COMPRESSIVE CHANNEL ESTIMATION FOR IRS-AIDED MILLIMETER-WAVE SYSTEMS VIA TWO-STAGE LAMP NETWORK

Wen-Chiao Tsai, Chi-Wei Chen, and An-Yeu (Andy) Wu
Graduate Institute of Electronics Engineering, National Taiwan University, Taipei, Taiwan
daniel@access.ee.ntu.edu.tw, wilbur@access.ee.ntu.edu.tw, andyw@ntu.edu.tw

ABSTRACT
In millimeter-wave (mmWave) systems aided by intelligent reflecting surfaces (IRSs), accurate channel estimation under low pilot overhead is challenging because of the large number of passive IRS elements. By exploiting the low-rank nature of mmWave channels in the virtual angular domain (VAD) and the powerful learned approximate message passing (LAMP) network, we propose a two-stage LAMP network with row compression (RCTS-LAMP). Specifically, the two LAMP networks jointly recover the VAD channel by solving two low-dimensional sparse signal recovery problems. Moreover, row compression is adopted between the two networks to further reduce the complexity according to the row sparsity structure. Numerical results show that the estimation performance is increased while the computational complexity can be significantly reduced, which achieves a better trade-off between the accuracy and the complexity.

Index Terms— Intelligent reflecting surface (IRS), compressive sensing, channel estimation, deep learning

1. INTRODUCTION
Intelligent reflecting surface (IRS), also known as reconfigurable intelligent surface (RIS), has been proposed recently to achieve a programmable wireless communication environment via reconfigurable reflection [1]-[4]. By smartly designing the reflection coefficients of the IRS elements, we can construct an additional link between the transceiver and provide passive beamforming gain. For example, IRS is able to create virtual line-of-sight (LoS) link to bypass obstacles between transceivers for robust millimeter-wave (mmWave) communications [5], [6]. To reach the full potential of the IRS, acquiring accurate channel state information (CSI) is crucial. However, the acquisition of CSI is challenging in IRS systems since the large number of reflecting elements without sensing capabilities will result in high pilot overhead for the estimation of the cascaded channel from the base station (BS) to the user equipment (UE) through the IRS. For instance, least square (LS) estimation methods [7]-[9] were proposed to estimate the cascaded channel, but these methods require the number of training subframes to be larger or equal to the number of IRS elements, which may hinder the potential performance gain from the IRS.

By exploiting the low-rank nature of mmWave channels due to limited scattering, the training overhead and complexity can be reduced with compressive sensing (CS) techniques [10]-[13], which recover the sparse virtual angular domain (VAD) channel. However, most of these works assume the true angle of arrivals (AoAs) and angles of departures (AoDs) lie perfectly on a grid, which may never be valid and will result in significant performance degradation due to grid mismatch. Although we can mitigate the mismatch by increasing the grid resolutions, the complexity increases significantly. To solve this issue, a two-stage channel estimation (TRICE) framework [14] was proposed, where the complexity can be reduced by decoupling the channel parameters. Nevertheless, the TRICE framework also relies on high grid resolutions since it cannot achieve satisfactory performance under low grid resolutions.

In this paper, by utilizing high-performance LAMP networks [15], we propose a two-stage LAMP network with row compression (RCTS-LAMP) to achieve a better trade-off between estimation accuracy and computational complexity. The main contributions are summarized as follows:

1) Proposed two-stage LAMP (TS-LAMP) network: To improve the channel estimation accuracy under low grid resolutions, we propose reconstructing the gains of all spatial frequencies with the LAMP networks while decoupling the channel parameters. Hence, the LAMP network can reconstruct the effect of the true spatial frequency by the combination of its surrounding grids instead of a single grid in the erroneous low-resolution dictionary. Simulation results show that TS-LAMP effectively improves performance under low grid resolutions and avoids using high-resolution dictionaries.

2) Row compression via row sparsity: Based on the proposed TS-LAMP network, we further exploit the row sparsity of the angular cascaded channel to reduce the computational complexity of the second LAMP in the second stage. Specifically, since the row indices of the angular cascaded channel represent the angle of arrivals (AoAs) at the BS and the first LAMP network has transformed the rows from spatial domain to angular domain, we can directly prune the rows of low power without performance loss. Simulation results show that the proposed RCTS-LAMP network can maintain the performance of TS-LAMP with an overall 98.7% less complexity than the joint method via one LAMP network.
2. SYSTEM MODEL AND PRIOR ARTS

2.1. System and Channel Models

In this paper, we consider an IRS-aided single-user mmWave communication system as illustrated in Fig. 1, where the BS equipped with the \(N_B\)-antenna uniform linear array (ULA) and \(N_{RF}^B \leq N_B\) RF chains cooperates with the IRS to serve a single-antenna UE. Although the direct link may be unavailable due to blockage, reliable communication can be achieved by the indirect link aided by the IRS equipped with the \(N_2\)-element uniform planar array (UPA) composed of \(N_2^H\) horizontal and \(N_2^V\) vertical elements, where \(N_2 = N_2^H N_2^V\).

The baseband equivalent channels from the IRS to the BS and from the UE to the IRS are respectively denoted by \(G \in \mathbb{C}^{N_B \times N_S}\) and \(h_r \in \mathbb{C}^{N_S \times 1}\). Similar to [10]-[15], we adopt the classical Saleh-Valenzuela channel model [16] to characterize the channels \(G\) and \(h_r\) as

\[
G = \sqrt{\frac{N_B N_S}{L_G}} \sum_{i=1}^{L_G} \alpha_i a_B(\phi_i^G) a_S^H(\phi_i^G, \gamma_i^G),
\]

(1)

\[
h_r = \frac{N_S}{L_r} \sum_{i=1}^{L_r} \alpha_i a_r(\theta_i^r, \gamma_i^r),
\]

(2)

where \(L_G\) and \(L_r\) are respectively the number of paths for \(G\) and \(h_r\), \(\alpha_i\) \((\alpha_i)\) denotes the complex path gain of the corresponding path, \(\phi_i^G\) denotes the AoA spatial frequency at the BS, \(\phi_i^G\) and \(\gamma_i^G\) \((\theta_i^r\) and \(\gamma_i^r)\) represent the azimuth and elevation AoDs (AoAs) spatial frequencies at the IRS, and \(a_B(\phi) \in \mathbb{C}^{N_B \times 1}\) and \(a_r(\theta, \gamma) \in \mathbb{C}^{N_S \times 1}\) are the array response vectors at the BS and the IRS, respectively.

Furthermore, we denote \(H_c = h_r \otimes G \in \mathbb{C}^{N_B \times N_S}\) as the cascaded channel for the indirect link, where \(\otimes\) denotes the Khatri-Rao product. Based on (1) and (2), \(H_c\) can be expressed as [13]

\[
H_c = \sqrt{\frac{N_B N_S}{L_G L_r}} \sum_{i=1}^{L_G} \sum_{j=1}^{L_r} \alpha_i \alpha_j a_B(\phi_i^G) a_r^S(\phi_i^G, \gamma_i^G) a_r^c(\theta_j^c, \gamma_j^c),
\]

(3)

where \(\phi_i^G = \phi_i^G - \phi_i^r\) \((\gamma_i^G = \gamma_i^G - \gamma_i^r)\) denotes the effective azimuth (elevation) spatial frequency at the IRS. Due to the limited scattering nature of mmWave channels, \(H_c\) exhibits sparsity in the VAD. By constructing two dictionary matrices \(A_B \in \mathbb{C}^{N_B \times K_B}\) and \(A_c \in \mathbb{C}^{N_S \times K_S}\) at the BS and the IRS, respectively, we can decompose \(H_c\) as

\[
H_c = A_B \tilde{H}_c A_c^H,
\]

(4)

where \(\tilde{H}_c \in \mathbb{C}^{K_B \times K_S}\) denotes the angular cascaded channel with grid resolutions \(K_B \geq N_B\) and \(K_S = K_B^H N_2^V \geq K_B^H N_2^V = N_2^V\) and \(\tilde{H}_c\) is a sparse matrix since \(L_G L_r \ll K_B K_S\).

For uplink channel estimation, the UE transmits known pilots to the BS over \(T\) time blocks. In the \(t\)-th time block, after being reflected by the IRS through a phase shift vector \(\theta_t \in \mathbb{C}^{N_S \times 1}\), the pilots are received at the BS as \(y_t \in \mathbb{C}^{Q \times 1}\) through a (random) combining matrix \(W \in \mathbb{C}^{N_B \times Q}\). By stacking \(y_t\) as the overall received signal \(Y = [y_1, ..., y_T] \in \mathbb{C}^{Q \times T}\) and denoting the reflection matrix as \(\Theta = [\theta_1, ..., \theta_T]\), we have [14]

\[
Y = WH_c \Theta + N,
\]

(5)

where \(N\) denotes the additive white Gaussian noise whose entries are independent and distributed as \(\mathcal{CN}(0, \sigma^2_N)\).

2.2. Compressive Channel Estimation (CCE) [10]-[15]

Channel estimation for IRS-aided systems is challenging since the large number of reflecting elements without signal processing capability results in high pilot overhead. To reduce the high pilot overhead, we can exploit the sparsity of the angular cascaded channel \(\tilde{H}_c\), and apply CS algorithms to solve the sparse signal recovery problem as follows [10]

\[
y = \text{vec}(Y) = (\Theta^T \otimes W^H)(A_c^* \circ A_B) \hat{h}_c + n,
\]

(6)

where \(\otimes\) denotes the Kronecker product, \(\hat{h}_c = \text{vec}(\hat{H}_c)\), \(n = \text{vec}(N)\), and \(\text{vec}(\cdot)\) is the vectorization operation.

However, conventional CS algorithms such as orthogonal matching pursuit (OMP) cannot achieve satisfactory estimation performance under low grid resolutions. Hence, the LAMP network is used in [15] to boost the channel estimation performance by unfolding conventional model-based algorithm (approximate message passing) into learnable layers, which are optimized through powerful deep learning (DL) techniques. For a \(L\)-layer LAMP network, the \(i\)-th layer outputs are shown below

\[
\hat{h}_{c,i} = \eta_{st} \left( \hat{h}_{c,i-1} + Z_i v_{i-1}; \frac{\alpha_i}{\sqrt{q_i}} \| v_{i-1} \| \right),
\]

(7)

\[
v_i = y - W_i \hat{h}_{c,i}^e + \frac{1}{q_i} \left\| \hat{h}_{c,i}^e \right\|_0 v_{i-1},
\]

(8)

where \([W_i, Z_i, \alpha_i]_{i=1}^L\) are the trainable parameters to be optimized, and \(\eta_{st}\) denotes the soft thresholding denoiser.

Finally, the estimated angular cascaded channel is the output of the last layer \(\hat{h}_{c,L}^e\).

Besides, considering that the joint estimation problem in (6) may be computationally prohibitive due to the 3D dictionary matrix \(A_c^* \circ A_B\), where each atom is a function of the AoA at the BS, the effective azimuth and elevation AoDs at the IRS. A two-stage channel estimation (TICE) framework [14] is proposed to reduce the computational complexity by decoupling the channel parameters.
3. PROPOSED TWO-STAGE LAMP NETWORK WITH ROW COMPRESSION

3.1 Two-Stage LAMP Network

Although the proposed TRICE framework [14] can effectively reduce the complexity by decoupling the channel parameters, it may not achieve satisfactory estimation performance under low grid resolutions since the deviation between the estimated spatial frequencies and the true spatial frequencies may be large. To improve the performance of channel estimation, we can adopt dictionaries with higher grid resolutions, but the performance improvement is limited while the complexity is significantly increased. In contrast to the TRICE framework based on the explicit estimation of the spatial frequencies, iterative thresholding algorithms like LAMP [15] directly recover the gains of all the spatial frequencies in the dictionary. Hence, LAMP can achieve better performance by using the combination of multiple networks is benefited from the high performance of the LAMP network. Specifically, instead of solving the joint channel estimation, we can adopt dictionaries with higher frequencies may be large. To improve the performance of the LAMP network, we propose a twostage LAMP (TS-LAMP) network for cascaded channel estimation in IRS-aided communications: (a) TS-LAMP without row compression, and (b) TS-LAMP with row compression (CMP) based on row sparsity.

Therefore, to achieve better estimation performance under low grid resolutions and strike a better trade-off between performance and complexity, we propose a two-stage LAMP (TS-LAMP) network for cascaded channel estimation as illustrated in Fig. 2(a). The proposed TS-LAMP network is benefited from the high performance of the LAMP network and also enjoys low complexity by decoupling the channel parameters. Specifically, instead of solving the joint recovery problem in (6), the TS-LAMP utilizes the structure resides in the overall received signal $Y$ as follows

$$Y = W H A_B \hat{H}_c A_S^T \Theta + N = A_1 \hat{H}_c A_2^T + N,$$

where we apply the VAD representation (4) into (5), and $A_1 = W H A_B \in \mathbb{C}^{Q \times K_B}$ and $A_2 = \Theta^T A_S \in \mathbb{C}^{T \times K_S}$ are the equivalent measurement matrices for $\hat{H}_c$ at the BS and the IRS, respectively. Note that based on (3), since each non-zero element of $\hat{H}_c$ is based on a cascaded path $(l,l')$, the row and column indexes of this element are characterized by $q_l^{l'}$ and $(\theta_{l,l'}^H, y_{l,l'}^H)$, respectively. Hence, $\hat{H}_c$ has $L_G$ non-zero rows, where each non-zero row has $L_r$ non-zero entries.

The proposed TS-LAMP is summarized in Algorithm 1, which can be explained as follows. In the first stage, we can observe that $H_{row}^e = \hat{H}_c A_S^T$ is a $L_G$-row sparse matrix, which can be recovered from $Y$ with a LAMP network using the equivalent measurement matrix $A_1$. Note that since each column of $H_{row}^e$ results in a single measurement vector (SMV) in $Y$, the LAMP network can recover the columns of $H_{row}^e$ simultaneously through the parallel computation of all the SMVs in $Y$. Denote $H_{row}^e$ as the estimated row sparse matrix, which is the output of the first LAMP network. Subsequently, in the second stage, we recover $\hat{H}_c$ from $H_{row}^e$ through another LAMP network based on the equivalent measurement matrix $A_2$, which equivalently solves the following sparse signal recovery problem

$$(H_{row}^e)^T = A_2 \hat{H}_c^e + Z,$$

where $Z$ is the residual noise after the estimation of the first LAMP network. Similar to the first LAMP network, the second LAMP network can also recover the rows of $\hat{H}_c$ simultaneously through the parallel computation of all the rows of $H_{row}^e$ (the resulting SMVs). Finally, we can obtain the estimated cascaded channel $H_c^e$ by transforming the angular channel into the spatial channel in Step 3.

3.2. Row Compression based on Row Sparsity

With the proposed TS-LAMP network, we can avoid the expensive 3D dictionary used in the joint estimation method by decoupling the channel parameters and benefit from the high performance of the LAMP network. However, since the second stage of the TS-LAMP network completely recovers all the $K_B K_S$ elements in the angular cascaded channel $\hat{H}_c$, the computational complexity is still high under high grid resolutions $K_B$ and $K_S$. Fortunately, based on the row sparsity of the $L_G$-row sparse matrix $H_{row}^e$, there exists redundancy in the second stage. We can reduce the complexity by only considering the rows whose spatial frequencies are effective in representing the true spatial frequencies $q_l^{l'}$.

The proposed TS-LAMP with row compression (RCTS-LAMP) is illustrated in Fig. 2(b), where the row compression block is introduced between the two LAMP networks to sift out the useful rows of $H_{row}$. To characterize the usefulness of a row $H_{row}(i,:)$, we exploit the $l_q$-norm of this row $||H_{row}(i,:)||_1$, which is an efficient indicator of the power level and will be large if the corresponding spatial frequency is close to the true spatial frequency. Furthermore, we utilize the property that the distribution of $l_q$-norms is positively skewed due to row sparsity, where most of the rows have

Algorithm 1: Proposed Two-Stage LAMP Network

**Input:** Overall received signal $Y$

1. **Stage 1:** Return the estimated row sparse matrix $H_{row}^e$ from $Y$ with the first LAMP network.
2. **Stage 2:** Return the estimated angular cascaded channel $H_c^e$ from $H_{row}^e$ with the second LAMP network.
3. $H_c^e = A_B H_{row}^e A_S^T$

**Output:** Estimated cascaded channel $H_c^e$
small $l_1$-norms and the few rows with large $l_1$-norms establish a long right tail of the distribution. Hence, we propose to use the mean of $l_1$-norms as a dynamic threshold for row compression, where the rows with large $l_1$-norms will be retained, and we can filter out more than 50% of the rows since the mean is larger than the median in the positively skewed distribution.

In contrast to the TS-LAMP network, RCTS-LAMP only utilizes the support rows $\Omega$ of $H_{\text{row}}$ to be the input of the second LAMP network, where $\Omega$ can be expressed as

$$\Omega = \{i: \|H_{\text{row}}(i,:)\|_1 > \sum_{j=1}^{K_{B}} \|H_{\text{row}}(j,:)\|_1 / K_{B}\}.$$ (11)

As a result, RCTS-LAMP can achieve complexity reduction by only recover $|\Omega|K_{S}$ of all the $K_{B}K_{S}$ elements in $H_{\text{c}}$.

### 4. SIMULATION RESULTS

This section shows the simulation results of our proposed TS-LAMP and RCTS-LAMP networks. For comparison, we also simulate the TRICE framework [14] denoted as TRICE-OMP, and the joint CS method with the LAMP network [15]. In our simulation, we consider that the BS has $N_{B} = 32$ antennas, the IRS has $N_{S} = 256$ ($N_{B}^S = N_{S}^S = 16$), the numbers of measurements are $Q = 16$ and $T = 64$, the numbers of paths are $L_{G} = L_{r} = 3$ where the angles of each path follow the uniform distribution $\mathcal{U}(-\pi/2, \pi/2)$ and the path gains are distributed as $a_i \sim \mathcal{CN}(0,1)$. The SNR is defined as $\mathbb{E}[\|W^{H}H_{\text{c}}\|_F^2 / \|N\|_F^2]$ from (5), where the elements of $W$ and $\Theta$ have unit gain and random phases which follow the $\mathcal{U}(0,2\pi)$ distribution.

To train the proposed TS-LAMP, the trainable parameters of the first LAMP $\{W_{i}^{1}, Z_{i}^{1}\}_{i=1}^{L_{1}}$ and the second LAMP $\{W_{i}^{2}, Z_{i}^{2}\}_{i=1}^{L_{2}}$ are initialized as $W_{i}^{1} = (Z_{i}^{1})^H = A_{1}$ and $W_{i}^{2} = (Z_{i}^{2})^H = A_{2}$, respectively. For the RCTS-LAMP, we initialize the network with the optimized parameters of TS-LAMP to guarantee performance. During training, Adam optimizer is used to update the parameters, and the NMSE metric is our loss function defined as $\text{NMSE} = \mathbb{E}[\|H_{\text{c}} - H_{\text{c}}^{\hat{}}\|_F^2 / \|H_{\text{c}}\|_F^2]$. Besides, to evaluate the performance under different grid resolutions, the grid resolution is indicated by the oversampling rate $\beta$ defined as $\beta = K_{G}/N_{B} = K_{B}^G / N_{B}^G = K_{S}^G / N_{S}^G$.

Fig. 3 shows the NMSE performance comparison against SNR. We can observe that the performance of TRICE-OMP under low grid resolution ($\beta = 1$) cannot achieve satisfactory estimation accuracy. Although the LAMP network outperforms TRICE-OMP under $\beta = 1$, the complexity is much higher due to the joint CS method. In contrast, our proposed TS-LAMP and RCTS-LAMP networks under $\beta = 1$ not only outperform TRICE-OMP under $\beta = 2$ but also enjoy low computational complexity from the decoupling of channel parameters. Furthermore, the proposed TS-LAMP and RCTS-LAMP networks under $\beta = 2$ can even outperform TRICE-OMP under high grid resolution ($\beta = 9$).

### 5. CONCLUSION

In this paper, we propose a RCTS-LAMP network for channel estimation in IRS-aided systems. By decoupling the channel parameters and exploiting the row sparsity, we can achieve a better trade-off between accuracy and complexity. Specifically, the RCTS-LAMP can not only improve the estimation accuracy but also reduce the computational complexity of LAMP by two low-dimensional dictionaries, which will be more suitable for practical IRS-aided systems.

<table>
<thead>
<tr>
<th>Algorithm</th>
<th>Computational Complexity</th>
</tr>
</thead>
<tbody>
<tr>
<td>TRICE-OMP</td>
<td>$O(L_{G}TQK_{B} + L_{G}L_{r}TK_{S})$</td>
</tr>
<tr>
<td>LAMP</td>
<td>$O(LTQK_{G}K_{S})$</td>
</tr>
<tr>
<td>TS-LAMP</td>
<td>$O(L_{1}TQK_{B} + L_{2}TK_{B}K_{S})$</td>
</tr>
<tr>
<td>RCTS-LAMP</td>
<td>$O(L_{1}TQK_{B} + L_{2}TK_{S})$</td>
</tr>
</tbody>
</table>

Besides, we can observe that RCTS-LAMP can maintain or even slightly improve the performance of TS-LAMP due to the noise suppression, which can further reduce the complexity and achieve a better trade-off between estimation accuracy and computational complexity.
REFERENCES


