

Analysis of Discrete-Time MIMO OFDM-Based Orthogonal Time Frequency Space Modulation

Ahmad RezazadehReyhani*, Arman Farhang[†], Mingyue Ji*, Rong Rong Chen* and Behrouz Farhang-Boroujeny*

*ECE Department, University of Utah, USA, [†]Trinity College Dublin, Ireland.

Email: {ahmad.rezazadeh, mingyue.ji, rchen, farhang}@utah.edu, farhanga@tcd.ie

Abstract—Orthogonal Time Frequency Space (OTFS) is a novel modulation scheme designed in the Doppler-delay domain to fully exploit time and frequency diversity of general time-varying channels. In this paper, we present a novel discrete-time analysis of OFDM-based OTFS transceiver with a concise and vectorized input-output relationship that clearly characterizes the contribution of each underlying signal processing block in such systems. When adopting cyclic prefix in the time domain, our analysis reveals that the proposed MIMO OTFS and OFDM systems have the same ergodic capacity despite the well-known fact that the former has great advantages in low-complexity receiver design for high Doppler channels. The proposed discrete-time vectorized formulation is applicable to general fast fading channels with arbitrary window functions. It also enables practical low-complexity receiver design for which such a concise formulation of the input-output relationship is of great benefits.

Index Terms—OTFS, OFDM, MIMO, Ergodic Capacity

I. INTRODUCTION

Orthogonal time frequency space (OTFS) modulation has recently been proposed as an effective waveform that takes advantage of the time diversity (i.e., variation of channel with time) to improve on the reliability of wireless links [1]. In OTFS, the transmit data symbols are treated as values of the grid points in a Doppler-delay space. A transformation step takes each data symbol and spread it over the entire space of time-frequency points. The result of this transformation is then passed to a multicarrier system for modulation and transmission. In this way, all data symbols are equally affected by the channel frequency selectivity and time diversity and, as a result, the time-varying channel, within a good approximation, converts to a unified time-invariant impulse response for all the data symbols.

To recover the transmitted (Doppler-delay space) data symbols, at the receiver, the respective multicarrier demodulation followed by a transformation takes the received signal back to the Doppler-delay space. In [1], it is proposed that an inverse symplectic finite Fourier transform (SFFT⁻¹) be used for transformation from the Doppler-delay space to the time-frequency space, at the transmitter, and the corresponding symplectic finite Fourier transform (SFFT) be used for the reverse operation, at the receiver.

In the above set-up, the equivalent channel that connects the transmitted data symbols and the received signal samples, both in the Doppler-delay space, is modeled by a time-invariant two dimensional (2-D) impulse response. Soft detectors/equalizers, e.g., some extensions to those presented in [2], may thus be

used for near-optimal recovery of the transmitted information. In order to simplify such detectors, [1] has proposed that proper windows should be applied to the time-frequency signals at both the transmitter and receiver sides to improve on the sparsity of the OTFS channel.

The transformation of data symbols from the Doppler-delay space to time-frequency at the transmitter and the corresponding inverse transformation at the receiver, clearly, leads to full diversity gain across both time and frequency. In multiple-input multiple-output (MIMO) channels, the space diversity gain will naturally be present because in a MIMO setup any signal going out of each antenna reaches all the receiver antennas with statistically similar gains.

The goal of this paper is to examine the details of OTFS in terms of the reliability of transmission brought up as a result of the addition of time diversity in the modulator. To allow simple derivations, OFDM is used for multicarrier transmission of time-frequency signals in OTFS. To this end, we present a novel discrete-time end-to-end formulation of an OFDM-based OTFS setup. Such a formulation provides a concise representation of the effect of each signal processing unit on the input-output relationship of an MIMO OFDM-based OTFS system. Based on this formulation, we show that the proposed MIMO OFDM-based OTFS system achieves the same ergodic capacity as that of an OFDM system, under the assumption that channel state information is known at the receiver and most importantly, with the use of cyclic prefix, a block of OTFS transmission can be implemented as N consecutive, non-interfering OFDM transmissions. Since this capacity analysis is derived assuming perfect channel knowledge, optimal receiver, and an infinite code length, it does not violate the known fact that in practical systems with higher Dopplers, OTFS outperforms OFDM and has lower receiver complexity due to simpler channel estimation and equalization in the Doppler-delay domain.

This work is inspired by the landmark paper [1], which was the first to propose a continuous-time formulation of a single-antenna OTFS modulation. Recently, [3] presents a discrete-time formulation of a single-antenna OTFS system by sampling the continuous-time channel for the case when the transmit and receive window functions are rectangular. In [4], a matrix-form discrete-time formulation of a single-antenna transceiver with rectangular windows is presented. However, this formulation is mostly based on two-dimensional signal matrices instead of vectors. Moreover, [5] studies a single-

antenna transceiver when the transmit and receive window functions are separable. In comparison, our formulation here uses a general form of window functions in an MIMO OFDM-based OTFS setup and presents a vectorized formulation of signals which is amicable to analytical analysis and practical implementation of MIMO OTFS systems.

The rest of this paper is organized as follows. In Section II, we summarize basic properties of the Kronecker products for which most of the derivations in this paper are based upon. In Sections III and IV, we describe the fundamentals of OTFS and present a vectorized formulation for the input-output relationship with general windowing functions for the single antenna case. Extensions of the derivations to the MIMO case are presented in Section V. The ergodic capacity analysis of the OFDM-based OTFS is developed in Section VI. The paper is concluded in Section VII.

Notations: Matrices, vectors, and scalar quantities are denoted by boldface uppercase, boldface lowercase, and normal letters, respectively. \mathbf{I}_N and $\mathbf{0}_N$ are the identity and null matrices of size N , respectively, and j is used to denote $\sqrt{-1}$. The superscripts $(\cdot)^T$, $(\cdot)^H$, and $(\cdot)^*$ indicate transpose, conjugate transpose, and conjugate operations, respectively. The N -point discrete Fourier transform (DFT) matrix is represented by \mathbf{F}_N and is assumed to be normalized such that $\mathbf{F}_N \mathbf{F}_N^H = \mathbf{I}_N$. $\mathbf{A} \in \mathcal{C}^{M \times N}$ denotes a $M \times N$ matrix with complex-valued elements and $a_{m,n}$ represents element of matrix \mathbf{A} at m -th row and n -th column.

II. KRONECKER PRODUCT PROPERTIES

In this paper, we make frequent use of the *Kronecker* product to arrive at a concise and elegant representation of the input-output relationship in OTFS systems. Hence, in this section we first review some of the relevant *Kronecker* properties. For an $m \times n$ matrix \mathbf{A} and a $p \times q$ matrix \mathbf{B} , the *Kronecker* product $\mathbf{A} \otimes \mathbf{B}$ is a $mp \times nq$ block matrix

$$\mathbf{A} \otimes \mathbf{B} = \begin{bmatrix} a_{11}\mathbf{B} & \cdots & a_{1n}\mathbf{B} \\ \vdots & \ddots & \vdots \\ a_{m1}\mathbf{B} & \cdots & a_{mn}\mathbf{B} \end{bmatrix}. \quad (1)$$

The *Kronecker* product is an associative but non-commutative operator

$$\mathbf{A} \otimes \mathbf{B} \neq \mathbf{B} \otimes \mathbf{A}, \quad (2)$$

$$(\mathbf{A} \otimes \mathbf{B}) \otimes \mathbf{C} = \mathbf{A} \otimes (\mathbf{B} \otimes \mathbf{C}). \quad (3)$$

The mixed-product property of the *Kronecker* product is

$$(\mathbf{A} \otimes \mathbf{B})(\mathbf{C} \otimes \mathbf{D}) = (\mathbf{AC}) \otimes (\mathbf{BD}). \quad (4)$$

Moreover, under transpose and Hermitian operations, order of the *Kronecker* product operands does not change

$$(\mathbf{A} \otimes \mathbf{B})^H = \mathbf{A}^H \otimes \mathbf{B}^H. \quad (5)$$

Finally, consider a matrix equation $\mathbf{AXB} = \mathbf{C}$, where \mathbf{A} , \mathbf{X} , \mathbf{B} , and \mathbf{C} are proper size matrices. We can rewrite this equation as

$$(\mathbf{B}^T \otimes \mathbf{A}) \text{vec}(\mathbf{X}) = \text{vec}(\mathbf{C}), \quad (6)$$

where $\text{vec}(\mathbf{X})$ denotes the vectorized version of the matrix \mathbf{X} formed by stacking the columns of \mathbf{X} into a single column vector.

III. OTFS WAVEFORM

An OTFS transmitter combines a set of complex-valued data symbols $\{d_{m,n}; m = 0, 1, \dots, M-1; n = 0, 1, \dots, N-1\}$ in the Doppler-delay domain to construct an OTFS signal. First, the transmitter maps data symbols on the Doppler-delay lattice, $\{d_{m,n}\}$, to a lattice in the time-frequency domain through an inverse symplectic finite Fourier transform operation (SFFT⁻¹)

$$x_{k,l} = \frac{1}{\sqrt{MN}} \sum_{m=0}^{M-1} \sum_{n=0}^{N-1} d_{m,n} e^{-j2\pi(\frac{mk}{M} - \frac{nl}{N})}, \quad (7)$$

where $k = 0, 1, \dots, M-1$ and $l = 0, 1, \dots, N-1$. Let $\mathbf{D} \in \mathcal{C}^{M \times N}$ denote the data matrix which contains elements $\{d_{m,n}\}$. A closer look into (7) reveals that the SFFT⁻¹ transform of \mathbf{D} can be obtained by applying an M -point discrete Fourier transform (DFT) and an N -point inverse DFT (IDFT) to the columns and rows of the matrix \mathbf{D} , respectively. Accordingly, (7) can be written in a compact form as

$$\mathbf{X} = \mathbf{F}_M \mathbf{D} \mathbf{F}_N^H, \quad (8)$$

where $\mathbf{X} \in \mathcal{C}^{M \times N}$.

Subsequently, the transmitter applies a transmit window, $\tilde{u}_{k,l}$, to the time-frequency signal as

$$\tilde{x}_{k,l} = x_{k,l} \tilde{u}_{k,l}. \quad (9)$$

Finally, the transmitter packs the time-frequency signal, $\tilde{x}_{k,l}$, to the transmitted signal, $s[i]$, using a set of time-frequency basis functions, $g_{k,l}[i]$,

$$s[i] = \sum_{k=0}^{M-1} \sum_{l=0}^{N-1} \tilde{x}_{k,l} g_{k,l}[i]. \quad (10)$$

The received signal samples after transmission over a linear time varying (LTV) channel can be obtained as

$$r[i] = \sum_{l=0}^{L-1} h[i, l] s[i-l] + w[i], \quad (11)$$

where $h[i, l]$ is the instantaneous channel impulse response with length L at time instant l , and $w[i]$ is the additive channel noise.

To obtain an estimate of $d_{m,n}$ from the received signal $r[i]$, first, we map $r[i]$ to the time-frequency lattice by projecting it onto another set of time-frequency basis functions $f_{k,l}[i]$ as

$$\tilde{y}_{k,l} = \langle f_{k,l}[i], r[i] \rangle. \quad (12)$$

The sets of transmit and receive basis functions are designed to satisfy the perfect reconstruction criterion

$$\langle f_{k,l}[i], g_{k',l'}[i] \rangle = \delta(k-k') \delta(l-l'). \quad (13)$$

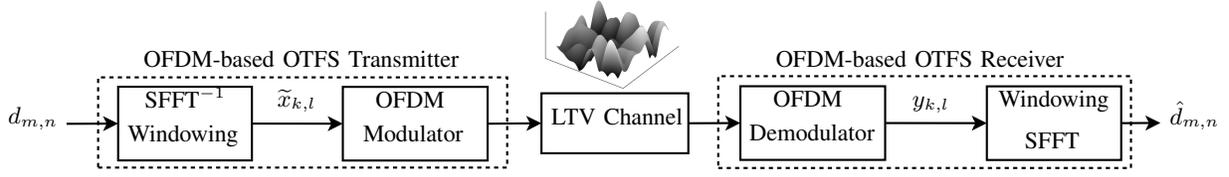


Fig. 1. OTFS transmitter and receiver structure.

Then, the receiver performs the windowing operation through multiplication of the receive window function coefficients $\tilde{v}_{k,l}$ to the time-frequency samples as

$$y_{k,l} = \tilde{y}_{k,l} \tilde{v}_{k,l}. \quad (14)$$

Finally, signals on the time-frequency lattice are transformed back to the Doppler-delay domain by applying SFFT as

$$\hat{d}_{m,n} = \frac{1}{\sqrt{MN}} \sum_{k=0}^{M-1} \sum_{l=0}^{N-1} y_{k,l} e^{j2\pi(\frac{mk}{M} - \frac{nl}{N})}. \quad (15)$$

Similar to SFFT^{-1} , the SFFT transform can be split into an M -point IDFT and an N -point DFT on rows and columns of its operand. Thus, (15) can be rearranged as

$$\hat{\mathbf{D}} = \mathbf{F}_M^H \mathbf{Y} \mathbf{F}_N, \quad (16)$$

where $\mathbf{Y} \in \mathcal{C}^{M \times N}$ and $\hat{\mathbf{D}} \in \mathcal{C}^{M \times N}$ contain elements $\{y_{k,l}\}$ and $\{\hat{d}_{m,n}\}$, respectively.

Typically, an OFDM-based OTFS modulation with M subcarriers and a cyclic prefix (CP) of length M_{cp} utilizes the following transmit and receive time-frequency basis functions

$$g_{k,0}[i] = \frac{1}{\sqrt{M}} e^{j\frac{2\pi}{M}ki}, \quad -M_{\text{cp}} \leq i < M$$

$$g_{k,l}[i] = g_{k,0}[i - (M + M_{\text{cp}})l] \quad (17)$$

$$f_{k,0}[i] = \frac{1}{\sqrt{M}} e^{-j\frac{2\pi}{M}ki}, \quad 0 \leq i < M$$

$$f_{k,l}[i] = f_{k,0}[i - (M + M_{\text{cp}})l]. \quad (18)$$

IV. OFDM-BASED OTFS: SINGLE ANTENNA

In this section, we present a vectorized formulation for the OFDM-based OTFS transceiver. The first step of the OTFS modulation in (8), i.e. the SFFT^{-1} operation, using (6) can be rearranged as

$$\mathbf{x} = (\mathbf{F}_N^H \otimes \mathbf{F}_M) \mathbf{d}, \quad (19)$$

where $\mathbf{x} = \text{vec}(\mathbf{X})$ and $\mathbf{d} = \text{vec}(\mathbf{D})$. Then, the vector \mathbf{x} is partitioned into blocks of length M , denoted by \mathbf{x}_n , $n = 0, 1, \dots, N-1$ and each block is multiplied to the corresponding transmit window as

$$\tilde{\mathbf{x}}_n = \mathbf{U}_n \mathbf{x}_n, \quad (20)$$

where the transmit window $\mathbf{U}_n \in \mathcal{C}^{M \times M}$ is a diagonal matrix consisting of diagonal elements $\tilde{u}_{k,n}$. Stacking the results as $\tilde{\mathbf{x}}$, we have

$$\tilde{\mathbf{x}} = \mathbf{U} \mathbf{x} \quad (21)$$

where $\mathbf{U} \in \mathcal{C}^{MN \times MN}$ is a diagonal matrix whose $(lM+k)$ -th diagonal element is $\tilde{u}_{k,l}$.

Each partition of $\tilde{\mathbf{x}}$, i.e. $\tilde{\mathbf{x}}_n$, is fed into an OFDM modulator. The OFDM modulator multiplies an IDFT matrix \mathbf{F}_M^H to each block and then appends a CP to each block as

$$\mathbf{s}_n = \mathbf{F}_M^H \tilde{\mathbf{x}}_n, \quad (22)$$

$$\tilde{\mathbf{s}}_n = \mathbf{A}_{\text{cp}} \mathbf{s}_n, \quad (23)$$

where M_{cp} is the CP length, $\mathbf{A}_{\text{cp}} = [\mathbf{G}_{\text{cp}}, \mathbf{I}_M]^T$ is the CP addition matrix and $\mathbf{G}_{\text{cp}} \in \mathcal{C}^{M \times M_{\text{cp}}}$ includes the last M_{cp} columns of the identity matrix \mathbf{I}_M . Note that \mathbf{A}_{cp} appends last M_{cp} samples of each OFDM block to its beginning. Stacking the results as vectors, (22) and (23) can be written as

$$\mathbf{s} = (\mathbf{I}_N \otimes \mathbf{F}_M^H) \tilde{\mathbf{x}}, \quad (24)$$

$$\tilde{\mathbf{s}} = (\mathbf{I}_N \otimes \mathbf{A}_{\text{cp}}) \mathbf{s}. \quad (25)$$

After the signal $\tilde{\mathbf{s}}$ is passed through the LTV channel, the received signal can be written as

$$\tilde{\mathbf{r}} = \mathbf{H} \tilde{\mathbf{s}} + \mathbf{w}, \quad (26)$$

where $\mathbf{H} \in \mathcal{C}^{N(M+M_{\text{cp}}) \times N(M+M_{\text{cp}})}$ is the channel impulse response matrix, and \mathbf{w} is the channel additive noise vector. The receiver partitions the received vector into N blocks, $\tilde{\mathbf{r}}_n$ and removes CP from each received block as

$$\mathbf{r}_n = \mathbf{R}_{\text{cp}} \tilde{\mathbf{r}}_n, \quad (27)$$

where $\mathbf{R}_{\text{cp}} \in \mathcal{C}^{M \times (M+M_{\text{cp}})}$ is the CP removal matrix which can be obtained by removing the first M_{cp} rows of $\mathbf{I}_{M+M_{\text{cp}}}$.

Stacking the output vectors \mathbf{r}_n into a length NM vector \mathbf{r} , (25) to (27) can be written as

$$\mathbf{r} = (\mathbf{I}_N \otimes \mathbf{R}_{\text{cp}}) \mathbf{H} (\mathbf{I}_N \otimes \mathbf{A}_{\text{cp}}) \mathbf{s} + \mathbf{w}$$

$$= \tilde{\mathbf{H}} \mathbf{s} + \mathbf{w}, \quad (28)$$

where $\tilde{\mathbf{H}} = (\mathbf{I}_N \otimes \mathbf{R}_{\text{cp}}) \mathbf{H} (\mathbf{I}_N \otimes \mathbf{A}_{\text{cp}})$ is an $MN \times MN$ block diagonal matrix

$$\tilde{\mathbf{H}} = \begin{bmatrix} \tilde{\mathbf{H}}_0 & \mathbf{0}_M & \cdots & \mathbf{0}_M \\ \mathbf{0}_M & \tilde{\mathbf{H}}_1 & \cdots & \mathbf{0}_M \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{0}_M & \mathbf{0}_M & \cdots & \tilde{\mathbf{H}}_{N-1} \end{bmatrix} \quad (29)$$

and $\tilde{\mathbf{H}}_n$ is the channel impulse response matrix of the n -th OFDM symbol.

The received signal is fed into an OFDM demodulator and the output signal is

$$\tilde{\mathbf{y}} = (\mathbf{I}_N \otimes \mathbf{F}_M) \mathbf{r}, \quad (30)$$

where $\tilde{\mathbf{y}}$ is a length MN time-frequency signal vector. Finally, after performing the receiver windowing operation, the time-frequency signal is mapped to the Doppler-delay domain and an estimate of the transmitted vector can be obtained as

$$\hat{\mathbf{d}} = (\mathbf{F}_N \otimes \mathbf{F}_M^H) \mathbf{y}, \quad (31)$$

where $\mathbf{y} = \mathbf{V}\tilde{\mathbf{y}}$, and the receive window $\mathbf{V} \in \mathcal{C}^{MN \times MN}$ is a diagonal matrix defined similarly to that of the transmit window.

We note that equations (19) to (31) describe the end-to-end relationship for a general OFDM-based OTFS setup which is summarized in a compact form shown in (32), at the top of the next page. It clearly demonstrates the contribution of each signal processing step of an OFDM-based OTFS setup from Doppler-delay domain to Doppler-delay domain where each step corresponds to one section of the transceiver block diagram shown in Fig. 1. It is clear that this compact form of (32) will allow simplified analysis and implementation of OFDM-based OTFS systems. Next, we will present how these equations simplify for some special setup considered below.

A. Separable Windows

In general, transmit and receive window functions can be arbitrary functions that are designed for various purposes. These include, for instance, zero-forcing or minimum mean squared error equalization, or shortening the channel response in Doppler-delay domain. For a wide range of applications, however, one can treat windowing across time axis and across frequency axis separately because of the independency between channel delay spread and Doppler spread. This separate windowing can lead to simplified transmitter and receiver structures as shown in [5].

Assume that the transmit and receive windows are separable functions, i.e. $\tilde{u}_{k,l} = a_l b_k$ and $\tilde{v}_{k,l} = p_l q_k$. We can write transmit and receive window matrices as $\mathbf{U} = \mathbf{A} \otimes \mathbf{B}$ and $\mathbf{V} = \mathbf{P} \otimes \mathbf{Q}$ where $\mathbf{A} \in \mathcal{C}^{N \times N}$, $\mathbf{P} \in \mathcal{C}^{N \times N}$, $\mathbf{B} \in \mathcal{C}^{M \times M}$, and $\mathbf{Q} \in \mathcal{C}^{M \times M}$ are diagonal matrices with elements $\{a_l\}$, $\{p_l\}$, $\{b_k\}$, and $\{q_k\}$, respectively. Substituting these window forms in (21) to (31) and after simplifications we have

$$\hat{\mathbf{d}} = (\mathbf{I}_N \otimes \mathbf{F}_M^H \mathbf{Q} \mathbf{F}_M) (\mathbf{F}_N \otimes \mathbf{I}_M) (\mathbf{P} \otimes \mathbf{I}_M) \tilde{\mathbf{H}} (\mathbf{A} \otimes \mathbf{I}_M) (\mathbf{F}_N^H \otimes \mathbf{I}_M) (\mathbf{I}_N \otimes \mathbf{F}_M^H \mathbf{B} \mathbf{F}_M) \mathbf{d} + \hat{\mathbf{w}}, \quad (33)$$

where $\hat{\mathbf{w}} = (\mathbf{F}_N \otimes \mathbf{F}_M^H) \mathbf{V} (\mathbf{I}_N \otimes \mathbf{F}_M) \mathbf{w}$. We note that (33) can be used for the simplification of OFDM-based OTFS transceiver.

B. Rectangular Window

When rectangular windows are used as the transmitter and receiver window functions, i.e. $\mathbf{U} = \mathbf{V} = \mathbf{I}_{MN}$, substituting (21) to (30) in (31), it simplifies to

$$\hat{\mathbf{d}} = (\mathbf{F}_N \otimes \mathbf{I}_M) \tilde{\mathbf{H}} (\mathbf{F}_N^H \otimes \mathbf{I}_M) \mathbf{d} + \hat{\mathbf{w}}, \quad (34)$$

where we used the mixed-product property of Kronecker product $(\mathbf{A} \otimes \mathbf{B})(\mathbf{C} \otimes \mathbf{D}) = (\mathbf{A}\mathbf{C}) \otimes (\mathbf{B}\mathbf{D})$.

C. Frequency Domain Representation

Let $\tilde{\mathbf{H}}_f$ denote the block diagonal frequency domain channel impulse response matrix. We see that $\tilde{\mathbf{H}}_f = (\mathbf{I}_N \otimes \mathbf{F}_M) \tilde{\mathbf{H}} (\mathbf{I}_N \otimes \mathbf{F}_M^H)$. Solving this for $\tilde{\mathbf{H}}$ we obtain $\tilde{\mathbf{H}} = (\mathbf{I}_N \otimes \mathbf{F}_M^H) \tilde{\mathbf{H}}_f (\mathbf{I}_N \otimes \mathbf{F}_M)$. Substituting $\tilde{\mathbf{H}}$ in (32), after simplification it boils down to

$$\hat{\mathbf{d}} = (\mathbf{F}_N \otimes \mathbf{F}_M^H) \mathbf{V} \tilde{\mathbf{H}}_f \mathbf{U} (\mathbf{F}_N^H \otimes \mathbf{F}_M) \mathbf{d} + \hat{\mathbf{w}}, \quad (35)$$

which shows that transmit and receive windows directly change the frequency domain channel impulse response matrix. In other words, the effective frequency domain channel is $\mathbf{V} \tilde{\mathbf{H}}_f \mathbf{U}$.

When the LTV channel varies slowly such that it is approximately invariant over each OFDM symbol, then the matrices $\{\tilde{\mathbf{H}}_n\}$ become circulant. Accordingly, $\tilde{\mathbf{H}}_f$ is an $MN \times MN$ diagonal matrix that contains N diagonal channel frequency response matrices, $\tilde{\mathbf{H}}_{f,n} = \mathbf{F}_M^H \tilde{\mathbf{H}}_n \mathbf{F}_M$. In this case, (35) is equivalent to the 2D circular convolution of the transmit data matrix with the effective 2D channel impulse response matrix.

V. OFDM-BASED OTFS: MIMO

Consider a $n_t \times n_r$ MIMO setup that utilizes OFDM-based OTFS modulation. Let us stack transmit data matrices of all antennas to form an $Mn_t \times N$ matrix $\bar{\mathbf{D}}$ as

$$\bar{\mathbf{D}} = \begin{bmatrix} \mathbf{D}^0 \\ \mathbf{D}^1 \\ \vdots \\ \mathbf{D}^{n_t-1} \end{bmatrix}, \quad (36)$$

where \mathbf{D}^t is the $M \times N$ data matrix of t -th antenna. Applying SFFT⁻¹ to each sub-matrix \mathbf{D}^t , we obtain

$$\bar{\mathbf{X}} = (\mathbf{I}_{n_t} \otimes \mathbf{F}_M) \bar{\mathbf{D}} \mathbf{F}_N^H, \quad (37)$$

where \mathbf{F}_M is repeated as $\mathbf{I}_{n_t} \otimes \mathbf{F}_M$ to cover all data matrices of all antennas. Using (6), (37) can be written as

$$\bar{\mathbf{x}} = (\mathbf{F}_N^H \otimes \mathbf{I}_{n_t} \otimes \mathbf{F}_M) \bar{\mathbf{d}}, \quad (38)$$

where $\bar{\mathbf{d}}$ is the vectorized version of $\bar{\mathbf{D}}$ defined as follows.

$$\bar{\mathbf{d}} = \begin{bmatrix} \mathbf{d}_0 \\ \mathbf{d}_1 \\ \vdots \\ \mathbf{d}_{N-1} \end{bmatrix}, \quad \mathbf{d}_n = \begin{bmatrix} \mathbf{d}_n^0 \\ \mathbf{d}_n^1 \\ \vdots \\ \mathbf{d}_n^{n_t-1} \end{bmatrix}, \quad \mathbf{d}_n^t = \begin{bmatrix} d_{0,n}^t \\ d_{1,n}^t \\ \vdots \\ d_{M-1,n}^t \end{bmatrix}, \quad (39)$$

and $d_{m,n}^t$; $t = 0, 1, \dots, n_t - 1$ is the data symbol of t -th transmit antenna.

After partitioning $\bar{\mathbf{x}}$ to N blocks, $\bar{\mathbf{x}}_n$, the window function is multiplied to each block. Assuming that all transmit antennas use the same window function, we get

$$\check{\mathbf{x}}_n = \bar{\mathbf{U}}_n \bar{\mathbf{x}}_n, \quad (40)$$

where $\bar{\mathbf{U}}_n = \mathbf{I}_{n_t} \otimes \mathbf{U}_n$. We stack these results to obtain

$$\check{\mathbf{x}} = \bar{\mathbf{U}} \bar{\mathbf{x}}, \quad (41)$$

$$\hat{\mathbf{d}} = \underbrace{(\mathbf{F}_N \otimes \mathbf{F}_M^H)}_{\text{SFFT}} \underbrace{\mathbf{V}}_{\text{Window}} \underbrace{(\mathbf{I}_N \otimes \mathbf{F}_M)}_{\text{OFDM Demod.}} \underbrace{(\mathbf{I}_N \otimes \mathbf{R}_{\text{cp}})}_{\text{Remove CP}} \underbrace{\mathbf{H}}_{\text{Channel}} \underbrace{(\mathbf{I}_N \otimes \mathbf{A}_{\text{cp}})}_{\text{Add CP}} \underbrace{(\mathbf{I}_N \otimes \mathbf{F}_M^H)}_{\text{OFDM Mod.}} \underbrace{\mathbf{U}}_{\text{Window}} \underbrace{(\mathbf{F}_N^H \otimes \mathbf{F}_M)}_{\text{Inv. SFFT}} \mathbf{d} + \hat{\mathbf{w}}. \quad (32)$$

where

$$\bar{\mathbf{U}} = \begin{bmatrix} \bar{\mathbf{U}}_0 & \mathbf{0}_{Mn_t} & \cdots & \mathbf{0}_{Mn_t} \\ \mathbf{0}_{Mn_t} & \bar{\mathbf{U}}_1 & \cdots & \mathbf{0}_{Mn_t} \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{0}_{Mn_t} & \mathbf{0}_{Mn_t} & \cdots & \bar{\mathbf{U}}_{N-1} \end{bmatrix}. \quad (42)$$

The OFDM modulator transforms the time-frequency domain signal of each antenna to the time domain signal as

$$\mathbf{s}_n = (\mathbf{I}_{n_t} \otimes \mathbf{F}_M^H) \check{\mathbf{x}}_n, \quad (43)$$

$$\mathbf{s} = (\mathbf{I}_{Nn_t} \otimes \mathbf{F}_M^H) \check{\mathbf{x}}. \quad (44)$$

The transmitter appends the CP to the transmit signal and sends the result through an LTV channel. At the receiver, after removing the CP, the received signal can be written in a compact form as

$$\begin{aligned} \mathbf{r}_n &= (\mathbf{I}_{n_r} \otimes \mathbf{R}_{\text{cp}}) \mathbf{H}_n (\mathbf{I}_{n_t} \otimes \mathbf{A}_{\text{cp}}) \mathbf{s}_n + \mathbf{w}_n \\ &= \bar{\mathbf{H}}_n \mathbf{s}_n + \mathbf{w}_n, \end{aligned} \quad (45)$$

where $\bar{\mathbf{H}}_n = (\mathbf{I}_{n_r} \otimes \mathbf{R}_{\text{cp}}) \mathbf{H}_n (\mathbf{I}_{n_t} \otimes \mathbf{A}_{\text{cp}})$ can be expanded as

$$\bar{\mathbf{H}}_n = \begin{bmatrix} \bar{\mathbf{H}}_n^{0,0} & \bar{\mathbf{H}}_n^{0,1} & \cdots & \bar{\mathbf{H}}_n^{0,n_t-1} \\ \bar{\mathbf{H}}_n^{1,0} & \bar{\mathbf{H}}_n^{1,1} & \cdots & \bar{\mathbf{H}}_n^{1,n_t-1} \\ \vdots & \vdots & \ddots & \vdots \\ \bar{\mathbf{H}}_n^{n_r-1,0} & \bar{\mathbf{H}}_n^{n_r-1,1} & \cdots & \bar{\mathbf{H}}_n^{n_r-1,n_t-1} \end{bmatrix}. \quad (46)$$

Here, $\bar{\mathbf{H}}_n^{r,t}$ is the n -th channel impulse response matrix between t -th transmit and r -th receive antenna.

We stack the output vectors to obtain

$$\mathbf{r} = \bar{\mathbf{H}} \mathbf{s} + \mathbf{w}, \quad (47)$$

where

$$\bar{\mathbf{H}} = \begin{bmatrix} \bar{\mathbf{H}}_0 & \mathbf{0}_{Mn_r \times Mn_t} & \cdots & \mathbf{0}_{Mn_r \times Mn_t} \\ \mathbf{0}_{Mn_r \times Mn_t} & \bar{\mathbf{H}}_1 & \cdots & \mathbf{0}_{Mn_r \times Mn_t} \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{0}_{Mn_r \times Mn_t} & \mathbf{0}_{Mn_r \times Mn_t} & \cdots & \bar{\mathbf{H}}_{N-1} \end{bmatrix}. \quad (48)$$

The received signal is fed to an OFDM demodulator and the output signal is

$$\check{\mathbf{y}} = (\mathbf{I}_{Nn_r} \otimes \mathbf{F}_M) \mathbf{r}, \quad (49)$$

where $\check{\mathbf{y}}$ is the length MNn_r time-frequency signal vector. Finally, after multiplying by the receiver window, the time-frequency signal is mapped to the Doppler-delay domain and the estimate of the transmitted vector can be obtained as

$$\hat{\mathbf{d}} = (\mathbf{F}_N \otimes \mathbf{I}_{n_r} \otimes \mathbf{F}_M^H) \bar{\mathbf{V}} \check{\mathbf{y}}, \quad (50)$$

where

$$\bar{\mathbf{V}} = \begin{bmatrix} \bar{\mathbf{V}}_0 & \mathbf{0}_{Mn_r} & \cdots & \mathbf{0}_{Mn_r} \\ \mathbf{0}_{Mn_r} & \bar{\mathbf{V}}_1 & \cdots & \mathbf{0}_{Mn_r} \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{0}_{Mn_r} & \mathbf{0}_{Mn_r} & \cdots & \bar{\mathbf{V}}_{N-1} \end{bmatrix} \quad (51)$$

and $\bar{\mathbf{V}}_n = \mathbf{I}_{n_r} \otimes \tilde{\mathbf{V}}_n$. Using (38) to (50), it is straightforward to find a similar end-to-end vectorized relationship for the general MIMO OFDM-based OTFS setup to that of (32).

VI. CAPACITY ANALYSIS OF MIMO OFDM-BASED OTFS

In this section, we examine the ergodic capacity of MIMO-OFDM and MIMO-OFDM-based OTFS systems. While the analysis here applies to general time-varying channels, it is important to note that due to the use of cyclic prefix, the transmission of an OTFS block of symbols consists of N consecutive transmissions of OFDM blocks, each with a block length $M + M_{\text{cp}}$ in the time domain. Assuming that the receiver knows the channel perfectly, one (larger) block of OTFS transmission is equivalent to N parallel transmissions of OFDM blocks. Furthermore, we assume that the channel is independent from one OTFS block to the next and the channel is ergodic. These form the underlining key assumptions of the capacity analysis below.

Consider an arbitrary OTFS transmission block. Since there is a one-to-one mapping between Doppler-delay domain data symbols $\bar{\mathbf{d}}$ and time-frequency data samples $\bar{\mathbf{x}}$, the mutual information between the OTFS received signal vector \mathbf{r} and transmit data vector $\bar{\mathbf{d}}$, i.e. $I(\bar{\mathbf{d}}; \mathbf{r})$, can be written as

$$I(\bar{\mathbf{d}}; \mathbf{r}) = I(\bar{\mathbf{x}}_0, \bar{\mathbf{x}}_1, \dots, \bar{\mathbf{x}}_{N-1}; \mathbf{r}_0, \mathbf{r}_1, \dots, \mathbf{r}_{N-1}). \quad (52)$$

Since the received signal \mathbf{r}_n only depends on $\bar{\mathbf{x}}_n$, i.e. $\mathbf{r}_n = \mathcal{H}_n \bar{\mathbf{x}}_n + \mathbf{w}_n$, where \mathcal{H}_n represents all signal processing steps that relates $\bar{\mathbf{x}}_n$ to \mathbf{r}_n including modulation and channel impact. Then, (52) can be written as

$$I(\bar{\mathbf{d}}; \mathbf{r}) = \sum_{n=0}^{N-1} I(\bar{\mathbf{x}}_n; \mathbf{r}_n), \quad (53)$$

due to the fact that each OTFS transmission is realized by transmissions over a set of N parallel channels. Furthermore, each parallel channel is an MIMO channel $\mathbf{r}_n = \mathcal{H}_n \bar{\mathbf{x}}_n + \mathbf{w}_n$ with known channel state information at the receiver. Since the capacity-achieving input distribution for such a channel is the zero-mean circularly symmetric complex Gaussian distribution $\mathcal{CN}(0, \mathbf{I}_M)$, [6], we conclude that such distribution is indeed capacity-achieving for the proposed MIMO-OFDM based OTFS system and the resulting ergodic capacity equals

to the average sum capacities of individual parallel channels, which equals with the ergodic capacity of the OFDM system considered here [6].

In the following, for an OFDM-based OTFS setup we calculate $I(\bar{\mathbf{d}}; \mathbf{r})$ and show that (53) holds for OFDM-based OTFS modulation and we proceed to calculate the capacity of OFDM-based OTFS setup. We recall that the differential entropy of an n -element complex Gaussian vector \mathbf{q} with covariance matrix \mathbf{C}_q is [7]

$$h(\mathbf{q}) = \log_2((2\pi e)^n |\mathbf{C}_q|) \quad (54)$$

where $|\cdot|$ is the determinant operator. Moreover, the mutual information between two complex Gaussian vectors \mathbf{q} and \mathbf{r} is the difference between the differential entropies $h(\mathbf{q})$ and $h(\mathbf{q}|\mathbf{r})$. This is written as

$$\begin{aligned} I(\mathbf{q}; \mathbf{r}) &= h(\mathbf{q}) - h(\mathbf{q}|\mathbf{r}) \\ &= \log_2 \frac{|\mathbf{C}_q|}{|\mathbf{C}_{q|\mathbf{r}}|}. \end{aligned} \quad (55)$$

Substituting (40) and (43) in (45), the received signal of the n -th OFDM transmission can be obtained as

$$\mathbf{r}_n = \bar{\mathbf{H}}_n (\mathbf{I}_{n_t} \otimes \mathbf{F}_M^H) \bar{\mathbf{U}}_n \bar{\mathbf{x}}_n + \mathbf{w}_n. \quad (56)$$

Then, the mutual information between the OFDM received signal vector \mathbf{r}_n and the transmit data vector $\bar{\mathbf{x}}_n$ is

$$I(\mathbf{r}_n; \bar{\mathbf{x}}_n) = \log_2 \left(\frac{|\mathbf{K}_n \mathbf{I}_{Mn_t} \mathbf{K}_n^H + \sigma^2 \mathbf{I}_{Mn_r}|}{|\sigma^2 \mathbf{I}_{Mn_r}|} \right) \quad (57)$$

where $\mathbf{K}_n = \bar{\mathbf{H}}_n (\mathbf{I}_{n_t} \otimes \mathbf{F}_M^H) \bar{\mathbf{U}}_n$.

Similarly, the received signal of OFDM-based OTFS can be obtained by substituting (38) and (41) in (47) as

$$\mathbf{r} = \bar{\mathbf{H}} (\mathbf{I}_{Nn_t} \otimes \mathbf{F}_M^H) \bar{\mathbf{U}} (\mathbf{F}_N^H \otimes \mathbf{I}_{n_t} \otimes \mathbf{F}_M) \bar{\mathbf{d}} + \mathbf{w}. \quad (58)$$

The mutual information between the OFDM-based OTFS received signal vector \mathbf{r} and the transmit data vector $\bar{\mathbf{d}}$ is given as

$$I(\mathbf{r}; \bar{\mathbf{d}}) = \log_2 \left(\frac{|\mathbf{K} \mathbf{I}_{MNn_t} \mathbf{K}^H + \sigma^2 \mathbf{I}_{MNn_r}|}{|\sigma^2 \mathbf{I}_{MNn_r}|} \right) \quad (59)$$

where $\mathbf{K} = \bar{\mathbf{H}} (\mathbf{I}_{Nn_t} \otimes \mathbf{F}_M^H) \bar{\mathbf{U}} (\mathbf{F}_N^H \otimes \mathbf{I}_{n_t} \otimes \mathbf{F}_M)$.

Note that $\mathbf{F}_N^H \otimes \mathbf{I}_{n_t} \otimes \mathbf{F}_M$ is a unitary matrix and $\bar{\mathbf{H}} (\mathbf{I}_{Nn_t} \otimes \mathbf{F}_M^H) \bar{\mathbf{U}}$ is a block diagonal matrix, thus we have

$$\mathbf{K} \mathbf{K}^H = \begin{bmatrix} \mathbf{K}_0 \mathbf{K}_0^H & \cdots & \mathbf{0}_{Mn_r} & \\ \mathbf{0}_{Mn_r} & \mathbf{K}_1 \mathbf{K}_1^H & \cdots & \mathbf{0}_{Mn_r} \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{0}_{Mn_r} & \mathbf{0}_{Mn_r} & \cdots & \mathbf{K}_{N-1} \mathbf{K}_{N-1}^H \end{bmatrix}. \quad (60)$$

Substituting (60) in (59), we have

$$\begin{aligned} I(\mathbf{r}; \bar{\mathbf{d}}) &= \sum_{n=0}^{N-1} \log_2 \left(\frac{|\mathbf{K}_n \mathbf{K}_n^H + \sigma^2 \mathbf{I}_{Mn_r}|}{|\sigma^2 \mathbf{I}_{Mn_r}|} \right) \\ &= \sum_{n=0}^{N-1} I(\mathbf{r}_n; \bar{\mathbf{x}}_n). \end{aligned} \quad (61)$$

Hence, the ergodic capacity can be obtained as

$$C_{\text{OTFS}} = C_{\text{OFDM}} = \frac{1}{M + M_{\text{cp}}} \mathbb{E} \left[\log_2 \left(\frac{|\mathbf{K}_n \mathbf{K}_n^H + \sigma^2 \mathbf{I}_{Mn_r}|}{|\sigma^2 \mathbf{I}_{Mn_r}|} \right) \right]. \quad (62)$$

While in the above analysis we reached the conclusion that OTFS has the same ergodic capacity as OFDM, it is important to note that such analysis assumes perfect knowledge of the (time-varying) channel, an optimal detector, and an infinite code length. In a practical receiver design, however, the sparsity and lower variability of the OTFS channel in the Doppler-delay domain yield great benefits over OFDM, especially for higher Doppler channels. In such scenarios, an OFDM receiver needs to track rapid channel variations in the time-domain. Here, because of the channel variation over each OFDM symbol, keeping track of such variations turns out to be a difficult task. In OTFS, on the other hand, channel variation in time averages out and translates to a much slower variation in the Doppler-delay domain. In addition, it results in a sparse channel which will be easier to estimate. These benefits of OTFS, clearly, enable a simpler channel estimation and equalization design and hence reduce overhead and complexity of the receiver.

VII. CONCLUSION

In this work, we conducted a discrete-time analysis for MIMO OFDM-based OTFS modulation. Such analysis led to a concise, vectorized input-output relationship that is applicable to general time-varying channels with arbitrary Dopplers and windowing functions. We provided an accurate characterization of the ergodic capacity which shows that both OFDM and OTFS achieve the same ergodic capacity despite great benefits of the latter in practical receiver design. The analysis developed here provides a strong theoretical foundation for the design of practical detectors/equalizers for OTFS systems.

REFERENCES

- [1] R. Hadani, S. Rakib, M. Tsatsanis, A. Monk, A. J. Goldsmith, A. F. Molisch, and R. Calderbank, "Orthogonal time frequency space modulation," in *2017 IEEE Wireless Communications and Networking Conference (WCNC)*, March 2017, pp. 1–6.
- [2] M. Tuchler, R. Koetter, and A. C. Singer, "Turbo equalization: principles and new results," *IEEE Transactions on Communications*, vol. 50, no. 5, pp. 754–767, May 2002.
- [3] P. Raviteja, K. T. Phan, Q. Jin, Y. Hong, and E. Viterbo, "Low-complexity iterative detection for orthogonal time frequency space modulation," *arXiv preprint arXiv:1709.09402*, 2017.
- [4] L. Li, H. Wei, Y. Huang, Y. Yao, W. Ling, G. Chen, P. Li, and Y. Cai, "A simple two-stage equalizer with simplified orthogonal time frequency space modulation over rapidly time-varying channels," *arXiv preprint arXiv:1709.02505*, 2017.
- [5] A. Farhang, A. RezaadehReyhani, L. E. Doyle, and B. Farhang-Boroujeny, "Low complexity modem structure for OFDM-based orthogonal time frequency space modulation," *arXiv preprint arXiv:1710.00655*, 2017.
- [6] D. Tse and P. Viswanath, *Fundamentals of wireless communication*. Cambridge university press, 2005.
- [7] T. M. Cover and J. A. Thomas, *Elements of information theory*. John Wiley & Sons, 2012.

This figure "LTV.png" is available in "png" format from:

<http://arxiv.org/ps/1710.07900v1>