Learning-Aided Joint Channel Estimation and Detection in OTFS with Graph Neural Networks

Xinwei Qu

Abstract—

***Index Terms—*OTFS joint channel estimation and symbol detection, graph neural network, 6G**

I. INTRODUCTION

A. Related Works

B. Contributions

The main contributions of this paper are summarized as follows.

- we propose an OTFS frame design that increase the transmission efficiency and overcomes the overspreading channel issue
- We propose a novel neural network-based framework for joint channel estimation and symbol detection for OTFS systems, referred as Joint Parallel Interference Cancellation Network (JPICNet) framework. The proposed framework integrates the joint PIC scheme and graph neural network.
- we incorporate attention mechanism into neural network-based framework to improve the performance.

C. Notations

a, $\boldsymbol{a}$ and $\boldsymbol{A}$ demote scalar, vector, and matrix respectively. $\mathbb{C}^{M \times N}$ denotes the set of $M \times N$ dimensional complex matrix. $(\cdot)^T$, $(\cdot)^H$, $(\cdot)^*$, and $[\cdot]_M$ represent the transpose, Hermitian transpose, conjugate, and mod-M operations.$\odot$ denotes Hadamard multiplication. $\text{diag}(\boldsymbol{a})$ denotes the operation to diagonize a vector $\boldsymbol{a}$, $\text{off}(\boldsymbol{A})$ forces all diagonal elements (A_{ii}) to zero. We define $\boldsymbol{a} = vec(\boldsymbol{A})$ as the column-wise vectorization of matrix $\boldsymbol{A}$.

II. SYSTEM MODEL

We consider an OTFS system that transmits symbols $\boldsymbol{X_{DD}} \in \mathbb{C}^{K \times L}$ over the delay-Doppler (DD) domain grids $\{(k\Delta\tau, l\Delta v)|k = 0, ..., K-1, l = 0, ..., L-1\}$, where $\Delta\tau$ and Δv denote the delay and Doppler resolutions, where:

- $\Delta\tau = T/L$ is the delay resolution, with T being the symbol duration of the OTFS system,
- $\Delta v = \Delta f/$ is the Doppler resolution, with $\Delta f = 1/T$ denoting the subcarrier spacing.

The system bandwidth is $L\Delta f$, and the OTFS frame duration is NT. The transmitter first maps the symbols $x[k,l]$ to the time-frequency (TF) domain grids $\{(l\Delta f, kT)|\}$ via the inverse finite symplectic Fourier transform (ISFFT). The time-domain signal is synthesized using a conventional OFDM modulator with a transmit shaping pulse yet employs a single initial cyclic prefix spanning the full delay-Doppler spread duration, contrasting with conventional OFDM's per-symbol cyclic prefix insertion. The time-domain signal is transmitted over a time-varying wireless channel characterized by the delay-Doppler impulse response $h(\tau, v)$ as [1],

$$h(\tau, v) = \sum_{i=1}^{P} h_i \delta(\tau - \tau_i)\delta(v - v_i), \quad (1)$$

where $\delta(\cdot)$ denotes the Dirac delta function, $h_i \sim \mathcal{N}(0, \frac{1}{P})$ is the gain of the i-th propagation path, and P represents the total number of paths. Each path is characterized by distinct delay and/or Doppler shifts, modeling the channel response between the receiver and either moving reflectors or the transmitting source. The delay and Doppler shifts are given as,

$$\tau_i = l_i \frac{T}{L}, v_i = k_i \frac{\Delta f}{K}, \quad (2)$$

respectively. Let the integers $l_i \in [0, l_{max}]$ and $k_i \in [-k_{max}, k_{max}]$ represent the delay and Doppler shift indices, respectively, where l_{max} and k_{max} denote the maximum delay index and maximum Doppler shift index across all propagation paths. Note that we restrict our consideration to integer-valued indices, as fractional delay and Doppler shifts can be equivalently represented through virtual integer taps in the delay-Doppler domain using the techniques described in [2–4].

At the receiver, the time-domain signal is first converted to the time-frequency (TF) domain through matched filtering and OFDM demodulation, then transformed to the delay-Doppler (DD) domain via inverse symplectic finite Fourier transform (ISFFT), yielding the received symbols $Y_{\text{DD}} \in \mathbb{C}^{K \times L}$. The DD domain input-output relationship can be formulated in vector form as [5],

$$\boldsymbol{y} = \boldsymbol{Hx} + \tilde{\boldsymbol{z}}, \quad (3)$$

where $\boldsymbol{x} = vec(\boldsymbol{X_{DD}^T})$, $\boldsymbol{y} = vec(\boldsymbol{Y_{DD}^T})$, and $\tilde{z} \sim \mathcal{CN}(0, \sigma_z^2)$ is an independent and identically distributed (i.i.d.) Gaussian noise [1, 4, 5].

III. JOINT PARALLEL INTERFERENCE CANCELLATION NETWORK

In this section, we propose a novel JPICNet framework that enhance the joint PIC scheme with graph neural network. Fig.**??** illustrates the process of the proposed framework. The proposed framework consists of channel parameters estimation (CPE), the GNN-aided symbol detection (GSD) and the GNN-aided channel estimation (GCHE) modules. CPE runs only once to generate initial input for GSD, while GSD and GCHE iteratively exchange outputs in a feedback loop.

A. Channel Parameters Estimation

Estimation area $k \in [0, K-1], l \in [0, lmax]$

Based on the transmission arrangement, we can express the received signal in the channel estimation area as,

$$
\begin{aligned}
y[k,l] =&\, h[k-k_p, l-l_{p0}]x_p \\
&+ \sum_i^P h_i \beta[k,l,k_i,l_i]x_d[[k-k_i]_K, [l-l_i]_L] \quad (4) \\
&+ \tilde{z}[k,l].
\end{aligned}
$$

The parameter $\beta[k,l,k_i,l_i]$ has distinct forms for different OTFS waveforms:

  i) *Ideal waveform*:

$$
\beta_{\text{ideal}}[k,l,k_i,l_i] = e^{-j2\pi \frac{k_i l_i}{KL}}, \quad (5)
$$

  ii) *Rectangular waveform*:

$$
\beta_{\text{rect}}[k,l,k_i,l_i] =
\begin{cases}
e^{j2\pi \frac{k_i(l-l_i)}{KL}}, & l_i \le l < L, \\
e^{-j2\pi\left(\frac{[k-k_i]_K}{K}\right)}e^{j2\pi\frac{k_i(l-l_i)}{KL}}, & 0 \le l < l_i.
\end{cases} \quad (6)
$$

From (4), the second term arises from the interference between pilot and data symbols. For simplicity, this cross-interference can be modeled as an additive noise term in subsequent analysis, i.e.,

$$
\mathcal{I}[k,l] = \sum_i^P h_i \beta[k,l,k_i,l_i]x_d[[k+k_i]_K, [l+l_i]_L]. \quad (7)
$$

The interference term $\mathcal{I}[k,l]$ is zero-mean ($\mathbb{E}[\mathcal{I}\{[k,l]\} = 0$) with its variance given by,

$$
\text{Var}(\mathcal{I}[k,l]) = \sigma_d^2. \quad (8)
$$

Proof: The proof is given in Appendix A. □

Based on the received signal model in (4), we propose a threshold-based channel estimation algorithm as follows. For each delay-Doppler bin $[k,l]$, if the received energy $|y[k,l]|^2$ falls below a predefined threshold $\mathcal{T}$, we declare an empty bin (no path present). Otherwise, we obtain an initial channel estimate as,

$$
\hat{h}[k-k_p, l-l_{p0}] = y[k,l]/x_p. \quad (9)
$$

Therefore, the threshold provides a mask vector to indicate whether there is a path locating at the delay vector $\boldsymbol{k}_m$ and Doppler vector $\boldsymbol{l}_m$, i.e,

$$
\begin{aligned}
\boldsymbol{h}_m =&\, [h_{m0}, \cdots, h_{mi}, \cdots, h_{mP_{max}}], \quad &(10\text{a}) \\
\boldsymbol{k}_m =&\, [-k_{max}, \cdots, 0, \cdots, k_{max}, \cdots, k_{mi} \quad &(10\text{b}) \\
&\, \cdots, -k_{max}, \cdots, 0, \cdots, k_{max}], \quad &(10\text{c}) \\
\boldsymbol{l}_m =&\, [0, 1, \cdots, l_{max}, \cdots, l_{mi}, \cdots, 0, 1, \cdots, l_{max}]. \quad &(10\text{d})
\end{aligned}
$$

where $h_{mi} = 1$ denotes a path at (k_{mi}, l_{mi}) and 0 otherwise, and $P_{max} = (l_{max}+1)(2k_{max}+1))$.

B. GNN-aided Symbol Detection

At each iteration t, the symbol vector $\boldsymbol{x}$ defined in (3) is treated as a random vector, with its statistical parameters estimated from the observation $\boldsymbol{y}$. based on the same equation. The conditional Gaussian probability distribution function (PDF) is subsequently derived as,

$$
p^{(t)} = (\boldsymbol{x}|\boldsymbol{y}) = \mathcal{N}(\boldsymbol{x}, \boldsymbol{\mu}_{\boldsymbol{x}}^{(t)}; \Sigma_{\boldsymbol{x}}^{(t)}), \quad (11)
$$

where

$$
\boldsymbol{\mu}_{\boldsymbol{x}}^{(t)} = (\boldsymbol{I} \odot \hat{\boldsymbol{H}}^{(t)^H} \hat{\boldsymbol{H}}^{(t)})^{-1} \hat{\boldsymbol{H}}^{(t)^H} y, \quad (12\text{a})
$$

$$
\Sigma_{\boldsymbol{x}}^{(t)} = (\sigma_z^2 + \sigma_x^2 D_{\Delta \hat{\boldsymbol{H}}})(\boldsymbol{I} \odot \hat{\boldsymbol{H}}^{(t)^H} \hat{\boldsymbol{H}}^{(t)})^{-1}. \quad (12\text{b})
$$

$$
(12\text{c})
$$

where

$$
D_{\Delta \hat{\boldsymbol{H}}} = diag(\sum_{i=0}^{KL-1} \boldsymbol{\Sigma}_{\hat{\boldsymbol{H}}}[:,i]), \quad (13)
$$

$\Sigma_{\hat{\boldsymbol{H}}}$ denotes the variance matrix of $\hat{\boldsymbol{H}}$, and $\sum_{i=0}^{KL-1} \boldsymbol{\Sigma}_{\hat{\boldsymbol{H}}}[:,i]$ means we sum all columns of $\Sigma_{\hat{\boldsymbol{H}}}$ together to build a column vector.

Proof: The proof is given in Appendix B. □

C. GNN-aided Channel Estimation

Rewriting (3) for channel estimation, we have

$$
\boldsymbol{y} = \boldsymbol{\Phi}\boldsymbol{h} + \tilde{z}, \quad (14)
$$

where $\boldsymbol{h}$ is the path gain in the time domain,

$$
\boldsymbol{\Phi} =
$$

$$
\begin{bmatrix}
\Phi_{k_{m0},l_{m0}}(1,1) & \cdots & \Phi_{k_{mP_{max}},l_{k_{mP_{max}}}}(KL, P_{max}) \\
\vdots & \ddots & \vdots \\
\Phi_{k_{m0},l_{m0}}(KL,1) & \cdots & \Phi_{k_{mP_{max}},l_{k_{mP_{max}}}}(KL, P_{max})
\end{bmatrix},
$$

where $(kL+l, i)$th entry of $\Phi \in \mathcal{C}^{(KL-1) \times P_{max}}$ is [8],

$$
\Phi_{k_{mi},l_{mi}}(kL+i, i) = X([k-k_{mi}]_K, [l - l_{mi}]_L)\beta[k,l,k_{mi},l_{mi}]. \quad (15)
$$

Inspired by [6, 7], we implement a linear minimum-mean-squared error (MMSE) channel estimation , i.e.,

$$
\begin{aligned}
\hat{\boldsymbol{h}}^{(t)} =&\, \boldsymbol{R_h}\boldsymbol{\Phi}^{(t)^H}(\boldsymbol{\Phi}^{(t)}R_{\boldsymbol{h}}\boldsymbol{\Phi}^{(t)^H} + \\
&\, \frac{1}{P}diag(\boldsymbol{h}_m)D_{\Delta\Phi^{(t)}} + \sigma_z^2 \boldsymbol{I})^{-1}\boldsymbol{y}, \quad (16\text{a})
\end{aligned}
$$

$$
\begin{aligned}
\text{Var}(\hat{\boldsymbol{h}}^{(t)}) =&\, \boldsymbol{R_h} - \boldsymbol{R_h}\boldsymbol{\Phi}^{(t)^H}(\boldsymbol{\Phi}^{(t)}R_{\boldsymbol{h}}\boldsymbol{\Phi}^{(t)^H} + \\
&\, \frac{1}{P}diag(\boldsymbol{h}_m)D_{\Delta\Phi^{(t)}} + \sigma_z^2 \boldsymbol{I}))^{-1}\Phi^{(t)}\boldsymbol{R_h}, \quad (16\text{b})
\end{aligned}
$$

$$
(16\text{c})
$$

where $\boldsymbol{\Phi}^{(t)}$ is the estimation of $\boldsymbol{\Phi}$ in t-th iteration, and,

$$
D_{\Delta\Phi^{(t)}} = diag(\sum_{i=0}^{KL-1} \boldsymbol{\Sigma}_{\Delta\Phi^{(t)}}[:,i]). \quad (17)
$$

$\Sigma_{\Delta\Phi^{(t)}}$ denotes the variance matrix of $\Phi^{(t)}$, and $\sum_{i=0}^{KL-1} \boldsymbol{\Sigma}_{\Delta\Phi^{(t)}}[:,i]$ means we sum all columns of $\Sigma_{\Delta\Phi^{(t)}}$ together to build a column vector.

Proof: The proof is given in Appendix C. □

APPENDIX A
CPE CROSS-INTERFERENCE

$$\mathbb{E}\{\mathcal{I}[k,l]\} = \sum_{i=1}^{P} \underbrace{\mathbb{E}\{h_i\}}_{0} \beta[k,l,k_i,l_i]$$
$$\underbrace{\mathbb{E}\{x_d[[k+k_i]_K,[l+l_i]_L]\}}_{0} = 0 \quad (18)$$

$$\mathrm{Var}(\mathcal{I}[k,l]) = \mathbb{E}\left\{|\mathcal{I}[k,l]|^2\right\}, \quad (19)$$

where

$$|\mathcal{I}[k,l]|^2 = \sum_{i=1}^{P} |h_i|^2 |\beta[k,l,k_i,l_i]|^2 |x_d[[k+k_i]_K,[l+l_i]_L]|^2$$
$$+ \sum_{i \neq j} h_i h_j^* \beta[k,l,k_i,l_i] \beta[k,l,k_j,l_j]^* \cdot$$
$$x_d[[k+k_i]_K,[l+l_i]_L] x_d[[k+k_j]_K,[l+l_j]_L]^* \quad (20)$$

The expectation of the second term in (20) vanishes because $\mathbb{E}\{h_i h_j^*\} = 0$, i.e.,

$$\mathrm{Var}(\mathcal{I}[k,l]) = \sum_{i=1}^{P} \underbrace{|h_i|^2}_{1/P} \cdot \underbrace{|\beta[k,l,k_i,l_i]|^2}_{1} \cdot$$
$$\underbrace{|x_d[[k+k_i]_K,[l+l_i]_L]|^2}_{\sigma_d^2} \quad (21)$$
$$= \sum_{i=1}^{P} 1/P \cdot \sigma_d^2 = \sigma_d^2$$

APPENDIX B
GSD VARIANCE

For simplification, we denote $\hat{H}^{(t)}$ as $\hat{H}$ and the power of x as σ_x^2. Therefore, the estimated symbol is

$$\mu_x = (I \odot \hat{H}^H \hat{H})^{-1} \hat{H}^H y,$$
$$= (I \odot \hat{H}^H \hat{H})^{-1} \hat{H}^H (Hx + \tilde{z}),$$
$$= \underbrace{(I \odot \hat{H}^H \hat{H})^{-1} \hat{H}^H}_{W_{GSD}} \times \quad (22)$$
$$(\hat{H}x + Hx - \hat{H}x + \tilde{z}),$$
$$= W_{GSD}(\hat{H}x + \Delta\hat{H}x + \tilde{z}),$$
$$= W_{GSD}\hat{H}x + W_{GSD}\Delta Hx + W_{GSD}\tilde{z}.$$

The estimation error is

$$e = \mu_x - x,$$
$$= (W_{GSD}\hat{H} - I)x + W_{GSD}\Delta Hx + W_{GSD}\tilde{z}. \quad (23)$$

Please note $\hat{H}$ is the estimation of H, i.e., $\mathbb{E}\{\Delta\hat{H}\} = 0$ and $\mathbb{E}\{\tilde{z}\} = 0$. Therefore, the corvariance of μ_x is,

$$\mathrm{Cov}(\mu_x) = \mathbb{E}\{ee^H\},$$
$$= \mathbb{E}\{(W_{GSD}\hat{H} - I)xx^H(W_{GSD}\hat{H} - I)^H\}$$
$$+ \mathbb{E}\{W_{GSD}\Delta\hat{H}xx^H\Delta\hat{H}^H W_{GSD}^H\}$$
$$+ \mathbb{E}\{W_{GSD}\tilde{z}\tilde{z}^H W_{GSD}^H\}, \quad (24)$$
$$= \sigma_x^2 \mathbb{E}\{(W_{GSD}\hat{H} - I)(W_{GSD}\hat{H} - I)^H\}$$
$$+ \sigma_x^2 \mathbb{E}\{W_{GSD}\Delta\hat{H}\Delta\hat{H}^H W_{GSD}^H\}$$
$$+ \sigma_z^2 \mathbb{E}\{W_{GSD}W_{GSD}^H\}.$$

Since the estimated channel coefficients are assumed to be mutually independent, the first term in (24) can thus be simplified as,

$$\sigma_x^2 \mathbb{E}\{(W_{GSD}\hat{H} - I)(W_{GSD}\hat{H} - I)^H\}$$
$$= \sigma_x^2 \mathbb{E}\{W_{GSD}\hat{H}\hat{H}^H W_{GSD}^H - I\},$$
$$= \sigma_x^2 \mathbb{E}\{(I \odot \hat{H}^H \hat{H})^{-1}\hat{H}^H \hat{H}\hat{H}^H \hat{H}(I \odot \hat{H}^H \hat{H})^{-1} \quad (25)$$
$$- I\},$$
$$= \sigma_x^2 \mathbb{E}\{I - I\} = 0.$$

Similarly, the last term in (24) can be simplified as,

$$\sigma_z^2 \mathbb{E}\{W_{GSD}W_{GSD}^H\}$$
$$= \sigma_z^2 \mathbb{E}\{(I \odot \hat{H}^H \hat{H})^{-1}\hat{H}^H \hat{H}(I \odot \hat{H}^H \hat{H})^{-1}\}, \quad (26)$$
$$= \sigma_z^2 (I \odot \hat{H}^H \hat{H})^{-1}.$$

Therefore, (24) can be written as,

$$\mathrm{Cov}(\mu_x) = \sigma_x^2 \mathbb{E}\{W_{GSD}D_{\Delta\hat{H}}W_{GSD}^H\}$$
$$+ \sigma_z^2 (I \odot \hat{H}^H \hat{H})^{-1}, \quad (27)$$

The variance is the diagonal of $Cov(\mu_x)$, i.e.,

$$\mathrm{Var}(\mu_{x_i}) = \sigma_z^2 \sum_{j=0}^{KL-1} |\hat{H}[j,i]|^{-2} +$$
$$\sigma_x^2 \sum_{j=0}^{KL-1} |\hat{H}[j,i]|^{-2} \sum_{j'=0}^{KL-1} \sigma_{\hat{H}}^2[i,j'], \quad (28)$$

where $\sigma_{\hat{H}}^2[j,i']$ is the (j,i')th entry of $\Sigma_{\hat{H}}$.

APPENDIX C
GCHE

In t-th iteration, (14) can be written as,

$$y = (\Phi^{(t)} + \Phi - \Phi^{(t)})h + \tilde{z},$$
$$= \Phi^{(t)}h + \Delta\Phi^{(t)}h + \tilde{z}, \quad (29)$$
$$= \Phi^{(t)}h + \tilde{z}^{(t)},$$

where $\mathbb{E}\{\Delta\Phi^{(t)}\} = 0$. The mean of $\tilde{z}^{(t)}$ is,

$$\mathbb{E}\{\tilde{z}^{(t)}\} = \mathbb{E}\{\Delta\Phi^{(t)}h\} + \mathbb{E}\{\tilde{z}\},$$
$$= \Delta\Phi^{(t)}\mathbb{E}\{h\} + \mathbb{E}\{\tilde{z}\}, \quad (30)$$
$$= 0.$$

The covariance of $\tilde{\boldsymbol{z}}^{(t)}$ is,

$$
\begin{aligned}
\text{Cov}(\tilde{\boldsymbol{z}}^{(t)}) &= \boldsymbol{R}_{\tilde{\boldsymbol{z}}^{(t)}}, \\
&= \boldsymbol{R}_{\Delta\boldsymbol{\Phi}^{(t)}\boldsymbol{h}} + \boldsymbol{R}_{\tilde{\boldsymbol{z}}}, \\
&= \mathbb{E}\{\Delta\boldsymbol{\Phi}^{(t)}\boldsymbol{h}\boldsymbol{h}^H\Delta\boldsymbol{\Phi}^{(t)^H}\} + \sigma_z^2\boldsymbol{I}, \\
&= \frac{1}{P}\mathbb{E}\{\Delta\boldsymbol{\Phi}^{(t)}diag(\boldsymbol{h}_m)\Delta\boldsymbol{\Phi}^{(t)^H}\}, \\
&= \frac{1}{P}diag(\boldsymbol{h}_m)D_{\Delta\Phi^{(t)}}.
\end{aligned}
\tag{31}
$$

Therefore, the linear MMSE estimation is,

$$
\begin{aligned}
\hat{\boldsymbol{h}}^{(t)} = \boldsymbol{R_h}\boldsymbol{\Phi}^{(t)^H}(\boldsymbol{\Phi}^{(t)}\boldsymbol{R_h}\boldsymbol{\Phi}^{(t)^H} + \frac{1}{P}diag(\boldsymbol{h}_m)D_{\Delta\Phi^{(t)}} \\
+ \sigma_z^2\boldsymbol{I})^{-1}\boldsymbol{y}.
\end{aligned}
\tag{32}
$$

The covariance of the estimation error is,

$$
\begin{aligned}
\text{Var}(\boldsymbol{e}) &= \mathbb{E}\{(\boldsymbol{h}-\hat{h}^{(t)})(\boldsymbol{h}-\hat{h}^{(t)})^H\}, \\
&= \mathbb{E}\{\boldsymbol{h}\boldsymbol{h}^H\} - \mathbb{E}\{\hat{\boldsymbol{h}}\hat{\boldsymbol{h}}^H\}, \\
&= \boldsymbol{R_h} - \boldsymbol{R_h}\Phi^{(t)^H}\boldsymbol{R_y^{-1}}R_y R_y^{-1}\Phi^{(t)}\boldsymbol{R_h}, \\
&= \boldsymbol{R_h} - \boldsymbol{R_h}\Phi^{(t)^H}R_y^{-1}\Phi^{(t)}\boldsymbol{R_h}, \\
&= \boldsymbol{R_h} - \boldsymbol{R_h}\Phi^{(t)^H}(\boldsymbol{\Phi}^{(t)}R_h\boldsymbol{\Phi}^{(t)^H} + \\
&\quad \frac{1}{P}diag(\boldsymbol{h}_m)D_{\Delta\Phi^{(t)}} + \sigma_z^2\boldsymbol{I}))^{-1}\Phi^{(t)}\boldsymbol{R_h},
\end{aligned}
\tag{33}
$$

REFERENCES

[1] R. Hadani, S. Rakib, M. Tsatsanis, A. Monk, A. J. Goldsmith, A. F. Molisch, and R. Calderbank, "Orthogonal time frequency space modulation," in *2017 IEEE Wireless Communications and Networking Conference (WCNC)*, 2017, pp. 1–6.

[2] A. Fish, S. Gurevich, R. Hadani, A. M. Sayeed, and O. Schwartz, "Delay-doppler channel estimation in almost linear complexity," *IEEE Transactions on Information Theory*, vol. 59, no. 11, pp. 7632–7644, 2013.

[3] P. Raviteja, K. T. Phan, Q. Jin, Y. Hong, and E. Viterbo, "Low-complexity iterative detection for orthogonal time frequency space modulation," in *2018 IEEE Wireless Communications and Networking Conference (WCNC)*, 2018, pp. 1–6.

[4] P. Raviteja, Y. Hong, E. Viterbo, and E. Biglieri, "Practical pulse-shaping waveforms for reduced-cyclic-prefix otfs," *IEEE Transactions on Vehicular Technology*, vol. 68, no. 1, pp. 957–961, 2019.

[5] H. Wen, W. Yuan, C. Yuen, and Y. Li, "Mf-oamp-based joint channel estimation and data detection for otfs systems," *IEEE Transactions on Vehicular Technology*, vol. 73, no. 2, pp. 2948–2953, 2024.

[6] S. Marinkovic, B. Vucetic, N. Ishii, S. Yoshida, and T. Ushirokawa, "Space-time iterative and multistage receiver structures for cdma mobile communication systems," in *IEEE VTS 53rd Vehicular Technology Conference, Spring 2001. Proceedings (Cat. No.01CH37202)*, vol. 3, 2001, pp. 1799–1803 vol.3.

[7] W. Hardjawana, R. Li, B. Vucetic, Y. Li, and X. Yang, "A new iterative channel estimation for high mobility mimo-ofdm systems," in *2010 IEEE 71st Vehicular Technology Conference*, 2010, pp. 1–5.

[8] X. Wang, W. Shen, C. Xing, J. An, and L. Hanzo, "Joint bayesian channel estimation and data detection for otfs systems in leo satellite communications," *IEEE Transactions on Communications*, vol. 70, no. 7, pp. 4386–4399, 2022.