IET Communications

Research Article

Hybrid precoding in point-to-point massive multiple-input multiple-output systems based on normalised matrix adaptive method

ISSN 1751-8628
Received on 19th January 2017
Revised 28th March 2017
Accepted on 24th April 2017
E-First on 8th September 2017
doi: 10.1049/iet-com.2017.0071
www.ietdl.org

Yongpan Feng[1], *Suk Chan Kim*[1] ✉

[1]*Department of Electronics Engineering, Pusan National University, Busan, Korea*

✉ *E-mail: sckim@pusan.ac.kr*

Abstract: The potential of millimetre wave frequencies band has motivated the study of large-scale antenna arrays to improve the spectrum efficiency of massive multiple-input multiple-output wireless communications systems. However, conventional fully digital precoding schemes, which require one radio frequency (RF) chain for each antenna, are cost prohibitive and have high-power consumption. To address these hardware limitations, this study considers a hybrid precoding scheme that provides high performance in terms of spectrum efficiency. By deducing the derivative of the spectrum efficiency function with respect to the phase matrix variable of the RF precoder, an efficient RF precoding scheme based on the normalised matrix adaptive method (NMAM) is proposed. Numerical simulations show that the NMAM achieves excellent performance approaching that of fully digital precoding, with relatively low computational complexity.

1 Introduction

Large-scale antenna techniques have been deployed in modern wireless networks to improve capacity and reliability [1–3]. In conventional multiple-input multiple-output (MIMO) systems, each antenna must be supported by a dedicated radio frequency (RF) chain, which includes digital-to-analogue converters, mixers and power amplifiers. For large-scale antenna arrays, this is prohibitive from both a cost perspective and a power consumption perspective.

To address the difficulty of limited numbers of RF chains, several precoding schemes for massive MIMO systems have been proposed recently [4–9]. The works in [4–6] developed several iterative hybrid beamforming algorithms based on partial channel knowledge and a variant of matching pursuit method. However, these algorithms are only limited to a certain channel model, i.e. the Saleh–Valenzuela model. In [7], a low-complexity RF precoding scheme was proposed, which considers the conjugate transpose of the channel phase array (CTCPA) as the RF precoder. This scheme has very low complexity, but its performance cannot match the performance of fully digital precoding, especially as the number of antennas and the signal-to-noise rate (SNR) increase. Gao *et al.* [8] proposed a successive interference cancelation-based hybrid precoding from the perspective of the energy efficiency. In [9], an iterative algorithm was proposed; each RF precoder element is sequentially updated according to its iterative coordinate descent algorithm. Although it approaches the performance of fully digital precoding, its computational complexity is very high.

In this paper, point-to-point multiple-input multiple-output (P2P MIMO) downlink wireless communication is considered. Our contributions in this paper focus on figuring out the sub-optimal RF precoder using the proposed efficient iterative algorithm based on the normalized matrix adaptive method (NMAM). In contrast to the previous works, it abstracts the phases array (a real matrix) of the RF precoder as the variable matrix; this eliminates the constant modulus constraint. The derivative of the spectrum efficiency with respect to each entry in the variable matrix is deduced, and can be used as the directional matrix to update the variable matrix. The digital precoding is then performed using the classical waterfilling algorithm, which is based on the equivalent channel obtained from the product of the channel matrix and the RF precoder. Numerical simulations show that the proposed algorithm closely matches the performance of the fully digital precoding and has relatively low computational complexity.

2 System model

Consider a narrowband downlink single-user MIMO system in which a base station (BS) is equipped with N antennas and K transmit RF chains, and the user is equipped with M antennas and outputs S data streams. For limited transmit RF chains, it is not possible to implement fully digital precoding which requires one dedicated RF chain per antenna element; thus, we assume that $K \ll N$. To clearly demonstrate the contributions of the RF precoding to spectrum efficiency of BS, in this study we aim to eliminate all other interference factors and only focus on hybrid precoding on the BS side, leaving the user side with only digital combiners. For receivers with both RF and digital combiners, the proposed method can also be applied through the strategy of updating the RF and digital precoders on the BS side and the RF and digital combiners on the user side iteratively with similar operations [9]. The block diagram of the system is shown in Fig. 1.

In the hybrid precoding structure, the transmitted signal streams are first processed by low-dimensional digital precoder $\boldsymbol{W} \in \mathbb{C}^{K\times S}$, which is in a similar manner as in conventional MIMO precoding; second, the processed signal streams go through an analogue phase array, named RF precoder $\boldsymbol{V} \in \mathbb{C}^{N\times K}$, in which the modulus of each entry equals a constant, i.e. $|V_{n,m}| = 1$.

For convenience, we assume that the digital precoders on both the BS and user sides are square matrices, i.e. $S = K = M$; meanwhile, we assume that $\mathbb{E}[\boldsymbol{s}\boldsymbol{s}^{\mathrm{H}}] = \boldsymbol{I}_M$. Then, the outputs of the receiver can be modelled as

$$\hat{\boldsymbol{s}} = \boldsymbol{Q}^{\mathrm{H}}\boldsymbol{H}\boldsymbol{V}\boldsymbol{W}\boldsymbol{s} + \boldsymbol{Q}^{\mathrm{H}}\boldsymbol{z} \tag{1}$$

where $\boldsymbol{s} \in \mathbb{C}^{M\times 1}$ is the original signal vector, $\boldsymbol{Q} \in \mathbb{C}^{M\times M}$ is the digital precoder of the receiver, $\boldsymbol{H} \in \mathbb{C}^{M\times N}$ is the matrix of complex channel from the BS to the user, and $\boldsymbol{z} \sim \mathcal{CN}(0, \sigma^2\boldsymbol{I}_M)$ denotes additive white Gaussian noise.

We consider the equivalent baseband channel as $\tilde{\boldsymbol{H}} = \boldsymbol{H}\boldsymbol{V}$, then the digital combiner $\boldsymbol{Q}$ can be obtained as the left singular vectors of $\tilde{\boldsymbol{H}}$. Hence, $\boldsymbol{Q}$ is a unitary matrix, which meets $\boldsymbol{Q}^{-1} = \boldsymbol{Q}^{\mathrm{H}}$. Then, the spectrum efficiency R can be written as [9]

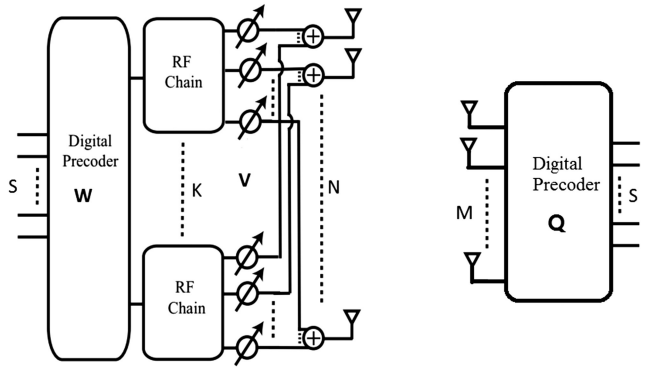

Fig. 1 *Block diagram of a P2P MIMO system*

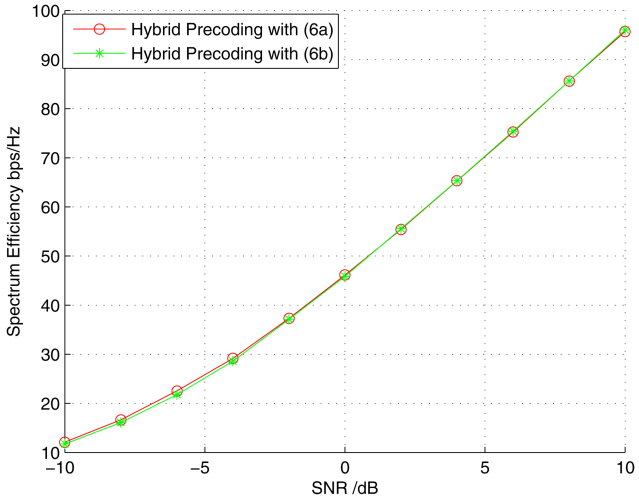

Fig. 2 *Hybrid precoding with (6a) and (6b)*

$$\begin{aligned} R &= \log_2|\boldsymbol{I} + \frac{1}{\sigma^2}\boldsymbol{Q}(\boldsymbol{Q}^{\mathrm{H}}\boldsymbol{Q})^{-1}\boldsymbol{Q}^{\mathrm{H}}\boldsymbol{H}\boldsymbol{V}\boldsymbol{W}\boldsymbol{W}^{\mathrm{H}}\boldsymbol{V}^{\mathrm{H}}\boldsymbol{H}^{\mathrm{H}}| \\ &= \log_2|\boldsymbol{I} + \frac{1}{\sigma^2}\boldsymbol{H}\boldsymbol{V}\boldsymbol{W}\boldsymbol{W}^{\mathrm{H}}\boldsymbol{V}^{\mathrm{H}}\boldsymbol{H}^{\mathrm{H}}|, \end{aligned} \tag{2}$$

where $(\cdot)^{\mathrm{H}}$ is Hermitian transpose operator.

3 Hybrid precoder design

The hybrid precoder design problem can be divided into two stages: RF precoder design and digital precoder design. Assuming perfect knowledge of the channel matrix, the precoder design problem for the BS side can be written as

$$\begin{aligned} \max_{\boldsymbol{W},\boldsymbol{V}} \quad & \log_2|\boldsymbol{I} + \frac{1}{\sigma^2}\boldsymbol{H}\boldsymbol{V}\boldsymbol{W}\boldsymbol{W}^{\mathrm{H}}\boldsymbol{V}^{\mathrm{H}}\boldsymbol{H}^{\mathrm{H}}| \\ s.t. \quad & |V_{n,m}| = 1, \quad \forall n,m \\ & \mathrm{tr}(\boldsymbol{V}\boldsymbol{W}\boldsymbol{W}^{\mathrm{H}}\boldsymbol{V}^{\mathrm{H}}) \le P \end{aligned} \tag{3}$$

where P is the power budget for the transmitting system.

3.1 RF precoder design

First, by initialising $\boldsymbol{W}\boldsymbol{W}^{\mathrm{H}} = \gamma^2\boldsymbol{I}$ [9], the RF precoder can be captured by the following optimisation problem

$$\begin{aligned} \max_{\boldsymbol{V}} \quad & \log_2|\boldsymbol{I} + \frac{\gamma^2}{\sigma^2}\boldsymbol{V}^{\mathrm{H}}\boldsymbol{H}^{\mathrm{H}}\boldsymbol{H}\boldsymbol{V}| \\ s.t. \quad & |V_{n,m}| = 1, \quad \forall n,m \end{aligned} \tag{4}$$

where γ^2 is the power of $\boldsymbol{W}\boldsymbol{W}^{\mathrm{H}}$.

In order to eliminate the constant modulus constraint for each RF precoder element, we define the RF precoder $\boldsymbol{V}$ as a function in terms of itself phase matrix

$$\boldsymbol{V} \triangleq \mathrm{e}^{\mathrm{i}\boldsymbol{X}} \tag{5}$$

where $\boldsymbol{X} \in \mathbb{R}^{N\times M}$ is the variable matrix and $\mathrm{i} = \sqrt{-1}$. The following work focuses on finding a directional matrix, along which the variable matrix $\boldsymbol{X}$ can be updated to provide better spectrum efficiency.

We now consider the solutions of the following two optimisation problems in (6):

$$\max_{\boldsymbol{V}} |\boldsymbol{V}^{\mathrm{H}}\boldsymbol{H}^{\mathrm{H}}\boldsymbol{H}\boldsymbol{V}| \tag{6a}$$

$$\max_{\boldsymbol{V}} |\boldsymbol{I} + \frac{\gamma^2}{\sigma^2}\boldsymbol{V}^{\mathrm{H}}\boldsymbol{H}^{\mathrm{H}}\boldsymbol{H}\boldsymbol{V}| \tag{6b}$$

For the reason that the solution in (6a) is try to make $\boldsymbol{V}^{\mathrm{H}}\boldsymbol{H}^{\mathrm{H}}\boldsymbol{H}\boldsymbol{V} \propto \boldsymbol{I}$, while the solution in (6b) has the same performance, the solutions of (6a) and (6b) are asymptotic consistent. In fact, the performances of hybrid precoding with (6a) and (6b) under the configurations of $(M, N) = (16, 128)$ are shown in Fig. 2.

Hence, we consider the two solutions are almost asymptotic equality, especially when the SNR or the $\min(M, N)$ is large. Thus, we focus on finding the solution of (6a) instead of (6b). For convenience, we define the objective function of (6a) as

$$f(\boldsymbol{X}) \triangleq |\boldsymbol{V}^{\mathrm{H}}\boldsymbol{H}^{\mathrm{H}}\boldsymbol{H}\boldsymbol{V}| \tag{7}$$

We first obtain the derivative of $f(\boldsymbol{X})$ with respect to $\boldsymbol{V}$ as follows (see the Appendix):

$$\begin{aligned} \boldsymbol{G} &\triangleq \frac{\partial f(\boldsymbol{X})}{\partial \boldsymbol{V}} \\ &= |\boldsymbol{V}^{\mathrm{H}}\boldsymbol{H}^{\mathrm{H}}\boldsymbol{H}\boldsymbol{V}|((\boldsymbol{V}^{\mathrm{H}}\boldsymbol{H}^{\mathrm{H}}\boldsymbol{H}\boldsymbol{V})^{-1}\boldsymbol{V}^{\mathrm{H}}\boldsymbol{H}^{\mathrm{H}}\boldsymbol{H})^{\mathrm{T}} \\ &= \alpha((\boldsymbol{H}\boldsymbol{V})^{-1}\boldsymbol{H})^{\mathrm{T}} \end{aligned} \tag{8}$$

where $\alpha = |\boldsymbol{V}^{\mathrm{H}}\boldsymbol{H}^{\mathrm{H}}\boldsymbol{H}\boldsymbol{V}| > 0$.

Then, according to the derivative chain rule for complex numbers, the derivative of the objective function with respect to the variable matrix $\boldsymbol{X}$, denoted by $\boldsymbol{D}$ can be obtained. The (n, m)th entry of $\boldsymbol{D}$, denoted by $D_{n,m}$, can be expressed as follows:

$$\begin{aligned} D_{n,m} &\triangleq \frac{\partial f(\boldsymbol{X})}{\partial X_{n,m}} \\ &= \mathrm{tr}\left(\left(\frac{\partial f(\boldsymbol{X})}{\partial \boldsymbol{V}}\right)^{\mathrm{T}}\frac{\partial \boldsymbol{V}}{\partial V_{n,m}}\right)\frac{\partial V_{n,m}}{\partial X_{n,m}} \\ &= \mathrm{tr}(\boldsymbol{G}^{\mathrm{T}}\boldsymbol{P}_{n,m})iV_{n,m} \\ &= G_{n,m}iV_{n,m} \end{aligned} \tag{9}$$

where $\boldsymbol{P}_{n,m} \in \mathbb{R}^{N\times M}$ is the single-entry matrix (1 at (n, m) and zero elsewhere), $A_{n,m}$ is the (n, m)th entry of the matrix $\boldsymbol{A}$ and $(\cdot)^{\mathrm{T}}$ is the transpose operator.

Then, the directional matrix $\boldsymbol{D}$ can be obtained easily as follows:

$$\boldsymbol{D} = \mathrm{i}\boldsymbol{G} \circ \boldsymbol{V}, \tag{10}$$

where $\circ$ is Hadamard (elementwise) product.

We have now obtained the directional matrix $\boldsymbol{D}$ for the variable matrix $\boldsymbol{X}$. Since the objective function is $f: \mathbb{R}^{N\times M} \to \mathbb{R}$, while its derivative $f' \in \mathbb{C}^{N\times M}$, the directional matrix should be projected to $\mathbb{R}^{N\times M}$ from $\mathbb{C}^{N\times M}$, which exactly is the real part of the directional matrix, so that the updated $\boldsymbol{X}$ will always be a real matrix.

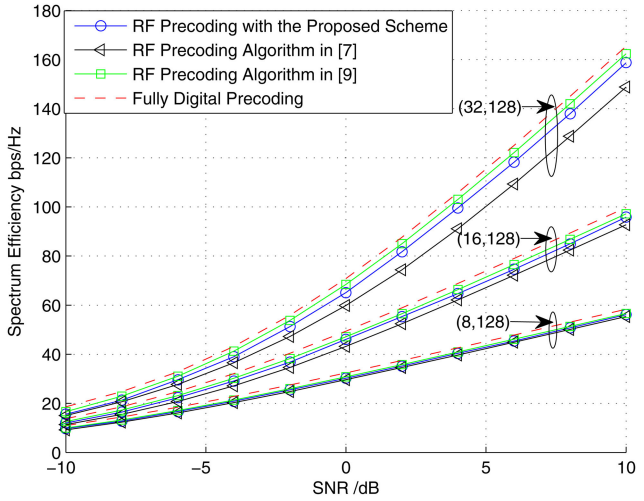

Fig. 3 *Performance of different precoders under different antenna configurations*

Table 1 Computational analysis

Algorithms	Computational complexities
proposed algorithm	$\mathcal{O}(MN^2 + 2M^3)$
algorithm in [9]	$\mathcal{O}(4M^2N^2 + 2M^3N + M^4)$

Meanwhile, in order to make each step not too large, we normalise the directional matrix by its two-norm value. Then, the update equation for the variable matrix can be formulated as follows:

$$\boldsymbol{X}_{t+1} = \boldsymbol{X}_t + \delta_t \frac{\Re \boldsymbol{D}_t}{||\boldsymbol{D}_t||}, \tag{11}$$

where δ_t is the tth step size, which can be obtained by the revised Armijo line search method (see Algorithm 1) and $\Re(\cdot)$ represents the real part.

The initialisation of $\boldsymbol{X}_0$ can be a random real matrix; the update can be stopped when the iteration time is greater than the threshold time, or when the two-norm of the directional matrix is sufficiently small, which can be expressed as $t > T_{\text{threshold}}$ or $||\boldsymbol{D}_t||_F^2 < \varepsilon$, respectively.

In brief, the design of RF precoder $\boldsymbol{V}$ is shown in Algorithm 1, where δ_0 is a small feasible step size and β is lightly >1, i.e. $\beta = 1.1$.

Algorithm 1: Design of RF precoder $\boldsymbol{V}$
Require: $\boldsymbol{H}$, δ_0, β

1: **Initialise:** $\boldsymbol{X}_0 \in \mathbb{R}^{N \times M}$ is a random matrix
2: **while** $t \leq T_{\text{threshold}}$ and $||\boldsymbol{D}_t||_F^2 \geq \varepsilon$ **do**
3: Calculate $\boldsymbol{V}_t = \mathrm{e}^{\mathrm{i}\boldsymbol{X}_t}$.
4: Calculate $\boldsymbol{D}_t = \mathrm{i}((\boldsymbol{HV})^{-1}\boldsymbol{H})^{\mathrm{T}} \circ \boldsymbol{V}_t$.
5: Initialise $\delta_t = \delta_0$.
6: // Find the optimal step size.
7: $F = f(\boldsymbol{X}_t)$, $F_1 = f\left(\boldsymbol{X}_t + \delta_t \frac{\Re \boldsymbol{D}_t}{||\boldsymbol{D}_t||_F^2}\right)$
8: **while** $F < F_1$ **do**
9: $\delta_t = \beta\delta_t$,
10: $F = F_1$,
11: $F_1 = f\left(\boldsymbol{X}_t + \delta_t \frac{\Re \boldsymbol{D}_t}{||\boldsymbol{D}_t||_F^2}\right)$.
12: **end while**
13: Update $\boldsymbol{X}_{t+1} = \boldsymbol{X}_t + \delta_t \frac{\Re \boldsymbol{D}_t}{||\boldsymbol{D}_t||_F^2}$.
14: **end while**

3.2 Digital precoder design

After obtaining the RF precoder $\boldsymbol{V}$, we consider the product of the real channel matrix and the RF precoder as the effective channel $\tilde{\boldsymbol{H}}$, which can be written as

$$\tilde{\boldsymbol{H}} = \boldsymbol{HV}, \tag{12}$$

Substituting the effective channel matrix into (2), we can obtain the digital precoder design problem as follows:

$$\begin{aligned} \max_{\boldsymbol{W}} \quad & \log_2|\boldsymbol{I} + \frac{1}{\sigma^2}\tilde{\boldsymbol{H}}\boldsymbol{W}\boldsymbol{W}^{\mathrm{H}}\tilde{\boldsymbol{H}}^{\mathrm{H}}| \\ \text{s.t.} \quad & \mathrm{tr}(\boldsymbol{VWW}^{\mathrm{H}}\boldsymbol{V}^{\mathrm{H}}) \leq P \end{aligned} \tag{13}$$

The problem has a well-known waterfilling solution [10] as

$$\boldsymbol{W} = \boldsymbol{U}\boldsymbol{\Gamma} \tag{14}$$

where $\boldsymbol{U}$ is the set of right singular vectors of $\tilde{\boldsymbol{H}}$, $\boldsymbol{\Gamma}$ is a diagonal matrix and the diagonal entries are the square roots of the power for each stream.

4 Simulation

We numerically compare our proposed NMAM precoding scheme against the CTCPA precoding scheme (which considers the Hermitian transpose of the channel phase array as the RF precoding, i.e. $V_{n,m}^{\text{CTCPA}} = H_{m,n}^*/|H_{m,n}|$) and Foad's method [9] in Fig. 3. As a reference of the performance, we also simulated the scheme of fully digital precoding, in which every antenna has a dedicated RF chain. In the simulation, the propagation environment between the BS and the user is modelled as a Rayleigh channel, in which each element of $\boldsymbol{H}$ is an independent and identically distributed (i.i.d.) complex Gaussian random variable with unit variance and zero mean, i.e. $H_{m,n} \sim \mathcal{CN}(0, 1)$.

4.1 Performance analysis

The performances of different schemes under three different pairs of the transmitting antennas N and the receiving antennas M, i.e. $(M, N) = (8, 128)$, $(16, 128)$ and $(32, 128)$, were simulated and we set $\delta_0 = 1$.

The simulation results in Fig. 3 show that compared with the performance of the RF precoding algorithm in [7], the proposed NMAM scheme is more closely emulates the spectrum efficiency of fully digital precoding; moreover, their gaps becomes larger as either the SNR or the number of $\min(M, N)$ increases. Meanwhile, the performance of the proposed NMAM scheme almost matches that of RF precoding algorithm in [9]. However, the algorithm in [9] uses an element-wise iteration to update the RF precoder, while the proposed NMAM scheme can update the entire RF precoder at one time; thus, the computational complexity of the method in [9] is significantly higher than that of the proposed method.

4.2 Computational complexity analysis

Assume that arithmetic with individual elements has complexity $\mathcal{O}(1)$, and we ignore the operations of addition and subtraction. The computational complexity of the matrix multiplication with one $n \times m$ matrix and one $m \times p$ matrix is $\mathcal{O}(nmp)$; the computational complexity of the $n \times n$ matrix inversion is $\mathcal{O}(n^3)$ (assume we adopt the Gauss–Jordan elimination method). Then the comparisons of computational complexity between the proposed algorithm and the algorithm in [9] are shown in Table 1.

Table 1 clearly shows that the computational complexity of the algorithm in [9] is at least M times higher than the proposed algorithm, while the performances of these algorithms are nearly the same. Thus, the efficiency of the proposed algorithm is at least M times higher than that of the algorithm in [9].

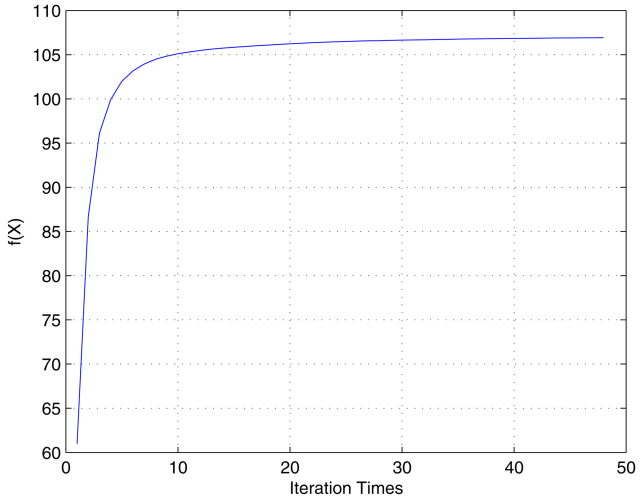

Fig. 4 *RF precoder objective function $f(\boldsymbol{X})$ versus iteration times*

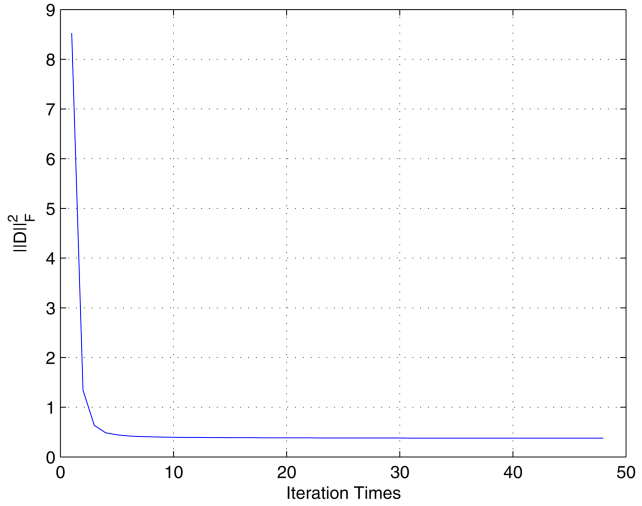

Fig. 5 *Two-norm of the directional matrix $\|\boldsymbol{D}\|_F^2$ versus iteration times*

4.3 Convergence analysis

To examine the convergence rate of the proposed algorithm, we considered a P2P MIMO system with the antenna configuration of $(M,\ N) = (16,\ 128)$. We set the SNR = 0 dB, $\delta_0 = 1$, $\beta = 1.1$, $T_{\text{threshold}} = 50$ and $\varepsilon = 0.1$. Besides, the channel model is same with that in the previous section. The RF precoder objective function $f(\boldsymbol{X})$ and two-norm of the directional matrix $\|\boldsymbol{D}\|_F^2$ versus iteration times are shown in Figs. 4 and 5, respectively.

From Fig. 4, we can see that the proposed algorithm reaches the convergence at about the 20th iteration, and Fig. 5 shows that the two-norm of the directional matrix almost keeps a constant value ($\simeq 0.4$) after the 10th iteration. It means that by taking this simulation as a prior experience, we next time only need to set the $T_{\text{threshold}} = 20$ and $\varepsilon = 0.4$, which yields a quite fast convergence rate.

5 Conclusion

This paper considered a hybrid precoding architecture for the P2P MIMO wireless communication system with large-scale antenna arrays. By deducing the derivative of the spectrum efficiency with respect to the phase matrix variable of the RF precoder, we proposed an efficient RF precoding iterative algorithm, NMAM, based on an array adaptive method. Simulation showed that the proposed method approaches the performance of fully digital precoding; moreover, its computational complexity is low, which makes it a candidate scheme for the design of the hybrid precoding.

6 References

[1] Yong, S.K., Chong, C.C.: 'An overview of multigigabit wireless through millimeter wave technology: potentials and technical challenges', *EURASIP J. Wirel. Commun. Netw.*, 2006, **2007**, (1), 078907, doi: 10.1155/2007/78907

[2] Pi, Z.Y., Khan, F.: 'An introduction to millimeter-wave mobile broadband systems', *IEEE Commun. Mag.*, 2011, **49**, (6), pp. 101–107

[3] Rappaport, T.S., Sun, S., Mayzus, R., *et al.*: 'Millimeter wave mobile communications for 5G cellular: it will work!', *IEEE Access*, 2013, **1**, pp. 335–349

[4] Alkhateeb, A., Leus, G., Heath, R.W.Jr.: 'Limited feedback hybrid precoding for multi-user millimeter wave systems', *IEEE Trans. Wirel. Commun.*, 2015, **14**, (11), pp. 6481–6494

[5] Alkhateeb, A., El Ayach, O., Leus, G., *et al.*: 'Hybrid precoding for millimeter wave cellular systems with partial channel knowledge'. IEEE Information Theory and Applications Workshop (ITA), February 2013, pp. 1–5

[6] El Ayach, O., Rajagopal, S., Abu-Surra, S., *et al.*: 'Spatially sparse precoding in millimeter wave MIMO systems', *IEEE Trans. Wirel. Commun.*, 2014, **13**, (3), pp. 1499–1513

[7] Liang, L., Xu, W., Dong, X.D.: 'Low-complexity hybrid precoding in massive multiuser MIMO systems', *IEEE Wirel. Commun. Lett.*, 2014, **3**, (6), pp. 653–656

[8] Gao, X., Dai, L., Han, S., *et al.*: 'Energy-efficient hybrid analog and digital precoding for mmWave MIMO systems with large antenna arrays', *IEEE J. Sel. Areas Commun.*, 2016, **34**, (4), pp. 998–1009

[9] Sohrabi, F., Yu, W.: 'Hybrid digital and analog beamforming design for large-scale antenna arrays', *IEEE J. Sel. Top. Signal Process.*, 2016, **10**, (3), pp. 501–513

[10] Jindal, N., Rhee, W., Vishwanath, S., *et al.*: 'Sum power iterative water-filling for multi-antenna Gaussian broadcast channels', *IEEE Trans. Inf. Theory*, 2005, **51**, (4), pp. 1570–1580

7 Appendix

The details of the proof of (8) are shown as follows:

$$\begin{aligned} \partial f(\boldsymbol{X}) &= |\boldsymbol{V}^{\mathrm{H}}\boldsymbol{H}^{\mathrm{H}}\boldsymbol{H}\boldsymbol{V}|\,\mathrm{tr}((\boldsymbol{V}^{\mathrm{H}}\boldsymbol{H}^{\mathrm{H}}\boldsymbol{H}\boldsymbol{V})^{-1}\partial(\boldsymbol{V}^{\mathrm{H}}\boldsymbol{H}^{\mathrm{H}}\boldsymbol{H}\boldsymbol{V})) \\ &= |\boldsymbol{V}^{\mathrm{H}}\boldsymbol{H}^{\mathrm{H}}\boldsymbol{H}\boldsymbol{V}|\,\mathrm{tr}((\boldsymbol{V}^{\mathrm{H}}\boldsymbol{H}^{\mathrm{H}}\boldsymbol{H}\boldsymbol{V})^{-1}(\partial\boldsymbol{V}^{\mathrm{H}})\boldsymbol{H}^{\mathrm{H}}\boldsymbol{H}\boldsymbol{V} \\ &\quad +(\boldsymbol{V}^{\mathrm{H}}\boldsymbol{H}^{\mathrm{H}}\boldsymbol{H}\boldsymbol{V})^{-1}\boldsymbol{V}^{\mathrm{H}}\boldsymbol{H}^{\mathrm{H}}\boldsymbol{H}\partial\boldsymbol{V}) \end{aligned} \tag{15}$$

First, the derivative is found with respect to the real part of $\boldsymbol{V}$

$$\begin{aligned} \frac{f(\boldsymbol{X})}{\partial\Re\boldsymbol{V}} &= |\boldsymbol{V}^{\mathrm{H}}\boldsymbol{H}^{\mathrm{H}}\boldsymbol{H}\boldsymbol{V}|\cdot\mathrm{tr}\left(\frac{(\boldsymbol{V}^{\mathrm{H}}\boldsymbol{H}^{\mathrm{H}}\boldsymbol{H}\boldsymbol{V})^{-1}(\partial\boldsymbol{V}^{\mathrm{H}})\boldsymbol{H}^{\mathrm{H}}\boldsymbol{H}\boldsymbol{V}}{\partial\Re\boldsymbol{V}}\right. \\ &\quad \left.+\frac{(\boldsymbol{V}^{\mathrm{H}}\boldsymbol{H}^{\mathrm{H}}\boldsymbol{H}\boldsymbol{V})^{-1}\boldsymbol{V}^{\mathrm{H}}\boldsymbol{H}^{\mathrm{H}}\boldsymbol{H}\partial\boldsymbol{V}}{\partial\Re\boldsymbol{V}}\right) \\ &= |\boldsymbol{V}^{\mathrm{H}}\boldsymbol{H}^{\mathrm{H}}\boldsymbol{H}\boldsymbol{V}|(\boldsymbol{H}^{\mathrm{H}}\boldsymbol{H}\boldsymbol{V}(\boldsymbol{V}^{\mathrm{H}}\boldsymbol{H}^{\mathrm{H}}\boldsymbol{H}\boldsymbol{V})^{-1} \\ &\quad +((\boldsymbol{V}^{\mathrm{H}}\boldsymbol{H}^{\mathrm{H}}\boldsymbol{H}\boldsymbol{V})^{-1}\boldsymbol{V}^{\mathrm{H}}\boldsymbol{H}^{\mathrm{H}}\boldsymbol{H})^{\mathrm{T}}) \end{aligned} \tag{16}$$

In addition, the derivative is found with respect to the imaginary part of $\boldsymbol{V}$

$$\begin{aligned} \mathrm{i}\frac{f(\boldsymbol{X})}{\partial\Im\boldsymbol{V}} &= |\boldsymbol{V}^{\mathrm{H}}\boldsymbol{H}^{\mathrm{H}}\boldsymbol{H}\boldsymbol{V}|\cdot\mathrm{tr}\left(\mathrm{i}\frac{(\boldsymbol{V}^{\mathrm{H}}\boldsymbol{H}^{\mathrm{H}}\boldsymbol{H}\boldsymbol{V})^{-1}(\partial\boldsymbol{V}^{\mathrm{H}})\boldsymbol{H}^{\mathrm{H}}\boldsymbol{H}\boldsymbol{V}}{\partial\Im\boldsymbol{V}}\right. \\ &\quad \left.+\mathrm{i}\frac{(\boldsymbol{V}^{\mathrm{H}}\boldsymbol{H}^{\mathrm{H}}\boldsymbol{H}\boldsymbol{V})^{-1}\boldsymbol{V}^{\mathrm{H}}\boldsymbol{H}^{\mathrm{H}}\boldsymbol{H}\partial\boldsymbol{V}}{\partial\Im\boldsymbol{V}}\right) \\ &= |\boldsymbol{V}^{\mathrm{H}}\boldsymbol{H}^{\mathrm{H}}\boldsymbol{H}\boldsymbol{V}|(\boldsymbol{H}^{\mathrm{H}}\boldsymbol{H}\boldsymbol{V}(\boldsymbol{V}^{\mathrm{H}}\boldsymbol{H}^{\mathrm{H}}\boldsymbol{H}\boldsymbol{V})^{-1} \\ &\quad -((\boldsymbol{V}^{\mathrm{H}}\boldsymbol{H}^{\mathrm{H}}\boldsymbol{H}\boldsymbol{V})^{-1}\boldsymbol{V}^{\mathrm{H}}\boldsymbol{H}^{\mathrm{H}}\boldsymbol{H})^{\mathrm{T}}) \end{aligned} \tag{17}$$

Hence, derivative yields

$$\begin{aligned} \frac{\partial f(\boldsymbol{X})}{\partial\boldsymbol{V}} &= \frac{1}{2}\left(\frac{\partial f(\boldsymbol{X})}{\partial\Re\boldsymbol{V}}-\mathrm{i}\frac{\partial f(\boldsymbol{X})}{\partial\Im\boldsymbol{V}}\right) \\ &= |\boldsymbol{V}^{\mathrm{H}}\boldsymbol{H}^{\mathrm{H}}\boldsymbol{H}\boldsymbol{V}|((\boldsymbol{H}\boldsymbol{V})^{-1}\boldsymbol{H})^{\mathrm{T}} \end{aligned} \tag{18}$$