

Channel Estimation for Wideband MmWave MIMO OFDM System Exploiting Block Sparsity

Yujie Wang, *Student Member, IEEE*, Chenhao Qi, *Senior Member, IEEE*,
Ping Li, Zhaohua Lu, and Ping Lu

Abstract

In this letter, we investigate time-domain channel estimation for wideband millimeter wave (mmWave) MIMO OFDM system. By transmitting frequency-domain pilot symbols as well as different beam-forming vectors, we observe that the time-domain mmWave MIMO channels exhibit channel delay sparsity and especially block sparsity among different spatial directions. Then we propose a time-domain channel estimation exploiting block sparsity (TDCEBS) scheme, which always aims at finding the best nonzero block achieving the largest projection of the residue at each iterations. In particular, we evaluate the system performance using the QuaDRiGa which is recommended by 5G New Radio to generate wideband mmWave MIMO channels. The effectiveness of the proposed TDCEBS scheme is verified by the simulation results, as the proposed scheme outperforms the existing schemes.

Index Terms

5G, channel estimation, mmWave communications, OFDM, sparse recovery.

I. INTRODUCTION

Since the fifth generation (5G) wireless communications has been standardized in 2020s, more concern is gathered by the industry, academia and government on beyond 5G (B5G) or even

Yujie Wang and Chenhao Qi are with the School of Information Science and Engineering, Southeast University, Nanjing 210096, China (Email: qch@seu.edu.cn).

Ping Li, Zhaohua Lu, and Ping Lu are with ZTE Corporation, Shenzhen 518057, China.

6G. To support high data rate transmission, wireless communications working on millimeter wave (mmWave) frequency band is on focus. Early works on mmWave MIMO communications and signal processing consider the narrow-band mmWave channels, by assuming the delay of different channel paths is the same [1]–[3]. But later on, practical testing results show that the mmWave channels are wideband, where the delay of different channel paths are different [4]–[6]. Since the delay spread causes frequency-selective fading, orthogonal frequency division multiplexing (OFDM) is introduced to mmWave MIMO communications, just like in sub-6GHz MIMO communications [7].

To acquire accurate channel state information for mmWave MIMO beamforming, efficient channel estimation is needed. The main-stream methods for wideband mmWave MIMO can be divided into two categories, including frequency-domain channel estimation methods [8], [9] and time-domain channel estimation methods [10], [11]. In [8], a simultaneous weighted orthogonal matching pursuit (SW-OMP) channel estimation method is proposed to estimate the frequency-domain mmWave MIMO channels by exploiting the spatial sparsity that is also called angle sparsity sometimes. In [9], the received training signal is treated as a low-rank three-dimensional tensor that fits a canonical polyadic model, where a structured canonical polyadic decomposition-based channel estimation method is proposed by utilizing the Vandermonde property of factor matrices that contains the channel parameters. Compared to frequency-domain channel estimation, time-domain channel estimation is challenging, since the standard least squared (LS) methods cannot be applied due to the low rank issue. But on the other hand, time-domain channel estimation can further exploit the sparse property of the channel delay spread. In [10], a two-step time-domain channel estimation scheme is proposed, where the effective channel composed of beamformer, channel steering vectors, and pulse shaping is firstly estimated by the LS method, and then the desired channel is estimated by the orthogonal matching pursuit (OMP) algorithm. But the block sparsity is not considered by [10]. In [11], a block sparse channel estimation method based on sparse Bayesian learning (SBL) is proposed by exploiting the channel delay sparsity and block sparsity among different angles.

In this letter, we consider the time-domain channel estimation for wideband mmWave MIMO OFDM system. By transmitting frequency-domain pilot symbols as well as different beamforming vectors, we observe that the time-domain mmWave MIMO channels exhibit channel delay sparsity and especially block sparsity among different spatial directions. Then we propose a time-

domain channel estimation exploiting block sparsity (TDCEBS) scheme, which always aims at finding the best nonzero block achieving the largest projection of the residue at each iterations. In particular, we evaluate the system performance using the QuaDRiGa which is recommended by 5G New Radio to generate wideband mmWave channels.

The notations used in this letter are as follows. Symbols for matrices and vectors are written in boldface. a , \mathbf{a} , \mathbf{A} denote a scalar, a vector and a matrix, respectively, while $(*)^T$, $(*)^H$, $\|*\|_2$ denote the transpose, the conjugate transpose and the ℓ_2 norm, respectively. $\mathbf{A}(m, n)$ represents the entry located in the m th row and n th column of matrix \mathbf{A} . \mathbf{I}_K denotes a K -by- K identity matrix. $\mathcal{CN}(m, \mathbf{R})$ represents the complex Gaussian distribution whose mean is m and covariance matrix is \mathbf{R} . \mathbb{C} and \emptyset represent the set of complex-valued numbers and the empty set, respectively. \cup denotes the union of sets.

II. SYSTEM MODEL

We consider a downlink mmWave MIMO OFDM system, as shown in Fig. 1. The base station (BS) is equipped with N_{BS} antennas in uniform linear arrays. To simplify the analysis, we assume that the users are equipped with a single antenna. To tackle the frequency-selective fading in wideband mmWave channels, OFDM is typically used for the signal transmission. The total number of OFDM subcarriers is denoted by N_c and the subcarriers spacing is denoted by Δf . Then the duration of each OFDM symbol is $1/\Delta f$, which is further divided into N_c OFDM samples and the duration of each OFDM sample is

$$T_s = \frac{1}{N_c \Delta f}. \quad (1)$$

In order to eliminate inter-symbol interference and inter-carrier interference, a cyclic prefix (CP) is placed at the head of each OFDM symbol. We denote the length of the CP by N_{cp} .

We denote the number of resolvable channel paths for the mmWave MIMO system by L , including a line-of-sight (LoS) path and $L - 1$ non-line-of-sight (NLoS) paths, as shown in Fig. 1. Generally, the LoS path is the path with the largest transmission gain. However, in some scenarios such as NLoS scenario, since the LoS path is blocked by channel obstacles, the transmission gain of the LoS path may be reduced to be comparable to that of the NLoS path, or even smaller than that of the NLoS path. That is to say, the LoS path cannot provide effective transmission gain and therefore can be treated a NLoS path. We denote the channel

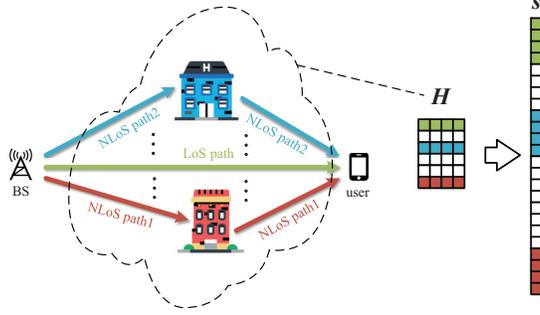


Fig. 1. Wideband mmWave MIMO system with block-sparse channels.

gain, delay and angle-of-departure (AoD) of each channel path respectively by γ_l , τ_l , and θ_l , for $l = 1, 2, \dots, L$. With these channel parameters, the Saleh-Valenzuela channel model widely used in the narrowband mmWave MIMO systems can be extended to the wideband mmWave MIMO systems [12]. Accordingly, the wideband mmWave MIMO channel can be written as

$$\mathbf{h}(t) = \sqrt{\frac{N_{\text{BS}}}{L}} \sum_{l=1}^L \gamma_l p(t - \tau_l) \boldsymbol{\alpha}^H(N_{\text{BS}}, \theta_l), \quad (2)$$

where $p(t)$ denotes the pulse shaping function and $\boldsymbol{\alpha}(N_{\text{BS}}, \theta_l)$ denotes the channel steering vector expressed as

$$\boldsymbol{\alpha}(N_{\text{BS}}, \theta_l) = \frac{1}{\sqrt{N_{\text{BS}}}} [1, e^{j\pi\theta_l}, \dots, e^{j(N_{\text{BS}}-1)\pi\theta_l}]^T. \quad (3)$$

After channel sampling, (2) can be written as

$$\mathbf{h}(n) = \sqrt{\frac{N_{\text{BS}}}{L}} \sum_{l=1}^L \gamma_l p(nT_s - \tau_l) \boldsymbol{\alpha}^H(N_{\text{BS}}, \theta_l), \quad (4)$$

for $n = 0, 1, \dots, N_{\text{cp}} - 1$, since the OFDM modulation requires the CP length larger than the maximum channel delay spread. We define a time-domain channel matrix $\mathbf{H} \in \mathbb{C}^{N_{\text{cp}} \times N_{\text{BS}}}$ as

$$\mathbf{H} \triangleq \{\mathbf{h}(0)^T, \mathbf{h}(1)^T, \dots, \mathbf{h}(N_{\text{cp}} - 1)^T\}^T. \quad (5)$$

To estimate \mathbf{H} , we transmit frequency-domain pilot symbols with the length of K ($K \leq N_c$), where the k th transmitted pilot symbol is denoted by $x(k)$ and the corresponding received pilot symbol is denoted by $y(k)$, $k = 1, 2, \dots, K$. Then we define the received pilot vector as

$$\mathbf{y} \triangleq [y(1), y(2), \dots, y(K)]^T. \quad (6)$$

We have

$$\mathbf{y} = \mathbf{XDH}\mathbf{f} + \boldsymbol{\eta}, \quad (7)$$

where $\mathbf{X} \in \mathbb{C}^{K \times K}$ is a diagonal matrix whose k th diagonal entry is $x(k)$, and $\boldsymbol{\eta} \in \mathbb{C}^K$ represents an additive white Gaussian noise (AWGN) vector with $\boldsymbol{\eta} \sim \mathcal{CN}(\mathbf{0}, \sigma^2 \mathbf{I}_K)$. $\mathbf{D} \in \mathbb{C}^{K \times N_{\text{cp}}}$ is a submatrix of the standard discrete Fourier transform (DFT) matrix. Given the standard DFT matrix $\mathbf{G} \in \mathbb{C}^{N_c \times N_c}$, we determine \mathbf{D} by extracting the first N_{cp} columns from \mathbf{G} and selecting the K rows corresponding to the pilot subcarriers from \mathbf{G} . In (7), $\mathbf{f} \in \mathbb{C}^{N_{\text{BS}}}$ is a beamforming vector to form directional signal transmission from the BS to the user.

To scan the whole angle space that the signal of the BS covers, we use N_b different beamforming vectors for pilot transmission. Note that if we only use one beam, it may occur that the beam is not aligned with the effective channel path and results in weak received signal strength and consequently poor channel estimation performance [13]. The N_b different beamforming vectors form a codebook matrix

$$\mathbf{F} \triangleq [\mathbf{f}_1, \mathbf{f}_2, \dots, \mathbf{f}_{N_b}] \in \mathbb{C}^{N_{\text{BS}} \times N_b}. \quad (8)$$

Typically we need $N_b \geq N_{\text{BS}}$. Based on (7), we have

$$\mathbf{Y} = \mathbf{XDHF} + \boldsymbol{\Psi}, \quad (9)$$

where $\mathbf{Y} \triangleq [\mathbf{y}_1, \mathbf{y}_2, \dots, \mathbf{y}_{N_b}] \in \mathbb{C}^{K \times N_b}$ is the received pilot matrix after using N_b different beamforming vectors, and $\boldsymbol{\Psi}$ is the consequent channel noise matrix.

III. CHANNEL ESTIMATION EXPLOITING BLOCK SPARSITY

By multiplying the right pseudo inverse of \mathbf{F} , denoted as $\mathbf{F}^H(\mathbf{F}\mathbf{F}^H)^{-1}$, on both sides of (9), we have

$$\mathbf{Y}\mathbf{F}^H(\mathbf{F}\mathbf{F}^H)^{-1} = \mathbf{XDH} + \boldsymbol{\Psi}\mathbf{F}^H(\mathbf{F}\mathbf{F}^H)^{-1}. \quad (10)$$

To simplify the notation, we define

$$\mathbf{A} \triangleq \mathbf{XD} \in \mathbb{C}^{K \times N_{\text{cp}}}, \quad (11)$$

$$\mathbf{Z} \triangleq \mathbf{Y}\mathbf{F}^H(\mathbf{F}\mathbf{F}^H)^{-1} \in \mathbb{C}^{K \times N_{\text{BS}}}, \quad (12)$$

$$\mathbf{N} \triangleq \boldsymbol{\Psi}\mathbf{F}^H(\mathbf{F}\mathbf{F}^H)^{-1} \in \mathbb{C}^{K \times N_{\text{BS}}}. \quad (13)$$

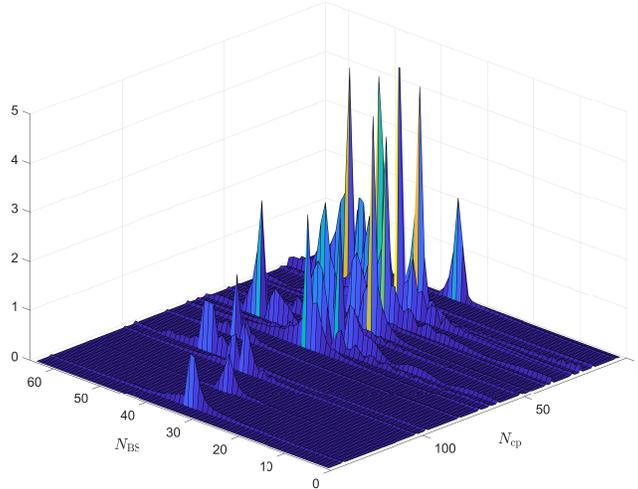


Fig. 2. Time-domain mmWave MIMO channels generated by QuaDRiGa.

Then (10) can be rewritten as

$$\mathbf{Z} = \mathbf{A}\mathbf{H} + \mathbf{N}. \quad (14)$$

The LS estimation of \mathbf{H} can be expressed as

$$\widehat{\mathbf{H}} = (\mathbf{A}^H \mathbf{A})^{-1} \mathbf{A}^H \mathbf{Z}. \quad (15)$$

However, (15) is based on the premise that \mathbf{A} is full column rank. In practice, even if $K \geq N_{cp}$, we cannot guarantee that \mathbf{A} is always full column rank, since the rank of \mathbf{A} is essentially determined by that of \mathbf{D} . If we use a small number of pilot symbols, i.e., K is small, it will occur with high probability that \mathbf{A} is a low-rank matrix. If we use $K < N_{cp}$ to reduce the number of pilot symbols and increase the spectral efficiency, \mathbf{A} will be a low-rank matrix. In this context, (15) cannot be used for channel estimation.

Some recent work on channel modeling and channel measurements shows that mmWave channels are typically sparse [11], [14]. In practice, the number of channel paths is much smaller than the length of the maximum channel delay spread, i.e., $L \ll N_{cp}$, which causes most $\mathbf{h}(n)$ to be zero vectors for $n = 0, 1, \dots, N_{cp} - 1$. Most rows of \mathbf{H} are zero and the number of nonzero rows of \mathbf{H} is L . In Fig. 2, we use the QuaDRiGa which is recommended by 5G New Radio (NR), to generate the time-domain mmWave MIMO channels [15], where the channels with $L = 21$ paths and maximum delay spread of $N_{cp} = 144$ taps are scanned by $N_b = N_{BS} = 64$ DFT codewords. Note that here we consider the NLoS scenario where the LoS path is blocked

and all the 21 paths are treated as NLoS paths. From Fig. 2, it is seen that along each spatial direction that scanned by a DFT codeword, the time-domain channel is a sparse vector with the length of $N_{\text{cp}} = 144$ taps but only $L = 21$ nonzero taps. In particular, $N_{\text{b}} = 64$ time-domain channel vectors pointing at different spatial directions share the common sparsity, although their channel gain might be substantially different.

Define

$$N_{\text{s}} \triangleq N_{\text{cp}} N_{\text{BS}}, \quad (16)$$

$$N_{\text{q}} \triangleq K N_{\text{BS}}. \quad (17)$$

By string together each row of \mathbf{H} , we convert \mathbf{H} to be a column vector \mathbf{s} as

$$\mathbf{s} \triangleq [\mathbf{h}(0), \mathbf{h}(1), \dots, \mathbf{h}(N_{\text{cp}} - 1)]^T \in \mathbb{C}^{N_{\text{s}}}. \quad (18)$$

As illustrated in Fig. 1, \mathbf{s} exhibits block sparsity, where the length of each block is N_{BS} .

Define the k th row of \mathbf{Z} and \mathbf{N} as \mathbf{z}_k and \mathbf{n}_k , respectively, for $k = 1, 2, \dots, K$. We have

$$\mathbf{Z} = [\mathbf{z}_1^T, \mathbf{z}_2^T, \dots, \mathbf{z}_K^T]^T, \quad (19)$$

$$\mathbf{N} = [\mathbf{n}_1^T, \mathbf{n}_2^T, \dots, \mathbf{n}_K^T]^T. \quad (20)$$

Then we define

$$\mathbf{q} \triangleq [\mathbf{z}_1, \mathbf{z}_2, \dots, \mathbf{z}_K]^T \in \mathbb{C}^{N_{\text{q}}}, \quad (21)$$

$$\mathbf{v} \triangleq [\mathbf{n}_1, \mathbf{n}_2, \dots, \mathbf{n}_K]^T \in \mathbb{C}^{N_{\text{q}}}. \quad (22)$$

We further define a stacked measurement matrix

$$\mathbf{B} \triangleq \begin{pmatrix} \mathbf{B}_{1,1} & \mathbf{B}_{1,2} & \cdots & \mathbf{B}_{1,N_{\text{cp}}} \\ \mathbf{B}_{2,1} & \mathbf{B}_{2,2} & \cdots & \mathbf{B}_{2,N_{\text{cp}}} \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{B}_{K,1} & \mathbf{B}_{K,2} & \cdots & \mathbf{B}_{K,N_{\text{cp}}} \end{pmatrix} \in \mathbb{C}^{N_{\text{q}} \times N_{\text{s}}} \quad (23)$$

where the (k, m) th block of \mathbf{B} , denoted as $\mathbf{B}_{k,m}$, is a diagonal matrix, with the definition of $\mathbf{B}_{k,m}$ as

$$\mathbf{B}_{k,m} \triangleq \mathbf{A}(k, m) \mathbf{I}_{N_{\text{BS}}} \in \mathbb{C}^{N_{\text{BS}} \times N_{\text{BS}}} \quad (24)$$

for $k = 1, 2, \dots, K$ and $m = 1, 2, \dots, N_{\text{cp}}$.

Finally, (14) can be rewritten as

$$\mathbf{q} = \mathbf{B}\mathbf{s} + \mathbf{v} \quad (25)$$

To fully exploit the block sparsity of \mathbf{s} , we can resort to the block sparse recovery algorithms.

First we define a residual vector $\mathbf{r} \in \mathbb{C}^{N_q}$, which is initialized to be $\mathbf{r} = \mathbf{q}$. We define a non-negative integer set Λ to keep the indices of the selected blocks of \mathbf{s} during the iterations, where Λ is initialized to be an empty set.

At each iteration, we always aim at finding the best nonzero block of \mathbf{s} , whose corresponding block of \mathbf{B} can achieve the largest projection of \mathbf{r} among all possible blocks. The projection is usually measured based on the LS estimation. Then the index of the best nonzero block of \mathbf{s} can be obtained by

$$I = \arg \max_{i \in \{1, 2, \dots, N_{cp}\} \setminus \Lambda} \|(\mathbf{P}_i^H \mathbf{P}_i)^{-1} \mathbf{P}_i^H \mathbf{r}\|_2, \quad (26)$$

where $\mathbf{P}_i \in \mathbb{C}^{N_q \times N_{BS}}$ is the i th column block of \mathbf{B} defined as

$$\mathbf{P}_i \triangleq [\mathbf{B}_{1,i}^T, \mathbf{B}_{2,i}^T, \dots, \mathbf{B}_{K,i}^T]^T \quad (27)$$

for $i = 1, 2, \dots, N_{cp}$. Since $\mathbf{B}_{k,m}$ is a diagonal matrix, different columns of \mathbf{P}_i are mutually orthogonal. In practice, the power of pilot symbols is in general the same. Therefore, $\mathbf{P}_i^H \mathbf{P}_i$ is a unit matrix multiplied by a constant. Then (26) can be simplified as

$$I = \arg \max_{i \in \{1, 2, \dots, N_{cp}\} \setminus \Lambda} \|\mathbf{P}_i^H \mathbf{r}\|_2. \quad (28)$$

Once I is selected, it is added to Λ to update Λ as

$$\Lambda \leftarrow \Lambda \cup \{I\}. \quad (29)$$

Then we also update the residue vector by removing all projection of \mathbf{q} on the selected blocks as

$$\mathbf{r} \leftarrow \mathbf{q} - \mathbf{P}_\Lambda (\mathbf{P}_\Lambda^H \mathbf{P}_\Lambda)^{-1} \mathbf{P}_\Lambda^H \mathbf{q} \quad (30)$$

where

$$\mathbf{P}_\Lambda \triangleq \{\mathbf{P}_i, i \in \Lambda\} \quad (31)$$

is a submatrix of \mathbf{B} by selecting the column blocks of \mathbf{B} indexed by Λ .

We iteratively run (28), (29) and (30) before satisfying a *stop condition*. Here are three different kinds of setting of the *stop condition*.

Algorithm 1 Time-domain Channel Estimation Exploiting Block Sparsity (TDCEBS)

Input: \mathbf{A} , \mathbf{Z} .

Output: $\widehat{\mathbf{H}}_{\Lambda}$.

- 1: Initialization: $\mathbf{r} \leftarrow \mathbf{q}$, $\Lambda \leftarrow \emptyset$.
 - 2: **while** *the stop condition is not satisfied* **do**
 - 3: Obtain the index of the best nonzero block via (28).
 - 4: Update the indice set via (29).
 - 5: Update the residue vector via (30).
 - 6: **end while**
 - 7: Estimate the nonzero row of \mathbf{H} via (32).
-

- 1) We can simply stop the iteration when the number of iterations exceeds the predetermined maximum number of iterations.
- 2) We can also set the *stop condition* to be a threshold, since the power of the residue vector is getting smaller with the increasing number of iterations. When $\|\mathbf{r}\|_2$ is smaller than the threshold, we stop the iteration.
- 3) We may also compare the power of the residue vector between the current iteration and the last iteration. If the power of the residue vector at current iteration is even larger than that of the last iteration, implying that the residue cannot get smaller any longer, we stop the iteration.

Finally, we estimate the nonzero rows of \mathbf{H} based on Λ and \mathbf{Z} by

$$\widehat{\mathbf{H}}_{\Lambda} = (\mathbf{A}_{\Lambda}^H \mathbf{A}_{\Lambda})^{-1} \mathbf{A}_{\Lambda}^H \mathbf{Z}. \quad (32)$$

where

$$\mathbf{A}_{\Lambda} \triangleq \{\mathbf{A}_i, i \in \Lambda\} \quad (33)$$

is a submatrix of \mathbf{A} by selecting the columns of \mathbf{A} indexed by Λ . Once the nonzero rows of \mathbf{H} is obtained, \mathbf{H} is also estimated. The detailed steps of the proposed TDCEBS scheme are summarized in **Algorithm 1**.

IV. SIMULATION RESULTS

To evaluate the system performance, we use the QuaDRiGa which is recommended by 5G NR to generate wideband mmWave channels [15]. We consider a BS equipped with $N_{\text{BS}} = 64$ antennas working at 28GHz mmWave frequency band serves several single-antenna users. As shown in Fig. 2, the NLoS scenario with $L = 21$ channel paths is generated by the QuaDRiGa. For the OFDM modulation in 5G NR high-frequency standard, the IFFT size is $N_c = 2048$; the CP length is $N_{\text{cp}} = 144$; and the OFDM subcarrier spacing is $\Delta f = 120\text{KHz}$. We set the *stop condition* to be a threshold 0.01 in the simulation. When $\|\mathbf{r}\|_2$ is smaller than 0.01, we stop the iteration. The detailed simulation parameters used for the QuaDRiGa are shown in Table I.

TABLE I
PARAMETERS OF QUADRIGA.

Parameter	value
Number of BS antennas	$N_{\text{BS}} = 64$
Number of paths	$L = 21$
Number of subcarriers	$N_c = 2048$
CP length	$N_{\text{cp}} = 144$
Subcarrier spacing	$\Delta f = 120\text{KHz}$
Center frequency	28GHz
Height of BS	25m
Height of UE	1.5m
Distance between UE and BS	25 – 250m

According to the 5G NR standard, the 132 resource blocks (RBs) occupying 1584 OFDM subcarriers near the center frequency are used for signal transmission, while the other OFDM subcarriers at two edges of the frequency band are used as null subcarriers. The pilot symbols are placed with equal interval on OFDM subcarriers. We consider two cases. In one case, we use $K = 132$ pilot symbols, where each RB transmits a pilot symbol. In the other case, we use $K = 88$ pilot symbols, where 3 RBs transmit 2 pilot symbols.

The performance of channel estimation can be evaluated by the normalized mean squared error (NMSE), which is defined as

$$\text{NMSE} = \frac{\|\widehat{\mathbf{H}} - \mathbf{H}\|_F^2}{\|\mathbf{H}\|_F^2} \quad (34)$$

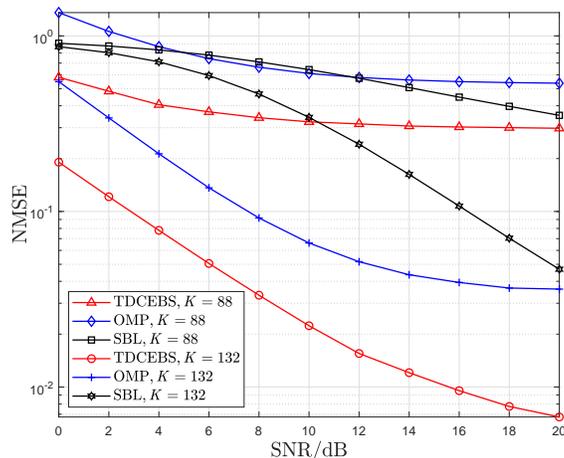


Fig. 3. Comparisons of channel estimation performance in terms of NMSE for different schemes.

where $\widehat{\mathbf{H}}$ is an estimate of the genuine channel matrix \mathbf{H} .

Fig. 3 shows the comparisons of the channel estimation performance in terms of NMSE for different schemes, including OMP, SBL [11] and the proposed TDCEBS scheme. It is seen that TDCEBS achieves the best performance, which is much better than both OMP and SBL. Either for $K = 132$ or $K = 88$, the NMSE of TDCEBS is significantly lower than that of OMP. When $\text{SNR} = 10\text{dB}$ and $K = 132$, the NMSE of TDCEBS is reduced by about 4.7dB compared to that of OMP. When $\text{SNR} = 10\text{dB}$ and $K = 88$, the NMSE of TDCEBS is about 2.7dB lower than that of OMP. Compared to TDCEBS, OMP makes sparse recovery independently for each column vector of \mathbf{H} , which does not exploit the block sparsity of \mathbf{H} . The SBL can only achieve the approximately-sparse channel estimation and there is channel power leakage for zero taps, which leads its performance to be worse than TDCEBS. For each scheme, using more pilots, e.g., from $K = 88$ to $K = 132$, can achieve better performance. In particular, such improvement is more apparent for TDCEBS and OMP than for SBL.

In addition to the NMSE performance, we also compare the bit error rate (BER) performance. Fig.4 shows the BER performance comparisons among TDCEBS, OMP and SBL. To indicate the performance upper bound, we also include the BER performance when the ideal channel is known. It is seen that among the three schemes, TDCEBS performs the best. When $K = 132$, both TDCEBS and OMP can well approach the performance upper bound. But when $K = 88$,

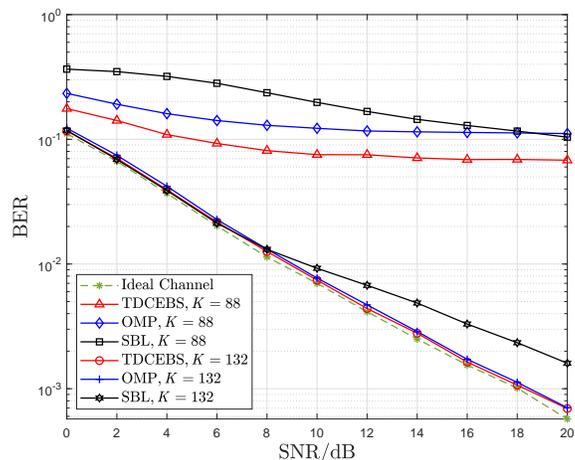


Fig. 4. Comparisons of BER performance for different schemes.

TDCEBS is substantially better than OMP. At $\text{SNR} = 10\text{dB}$, the improvement of TDCEBS over OMP is around 2.1dB, where the BER of TDCEBS might be reduced to zero if a proper channel coding scheme is applied. However, since OMP cannot achieve BER of 0.1 even at high SNR region, the channel coding scheme cannot be well applied for OMP in this context. Therefore, TDCEBS can function more effectively than OMP for channel estimation when the number of used pilot symbols is smaller.

V. CONCLUSION

In this letter, we have investigated time-domain channel estimation for wideband mmWave MIMO OFDM system. By transmitting frequency-domain pilot symbols as well as different beamforming vectors, we have observed that the time-domain mmWave MIMO channels exhibit channel delay sparsity and especially block sparsity among different spatial directions. We have proposed the TDCEBS scheme, which always aims at finding the best nonzero block achieving the largest projection of the residue at each iterations. In particular, we have evaluated the system performance using the QuaDRiGa which is recommended by 5G NR to generate wideband mmWave MIMO channels. Simulation results have verified the effectiveness of the proposed TDCEBS scheme and have shown that it can outperform the existing schemes. In the future work, we will continue to focus on the study of the sparse channel estimation exploiting joint delay and angle sparsity.

REFERENCES

- [1] R. W. Heath, N. Gonzalez-Prelcic, S. Rangan, W. Roh, and A. Sayeed, "An overview of signal processing techniques for millimeter wave MIMO systems," *IEEE J. Sel. Top. Signal Process.*, vol. 10, no. 3, pp. 436–453, Apr. 2016.
- [2] X. Yang, M. Matthaiou, J. Yang, C.-K. Wen, F. Gao, and S. Jin, "Hardware-constrained millimeter-wave systems for 5G: Challenges, opportunities, and solutions," *IEEE Commun. Mag.*, vol. 57, no. 1, pp. 44–50, Jan. 2019.
- [3] W. Ma, C. Qi, Z. Zhang, and J. Cheng, "Sparse channel estimation and hybrid precoding using deep learning for millimeter wave massive MIMO," *IEEE Trans. Commun.*, vol. 68, no. 5, pp. 2838–2849, May 2020.
- [4] B. Wang, F. Gao, S. Jin, H. Lin, G. Y. Li, S. Sun, and T. S. Rappaport, "Spatial-wideband effect in massive MIMO with application in mmWave systems," *IEEE Commun. Mag.*, vol. 56, no. 12, pp. 134–141, Dec. 2018.
- [5] I.-S. Kim and J. Choi, "Spatial wideband channel estimation for mmWave massive MIMO systems with hybrid architectures and low-resolution ADCs," *IEEE Trans. Wireless Commun.*, vol. 20, no. 6, pp. 4016–4029, June 2021.
- [6] Z. Gao, C. Hu, L. Dai, and Z. Wang, "Channel estimation for millimeter-wave massive MIMO with hybrid precoding over frequency-selective fading channels," *IEEE Commun. Lett.*, vol. 20, no. 6, pp. 1259–1262, Oct. 2016.
- [7] F. Shu, J. Wang, J. Li, R. Chen, and W. Chen, "Pilot optimization, channel estimation, and optimal detection for full-duplex OFDM systems with IQ imbalances," *IEEE Trans. Veh. Technol.*, vol. 66, no. 8, pp. 6993–7009, Aug. 2017.
- [8] J. Rodríguez-Fernández, N. González-Prelcic, K. Venugopal, and R. W. Heath, "Frequency-domain compressive channel estimation for frequency-selective hybrid millimeter wave MIMO systems," *IEEE Trans. Wireless Commun.*, vol. 17, no. 5, pp. 2946–2960, May 2018.
- [9] Y. Lin, S. Jin, M. Matthaiou, and X. You, "Tensor-based channel estimation for millimeter wave MIMO-OFDM with dual-wideband effects," *IEEE Trans. Commun.*, vol. 68, no. 7, pp. 4218–4232, July 2020.
- [10] H. Kim, G.-T. Gil, and Y. H. Lee, "Two-step approach to time-domain channel estimation for wideband millimeter wave systems with hybrid architecture," *IEEE Trans. Commun.*, vol. 67, no. 7, pp. 5139–5152, July 2019.
- [11] S. Srivastava, A. Mishra, A. Rajoriya, A. K. Jagannatham, and G. Ascheid, "Quasi-static and time-selective channel estimation for block-sparse millimeter wave hybrid MIMO systems: Sparse Bayesian learning (SBL) based approaches," *IEEE Trans. Signal Process.*, vol. 67, no. 5, pp. 1251–1266, Mar. 2019.
- [12] W. Ma, C. Qi, and G. Y. Li, "High-resolution channel estimation for frequency-selective mmWave massive MIMO systems," *IEEE Trans. Wireless Commun.*, vol. 19, no. 5, pp. 3517–3529, May 2020.
- [13] C. Qi, K. Chen, O. A. Dobre, and G. Y. Li, "Hierarchical codebook-based multiuser beam training for millimeter wave massive MIMO," *IEEE Trans. Wireless Commun.*, vol. 19, no. 12, pp. 8142–8152, Dec. 2020.
- [14] A. Liao, Z. Gao, H. Wang, S. Chen, M.-S. Alouini, and H. Yin, "Closed-loop sparse channel estimation for wideband millimeter-wave full-dimensional MIMO systems," *IEEE Trans. Commun.*, vol. 67, no. 12, pp. 8329–8345, Dec. 2019.
- [15] S. Jaeckel, L. Raschkowski, K. Börner, and L. Thiele, "QuaDRiGa: A 3-D multi-cell channel model with time evolution for enabling virtual field trials," *IEEE Trans. Antennas Propagat.*, vol. 62, no. 6, pp. 3242–3256, Mar. 2014.