Performance Analysis of OTFS with Imperfect Delay-Doppler Channel State Information

Ashwitha Naikoti and Ananthanarayanan Chockalingam
Department of ECE, Indian Institute of Science, Bangalore 560012

***Abstract*—In this paper, we analyze the effect of imperfect delay-Doppler channel state information (CSI) on the bit error rate (BER) performance of orthogonal time frequency space (OTFS) modulation. We carry out the BER analysis when a mismatched maximum-likelihood (ML) detector is used, i.e., when an estimated channel matrix is used for detection in place of the true channel matrix. We derive an exact expression for the pairwise error probability (PEP) using the characteristic function of the decision statistic. Using the PEP expression, we obtain an upper bound on the BER. Our results show that the BER bound is tight at high SNRs. We also obtain the decision rule for the true ML detector with imperfect CSI, which takes into account the delay-Doppler channel estimation error statistics. We quantify the performance gap between the true ML detector and the mismatched ML detector through simulations.**

***Index Terms*—OTFS modulation, delay-Doppler domain, channel estimation error, mismatched and true ML detectors, pairwise error probability, BER analysis.**

I. Introduction

Orthogonal time frequency space (OTFS) modulation is a promising, recently introduced modulation scheme [1]. In OTFS, information symbols are multiplexed in the delay-Doppler (DD) domain rather than in the time-frequency (TF) domain. Also, the time-varying channel is represented in the DD domain, where the channel exhibits a more static behavior. There has been growing interest in this modulation because of its superior performance compared to orthogonal frequency division multiplexing (OFDM) in high-mobility/high-Doppler scenarios [2]-[7]. In terms of bit error performance, OTFS has been shown to significantly outperform OFDM for vehicle speeds up to 500 km/h in 4 GHz band [2] and also in mmWave frequency bands [3], [4]. Also, OTFS is amenable to be implemented as an overlay on existing multicarrier modulation schemes such as OFDM.

Signal detection is crucial in any communication receiver. The task of signal detection requires the knowledge of the channel. Hence, channel estimation at the receiver is essential. Typically, the channel is estimated by transmitting pilot symbols. The received signal corresponding to the pilot symbols is used to estimate the channel. This results in estimation errors depending on the pilot SNR and the method used for channel estimation. Several DD channel estimation methods for OTFS using pilot symbols have been studied [8]-[14]. Embedded pilot based methods where an OTFS frame is embedded with pilot symbols and data symbols are studied in [10]-[12]. In [13],[14], superimposed pilot based methods have been studied, where pilot symbols are superimposed on data symbols. The DD channel state information (CSI) obtained from these methods will be imperfect due to channel estimation errors. It is of interest to analyze the degrading effects of imperfect DD CSI on the bit error performance of OTFS. The above works on OTFS channel estimation focus primarily on algorithmic and system level aspects of DD channel estimation, and performance evaluation is carried out through simulations. To our knowledge, an analytical derivation of the BER performance of OTFS with imperfect DD CSI has not been reported.

Based on the observations made above, in this paper, we present an analytical derivation of the BER performance of OTFS with imperfect DD CSI. We study the performance of the following two detectors: 1) mismatched maximum-likelihood (ML) detector, and 2) true ML detector. The decision rule in the mismatched ML detector is same as that of the conventional ML detector, except that the imperfectly estimated DD channel matrix is used in place of the perfect DD channel matrix in the decision rule. In the true ML detector, the decision rule is obtained by maximizing the likelihood function by considering the imperfectly estimated DD channel statistics. For the mismatched ML detector, we obtain an exact expression for the pairwise error probability (PEP) using the characteristic function of the decision statistic. We verify that the PEP expression is exact by comparing the analytical PEP with the simulated PEP. Using the PEP expression, we obtain an upper bound on the BER, which is found to be tight for large SNR values. For the true ML detector, we obtain the decision rule with imperfect DD CSI and quantify the performance gap between the true and mismatched ML detectors. We also demonstrate the performance of OTFS with phase rotation in the presence of imperfect CSI.

The rest of this paper is organized as follows. Section II presents the OTFS system model with imperfect DD CSI. Section III presents the mismatched ML detector and its performance analysis. Section IV presents the true ML detector and obtains its decision rule. Results and discussions are presented in Sec. V. Conclusions are presented in Sec. VI.

II. OTFS system model

The block diagram of the OTFS modulation scheme is shown in Fig. 1. The OTFS transmitter multiplexes MN information symbols (e.g., PSK/QAM symbols from a modulation alphabet $\mathbb{A}$), denoted by $x[k,l]$, $k = 0, \cdots, N-1$, $l = 0, \cdots, M-1$, onto a $N \times M$ DD grid, where N and M

This work was supported in part by the J. C. Bose National Fellowship, Department of Science and Technology, Government of India.

Fig. 1: OTFS modulation scheme.

are the number of Doppler and delay bins, respectively. These MN symbols are sent over a frame duration of NT seconds using $M\Delta f$ bandwidth, where $\Delta f = 1/T$.

A. System model

The DD domain symbols $x[k,l]$s are mapped to TF domain using inverse symplectic finite Fourier transform (ISFFT) as

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

Using Heisenberg transform, the TF signal in (1) is mapped to a time-domain signal as

$$x(t) = \sum_{n=0}^{N-1}\sum_{m=0}^{M-1} X[n,m]g_{tx}(t-nT)e^{j2\pi m\Delta f(t-nT)}, \quad (2)$$

where $g_{tx}(t)$ defines the shape of the transmit pulse. The transmitted signal is passed through the channel whose DD domain response is given by $h(\tau,\nu) = \sum_{i=1}^{L} h_i\delta(\tau-\tau_i)\delta(\nu-\nu_i)$, where τ_i, ν_i, and h_i denote the delay, Doppler, and channel gain of the ith path, respectively, and L denotes the number of resolvable DD paths. The received signal is given by

$$y(t) = \int_\nu\int_\tau h(\tau,\nu)x(t-\tau)e^{j2\pi\nu(t-\tau)}\mathrm{d}\tau\mathrm{d}\nu + w(t), \quad (3)$$

where $w(t)$ is the additive white Gaussian noise (AWGN). The received signal $y(t)$ is converted into a TF domain signal using Wigner transform as $Y[n,m] = A_{g_{rx},y}(t,f)|_{t=nT,f=m\Delta f}$, where $A_{g_{rx},y}(t,f) = \int g^*_{rx}(t'-t)y(t)e^{-j2\pi f(t'-t)}dt'$ and $g_{rx}(t)$ defines the shape of the receive pulse. When the transmit and receive pulses satisfy the bi-orthogonality condition, the TF input-output relation is obtained as [5]

$$Y[n,m] = H[n,m]X[n,m] + W[n,m], \quad (4)$$

where $H[n,m] = \int_\tau\int_\nu h(\tau,\nu)e^{j2\pi\nu nT}e^{-j2\pi(\nu+m\Delta f)\tau}d\nu d\tau$ and $W[n,m]$ is the TF domain noise. The TF signal in (4) is transformed to DD domain using SFFT as

$$y[k,l] = \sum_{n=0}^{N-1}\sum_{m=0}^{M-1} Y[n,m]e^{-j2\pi(\frac{nk}{N}-\frac{ml}{M})}. \quad (5)$$

From (1)-(5), the DD domain input-output relation can be written as [5]

$$y[k,l] = \frac{1}{MN}\sum_{l'=0}^{N-1}\sum_{k'=0}^{M-1} x[k',l']h_u(\frac{k-k'}{NT},\frac{l-l'}{M\Delta f}) + w[k,l], \quad (6)$$

where $h_u(\nu,\tau)$ denotes the circular convolution of the channel response with a windowing function $u(\tau,\nu)$ and $h_u(\frac{k-k'}{NT},\frac{l-l'}{M\Delta f}) = h_u(\nu,\tau)|_{\nu=\frac{k-k'}{NT},\tau=\frac{l-l'}{M\Delta f}}$. The DD domain input-output relation in (6) can be written as [5]

$$y[k,l] = \sum_{i=1}^{L} h'_i x[(k-b_i)_N,(l-a_i)_M] + w[k,l], \quad (7)$$

where b_i and a_i are integers corresponding to indices of Doppler and delay[1], respectively, for the ith path, i.e., $\tau_i \triangleq \frac{a_i}{M\Delta f}$ and $\nu_i \triangleq \frac{b_i}{NT}$, $w[k,l]$ is the AWGN, and $h'_i = h_i e^{-j2\pi\nu_i\tau_i}$, where h_is are assumed to be i.i.d $\mathcal{CN}(0,1/L)$. The input-output relation in (7), $\forall k,l$, can be written in a vectorized form as

$$\mathbf{y} = \mathbf{H}\mathbf{x} + \mathbf{w}, \quad (8)$$

where $\mathbf{x},\mathbf{y},\mathbf{w} \in \mathbb{C}^{MN\times 1}$, the $(k+Nl)$th entry of $\mathbf{x}$, $x_{k+Nl} = x[k,l]$, $k = 0,1,\cdots,N-1$, $l = 0,1,\cdots,M-1$. Similarly, $y_{k+Nl} = y[k,l]$ and $w_{k+Nl} = w[k,l]$, $k = 0,1,\cdots,N-1, l = 0,1,\cdots,M-1$, and $\mathbf{H} \in \mathbb{C}^{MN\times MN}$ is the effective channel matrix, whose jth row ($j = k+Nl$), denoted by $\mathbf{H}[j]$, is given by $\mathbf{H}[j] = [\tilde{h}((k-0)_N,(l-0)_M)\ \ \tilde{h}((k-1)_N,(l-0)_M)\ \cdots \tilde{h}((k-N-1)_N,(l-M-1)_M)]$, where $\tilde{h}(k,l)$ is given by

$$\tilde{h}(k,l) = \begin{cases} h'_i & \text{if } k = b_i,\ l = a_i,\ i \in \{1,\cdots,L\} \\ 0 & \text{otherwise.} \end{cases} \quad (9)$$

Note that the $\mathbf{H}$ matrix has only L non-zero entries in each row and column, i.e., there are only MNL non-zero elements in $\mathbf{H}$. The linear vector channel model in (8) is used for OTFS signal detection/equalization and DD channel estimation. If the receiver is assumed to have perfect knowledge of $\mathbf{H}$, the ML detection rule is given by

$$\hat{\mathbf{x}} = \underset{\mathbf{x}\in\mathbb{A}^{MN}}{\arg\min}\ \|\mathbf{y}-\mathbf{H}\mathbf{x}\|^2. \quad (10)$$

B. Channel estimation error

The knowledge of $\mathbf{H}$ is needed at the receiver to detect the transmitted OTFS signal vector $\mathbf{x}$. An estimate of $\mathbf{H}$ can be obtained by transmitting pilot symbols and using suitable channel estimation schemes at the receiver. The estimate $\hat{h}_i$ of the channel gain h_i of the ith path obtained using pilot based estimation techniques can be written as

$$\hat{h}_i = h_i + e_i, \quad (11)$$

where e_i is the channel estimation error corresponding to h_i. We assume that e_is are independent of each other and also independent of h_is, and are identically distributed as $\mathcal{CN}(0,\sigma_e^2)$, where σ_e^2 is the variance of the channel estimation error. The statistics of the channel estimation error depend on the estimation scheme used. From (11), note that $\hat{h}_i$s are i.i.d and distributed as $\mathcal{CN}(0,\sigma_h^2+\sigma_e^2)$, where $\sigma_h^2 = 1/L$. Let $\hat{\mathbf{H}}$ denote the estimated channel matrix comprising of the estimated channel gains $\hat{h}_i$s and $\mathbf{E}$ denote the estimation error matrix comprising of e_is. From (11), $\hat{\mathbf{H}}$ can be written as

$$\hat{\mathbf{H}} = \mathbf{H} + \mathbf{E}. \quad (12)$$

The σ_e^2 value is an indicator of the channel estimate's. Lower is the value of σ_e^2, better is the estimate. When $\sigma_e^2 = 0$, we have $\hat{\mathbf{H}} = \mathbf{H}$, i.e., the channel matrix is perfectly known.

[1]A similar input-output relation can be obtained for fractional Dopplers, as described in [5].

III. Mismatched ML Detector

In this section, we analyze the BER performance of the mismatched ML detector. The mismatched ML detector employs the decision rule in (10) by using the estimated matrix $\hat{\mathbf{H}}$ instead of $\mathbf{H}$, i.e., the decision rule for this detector is

$$\hat{\mathbf{x}} = \underset{\mathbf{x}\in\mathbb{A}^{MN}}{\arg\min} \|\mathbf{y} - \hat{\mathbf{H}}\mathbf{x}\|^2. \tag{13}$$

We first derive the PEP expression for mismatched ML detector using characteristic function approach. Assuming that all the transmit vectors are equally probable, the probability of the transmitted signal vector $\mathbf{x}$ being detected as $\mathbf{x}'$ using the rule in (13) is given by

$$\mathrm{P}(\mathbf{x}\to\mathbf{x}') = \mathrm{P}(\|\mathbf{y} - \hat{\mathbf{H}}\mathbf{x}'\|^2 < \|\mathbf{y} - \hat{\mathbf{H}}\mathbf{x}\|^2). \tag{14}$$

Using (8) and (12), (14) can be written as

$$\begin{aligned}\mathrm{P}(\mathbf{x}\to\mathbf{x}') &= \mathrm{P}\left(\|\mathbf{H}(\mathbf{x}-\mathbf{x}') + \mathbf{w} - \mathbf{E}\mathbf{x}'\|^2 < \|\mathbf{w}-\mathbf{E}\mathbf{x}\|^2\right)\\ &= \mathrm{P}(\|\mathbf{H}(\mathbf{x}-\mathbf{x}') + \mathbf{w} - \mathbf{E}\mathbf{x}'\|^2 - \|\mathbf{w}-\mathbf{E}\mathbf{x}\|^2 < 0)\\ &= \mathrm{P}(S<0),\end{aligned} \tag{15}$$

where $S \triangleq \|\mathbf{H}(\mathbf{x}-\mathbf{x}') + \mathbf{w} - \mathbf{E}\mathbf{x}'\|^2 - \|\mathbf{w}-\mathbf{E}\mathbf{x}\|^2$ is the decision statistic. Now, defining $\mathbf{a} \triangleq \mathbf{H}(\mathbf{x}-\mathbf{x}') + \mathbf{w} - \mathbf{E}\mathbf{x}'$ and $\mathbf{b} \triangleq \mathbf{w} - \mathbf{E}\mathbf{x}$, S can be written as

$$S = \|\mathbf{a}\|^2 - \|\mathbf{b}\|^2, \tag{16}$$

which can be further written in Hermitian quadratic form as

$$S = \mathbf{u}^H\mathbf{Q}\mathbf{u}, \tag{17}$$

where $\mathbf{u} = [\mathbf{a}^T\ \mathbf{b}^T]^T$ and $\mathbf{Q} = \mathrm{diag}(\mathbf{I}_{MN}, -\mathbf{I}_{MN})$. Now, since the channel matrix $\mathbf{H}$ has only L non-zero elements in each row and column (due to modulo operations), we can write the vectorized relation in (8) in an alternate form as [6]

$$\mathbf{y} = \mathbf{X}\mathbf{h} + \mathbf{w}, \tag{18}$$

where $\mathbf{h} = [h_1, \cdots, h_L]^T$ and $\mathbf{X}$ is an $MN \times L$ matrix whose ith row, denoted by $\mathbf{X}[i]$, $i = k + Nl,\ \forall k, l$, is given by

$$\mathbf{X}[i] = \begin{bmatrix} x_{(k-b_1)_N + N(l-a_1)_M} \\ x_{(k-b_2)_N + N(l-a_2)_M} \\ \vdots \\ x_{(k-b_L)_N + N(l-a_L)_M} \end{bmatrix}^T. \tag{19}$$

From the system model in (18), $\mathbf{H}\mathbf{x} = \mathbf{X}\mathbf{h}$, $\mathbf{H}\mathbf{x}' = \mathbf{X}'\mathbf{h}$, $\mathbf{E}\mathbf{x} = \mathbf{X}\mathbf{e}$, and $\mathbf{E}\mathbf{x}' = \mathbf{X}'\mathbf{e}$, where $\mathbf{e} = [e_1, \cdots, e_L]^T$. Then, we can rewrite $\mathbf{a}$ and $\mathbf{b}$ as

$$\mathbf{a} = (\mathbf{X}-\mathbf{X}')\mathbf{h} + \mathbf{w} - \mathbf{X}'\mathbf{e},\ \ \mathbf{b} = \mathbf{w} - \mathbf{X}'\mathbf{e}. \tag{20}$$

From (20), $\mathbf{u}$ can be written as

$$\begin{aligned}\mathbf{u} &= \begin{bmatrix}\mathbf{a}\\ \mathbf{b}\end{bmatrix} = \underbrace{\begin{bmatrix}\mathbf{X}-\mathbf{X}' & -\mathbf{X}' & \mathbf{I}_{MN}\\ \mathbf{0}_{MN\times L} & -\mathbf{X} & \mathbf{I}_{MN}\end{bmatrix}}_{\triangleq\mathbf{G}} \underbrace{\begin{bmatrix}\mathbf{h}\\ \mathbf{e}\\ \mathbf{w}\end{bmatrix}}_{\triangleq\mathbf{n}}\\ &= \mathbf{G}\mathbf{n},\end{aligned} \tag{21}$$

where $\mathbf{0}_{MN\times L}$ is an all-zero matrix with MN rows and L columns, $\mathbf{G}$ is a $2MN\times(MN+2L)$ matrix, and $\mathbf{n}$ is a $(MN+2L)\times 1$ vector. It is noted that $\mathbf{n}$ is a circularly-symmetric complex Gaussian random vector, and its covariance $\mathbf{K_n}$ is given by

$$\mathbf{K_n} = \mathrm{diag}(\sigma_h^2\mathbf{I}_L, \sigma_e^2\mathbf{I}_L, \sigma_w^2\mathbf{I}_{MN}), \tag{22}$$

Therefore, the vector $\mathbf{u}$ is also a circularly-symmetric Gaussian random vector whose covariance $\mathbf{K_u}$ is given by [15]

$$\begin{aligned}\mathbf{K_u} &= \mathbf{G}\mathbf{K_n}\mathbf{G}^H\\ &= \begin{bmatrix} \sigma_h^2(\mathbf{X}-\mathbf{X}')(\mathbf{X}-\mathbf{X}')^H + \sigma_e^2\mathbf{X}'\mathbf{X}'^H + \sigma_w^2\mathbf{I}_{MN} & \sigma_e^2\mathbf{X}\mathbf{X}^H + \sigma_w^2\mathbf{I}_{MN} \\ \sigma_e^2\mathbf{X}\mathbf{X}'^H + \sigma_w^2\mathbf{I}_{MN} & \sigma_e^2\mathbf{X}\mathbf{X}^H + \sigma_w^2\mathbf{I}_{MN}\end{bmatrix}.\end{aligned} \tag{23}$$

Thus, the decision statistic S is a Hermitian quadratic form that allows its characteristic function $\Phi_S(j\omega)$ to be written in closed-form as [16]

$$\begin{aligned}\Phi_S(j\omega) = \mathbb{E}[e^{j\omega S}] &= \frac{1}{\det\left(\mathbf{I}_{2MN} - j\omega\mathbf{K_u}\mathbf{Q}\right)}\\ &= \frac{1}{\det\left(\mathbf{I}_{2MN} - j\omega\mathbf{A}\right)},\end{aligned} \tag{24}$$

where $\mathbb{E}[.]$ denotes the expectation operator, $j = \sqrt{-1}$, and $\mathbf{A} \triangleq \mathbf{K_u}\mathbf{Q}$. Let $\Lambda_1, \Lambda_2, \cdots, \Lambda_{2MN}$ denote the eigenvalues of $\mathbf{A}$. The characteristic function in (24) can be simplified as

$$\Phi_S(j\omega) = \frac{1}{\prod_{i=1}^{2MN}(1 - j\omega\Lambda_i)}. \tag{25}$$

After changing the variable to $z = j\omega$, the PEP expression can be obtained from the characteristic function using inversion theorem as [17]

$$\begin{aligned}\mathrm{P}(\mathbf{x}\to\mathbf{x}') &= -\text{sum of the residues of } \frac{\Phi_S(z)}{z} \text{ computed at poles on negative } z\text{-plane}\\ &= -\sum_k \frac{1}{(m_k-1)!}\frac{d^{m_k-1}}{dz^{m_k-1}}\left\{(z-\Lambda_k)^{m_k}\frac{\Phi_S(z)}{z}\right\},\end{aligned} \tag{26}$$

where Λ_ks are the negative eigenvalues of $\mathbf{A}$ (i.e., $\Re\{\Lambda_k\} < 0$) having multiplicity m_k. Now, denoting the negative eigenvalues by λ_i having multiplicities c_i, $i = 1, 2, \cdots, K_n$ and denoting the non-negative eigenvalues by ν_j having multiplicities d_j, $j = 1, 2, \cdots, K_p$, such that $\sum_i c_i + \sum_j d_j = 2MN$, and using Faa di Bruno's formula, the $m_k - 1$th derivative in (26) can be obtained and a closed-form expression for the PEP in (26) is obtained as [18],[19]

$$\begin{aligned}\mathrm{P}(\mathbf{x}\to\mathbf{x}') &= \sum_{i=1}^{K_n}\frac{(-\lambda_i)^{2MN-c_i}}{\prod_{j=1}^{K_p}(\mu_j-\lambda_i)^{d_j}\prod_{k\neq i}(\lambda_k-\lambda_i)^{c_k}}\\ &\cdot \sum_{(l_1,l_2,\cdots,l_{c_i-1})}\prod_{m=1}^{c_i-1}\frac{1}{l_m!}\cdot\frac{1}{m}\left[1 + \left(\sum_{j=1}^{K_p}\frac{d_j\nu_j^m}{(\nu_j-\lambda_i)^m} + \sum_{k\neq i}\frac{c_k\lambda_k^m}{(\lambda_k-\lambda_i)^m}\right)\right]^{l_m},\end{aligned} \tag{27}$$

where $l_1, l_2, \cdots, l_{c_i-1}$ are such that $0 \le l_1, l_2, \cdots, l_{c_i-1} \le c_i - 1$ and $\sum_{n=1}^{c_i-1} n l_n = c_i - 1$. The possible enumerations of $(l_1, l_2, \cdots, l_{c_i-1})$ satisfying the above conditions can be precomputed and kept in a lookup table. With the enumerations

readily available, the PEP expression in (27) can be easily calculated. Now, using the PEP expression, an upper bound on the BER can be obtained using union bound as

$$BER \leq \frac{1}{2^{MN}} \sum_{\mathbf{x}} \sum_{\mathbf{x}' \neq \mathbf{x}} \mathrm{P}(\mathbf{x} \to \mathbf{x}') \frac{\eta_{\mathbf{x},\mathbf{x}'}}{MN}, \tag{28}$$

where $\eta_{\mathbf{x},\mathbf{x}'}$ denotes the number of bits in which $\mathbf{x}$ and $\mathbf{x}'$ differ.

IV. True ML Detector

In this section, we obtain the decision rule for the true ML detector for OTFS with imperfect DD CSI and compare its bit error performance with that of the mismatched ML detector. The true ML criterion when the channel estimate $\hat{\mathbf{H}}$ is available at the receiver is

$$\hat{\mathbf{x}} = \arg \max_{\mathbf{x} \in \mathbb{A}^{MN}} \mathrm{P}(\mathbf{y}|\hat{\mathbf{H}}, \mathbf{x}). \tag{29}$$

To obtain the decision statistic for the true ML detector, the probability distribution of $\mathbf{y}$ conditioned on $\hat{\mathbf{H}}$ and $\mathbf{x}$ has to be obtained. It can be seen in (11) that h_i depends only on $\hat{h}_i$ and $\hat{h}_i$s are independent of each other. Thus, h_is conditioned on $\hat{\mathbf{H}}$ are independent. The jth element y_j of $\mathbf{y}$ is given by

$$y_j = \sum_{i=1}^{L} h_i x_{(k-b_i)_N + N(l-a_i)_M} + w_j, \quad j = k + Nl. \tag{30}$$

The linear combination in (30) is a Gaussian r.v. when conditioned on $\hat{\mathbf{H}}$ [20]. Therefore, conditioned on $\hat{\mathbf{H}}$ and $\mathbf{x}$, the received vector $\mathbf{y}$ is a complex Gaussian random vector. Its mean vector $\boldsymbol{\mu}_{\mathbf{y}|\hat{\mathbf{H}},\mathbf{x}}$ is obtained as

$$\begin{aligned} \boldsymbol{\mu}_{\mathbf{y}|\hat{\mathbf{H}},\mathbf{x}} &= \mathbb{E}[\mathbf{y}|\hat{\mathbf{H}}, \mathbf{x}] \\ &= \mathbb{E}[\mathbf{H}|\hat{\mathbf{H}}]\mathbf{x} + \mathbb{E}[\mathbf{w}] \\ &= \rho^2 \hat{\mathbf{H}}\mathbf{x}, \end{aligned} \tag{31}$$

where $\rho^2 = \sigma_h^2/(\sigma_h^2 + \sigma_e^2)$. The covariance matrix $\mathbf{K}_{\mathbf{y}|\hat{\mathbf{H}},\mathbf{x}}$ is obtained as

$$\begin{aligned} \mathbf{K}_{\mathbf{y}|\hat{\mathbf{H}},\mathbf{x}} &= \mathbb{E}[(\mathbf{y} - \boldsymbol{\mu}_{\mathbf{y}|\hat{\mathbf{H}},\mathbf{x}})(\mathbf{y} - \boldsymbol{\mu}_{\mathbf{y}|\hat{\mathbf{H}},\mathbf{x}})^H | \hat{\mathbf{H}}, \mathbf{x}] \\ &= \mathbb{E}[(\mathbf{H}\mathbf{x} + \mathbf{w} - \rho^2 \hat{\mathbf{H}}\mathbf{x})(\mathbf{H}\mathbf{x} + \mathbf{w} - \rho^2 \hat{\mathbf{H}}\mathbf{x})^H | \hat{\mathbf{H}}, \mathbf{x}] \\ &= \sigma_w^2 \mathbf{I}_{MN} + \mathbb{E}[(\mathbf{H} - \rho^2 \hat{\mathbf{H}})\mathbf{x}\mathbf{x}^H (\mathbf{H} - \rho^2 \hat{\mathbf{H}})^H | \hat{\mathbf{H}}, \mathbf{x}] \\ &= \sigma_w^2 \mathbf{I}_{MN} + \mathbb{E}[\mathbf{X}(\mathbf{h} - \rho^2 \hat{\mathbf{h}})(\mathbf{h} - \rho^2 \hat{\mathbf{h}})^H \mathbf{X}^H | \hat{\mathbf{H}}, \mathbf{x}] \\ &= \sigma_w^2 \mathbf{I}_{MN} + \mathbf{X}\mathbb{E}[(\mathbf{h} - \rho^2 \hat{\mathbf{h}})(\mathbf{h} - \rho^2 \hat{\mathbf{h}})^H | \hat{\mathbf{H}}] \mathbf{X}^H \\ &= \sigma_w^2 \mathbf{I}_{MN} + \sigma_h^2 (1 - \rho^2) \mathbf{X}\mathbf{X}^H, \end{aligned} \tag{32}$$

where $\mathbf{X}$ and $\mathbf{h}$ are as in (18), $\hat{\mathbf{h}} = [\hat{h}_1, \cdots, \hat{h}_L]^T$, and $\mathbb{E}[(\mathbf{h} - \rho^2 \hat{\mathbf{h}})(\mathbf{h} - \rho^2 \hat{\mathbf{h}})^H | \hat{\mathbf{H}}] = \sigma_h^2(1 - \rho^2)$ is due to the independence among the elements of $\mathbf{H}$ having covariance $\sigma_h^2(1 - \rho^2)$ conditioned on $\hat{\mathbf{H}}$. The pseudo-covariance matrix $\mathbf{C}_{\mathbf{y}|\hat{\mathbf{H}},\mathbf{x}}$ is given by

$$\begin{aligned} \mathbf{C}_{\mathbf{y}|\hat{\mathbf{H}},\mathbf{x}} &= \mathbb{E}[(\mathbf{y} - \boldsymbol{\mu}_{\mathbf{y}|\hat{\mathbf{H}},\mathbf{x}})(\mathbf{y} - \boldsymbol{\mu}_{\mathbf{y}|\hat{\mathbf{H}},\mathbf{x}})^T | \hat{\mathbf{H}}, \mathbf{x}] \\ &= \mathbb{E}[(\mathbf{H}\mathbf{x} + \mathbf{w} - \rho^2 \hat{\mathbf{H}}\mathbf{x})(\mathbf{H}\mathbf{x} + \mathbf{w} - \rho^2 \hat{\mathbf{H}}\mathbf{x})^T | \hat{\mathbf{H}}, \mathbf{x}] \\ &= \mathbb{E}[\mathbf{w}\mathbf{w}^T] + \mathbf{X}\mathbb{E}[(\mathbf{h} - \rho^2 \hat{\mathbf{h}})(\mathbf{h} - \rho^2 \hat{\mathbf{h}})^T | \hat{\mathbf{H}}] \mathbf{X}^T \\ &= \mathbf{0}_{MN \times MN}, \end{aligned} \tag{33}$$

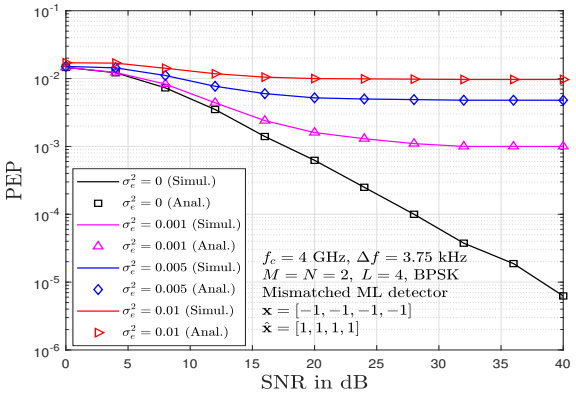

Fig. 2: Analytical PEP versus simulated PEP of the mismatched ML detector for $\mathbf{x} = [-1, -1, -1, -1]$ and $\hat{\mathbf{x}} = [1, 1, 1, 1]$ with imperfect DD CSI.

where (33) is due to the independence among the elements of $\mathbf{h}$ having zero pseudo-covariance when conditioned on $\hat{\mathbf{H}}$. Therefore, conditioned on $\hat{\mathbf{H}}$ and $\mathbf{x}$, the received vector $\mathbf{y}$ is a complex Gaussian random vector with mean vector $\boldsymbol{\mu}_{\mathbf{y}|\hat{\mathbf{H}},\mathbf{x}} = \rho^2 \hat{\mathbf{H}}\mathbf{x}$, covariance matrix $\mathbf{K}_{\mathbf{y}|\hat{\mathbf{H}},\mathbf{x}} = \sigma_w^2 \mathbf{I}_{MN} + \sigma_h^2(1 - \rho^2)\mathbf{X}\mathbf{X}^H$, and pseudo-covariance matrix $\mathbf{C}_{\mathbf{y}|\hat{\mathbf{H}},\mathbf{x}} = \mathbf{0}_{MN \times MN}$. With the knowledge of the statistics of $\mathbf{y}$ conditioned on $\hat{\mathbf{H}}$ and $\mathbf{x}$, we can write the decision rule in (29) as

$$\hat{\mathbf{x}} = \arg \max_{\mathbf{x} \in \mathbb{A}^{MN}} \frac{1}{(2\pi)^{MN} \sqrt{\det(\mathbf{K}_{\mathbf{y}})}} e^{-\frac{1}{2} \mathbf{y}_\mu^T \mathbf{K}_{\mathbf{y}}^{-1} \mathbf{y}_\mu}, \tag{34}$$

where

$$\mathbf{y}_\mu = \left[\Re\{\mathbf{y} - \boldsymbol{\mu}_{\mathbf{y}|\hat{\mathbf{H}},\mathbf{x}}\}, \Im\{\mathbf{y} - \boldsymbol{\mu}_{\mathbf{y}|\hat{\mathbf{H}},\mathbf{x}}\}\right]^T,$$

$$\mathbf{K}_{\mathbf{y}} = \frac{1}{2} \begin{bmatrix} \Re\{\mathbf{K}_{\mathbf{y}|\hat{\mathbf{H}},\mathbf{x}} + \mathbf{C}_{\mathbf{y}|\hat{\mathbf{H}},\mathbf{x}}\} & -\Im\{\mathbf{K}_{\mathbf{y}|\hat{\mathbf{H}},\mathbf{x}} - \mathbf{C}_{\mathbf{y}|\hat{\mathbf{H}},\mathbf{x}}\} \\ \Im\{\mathbf{K}_{\mathbf{y}|\hat{\mathbf{H}},\mathbf{x}} + \mathbf{C}_{\mathbf{y}|\hat{\mathbf{H}},\mathbf{x}}\} & \Re\{\mathbf{K}_{\mathbf{y}|\hat{\mathbf{H}},\mathbf{x}} - \mathbf{C}_{\mathbf{y}|\hat{\mathbf{H}},\mathbf{x}}\} \end{bmatrix},$$

and $\Re(\cdot)$ and $\Im(\cdot)$, are the real part and imaginary part, respectively.

V. Results and Discussions

In this section, we present the BER performance of the mismatched ML detector and the true ML detector with imperfect DD CSI. The following OTFS system parameters are used in the numerical evaluation and simulations. A carrier frequency of 4 GHz and a subcarrier spacing of $\Delta f = 3.75$ KHz are used. The OTFS frame size is $M = N = 2$ and the modulation used is BPSK. The number of DD channel paths is $L = 4$ and the profile of DD values (τ, ν) of the paths is given by $\{(\frac{1}{M\Delta f}, 0), (\frac{1}{M\Delta f}, \frac{1}{NT}), (\frac{1}{M\Delta f}, 0), (\frac{1}{M\Delta f}, \frac{1}{NT})\}$. Uniform power delay profile is assumed.

In Fig. 2, we compare the analytically derived PEP for the mismatched ML detector with the corresponding simulated PEP. The transmitted signal vector is taken to be $\mathbf{x} = [-1, -1, -1, -1]$. The probability of this transmitted vector being wrongly decoded as $\hat{\mathbf{x}} = [1, 1, 1, 1]$ is obtained by using (27) as well as by simulation. This is one of the pairs which have a dominant PEP causing the BER to floor. The PEP curves for $\sigma_e^2 = 0, 0.001, 0.005, 0.01$ are obtained by varying the SNR. It can been seen in Fig. 2 that the

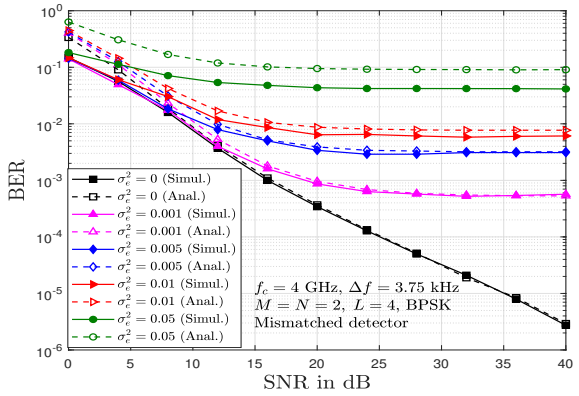

Fig. 3: Analytical BER upper bound versus simulated BER performance of the mismatched ML detector with imperfect DD CSI.

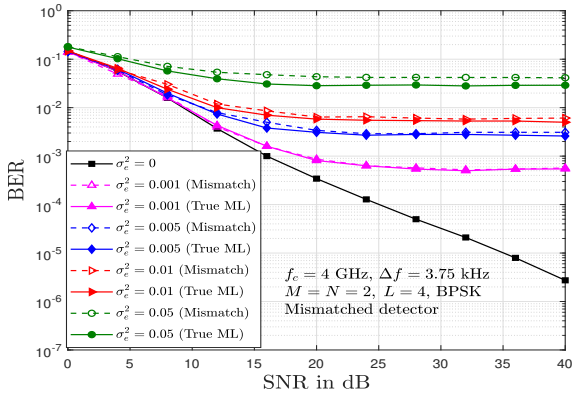

Fig. 4: Comparison between the BER performance of the mismatched ML detector and the true ML detector with imperfect DD CSI.

PEP obtained through simulation is same as the PEP obtained through analysis, thus validating the analysis.

Figure 3 demonstrates the closeness of the BER upper bound with the simulated BER of the mismatched ML detector. We see that, as the SNR increases, the analytical BER upper bound gets close to the simulated BER. We further see that the BER floors at higher SNRs and the flooring occurs at higher BER values for increasing values of σ_e^2. For example, at $\sigma_e^2 = 0.001$, the BER starts flooring at 24 dB SNR at a BER value of about 5×10^{-4}. On the other hand, when $\sigma_e^2 = 0.01$, the BER starts flooring at 20 dB SNR at a BER value of about 7×10^{-2}.

Figure 4 shows the comparison between BER performance of mismatched ML detector and true ML detector with imperfect CSI. It can be seen that both the mismatched and true ML detectors have almost the same performance for small channel estimation errors. However, the true ML detector outperforms the mismatched ML detector when the channel estimation error increases. For example, when $\sigma_e^2 = 0.001, 0.01$, both the detectors have almost same performance, whereas the true ML detector has superior performance when $\sigma_e^2 = 0.05$. This is because of the suboptimality of the mismatched ML detector in the presence of imperfect CSI.

A. OTFS with phase rotation

It has been shown in [6] that the basic OTFS scheme achieves an asymptotic diversity order of one. It has also been shown that OTFS, when used with phase rotation, achieves the full diversity of L [6]. Here, we are interested in assessing how OTFS with phase rotation performs in the presence of imperfect DD CSI.

In OTFS with phase rotation, instead of transmitting the vector $\mathbf{x}$ as in the case of the basic OTFS scheme, a phase rotated version of the $\mathbf{x}$ vector is transmitted. That is, the transmitted vector in OTFS with phase rotation, denoted by $\tilde{\mathbf{x}}$, is given by

$$\tilde{\mathbf{x}} = \mathbf{\Phi}\mathbf{x}, \tag{35}$$

where the phase rotation matrix $\mathbf{\Phi}$ is given by

$$\mathbf{\Phi} = \text{diag}(\phi_0^{(0)}, \cdots, \phi_{N-1}^{(0)}, \phi_0^{(1)}, \cdots, \phi_{N-1}^{(1)}, \cdots, \phi_{N-1}^{(M-1)}), \tag{36}$$

where $\phi_r^{(s)} = e^{jq_r^{(s)}}$, $r = 0, \cdots, N-1$, $s = 0, \cdots, M-1$ are transcendental numbers with $q_r^{(s)}$ real and distinct.

Figure 5 compares the BER performance of OTFS without and with phase rotation using mismatched ML detector. The phase rotation matrix used is $\mathbf{\Phi} = \text{diag}(1, e^{j\frac{1}{MN}}, \cdots, e^{j\frac{MN-1}{MN}})$. It can be seen that the BER performance with phase rotation is superior compared to that without phase rotation. For example, at $\sigma_e^2 = 0.001$ and SNR = 40 dB, the BER without phase rotation is in the order of 10^{-4} while it is in the order of 10^{-6} with phase rotation. In Fig. 6, a similar performance improvement in OTFS with phase rotation is observed when true ML detector is used.

VI. Conclusions

We analyzed the BER performance of OTFS modulation in the presence of imperfect DD CSI, which has not been reported before. Using the characteristic function approach, we derived an exact closed-form expression for the pairwise error probability of the mismatched ML detector for OTFS, and obtained a BER upper bound which is tight in the medium to high SNR regime. Further, we considered the true ML detector and obtained its decision rule by considering the statistics of channel estimation error. We showed that true ML detector has better performance compared to the mismatched ML detector. We also showed the performance of OTFS with phase rotation in the presence of imperfect DD CSI.

References

[1] R. Hadani *et al.*, "Orthogonal time frequency space modulation," *Proc. IEEE WCNC'2017*, pp. 1-6, Mar. 2017.

[2] R. Hadani *et al.*, "Orthogonal time frequency space modulation," Aug. 2018. [Online]. Available: https://arxiv.org/abs/1808.00519v1

[3] F. Wiffen *et al.*, 'Comparison of OTFS and OFDM in ray launched sub-6 GHz and mmWave line-of-sight mobility channels," *Proc. IEEE PIMRC'2018*, pp. 73-79, Sep. 2018.

[4] R. Hadani *et al.*, "Orthogonal time frequency space (OTFS) modulation for millimeter-wave communications systems," *Proc. IEEE MTT-S Intl. Microwave Symp.*, pp. 681-683, Jun. 2017.

[5] 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.

[6] 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.

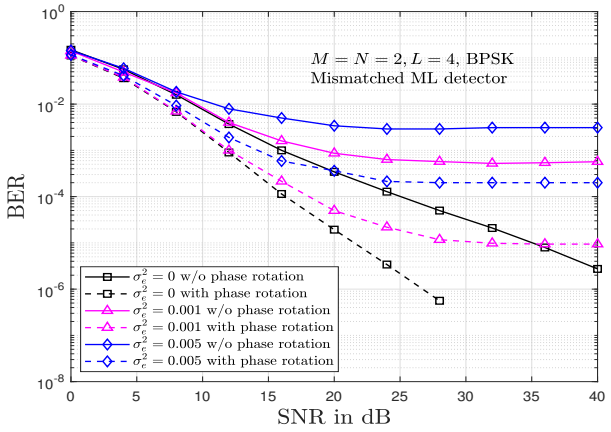

Fig. 5: BER performance of OTFS without and with phase rotation using mismatched ML detector with imperfect DD CSI.

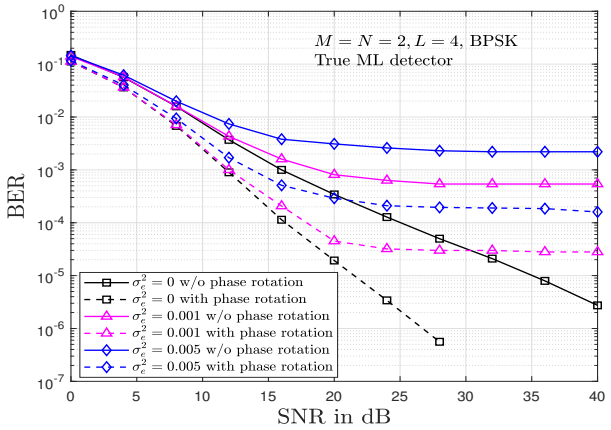

Fig. 6: BER performance of OTFS without and with phase rotation using true ML detector with imperfect DD CSI.

[7] S. K. Mohammed, "Derivation of OTFS modulation from first principles," *IEEE Trans. Veh. Tech.*, vol. 70, no. 8, pp. 7619-7636, Mar. 2021.

[8] K. R. Murali and A. Chockalingam, "On OTFS modulation for high-Doppler fading channels," *Proc. ITA'2018*, pp. 1-10, Feb. 2018.

[9] M. K. Ramachandran and A. Chockalingam, "MIMO-OTFS in high-Doppler fading channels: signal detection and channel estimation, *Proc. IEEE GLOBECOM'2018*, Dec. 2018.

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

[11] L. Zhao, W-J. Gao, and W. Guo, "Sparse Bayesian learning of delay-Doppler channel for OTFS system," *IEEE Commun. Lett.*, vol. 24, no. 12, pp. 2766-2769, Dec. 2020.

[12] S. Srivastava, R. K. Singh, A. K. Jagannatham, and L. Hanzo, "Bayesian learning aided sparse channel estimation for orthogonal time frequency space modulated systems," *Proc. IEEE Trans. Veh. Tech.*, vol. 70, no. 8, pp. 8343-8348, Aug. 2021.

[13] H. B. Mishra, P. Singh, A. K. Prasad, and R. Budhiraja, "OTFS channel estimation and data detection designs with superimposed pilots," *IEEE Trans. Wireless Commun.*, 2021, (early access) doi: 10.1109/TWC.2021.3110659.

[14] W. Yuan *et al.*, "Data-aided channel estimation for OTFS systems with a superimposed pilot and data transmission Scheme," *IEEE Wireless Commun. Lett.*, vol. 10, no. 9, pp. 1954-1958, Sep. 2021.

[15] R. G. Gallager, "Circularly-symmetric Gaussian random vectors," Jan. 2008. [Online]. Available: http://www.rle.mit.edu/rgallager/documents/CircSymGauss.pdf

[16] G. L. Turin, "The characteristic function of Hermitian quadratic forms in complex normal variables," *Biometrika*, vol. 47, nos. 1/2, pp. 199–201, Jun. 1960.

[17] J. Gil-Pelaez, "Note on the inversion theorem," *Biometrika*, vol. 38, pp. 481–482, 1951.

[18] C.-J. de la Vallee Poussin, Cours D'Analyze Infinitesimale, 12th ed, Paris: Gauthier-Villars, Libraire Universitaire Louvain, 1959, vol. 1.

[19] P. Garg, R. K. Mallik and H. M. Gupta, "Performance analysis of space-time coding with imperfect channel estimation," *IEEE Trans. Wireless Commun.*, vol. 4, no. 1, pp. 257-265, Jan. 2005.

[20] D. P. Bertsekas and J. N. Tsitsiklis, *Introduction to Probability*, 2nd ed., Belmont, MA, USA: Athena Sci., 2002.