Multi-Carrier Modulation: An Evolution from Time-Frequency Domain to Delay-Doppler Domain

Hai Lin, Senior Member, IEEE, Jinhong Yuan, Fellow, IEEE, Wei Yu, Fellow, IEEE,
Jingxian Wu, Senior Member, IEEE, Lajos Hanzo, Life Fellow, IEEE

Abstract—The recently proposed orthogonal delay-Doppler division multiplexing (ODDM) modulation, which is based on the new delay-Doppler (DD) domain orthogonal pulse (DDOP), is studied. A substantial benefit of the DDOP-based ODDM or general delay-Doppler domain multi-carrier (DDMC) modulation is that it achieves orthogonality with respect to the fine time and frequency resolutions of the DD domain. We first revisit the family of wireless channel models conceived for linear time-varying (LTV) channels, and then review the conventional multi-carrier (MC) modulation schemes and their design guidelines for both linear time-invariant (LTI) and LTV channels. Then we discuss the time-varying property of the LTV channels' DD domain impulse response and propose an impulse function based transmission strategy for equivalent sampled DD domain (ESDD) channels. Next, we take an in-depth look into the DDOP and the corresponding ODDM modulation to unveil its unique input-output relation for transmission over ESDD channels. Then, we point out that the conventional MC modulation design guidelines based on the Wely-Heisenberg (WH) frame theory can be relaxed without compromising its orthogonality or without violating the WH frame theory. More specifically, for a communication system having given bandwidth and duration, MC modulation signals can be designed based on a WH subset associated with sufficient (bi)orthogonality, which governs the (bi)orthogonality of the MC signal within the bandwidth and duration. This is different from the conventional MC modulation design guidelines based on the global (bi)orthogonality governed by the WH frame theory, which corresponds to the MC signals being (bi)orthogonal across the whole TF domain. This novel design guideline could potentially open up opportunities for developing future waveforms required by new applications such as communication systems associated with high delay and/or Doppler shifts, as well as integrated sensing and communications, etc.

Index Terms—Multi-carrier modulation, Orthogonal frequency division multiplexing (OFDM), Doubly-selective channel, Orthogonal delay-Doppler division multiplexing (ODDM), Delay-Doppler domain multi-carrier (DDMC), Delay-Doppler domain orthogonal pulse (DDOP), Pulse-train-shaped OFDM (PTS-OFDM), Sufficient (bi)orthogonality, Global (bi)orthog-

onality, Gabor limit, Heisenberg uncertainty principle, Orthogonal time frequency space (OTFS), Zak transform, Integrated sensing and communications (ISAC)

I. Introduction

A wireless channel typically introduces both time and frequency dispersions, which correspond to the channel's frequency and time selectivity, respectively. Usually, such a doubly-selective channel can be modeled as a linear time-varying system, and it is represented by its time-varying channel impulse response or delay-Doppler spread function [?], a.k.a the spreading function [?]. Within the channel's coherence time, the channel model can be simplified to a linear time-invariant system, which only has time dispersion. The channel-induced dispersions have a crucial impact on signal transmission and therefore become the primary concern in the design of modulation schemes.

In digital communications, a modulation technique is essentially a scheme of using analog pulses or mathematically equivalent continuous-time functions to synthesize transmit signal waveforms, where each pulse carries an information-bearing digital symbol drawn from a signal constellation diagram [?]. In other words, a "symbol" is represented by a continuous-time function, which usually is the product of a digital symbol and an analog pulse. At the receiver, demodulation is often performed first by receive filtering based on matched filters or correlators corresponding to these analog pulses. Then, the extracted signal components are fed into a channel equalizer to recover the transmitted digital symbols. To avoid interference among the symbols and consequently ease the channel equalization, it is expected that these analog pulses do not interfere with each other, if possible, even in the presence of channel dispersions. As a result, (bi)orthogonal pulses or functions, which prevent mutual interference upon obeying the (bi)orthogonality among themselves, lie at the very heart of modulation techniques.

Since the eigenfunctions of a linear system are orthogonal functions and excite simple scaled system outputs, using eigenfunctions as the pulses for signal transmission seems to be an ideal strategy. For LTI systems, the eigenfunctions are complex sinusoids with infinite duration [?], while any practical pulse must have a finite duration. Fortunately, complex sinusoids are periodic functions, which implies that if we choose their frequencies appropriately, truncated complex sinusoids can still be orthogonal to each other and exhibit a scalar system input-output relation. The corresponding modulation scheme conceived for LTI

H. Lin is with the Department of Electrical and Electronic Systems Engineering, Graduate School of Engineering, Osaka Metropolitan University, Sakai, Osaka 599-8531, Japan (e-mail: hai.lin@ieee.org).

J. Yuan is with the School of Electrical Engineering and Telecommunications, the University of New South Wales, Sydney, Australia (e-mail: j.yuan@unsw.edu.au).

W. Yu is with the Edward S. Rogers Sr. Department of Electrical and Computer Engineering, the University of Toronto, Toronto, Canada (e-mail: weiyu@comm.utoronto.ca).

J. Wu is with the Department of Electrical Engineering, University of Arkansas, Fayetteville, Arkansas, USA (e-mail: wuj@uark.edu).

L. Hanzo is with the School of Electronics and Computer Science, University of Southampton, Southampton SO17 1BJ, U.K (E-mail: lh@ecs.soton.ac.uk).
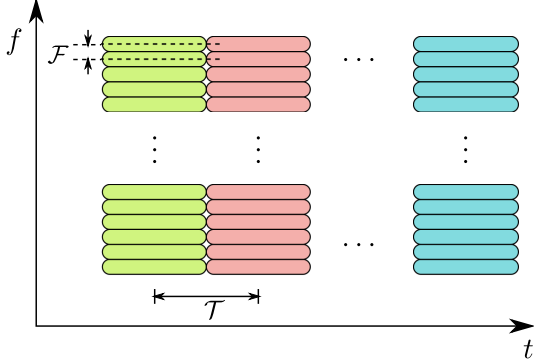
Acronyms

Fig. 1. TF grid in MC modulation

channels is the popular time-frequency domain multi-carrier modulation, typically, the orthogonal frequency division multiplexing [?], [?], [?], [?][1], which has been widely adopted in wireless standards, such as the Wi-Fi [?], [?], the fourth generation mobile communication system [?], the fifth generation mobile communication system [?], and so on.

In OFDM, to achieve eigenfunction-based transmission, the transmit pulses are exactly complex sinusoids termed as subcarriers and truncated by a common time-domain window function named as prototype pulse [?]. Therefore, a popular way of thinking about the transmit pulses in the MC modulation schemes is to treat them as TF-shifted prototype pulses along a TF grid. In fact, MC modulation can be defined by its prototype pulse and TF grid parameterized by a specific frequency resolution (subcarrier spacing) $\mathcal{F}$ and a specific time resolution (symbol interval) $\mathcal{T}$, where the minimum "distance" among these pulses may be quantified by the joint time-frequency resolution given as $\mathcal{R} = \mathcal{T}\mathcal{F}$.

An example of the TF grid in MC modulation is shown in Fig. ??, where each color block presents a transmit pulse obtained by TF-shifting the prototype pulse. Given $\mathcal{T}$ and $\mathcal{F}$, the two-dimensional TF domain in Fig. ?? is a gridded two-dimensional domain dedicated for representing the discrete TF structure and the signal (energy) localization of the MC modulation. Then, the fundamental issue of designing an MC modulation scheme becomes that of finding the (bi)orthogonal prototype pulse with respect to $\mathcal{T}$ and $\mathcal{F}$. Traditionally, the transmit pulses in Fig. ?? are considered as a Wely-Heisenberg or Gabor function set [?], which is the principal tool of the short-time Fourier transform based time-frequency analysis [?], [?], [?]. According to the WH frame theory, the (bi)orthogonal WH function sets only exist for $\mathcal{R} \geq 1$ [?], [?], and therefore most MC modulation schemes are designed for $\mathcal{R} \geq 1$ [?],

[?].

Things become more complicated when the channel's frequency dispersion is significant, for example, for an LTV channel encountered in a high-mobility environment. In contrast to LTI channels, which have elegant and common eigenfuctions to ease the pulse design, the underspead LTV channels at best have a structured set of approximate eigenfunctions, which depend on the spreading or scattering function of the channel [?]. In fact, these approximate eigenfunctions require not only a channel-dependent prototype pulse but also a much coarser JTFR than that of the classic OFDM [?], [?]. The channel-dependency and the coarse JTFR make it impracticable to realize eigenfunction-based transmission over LTV channels.

Recently, delay-Doppler domain modulation schemes – including the popular orthogonal time frequency space [?], [?] and the orthogonal delay-Doppler division multiplexing [?], [?] – have been proposed to address the challenges of waveform design for LTV or doubly-selective channels. The rationale behind the DD domain modulation schemes is to explore the fact that LTV channels often have a much longer stationary time[2] over which the DD state of the channel remains approximately constant, as compared to their coherence time over which the TF state is approximately constant [?]. Although the channel's TV-CIR still needs a large number of parameters within the stationary time, its DD spread function becomes deterministic and has a compact representation relying on much fewer parameters. Since each propagation path imposes not only attenuation and delay but also the Doppler spread, the TV-CIR of an LTV channel is given by the superposition of path-wise complex sinusoids associated with Doppler frequencies, and therefore presents fading. On the other hand, its DD spread function is simply characterized by each paths attenuation, delay and Doppler. Then, the destructive multi-path fading exhibited by the TV-CIR turns into separable paths in the DD domain, where path diversity can be harvested as a benefit of independent paths [?]. Based on these observations, the DD domain modulation schemes aim for coupling [?] the modulated signal with the DD domain channel or equivalently with the DD spread function of the LTV channel. Here, coupling represents the match between the TF grid namely the TF resolutions of the modulated signal and those of the DD domain channel [?]. Given an ideal coupling, the stability and sparsity of DD domain channels can be exploited to obtain diversity gains with minimum interference, and

---

[1]MC modulation is a general term, while the OFDM is a special form of MC modulation [?].

[2]Although the term of stationary time was originally used in the stochastic modelling of LTV channel, here we refer it as the time span over which the channel's delay and Doppler shifts remain almost the same.

hence achieve reliable communications with low pilot overhead and low processing complexity [?].

As for any modulation scheme, the fundamental issue of a DD domain modulation also resides in the pulse design. Intuitively, to achieve the signal-channel coupling with matched resolutions, the DD domain modulation requires a localized pulse in the DD domain. However, such a delay-Doppler domain localized pulse would violate the Heisenberg uncertainty principle [?], and therefore does not exist. To circumvent this issue, OTFS first converts the digital DD domain symbols to the digital TF domain signals via an inverse symplectic finite Fourier transform based digital precoder. Then, realizable orthogonal pulses in the TF domain are employed to carry these precoded signals and synthesize a conventional time-frequency domain multi-carrier modulation waveform, typically OFDM [?]. As a result, the transmit pulses of OTFS are still the time-frequency domain orthogonal pulse, and the OTFS is essentially a precoded OFDM [?]. In OTFS, the ideal pulse assumed has to be bi-orthogonal robust against the channel-induced delay and Doppler. Unfortunately, such an ideal pulse does not exist either [?], as none of the TFOPs can meet this requirement. Meanwhile, the widely adopted rectangular prototype pulse of popular OTFS studies is still a TFOP and therefore non-ideal, which will face practical challenges in implementation, such as high out-of-band emission and severe interference [?].

On the other hand, based on the newly discovered delay-Doppler domain orthogonal pulse, ODDM represents a novel delay-Doppler domain multi-carrier modulation scheme that can avoid the impediments of OTFS mentioned above [?]. Note that the time (duration) and frequency (bandwidth) constraints of any practical waveform define a TF region of interests, which results in

1) an equivalent sampled delay-Doppler domain channel associated with its specific delay and Doppler resolutions [?];
2) the DD domain in practical systems becomes a gridded DD domain, whose grids are defined by the delay and Doppler resolutions;
3) the gridded DD domain is essentially a gridded TF domain but with much finer resolutions, because the delay and Doppler have the physical units of time and frequency, respectively.

Thus, considering the presence of frequency resolution, a DD domain modulation is naturally a multi-carrier modulation, which requires a DDLP or a DDOP. Although the DDLP does not exist, the DDOP introduced in [?], [?], [?] consists of a train of square-root Nyquist pulses, and behaves like the "nonexistent" DDLP in the TF region of interests, and it achieves perfect coupling between the modulated signal and the ESDD channel. Specifically, without violating the Heisenberg uncertainty principle, the DDOP has an equally-spaced signal localization in the TF region of interests [?] to satisfy the orthogonality with respect to the specific delay and Doppler resolutions. It should be noted that the ESDD channel actually obeys the classic equivalent sampled channel model [?], [?], which is the on-grid equivalent of the effective channel corresponding to the cascaded transmit filter, propagation channel, and receive filter. Naturally, it is the ESDD channel that really matters for transceiver design.

In comparison to the conventional TFMC modulation schemes associated with $\mathcal{R} \geq 1$, the DDOP-based ODDM or general DDMC modulation is a new type of MC modulation having a much reduced JTFR of $\mathcal{R} \ll 1$, as we will show later. The ambitious objective of this paper is to analytically appraise this new modulation.

We commence by revisiting the LTV and LTI channels, and then review the conventional TFMC modulation schemes in terms of their transmission strategy, pulse design principles, implementation methods, and performance over LTV channels. Then we discuss MC modulation schemes designed for ESDD channels. We characterize the time-varying property of the DD domain impulse response of the channel, and propose an impulse function based transmission strategy for ESDD channels, which requires a pulse orthogonal with respect to the delay and Doppler resolutions of the ESDD channel, namely the DDOP. Next, we take an in-depth look into the DDOP's characteristics and the corresponding ODDM modulation for unveiling its unique IO relation over ESDD channels. Then, we point out that the conventional MC modulation design guidelines based on the WH frame theory can be relaxed without compromising its orthogonality. In particular, instead of the global (bi)orthogonality governed by the WH frame theory, the MC pulse design can be redefined by exploiting the WH subset based sufficient (bi)orthogonality. This new interpretation of the sufficient (bi)orthogonality actually relaxes the JTFR constraint of $\mathcal{R} \geq 1$ for (bi)orthogonal pulse design and leads to more general DDMC modulation schemes. This novel design guideline may open up opportunities for developing future waveforms required by new applications such as, integrated communications and sensing, high-mobility communications, etc.

The rest of the paper is organized as follows: Section II revisits the LTV and LTI channel models, especially the ESDD channel model taking into account the time and frequency constraints of practical signal waveforms. The family of classic TFMC modulation schemes is reviewed in Section III and Section IV for LTI channels and LTV channels, respectively, focusing on their transmission strategies, pulse designs and implementation methods. Then, based on the continuous-time channel IO relation, Section V investigates the properties of the ESDD channel, clarifies the corresponding time-varying DD domain impulse response, and proposes an impulse function based transmission strategy. The DDOP and its (bi)orthogonality are analyzed in details in Section VI, where the relation to the WH frame theory is explained and new pulse design guidelines are proposed. The important properties of the ODDM modulation are unveiled in Section VII, including its signal localization, bandwidth efficiency, implementation methods, and ISAC potentials.

Our simulation results are provided in Section VIII and finally Section IX concludes the paper.

Notations: In this paper, uppercase boldface letters are used to represent matrices, and lowercase boldface letters are used for column vectors. Furthermore, $\Pi_T(t)$ denotes the rectangular function of unit energy and TD support of $[0, T]$. The superscript T denotes the transpose operator, while $[\cdot]_M$ stands for the mod M operator. Finally, $\mathcal{A}_{g,\tilde{g}}(\tau, \nu)$ is the cross ambiguity function between two pulses $g(t)$ and $\tilde{g}(t)$, given by

$$\mathcal{A}_{g,\tilde{g}}(\tau, \nu) = \langle g(t), \tilde{g}(t-\tau)e^{j2\pi\nu(t-\tau)} \rangle,$$
$$= \int_{-\infty}^{\infty} g(t)\tilde{g}^*(t-\tau)e^{-j2\pi\nu(t-\tau)}dt.$$

II. Wireless Channel Models

A. Propagation Channel Models

Let us consider a complex-valued baseband signal $x(t)$ in the band of $[-B/2, B/2]$. Due to its limited bandwidth, $x(t)$ cannot be strictly time-limited. Similarly, a time-limited signal cannot be strictly frequency-limited. To accommodate practical time-limited signals having finite duration, the bandwidth here is defined in an essential sense of [?][3]. Then, $x(t)$ having a duration of T_x may be considered as both time- and band-limited.

For a wireless system communicating over an LTV channel, given the carrier frequency f_c, the passband radio frequency signal $\mathsf{x}(t) = \Re\{x(t)e^{j2\pi f_c t}\}$ is amplified and then sent through the LTV channel. We assume that the LTV channel is composed of $\tilde{P}$ paths corresponding to $\tilde{P}$ discrete specular scatters. Under the "narrowband" assumption of $B \ll f_c$, the variation of path attenuations and propagation delays vs. frequency can be omitted. Furthermore, noise terms are ignored in the following discussion for the sake of simplicity. Then, we have the received real-valued passband signal [?]

$$\mathsf{y}(t) = \sum_{p=1}^{\tilde{P}} \mathsf{a}_p(t)\mathsf{x}(t - \tilde{\tau}_p(t)), \tag{1}$$

where $\mathsf{a}_p(t)$ and $\tilde{\tau}_p(t)$ are the time-varying attenuation and delay of the p-th path, respectively. The corresponding received complex-valued baseband signal is

$$y(t) = \sum_{p=1}^{\tilde{P}} \tilde{h}_p(t)x(t - \tilde{\tau}_p(t)), \tag{2}$$

where $\tilde{h}_p(t) = \mathsf{a}_p(t)e^{-j2\pi f_c \tilde{\tau}_p(t)}$ represents the "gain" or the attenuation of the p-th path. Thus, the baseband TV-CIR can be written as

$$\tilde{h}(\tau, t) = \sum_{p=1}^{\tilde{P}} \tilde{h}_p(t)\delta(\tau - \tilde{\tau}_p(t)), \tag{3}$$

where δ denotes the Kronecker delta function and τ is the delay domain variable.

[3]The bandwidth may be defined by ignoring the negligibly small high-frequency tails beyond $[-B/2, B/2]$.

During the channel's stationary time when the time-variation of $\tilde{\tau}_p(t)$ accounting for "delay drift" can be neglected and the time variation of $\mathsf{a}_p(t)$ is caused by a Doppler spread $\tilde{\nu}_p$ [?], we have $\tilde{\tau}_p(t) = \tilde{\tau}_p$ and $\tilde{h}_p(t) = \tilde{h}_p e^{j2\pi\tilde{\nu}_p t}$. Then, we can rewrite the received complex-valued baseband signal in (??) as

$$y(t) = \sum_{p=1}^{\tilde{P}} \tilde{h}_p x(t - \tilde{\tau}_p)e^{j2\pi\tilde{\nu}_p t}, \tag{4}$$

and the baseband TV-CIR in (??) as

$$\tilde{h}(\tau, t) = \sum_{p=1}^{\tilde{P}} \tilde{h}_p e^{j2\pi\tilde{\nu}_p t}\delta(\tau - \tilde{\tau}_p). \tag{5}$$

The corresponding DD domain representation of the LTV channel is

$$\tilde{h}(\tau, \nu) = \sum_{p=1}^{\tilde{P}} \tilde{h}_p \delta(\tau - \tilde{\tau}_p)\delta(\nu - \tilde{\nu}_p), \tag{6}$$

where ν is the Doppler domain variable.

During the coherence time, when the channel's time-variation caused by $\tilde{\nu}_p$ can be further neglected, (??) becomes our familiar LTI channel's impulse response

$$\tilde{h}(\tau) = \sum_{p=1}^{\tilde{P}} \tilde{h}_p \delta(\tau - \tilde{\tau}_p), \tag{7}$$

which only introduces time dispersion. The IO relation of the LTI channel can be exactly written as a one-dimensional convolution between $x(t)$ and the time-invariant impulse response $\tilde{h}(\tau)$ in (??), given by

$$y(t) = \sum_{p=1}^{\tilde{P}} \tilde{h}_p x(t - \tilde{\tau}_p) = \int_{-\infty}^{\infty} x(t - \tau)\tilde{h}(\tau)d\tau. \tag{8}$$

Being a two-dimensional discrete function in the DD domain, $\tilde{h}(\tau, \nu)$ in (??) represents a deterministic channel model and it is known as a special case of the spreading function $\mathcal{S}(\tau, \nu)$, which usually is a 2D continuous function characterizing a continuum of scatters [?]. The counterpart of the spreading function $\mathcal{S}(\tau, \nu)$ in the statistical description of LTV channel is the scattering function given by [?]

$$\mathcal{C}(\tau, \dot{\tau}; \nu, \dot{\nu}) = \mathbb{E}\{S(\tau, \nu)S^*(\dot{\tau}, \dot{\nu})\}. \tag{9}$$

When $\mathcal{C}(\tau, \dot{\tau}; \nu, \dot{\nu}) = \mathsf{C}(\tau, \nu)\delta(\tau - \dot{\tau})\delta(\nu - \dot{\nu})$, we obtain the well-known wide-sense stationary uncorrelated scattering channel. In this paper, we assume that communications occur in the stationary time interval with a deterministic channel model represented by the spreading function, instead of a random LTV channel model represented by an ensemble of spreading functions.

Given a general spreading function $\mathcal{S}(\tau, \nu)$, the received complex-valued baseband signal in (??) can be generalized to

$$y(t) = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} \mathcal{S}(\tau, \nu)x(t - \tau)e^{j2\pi\nu t}d\tau d\nu, \tag{10}$$

whose frequency-domain representation is

$$Y(f) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \mathcal{S}(\tau,\nu) X(f-\nu) e^{-j2\pi\tau(f-\nu)} d\tau d\nu. \quad (11)$$

Meanwhile, given a normalized analysis window $g(t)$, the STFT of the transmitted signal $x(t)$ is defined as

$$X^{(g)}(t,f) \triangleq \int_{-\infty}^{\infty} x(\dot{t}) g^*(\dot{t}-t) e^{-j2\pi f\dot{t}} d\dot{t}. \quad (12)$$

Then, the STFT of the received signal $y(t)$ becomes [?]

$$Y^{(g)}(t,f) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \mathcal{S}(\tau,\nu) X^{(g)}(t-\tau, f-\nu) e^{-j2\pi\tau(f-\nu)} d\tau d\nu. \quad (13)$$

It is interesting to observe that except for the phase term $e^{-j2\pi\tau(f-\nu)}$, the above STFT-based IO relation of the LTV channel is a 2D convolution.

In mobile communications, the LTV channel is typically underspread with a spreading function confined to a small region in the DD domain. In particular, let $\tau_{\max}$ and $\nu_{\max}$ denote the channel's delay spread and Doppler spread, respectively. Then, an LTV channel is said to be underspread, when $4\tau_{\max}\nu_{\max} \leq 1$ [?]. Also, it should be noted that $\tilde{\tau}_p$ and $\tilde{\nu}_p$ are generally off-grid and the LTV channel $\mathcal{S}(\tau,\nu)$ or $\tilde{h}(\tau,\nu)$ is neither time- nor band-limited, as evidenced by its TF representation [?]

$$\mathcal{L}_{\mathcal{S}}(t,f) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \mathcal{S}(\tau,\nu) e^{j2\pi(t\nu-f\tau)} d\tau d\nu. \quad (14)$$

B. Equivalent Sampled Channel Models

Fig. ?? illustrates the block diagram of a typical wireless communication link. At the transmitter, even though the baseband signal $x(t)$ is designed to be band-limited, to mitigate the OOBE induced by the non-linearity of power amplifier, the output of the PA is usually filtered by a bandpass filter to avoid leakage to adjacent channels and to ensure compliance with the spectrum masks. On the other hand, using a BPF is a must at the receiver to select the channel and reject adjacent channel interference and out-of-band noise. In direct-conversion receivers, channel selection is realized by the combination of a BPF and low-pass filter after RF down-conversion. For heterodyne receivers, a cascade of an image-rejection filter, mixer, and channel selection filter after the low-noise amplifier comes out the channel selection [?]. It is clear that regardless of the transceiver design, the band-limited nature of $x(t)$ and the filtering operations lead to a band-limited channel and system.

Bearing in mind that $x(t)$ is also time-limited, in the communication link of Fig. ??, we can only observe an effective channel that is the time- and band-limited version of the propagation channel. In other words, the channel that really matters is not the propagation channel but the effective channel, which is the cascade of the transmit filter, propagation channel, and receive filter. The concept of effective channel is straightforward for single-carrier modulation, since its transmit/receive filter

is just the transmit/receive pulse. For MC modulation, the effective channel model is also valid, provided that the transmit/receive filter represents the overall effect of the subcarrier-wise transmit/receive pulses/filter, for example, an ideal LPF having a passband bandwidth of B.

Note that baseband signal processing is typically conducted by digital signal processor, and therefore requires an appropriate sampling of the received basedband signal $y(t)$. The frequency dispersion of the LTV channel will expand the bandwidth of $y(t)$ beyond B, but this is usually ignored in practice, since the Doppler spread is relatively small (on the order of tens to hundreds Hz) compared to the bandwidth B (on the order of MHz) [?].

For SC modulation, the LPF can act as the receive matched filter, the output of which is sampled at the symbol rate. The symbol rate (or sampling rate) is usually lower than B to intentionally cause aliasing and then achieve zero inter-symbol-interference over an ideal band-limited channel. For MC modulations, a DSP-based low-complexity implementation is preferred, because the direct implementations of subcarrier-wise receive pulses/filters are expensive. In this case, usually a Nyquist sampling rate is adopted to strike a balance between the link performance and its implementation cost. Then, the LPF acts as an anti-aliasing filter, which may be omitted if $y(t)$ has been appropriately filtered in the in-phase and quadrature demodulator.

It is clear now that although the LTV channel $\mathcal{S}(\tau,\nu)$ or $\tilde{h}(\tau,\nu)$ is of off-grid nature and neither time- nor band-limited, the combined duration and bandwidth constraints of $x(t)$ lead to a finite number of discrete-time samples at the output of the effective channel, which results in an equivalent on-grid channel, namely the ESDD channel, as shown in the dashed block of Fig. ??. In other words, the DD domain considering a practical time- and band-limited signal is always a gridded DD domain associated with the specific delay (time) and Doppler (frequency) resolutions.

Let the received signal's sampling rate and the sample duration be W and T, respectively. The ESDD channel corresponding to the DD domain channel in (??) may be written as [?]

$$h(\tau,\nu) = \sum_{p=1}^{P} h_p \delta(\tau-\tau_p) \delta(\nu-\nu_p), \quad (15)$$

where we have $P \geq \tilde{P}$, $\tau_p = l_p/\mathsf{W}$, $\nu_p = k_p/\mathsf{T}$, $l_p, k_p \in \mathbb{Z}$, and $1/\mathsf{W}$ as well as $1/\mathsf{T}$ are known as the delay and Doppler resolutions, respectively. The TF representation of (??) is given by

$$\mathcal{L}_h(t,f) = \sum_{p=1}^{P} h_p e^{j2\pi(t\nu_p - f\tau_p)} \quad (16)$$

for $t \in [0,\mathsf{T}]$ and $f \in [-\frac{\mathsf{W}}{2}, \frac{\mathsf{W}}{2}]$. Let us assume that the paths are arranged in ascending order of delay and $\tau_1 = 0$.
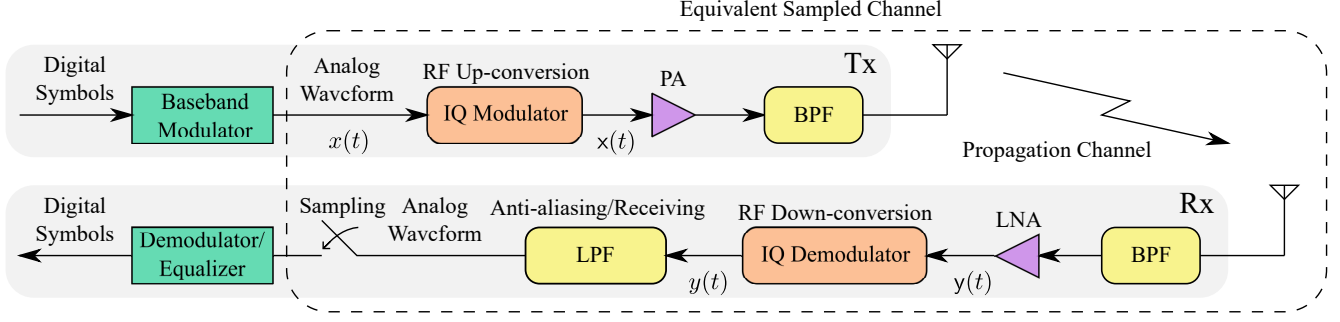
Fig. 2. Block diagram of a wireless communication link

TABLE I
MC Modulation Parameters

Notation	Parameter
$\mathcal{F}$	frequency resolution, subcarrier spacing, fundamental frequency
T	symbol period, $T = 1/\mathcal{F}$
$\mathcal{T}$	time resolution, symbol interval
$\mathcal{R}$	JTFR, $\mathcal{R} = \mathcal{T}\mathcal{F}$
N	number of subcarriers
M	number of symbols
$g(t)$	transmit prototype pulse
T_g	duration of $g(t)$, symbol duration
$G(f)$	Fourier transform of $g(t)$
B_g	bandwidth of $g(t)$, span of $G(f)$
$\gamma(t)$	receive prototype pulse

Then the delay spread of the ESDD channel is given by $\tau_{\max} = \tau_P$, and the LTI version of (??) is given by

$$h(\tau) = \sum_{p=1}^{P} h_p \delta(\tau - \tau_p), \qquad (17)$$

whose FD representation, i.e. the channel's transfer function is

$$H(f) = \sum_{p=1}^{P} h_p e^{-j2\pi\tau_p f}, \qquad (18)$$

for $f \in [-\frac{\mathsf{W}}{2}, \frac{\mathsf{W}}{2}]$. Note that the relation between (??) and (??) can be found in [?] when the transmit filter is an ideal LPF having a passband bandwidth of W.

The impact of the LTV and the LTI channels imposed on a practical modulation waveform is characterized by the equivalent sampled channel models in (??) and (??), respectively. Bearing in mind these impacts, a modulation scheme entails an appropriate (bi)orthogonal pulse design to strike a comprise between the bandwidth efficiency and the complexity of demodulation and equalization. In the following sections, we will discuss the modulation designs conceived for these channels.

III. TFMC modulation schemes designed for LTI channels

Starting from 1950s, MC modulation techniques including the OFDM have been developed for more than half a century. In the literature, there are many technical reviews of OFDM and its applications, such as [?], [?], [?], [?], [?], [?], [?]. The interested reader can find a comprehensive historic evolution of OFDM in [?, Table II].

The main purpose of this paper is to study the new ODDM/DDMC modulation, the corresponding new pulse design, and the unique transmission strategy for LTV channels. It is known that the (bi)orthogonal pulse design in conventional TFMC modulation schemes is closely related to the properties of a (bi)orthogonal WH function set, which is governed by the WH frame theory [?]. In this section, we will review the conventional TFMC modulation schemes, in particular, from the perspective of the eigenfunction-based transmission strategy, the WH frame theory based (bi)orthogonal pulse design principles, and the implementation methods.

A. Pulse Design Principles

The transmit pulses in an MC modulation can be represented by a function set [?]

$$(g, \mathcal{T}, \mathcal{F}) = \{g_{m,n}\}_{m,n\in\mathbb{Z}}, \qquad (19)$$

where $g_{m,n} \triangleq g(t - m\mathcal{T})e^{j2\pi n\mathcal{F}(t-m\mathcal{T})}$ and $g(t)$ is the prototype pulse. Similarly, we can form the receive pulses $(\gamma, \mathcal{T}, \mathcal{F})$ using another prototype pulse $\gamma(t)$ having the same time and frequency resolutions. Meanwhile, the inverse of the frequency resolution is known as the symbol period denoted by $T = 1/\mathcal{F}$.

Then, the transmit waveform of MC modulation synthesized by the transmit pulses in (??) is given by

$$x(t) = \sum_{m=0}^{M-1} \sum_{n=-N/2}^{N/2-1} X[m,n]g(t - m\mathcal{T})e^{j2\pi n\mathcal{F}(t-m\mathcal{T})}, \qquad (20)$$

where M is the number of MC symbols contained by $x(t)$, and the number of subcarriers N is usually supposed to be an even number. Furthermore, $X[m,n]$ for $-N/2 \le n \le N/2 - 1, 0 \le m \le M - 1$ represent the information-bearing digital symbols drawn from a signal constellation diagram, for example, quadrature amplitude modulation. One can see from (??) that single-carrier modulation is a special case of MC modulation associated with $N = 1$ or $n = 0$.
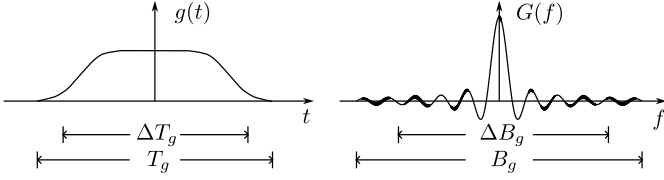
Fig. 3. Metrics of pulse's TF occupancy

Let T_g and $G(f)$ denote the duration and the Fourier transform of $g(t)$, respectively. The main parameters of MC modulation are listed in Table ??, where B_g, the bandwidth of $g(t)$ or the span of $G(f)$, is also defined in an essential sense of [?]. Meanwhile, it is noteworthy that apart from T_g and B_g, a classic metric of the occupancy of $g(t)$ in the TF domain is its time-frequency area $A_g = \Delta T_g \Delta B_g$, where the pulse's effective duration ΔT_g and effective bandwidth ΔB_g are defined as the standard deviations of its TD and FD shapes, respectively [?], [?]. As an example, the differences between T_g and ΔT_g, B_g and ΔB_g are shown in Fig. ??. Due to the Heisenberg uncertainty principle, the TFA obeys a lower bound $A_g \geq 1/(4\pi)$ known as the Gabor limit, which is attained by the Gaussian pulse [?][4]. Usually, $g(t)$ is said to be well-localized in the sense of minimum TF energy spread, when its signal energy is concentrated around its centre to have a small TFA [?].

Given $\mathcal{T}$ and $\mathcal{F}$, the fundamental issue of MC modulation is to find $g(t)$ and $\gamma(t)$ satisfying the orthogonal condition of

$$\langle g_{m,n}, g_{\dot{m},\dot{n}} \rangle = \delta(m - \dot{m})\delta(n - \dot{n}), \quad (21)$$

or the biorthogonal condition of

$$\langle g_{m,n}, \gamma_{\dot{m},\dot{n}} \rangle = \delta(m - \dot{m})\delta(n - \dot{n}). \quad (22)$$

By considering the TF domain as a 2D phase space, the function set in (??) forms a discrete grid "sampling" the phase space [?], [?], where the sampling resolution is the JTFR $\mathcal{R} = \mathcal{T}\mathcal{F}$. Then, the function set in (??) is treated as a WH set, the density of which is given by the inverse of the JTFR as $\mathcal{D} = \mathcal{R}^{-1}$. From the WH frame theory, the existence of (bi)orthogonal WH sets depends on the sampling resolution $\mathcal{R}$, which can be summarized as [?], [?], [?], [?], [?], [?], [?], [?], [?]:

- Critical sampling ($\mathcal{R} = 1$) : Orthogonal WH sets exist. However, they have either infinite TD or FD energy spread according to the Balian-Low theory [?]. Therefore, they are not well-localized in the TF domain.
- Under-critical sampling ($\mathcal{R} > 1$) : Well-localized orthogonal or biorthogonal WH sets exist, if $\mathcal{R}$ is sufficiently larger than 1.
- Over-critical sampling ($\mathcal{R} < 1$) : Neither orthogonal nor biorthogonal WH set exists.

---

[4]Suffice to say that the Gaussian pulse exhibits the smallest occupancy in the TF domain. As a result, it was adopted by the second generation mobile communication system global system of mobile communications across about 150 countries.

Here, we define the bandwidth efficiency of the transmit signal $x(t)$ as

$$\eta = \frac{MN}{BT_x}. \quad (23)$$

Since the functions $g_{m,n}$s in (??) are only different in terms of their TF centre, it is clear that η depends not only on the density of $(g, \mathcal{T}, \mathcal{F})$ but also on the duration T_g of $g(t)$, and on the bandwidth B_g of $g(t)$. In other words, to achieve high bandwidth efficiency, we have to place the pulses as densely as possible while keeping them (bi)orthogonal, which subsequently requires a fine JTFR and a well-localized $g(t)$. In fact, the highest bandwidth efficiency corresponds to the best use of the available dimension, which is also known as the degree of freedom of time- and band-limited signals [?], [?], [?].

B. Eigenfunction-based Transmission over LTI Channels

Passing the MC signal $x(t)$ of (??) through the LTI channel of (??), the received waveform is given by

$$y(t) = \int_{-\infty}^{\infty} x(t - \tau)h(\tau)d\tau = \sum_{p=1}^{P} h_p x(t - \tau_p). \quad (24)$$

Then, the receive pulse $\gamma_{m,n}$ is applied to $y(t)$ for extracting the signal component at the (m, n)-th TF grid point

$$Y[m,n] = \int_{-\infty}^{\infty} y(t)\gamma^*(t - m\mathcal{T})e^{-j2\pi n\mathcal{F}(t - m\mathcal{T})}dt, \quad (25)$$

which are fed into the channel equalizer of Fig. ?? to recover the transmitted digital symbols.

For SC modulation associated with $N = 1$ or $n = 0$, the transmit pulses occupy the whole bandwidth. Hence, the channel equalization is generally expensive, especially for channels exhibiting severe frequency selectivity. On the other hand, by slicing a wideband frequency selective channel into multiple narrowband frequency-flat subchannels, simple single-tap FD equalizers may be used, as in OFDM [?].

The rationale behind the single-tap equalization of OFDM is exactly its eigenfunction-based transmission strategy [?]. Recall that the subcarriers in OFDM or in general MC modulations are complex sinusoids, which are the eigenfunctions of LTI systems. Their frequencies are deliberately selected to be integer multiples of the frequency resolution $\mathcal{F}$, which leads to the term $n\mathcal{F}$ in (??). Because of the discrete upper harmonic frequencies of the subcarriers, OFDM is also known as discrete multi-tone modulation, especially in the digital subscriber line technology [?]. Meanwhile, observe that the symbol period $T = 1/\mathcal{F}$ is the common or fundamental period of these sinusoids, the orthogonal pulses can be obtained via truncating the subcarriers or equivalently modulating the subcarriers with the prototype pulse $g(t) = \Pi_T(t)$ [?], [?], [?]. Clearly, these truncated subcarriers are orthogonal to each other within the symbol period T, and each MC symbol is basically a single cycle of a periodic signal having a period of T.
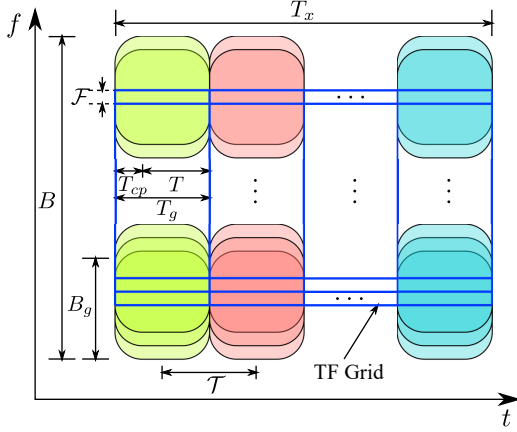
Fig. 4. TF grid and signal localization of CP-OFDM

Considering the time dispersion of the channel, the symbol duration T_g, which is also the length of the prototype pulse, is extended from T to $T + \tau_{\max}$. The pulses' extra length corresponding to the channel's delay spread $\tau_{\max}$ is a necessary TD redundancy, which prevents ISI induced by the time dispersion of the channel and then guarantees the orthogonality across a symbol period, namely the effective part of these truncated eigenfunctions. Since each MC symbol is a part of a periodic signal, the extension of the transmit pulses is equivalent to prepending a $T_{\rm cp}$-length cyclic prefix [?], corresponding to $g(t) = \Pi_{T_g}(t)$ where $T_g = T_{\rm cp} + T$ and $T_{\rm cp} \geq \tau_{\max}$. This type of OFDM is called CP-OFDM or just OFDM for simplicity. Furthermore, since the CP-OFDM symbols are sent one after another, we have $\mathcal{T} = T_g$. Given these parameter settings, each CP-OFDM symbol emulates an eigenfunction-based input that results in a scalar channel IO relation. In particular, given the receive prototype pulse $\gamma(t) = \Pi_T(t - T_{\rm cp})$, we have $Y[m,n] = H(n\mathcal{F})X[m,n]$, which is free of ISI and of inter-carrier-interference, enabling an easy recovery of $X[m,n]$.

Let a $B_g \times T_g$ block represent the occupancy of $g(t)$ in the TF domain. The corresponding TF grid and the TF signal localization of CP-OFDM with $g(t) = \Pi_{\mathcal{T}}(t)$ is illustrated in Fig. ??, where we obtain $T_x = MT_g = M\mathcal{T}$ and $B = (N-1)\mathcal{F} + B_g > N\mathcal{F}$, because $B_g > \mathcal{F}$. Given a fixed T_x, the TD redundancy causes the reduction of the number of symbols and consequently erodes the bandwidth efficiency. Meanwhile, without explicitly presenting the bandwidth and duration of each pulse, the TF grid is usually treated as a simplified version of the TF signal localization [?], as shown earlier in Fig. ??.

C. Classification of Pulses and TFMC Schemes

Given that the pulse density obeys $\mathcal{D} = \mathcal{R}^{-1}$, it is clear that $g(t)$ satisfying the orthogonal condition of (??) for the critical sampling of $\mathcal{R} = 1$ achieves the highest pulse density. A plausible choice of such a $g(t)$ is the aforementioned rectangular pulse $\Pi_T(t)$ first proposed in [?], which corresponds to the CP-free OFDM. Since its duration is constrained to the symbol period T, $\Pi_T(t)$ belongs to the time-limited orthogonal pulse set [?], where

T_g is usually less than $2T$. However, $\Pi_T(t)$ theoretically has an infinite energy spread in the FD caused by its slowly decaying Sinc shaped spectrum, which agrees with the Balian-Low theory [?]. In practical OFDM systems having stringent spectral restrictions, some edge subcarriers are unloaded to suppress the OOBE and to ease the transmit and receive filtering [?], [?], [?]. Moreover, we may have to further sharpen the signal spectrum by applying a frequency localized TD window after cyclically extending the OFDM symbol by both a CP and a cyclic suffix [?], [?], [?], [?], [?], [?], [?]. Upon further taking the necessary TD redundancy into account, the cyclic extension of each OFDM symbol results in a lower time resolution namely an extended symbol interval of $\mathcal{T} > T = 1/\mathcal{F}$, implying the spectrally inefficient under-critical sampling of $\mathcal{R} > 1$. As a result, by considering the vacant subcarriers in the FD and the cyclic extension in the TD, the TLOP-based OFDM suffers from a considerable loss of bandwidth efficiency in practice, due to the lack of well-localized orthogonal pulse in the case of critical sampling.

The somewhat disappointing spectral containment of the TLOP has motivated the design of a band-limited orthogonal pulse [?], [?], [?], [?], [?], [?], [?], [?], where the constraint imposed on the pulse's duration T_g is relaxed from $2T$ to multiple symbol periods, for better approximating their theoretically infinite duration. Since the pulse's duration T_g becomes much longer than the symbol interval $\mathcal{T}$, a heavy overlap of pulses occurs in the TD. As a result, the BLOP-based OFDM is also known as staggered multi-tone modulation [?], [?]. In BLOP-based OFDM, in exchange for the TD overlapping, only the neighboring FD subchannels are overlapped with each other [?]. This is in contrast to the densely overlapped subchannels of TLOP-based OFDM shown in Fig. ??. Also, in contrast to the classic QAM signaling in TLOP-based OFDM, BLOP-based OFDM usually employs offset quadrature amplitude modulation signaling and is termed as OFDM/OQAM [?]. Furthermore, if we treat $g(t)$ as the impulse response of a filter, the frequency-shifted pulses actually form a filter bank. Therefore, the OFDM/OQAM is also known as filter bank multi-carrier with OQAM (FBMC/OQAM). Similarly, the TLOP-based OFDM associated with a well-localized $g(t)$ rather than with the ordinary rectangular one is called FBMC or pulse-shaped OFDM [?].

With a well-localized pulse $g(t)$, OFDM/OQAM achieves excellent FD containment in the case of critical sampling. The "little magic" of the OFDM/OQAM in terms of circumventing the Balian-Low theory [?] involves shortening the symbol interval to $T/2$, and at the same time replacing the complex-valued orthogonality in (??) by real-valued orthogonality, which can be written as [?]

$$\langle g_{m,n}, g_{\dot{m},\dot{n}}\rangle_{\Re} \triangleq \Re\{\langle g_{m,n}, g_{\dot{m},\dot{n}}\rangle\} = \delta_{m-\dot{m}, n-\dot{n}}, \qquad (26)$$

where we have $g_{m,n} := g(t - m\frac{T}{2})e^{j2\pi n\mathcal{F}(t - m\frac{T}{2})}e^{j\phi_{m,n}}$. In addition, due to the pulse's long duration, the channel-induced ISI is usually negligible, and therefore the CP used for ISI mitigation can be omitted. As consequence, OFDM/OQAM has the highest bandwidth efficiency and

TABLE II
Comparison of Popular TFMC Modulation Schemes

Scheme	Pulse	Cyclic extension, pulse duration, symbol interval	Signaling
OFDM (CP-OFDM), Filtered OFDM[5]	Ordinary rectangular pulse with vacant edge subcarriers	CP, $T < T_g < 2T$, $\mathcal{T} = T_g$	QAM
PS-OFDM, FBMC	TF well-localized TLOP	CP+CS, $T < T_g < 2T$, $T < \mathcal{T} \leq T_g$	QAM
OFDM/OQAM, FBMC/OQAM, SMT	TF well-localized BLOP	No CP, $T_g > 2T$, $\mathcal{T} = T/2$	OQAM

Fig. 5. Analog implementation with $g(t)$ as transmit pulse/filter

still exhibits robustness against the channel's time dispersion.

In summary, TFMC modulations designed for LTI channels adopt the eigenfunction-based transmission strategy to ease the equalization of Fig. ??. According to the different eigenfunction-based pulse designs, they may be categorized as the TLOP-based OFDM (OFDM or CP-OFDM, FBMC or PS-OFDM) and the BLOP-based OFDM (OFDM/OQAM, FBMC/OQAM or SMT). A comparison of these popular TFMC modulation schemes is shown in Table ??. Hereafter, we will use the terms of OFDM, PS-OFDM and OFDM/OQAM, for the sake of simplicity.

It is noteworthy that in contrast to the eigenfunction-based transmission of TFMC modulation, the conventional SC modulation is an impulse function based transmission. In particular, its transmit pulses are basically the variants of band-limited impulses, and lead to $T_g > \mathcal{T}$. Usually, these transmit pulses will lose mutual orthogonality in the presence of the channel's time dispersion, and result in a convolutional system IO relation. Meanwhile, it has been known that the SC modulation can also be combined with TD redundancy to achieve simple FD equalization [?] and flexible multi-user band allocation [?]. Given its benefits, it was adopted in the 4G standard [?].

Now one can see that for LTI channels, generally we have two choices of transmission strategy: either the eigenfunction-based one having reduced bandwidth efficiency and simple channel equalization, or the impulse function based one having higher bandwidth efficiency but relatively complex channel equalization. In other words, there is always a trade-off between the bandwidth efficiency and the equalization complexity.

D. Implementation Methods

Given $\mathcal{T}$, $\mathcal{F}$ and $g(t)$, how to generate the waveform in (??) for MC modulation at low complexity is of pivotal practical importance, especially when the number of subcarriers N is large. In theory, we have two direct analog approaches using N modulators associated with
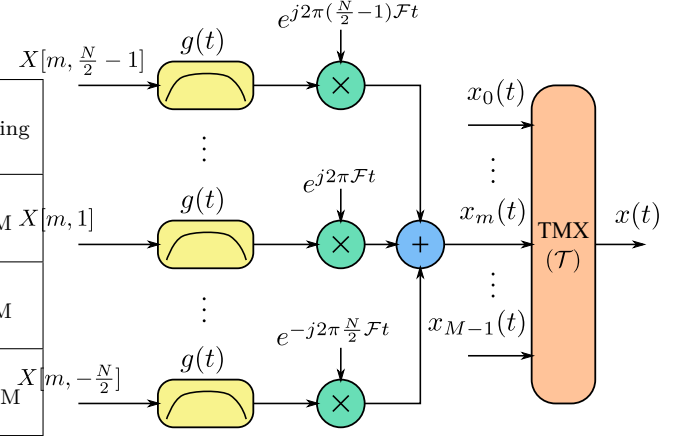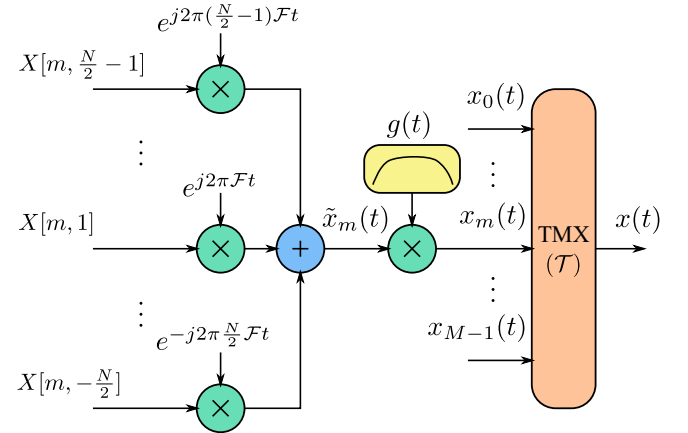
[5]This is explained in Section ??.

Fig. 6. Analog implementation with $g(t)$ as prototype pulse/winow function

carrier frequencies of $n\mathcal{F}, n = -\frac{N}{2}, \cdots \frac{N}{2} - 1$. As shown in Fig. ??, one of them is to generate $X[m,n]g(t - m\mathcal{T})$, modulate it according to $e^{j2\pi n\mathcal{F}(t-m\mathcal{T})}$, and then add them up to obtain

$$x_m(t) = \sum_{n=-N/2}^{N/2-1} X[m,n]g(t - m\mathcal{T})e^{j2\pi n\mathcal{F}(t-m\mathcal{T})}, \quad (27)$$

where $g(t)$ is treated as a transmit pulse/filter. The other one shown in Fig. ?? is to generate $X[m,n]e^{j2\pi n\mathcal{F}(t-m\mathcal{T})}$, add them up to have

$$\tilde{x}_m(t) = \sum_{n=-N/2}^{N/2-1} X[m,n]e^{j2\pi n\mathcal{F}(t-m\mathcal{T})}, \quad (28)$$

and then truncate the result by the pulse $g(t - m\mathcal{T})$ to obtain $x_m(t) = \tilde{x}_m(t)g(t - m\mathcal{T})$, where $g(t)$ is treated as a prototype pulse or window function. Once $x_m(t), m = 0, \cdots, M - 1$, are available, we can send them to a time division multiplexer parameterized by the symbol interval $\mathcal{T}$ to obtain $x(t) = \sum_{m=0}^{M-1} x_m(t)$, as in (??). It is noteworthy that the filter bank of Fig. ?? and the pulse-shaping seen in Fig. ?? exactly tally the terminology of FBMC and PS-OFDM, respectively.

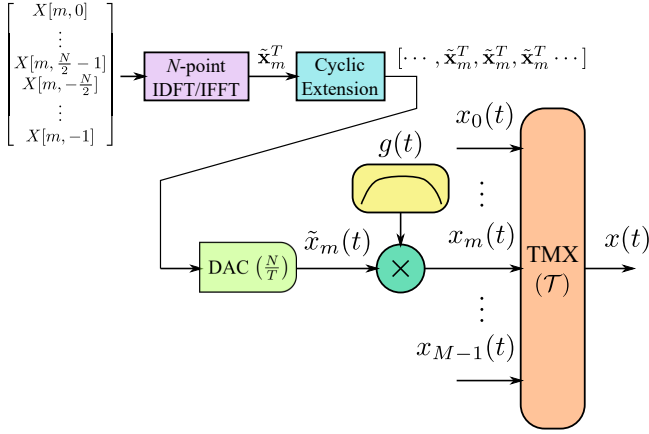Fig. 7. Digital implementation with $g(t)$ as prototype pulse/window function

Fig. 8. Oversampling based digital implementation

However, these two methods require N modulators, and therefore have prohibitively high implementation complexity. The implementation of MC modulation was considered unreasonable, until the connection between the inverse discrete Fourier transform, the discrete Fourier transform, and the MC modulation/demodulation was found in [?], where an analog hardware based implementation of the IDFT was employed to generate the MC waveform. After that, the inverse fast Fourier transform and fast Fourier transform algorithms [?] were introduced in [?] to realize the IDFT and DFT, which leads to the widely adopted IFFT/FFT-based implementation of MC modulation [?].

As shown in Fig. ??, the generation of $x(t)$ or $x_m(t)$ using the IFFT relies on a digital implementation of the second analog approach of Fig. ??, which is based on the following two observations. The first observation is that although $x_m(t)$ has a bandwidth of $B > N\mathcal{F}$, $\tilde{x}_m(t)$ in (??) is strictly band-limited to $[-\frac{N}{2}\mathcal{F}, (\frac{N}{2}-1)\mathcal{F}]$. Because the highest frequency is $\frac{N}{2}\mathcal{F}$, sampling $\tilde{x}_m(t)$ at the Nyquist rate of $N\mathcal{F} = N/T$ becomes feasible, by obeying the sampling theorem. Then, the N samples of $\tilde{x}_m(t)$ within one period $T = \frac{1}{\mathcal{F}}$ are given by

$$\tilde{x}_m[\dot{n}] \triangleq \tilde{x}_m\left(m\mathcal{T} + \dot{n}\frac{T}{N}\right) = \sum_{n=-N/2}^{N/2-1} X[m,n]e^{j2\pi\frac{n\dot{n}}{N}} \quad (29)$$

for $0 \le \dot{n} \le N-1$, which exactly represent the IDFT of $[X[m,0], \cdots, X[m, \frac{N}{2}-1], X[m, -\frac{N}{2}], \cdots, X[m, -1]]^T$. The second observation is that $\tilde{x}_m(t)$ is an infinite-length periodic signal, which indicates that we can repeat N samples in one period to obtain the samples of $\tilde{x}_m(t)$. Let $\tilde{\mathbf{x}}_m = [\tilde{x}_m[0], \cdots, \tilde{x}_m[N-1]]^T$. Then $\tilde{x}_m(t)$ can be generated by passing the cyclic extension of $\tilde{\mathbf{x}}_m^T$ namely $[\cdots, \tilde{\mathbf{x}}_m^T, \tilde{\mathbf{x}}_m^T, \tilde{\mathbf{x}}_m^T \cdots]$ through an ideal LPF with passband bandwidth $\frac{N}{T}$ [?], which is actually the interpolation filter in the digital-to-analog converter having a rate of $\frac{N}{T}$. After that, the $g(t)$-based windowing is applied to $\tilde{\mathbf{x}}_m(t)$ to obtain $x_m(t)$, whose bandwidth is then expanded to $B = (N-1)\mathcal{F} + B_g > N\mathcal{F}$ when $B_g > \mathcal{F}$. It should be noted that the windowing and time multiplexing can
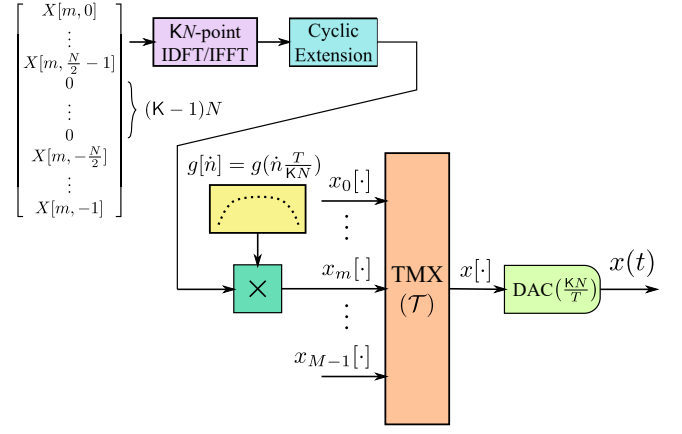
also be conducted in the digital domain before the DAC, if their implementation is performed at an oversampling rate of $\mathsf{K}\frac{N}{T}$, where $\mathsf{K} > 1$ is an integer representing the oversampling factor. The oversampling based implementation of MC modulation is shown in Fig. ??.

In practice, given the appropriate parameter settings, the cyclic extension of $\tilde{\mathbf{x}}_m^T$ and the subsequent windowing can be omitted for further simplifying the implementation. Recall that $B = (N-1)\mathcal{F} + B_g > N\mathcal{F}$ and N is usually a power of 2 for the IFFT. Then $x_m(t)$ becomes approximately band-limited to $[-\frac{N}{2}\mathcal{F}, (\frac{N}{2}-1)\mathcal{F}]$, if we have some unloaded subcarriers at the band edge. In fact, a practical OFDM system often only has $\bar{N} < N$ subcarriers, resulting in

$$x_m(t) = \sum_{\substack{n=-\bar{N}/2, \\ n\ne0}}^{\bar{N}/2} X[m,n]g(t-m\mathcal{T})e^{j2\pi n\mathcal{F}(t-m\mathcal{T})}, \quad (30)$$

where the direct-current subcarrier is also usually left unloaded to ease the RF circuit design [?]. With $N - \bar{N}$ vacant subcarriers, the bandwidth of $x_m(t)$ becomes $B = \bar{N}\mathcal{F} + B_g \le N\mathcal{F}$, where the setting of $B < N\mathcal{F}$ actually corresponds to the aforementioned oversampling-based implementation of Fig. ??. Then, we can pass $\tilde{\mathbf{x}}_m^T$ (after prepending a CP for channel delay spread) through the ideal LPF to obtain $x_m(t)$ directly, as long as $g(t)$ is $\Pi_\mathcal{T}(t)$, implying that the simple truncation is no need for $\tilde{\mathbf{x}}_m^T$ which is already time-limited. This vacant subcarriers based implementation, which has been widely adopted in practice because of its bandwidth flexibility, is shown in Fig. ??.

Filtering or equivalently sample-wise pulse-shaping $\tilde{\mathbf{x}}_m^T$ by an LPF results in filtered OFDM [?], [?], where the orthogonality among subcarriers will be eroded, especially for those at the band edge. This problem can be avoided by leaving edge subcarriers blank [?]. As a result, filtered OFDM can be considered as OFDM associated with rectangular pulse and vacant edge subcarriers, as shown in Table ??.

It is also noteworthy that since the interpolation filter in a DAC is exactly an LPF, practical OFDM systems with
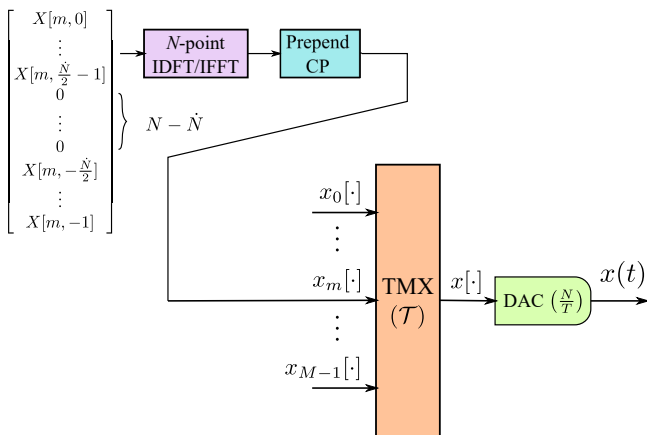
Fig. 9. Vacant edge subcarriers based digital implementation (Filtered OFDM)

the vacant subcarriers based implementation are actually filtered OFDM systems. The simulations of such OFDM systems are usually performed based on TD samples with $\frac{1}{N\mathcal{F}}$-interval, which is appropriate only when there is a sufficiently large number of vacant edge subcarriers in the system.

In comparison to the convenient IFFT-based implementation of the OFDM and PS-OFDM, the implementation of OFDM/OQAM has a considerably higher complexity, owing to the overlapping of MC symbols shaped by the long-duration $g(t)$. An elegant approach to alleviate this difficulty is to amalgamate the IFFT with a polyphase network for efficiently realizing a filter bank [?], [?]. This approach is essentially a FD equivalent of the aforementioned TD operations, including the cyclic extension and the windowing/pulse-shaping.

IV. TFMC modulation schemes for LTV channels

For doubly-selective LTV channels, unfortunately, neither OFDM nor OFDM/OQAM performs well. However, considering their wide deployment in practice, numerous efforts have been devoted to studying OFDM for transmission over doubly-selective channels. Based on different approximations of the TV-CIR and the resultant nondiagonal but banded FD channel matrix, the LTV channel estimation and equalization techniques can be found in [?], [?], [?], [?], [?], [?], [?], [?], [?], [?], [?], [?] and the references therein. It was pointed out that the off-diagonal elements of the FD channel matrix actually offer time diversity or Doppler diversity [?], [?]. However, these complex TV-CIR channel models often suffer from expensive yet somewhat inaccurate channel estimation. This is because the TV-CIR has a large number of parameters even along with a low-rank approximation. For OFDM/OQAM, due to its intrinsic imaginary interference caused by the real-valued orthogonality constraint, the channel-induced ISI and ICI remain cumbersome. To the best of our knowledge, there is no potent equalization solution for OFDM/OQAM operating in high-mobility environments, where single-tap equalizers fail to achieve satisfactory performance [?].

On the other hand, in the spirit of eigenfunction-based transmission, PS-OFDM schemes for achieving a scalar IO relation for transmission over LTV channels can be found in [?], [?], [?], [?], [?] and the references therein. It is widely recognized that the underspead LTV channels at best have a structured set of approximate eigenfunctions [?]. Owing to the necessity to consider both time and frequency dispersions, these approximate eigenfunctions not only depend on the channel's spreading or scattering functions, bust also require a well-localized prototype pulse $g(t)$ and a much coarser JTFR than that of OFDM [?], [?]. In fact, the corresponding pulses can be considered as an approximation of the (bi)orthogonal WH sets in the case of under-critical sampling of $\mathcal{R} > 1$, and therefore require not only TD but also FD redundancy. Moreover, the approximate nature of these eigenfunctions leads to residual ISI and ICI, which may still remain cumbersome and have to be equalized carefully. As a result, there are two main challenges in realizing eigenfunction-based transmission in LTV channels. The first one is the adaption of transmit pulses to the channel, which is barely possible in practical transmitters. The second one is the considerably reduced bandwidth efficiency associated with the rather coarse JTFR.

In summary, the conventional TFMC modulations designed for eigenfunction-based transmissions in LTI channels suffer not only from costly and inaccurate channel estimation, but also from complex equalization in LTV channels. On the other hand, PS-OFDM designed for eigenfunction-based transmission in LTV channels have severe practical challenges and exhibit low bandwidth efficiency.

V. MC modulations for ESDD channels

The challenges imposed by eigenfunction-based transmission in LTV channels motivate us to consider new transmission strategies and develop alternative modulation schemes. To this end, it is necessary to reconsider the properties of LTV channels and design specifically "tailored" channel-oriented pulses.

A. Common Properties of ESDD Channels

In practical transceivers, the transmit and receive pulses are usually fixed. Therefore, these pulses should be designed according to the common properties of LTV channels. However, in contrast to LTI channels relying on the complex sinusoidal eigenfunctions as their common properties, there is basically no common property for LTV or DD channels, due to the different propagation environments.

Recall that due to the limited bandwidth and duration of the signal, we only observe an ESDD channel at the receiver. Thus the ESDD channel is the one that matters for signal transmissions. Then, what we really care about is the common properties of ESDD channels, rather than those of DD channels. As shown in (??), the spreading functions of ESDD channels are discretized

with the delay resolution $1/\mathsf{W}$ and the Doppler resolution $1/\mathsf{T}$, corresponding to the signal's sampling rate and duration, respectively. As a result, although the general LTV or DD channels do not have common properties, the corresponding ESDD channels do have common delay and Doppler resolutions, which are determined either by the signal or by the system.

Since the delay and Doppler have the same unit with the time and frequency, respectively, we consider MC modulation in the DD domain with the delay resolution of $1/\mathsf{W}$ and the Doppler resolution of $1/\mathsf{T}$, and design the corresponding pulse.

B. DD Domain 2D Impulse Response

The discretized spreading function in (??) can be viewed as the ESDD channel's 2D impulse response, which is literally time-invariant. Recall that the gridded DD domain has the physical units of time and frequency, so it is also a gridded TF domain, but the DD domain is able to resolve delay and Doppler with much finer resolution. As we will show below, the complex ICI of OFDM or OFDM/OQAM in LTV channels is exactly due to the mismatch between the modulation and the channel.

Let us consider passing an MC signal $x(t)$ through the ESDD channel of (??). The received waveform is given by

$$y(t) = \sum_{p=1}^{P} h_p x(t - \tau_p) e^{j2\pi\nu_p(t-\tau_p)}, \qquad (31)$$

where $x(t)$ in (??) can be rewritten as $x(t) = \sum_{m=0}^{M-1} x_m(t)$. Assume that we set a large enough $\mathcal{T} \geq T_g + \tau_{\max}$ for completely isolating MC symbols from each other and preventing ISI. Then, from (??), the mth received MC symbol is given by

$$
\begin{aligned}
y_m(t) =& \sum_{p=1}^{P} h_p \sum_{n=0}^{N-1} X[m,n] g(t - m\mathcal{T} - \tau_p) \\
&\times e^{j2\pi n \mathcal{F}(t - m\mathcal{T} - \tau_p)} e^{j2\pi\nu_p(t-\tau_p)} \\
=& \sum_{p=1}^{P} h_p e^{j2\pi\nu_p m\mathcal{T}} \sum_{n=0}^{N-1} X[m,n] g(t - m\mathcal{T} - \tau_p) \\
&\times e^{j2\pi n \mathcal{F}(t - m\mathcal{T} - \tau_p)} e^{j2\pi\nu_p(t - m\mathcal{T} - \tau_p)}, \qquad (32)
\end{aligned}
$$

which is contaminated by ICI due to the Doppler-induced frequency dispersion. A well-known interpretation of this result is to consider ν_p as a carrier frequency offset, as in conventional OFDM systems. Since usually $\nu_p/\mathcal{F} \notin \mathbb{Z}$, this kind of fractional CFO will cause severe ICI [?].

Eq. (??) also implies that for $x_m(t)$, the ESDD channel is

$$h_m(\tau, \nu) = \sum_{p=1}^{P} h_p e^{j2\pi\nu_p m\mathcal{T}} \delta(\tau - \tau_p) \delta(\nu - \nu_p), \qquad (33)$$

rather than $h(\tau, \nu)$ of (??). In fact, as long as we have a symbol interval of $\mathcal{T} > 0$, (??) holds regardless of whether the ISI exists or not. In the context of OFDM, the phase term $e^{j2\pi\nu_p m\mathcal{T}}$ in (??) has similar effects as the phase

difference between common phase errors of two adjacent OFDM symbols corrupted by the same CFO [?]. As a result, for a signal consisting of time-multiplexed symbols, different symbols experience different ESDD channels, where these ESDD channels have the same number of paths, the delay and Doppler shifts, but different path gains. In other words, from the perspective of signal transmission, the 2D impulse response of an ESDD channel is still time-varying, due to the phase term discussed above. Similar results can also be obtained straightforwardly for the DD channel in (??).

C. Transmission Strategy for ESDD Channels

Bearing in mind that $\tau_p = l_p/\mathsf{W}$, $\nu_p = k_p/\mathsf{T}$, and $l_p, k_p \in \mathbb{Z}$ in (??), one can see from (??) that if $\mathcal{F} = 1/\mathsf{T}$, we have $\nu_p/\mathcal{F} = k_p \in \mathbb{Z}$, and then the ICI becomes aligned with the fine frequency resolution $1/\mathsf{T}$, which has a similar effect to that of integer CFO [?]. The same thing happens to the ISI if $\mathcal{T} = 1/\mathsf{W}$. In other words, if we can design a DDMC modulation as

$$x(t) = \sum_{m=0}^{M-1} \sum_{n=-N/2}^{N/2-1} X[m,n] g\left(t - \frac{m}{\mathsf{W}}\right) e^{j2\pi n \frac{1}{\mathsf{T}}\left(t - \frac{m}{\mathsf{W}}\right)}, \quad (34)$$

the time and frequency resolutions of which are identical to those of the ESDD channel, the ISI and ICI will be aligned to the ESDD channel grid. As a result, each transmitted symbol is only interfered by a minimum number of its neighbor symbols, and the pattern of the whole interference become compact. Further considering that the delay resolution $1/\mathsf{W}$ and the Doppler resolution $1/\mathsf{T}$ are common properties of ESDD channels, the DDMC modulation in (??) seems to be an appropriate modulation scheme for the ESDD channels.

The aligned ISI and ICI can also be interpreted using the relation between the STFT of $x(t)$ and $y(t)$ in (??). Note that the STFT is equivalent to applying receive pulses (analysis window) to extract the signal components in the MC modulation. Similar to (??), upon substituting $\mathcal{S}(\tau, \nu) = h(\tau, \nu)$ and $t = m/\mathsf{W}$, $f = n/\mathsf{T}$ into (??) to extract the signal component at the (m, n)-th TF grid point, we have

$$
\begin{aligned}
Y[m,n] =& Y^{(g)}\left(\frac{m}{\mathsf{W}}, \frac{n}{\mathsf{T}}\right) \\
=& \sum_{p=1}^{P} h\left(\frac{l_p}{\mathsf{W}}, \frac{k_p}{\mathsf{T}}\right) X^{(g)}\left(\frac{m - l_p}{\mathsf{W}}, \frac{n - k_p}{\mathsf{T}}\right) \\
&\times e^{-j2\pi \frac{l_p(n - k_p)}{\mathsf{WT}}}. \qquad (35)
\end{aligned}
$$

Then, we have the following proposition.

Proposition 1. If the transmit pulse $g(t)$ is an orthogonal pulse with respect to the delay resolution $1/\mathsf{W}$ and

Doppler resolution $1/\mathsf{T}$ of the ESDD channel, the received signal component at the (m,n)-th TF grid point obeys

$$
Y[m,n] = \sum_{p=1}^{P} h\left(\frac{l_p}{\mathsf{W}}, \frac{k_p}{\mathsf{T}}\right) X[m-l_p, n-k_p]
$$
$$
\times e^{-j2\pi \frac{l_p(n-k_p)}{\mathsf{WT}}}. \tag{36}
$$

Proof: By substituting the orthogonality property of $g(t)$ into (??), (??) can be obtained straightforwardly. ∎

Proposition ?? reveals the basic form of the IO relation for DDMC over ESDD channels. From (??), one can see that except for some phase terms, the extracted signal component at each TF grid point in the DDMC can be described as a 2D convolution between the transmitted digital symbols and the ESDD channel. Upon considering the impulse function based transmission strategy of SC modulation and the resultant 1D convolutional IO relation over LTI channels, it becomes clear that the DDMC modulation is an impulse function based transmission for ESDD channels.

Apparently, an appropriately designed $g(t)$ is necessary to realize such an impulse function based transmission. However, the impulse in the DD domain or DDLP has a TFA less than $1/(4\pi)$, which violates the Heisenberg uncertainty principle. Therefore, DDLP does not exist [?]. Meanwhile, because the JTFR is now $\mathcal{R}_{\mathrm{DD}} = 1/(\mathsf{WT}) \ll 1$, according to the WH frame theory, (bi)orthogonal WH sets do not exist either. Hence, it seems impossible to perform DD domain modulation, due to the lack of pulses. In the following sections, we will discuss progress in the design of DD domain modulation.

D. OTFS Modulation

Modulating information-bearing symbols in the DD domain was first considered in form of the OTFS modulation [?], [?]. To avoid the aforementioned pulse design challenges, the OTFS modulation transforms the signals from the DD domain to the TF domain by the ISFFT precoder. Then it modulates the transformed signals using the symbol-wise CP-free OFDM, and prepends a frame based CP for the whole OTFS frame.

Let $\mathcal{F} = F_0$ and $T_0 = 1/F_0$. Then, the TF domain grid of OTFS obeys $\{\hat{n}T_0, \hat{m}\frac{1}{T_0}\}$ for $\hat{n} = 0, \ldots, N-1$ and $\hat{m} = 0, \ldots M-1$, while the corresponding DD domain grid is defined as $\left\{m\frac{T_0}{M}, n\frac{1}{NT_0}\right\}$ for $m = 0, \ldots, M-1$ and $n = 0, \ldots N-1$. Then, the waveform of an OTFS frame without the frame-wise CP can be written as [?]

$$
\hat{x}(t) = \sum_{\hat{n}=0}^{N-1} \sum_{\hat{m}=-\frac{M}{2}}^{\frac{M}{2}-1} \mathcal{X}[\hat{n}, [\hat{m}]_M] g(t - \hat{n}T_0) e^{j2\pi \hat{m} F_0(t-\hat{n}T_0)},
$$
$$
\tag{37}
$$

where $\mathcal{X}[\hat{n}, \hat{m}]$ is obtained by the ISFFT as

$$
\mathcal{X}[\hat{n}, \hat{m}] = \frac{1}{\sqrt{MN}} \sum_{m=0}^{M-1} \sum_{n=0}^{N-1} X[m,n] e^{j2\pi(\frac{\hat{n}n}{N} - \frac{\hat{m}m}{M})}. \tag{38}
$$

Note that in OTFS, M and N are also the number of subcarriers and symbols of the underlying OFDM modulation, respectively. Observe from (??), the OTFS waveform is designed exactly with $\mathcal{R} = \mathcal{TF} = 1$, because $\mathcal{T} = T = T_0$.

The ideal pulse of OTFS in (??) is said to satisfy the biorthogonal robustness property [?]. Roughly speaking, this means that the pulses $g(t - \hat{n}T_0) e^{j2\pi \hat{m} F_0(t-\hat{n}T_0)}$ for different m or n in (??) are (bi)orthogonal with each other, even after experiencing the time and frequency dispersion induced by the channel. Given this ideal pulse, OTFS expects to achieve a 2D convolution between the transmit symbols $X[m,n]$ and the channel $h(\tau, \nu)$ at the channel output [?]. Unfortunately, the assumed ideal pulse cannot be realized in practice [?]. The underlying reason may be that the pulses satisfying the biorthogonal robustness property essentially correspond to the eigenfunction-based transmission that achieves a scalar IO relation over the LTV channel, which requires $\mathcal{R} > 1$, as mentioned in Section ??. Since the OTFS waveform is designed for $\mathcal{R} = 1$, the corresponding TLOP cannot be biorthogonal robust.

Due to the lack of ideal pulse, the rectangular pulse $\Pi_{T_0}(t)$ has been widely adopted in current OTFS studies [?], which is the TLOP for the CP-free OFDM with $\mathcal{R} = 1$. Recall that for the CP-free OFDM, some vacant edge subcarriers are necessary to suppress the OOBE, as mentioned in Section ??. Since OTFS relies on CP-free OFDM, OOBE is also an inevitable practical issue for OTFS [?]. However, letting $\mathcal{X}[\hat{n}, \hat{m}] = 0$ for a part of $\hat{m}, 0 \leq \hat{m} \leq M - 1$ may break the inherent connection between $X[m,n]$ and $\mathcal{X}[\hat{n}, \hat{m}]$ governed by the ISFFT precoder in (??), as $X[m,n]$ is drawn from a QAM constellation. Also, the absence of CP and CS in CP-free OFDM makes the windowing-based OOBE mitigation methods infeasible, while letting $g(t)$ in (??) be a spectrally compact pulse leads to severe performance degradation [?] due to loss of orthogonality. Furthermore, it is noteworthy that from (??), the expected 2D convolution may be unachievable, considering the phase terms induced by the time-varying path gains.

E. DDMC/ODDM Modulation

Motivated by OTFS's concepts of modulating information-bearing symbols in the DD domain and DD grid, we present a general DDMC signal design in this section. Considering the DD domain grid in OTFS, let us substitute $\frac{1}{\mathsf{W}} = \frac{T_0}{M}$ and $\frac{1}{\mathsf{T}} = \frac{1}{NT_0}$ into (??), for the sake of comparison. Then, the DDMC modulation waveform becomes

$$
x(t) = \sum_{m=0}^{M-1} \sum_{n=-N/2}^{N/2-1} X[m, [n]_N] g\left(t - m\frac{T_0}{M}\right) e^{j2\pi n \frac{1}{NT_0}(t - m\frac{T_0}{M})},
$$
$$
\tag{39}
$$

where the time and frequency resolutions are $\mathcal{T} = \frac{T_0}{M} = \frac{1}{MF_0}$ and $\mathcal{F} = \frac{1}{NT_0} = \frac{F_0}{N}$, respectively. The main differences between the OTFS waveform in (??) and the DDMC
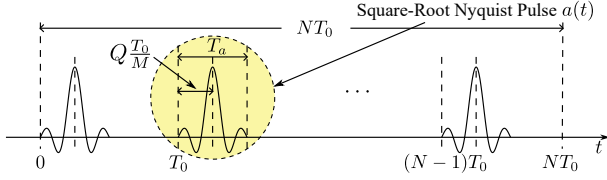
Fig. 10. Delay Doppler domain orthogonal pulse (DDOP) $u(t)$.

waveform in (??) are the resolutions in the time (dealy) domain (T_0 in time v.s. $\frac{T_0}{M}$ in delay) and the frequency (Doppler) domain ($\frac{1}{T_0}$ in frequency v.s. $\frac{1}{NT_0}$ in Doppler).

Now the crucial question is whether there exists a DDOP denoted by $u(t)$, that is orthogonal with respect to both the time (delay) resolution $\mathcal{T} = \frac{T_0}{M}$ and the frequency (Doppler) resolution $\mathcal{F} = \frac{1}{NT_0}$. If such a DDOP does exist, we can realize this hypothetical ODDM modulation by letting $g(t) = u(t)$ in (??). According to [?], such a DDOP does exist, if we let $u(t)$ be a pulse-train given by

$$u(t) = \sum_{\dot{n}=0}^{N-1} a(t - \dot{n}T_0), \tag{40}$$

where the subpulse $a(t)$ is a square-root Nyquist pulse parameterized by its zero-ISI interval T_0/M and its duration $T_a = 2Q\frac{T_0}{M}$, where Q is a positive integer.

Figure ?? illustrates the unique structure of the DDOP $u(t)$. When $2Q \ll M$ and therefore $T_a \ll T_0$, it has been proved that $u(t)$ satisfies the orthogonality property of

$$\mathcal{A}_{u,u}\left(m\frac{T_0}{M}, n\frac{1}{NT_0}\right) = \delta(m)\delta(n), \tag{41}$$

for $|m| \leq M - 1$ and $|n| \leq N - 1$, where $\mathcal{A}_{u,u}(\cdot)$ is the ambiguity function of $u(t)$ defined as

$$\mathcal{A}_{u,u}(\tau, \nu) = \langle u(t), u(t - \tau)e^{j2\pi\nu(t-\tau)} \rangle, \tag{42}$$
$$= \int_{-\infty}^{\infty} u(t)u^*(t - \tau)e^{-j2\pi\nu(t-\tau)}dt.$$

In other words, $u(t)$ is orthogonal with respect to the delay and Doppler resolutions of the ESDD channel within the range of interest, $|m| \leq M - 1$ and $|n| \leq N - 1$. After prepending an appropriate frame based CP to $x(t)$ in (??), the ODDM frame is sent through the channel.

Let us assume that the maximum delay and Doppler of the ESDD channel in (??) are $(L - 1)\frac{T_0}{M}$ and $K\frac{1}{NT_0}$, respectively. The P paths can be arranged in a $(2K+1) \times L$ DD domain channel matrix $\boldsymbol{\Theta}$, where each row and each column of $\boldsymbol{\Theta}$ correspond to a Doppler and delay index, respectively. For example, let $\hat{k} = k - K - 1$, a nonzero element of $\boldsymbol{\Theta}$, denoted by $\theta(\hat{k} + K + 1, l)$, be equal to the gain h_p of the pth path, whose delay and Doppler are $l\frac{T_0}{M}$ and $\hat{k}\frac{1}{NT_0}$, respectively. Note that the total number of nonzero elements in $\boldsymbol{\Theta}$ is P.

At the receiver, the matched filtering based on $u(t - m\frac{T_0}{M})e^{-j2\pi n\frac{1}{NT_0}(t-m\frac{T_0}{M})}$ is performed to obtain $Y[m,n]$, the signal component at the (m,n)-th TF grid point,

namely at the n-th subcarrier of the m-th ODDM symbol. Let us arrange $Y[m,n]$ and $X[m,n]$ in two vectors as

$$\mathbf{y} = [\mathbf{y}_0^T, \mathbf{y}_1^T, \cdots, \mathbf{y}_{M-1}^T]^T,$$
$$\mathbf{x} = [\mathbf{x}_0^T, \mathbf{x}_1^T, \cdots, \mathbf{x}_{M-1}^T]^T,$$

where

$$\mathbf{y}_m = [Y[m,0], Y[m,1], \cdots, Y[m,N-1]]^T,$$
$$\mathbf{x}_m = [X[m,0], X[m,1], \cdots, X[m,N-1]]^T,$$

for $0 \leq m \leq M - 1$. Then, the noise-free IO relation of ODDM over the ESDD channel can be written as

$$\mathbf{y} = \mathbf{H}\mathbf{x}, \tag{43}$$

where the DD domain channel matrix $\mathbf{H}$ is given by

$$\begin{bmatrix} \mathbf{H}_0^0 & & & & \mathbf{H}_{L-1}^0\mathbf{D} & \cdots & \cdots & \mathbf{H}_1^0\mathbf{D} \\ \vdots & \ddots & & & & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & & & & \ddots & \vdots \\ \mathbf{H}_{L-2}^{L-2} & \ddots & \ddots & \mathbf{H}_0^{L-2} & & \mathbf{0} & & \mathbf{H}_{L-1}^{L-2}\mathbf{D} \\ \mathbf{H}_{L-1}^{L-1} & \ddots & \ddots & \ddots & \mathbf{H}_0^{L-1} & & & \\ & \ddots & \ddots & \ddots & \ddots & \ddots & & \\ & & \ddots & \ddots & \ddots & \ddots & \ddots & \\ \mathbf{0} & & & \mathbf{H}_{L-1}^{M-1} & \ddots & & \ddots & \mathbf{H}_0^{M-1} \end{bmatrix} \tag{44}$$

with

$$\mathbf{H}_l^m = \sum_{\hat{k}=-K}^{K} \theta(\hat{k} + K + 1, l)e^{j2\pi\frac{\hat{k}(m-l)}{MN}}\mathbf{C}^{\hat{k}}, \tag{45}$$

$$\mathbf{D} = \text{diag}\left\{1, e^{-j\frac{2\pi}{N}}, \ldots, e^{-j\frac{2\pi(N-1)}{N}}\right\}, \tag{46}$$

and the $N \times N$ cyclic permutation matrix is formulated as

$$\mathbf{C} = \begin{bmatrix} 0 & \ldots & 0 & 1 \\ 1 & \ddots & 0 & 0 \\ \vdots & \ddots & \ddots & \vdots \\ 0 & \ldots & 1 & 0 \end{bmatrix}. \tag{47}$$

As an $MN \times MN$ block-circulant-like matrix, the DD domain channel matrix $\mathbf{H}$ in (??) represents the linear combination between $X[m,n]$ and the ESDD channel $h(\tau, \nu)$. Upon using the DDOP $u(t)$ of Fig. ?? as the transmit and receive prototype pulses, ODDM becomes capable of outperforming the OTFS in terms of both its OOBE and bit error rate [?], [?]. It is noteworthy that with the DDOP namely the pulse-train $u(t)$, the ODDM can be viewed as a PS-OFDM or more precisely a pulse-train-shaped OFDM.

The existence of the DDOP of (??) comes as a bit of surprise, because the conventional understanding of (bi)orthogonal pulses indicates that orthogonal WH sets do not exist with such fine JTFR $\mathcal{R}_{\text{DD}} = T_0/M \times 1/(NT_0) = 1/(MN) \ll 1$. The existence of $u(t)$ thus motivates us to reconsider the fundamental (bi)orthogonal pulse design

principles based on the WH frame theory, which will be discussed in the next section.

VI. Pulse Design Subject to the Signal's TF Constraints

Observe from (??), this orthogonality is constrained within M symbols each having N subcarriers. Therefore it only applies to a local region in the TF domain. Since MC modulation has a limited number of symbols and subcarriers, the orthogonality within this local TF region defined by the bandwidth and duration of the signal is sufficient. As a result, we can reformulate the pulse design problem for MC modulation by taking the TF constraints of the practical signal into account.

Without loss of generality, let us consider the TF region in Fig. ??, where the bandwidth and duration are B and T_x, respectively. It is widely understood that for a signal contained in this region, its DoF is bounded by BT_x, which can be achieved by using the prolate spheroidal wave functions [?]. In other words, we can transmit up to $\lceil BT_x \rceil$ digital symbols, by carrying them using the PSW functions. However, the PSW functions neither have a complex sinusoidal based structure required by MC modulations, nor can they benefit the channel equalization. Upon further considering the dispersive channel effects and the DoF of the received signal, it might be necessary to relax the sampling rate W and samples duration T to $\mathsf{W} \leq B$ and/or $\mathsf{T} \leq T_x$ to design an MC modulation associated with other (bi)orthogonal functions and transmit up to $\lceil \mathsf{WT} \rceil \leq \lceil BT_x \rceil$ digital symbols, at the cost of a modest erosion of bandwidth efficiency.

A. Global and Sufficient (Bi)orthogonality

Analogous to (??) and (??), the (bi)orthogonal pulse design problem taking the TF constraints of the signal, namely the limited number of symbols and subcarriers, into account is to find specific WH subsets $(g, \mathcal{T}, \mathcal{F}, M, N)$ and $(\gamma, \mathcal{T}, \mathcal{F}, M, N)$ that satisfy the orthogonal condition of

$$\langle g_{m,n}, g_{\dot{m},\dot{n}} \rangle = \delta(m - \dot{m})\delta(n - \dot{n}), \ m, \dot{m} \in \mathbb{Z}_m, n, \dot{n} \in \mathbb{Z}_N, \tag{48}$$

or the biorthogonal condition of

$$\langle g_{m,n}, \gamma_{\dot{m},\dot{n}} \rangle = \delta(m - \dot{m})\delta(n - \dot{n}), \ m, \dot{m} \in \mathbb{Z}_m, n, \dot{n} \in \mathbb{Z}_N, \tag{49}$$

where

$$\mathbb{Z}_M = \{0, \cdots, M - 1\}, \ \mathbb{Z}_N = \{0, \cdots, N - 1\}. \tag{50}$$

Here, the index of subcarriers $\{-N/2, \cdots 0, \cdots, N/2 - 1\}$ is changed to $\{0, \cdots, N - 1\}$ for simplifying the notation, which corresponds to a half-bandwidth shift of the carrier frequency f_c. This will not affect the analysis of (bi)orthogonality.

Since (??) and (??) only consider a local region in the TF domain, we term them as the sufficient orthogonal condition and sufficient biorthogonal condition, respectively. Furthermore, because of

$$\langle g_{m,n}, g_{\dot{m},\dot{n}} \rangle = \mathcal{A}_{g,g}(\bar{m}\mathcal{T}, \bar{n}\mathcal{F})e^{j2\pi n\bar{m}\mathcal{F}\mathcal{T}}, \tag{51}$$

where $\bar{m} = \dot{m} - m$ and $\bar{n} = \dot{n} - n$, the sufficient orthogonal condition in (??) is equivalent to

$$\mathcal{A}_{g,g}(\bar{m}\mathcal{T}, \bar{n}\mathcal{F}) = \delta(\bar{m})\delta(\bar{n}), \tag{52}$$

for $|\bar{m}| \leq M - 1, |\bar{n}| \leq N - 1$. Similar results can be obtained for the sufficient biorthogonal condition in (??).

In the context of TFA, WH sets are used for analyzing finite-energy signals lying in the space of $L^2(\mathbb{R})$. For accurate analysis, the WH sets have to be WH frames, which are complete or overcomplete WH sets with a certain guaranteed numerical stability of reconstruction and this requires $\mathcal{R} \leq 1$ [?], [?]. When a WH set $(g, \mathcal{T}, \mathcal{F})$ is a WH frame, we can denote it as $\{g, \mathcal{T}, \mathcal{F}\}$ by replacing round brackets with curly brackets. Let $\mathcal{T}_\dagger = 1/\mathcal{F}$ and $\mathcal{F}_\dagger = 1/\mathcal{T}$. From the duality and biorthogonality theory for WH frames [?], [?], [?], we know that $(g, \mathcal{T}, \mathcal{F})$ and $(\gamma, \mathcal{T}, \mathcal{F})$ are biorthogonal if and only if the associated WH sets $(g, \mathcal{T}_\dagger, \mathcal{F}_\dagger)$ and $(\gamma, \mathcal{T}_\dagger, \mathcal{F}_\dagger)$ are dual frames, while $(g, \mathcal{T}, \mathcal{F})$ is orthogonal if and only if the associated WH set $(g, \mathcal{T}_\dagger, \mathcal{F}_\dagger)$ is a tight frame[6]. To obtain the dual frames $\{g, \mathcal{T}_\dagger, \mathcal{F}_\dagger\}$ and $\{\gamma, \mathcal{T}_\dagger, \mathcal{F}_\dagger\}$ or the tight frame $\{g, \mathcal{T}_\dagger, \mathcal{F}_\dagger\}$, the corresponding JTFR has to satisfy $\mathcal{R}_\dagger = \mathcal{T}_\dagger \mathcal{F}_\dagger \leq 1$ and consequently $\mathcal{R} = \mathcal{T}\mathcal{F} \geq 1$. Therefore, (bi)orthogonal WH sets do not exist for $\mathcal{R} < 1$.

The WH frame theory based results regarding (bi)orthogonal WH sets are rigorous. However, since a WH set is originally a TFA tool conceived for functions in $L^2(\mathbb{R})$, it considers the whole TF domain where $m, n \in \mathbb{Z}$, and corresponds to the signal without the limit of bandwidth and duration. To make this possible, given $\mathcal{T}$ and $\mathcal{F}$, $g(t)$ must be independent of the number of symbols M and the number of subcarriers N, to be shifted freely over the whole TF domain. As a result, $g(t)$ is designed based only on $\mathcal{F}$ and $\mathcal{T}$ to achieve the global (bi)orthogonality in (??) and (??), and therefore bounded by the JTFR limit of $\mathcal{R} = 1$.

On the other hand, for MC modulations, in contrast to achieving the global (bi)orthogonality in (??) and (??), we only have to consider the sufficient (bi)orthogonality in (??) and (??) to satisfy the perfect reconstruction condition for M MC symbols with N subcarriers, corresponding to a WH subset under the TF constraints of the signal. Apparently, $g(t)$ that achieves the global (bi)orthogonality can form a (bi)orthogonal WH subset. However, since we really only require a WH subset to satisfy the sufficient (bi)orthogonality, it is not necessarily bound by the WH frame theory for the WH set. In fact, the pulses parameterized by not only $\mathcal{T}$ and $\mathcal{F}$ but also by M and N can also achieve the sufficient orthogonality.

B. Orthogonality with Respect to $\mathcal{F}$

Let us consider a fixed m and variable n in $g_{m,n}$, and investigate the orthogonality with respect to the frequency resolution $\mathcal{F}$ first. We want to find $g(t)$ that can achieve the orthogonality among $g(t - m\mathcal{T})e^{j2\pi n\mathcal{F}(t - m\mathcal{T})}$ with a

---

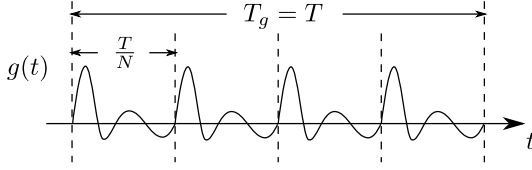[6]A WH frame is called tight if its lower and upper frame bounds are the same.

Fig. 11. $g(t)$ orthogonal w.r.t $\mathcal{F} = \frac{1}{T}$ for $|n| \leq N - 1$ and fixed m

given m but variable n, where $0 \leq t \leq T_g$ and $T_g = T = 1/\mathcal{F}$. Without loss of generality, let $m = 0$, we can obtain the following results:

F1) Unbounded n ($n \in \mathbb{Z}$): $g(t)$ is the rectangular pulse $\Pi_T(t)$, which is independent of N.

F2) Bounded n ($|n| \leq N - 1$): We have the following proposition:

Proposition 2. When $g(t)$ is a unit energy periodic function with a period of $\frac{T}{N}$ for $0 \leq t \leq T_g$ and $T_g = T$, it satisfies the orthogonal property of

$$\mathcal{A}_{g,g}(0, n\mathcal{F}) = \langle g(t), g(t)e^{j2\pi n\mathcal{F}t} \rangle = \delta(n). \quad (53)$$

for $|n| \leq N - 1$.

Proof: See Appendix A. ∎

Proposition ?? indicates that once there is a constraint imposed on the number of subcarriers, surprisingly there are an infinite number of pulses satisfying the orthogonality with respect to $\mathcal{F}$ within the symbol period of $T = 1/\mathcal{F}$. In other words, as long as $g(t)$ is a periodic function satisfying the above conditions, regardless of its bandwidth B_g, it can achieve the orthogonality among N subcarriers. Considering the case of F1) where B_g is proportional to $\mathcal{F}$ and the total bandwidth of the signal is about $N\mathcal{F}$, Proposition ?? actually decouples the relation between B_g and $\mathcal{F}$. An example of such a function $g(t)$ for $N = 4$ is shown in Fig. ??.

C. Orthogonality with Respect to $\mathcal{T}$

Similarly, we can consider a fixed n in $g_{m,n}$, and investigate the orthogonality with respect to the time resolution $\mathcal{T}$. Our target now becomes that of finding a specific $g(t)$ that can achieve the orthogonality among $g(t - m\mathcal{T})e^{j2\pi n\mathcal{F}(t - m\mathcal{T})}$ with a fixed n but variable m. When $n \neq 0$, we have the following straightforward answer with temporally isolated pulses/subpulses:

T1) Unbounded m ($m \in \mathbb{Z}$) : Any $g(t)$ with duration $T_g \leq \mathcal{T}$, which is independent of M.

T2) Bounded m ($|m| \leq M - 1$): $g(t)$ is a pulse-train being made up of $\dot{N} > 1$ subpulses $b_{\dot{n}}(t), 0 \leq \dot{n} \leq \dot{N} - 1$, where these subpulses are temporally spaced by $M\mathcal{T}$ and each subpulse has a duration $T_{b_{\dot{n}}} \leq \mathcal{T}$.

Meanwhile, when $n = 0$, we have another solution associated with temporally overlapped pulse/subpulses:

T3) Unbounded m ($m \in \mathbb{Z}$) : square-root Nyquist pulse $a_\mathcal{T}(t)$ with $\mathcal{T}$ being the zero-ISI interval, which is also independent of M.

T4) Bounded m ($|m| \leq M-1$) : $g(t)$ is a pulse-train being made up of $\dot{N} > 1$ square-root Nyquist subpulses $a_\mathcal{T}(t)$, where these subpulses are temporally spaced by $M\mathcal{T}$.

D. Sufficient Orthogonality with Respect to $\mathcal{F}$ and $\mathcal{T}$

From T1)-T4), we know that given the time resolution $\mathcal{T}$, the key to achieving the orthogonality among M symbols is to either limit the pulse duration to be no greater than $\mathcal{T}$ or to employ square-root Nyquist pulses whose zero-ISI interval is $\mathcal{T}$. Also, for the case of the pulse-train in T2) and T4) corresponding to the sufficient orthogonality, the subpulses have to be temporally spaced by $M\mathcal{T}$. At the same time, we know from F2) that given the frequency resolution $\mathcal{F}$, the key to achieving the sufficient orthogonality among N subcarriers is to form a periodic function with period $\frac{1}{N\mathcal{F}}$. Therefore, to achieve the sufficient orthogonality with respect to $\mathcal{F}$ and $\mathcal{T}$ concurrently, we can consider a combination of the conditions in T2), T4) and F2). Furthermore, because the subpulses in T2) have much shorter duration and therefore much wider bandwidth than those in T4), the combination of T4) and F2) is preferred.

Given the time resolution $\mathcal{T}$ and assuming that the duration of $a_\mathcal{T}(t)$ obeys $T_{a_\mathcal{T}} < M\mathcal{T}$, it is interesting to observe that by letting the number of subpulses to be $\dot{N} = N$, the pulse-train in T4) is a periodic function that satisfies F2), if we let $T_g = T = MN\mathcal{T}$ and $\mathcal{F} = 1/T = 1/(MN\mathcal{T})$ in F2). Further, upon substituting $\frac{T_0}{M}$ into $\mathcal{T}$, the pulse-train becomes the DDOP $u(t)$ of (??) and achieves the orthogonality with respect to $\mathcal{T} = \frac{T_0}{M}$ and $\mathcal{F} = 1/(NT_0)$, where $\mathcal{R} = 1/(MN) \ll 1$.

As a result, by combining the pulse-train structure required by the orthogonality with respect to the Doppler resolution, and the square-root Nyquist pulse required by the orthogonality with respect to the delay resolution, we can bypass the JTFR limit of $\mathcal{R} \geq 1$ for global (bi)orthogonality to achieve sufficient orthogonality with $\mathcal{R} \ll 1$. Compared to the traditional WH set based principles, the WH subset based principles and the resultant pulse-train structure may pave the way for conceiving new pulse designs for MC modulations.

E. General DDOP and Sufficient Biorthogonality

The above result of sufficient orthogonality is based on the appropriate duration of $a_\mathcal{T}(t)$ namely $T_{a_\mathcal{T}} < M\mathcal{T}$. Recall that the orthogonality of the DDOP in (??) is also subject to a similar duration constraint of $T_a \ll T_0$ (equivalently $2Q \ll M$). In practice, it is desirable to relax this constraint for the sake of flexible design. In the following, we will show that this duration constraint can be relaxed by introducing the cyclic extension, which leads to a general DDOP design.

When $T_a > T_0$, $u(t)$ is no longer a period function with a period of T_0 during $[0, NT_0]$, which is required to satisfy the orthogonality with respect to $\mathcal{F} = \frac{1}{NT_0}$. This observation inspires us to use a cyclically extended

version of $u(t)$ namely $u_{ce}(t)$, as the transmit pulse, while the receive pulse is still $u(t)$, corresponding to a biorthogonality condition. Furthermore, because the cross ambiguity function $\mathcal{A}_{u_{ce},u}(\cdot)$ is calculated between $u_{ce}(t)$ and $u(t - m\frac{T_0}{M})e^{j2\pi\frac{n}{NT_0}(t-m\frac{T_0}{M})}$, the problem to satisfy the orthogonality with respect to $\mathcal{F} = \frac{1}{NT_0}$ becomes how $u_{ce}(t)$ can have the specified periodicity within the range of $u(t - m\frac{T_0}{M})e^{j2\pi\frac{n}{NT_0}(t-m\frac{T_0}{M})}$ for $|m| \leq M - 1$. We then have the following proposition:

Proposition 3. Let a pulse-train $u(t)$ be made up of N square-root Nyquist pulses $a(t)$, which are temporally spaced by T_0. The resultant pulse train satisfies the biorthogonality property of

$$\mathcal{A}_{u_{ce},u}\left(m\frac{T_0}{M}, n\frac{1}{NT_0}\right) = \delta(m)\delta(n), \quad (54)$$

for $|m| \leq M - 1$ and $|n| \leq N - 1$, where $u_{ce}(t)$ is a cyclically extended version of $u(t)$. Specifically, $u_{ce}(t)$ is a periodic function with period T_0 during $-(M-1)\frac{T_0}{M} \leq t \leq (MN-1)\frac{T_0}{M} + T_a$.

Proof: See Appendix B. ∎

Note that the proof of Proposition ?? does not depend on T_a, which indicates that the duration constraint of $a(t)$ in $u(t)$ can be removed. Once the appropriate CP and CS are added in accordance with (??), where the extension parameter for CP and CS is $D = \lceil T_a/T_0 \rceil = \lceil 2Q/M \rceil$, the desired sufficient biorthogonality can be achieved as well.

VII. Important Properties of ODDM Modulation

Given the general DDOP design, the transmit pulse of ODDM modulation becomes $u_{ce}(t)$. When $M \gg 2Q$, we have $2Q/M \approx 0$. Then, as proved in [?], ODDM can employ the DDOP $u(t)$ without cyclic extension ($D = 0$).

A. TF Signal Localization

The TF signal localization plays an important role in the analysis of modulation waveforms. For example, the TF signal localization of CP-OFDM given in Fig. ?? explains its orthogonality with respect to the coarse JTFR in the conventional TF domain. To understand the TF signal localization of an MC modulation, we need both the TD and FD representations of its transmit pulse. In the following, we will derive $U(f)$, namely the FD representation of $u(t)$ corresponding to the case of $D = 0$, as we usually have $M \gg 2Q$ in practice.

1) Frequency Domain Representation of DDOP: It is widely exploited that the FD representation of an impulse train

$$\ddot{u}(t) = \sum_{n=-\infty}^{\infty} \delta(t - nT_0), \quad (55)$$

is a Fourier series, which can also be written as an impulse train in the FD

$$\ddot{U}(f) = \frac{1}{T_0} \sum_{m=-\infty}^{\infty} \delta(f - \frac{m}{T_0}). \quad (56)$$

It is interesting to observe that the DDOP can be obtained from $\dot{u}(t)$ by applying a rectangular window $\Pi_{NT_0}\left(t + \frac{T_0}{2}\right)$ followed by a filter with the impulse response $a(t)$. Then, we have

$$u\left(t + \frac{T_a}{2}\right) = \dot{u}(t) * a(t), \quad (57)$$

where $\dot{u}(t) = \ddot{u}(t) \times \Pi_{NT_0}\left(t + \frac{T_0}{2}\right)$ and $*$ denotes convolution. Since the multiplication and convolution in the TD correspond to the convolution and multiplication in the FD, respectively, we have

$$\begin{aligned}
U(f) &= e^{-j2\pi f \frac{T_a}{2}} A(f)\dot{U}(f) \\
&= e^{-j2\pi f \frac{T_a}{2}} A(f)\left(\ddot{U}(f) * e^{-j2\pi f \frac{(N-1)T}{2}} \mathrm{Sinc}(fNT_0)\right), \\
&= \frac{e^{-j2\pi f \tilde{T}}}{T_0} A(f) \sum_{m=-\infty}^{\infty} e^{j2\pi \frac{m(N-1)}{2}} \mathrm{Sinc}(fNT_0 - mN),
\end{aligned}$$
(58)

where $\dot{U}(f) = \ddot{U}(f) * e^{-j2\pi f \frac{(N-1)T}{2}} \mathrm{Sinc}(fNT_0)$, $\tilde{T} = (T_a + (N-1)T_0)/2$ and $A(f)$ is the Fourier transform of $a(t)$.

Without loss of generality, let M be an even number. Then, the derivation of $U(f)$ is graphically illustrated in Fig. ??, where the phase terms are omitted and the shapes of $\mathrm{Sinc}(fNT_0 - mN)$ are truncated for the purpose of display [?], [?]. Now, it becomes clear that $\mathrm{Sinc}(fNT_0 - mN)$ and $A(f)$ correspond to the orthogonality with respect to $\mathcal{F} = \frac{1}{NT_0}$ and $\mathcal{T} = \frac{T_0}{M}$, respectively.

The operations in Fig. ?? can be extended straightforwardly to other pulse-trains with different subpulses $a(t)$. In fact, we can choose the subpulse to determine the envelope of the FD representation of the pulse-train, which corresponds to the bandwidth of the pulse-train. Meanwhile, we are also free to choose N and T_0 for beneficially manipulating the FD representation of the pulse train under the envelope, which is $\dot{U}(f)$.

2) DDOP as Virtual 2D Pulse: Being a continuous-time function, a pulse can be described either by its TD representation $g(t)$ or FD representation $G(f)$, where these two representations are tightly bound by the (inverse) Fourier transform and therefore dependent on each other. This dependency is exactly the reason why the TFA of the pulse has a lower bound (Gabor limit) of $1/(4\pi)$ corresponding to the Heisenberg uncertainty principle. As a result, although we may be able to present the TF localization of a pulse in the TF domain by illustrating its time and frequency representations together as shown in Fig. ??, the time and frequency variables of the pulse's TF domain are inter-dependent. Therefore, they cannot form a real 2D domain (t, f) to let us design a 2D pulse/filter that can be denoted by $\mathbf{g}(t, f)$. In other words, a pulse is always a 1D function.

On the other hand, we would point out that the delay variable τ and the Doppler variable ν, namely the time and frequency variables of the channel's TF domain, are independent. Therefore, they can form a real 2D domain, and we do have 2D impulse responses for LTV channels, represented by the spreading function $h(\tau, \nu)$.
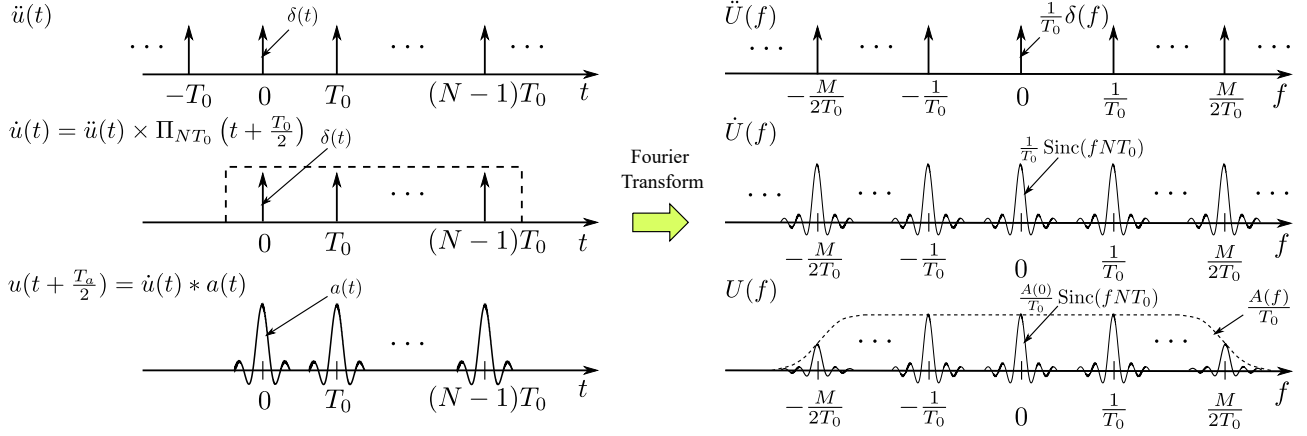
Fig. 12. Derivation of $U(f)$

If we can artificially introduce an extra time variable, a pulse having a pulse-train structure may be considered as a "virtual" 2D pulse to match the signal to the channel. For example, the intervals between the subpulses and the time during an interval may be considered as two potential variables of the pulse-train or the virtual 2D pulse. This is similar to slow time and fast time, which have been widely used in radar waveform design [?], [?]. In particular, for the pulse-train $u(t)$, although its subpulse $a(t)$ has a tightly bound pair of time variable t and frequency variable f, we can repeat the subpulse $a(t)$ and introduce an extra time variable $\dot{t}$, the minimum unit of which is T_0. By letting $t = \tau + \dot{t}$, we can virtually represent $u(t)$ as a 2D function $\mathsf{u}(\tau, \dot{t})$ subject to the following constraints

$$\begin{cases} \tau \in [0, T_0), \\ \dot{t} = nT_0, n = 0, \cdots, N-1, \\ \mathsf{u}(\tau, \dot{t}) = \mathsf{u}(\tau, 0). \end{cases} \quad (59)$$

Note that $\dot{t}$ is independent of τ when (??) holds. Now we may be able to apply the Fourier transform to $\dot{t}$ and obtain a "virtual" 2D domain (τ, ν), where $u(t)$ may be written as $\mathsf{u}(\tau, \nu)$ representing a virtual 2D pulse. The last two constraints in (??) actually correspond to $\dot{U}(f)$, the FD representation of the pulse train under the envelope we mentioned before.

It should be noted that introducing $\dot{t}$ does not mean that we can escape from the tight relation between the unconstrained t and f to create a new dimension for pulse design. As we can see from $u(t)$ and $U(f)$, it just causes the signal to be equally-spaced distributed in both the TD and FD, and partitions the TF region of the signal into MN small "effective" regions. Inside an effective TF region, the orthogonality with respect to the delay resolution $\frac{T_0}{M}$ and the Doppler resolution $\frac{1}{NT_0}$ is achieved by $a(t)$ in $u(t)$ and $\mathrm{Sinc}(fNT_0)$ in $U(f)$, respectively. With the aid of $u(t)$ and $U(f)$, the simplified TF signal localization of DDOP is illustrated in Fig. ??.

3) Comparison of TF Signal Localization: To transmit MN QAM symbols, an MC modulation scheme employs MN orthogonal pulses corresponding to its JTFR and
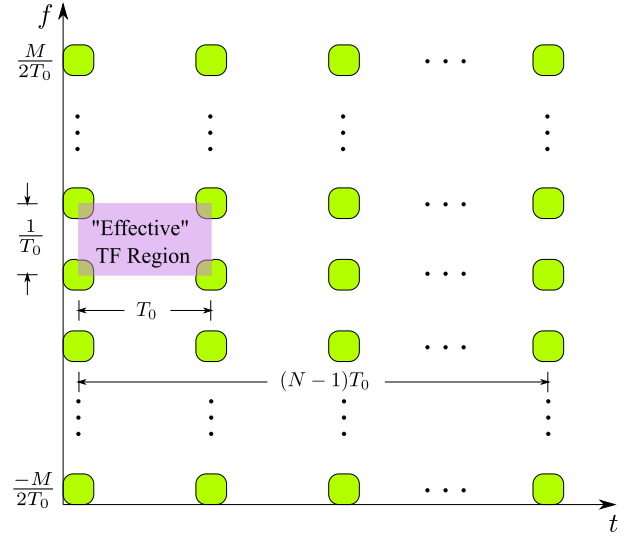

Fig. 13. Simplified TF signal localization of DDOP

results in its own TF signal localization. Based on Fig. ??, the comparison between ODDM and other modulation schemes in terms of their simplified TF signal localization can be schematically illustrated in Fig. ??, where $M = 4$, $N = 2$. It should be noted that for ODDM, M has to be a large enough integer to have a reasonable extension parameter $D = \lceil 2Q/M \rceil$ for the general DDOP. From Fig. ??, one can observe that:

1) For SC modulation, which is a time-division multiplexing scheme, the MN QAM symbols are conveyed by MN square-root Nyquist pulses for the zero-ISI interval $\frac{T_0}{M}$. The pulses are orthogonally overlapped in the TD.

2) For a frequency-division multiplexing (FDM) scheme, such as for example OFDM associated with frequency resolution of $\frac{1}{NT_0}$, MN QAM symbols are conveyed by MN rectangular pulses $\Pi_{NT_0}(t)$ modulated by MN subcarriers, respectively. The pulses are inseparably overlapped in the TD however they are orthogonally overlapped in the FD.
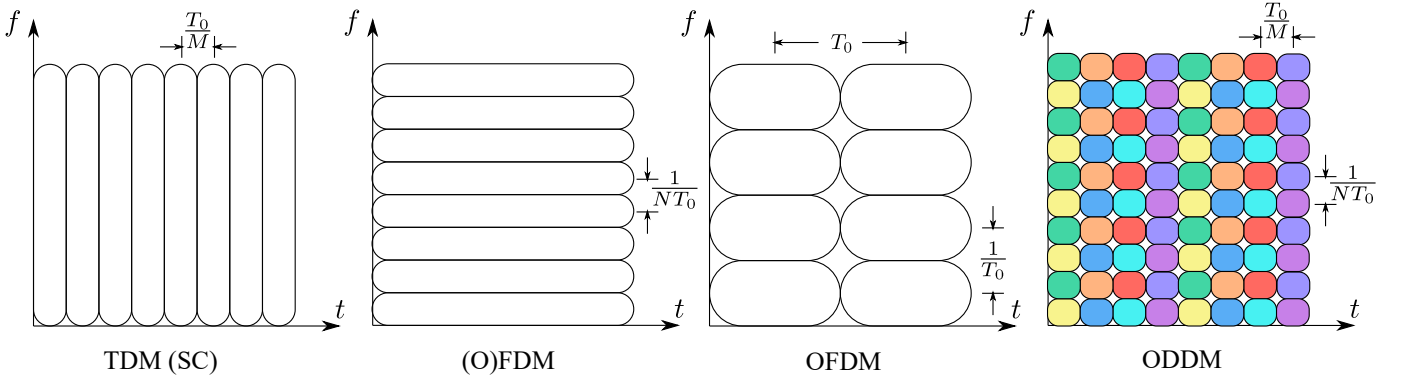
Fig. 14. Comparison of simplified TF signal localization

3) For conventional OFDM having a frequency resolution of $\frac{1}{T_0}$ and time resolution T_0, MN QAM symbols are conveyed by N OFDM symbols, where each OFDM symbol has M rectangular pulses $\Pi_{T_0}(t)$ modulated by M subcarriers, respectively. Since N OFDM symbols are isolated in the TD, the inter-symbol pulses are not overlapped either in the TD or in the FD, while the intra-symbol pulses are inseparably and orthogonally overlapped in the TD and FD, respectively.

4) For ODDM having a frequency resolution of $\frac{1}{NT_0}$ and a time resolution of $\frac{T_0}{M}$, MN QAM symbols are conveyed by M pulse trains $u(t)$ modulated by N subcarriers, respectively. These pulses are overlapped orthogonally in both the TD and FD.

Since the overlapping of pulses is the key to high bandwidth efficiency [?], [?], it is meaningful to investigate the bandwidth efficiency of the ODDM modulation, the pulses of which are overlapped in both the TD and FD. Recall that for the TF region is bounded by B and T_x, we have a DoF around BT_x. Then, for a given M and N, we can calculate the necessary B and T_x for each modulation scheme and obtain their bandwidth efficiencies accordingly.

B. Bandwidth Efficiency

Let the square-root Nyquist pulse be a root raised cosine pulse[7] with roll-off factor ρ. Upon recalling that $T_a = 2Q\frac{T_0}{M}$, we have

$$\eta_{\text{TDM}} = \frac{MN}{\frac{M}{T_0}(1+\rho)((MN-1)\frac{T_0}{M}+T_a)}$$
$$= \frac{1}{(1+\rho)(1+\frac{2Q-1}{MN})}. \quad (60)$$

For (O)FDM, because $|\text{Sinc}(fNT_0)|$ decays as $1/f$, it can be treated as negligibly small beyond the $\mathcal{K}$th zero-crossing

---

[7]Due to its good TF localization, the RRC pulses modulated by spreading code/sequence were adopted by the third generation mobile communication system code-division multiple access communications.

on both sides of the main lobe. In other words, the band-width of $g(t) = \Pi_{NT_0}(t)$ is considered as $B_g = 2\mathcal{K}\frac{1}{NT_0}$. Then, we have

$$\eta_{\text{(O)FDM}} = \frac{MN}{(\frac{MN-1}{NT_0}+\frac{2\mathcal{K}}{NT_0})NT_0} = \frac{1}{1+\frac{2\mathcal{K}-1}{MN}}. \quad (61)$$

Let us now consider CP-OFDM. Because of the delay spread of the channel, $g(t)$ is $\Pi_{T_0+T_{cp}}(t)$ and $T_{cp} = L\frac{T_0}{M}$. Therefore, we have

$$\eta_{\text{CP-OFDM}} = \frac{MN}{(\frac{M-1}{T_0}+\frac{2\mathcal{K}}{T_0})NT_0(1+\frac{L}{M})},$$
$$= \frac{1}{(1+\frac{2\mathcal{K}-1}{M})(1+\frac{L}{M})}. \quad (62)$$

For ODDM, because the cyclic extension of $u(t)$ is equivalent to the frame based CP and CS, we have

$$\eta_{\text{ODDM}} = \frac{MN}{\frac{MN(1+\rho)+N-1}{NT_0}((N-1+2D)T_0+\frac{(M-1+L)T_0}{M}+T_a)},$$
$$= \frac{1}{(1+\rho+\frac{N-1}{MN})(1+\frac{2DM+L+2Q-1}{MN})}. \quad (63)$$

When $2Q \ll M$, we can let $D = 0$. Then, the ODDM frame only needs a single CP corresponding to the delay spread of the channel. For such an ODDM frame with CP only (CP-ODDM), we have

$$\eta_{\text{CP-ODDM}} = \frac{1}{(1+\rho+\frac{N-1}{MN})(1+\frac{L+2Q-1}{MN})}. \quad (64)$$

Considering $\frac{1}{T_0} = 15\text{kHz}$ and the Extended Vehicular A channel model of [?] associated with delay spread 2510ns, a bandwidth efficiency comparison of these modulation schemes is shown in Fig. ??, where $M = 512$, $L = 20$, and $\mathcal{K} = 11$ corresponding to the 99% fractional power containment bandwidth. One can see that ODDM has similar bandwidth efficiency to SC modulation. For a moderately large N and an appropriate ρ, ODDM has better bandwidth efficiency than CP-OFDM, while the (O)FDM with $\mathcal{F} = \frac{1}{NT_0}$ has the highest bandwidth efficiency. Note that the (O)FDM associated with $\mathcal{F} = \frac{1}{NT_0}$ has MN subcarriers and therefore it is extremely "expensive" to implement.
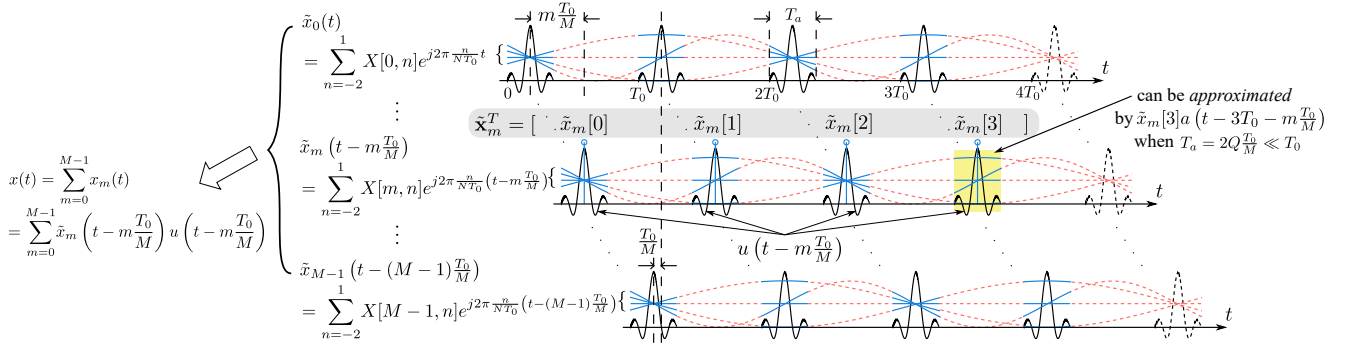
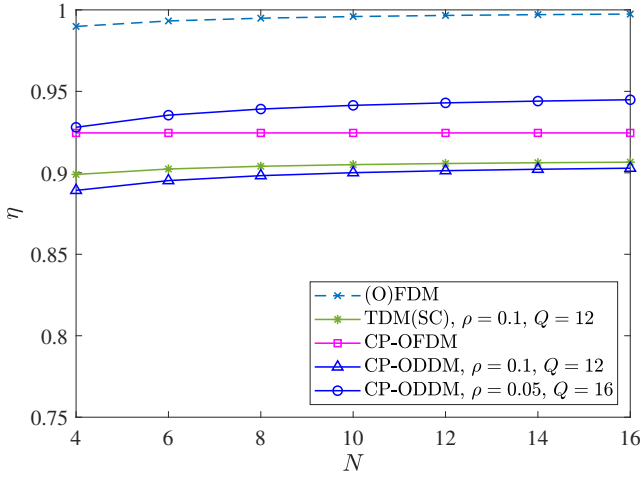Fig. 16. ODDM waveform without CP and CS, when $N = 4$.

Fig. 15. bandwidth efficiency of modulation schemes, $L = 20$.

Fig. 17. Approximate implementation of ODDM

Fig. 18. Simplified approximate implementation of ODDM

C. Implementation Methods

Being a standard MC modulation, ODDM can be implemented straightforwardly via the analog and digital approaches mentioned in Section ??. In particular, we can substitute $\mathcal{T} = \frac{T_0}{M}$ and $\mathcal{F} = \frac{1}{NT_0}$ or $T = NT_0$ into Figs. ??-??, and replace $g(t)$ by $u(t)$. However, due to the long duration of $u(t)$, these direct implementations have a high complexity, even if we generate $\tilde{x}_m(t)$ using the IFFT.

Figure ?? shows an ODDM waveform without CP and CS, when $N = 4$. One can see that due to the limited duration of $a(t)$, $x_m(t)$ namely $\tilde{x}_m(t - m\frac{T_0}{M})$ shaped (multiplied) by $u(t - m\frac{T_0}{M})$ becomes discontinuous with N segments, each with a length T_a. When $T_a \ll T_0$, it has been proved in [?, Appendix A] that instead of using $u(t - m\frac{T_0}{M})$-based pulse shaping, we can generate the discrete samples $\tilde{\mathbf{x}}_m^T$ and then filter them with $a(t)$ to approximate $x_m(t)$. For example, the fourth segment of $x_m(t)$ can be approximately generated by filtering/pulse-shaping $\tilde{x}_m[3]$ with $a(t)$, as shown in Fig. ??. This approximation is actually valid for any other subpulses, as long as $T_a \ll T_0$. As a result, we have ODDM approximated by a filtered OFDM, where the filter is a wideband filter to retain the frequency diversity created by the sampling-induced aliasing, see detailed explanations in Remarks 2
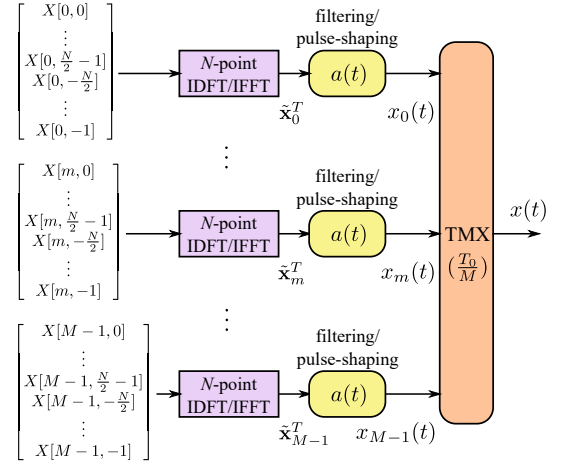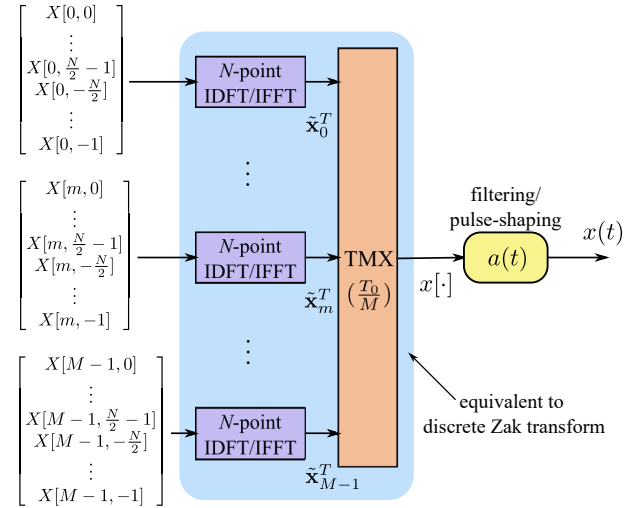
and 4 of [?].

The wideband filtered OFDM based approximation leads to the low-complexity implementation of ODDM shown in Fig. ??, where a sample-wise filtering/pulse-shaping is employed [?]. Moreover, because the M branches of Fig. ?? share the same filter $a(t)$, we can exchange the order of the TMX and the filters to further simplify the implementation. The resultant simplified

approximate implementation of ODDM using a digital TMX and a single wideband filter is shown in Fig. ??.

It is noteworthy that the combination of M branches of N-point IDFT/IFFT having T_0-interval output samples and the $\frac{T_0}{M}$-interval digital TMX is equivalent to a discrete Zak transform [?]. Because a discrete-time OTFS sequence can be generated using the discrete Zak transform [?], as indicated in Remark 4 of [?], a digital namely discrete-time OTFS signal filtered or pulse-shaped by $a(t)$ approximates the ODDM waveform. Meanwhile, it should be noted that the wideband filter $a(t)$ can also be implemented digitally, followed by a DAC having a high enough sampling rate.

D. Potentials for ISAC

At the time of writing, ISAC is regarded as a promising technology for next-generation wireless communications to intelligently utilize the precious spectrum. An ISAC system is essentially a dual-functional radar-communication (DFRC) system, where the backscattered signals are used for estimating object or scattering parameters, such as range and velocity corresponding to delay and Doppler, respectively. Because radar sensing and communication functions have different sometimes even opposing requirements on waveform properties, the main challenge in the ISAC system development is to design a suitable waveform that can simultaneously perform these two tasks well [?].

A waveform is characterized by its format and diverse other parameters including bandwidth, duration, TF resolutions, etc. Regarding the format, because the conventional radar waveforms, such as frequency modulated continuous waveforms and chirp signals, only have limited communication capability [?], the communication waveforms including the SC and MC modulations become the primary choice for DFRC systems [?], [?], [?]. Regarding the parameters, from a radar sensing perspective, the waveform is expected to be wideband with a long duration, which corresponds to both high delay and Doppler resolutions and subsequently good sensing performance. On the other hand, from the communications perspective, a high throughput and low latency require a wideband waveform having a short duration, while a narrowband signal with a relatively long duration can ease the channel equalization. Therefore, first we may have to determine the bandwidth and duration of the ISAC waveform, according to the radar sensing and communication applications under consideration.

Once the bandwidth and duration constraints are given, we can design a communication signal based ISAC waveform, by taking into account the performance metrics of both functions. For communications, the metrics include the achievable rate, the bandwidth efficiency, the equalization complexity, etc. For sensing, considering the classic correlation-based approaches, popular metrics include different characteristics of the TD autocorrelation function, for example, the main-lobe width, the peak to side-lobe level, and the integrated side-lobe level [?]. Without exploiting the information content of the signal,

the correlation based sensing approaches have limited performance, especially in the presence of large Doppler shifts [?]. In particular, for the SC waveform (in combination with spread-spectrum techniques) designed for optimizing the time-domain autocorrelation, the estimation of Doppler/velocity is difficult [?]. Meanwhile, by explicitly exploiting the information content of the signals in MC modulations, radar sensings in the "Modulation Symbol" domain [?], [?] can achieve superior performance over the correlation based approaches.

It should be noted that the radar sensing is exactly constituted by the estimation of the backscattered channel [?], which is also an ESDD channel. The rationale behind the "Modulation Symbol" domain based sensing approaches is simply that the transmit information-bearing symbols are known at the radar receiver and therefore can be used as pilots to perform pilot-based channel estimation. By contrast, the correlation based sensing can be viewed as a blind channel estimation, which usually has inferior performance.

Recall that ODDM is an impulse function based transmission technique designed for ESDD channels. The estimation of the forward communication ESDD channel can be performed straightforwardly with the aid of DD domain pilots. On the other hand, radar sensing or the estimation of the backscatter ESDD channel only requires an appropriately extended frame based CP corresponding to the longer delay spread of the backscatter ESDD channel. Meanwhile, because the ESDD channel estimation is a necessary part of an ODDM receiver, an ODDM system can be interpreted as an ISAC system, where the communication and radar sensing have been seamlessly integrated.

An interesting interpretation for the ISAC capability of the ODDM waveform can be obtained from the characteristics of the ambiguity function of DDOP. Notice that the TD autocorrelation function is an ambiguity function without frequency shift. A more appropriate metric for radar sensing may be the normalized ISL of the ambiguity function defined as [?]

$$\text{ISL}_u = \frac{\int \int_{\mathcal{R}_s} |A_{u,u}(\tau,\nu)|^2 d\tau d\nu}{|\mathcal{R}_s||A_{u,u}(0,0)|^2}, \tag{65}$$

where $\mathcal{R}_s$ denotes the side-lobe region in the DD domain. Further considering the bandwidth and duration constraints and the corresponding delay and Doppler resolutions, we can modify (??) to define a normalized sampled ISL of the ambiguity function as

$$\text{SISL}_u = \frac{\sum_{m=0}^{L-1} \sum_{n=-K}^{K} |A_{u,u}(\frac{m}{\text{W}}, \frac{n}{\text{T}})|^2}{|A_{u,u}(0,0)|^2}, \tag{66}$$

the minimum of which can be achieved by the DDOP.

Meanwhile, because the repeatedly sent pulses in a pulse-Doppler radar can be viewed as a pulse-train [?], one can see that the DDOP in Fig. ?? is exactly a kind of pulse-Doppler radar waveform. Note that the measurement of Doppler/velocity depends on the repeat subpulses or the signal structure of pulse-train. By using

TABLE III
Simulation Parameters

Parameter	Value
Carrier frequency f_c	5 GHz
$1/T_0$	15 kHz
M	512
CP length	$3.125\mu s$
Modulation alphabet	4-QAM
UE speed (km/h)	80, 120, 500

the root Nyquist pulse as the subpulse in the pulse train, we can also benefit the measurement of delay/range in radar sensing function, as well as the alignment of interferences in communication function. As a result, the DDOP-based ODDM combines the key characteristics of radar and communication waveforms and becomes a natural waveform choice for ISAC.

VIII. Numerical results

In this section, simulations are conducted to verify the performance of the ODDM modulation. The simulation parameters are shown in Table ??. For the doubly-selective channel, we adopt the EVA model [?], where each path has a single Doppler generated using Jake's formula $\nu_p = \nu_{\max}\cos(\phi_p)$, the maximum Doppler $\nu_{\max}$ is determined by the user equipment speed and ϕ_p is uniformly distributed over $[-\pi, \pi]$. It is noteworthy that the EVA channel has not only off-grid channel taps on the delay axis, but also possible off-grid Dopplers. Also, a RRC pulse with a roll-off factor of ρ and a duration of $2Q\frac{T_0}{M}$ is employed as $a(t)$.

The power spectral density comparison of the modulated signals with various RRC pulse parameters is shown in Fig. ??. From this figure, we can see that the PSD of the proposed ODDM signals can maintain much low OOBE. In addition, we also see that by tuning the roll-off factor, a trade-off between the excess bandwidth and OOBE can be struck to achieve the desirable bandwidth efficiency.

Let $\dot{x}(t)$ denote the approximated ODDM waveform generated using the approximate implementations in Fig. ?? and Fig. ??. The normalized mean square error between the approximated ODDM waveform and the exact ODDM signal is defined as

$$\text{NMSE}_{\dot{x}} = \frac{\int |\dot{x}(t) - x(t)|^2 dt}{\int |x(t)|^2 dt}, \qquad (67)$$

where $x(t)$ is the exact ODDM signal, given in (??) with $g(t) = u(t)$. The NMSE results for various roll-off factor ρ and duration of pulses is shown in Fig. ??. We can see form this figure that the NMSE is not significantly affected by the parameter Q of the pulse, and the NMSE decreases with the roll-off factor of the employed pulse. The figure also demonstrates that the simplified approximate implementations in Fig. ?? and Fig. ?? can generate very close ODDM waveform as the NMSE between them is below -40dB.
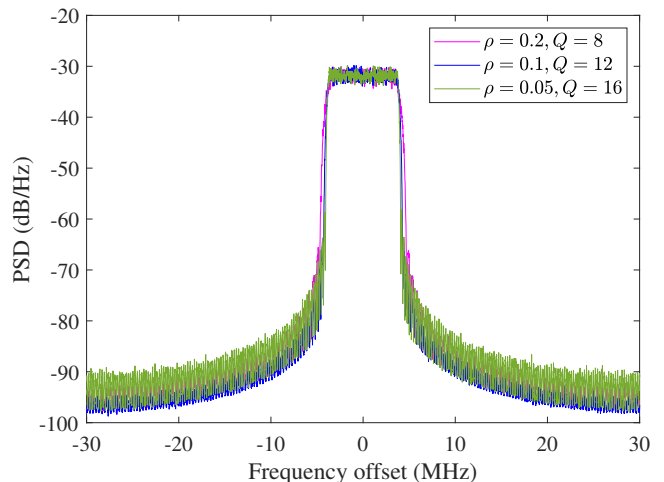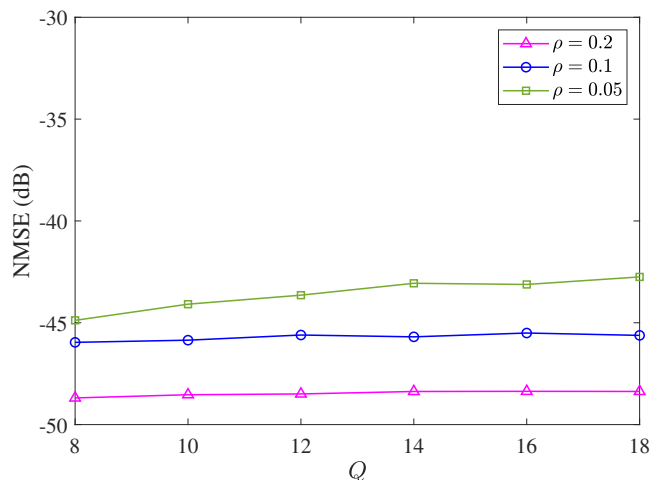

Fig. 19. PSD, $M = 512$, $N = 32$, 4-QAM.

Fig. 20. NMSE of approximated ODDM waveform, $M = 512$, $N = 32$, 4-QAM.

We now evaluate the BER performance of the uncoded ODDM modulation. The signal detection is based on the message passing algorithm [?] and the DD domain channel matrix $\mathbf{H}$ in (??). Fig. ?? shows the BER of the ODDM signals with $M = 512$, $N = 32$ and 4-QAM. In the simulation, the maximum UE speed is 500km/h and the roll-off factor of the pulse is chosen as 0.05, 0.1 and 0.2. This figure demonstrates that the ODDM signal achieves almost the same BER for various roll-off factor of the pulse.

Fig. ?? shows the BER of the proposed ODDM system with 4-QAM signals, $M = 512$, $N = 32$, and UE speed of 80km/h, 120km/h and 500km/h. The figure demonstrates that the ODDM signals achieve almost the same BER performance over the high-mobility channels regardless of the UE speed, which means that ODDM signals are robust against Doppler shifts. Meanwhile, Fig. ?? illustrates the BER of the proposed ODDM system with 4-QAM signals, $M = 512$, $N = \{16, 32, 64\}$ and UE speed of 500km/h. The figure shows that the BER performance of the ODDM
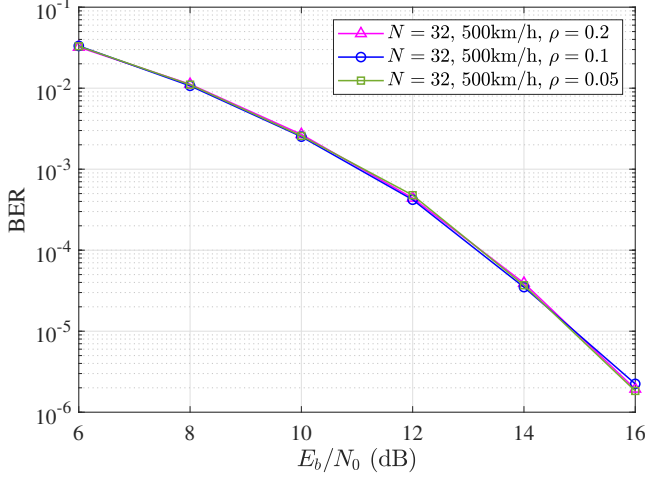
Fig. 21. BER comparison, $M = 512$, $N = 32$, 4-QAM.
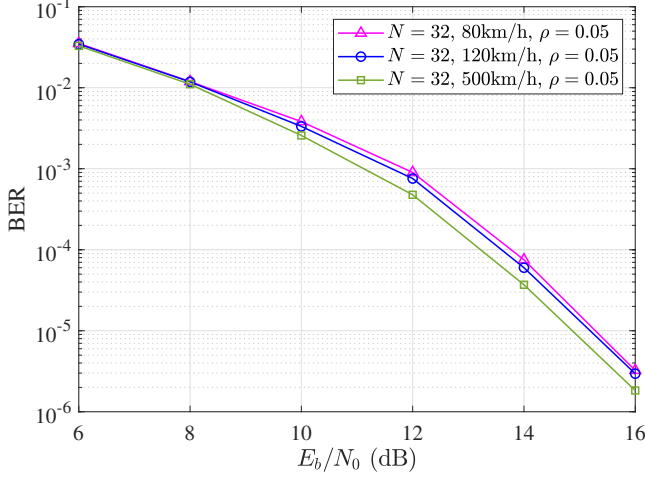
Fig. 22. BER comparison, $M = 512$, $N = 32$, 4-QAM, $\rho = 0.05$, $Q = 16$.
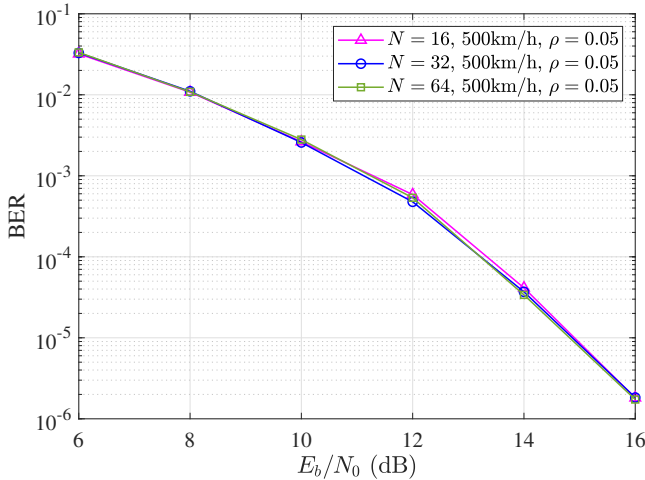
Fig. 23. BER comparison, $M = 512$, 4-QAM, $\rho = 0.05$, $Q = 16$.
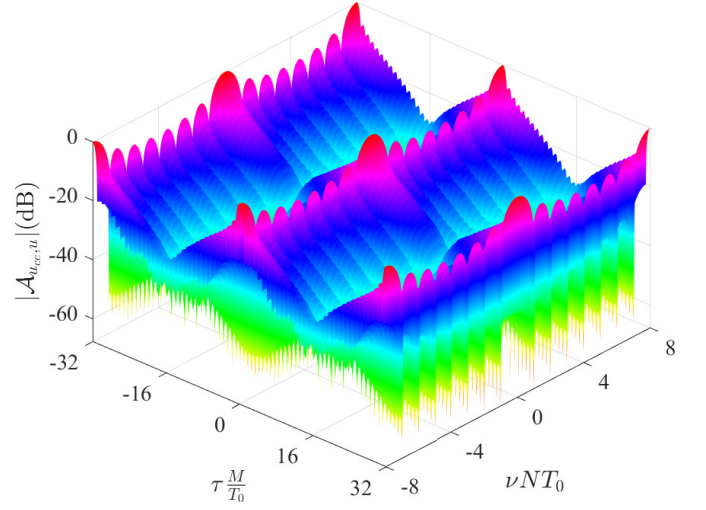
Fig. 24. $|\mathcal{A}_{u_{ce},u}(\tau, \nu)|$ for $M = 32$ and $N = 8$.
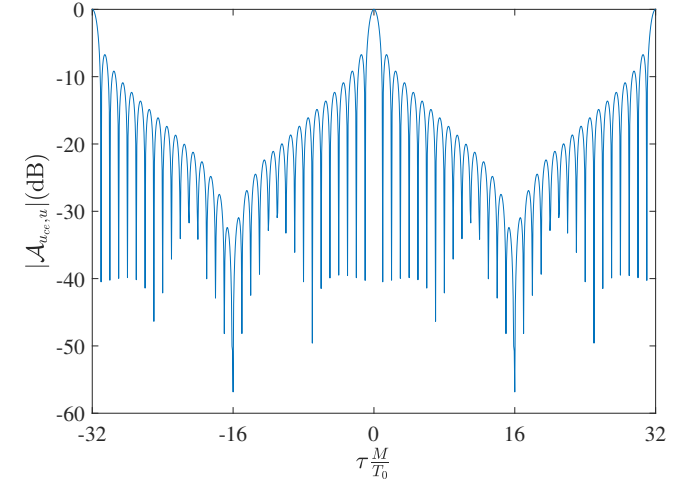
Fig. 25. $|\mathcal{A}_{u_{ce},u}(\tau, 0)|$ for $M = 32$ and $N = 8$.

signals also remains almost the same for various values of N.

A three-dimensional plot of the ambiguity function in (??), a.k.a ambiguity surface [?], is shown in Fig. ??, where $\mathcal{F} = \frac{1}{NT_0}$, $\mathcal{T} = \frac{T_0}{M}$ with $M = 32$, $N = 8$, and the RRC pulse $a(t)$ has the roll-off factor $\rho = 0.1$ and $Q = 20$. Because for this parameter setting, the extension parameter for CP and CS is $D = 2$, and we adopt the general DDOP design. The corresponding 2D plots of $|\mathcal{A}_{u_{ce},u}(\tau, \nu)|$ with $\nu = 0$ and $\tau = 0$ are also given in Figs. ?? and ??, respectively. One can see that with appropriate CP and CS, the DDOP can achieve the sufficient orthogonality within $|m| \leq M - 1$ and $|n| \leq N - 1$. For $|m| \geq M$ or $|n| \geq N$, the ambiguity function repeats with time period T_0 and frequency period $\frac{1}{T_0}$, if we further extend the CP and CS. These figures indicate the great potential of the DDOP-based ODDM signals for ISAC applications.
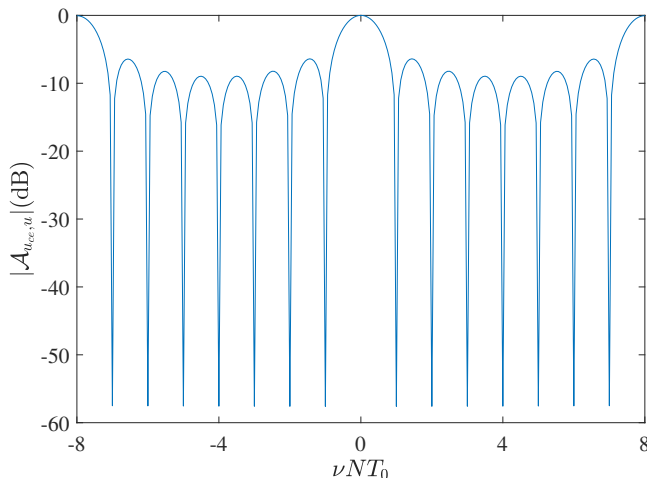
Fig. 26. $|\mathcal{A}_{u_{ce},u}(0,\nu)|$ for $M = 32$ and $N = 8$.

IX. Conclusions and Future Research

An in-depth look into DDOP and the corresponding ODDM modulation was offered to unveil their unique characteristics. We first revisited the conventional time-frequency domain multi-carrier modulation schemes in terms of their transmission strategy, the channel-oriented orthogonal or biorthogonal pulses and the resultant bandwidth efficiency. Next, we addressed the time-varying property of the DD domain channel's 2D impulse response and clarified the unique and innovative transmission strategy of ODDM. Conventional TFMC modulation pulse/waveform design principles are governed by the WH frame theory, which ensures global (bi)orthogonality across the whole TF domain. For practical systems having a limited bandwidth and frame duration, the MC modulation pulse design just needs to satisfy the sufficient (bi)orthogonality inside the TF region of interest. Then, by reformulating the (bi)orthogonal pulse design problem, we revealed the "mystery" of DDOP, and justified its hitherto unknown benefits, which are achieved without violating the classic WH frame theory. Finally, we presented the salient properties of the ODDM modulation, including its signal localization, bandwidth efficiency, implementation benefits, and its ISAC potentials.

In the conventional TFMC modulation schemes such as OFDM, the orthogonality among OFDM subcarriers facilitates the following benefits, which justify its wide adoption in many communications systems [?], [?], [?], [?], [?], [?], [?]: (A.1) the immunity to frequency-selective multipath effects; (A.2) the low-complexity single-tap equalization; (A.3) the trivial bandwidth partitioning; (A.4) the straightforward adoption to multiple-input multiple-output systems. Nonetheless, the following OFDM deficiencies are also widely recognized: (D.1) the high peak-to-average power ratio that encumbers the power amplifier design; (D.2) the bandwidth efficiency erosion due to the CP overhead and owing to the unloaded subcarriers inserted at the band edge for controlling the OOBE; (D.3) the sensitivity to CFO that includes Doppler shift and oscillator mismatch; (D.4) the OOBE that affects the co-existence of asynchronous users. Over the past six decades, a variety of transmission and reception techniques have been developed, including orthogonal frequency division multiplexing combined with index modulation [?] and the recent multi-band discrete Fourier transform-spread-orthogonal frequency division multiplexing amalgamating with index modulation [?], in order to mitigate the OFDM deficiencies. However, they tend to compromise some of the key OFDM benefits.

In the proposed DDMC modulation scheme such as ODDM, modulating information in the DD domain and orthogonality between the ODDM subcarriers with respect to the Doppler resolution and between multiple ODDM symbols with respect to the delay resolution brings about a number of benefits: (A.1) both time- and frequency-diversity can be attained in doubly-selective channels, leading to reliable transmissions in these hostile channels; (A.2) low channel estimation pilot overhead and less frequent channel estimation; (A.3) having a preferable ambiguity function, which is attractive for future ISAC applications; (A.4) reduced CP cost, since only one CP is required for a single data frame, resulting in an improved bandwidth efficiency; (A.5) having a moderate PAPR for an appropriate pulse $u(t)$ and for suitable values of M and N. We remark that ODDM does not achieve a common decomposition of arbitrary LTV channels into independent subchannels (as OFDM does for LTI channels). However, because of its orthogonality with respect to the delay and Doppler resolutions of the channel, ODDM better matches the delay and Doppler characteristics of the channel, and it is expected to lead to lower implementation complexity for both communications and sensing applications. As a novel and fundamentally new waveform, ODDM or DDMC in general is still in its early stage of development. There are many challenging open questions to be answered in future research. Some of them are listed in the following.

- To detect ODDM signals over doubly selective channels, the conventional low-complexity single-tap equalization does not provide satisfactory performance. Existing OTFS detectors based on message passing [?], [?], on linear maximal-ratio combining [?], and on minimum mean squared error [?] can be extended to ODDM [?], which can offer good performance, but at a high computational cost. Receiver designs based on deep-learning to explore the DD domain channel properties and signal structures are also of interest. In this regard, receivers exhibiting performance vs. complexity trade-offs are essential for practical systems.
- Circuit impairments including CFO, DC offset, IQ imbalance, phase noise, etc., constitute a critical issue in practical transceiver designs [?]. Due to its orthogonality with respect to the fine JTFR, the ODDM performance erosion under realistic circuit impairments requires further investigation. Also, pilot or signal designs conceived for compensating these

impairments in OFDM systems [?], [?], [?] may be also extended to ODDM systems.

- Similar to conventional OFDM waveform and its relatives, ODDM can also have diverse beneficial variants. For example, ODDM may be further evolved to DFT-S-ODDM, as a bridge between SC transmission and DDMC transmission. Like OFDM-IM [?] and MC-CDMA [?], [?], ODDM can also be combined with index modulation or conventional CDMA technologies achieving good bandwidth-/power-/energy-efficiencies.

- For ODDM to be applicable to practical systems, its transmitter and receiver must be flexibly scalable both in terms of antennas and users. The implementation architecture of MIMO-ODDM systems, as well as the associated resource allocation strategies for both multi-user systems and multi-cell systems, are relevant future directions.

- Performance analysis in terms of its the achievable rates, error probability and channel code design for doubly-selective channels constitute further research challenges to be tackled.

- To provide flexibility in term of resource allocation and system optimization, it may be of interest to combine ODDM with interference cancellation based strategies, such as rate-splitting multiple access [?] and non-orthogonal multiple access [?] for both multiple-input single-output and MIMO systems.

- Security is at essence in wireless systems, where the unique physical layer channel properties of legitimate users can be exploited to provide security to complement upper-layer security relying on secret keys. For ODDM signals, due to the associated spreading of signal, there is an improved grade of physical layer security. This is another compelling research item.

- Furthermore, the popular reconfigurable intelligent surfaces can also be combined with ODDM as they are capable of improving the coverage of space–air–ground integrated networks in the era of the sixth generation mobile communication system.

- Note that ISAC is an emerging service in future systems. Although much progress has been reported in the last decade, there are many open research problems to be solved in this research area. On the other hand, as a novel MC modulation waveform, the application of ODDM and DDMC to ISAC will also have many fundamental and practical questions to be answered, such as the capacity or achievable rates of the signals over various channels, sensing limits or performance bound, potential trade-offs between communications and sensing, practical considerations including the effect of imperfect frequency offsets, timing and frequency synchronizations, etc., on system performance. Solving these problems will help pave the way of practical applications of ODDM or DDMC in future ISAC systems.
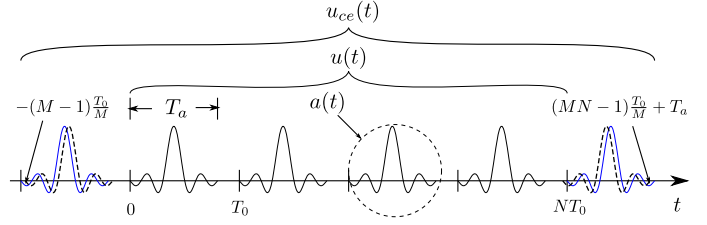

Fig. 27. $u_{ce}(t)$ for $D = 1$

Appendix A
Proof of Proposition 2

Since the period of $g(t)$ is $\frac{T}{N}$, we have

$$g(t) = g(t + \dot{n}\frac{T}{N}), 0 \leq \dot{n} \leq N - 1, \qquad (68)$$

for $0 \leq t < \frac{T}{N}$. Then, bearing in mind that $T = 1/\mathcal{F}$, the ambiguity function of $g(t)$ is given by

$$
\begin{aligned}
\mathcal{A}_{g,g}(0, n\mathcal{F}) &= \int_0^{T_g} g(t)g^*(t)e^{-j2\pi n\mathcal{F}t}dt, \\
&= \sum_{\dot{n}=0}^{N-1} \int_{\dot{n}\frac{T}{N}}^{(\dot{n}+1)\frac{T}{N}} g(t)g^*(t)e^{-j2\pi n\mathcal{F}t}dt, \\
&= \sum_{\dot{n}=0}^{N-1} e^{-j2\pi \frac{n\dot{n}}{N}} \times \int_0^{\frac{T}{N}} g(t)g^*(t)e^{-j2\pi n\mathcal{F}t}dt, \\
&= \delta(n), \qquad (69)
\end{aligned}
$$

for $|n| \leq N - 1$, and the last equality is based on the fact that $g(t)$ is normalized to unit energy. This completes the proof.

Appendix B
Proof of Proposition 3

Let us first check the periodicity of $u_{ce}(t)$ within the range of $-(M - 1)\frac{T_0}{M} \leq t \leq (MN - 1)\frac{T_0}{M} + T_a$, which corresponds to the start of the first subpulse of $u(t+(M-1)\frac{T_0}{M})$ and the end of the last subpulse of $u(t-(M-1)\frac{T_0}{M})$, respectively. Recall that

$$u(t) = \sum_{n=0}^{N-1} a(t - nT_0), \qquad (70)$$

we can divide $u(t)$ into N segments, where $u(t) = \sum_{n=0}^{N-1} u_n(t)$ and the nth segment is given by

$$u_n(t) = \begin{cases} u(t) & nT_0 \leq t < (n+1)T_0 \\ 0 & \text{otherwise} \end{cases}. \qquad (71)$$

Let $D = \lceil T_a/T_0 \rceil$. If $D = 1$, we have

$$u_n(t) = a(t - nT_0), \qquad (72)$$

which implies that the periodicity within $-(M - 1)\frac{T_0}{M} \leq t \leq (MN - 1)\frac{T_0}{M} + T_a$ can be obtained by cyclically extending $u(t)$ in (??) to

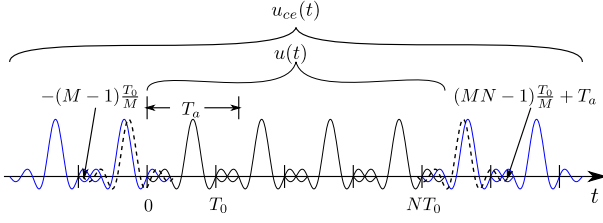$$u_{ce}(t) = \sum_{n=-1}^{N} a(t - nT_0). \qquad (73)$$

Fig. 28. $u_{ce}(t)$ for $D = 2$

Similarly, when $D > 1$, the periodicity can be obtained by cyclically extending $u(t)$ in (??) to

$$u_{ce}(t) = \sum_{n=-D}^{N-1+D} a(t - nT_0). \tag{74}$$

Two examples of $u_{ce}(t)$ associated with $D = 1, 2$ are shown in Fig. ?? and Fig. ??, respectively, where the first subpulse of $u(t+(M-1)\frac{T_0}{M})$ and the last subpulse of $u(t-(M-1)\frac{T_0}{M})$ are also plotted with dashed lines.

Next, let us verify the ambiguity functions. Due to the aforementioned periodicity of $u_{ce}(t)$, we have

$$u_{ce}(t) = u_{ce}(t + \dot{n}T_0), 0 \le \dot{n} \le N - 1. \tag{75}$$

for $m\frac{T_0}{M} \le t \le m\frac{T_0}{M} + T_u$, where $|m| \le M - 1$ and $T_u = (N-1)T_0 + T_a$. Then, the cross ambiguity function between $u_{ce}(t)$ and $u(t)$ for $|n| \le N - 1$ and $|m| \le M - 1$ is given by

$$\mathcal{A}_{u_{ce},u}\left(m\frac{T_0}{M}, n\frac{1}{NT_0}\right)$$

$$= \int_{m\frac{T_0}{M}}^{m\frac{T_0}{M}+T_u} u_{ce}(t)u^*\left(t - m\frac{T_0}{M}\right) e^{-j2\pi n\frac{1}{NT_0}(t-m\frac{T_0}{M})}dt,$$

$$= \int_0^{T_u} u_{ce}\left(t + m\frac{T_0}{M}\right) u^*(t)e^{-j\frac{2\pi n}{NT_0}t}dt,$$

$$\stackrel{(a)}{=} \sum_{\dot{n}=0}^{N-1} \int_{\dot{n}T_0}^{\dot{n}T_0+T_a} u_{ce}\left(t + m\frac{T_0}{M}\right) a^*(t - \dot{n}T_0)e^{-j\frac{2\pi n}{NT_0}t}dt,$$

$$\stackrel{(b)}{=} \sum_{\dot{n}=0}^{N-1} e^{-j2\pi\frac{n\dot{n}}{N}} \int_0^{T_a} u_{ce}\left(t + m\frac{T_0}{M}\right) a^*(t)e^{-j\frac{2\pi n}{NT_0}t}dt,$$

$$= \delta(n) \int_0^{T_a} u_{ce}\left(t + m\frac{T_0}{M}\right) a^*(t)dt$$

$$= \delta(n)\delta(m), \tag{76}$$

where $\stackrel{(a)}{=}$ and $\stackrel{(b)}{=}$ are due to (??) and (??), respectively. This completes the proof.