



Article A Structured Sparse Bayesian Channel Estimation Approach for Orthogonal Time–Frequency Space Modulation

Mi Zhang ^{1,†}, Xiaochen Xia ^{1,†}, Kui Xu ^{1,*}, Xiaoqin Yang ¹, Wei Xie ¹, Yunkun Li ² and Yang Liu ¹

¹ School of Communication Engineering, Army Engineering University of PLA, Nanjing 210007, China

² Unit 31105 of PLA, Nanjing 210042, China

* Correspondence: lgdxxukui@sina.com

+ These authors contributed equally to this work.

Abstract: Orthogonal time–frequency space (OTFS) modulation has been advocated as a promising waveform for achieving integrated sensing and communication (ISAC) due to its superiority in high-mobility adaptability and spectral efficiency. In OTFS modulation-based ISAC systems, accurate channel acquisition is critical for both communication reception and sensing parameter estimation. However, the existence of the fractional Doppler frequency shift spreads the effective channels of the OTFS signal significantly, making efficient channel acquisition very challenging. In this paper, we first derive the sparse structure of the channel in the delay Doppler (DD) domain according to the input and output relationship of OTFS signals. On this basis, a new structured Bayesian learning approach is proposed for accurate channel estimation, which includes a novel structured prior model for the delay-Doppler channel and a successive majorization–minimization (SMM) algorithm for efficient posterior channel estimate computation. Simulation results show that the proposed approach significantly outperforms the reference schemes, especially in the low signal-to-noise ratio (SNR) region.



Citation: Zhang, M.; Xia, X.; Xu, K.; Yang, X.; Xie, W.; Li, Y.; Liu, Y. A Structured Sparse Bayesian Channel Estimation Approach for Orthogonal Time–Frequency Space Modulation. *Entropy* 2023, *25*, 761. https:// doi.org/10.3390/e25050761

Academic Editors: Weijie Yuan, Zhiqiang Wei, Baoming Bai and Shuangyang Li

Received: 3 April 2023 Revised: 3 May 2023 Accepted: 4 May 2023 Published: 6 May 2023



Copyright: © 2023 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/). **Keywords:** orthogonal time–frequency space modulation; integrated sensing and communication; channel estimation; structured sparse Bayesian learning

1. Introduction

Future mobile communication systems must be better suited to serve a variety of developing applications in environments with high levels of mobility, including loworbit satellites, manned and unmanned aircraft, manned and unmanned vehicles, and high-speed trains [1–3]. In these high-mobility scenarios, mobile channels exhibit the characteristics of doubly dispersive channels due to the influence of the Doppler effect and the multipath propagation effect. Orthogonal frequency division multiplexing (OFDM) technology, which is currently widely used in fourth-generation mobile communication systems, fifth-generation mobile communication systems [4,5], wireless local area networks (WLANs), digital video broadcasting (DVB), and other broadband transmission systems, can overcome intersymbol interference (ISI) caused by time dispersion [6]. Of course, the studies on OFDM systems over doubly dispersive channel are also plentiful. Some researchers [7,8] have studied the channel estimation and data detection problems of OFDM systems over doubly dispersive channels, and the simulation results showed that the performance of the proposed method is close to the ideal case with perfect channel state information. However, due to the strict orthogonality among the subcarriers of OFDM systems, the orthogonality will be greatly destroyed by the channel frequency dispersion effect caused by frequency bias and fast time-varying channels. This results in serious intercarrier interference (ICI), which greatly degrades system performance and makes it impossible to provide efficient and reliable services. Therefore, OFDM is highly sensitive to frequency offset and channel frequency dispersion [9].

Orthogonal time–frequency space (OTFS) is a new modulation technology that can adapt to the time–frequency double dispersion propagation characteristics of channels in a high-speed mobile environment; it has attracted extensive attention in industry [10–12]. A time-varying channel in the time–frequency (TF) domain is changed into a time-invariant channel in the DD domain using OTFS modulation. After such a two-dimensional transformation, the transmitted data symbols are subjected to minor subcarrier interference, and all symbols in the same OTFS data frame experience slow fading independent of time selectivity. OTFS modulation is carried out in the DD domain.

At present, OTFS modulation has made certain research progress in many aspects, such as system performance analysis and modulation and demodulation algorithms. It is assumed in [10] that the ideal pulse-forming waveform satisfies orthogonal criteria in both time and frequency, but in reality, this is not possible because of Heisenberg's uncertainty principle. As a result, the biorthogonality assumption was loosened, and an OFDM-based OTFS modulation system was presented in [13,14]. This system takes into account the cyclic prefix (CP) of each OFDM symbol in the OTFS frame as well as the non-ideal rectangular pulse-shaping waveform. This benefit stems from the simplicity of implementation, but it also introduces the issue of high out-of-band radiation causing interference with adjacent channels. At the moment, frequency local pulse shaping [15] and time-domain windowing technology [16] are the two major techniques for maximizing out-of-band attenuation by lowering out-of-band radiation. Compared with the TF domain, the advantage of the DD domain is that only a small area, which is based on the channel response that will be introduced in the second part of the system model, is needed to describe the channel, and it has stronger sparsity for better channel estimation.

Currently, the main technical research on OTFS modulation involves the estimation of channel state information and data detection after estimation. Ref. [17] proposed an embedded pilot-assisted channel estimation scheme. The delay-Doppler plane's arrangement rules for data symbols, pilot symbols, and guard symbols were created to successfully prevent pilot symbols and data symbols from interfering with one another during each OTFS frame. The pilots were created for multipath channels with integer and fractional Doppler shifts using the OTFS modulation technology. However, if only one pilot symbol was inserted, it caused high peak-to-average-power ratio (PAPR) problems. For example, in this article, pilot power is greater than data power. In [18], a channel estimation and data detection framework based on a superposition pilot was proposed. This framework super-imposes a low-power pilot on data symbols in the DD domain, thus increasing the space of data transmission and effectively improving the spectral efficiency of OTFS modulation transmission. Such a design induces high computation complexity, and the interference between data and pilot symbol also reduces the estimated accuracy.

As emphasized in [19], integer approximation cannot adequately simulate the channel's noninteger delay and Doppler shift. Therefore, using an orthogonal matching pursuit approach based on binary refinement estimation, which has a tolerable computing complexity and much lower estimated normalized mean square error, has been suggested. According to the sparseness of the channel in the delay-Doppler domain, ref. [20] suggested a new pilot mode and a channel estimate method based on sparse Bayesian learning (SBL). In this new pilot pattern, the guard interval was not set, and the data and pilot had the same energy. To update the parameters in the previous model using the expectation maximum (EM) algorithm, a sparse Bayesian learning framework was provided. Channel estimation was treated as a sparse recovery problem. To a certain extent, the pilot symbol power cost was reduced, as was the pilot symbol power consumption, and the noise interference was weakened.

In this paper, to reduce the pilot cost and improve the accuracy of channel estimation, we treat the channel problem of estimating the DD domain as a sparse recovery problem and derive the sparse model expression from the input–output relationship of the DD domain. We provide a new sparse Bayesian framework based on the channel's sparse features, and the block successive majorization–minimization (SMM) technique is employed to handle

the problem of the new prior structure. The simulation results show that, in comparison to existing reference approaches, our suggested method greatly reduces the noise interference in the situation of a low signal-to-noise ratio.

The rest of the paper is organized as follows. The OTFS modulation system model is described in Section 2. The structured sparse Bayesian approach for channel estimation is presented in Section 3. The simulation results are reported in Section 4, and Section 5 concludes the paper.

2. System Model

We consider an OTFS modulation system with *M* symbols and *N* subcarriers. The transmitter and receiver are equipped with a single antenna. The OTFS modulation frame is shown in Figure 1.



Figure 1. OTFS modulation frame.

2.1. OTFS Modulation and Demodulation

The transmitter transforms the information symbol $x[k, l] \in \mathbb{C}^{M \times N}$ from the DD domain to the time–frequency domain using the inverse symplectic finite Fourier transform (ISFFT).

$$X[n,m] = \frac{1}{\sqrt{MN}} \sum_{k=0}^{N} \sum_{l=0}^{M} x[k,l] e^{j2\pi (\frac{nk}{N} - \frac{ml}{M})}$$
(1)

At this time, X[n, m] generates a time waveform s(t) using a transmission pulse $g_{tx}(t)$. This transform is called the Heisenberg transform.

$$s(t) = \sum_{n=0}^{N} \sum_{m=0}^{M} X[n,m] g_{tx}(t) e^{j2\pi m\Delta f(t-nT)}$$
(2)

where Δf and *T* are the subcarrier spacing and symbol period, respectively.

After that, the signal s(t) is transmitted through a fast time-varying channel with a complex baseband channel impulse response $h(\tau, \nu)$, which specifies the channel response to an impulse with delay τ and Doppler ν .

$$r(t) = \int \int h(\tau, \nu) s(t-\tau) e^{j2\pi\nu(t-\tau)}$$
(3)

The channel only has a few parameters in the delay-Doppler domain. The sparse representation of channel $h(\tau, \nu)$ is given as

$$h(\tau,\nu) = \sum_{i=1}^{p} h_i \delta(\tau - \tau_i) \delta(\nu - \nu_i)$$
(4)

where *P* denotes the number of paths for propagation. $\delta()$ denotes the Dirac delta function, whereas h_i , τ_i and ν_i stand for the path gain, delay, and Doppler shift associated with the *i*-th path, respectively. We choose the delay and Doppler taps for the *i*-th path as follows:

$$\tau_i = \frac{l_{\tau_i}}{M\Delta f}, \nu_i = \frac{k_{\nu_i} + \kappa_{\nu_i}}{NT}$$
(5)

where the integers l_{τ_i} , k_{ν_i} represent the delay tap and Doppler tap indices corresponding to the delay τ_i and Doppler frequency ν_i , respectively, and $\kappa_{\nu_i} \in [-0.5, 0.5]$. We do not need to examine the impact of the fractional delay because the resolution of the time axis is sufficient to approximate the path delay to the nearest integer grid point.

The received signal r(t) is given by the Wigner transform and is realized by the receiving matching filter of impulse response g_{rx} , obtaining the time–frequency domain symbol Y[n, m].

$$Y[n,m] = \int r(t)g_{rx}(t-nT)e^{j2\pi m\Delta f(t-nT)}dt$$
(6)

Finally, the symplectic finite Fourier transform (SFFT) is used to convert the information symbols Y[n, m] in the TF domain into the symbols y[k, l] in the DD domain .

When the transmitter pulse g_{tx} and the receiver pulse g_{rx} are ideal pulse functions, the relationship between the transmitted signal and the received signal in the delay-Doppler domain can be written as

$$y[k,l] = \sum_{i=1}^{P} \sum_{k'=[k-k_{\nu_{i}}+Q]_{N}}^{[k-k_{\nu_{i}}+Q]_{N}} \left(\frac{e^{-j2\pi \left(k-k_{\nu_{i}}-k'-\kappa_{\nu_{i}}\right)}-1}{Ne^{-j\frac{2\pi}{N}\left(k-k_{\nu_{i}}-k'-\kappa_{\nu_{i}}\right)}-N} \right) + v[k,l]$$

$$= \sum_{i=1}^{P} \sum_{q=-Q}^{Q} \left(\frac{e^{-j2\pi \left(-q-\kappa_{\nu_{i}}\right)}-1}{Ne^{-j\frac{2\pi}{N}\left(-q-\kappa_{\nu_{i}}\right)}-N} \right) h_{i}e^{-j2\pi\nu_{i}\tau_{i}}$$

$$\times x[[k-k_{\nu_{i}}+q]_{N},[l-l_{\tau_{i}}]_{M}] + v[k,l]$$

$$(7)$$

where v[k, l] denotes the system noise in the DD domain. We set the proper Q to counteract the impact of fractional Doppler. The delay-Doppler plane is discretized to an $M \times N$ grid. According to Equation (7), due to the influence of the doubly dispersive channel, the signal received at the *k*-th Doppler point is subjected to the superposition of signals from the 2Q + 1 (k - Q to k + Q) Doppler points. In other words, the received signal y[k, l] is a linear combination of the sent signal $\Omega = \sum_{i=1}^{p} 2Q + 1$.

2.2. Sparse Delay-Doppler Domain Channel Model

To implement channel estimation, we place $L = (2k_{max} + 2Q + 1)(l_{max} + 1)$ pilot symbols $x_p[k, l]$ in the DD domain. The greatest delay τ_{max} and Doppler ν_{max} related to the delay and Doppler taps are defined as k_{max} and l_{max} . In addition, we formulate Equation (7) in a different form, as follows:

$$y[k,l] = \sum_{k'=-k_{max}-Q}^{k_{max}+Q} \sum_{l'=0}^{l_{max}} h[k',l'] x[k-k',l-l']$$
(8)

After a series of formula calculations, we finally obtain the matrix form of the formula as

$$y = \Phi h + V \tag{9}$$

where

$$\boldsymbol{h}[l'] = \begin{bmatrix} h[-k_{max} - Q, l'] \\ \vdots \\ h[k_{max} + Q, l'] \end{bmatrix}$$
(10a)

where $l' \in [0, l_{max}]$

$$\Phi = \sum_{l'=0}^{l_{max}} x_{k,l-l'}^{T}$$

$$= \sum_{l'=0}^{l_{max}} \begin{bmatrix} x[k - (-k_{max} - Q), l - l'] \\ \vdots \\ x[k - (k_{max} + Q), l - l'] \end{bmatrix}$$
(11)
$$= \begin{bmatrix} X_{k',l-0}^{T} \\ \vdots \\ X_{k',l-l_{max}}^{T} \end{bmatrix} = \begin{bmatrix} X_{k',l'}^{T} \\ \vdots \\ X_{k',l'}^{T} \end{bmatrix}$$

where $\Phi \in \mathbb{C}^{S \times L}$, $h \in \mathbb{C}^{L \times 1}$, and $V \in \mathbb{C}^{S \times 1}$ are the received signal, transmitted signal. and Gaussian noise in the DD domain, respectively. *S* is the size of the channel estimation area. A simple proof is provided in Appendix A.

3. Structured Sparse Bayesian Approach for Channel Estimation

In this section, according to the sparse characteristics of the channel in the DD domain, we use Equation (8) to introduce the improved sparse Bayesian algorithm in detail. First, we briefly introduce the design of the pilot pattern at the receiver and transmitter. Second, we consider the sparse characteristics of the channel and design an improved sparse Bayesian algorithm. Finally, the complexity of the proposed algorithm is discussed.

3.1. Pilot Placement and Pattern Design

The power of the pilot symbol in the embedded pilot channel estimate approach is frequently substantially larger than that of the data symbol, which is unrealizable in actual applications. Only inserting one pilot symbol will result in a high PARP issue [17]. Meanwhile, unlike [20], we have set up a protection interval symbol to avoid interference between data and pilot symbols. Here, we set the pilot symbol power to be the same as the data symbol power, transforming the channel estimation problem into a sparse channel recovery problem. This not only effectively solves the problem of high pilots power but also avoids interference between data that affects estimation results.

We arrange the pilot and data symbols in the delay-Doppler grid for OTFS frame transmission, as shown in Figure 2.



Figure 2. Tx symbols pattern.

We select $k \in [k_p - k_{\max}, k_p - k_{\max}]$ and $l \in [l_p, l_p + l_{\tau}]$ as the pilot symbols. The other fields are guard symbols and data symbols.

$$x[k,l] = \begin{cases} x_{p}[k,l] & k_{p} - k_{\max} \le k \le k_{p} - k_{\max} \\ l_{p} \le l \le l_{p} + l_{\tau} \\ 0 & k_{p} - 3k_{\max} \le k \le k_{p} - 3k_{\max} \\ l_{p} - l_{\max} \le l \le l_{p} + l_{\tau} + l_{\max} \\ x_{d}[k,l] & otherwise \end{cases}$$
(12)

Usually, (k_p, l_p) are placed onto $(\frac{N}{2}, \frac{M}{2})$. l_{τ} is used to control the proportion of pilots in the whole transmission symbol. The final channel estimation region is set to $k \in [k_p - k_{max} - Q, k_p - k_{max} + Q]$ and $l \in [l_p, l_p + 2l_{\tau}]$. In other words, the size of the channel estimation section $S = (4k_{max} + 2Q + 1)(l_{max} + l_{\tau} + 1)$, as shown in Figure 3.



Figure 3. Rx symbols pattern.

We present the channel coefficient characteristics in the DD domain in Figure 4. According to the characteristics of channel sparsity, we propose a new estimation algorithm.



Figure 4. Characteristics of sparsity.

We provide an example of the DD domain channel corresponding to (7). As shown in Figure 4, the DD domain channel only has 4 non-zero responses over the whole DD channel, whose coordinates align with the associated delay and Doppler shifts. Therefor, estimating channel **h** can be viewed as a sparse recovery problem. Although the traditional sparse Bayesian algorithm can be used here, there are only values on each path according to the channel sparsity in the DD domain, and when a value is detected, it means that there are values before and after the value, so the traditional algorithm does not reflect this feature.

According to the sparse characteristics of the channel, we propose an improved sparse Bayesian algorithm, as follows:

$$p(\mathbf{h}|\alpha_{l,k},\gamma) = \mathcal{N}c(\mathbf{h}|0, diag(\gamma) \times diag(\mathbf{D}_l))$$
(13)

where $\gamma = {\gamma_1, \gamma_2, ..., \gamma_{l_{max}+1}}$, D_l is a diagonal matrix whose n-th diagonal element is given by

$$[D_{l}]_{k,k} = \begin{cases} \alpha_{l,1} + \alpha_{l,2} & k = 1\\ \alpha_{l,k-1} + \alpha_{l,k} + \alpha_{l,k+1} & 2 \le k \le 2k_{\max} + 2Q\\ \alpha_{l,k-1} + \alpha_{l,k+1} & k = 2k_{\max} + 2Q + 1 \end{cases}$$
(14)

where $\alpha_l = \{ \alpha_{l,1}, \alpha_{l,2}, \dots, \alpha_{l,2k_{\max}+2Q+1} \}, D_l = \{ D_1, D_2, \dots, D_{l_{\max}+1} \}].$

As above, in the design of the diagonal matrix D_l , the design of the parameter $\alpha_{l,k}$ not only affects the *k*th element of channel **h** to be estimated; it also affects its neighbouring elements. As $\alpha_{l,k}$ tends to infinity, the *k*-th, (k + 1)th, and (k - 1)th elements are simultaneously driven to zero. The distribution of zero elements exhibits a blocky distribution. On the other hand, by dividing the whole h into l_{max} column vectors, the function of γ is used to distinguish the difference between the different column vectors.

According to Equation (9), the posterior distribution of **h** can be expressed as

$$p(\mathbf{h}|\boldsymbol{y},\boldsymbol{\alpha}_{l,k},\boldsymbol{\gamma}) \propto p(\boldsymbol{y}|\mathbf{h})p(\mathbf{h}|\boldsymbol{\alpha}_{l,k},\boldsymbol{\gamma})$$
(15)

where the likelihood function of *y* is $p(y|\mathbf{h})$, which is

$$p(\boldsymbol{y}|\boldsymbol{h}) = \mathcal{N}c\left(\boldsymbol{y}|\boldsymbol{\Phi}\boldsymbol{h},\sigma^{2}\boldsymbol{I}\right)$$
(16)

By substituting Equations (13) and (16) into Formula (15), it is not difficult to prove that this is a posterior distribution that follows a complex Gaussian distribution through mathematical calculation, and the mean and covariance are as follows:

$$\mu = \frac{1}{\sigma^2} \sum \mathbf{\Phi}^H \mathbf{y}$$

$$\sum = (diag\{\gamma\} \times diag\{\mathbf{D}_l\} + \frac{1}{\sigma^2} \mathbf{\Phi}^H \mathbf{\Phi})^{-1}$$
(17)

Here, as in the traditional sparse Bayesian algorithm, we express the Gamma hyperprior with respect to α as follows:

$$p(\boldsymbol{\alpha}) = \prod_{k=1}^{2k_{\max}+2Q+1} Gam(\alpha_{l,k}|a,b)$$

$$= \prod_{k=1}^{2k_{\max}+2Q+1} \frac{b^a}{\Gamma(a)} \alpha_{l,k}{}^a \exp(-b\alpha_{l,k})$$
(18)

where *a* and *b* are fixed parameters. In this section, we introduce the small positive numbers a and b $(a = b = 10^{-4})$. Such a configuration encourages a flat prior for α .

We also select a Dirichlet hyperprior for γ ,

$$p(\boldsymbol{\gamma}) = C(\boldsymbol{u}) \prod_{l=1}^{l_{\max}+1} \boldsymbol{\gamma}^{\boldsymbol{u}_l}$$
(19)

where $\boldsymbol{u} = [u_1, u_2, \dots, u_{l_{max}}]$ is a fixed parameter. $C(\boldsymbol{u}) = \frac{\Gamma(\sum_{l=1}^{l_{max}} u_l)}{\Gamma(u_1)\cdots\Gamma(u_l)}$ is the normalization constant. The role of γ is to model the relative difference between different sub-vectors \boldsymbol{u} . By exploiting the common sparsity structure, the suboptimal solutions in the conventional methods that predict a value of zero for one element of $\mathbf{h}_{[l]}$ and predict a nonzero value for the element of $\mathbf{h}_{[l' \neq l]}$ in the same position can be eliminated. Since the expectation of γ with respect to the distribution (19) is given by $\mathbb{E}[\gamma] = u_l / \sum_{l'=1}^{l_{max}} u_{l'}$, we can interpret the parameter \boldsymbol{u} , which provides a initial guess as to the relative difference between the submodules of \mathbf{h} can be interpreted as \boldsymbol{u} . Note that a default setup for \boldsymbol{u} can be $u_l = 1$ for $l = 1, \dots, l_{max}$, which indicates an initial assumption of different delay modules.

The maximum posterior (MAP) estimation of **h** can be derived by its posterior mean in Equation (17) as long as $\alpha_{l,k}$. As a result, in the sections that follow, we concentrate on obtaining the ideal $\alpha_{l,k}$ and γ by solving the MAP issue.

$$\left\{ \alpha_{l,k}^{opt}, \boldsymbol{\gamma}^{opt} \right\} = \arg \max_{\boldsymbol{\alpha}_{l,k}, \boldsymbol{\gamma}} p(\boldsymbol{\alpha}_{l,k}, \boldsymbol{\gamma} | \boldsymbol{y})$$

= $\arg \min_{\boldsymbol{\alpha}_{l,k}, \boldsymbol{\gamma}} - \log p(\boldsymbol{\alpha}_{l,k}, \boldsymbol{\gamma}, \boldsymbol{y})$ (20)

Since we proposed a new structured prior model for **h**, it is clear that the above problems cannot be solved if we continue to use the traditional sparse Bayesian algorithm. Therefore, we provide a solution using the block SMM algorithm.

SMM Algorithm for Posterior Channel Estimate Evaluation

The proposed block SMM algorithm is similar to the expected EM algorithm. It has two main steps. The first step is to replace the optimization step with the expected step, and its purpose is to find the alternative function of the objective function, that is, the local approximate solution. The second step is the minimization step, in which the alternative function found in the first step is minimized. While the block SMM technique divides the parameters to be estimated into many blocks for alternate optimization, the EM algorithm updates all parameters concurrently. This is helpful when solving structured SBL is the only obstacle to successfully minimizing all of the proxy function's parameters.

At this point, we assume that $\mathcal{G}(\alpha_{l,k}, \gamma, y | \hat{\alpha}_{l,k}, \hat{\gamma})$ is the alternative function in Step 1, where $(\hat{\alpha}_{l,k}, \hat{\gamma})$ is the estimation of $(\alpha_{l,k}, \gamma)$. If $\mathcal{G}(\alpha_{l,k}, \gamma, y | \hat{\alpha}_{l,k}, \hat{\gamma})$ satisfies the following condition:

$$\begin{aligned}
\mathcal{G}(\alpha_{l,k},\gamma,\boldsymbol{y}|\hat{\alpha}_{k},\hat{\gamma}) &\geq -\log p(\alpha_{l,k},\gamma,\boldsymbol{y}), \forall \alpha_{l,k},\gamma \\
\mathcal{G}(\hat{\alpha}_{l,k},\hat{\gamma},\boldsymbol{y}|\hat{\alpha}_{l,k},\hat{\gamma}) &= -\log p(\hat{\alpha}_{l,k},\hat{\gamma},\boldsymbol{y}) \\
\frac{\partial \mathcal{G}(\alpha_{l,k},\hat{\gamma},\boldsymbol{y}|\hat{\alpha}_{l,k},\hat{\gamma})}{\partial \alpha_{l,k}}\Big|_{\alpha_{l,k}=\hat{\alpha}_{l,k}} &= \frac{\partial \{-\log p(\alpha_{l,k},\hat{\gamma},\boldsymbol{y})\}}{\partial \alpha_{l,k}}\Big|_{\alpha_{l,k}=\hat{\alpha}_{l,k}} \\
\frac{\partial \mathcal{G}(\hat{\alpha}_{l,k},\gamma,\boldsymbol{y}|\hat{\alpha}_{l,k},\hat{\gamma})}{\partial \gamma}\Big|_{\gamma=\hat{\gamma}} &= \frac{\partial \{-\log p(\hat{\alpha}_{l,k},\gamma,\boldsymbol{y})\}}{\partial \gamma}\Big|_{\gamma=\hat{\gamma}}
\end{aligned}$$
(21)

In addition, $\{\alpha_{l,k}, \gamma\}$ follows the following rules in each iteration:

$$\hat{\alpha}_{l,k}^{(i+1)} = \arg\max_{\alpha_{l,k}} \mathcal{G}(\alpha_{l,k}, \hat{\gamma}^{(i)}, \boldsymbol{y} | \hat{\alpha}_{l,k}^{(i)}, \hat{\gamma}^{(i)})
\hat{\gamma}^{(i+1)} = \arg\max_{\alpha_{l,k}} \mathcal{G}(\hat{\alpha}_{l,k}^{(i+1)}, \boldsymbol{\gamma}, \boldsymbol{y} | \hat{\alpha}_{l,k}^{(i)}, \hat{\gamma}^{(i)})$$
(22)

where $\{\alpha_{l,k}^{(i)}, \hat{\gamma}^{(i)}\}$ represents $\alpha_{l,k}$, *the* γ estimated value in the *i*-th iteration.

After a series of mathematical calculations, the specific algorithm is as follows.

(1) Step 1: A substitute for the objective function $-\log p(\alpha_{l,k}.\gamma,y)$ is required at the corresponding step of the (i + 1)th iteration. In this study, we make use of the surrogate function proposed by [21], which meets the requirements stated in (21), that is, a simple proof is provided in Appendix B.

$$\mathcal{G}(\alpha_{l,k}, \boldsymbol{\gamma}, \boldsymbol{y} | \hat{\alpha}_{l,k}^{(l)}, \hat{\boldsymbol{\gamma}}^{(i)}) = -\int p(\mathbf{h} | \boldsymbol{y}, \hat{\alpha}_{l,k}^{(i)}, \hat{\boldsymbol{\gamma}}^{(i)}) \log \frac{p(\alpha_{l,k}, \boldsymbol{\gamma}, \mathbf{h}, \boldsymbol{y})}{p(\mathbf{h} | \boldsymbol{y}, \hat{\alpha}_{l,k}^{(i)}, \hat{\boldsymbol{\gamma}}^{(i)})} d\mathbf{h}$$

$$\geq -\log p(\alpha_{l,k}, \boldsymbol{\gamma}, \boldsymbol{y})$$
(23)

where the final inequation obeys Jensen's inequality, and when the inequation satisfies $(\alpha_{l,k}, \gamma) = (\alpha_{l,k}^{(i)}, \hat{\gamma}^{(i)})$, the equal sign is true. By substituting (15) into (23), we can obtain

$$\mathcal{G}(\alpha_{l,k}, \gamma, \boldsymbol{y} | \hat{\alpha}_{l,k}^{(i)}, \hat{\gamma}^{(i)}) = -\log p(\alpha_{l,k}) - \log p(\gamma) - \mathbb{E}_{[\log p(\mathbf{h} | \alpha_{l,k}, \gamma)]} + const$$
(24)

where the expectation is in relation to $p(\mathbf{h}|y, \hat{\alpha}_{l,k}^{(i)}, \hat{\gamma}^{(i)})$. *const* denotes that it is not related to $\alpha_{l,k}$ or γ .

(2) Step 2: To minimize the surrogate function, the parameters $\alpha_{l,k}$ are changed sequentially in the corresponding step of the (i + 1)th iteration.

To update $\alpha_{l,k}$; γ is regarded as a constant when $\alpha_{l,k}$ is estimated. By substituting (13) and (18) into (24), we obtain

$$\begin{aligned} \mathcal{G}(\alpha_{l,k}, \hat{\boldsymbol{\gamma}}^{(i)}, \boldsymbol{y} | \hat{\alpha}_{l,k}^{(i)}, \hat{\boldsymbol{\gamma}}^{(i)}) \\ &= -\sum_{k=1}^{2k_{\max}+2Q+1} (a \log \alpha_{l,k} - b \alpha_{l,k}) \\ &- \log |diag(\boldsymbol{\gamma}) \times diag(\boldsymbol{D}_{l})| \\ &+ tr((diag(\boldsymbol{\gamma}) \times diag(\boldsymbol{D}_{l})) \times (\hat{\boldsymbol{\Sigma}} + \hat{\boldsymbol{\mu}} \hat{\boldsymbol{\mu}}^{H})) \\ &+ const \end{aligned} \tag{25}$$

where Equation (17) is used to compute $\hat{\Sigma}$ and $\hat{\mu}$, and $(\alpha_{l,k}, \gamma)$ is the result obtained from the *i*-th iteration. Next, we consider the derivative of (25) with regard to $\alpha_{l,k}$ to analyse the optimality condition.

$$\frac{\partial \mathcal{G}(\alpha_{l,k}, \hat{\boldsymbol{\gamma}}^{(i)}, \boldsymbol{y} | \hat{\boldsymbol{\alpha}}_{l,k}^{(i)}, \hat{\boldsymbol{\gamma}}^{(i)})}{\partial \alpha_{l,k}} = \frac{a}{\alpha_{l,k}} - b - \hat{\boldsymbol{\gamma}}^{(i)} \boldsymbol{g} + (\xi_{k-1} + \xi_k + \xi_{k+1})$$
(26)

where

$$g = \sum_{j=1}^{L} \left\{ \begin{array}{l} \left[\tilde{\Sigma}\right]_{[j,j]} + \left|\left[\tilde{\mu}\right]_{j}\right|^{2} \\ + \left[\tilde{\Sigma}\right]_{[[j+MAX]_{L'}[j+MAX]_{L}]} + \left|\left[\tilde{\mu}\right]_{[j+MAX]_{L}}\right|^{2} \\ + \left[\tilde{\Sigma}\right]_{[[l_{\max}\cdot MAX+j]_{L'}[l_{\max}\cdot MAX+j]_{L}]} \\ + \left|\left[\tilde{\mu}\right]_{[l_{\max}\cdot MAX+j]_{L}}\right|^{2} \end{array} \right\}$$

$$(27)$$

$$\xi_k = \frac{1}{\alpha_{l,k-1} + \alpha_{l,k} + \alpha_{l,k+1}}$$
(28)

Specifically, when k = 1

$$\xi_k = \frac{1}{\alpha_1 + \alpha_2} \tag{28a}$$

when k = MAX

$$\xi_k = \frac{1}{\alpha_{MAX-1} + \alpha_{MAX}} \tag{28b}$$

where $MAX = 2k_{max} + 2Q + 1$.

Because $\alpha_{l,k}$ is coupled with the adjacent terms $\alpha_{l,k-1}$ and $\alpha_{l,k+1}$ in the term ξ_k , we are unable to acquire $\alpha_{l,k}$ by setting (26) to zero. Here, we have $\alpha_{l,k}$, $\alpha_{l,k+1}$, $\alpha_{l,k-1}$. The expression for Equation (26) is

$$\frac{\partial \mathcal{G}(\alpha_{l,k}, \hat{\gamma}^{(i)}, \boldsymbol{y} | \hat{\alpha}_{l,k}^{(i)}, \hat{\gamma}^{(i)})}{\partial \alpha_{l,k}} = \frac{a}{\alpha_{l,k}} - b + \frac{1}{\alpha_{l,k}} - \hat{\gamma}^{(i)} \boldsymbol{g}$$
(29)

By setting (29) to zero, we can obtain

$$\hat{x}_{l,k}^{(i+1)} = \frac{a+1}{\hat{\gamma}^{(i)}g + b}$$
(30)

Here, we do not address special cases (when k = 1 and k = MAX).

To update γ , by leaving out the terms that are unrelated to γ and inserting (13) and (19) into (24), we obtain

$$\mathcal{G}(\hat{a}_{l,k}^{(i+1)}, \boldsymbol{\gamma}, \boldsymbol{y} | \hat{a}_{l,k}^{(i)}, \hat{\boldsymbol{\gamma}}^{(i)}) = -\sum_{l=0}^{l_{\max}} \boldsymbol{u} \log \boldsymbol{\gamma} - \log |diag(\boldsymbol{\gamma}) \times diag(\boldsymbol{D}_{l})| + tr((diag(\boldsymbol{\gamma}) \times diag(\boldsymbol{D}_{l})) \times (\hat{\boldsymbol{\Sigma}} + \hat{\boldsymbol{\mu}} \hat{\boldsymbol{\mu}}^{H})) + const$$
(31)

where $\alpha_{l,k}$ is fixed at its latest estimation $\alpha_{l,k} = \alpha_{l,k}^{(i+1)}$. In relation to γ , the derivative of (31) can be written as

$$\frac{\partial \mathcal{G}(\hat{\boldsymbol{x}}_{l,k}^{(i+1)}, \boldsymbol{\gamma}, \boldsymbol{y} | \hat{\boldsymbol{x}}_{l,k}^{(i)}, \hat{\boldsymbol{\gamma}}^{(i)})}{\partial \boldsymbol{\alpha}_{l,k}} = \frac{l_{\max} + \boldsymbol{u}}{\boldsymbol{\gamma}} - \left(\hat{\boldsymbol{x}}_{l,k}^{(i+1)} + \hat{\boldsymbol{x}}_{l,k-1}^{(i+1)} + \hat{\boldsymbol{x}}_{l,k+1}^{(i+1)} \right) \times \left(\sum_{j=1}^{L} \left[\hat{\boldsymbol{\Sigma}} \right]_{[j,j]} + \left| [\hat{\boldsymbol{\mu}}]_{j} \right|^{2} \right)$$
(32)

The best value of γ can be attained by setting (32) to zero as

$$\hat{\gamma}^{(i+1)} = \frac{l_{\max} + u}{\hat{\alpha}_{l,k}^{(i+1)} + \hat{\alpha}_{l,k-1}^{(i+1)} + \hat{\alpha}_{l,k+1}^{(i+1)}} \times \left(\sum_{j=1}^{L} \left[\hat{\Sigma}\right]_{[j,j]} + \left| [\hat{\mu}]_{j} \right|^{2} \right)^{-1}$$
(33)

Step 1 and Step 2 are performed iteratively until convergence. Algorithm 1 is a summary of the SMM algorithm presented in this paper.

Algorithm 1 SMM algorithm for structured SBL

Input: threshold *ε*, maximum iterations *iter*_{max}

Setting the initial value α_k and γ as the first loop value.

- **Output:** Channel estimation **h**.
- 1: Step 1: the *i*-th iteration $\alpha_{l,k}^{(i)}$, $\hat{\gamma}^{(i)}$ as fixed parameters $\alpha_{l,k}$, γ . Update the posterior mean and covariance matrix of **h** using (16).
- 2: Step 2: Update $\alpha_{l,k}$ by (29). Update γ by (32).
- 3: Convergence of judgment: Let h^{i+1} denote the estimation of h after the (i+1)th iteration.

4: if
$$\|\mathbf{h}^{i+1} - \mathbf{h}^{i}\|^{2} \leq \varepsilon$$
 or $i + 1 = iter_{max}$ then
5: Output $\hat{\mathbf{h}}$
6: else
7: Set $i = i + 1$ and runtun Step 1
8: end if

Afterwards, the algorithm will be further studied and introduced into the MIMO system. The dimensions of the channel change from the SISO system to the MIMO system, and the channel dimension increases linearly with the increase of the number of antennas. The sparse structure of the channel will therefore expand from one-dimensional to multi-dimensional.

The calculation of the posterior μ and covariance matrix Σ of **h** in the MA-step in Algorithm 1 accounts for the majority of the computing effort for each iteration *iter*_{max}. Therefore, the maximum computational complexity order of the structured SBL-based channel estimation algorithm is $O(iter_{max}L^3S^3)$.

4. Simulation Results

In this section, we verify the performance of our proposed algorithm by the normalized mean square error (NMSE). The NMSE was defined as $\|\mathbf{h} - \hat{\mathbf{h}}\|_2^2 / \|\mathbf{h}\|_2^2$. We set the parameters as follows (see Table 1): the number of symbols N = 64, the number of subcarriers M = 128, the carrier frequency was 15 kHz, and the subcarrier frequency was 4×10^9 Hz. The maximum Doppler shift varied according to the speed, and the maximum delay value was $l_{\tau} = 10$. k_{max} was a parameter related to speed.

Table 1. Simulation parameters.

Parameters	Value
Symbols N	64
Subcarries M	128
Carrier frequency Δf	15 kHz
Subcarrier frequency	$4 imes 10^9~{ m Hz}$
The maximum delay l_{max}	10

First, we compared the performance of different algorithms under different signalto-noise ratios (SNR). Here, we used the reference algorithm for the least square (LS) algorithm, the orthogonal matching pursuit (OMP) algorithm, and the conventional SBL algorithm [22]. The user speed was 250 km/h, and $l_{\tau} = 10$. As shown in Figure 5, in the case of a low signal-to-noise ratio, our proposed structured SBL algorithms have a significant advantage over other algorithms. Specifically, our algorithm better reduces interference.



Figure 5. Normalized MSE versus SNR.

Then, we compared the cases where l_{τ} is constantly changing in different delay regions, as shown in Figure 6. In this OTFS modulation system, we set the SNR to 20 dB, the user speed to 250 km/h, and Q = 4. The figure shows that the structured SBL algorithm we proposed maintains certain advantages. As the delay range l_{τ} continues to expand, the

estimation results become increasingly accurate. We can see that under the same estimation error, compared to other algorithms, the structured SBL has the smallest value in delay dimension, which means that the pilot proportion is the smallest. For example, under the condition of a mean square error of 10^{-2} , the values of the delay dimension l_{τ} of the structured SBL and conventional SBL algorithm are 7 and 8, respectively. The pilot overhead (except guard symbols) is $\eta = \frac{(2k_{\max}+1)(l_{\tau}+1)}{M \times N}$. The pilot overheads of the structured SBL are about 1.07% and 1.21%, respectively.



Figure 6. Delay range l_{τ} versus normalized MSE.

Next, we set the different speeds for performance comparison in Figure 7. Here, we set the speed variation range to [150, 350]; the SNR was 20 dB, Q = 4, and $l_{\tau} = 5$. The figure shows that the channel estimation results are affected to some extent with increasing speed. The higher the speed is, the slightly worse the accuracy of the estimation becomes. However, in terms of the trend of the curve, the curve is smooth overall with little change, and the LS algorithm is highly sensitive to changes in speed. On the other hand, the structured SBL algorithm maintains high robustness.



Figure 7. Velocity versus normalized MSE.

5. Conclusions

In this paper, we derived a sparse structure model formula from the input–output relationship of the DD domain. Based on the characteristics of the channel's sparse structure, we treated channel estimation as a sparse recovery problem. In addition, we used block and continuous alternating optimization to solve the parameter problem of block distribution, which further improved the accuracy of channel estimation. Finally, through simulation analysis, we verified that the proposed algorithm has certain advantages over the traditional SBL algorithm. Under the condition of a low signal-to-noise ratio, the estimation accuracy was higher, and the influence of interference was diminished.

Author Contributions: Conceptualization, M.Z. and X.X.; methodology, M.Z.; writing—original draft preparation, M.Z.; writing—review and editing, X.X., K.X., W.X., X.Y., Y.L. (Yunkun Li) and Y.L. (Yang Liu). All authors have read and agreed to the published version of the manuscript.

Funding: This work was supported in part by the National Natural Science Foundation of China under Grant 62271503, Grant 62071485, and Grant 62001513, in part by the Basic Research Project of Jiangsu Province under Grant BK 20192002 and the Natural Science Foundation of Jiangsu Province under Grant BK 20201334, and BK 20200579.

Data Availability Statement: Not applicable.

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

Appendix A

In an OTFS system, the input-output relation of DD domain is denoted as

$$y[k,l] = \underbrace{\sum_{i=1}^{P} \sum_{q=-Q}^{Q} \left(\frac{e^{-j2\pi}(-q-\kappa_{\nu_{i}}) - 1}{Ne^{-j\frac{2\pi}{N}}(-q-\kappa_{\nu_{i}}) - N} \right) h_{i}e^{-j2\pi\nu_{i}\tau_{i}} \times x[[k-k_{\nu_{i}}+q]_{N'}[l-l_{\tau_{i}}]_{M}] + v[k,l]}_{Real \ channel}$$
(A1)

Another equation form is

$$y[k,l] = \sum_{k'=-k_{max}-Q}^{k_{max}+Q} \sum_{l'=0}^{l_{max}} h[k',l']x[[k-k']_N,[l-l']_M]$$
(A2)

When k' is fixed as constant, we have

$$y[k,l] = \sum_{l'=0}^{l_{\text{max}}} \mathbf{x}_{k,l-l'}^T \mathbf{h}[l']$$
(A3)

where

$$\mathbf{x}_{k,l-l'} = \begin{bmatrix} x[k - (-k_{max} - Q), l - l'] \\ \vdots \\ x[k - (k_{max} + Q), l - l'] \end{bmatrix}$$
(A4)

$$\mathbf{h}[l'] = \begin{bmatrix} h[-k_{max} - Q, l'] \\ \vdots \\ h[k_{max} + Q, l'] \end{bmatrix}$$
(A5)

Finally,

$$\mathbf{y} = \begin{bmatrix} y[k^-, l^-] \\ \vdots \\ y[k^+, l^+] \end{bmatrix} = \begin{bmatrix} \mathbf{x}_{k^-, l^-} \\ \vdots \\ \mathbf{x}_{k^+, l^+}^T \end{bmatrix} \mathbf{h} = \Phi \mathbf{h} + \mathbf{w}$$
(A6)

where $[k^-, l^-]$ is the part of the estimation region.

Appendix **B**

$$\begin{aligned} \mathcal{G}(\alpha_{k},\gamma_{l},y|\hat{\alpha}_{k}^{(i)},\hat{\gamma}_{l}^{(i)}) &= -\int p(h|y,\hat{\alpha}_{k}^{(i)},\hat{\gamma}_{l}^{(i)})\log\frac{p(\alpha_{k},\gamma_{l},h,y)}{p(h|y,\hat{\alpha}_{k}^{(i)},\hat{\gamma}_{l}^{(i)})}dh \\ &= {}_{h|y,\hat{\alpha}_{k}^{(i)},\hat{\gamma}_{l}^{(i)}}[-\log\frac{p(\alpha_{k},\gamma_{l},h,y)}{p(h|y,\hat{\alpha}_{k}^{(i)},\hat{\gamma}_{l}^{(i)})}] \\ &\geq -\log p(\alpha_{k},\gamma_{l},y) \\ &= -\log\left(\int p(h|y,\hat{\alpha}_{k}^{(i)},\hat{\gamma}_{l}^{(i)})\log\frac{p(\alpha_{k},\gamma_{l},h,y)}{p(h|y,\hat{\alpha}_{k}^{(i)},\hat{\gamma}_{l}^{(i)})}\right) \\ &= -\log p(\alpha_{k},\gamma_{l},y) \end{aligned}$$
(A7)

References

- Joo, C.; Choi, J. Low-delay broadband satellite communications with high-altitude unmanned aerial vehicles. *J. Commun. Netw.* 2018, 20, 102–108. [CrossRef]
- Zhang, X.; Niu, Y.; Mao, S.; Cai, Y.; He, R.; Ai, B.; Zhong, Z.; Liu, Y. Resource Allocation for Millimeter-Wave Train-Ground Communications in High-Speed Railway Scenarios. *IEEE Trans. Veh. Technol.* 2021, 70, 4823–4838. [CrossRef]
- Hui, L.Z.; Yang, G.Q. Uplink User Power Control for Low-Orbit Satellite Communication Systems. In Proceedings of the 2018 IEEE 4th International Conference on Computer and Communications (ICCC), Chengdu, China, 7–10 December 2018; pp. 634–639.
- 4. Yli-Kaakinen, J.; Loulou, A.; Levanen, T.; Pajukoski, K.; Palin, A.; Renfors, M.; Valkama, M. Frequency-Domain Signal Processing for Spectrally-Enhanced CP-OFDM Waveforms in 5G New Radio. *IEEE Trans. Wirel. Commun.* 2021, 20, 6867–6883. [CrossRef]
- Dong, Y.X.; Jin, W.; Giddings, R.P.; O'Sullivan, M.; Tipper, A.; Durrant, T.; Tang, J.M. Hybrid DFT-Spread OFDM-Digital Filter Multiple Access PONs for Converged 5G Networks. J. Opt. Commun. Netw. 2019, 11, 347–353. [CrossRef]
- Wang, S.; Thompson, J.S.; Grant, P.M. Closed-Form Expressions for ICI/ISI in Filtered OFDM Systems for Asynchronous 5G Uplink. *IEEE Trans. Commun.* 2017, 65, 4886–4898. [CrossRef]
- He, L.; Ma, S.; Wu, Y.-C.; Zhou, Y.; Ng, T.-S.; Poor, H.V. Pilot-Aided IQ Imbalance Compensation for OFDM Systems Operating Over Doubly Selective Channels. *IEEE Trans. Signal Process.* 2011, 59, 2223–2233. [CrossRef]
- 8. He, L.; Wu, Y.-C.; Ma, S.; Ng, T.-S.; Poor, H.V. Superimposed Training-Based Channel Estimation and Data Detection for OFDM Amplify-and-Forward Cooperative Systems Under High Mobility. *IEEE Trans. Signal Process.* **2012**, *60*, 274–284. [CrossRef]
- 9. Wang, T.; Proakis, J.G.; Masry, E.; Zeidler, J.R. Performance degradation of OFDM systems due to Doppler spreading. *IEEE Trans. Wirel. Commun.* **2006**, *5*, 1422–1432. [CrossRef]
- Hadani, R.; Rakib, S.; Tsatsanis, M.; Monk, A.; Goldsmith, A.J.; Molisch, A.F.; Calderbank, R. Orthogonal Time Frequency Space Modulation. In Proceedings of the 2017 IEEE Wireless Communications and Networking Conference (WCNC), San Francisco, CA, USA, 19–22 March 2017; pp. 1–6.
- 11. Raviteja, P.; Phan, K.T.; Hong, Y.; Viterbo, E. Interference Cancellation and Iterative Detection for Orthogonal Time Frequency Space Modulation. *IEEE Trans. Wirel. Commun.* **2018**, *17*, 6501–6515. [CrossRef]
- 12. Wei, Z.; Yuan, W.; Li, S.; Yuan, J.; Bharatula, G.; Hadani, R.; Hanzo, L. Orthogonal Time-Frequency Space Modulation: A Promising Next-Generation Waveform. *IEEE Wirel. Commun.* **2021**, *28*, 136–144. [CrossRef]
- 13. Farhang, A.; RezazadehReyhani, A.; Doyle, L.E.; Farhang-Boroujeny, B. Low Complexity Modem Structure for OFDM-Based Orthogonal Time Frequency Space Modulation. *IEEE Wirel. Commun. Lett.* **2018**, *7*, 344–347. [CrossRef]
- RezazadehReyhani, A.; Farhang, A.; Ji, M.; Chen, R.R.; Farhang-Boroujeny, B. Analysis of Discrete-Time MIMO OFDM-Based Orthogonal Time Frequency Space Modulation. In Proceedings of the 2018 IEEE International Conference on Communications (ICC), Kansas City, MO, USA, 20–24 May 2018; pp. 1–6.
- 15. Tiwari, S.; Das, S.S. Circularly pulse-shaped orthogonal time frequency space modulation. *Electron. Lett.* **2020**, *56*, 157–160. [CrossRef]
- Hossain, M.N.; Sugiura, Y.; Shimamura, T.; Ryu, H.-G. Waveform Design of Low Complexity WR-OTFS System for the OOB Power Reduction. In Proceedings of the 2020 IEEE Wireless Communications and Networking Conference Workshops (WCNCW), Seoul, Republic of Korea, 6–9 April 2020; pp. 1–5.
- 17. Raviteja, P.; Phan, K.T.; Hong, Y. Embedded Pilot-Aided Channel Estimation for OTFS in Delay–Doppler Channels. *IEEE Trans. Veh. Technol.* **2019**, *68*, 4906–4917. [CrossRef]
- 18. Mishra, H.B.; Singh, P.; Prasad, A.K.; Budhiraja, R. OTFS Channel Estimation and Data Detection Designs With Superimposed Pilots. *IEEE Trans. Wirel. Commun.* **2022**, *21*, 2258–2274. [CrossRef]
- 19. Gómez-Cuba, F. Compressed Sensing Channel Estimation for OTFS Modulation in Non-Integer Delay-Doppler Domain. In Proceedings of the 2021 IEEE Global Communications Conference (GLOBECOM), Madrid, Spain, 7–11 December 2021; pp. 1–6.
- Zhao, L.; Gao, W.-J.; Guo, W. Sparse Bayesian Learning of Delay-Doppler Channel for OTFS System. *IEEE Commun. Lett.* 2020, 24, 2766–2769. [CrossRef]

- 21. Sun, Y.; Babu, P.; Palomar, D.P. Majorization-Minimization Algorithms in Signal Processing, Communications, and Machine Learning. *IEEE Trans. Signal Process.* 2017, 65, 794–816. [CrossRef]
- 22. Tipping, M.E. Sparse Bayesian Learning and the Relevance Vector Machine. J. Mach. Learn. Res. 2011, 1, 211–244.

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.