Interference Cancellation and Iterative Detection for Orthogonal Time Frequency Space Modulation

P. Raviteja, *Student Member, IEEE*, Khoa T. Phan, *Member, IEEE*, Yi Hong, *Senior Member, IEEE*, and Emanuele Viterbo, *Fellow, IEEE*

Abstract— The recently proposed orthogonal time–frequency–space (OTFS) modulation technique was shown to provide significant error performance advantages over orthogonal frequency division multiplexing (OFDM) over delay-Doppler channels. In this paper, we first derive the explicit input–output relation describing OTFS modulation and demodulation (mod/demod). We then analyze the cases of: 1) ideal pulse-shaping waveforms that satisfy the bi-orthogonality conditions and 2) rectangular waveforms which do not. We show that while only *inter-Doppler interference* (IDI) is present in the former case, additional *inter-carrier interference* (ICI) and *inter-symbol interference* (ISI) occur in the latter case. We next characterize the interferences and develop a novel low-complexity yet efficient message passing (MP) algorithm for joint interference cancellation (IC) and symbol detection. While ICI and ISI are eliminated through appropriate phase shifting, IDI can be mitigated by adapting the MP algorithm to account for only the largest interference terms. The MP algorithm can effectively compensate for a wide range of channel Doppler spreads. Our results indicate that OTFS using practical rectangular waveforms can achieve the performance of OTFS using ideal but *non-realizable* pulse-shaping waveforms. Finally, simulation results demonstrate the superior error performance gains of the proposed *uncoded* OTFS schemes over OFDM under various channel conditions.

Index Terms— Delay–Doppler channel, OTFS, message passing, time–frequency modulation.

I. INTRODUCTION

FIFTH-GENERATION (5G) mobile systems are expected to accommodate an enormous number of emerging wireless applications with high data rate requirements such as real-time video streaming, and online gaming, connected and autonomous vehicles. While the orthogonal frequency division multiplexing (OFDM) modulation scheme currently deployed in fourth-generation (4G) mobile systems can achieve high spectral efficiency for time-invariant frequency selective channels, it is not robust to time-varying channels with high Doppler spread (e.g., high-speed railway mobile communications). Hence, new modulation techniques that are robust to channel time-variations have been extensively explored.

To cope with time-varying channels, one existing approach is to shorten the OFDM symbol duration so that the channel variations over each symbol appear inconsequential [1]. However, one major drawback is the reduced spectral efficiency due to cyclic prefix (CP). Another approach exploits time–frequency signaling [2], [3]. In [4] the authors have introduced the frequency-division multiplexing with frequency-domain cyclic prefix (FDM-FDCP), which can efficiently compensate for channel Doppler spread. In high-Doppler low-delay spread channels, FDM-FDCP is shown to outperform OFDM at the same spectral efficiency. The performance of FDM-FDCP under other channel conditions are yet to be studied.

A new time–frequency modulation technique called orthogonal time frequency space (OTFS) was recently proposed in [5] and [6], which shows significant advantages over OFDM in delay–Doppler channels. The delay–Doppler domain provides an alternative representation of a time-varying channel geometry modeling mobile terminals and reflectors [7], [8]. Leveraging on this representation, the OTFS modulator spreads each information symbol over a two dimensional (2D) orthogonal basis function, which spans across the entire time–frequency domain required to transmit a frame. The set of basis functions is specifically designed to combat the dynamics of the time-varying multi-path channels. In [5], a general framework of OTFS based on ideal pulse-shaping waveforms was introduced. A *coded* OTFS system with turbo equalization was compared with coded OFDM, showing remarkable gains. In [9], since mm-wave channels incur high frequency dispersion, OTFS is shown to outperform OFDM significantly.

In this paper, we analyze the input–output relation describing OTFS mod/demod for delay–Doppler channels using general pulse-shaping waveforms. The relation reveals the effects of the inverse symplectic finite Fourier transform (ISFFT) and SFFT operations interpreted as pre- and post-processing blocks applied to a time–frequency signaling scheme. We then analyze the cases of (*i*) ideal pulse-shaping waveforms that satisfy the bi-orthogonality conditions, and (*ii*) practical rectangular waveforms which do not. Unlike previous works [10], [11], [12], we assume *no* CP in the second case. While the OTFS input-output relation derived in [12] is complex, and does not provide insights, our work presents a simple relation characterizing the interference. Specifically, we show

that, while only *inter-Doppler interference* (IDI) is present in the ideal waveform case due to unavoidable fractional Doppler effects, additional *inter-carrier interference* (ICI) and *inter-symbol interference* (ISI) occur in the rectangular waveform case due to imperfect *bi-orthogonality* in time–frequency domain of the rectangular waveforms. Our interference analysis enables the development of an efficient algorithm for OTFS detection, which is the next main contribution of the work.

The delay–Doppler channel model with a small number of paths, with varying delay and Doppler values, provides a sparse representation of the communication channel. We then propose a low-complexity yet efficient message passing (MP) algorithm for a joint interference cancellation (IC) and detection, which takes advantage of the inherent delay–Doppler channel sparsity. The MP algorithm is based on a sparse factor graph and uses Gaussian approximation of the interference terms to further reduce the complexity. The approach is similar to [13], where it was applied to massive MIMO without the advantage of channel sparsity. The complexity and convergence of the MP algorithm are analyzed. In the MP algorithm, while the ICI and ISI can be eliminated by suitable phase shifting, the IDI can be mitigated by adapting the MP algorithm to account for only the largest interference terms. Consequently, the proposed MP algorithm can effectively compensate for a wide range of channel Doppler spreads. Further, our results show that OTFS using practical rectangular waveforms can achieve the performance of OTFS using ideal but *non-realizable* pulse-shaping waveforms. Simulations results illustrate the superior performance gains of the proposed *uncoded* OTFS schemes over OFDM under various channel conditions. Note that by considering uncoded systems in the simulation results, the focus is to demonstrate the performance gains of OTFS over OFDM due to specific OTFS mod/demod architecture (i.e., ISFFT and SFFT operations as pre- and post-processing blocks) and the effect of the proposed MP algorithm. The readers are referred to Hadani *el al.* works for the performance of coded OTFS.

The rest of the manuscript is organized as follows. Section II recalls the OTFS mod/demod and derives the corresponding input–output relation. In Section III, we analyze the time–frequency domain and delay–Doppler domain relations for the ideal waveform case. Section IV is dedicated to the case of OTFS using rectangular waveforms. Section V proposes the MP algorithm for the joint IC and detection. Simulation results are presented in Section VI followed by the conclusions in Section VII. The proofs are relegated to Appendix at the end of the manuscript.

II. System Model

In this section, we first recall the basic concepts in OTFS and then present the explicit analysis of OTFS mod/demod. More importantly, we derive the input–output relation of OTFS mod/demod for delay–Doppler channels.

A. Basic OTFS Concepts/Notations

We follow the notations in [5] and [14] summarized below.

– The *time–frequency signal plane* is discretized to a grid by sampling time and frequency axes at intervals T (seconds) and Δf (Hz), respectively, i.e.,

$$\Lambda = \big\{ (nT, m\Delta f),\ n = 0, \ldots, N-1, m = 0, \ldots, M-1 \big\}$$

for some integers $N, M > 0$.

– Modulated *time–frequency samples* $X[n,m]$, $n = 0, \ldots, N-1$, $m = 0, \ldots, M-1$ are transmitted over an OTFS frame with duration $T_f = NT$ and occupy a bandwidth $B = M\Delta f$.

– Transmit and receive pulses (or waveforms) are denoted by $g_{\text{tx}}(t)$ and $g_{\text{rx}}(t)$. Let $A_{g_{\text{rx}}, g_{\text{tx}}}(t, f)$ denote the *cross-ambiguity function* between $g_{\text{tx}}(t)$ and $g_{\text{rx}}(t)$, i.e.,

$$A_{g_{\text{rx}}, g_{\text{tx}}}(t, f) \triangleq \int g_{\text{rx}}^*(t' - t) g_{\text{tx}}(t') e^{-j2\pi f(t' - t)} dt'. \quad (1)$$

– The delay–Doppler plane is discretized to an information grid

$$\Gamma = \left\{ \left(\frac{k}{NT}, \frac{l}{M\Delta f} \right),\ k = 0, \ldots, N-1, l = 0, \ldots, M-1 \right\},$$

where $1/M\Delta f$ and $1/NT$ represent the quantization steps of the delay and Doppler frequency, respectively.[1]

Remark 1 (Choice of Parameters in OTFS Systems): Given a communications system with total bandwidth $B = M\Delta f$ and latency $T_f = NT$ constraints, we may choose N, M, T (since $\Delta f = 1/T$) to support communications over a time-varying channel with maximum delay $\tau_{\max}$ and maximum Doppler $\nu_{\max}$, among all channel paths. We can see that T and Δf determine the maximum supportable Doppler (i.e., $1/T$) and delay (i.e., $1/\Delta f$). Hence, it is required that $\nu_{\max} < 1/T$ and $\tau_{\max} < 1/\Delta f$ so that N and M are determined. To support a fixed data rate of NM symbols per frame, depending on the channel conditions, we can choose a larger T and smaller Δf, which results in a smaller N and larger M, respectively, or vice versa.

B. General OTFS Mod/Demod Block Diagram

The OTFS system diagram is given in Fig. 1. OTFS modulation is produced by a cascade of a pair of 2D transforms at both transmitter and receiver. The modulator first maps the information symbols $x[k, l]$ in the delay–Doppler domain to samples $X[n, m]$ in the time–frequency domain using the *inverse symplectic finite Fourier transform* (ISFFT). Next, the *Heisenberg transform* is applied to $X[n, m]$ to create the time domain signal $s(t)$ transmitted over the wireless channel. At the receiver, the time-domain signal $r(t)$ is mapped to the time–frequency domain through the *Wigner transform* (the inverse of the Heisenberg transform), and then to the delay–Doppler domain using SFFT for symbol demodulation.

C. OTFS Modulation

Consider a set of NM information symbols $\{x[k, l], k = 0, \ldots, N-1, l = 0, \ldots, M-1\}$ from a modulation alphabet

---

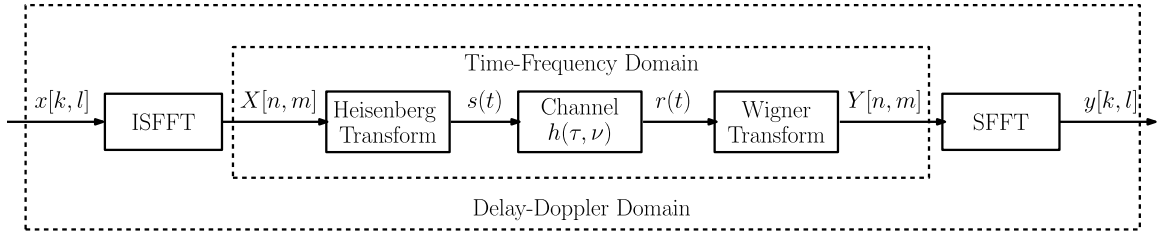[1]Note that the first and second indexes, k and l, in Γ represent the Doppler and delay axis, respectively.

Fig. 1. OTFS mod/demod.

of size Q $\mathbb{A} = \{a_1, \cdots, a_Q\}$ (e.g. QAM symbols), which are arranged on the delay–Doppler grid Γ.

The OTFS transmitter first maps symbols $x[k, l]$ to NM samples $X[n, m]$ on the time–frequency grid Λ using the ISFFT as follows

$$X[n, m] = \frac{1}{\sqrt{NM}} \sum_{k=0}^{N-1} \sum_{l=0}^{M-1} x[k, l] e^{j2\pi\left(\frac{nk}{N} - \frac{ml}{M}\right)} \quad (2)$$

for $n = 0, \ldots, N-1, m = 0, \ldots, M-1$.

Next, a time–frequency modulator converts the samples $X[n, m]$ to a continuous time waveform $s(t)$ using a transmit waveform $g_{\text{tx}}(t)$ as

$$s(t) = \sum_{n=0}^{N-1} \sum_{m=0}^{M-1} X[n, m] g_{\text{tx}}(t - nT) e^{j2\pi m\Delta f(t-nT)}. \quad (3)$$

As noted in [5], (3) is also referred to in the mathematical literature as the (discrete) Heisenberg transform [15], parametrized by $g_{\text{tx}}(t)$.

D. Wireless Transmission and Reception

The signal $s(t)$ is transmitted over a time-varying channel with complex baseband channel impulse response $h(\tau, \nu)$, which characterizes the channel response to an impulse with delay τ and Doppler ν [8]. The received signal $r(t)$ is given by (disregarding the noise to simplify notation)

$$r(t) = \int \int h(\tau, \nu) s(t - \tau) e^{j2\pi\nu(t-\tau)} d\tau d\nu. \quad (4)$$

Equation (4) represents a continuous Heisenberg transform parametrized by $s(t)$ [5]. Since typically there are only a small number of reflectors in the channel with associated delays and Dopplers, very few parameters are needed to model the channel in the delay–Doppler domain. The sparse representation of the channel $h(\tau, \nu)$ is given as

$$h(\tau, \nu) = \sum_{i=1}^{P} h_i \delta(\tau - \tau_i) \delta(\nu - \nu_i) \quad (5)$$

where P is the number of propagation paths; h_i, τ_i, and ν_i represent the path gain, delay, and Doppler shift (or frequency) associated with i-th path, respectively, and $\delta(\cdot)$ denotes the Dirac delta function. We denote the delay and Doppler taps for i-th path as follows

$$\tau_i = \frac{l_{\tau_i}}{M\Delta f}, \quad \nu_i = \frac{k_{\nu_i} + \kappa_{\nu_i}}{NT} \quad (6)$$

for integers l_{τ_i}, k_{ν_i} and real $-\frac{1}{2} < \kappa_{\nu_i} \leq \frac{1}{2}$. Specifically, l_{τ_i} and k_{ν_i} represent the indexes of the delay tap and Doppler tap corresponding to (continuous) delay τ_i and Doppler frequency ν_i, respectively. We will refer to κ_{ν_i} as the *fractional* Doppler since it represents the fractional shift from the nearest Doppler tap k_{ν_i}. We do not need to consider fractional delays since the resolution of the sampling time $\frac{1}{M\Delta f}$ is sufficient to approximate the path delays to the nearest sampling points in typical wide-band systems [16].

E. OTFS Demodulation

At the receiver, a matched filter computes the cross-ambiguity function $A_{g_{\text{rx}},r}(t, f)$ as follows

$$Y(t, f) = A_{g_{\text{rx}},r}(t, f) \triangleq \int g_{\text{rx}}^*(t' - t) r(t') e^{-j2\pi f(t'-t)} dt'. \quad (7)$$

The matched filter output is obtained by sampling $Y(t, f)$ as

$$Y[n, m] = Y(t, f)|_{t=nT, f=m\Delta f} \quad (8)$$

for $n = 0, \ldots, N-1$ and $m = 0, \ldots, M-1$. Operations (7) and (8) are referred as the *Wigner transform*. In the following theorem, we characterize the relationship between time–frequency output samples $Y[n, m]$ and input samples $X[n, m]$.

Theorem 1 (OTFS Time–Frequency Domain Analysis): The following input–output relation of OTFS in time–frequency domain is given by

$$Y[n, m] = \sum_{n'=0}^{N-1} \sum_{m'=0}^{M-1} H_{n,m}[n', m'] X[n', m'], \quad (9)$$

where

$$H_{n,m}[n', m']$$
$$= \int \int h(\tau, \nu) A_{g_{\text{rx}},g_{\text{tx}}}((n - n')T - \tau, (m - m')\Delta f - \nu)$$
$$\times e^{j2\pi(\nu+m'\Delta f)((n-n')T-\tau)} e^{j2\pi\nu n'T} d\tau d\nu. \quad (10)$$

Proof: The proof is given in Appendix A. ∎

We can see that the terms $H_{n,m}[n', m']$ include the combined effects of the transmit pulse, channel response, and receive pulse. Note that similar results have been presented for the case of pulse-shaped (PS) OFDM [17], [18].

Next, the SFFT is applied on the samples $Y[n, m]$ to obtain symbols $y[k, l]$ in the delay–Doppler domain

$$y[k, l] = \frac{1}{\sqrt{NM}} \sum_{n=0}^{N-1} \sum_{m=0}^{M-1} Y[n, m] e^{-j2\pi\left(\frac{nk}{N} - \frac{ml}{M}\right)}. \quad (11)$$

Theorem 1 provides the basis of the study of OTFS in two special cases, namely using *ideal waveforms* (Section III) and more practical *rectangular waveforms* (Section IV). We will obtain explicit input-output relations using the delay–Doppler channel model (5) for both cases.

III. OTFS WITH IDEAL WAVEFORMS

The $g_{\mathrm{rx}}(t)$ and $g_{\mathrm{tx}}(t)$ pulses are said to be ideal if they satisfy the *bi-orthogonal property* [5]

$$A_{g_{\mathrm{rx}},g_{\mathrm{tx}}}(t,f)|_{t=nT+(-\tau_{\max},\tau_{\max}),f=m\Delta f+(-\nu_{\max},\nu_{\max})}$$
$$= \delta[n]\delta[m]q_{\tau_{\max}}(t)q_{\nu_{\max}}(f) \qquad (12)$$

where $q_a(x) = 1$ for $x \in (-a,a)$ and zero otherwise. Equivalently, the cross-ambiguity function $A_{g_{\mathrm{rx}},g_{\mathrm{tx}}}(t,f) = 0$ for $t \in (nT-\tau_{\max}, nT+\tau_{\max})$ and $f \in (m\Delta f - \nu_{\max}, m\Delta f + \nu_{\max})$, $\forall n,m$ except for $n = 0, m = 0$, where $A_{g_{\mathrm{rx}},g_{\mathrm{tx}}}(t,f) = 1$ for $t \in (-\tau_{\max}, \tau_{\max})$ and $f \in (-\nu_{\max}, \nu_{\max})$.

Unfortunately, ideal pulses cannot be realized in practice but can be approximated by waveforms with a support concentrated as much as possible in time and in frequency, given the constraint imposed by the uncertainty principle. Nevertheless, it is important to study the error performance of OTFS with ideal waveforms since it serves as a lower bound on the performance of OTFS with practically realizable waveforms such as rectangular waveforms, etc.

A. Time–Frequency Domain Analysis

For ideal waveforms, the following result was given in [5] without proof. Here, we show that it can be obtained as a special case of Theorem 1.

Proposition 1: For ideal pulses, the following result can be obtained

$$Y[n,m] = H_{n,m}[n,m]X[n,m] \qquad (13)$$

where

$$H_{n,m}[n,m] = \int\int h(\tau,\nu)e^{j2\pi\nu nT}e^{-j2\pi(\nu+m\Delta f)\tau}d\tau d\nu.$$

Proof: From (10), we observe that the value of $H_{n,m}[n',m']$ is non-zero only at $n' = n$ and $m' = m$ for the ideal pulses satisfying the bi-orthogonal property (12). Hence, the result in (13) follows from (9) by considering only the term with $n' = n$, $m' = m$ in the summations. ∎

B. Delay–Doppler Domain Analysis

1) Input–Output Relationship: We now apply SFFT on $Y[n,m]$ in (13) to obtain the symbols $y[k,l]$ in the delay–Doppler domain. The following proposition, given in [5] without proof, describes the input–output relation in delay–Doppler domain.

Proposition 2: For ideal pulses, the following input-output relation holds

$$y[k,l] = \frac{1}{NM}\sum_{k'=0}^{N-1}\sum_{l'=0}^{M-1} x[k',l']h_w[k-k',l-l'], \qquad (14)$$

where $h_w[\cdot,\cdot]$ is a sampled version of the impulse response function

$$h_w[k-k',l-l'] = h_w(\nu,\tau)|_{\nu=\frac{k-k'}{NT},\tau=\frac{l-l'}{M\Delta f}}$$

for $h_w(\nu,\tau)$ being the circular convolution of the channel response with the SFFT of a rectangular windowing function in the time-frequency domain

$$h_w(\nu,\tau) = \int\int h(\tau',\nu')w(\nu-\nu',\tau-\tau')e^{-j2\pi\nu\tau}d\tau'd\nu', \qquad (15)$$

$$w(\nu,\tau) = \sum_{n=0}^{N-1}\sum_{m=0}^{M-1} 1 \cdot e^{-j2\pi(\nu nT-\tau m\Delta f)}. \qquad (16)$$

Proof: The proof is relegated to the Appendix B. ∎

2) Inter-Doppler Interference (IDI) Analysis: From (14), we can see that a received signal $y[k,l]$ is a linear combination of all the transmitted signals $x[k',l']$, $k' = 0,\ldots,N-1, l' = 0,\ldots,M-1$. Consequently, the input-output relation (14) can be represented as a linear system with NM variables $x[k',l']$. Since N and M tend to be very large for practical OTFS systems, the detection complexity can be prohibitive. In the following, by using (5) as the sparse representation of the delay–Doppler channel, (14) reduces to a sparse linear system, where each received signal can be approximately expressed as a linear combination of only a few transmitted signals. Such sparsity will then be exploited in Section V to devise a low-complexity yet efficient iterative detection algorithm based on message passing on the factor graph representation.

By substituting (5) and (16) into (15), we obtain

$$h_w(\nu,\tau) = \sum_{i=1}^{P} h_i e^{-j2\pi\nu_i\tau_i}\,w(\nu-\nu_i,\tau-\tau_i)$$
$$= \sum_{i=1}^{P} h_i e^{-j2\pi\nu_i\tau_i}\,\mathcal{G}(\nu,\nu_i)\,\mathcal{F}(\tau,\tau_i),$$

where we have denoted

$$\mathcal{F}(\tau,\tau_i) \triangleq \sum_{m'=0}^{M-1} e^{j2\pi\,(\tau-\tau_i)m'\Delta f},$$

$$\mathcal{G}(\nu,\nu_i) \triangleq \sum_{n'=0}^{N-1} e^{-j2\pi(\nu-\nu_i)n'T}.$$

Let us first evaluate $\mathcal{F}(\tau,\tau_i)$ at $\tau = \frac{l-l'}{M\Delta f}$ as

$$\mathcal{F}\left(\frac{l-l'}{M\Delta f},\tau_i\right) = \sum_{m'=0}^{M-1} e^{j\frac{2\pi}{M}(l-l'-l_{\tau_i})m'} = \frac{e^{j2\pi(l-l'-l_{\tau_i})}-1}{e^{j\frac{2\pi}{M}(l-l'-l_{\tau_i})}-1} \qquad (17)$$

recalling that l_{τ_i} is the delay tap of i-th path with a delay τ_i defined in (6). From (17), we see that

$$\mathcal{F}\left(\frac{l-l'}{M\Delta f},\tau_i\right) = \begin{cases} M, & [l-l'-l_{\tau_i}]_M = 0, \\ 0, & \text{otherwise,} \end{cases}$$

where $[\cdot]_M$ represents mod M operation, i.e., $\mathcal{F}\left(\frac{l-l'}{M\Delta f},\tau_i\right)$ equals to M for $l' = [l-l_{\tau_i}]_M$ and is zero otherwise.

Similarly, we can evaluate

$$\mathcal{G}\left(\frac{k-k'}{NT}, \nu_i\right) = \frac{e^{-j2\pi(k-k'-k_{\nu_i}-\kappa_{\nu_i})} - 1}{e^{-j\frac{2\pi}{N}(k-k'-k_{\nu_i}-\kappa_{\nu_i})} - 1}. \quad (18)$$

Due to the fractional κ_{ν_i}, we can see that for a given k, $\mathcal{G}\left(\frac{k-k'}{NT}, \nu_i\right) \neq 0$, for all k'.

We will show that the magnitude of $\frac{1}{N}\mathcal{G}\left(\frac{k-k'}{NT}, \nu_i\right)$ has a peak at $k' = k - k_{\nu_i}$ and decreases as k' moves away from $k - k_{\nu_i}$. From (18), after some manipulations, we have

$$\left|\frac{1}{N}\mathcal{G}\left(\frac{k-k'}{NT}, \nu_i\right)\right| = \left|\frac{\sin(N\theta)}{N\sin(\theta)}\right|$$

where we set $\theta \triangleq -\frac{\pi}{N}(k - k' - k_{\nu_i} - \kappa_{\nu_i})$. It can be easily shown that

$$\left|\frac{\sin(N\theta)}{N\sin(\theta)}\right| = \left|\frac{\sin((N-1)\theta)\cos(\theta) + \sin(\theta)\cos((N-1)\theta)}{N\sin(\theta)}\right|$$

$$\leq \frac{N-1}{N}|\cos(\theta)| + \frac{1}{N}. \quad (19)$$

Here, we used the inequality, $|\sin(N\theta)| \leq N|\sin(\theta)|$, which can be proven by induction. The upper bound (19) is tight for small values of θ (when both sides are close to 1) and it has a peak at the smallest value of θ when $k' = k - k_{\nu_i}$. As $|\theta|$ increases (due to k' moving away from $k - k_{\nu_i}$), the upper bound decreases with (approximate) slope of $\frac{\pi}{N}(k - k' - k_{\nu_i} - \kappa_{\nu_i})$. Since N is quite large in OTFS, the function decreases rapidly.

From the above analysis, we need to consider only a small number $2N_i + 1$, for some $N_i > 0$, of significant values of $\mathcal{G}\left(\frac{k-k'}{NT}, \nu_i\right)$ in (18) around the peak $k - k_{\nu_i}$, i.e., $[k - k_{\nu_i} - N_i]_N \leq k' \leq [k - k_{\nu_i} + N_i]_N$, where $N_i \ll N$. Using this approximation, we can now express the receive signal $y[k, l]$ in (14) as

$$y[k, l] \approx \sum_{i=1}^{P} \sum_{k'=[k-k_{\nu_i}-N_i]_N}^{[k-k_{\nu_i}+N_i]_N} \left(\frac{e^{-j2\pi(k-k_{\nu_i}-k'-\kappa_{\nu_i})} - 1}{Ne^{-j\frac{2\pi}{N}(k-k_{\nu_i}-k'-\kappa_{\nu_i})} - N}\right)$$
$$\times h_i e^{-j2\pi\nu_i\tau_i} x[k', [l - l_{\tau_i}]_M]$$
$$\approx \sum_{i=1}^{P} \sum_{q=-N_i}^{N_i} \left(\frac{e^{-j2\pi(-q-\kappa_{\nu_i})} - 1}{Ne^{-j\frac{2\pi}{N}(-q-\kappa_{\nu_i})} - N}\right) h_i e^{-j2\pi\nu_i\tau_i}$$
$$\times x[[k - k_{\nu_i} + q]_N, [l - l_{\tau_i}]_M]. \quad (20)$$

In the simulation result section, we will demonstrate that for $N = 128$, by choosing $N_i = 10$, negligible performance loss is incurred. From (20), we can see that the received signal $y[k, l]$ is a linear combination of $S = \sum_{i=1}^{P} 2N_i + 1$ transmitted signals. Out of $2N_i + 1$ transmitted signals in i-th path, the signal corresponding to $q = 0$, $x[[k - k_{\nu_i}]_N, [l - l_{\tau_i}]_M]$, contributes the most and all the other $2N_i$ signals can be seen as interference. Such interference is due to the transmitted signals that are neighboring $x[[k - k_{\nu_i}]_N, [l - l_{\tau_i}]_M]$ in the Doppler domain and we refer to this interference as *inter-Doppler interference (IDI)*. Further, the number of transmitted signals S affecting the received signal $y[k, l]$ in (20) is much smaller than NM in (14). Hence, the graph (or linear system) describing (20) is sparsely-connected.

3) Special Channel Model Cases: The above input-output expression simplifies for the following special cases.

i) Ideal channel: Assuming $h(\tau, \nu) = \delta(\tau)\delta(\nu)$, the received signal becomes

$$y[k, l] = x[k, l]$$

and behaves as an AWGN channel as expected.

ii) No fractional Doppler (i.e., $\kappa_{\nu_i} = 0, \forall i$): Assuming that Doppler frequencies exactly coincide with Doppler taps, the received signal can be obtained by replacing $N_i = 0$ in (20), i.e.,

$$y[k, l] = \sum_{i=1}^{P} h_i e^{-j2\pi\nu_i\tau_i} x[[k - k_{\nu_i}]_N, [l - l_{\tau_i}]_M]. $$

For each path, the transmitted signal is circularly shifted by the delay and Doppler taps and scaled by the associated channel gain.

IV. OTFS WITH RECTANGULAR WAVEFORMS

Since the ideal pulses cannot be realized in practice, we now analyze the OTFS with the rectangular pulses at both the transmitter and receiver. These pulses do not satisfy the bi-orthogonality conditions and generate some interference which degrades the system performance. Here, we analyze the effect of such interference and show that it can be compensated to achieve the ideal pulses performance.

We assume the rectangular pulse has amplitude $1/\sqrt{T}$ for $t \in [0, T]$ and 0 at all other values, to have unit energy.

A. Time–Frequency Domain Analysis

For the rectangular pulses, we can see that the cross-ambiguity term in the time–frequency relation of Theorem 1, $A_{g_{rx}, g_{tx}}((n - n')T - \tau, (m - m')\Delta f - \nu)$ is non-zero for $|\tau| < \tau_{max}$ $|\nu| < \nu_{max}$ only when $n' = n$ and $n' = n - 1$, since $g_{tx}(t)$ and $g_{rx}(t)$ are pulses of duration T and $\tau_{max} \ll T$. Hence, the time–frequency relation (9) becomes

$$Y[n, m] = \sum_{n'=n-1}^{n} \sum_{m'=0}^{M-1} H_{n,m}[n', m']X[n', m']$$
$$= H_{n,m}[n, m]X[n, m]$$
$$+ \sum_{m'=0, m'\neq m}^{M-1} H_{n,m}[n, m']X[n, m']$$
$$+ \sum_{m'=0}^{M-1} H_{n,m}[n - 1, m']X[n - 1, m']. \quad (21)$$

The second term in (21) can be seen as the total interference from the samples $X[n, m']$ at different frequencies $m' \neq m$, but in the same time slot n as the current sample $X[n, m]$. On the other hand, the third term in (21) accumulates the interference from the samples $X[n - 1, m']$ in the previous time slot $n - 1$. Hence, we call the second and third terms as the *inter carrier interference (ICI)* and *inter symbol interference (ISI)*, respectively. The interference depends on the delay τ and Doppler ν of the channel. In particular, they are affected by the value of the cross-ambiguity function

$A_{g_{\mathrm{rx}},g_{\mathrm{tx}}}((n-n')T-\tau,(m-m')\Delta f - \nu)$ in $H_{n,m}[n',m']$. In the following, we focus on the cross-ambiguity function for ICI and ISI.

1) ICI Analysis: Fix n,m. We note that the cross-ambiguity function in the $H_{n,m}[n,m']$, $m' \neq m$ term of ICI, $A_{g_{\mathrm{rx}},g_{\mathrm{tx}}}(-\tau,(m-m')\Delta f - \nu)$, is independent of n, and is computed for the i-th channel path with delay τ_i and Doppler ν_i (i.e., see (5)) as

$$A_{\mathrm{ici}} \triangleq \int g_{\mathrm{rx}}^*(t'+\tau_i)g_{\mathrm{tx}}(t')e^{-j2\pi((m-m')\Delta f - \nu_i))(t'+\tau_i)}dt'.$$

We discard the dependency of A_{ici} on (m,m',τ_i,ν_i) for simplicity. Since the received signal $r(t)$ is sampled at intervals of T/M (or $1/(M\Delta f)$), we can compute A_{ici} as

$$A_{\mathrm{ici}} = \frac{1}{M}\sum_{p=0}^{M-1-l_{\tau_i}} e^{-j2\pi((m-m')\Delta f - \nu_i)\left(\frac{p}{M\Delta f}+\tau_i\right)}. \quad (22)$$

Recall that the pulses $g_{\mathrm{tx}}(t)$ and $g_{\mathrm{rx}}(t)$ have duration T, and l_{τ_i} is the delay tap defined in (6). The amplitude of A_{ici} is

$$|A_{\mathrm{ici}}| = \frac{1}{M}\left|\sum_{p=0}^{M-1-l_{\tau_i}} e^{-j2\pi((m-m')\Delta f - \nu_i)\frac{p}{M\Delta f}}\right|$$

$$= \frac{\left|e^{-j2\pi\left(m-m'-\frac{k_{\nu_i}+\kappa_{\nu_i}}{N}\right)\frac{M-l_{\tau_i}}{M}}-1\right|}{\left|Me^{-j2\pi\left(m-m'-\frac{k_{\nu_i}+\kappa_{\nu_i}}{N}\right)\frac{1}{M}}-M\right|}.$$

Similar to the analysis of (18), we can observe that $|A_{\mathrm{ici}}|$ decreases as m' moves away from m. It implies that the ICI becomes less as the interfering subcarriers are further away from the interfered subcarrier. We can also see that an increase in Doppler (i.e., $k_{\nu_i}+\kappa_{\nu_i}$) increases the number of neighboring subcarriers that interfere with the present subcarrier. This is similar to the fractional Doppler effect studied for (18).

2) ISI Analysis: Similar to the ICI analysis, the cross-ambiguity function in the $H_{n,m}[n-1,m']$ term of ISI, $A_{\mathrm{isi}} \triangleq A_{g_{\mathrm{rx}},g_{\mathrm{tx}}}(T-\tau,(m-m')\Delta f - \nu)$, is computed for the i-th channel path as

$$A_{\mathrm{isi}} = \frac{1}{M}\sum_{p=M-l_{\tau_i}}^{M-1} e^{-j2\pi((m-m')\Delta f - \nu_i)\left(\frac{p}{M\Delta f}+\tau_i - T\right)}. \quad (23)$$

The amplitude $|A_{\mathrm{isi}}|$ also has similar properties of $|A_{\mathrm{ici}}|$, where it reduces as m' moves away from m implying that the ISI is smaller for interfering symbols further away (in the frequency axis) from the interfered symbol.

Note that the terms that affect the ICI and ISI in the summations (22) and (23) are mutually exclusive, i.e., $p=0$ to $M-1-l_{\tau_i}$ contributes to ICI whereas $p = M-l_{\tau_i}$ to $M-1$ contributes to ISI. This property helps in differentiating the ICI and ISI effects in delay–Doppler domain, which will be studied below.

B. Delay–Doppler Domain Analysis

We now characterize the input–output relation in delay–Doppler domain for OTFS with rectangular pulses.

Theorem 2: The received signal $y[k,l]$ in delay–Doppler domain with the rectangular pulses can be written as

$$y[k,l] \approx \sum_{i=1}^{P}\sum_{q=-N_i}^{N_i} h_i e^{j2\pi\left(\frac{l-l_{\tau_i}}{M}\right)\left(\frac{k_{\nu_i}+\kappa_{\nu_i}}{N}\right)}\alpha_i(k,l,q)$$

$$\times x\left[[k-k_{\nu_i}+q]_N,[l-l_{\tau_i}]_M\right] \quad (24)$$

where we have

$$\alpha_i(k,l,q) = \begin{cases} \dfrac{1}{N}\beta_i(q) & l_{\tau_i} \leq l < M \\ \dfrac{1}{N}\left(\beta_i(q)-1\right)e^{-j2\pi\frac{[k-k_{\nu_i}+q]_N}{N}} & 0 \leq l < l_{\tau_i} \end{cases}$$

$$\beta_i(q) = \frac{e^{-j2\pi(-q-\kappa_{\nu_i})}-1}{e^{-j\frac{2\pi}{N}(-q-\kappa_{\nu_i})}-1}. \quad (25)$$

Proof: The proof is relegated to the Appendix C. ∎

Note that the approximation error in (24) is very small and it reduces by increasing N (see (62) in Appendix C). Theorem 2 implies that the ICI and ISI in time–frequency domain are converted to simple phase shifts in the delay–Doppler domain. Moreover, from (20) and (24), we can observe that the number of transmitted signals that affects a received signal is the same for both ideal and rectangular pulse cases. The only difference is that the channel is shifted by an additional phase that depends on the location of the transmitted signal in the delay–Doppler plane (i.e., k and l).

Special channel model cases: Let us consider the above input-output expression (24) for the special cases mentioned in Sec. III-B.3.

i) Ideal channel: The received signal becomes

$$y[k,l] = x[k,l],$$

which is the same as the ideal pulses case since the rectangular pulses satisfy the bi-orthogonal property in (12) when the channel is ideal (i.e., $\tau_{\max} = 0$ and $\nu_{\max} = 0$). This can be seen easily by observing (1) at $t = nT$ and $f = m\Delta f$.

ii) No fractional Doppler (i.e., $\kappa_{\nu_i} = 0, \forall i$): Equation (24) simplifies to

$$y[k,l] \approx \sum_{i=1}^{P} h_i e^{j2\pi\left(\frac{l-l_{\tau_i}}{M}\right)\frac{k_{\nu_i}}{N}}\alpha_i(k,l)x[[k-k_{\nu_i}]_N,[l-l_{\tau_i}]_M],$$

where

$$\alpha_i(k,l) = \begin{cases} 1 & l_{\tau_i} \leq l < M \\ \dfrac{N-1}{N}e^{-j2\pi\left(\frac{[k-k_{\nu_i}]_N}{N}\right)} & 0 \leq l < l_{\tau_i}. \end{cases}$$

In this case, IDI does not appear as in the case of ideal pulses.

V. MESSAGE PASSING ALGORITHM FOR JOINT INTERFERENCE CANCELLATION AND DETECTION

We now propose a message passing (MP) algorithm for OTFS using the input-output relation in (20) (or (24)).

A. Low-Complexity MP Detection Algorithm

The received signal in vectorized form can be written as

$$\mathbf{y} = \mathbf{H}\,\mathbf{x} + \mathbf{z} \qquad (26)$$

where $\mathbf{y}$ and $\mathbf{z}$ are complex vectors of dimension $NM \times 1$ with elements denoted by $y[d]$ and $z[d], 1 \leq d \leq NM$, respectively; $\mathbf{H}$ is a $NM \times NM$ complex matrix with elements $H[d,c], 1 \leq d,c \leq NM$; $\mathbf{x}$ is the information vector of dimension $NM \times 1$ with elements $x[c] \in \mathbb{A}, 1 \leq c \leq NM$.[2] The elements of $\mathbf{y}, \mathbf{x}$, and $\mathbf{H}$ are determined from (20) (or (24)) and $\mathbf{z}$ is the noise vector. Due to mod N and mod M operations in (20), we observe that only $S = \sum_{i=1}^{P}(2N_i + 1)$ elements out of NM are non-zero in each row and column of $\mathbf{H}$. Recall that P is the number of propagation paths. We can see that since S is much smaller than NM, $\mathbf{H}$ is a sparse matrix. Let $\mathcal{I}(d)$ and $\mathcal{J}(c)$ denote the sets of indexes with non-zero elements in the d-th row and c-th column, respectively, then $|\mathcal{I}(d)| = |\mathcal{J}(c)| = S$ for all rows and columns. Note that although (26) applies to both ideal pulses case in (20) and rectangular pulses case in (24), with different matrices $\mathbf{H}$, the number of non-zero elements S in each row and column of $\mathbf{H}$ remains the same for both cases. This condition helps in compensating ICI and ISI of rectangular pulses with the same complexity detection algorithm of ideal pulses.

Based on (26), we model the system as a sparsely-connected factor graph with NM variable nodes corresponding to $\mathbf{x}$ and NM observation nodes corresponding to $\mathbf{y}$. In this factor graph, each observation node $y[d]$ is connected to the set of S variable nodes $\{x[c], c \in \mathcal{I}(d)\}$. Similarly, each variable node $x[c]$ is connected to the set of S observation nodes $\{y[d], d \in \mathcal{J}(c)\}$.

From (26), the joint maximum a posterior probability (MAP) detection rule for estimating the transmitted signals is given by

$$\widehat{\mathbf{x}} = \underset{\mathbf{x} \in \mathbb{A}^{NM \times 1}}{\arg\max} \; \Pr\left(\mathbf{x}\,|\,\mathbf{y}, \mathbf{H}\right),$$

which has a complexity exponential in NM. Since the joint MAP detection can be intractable for practical values of N and M, we consider the symbol-by-symbol MAP detection rule for $c = 1, \ldots, NM$

$$\widehat{x}[c] = \underset{a_j \in \mathbb{A}}{\arg\max}\, \Pr\left(x[c] = a_j | \mathbf{y}, \mathbf{H}\right)$$

$$= \underset{a_j \in \mathbb{A}}{\arg\max}\, \frac{1}{Q} \Pr\left(\mathbf{y}|x[c] = a_j, \mathbf{H}\right) \qquad (27)$$

$$\approx \underset{a_j \in \mathbb{A}}{\arg\max} \prod_{d \in \mathcal{J}_c} \Pr\left(y[d]|x[c] = a_j, \mathbf{H}\right). \qquad (28)$$

In (27), we assume all the transmitted symbols $a_j \in \mathbb{A}$ are equally likely and in (28) we assume the components of $\mathbf{y}$ are approximately independent for a given $x[c]$, due to the sparsity of $\mathbf{H}$. That is, we assume the interference terms $\zeta_{d,c}^{(i)}$ defined in (29) are independent for a given c. In order to solve the approximate symbol-by-symbol MAP detection in (28),

---

[2]The proposed algorithm can also be applied for the corresponding real valued systems which is beneficial for higher order QAM modulation. Here, for the general case, we consider a complex valued system.
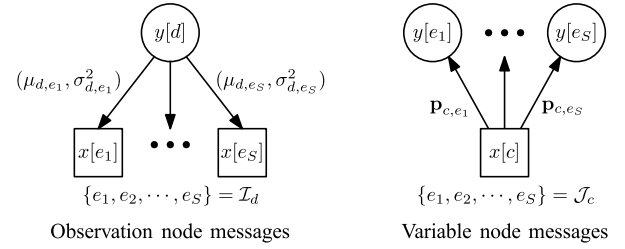

Fig. 2. Messages in factor graph.

we propose a MP detector which has a linear complexity in NM. Similarly to [13], for each $y[d]$, a variable $x[c]$ is isolated from the other interference terms, which are then approximated as Gaussian noise with an easily computable mean and variance.

In the MP algorithm, the mean and variance of the interference terms are used as messages from observation nodes to variable nodes. On the other hand, the message passed from a variable node $x[c]$ to the observation nodes $y[d], d \in \mathcal{J}(c)$, is the probability mass function (pmf) of the alphabet $\mathbf{p}_{c,d} = \{p_{c,d}(a_j)|a_j \in \mathbb{A}\}$. Fig. 2 shows the connections and the messages passed between the observation and variable nodes. The MP algorithm is described in **Algorithm 1**.

Algorithm 1 MP Algorithm for OTFS Symbol Detection

Input: Received signal $\mathbf{y}$, channel matrix $\mathbf{H}$.
Initialization: pmf $\mathbf{p}_{c,d}^{(0)} = 1/Q, c = 1, \cdots, NM, d \in \mathcal{J}(c)$, iteration count $i = 1$.
repeat

- Observation nodes $y[d]$ compute the means $\mu_{d,c}^{(i)}$ and variances $(\sigma_{d,c}^{(i)})^2$ of Gaussian random variables $\zeta_{d,c}^{(i)}$ using $\mathbf{p}_{c,d}^{(i-1)}$ and pass them to variables nodes $x[c], c \in \mathcal{I}(d)$.
- Variable nodes $x[c]$ update $\mathbf{p}_{c,d}^{(i)}$ using $\mu_{d,c}^{(i)}$, $(\sigma_{d,c}^{(i)})^2$, and $\mathbf{p}_{c,d}^{(i-1)}$ and pass them to observation nodes $y[d], d \in \mathcal{J}(c)$.
- Compute convergence indicator $\eta^{(i)}$.
- Update the decision on the transmitted symbols $\widehat{x}[c], c = 1, \ldots, NM$ if needed.
- $i = i + 1$

until *Stopping criteria*;
Output: The decision on transmitted symbols $\widehat{x}[c]$.

The details of the steps in iteration i in the MP algorithm are detailed below. **Message passings from observation nodes** $y[d]$ **to variable nodes** $x[c], c \in \mathcal{I}(d)$: The mean $\mu_{d,c}^{(i)}$ and variance $(\sigma_{d,c}^{(i)})^2$ of the interference, approximately modeled as a Gaussian random variable $\zeta_{d,c}^{(i)}$ defined as

$$y[d] = x[c]H[d,c] + \underbrace{\sum_{e \in \mathcal{I}(d), e \neq c} x[e]H[d,e] + z[d]}_{\zeta_{d,c}^{(i)}}, \quad (29)$$

can be computed as

$$\mu_{d,c}^{(i)} = \sum_{e \in \mathcal{I}(d), e \neq c} \sum_{j=1}^{Q} p_{e,d}^{(i-1)}(a_j) a_j H[d,e], \qquad (30)$$

and

$$(\sigma_{d,c}^{(i)})^2 = \sum_{e \in \mathcal{I}(d), e \neq c} \left(\sum_{j=1}^{Q} p_{e,d}^{(i-1)}(a_j)|a_j|^2 |H[d,e]|^2 \right.$$
$$\left. - \left| \sum_{j=1}^{Q} p_{e,d}^{(i-1)}(a_j)a_j H[d,e] \right|^2 \right) + \sigma^2. \quad (31)$$

Message passings from variable nodes $x[c]$ to observation nodes $y[d], d \in \mathcal{J}(c)$: The pmf vector $\mathbf{p}_{c,d}^{(i)}$ can be updated as

$$p_{c,d}^{(i)}(a_j) = \Delta \cdot \tilde{p}_{c,d}^{(i)}(a_j) + (1-\Delta) \cdot p_{c,d}^{(i-1)}(a_j), \quad a_j \in \mathbb{A} \quad (32)$$

where $\Delta \in (0,1]$ is the *damping factor* used to improve the performance by controlling the convergence speed [19], and

$$\tilde{p}_{c,d}^{(i)}(a_j) \propto \prod_{e \in \mathcal{J}(c), e \neq d} \Pr\left(y[e] \Big| x[c] = a_j, \mathbf{H}\right)$$
$$= \prod_{e \in \mathcal{J}(c), e \neq d} \frac{\xi^{(i)}(e,c,j)}{\sum_{k=1}^{Q} \xi^{(i)}(e,c,k)}, \quad (33)$$

where $\xi^{(i)}(e,c,k) = \exp\left(\frac{-|y[e]-\mu_{e,c}^{(i)}-H_{e,c}a_k|^2}{(\sigma_{e,c}^{(i)})^2} \right).$

Convergence indicator: Compute the convergence indicator $\eta^{(i)}$ as

$$\eta^{(i)} = \frac{1}{NM} \sum_{c=1}^{NM} \mathbb{I}\left(\max_{a_j \in \mathbb{A}} p_c^{(i)}(a_j) \geq 1 - \gamma \right), \quad (34)$$

for some small $\gamma > 0$ and where

$$p_c^{(i)}(a_j) = \prod_{e \in \mathcal{J}(c)} \frac{\xi^{(i)}(e,c,j)}{\sum_{k=1}^{Q} \xi^{(i)}(e,c,k)} \quad (35)$$

and $\mathbb{I}(\cdot)$ is an indicator function which gives a value of 1, if the expression in the argument is true, and 0 otherwise.

Update decision: If $\eta^{(i)} > \eta^{(i-1)}$, then we update the decision of the transmitted symbol as

$$\widehat{x}[c] = \arg\max_{a_j \in \mathbb{A}} p_c^{(i)}(a_j), \quad c = 1, \cdots, NM. \quad (36)$$

We update the decision on the transmitted symbols only when the current iteration can provide better estimates than the previous iteration.

Stopping criteria. The MP algorithm stops when at least one of the following conditions is satisfied.
1) $\eta^{(i)} = 1$.
2) $\eta^{(i)} < \eta^{(i^*)} - \epsilon$, where i^* is the iteration index from $\{1, \cdots, (i-1)\}$ for which $\eta^{(i^*)}$ is maximum.
3) Maximum number n_{iter} of iterations is reached.

We select $\epsilon = 0.2$ to disregard small fluctuations of η. Here, the first condition occurs in the best case, where all the symbols have converged. The second condition is useful to stop the algorithm if the current iteration provides a worse decision than the one in previous iterations.

Remark 2 (Complexity of the Proposed MP Algorithm): The complexity of one iteration involves the computation of (30), (31), (32), (34), and (36). More specifically, each

of (30), (31), and (32),[3] has a complexity order $\mathcal{O}(NMSQ)$. Furthermore, (34) and (36) can be computed with a complexity order $\mathcal{O}(NMQ)$.[4] Therefore, the overall complexity order is $\mathcal{O}(n_{\text{iter}}NMSQ)$. In simulations, we observed that the algorithm converges typically within 20 iterations (i.e., see Figure 4 in the illustrative result section for more references). We conclude that the IDI analysis, which includes the smart approximation of IDI, to exploit the sparsity of the delay-Doppler channel representation is a key factor in reducing the complexity of the detector (due to relatively small S). The memory requirement is dominated by the storage of $2NMSQ$ real values for $\mathbf{p}_{c,d}^{(i)}$ and $\mathbf{p}_{c,d}^{(i-1)}$. In addition, we have the messages $(\mu_{d,c}^{(i)}, (\sigma_{d,c}^{(i)})^2)$, requiring NMS complex values and NMS real values, respectively.

B. Application of MP Detection Algorithm for OFDM Over Delay–Doppler Channels

In the simulation result section, we will compare the performance of OTFS and OFDM over delay–Doppler channels. In this section, we demonstrate that it is also possible to utilize the above MP algorithm to compensate the Doppler effects in OFDM systems.

Consider OFDM system with OFDM symbol of duration T and M subcarriers. Hence, the received signal and noise are sampled at T/M. Then, the frequency-domain signal after fast Fourier transform (FFT) operation is given by

$$\mathbf{y} = \mathbf{W}\mathbf{H}_{\text{t}}\mathbf{W}^{\text{H}}\mathbf{x} + \mathbf{z} \quad (37)$$

where $(\cdot)^{\text{H}}$ denotes Hermitian transpose, $\mathbf{W}$ is M-point FFT matrix, and $\mathbf{x} \in \mathbb{A}^{M \times 1}$ is the transmitted OFDM symbol. The elements $H_{\text{t}}[p,q]$ of time-domain channel matrix $\mathbf{H}_{\text{t}}$ are given in [20] as

$$H_{\text{t}}[p,q] = \sum_{i=1}^{P} h_i \delta\left[\left[p - q - \frac{\tau_i M}{T} \right]_M \right] e^{j\frac{2\pi (q-1)\nu_i}{M}},$$
$$p, q = 1, \ldots, M.$$

Using the $M \times M$ frequency-domain channel matrix $\mathbf{H}_{\text{ofdm}} = \mathbf{W}\mathbf{H}_{\text{t}}\mathbf{W}^{\text{H}}$, we can re-write (37) as

$$\mathbf{y} = \mathbf{H}_{\text{ofdm}}\mathbf{x} + \mathbf{z}. \quad (38)$$

Since (38) has the form similar to (26), the MP previously developed for OTFS can also be applied for OFDM symbol detection. We note that $\mathbf{H}_{\text{ofdm}}$ is *diagonally dominant* and the values of off-diagonal elements in each row decay as we move away from the diagonal entry as explained in [20]. Hence, the $\mathbf{H}_{\text{ofdm}}$ matrix is also sparse enabling the use of the proposed low-complexity MP detection algorithm.

---

[3]In computing (33), first we find the $p_c^{(i)}(a_j)$ in (35) which requires $\mathcal{O}(NMQ)$ complexity and then we obtain (33) by dividing (35) with the term related to $e = d$ for all d, which requires $\mathcal{O}(S)$ complexity for each c. Hence, the over all complexity of (33) becomes $\mathcal{O}(NMSQ)$.

[4]The computation of (34) and (36) require to find the maximum element out of Q elements for each c. As (35) is already computed for (33), finding the maximum element requires $\mathcal{O}(Q)$ complexity for each c, which leads to an overall complexity of $\mathcal{O}(NMQ)$ to compute (34) and (36).

Parameter	Value
Carrier frequency	4 GHz
No. of subcarriers (M)	512
No. of OTFS symbols (N)	128
Subcarrier spacing	15 KHz
Cyclic prefix of OFDM	2.6 μs
Modulation alphabet	4-QAM, 16-QAM
UE speed (Kmph)	30, 120, 500
Channel estimation	Ideal
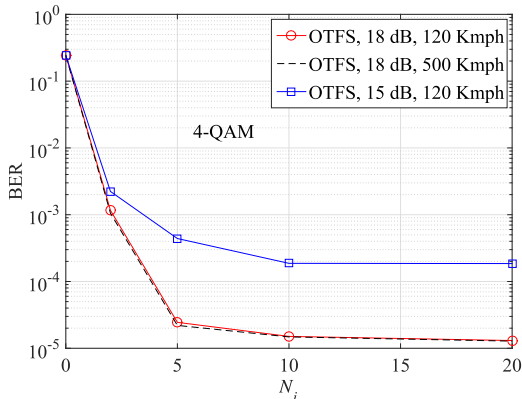
Fig. 3. The BER performance of OTFS for different number of interference terms N_i with 4-QAM.

Fig. 4. The variation of BER and average no. of iterations with Δ.

VI. ILLUSTRATIVE RESULTS AND DISCUSSIONS

In this section, we simulate the error performance of uncoded OTFS and OFDM systems over delay-Doppler channels. In particular, we first study the bit-error-rate (BER) performance of OTFS for ideal pulses with the number of IDI interference terms N_i and MP parameter Δ. We next study the BER performance of OTFS with ideal pulses and rectangular pulses, and its comparison with OFDM.
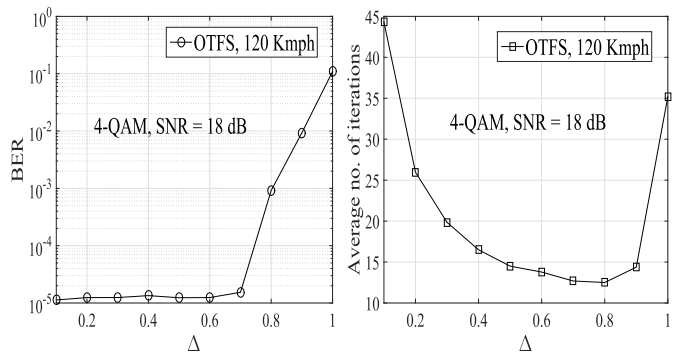
All relevant simulation parameters are given in Table I. First, ideal channel estimation is assumed, i.e., the channel impulse function $h(\tau, \nu)$ is perfectly known at the receiver. Then, we consider the effect of imperfect channel estimation on OTFS performance. For both OTFS and OFDM systems, Extended Vehicular A model [22] is adopted as the channel model for the path delays (or delay taps) and each delay tap has a single Doppler shift generated using Jakes' formula

$$\nu_i = \nu_{\max} \cos(\theta_i)$$

where $\nu_{\max}$ is the maximum Doppler shift determined by the UE speed and θ_i is uniformly distributed over $[-\pi, \pi]$. In general, the channel can have multiple paths with the same delay but different Doppler shifts. In our simulations, we consider channels with only one path for given delay. However, it is straightforward to consider multiple paths with different Doppler shifts for a given delay, and our main theoretical result in Theorem 2 is still valid.

In order to obtain BER values, we consider 7×10^4 different channel realizations in the Monte-Carlo simulations.

We first demonstrate the effects of IDI in OTFS. Fig. 3 shows the BER performance of OTFS system with ideal pulses

using the proposed MP detector for different number of IDI interference terms N_i with 4-QAM signaling over the delay–Doppler channel with different Doppler frequencies (i.e., UE speeds of 120, 500 Kmph) and SNRs. Note that ICI and ISI are not present for the ideal pulses case. We consider the same N_i for all paths. We can see that there is a significant BER improvement when N_i increases from 0 to 10 and saturation thereafter. Note that $N_i = 0$ corresponds to the case when IDI is not taken into account. The results imply that fewer neighboring interference terms are sufficient to consider in the MP algorithm (e.g. $N_i = 10$) without incurring performance loss. We also observe that if IDI is not taken into account at all or an insufficient number of IDI terms is considered (i.e., $N_i \leq 5$), the error performance worsens significantly. These observations demonstrate the importance of our previous IDI analysis for suitable interference cancellation algorithm development. Also, it can be observed that for given SNR (e.g., SNR = 18 dB), the error performances of OTFS with different Doppler frequencies are similar since the proposed MP algorithm can effectively compensate for a wide range of channel Doppler variations as demonstrated further in the following.

In Fig. 4, we illustrate the BER performance and average number of iterations with ideal pulses using the MP algorithm. We fix $N_i = 10$ and vary the damping factor Δ. We consider 4-QAM signaling, SNR = 18 dB, and UE speed of 120 Kmph. We observe that, when $\Delta \leq 0.7$, the BER remains almost the same, but deteriorates thereafter. Further, when $\Delta = 0.7$, the MP algorithm converges with the least number of iterations. Hence, we choose $\Delta = 0.7$ as the optimum damping factor in terms of performance and complexity.

In Fig. 5, we compare the BER performance of OTFS with ideal pulses and OFDM using 4-QAM signaling over the delay-Doppler channels of different Doppler frequencies (i.e., UE speeds of 30, 120, 500 Kmph). Note that the MP algorithm proposed in Section V.B is used for OFDM detection. We observe that OTFS outperforms OFDM by approximately 15 dB at BER of 10^{-4} thanks to the constant channel gain over all transmitted symbols in OTFS, whereas in OFDM, the overall error performance is limited by the subcarrier(s) experiencing the worse channel conditions. Moreover, OTFS exhibits the same performance at different Doppler frequencies thanks to the IDI cancellation provided by the MP detector.
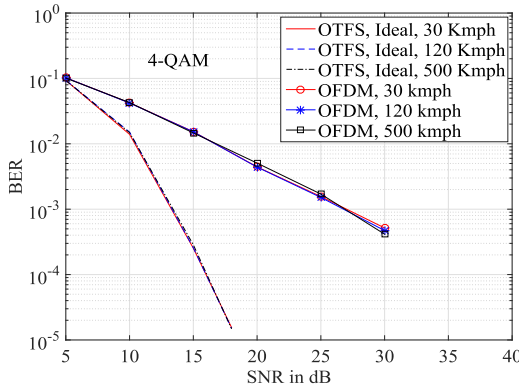
Fig. 5. The BER performance comparison between OTFS with ideal pulses and OFDM systems at different Doppler frequencies.
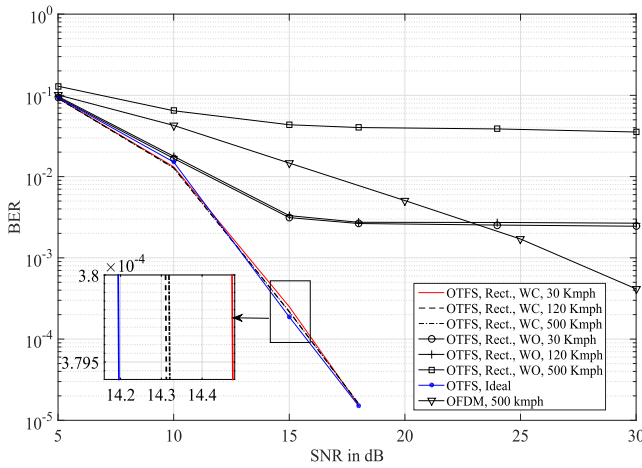

Fig. 6. The BER performance of OTFS with rectangular and ideal pulses at different Doppler frequencies for 4-QAM.
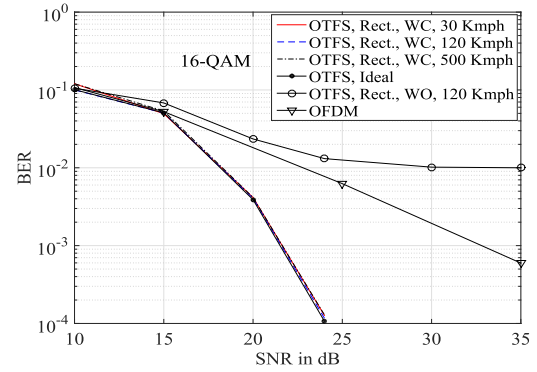

Fig. 7. The BER performance of OTFS with rectangular and ideal pulses for 16-QAM.
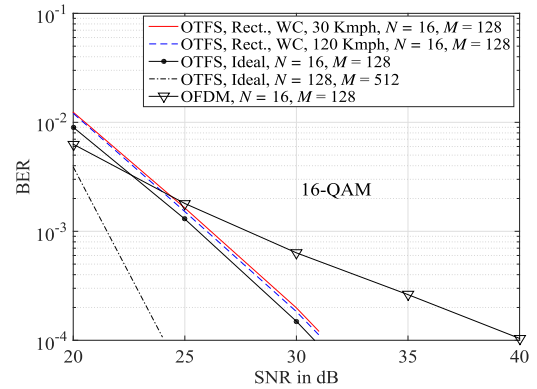
Fig. 8. The BER performance of OTFS with rectangular pulses and low latency ($N = 16, T_f \approx 1.1$).

Similar behavior applies to OFDM, since the ICI can be removed by the MP detector. We can conclude that the performance of OTFS under the proposed MP algorithm is robust to Doppler variations and is much better than that of OFDM.

Fig. 6 shows the BER performance of OTFS with rectangular pulses using 4-QAM signaling for two scenarios: one with ICI and ISI cancellations (WC) and the other without (WO). In the second scenario, we observe that OTFS with rectangular pulses has an error floor incurred by the ICI and ISI. The performance degradation becomes more severe at high Doppler (e.g., 500 Kmph) due to large ICI and ISI. On the other hand, the BER performance of OTFS with rectangular pulses approaches that of OTFS with ideal pulses, when ISI and ICI are mitigated. Moreover, we can see that the proposed MP algorithm can effectively mitigate ISI and ICI and thus OTFS performance remains almost constant regardless of the Doppler frequencies. These results show that it is possible to achieve the performance of OTFS with ideal waveforms even with the more practical rectangular waveforms by using our MP algorithm together with appropriate IDI, ICI and ISI cancellation.

In Fig. 7, we compare the BER performance of OTFS and OFDM at a Doppler of 120 Kmph using 16-QAM signaling. We observe that OTFS with ICI and ISI cancellation outperforms OFDM by 11 dB at BER = 10^{-3}. We also simulate OTFS at different Doppler frequencies of 30 and 500 Kmph and we observe the BER performances are similar to that of 120 Kmph.

In the next experiments, we study the performance of OTFS under different practical constraints, namely low-latency and non-ideal channel estimation.

Fig. 8 shows the OTFS performance under low-latency constraint where $N = 16$, $M = 128$, and 16-QAM. In this experiment, the frame duration is $T_f = NT \approx 1.1$ ms, which is much smaller than the previous case with $N = 128$ and $T_f \approx 8.8$ ms. We can observe that the OTFS performance is the same for different Dopplers. Further, the performance of OTFS degrades with low-latency because the delay–Doppler grid has lower resolution on the Doppler axis, and hence, the receiver resolves a smaller number of paths in the channel. This leads to the diversity loss over the high-latency case. When compared to OFDM, we observe that OTFS outperforms OFDM below BER = 2×10^{-3} with a significant diversity gain.

Fig. 9 shows the effect of imperfect channel estimation on the performance of OTFS with $N = 128$, $M = 512$, 16-QAM, and a Doppler of 120 Kmph. Here, we introduce the error in
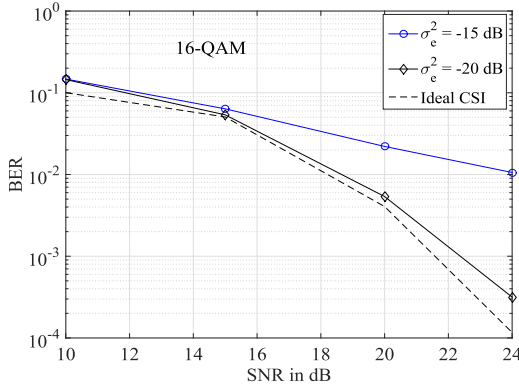
Fig. 9. The BER performance of OTFS with channel estimation errors.

the channel estimation using the model [21]

$$h_i' = h_i + n_e, 1 \le i \le P$$

where $n_e \sim \mathcal{CN}(0, \sigma_e^2)$ and we assume the delay and Doppler taps are perfectly estimated. From Fig. 9, we see that as the noise variance reduces, performance of OTFS is approaching the ideal system, particularly for a noise variance of -20 dB it is very close to the ideal system. A comprehensive study on the performance of OTFS under non-ideal channel estimation is out of the scope of this paper, and will be a topic of future work.

VII. CONCLUSION

In this paper, we have analyzed the input–output relation describing OTFS mod/demod over delay–Doppler channels. We have studied in detail the cases of ideal waveforms and rectangular waveforms. In particular, we have characterized the inter-Doppler interference (IDI), inter-carrier interference (ICI), and inter-symbol interference (ISI) using sparse representation of the channel in the delay–Doppler domain. A low-complexity yet efficient message passing (MP) algorithm for joint IC and symbol detection was proposed, which is suitable for large-scale OTFS with inherent channel sparsity. In the MP algorithm, the ISI and ICI can be canceled by using appropriate phase shifting, while the IDI can be mitigated by accounting for a small number of significant interference terms only. The proposed MP algorithm can effectively compensate for a wide range of channel Doppler spreads. Moreover, we have demonstrated that it is

possible to achieve the performance of OTFS with ideal yet *non-realizable* waveforms using *practical* rectangular waveforms. Through simulations, we have shown that OTFS has significant error performance gains over OFDM under various communications scenarios including ideal and non-ideal channel estimation, low-latency communications etc.

APPENDIX A
PROOF OF THEOREM 1: OTFS INPUT–OUTPUT RELATION IN TIME–FREQUENCY DOMAIN

The received signal after Wigner transform $Y(t, f)$, from (4), can be written as in (39), as shown at the bottom of this page. It can be further expanded as in (40) and (41), as shown at the bottom of this page, using the transmitted signal $s(t)$ in (3) and some re-ordering of summations and integrations. Therefore, the sampled version of $Y(t, f)$, i.e., $Y[n, m]$, can be written as

$$Y[n, m] = \sum_{n'=0}^{N-1} \sum_{m'=0}^{M-1} X[n', m']H_{n,m}[n', m'],$$

where $H_{n,m}[n', m']$ is given in (42), as shown at the top of the next page. By applying the change of variable $t' - \tau - n'T \to t''$ in the inner integral and some simple algebraic calculations, we can write $H_{n,m}[n', m']$ as in (43) and (44), as shown at the top of the next page, respectively. Finally, we obtain $H_{n,m}[n', m']$ as in (45), as shown at the top of the next page, by replacing the square bracket in (44) with cross-ambiguity function in (1), which completes the proof.

APPENDIX B
PROOF OF PROPOSITION 2: OTFS INPUT–OUTPUT RELATION IN DELAY–DOPPLER DOMAIN FOR IDEAL PULSES

The received signal $y[k, l]$ for the ideal pulses, from (11) and (13), can be written as

$$y[k, l] = \frac{1}{\sqrt{NM}} \sum_{n=0}^{N-1} \sum_{m=0}^{M-1} H_{n,m}[n, m]X[n, m]e^{-j2\pi\left(\frac{nk}{N} - \frac{ml}{M}\right)}.$$

By substituting the ISFFT equation from (2), $y[k, l]$ can be expanded as in from (46) to (48), as shown at the top of the next page. Here, $h_w[k - k', l - l']$ can be seen as the value of $h_w(\nu, \tau)$ sampled at $\nu = \frac{k-k'}{NT}, \tau = \frac{l-l'}{M\Delta f}$. The value of

$$Y(t, f)$$
$$= \int_{t'} g_{rx}^*(t' - t) \left[\int_\tau \int_\nu h(\tau, \nu)s(t' - \tau)e^{j2\pi\nu(t'-\tau)}d\tau d\nu \right] e^{-j2\pi f(t'-t)}dt' \tag{39}$$
$$= \int_{t'} g_{rx}^*(t' - t) \left[\int_\tau \int_\nu h(\tau, \nu) \left\{ \sum_{n'=0}^{N-1} \sum_{m'=0}^{M-1} X[n', m']g_{tx}(t' - \tau - n'T)e^{j2\pi m'\Delta f(t'-\tau-n'T)} \right\} e^{j2\pi\nu(t'-\tau)}d\tau d\nu \right]$$
$$\times e^{-j2\pi f(t'-t)}dt'. \tag{40}$$
$$= \sum_{n'=0}^{N-1} \sum_{m'=0}^{M-1} X[n', m'] \left[\int_\tau \int_\nu h(\tau, \nu) \left\{ \int_{t'} g_{rx}^*(t' - t)g_{tx}(t' - \tau - n'T)e^{j2\pi m'\Delta f(t'-\tau-n'T)}e^{j2\pi\nu(t'-\tau)}e^{-j2\pi f(t'-t)}dt' \right\}d\tau d\nu \right]. \tag{41}$$

$$H_{n,m}[n',m']$$

$$= \int_\tau \int_\nu h(\tau,\nu) \left[\int_{t'} g_{\text{rx}}^*(t'-nT) g_{\text{tx}}(t'-\tau-n'T) e^{j2\pi m'\Delta f(t'-\tau-n'T)} e^{j2\pi\nu(t'-\tau)} e^{-j2\pi m\Delta f(t'-nT)} dt' \right] d\tau d\nu \tag{42}$$

$$= \int_\tau \int_\nu h(\tau,\nu) \left[\int_{t''} g_{\text{rx}}^*(t''-(n-n')T+\tau) g_{\text{tx}}(t'') e^{j2\pi m'\Delta f t''} e^{j2\pi\nu(t''+n'T)} e^{-j2\pi m\Delta f(t''+(n-n')T+\tau)} dt'' \right] d\tau d\nu \tag{43}$$

$$= \int_\tau \int_\nu h(\tau,\nu) \left[\int_{t''} g_{\text{rx}}^*(t''-(n-n')T+\tau) g_{\text{tx}}(t'') e^{-j2\pi\left((m-m')\Delta f-\nu\right)\left(t''-(n-n')T+\tau\right)} dt'' \right]$$
$$e^{j2\pi(\nu+m'\Delta f)((n-n')T-\tau)} e^{j2\pi\nu n'T} d\tau d\nu \tag{44}$$

$$= \int_\tau \int_\nu h(\tau,\nu) A_{g_{\text{rx}},g_{\text{tx}}}((n-n')T-\tau,(m-m')\Delta f-\nu) e^{j2\pi(\nu+m'\Delta f)((n-n')T-\tau)} e^{j2\pi\nu n'T} d\tau d\nu. \tag{45}$$

$$y[k,l] = \frac{1}{NM} \sum_{n=0}^{N-1} \sum_{m=0}^{M-1} H_{n,m}[n,m] \left[\sum_{k'=0}^{N-1} \sum_{l'=0}^{M-1} x[k',l'] e^{j2\pi\left(\frac{nk'}{N}-\frac{ml'}{M}\right)} \right] e^{-j2\pi\left(\frac{nk}{N}-\frac{ml}{M}\right)} \tag{46}$$

$$= \frac{1}{NM} \sum_{k'=0}^{N-1} \sum_{l'=0}^{M-1} x[k',l'] \left[\sum_{n=0}^{N-1} \sum_{m=0}^{M-1} H_{n,m}[n,m] e^{-j2\pi nT\left(\frac{k-k'}{NT}\right)} e^{j2\pi m\Delta f\left(\frac{l-l'}{M\Delta f}\right)} \right] \tag{47}$$

$$= \frac{1}{NM} \sum_{k'=0}^{N-1} \sum_{l'=0}^{M-1} x[k',l'] h_w[k-k',l-l']. \tag{48}$$

$$h_w(\nu,\tau) = \sum_{n=0}^{N-1} \sum_{m=0}^{M-1} \left[\int_{\tau'} \int_{\nu'} h(\tau',\nu') e^{j2\pi\nu'nT} e^{-j2\pi(\nu'+m\Delta f)\tau'} d\tau' d\nu' \right] e^{-j2\pi nT\nu} e^{j2\pi m\Delta f\tau} \tag{49}$$

$$= \int_{\tau'} \int_{\nu'} h(\tau',\nu') \left[\sum_{n=0}^{N-1} \sum_{m=0}^{M-1} e^{-j2\pi(\nu-\nu')nT} e^{j2\pi(\tau-\tau')m\Delta f} \right] e^{-j2\pi\tau'\nu'} d\tau' d\nu' \tag{50}$$

$$= \int_{\tau'} \int_{\nu'} h(\tau',\nu') w(\nu-\nu',\tau-\tau') e^{-j2\pi\tau'\nu'} d\tau' d\nu'. \tag{51}$$

$$y[k,l] = \frac{1}{\sqrt{NM}} \sum_{n=0}^{N-1} \sum_{m=0}^{M-1} \left[\sum_{m'=0}^{M-1} H_{n,m}[n,m'] X[n,m'] + \sum_{m'=0}^{M-1} H_{n,m}[n-1,m'] X[n-1,m'] \right] e^{-j2\pi\left(\frac{nk}{N}-\frac{ml}{M}\right)}. \tag{52}$$

$$y_{\text{ici}}[k,l] = \frac{1}{NM} \sum_{n=0}^{N-1} \sum_{m=0}^{M-1} \sum_{m'=0}^{M-1} H_{n,m}[n,m'] \left[\sum_{k'=0}^{N-1} \sum_{l'=0}^{M-1} x[k',l'] e^{j2\pi\left(\frac{nk'}{N}-\frac{m'l'}{M}\right)} \right] e^{-j2\pi\left(\frac{nk}{N}-\frac{ml}{M}\right)}$$

$$= \frac{1}{NM} \sum_{k'=0}^{N-1} \sum_{l'=0}^{M-1} x[k',l'] \left[\sum_{n=0}^{N-1} \sum_{m=0}^{M-1} \sum_{m'=0}^{M-1} H_{n,m}[n,m'] e^{-j2\pi n\left(\frac{k-k'}{N}\right)} e^{j2\pi\left(\frac{ml-m'l'}{M}\right)} \right]$$

$$= \frac{1}{NM} \sum_{k'=0}^{N-1} \sum_{l'=0}^{M-1} x[k',l'] h_{k,l}^{\text{ici}}[k',l']. \tag{53}$$

$h_w(\nu,\tau)$ can be obtained as from (49) to (51), as shown at the top of this page, by substituting $H_{n,m}[n,m]$ from (10), which completes the proof.

APPENDIX C
PROOF OF THEOREM 2: OTFS INPUT–OUTPUT RELATION IN DELAY–DOPPLER DOMAIN FOR RECTANGULAR PULSES

We start with expanding $y[k,l]$ in (11) using the $Y[n,m]$ for rectangular pulses in (21) as in (52), as shown at the top of this page. We write $y[k,l]$ as

$$y[k,l] = y_{\text{ici}}[k,l] + y_{\text{isi}}[k,l],$$

where $y_{\text{ici}}[k,l]$ and $y_{\text{isi}}[k,l]$ contains the first term and the second term of the summation in square brackets

of (52), respectively. We analyze these ICI and ISI terms as below.

Analysis of $y_{\text{ici}}[k,l]$: The value of $y_{\text{ici}}[k,l]$ can be written as in (53), as shown at the top of this page, using the ISFFT of $X[n,m]$ given in (2). Now, $h_{k,l}^{\text{ici}}[k',l']$ is expanded in (54), as shown on the next page, by using the $H_{n,m}[n,m']$ value in (10). This can be further written as in (55), as shown on the next page, from the channel assumption in (5) and the cross-ambiguity function in (22).

To write the expression in (55) to a simple form, let us separate the terms related to n, m, m', and p. The terms related to n are

$$\zeta_n = e^{-j2\pi n\left(\frac{k-k'}{N}\right)} e^{j2\pi\nu_i nT}$$
$$= e^{-j2\pi n\left(\frac{k-k'-k_{\nu_i}-\kappa_{\nu_i}}{N}\right)}.$$

$$h_{k,l}^{\mathrm{ici}}[k',l']$$

$$= \sum_{n=0}^{N-1}\sum_{m=0}^{M-1}\sum_{m'=0}^{M-1}\left[\int_{\tau}\int_{\nu}h(\tau,\nu)A_{g_{\mathrm{rx}},g_{\mathrm{tx}}}(-\tau,(m-m')\Delta f-\nu)e^{-j2\pi\ (\nu+m'\Delta f)\tau}e^{j2\pi\nu nT}d\tau d\nu\right]e^{-j2\pi n\left(\frac{k-k'}{N}\right)}e^{j2\pi\left(\frac{ml-m'l'}{M}\right)}$$

$$(54)$$

$$= \frac{1}{M}\sum_{n=0}^{N-1}\sum_{m=0}^{M-1}\sum_{m'=0}^{M-1}\left[\sum_{i=1}^{P}h_i\sum_{p=0}^{M-1-l_{\tau_i}}e^{-j2\pi((m-m')\Delta f-\nu_i)(p(T/M)+\tau_i)}e^{-j2\pi\ (\nu_i+m'\Delta f)\tau_i}e^{j2\pi\nu_i nT}\right]e^{-j2\pi n\left(\frac{k-k'}{N}\right)}e^{j2\pi\left(\frac{ml-m'l'}{M}\right)}$$

$$(55)$$

$$= \sum_{i=1}^{P}h_i\left[\sum_{n=0}^{N-1}e^{-j2\pi n\left(\frac{k-k'-k_{\nu_i}-\kappa_{\nu_i}}{N}\right)}\right]\left[\frac{1}{M}\sum_{p=0}^{M-1-l_{\tau_i}}e^{j2\pi\frac{p}{M}\left(\frac{k_{\nu_i}+\kappa_{\nu_i}}{N}\right)}\sum_{m=0}^{M-1}e^{-j2\pi\ (p+l_{\tau_i}-l)\frac{m}{M}}\sum_{m'=0}^{M-1}e^{j2\pi\ (p-l')\frac{m'}{M}}\right].$$

$$= \sum_{i=1}^{P}h_i\mathcal{G}^{\mathrm{ici}}(\nu_i)\mathcal{F}^{\mathrm{ici}}(\tau_i,\nu_i). \tag{56}$$

$$\mathcal{F}^{\mathrm{ici}}(\tau_i,\nu_i) = M\sum_{p=0}^{M-1-l_{\tau_i}}e^{j2\pi\frac{p}{M}\left(\frac{k_{\nu_i}+\kappa_{\nu_i}}{N}\right)}\delta([p+l_{\tau_i}-l]_M)\delta([p-l']_M). \tag{57}$$

$$y_{\mathrm{ici}}[k,l] = \frac{1}{N}\sum_{i=1}^{P}h_i\left[\sum_{l'=0}^{M-1}\sum_{p=0}^{M-1-l_{\tau_i}}e^{j2\pi\frac{p}{M}\left(\frac{k_{\nu_i}+\kappa_{\nu_i}}{N}\right)}\delta([p+l_{\tau_i}-l]_M)\delta([p-l']_M)\sum_{k'=0}^{N-1}\mathcal{G}^{\mathrm{ici}}(\nu_i)x[k',l']\right]$$

$$\approx \frac{1}{N}\sum_{i=1}^{P}h_i\left[\sum_{p=0}^{M-1-l_{\tau_i}}e^{j2\pi\frac{p}{M}\left(\frac{k_{\nu_i}+\kappa_{\nu_i}}{N}\right)}\delta([p+l_{\tau_i}-l]_M)\sum_{q=-N_i}^{N_i}\left(\frac{-e^{j2\pi(-q-\kappa_{\nu_i})}-1}{e^{-j\frac{2\pi}{N}(-q-\kappa_{\nu_i})}-1}\right)x[[k-k_{\nu_i}+q]_N,p]\right]. \tag{58}$$

$$y_{\mathrm{ici}}[k,l] \approx \begin{cases}\displaystyle\sum_{i=1}^{P}\sum_{q=-N_i}^{N_i}h_i\left[\frac{1}{N}\beta_i(q)\right]e^{j2\pi\left(\frac{l-l_{\tau_i}}{M}\right)\left(\frac{k_{\nu_i}+\kappa_{\nu_i}}{N}\right)}x\left[[k-k_{\nu_i}+q]_N,[l-l_{\tau_i}]_M\right] & l\geq l_{\tau_i},\\[4mm]0 & \text{otherwise.}\end{cases} \tag{59}$$

$$y_{\mathrm{isi}}[k,l] = \frac{1}{NM}\sum_{n=0}^{N-1}\sum_{m=0}^{M-1}\sum_{m'=0}^{M-1}H_{n,m}[n-1,m']\left[\sum_{k'=0}^{N-1}\sum_{l'=0}^{M-1}x[k',l']e^{j2\pi\left(\frac{(n-1)k'}{N}-\frac{m'l'}{M}\right)}\right]e^{-j2\pi\left(\frac{nk}{N}-\frac{ml}{M}\right)}$$

$$= \frac{1}{NM}\sum_{k'=0}^{N-1}\sum_{l'=0}^{M-1}e^{-j2\pi\frac{k'}{N}}x[k',l']\left[\sum_{n=0}^{N-1}\sum_{m=0}^{M-1}\sum_{m'=0}^{M-1}H_{n,m}[n-1,m']e^{-j2\pi n\left(\frac{k-k'}{N}\right)}e^{j2\pi\left(\frac{ml-m'l'}{M}\right)}\right]$$

$$= \frac{1}{NM}\sum_{k'=0}^{N-1}\sum_{l'=0}^{M-1}e^{-j2\pi\frac{k'}{N}}x[k',l']h_{k,l}^{\mathrm{isi}}[k',l']. \tag{60}$$

$$h_{k,l}^{\mathrm{isi}}[k',l'] = \sum_{i=1}^{P}h_i\left[\sum_{n=1}^{N-1}e^{-j2\pi n\left(\frac{k-k'-k_{\nu_i}-\kappa_{\nu_i}}{N}\right)}\right]\left[\frac{1}{M}\sum_{p=M-l_{\tau_i}}^{M-1}e^{j2\pi\left(\frac{p-M}{M}\right)\left(\frac{k_{\nu_i}+\kappa_{\nu_i}}{N}\right)}\sum_{m=0}^{M-1}e^{-j2\pi\ (p+l_{\tau_i}-l+M)\frac{m}{M}}\sum_{m'=0}^{M-1}e^{j2\pi\ (p-l')\frac{m'}{M}}\right]$$

$$= \sum_{i=1}^{P}h_i\mathcal{G}^{\mathrm{isi}}(\nu_i)\mathcal{F}^{\mathrm{isi}}(\tau_i,\nu_i). \tag{61}$$

$$y_{\mathrm{isi}}[k,l] = \frac{1}{N}\sum_{i=1}^{P}h_i\left[\sum_{l'=0}^{M-1}\sum_{p=M-l_{\tau_i}}^{M-1}e^{j2\pi\left(\frac{p-M}{M}\right)\left(\frac{k_{\nu_i}+\kappa_{\nu_i}}{N}\right)}\delta([p+l_{\tau_i}-l]_M)\delta([p-l']_M)\sum_{k'=0}^{N-1}\mathcal{G}^{\mathrm{isi}}(\nu_i)e^{-j2\pi\frac{k'}{N}}x[k',l']\right]$$

$$= \frac{1}{N}\sum_{i=1}^{P}h_i\left[\sum_{p=M-l_{\tau_i}}^{M-1}e^{j2\pi\left(\frac{p-M}{M}\right)\left(\frac{k_{\nu_i}+\kappa_{\nu_i}}{N}\right)}\delta([p+l_{\tau_i}-l]_M)\sum_{k'=0}^{N-1}\mathcal{G}^{\mathrm{isi}}(\nu_i)e^{-j2\pi\frac{k'}{N}}x[k',p]\right]$$

$$\approx \frac{1}{N} \sum_{i=1}^{P} h_i \left[\sum_{p=M-l_{\tau_i}}^{M-1} e^{j2\pi \left(\frac{p-M}{M} \right) \left(\frac{k_{\nu_i} + \kappa_{\nu_i}}{N} \right)} \delta([p + l_{\tau_i} - l]_M) \left\{ \sum_{q=-N_i}^{N_i} (\beta_i(q) - 1) e^{-j2\pi \frac{[k-k_{\nu_i}+q]_N}{N}} x[[k - k_{\nu_i} + q]_N, p] \right. \right.$$

$$\left. \left. - \sum_{\substack{k'=0, \\ k' \neq [k-k_{\nu_i}+q]_N, q \in [-N_i, N_i]}}^{N-1} e^{-j2\pi \frac{k'}{N}} x[k', p] \right\} \right] \tag{62}$$

$$\approx \frac{1}{N} \sum_{i=1}^{P} h_i \left[\sum_{p=M-l_{\tau_i}}^{M-1} e^{j2\pi \left(\frac{p-M}{M} \right) \left(\frac{k_{\nu_i} + \kappa_{\nu_i}}{N} \right)} \delta([p + l_{\tau_i} - l]_M) \sum_{q=-N_i}^{N_i} (\beta_i(q) - 1) e^{-j2\pi \frac{[k-k_{\nu_i}+q]_N}{N}} x[[k - k_{\nu_i} + q]_N, p] \right]. \tag{63}$$

$$y_{\text{isi}}[k, l] \approx \begin{cases} \sum_{i=1}^{P} \sum_{q=-N_i}^{N_i} h_i \left[\frac{1}{N} (\beta_i(q) - 1) \right] e^{-j2\pi \frac{[k-k_{\nu_i}+q]_N}{N}} e^{j2\pi \left(\frac{l-l_{\tau_i}}{M} \right) \left(\frac{k_{\nu_i} + \kappa_{\nu_i}}{N} \right)} x[[k - k_{\nu_i} + q]_N, [l - l_{\tau_i}]_M] & l < l_{\tau_i}, \\ 0 & \text{otherwise.} \end{cases} \tag{64}$$

Here, we used the delay and Doppler taps defined in (6). Similarly, the terms related to m and m' are

$$\zeta_m = e^{-j2\pi m \Delta f \ (p(T/M) + \tau_i)} e^{j2\pi l \frac{m}{M}}$$
$$= e^{-j2\pi \ (p + l_{\tau_i} - l) \frac{m}{M}}.$$
$$\zeta_{m'} = e^{j2\pi m' \Delta f \ (p(T/M) + \tau_i)} e^{-j2\pi m' \Delta f \tau_i} e^{-j2\pi l' \frac{m'}{M}}$$
$$= e^{j2\pi \ (p - l') \frac{m'}{M}}.$$

Finally, the terms related to p are

$$\zeta_p = e^{j2\pi \nu_i \ (p(T/M) + \tau_i)} e^{-j2\pi \nu_i \tau_i}$$
$$= e^{j2\pi \frac{p}{M} \left(\frac{k_{\nu_i} + \kappa_{\nu_i}}{N} \right)}.$$

Therefore, from the above terms, the value of $h_{k,l}^{\text{ici}}[k', l']$ can be written as in (56), as shown on the previous page, where $\mathcal{G}^{\text{ici}}(\nu_i)$ and $\mathcal{F}^{\text{ici}}(\tau_i, \nu_i)$ denote the terms in the first and second square brackets. The value of $\mathcal{G}^{\text{ici}}(\nu_i)$ is the same as the one studied in (18) for ideal pulses case. Similar to the analysis of (17), $\mathcal{F}^{\text{ici}}(\tau_i, \nu_i)$ can be written as in (57), as shown on the previous page. Hence, by substituting (56) and (57) in (54), $y_{\text{ici}}[k, l]$ can be approximated as in (58), as shown on the previous page. From (58), we can easily see that it is non-zero only if the following conditions satisfied

$$p = [l - l_{\tau_i}]_M \text{ and } 0 \leq p \leq M - 1 - l_{\tau_i}.$$

These conditions are satisfied only if $l \geq l_{\tau_i}$ and $p = l - l_{\tau_i}$. Finally, with the conditions on l and p, $y_{\text{ici}}[k, l]$ can be obtained as in (59), as shown on the previous page, where $\beta_i(q)$ is defined in (25).

Analysis of $y_{\text{isi}}[k, l]$: Similar to $y_{\text{ici}}[k, l]$ in (53), $y_{\text{isi}}[k, l]$ can be expanded as in (60), as shown on the previous page. By substituting the value of $H_{n,m}[n-1, m']$ from (10), cross-ambiguity function in (23), and similar analysis of separating terms for $h_{k,l}^{\text{ici}}[k', l']$, the value of $h_{k,l}^{\text{isi}}[k', l']$ can be obtained as in (61), as shown on the previous page. Here, the summation n starts from 1 as the first symbol does not have previous symbol to experience ISI. Therefore, the value of $\mathcal{G}^{\text{isi}}(\nu_i)$ is equal to $\mathcal{G}^{\text{ici}}(\nu_i) - 1$. Using the value of $\mathcal{G}^{\text{isi}}(\nu_i)$, $y_{\text{isi}}[k, l]$ can be approximated as in (62), as shown at the top of this page.

Further, the expression in (62) can be approximated as in (63), as shown at the top of this page, by neglecting the signals $x[k', p]$ for which $k' \neq [k - k_{\nu_i} + q]_N, q \in [-N_i, N_i]$, as their coefficients are very small $(1/N)$ for practical values of N (e.g., $N = 64, 128$).

Now, (63) is non-zero only if the following conditions are satisfied

$$p = [l - l_{\tau_i}]_M \text{ and } M - l_{\tau_i} \leq p \leq M - 1.$$

These conditions are satisfied only if $l < l_{\tau_i}$ and $p = l - l_{\tau_i} + M$. With these conditions, the value of $y_{\text{isi}}[k, l]$ is written in (64).

Finally, by combining (59) and (64), as shown at the top of this page, the value of $y[k, l]$ in (52) can be obtained as in (24), which completes the proof.

ACKNOWLEDGEMENT

REFERENCES

[1] T. Wang, J. G. Proakis, E. Masry, and J. R. Zeidler, "Performance degradation of OFDM systems due to Doppler spreading," *IEEE Trans. Wireless Commun.*, vol. 5, no. 6, pp. 1422–1432, Jun. 2006.

[2] G. Matz, H. Bolcskei, and F. Hlawatsch, "Time-frequency foundations of communications: Concepts and tools," *IEEE Signal Process. Mag.*, vol. 30, no. 6, pp. 87–96, Nov. 2013.

[3] K. Liu, T. Kadous, and A. M. Sayeed, "Orthogonal time-frequency signaling over doubly dispersive channels," *IEEE Trans. Inf. Theory*, vol. 50, no. 11, pp. 2583–2603, Nov. 2004.

[4] T. Dean, M. Chowdhury, and A. Goldsmith, "A new modulation technique for Doppler compensation in frequency-dispersive channels," in *Proc. IEEE 28th Annu. Int. Symp. Pers., Indoor, Mobile Radio Commun. (PIMRC)*, Oct. 2017, pp. 1–7. [Online]. Available: https://web.stanford.edu/~trdean/papers/dcfdm.pdf

[5] R. Hadani *et al.*, "Orthogonal time frequency space modulation," in *Proc. IEEE Wireless Commun. Netw. Conf. (WCNC)*, San Francisco, CA, USA, Mar. 2017, pp. 1–6.

[6] R. Hadani and A. Monk. (2018). "OTFS: A new generation of modulation addressing the challenges of 5G." [Online]. Available: https://arxiv.org/pdf/1802.02623.pdf

[7] F. Hlawatsch and G. Matz, Eds., *Wireless Communications Over Rapidly Time-Varying Channels*. New York, NY, USA: Academic, 2011.

[8] W. C. Jakes, Jr., *Microwave Mobile Communications*. New York, NY, USA: Wiley, 1974.

[9] R. Hadani *et al.*, "Orthogonal time frequency space (OTFS) modulation for millimeter-wave communications systems," in *Proc. IEEE MTT-S Int. Microw. Symp. Dig.*, Honolulu, HI, USA, Jun. 2017, pp. 681–683.

[10] L. Li *et al.* (2017). "A simple two-stage equalizer with simplified orthogonal time frequency space modulation over rapidly time-varying channels." [Online]. Available: https://arxiv.org/abs/1709.02505

[11] A. Farhang, A. RezazadehReyhani, L. E. Doyle, and B. Farhang-Boroujeny. (2017). "Low complexity modem structure for OFDM-based orthogonal time frequency space modulation." [Online]. Available: https://arxiv.org/pdf/1710.00655.pdf

[12] A. RezazadehReyhani, A. Farhang, M. Ji, R. R. Chen, and B. Farhang-Boroujeny. (2017). "Analysis of discrete-time MIMO OFDM-based orthogonal time frequency space modulation." [Online]. Available: https://arxiv.org/pdf/1710.07900.pdf

[13] P. Som, T. Datta, N. Srinidhi, A. Chockalingam, and B. S. Rajan, "Low-complexity detection in large-dimension MIMO-ISI channels using graphical models," *IEEE J. Sel. Topics Signal Process.*, vol. 5, no. 8, pp. 1497–1511, Dec. 2011.

[14] P. Raviteja, K. T. Phan, Q. Jin, Y. Hong, and E. Viterbo, "Low-complexity iterative detection for orthogonal time frequency space modulation," in *Proc. IEEE Wireless Commun. Netw. Conf. (WCNC)*, Barcelona, Spain, Apr. 2018, pp. 1–6.

[15] W. Mecklenbräuker, "A tutorial on non-parametric bilinear time-frequency signal representations," in *Time and Frequency Representation of Signals and Systems*, vol. 309, G Longo and B. Picinbono, Eds. pp. 11–68, 1989.

[16] D. N. C. Tse and P. Viswanath, *Fundamentals of Wireless Communication*. Cambridge, U.K.: Cambridge Univ. Press, 2005.

[17] W. Kozek and A. F. Molisch, "Nonorthogonal pulseshapes for multicarrier communications in doubly dispersive channels," *IEEE J. Sel. Areas Commun.*, vol. 16, no. 8, pp. 1579–1589, Oct. 1998.

[18] G. Durisi, U. G. Schuster, H. Bolcskei, and S. Shamai (Shitz), "Non-coherent capacity of underspread fading channels," *IEEE Trans. Inf. Theory*, vol. 56, no. 1, pp. 367–395, Jan. 2010.

[19] M. Pretti, "A message-passing algorithm with damping," *J. Statist. Mech., Theory Exp.*, Nov. 2005, p. P11008.

[20] Y. Zhao and S.-G. Haggman, "Sensitivity to Doppler shift and carrier frequency errors in OFDM systems-the consequences and solutions," in *Proc. IEEE Veh. Technol. Conf.*, Atlanta, GA, USA, Apr. 1996, pp. 1564–1568.

[21] T. Yoo and A. Goldsmith, "Capacity of fading MIMO channels with channel estimation error," in *Proc. IEEE Int. Conf. Commun. (ICC)*, Paris, France, vol. 2, Jun. 2004, pp. 808–813.

[22] *Evolved Universal Terrestrial Radio Access (E-UTRA); Base Station (BS) Radio Transmission and Reception, Version 8.6.0*, document 3GPP TS 36.104, Jul. 2009.

Khoa T. Phan (M'17) received the B.Eng. degree (Hons.) in telecommunications from the University of New South Wales (UNSW), Sydney, NSW, Australia, in 2006, the M.Sc. degree in electrical engineering from the University of Alberta, Edmonton, AB, Canada, in 2008, the M.Sc. degree in electrical engineering from the California Institute of Technology (Caltech), Pasadena, CA, USA, in 2009, and the Ph.D. degree in electrical engineering from McGill University, Montreal, QC, Canada, in 2017. He was a Researcher with the Electrical Engineering Department, University of California at Los Angeles, Los Angeles, CA, USA, from 2009 to 2011. Since 2017, he has been a Research Fellow with the Electrical and Computer Systems Engineering Department, Monash University, Clayton, VIC, Australia. His research interests include design, control, optimization, and operation of 5G mobile communications networks with contemporary applications in the Internet of Things, machine-type communications, smart grids, and cloud computing.

Dr. Phan received several prestigious fellowships, including the Australian Development Scholarship from UNSW, the Alberta Ingenuity Fund Student Fellowship and the iCORE Graduate Student Award from the University of Alberta, the Atwood Fellowship from Caltech, the McGill Doctoral Engineering Award and the SRTelecom Doctoral Award from McGill University.

Yi Hong (S'00–M'05–SM'10) received the Ph.D. degree in electrical engineering and telecommunications from the University of New South Wales, Sydney. She is currently a Senior Lecturer with the Department of Electrical and Computer Systems Engineering, Monash University, Clayton, VIC, Australia. She currently serves on the Australian Research Council College of Experts from 2018 to 2020. She received the NICTA-ACoRN Earlier Career Researcher Award at the Australian Communication Theory Workshop, Adelaide, Australia, in 2007.

Her research interests include communication theory and coding and information theory with applications to telecommunication engineering. She was a technical program committee member for many IEEE leading conferences. She was the General Co-Chair of the IEEE Information Theory Workshop 2014, Hobart; the Technical Program Committee Chair of the Australian Communications Theory Workshop 2011, Melbourne; and the Publicity Chair at the IEEE Information Theory Workshop 2009, Sicily. She was an Associate Editor of IEEE WIRELESS COMMUNICATION LETTERS and *Transactions on Emerging Telecommunications Technologies*.

P. Raviteja (S'15) received the B.Tech. degree in electronics and communication from JNTU University, Hyderabad, India, in 2012, and the M.E. degree in telecommunications from the Indian Institute of Science, Bengaluru, India, in 2014. He is currently pursuing the Ph.D. degree with the Department of Electrical and Computer Systems Engineering, Monash University, Australia. From 2014 to 2015, he was with Qualcomm India Private Limited, Bengaluru, where he was involved in WLAN systems design. His current research interests include new waveform designs for next generation wireless systems. He was a recipient of the Prof. S. V. C. Aiya Medal from the Indian Institute of Science, in 2014.

Emanuele Viterbo (M'95–SM'04–F'11) received the Ph.D. degree in electrical engineering from the Politecnico di Torino, Torino, Italy, in 1995. He is currently a Professor with the ECSE Department and an Associate Dean in Graduate Research at Monash University, Clayton, VIC, Australia. From 1990 to 1992, he was a Patent Examiner in dynamic recording and error-control coding with the European Patent Office, The Hague, The Netherlands. From 1995 to 1997, he held a post-doctoral position at the Dipartimento di Elettronica, Politecnico di Torino. From 1997 to 1998, he was a Post-Doctoral Research Fellow at the Information Sciences Research Center of AT&T Research, Florham Park, NJ, USA. From 1998 to 2005, he was an Assistant Professor and then an Associate Professor with the Dipartimento di Elettronica, Politecnico di Torino. From 2006 to 2009, he was a Full Professor with DEIS, University of Calabria, Italy. His main research interests are in lattice codes for the Gaussian and fading channels, algebraic coding theory, algebraic space–time coding, digital terrestrial television broadcasting, digital magnetic recording, and irregular sampling. He received the NATO Advanced Fellowship in 1997 from the Italian National Research Council. He has been an ISI Highly Cited Researcher since 2009. He is an Associate Editor of the IEEE TRANSACTIONS ON INFORMATION THEORY, *European Transactions on Telecommunications*, and *Journal of Communications and Networks*, and a Guest Editor of the IEEE JOURNAL OF SELECTED TOPICS IN SIGNAL PROCESSING Special Issue on Managing Complexity in Multiuser MIMO Systems.