf{y}_1 = \mathbf{A}_1 \mathbf{H}_1 \mathbf{V}_{22} \mathbf{G}_1 \mathbf{s}_1 + \frac{\mathbf{A}_1}{\beta_1} \mathbf{z}_1. \tag{12}$$

The mean-squared error (MSE), J_1, for group 1 can be defined as follows:

$$J_1 = E_{\mathbf{s}_1, \mathbf{z}_1} \left[\left\| \frac{1}{\beta_1} \mathbf{A}_1 \mathbf{y}_1 - \mathbf{s}_1 \right\|^2 \right] = E_{\mathbf{s}_1, \mathbf{z}_1} \left[\left\| \mathbf{A}_1 \mathbf{H}_1 \mathbf{V}_{22} \mathbf{G}_1 \mathbf{s}_1 + \frac{1}{\beta_1} \mathbf{A}_1 \mathbf{z}_1 - \mathbf{s}_1 \right\|^2 \right]. \tag{13}$$

If we calculate the expectation for $\mathbf{s}_1$ and $\mathbf{z}_1$ in this equation, J_1 can be rewritten as follows:

$$J_1 = \mathrm{tr}\left\{ (\mathbf{A}_1 \mathbf{H}_1 \mathbf{V}_{22} \mathbf{G}_1 - \mathbf{I}_K)(\mathbf{A}_1 \mathbf{H}_1 \mathbf{V}_{22} \mathbf{G}_1 - \mathbf{I}_K)^H \right\} + \frac{\sigma_z^2}{\beta_1^2} \mathrm{tr}(\mathbf{A}_1 \mathbf{A}_1^H). \tag{14}$$

Meanwhile, by defining $\mathbf{Q}_1 = \mathbf{A}_1 \mathbf{H}_1 \mathbf{V}_{22}$ and using the relationship $\beta_1^2 = \frac{P_1}{\mathrm{tr}(\mathbf{W}_1 \mathbf{W}_1^H)} = \frac{P_1}{\mathrm{tr}(\mathbf{G}_1 \mathbf{G}_1^H)}$, J_1 can be expressed as follows:

$$J_1 = \mathrm{tr}\left\{ (\mathbf{Q}_1 \mathbf{G}_1 - \mathbf{I}_K)(\mathbf{Q}_1 \mathbf{G}_1 - \mathbf{I}_K)^H \right\} + \frac{\sigma_z^2}{P_1} \mathrm{tr}(\mathbf{A}_1 \mathbf{A}_1^H) \, \mathrm{tr}(\mathbf{G}_1 \mathbf{G}_1^H). \tag{15}$$

Now, if we design the linear precoding matrix, $\mathbf{G}_1$, to minimize J_1, the optimization problem can be expressed as follows:

$$\min_{\mathbf{G}_1} J_1 = \min_{\mathbf{G}_1} \, \mathrm{tr}\left\{ (\mathbf{Q}_1 \mathbf{G}_1 - \mathbf{I}_K)(\mathbf{Q}_1 \mathbf{G}_1 - \mathbf{I}_K)^H \right\} + \frac{\sigma_z^2}{P_1} \mathrm{tr}(\mathbf{A}_1 \mathbf{A}_1^H) \, \mathrm{tr}(\mathbf{G}_1 \mathbf{G}_1^H). \tag{16}$$

In order to find a solution to this optimization problem, the differentiation of J_1 by $\mathbf{G}_1^*$ is designed to be $\mathbf{0}$, where $\mathbf{G}_1^*$ represents the complex conjugate of $\mathbf{G}_1$.

$$\begin{aligned}
\frac{\partial}{\partial \mathbf{G}_1^*} J_1 &= \mathbf{Q}_1^H (\mathbf{Q}_1 \mathbf{G}_1 - \mathbf{I}_K) + \frac{\sigma_z^2}{P_1} \mathrm{tr}(\mathbf{A}_1 \mathbf{A}_1^H) \, \mathbf{G}_1 \\
&= \left\{ \mathbf{Q}_1^H \mathbf{Q}_1 + \frac{\sigma_z^2}{P_1} \mathrm{tr}(\mathbf{A}_1 \mathbf{A}_1^H) \mathbf{I}_K \right\} \mathbf{G}_1 - \mathbf{Q}_1^H = \mathbf{0}.
\end{aligned} \tag{17}$$

By solving this equation, the solution to the optimization problem can be obtained as follows:

$$\mathbf{G}_{1,\mathrm{opt}} = \left\{ \mathbf{Q}_1^H \mathbf{Q}_1 + \frac{\sigma_z^2}{P_1} \mathrm{tr}(\mathbf{A}_1 \mathbf{A}_1^H) \mathbf{I}_K \right\}^{-1} \mathbf{Q}_1^H. \tag{18}$$

3.2. Non-Linear OSLP Design

Here, we explain how to design the symbol-level precoding vector, $\mathbf{f}_2$. Since $\mathbf{H}_2\mathbf{W}_1 = \mathbf{0}$, the received signal vector, $\mathbf{y}_2$, for group 2 can be simply written as follows:

$$\mathbf{y}_2 = \beta_2\mathbf{H}_2\mathbf{x}_2 + \mathbf{z}_2 = \beta_2\mathbf{H}_2\mathbf{V}_{12}\mathbf{f}_2 + \mathbf{z}_2. \tag{19}$$

If we define the equalizer for the received signal in group 2 as $\mathbf{A}_2 = \mathrm{diag}(a_{2,1}, \dots, a_{2,J})$, the estimate $\hat{\mathbf{s}}_2$ for $\mathbf{s}_2$ can be obtained as follows:

$$\hat{\mathbf{s}}_2 = \frac{1}{\beta_2}\mathbf{A}_2\mathbf{y}_2 = \mathbf{A}_2\mathbf{H}_2\mathbf{V}_{12}\mathbf{f}_2 + \frac{1}{\beta_2}\mathbf{A}_2\mathbf{z}_2. \tag{20}$$

If we represent the vector obtained by moving $\mathbf{s}_2$ to the CIR as $\tilde{\mathbf{s}}_2$, $\tilde{\mathbf{s}}_2$ can be expressed as follows:

$$\tilde{\mathbf{s}}_2 = \mathbf{s}_2 + \begin{bmatrix} \theta_{\alpha,1}r_{\alpha,1} + \theta_{\mu,1}r_{\mu,1} \\ \vdots \\ \theta_{\alpha,J}r_{\alpha,J} + \theta_{\mu,J}r_{\mu,J} \end{bmatrix} = \mathbf{s}_2 + \mathbf{R}\,\boldsymbol{\theta}, \tag{21}$$

where $r_{\alpha,j}$ and $r_{\mu,j}$ are parameters representing the decision boundary for the CIR of $\mathbf{s}_2$, and $\theta_{\alpha,j}$ and $\theta_{\mu,j}$ represent the weights for $r_{\alpha,j}$ and $r_{\mu,j}$ [10–12]. Here, decision boundary parameters can be easily obtained once the constellation of the transmission symbol is determined, and weight parameters must be optimized. The CIR matrix, $\mathbf{R}$, and weight vector, $\boldsymbol{\theta}$, are defined as follows:

$$\mathbf{R} = \begin{bmatrix} r_{\alpha,1} & 0 & & \cdots & 0 & r_{\mu,1} & 0 & & \cdots & 0 \\ & & \ddots & & & & & \ddots & & \\ 0 & \cdots & & 0 & r_{\alpha,J} & 0 & \cdots & & 0 & r_{\mu,J} \end{bmatrix}, \tag{22}$$

$$\boldsymbol{\theta} = [\theta_{\alpha,1} \cdots \theta_{\alpha,J} \ \theta_{\mu,1} \cdots \theta_{\mu,J}]^T. \tag{23}$$

Now, the MSE, J_2, for group 2 is defined as follows:

$$J_2 = E_{\mathbf{z}_2}\left[\left\| \frac{1}{\beta_2}\mathbf{A}_2\mathbf{y}_2 - \tilde{\mathbf{s}}_2 \right\|^2\right] = E_{\mathbf{z}_2}\left[\left\| \mathbf{A}_2\mathbf{H}_2\mathbf{V}_{12}\mathbf{f}_2 + \frac{1}{\beta_2}\mathbf{A}_2\mathbf{z}_2 - \tilde{\mathbf{s}}_2 \right\|^2\right]. \tag{24}$$

If we calculate the expectation for $\mathbf{z}_2$ in this equation, J_2 can be expressed as follows:

$$\begin{aligned} J_2 &= \|\mathbf{A}_2\mathbf{H}_2\mathbf{V}_{12}\mathbf{f}_2 - \tilde{\mathbf{s}}_2\|^2 + \frac{1}{\beta_2^2}E\left[\mathrm{tr}\left(\mathbf{A}_2\mathbf{z}_2\mathbf{z}_2^H\mathbf{A}_2^H\right)\right] \\ &= \|\mathbf{A}_2\mathbf{H}_2\mathbf{V}_{12}\mathbf{f}_2 - \tilde{\mathbf{s}}_2\|^2 + \frac{\sigma_z^2}{\beta_2^2}\mathrm{tr}\left(\mathbf{A}_2\mathbf{A}_2^H\right). \end{aligned} \tag{25}$$

By substituting $\beta_2 = \sqrt{\frac{P_2}{\|\mathbf{x}_2\|^2}} = \sqrt{\frac{P_2}{\|\mathbf{V}_{12}\mathbf{f}_2\|^2}} = \sqrt{\frac{P_2}{\|\mathbf{f}_2\|^2}}$ in this equation and using the relationship $\mathbf{V}_{12}^H\mathbf{V}_{12} = \mathbf{I}_K$, the following equation can be obtained:

$$J_2 = \|\mathbf{A}_2\mathbf{H}_2\mathbf{V}_{12}\mathbf{f}_2 - \tilde{\mathbf{s}}_2\|^2 + \frac{\sigma_z^2}{P_2}\mathrm{tr}\left(\mathbf{A}_2\mathbf{A}_2^H\right)\mathbf{f}_2^H\mathbf{f}_2. \tag{26}$$

To simplify the equation, if we define $\mathbf{B} \triangleq \mathbf{A}_2\mathbf{H}_2\mathbf{V}_{12}$ and $\gamma = \mathrm{tr}\left(\mathbf{A}_2\mathbf{A}_2^H\right)$, J_2 can be expressed as follows:

$$J_2 = \|\mathbf{B}\mathbf{f}_2 - \tilde{\mathbf{s}}_2\|^2 + \frac{\sigma_z^2}{P_2}\gamma\,\mathbf{f}_2^H\mathbf{f}_2. \tag{27}$$

Now, if we design the precoding vector, $\mathbf{f}_2$, and the CIR weight vector, θ, to minimize J_2, the optimization problem can be expressed as follows:

$$\min_{\mathbf{f}_2,\,\theta} \quad \|\mathbf{B}\mathbf{f}_2 - \tilde{\mathbf{s}}_2\|^2 + \frac{\sigma_z^2}{P_2}\gamma\,\mathbf{f}_2^H\mathbf{f}_2. \tag{28}$$

To find a solution to this optimization problem, we first design $\mathbf{f}_2$ so that the differentiation of J_2 by $\mathbf{f}_2^*$ is $\mathbf{0}$:

$$\frac{\partial}{\partial\,\mathbf{f}_2^*}J_2 = \mathbf{B}^H\mathbf{B}\mathbf{f}_2 - \mathbf{B}^H(\mathbf{s}_2 + \mathbf{R}\theta) + \frac{\sigma_z^2\gamma}{P_2}\mathbf{f}_2 = \mathbf{0}. \tag{29}$$

The solution to this equation can be obtained as follows:

$$\mathbf{f}_{2,\mathrm{opt}} = \left(\mathbf{B}^H\mathbf{B} + \frac{\sigma_z^2\gamma}{P_2}\mathbf{I}_{K\times K}\right)^{-1}\mathbf{B}^H(\mathbf{s}_2 + \mathbf{R}\theta). \tag{30}$$

Meanwhile, using only the parts related to θ in J_2, another optimization problem to find the optimal solution for θ can be expressed as follows:

$$\min_{\theta\,\geq\,0} \quad \|\mathbf{B}\,\mathbf{f}_2 - (\mathbf{s}_2 + \mathbf{R}\theta)\|^2. \tag{31}$$

After substituting $\mathbf{f}_{2,\mathrm{opt}}$ into this equation and then manipulating the equation, it can be written as follows:

$$\min_{\theta} \quad \left\|\left\{\mathbf{B}\left(\mathbf{B}^H\mathbf{B} + \frac{\sigma_z^2\gamma}{P_2}\mathbf{I}_{K\times K}\right)^{-1}\mathbf{B}^H - \mathbf{I}_{J\times J}\right\}(\mathbf{s}_2 + \mathbf{R}\theta)\right\|^2. \tag{32}$$

In this equation, if we define $\mathbf{C} \equiv \mathbf{B}(\mathbf{B}^H\mathbf{B} + \frac{\sigma_z^2\gamma}{P_2}\mathbf{I}_K)^{-1}\mathbf{B}^H - \mathbf{I}_J$ and $\mathbf{d} \equiv -\mathbf{C}\mathbf{s}_2$, Equation (32) can be simply expressed as follows:

$$\min_{\theta\,\geq\,\mathbf{0}} \quad \|\mathbf{C}\mathbf{R}\theta - \mathbf{d}\|^2. \tag{33}$$

Since this problem is a non-negative least-squares problem, θ can be easily obtained using the optimization function *lsqnonneg* in the MATLAB 2019 program.

4. Simulation Results

Computer simulations were performed to compare the performance of the proposed hybrid precoding scheme to existing ones [21,22]. Assume that the number of transmitting antennas at the base station is M = 14, 28, and the number of receiving antennas for each user is one. Assume that the number of users belonging to group 1 applying average-based linear precoding is K, and the number of users belonging to group 2 applying symbol-based precoding is J.

If the total transmission power for all users is P_T and the transmission power for each group is allocated in proportion to the number of users belonging to each group, the transmission power for each group can be expressed as $P_1 = \frac{K}{K+J}P_T$, $P_2 = \frac{J}{K+J}P_T$. It was assumed that quadrature phase-shift keying (QPSK) was applied as the modulation method for each symbol. The Rayleigh flat-fading channel model was used for the channel between the base station and each user, and it was assumed that the channel did not change during the number of symbol intervals, L. And since the design of the equalizer matrices $\mathbf{A}_1$ and $\mathbf{A}_2$ is beyond the scope of this paper, the special cases of $\mathbf{A}_1 = \mathbf{I}_K$ and $\mathbf{A}_2 = \mathbf{I}_J$ were considered.

Figure 2 compares the SER performance of the proposed scheme, a linear MMSE precoding scheme, and two SLP schemes [21,22] for the case where the number of base station antennas is M = 14 and (K + J) = 12. In the figure, the black, green, and red

lines represent the conventional MMSE precoding scheme and the two conventional SLP schemes (SLP-1 and SLP-2) [21,22], respectively. Also, the blue lines represent the proposed hybrid schemes. The performance of the proposed scheme was compared for K = 2, 6, and 10. In this figure, the horizontal axis represents the signal-to-noise ratio (SNR), while the vertical axis represents the symbol error rate (SER). To calculate the SER, 1000 independent channels were generated, and for each channel, 1000 independent received signals were generated to calculate the average error probability. From these results, it can be seen that the proposed scheme has slightly worse SER performance than the SLP-2 scheme [22] when K is 2, but it has better performance than the linear MMSE precoding and SLP-1 schemes [21]. And it can be seen that, as K decreases, the SER performance of the proposed technique approaches that of the SLP-2 scheme.
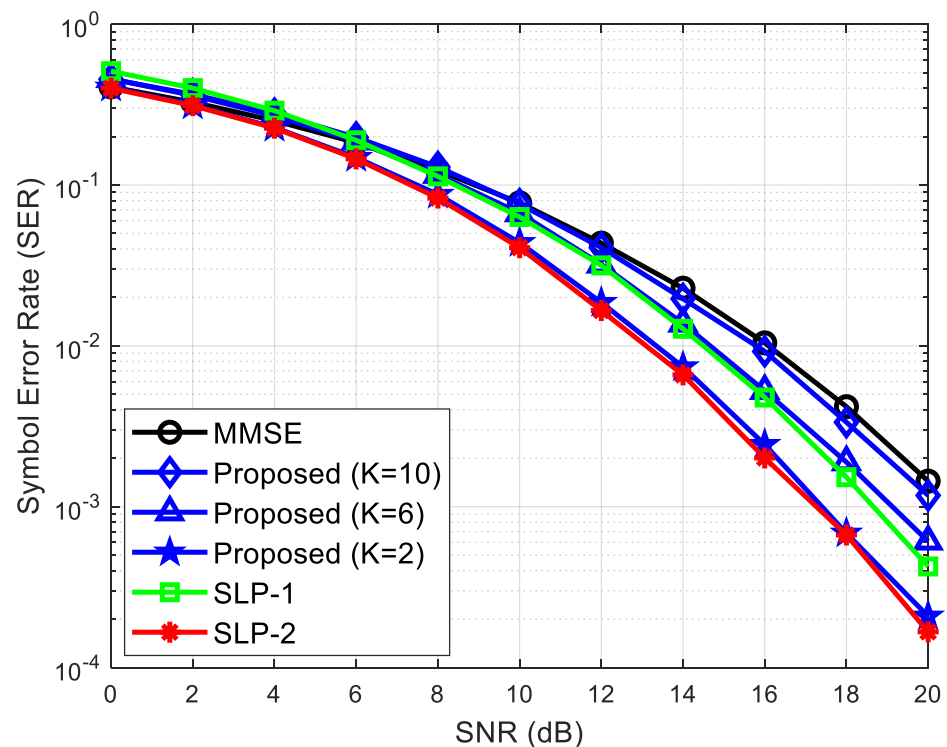

Figure 2. SER comparison of the proposed scheme and conventional ones when M = 14 and K + J = 12. (SLP-1 and SLP-2 refer to the schemes in [21] and [22], respectively).

Figure 3 compares the performance of the proposed scheme and the conventional ones for the case where the number of base station antennas is M = 28 and (K + J) = 26. The performance of the proposed scheme was compared for K = 4, 10, 16, and 22. As in Figure 2, it can be seen that, as K increases, the performance of the proposed scheme approaches that of the linear MMSE precoding scheme, and as K decreases, it approaches the performance of the SLP-2 scheme. Therefore, by adjusting the value of K, we can achieve the desired SER performance.

Tables 1 and 2 show the results from comparing the computational complexity performance of the proposed scheme and conventional ones for various K values, and the units of the numbers are seconds. Here, the time required to calculate the precoding matrix and vector was defined as the computational complexity performance. The execution time taken to calculate precoding for each scheme was compared using a computer with an Intel Core (TM) i7-8700 3.19 GHz processor with 16 GB of memory. The period during which the channel does not change and maintains a constant value is expressed as the block length, L, and the performance was compared for the case where L is 1000.
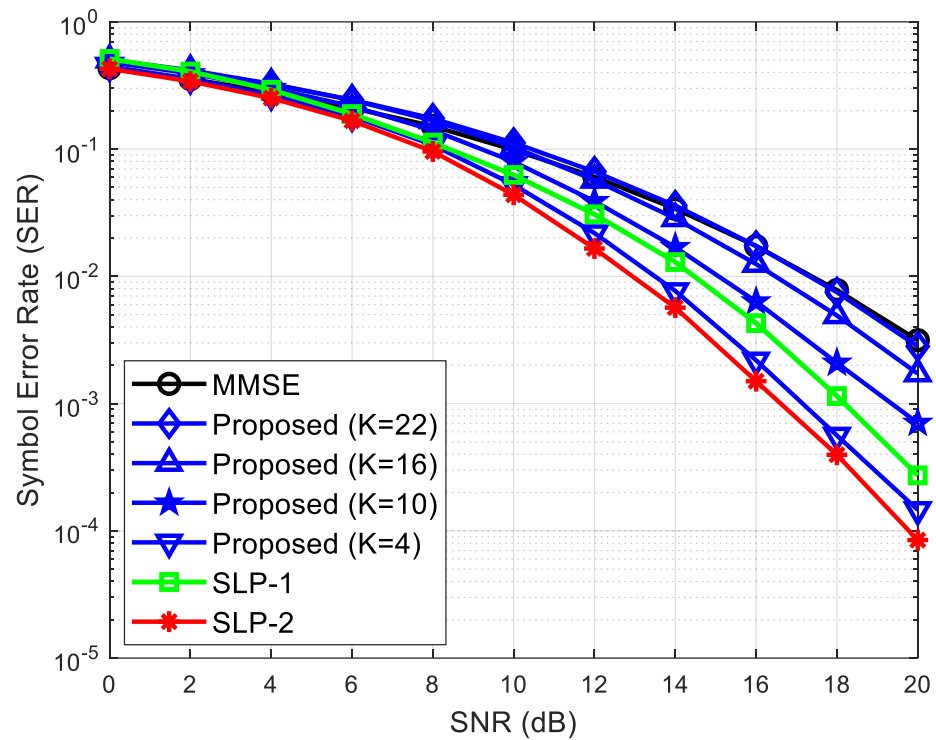
Figure 3. SER comparison of the proposed scheme and conventional ones when M = 28 and K + J = 26. (SLP-1 and SLP-2 refer to the schemes in [21] and [22], respectively).

Table 1. Computation complexity comparison in terms of the execution time when block length L = 1000, M = 14, and K + J = 12.

K	MMSE (s)	SLP-2 (s)	Proposed (s)	Pro./SLP-2 (%)
2	3.1369×10^{-5}	2.1784×10^{-1}	2.1581×10^{-1}	99.1
4	3.1369×10^{-5}	2.1784×10^{-1}	2.0712×10^{-1}	95.1
6	3.1369×10^{-5}	2.1784×10^{-1}	1.9234×10^{-1}	88.3
8	3.1369×10^{-5}	2.1784×10^{-1}	1.7032×10^{-1}	78.2
10	3.1369×10^{-5}	2.1784×10^{-1}	1.5319×10^{-1}	70.3

Table 2. Computation complexity comparison in terms of the execution time when block length L = 1000, M = 28, and K + J = 26.

K	MMSE (s)	SLP-2 (s)	Proposed (s)	Pro./SLP-2 (%)
4	8.0797×10^{-5}	6.1286×10^{-1}	5.1434×10^{-1}	83.9
10	8.0797×10^{-5}	6.1286×10^{-1}	3.8130×10^{-1}	62.2
16	8.0797×10^{-5}	6.1286×10^{-1}	2.7781×10^{-1}	45.3
22	8.0797×10^{-5}	6.1286×10^{-1}	2.1152×10^{-1}	34.5

Table 1 shows the results from comparing the computational complexity for the case of M = 14 and K + J = 12. From these results, it can be seen that the computation time of the SLP-2 scheme is much longer than that of the linear MMSE precoding technique and that the computation time of the proposed scheme is longer than that of the linear MMSE precoding scheme and shorter than that of the SLP-2 scheme. The reason for this is that the MMSE precoding scheme is an average-based precoding technique, so the precoding matrix is obtained only once during the block length period, but the SLP-2 scheme is a symbol-based precoding technique, so the precoding result must be obtained for each

symbol, so the calculation time of the SLP-2 technique takes much longer. Meanwhile, since the proposed technique applies the MMSE precoding technique to the data for K users belonging to group 1, the precoding matrix is obtained only once during the block length period, so the calculation time is reduced compared to that of the SLP-2 technique.

In this table, Pro./SLP-2 (%) is the result of comparing the calculation time of the proposed method with that of the SLP-2 method. From these results, it can be seen that as K increases, the calculation time of the proposed method decreases significantly compared to that of the SLP-2 method. Table 2 shows the performance for the case of M = 28 and K + J = 26. In this case, similar results are shown in Table 2.

Combining the results in Figures 1 and 2 and Tables 1 and 2, it can be seen that the proposed technique shows better SER performance than the MMSE precoding technique and requires much less computation than the SLP-2 technique. It can be seen that the SER performance and computational complexity performance of the proposed method can be appropriately controlled by adjusting the value of K.

5. Conclusions

This paper proposes a hybrid technique that simultaneously applies a linear orthogonal MMSE precoding scheme and a non-linear orthogonal SLP scheme to control the SER performance and computational complexity of the entire system. If the SLP technique is applied to all users, the computational complexity increases too much, and if linear precoding is applied to all users, the SER performance deteriorates. Therefore, to solve this problem, we proposed a hybrid precoding technique that divides users into two groups and applies linear OMMSE precoding to users in group 1 and the non-linear OSLP technique to users in group 2. Because the SER performance deteriorates significantly when the signals of users applying different types of precoding interfere with each other, the precoding matrix and vector are designed to prevent interference between the two groups by utilizing the channel information for the two groups. Through computer simulation, it was confirmed that the SER performance of the proposed method is superior to that of the linear MMSE precoding method, and the computational complexity performance of the proposed method is superior to that of the SLP-2 method. Additionally, the proposed technique suggests that the SER performance and computational complexity performance can be appropriately controlled by adjusting the number of users belonging to group 1.

This paper assumed a perfect channel estimation scenario, hence presuming the absence of channel estimation errors. Research on designing precoding methods considering partial information about the channel or the presence of channel estimation errors could be a significant area of study. However, due to the scope limitations of this paper, such considerations are planned for future work.

Conflicts of Interest: The authors declare no conflicts of interest.

References

1. Wang, C.X.; You, X.; Gao, X.; Zhu, X.; Li, Z.; Zhang, C.; Wang, H.; Huang, Y.; Chen, Y.; Haas, H.; et al. On the road to 6G: Visions, requirements, key technologies, and testbeds. *IEEE Commun. Surv. Tutor.* **2023**, *25*, 905–974. [CrossRef]
2. Jiang, W.; Han, B.; Habibi, M. The road towards 6G: A comprehensive survey. *IEEE Open J. Commun. Soc.* **2021**, *2*, 334–366. [CrossRef]
3. Goldsmith, A.; Jafar, S.A.; Jindal, N.; Vishwanath, S. Capacity limits of MIMO channels. *IEEE J. Sel. Areas Commun.* **2003**, *21*, 684–702. [CrossRef]
4. Paulraj, A.J.; Gore, D.A.; Nabar, R.U.; Bolcskei, H. An overview of MIMO communications—A key to gigabit wireless. *Proc. IEEE* **2004**, *92*, 198–218. [CrossRef]

5. Jindal, N. MIMO broadcast channels with finite-rate feedback. *IEEE Trans. Inf. Theory* **2006**, *52*, 5045–5060. [CrossRef]
6. Honig, M.; Madhow, U.; Verdu, S. Blind adaptive multiuser detection. *IEEE Trans. Inf. Theory* **1995**, *41*, 944–960. [CrossRef]
7. Poor, H.V.; Verdu, S. Probability of error in MMSE multiuser detection. *IEEE Trans. Inf. Theory* **1997**, *43*, 858–871. [CrossRef]
8. Zheng, L.; Tse, D.N.C. Diversity and multiplexing: A fundamental tradeoff in multiple-antenna channels. *IEEE Trans. Inf. Theory* **2003**, *49*, 1073–1096. [CrossRef]
9. Johan, M.; Utschick, W.; Nossek, J.A. Linear transmit processing in MIMO communication systems. *IEEE Trans. Signal Process.* **2005**, *53*, 2700–2712. [CrossRef]
10. Wiesel, A.; Eldar, Y.C.; Shitz, S. Linear precoding via conic optimization for fixed MIMO receivers. *IEEE Trans. Signal Process.* **2006**, *54*, 161–176. [CrossRef]
11. Vu, M.; Paulraj, A.J. MIMO wireless linear precoding. *IEEE Signal Process. Mag.* **2007**, *24*, 86–105. [CrossRef]
12. Costa, M. Writing on dirty paper. *IEEE Trans. Inf. Theory* **1893**, *29*, 439–441. [CrossRef]
13. Peel, C.B.; Hochwald, B.M.; Swindlehurst, A.L. A vector-perturbation technique for near-capacity multiantenna multiuser communication-Part 2: Channel inversion and regularization. *IEEE Trans. Commun.* **2005**, *53*, 537–544. [CrossRef]
14. Li, A.; Masouros, C. A constellation scaling approach to vector perturbation for adaptive modulation in MU-MIMO. *IEEE Wirel. Commun. Lett.* **2015**, *4*, 289–292. [CrossRef]
15. Li, A.; Masouros, C. A two-stage vector perturbation scheme for adaptive modulation in downlink MU-MIMO. *IEEE Trans. Veh. Technol.* **2016**, *65*, 7785–7791. [CrossRef]
16. Li, A.; Spano, D.; Krivochiza, J.; Domouchtsidis, S.; Tsinos, C.C.; Masouros, C.; Chatzinotas, S.; Li, Y.; Vucetic, B.; Ottersten, B. A tutorial on interference exploitation via symbol-level precoding: Overview, state-of-the-art and future directions. *IEEE Commun. Surv. Tutor.* **2020**, *22*, 796–839. [CrossRef]
17. Alodeh, M.; Spano, D.; Kalantari, A.; Tsinos, C.; Christopoulos, D.; Chatzinotas, S.; Ottersten, B. Symbol-level and multicast precoding for multiuser multiantenna downlink: A survey classification and challenges. *IEEE Commun. Surv. Tutor.* **2018**, *20*, 1733–1757. [CrossRef]
18. Li, A.; Masouros, C. Interference exploitation precoding made practical: Optimal closed-form solutions for PSK modulations. *IEEE Trans. Wirel. Commun.* **2018**, *17*, 7661–7676. [CrossRef]
19. Haqiqatnejad, A.; Kayhan, F.; Ottersten, B. Symbol-level precoding design based on distance preserving constructive interference regions. *IEEE Trans. Signal Process.* **2018**, *66*, 5817–5832. [CrossRef]
20. Masouros, C.; Alsusa, E. Dynamic linear precoding for the exploitation of known interference in MIMO broadcast systems. *IEEE Trans. Wirel. Commun.* **2009**, *8*, 1396–1404. [CrossRef]
21. Haqiqatnejad, A.; Kayhan, F.; Ottersten, B. An approximate solution for symbol-level multiuser precoding using support recovery. In Proceedings of the 2019 IEEE 20th International Workshop on Signal Processing Advances in Wireless Communications (SPAWC), Canes, France, 2–5 July 2019; pp. 1–5.
22. Wang, Y.; Wang, W.; You, L.; Tsinos, C.G.; Jin, S. Weighted MMSE precoding for constructive interference region. *IEEE Wireless Commun. Lett.* **2022**, *11*, 2605–2609. [CrossRef]
23. Liu, R.; Li, H.; Ming, L. Symbol-level hybrid precoding in mmWave multiuser MISO systems. *IEEE Commun. Lett.* **2019**, *23*, 1636–1639. [CrossRef]
24. Shao, M.; Li, Q.; Ma, W. Minimum symbol-error probability symbol-level precoding with intelligent reflecting surface. *IEEE Wirel. Commun. Lett.* **2020**, *10*, 1601–1605. [CrossRef]
25. Liu, Y.; Shao, M.; Ma, W.; Li, Q. Symbol-level precoding through the lens of zero forcing and vector perturbation. *IEEE Trans. Signal Process.* **2022**, *70*, 687–1703. [CrossRef]
26. Argenti, F.; Biagini, M.; Del Re, E.; Morosi, S. Time-frequency MSE analysis of linear channel estimation methods for the LTE downlink. *Trans. Emerg. Telecommun. Technol.* **2015**, *26*, 704–717. [CrossRef]