

Uplink Sparse Channel Estimation for Hybrid Millimeter Wave Massive MIMO Systems by UTAMP-SBL

Shuai Hou, Yafeng Wang * and Chao Li

School of Information and Communication Engineering, Beijing University of Posts and Telecommunications, Beijing 100876, China; houshuai@bupt.edu.cn (S.H.); lcay610@bupt.edu.cn (C.L.)

* Correspondence: wangyf@bupt.edu.cn

Abstract: The compressive sensing (CS)-based sparse channel estimator is recognized as the most effective solution to the excessive pilot overhead in massive MIMO systems. However, due to the complex signal processing in the wireless communication systems, the measurement matrix in the CS-based channel estimation is sometimes “unfriendly” to the channel recovery. To overcome this problem, in this paper, the state-of-the-art sparse Bayesian learning using approximate message passing with unitary transformation (UTAMP-SBL), which is robust to various measurement matrices, is leveraged to address the multi-user uplink channel estimation for hybrid architecture millimeter wave massive MIMO systems. Specifically, the sparsity of channels in the angular domain is exploited to reduce the pilot overhead. Simulation results demonstrate that the UTAMP-SBL is able to achieve effective performance improvement than other competitors with low pilot overhead.

Keywords: millimeter-wave massive MIMO; channel estimation; sparse Bayesian learning; compressive sensing



Citation: Hou, S.; Wang, Y.; Li, C. Uplink Sparse Channel Estimation for Hybrid Millimeter Wave Massive MIMO Systems by UTAMP-SBL. *Sensors* **2021**, *21*, 4760. <https://doi.org/10.3390/s21144760>

Academic Editor: Luís Castedo Ribas

Received: 3 June 2021

Accepted: 11 July 2021

Published: 12 July 2021

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Copyright: © 2021 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

Massive MIMO is a key technology for the fifth-generation (5G) communication systems [1]. Thanks to the shorter wavelength of the millimeter wave (mmWave) signal, the numerous antennas are packed into a compact-size array, which facilitates the commercial deployment of massive MIMO systems [2], but under the consideration of the high power consumption of Analog-to-Digital Converters (ADCs), the hybrid beamforming architecture, that divides the precoder and the combiner into the analog and digital domains, is developed to solve it and regarded as an effective alternative [3,4]. In this hybrid architecture system, the precoding is crucial and dependent on the acquisition of accurate channel state information (CSI). However, the number of pilots used for channel estimation will increase linearly with that of antennas and users in the wireless systems. As a result, the channel estimation with fewer pilots is a significant challenge in the hybrid mmWave massive MIMO systems [3].

The compressive sensing (CS)-based channel estimator is capable of exploiting the sparsity of channels to reduce the pilot overhead greatly and a lot of studies on this topic have been done. Some studies focus on the sparsity in the delay domain. In [5], the orthogonal matching pursuit (OMP) algorithm is utilized to estimate the wideband mmWave delay-domain sparse channels. In [6], an adaptive structured subspace pursuit algorithm at the user is proposed to estimate the delay-domain MIMO channels with the spatio-temporal common sparsity. The authors of [7] propose a sparse channel recovery algorithm based on the vector approximate message passing (VAMP) for the massive MIMO delay-domain channels. Some works concentrate on the sparsity in the angular domain. For example, a distributed sparsity adaptive matching pursuit (DSAMP) algorithm is proposed in [8], whereby the spatially common sparsity is exploited. In [9], a novel sparse Bayesian learning (SBL) approach is utilized to recover the angular-domain block-sparse

channels. In [10], a structured turbo-CS algorithm is proposed to estimate the angular-domain massive MIMO channels which are modeled by a Markov chain prior. Furthermore, the work [11] even jointly exploits the 3-D clustered structure of channels in the angular-delay domain and proposes an approximate message passing (AMP)-based estimation algorithm. However, all of above angular-domain sparse channel estimation schemes are based on the assumption that the angles exactly lie on the grids, or ignore the power leakage caused by grid mismatch. Recently, there are many off-grid channel estimation techniques. In [12], the authors utilize the low-rank structure along with the sparsity in angular domain to improve the channel estimation performance, and the off-grid angles can be recovered with their algorithm successfully. In [13], the angles are treated as random parameters and the grid-less quantized variational Bayesian channel estimation algorithm is proposed for antenna array systems with low resolution ADCs. Besides, the multi-dimensional variational line spectral estimation algorithm proposed in [14] can be effectively applied for multi-dimensional off-grid angle estimation. In addition, a novel super-resolution downlink channel estimation approach developed from the SBL is provided in [15], where the sampled angular grid points are treated as the underlying parameters.

The performance of these CS-based algorithms is affected by the measurement matrix which is usually associated with the signal processing operations in the system. If these operations are not carefully designed, the measurement matrix perhaps becomes detrimental to channel recovery. Recently, a CS algorithm, termed the SBL using approximate message passing with unitary transformation (UTAMP-SBL), is proposed by Luo et al. and it outperforms the state-of-the-art AMP-based SBL algorithms, the Gaussian generalized AMP-based SBL (GGAMP-SBL), in terms of robustness, speed and recovery accuracy for difficult measurement matrices [16]. In this paper, we apply UTAMP-SBL to the sparse channel estimation to improve the performance. Specifically, the multi-user uplink channel estimation for hybrid architecture mmWave massive MIMO systems is studied as an example. And the angular-domain sparsity of the mmWave channels is fully exploited in our estimation. Additionally, this algorithm can be extended to other sparse channel estimation for the performance improvement. Simulation results verify that the UTAMP-SBL outperforms the other competitors.

The remainder of this paper is organized as follows. In Section 2, the hybrid millimeter wave massive MIMO system model is introduced. In Section 3, the uplink sparse channel estimation with UTAMP-SBL is described. In Section 4, the Cramér-Rao bound (CRB) is provided as a performance benchmark. Simulation results are provided and the performance is discussed in Section 5. Conclusions are given in Section 6.

Notations: $\mathbf{H}$ and $\otimes$ denote the conjugate transpose operation and the Kronecker product. $\odot$ and $\oslash$ represent the componentwise vector multiplication and the componentwise vector division. $\text{vec}(\mathbf{A})$ denotes vectorizing the matrix $\mathbf{A}$ as a vector. $\mathbf{I}$, $\mathbf{1}$ and $\mathbf{0}$ denote the identity matrix, the all-one column vector and the all-zero column vector, respectively. $\text{diag}(\mathbf{a})$ returns a diagonal matrix with the elements of vector $\mathbf{a}$ on its main diagonal. $\mathcal{N}_{\mathbb{C}}(\mathbf{x}; \boldsymbol{\mu}_x, \boldsymbol{\Sigma}_x)$ represents the Gaussian distribution of the complex vector $\mathbf{x}$ with mean $\boldsymbol{\mu}_x$ and covariance matrix $\boldsymbol{\Sigma}_x$. $\text{Ga}(\gamma; \epsilon, \eta)$ denotes a Gamma distribution with shape parameter ϵ and rate parameter η . $\mathbb{E}_{b(\mathbf{x})}[f(\mathbf{x})]$ denotes the expectation of the function $f(\mathbf{x})$ with respect to probability density $b(\mathbf{x})$. Finally, $\propto$ denotes equality up to a constant scale factor.

2. System Model

2.1. Millimeter Wave MIMO Channel Model

We consider a mmWave MIMO-OFDM system where the base station (BS) is equipped with N_{BS} antennas and serves K multi-antenna user equipments (UEs), and each UE has N_{UE} antennas. The frequency-domain channel between the BS and the k th UE at the p th ($p = 1, 2, \dots, P$) subcarrier can be modeled as [17]

$$\mathbf{H}_{p,k} = \sum_{l=1}^L \alpha_{l,k} \mathbf{a}_{\text{BS}}(\theta_{l,k}) \mathbf{a}_{\text{UE}}^H(\phi_{l,k}) e^{-j2\pi f_s \tau_{l,k} p/P}, \quad (1)$$

where L is the number of physical paths, $\alpha_{l,k}$ is the complex path gain, f_s is the sampling rate, $\tau_{l,k}$ is the path delay, $\mathbf{a}_{\text{BS}}(\theta_{l,k})$ and $\mathbf{a}_{\text{UE}}(\phi_{l,k})$ are the steering vectors at the BS and the UE, respectively, $\theta_{l,k}$ and $\phi_{l,k}$ are the azimuth angles of departure and arrival (AoD/AoA) uniformly distributed in $[0, 2\pi)$.

For each UE in the system, we generate one line of sight (LoS) path and $L - 1$ non-LoS (NLoS) paths. The path gains follow the complex Gaussian distribution and the ratio of the power of LOS path to that of NLoS path is 10 dB [18]. In the OFDM system, the cyclic prefix (CP) is introduced to mitigate the inter-symbol interference and we have $\tau_{\max} f_s \leq N_{\text{cp}}$, where $\tau_{\max}$ is the maximum path delay and N_{cp} is the length of CP. So we generate $\tau_{1,k} f_s = 1$ and $\tau_{l,k} f_s \in \{2, 3, \dots, N_{\text{cp}}\}$, $l \neq 1$. For the steering vectors, assuming the uniform linear arrays are adopted at the BS and the UE, they can be shown as

$$\begin{aligned}\mathbf{a}_{\text{BS}}(\theta_{l,k}) &= [1, e^{j\frac{2\pi}{\lambda}d \sin(\theta_{l,k})}, \dots, e^{j(N_{\text{BS}}-1)\frac{2\pi}{\lambda}d \sin(\theta_{l,k})}]^T \\ \mathbf{a}_{\text{UE}}(\phi_{l,k}) &= [1, e^{j\frac{2\pi}{\lambda}d \sin(\phi_{l,k})}, \dots, e^{j(N_{\text{UE}}-1)\frac{2\pi}{\lambda}d \sin(\phi_{l,k})}]^T,\end{aligned}\quad (2)$$

where λ is the wavelength and $d = \lambda/2$ is the antenna spacing.

2.2. Uplink Pilot Transmission

We consider the typical hybrid analog digital precoding and combining architecture in the mmWave massive MIMO systems, where the BS and each UE are assumed to have $N_{\text{BS}}^{\text{RF}}$ RF chains and $N_{\text{UE}}^{\text{RF}}$ RF chains, respectively. We focus on the uplink pilot training. The k th UE transmits the pilot $\mathbf{s}_{p,t,k} \in \mathbb{C}^{N_s \times 1}$ at the p th pilot subcarrier and the t th OFDM symbol to the BS, where N_s is the number of data streams. At the BS, the receiver receives pilots from all UEs in the same time-frequency resources and the received signal can be written as

$$\mathbf{y}_{p,t} = (\mathbf{Z}_t^{\text{RF}} \mathbf{Z}_{p,t}^{\text{BB}})^H \left(\sum_{k=1}^K \mathbf{H}_{p,k} \mathbf{F}_{t,k}^{\text{RF}} \mathbf{F}_{p,t,k}^{\text{BB}} \mathbf{s}_{p,t,k} + \mathbf{n}_{p,t} \right), \quad (3)$$

where $\mathbf{F}_{p,t,k}^{\text{BB}} \in \mathbb{C}^{N_{\text{UE}}^{\text{RF}} \times N_s}$ and $\mathbf{F}_{t,k}^{\text{RF}} \in \mathbb{C}^{N_{\text{UE}} \times N_{\text{UE}}^{\text{RF}}}$ are the digital precoder and the analog precoder of the k th user, $\mathbf{Z}_{p,t}^{\text{BB}} \in \mathbb{C}^{N_{\text{BS}}^{\text{RF}} \times N_s}$ and $\mathbf{Z}_t^{\text{RF}} \in \mathbb{C}^{N_{\text{BS}} \times N_{\text{BS}}^{\text{RF}}}$ are the digital combiner and the analog combiner of the BS and $\mathbf{n}_{p,t} \in \mathbb{C}^{N_{\text{BS}} \times 1}$ is the additive white Gaussian noise (AWGN). Note that the analog precoder and the analog combiner are frequency-flat, while the digital precoder and the digital combiner are different for each subcarrier. This is because the RF phase shifters can provide constant response over the wide frequency band when the fully connected network is used.

2.3. Sparse Channel Estimation Formulation

In the multi-user uplink channel estimation, the channels of K users, $\mathbf{H}_{p,k} \in \mathbb{C}^{N_{\text{BS}} \times N_{\text{UE}}}$, $k = 1, 2, \dots, K$, need to be estimated from the received signal $\mathbf{y}_{p,t}$. It is clear that there are $KN_{\text{BS}}N_{\text{UE}}$ non-zero elements in these K channel matrices. Because there are only N_s elements in $\mathbf{y}_{p,t}$, we have to arrange $KN_{\text{BS}}N_{\text{UE}}/N_s$ pilots for channel estimation and the pilot overhead is unacceptable. Fortunately, we benefit from the inherent sparse characteristic of millimeter wave channels and apply the CS algorithms in the channel estimation to reduce the pilot overhead.

In order to formulate the sparse channel estimation problem, we first rewrite the frequency-domain channel matrix $\mathbf{H}_{p,k} \in \mathbb{C}^{N_{\text{BS}} \times N_{\text{UE}}}$ as

$$\mathbf{H}_{p,k} = \mathbf{A}_{\text{BS}} \mathbf{H}_{p,k}^{\omega} \mathbf{A}_{\text{UE}}^H, \quad (4)$$

where $\mathbf{H}_{p,k}^{\omega} \in \mathbb{C}^{G_{\text{BS}} \times G_{\text{UE}}}$ is the virtual angle-domain channel matrix, $\mathbf{A}_{\text{BS}} \in \mathbb{C}^{N_{\text{BS}} \times G_{\text{BS}}}$ and $\mathbf{A}_{\text{UE}} \in \mathbb{C}^{N_{\text{UE}} \times G_{\text{UE}}}$ are the partial discrete Fourier transform (DFT) matrices, G_{BS} and G_{UE} are the numbers of grids at the BS and the UE. Because of the domination of the LOS path in the multipath channel, the virtual angle-domain mmWave channel appears

the sparsity, i.e., there are only a few non-zero elements in the channel matrix $\mathbf{H}_{p,k}^\omega$ and other components are zero. Actually, these zeros are not strictly equal to zero but are only close to zero owing to the power leakage problem [19], which is ignored in this paper for simplification and it will be discussed in the future research. The frequency-domain channel $\mathbf{H}_{p,k}$ and the transformed virtual angle-domain channel $\mathbf{H}_{p,k}^\omega$ are shown in Figure 1, where the frequency-domain channel is generated according to (1) and the virtual angle-domain channel is derived from (4). Obviously, from Figure 1b, the power of the channel is concentrated in a few virtual angles and these channel components can be estimated from the low-overhead pilot.

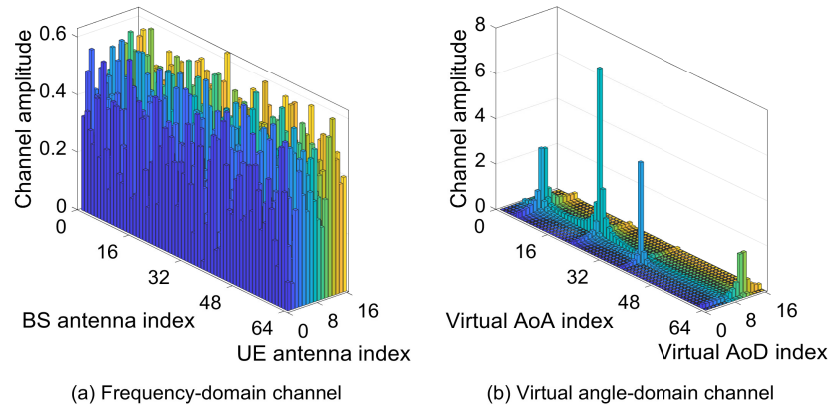


Figure 1. The illustration of the frequency-domain channel and the virtual angle-domain channel.

Substituting (4) in (3), we have

$$\mathbf{y}_{p,t} = \mathbf{Z}_{p,t}^H \sum_{k=1}^K \mathbf{A}_{BS} \mathbf{H}_{p,k}^\omega \mathbf{A}_{UE}^H \mathbf{x}_{p,t,k} + \bar{\mathbf{n}}_{p,t}, \quad (5)$$

where $\mathbf{Z}_{p,t} = \mathbf{Z}_t^{\text{RF}} \mathbf{Z}_{p,t}^{\text{BB}} \in \mathbb{C}^{N_{BS} \times N_s}$ is the combiner at the BS, $\mathbf{x}_{p,t,k} = \mathbf{F}_{t,k}^{\text{RF}} \mathbf{F}_{p,t,k}^{\text{BB}} \mathbf{s}_{p,t,k} \in \mathbb{C}^{N_{UE} \times 1}$ is the pilot after precoding at the UE and $\bar{\mathbf{n}}_{p,t} = \mathbf{Z}_{p,t}^H \mathbf{n}_{p,t} \in \mathbb{C}^{N_s \times 1}$ is the noise after combining at the BS. Note that usually the channel $\mathbf{H}_{p,k}^\omega$ is invariant in some successive OFDM symbols. Rewrite (5) in matrix form as

$$\mathbf{y}_{p,t} = \mathbf{Z}_{p,t}^H \mathbf{A}_{BS} \bar{\mathbf{H}}_p^\omega \bar{\mathbf{A}}_{UE}^H \bar{\mathbf{x}}_{p,t} + \bar{\mathbf{n}}_{p,t}, \quad (6)$$

where $\bar{\mathbf{H}}_p^\omega = [\mathbf{H}_{p,1}^\omega, \mathbf{H}_{p,2}^\omega, \dots, \mathbf{H}_{p,K}^\omega] \in \mathbb{C}^{G_{BS} \times K G_{UE}}$, $\bar{\mathbf{A}}_{UE}^H = \text{diag}(\mathbf{A}_{UE}^H, \mathbf{A}_{UE}^H, \dots, \mathbf{A}_{UE}^H) \in \mathbb{C}^{K G_{UE} \times K N_{UE}}$, and $\bar{\mathbf{x}}_{p,t} = [\mathbf{x}_{p,t,1}^T, \mathbf{x}_{p,t,2}^T, \dots, \mathbf{x}_{p,t,K}^T]^T \in \mathbb{C}^{K N_{UE} \times 1}$. Using the result $\text{vec}(\mathbf{ABC}) = (\mathbf{C}^T \otimes \mathbf{A}) \text{vec}(\mathbf{B})$, (6) can be equivalently expressed as

$$\begin{aligned} \mathbf{y}_{p,t} &= \text{vec}(\mathbf{Z}_{p,t}^H \mathbf{A}_{BS} \bar{\mathbf{H}}_p^\omega \bar{\mathbf{A}}_{UE}^H \bar{\mathbf{x}}_{p,t}) + \bar{\mathbf{n}}_{p,t} \\ &= ((\bar{\mathbf{A}}_{UE}^H \bar{\mathbf{x}}_{p,t})^T \otimes (\mathbf{Z}_{p,t}^H \mathbf{A}_{BS})) \text{vec}(\bar{\mathbf{H}}_p^\omega) + \bar{\mathbf{n}}_{p,t} \\ &= \Phi_{p,t} \mathbf{h}_p^\omega + \bar{\mathbf{n}}_{p,t}, \end{aligned} \quad (7)$$

where $\Phi_{p,t} = ((\bar{\mathbf{A}}_{UE}^H \bar{\mathbf{x}}_{p,t})^T \otimes (\mathbf{Z}_{p,t}^H \mathbf{A}_{BS})) \in \mathbb{C}^{N_s \times K G_{BS} G_{UE}}$, $\mathbf{h}_p^\omega = \text{vec}(\bar{\mathbf{H}}_p^\omega) \in \mathbb{C}^{K G_{BS} G_{UE} \times 1}$. We stack the received signals that are transmitted on G successive symbols, and we have

$$\mathbf{y}_p = \Phi_p \mathbf{h}_p^\omega + \bar{\mathbf{n}}_p, \quad (8)$$

where $\mathbf{y}_p = [\mathbf{y}_{p,1}^T, \mathbf{y}_{p,2}^T, \dots, \mathbf{y}_{p,G}^T]^T \in \mathbb{C}^{GN_s \times 1}$, $\Phi_p = [\Phi_{p,1}^T, \Phi_{p,2}^T, \dots, \Phi_{p,G}^T]^T \in \mathbb{C}^{GN_s \times KG_{BS}}$
 G_{UE} and $\bar{\mathbf{n}}_p = [\mathbf{n}_{p,1}^T, \mathbf{n}_{p,2}^T, \dots, \mathbf{n}_{p,G}^T]^T \in \mathbb{C}^{GN_s \times 1}$.

As shown above, in the channel estimation, the measurement matrix Φ_p derives from pilots, precoders, combiners and the partial DFT matrices, so it is a difficult matrix in all probability. With the difficult measurement matrix, the existing sparse channel estimation algorithms can fail to recover channels. As a result, the UTAMP-SBL which is robust to difficult measurement matrix is a wise choice.

3. Uplink Channel Estimation with UTAMP-SBL

Focus on the channel estimation of the single subcarrier and the subscript p is omitted for conciseness,

$$\mathbf{y} = \Phi \mathbf{h}^\omega + \bar{\mathbf{n}}. \quad (9)$$

Let $M = GN_s$, $N = KG_{BS}G_{UE}$, so the sizes of $\mathbf{y}$, Φ and $\mathbf{h}^\omega$ are $M \times 1$, $M \times N$ and $N \times 1$, respectively, where $M \ll N$. In order to characterize the sparsity of $\mathbf{h}^\omega$, the Gaussian prior model is adopted, i.e.,

$$p(\mathbf{h}^\omega) = \int p_{\mathbf{h}^\omega}(\mathbf{h}^\omega | \gamma) p_\gamma(\gamma) d\gamma, \quad (10)$$

where $p_{\mathbf{h}^\omega}(\mathbf{h}^\omega | \gamma) = \mathcal{N}_{\mathbb{C}}(\mathbf{h}^\omega; \mathbf{0}, \text{diag}(\gamma))$ and $p_\gamma(\gamma) = \text{Ga}(\gamma; \epsilon, \eta)$. Define the singular value decomposition (SVD) of Φ is $\Phi = \mathbf{U} \Sigma \mathbf{V}$, where $\mathbf{U}$ is an orthogonal matrix and $\mathbf{U}^H \mathbf{U} = \mathbf{I}$. Perform unitary transformation to (9) and we have

$$\begin{aligned} \mathbf{U}^H \mathbf{y} &= \mathbf{U}^H \Phi \mathbf{h}^\omega + \mathbf{U}^H \bar{\mathbf{n}} \\ \mathbf{r} &= \Sigma \mathbf{V} \mathbf{h}^\omega + \mathbf{w}, \end{aligned} \quad (11)$$

where $\mathbf{r} = \mathbf{U}^H \mathbf{y} \in \mathbb{C}^{M \times 1}$, $\mathbf{w} = \mathbf{U}^H \bar{\mathbf{n}} \in \mathbb{C}^{M \times 1}$ and $\mathbf{w}$ follows $\mathcal{N}_{\mathbb{C}}(\mathbf{w}; \mathbf{0}, \sigma^2 \mathbf{I})$. Define $\mathbf{z} = \Sigma \mathbf{V} \mathbf{h}^\omega$ and the likelihood function of $\mathbf{z}$ is

$$p_{\mathbf{r}}(\mathbf{r} | \mathbf{z}; \sigma) = \mathcal{N}_{\mathbb{C}}(\mathbf{r}; \mathbf{z}, \sigma^2 \mathbf{I}). \quad (12)$$

According to the Bayesian rule, the joint probability density function is

$$p(\mathbf{r}, \mathbf{z}, \mathbf{h}^\omega; \sigma, \epsilon, \eta) = p_{\mathbf{r}}(\mathbf{r} | \mathbf{z}; \sigma) \delta(\mathbf{z} - \Sigma \mathbf{V} \mathbf{h}^\omega) p(\mathbf{h}^\omega), \quad (13)$$

where the Dirac function factor $\delta(\cdot)$ is adopted first in [20] to facilitate the derivation of the algorithm. The factor graph of the factorization (13) is shown in Figure 2. The UTAMP-SBL algorithm will be derived from message-passing on this graph and applied to the channel estimation.

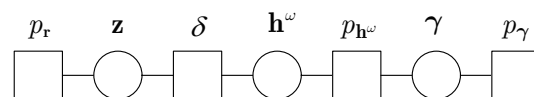


Figure 2. The factor graph of uplink sparse channel estimation.

Firstly, we consider the messages between node $p_{\mathbf{r}}$ and node $\mathbf{z}$. According to the sum-product algorithm, the message from function node $p_{\mathbf{r}}$ to variable node $\mathbf{z}$ is $m_{p_{\mathbf{r}} \rightarrow \mathbf{z}}(\mathbf{z}) = p_{\mathbf{r}}(\mathbf{r} | \mathbf{z}; \sigma) = \mathcal{N}_{\mathbb{C}}(\mathbf{z}; \mathbf{r}, \sigma^2 \mathbf{I})$. The message from variable node $\mathbf{z}$ to function node $p_{\mathbf{r}}$ is defined as $m_{\mathbf{z} \rightarrow p_{\mathbf{r}}}(\mathbf{z}) = \mathcal{N}_{\mathbb{C}}(\mathbf{z}; \boldsymbol{\mu}, \text{diag}(\boldsymbol{\tau}_{\boldsymbol{\mu}}))$. So the belief $b(\mathbf{z}) = m_{\mathbf{z} \rightarrow p_{\mathbf{r}}}(\mathbf{z}) m_{p_{\mathbf{r}} \rightarrow \mathbf{z}}(\mathbf{z})$ can be calculated as $b(\mathbf{z}) = \mathcal{N}_{\mathbb{C}}(\mathbf{z}; \hat{\mathbf{z}}, \text{diag}(\hat{\mathbf{v}}_{\mathbf{z}}))$, where the mean and variance are

$$\begin{aligned} \hat{\mathbf{v}}_{\mathbf{z}} &= \mathbf{1} \odot \left(\mathbf{1} \odot \boldsymbol{\tau}_{\boldsymbol{\mu}} + \frac{1}{\hat{\sigma}^2} \mathbf{1} \right), \\ \hat{\mathbf{z}} &= \hat{\mathbf{v}}_{\mathbf{z}} \odot \left(\frac{\mathbf{r}}{\hat{\sigma}^2} + \boldsymbol{\mu} \odot \boldsymbol{\tau}_{\boldsymbol{\mu}} \right). \end{aligned} \quad (14)$$

Then, we focus on the messages between node $p_{\mathbf{h}^\omega}$ and node $\mathbf{h}^\omega$. Similarly, the message from function node $p_{\mathbf{h}^\omega}$ to variable node $\mathbf{h}^\omega$ is defined as $m_{p_{\mathbf{h}^\omega} \rightarrow \mathbf{h}^\omega}(\mathbf{h}^\omega) = \mathcal{N}_{\mathbb{C}}(\mathbf{h}^\omega; \mathbf{0}, \text{diag}(\hat{\gamma}))$. And the message from variable node $\mathbf{h}^\omega$ to function node $p_{\mathbf{h}^\omega}$ is defined as $m_{\mathbf{h}^\omega \rightarrow p_{\mathbf{h}^\omega}}(\mathbf{h}^\omega) = \mathcal{N}_{\mathbb{C}}(\mathbf{h}^\omega; \boldsymbol{\pi}, \tau_\pi \mathbf{I})$. So the belief $b(\mathbf{h}^\omega) = m_{\mathbf{h}^\omega \rightarrow p_{\mathbf{h}^\omega}}(\mathbf{h}^\omega) m_{p_{\mathbf{h}^\omega} \rightarrow \mathbf{h}^\omega}(\mathbf{h}^\omega)$ is calculated as $b(\mathbf{h}^\omega) = \mathcal{N}_{\mathbb{C}}(\mathbf{h}^\omega; \hat{\mathbf{h}}^\omega, v_{h^\omega} \mathbf{I})$, where the mean and variance are

$$\begin{aligned} v_{h^\omega} &= \frac{1}{N} \mathbf{1}^\top (\mathbf{1} \odot (\mathbf{1}/\tau_\pi + \mathbf{1} \odot \hat{\gamma})), \\ \hat{\mathbf{h}}^\omega &= \boldsymbol{\pi} \odot (\mathbf{1} + \tau_\pi \mathbf{1} \odot \hat{\gamma}). \end{aligned} \quad (15)$$

For the sparse channel estimation, the mean of the belief $b(\mathbf{h}^\omega)$ is the estimated channel.

Next, the variable $\boldsymbol{\mu}$, τ_μ , $\boldsymbol{\pi}$ and τ_π mentioned above can be updated with the UTAMP which is derived from the AMP. In the UTAMP, the means and variances are calculated by [16]

$$\begin{aligned} \tau_\mu &= v_{h^\omega} \boldsymbol{\Sigma} \boldsymbol{\Sigma}^H \mathbf{1}, \\ \boldsymbol{\mu} &= \boldsymbol{\Sigma} \mathbf{V} \hat{\mathbf{h}}^\omega - \tau_\mu \odot \mathbf{s}, \\ \tau_\pi &= \left(\frac{1}{N} \left(\left(\boldsymbol{\Sigma} \boldsymbol{\Sigma}^H \mathbf{1} \right)^H \boldsymbol{\tau}_s \right) \right)^{-1}, \\ \boldsymbol{\pi} &= \hat{\mathbf{h}}^\omega + \tau_\pi \left((\boldsymbol{\Sigma} \mathbf{V})^H \mathbf{s} \right), \end{aligned} \quad (16)$$

where the intermediate variables $\mathbf{s}$ and $\boldsymbol{\tau}_s$ are given by [16]

$$\begin{aligned} \boldsymbol{\tau}_s &= \mathbf{1} \odot (\boldsymbol{\tau}_\mu + \hat{\sigma} \mathbf{1}), \\ \mathbf{s} &= \boldsymbol{\tau}_s \odot (\mathbf{r} - \boldsymbol{\mu}). \end{aligned} \quad (17)$$

Besides, there are several model parameters need to be learned at each iteration. For the noise precision $\hat{\sigma}$, it can be updated with the expectation-maximization (EM) algorithm [16]

$$\hat{\sigma} = \frac{1}{M} \left(\|\mathbf{r} - \hat{\mathbf{z}}\|^2 + \mathbf{1}^\top \mathbf{v}_z \right). \quad (18)$$

For the variable γ , according to the sum-product algorithm, the componentwise message from function node $p_{h_n^\omega}$ to variable node γ_n is

$$\begin{aligned} m_{p_{h_n^\omega} \rightarrow \gamma_n}(\gamma_n) &= p_{h_n^\omega}(\gamma_n) m_{h_n^\omega \rightarrow p_{h_n^\omega}}(\gamma_n) \\ &= \mathcal{N}_{\mathbb{C}}(h_n^\omega; 0, \gamma_n) \mathcal{N}_{\mathbb{C}}(h_n^\omega; \pi_n, \tau_\pi) \\ &\propto \sqrt{\frac{1}{\gamma_n}} \exp \left\{ -\frac{1}{2\gamma_n} \left(|\hat{h}_n^\omega|^2 + v_{h^\omega} \right) \right\}. \end{aligned} \quad (19)$$

The message from variable node γ to function node $p_{\mathbf{h}^\omega}$ is $m_{\gamma \rightarrow p_{\mathbf{h}^\omega}}(\gamma) = m_{p_\gamma \rightarrow \gamma}(\gamma) = \text{Ga}(\gamma; \epsilon, \eta)$. Consequently, the belief $b(\gamma_n)$ is expressed as

$$\begin{aligned} b(\gamma_n) &= m_{\gamma_n \rightarrow p_{h_n^\omega}}(\gamma_n) m_{p_{h_n^\omega} \rightarrow \gamma_n}(\gamma_n) \\ &\propto \gamma_n^{\frac{1}{2} - \frac{1}{2}\hat{\epsilon}} \exp \left\{ -\frac{1}{2\gamma_n} \left(|\hat{h}_n^\omega|^2 + v_{h^\omega} + 2\eta \right) \right\}. \end{aligned} \quad (20)$$

So, $\hat{\gamma}_n$ is the mean of this distribution, i.e.,

$$\hat{\gamma}_n = \frac{|\hat{h}_n^\omega|^2 + v_{h^\omega}}{\hat{\epsilon} + 1}, \quad (21)$$

when we set $\eta = 0$. For the parameter $\hat{\epsilon}$, it is difficult to renewed, but there is an effective way found in [16] and we use it directly. As a result, $\hat{\epsilon}$ is calculated by

$$\hat{\epsilon} = \frac{1}{2} \sqrt{\log \left(\frac{1}{N} \sum_n \frac{1}{\hat{\gamma}_n} \right) - \frac{1}{N} \sum_n \log \frac{1}{\hat{\gamma}_n}}. \quad (22)$$

The process of the channel estimation with UTAMP-SBL is given in Algorithm 1. In this algorithm, the iteration will be halted if the number of it reaches the maximum value $T_{\max}$ or $\|\hat{\mathbf{h}}^{\omega(t)} - \hat{\mathbf{h}}^{\omega(t-1)}\| / \|\hat{\mathbf{h}}^{\omega(t)}\| < 10^{-12}$.

Algorithm 1 Uplink channel estimation with UTAMP-SBL.

Input: The received signal $\mathbf{y}$ and the matrix Φ .

Output: The estimated sparse channel $\hat{\mathbf{h}}^{\omega}$.

- 1: **Initialization:** SVD $\Phi = \mathbf{U}\Sigma\mathbf{V}$, $\mathbf{r} = \mathbf{U}^H\mathbf{y}$, $v_{h^\omega} = 1$, $\hat{\mathbf{h}}^{\omega} = \mathbf{0}$, $\mathbf{s} = \mathbf{0}$, $\hat{\sigma} = 1$, $\hat{\gamma} = \mathbf{1}$, $\hat{\epsilon} = 0.001$
 - 2: **repeat**
 - 3: Calculate the mean $\boldsymbol{\mu}$ and variance $\boldsymbol{\tau}_{\mu}$ as (16).
 - 4: Calculate the mean $\hat{\mathbf{z}}$ and variance $\mathbf{v}_z$ as (14).
 - 5: Update the noise variance $\hat{\sigma}$ as (18).
 - 6: Calculate the mean $\mathbf{s}$ and variance $\boldsymbol{\tau}_s$ as (17).
 - 7: Calculate the mean $\boldsymbol{\pi}$ and variance $\boldsymbol{\tau}_{\pi}$ as (16).
 - 8: Calculate the mean $\hat{\mathbf{h}}^{\omega}$ and variance v_{h^ω} as (15).
 - 9: Calculate the variance $\hat{\gamma}$ according to (21).
 - 10: Update the parameter $\hat{\epsilon}$ according to (22).
 - 11: **until halt**
-

The computational complexity of the UTAMP-SBL is dominated by the complex multiplications required for matrix-vector operations. At the stage of initialization, the complexity of SVD is $\mathcal{O}(M^2N)$ [21]. The matrix-vector product of $\mathbf{U}^H\mathbf{y}$ is $\mathcal{O}(M^2)$. At the stage of iteration, the complexity of calculating $\boldsymbol{\tau}_{\mu}$ and $\boldsymbol{\mu}$ are respectively $\mathcal{O}(M^2 + 2M)$ and $\mathcal{O}(2MN + M)$ in the step 3 of Algorithm 1. The calculations of remaining steps only involve the component-wise vector multiplication or scalar operations which bring a small amount of complex multiplication compared with matrix-vector product. Therefore, the computational complexity of them is omitted. When the iteration reaches T , the total complexity is $\mathcal{O}(M^2N + M^2 + T(M^2 + 2M + 2MN + M))$. Discarding the low-order terms, the total complexity is $\mathcal{O}(M^2N)$.

4. Cramér-Rao Bound Analysis

In this section, the CRB is provided as a performance benchmark of the channel estimation. From Section 3, the sparse channel estimation is modeled as an SBL problem and the CRBs for SBL derived in [22] can be adopted directly.

According to [22], the lower bound on the mean squared error (MSE) matrix of $\mathbf{h}^{\omega}$ is

$$\mathbf{E} \succeq \left(\frac{\Phi^H \Phi}{\hat{\sigma}^2} + \text{diag}(\hat{\gamma}) \right)^{-1}, \quad (23)$$

where $\hat{\sigma}$ is from (18) and the n th entry of $\hat{\gamma} = [\hat{\gamma}_1, \hat{\gamma}_2, \dots, \hat{\gamma}_N]^T$ is gotten from (21). Thus, the CRB of $\hat{\mathbf{h}}^{\omega}$ is

$$\text{CRB}(\hat{\mathbf{h}}^{\omega}) = \frac{1}{N} \text{tr}(\mathbf{E}). \quad (24)$$

5. Simulation Results

The simulation parameters are chosen as follows, $N_{\text{BS}} = G_{\text{BS}} = 64$, $N_{\text{UE}} = G_{\text{UE}} = 16$, $N_{\text{BS}}^{\text{RF}} = N_{\text{UE}}^{\text{RF}} = 4$, $N_s = 4$, $K = 4$, $P = 32$ and $L = 4$. The pilots, precoders and combiners

are generated according to [17]. In order to evaluate the performance, the normalized MSE (NMSE) $E\left\{\frac{\|\mathbf{h}^\omega - \hat{\mathbf{h}}^\omega\|^2}{\|\mathbf{h}^\omega\|^2}\right\}$ between the estimated channel $\hat{\mathbf{h}}^\omega$ and the real channel $\mathbf{h}^\omega$ is introduced. We compare the NMSE performance for the OMP [5], SBL [9], EM-BBG-VAMP [7], SuRe-CSBL [15], TurboCS [10], UTAMP-SBL, and the CRB derived in Section 4.

Figure 3 shows the NMSE versus the pilot overhead at SNR = 20 dB. As defined in Section 3, N is the dimension of the received signal $\mathbf{y}$ and M is the dimension of the sparse channel $\mathbf{h}^\omega$. In this simulation, the pilot overhead is denoted as M/N . We can find that the off-grid algorithm, SuRe-CSBL, could not achieve the desired performance when the pilot is insufficient in spite of it is the winners in [15]. The SBL, EM-BBG-VAMP and UTAMP-SBL work well with 9.38% pilot overhead compared with other algorithms. Thus, when the pilot overhead is extremely low, the SuRe-CSBL, TurboCS and OMP are not sensible choices. It is also shown that the UTAMP-SBL performs distinctly better than the other estimators with the same pilot overhead. From another side, it is also concluded that the UTAMP-SBL can accurately estimate the channels with fewer pilots. For example, when the NMSE performance meets −10 dB, the OMP, SBL and EM-BBG-VAMP respectively need 22.66%, 21.88%, 16.41% pilot overhead. However, the UTAMP-SBL only needs 11.72% overhead, which means the pilot overhead is reduced by 28.58% compared with the EM-BBG-VAMP.

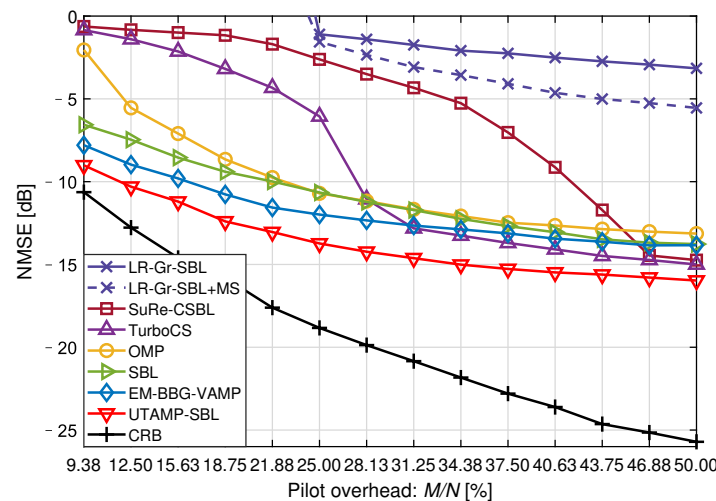


Figure 3. The NMSE versus the pilot overhead at SNR = 20 dB.

Figure 4 shows the NMSE versus the SNR with 25.00% pilot overhead. Due to insufficient of pilots, the SuRe-CSBL and TurboCS could not estimate the channels well, even if their NMSEs are lower than CRB at low SNRs. It is observed that the UTAMP-SBL has the best performance compared with the other competitors. This is because the UTAMP-SBL is robust to different types of difficult measurement matrices, such as non-zero mean, rank-deficient, correlated, or ill-conditioned matrix [16]. This feature is crucial for the practical application of this algorithm. In the channel estimation, the measurement matrix Φ_p is usually related to the complex signal processing operations as described in Section 2.3, so it can be a difficult matrix. Therefore, the UTAMP-SBL with strong robustness performs better in the channel estimation. Moreover, the matrix inversion step of SBL causing the high complexity is avoided by replacing the E-step in the EM with UTAMP, and the algorithm complexity is reduced [16].

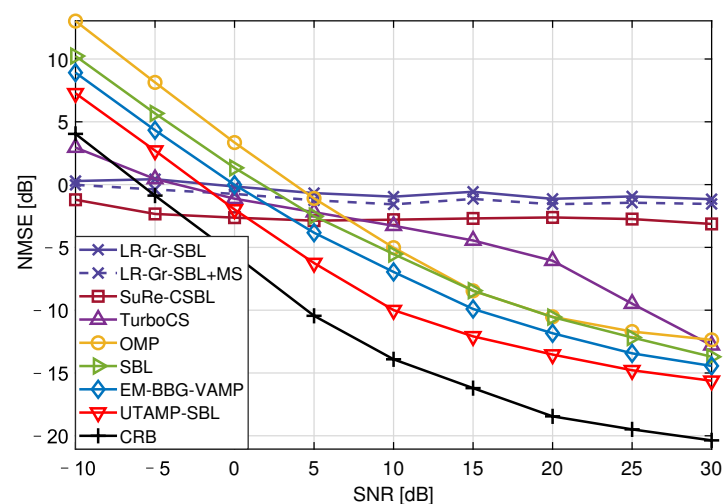


Figure 4. The NMSE versus the SNR with 25.00% pilot overhead.

Figure 5 provides the NMSE versus the number of iterations at different SNRs with 25.00% pilot overhead, which illustrates the convergences of the algorithms. Since the algorithms, SuRe-CSBL, TurboCS, are not suitable for our channel estimation when the pilot overhead is 25.00%, only the convergences of the OMP, SBL, EM-BBG-VAMP and UTAMP-SBL are compared. When SNR = −10 dB, the OMP fails to converge even if the number of iterations reaches 300, and the NMSE performance decreases with the increasing iterations. This shows that the OMP is not suitable for the low SNR cases. The SBL, EM-BBG-VAMP and UTAMP-SBL require about 100 iterations to converge. When SNR = 10 dB, the OMP is still with the worst convergence and the UTAMP-SBL is with the best convergence. However, when SNR = 30 dB, the EM-BBG-VAMP converges slightly faster than the UTAMP-SBL. From the results, the EM-BBG-VAMP converges within 20 iterations while the UTAMP-SBL requires about 40 iterations. And these two algorithms are both faster than the SBL and the OMP. Additionally, in this figure, only the performance of the UTAMP-SBL becomes better with the iteration at all SNRs, while that of other algorithms sometimes deteriorates with the iteration. This further demonstrates the advantage of the UTAMP-SBL in convergence. From above observation and analysis, the UTAMP-SBL is superior to the comparison algorithms in the convergence speed and the NMSE performance.

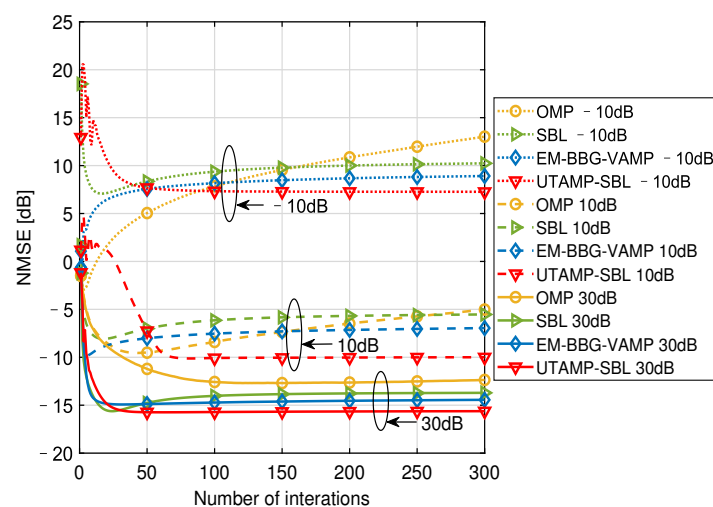


Figure 5. The NMSE versus the number of iterations at different SNRs with 25.00% pilot overhead.

6. Conclusions

In this paper, the UTAMP-SBL algorithm is applied to estimate the sparse channels for improving the performance. In order to evaluate the effect of this algorithm in practical

channel estimation, we study the multi-user uplink channel estimation for hybrid architecture mmWave massive MIMO systems. Simulation results demonstrate the superiority of the UTAMP-SBL in terms of NMSE performance and pilot overhead reduction.

Author Contributions: Conceptualization, S.H. and Y.W.; methodology, S.H.; software, S.H.; validation, S.H., Y.W. and C.L.; formal analysis, S.H. and C.L.; investigation, S.H.; data curation, S.H. and C.L.; writing—original draft preparation, S.H.; writing—review and editing, Y.W. and C.L.; funding acquisition, Y.W. All authors have read and agreed to the published version of the manuscript.

Funding: This work was supported by the Ministry of Education and China Mobile Joint Scientific Research Fund (No. MCM20180201).

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: Not applicable.

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

References

- Andrews, J.G.; Buzzi, S.; Choi, W.; Hanly, S.V.; Lozano, A.; Soong, A.C.K.; Zhang, J.C. What Will 5G Be? *IEEE J. Sel. Areas Commun.* **2014**, *32*, 1065–1082. [\[CrossRef\]](#)
- Xiao, M.; Mumtaz, S.; Huang, Y.; Dai, L.; Li, Y.; Matthaiou, M.; Karagiannidis, G.K.; Björnson, E.; Yang, K.; Chih-Lin, I.; et al. Millimeter Wave Communications for Future Mobile Networks. *IEEE J. Sel. Areas Commun.* **2017**, *35*, 1909–1935. [\[CrossRef\]](#)
- Heath, R.W.; González-Prelcic, N.; Rangan, S.; Roh, W.; Sayeed, A.M. An Overview of Signal Processing Techniques for Millimeter Wave MIMO Systems. *IEEE J. Sel. Topics Signal Process.* **2016**, *10*, 436–453. [\[CrossRef\]](#)
- Alkhateeb, A.; Ayach, O.E.; Leus, G.; Heath, R.W. Channel Estimation and Hybrid Precoding for Millimeter Wave Cellular Systems. *IEEE J. Sel. Topics Signal Process.* **2014**, *8*, 831–846. [\[CrossRef\]](#)
- Venugopal, K.; Alkhateeb, A.; Prelcic, N.G.; Heath, R.W. Channel Estimation for Hybrid Architecture-Based Wideband Millimeter Wave Systems. *IEEE J. Sel. Areas Commun.* **2017**, *35*, 1996–2009. [\[CrossRef\]](#)
- Gao, Z.; Dai, L.; Dai, W.; Shim, B.; Wang, Z. Structured Compressive Sensing-Based Spatio-Temporal Joint Channel Estimation for FDD Massive MIMO. *IEEE Trans. Commun.* **2016**, *64*, 601–617. [\[CrossRef\]](#)
- Wu, S.; Yao, H.; Jiang, C.; Chen, X.; Kuang, L.; Hanzo, L. Downlink Channel Estimation for Massive MIMO Systems Relying on Vector Approximate Message Passing. *IEEE Trans. Veh. Technol.* **2019**, *68*, 5145–5148. [\[CrossRef\]](#)
- Gao, Z.; Dai, L.; Wang, Z.; Chen, S. Spatially Common Sparsity Based Adaptive Channel Estimation and Feedback for FDD Massive MIMO. *IEEE Trans. Signal Process.* **2015**, *63*, 6169–6183. [\[CrossRef\]](#)
- Srivastava, S.; Mishra, A.; Rajoriya, A.; Jagannatham, A.K.; Ascheid, G. 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.* **2018**, *67*, 1251–1266. [\[CrossRef\]](#)
- Chen, L.; Liu, A.; Yuan, X. Structured Turbo Compressed Sensing for Massive MIMO Channel Estimation Using a Markov Prior. *IEEE Trans. Veh. Technol.* **2018**, *67*, 4635–4639. [\[CrossRef\]](#)
- Lin, X.; Wu, S.; Jiang, C.; Kuang, L.; Yan, J.; Hanzo, L. Estimation of Broadband Multiuser Millimeter Wave Massive MIMO-OFDM Channels by Exploiting Their Sparse Structure. *IEEE Trans. Wireless Commun.* **2018**, *17*, 3959–3973. [\[CrossRef\]](#)
- Zhu, J.; Liu, Z.; Song, C.; Xu, Z.; Zhong, C. Low-rank and Angular Structures aided mmWave MIMO Channel Estimation with Few-bit ADCs. In Proceedings of the 2020 IEEE 11th Sensor Array and Multichannel Signal Processing Workshop (SAM), Hangzhou, China, 8–11 June 2020; pp. 1–5.
- Zhu, J.; Wen, C.-K.; Tong, J.; Xu, C.; Jin, S. Grid-Less Variational Bayesian Channel Estimation for Antenna Array Systems With Low Resolution ADCs. *IEEE Trans. Wireless Commun.* **2020**, *19*, 1549–1562. [\[CrossRef\]](#)
- Zhang, Q.; Zhu, J.; Zhang, N.; Xu, Z. Multidimensional Variational Line Spectra Estimation. *IEEE Signal Process. Lett.* **2020**, *27*, 945–949. [\[CrossRef\]](#)
- He, Z.-Q.; Yuan, X.; Chen, L. Super-Resolution Channel Estimation for Massive MIMO via Clustered Sparse Bayesian Learning. *IEEE Trans. Veh. Technol.* **2019**, *68*, 6156–6160. [\[CrossRef\]](#)
- Luo, M.; Guo, Q.; Huang, D.; Xi, J. Sparse Bayesian Learning Based on Approximate Message Passing with Unitary Transformation. In Proceedings of the 2019 IEEE VTS Asia Pacific Wireless Communications Symposium (APWCS), Singapore, 28–30 August 2019; pp. 1–5.
- Gao, Z.; Hu, C.; Dai, L.; Wang, Z. Channel Estimation for Millimeter-Wave Massive MIMO With Hybrid Precoding Over Frequency-Selective Fading Channels. *IEEE Commun. Lett.* **2016**, *20*, 1259–1262. [\[CrossRef\]](#)
- Muhi-Eldeen, Z.; Ivrisimtzis, L.P.; Al-Nuaimi, M. Modelling and Measurements of Millimetre Wavelength Propagation in Urban Environments. *IET Microwaves Antennas Propag.* **2010**, *4*, 1300–1309. [\[CrossRef\]](#)

-
19. Hu, C.; Dai, L.; Mir, T.; Gao, Z.; Fang, J. Super-Resolution Channel Estimation for MmWave Massive MIMO With Hybrid Precoding. *IEEE Trans. Veh. Technol.* **2018**, *67*, 8954–8958. [[CrossRef](#)]
 20. Meng, X.; Wu, S.; Zhu, J. A Unified Bayesian Inference Framework for Generalized Linear Models. *IEEE Signal Process. Lett.* **2018**, *25*, 398–402. [[CrossRef](#)]
 21. Rangan, S.; Schniter, P.; Fletcher, A.K. Vector Approximate Message Passing. *IEEE Trans. Inf. Theory* **2019**, *65*, 6664–6684. [[CrossRef](#)]
 22. Prasad, R.; Murthy, C.R. Cramér-Rao-Type Bounds for Sparse Bayesian Learning. *IEEE Trans. Signal Process.* **2013**, *61*, 622–632. [[CrossRef](#)]