An adaptive channel estimation algorithm for millimeter wave cellular systems

LU Wenlü^{1,2}, ZOU Weixia^{1,2}, LIU Xuefeng^{1,2}

1. Key Laboratory of Universal Wireless Communications, Ministry of Education, Beijing 100876, China

2. State Key Laboratory of Millimeter Waves, Southeast University, Nanjing 210096, China

Abstract: The large bandwidth available with mmWave (millimeter Wave) makes it a promising candidate for 5th generation cellular networks. Proper channel estimation algorithms must be developed to enable beamforming in mmWave systems. In this paper, we propose an adaptive channel estimation algorithm that exploits the poor scattering nature of the mmWave channel and adjusts the training overhead adaptively with the change of channel quality for mmWave cellular systems. First, we use a short training sequence to estimate the channel parameters based on the two-dimensional discrete Fourier transform method. Then, we design a feedback scheme to adjust the length of the training sequence under the premise of ensuring the accuracy of the channel estimation. The key threshold in the feedback scheme is derived and its influence on the accuracy of the estimation results is analyzed. Simulation results confirm that the proposed algorithm can adjust the length of the training sequence adaptively according to the current channel condition maintaining a stable estimation accuracy.

Key words: millimeter wave, 5G, cellular system, large antenna array, adaptive channel estimation.

Citation: LU W L, ZOU W X, LIU X F, et al. An adaptive channel estimation algorithm for millimeter wave cellular systems[J]. Journal of communications and information networks, 2016, 1(2): 37-44

1 Introduction

The data rate requirements of 5G (the next generation) cellular communication systems will approach or surpass gigabits per second^[1]. Such a high data rate inevitably relies on an enormous available bandwidth^[2]. The huge bandwidth available at mmWave frequencies makes it one of the best candidates for future 5G cellular systems^[3-4]. However, extreme path fading restricts the propagation distance of mmWave. Directional beamforming with large antenna arrays

appears to be inevitable to support longer outdoor links and provide sufficient received signal power^[5]. Moreover, channel estimation must be performed to obtain the perfect CSI (Channel State Information) necessary to support the directional beamforming. However, the channel matrix can be considerably large in mmWave systems owing to the large number of antennas, making the classical approach unfeasible for the estimation of the mmWave channel^[6,7]. Therefore, a major challenge is to develop accurate and reliable channel estimation algorithms for

Manuscript received Jun. 23, 2016; accepted Aug. 3, 2016

This work is supported by The National High Technology Research and Development Program of China (863 Program) (No.2015AA01A703), The Fundamental Research Funds for the Central Universities (No.2014ZD03-02), The National Natural Science Foundation of China (No.61571055), fund of State Key Laboratory of Millimeter Wave (No. K201501).

mmWave cellular systems^[8].

Considering the high cost and power consumption of devices and enabling multiple data streams, the hybrid structure of large scale antennas that enable hybrid analog-digital beamforming have become a new trend^[2,9]. Several estimation algorithms for hybrid structures have been developed. In Refs.[6] and [10], a novel multiresolution codebook, which is divided into analog and digital domains, is designed and a channel estimation algorithm is proposed. In Refs. [11] and [12], the authors use a compressed sensing tool to compress the training overhead and propose corresponding algorithms. A matrix-block algorithm (DFT-CEA) based on 2D-DFT (Two-Dimensional Discrete Fourier Transform) with a short training overhead to estimate the channel parameters is proposed in Ref.[7]. All these algorithms can save the training sequence, however, they cannot adjust the training overhead adaptively when the channel condition (such as SNR) changes.

In this paper, considering that the locations of the MS (Mobile Stations) relative to the BS (Base Station) are random and channel qualities at different locations differ significantly, we develop a channel estimation algorithm that adjusts the training sequence length according to the channel condition. First, the BS sends a short training sequence; after computing the 2D-DFT of the received signal, the MS estimates the path parameters iteratively, namely path gain, AoA (Angle of Arrival), and AoD (Angle of Departure). Then, we design a feedback scheme that allows the MS to distinguish if the estimated path parameter is noise or a true path by comparing with a threshold and feedbacks “Yes/No” to decide whether to continue to send the training sequence. We emphasize that the proposed first step refers to the DFT-ACE method in Ref.[7]; the difference is that we design an adaptive feedback scheme by which the system adjusts the training overhead. Numerical results confirm that the proposed algorithm provides

effective performance in mmWave systems.

Notation: $\mathbf{A}$ is a matrix, $\mathbf{a}$ is a vector, and a is a scalar. $\mathbf{A}^T$, $\mathbf{A}^H$, $\mathbf{A}^{-1}$, and $\|\mathbf{A}\|$ represent the transpose, conjugate transpose, inverse, and Frobenius norm, respectively. $[\mathbf{y}]_{a:b}$ denotes a vector obtained by extracting elements of vector $\mathbf{y}$ from index a to index b . $\mathbf{0}_{n \times m}$ denotes a zero matrix of size $N \times M$.

2 System model

We consider the single user mmWave system with hybrid analog-digital beamforming illustrated in Fig.1. A BS with M antennas and M_{RF} RF (Radio Frequency) chains communicates with the MS with N antennas and N_{RF} RF chains. The BS transmits M_s streams to the MS and the MS obtains N_s streams after processing; in general, $N_s = M_s$. According to Refs.[2,13], we know that $M_s \leq M_{\text{RF}} \leq M$ and $N_s \leq N_{\text{RF}} \leq N$. We focus on the downlink transmission. The BS is assumed to apply an $M_{\text{RF}} \times M_s$ baseband beamformer $\mathbf{F}_{\text{BB}}$ followed by an $M \times M_{\text{RF}}$ RF beamformer $\mathbf{F}_{\text{RF}}$. Similarly, the MS is constituted of an $N_{\text{RF}} \times N_s$ baseband combiner $\mathbf{W}_{\text{BB}}$ and an $N \times N_{\text{RF}}$ RF combiner $\mathbf{W}_{\text{RF}}$. To simplify the hardware implementation, each element of $\mathbf{F}_{\text{RF}}$ and $\mathbf{W}_{\text{RF}}$ has unitary magnitude; however, it may have an arbitrary phase.

$\mathbf{H}$ denotes the $N \times M$ channel matrix and $\mathbf{x}_{\text{BB}} = \mathbf{F}_{\text{BB}} \mathbf{x}$ and $\mathbf{x}_{\text{RF}} = \mathbf{F}_{\text{RF}} \mathbf{x}_{\text{BB}}$, where $\mathbf{x}$ is the $M_s \times 1$ vector of the transmitted symbols. We adopt a narrowband block-fading channel model and $\mathbf{y}_{\text{BB}}$ at the MS is

$$\mathbf{y}_{\text{BB}} = \mathbf{W}_{\text{RF}}^H (\mathbf{H} \mathbf{F}_{\text{RF}} \mathbf{x}_{\text{BB}} + \mathbf{n}), \quad (1)$$

where $\mathbf{n} \sim \mathcal{CN}(\mathbf{0}_{N \times 1}, \sigma^2 \mathbf{I}_N)$ is the thermal noise with variance σ^2 . The final processed data at the MS is $\mathbf{y} = \mathbf{W}_{\text{BB}} \mathbf{y}_{\text{BB}}$.

Considering the limited scattering nature of the mmWave channel^[14], we adopt a geometric channel model with L dominant scatterers where each scatterer is assumed to contribute a single propagation path between the BS and MS^[2]. L is a statistics mean

Figure 1 Block diagram of mmWave system that uses RF and baseband beamformers at both ends^[2,9]

and we assume it is resolved. $\mathbf{H}$ can be expressed as

$$\mathbf{H} = \frac{1}{\sqrt{L}} \sum_{l=1}^L g_l \mathbf{a}_{MS}(\phi_l^r) \mathbf{a}_{BS}^H(\phi_l^t), \quad (2)$$

where $\phi_l^t \in [0, 2\pi]$, $\phi_l^r \in [0, 2\pi]$, and $g_l \sim \mathcal{CN}(0, 1)$ are the l -th path's azimuth AoD (Angle of Departure), azimuth AoA (Angle of Arrival), and complex random gain, respectively. $\mathbf{a}_{BS}(\phi_l^t)$ and $\mathbf{a}_{MS}(\phi_l^r)$ are the antenna array response vectors at the BS and MS, respectively. If a ULA (Uniform Linear Array) with M antennas is assumed

$$\mathbf{a}_{BS}(\phi_l^t) = \left[1, e^{j(\frac{2\pi}{\lambda})d \sin \phi_l^t}, \dots, e^{j(\frac{2\pi}{\lambda})(M-1)d \sin \phi_l^t} \right]^T, \quad (3)$$

where λ is the carrier wavelength and d is the distance between antenna elements. In this paper, we assume $d = \lambda/2$. We rewrite Eq.(2) in a more compact form as

$$\mathbf{H} = \frac{1}{\sqrt{L}} \mathbf{A}_{MS} \text{diag}(\mathbf{g}) \mathbf{A}_{BS}^H, \quad (4)$$

where $\mathbf{g} = [g_1, \dots, g_L]^T$, $\mathbf{A}_{BS} = [\mathbf{a}_{BS}(\phi_1^t), \dots, \mathbf{a}_{BS}(\phi_L^t)]$, and $\mathbf{A}_{MS} = [\mathbf{a}_{MS}(\phi_1^r), \dots, \mathbf{a}_{MS}(\phi_L^r)]$.

In mmWave systems, the dimension of $\mathbf{H}$ can be extremely large, up to hundreds of columns and rows. However, its path parameters are only $2L$ angles and L complex gains. If only these parameters are estimated, the entire matrix can be constructed according to Eq.(2).

3 Adaptive channel estimation

A. Estimate parameters based on DFT-CEA method

Eq.(2) indicates that each element of $\mathbf{H}$ contains all the path parameter information. We can estimate a part of $\mathbf{H}$ and extract the path parameters. We represent the submatrix as

$$\tilde{\mathbf{H}} = [\mathbf{H}]_{0:N_d-1, 0:M_d-1}, \quad (5)$$

where $N_d \leq N$, $M_d \leq M$, and N_d and M_d are two suitable integer parameters depending on the length of the training sequence. Further $\tilde{\mathbf{H}}$ can be divided into submatrix blocks

$$\tilde{\mathbf{H}}^{(p,q)} = [\tilde{\mathbf{H}}]_{(p-1)N_{RF}:pN_{RF}-1, (q-1)M_{RF}:qM_{RF}-1}, \quad (6)$$

where $p=1, 2, \dots, \lceil N_d/N_{RF} \rceil$ and $q=1, 2, \dots, \lceil M_d/M_{RF} \rceil$. To estimate a submatrix block $\tilde{\mathbf{H}}^{(p,q)}$, BS sends the vectors $\mathbf{x}_{BB}^{(m)}$, $m=0, 1, 2, \dots, M_{RF}-1$ successively. $\mathbf{x}_{BB}^{(m)}$ has $1/\sqrt{M_{RF}}$ in position and zero otherwise. For each vector of this sequence, the analog RF beamformer $\mathbf{F}_{RF}^{(q)}$ is designed as follows:

$$\mathbf{F}_{RF}^{(q)} = \begin{bmatrix} \mathbf{0}_{(q-1)M_{RF} \times M_{RF}} \\ \bar{\mathbf{F}}_{RF} \\ \mathbf{0}_{(M-qM_{RF}) \times M_{RF}} \end{bmatrix}, \quad (7)$$

where $\bar{\mathbf{F}}_{\text{RF}}$ is an $M_{\text{RF}} \times M_{\text{RF}}$ square matrix. Similarly, the analog combiner is designed as

$$\mathbf{W}_{\text{RF}}^{(p)} = \begin{bmatrix} \mathbf{0}_{(p-1)N_{\text{RF}} \times N_{\text{RF}}} \\ \bar{\mathbf{W}}_{\text{RF}} \\ \mathbf{0}_{(M-p)N_{\text{RF}} \times N_{\text{RF}}} \end{bmatrix}, \quad (8)$$

where $\bar{\mathbf{W}}_{\text{RF}}$ is an $N_{\text{RF}} \times N_{\text{RF}}$ square matrix. $\bar{\mathbf{F}}_{\text{RF}}$ and $\bar{\mathbf{W}}_{\text{RF}}$ are full-rank and have elements with unitary magnitude and arbitrary phases. According to Eq.(1), the baseband signal at MS is

$$\begin{aligned} \mathbf{y}_{\text{BB}}^{(p,q,m)} &= (\mathbf{W}_{\text{RF}}^{(p)})^H (\mathbf{H} \mathbf{F}_{\text{RF}}^{(q)} \mathbf{x}_{\text{BB}}^{(m)} + \mathbf{n}^{(m)}) \\ &= \bar{\mathbf{W}}_{\text{RF}}^H (\tilde{\mathbf{H}}^{(p,q)} \bar{\mathbf{F}}_{\text{RF}} \mathbf{x}_{\text{BB}}^{(m)} + \bar{\mathbf{n}}^{(m)}), \end{aligned} \quad (9)$$

where $\bar{\mathbf{n}}^{(m)} = [\mathbf{n}^{(m)}]_{(p-1)N_{\text{RF}}:pN_{\text{RF}}-1}$

When all the vectors $\mathbf{x}_{\text{BB}}^{(m)}$, $m = 0, 1, 2, \dots, M_{\text{RF}} - 1$ are sent, we write the received signal in matrix form:

$$\mathbf{Y}_{\text{BB}}^{(p,q)} = [\mathbf{y}_{\text{BB}}^{(p,q,0)}, \mathbf{y}_{\text{BB}}^{(p,q,1)}, \dots, \mathbf{y}_{\text{BB}}^{(p,q,M_{\text{RF}}-1)}]. \quad (10)$$

According to Eqs.(9) and (10), a submatrix block is estimated by

$$\begin{aligned} \hat{\mathbf{H}}^{(p,q)} &= \sqrt{M_{\text{RF}}} (\bar{\mathbf{W}}_{\text{RF}}^H)^{-1} \mathbf{Y}_{\text{BB}}^{(p,q)} (\bar{\mathbf{F}}_{\text{RF}})^{-1} \\ &= \tilde{\mathbf{H}}^{(p,q)} + \tilde{\mathbf{n}}, \end{aligned} \quad (11)$$

where $\tilde{\mathbf{n}} = \sqrt{M_{\text{RF}}} [\bar{\mathbf{n}}^{(0)}, \bar{\mathbf{n}}^{(1)}, \dots, \bar{\mathbf{n}}^{(M_{\text{RF}}-1)}] (\bar{\mathbf{F}}_{\text{RF}})^{-1}$. To ensure that $\bar{\mathbf{F}}_{\text{RF}}$ and $\bar{\mathbf{W}}_{\text{RF}}$ are full-rank and have elements with unitary magnitude, they are assumed to be Hadamard matrices; this also ensure that the noise in Eq.(11) is white with variance $\sigma^2[7]$. Estimating and stacking several submatrix blocks $\hat{\mathbf{H}}^{(p,q)}$, we can obtain an $N_d \times M_d$ estimated submatrix $\hat{\mathbf{H}}$ of $\tilde{\mathbf{H}}$.

To extract the path parameters ϕ_l^t, ϕ_l^r, g_l , $l=1, 2, \dots, L$, we compute the 2D-DFT of $\hat{\mathbf{H}}$ on $N_{\text{DFT}} \times N_{\text{DFT}}$ samples

$$T(k, i) = \sum_{m=0}^{M_d-1} \sum_{n=0}^{N_d-1} [\hat{\mathbf{H}}]_{n,m} e^{-j2\pi(im+kn)/N_{\text{DFT}}}, \quad (12)$$

Then, the path parameters can be obtained using the DFT-CEA method in Ref.[7]. For convenience, the DFT-CEA method is presented as follows. $D(k, i)$, $k, i=0, 1, \dots, N_{\text{DFT}}$ is the 2D-DFT of the $N_d \times M_d$ ones matrix on $N_{\text{DFT}} \times N_{\text{DFT}}$ samples.

Algorithm 1 DFT-CEA method

Input: $T(k, i)$, $k, i=0, 1, \dots, N_{\text{DFT}}$

1. $T_1(k, i) = T(k, i)$
 2. for $l=1$ to L
 3. $(\hat{k}_l, \hat{i}_l) = \arg \max_{k,i} |T_l(k, i)|$
 4. $\hat{g}_l = T_l(\hat{k}_l, \hat{i}_l) \sqrt{L} / (N_d M_d)$
 5. $T_{l+1}(k, i) = T_l(k, i) - \hat{g}_l D(k - \hat{k}_l, i - \hat{i}_l) / \sqrt{L}$
 6. end
 7. $\hat{\phi}_l^t = \sin^{-1}(-2\hat{i}_l / N_{\text{DFT}})$, $\hat{\phi}_l^r = \sin^{-1}(2\hat{k}_l / N_{\text{DFT}})$
-

B. Adaptive feedback scheme

Owing to the noise interference, the estimated paths are not all sufficiently accurate, especially for the paths whose gains are small. Under the same noise conditions, we can improve the estimation accuracy by increasing the dimensions of submatrix $\tilde{\mathbf{H}}$. The larger dimensions of $\tilde{\mathbf{H}}$ require a longer length training sequence; however, this results in a more accurate estimation. First, we send a short training sequence to estimate the parameters. Then, we design a feedback scheme: 1) MS determines if the estimated paths are noise jamming or real paths by comparing with a threshold. 2) Feedback “YES” is sent to BS and BS stops sending training sequences if all L estimated paths are evaluated to be real paths, otherwise feedback “NO” is sent and BS continues to send training sequences. Then, we adjust the length of the training sequence adaptively with the feedback scheme and ensure the estimation accuracy remains stable.

The key step of the DFT-CEA method is the third step, however, the point selected is probably not the real path, rather, it could be noise jamming. To distinguish the noise and real path, we explore the characteristics of the noise in $|T_l(k, i)|$.

The estimated submatrix in 3(A) is

$$\hat{\mathbf{H}} = \tilde{\mathbf{H}} + \tilde{\mathbf{n}}, \quad (13)$$

where $\tilde{\mathbf{n}}$ is the $N_d \times M_d$ noise matrix and each element obeys the complex Gaussian distribution $\mathcal{CN}(0, \sigma^2)^{[7]}$. We rewrite Eq.(12) as follows:

$$\begin{aligned}
T(k, i) &= \sum_{m=0}^{M_d-1} \sum_{n=0}^{N_d-1} ([\tilde{\mathbf{H}}]_{n,m} + [\tilde{\mathbf{n}}]_{n,m}) e^{-j2\pi(im+kn)/N_{\text{DFT}}} \\
&= \sum_{m=0}^{M_d-1} \sum_{n=0}^{N_d-1} [\tilde{\mathbf{H}}]_{n,m} e^{-j2\pi(im+kn)/N_{\text{DFT}}} + \\
&\quad \sum_{m=0}^{M_d-1} \sum_{n=0}^{N_d-1} [\tilde{\mathbf{n}}]_{n,m} e^{-j2\pi(im+kn)/N_{\text{DFT}}} \\
&= H(k, i) + \tilde{N}(k, i),
\end{aligned} \tag{14}$$

where $|\tilde{N}(k, i)|$ is the 2D-DFT of $\tilde{\mathbf{n}}$ and each element obeys the complex Gaussian distribution

$$\tilde{N}(k, i) \sim \mathcal{CN}(0, N_d M_d \sigma^2), \tag{15}$$

The mode $|\tilde{N}(k, i)|$ obeys the Rayleigh distribution; the cumulative distribution function is

$$F(x) = 1 - e^{-x^2/2N_d M_d \sigma^2}, x \geq 0, \tag{16}$$

According to the poor scattering nature of the mmWave channel^[6], only few angles, namely the AoAs and AoDs of the paths, have large energies; other angles are near zero energy. Hence, the matrix $|H(k, i)|$, $k, i = 0, 1, \dots, N_{\text{DFT}} - 1$ has the following features: 1) the values at the positions corresponding to the AoAs and AoDs are large, appearing to be “peaks”; 2) other values are near zero.

According to the distribution features of $|H(k, i)|$ and $|\tilde{N}(k, i)|$ described above, we obtain

$$|T(k, i)| = |H(k, i) + \tilde{N}(k, i)| \approx |H(k, i)| + |\tilde{N}(k, i)|. \tag{17}$$

The third step of the DFT-CEA method is

$$\begin{aligned}
(\hat{k}_l, \hat{i}_l) &= \max_{k,i} |T(k, i)| \\
&\approx \max_{k,i} (|H(k, i)| + |\tilde{N}(k, i)|)
\end{aligned} \tag{18}$$

To ensure the accuracy and eliminate the interference of noise, the energy of the selected point must be greater than the noise jamming, namely

$$|T(\hat{k}_l, \hat{i}_l)| \geq \max_{k,i} |T(k, i)| \approx \max_{k,i} |\tilde{N}(k, i)|. \tag{19}$$

However, Eq.(19) is overly difficult to satisfy. In this paper, we adopt a proper threshold N_{thres} to replace $\max_{k,i} (|\tilde{N}(k, i)|)$, and the determining condition

becomes

$$|T(\hat{k}_l, \hat{i}_l)| > N_{\text{thres}}. \tag{20}$$

$|\tilde{N}(k, i)|$ obeys the Rayleigh distribution and from Eq.(16) we take a proper value x_0 satisfying

$$F(x_0) = 1 - e^{-x_0^2/2N_d M_d \sigma^2} \rightarrow 1. \tag{21}$$

In this paper, we let

$$x_0 = 3\sigma\sqrt{N_d M_d}, \tag{22}$$

and $F(x_0) = 1 - e^{-x_0^2/2N_d M_d \sigma^2} = 0.9889 \rightarrow 1$, ensures that x_0 is a proper threshold. To adjust the threshold conveniently, we append an adjustable coefficient

$$N_{\text{thres}} = \mu x_0, \tag{23}$$

where μ is a variable and we can tune N_{thres} by adjusting the value of μ .

MS can determine if the estimated results are reliable by comparing $|T(\hat{k}_l, \hat{i}_l)|$ and N_{thres} . If $|T(\hat{k}_l, \hat{i}_l)| > N_{\text{thres}}$, it is considered to be a reliable path; otherwise it is likely to be noise jamming.

After designing the feedback scheme, we design the proposed estimation algorithm. Initially, the dimensions of the submatrix are the minimum, namely $M_d = M_{\text{RF}}$ and $N_d = N_{\text{RF}}$. BS and MS perform the transmitting and receiving training sequence operation. MS obtains an estimated submatrix block $\hat{\mathbf{H}}^{(1,1)}$ and stacks this to obtain an $N_d \times M_d$ estimated submatrix $\hat{\mathbf{H}}$. Then, the DFT-CEA method is used to estimate the path parameters and MS determines if the estimated results are reliable. Once a path is judged to be noise jamming, BS and MS continue the transmitting and receiving operation. MS obtains several new submatrix blocks $\hat{\mathbf{H}}^{(p,q)}$ and stacks these with the previous blocks to realize a larger submatrix $\hat{\mathbf{H}}$; that is, M_d and N_d increase. Then, the larger submatrix is used to estimate the path parameters until all the estimated paths are reliable.

Once all the paths are estimated, the channel

matrix $\mathbf{H}$ can be reconstructed by Eq.(2) and we obtain the estimated channel matrix $\hat{\mathbf{H}}$.

Algorithm 2 Adaptive channel estimation algorithm with feedback scheme

Input: BS and MS know $M, N, M_{\text{RF}}, N_{\text{RF}}, L, N_{\text{DFT}}, \sigma, \rho$, and have $\bar{\mathbf{F}}_{\text{RF}}, \bar{\mathbf{W}}_{\text{RF}}$

Initialization: $M_d = M_{\text{RF}}, N_d = N_{\text{RF}}$

1. $q = M_d / M_{\text{RF}}$
 2. for $p = 1 : N_d / N_{\text{RF}}$
 3. for $m = 0 : M_{\text{RF}} - 1$
 4. BS sends $\mathbf{x}_{\text{BB}}^{(m)}$ and uses $\mathbf{F}_{\text{RF}}^{(q)}$, MS uses $\mathbf{W}_{\text{RF}}^{(p)}$
 5. calculate $\hat{\mathbf{H}}^{(p,q)}$ and stack them in $\hat{\mathbf{H}}$.
 6. if $M_d / M_{\text{RF}} > 1$
 7. $p = N_d / N_{\text{RF}}$
 8. for $q = 1 : N_d / N_{\text{RF}} - 1$
 9. for $m = 0 : M_{\text{RF}} - 1$
 10. BS sends $\mathbf{x}_{\text{BB}}^{(m)}$ and uses $\mathbf{F}_{\text{RF}}^{(q)}$, MS uses $\mathbf{W}_{\text{RF}}^{(p)}$
 11. calculate $\hat{\mathbf{H}}^{(p,q)}$, and stack them $\hat{\mathbf{H}}$
 12. calculate $T(k, i)$ and $T_1(k, i) = T(k, i)$
 13. $N_{\text{thres}} = 3\mu\sigma\sqrt{N_d M_d}$
 14. for $l = 1 : L$
 15. $(\hat{k}_l, \hat{i}_l) = \arg \max_{k, i} |T_l(k, i)|$
 16. if $T_l(\hat{k}_l, \hat{i}_l) < N_{\text{thres}} \& \& M_d < M \& \& N_d < N$
 17. $M_d = M_d + M_{\text{RF}}, N_d = N_d + N_{\text{RF}}$
 18. goto step 1.
 19. $\hat{g}_l = T_l(\hat{k}_l, \hat{i}_l) \sqrt{L} / (N_d M_d)$
 20. $T_{l+1}(k, i) = T_l(k, i) - \hat{g}_l D(k - \hat{k}_l, i - \hat{i}_l) / \sqrt{L}$
 21. $\hat{\phi}_l^i = \sin^{-1}(-2\hat{i}_l / N_{\text{DFT}}), \hat{\phi}_l^r = \sin^{-1}(2\hat{k}_l / N_{\text{DFT}})$
-

The proposed algorithm requires the length of the training sequence

$$N_{\text{TS}} = \lceil N_d / N_{\text{RF}} \rceil \lceil M_d / M_{\text{RF}} \rceil M_{\text{RF}}. \quad (24)$$

The complexity of the proposed algorithm is mainly determined by the 2D-DFT. The proposed algorithm must calculate $T(k, i)$ for M_d / M_{RF} times and each time requires a complexity of $O(N_{\text{DFT}} \log N_{\text{DFT}})$ using the fast Fourier transform algorithm.

4 Numerical results

In this section, we evaluate the performance of the proposed algorithm in a typical mmWave scenario. We consider the system model in Section 2. BS employs a ULA of $M=128$ antennas and $M_{\text{RF}}=16$ RF chains; MS employs a ULA of $N=64$ antennas and $N_{\text{RF}}=8$ RF chains. The distance between antenna elements is $d=\lambda/2$. MmWave channels with several scatterers between are considered. We set $N_{\text{DFT}}=512$ when calculating the 2D-DFT. As a comparison, we also simulate the DFT-ACE method with $M_d=80, N_d=40$; the other parameters are the same as above. Further, owing to the fact that $\mathbf{F}_{\text{RF}}^{(q)}$ switches on only M_{RF} antennas, the signal-to-noise ratio is: $SNR_d = SNR + 10 \times \lg(M_{\text{RF}}/M)$ where SNR is the signal-to-noise ratio when all M antennas are used. Since the signal power of the training vector is one, the noise variance $\sigma = 10^{-(SNR + 10 \times \lg(M_{\text{RF}}/M)/10)}$.

Fig.2 presents the curves of the average length of the training sequence versus the SNR. The average length of the training sequence is calculated by averaging Eq.(24) for 500 realizations in the same channel condition. Fig.3 displays the curves of the NMSE (Normalized Mean Square Error) versus SNR. NMSE is given by $E[\|\hat{\mathbf{H}} - \mathbf{H}\|^2 / \|\mathbf{H}\|^2]$. In Figs.2 and 3, the curves of the DFT-CEA method and proposed algorithm with four values of μ are given. From these figures, the following conclusions can be drawn:

1) The curves for $\mu=1$ and L equals 3, 4, 5, and 6 indicate that for different values of L , the trend of the curves are the same, namely that the proposed algorithm is robust for different values of L , that the training overhead increases with an increase of L , and the NMSE improves.

2) With and the $SNR > -6$ dB increasing gradually, the curves of the proposed algorithm in Fig.2 decline, indicating that the lengths of the training sequence decrease gradually. Simultaneously, the NMSE curves

Figure 2 Curves of average length of training sequence versus SNR

Figure 2 Curves of NMSE versus SNR

of the proposed algorithm in Fig.3 remain basically stable. This is consistent with our expectations: that is, as the SNR increases, the length of the training sequence decreases and the estimation accuracy remains basically stable. With $SNR < -6$ dB and the SNR decreasing, both the lengths of the training sequence and NMSE increase, meaning that the increase of the training sequence cannot, to a certain extent, offset the decrease of the estimation accuracy when the

channel condition is bad.

3) Comparing the curves $\mu=1, 1.1, 1.2, 1.5$, as μ increases, the threshold N_{thres} increases relatively, leading to an increase of the lengths of the training sequence and a decrease of NMSE. We can set a tradeoff between the length of the training sequence and the estimation accuracy by adjusting the value of μ .

4) Comparing the curves of DFT-CEA with the curves of the proposed algorithm with $\mu=1$, when SNR is near -5 dB, both the average length of the training sequence and NMSEs of the two algorithms are approximately equal. When $SNR > -5$ dB, the DFT-CEA method has a longer training sequence length than the proposed algorithm, which may be redundant. When $SNR < -5$ dB, the proposed algorithm has a longer length training sequence and superior NMSE compared to the DFT-CEA method, which may be necessary in that channel condition. Conclusively, the proposed algorithm adjusts the length of the training sequence adaptively with the change of the channel and the estimation accuracy remains stable.

5 Conclusion

In this paper, we considered the problem of channel estimation for mmWave cellular systems and proposed a novel adaptive channel estimation algorithm. First, the BS transmits a short training sequence to estimate the channel parameters based on the DFT-CEA method. Then, we designed a feedback scheme to adjust the length of training sequence, under the premise that the estimation accuracy remains stable. Simulation results confirm that the proposed algorithm has excellent performance in a single user mmWave system. In a certain range of SNR, the algorithm can adjust the length of the training sequence adaptively with the change of the channel and the estimation accuracy remains stable.

Moreover, we can set a tradeoff between the length of the training sequence and the estimation accuracy by adjusting the coefficient .

References

- [1] ROH W, SEOL J Y, PARK J, et al. Millimeter-wave beamforming as an enabling technology for 5G cellular communications: theoretical feasibility and prototype results[J]. Communications magazine, 2014, 52(2): 106-113.
- [2] AYACH O E L, RAJAGOPAL S, ABU-SURRA S, et al. Spatially sparse precoding in millimeter wave MIMO systems[J]. Wireless communications, IEEE transactions on, 2014, 13(3): 1499-1513.
- [3] PI Z, KHAN F. An introduction to millimeter-wave mobile broadband systems[J]. Communications magazine, IEEE, 2011, 49(6): 101-107.
- [4] RAPPAPORT T S, SUN S, MAYZUS R, et al. Millimeter wave mobile communications for 5G cellular: It will work![J]. Access, IEEE, 2013, 1: 335-349.
- [5] BIGLARBEIGIAN B, FAKHARZADEH M, BUSUIOC D, et al. Optimized microstrip antenna arrays for emerging millimeter-wave wireless applications[J]. Antennas and propagation, IEEE transactions on, 2011, 59(5): 1742-1747.
- [6] ALKHATEEB A, EL AYACH O, LEUS G, et al. Channel estimation and hybrid precoding for millimeter wave cellular systems[J]. Selected topics in signal processing, IEEE journal of, 2014, 8(5): 831-846.
- [7] MONTAGNER S, BENVENUTO N, BARACCA P. Channel estimation using a 2D DFT for millimeter-wave systems[C]// The IEEE 81st Vehicular Technology Conference (VTC Spring), 2015: 1-5.
- [8] RANGAN S, RAPPAPORT T S, ERKIP E. Millimeter-wave cellular wireless networks: potentials and challenges[J]. Proceedings of the IEEE, 2014, 102(3): 366-385.
- [9] ALKHATEEB A, AYACH O E L, LEUS G, et al. Hybrid precoding for millimeter wave cellular systems with partial channel knowledge[C]//Information Theory and Applications Workshop (ITA), 2013: 1-5.
- [10] ALKHATEEB A, EL AYACH O, LEUS G, et al. Single-sided adaptive estimation of multi-path millimeter wave channels[C]// IEEE 15th International Workshop on Signal Processing Advances in Wireless Communications (SPAWC), 2014: 125-129.
- [11] ALKHATEEBBY A, LEUSZ G, HEATH R W. Compressed sensing based multi-user millimeter wave systems: how many measurements are needed?[C]//IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), 2015: 2909-2913.
- [12] MO J, SCHNITER P, GONZALEZ PRELCIC N, et al. Channel estimation in millimeter wave MIMO systems with one-bit quantization[C]//IEEE 48th Asilomar Conference on Signals, Systems and Computers, 2014: 957-961.
- [13] XIA P, YONG S K, OH J, et al. Multi-stage iterative antenna training for millimeter wave communications[C]//Global Telecommunications Conference, 2008: 1-6.
- [14] MURDOCK J N, BEN-DOR E, QIAO Y, et al. A 38 GHz cellular outage study for an urban outdoor campus environment [C]// Wireless Communications and Networking Conference (WCNC), 2012: 3085-3090.

About the authors

LU Wenli is now studying at Beijing University of Posts and Telecommunications for a master degree in information and communication engineering. He received his B.S. degree in electronic information engineering from Tianjin University in 2014. His current research is in short-range wireless communication. (Email: luwenlv@bupt.edu.cn.)

ZOU Weixia [corresponding author] is currently an Associate Professor in Beijing University of Posts and Telecommunications Beijing, China. She received her B.S. degree in electric traction and drive control from Tongji University in 1994, and her M.S. degree in circuits and systems from Shandong University, in 2002, and her Ph.D.

LIU Xuefeng is now studying at Beijing University of Posts and Telecommunications for a Ph.D. degree in communication and information system. His current researches are short-range wireless communication and MIMO transmission technology.

degree in communication and information system from Beijing University of Posts and Telecommunications, in 2006. Her current research is in new technologies of short-range wireless communication and 60GHz. (Email: zwx0218@bupt.edu.cn.)