II Всероссийская (национальная) научная
конференция с международным участием
«Российская наука, инновации, образование»
(РОСНИО-II-2023)

8 (2023)

УДК 621.376

EDN HZFXUL

Эффективный гибридный метод прекодирования для массовых MIMO-систем миллиметрового диапазона радиоволн

М. Асаф*, О.Г. Пономарев

Томский государственный университет, Томск, Россия

*E-mail: md.moh1600@gmail.com

Аннотация. В этой статье мы предлагаем новую гибридную методику прекодирования, которая может значительно улучшить производительность приемопередающих систем MIMO миллиметрового диапазона волн. В общем случае гибридное прекодирование использует радиочастотные цепи, которые больше, чем количество потоков данных, чтобы максимизировать спектральную эффективность. Существующие методы гибридного прекодирования не эффективно используют преимущества дополнительных радиочастотных цепей для улучшения производительности гибридной системы. В данном исследовании мы покажем, как эффективно использовать эти радиочастотные цепи для воспроизведения идеальной диаграммы направленности луча, которая представляет собой верхнюю границу производительности гибридной системы прекодирования. Результаты моделирования показывают высокую степень эффективности предложенного метода по сравнению с существующими методами с точки зрения спектральной эффективности.

Ключевые слова: аналоговое/цифровое прекодирование, связь на миллиметровых волнах, массовое MIMO, гибридное формирование луча.

Efficient Hybrid Precoding method for Millimeter-Wave Massive MIMO Systems

M. Assaf *, O.G. Ponomarev

Tomsk State University, Tomsk, Russia

*E-mail: md.moh1600@gmail.com

Abstract. In this study, we provide a novel hybrid precoding technique that can significantly improve the performance of millimeter-wave MIMO transceiver systems. The general case of hybrid precoding employs radio frequency (RF) chains that are larger than the number of data streams in order to maximize spectral efficiency. Existing hybrid precoding techniques do not take advantage of the additional RF chains effectively to improve the performance of the hybrid system. In this study, we will show how to exploit these RF chains to reproduce the ideal beam pattern, which represents the upper performance bound of the hybrid precoding system. The simulation results show the effectiveness of the proposed method compared to the current methods in terms of spectral efficiency.

Keywords: analog/ digital precoding**,** Millimeter wave communication, massive MIMO, hybrid beamforming.

1. Introduction

One of the important technologies for achieving high capacity performance in the 5G/6G mobile cellular networks is the utilization of the massive multiple-input multiple-output (MIMO) millimeter (mm) wave technology [1,2]. In conventional MIMO systems, also known as full-digital MIMO systems, the number of RF chains is exactly equal to the number of antenna elements, so it is easy to perform digital precoding at the transmitter and digital combining at the receiver in order to achieve maximum spectral efficiency. In massive MIMO, digital beamforming cannot be applied due to the high manufacturing cost and energy consumption of the RF circuits. This problem is solved by using the concept of hybrid processing [3], in which baseband digital precoding is combined with analogue precoding performed using cost-effective variable phase shifters. The hybrid system presents a challenge with regard to choosing the precoding and combining weights in different conditions to maximize spectral efficiency. In [4], the scattering properties of mmWave channels are exploited to design RF and baseband precoders. The authors of [5] suggested designing the RF precoding matrix directly based on the right singular vectors of the channel matrix. The concept of "equivalent channel" has been exploited in [6] to design a hybrid precoder and combiner jointly to improve the system's spectral efficiency. A two-stage successive method is proposed in [7] to design the analog precoder and combiner jointly.

The common problem with the existing precoding approaches is that they do not provide high performance in cases where the number of RF chains is higher than the number of data streams. In this study, we will explain how we can reconstruct the optimum beam pattern using the additional RF chains without requiring full channel knowledge and complex decomposition techniques, which reduce complexity and the amount of feedback information.

2. A. System Model

Figure 1. depicts an mmWave hybrid MIMO system for a single user. The transmitter is assumed to have N_s spatial streams transmitted via N_t antennas using N_t^{RF} RF chains, while the receiver is assumed to use N_r receive antennas and N_r^{RF} RF chains for reception. To achieve the hybrid structure's main objective of low complexity, the number of RF chains should be less than the number of antennas, and hence $N_s \le N_t^{RF}(N_r^{RF}) \ll N_t(N_r)$. Each RF chain is connected

to the $N_t\left(N_r\right)$ phase shifters in the transmitter (receiver) and can be utilized to create a single beam to transmit (receive) an independent data stream.

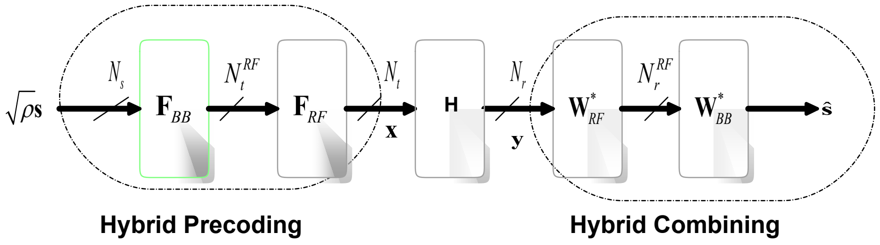


Figure 1. MIMO hybrid precoding/combining Block diagram.

The resultant signal at the receiver can be written as:

$$\hat{\mathbf{y}} = \sqrt{\rho}\mathbf{W}_{\mathbf{BB}}^*\mathbf{W}_{\mathbf{RF}}^*\mathbf{H}\mathbf{F}_{\mathbf{RF}}\mathbf{F}_{\mathbf{BB}}\mathbf{s} + \mathbf{W}_{\mathbf{BB}}^*\mathbf{W}_{\mathbf{RF}}^*\mathbf{n} \tag{1}$$

Where **s** is the symbol vector, ρ represents the average received signal power. $\mathbf{F}_{\mathbf{BB}}$ is $N_t^{RF} \times N_s$ baseband precoder matrix, and $\mathbf{F}_{\mathbf{RF}}$ is $N_t \times N_t^{RF}$ RF precoder matrix. **H** represents mmWave propagation channel. The noise vector $\mathbf{n} \in \mathbb{C}^{N_r \times 1}$ is assumed to have zero-mean and covariance matrix $\sigma^2\mathbf{I}_{N_r}$. The combiner matrix $\mathbf{W}^* = \mathbf{W}_{\mathbf{BB}}^*\mathbf{W}_{\mathbf{RF}}^*$, where $\mathbf{W}_{\mathbf{BB}} \in \mathbb{C}^{N_r^{RF} \times N_s}$ is the baseband combiner and $\mathbf{W}_{\mathbf{RF}} \in \mathbb{C}^{N_r \times N_r^{RF}}$ is the RF combiner.

We adopt the expanded Saleh-Valenzuela geometric model as in [8] to model massive MIMO mmWave channel. This model is based on the concept that N_{cl} scattering clusters, each of which contributes N_{ray} propagation paths, add together to form the channel matrix **H**. Therefore, the discrete-time narrowband channel **H** can be written as:

$$\mathbf{H} = \sqrt{\frac{N_t N_r}{N_{cl} N_{ray}}} \sum_{i=1}^{N_{cl}} \sum_{l=1}^{N_{ray}} \alpha_{i,l} \mathbf{a}_{\mathbf{r}}\left(\phi_{i,l}^r\right) \mathbf{a}_t\left(\phi_{i,l}^r\right)^* \tag{2}$$

where $\alpha_{i,l}$ is the complex gain of the l^{th} ray in the i^{th}cluster. $\mathbf{a}_{\mathbf{r}}\left(\phi_{i,l}^r\right)$ and $\mathbf{a}_t\left(\phi_{i,l}^r\right)$ are the array response vectors at the receiver and transmitter of the l^{th} ray in the i^{th} cluster, with azimuth angles $\phi_{i,l}^r$ and $\phi_{i,l}^r$, respectively. As we observe, the channel matrix $\mathbf{H} \in \mathbb{C}^{N_r \times N_t}$ and normalized to satisfy $\mathrm{E}\{\|\mathbf{H}\|_F^2\} = N_t N_r$. To simplify representing channel matrix, we rewrite it in a more compact form as:

$$\mathbf{H} = \mathbf{A}_r diag(\boldsymbol{\alpha})\mathbf{A}_t^* \tag{3}$$

Where $\boldsymbol{\alpha} = \sqrt{\frac{N_t N_r}{N_{cl} N_{ray}}} \left[\alpha_{1,1}, \alpha_{1,2}, \ldots, \alpha_{N_{cl}, N_{ray}} \right]^T$ contains the complex gains of all paths, and the matrices $\mathbf{A}_r = \left[\mathbf{a}_\mathbf{r}\left(\phi_{1,1}^r\right), \mathbf{a}_\mathbf{r}\left(\phi_{1,2}^r\right), \ldots, \mathbf{a}_\mathbf{r}\left(\phi_{N_{cl},N_{ray}}^r\right) \right]$ and $\mathbf{A}_t = \left[\mathbf{a}_t\left(\phi_{1,1}^t\right), \mathbf{a}_t\left(\phi_{1,2}^t\right), \ldots, \mathbf{a}_t\left(\phi_{N_{cl},N_{ray}}^t\right) \right]$ contain the array response vectors. We could find that the number of the paths is the upper bound of the rank of the mmWave channel matrix.

The main challenge in hybrid structure is to construct precoding and combining matrices in order to maximize the spectral efficiency:

$$R = \log_2 \left(\left| \begin{array}{c} \mathbf{I}_{Ns} + \frac{\rho}{N_s} \mathbf{R}_n^{-1} \mathbf{W}_{\mathbf{BB}}^* \mathbf{W}_{\mathbf{RF}}^* \mathbf{H} \mathbf{F}_{\mathbf{RF}} \mathbf{F}_{\mathbf{BB}} \\ \times \mathbf{F}_{\mathbf{BB}}^* \mathbf{F}_{\mathbf{RF}}^* \mathbf{H}^* \mathbf{W}_{\mathbf{RF}} \mathbf{W}_{\mathbf{BB}} \end{array} \right| \right) \tag{4}$$

3. The proposed method

The optimal precoding matrix at the transmitter and the optimal combining matrix at the receiver can be obtained by decomposing the channel matrix and choosing the N_s columns of the left and right singular vectors that correspond to the highest singular values as in:

$$\mathbf{H} = \mathbf{U}\boldsymbol{\Sigma}\mathbf{V}^* \text{ and} \begin{cases} \tilde{\mathbf{V}} \leftarrow N_s \text{ principal right singular vectors of } \mathbf{H} \\ \tilde{\mathbf{U}} \leftarrow N_s \text{ principal left singular vectors of } \mathbf{H} \end{cases} \tag{5}$$

where $\mathbf{U}$ and $\mathbf{V}$ are $N_r \times N_r$ and $N_t \times N_t$ unitary matrices, respectively, and $\boldsymbol{\Sigma}$ is $N_r \times N_t$ diagonal matrix with diagonal entries arranged in decreasing order. Then the optimal precoding matrix $\mathbf{F}_{opt} = \tilde{\mathbf{V}}\tilde{\Sigma}$, and optimal combining matrix $\mathbf{W}_{opt} = \tilde{\mathbf{U}}$, where $\tilde{\Sigma}$ contains the power allocation for each stream.

The optimization problem in (4) will be separated into transmitter and receiver, with the precoding matrices $\mathbf{F}_{\mathbf{BB}}$ and $\mathbf{F}_{\mathbf{RF}}$ designed to be as close to the optimal precoding matrix $\mathbf{F}_{opt}$ as possible, and then the same process will be repeated in the receiver to optimize combining matrices. On the transmitter side, the objective of hybrid design is to optimize the two matrices $\mathbf{F}_{\mathbf{BB}}$ and $\mathbf{F}_{\mathbf{RF}}$ jointly that approaches the targeting matrix $\mathbf{F}_{opt}$ as in:

$$\begin{gathered} \left(\boldsymbol{F}_{RF}^{opt}, \boldsymbol{F}_{BB}^{opt}\right) = arg \min_{\boldsymbol{F_{RF}}, \boldsymbol{F_{BB}}} \left\| \boldsymbol{F}_{opt} - \boldsymbol{F_{RF}} \boldsymbol{F_{BB}} \right\|_F^2 \\ \text{S.t.} \quad |\boldsymbol{F_{RF}}(i,j)| = 1 \quad \vee\, i,j \\ \|\boldsymbol{F_{RF}} \boldsymbol{F_{BB}}\|_F^2 = N_s \end{gathered} \tag{6}$$

It is worth noting that without the constraint of $\left|\mathbf{F_{RF}}(i,j)\right|=1 \;\; \vee i,j$, it is very easy to achieve the optimal performance of a fully digital system using a hybrid structure; for example, if $N_t^{RF}=N_s$, the solution of (6) can be chosen as $\mathbf{F_{RF}}=\mathbf{F}_{opt}, \mathbf{F_{BB}}=\mathbf{I}_{N_s}$. This case will be referred to as "optimal unconstraint precoding solution".

The solution of (6) presented in [4] depends on choosing N_r^{RF} vectors from the array response matrix $\mathbf{A}_t$ that have maximum projection on the optimum precoding matrix $\mathbf{F}_{opt}$, after that the baseband precoding matrix can be found using least square solution as in:

$$\mathbf{F_{BB}}=\left(\mathbf{F_{RF}^{*}}\mathbf{F_{RF}}\right)^{-1}\mathbf{F_{RF}^{*}}\mathbf{F}_{opt} \tag{7}$$

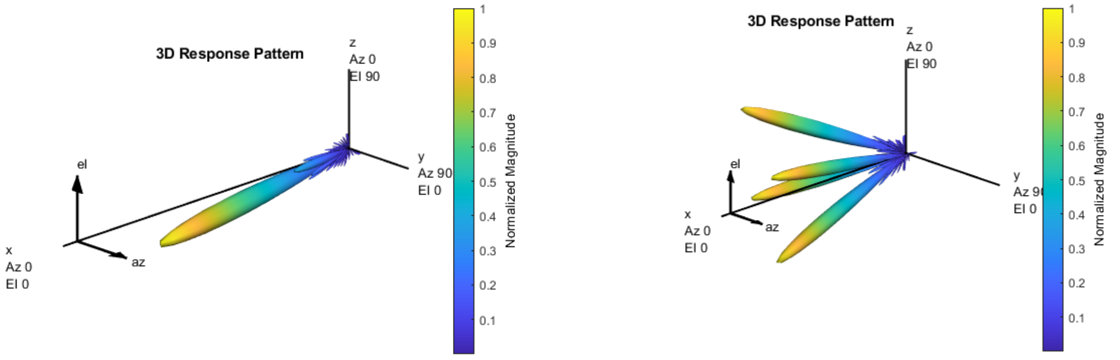
The authors in [9] investigated an important lemma: when the channel has total paths much less than the number of antennas $N_{ray}N_{cl} \ll \min(N_t, N_r)$, then in the terms of chordal distance, the left and right singular vectors of the matrix channel converge to the array response vectors.

Lemma 1 demonstrates that the array response vectors in $\mathbf{A}_t$ converges to left singular vectors $\mathbf{V}$ of the channel $\mathbf{H}$, and using the dominant vectors of $\mathbf{A}_t$ to construct the RF matrix $\mathbf{F_{RF}}$ as shown in [4] converges to optimum precoding matrix $\mathbf{F}_{opt}$.

In the hybrid configuration, we can distinct between two different cases, the first is when the number of streams is equal to the number of RF chains $N_t^{RF}=N_s$ and the second is when the number of streams is less than the number of RF chains $N_t^{RF}>N_s$. When $N_t^{RF}=N_s$, the current approaches, as in [4][5][6], transmit each data stream via a dominant path with an angle of departure obtained from the array response matrix $\mathbf{A}_t$ that has maximum project onto the optimum precoding matrix $\mathbf{F}_{opt}$. On the receiver side, N_s dominant arrival paths are selected from the array response matrix $\mathbf{A}_r$, and then the phase shifters of the antenna elements are weighted to form N_s directed beams in order to maximize spectral efficiency as much as possible. In other words, N_s independent directed beams that correspond to the strongest paths are formed to send and receive N_s data streams. When the number of RF chains exceeds the number of streams $N_t^{RF}>N_s$, the present techniques continue looking for additional directed beams to send and receive the data streams via them. As a result, the number of users that can be serviced at the same physical resource blocks, i.e., the same time/frequency resources but distinct spatial beams, will be reduced.

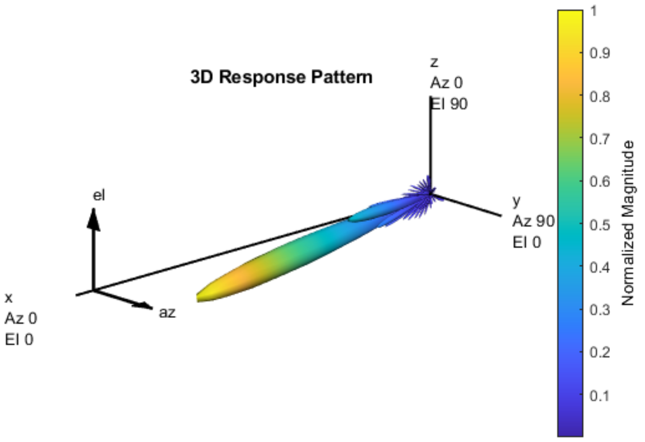
In order to address this issue, we propose a new method that employs a number of directed beams equal to the number of data streams until $N_t^{RF} > N_s$, where the additional RF chains $N_t^{RF} - N_s$ are used to reduce the residual error between the product precoding matrices $\mathbf{F}_{\mathbf{RF}}\mathbf{F}_{\mathbf{BB}}$ and the optimum matrix $\mathbf{F}_{opt}$ caused by the RF precoding matrix's constant magnitude $\left|\mathbf{F}_{\mathbf{RF}}(i,j)\right| = 1 \quad \vee i,j$.

In the proposed method, the first N_s columns of the RF precoder are computed by using the angle information of the first N_s dominant paths, and then the first N_S rows of the matrix $\mathbf{F}_{\mathbf{BB}}$ are calculated based on equation (7). The remaining $N_t^{RF} - N_s$ columns of the RF and baseband matrices are used to minimize the residual error between the matrix $\mathbf{F}_{opt}$ and the product $\mathbf{F}_{\mathbf{RF}}\mathbf{F}_{\mathbf{BB}}$. To get a better understanding of the suggested precoding method, figure 2. shows the beam pattern for optimum unconstrained precoding, the OMP method in [4] and the proposed method. In figure 2 one data stream is transmitted using the following hybrid configuration $N_t^{RF} = N_r^{RF} = 4, N_t = 256, N_r = 64$, the channel is supposed to have 8 scattering clusters randomly distributed in space, for simplicity suppose that each cluster has one path.



(a)

(b)



(c)

Figure 2. Beam pattern generated a 256-element square array in an example channel realization with 8 rays (or equivalently 8 clusters with 0 angular spread) with 4 RF chains and one data stream using (a) optimal unconstrained beamforming, (b) sparse precoding solution and (c) the proposed method.

We note that whereas the suggested approach in [4] creates $N_t^{RF} = 4$ various directed beams to construct its response pattern, the optimum precoder and the proposed method only generate one beam to carry the data stream. As a result, the proposed approach takes advantage of the extra RF chains to rebuild an optimal unconstrained beam, which maximizes spectral efficiency from one side while using less spatial resources from the other.

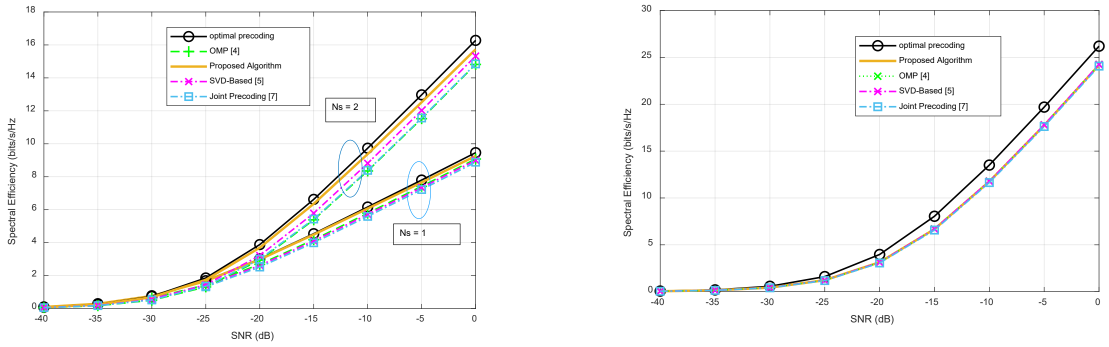
Algorithm 1. The proposed hybrid precoding algorithm

```
Require $\mathbf{F}_{opt}$
$\mathbf{F}_{\mathbf{RF}}, \mathbf{F}_{\mathbf{BB}}$ are empty matrices, $L = \left\lceil \frac{N_t^{RF}}{N_s} \right\rceil - 1$
Use OMP algorithm to compute first $N_s$ dominant path to create first $N_s$ columns and rows of $\mathbf{F}_{\mathbf{RF}}, \mathbf{F}_{\mathbf{BB}}$
For $i = 1{:}L$
    $\mathbf{F}_{res} = \mathbf{F}_{opt} - \mathbf{F}_{\mathbf{RF}}\mathbf{F}_{\mathbf{BB}}$
    $\hat{\mathbf{F}}_{\mathbf{RF}} = \mathbf{F}_{res}./\,\mathrm{abs}\left(\mathbf{F}_{res}\right)$
    $\hat{\mathbf{F}}_{\mathbf{BB}} = \left(\hat{\mathbf{F}}_{\mathbf{RF}}^{*}\hat{\mathbf{F}}_{\mathbf{RF}}\right)^{-1}\hat{\mathbf{F}}_{\mathbf{RF}}^{*}\mathbf{F}_{res}$
    $\mathbf{F}_{\mathbf{RF}} = \left[\mathbf{F}_{\mathbf{RF}} \mid \hat{\mathbf{F}}_{\mathbf{RF}}\right], \mathbf{F}_{\mathbf{BB}} = \left[\mathbf{F}_{\mathbf{BB}} \mid \hat{\mathbf{F}}_{\mathbf{BB}}\right]$
```

```
8.  end for
9.  $\mathbf{F}_{\mathbf{BB}} = \sqrt{N_s}\,\frac{\mathbf{F}_{\mathbf{BB}}}{\|\mathbf{F}_{\mathbf{RF}}\mathbf{F}_{\mathbf{BB}}\|_F}$
return $\mathbf{F}_{\mathbf{RF}}, \mathbf{F}_{\mathbf{BB}}$
```

Algorithm 1 shows the pseudo-code for the proposed precoder. In summary, the first N_s columns of the RF precoder and first N_s rows of the $\mathbf{F}_{\mathbf{BB}}$ are calculated in step 2 using the method proposed in [4], where each data stream is linked to its beam. The steps from 3 to 9 correspond to finding the remaining $N_t^{RF} - N_s$ columns of the precoding matrices, which are computed iteratively to minimize the residual error between the matrix $\mathbf{F}_{opt}$ and the product $\mathbf{F}_{\mathbf{RF}}\mathbf{F}_{\mathbf{BB}}$ as in step 4. It is worth noting that the hybrid RF and baseband combiners on the receiver side can be obtained in the same way as the hybrid precoder design in Algorithm 1.

To show the spectral efficiency improvement of our proposed hybrid precoding scheme, our results are compared with the OMP algorithm in [4], SVD-Based Low-Complexity algorithm in [5], and Joint Hybrid Precoding/Combining in [7]. Figure 3 illustrate the spectral efficiencies versus the varying SNRs under the conditions of $\left(N_t = 64, N_r = 16, N_t^{RF} = N_r^{RF} = 4\right)$, for different values of N_s.



(a) (b)

Figure 3. Spectral efficiencies with varying SNRs when $N_t = 64$ and $N_r = 16$. (a) Data stream $N_s = 1$ and 2. (b) Data stream $N_s = 4$.

From the simulation results in figure 3, when $N_s = 1$, the proposed scheme performs the same as the optimal precoding scheme, when $N_s = 2$, the gap between our proposed scheme and the optimal scheme increases, but is still less than the gap with the current schemes, and when $N_s = N_t^{RF} = 4$, the gap increases and the achieved performance is similar to the current schemes

4. Conclusion

In this paper, we discuss the issue of beamforming for a single-user mmWave massive MIMO hybrid system that provides a reasonable compromise between performance and complexity. The first N_s dominating beams are located using the approach suggested in [4], and these beams are adjusted using the remaining $N_t^{RF} - N_s$ RF chains to reduce the residual error between the matrix $\mathbf{F}_{opt}$ and the product $\mathbf{F}_{\mathbf{RF}}\mathbf{F}_{\mathbf{BB}}$. Simulation results showed that the proposed scheme can almost match the performance of the upper bound obtained by the optimal full-baseband design when $N_s < N_t^{RF}$.

References

1. Xiao Z. A survey on millimeter-wave beamforming enabled UAV communications and networking / Z Xiao et al. // IEEE Commun. Surveys Tuts. – 2022. – Vol. 24(1). – P. 557-610.
2. Busari S.A. Millimeter-wave massive MIMO communication for future wireless systems: A survey / S.A. Busari, K.M.S. Huq, S. Mumtaz, L. Dai, J. Rodriguez // IEEE Commun. Surveys Tuts. – 2017. – Vol. 20(2). – P. 836-869.
3. Molisch A.F. Hybrid beamforming for massive MIMO: A survey, / A.F. Molisch et al. // IEEE Commun. Mag. – 2017. – Vol. 55(9). – P. 134-141.
4. Ayach O.E. Spatially Sparse Precoding in Millimeter Wave MIMO Systems / O.E. Ayach, S. Rajagopal, S. Abu-Surra, Z. Pi, R.W. Heath // IEEE Transactions on Wireless Communications. – 2014. – Vol. 13(3). – P. 1499-1513.
5. Zhang D. SVD-Based Low-Complexity Hybrid Precoding for Millimeter-Wave MIMO Systems / D. Zhang, P. Pan, R. You, H. Wang // IEEE Communications Letters. – 2018. – Vol. 22(10). – P. 2176-2179.
6. Wang S. A Family of Hybrid Precoding Schemes for Millimeter-Wave Massive MIMO Systems / S. Wang, M. He, J. Wang, R. Ran, H. Ji, V.C.M. Leung // IEEE Systems Journal. – 2022. – Vol. 16(3). – P. 4881-4891.

7. Wang S. A Joint Hybrid Precoding/Combining Scheme Based on Equivalent Channel for Massive MIMO Systems / S. Wang et al. // IEEE Journal on Selected Areas in Communications. – 2022. – Vol. 40(10). – P. 2882-2893.
8. Ayach O. El, The capacity optimality of beam steering in large millimeter wave MIMO systems / Ayach O. El, Heath R.W., Abu-Surra S., Rajagopal S., Pi Z. // Proc. IEEE 13th Int. Workshop Signal Process. Adv. Wireless Commun. – 2012. – P. 100-104.
9. Raghavan V. Sublinear capacity scaling laws for sparse MIMO channels / V. Raghavan, A.M. Sayeed // IEEE Trans. Inf. Theory. – 2011. – Vol. 57(1). – P. 345-364.