Performance Analysis of Coded OTFS Systems Over High-Mobility Channels

Shuangyang Li, *Graduate Student Member, IEEE*, Jinhong Yuan, *Fellow, IEEE*, Weijie Yuan, *Member, IEEE*, Zhiqiang Wei, *Member, IEEE*, Baoming Bai, *Senior Member, IEEE*, and Derrick Wing Kwan Ng, *Fellow, IEEE*

***Abstract*— Orthogonal time frequency space (OTFS) modulation is a recently developed multi-carrier multi-slot transmission scheme for wireless communications in high-mobility environments. In this paper, the error performance of coded OTFS modulation over high-mobility channels is investigated. We start from the study of conditional pairwise-error probability (PEP) of the OTFS scheme, based on which its performance upper bound of the coded OTFS system is derived. Then, we show that the coding improvement for OTFS systems depends on the squared Euclidean distance among codeword pairs and the number of independent resolvable paths of the channel. More importantly, we show that there exists a fundamental trade-off between the coding gain and the diversity gain for OTFS systems, i.e., the diversity gain of OTFS systems improves with the number of resolvable paths, while the coding gain declines. Furthermore, based on our analysis, the impact of channel coding parameters on the performance of the coded OTFS systems is unveiled. The error performance of various coded OTFS systems over high-mobility channels is then evaluated. Simulation results demonstrate a significant performance improvement for OTFS modulation over the conventional orthogonal frequency division multiplexing (OFDM) modulation over high-mobility channels. Analytical results and the effectiveness of the proposed code design are also verified by simulations with the application of both classical and modern codes for OTFS systems.**

***Index Terms*— OTFS modulation, diversity analysis, code design, high-mobility, fading channels.**

Manuscript received June 18, 2020; revised December 20, 2020 and March 1, 2021; accepted April 2, 2021. Date of publication April 14, 2021; date of current version September 10, 2021. The work of Derrick Wing Kwan Ng was supported in part by the UNSW Digital Grid Futures Institute, UNSW, Sydney, under a cross-disciplinary fund scheme and in part by the Australian Research Council's Discovery Project under Grant DP210102169. The article will be presented in part at the IEEE International Conference on Communication Workshops 2021. The associate editor coordinating the review of this article and approving it for publication was D. R. Brown. *(Corresponding author: Weijie Yuan.)*

Shuangyang Li, Jinhong Yuan, and Derrick Wing Kwan Ng are with the School of Electrical Engineering and Telecommunications, University of New South Wales, Sydney, NSW 2052, Australia (e-mail: shuangyang.li@unsw.edu.au; j.yuan@unsw.edu.au; w.k.ng@unsw.edu.au).

Weijie Yuan is with the Department of Electrical and Electronic Engineering, Southern University of Science and Technology, Shenzhen 518055, China (e-mail: yuanwj@sustech.edu.cn).

Zhiqiang Wei was with the School of Electrical Engineering and Telecommunications, University of New South Wales, Sydney, NSW 2052, Australia. He is now with the Department of Electrical Engineering-Electronics-Information Technology, Friedrich-Alexander University Erlangen-Nuremberg, 91058 Erlangen, Germany (e-mail: zhiqiang.wei@fau.de).

Baoming Bai is with the State Key Laboratory of ISN, Xidian University, Xi'an 710071, China (e-mail: bmbai@mail.xidian.edu.cn).

Color versions of one or more figures in this article are available at https://doi.org/10.1109/TWC.2021.3071493.

Digital Object Identifier 10.1109/TWC.2021.3071493

I. Introduction

BEYOND the fifth-generation (B5G) wireless communication systems are required to accommodate various emerging applications in high-mobility environments, such as mobile communications on board aircraft (MCA), low-earth-orbit satellites (LEOSs), high speed trains, and unmanned aerial vehicles (UAVs) [2]–[4]. Currently deployed orthogonal frequency division multiplexing (OFDM) modulation may not be able to support efficient and reliable communications in such scenarios [5]. Therefore, as a potential solution to supporting heterogeneous requirements of B5G wireless systems in high-mobility scenarios, the recently proposed orthogonal time frequency space (OTFS) modulation has attracted substantial attention [6].

In high-mobility scenarios, wireless channels are usually doubly-dispersive in the time-frequency (TF) domain [7], [8]. In specific, the time dispersion is caused by the effect of multi-path, while the frequency dispersion is caused by the Doppler shifts. Conventionally, OFDM modulation can efficiently mitigate the intersymbol interference (ISI) induced by the time dispersion by introducing a cyclic prefix (CP). However, the success of OFDM modulation relies deeply on maintaining the orthogonality among all the sub-carriers. Note that perfect orthogonality is highly impractical at the receiver, especially in high-mobility environments, due to the exceedingly large frequency dispersion, and consequently, the performance of conventional OFDM systems is unsatisfactory in such scenarios [5]. On the other hand, by invoking the two-dimensional (2D) symplectic finite Fourier transform (SFFT), OTFS modulates the information symbols in the delay-Doppler (DD) domain, where the channel parameters are relatively stable compared to those in the TF domain [7]. More importantly, it can be shown that by modulating the information symbols in the DD domain rather than the TF domain, each symbol principally experiences the whole fluctuations of the TF channel over an OTFS frame and thus OTFS modulation offers the potential of exploiting the full channel diversity, achieving a better error performance compared to that of the conventional OFDM modulation in a high-mobility environment [6]. Furthermore, with the domain transformation performed in OTFS modualtion, one can represent the high dynamic channel parameters in the TF domain equivalently by a sparse presentation in the DD domain. This unique property suggests that the acquisition of channel state information (CSI) for OTFS systems can be performed with a low pilot signaling

overhead and that the symbol detection for OTFS systems can be carried out with a low complexity [9], [10].

To unleash the potential of OTFS modulation, some recent works have focused on the implementation of practical OTFS systems. For example, a low-complexity modem structure for OTFS systems was proposed in [11]. In particular, it showed that the OTFS modulation can be efficiently implemented with simple pre- and post-processing units based on the conventional OFDM modulator. Besides, to reduce the signaling overhead, an embedded pilot-aided channel estimation method for OTFS modulation was proposed in [9]. By taking advantage of the sparse representation of the wireless channel in the DD domain, this channel estimation method only requires one pilot symbol with a small number of guard symbols in the DD domain. Furthermore, in order to reduce the detection complexity, various variations of sum-product algorithms (SPAs) have been proposed to facilitate the OTFS detection. In specific, Raviteja *et. al* proposed an SPA-based OTFS detector, where the inter-Doppler interference (IDI) was approximated as a Gaussian variable to reduce the detection complexity [10]. A variational Bayes detector for OTFS systems was proposed in [12]. The basic idea of this detector is to approximate the corresponding *a posteriori* distribution of the optimal detection by exploiting the Kullback-Leibler (KL) divergence. In [13], a cross domain iterative detection for OTFS modulation was proposed, where the algorithm applies basic estimation/detection approaches to both the time domain and DD domain and iteratively updates the extrinsic information from two domains with the unitary transformation.

Although the aforementioned excellent works have provided guidelines for practical OTFS system designs, the theoretical error performance advantages of OTFS systems over the conventional OFDM systems have not been thoroughly studied yet, especially for coded cases. We note that there are several previous works [14]–[16] on the error performance analysis of OTFS systems. However, these works mainly considered the uncoded case and their analysis may not be directly extended to the coded cases. Consequently, the error performance analysis of coded OTFS systems is still missing in the literature to the best of our knowledge. As commonly recognized, channel coding is an efficient tool to combat fading and channel impairments and thus is a key enabler for reliable communications between users with high-mobility [17]. For OTFS modulation, the 2D transformation from the TF domain to the DD domain provides a potential of exploiting the full TF diversity. In this case, a good channel code needs to couple the coded symbols to the 2D OTFS modulation, in order to exploit the full diversity and in the meantime maximize the coding gain. However, it is still unknown what is the key coding parameter determining the coding gain for OTFS modulation. Therefore, in order to facilitate the implementation of practical OTFS systems, the error performance analysis for coded OTFS systems needs to be investigated.

In this paper, we aim to analyze the error performance of coded OTFS systems over general Rayleigh-fading *underspread wide-sense stationary uncorrelated scattering (WSSUS)* channels. To this end, we start from the study of conditional pairwise-error probability (PEP) [18], [19] for a given channel realization. In order to obtain an accurate performance analysis, we consider two cases of the OTFS transmission depending on the number of independent resolvable paths of the channel and derive the corresponding conditional PEPs. Since the exact unconditional PEP is generally intractable [20], we resort to the application of some proper bounding techniques to study the conditional PEP and derive the unconditional performance upper bounds. Based on the unconditional performance bound, the impact of channel coding parameters on the performance of OTFS modulation is unveiled. In particular, we find that the squared Euclidean distance between a pair of codewords is the key parameter that determines the coding gain for coded OTFS systems, given the number of independent resolvable paths. Therefore, the code design criterion is formulated to optimize the coding gain by maximizing the minimum squared Euclidean distance between all codeword pairs. The main contributions of this paper can be summarized as follows.

- We investigate the conditional PEP of OTFS systems for a given channel realization by studying the pairwise Euclidean distance between OTFS codewords. Based on the conditional PEP, we derive the unconditional performance upper bounds for OTFS systems, according to the number of independent resolvable paths. We also show a few important properties of the codeword difference matrix. Furthermore, according to the derived bounds, we define the coding gain and diversity gain of OTFS systems. More importantly, we show that the coding gain of OTFS systems depends on the squared Euclidean distance and the number of independent resolvable paths.
- According to the derived performance bounds, we show that there is a fundamental trade-off between the diversity gain and the coding gain for OTFS systems. In particular, the diversity gain of OTFS systems improves with the number of independent resolvable paths, while the coding gain declines.
- Based on the derived performance bounds, we propose our code design criterion to optimize the coding gain, which is to maximize the minimum squared Euclidean distance of among all codeword pairs. In order words, traditional good codes with a large minimum Euclidean distance can be directly applied to OTFS systems.
- We demonstrate a significant performance improvement achieved by the coded OTFS modulation over the coded OFDM modulation over high-mobility channels by numerical simulations. We also provide numerical results of coded OTFS systems over high-mobility channels with various channel codes, such as classical convolutional codes and state-of-art low-density parity-check (LDPC) codes. Our performance analysis and code design are explicitly verified by these results.

Notations: The blackboard bold letter $\mathbb{A}$ and $\mathbb{H}$ denote the constellation set and an arbitrary subspace, respectively; The notations $(\cdot)^{\mathrm{T}}$, $(\cdot)^*$, $\|\cdot\|$, $(\cdot)^{-1}$, and $(\cdot)^{\mathrm{H}}$ denote the transpose, the conjugate, the Euclidean norm, the inverse, and the Hermitian operations for a matrix, respectively; $\mathbb{E}[\cdot]$ denote the expectation; $\det(\cdot)$, $\mathrm{tr}(\cdot)$, and $\mathrm{vec}(\cdot)$ denote the determinant, the trace, and the vectorization operation; $\mathrm{diag}\{\cdot\}$

denotes the diagonal matrix; "$\otimes$" denotes the Kronecker product operator; $\mathbf{F}_N$ and $\mathbf{I}_M$ denote the discrete Fourier transform (DFT) matrix of size $N \times N$ and the identity matrix of size $M \times M$, respectively; $Q(\cdot)$ denotes the tail distribution function of the standard normal distribution, $\delta(\cdot)$ denotes the Dirac delta function, $I_0(\cdot)$ denotes the zero-order modified Bessel function of the first kind, respectively; $\Pr(\cdot)$ denotes the probability of an event and $p(\cdot)$ denotes the probability density function (PDF), respectively; $f(\cdot)$ denotes an arbitrary function; $(f(\cdot))_{\max}$ and $(f(\cdot))_{\min}$ denote the maximum and minimum values of function $f(\cdot)$, respectively; $\lceil \cdot \rceil$ denotes the round up operation.

II. System Model

In this section, we first review the underspread WSSUS channel property and the OTFS concept, and then introduce the considered system model.

A. Underspread WSSUS Channel Model

The underspread WSSUS channel model is a commonly adopted model for real-world wireless channels [7], [21], whose stochastic properties can match well with the actual channel measurement results [7], [22]. Generally speaking, the second-order statistics of a given linear time-varying channel depend on four variables, i.e., time, frequency, delay, and Doppler, respectively. In the pioneered paper from Bello [23], this complex dependence relationship is simplified to only two variables by introducing the assumption of WSSUS. In particular, the US assumption implies that the different delay shifts associated with different resolvable paths are uncorrelated, while the WSS assumption indicates that the time domain channel taps are jointly wide-sense stationary with respect to the time variable, which implies that the different Doppler shifts associated with different resolvable paths are uncorrelated. The rationale behind the WSSUS assumption is that each DD pair corresponds to a scatterer with a specific reflectivity and the reflectivities of any two distinct scatterers are generally uncorrelated [7]. In specific, let us consider the DD domain representation of a linear time-varying channel [6], [7],

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

In (1), P is the number of resolvable paths, while h_i, τ_i, and ν_i are the channel coefficients, delay shifts, and Doppler shifts corresponding to the i-th path, respectively. Based on the WSSUS assumption, it can be shown that the delay shift τ_i, Doppler shift ν_i, and channel coefficient h_i associated with the i-th path are uncorrelated to those from the j-th path, if $i \neq j$. Furthermore, since the delay shift and Doppler shift correspond to the relative distance and relative velocity, which are two different physical attributes of the scatterer, they are usually modeled as uncorrelated variables for each path [7]. However, the channel coefficient h_i strongly dependents on the statistical characterization of the reflectivity strength of the underlying path, which is usually related to the corresponding delay shift τ_i and Doppler shift ν_i.

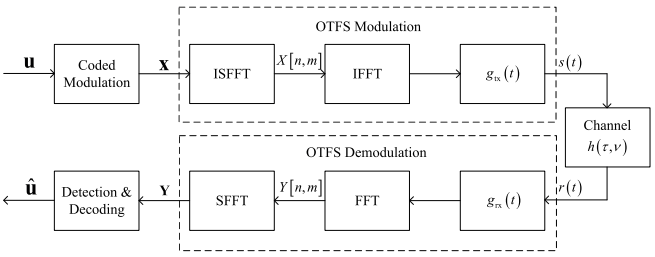


Fig. 1. The block diagram of an OTFS system.

On the other hand, the concept of the underspread channel is also commonly considered for wireless communication designs. Let $\tau_{\max}$ and $\nu_{\max}$ be the maximum delay and Doppler values among all resolvable paths.[1] Then, a channel is said to be *underspread*, if $4\tau_{\max}\nu_{\max} \leq 1$ [7]. The underspread property implies that the TF domain channel response has to be constant within a region that is no smaller than 1, where the corresponding time domain and frequency domain intervals are commonly referred to as the *coherence time* and *coherence bandwidth*, respectively. It is of practical importance that we study the underspread channel because all real-world wireless channels are virtually underspread and most of the multi-carrier systems are designed according to the underspread property, including both OFDM and OTFS systems [7].

In this paper, we adopt the underspread WSSUS channel model and investigate the error performance of OTFS modulation according to the channel properties discussed above.

B. Coded OTFS System Model

Without loss of generality, we consider a coded OTFS system as shown in Fig. 1. Let M be the number of sub-carriers and N be the number of time slots for each OTFS symbol, respectively. An information sequence $\mathbf{u}$ is channel-encoded and then modulated into $\mathbf{x} \in \mathbb{A}^{MN}$ with length MN. Let us arrange $\mathbf{x}$ into a 2D matrix $\mathbf{X} \in \mathbb{A}^{M\times N}$, i.e., $\mathbf{x} \overset{\Delta}{=} \mathrm{vec}(\mathbf{X})$, representing the symbols in the DD domain, whose (k,l)-th element $x[k,l]$ is the modulated signal in the k-th Doppler and l-th delay grid [6], for $0 \leq k \leq N-1, 0 \leq l \leq M-1$. Each transmitted symbol in the TF domain $X[n,m], 0 \leq n \leq N-1, 0 \leq m \leq M-1$, is then obtained based on $\mathbf{X}$ via the inverse symplectic fast Fourier transform (ISFFT) as follows [6]

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

A brief diagram regarding the DD and TF domain transformation is shown in Fig. 2, where Δf is the frequency spacing between adjacent sub-carriers and $T = 1/\Delta f$ is the corresponding TF domain time slot duration. Therefore, each OTFS system takes the total bandwidth of $M\Delta f$ and symbol duration NT. In particular, the sampling time $1/(M\Delta f)$ and sampling

[1]The maximum Doppler value is given by $\nu_{\max} = \frac{v}{c} f_c$, where v is the relative speed, c is the speed of light, and f_c is the carrier frequency, respectively. The maximum delay shift is given by $\tau_{\max} = \frac{d_{\max}}{c}$, where $d_{\max}$ is the maximum distance difference among the P channel paths.

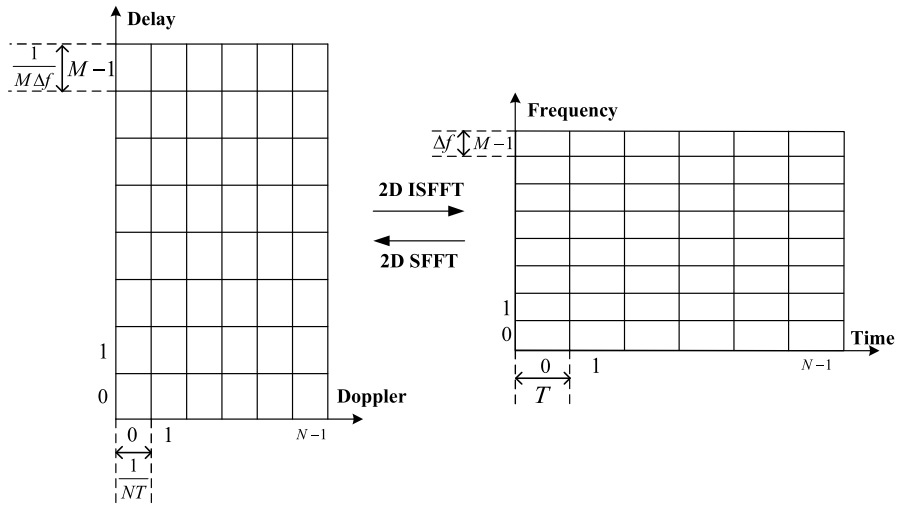


Fig. 2. The diagram of the transformation between the DD domain and the TF domain.

frequency $1/(NT)$ are referred to as the *delay resolution* and the *Doppler resolution* of the DD grid, respectively [10], which indicate how precise the acquisition of the channel delay and Doppler can be for the underlying OTFS system. After the domain transformation (ISFFT), the TF domain transmitted symbol $X[n,m]$ is then modulated via a conventional OFDM modulator. The time domain OTFS signal $s(t)$ is written as

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

where $g_{\rm tx}(t)$ denotes the pulse shaping filter. As described above, for OFDM-based OTFS implementation, the OTFS modulator can be viewed as a concatenation of a precoder (ISFFT) and a conventional OFDM modulator [11], where the OFDM modulator consists of an inverse FFT (IFFT) block and a pulse shaping filter $g_{\rm tx}(t)$. Let $\bar{w}(t)$ be the additive white Gaussian noise process with one-sided PSD N_0. According to the DD domain channel representation in (1), the received signal can be written as

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

Let $g_{\rm rx}(t)$ be the filter adopted at the receiver side. The received symbols $Y[n,m]$ in the TF domain are then obtained by

$$Y[n,m]=\int r(t)\,g_{\rm rx}^{*}(t-nT)\,e^{-j2\pi m\Delta f(t-nT)}dt. \quad (5)$$

Substituting (4) into (5) and after similar manipulations as in [10], (5) can be simplified as

$$Y[n,m]=\sum_{n'=0}^{N-1}\sum_{m'=0}^{M-1}H_{n,m}[n',m']X[n',m']+\bar{w}[n,m], \quad (6)$$

where $\bar{w}[n,m]$ is the corresponding TF domain noise sample and $H_{n,m}[n',m']$ is the TF domain channel response, given by

$$\begin{aligned}&H_{n,m}[n',m']\\&=\int\int h(\tau,\nu)\,A_{g_{\rm tx},g_{\rm rx}}\Big((n-n')T-\tau,(m-m')\Delta f-\nu\Big)\\&\quad\times e^{j2\pi\left(\nu+m'\Delta f\right)\left((n-n')T-\tau\right)}e^{j2\pi\nu n'T}d\tau d\nu. \quad (7)\end{aligned}$$

In (7), the function $A_{g_{\rm tx},g_{\rm rx}}(\tau_\Delta,\nu_\Delta)$ is the so-called *cross-ambiguity* function, which indicates the interference level between the TF domain symbols due to the channel dispersion and is given by [10]

$$A_{g_{\rm tx},g_{\rm rx}}(\tau_\Delta,\nu_\Delta)\triangleq\int g_{\rm tx}(t)\,g_{\rm rx}^{*}(t-\tau_\Delta)\,e^{j2\pi\nu_\Delta\Delta ft}dt. \quad (8)$$

Hence, the received symbols $y[k,l]$ in the DD domain are obtained by performing the SFFT on the TF domain received symbols $Y[n,m]$ which are written as

$$y[k,l]=\frac{1}{\sqrt{NM}}\sum_{n=0}^{N-1}\sum_{m=0}^{M-1}Y[n,m]\,e^{-j2\pi\left(\frac{nk}{N}-\frac{ml}{M}\right)}+w[k,l], \quad (9)$$

where $w[k,l]$ denotes the equivalent AWGN samples in the DD domain. In specific, the DD domain received symbols $y[k,l]$ can be arranged into the 2D received symbol matrix $\mathbf{Y}$ according to the DD grid, whose (k,l)-th element is $y[k,l]$. For the ease of presentation and analysis, we consider the vector form representation of the input-output relationship of OTFS system in the DD domain based on (9) in the sequel.

C. Vector Form Representation of OTFS

Let $\mathbf{x}\triangleq{\rm vec}(\mathbf{X})\in\mathbb{A}^{MN}$ and $\mathbf{y}\triangleq{\rm vec}(\mathbf{Y})\in\mathbb{A}^{MN}$ denote the vector forms of the transmitted symbols $\mathbf{X}$ and the received symbols $\mathbf{Y}$ in the DD domain, respectively. According to (9), we have

$$\mathbf{y}=\mathbf{H}_{\rm eff}\mathbf{x}+\mathbf{w}, \quad (10)$$

where $\mathbf{w}$ is the corresponding noise vector and $\mathbf{H}_{\rm eff}$ of size $MN\times MN$ is the *effective channel matrix* in the DD domain.

Assuming that both $g_{\rm tx}(t)$ and $g_{\rm rx}(t)$ are rectangular pulses, with a reduced CP frame format, the effective channel matrix $\mathbf{H}_{\rm eff}$ is given by [24]

$$\mathbf{H}_{\rm eff} = \sum_{i=1}^{P} h_i \left(\mathbf{F}_N \otimes \mathbf{I}_M\right) \mathbf{\Pi}^{l_i+\iota_i} \mathbf{\Delta}^{k_i+\kappa_i} \left(\mathbf{F}_N^{\rm H} \otimes \mathbf{I}_M\right), \quad (11)$$

where $\mathbf{\Pi}$ is the permutation matrix (forward cyclic shift), i.e.,

$$\mathbf{\Pi} = \begin{bmatrix} 0 & \cdots & 0 & 1 \\ 1 & \ddots & 0 & 0 \\ \vdots & \ddots & \ddots & \vdots \\ 0 & \cdots & 1 & 0 \end{bmatrix}_{MN \times MN}, \quad (12)$$

and $\mathbf{\Delta} = \text{diag}\{z^0, z^1, \ldots, z^{MN-1}\}$ is a diagonal matrix with $z \triangleq e^{\frac{j2\pi}{MN}}$ [24]. In (11), l_i and k_i are the delay and Doppler indices corresponding to the i-th path, respectively, and we have

$$\tau_i = \frac{l_i + \iota_i}{M\Delta f}, \quad \nu_i = \frac{k_i + \kappa_i}{NT}. \quad (13)$$

Notice that the term $-1/2 \le \iota_i \le 1/2$ and $-1/2 \le \kappa_i \le 1/2$ denote the fractional delay and Doppler shifts which correspond to the fractional shifts from the nearest delay and Doppler indices [10]. It should be noted that the typical value of the sampling time $1/M\Delta f$ in the delay domain is usually sufficiently small. Therefore, the impact of fractional delays in typical wide-band systems can be neglected[2] [8]. However, in this paper we will still consider the fractional delay case in order to obtain some general conclusions on OTFS modulation. Specifically, we have $0 \le l_i + \iota_i \le l_{\max}$, where $l_{\max}$ is the maximum delay index satisfying $l_{\max} = \lceil M\Delta f \tau_{\max} \rceil$ [10]. We also have $-k_{\max} \le k_i + \kappa_i \le k_{\max}$, where $k_{\max}$ is the maximum Doppler index satisfying $k_{\max} = \lceil NT\nu_{\max} \rceil$ [10]. In particular, let $\boldsymbol{\omega}_\tau$ and $\boldsymbol{\omega}_\nu$ denote the vectors of delay indices and Doppler indices, respectively, i.e., $\boldsymbol{\omega}_\tau = [l_1 + \iota_1, l_2 + \iota_2, \ldots, l_P + \iota_P]^{\rm T}$, $\boldsymbol{\omega}_\nu = [k_1 + \kappa_1, k_2 + \kappa_2, \ldots, k_P + \kappa_P]^{\rm T}$, respectively. Therefore, according to [14], (10) can be rewritten as

$$\mathbf{y} = \mathbf{\Phi}_{\boldsymbol{\omega}_\tau, \boldsymbol{\omega}_\nu}(\mathbf{x})\, \mathbf{h} + \mathbf{w}, \quad (14)$$

where $\mathbf{\Phi}_{\boldsymbol{\omega}_\tau, \boldsymbol{\omega}_\nu}(\mathbf{x})$ is referred to as the *equivalent codeword matrix* and it is a concatenated matrix of size $MN \times P$ constructed by the column vector $\mathbf{\Xi}_i \mathbf{x}$, i.e.,

$$\mathbf{\Phi}_{\boldsymbol{\omega}_\tau, \boldsymbol{\omega}_\nu}(\mathbf{x}) = \left[\mathbf{\Xi}_1 \mathbf{x} \quad \mathbf{\Xi}_2 \mathbf{x} \quad \cdots \quad \mathbf{\Xi}_P \mathbf{x}\right], \quad (15)$$

and $\mathbf{\Xi}_i$ is given by

$$\mathbf{\Xi}_i \triangleq \left(\mathbf{F}_N \otimes \mathbf{I}_M\right) \mathbf{\Pi}^{l_i+\iota_i} \mathbf{\Delta}^{k_i+\kappa_i} \left(\mathbf{F}_N^{\rm H} \otimes \mathbf{I}_M\right), \quad 1 \le i \le P. \quad (16)$$

In (14), $\mathbf{h}$ is the channel coefficient vector of size $P \times 1$, i.e., $\mathbf{h} = [h_1, h_2, \ldots, h_P]^{\rm T}$, where the elements in $\mathbf{h}$ are assumed to be independent and identically distributed complex Gaussian random variables. Besides, we assume a uniform power delay and Doppler profile of the channel such that the channel coefficient h_i has mean μ and variance $1/(2P)$ per real dimension for $1 \le i \le P$ and is independent from the delay and Doppler indices [21]. In particular, we note that if $\mu = 0$, $|h_i|$ follows the Rayleigh distribution, which will be considered as a special case in our error performance analysis and code design. Based on (14), the error performance analysis of the OTFS systems will be conducted in the next section.

[2]It should be noted that the effect of fractional Doppler shifts can be mitigated by adding TF domain windows [25].

III. Error Performance Analysis

In order to investigate the theoretical error performance of the coded OTFS systems, we assume that ideal channel state information (CSI) is available at the receiver, including $\mathbf{h}$, $\boldsymbol{\omega}_\tau$, and $\boldsymbol{\omega}_\nu$.[3] We note that matrix $\mathbf{\Phi}_{\boldsymbol{\omega}_\tau, \boldsymbol{\omega}_\nu}(\mathbf{x})$ depends on $\boldsymbol{\omega}_\tau$, $\boldsymbol{\omega}_\nu$, and the transmitted symbol vector $\mathbf{x}$. Therefore, for a given channel realization, we define the *conditional Euclidean distance* $d^2_{\mathbf{h}, \boldsymbol{\omega}_\tau, \boldsymbol{\omega}_\nu}(\mathbf{x}, \mathbf{x}')$ between a pair of codewords $\mathbf{x}$ and $\mathbf{x}'$ ($\mathbf{x} \ne \mathbf{x}'$) as

$$d^2_{\mathbf{h}, \boldsymbol{\omega}_\tau, \boldsymbol{\omega}_\nu}(\mathbf{x}, \mathbf{x}') = d^2_{\mathbf{h}, \boldsymbol{\omega}_\tau, \boldsymbol{\omega}_\nu}(\mathbf{e}) \triangleq \left\|\mathbf{\Phi}_{\boldsymbol{\omega}_\tau, \boldsymbol{\omega}_\nu}(\mathbf{e})\mathbf{h}\right\|^2 = \mathbf{h}^{\rm H} \mathbf{\Omega}_{\boldsymbol{\omega}_\tau, \boldsymbol{\omega}_\nu}(\mathbf{e})\, \mathbf{h}, \quad (17)$$

where $\mathbf{e} = \mathbf{x} - \mathbf{x}'$ is the corresponding codeword difference (error) sequence and $\mathbf{\Omega}_{\boldsymbol{\omega}_\tau, \boldsymbol{\omega}_\nu}(\mathbf{e}) = \left(\mathbf{\Phi}_{\boldsymbol{\omega}_\tau, \boldsymbol{\omega}_\nu}(\mathbf{e})\right)^{\rm H} \left(\mathbf{\Phi}_{\boldsymbol{\omega}_\tau, \boldsymbol{\omega}_\nu}(\mathbf{e})\right)$ is referred to as the *codeword difference matrix*. Without loss of generality and for notational simplicity, we henceforth drop the subscript of $\mathbf{\Omega}_{\boldsymbol{\omega}_\tau, \boldsymbol{\omega}_\nu}(\mathbf{e})$ and we now have

$$\mathbf{\Omega}(\mathbf{e}) = \begin{bmatrix} \mathbf{e}^{\rm H}\mathbf{\Xi}_1^{\rm H}\mathbf{\Xi}_1\mathbf{e} & \mathbf{e}^{\rm H}\mathbf{\Xi}_1^{\rm H}\mathbf{\Xi}_2\mathbf{e} & \cdots & \mathbf{e}^{\rm H}\mathbf{\Xi}_1^{\rm H}\mathbf{\Xi}_P\mathbf{e} \\ \mathbf{e}^{\rm H}\mathbf{\Xi}_2^{\rm H}\mathbf{\Xi}_1\mathbf{e} & \mathbf{e}^{\rm H}\mathbf{\Xi}_2^{\rm H}\mathbf{\Xi}_2\mathbf{e} & & \vdots \\ \vdots & & \ddots & \vdots \\ \mathbf{e}^{\rm H}\mathbf{\Xi}_P^{\rm H}\mathbf{\Xi}_1\mathbf{e} & \cdots & \cdots & \mathbf{e}^{\rm H}\mathbf{\Xi}_P^{\rm H}\mathbf{\Xi}_P\mathbf{e} \end{bmatrix}. \quad (18)$$

The conditional PEP is upper-bounded by [18], [19]

$$\Pr\left(\mathbf{x}, \mathbf{x}' | \mathbf{h}, \boldsymbol{\omega}_\tau, \boldsymbol{\omega}_\nu\right) \le \exp\left(-\frac{E_s}{4N_0} d^2_{\mathbf{h}, \boldsymbol{\omega}_\tau, \boldsymbol{\omega}_\nu}(\mathbf{x}, \mathbf{x}')\right), \quad (19)$$

where E_s is the average symbol energy. Note that the codeword difference matrix $\mathbf{\Omega}(\mathbf{e})$ is positive semidefinite Hermitian with a rank r, where $r \le P$. Let us denote by $\{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_P\}$ the eigenvectors of $\mathbf{\Omega}(\mathbf{e})$ and $\{\lambda_1, \lambda_2, \ldots, \lambda_P\}$ the corresponding nonnegative real eigenvalues sorted in the descending order, where $\lambda_i > 0$ for $1 \le i \le r$ and $\lambda_i = 0$ for $r+1 \le i \le P$. Thus, (19) can be further expanded as [18], [19]

$$\Pr\left(\mathbf{x}, \mathbf{x}' | \mathbf{h}, \boldsymbol{\omega}_\tau, \boldsymbol{\omega}_\nu\right) \le \exp\left(-\frac{E_s}{4N_0} \sum_{i=1}^{r} \lambda_i \left|\tilde{h}_i\right|^2\right), \quad (20)$$

where $\tilde{h}_i = \mathbf{h} \cdot \mathbf{v}_i$, for $1 \le i \le r$. It can be shown that $\left\{\tilde{h}_1, \tilde{h}_2, \ldots, \tilde{h}_r\right\}$ are independent complex Gaussian random variables with mean $\mu_{\tilde{h}_i} = \mathbb{E}[\mathbf{h}] \cdot \mathbf{v}_i$ and variance $1/(2P)$ per real dimension. Thus, it is obvious that $|\tilde{h}_i|$ follows the Rician distribution with a Rician factor $K_i = \left|\mu_{\tilde{h}_i}\right|^2$ [18], and

[3]Some references regarding the channel estimation for OTFS modulation can be found in [9]. Meanwhile, recent advances of machine learning based methods can also potentially improve the channel estimation and detection performance of OTFS modulation [26]–[29].

its PDF is given by

$$p\left(\left|\tilde{h}_i\right|\right) = 2P\left|\tilde{h}_i\right|\exp\left(-P\left|\tilde{h}_i\right|^2 - PK_i\right) I_0\left(2P\left|\tilde{h}_i\right|\sqrt{K_i}\right). \quad (21)$$

In the following, we will target on the analysis of the unconditional PEP. To this end, we aim to calculate the average of (20) over the channel distribution according to (21). More specifically, we will discuss two important cases depending on the number of independent resolvable paths P.

Remark 1: It has been defined in the previous works [14]–[16], [30] that the rank of $\mathbf{\Omega}(\mathbf{e})$ is the **diversity** gain of the OTFS modulation. Specifically, it has been shown in [16] that the diversity gain of uncoded OTFS modulation systems can be one but the full diversity can be obtained by suitable precoding schemes. Furthermore, [14] has shown that the full diversity can be achieved almost surely for the case of $P = 2$ when the frame size is sufficiently large, even for uncoded OTFS modulation systems.

A. Error Performance Analysis for Coded OTFS Systems

Notice that $\mathbf{h}$, $\boldsymbol{\omega}_\tau$, and $\boldsymbol{\omega}_\nu$ are independent from each other with the assumptions of WSSUS channel and the uniform power delay profile [7]. Therefore, the unconditional PEP can be derived by firstly averaging (20) over $|\tilde{h}_i|$ term by term which results in

$$\Pr\left(\mathbf{x}, \mathbf{x}' | \boldsymbol{\omega}_\tau, \boldsymbol{\omega}_\nu\right) \leq \prod_{i=1}^{r} \frac{1}{1 + \frac{E_s}{4N_0}\cdot\frac{\lambda_i}{P}} \exp\left(-\frac{K_i \frac{E_s}{4N_0}\cdot\frac{\lambda_i}{P}}{1 + \frac{E_s}{4N_0}\cdot\frac{\lambda_i}{P}}\right). \quad (22)$$

Furthermore, we consider a special case where $K_i = 0$ and $|\tilde{h}_i|$ follows the Rayleigh distribution, i.e., $|h_i|$ also follows the Rayleigh distribution. In the case of **Rayleigh fading**, (22) can be further simplified as

$$\Pr\left(\mathbf{x}, \mathbf{x}' | \boldsymbol{\omega}_\tau, \boldsymbol{\omega}_\nu\right) \leq \left(\prod_{i=1}^{r} \lambda_i / P\right)^{-1} \left(\frac{E_s}{4N_0}\right)^{-r} = \frac{1}{\prod_{i=1}^{r}\lambda_i}\left(\frac{E_s}{4N_0 P}\right)^{-r}, \quad (23)$$

It should be noted that (23) is consistent with the analysis in [15]. On the other hand, the PEP in (23) depends on the delay and Doppler indices $\boldsymbol{\omega}_\tau$ and $\boldsymbol{\omega}_\nu$. In order to derive the unconditional PEP, we need to find the statistical distribution for the term $\prod_{i=1}^{r}\lambda_i$ regarding the delay and Doppler indices. Unfortunately, such a task is generally intractable [20] and is normally handled by applying the Monte Carlo method without providing any important insight. Instead of resorting to the Monte Carlo method, we apply proper bounding techniques to evaluate the value of $\prod_{i=1}^{r}\lambda_i$ in order to obtain some general results about the unconditional PEP. Specifically, we have the following property.

Property 1 (Gram Matrix [31]): Let $\bar{\mathbf{u}}_i \triangleq \mathbf{\Xi}_i\mathbf{e}$, for $1 \leq i \leq P$, where $\mathbf{\Xi}_i$ is given by (16). Then, $\{\bar{\mathbf{u}}_1, \bar{\mathbf{u}}_2, \ldots \bar{\mathbf{u}}_P\}$ form a list of vectors chosen from the P-dimensional complex inner-product subspace $\mathbb{H}^P$. Thus, the codeword difference matrix $\mathbf{\Omega}(\mathbf{e})$ is positive semidefinite Hermitian and it is a *Gram matrix* corresponding to the vectors $\{\bar{\mathbf{u}}_1, \bar{\mathbf{u}}_2, \ldots \bar{\mathbf{u}}_P\}$.

Based on the property of $\mathbf{\Omega}(\mathbf{e})$, we now introduce four important Lemmas, which will be served as the building blocks for our error performance analysis for coded OTFS systems.

Lemma 1 (Main Diagonal Elements of $\mathbf{\Omega}(\mathbf{e})$): The main diagonal elements of the codeword difference matrix $\mathbf{\Omega}(\mathbf{e})$ are of the same value $d_{\mathrm{E}}^2(\mathbf{e})$, where $d_{\mathrm{E}}^2(\mathbf{e}) = \mathbf{e}^{\mathrm{H}}\mathbf{e}$ is the squared Euclidean distance for a pair of codewords $\mathbf{x}$ and $\mathbf{x}'$ with the corresponding to the error sequence $\mathbf{e}$.

Proof: By considering (15), the i-th diagonal element of $\mathbf{\Omega}(\mathbf{e})$ is given by $\mathbf{e}^{\mathrm{H}}\mathbf{\Xi}_i^{\mathrm{H}}\mathbf{\Xi}_i\mathbf{e}$, and it is equal to the inner product of $\bar{\mathbf{u}}_i$. Furthermore, we have

$$\begin{aligned}\mathbf{\Xi}_i^{\mathrm{H}}\mathbf{\Xi}_i &= \left(\mathbf{F}_N \otimes \mathbf{I}_M\right)\left(\mathbf{\Delta}^{k_i+\kappa_i}\right)^{\mathrm{H}} \times \left(\mathbf{\Pi}^{l_i+\iota_i}\right)^{\mathrm{H}}\mathbf{\Pi}^{l_i+\iota_i}\mathbf{\Delta}^{k_i+\kappa_i}\left(\mathbf{F}_N^{\mathrm{H}} \otimes \mathbf{I}_M\right) \\ &= \left(\mathbf{F}_N \otimes \mathbf{I}_M\right)\mathbf{\Delta}^{-k_i-\kappa_i}\mathbf{\Pi}^{-l_i-\iota_i} \times \mathbf{\Pi}^{l_i+\iota_i}\mathbf{\Delta}^{k_i+\kappa_i}\left(\mathbf{F}_N^{\mathrm{H}} \otimes \mathbf{I}_M\right) \quad (24) \\ &= \mathbf{I}_{MN}, \quad (25)\end{aligned}$$

where (24) is due to the properties of $\mathbf{\Pi}$ and $\mathbf{\Delta}$, respectively, and (25) is due to the property of the Kronecker product. Therefore, the term $\mathbf{e}^{\mathrm{H}}\mathbf{\Xi}_i^{\mathrm{H}}\mathbf{\Xi}_i\mathbf{e}$ is further simplified as $d_{\mathrm{E}}^2(\mathbf{e})$. This completes the proof of Lemma 1.

Lemma 2 (Trace of $\mathbf{\Omega}(\mathbf{e})$): The trace of the codeword difference matrix $\mathbf{\Omega}(\mathbf{e})$ is $Pd_{\mathrm{E}}^2(\mathbf{e})$. Equivalently, the summation of the first r eigenvalues satisfies $\sum_{i=1}^{r}\lambda_i = Pd_{\mathrm{E}}^2(\mathbf{e})$.

Proof: This lemma is a straightforward extension of Lemma 1 by noticing that the $\lambda_i = 0$ for $r+1 \leq i \leq P$.

Lemma 3 (Lower Bound on $\sum_{i=1}^{r} 1/\lambda_i$): The summation of first r eigenvalue inverses of $\mathbf{\Omega}(\mathbf{e})$ is lower-bounded by

$$\sum_{i=1}^{r}\frac{1}{\lambda_i} \geq \frac{r^2}{Pd_{\mathrm{E}}^2(\mathbf{e})}, \quad (26)$$

where the equality holds if $\mathbf{\Omega}(\mathbf{e})$ is a diagonal matrix, i.e., $\mathbf{\Omega}(\mathbf{e}) = \mathrm{diag}\left\{d_{\mathrm{E}}^2(\mathbf{e}), \ldots, d_{\mathrm{E}}^2(\mathbf{e})\right\}$.

Proof: The proof is given in Appendix A.

Lemma 4 (Lower Bound on $\sum_{i=1}^{r}\lambda_i^2$): The summation of the first r squared eigenvalues of $\mathbf{\Omega}(\mathbf{e})$ is lower-bounded by

$$\sum_{i=1}^{r}\lambda_i^2 \geq \frac{P^2}{r}\left(d_{\mathrm{E}}^2(\mathbf{e})\right)^2, \quad (27)$$

where the equality holds if $\mathbf{\Omega}(\mathbf{e})$ is a diagonal matrix, i.e., $\mathbf{\Omega}(\mathbf{e}) = \mathrm{diag}\left\{d_{\mathrm{E}}^2(\mathbf{e}), \ldots, d_{\mathrm{E}}^2(\mathbf{e})\right\}$.

Proof: Notice that the first r eigenvalues of $\mathbf{\Omega}(\mathbf{e})$ are all positive. Therefore, we apply the Cauchy-Schwarz inequality and the following holds

$$\sum_{i=1}^{r}\lambda_i^2 \geq \frac{1}{r}\left(\sum_{i=1}^{r}\lambda_i\right)^2 = \frac{P^2}{r}\left(d_{\mathrm{E}}^2(\mathbf{e})\right)^2, \quad (28)$$

where the equality is achieved when the eigenvalues of $\mathbf{\Omega}(\mathbf{e})$ are of the same value, e.g., $\mathbf{\Omega}(\mathbf{e}) = \mathrm{diag}\left\{d_{\mathrm{E}}^2(\mathbf{e}), \ldots, d_{\mathrm{E}}^2(\mathbf{e})\right\}$. This completes the proof of Lemma 4.

The above Lemmas show some important properties of the codeword difference matrix $\mathbf{\Omega}(\mathbf{e})$. Based on these properties of $\mathbf{\Omega}(\mathbf{e})$, we can now consider the following lower bounds of the eigenvalue product.

Theorem 1 (Lower Bound on $\prod_{i=1}^{r}\lambda_i$): The product of the first r eigenvalues of $\mathbf{\Omega}(\mathbf{e})$ is lower-bounded by

$$\prod_{i=1}^{r}\lambda_i \geq \left(d_{\mathrm{E}}^2(\mathbf{e})\right)^r \exp\left(r - d_{\mathrm{E}}^2(\mathbf{e})\sum_{i=1}^{r}\frac{1}{\lambda_i}\right), \tag{29}$$

where the equality holds if $\mathbf{\Omega}(\mathbf{e})$ is a diagonal matrix, i.e., $\mathbf{\Omega}(\mathbf{e}) = \mathrm{diag}\left\{d_{\mathrm{E}}^2(\mathbf{e}),\ldots,d_{\mathrm{E}}^2(\mathbf{e})\right\}$.

Proof: The proof is given in Appendix B.

It should be noted that (29) still depends on the channel parameters $\boldsymbol{\omega}_\tau$ and $\boldsymbol{\omega}_\nu$. To obtain an unconditional lower bound, we apply an approximation to the lower bound in (29), which is summarized in the following Theorem.

Theorem 2 (Approximated Lower Bound on $\prod_{i=1}^{r}\lambda_i$): The product of the first r eigenvalues of $\mathbf{\Omega}(\mathbf{e})$ can be approximately lower-bounded by

$$\prod_{i=1}^{r}\lambda_i \gtrsim \left(d_{\mathrm{E}}^2(\mathbf{e})\right)^r, \tag{30}$$

where the approximation is exact if $\mathbf{\Omega}(\mathbf{e})$ is a diagonal matrix, i.e., $\mathbf{\Omega}(\mathbf{e}) = \mathrm{diag}\left\{d_{\mathrm{E}}^2(\mathbf{e}),\ldots,d_{\mathrm{E}}^2(\mathbf{e})\right\}$.

Proof: The proof is given in Appendix C.

Based on Theorem 2, it is not hard to verify that the eigenvalue product can be approximated by the term $\left(d_{\mathrm{E}}^2(\mathbf{e})\right)^r$, regardless of the specific distributions of the delay and Doppler indices. Meanwhile, the above approximation is quite insightful in the sense that it only relates to the rank of the codeword difference matrix $\mathbf{\Omega}(\mathbf{e})$ (this is actually the diversity gain as we will introduce later) and the squared Euclidean distance $d_{\mathrm{E}}^2(\mathbf{e})$, which does not depend on the exact value of the delay and Doppler indices. In particular, the approximation becomes exact if $\mathbf{\Omega}(\mathbf{e}) = \mathrm{diag}\left\{d_{\mathrm{E}}^2(\mathbf{e}),\ldots,d_{\mathrm{E}}^2(\mathbf{e})\right\}$, which can be interpreted as the projections of the error sequence $\mathbf{e}$ onto each independent resolvable path, i.e., $\bar{\mathbf{u}}_i$, are orthogonal to each other. According to Theorem 2, we approximate the unconditional PEP by

$$\Pr(\mathbf{x},\mathbf{x}') \lesssim \left(\frac{d_{\mathrm{E}}^2(\mathbf{e})}{P}\right)^{-r}\left(\frac{E_s}{4N_0}\right)^{-r}. \tag{31}$$

Based on (31), we note that the unconditional PEP for OTFS modulation only depends on $d_{\mathrm{E}}^2(\mathbf{e})$, the rank of $\mathbf{\Omega}(\mathbf{e})$, and number of independent resolvable paths P, and is independent from the specific distribution of delay and Doppler indices. Our derived result is a generalized framework of the unconditional PEP for small P that is suitable for various OTFS transmission cases, including the case where $r < P$ and the delay and Doppler indices are not uniformly distributed. In particular, the term P in the denominator can be interpreted as the energy averaging with respect to the number of independent paths, while the term $\left(d_{\mathrm{E}}^2(\mathbf{e})\right)^{-r}$ indicates the potential improvement of the error performance introduced by channel coding. Furthermore, according to (31), we refer to the power of the signal-to-noise ratio (SNR) as the **diversity gain**, which dominates the exponential behaviour of the error performance for OTFS systems against the average SNR. On the other hand, the term $d_{\mathrm{E}}^2(\mathbf{e})/P$ is referred to as the **coding gain**, which characterizes the approximate improvement of coded OTFS systems over the uncoded counterpart with the same diversity gain, i.e., the same exponent $-r$ [18]. Considering the diversity property of OTFS modulation discussed in the previous remarks, it is interesting to see from (31) that there exists a fundamental trade-off between the diversity gain and the coding gain which is formally stated in the following.

Corollary 1 (Trade-Off Between Diversity and Coding Gain): For a given channel code, the diversity gain of OTFS systems improves with the number of independent resolvable paths P, while the coding gain declines.

Based on Corollary 1, we note that when P (the rank of $\mathbf{\Omega}(\mathbf{e})$) is small, the diversity gain is small. In this case, the squared Euclidean distance between codewords is crucial for OTFS systems as an optimized code can greatly improve the error performance. On the other hand, when P (the rank of $\mathbf{\Omega}(\mathbf{e})$) is large, there is a large diversity gain. In this case, it is expected that the code design can only offer a limited error performance improvement. However, it should also be noted that the coding gain always improves with the increase of $d_{\mathrm{E}}^2(\mathbf{e})$, regardless of the value of the diversity gain according to (31). Therefore, a preliminary guideline for the code design for the OTFS systems is to maximize the minimum value of $d_{\mathrm{E}}^2(\mathbf{e})$ among all pairs of codewords of the code.

To verify the accuracy of the derived unconditional PEP bound, we numerically compare the coding gain and the derived bound corresponding to (31) and (23). In particular, recalling (23), after some manipulations, we obtain

$$\Pr(\mathbf{x},\mathbf{x}'|\boldsymbol{\omega}_\tau,\boldsymbol{\omega}_\nu) \leq \left(\frac{\left(\prod_{i=1}^{r}\lambda_i\right)^{\frac{1}{r}}}{P}\right)^{-r}\left(\frac{E_s}{4N_0}\right)^{-r}. \tag{32}$$

Hence, we refer to the term $\left(\prod_{i=1}^{r}\lambda_i\right)^{\frac{1}{r}}\Big/P$ as the *conditional coding gain* of the OTFS systems for given channel parameters $\boldsymbol{\omega}_\tau$ and $\boldsymbol{\omega}_\nu$, as the eigenvalues are related to $\boldsymbol{\omega}_\tau$ and $\boldsymbol{\omega}_\nu$. Based on the conditional coding gain, we can also obtain the *average coding gain* with respect to various channel parameters $\boldsymbol{\omega}_\tau, \boldsymbol{\omega}_\nu$ and error sequences $\mathbf{e}$ by means of Monte Carlo simulation. On the other hand, from the unconditional PEP upper bound (31), we call the function $f\left(d_{\mathrm{E}}^2(\mathbf{e})\right) = d_{\mathrm{E}}^2(\mathbf{e})/P$ the *coding gain bound* of the OTFS systems. Now let us compare the average coding gain and the coding gain obtained from the performance bound via simulations. We numerically average the conditional coding gains subjected to all error sequences with $d_{\mathrm{E}}^2(\mathbf{e})$ and channel parameters $\boldsymbol{\omega}_\tau, \boldsymbol{\omega}_\nu$ to obtain the average coding gain. Without loss of generality, we consider binary phase shift keying (BPSK) signals and error sequences with $d_{\mathrm{E}}^2(\mathbf{e})$ for OTFS systems with integer delay and Doppler shifts.

Fig. 3 shows the comparison between the average coding gain and the corresponding coding gain bound in decibels for $P = 2$ with various maximum delay and Doppler indices, where the OTFS system with $N = 5$ and $M = 2$ is considered. As shown in the figure, different values of maximum delay and

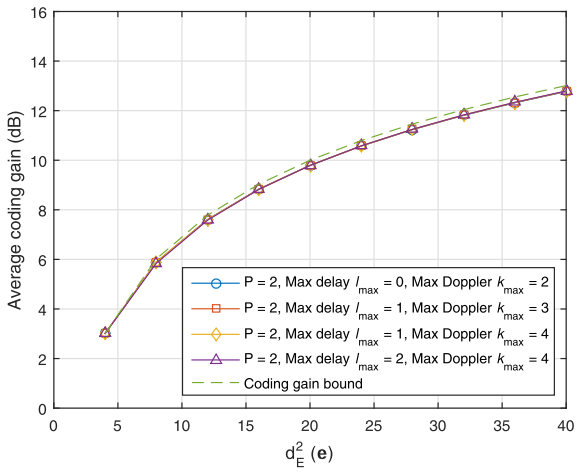


Fig. 3. Average coding gain for error sequences with $P = 2$ and $d_{\mathrm{E}}^2(\mathbf{e})$ in terms of various delay and Doppler indices, comparing with the coding gain bound.

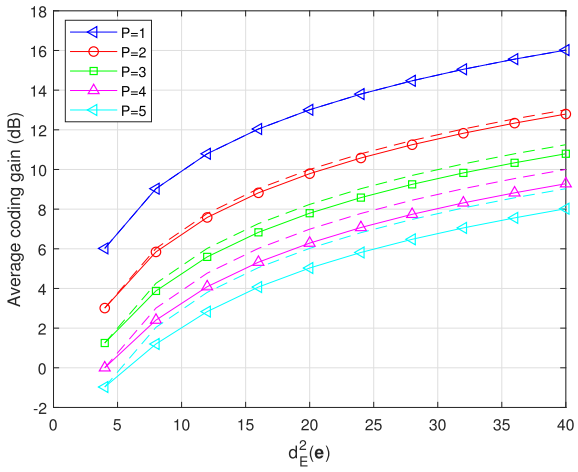


Fig. 4. Average coding gain for error sequences with $d_{\mathrm{E}}^2(\mathbf{e})$ in terms of different numbers of independent resolvable paths, comparing with the coding gain bounds (in dashed lines), where the maximum delay and Doppler indices are set to be $l_{\max} = 2$ and $k_{\max} = 4$, respectively.

Doppler indices do not have a strong impact on the average coding gain. Meanwhile, it can be observed in the figure that the average coding gain improves with the increase of the squared Euclidean distance $d_{\mathrm{E}}^2(\mathbf{e})$. Furthermore, the derived coding gain bound shows a close match with the overall average coding gain, especially when $d_{\mathrm{E}}^2(\mathbf{e})$ is small.

Fig. 4 illustrates the average coding gains and the corresponding coding gain bounds (in dashed lines) in decibels with respect to different values of P, where the OTFS system with $N = 5$ and $M = 2$ is considered. As shown in the figure, given $d_{\mathrm{E}}^2(\mathbf{e})$, the average coding gain decreases with the increase of the number of paths P, which is consistent with Corollary 1. Similar to the previous figure, the coding gain bounds match well with the general behaviour of average coding gains, especially when P is small, which verifies the correctness of our derivation. On the other hand, we notice that the derived coding gain bound slightly diverge from the average coding gains, when P is large. To verify this observation, we in the following consider the case with a large number of paths and a larger OTFS frame.

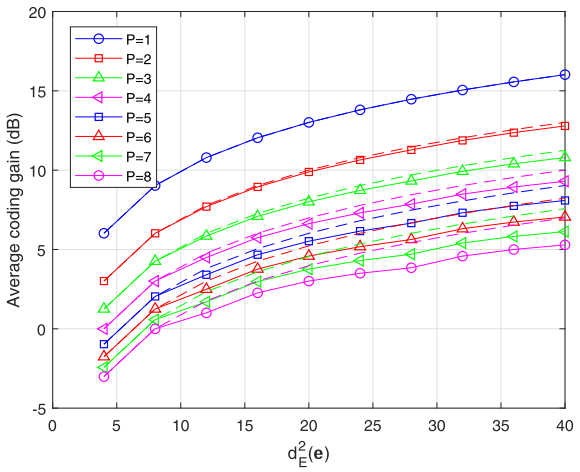


Fig. 5. Average coding gain for error sequences with $d_{\mathrm{E}}^2(\mathbf{e})$ in terms of different numbers of independent resolvable paths, comparing with the coding gain bounds (in dashed lines), where the maximum delay and Doppler indices are set to be $l_{\max} = 2$ and $k_{\max} = 2$, respectively.

We consider an OTFS system with $M = 32$ and $N = 16$ in Fig. 5, where the maximum delay and Doppler indices are set to be $l_{\max} = 2$ and $k_{\max} = 2$, respectively. Since numerically emulating all the error sequences with such a frame size is generally intractable in a reasonable time frame even with BPSK mapping, we consider the comparison between the average coding gains and the corresponding coding gain bounds with given error sequences.[4] Similar to the previous figure, we observe that the derived coding gain bounds diverge from the average coding gains for a large number of paths and the gap between them becomes wider with the increase of P. This observation motivates us to derive a more suitable approximation for the coding gain for the case of a large P, the details of which will be given in the following subsection.

B. Error Performance Analysis for Large Values of P

When there is a large number of independent resolvable paths of the channel, i.e., the value of P is large, the unconditional PEP can be more accurately bounded by considering the strong law of large number [19], [32]. In specific, the term $\sum_{i=1}^{r} \lambda_i |\tilde{h}_i|^2$ in (20) approaches a Gaussian random variable, due to the central limit theorem [33] and the fact that OTFS modulation can almost surely achieve the full diversity [6].

Notice that $\tilde{h}_i$ follows the complex Gaussian distribution, with mean $\mu_{\tilde{h}_i} = \mathbb{E}[\mathbf{h}] \cdot \mathbf{v}_i$ and variance $1/(2P)$ per real dimension. Therefore, for the ease of derivation, we normalize the variance and rewrite (20) as

$$\Pr(\mathbf{x}, \mathbf{x}' | \mathbf{h}, \boldsymbol{\omega}_\tau, \boldsymbol{\omega}_\nu) \leq \exp\left(-\frac{E_s}{4N_0} \sum_{i=1}^{r} \frac{\lambda_i}{P} |\bar{h}_i|^2\right), \quad (33)$$

where $\bar{h}_i = \sqrt{P}\tilde{h}_i$. Note that $\{\bar{h}_1, \bar{h}_2, \ldots, \bar{h}_r\}$ are independent complex Gaussian random variables with mean $\mu_{\bar{h}_i} = \sqrt{P}\mu_{\tilde{h}_i}$ and variance $1/2$ per real dimension. Furthermore, it is obvious that $|\bar{h}_i|$ follows the Rician distribution with a

[4]Without loss of generality, the error sequences are of the form $\mathbf{e} = [2, 0, -2, 2, 0, -2, \ldots, 0 \ldots 0]^{\mathrm{T}}$.

Rician factor $\bar{K}_i = \left|\mu_{\bar{h}_i}\right|^2$ and a unit variance. Thus, it can be shown that $\left|\bar{h}_i\right|^2$ follows a noncentral chi-squared distribution with one degree of freedom (DoF) and noncentrality parameter $S = \bar{K}_i$, whose mean and variance are given by [34]

$$\mu_{\left|\bar{h}_i\right|^2} = 1 + \bar{K}_i, \tag{34}$$

$$\sigma^2_{\left|\bar{h}_i\right|^2} = 2 + 4\bar{K}_i. \tag{35}$$

Next, we derive the unconditional PEP by means of Gaussian approximation. To start with, let $\psi = \sum_{i=1}^{r} \lambda_i \left|\bar{h}_i\right|^2$. According to (34) and (35), we approximate ψ as a Gaussian random variable, whose mean is $\mu_\psi = \sum_{i=1}^{r} \lambda_i \left(1 + \bar{K}_i\right)$ and variance is $\sigma^2_\psi = \sum_{i=1}^{r} \lambda_i^2 \left(2 + 4\bar{K}_i\right)$. Thus, according to the Gaussian distribution of ψ, the conditional PEP in (20) is upper-bounded by

$$\Pr\left(\mathbf{x}, \mathbf{x}' | \boldsymbol{\omega}_\tau, \boldsymbol{\omega}_\nu\right) \leq \int_0^{+\infty} \exp\left(-\frac{E_s}{4N_0 P}\psi\right) p\left(\psi\right) d\psi. \tag{36}$$

Considering

$$\int_0^{+\infty} \exp\left(-\gamma\psi\right) p\left(\psi\right) d\psi = \exp\left(\frac{1}{2}\gamma^2\sigma^2_\psi - \gamma\mu_\psi\right) Q\left(\frac{\gamma\sigma^2_\psi - \mu_\psi}{\sigma_\psi}\right), \quad \gamma > 0, \tag{37}$$

we obtain

$$\Pr\left(\mathbf{x}, \mathbf{x}' | \boldsymbol{\omega}_\tau, \boldsymbol{\omega}_\nu\right) \leq \exp\left(\frac{1}{2}\left(\frac{E_s}{4N_0}\right)^2 \cdot \frac{\sigma^2_\psi}{P^2} - \frac{E_s}{4N_0} \cdot \frac{\mu_\psi}{P}\right) \times Q\left(\frac{E_s}{4N_0} \cdot \frac{\sigma_\psi}{P} - \frac{\mu_\psi}{\sigma_\psi}\right). \tag{38}$$

Similar to the previous subsection, we consider the special case of Rayleigh fading.

In the case of **Rayleigh fading**, i.e., $|\bar{h}_i|$ and $|h_i|$ follow the Rayleigh distribution, we have $\mu_\psi = \sum_{i=1}^{r} \lambda_i$ and $\sigma^2_\psi = 2\sum_{i=1}^{r} \lambda_i^2$. Therefore, the right hand side of (38) is given by

$$\Pr\left(\mathbf{x}, \mathbf{x}' | \boldsymbol{\omega}_\tau, \boldsymbol{\omega}_\nu\right) \leq \exp\left(\left(\frac{E_s}{4N_0}\right)^2 \sum_{i=1}^{r} \frac{\lambda_i^2}{P^2} - \frac{E_s}{4N_0} \sum_{i=1}^{r} \frac{\lambda_i}{P}\right) \times Q\left(\frac{E_s}{4N_0 P}\sqrt{2\sum_{i=1}^{r} \lambda_i^2} - \frac{\sum_{i=1}^{r} \lambda_i}{\sqrt{2\sum_{i=1}^{r} \lambda_i^2}}\right). \tag{39}$$

Furthermore, we consider the Chernoff bound of the Q-function [32]

$$Q\left(\gamma\right) \leq \exp\left(-\frac{1}{2}\gamma^2\right), \quad \gamma > 0, \tag{40}$$

and (39) can be further upper-bounded by

$$\begin{aligned}
&\Pr\left(\mathbf{x}, \mathbf{x}' | \boldsymbol{\omega}_\tau, \boldsymbol{\omega}_\nu\right) \\
&\leq \exp\left(\left(\frac{E_s}{4N_0}\right)^2 \sum_{i=1}^{r} \frac{\lambda_i^2}{P^2} - \frac{E_s}{4N_0} \sum_{i=1}^{r} \frac{\lambda_i}{P}\right) \\
&\quad \times \exp\left(-\left(\frac{E_s}{4N_0}\right)^2 \sum_{i=1}^{r} \frac{\lambda_i^2}{P^2} - \frac{\left(\sum_{i=1}^{r} \lambda_i\right)^2}{4\sum_{i=1}^{r} \lambda_i^2} + \frac{E_s}{4N_0} \sum_{i=1}^{r} \frac{\lambda_i}{P}\right) \\
&= \exp\left(-\frac{\left(\sum_{i=1}^{r} \lambda_i\right)^2}{4\sum_{i=1}^{r} \lambda_i^2}\right),
\end{aligned} \tag{41}$$

when

$$\frac{E_s}{4N_0} \geq \frac{P\sum_{i=1}^{r} \lambda_i}{2\sum_{i=1}^{r} \lambda_i^2}. \tag{42}$$

Based on (41), the unconditional PEP can be approximately upper-bounded as shown in the following Theorem.

Theorem 3 (Unconditional PEP Upper Bound for Large P): For a large value of P and a reasonably high SNR, i.e., $\frac{E_s}{4N_0} \geq \frac{r}{2d^2_{\mathrm{E}}(\mathbf{e})}$, the unconditional PEP of OTFS systems can be approximately upper-bounded by

$$\Pr\left(\mathbf{x}, \mathbf{x}'\right) \lesssim \exp\left(-\frac{E_s}{16N_0} d^2_{\mathrm{E}}\left(\mathbf{e}\right)\right). \tag{43}$$

Proof: The proof is given in Appendix D.

It should be noted that, for $r \geq 4$, the approximation in Theorem 3 is sufficiently accurate owing to the strong law of the large number [19], [32]. On the other hand, the value of r is usually smaller than the squared Euclidean distance $d^2_{\mathrm{E}}\left(\mathbf{e}\right)$ for practical wireless transmissions with good channel codes.[5] Therefore, our SNR assumption is reasonable. Compared with Corollary 1, it is not surprising that the unconditional PEP only depends on the squared Euclidean distance $d^2_{\mathrm{E}}\left(\mathbf{e}\right)$, regardless of the delay and Doppler indices. Furthermore, we note that (43) is of the similar form of the error performance for AWGN channels [36]. This is because the impact of fading is mitigated by a large number of diversity branches and consequently, in other words, the channel with a large number of diversity paths approaches an AWGN model [36].

Notice that the upper bound in Theorem 3 is based on the PEP analysis, which does not directly indicate the average error performance of the coded system [37]. But it can be used to approximate the coding gain for large P. Based on the results from Theorem 3, we apply the commonly adopted coding gain approximation for AWGN channels [35] to evaluate the coding gain of OTFS systems with a large P. In specific, we have

$$\text{Coding gain} \simeq 10\log_{10}\left(\frac{d^2_{\mathrm{c,min}}}{d^2_{\mathrm{u,min}}}\right) \text{dB}, \tag{44}$$

[5]For example, a popular industry-standard rate-$1/2$ convolutional code with code memory of 6 has a minimum squared Euclidean distance $d^2_{\mathrm{E}}\left(\mathbf{e}\right) = 40$ [35].

TABLE I
CODE PARAMETERS

Code structure	Generator matrix	Memory length	Minimum $d_{\mathrm{E}}^2(\mathbf{e})$
A	$[1+D, D]$	1	12
B	$[1+D^2, 1+D+D^2]$	2	20
C	$[1+D^2+D^5, 1+D+D^2+D^3+D^4+D^5]$	5	32
D	$[1+D+D^2+D^5+D^6, 1+D^2+D^3+D^4+D^6]$	6	40

where $d_{\mathrm{c,min}}^2$ and $d_{\mathrm{u,min}}^2$ are the minimum squared Euclidean distances for coded and uncoded OTFS systems, respectively.

In the previous subsections, we have derived the error performance analysis of coded OTFS systems, which is valid for general underspread WSSUS channels including both the integer and fractional delay and Doppler cases. In the following, we will discuss the design of channel codes according to our error performance analysis.

C. Code Design Issues

According to the derived analysis, the rule-of-thumb channel code design criterion is discussed in this section. Without loss of generality, we only consider the Rayleigh fading channel in the following. We can see from the previous analysis that the code design criterion for the coded OTFS system is to maximize the minimum squared Euclidean distance $d_{\mathrm{E}}^2(\mathbf{e})$.

Proposition 1 (The Squared Euclidean Distance Criterion): The channel code should be designed to maximize the minimum squared Euclidean distance among all pairs of possible codewords.

We note that even with the designed code, the error performance of coded OTFS systems may still vary with different channel parameters e.g., $\boldsymbol{\omega}_\tau$ and $\boldsymbol{\omega}_\nu$. This detrimental effect due to channel realizations is widely observed in the system designs for fading channels, such as in [20], [38]. In specific, with different channel parameters, the value of the conditional coding gain can be different even with the same error sequence, which may potentially jeopardize the overall error performance of OTFS systems. In order to obtain a more robust performance, it is desirable to apply an interleaver to permute the coded symbols before sending to the constellation mapper or the OTFS modulator in the DD domain. As pointed out in [38], such an interleaver can "whiten" the transmitted symbols from the information theoretic point of view and the detrimental effect on the error performance due to the channel parameters can thus be alleviated.

To examine our performance analysis of the coded OTFS systems, we perform numerical simulations for OTFS systems over high-mobility channels, the results of which will be shown in the next section.

IV. NUMERICAL RESULTS

In this section, the error performance of the coded OTFS system with various channel codes is evaluated via numerical simulations. We consider the BPSK signal for the OTFS system, where the data sequence is firstly encoded and interleaved, and then BPSK mapped, according to Fig. 1. Without loss of generality, we consider the sum-product algorithm [39], [40] for OTFS detection, where the details can be found in [41]. In order to verify the accuracy of the analytical results, we consider four different convolutional codes (with trellis termination) with different minimum squared Euclidean distance $d_{\mathrm{E}}^2(\mathbf{e})$ among all possible codeword pairs. The details of the code parameters are given in Table I, including the generator matrix and the memory length. In particular, we also show the smallest value of squared Euclidean distance $d_{\mathrm{E}}^2(\mathbf{e})$ among all possible codeword pairs for each code. In specific, we consider a coded OTFS system with $N = 8$ and $M = 16$ and correspondingly the codeword length for all considered simulations is 128 bits unless otherwise specified. The channel decoder adopts the logarithm domain Bahl-Cocke-Jelinek-Raviv (BCJR) algorithm [42]. Furthermore, we consider the Rayleigh fading case. If not otherwise specified, we only consider the integer delay and Doppler case and set the maximum delay index as $l_{\max} = 3$ and the maximum Doppler index as $k_{\max} = 5$, which is corresponding to a relative speed around 250 km/h with 4 GHz carrier frequency and 1.5 kHz sub-carrier spacing. For each channel realization, we randomly select the delay and Doppler indices according to the uniform distribution, such that we have $-k_{\max} \leq k_i \leq k_{\max}$ and $0 \leq l_i \leq l_{\max}$.

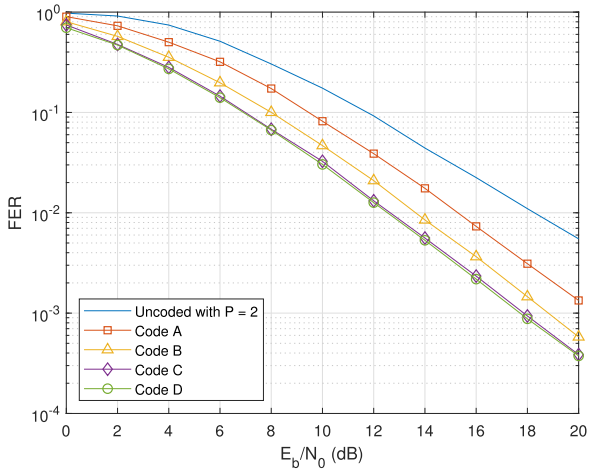


Fig. 6. FER performance for OTFS modulation with different codes for $P = 2$, where the relative UE speed is 250 km/h.

The frame-error-rate (FER) performance of the OTFS systems with $P = 2$ is shown in Fig. 6. It can be observed that the slope of the FER curve for the uncoded OTFS system is slightly worse than that for coded OTFS systems. This indicates that the uncoded OTFS system with $P = 2$ does not guarantee the full diversity for all possible channel realizations, which is consistent with the analysis in [14]. More importantly, this observation also shows that the application of channel coding can improve the diversity gain of OTFS systems in

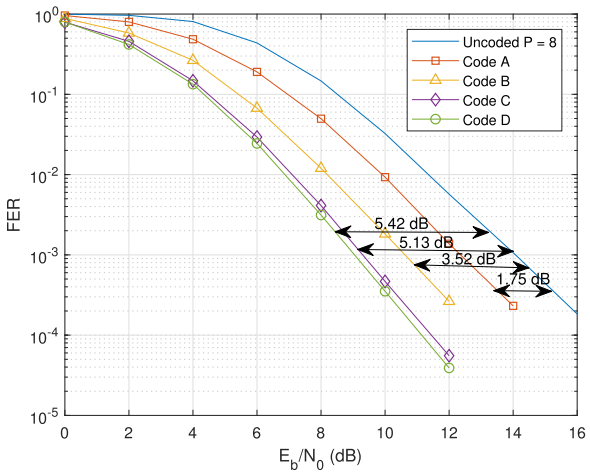


Fig. 7. FER performance for OTFS modulation with different codes for $P = 8$, where the relative UE speed is 250 km/h.

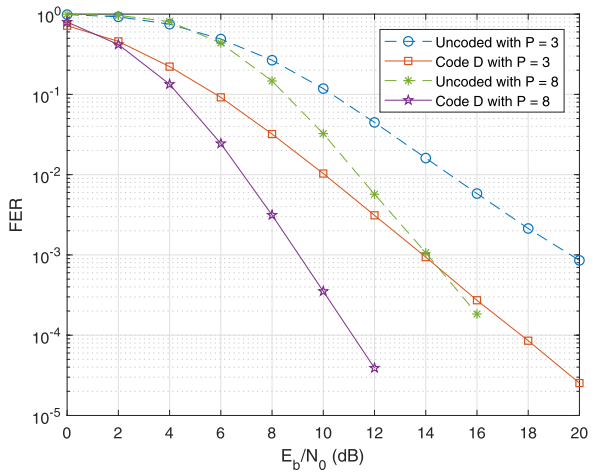


Fig. 8. FER performance of code D for OTFS modulation with $P = 3$ and $P = 8$, comparing with that of uncoded OTFS systems, where the relative UE speed is 250 km/h.

the case where the OTFS modulation fails to achieve the full diversity. Moreover, we observe that employing the channel code with a larger minimum squared Euclidean distance $d_{\mathrm{E}}^2(\mathbf{e})$ indeed leads to a larger coding gain. In specific, we observe from the figure that for FER $\approx 10^{-2}$, the required SNRs for codes A, B, C and D, are 15.28 dB, 13.64 dB, 12.65 dB, and 12.54 dB, respectively. Compared to uncoded OTFS systems, these four coded OTFS systems achieve coding gains roughly 2.99 dB, 4.63 dB, 5.62 dB, and 5.73 dB, respectively. This observation clearly substantiates the proposed performance analysis and code design criterion in Proposition 1.

Fig. 7 shows the FER performance of the OTFS systems with $P = 8$. We observe that the channel code with a larger minimum squared Euclidean distance $d_{\mathrm{E}}^2(\mathbf{e})$ enjoys a larger coding gain compared with the uncoded OTFS system, which is consistent with the analysis of Theorem 3. Furthermore, we can calculate the approximated coding gains according to (44) for the adopted channel codes, which are 1.83 dB, 4.12 dB, 6.37 dB, and 7.42 dB for code A to D, respectively, while we also obtain the actual coding gains based on the simulation results that are 1.75 dB, 3.52 dB, 5.13 dB, and 5.42 dB for code A to D, respectively. These observations verify the accuracy of our analysis for large P in Theorem 3 and the coding gain approximation. Together with the observations from Fig. 6, we can also conclude that our proposed code design criterion in Proposition 1 is universal for general OTFS systems regardless of the channel parameters and the number of paths P.

Fig. 8 presents the trade-off between the diversity and coding gain. In particular, we consider the FER performance of code D with $P = 3$ and $P = 8$, comparing with that of the corresponding uncoded OTFS systems. As shown in figure, at FER $\approx 10^{-3}$, the coded OTFS system with $P = 3$ exhibits around 5.7 dB coding gain compared to that of the uncoded OTFS system with the same FER, while only around 5.0 dB coding gain is obtained for the coded OTFS system with $P = 8$. This observation matches the prediction in Corollary 1, which indicates that the coding gain reduces with the increase of P, given the same channel code.

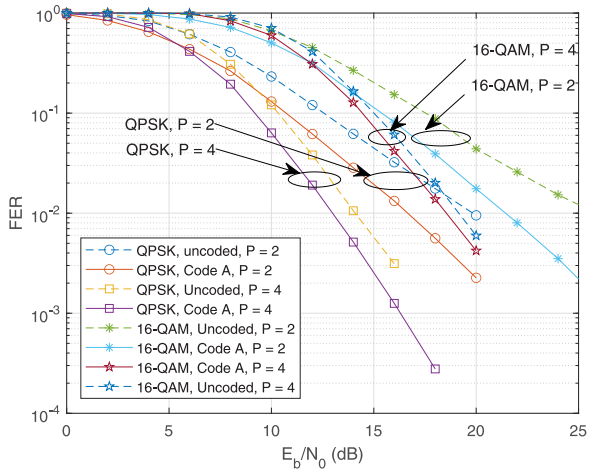


Fig. 9. FER performance of code A for QPSK/16-QAM mapped OTFS systems with different number of paths and the relative UE speed of 250 km/h.

Fig. 9 presents the trade-off between the diversity and coding gain with practical mappings, namely, quadrature phase shift keying (QPSK) and 16-quadrature amplitude modulation (16-QAM) mappings. In particular, we consider the FER performance of code A with $P = 2$ and $P = 4$, comparing with that of the corresponding uncoded OTFS systems. Similar to the previous figure, the trade-off between the diversity and coding gain is clearly demonstrated by the simulation results.

The FER performance of the OTFS modulation with code D and different number of paths P is illustrated in Fig. 9. It can be observed from the figure that given the same code, the error performance of the coded OTFS systems improves with the increase of number of distinguishable paths. Furthermore, it is obvious that with the same code, a larger value of P corresponds to a larger diversity advantage as indicated in the figure, which is consistent with our analysis.

We compare the FER performance of code A with $P = 3$ and $P = 8$, and that of the corresponding OFDM systems in Fig. 10. As observed from the figure, the OTFS system enjoy better error performance than that of the OFDM system with the same code, for both $P = 3$ and $P = 8$.

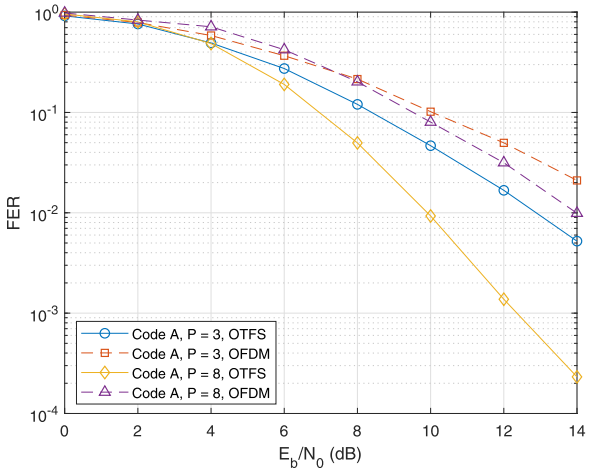


Fig. 10. FER performance of code A for OTFS modulation with $P = 3$ and $P = 8$, compared to an OFDM system, where the relative UE speed is 250 km/h.

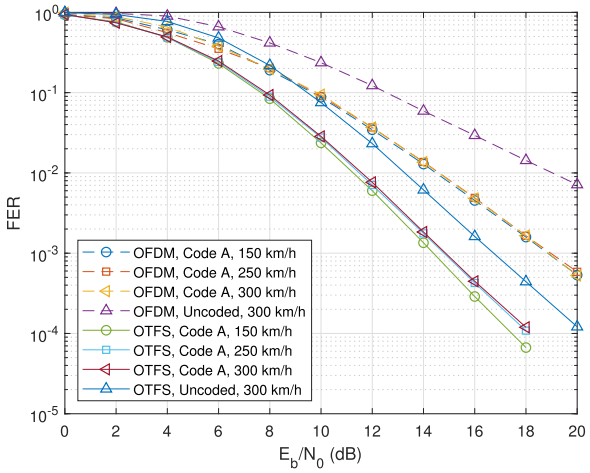


Fig. 11. FER performance of code A for OTFS modulation and OFDM modulation with $P = 4$, where the values of the relative UE speed are 150 km/h, 250 km/h, and 300 km/h, respectively.

Furthermore, the FER curve of the OFDM system shares almost the same slope as that of the OTFS system, for $P = 3$. As for $P = 8$, the achieved diversity gain of the OTFS system is clearly higher than that of the OFDM system. Note that the diversity gain of coded OFDM systems is determined by the smaller value of the minimum symbol-wise Hamming distance of the code δ_{H} and the number of paths P[6] [43], while OTFS systems can obtain the full diversity almost surely regardless of the employment of channel codes. Therefore, this observation clearly shows the advantage of the OTFS systems over the OFDM systems.

The FER performance of both OTFS systems and OFDM systems with various values of the relative UE speed is compared in Fig. 11. We consider $l_{\max} = 3$ and $k_{\max} = 3, 5, 6$, which corresponds to the cases where the relative UE speeds are 150 km/h, 250 km/h, and 300 km/h, respectively. We apply the near-optimal ML detection [41] for both the OTFS and OFDM systems to have a fair comparison. We observe that

[6]According to [43], we have $r_{\mathrm{OFDM}} = \min(\delta_{\mathrm{H}}, P)$, where r_{OFDM} is the achievable diversity gain of a coded OFDM system. In specific, we have $r_{\mathrm{OFDM}} = 3$ for both $P = 3$ and $P = 8$ with Code A.

TABLE II
CODE PARAMETERS FOR FIG. 12

Code	Data length	Codeword length	Code rate
Convolutional Code D	250	512	0.488
5G LDPC	256	512	0.5
LTE Turbo	250	512	0.488

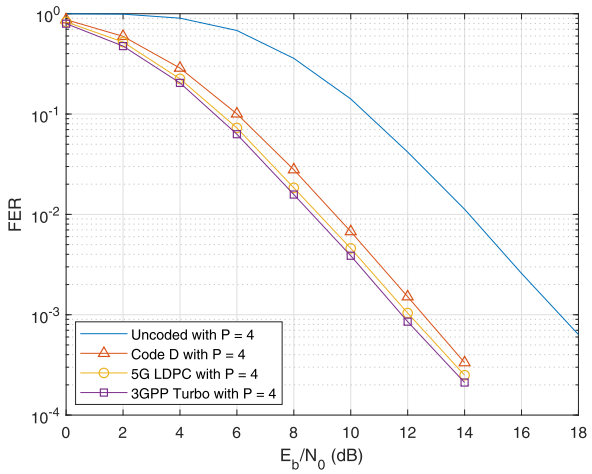


Fig. 12. FER performance of 5G LDPC code and 3GPP Turbo code for OTFS modulation with $P = 4$ and a relative UE speed of 250 km/h, where the codeword length is 512 bits.

the error performance of both OTFS and OFDM systems with ML detection does not change much with different relative UE speeds, which is consistent with the observations in [10]. Note that, the ML detection is not practically feasible for conventional OFDM systems due to the high detection complexity. Therefore, frequency domain equalization is usually deployed for OFDM systems [44]. In this case, the error performance of OFDM systems will degrade dramatically with the increase of the speed [44] due to the severe inter-carrier interference (ICI) induced by the Doppler spread. Furthermore, the FER performance of the OTFS systems outperform that of the OFDM systems, including both coded and uncoded cases with various values of the speed. Similar to the previous figure, we also observe that the achieved diversity gain of the OTFS system is higher than that of the OFDM system, which agrees with our analysis.

We present the FER results of coded OTFS systems with modern codes in Fig. 12. We consider an LDPC code from the 5G communication standard [45] (referred to as the 5G LDPC code) and the Turbo code from the 3GPP long term evolution (LTE) standard [46] (referred to as the LTE Turbo code), where code parameters are given in Table II and we have $N = 16$ and $M = 32$ for OTFS modulation. As a benchmark, the FER performance of the convolutional code D is also given in Fig. 12. We observe that, the LTE Turbo code achieves the best error performance compared to the 5G LDPC code and the convolutional code D with $P = 4$, although they all share the same diversity gain. More specifically, at FER $\approx 10^{-3}$, 5G LDPC code and LTE Turbo code show around 0.5 dB and 0.7 dB SNR gain compared to the convolutional code D. Note that a similar observation of the error performance can be observed over the AWGN channel, where the LTE Turbo

code has the best performance while the convolutional code D has the worst performance. Therefore, this observation indicate that codes optimized for AWGN channels can also achieve a good error performance in the OTFS systems, which is consistent with our analysis.

Remark 2: Based on the above simulation results, we summarize that the slopes of FER curves for both uncoded and coded OTFS systems are generally different. This is because the diversity of uncoded OTFS systems is determined only by channel parameters while the diversity of coded OTFS systems is determined by both the channel and code parameters. In particular, uncoded OTFS systems cannot guarantee the exploitation of the full diversity in all possible channel coefficients and transmitted symbols realizations, while coded OTFS systems can achieve the full diversity with a careful code design. Note that how to design practical codes to guarantee full diversity is an interesting research work and one initial idea is to ensure the *codeword difference matrix* in (18) to be a diagonal matrix for any pair of codewords. However, we will explore the details on the code design in our future work due to the space limitation.

V. Conclusion

In this paper, we studied the performance analysis of coded OTFS systems over high-mobility channels. We first derived the conditional PEP for a given channel realization and then obtained the unconditional PEP by leveraging proper bounding techniques. We discussed two cases of the OTFS transmission according to the number of independent resolvable paths of the channel, where we showed that the coding improvement of OTFS systems depends on the squared Euclidean distance between a pair of codewords. More importantly, we demonstrated the fundamental trade-off between the diversity gain and the coding gain for OTFS systems. Furthermore, we proposed a code design criterion based on the derived unconditional bound. The analysis and the code design criterion are verified by numerical simulations with various channel codes. Our future work will focus on the modern code design for OTFS systems by considering analytical tools such as the density evolution and the extrinsic information transfer (EXIT) chart.

Appendix A
Proof of Lemma 3

Notice that, $\mathbf{\Omega}(\mathbf{e})$ is positive semidefinite Hermitian. Hence, the eigenvalues $\{\lambda_i\}$ of $\mathbf{\Omega}(\mathbf{e})$ for $1 \leq i \leq r$ are all positive. Considering the arithmetic mean and geometric mean (AM-GM) inequality, we obtain

$$\sum_{i=1}^{r} \frac{1}{\lambda_i} \geq r \left(\prod_{i=1}^{r} \frac{1}{\lambda_i} \right)^{\frac{1}{r}} = \frac{r}{\left(\prod_{i=1}^{r} \lambda_i \right)^{\frac{1}{r}}}. \tag{45}$$

Then, we apply the Cauchy-Schwarz inequality to the denominator of the right hand side of (45), which yields

$$\frac{r}{\left(\prod_{i=1}^{r} \lambda_i \right)^{\frac{1}{r}}} \geq \frac{r^2}{\sum_{i=1}^{r} \lambda_i} = \frac{r^2}{P d_{\mathrm{E}}^2(\mathbf{e})}. \tag{46}$$

It is obvious that the equality only holds when the eigenvalues $\{\lambda_i\}$ are of the same value, e.g., the codeword difference matrix $\mathbf{\Omega}(\mathbf{e})$ is a diagonal matrix, i.e., $\mathbf{\Omega}(\mathbf{e}) = \mathrm{diag}\left\{d_{\mathrm{E}}^2(\mathbf{e}), \ldots, d_{\mathrm{E}}^2(\mathbf{e})\right\}$. This completes the proof of Lemma 3.

Appendix B
Proof of Theorem 1

The product of the first r eigenvalues of $\mathbf{\Omega}(\mathbf{e})$ can be written as

$$\prod_{i=1}^{r} \lambda_i = \exp\left(\ln\left(\prod_{i=1}^{r} \lambda_i \right) \right) = \exp\left(\sum_{i=1}^{r} \ln(\lambda_i) \right). \tag{47}$$

Let us consider the inequality $\ln(\gamma) \geq 1 - \frac{1}{\gamma}, \gamma \in (0, +\infty)$, where the equality only holds when $\gamma = 1$. Therefore, we have

$$\begin{aligned} \prod_{i=1}^{r} \lambda_i &= \exp\left(\sum_{i=1}^{r} \ln(\lambda_i) \right) \\ &= \exp\left(r \ln\left(d_{\mathrm{E}}^2(\mathbf{e}) \right) + \sum_{i=1}^{r} \ln\left(\frac{\lambda_i}{d_{\mathrm{E}}^2(\mathbf{e})} \right) \right) \\ &\geq \exp\left(r \ln\left(d_{\mathrm{E}}^2(\mathbf{e}) \right) + \sum_{i=1}^{r} \left(1 - \frac{d_{\mathrm{E}}^2(\mathbf{e})}{\lambda_i} \right) \right) \qquad (48) \\ &= \left(d_{\mathrm{E}}^2(\mathbf{e}) \right)^r \exp\left(r - d_{\mathrm{E}}^2(\mathbf{e}) \sum_{i=1}^{r} \frac{1}{\lambda_i} \right). \qquad (49) \end{aligned}$$

It is obvious that the equality holds in (48) only if the first r eigenvalues $\{\lambda_i\}$ of $\mathbf{\Omega}(\mathbf{e})$ equal to $d_{\mathrm{E}}^2(\mathbf{e})$. Notice that the eigenvalues $\{\lambda_i\}$ of $\mathbf{\Omega}(\mathbf{e})$ equal to the main diagonal elements when $\mathbf{\Omega}(\mathbf{e})$ is a diagonal matrix, i.e., $\mathbf{\Omega}(\mathbf{e}) = \mathrm{diag}\left\{d_{\mathrm{E}}^2(\mathbf{e}), \ldots, d_{\mathrm{E}}^2(\mathbf{e})\right\}$, in which case we have $\det(\mathbf{\Omega}(\mathbf{e})) = \left(d_{\mathrm{E}}^2(\mathbf{e})\right)^P$. This completes the proof of Theorem 1.

Appendix C
Proof of Theorem 2

Following Theorem 1, we note that value of the term $r - d_{\mathrm{E}}^2(\mathbf{e}) \sum_{i=1}^{r} \frac{1}{\lambda_i}$ is smaller than 1 according to Lemma 3. Notice that the value of $f(x) = \exp(-x)$ decays relatively slowly with the increase of x, if $x < 1$. This fact motives us to consider the approximation based on Lemma 3 as considered in [19], [32], such as

$$\begin{aligned} \prod_{i=1}^{r} \lambda_i &\geq \left(d_{\mathrm{E}}^2(\mathbf{e}) \right)^r \exp\left(r - d_{\mathrm{E}}^2(\mathbf{e}) \sum_{i=1}^{r} \frac{1}{\lambda_i} \right) \\ &\gtrsim \left(d_{\mathrm{E}}^2(\mathbf{e}) \right)^r \exp\left(r - d_{\mathrm{E}}^2(\mathbf{e}) \left(\sum_{i=1}^{r} \frac{1}{\lambda_i} \right)_{\min} \right) \\ &= \left(d_{\mathrm{E}}^2(\mathbf{e}) \right)^r \exp\left(r - d_{\mathrm{E}}^2(\mathbf{e}) \frac{r^2}{P d_{\mathrm{E}}^2(\mathbf{e})} \right) \\ &= \left(d_{\mathrm{E}}^2(\mathbf{e}) \right)^r \exp\left(r - \frac{r^2}{P} \right), \\ &\geq \left(d_{\mathrm{E}}^2(\mathbf{e}) \right)^r, \end{aligned} \tag{50}$$

where the equality holds if $\mathbf{\Omega}(\mathbf{e})$ is a diagonal matrix, i.e., $\mathbf{\Omega}(\mathbf{e}) = \text{diag}\left\{d_{\text{E}}^2(\mathbf{e}), \ldots, d_{\text{E}}^2(\mathbf{e})\right\}$. Mathematically, the above approximation may be loose if $\mathbf{\Omega}(\mathbf{e})$ is ill-conditioned. Therefore, we justify the accuracy of our approximation as follows. A commonly adopted approach in testifying if a matrix is in ill-condition is the $\mathbf{P}$-condition number [47]. In specific, we consider the lower bound on the $\mathbf{P}$-condition number of a Gram matrix [47]. The $\mathbf{P}$-condition number of $\mathbf{\Omega}(\mathbf{e})$ is defined by

$$\mathbf{P}(\mathbf{\Omega}(\mathbf{e})) \triangleq \text{Radius}(\mathbf{\Omega}(\mathbf{e}))\,\text{Radius}\left((\mathbf{\Omega}(\mathbf{e}))^{-1}\right), \quad (51)$$

where $\text{Radius}(\mathbf{\Omega}(\mathbf{e}))$ is the *spectral radius* of $\mathbf{\Omega}(\mathbf{e})$, i.e., $\text{Radius}(\mathbf{\Omega}(\mathbf{e}))$ equals to the largest eigenvalues of $\mathbf{\Omega}(\mathbf{e})$. In particular, the matrix $\mathbf{\Omega}(\mathbf{e})$ is said to be ill-conditioned if $\mathbf{P}(\mathbf{\Omega}(\mathbf{e}))$ is large and is to be well-conditioned if $\mathbf{P}(\mathbf{\Omega}(\mathbf{e}))$ is small. According to [47], we have $\mathbf{P}(\mathbf{\Omega}(\mathbf{e})) \geq \left(\|\mathbf{u}_i\|^2\right)_{\max} \Big/ \left(\|\mathbf{u}_j\|^2\right)_{\min}$ for $1 \leq i, j \leq P$, which yields

$$\mathbf{P}(\mathbf{\Omega}(\mathbf{e})) \geq d_{\text{E}}^2(\mathbf{e}) / d_{\text{E}}^2(\mathbf{e}) = 1. \quad (52)$$

We can see that the $\mathbf{P}$-condition number of $\mathbf{\Omega}(\mathbf{e})$ always greater than or equal to 1, which indicates that $\mathbf{\Omega}(\mathbf{e})$ is generally well-conditioned. This completes the proof of Theorem 2.

Appendix D
Proof of Theorem 3

Recalling (42), we note that (41) only holds if

$$\sum_{i=1}^{r} \lambda_i^2 \geq \frac{P\sum_{i=1}^{r}\lambda_i}{E_s/(4N_0)} = \frac{P^2 d_{\text{E}}^2(\mathbf{e})}{E_s/(4N_0)}. \quad (53)$$

Therefore, we consider the approximation of (41) as follows

$$\begin{aligned}
\Pr(\mathbf{x}, \mathbf{x}') &\leq \exp\left(-\left(\sum_{i=1}^{r}\lambda_i\right)^2 \Big/ \left(4\sum_{i=1}^{r}\lambda_i^2\right)\right) \\
&\lesssim \exp\left(-\left(\sum_{i=1}^{r}\lambda_i\right)^2 \Big/ \left(4\sum_{i=1}^{r}\lambda_i^2\right)_{\min}\right) \\
&= \exp\left(-\left(P d_{\text{E}}^2(\mathbf{e})\right)^2 \Big/ \left(\frac{4P^2 d_{\text{E}}^2(\mathbf{e})}{E_s/(4N_0)}\right)\right) \\
&= \exp\left(-\frac{E_s}{16N_0} d_{\text{E}}^2(\mathbf{e})\right). \quad (54)
\end{aligned}$$

The above approximation is reasonable because the value of $\exp\left(-\left(\sum_{i=1}^{r}\lambda_i\right)^2 \Big/ \left(4\sum_{i=1}^{r}\lambda_i^2\right)\right)$ changes relatively slowly with the increase of $4\sum_{i=1}^{r}\lambda_i^2$ [19], [32]. On the other hand, the justification of the SNR assumption of (42) is necessary. According to Lemma 4, we have

$$\frac{P\sum_{i=1}^{r}\lambda_i}{2\sum_{i=1}^{r}\lambda_i^2} \leq \frac{P^2 d_{\text{E}}^2(\mathbf{e})}{\frac{2P^2}{r}\left(d_{\text{E}}^2(\mathbf{e})\right)^2} = \frac{r}{2d_{\text{E}}^2(\mathbf{e})}, \quad (55)$$

where the equality only holds when the eigenvalues of $\mathbf{\Omega}(\mathbf{e})$ share the same value, e.g., $\mathbf{\Omega}(\mathbf{e}) = \text{diag}\left\{d_{\text{E}}^2(\mathbf{e}), \ldots, d_{\text{E}}^2(\mathbf{e})\right\}$. Therefore, we can see that the term $\frac{P\sum_{i=1}^{r}\lambda_i}{2\sum_{i=1}^{r}\lambda_i^2}$ is upper-bounded by $\frac{r}{2d_{\text{E}}^2(\mathbf{e})}$. Hence, the assumption of SNR of (42) can be further restricted as $\frac{E_s}{4N_0} \geq \frac{r}{2d_{\text{E}}^2(\mathbf{e})}$. This completes the proof of Theorem 3.

References

[1] S. Li, J. Yuan, W. Yuan, Z. Wei, B. Bai, and D. W. K. Ng, "On the performance of coded OTFS modulation over high-mobility channels," in *Proc. IEEE Int. Conf. Commun. Workshops (ICC Workshops)*, 2021, pp. 1–6.

[2] G. Meyer and S. Beiker, *Road Vehicle Automation*. Cham, Switzerland: Springer, 2019.

[3] Y. Cai, Z. Wei, R. Li, D. W. K. Ng, and J. Yuan, "Joint trajectory and resource allocation design for energy-efficient secure UAV communication systems," *IEEE Trans. Commun.*, vol. 68, no. 7, pp. 4536–4553, Jul. 2020.

[4] A. Fotouhi *et al.*, "Survey on UAV cellular communications: Practical aspects, standardization advancements, regulation, and security challenges," *IEEE Commun. Surveys Tuts.*, vol. 21, no. 4, pp. 3417–3442, 4th Quart., 2019.

[5] T. Hwang, C. Yang, G. Wu, S. Li, and G. Ye Li, "OFDM and its wireless applications: A survey," *IEEE Trans. Veh. Technol.*, vol. 58, no. 4, pp. 1673–1694, May 2009.

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

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

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

[9] P. Raviteja, K. T. Phan, and Y. Hong, "Embedded pilot-aided channel estimation for OTFS in Delay–Doppler channels," *IEEE Trans. Veh. Technol.*, vol. 68, no. 5, pp. 4906–4917, May 2019.

[10] P. Raviteja, K. T. Phan, Y. Hong, and E. Viterbo, "Interference cancellation and iterative detection for orthogonal time frequency space modulation," *IEEE Trans. Wireless Commun.*, vol. 17, no. 10, pp. 6501–6515, Oct. 2018.

[11] A. Farhang, A. Rezazadehreyhani, L. E. Doyle, and B. Farhang-Boroujeny, "Low complexity modem structure for OFDM-based orthogonal time frequency space modulation," *IEEE Wireless Commun. Lett.*, vol. 7, no. 3, pp. 344–347, Jun. 2018.

[12] W. Yuan, Z. Wei, J. Yuan, and D. W. K. Ng, "A simple variational Bayes detector for orthogonal time frequency space (OTFS) modulation," *IEEE Trans. Veh. Technol.*, vol. 69, no. 7, pp. 7976–7980, Jul. 2020.

[13] S. Li, W. Yuan, Z. Wei, and J. Yuan, "Cross domain iterative detection for orthogonal time frequency space modulation," 2021, *arXiv:2101.03822*. [Online]. Available: http://arxiv.org/abs/2101.03822

[14] P. Raviteja, Y. Hong, E. Viterbo, and E. Biglieri, "Effective diversity of OTFS modulation," *IEEE Wireless Commun. Lett.*, vol. 9, no. 2, pp. 249–253, Feb. 2020.

[15] E. Biglieri, P. Raviteja, and Y. Hong, "Error performance of orthogonal time frequency space (OTFS) modulation," in *Proc. IEEE Int. Conf. Commun. Workshops (ICC Workshops)*, May 2019, pp. 1–6.

[16] G. D. Surabhi, R. M. Augustine, and A. Chockalingam, "On the diversity of uncoded OTFS modulation in doubly-dispersive channels," *IEEE Trans. Wireless Commun.*, vol. 18, no. 6, pp. 3049–3063, Jun. 2019.

[17] J. Wu and P. Fan, "A survey on high mobility wireless communications: Challenges, opportunities and solutions," *IEEE Access*, vol. 4, pp. 450–476, Jan. 2016.

[18] V. Tarokh, N. Seshadri, and A. R. Calderbank, "Space-time codes for high data rate wireless communication: Performance criterion and code construction," *IEEE Trans. Inf. Theory*, vol. 44, no. 2, pp. 744–765, Mar. 1998.

[19] B. Vucetic and J. Yuan, *Space-Time Coding*. Hoboken, NJ, USA: Wiley, 2003.

[20] B. Lu and X. Wang, "Space-time code design in OFDM systems," in *Proc. IEEE. Global Telecommun. Conference. Conf. Rec. (Globecom)*, vol. 2, Nov. 2000, pp. 1000–1004.

[21] A. F. Molisch, *Wireless Communications*, vol. 34. Hoboken, NJ, USA: Wiley, 2012.

[22] U. Schuster and H. Bolcskei, "Ultrawideband channel modeling on the basis of information-theoretic criteria," *IEEE Trans. Wireless Commun.*, vol. 6, no. 7, pp. 2464–2475, Jul. 2007.

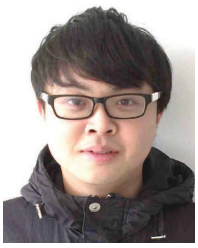
[23] P. Bello, "Characterization of randomly time-variant linear channels," *IEEE Trans. Commun.*, vol. 11, no. 4, pp. 360–393, Dec. 1963.

[24] P. Raviteja, Y. Hong, E. Viterbo, and E. Biglieri, "Practical pulse-shaping waveforms for Reduced-Cyclic-Prefix OTFS," *IEEE Trans. Veh. Technol.*, vol. 68, no. 1, pp. 957–961, Jan. 2019.

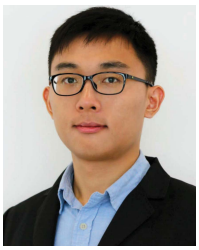
[25] Z. Wei, W. Yuan, S. Li, J. Yuan, and D. W. K. Ng, "Transmitter and receiver window designs for orthogonal time frequency space modulation," *IEEE Trans. Commun.*, early access, Jan. 13, 2021, doi: 0.1109/TCOMM.2021.3051386.
[26] X. Liu, C. Liu, Y. Li, B. Vucetic, and D. W. K. Ng, "Deep residual learning-assisted channel estimation in ambient backscatter communications," *IEEE Wireless Commun. Lett.*, vol. 10, no. 2, pp. 339–343, Feb. 2021.
[27] C. Liu, X. Liu, D. Wing Kwan Ng, and J. Yuan, "Deep residual learning for channel estimation in intelligent reflecting surface-assisted multi-user communications," 2020, *arXiv:2009.01423*. [Online]. Available: http://arxiv.org/abs/2009.01423
[28] C. Liu, Z. Wei, D. W. K. Ng, J. Yuan, and Y.-C. Liang, "Deep transfer learning for signal detection in ambient backscatter communications," *IEEE Trans. Wireless Commun.*, vol. 20, no. 3, pp. 1624–1638, Mar. 2021.
[29] C. Liu, J. Wang, X. Liu, and Y.-C. Liang, "Deep CM-CNN for spectrum sensing in cognitive radio," *IEEE J. Sel. Areas Commun.*, vol. 37, no. 10, pp. 2306–2321, Oct. 2019.
[30] Z. Wei *et al.*, "Orthogonal time-frequency space modulation: A promising next-generation waveform," 2020, *arXiv:2010.03344*. [Online]. Available: http://arxiv.org/abs/2010.03344
[31] R. A. Horn and C. R. Johnson, *Matrix Analysis*. Cambridge, U.K.: Cambridge Univ. Press, 2012.
[32] J. Yuan, Z. Chen, B. Vucetic, and W. Firmanto, "Performance and design of space-time coding in fading channels," *IEEE Trans. Commun.*, vol. COM-51, no. 12, pp. 1991–1996, Dec. 2003.
[33] A. Papoulis and S. U. Pillai, *Probability, Random Variables, and Stochastic Processes*. New York, NY, USA: McGraw-Hill, 2002.
[34] S. M. Kay, *Fundamentals of Statistical Signal Processing*. Upper Saddle River, NJ, USA: Prentice-Hall, 1993.
[35] W. Ryan and S. Lin, *Channel Codes: Classical Modern*. Cambridge, U.K.: Cambridge Univ. Press, 2009.
[36] J. Ventura-Traveset, G. Caire, E. Biglieri, and G. Taricco, "Impact of diversity reception on fading channels with coded modulation. I. Coherent detection," *IEEE Trans. Commun.*, vol. 45, no. 5, pp. 563–572, May 1997.
[37] A. Stefanov and T. M. Duman, "Performance bounds for space-time trellis codes," *IEEE Trans. Inf. Theory*, vol. 49, no. 9, pp. 2134–2140, Sep. 2003.
[38] E. Biglieri, J. Proakis, and S. Shamai (Shitz), "Fading channels: Information-theoretic and communications aspects," *IEEE Trans. Inf. Theory*, vol. 44, no. 6, pp. 2619–2692, Oct. 1998.
[39] F. R. Kschischang, B. J. Frey, and H.-A. Loeliger, "Factor graphs and the sum-product algorithm," *IEEE Trans. Inf. Theory*, vol. 47, no. 2, pp. 498–519, Feb. 2001.
[40] W. Yuan, N. Wu, A. Zhang, X. Huang, Y. Li, and L. Hanzo, "Iterative receiver design for FTN signaling aided sparse code multiple access," *IEEE Trans. Wireless Commun.*, vol. 19, no. 2, pp. 915–928, Feb. 2020.
[41] S. Li *et al.*, "Hybrid MAP and PIC detection for OTFS modulation," 2020, *arXiv:2010.13030*. [Online]. Available: http://arxiv.org/abs/2010.13030
[42] L. Bahl, J. Cocke, F. Jelinek, and J. Raviv, "Optimal decoding of linear codes for minimizing symbol error rate (Corresp.)," *IEEE Trans. Inf. Theory*, vol. IT-20, no. 2, pp. 284–287, Mar. 1974.
[43] J. Yuan and J. Choi, "Channel adaptive space-time (CAST) coding and precoding for wireless downlink packet services," ARC Discovery, Canberra, ACT, Australia, Interim Rep. DP0452151, 2006.
[44] Y. Mostofi and D. C. Cox, "ICI mitigation for pilot-aided OFDM mobile systems," *IEEE Trans. Wireless Commun.*, vol. 4, no. 2, pp. 765–774, Mar. 2005.
[45] *NR; Multiplexing and Channel Coding, Release 15*, document TS 38.212, , 2017.
[46] *Group Radio Access Network, Evolved Universal Terrestrial Radio Access, Multiplexing and Channel Coding, Release 8*, document TS 36.212, 3GPP, 2007.
[47] J. M. Taylor, "The condition of gram matrices and related problems," *Proc. Roy. Soc. Edinburgh, Sect. A Math.*, vol. 80, nos. 1–2, pp. 45–56, 1978.

Shuangyang Li (Graduate Student Member, IEEE) received the B.S. and M.S. degrees from Xidian University, China, in 2013 and 2016, respectively. He is currently pursuing the Ph.D. degree with Xidian University and the University of New South Wales. His research interests include signal processing, channel coding and their applications to communication systems.

Jinhong Yuan (Fellow, IEEE) received the B.E. and Ph.D. degrees in electronics engineering from the Beijing Institute of Technology, Beijing, China, in 1991 and 1997, respectively. From 1997 to 1999, he was a Research Fellow with the School of Electrical Engineering, The University of Sydney, Sydney, NSW, Australia. In 2000, he joined the School of Electrical Engineering and Telecommunications, University of New South Wales, Sydney, NSW, Australia, where he is currently a Professor and the Head of Telecommunication Group. He has published two books, five book chapters, over 300 papers in telecommunications journals and conference proceedings, and 50 industrial reports. He is a co-inventor of one patent on MIMO systems and two patents on low-density-parity-check codes. His current research interests include error control coding and information theory, communication theory, and wireless communications. He has coauthored four Best Paper Awards and one Best Poster Award, including the Best Paper Award from the IEEE International Conference on Communications, Kansas City, USA, in 2018, the Best Paper Award from the IEEE Wireless Communications and Networking Conference, Cancun, Mexico, in 2011, and the Best Paper Award from the IEEE International Symposium on Wireless Communications Systems, Trondheim, Norway, in 2007. He served as the IEEE NSW Chapter Chair for Joint Communications/Signal Processions/Ocean Engineering Chapter from 2011 to 2014 and an Associate Editor for the IEEE TRANSACTIONS ON COMMUNICATIONS from 2012 to 2017. He is currently serving as an Associate Editor for the IEEE TRANSACTIONS ON WIRELESS COMMUNICATIONS.

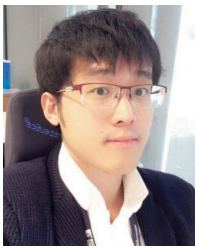


Weijie Yuan (Member, IEEE) received the Ph.D. degree from the Beijing Institute of Technology, China, and the Ph.D. degree from the University of Technology Sydney, Australia, in 2019. He is currently an Assistant Professor with the Department of Electrical and Electronics Engineering, Southern University of Science and Technology. Prior to that, he was a Research Associate with the School of Electrical Engineering and Telecommunications, University of New South Wales, Sydney, NSW, Australia. He has served as a Research Assistant with the University of Sydney, a Visiting Associate Fellow with the University of Wollongong, and a Visiting Fellow with the University of Southampton, from 2017 to 2019. In 2016, he was a Visiting Ph.D. Student with the Institute of Telecommunications, Vienna University of Technology, Austria. He serves as the TPC Co-Chair for the IEEE ICC 2021 Workshop on orthogonal time frequency space (OTFS) and the Co-organizer for SPAWC 2021 Special session on integrated sensing and communication. His research interests include statistical signal processing on graphical models and intelligent transportation systems.



Zhiqiang Wei (Member, IEEE) received the B.E. degree in information engineering from Northwestern Polytechnical University (NPU), Xi'an, China, in 2012, and the Ph.D. degree in electrical engineering and telecommunications from the University of New South Wales, Sydney, NSW, Australia, in 2019. From 2019 to 2020, he was a Post-Doctoral Research Fellow with the University of New South Wales. He is currently a Humboldt Post-Doctoral Research Fellow with the Friedrich-Alexander University Erlangen-Nuremberg. His current research interests include convex/non-convex optimization, resource allocation design, intelligent reflecting surface, millimeter-wave communications, and orthogonal time-frequency space (OTFS) modulation. He received the Best Paper Award from the IEEE International Conference on Communications (ICC), in 2018.

Baoming Bai (Senior Member, IEEE) received the B.S. degree from Northwest Telecommunications Engineering Institute, China, in 1987, and the M.S. and Ph.D. degrees in communication engineering from Xidian University, China, in 1990 and 2000, respectively. From 2000 to 2003, he was a Senior Research Assistant with the Department of Electronic Engineering, City University of Hong Kong. Since April 2003, he has been with the State Key Laboratory of Integrated Services Networks (ISN), School of Telecommunication Engineering, Xidian University, where he is currently a Professor. In 2005, he was with the University of California at Davis, as a Visiting Scholar. His research interests include information theory and channel coding, wireless communication, and quantum communication.



Derrick Wing Kwan Ng (Fellow, IEEE) received the bachelor's degree (Hons.) and the Master of Philosophy (M.Phil.) degree in electronic engineering from The Hong Kong University of Science and Technology (HKUST), in 2006 and 2008, respectively, and the Ph.D. degree from The University of British Columbia (UBC) in 2012.

He was a Senior Post-Doctoral Fellow with the Institute for Digital Communications, Friedrich-Alexander-University Erlangen-Nürnberg (FAU), Germany. He is currently working as a Senior Lecturer and a Scientia Fellow with the University of New South Wales, Sydney, NSW, Australia. His research interests include convex and non-convex optimization, physical layer security, IRS-assisted communication, UAV-assisted communication, wireless information and power transfer, and green (energy-efficient) wireless communications. He received the Australian Research Council (ARC) Discovery Early Career Researcher Award 2017, the Best Paper Award from the WCSP 2020, the IEEE TCGCC Best Journal Paper Award 2018, INISCOM 2018, the IEEE International Conference on Communications (ICC) 2018, the IEEE International Conference on Computing, Networking and Communications (ICNC) 2016, the IEEE Wireless Communications and Networking Conference (WCNC) 2012, the IEEE Global Telecommunication Conference (Globecom) 2011, and the IEEE Third International Conference on Communications and Networking in China 2008. He has been serving as an Editorial Assistant to the Editor-in-Chief for the IEEE TRANSACTIONS ON COMMUNICATIONS from January 2012 to December 2019. He is currently serving as an Editor for the IEEE TRANSACTIONS ON COMMUNICATIONS, and the IEEE TRANSACTIONS ON WIRELESS COMMUNICATIONS, and an Area Editor for the IEEE OPEN JOURNAL OF THE COMMUNICATIONS SOCIETY. He has been listed as a Highly Cited Researcher by Clarivate Analytics since 2018.