Phase-Noise Compensating Receiver for Zak-OTFS

Naveed Bin Nazir and Ananthanarayanan Chockalingam

***Abstract*—In this letter, we investigate the effect of oscillator phase noise (PN) on the performance of Zak-OTFS receiver and propose a novel PN compensation scheme using a learning framework. Towards this end, we first derive a closed-form expression for the input-output (I/O) relation for Zak-OTFS in the presence of PN, which has not been reported before. Next, leveraging the structure of the PN-induced terms in the I/O relation, we develop a PN-compensating receiver with learning-aided refinement. In the proposed receiver, a lightweight convolutional neural network (CNN) estimates the effective channel, whose output is used by a lightweight deep neural network (DNN) to estimate the PN derivative. This estimate is then used to construct a PN-aware effective channel for minimum mean square error (MMSE) detection. Simulation results with Veh-A channel model with fractional delays and Dopplers show that the proposed receiver is able to effectively compensate the PN and achieve robust performance over a wide range of PN conditions.**

***Index Terms*—Zak-OTFS modulation, delay-Doppler domain, oscillator phase noise, I/O relation, phase noise compensation.**

I. Introduction

Orthogonal time frequency space (OTFS) modulation is a delay-Doppler (DD) modulation suited for high-mobility channels [1]. The Zak-OTFS framework, based on Zak transform theory, allows for a direct mapping between the DD and time domains [2]. This framework yields a predictable, non-fading input-output (I/O) relation characterized by twisted convolution. Consequently, Zak-OTFS provides a systematic framework for theoretical analysis, waveform design, and receiver algorithm development. While research interest on Zak-OTFS is on the rise, the topic of sensitivity of Zak-OTFS to practical system impairments (e.g., phase noise (PN), IQ imbalance, non-linearity, carrier frequency/timing offset [3]) and their compensation remains largely unexplored. This forms the basic motivation behind the work in this letter. Among these practical impairments, PN is particularly critical as it can severely affect the performance of coherent receivers, especially at high carrier frequencies (phase noise $\propto f_c^2$, where f_c is the carrier frequency [3]). Generated by oscillators used for up/down conversion at the transmitter/receiver, PN is a source of signal distortion that leads to degraded performance [3]. Techniques for compensating PN for modulation waveforms like OFDM have been widely investigated in the literature [4], [5]. Such investigations for Zak-OTFS have not been reported. The current work addresses this gap and makes the following new and novel contributions.

- We study the effect of oscillator PN on the performance of Zak-OTFS receiver and propose a novel scheme for PN compensation in Zak-OTFS, using a learning framework.
- Towards this end, we first derive a closed-form expression for the I/O relation for Zak-OTFS in the presence of PN, which has not been reported before. This closed-form characterization is important for the sensitivity analysis of PN and the development of PN compensation algorithms.
- Next, leveraging the PN-induced structure in the I/O relation, we develop a learning-aided PN-compensating receiver in which a lightweight convolutional neural network (CNN) estimates the effective channel, whose output is used by a lightweight deep neural network (DNN) to estimate the PN derivative. This estimate is then used to construct a PN-aware effective channel for minimum mean square error (MMSE) detection.
- Simulation results with Vehicular-A channel model with fractional DDs show that the proposed receiver is able to effectively compensate the PN and achieve robust performance over a wide range of PN conditions.

The letter is organized as follows. In Sec. II, Zak-OTFS system model with PN is presented. In Sec. III, we present the proposed PN-compensating receiver. Simulation results are presented in Sec. IV, followed by conclusions in Sec. V.

II. Zak-OTFS System Model with Phase Noise

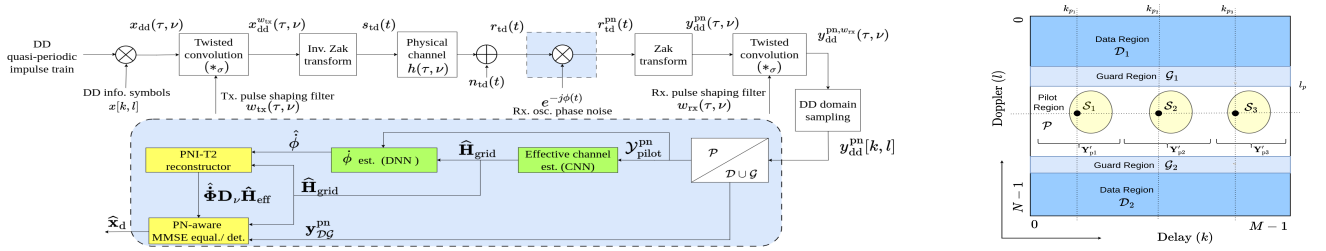
Figure 1 (a) shows the Zak-OTFS system with the proposed PN-compensating receiver. Let $\tau_{\rm p}$ and $\nu_{\rm p}$ denote the delay and Doppler periods satisfying $\tau_{\rm p}\nu_{\rm p} = 1$. Define integers $M = B\tau_{\rm p}$ and $N = T\nu_{\rm p}$, so that one Zak-OTFS frame of duration $T = N\tau_{\rm p}$ and bandwidth $B = M\nu_{\rm p}$ carries MN information symbols $x[k,l] \in \mathbb{A}$ for $k = 0, \ldots, M-1$ and $l = 0, \ldots, N-1$. The symbols are extended quasi-periodically over the DD lattice as $x_{\rm dd}[k+nM, l+mN] = x[k,l]e^{j2\pi n\frac{l}{N}},\ n,m \in \mathbb{Z}$. The resulting discrete DD signal is mapped to the continuous DD domain through impulse modulation, $x_{\rm dd}(\tau,\nu) = \sum_{k,l\in\mathbb{Z}} x_{\rm dd}[k,l]\,\delta\big(\tau - \frac{k\tau_{\rm p}}{M}\big)\delta\big(\nu - \frac{l\nu_{\rm p}}{N}\big)$. To confine the signal within the intended bandwidth B and duration T, a DD-domain transmit shaping filter $w_{\rm tx}(\tau,\nu)$ is applied via twisted convolution to obtain the filtered output as $x_{\rm dd}^{w_{\rm tx}}(\tau,\nu) = w_{\rm tx}(\tau,\nu) *_\sigma x_{\rm dd}(\tau,\nu)$, where $*_\sigma$ denotes twisted convolution operation[1]. The corresponding time-domain transmit signal is obtained using the inverse Zak transform (IZT)[2] as $s_{\rm td}(t) = \mathcal{Z}_t^{-1}\left(x_{\rm dd}^{w_{\rm tx}}(\tau,\nu)\right) = \sqrt{\tau_{\rm p}}\int_0^{\nu_{\rm p}} x_{\rm dd}^{w_{\rm tx}}(t,\nu)\,d\nu$. Assuming a doubly-dispersive channel with DD response function $h(\tau,\nu) = \sum_{i=1}^{P} h_i\,\delta(\tau-\tau_i)\delta(\nu-\nu_i)$, where P is the number of paths and h_i, τ_i, and ν_i denote ith path's

Manuscript received Aug. 12, 2026; revised Sep. 23, 2026; accepted Oct. 07, 2026. This work was supported in part by the J. C. Bose National Fellowship and the NSF-DST project (International Cooperation Division), Department of Science and Technology, Government of India. The associate editor coordinating the review of this letter and approving it for publication was Dr. Zhe Wang (*Corresponding author: Ananthanarayanan Chockalingam.)*

The authors are with the Department of ECE, Indian Institute of Science, Bangalore 560012, India (e-mail: naveedbin@iisc.ac.in, achockal@iisc.ac.in).

[1] $(u *_\sigma v)(\tau,\nu) = \iint u(\tau',\nu')\,v(\tau-\tau',\nu-\nu')\,e^{j2\pi\nu'(\tau-\tau')}d\tau'd\nu'$.

[2] IZT is defined as $\mathcal{Z}_t^{-1}\left(x(\tau,\nu)\right) \triangleq \sqrt{\tau_{\rm p}}\int_0^{\nu_{\rm p}} x(t,\nu)d\nu$. Zak transform (ZT) is defined as $\mathcal{Z}_t(x(t)) = \sqrt{\tau_{\rm p}}\sum_{k\in\mathbb{Z}} x(\tau+k\tau_{\rm p})\,e^{-j2\pi\nu k\tau_{\rm p}}$.



(a) (b)

Fig. 1: (a): Zak-OTFS system with the proposed PN-compensating receiver; (b) Multi-pilot embedded frame.

complex gain, delay, and Doppler, the received time-domain signal is $r_{\rm td}(t) = \iint h(\tau,\nu)\, s_{\rm td}(t-\tau)\, e^{j2\pi\nu(t-\tau)} d\tau d\nu + n_{\rm td}(t)$, where $n_{\rm td}(t)$ denotes additive white Gaussian noise (AWGN). In the presence of receiver oscillator PN, it becomes $r^{\rm pn}_{\rm td}(t) = e^{-j\phi(t)}\iint h(\tau,\nu)\, s_{\rm td}(t-\tau)\, e^{j2\pi\nu(t-\tau)}\, d\tau d\nu + \tilde{n}(t)$, where $\tilde{n}(t) = n_{\rm td}(t)e^{-j\phi(t)}$, whose statistics remain unchanged because $n_{\rm td}(t)$ is circularly symmetric complex Gaussian. The process $\phi(t)$ denotes the receiver oscillator PN, modeled as a random phase whose statistics follow the PN power spectral density in [3]. Using the small-angle approximation, i.e., $\mathbb{E}[\phi^2(t)] \ll 1$, we have $e^{-j\phi(t)} \approx 1 - j\phi(t)$ and

$$r^{\rm pn}_{\rm td}(t) \approx \iint h(\tau,\nu)\, s_{\rm td}(t-\tau)\, e^{j2\pi\nu(t-\tau)}\, d\tau d\nu - j\phi(t)\iint h(\tau,\nu)\, s_{\rm td}(t-\tau)\, e^{j2\pi\nu(t-\tau)}\, d\tau d\nu + \tilde{n}(t). \quad (1)$$

Applying Zak transform to (1), we have

$$y^{\rm pn}_{\rm dd}(\tau',\nu') = \big(h *_\sigma x^{w_{\rm tx}}_{\rm dd}\big)(\tau',\nu') \underbrace{-\left(j\,\mathcal{Z}_t\left\{\phi(t)\iint h(\tau,\nu)\, s_{\rm td}(t-\tau)\, e^{j2\pi\nu(t-\tau)} d\tau d\nu\right\}(\tau',\nu')\right)}_{\triangleq\, T_{\rm pn}(\tau',\nu')} + \tilde{n}_{\rm dd}(\tau',\nu'), \quad (2)$$

$$T_{\rm pn}(\tau',\nu') = j\sqrt{\tau_{\rm p}}\sum_k \phi(\tau'+k\tau_{\rm p})\iint h(\tau,\nu) s_{\rm td}(\tau'+k\tau_{\rm p}-\tau) \times e^{j2\pi\nu(\tau'+k\tau_{\rm p}-\tau)} e^{-j2\pi\nu' k\tau_{\rm p}}\, d\tau d\nu. \quad (3)$$

Applying the first-order Taylor approximation $\phi(\tau'+k\tau_{\rm p}) \approx \phi(\tau') + k\tau_{\rm p}\dot{\phi}(\tau')$ to (3) yields[3]

$$\begin{aligned} T_{\rm pn}(\tau',\nu') \;\approx\; & j\iint h(\tau,\nu)\sqrt{\tau_{\rm p}}\sum_k \big(\phi(\tau') + k\tau_{\rm p}\dot{\phi}(\tau')\big) \\ & \times s_{\rm td}(\tau'+k\tau_{\rm p}-\tau) e^{j2\pi\nu(\tau'+k\tau_{\rm p}-\tau)} e^{-j2\pi\nu' k\tau_{\rm p}}\, d\tau d\nu \\ =\; & j\phi(\tau')\iint h(\tau,\nu) e^{j2\pi\nu(\tau'-\tau)} \\ & \times\sqrt{\tau_{\rm p}}\sum_k s_{\rm td}(\tau'+k\tau_{\rm p}-\tau) e^{-j2\pi(\nu'-\nu)k\tau_{\rm p}}\, d\tau d\nu \\ & + j\tau_{\rm p}\dot{\phi}(\tau')\iint h(\tau,\nu) e^{j2\pi\nu(\tau'-\tau)} \\ & \times\sqrt{\tau_{\rm p}}\sum_k k\, s_{\rm td}(\tau'+k\tau_{\rm p}-\tau) e^{-j2\pi(\nu'-\nu)k\tau_{\rm p}}\, d\tau d\nu. \end{aligned} \quad (4)$$

Using the Zak transform property

$$\sqrt{\tau_{\rm p}}\sum_k k\, x(\tau+k\tau_{\rm p}) e^{-j2\pi\nu k\tau_{\rm p}} = \frac{-1}{j2\pi\tau_{\rm p}}\frac{\partial}{\partial\nu}\mathcal{Z}_t(x(t))(\tau,\nu),$$

obtained by differentiating its definition with respect to ν,

$$T_{\rm pn}(\tau',\nu') = j\phi(\tau')\big(h *_\sigma x^{w_{\rm tx}}_{\rm dd}\big)(\tau',\nu') - \frac{\dot{\phi}(\tau')}{2\pi}\frac{\partial}{\partial\nu'}\big(h *_\sigma x^{w_{\rm tx}}_{\rm dd}\big)(\tau',\nu'). \quad (5)$$

Next, the received DD-domain signal is filtered by the receive DD filter $w_{\rm rx}(\tau,\nu)$ matched to the transmit filter, i.e., $w_{\rm rx}(\tau,\nu) = w^*_{\rm tx}(-\tau,-\nu)e^{j2\pi\tau\nu}$, yielding

$$y^{{\rm pn},w_{\rm rx}}_{\rm dd}(\tau',\nu') = \big(w_{\rm rx} *_\sigma (h *_\sigma x^{w_{\rm tx}}_{\rm dd} + T_{\rm pn} + \tilde{n}_{\rm dd})\big)(\tau',\nu'). \quad (6)$$

Under the assumption that the PN varies slowly over the effective delay support of the receive window[4], we approximate $\phi(\tau'-\tau_1) \approx \phi(\tau')$ and $\dot{\phi}(\tau'-\tau_1) \approx \dot{\phi}(\tau')$, yielding

$$y^{{\rm pn},w_{\rm rx}}_{\rm dd}(\tau',\nu') = (h_{\rm eff} *_\sigma x_{\rm dd})(\tau',\nu') - j\phi(\tau')(h_{\rm eff} *_\sigma x_{\rm dd})(\tau',\nu') + \frac{\dot{\phi}(\tau')}{2\pi}\frac{\partial}{\partial\nu'}(h_{\rm eff} *_\sigma x_{\rm dd})(\tau',\nu') + \tilde{n}^{w_{\rm rx}}_{\rm dd}(\tau',\nu'), \quad (7)$$

where $h_{\rm eff}(\tau,\nu) = w_{\rm rx}(\tau,\nu) *_\sigma h(\tau,\nu) *_\sigma w_{\rm tx}(\tau,\nu)$. Sampling $y^{{\rm pn},w_{\rm rx}}_{\rm dd}(\tau',\nu')$ at $\tau' = \frac{k\tau_{\rm p}}{M}$ and $\nu' = \frac{l\nu_{\rm p}}{N}$ yields the discrete signal $y^{\rm pn}_{\rm dd}[k,l]$. The discretized $(h_{\rm eff} *_\sigma x_{\rm dd})(\tau',\nu')$ is $(h_{\rm eff} *_{\sigma{\rm d}} x_{\rm dd})[k,l]$, where $*_{\sigma{\rm d}}$ denotes discrete twisted convolution defined as $(h *_{\sigma{\rm d}} x)[k,l] = \sum_{k',l'} h[k-k', l-l']x[k',l']e^{j2\pi\frac{k'(l-l')}{MN}}$. Defining $z_{\rm dd}[k,l] \triangleq (h_{\rm eff} *_{\sigma{\rm d}} x_{\rm dd})[k,l]$. For derivatives, we use the central-difference scheme approximation[5] with $\Delta\nu = \nu_{\rm p}/N$ and $\Delta\tau = \tau_{\rm p}/M$, exploiting the inherent periodicity of the Zak grid. Thus,

$$\frac{\partial z_{\rm dd}}{\partial\nu'} \approx \frac{z_{\rm dd}[k,l+1] - z_{\rm dd}[k,l-1]}{2\Delta\nu}, \dot{\phi}[k] \approx \frac{\phi[k+1]-\phi[k-1]}{2\Delta\tau}.$$

The resulting discrete DD-domain I/O relation becomes

$$y^{\rm pn}_{\rm dd}[k,l] \approx z_{\rm dd}[k,l] - j\phi[k]\, z_{\rm dd}[k,l] + \frac{1}{2\pi}\dot{\phi}[k]\frac{z_{\rm dd}[k,l+1]-z_{\rm dd}[k,l-1]}{2\Delta\nu} + \tilde{n}_{\rm dd}[k,l]. \quad (8)$$

In vectorizing the discrete model above, $z_{\rm dd}[k,l] = (h_{\rm eff} *_{\sigma{\rm d}} x_{\rm dd})[k,l]$ corresponds to the matrix form $\mathbf{H}_{\rm eff}\mathbf{x}$. Define $\mathbf{D}_N \in \mathbb{R}^{N\times N}$ as the circulant central-difference matrix implementing $\frac{z_{\rm dd}[k,l+1]-z_{\rm dd}[k,l-1]}{2}$, with first row $\frac{1}{2}[0\ 1\ 0\ \cdots\ 0\ -1]$. Since

[3]Since $s_{\rm td}(t)$ is time-limited to $T = N\tau_{\rm p}$, only $K_{\max} = \lfloor T/\tau_{\rm p}\rfloor = N-1$ shifts contribute to the Zak summation. By Taylor's theorem with the Lagrange form of the remainder, $\phi(\tau'+k\tau_{\rm p}) = \phi(\tau') + k\tau_{\rm p}\dot{\phi}(\tau') + (k\tau_{\rm p})^2\ddot{\phi}(\xi)/2$, where $\xi \in (\tau', \tau'+k\tau_{\rm p})$, and the truncation error is $R_2 = (k\tau_{\rm p})^2\ddot{\phi}(\xi)/2$. Since the phase noise evolves on the time scale $1/B_{\rm PLL}$, where $B_{\rm PLL}$ is the phase-locked loop (PLL) bandwidth, we have $\ddot{\phi}(t) = \mathcal{O}(B^2_{\rm PLL})$, yielding $R_2 = \mathcal{O}((k\tau_{\rm p}B_{\rm PLL})^2)$. Hence, the first-order approximation is accurate for $K_{\max}\tau_{\rm p}B_{\rm PLL} < 2$.

[4]The twisted convolution with $w_{\rm rx}(\tau,\nu)$ involves delay shifts τ_1 restricted to the effective delay support of the receive window, which is on the order of $1/B$. Since the PLL phase evolves over the characteristic time scale $1/B_{\rm PLL}$ and typically $B_{\rm PLL} \ll B$, we have $1/B \ll 1/B_{\rm PLL}$. Consequently, both $\phi(\tau'-\tau_1)$ and $\dot{\phi}(\tau'-\tau_1)$ undergo negligible variation over the receive-window delay support and can therefore be regarded as locally constant.

[5]The second-order centered difference incurs a truncation error $\epsilon_{\rm FD} = \frac{\Delta\nu^2}{6}\frac{\partial^3 z_{\rm dd}}{\partial\nu^3} + \mathcal{O}(\Delta\nu^4)$, resulting in a DD-domain model mismatch $\Delta y_{\rm dd} = \frac{\dot{\phi}}{2\pi}\epsilon_{\rm FD}$ attenuated by the small PN slope. The normalized RMSE (NRMSE) between the PNI-T2 terms computed via 2nd- and 4th-order approximations for $N = 18$ is 6.5×10^{-4}, and their resulting BER performances are virtually indistinguishable (see Fig. 3(a)). The NRMSE remains small at 9.3×10^{-4} for $N = 9$ ($K_{\max}\tau_{\rm p}B_{\rm PLL} = 0.4 < 2$) and 3.1×10^{-4} for $N = 36$ ($K_{\max}\tau_{\rm p}B_{\rm PLL} = 1.75 < 2$), demonstrating the generality of the approximation across different Doppler grid sizes.

Algorithm 1 Proposed PN-compensating receiver

Require: Received DD-domain signal $\mathbf{y}$, pilot symbols $\mathbf{x}_p$
Ensure: Estimated data symbols $\widehat{\mathbf{x}}_{\rm d}$
1: Extract the pilot responses from $\mathbf{y}$.
2: Estimate eff. channel coefficients $\hat{h}_{\rm eff}[k,l]$ using CNN.
3: Construct the estimated eff. channel matrix $\hat{\mathbf{H}}_{\rm eff}$.
4: Estimate the PN derivative $\hat{\dot{\phi}}[k]$ using DNN.
5: Construct the PN-aware effective channel as: $\tilde{\mathbf{H}}_{\rm eff} = \hat{\mathbf{H}}_{\rm eff} + \alpha\hat{\dot{\mathbf{\Phi}}}\mathbf{D}_\nu\hat{\mathbf{H}}_{\rm eff}$.
6: Perform PN-aware MMSE equal. using $\bar{\mathbf{H}}_{\rm eff}$ to obtain $\widehat{\mathbf{x}}_{\rm d}$.

the derivative operates along the Doppler index l, the corresponding operator is given by $\mathbf{D}_\nu = \mathbf{I}_M \otimes \mathbf{D}_N \in \mathbb{R}^{MN\times MN}$, where $\otimes$ denotes the Kronecker product, and the vectorized discrete model becomes

$$\mathbf{y} \approx \Big(\mathbf{H}_{\rm eff} - \underbrace{j\mathbf{\Phi}\mathbf{H}_{\rm eff}}_{\text{PNI-T1}} + \underbrace{\alpha\dot{\mathbf{\Phi}}\mathbf{D}_\nu\mathbf{H}_{\rm eff}}_{\text{PNI-T2}}\Big)\mathbf{x} + \tilde{\mathbf{n}}, \tag{9}$$

where $\alpha \triangleq N/(2\pi\nu_{\rm p})$. Here, PNI-T1 and PNI-T2 denote PN-induced terms 1 and 2, respectively, $\mathbf{x}, \mathbf{y}, \tilde{\mathbf{n}} \in \mathbb{C}^{MN\times 1}$ are formed by stacking $x_{\rm dd}[k,l]$, $y^{\rm pn}_{\rm dd}[k,l]$, and $\tilde{n}_{\rm dd}[k,l]$, respectively, such that $x_{kN+l+1} = x_{\rm dd}[k,l]$. The effective channel matrix $\mathbf{H}_{\rm eff} \in \mathbb{C}^{MN\times MN}$ captures the DD-domain channel interaction and is constructed using $h_{\rm eff}[k,l]$s as [2]

$$\mathbf{H}_{\rm eff}[k'N+l'+1, kN+l+1] = \sum_{m,n\in\mathbb{Z}} h_{\rm eff}[k'-k-nM, l'-l-mN]e^{j2\pi\frac{nl}{N}}e^{j2\pi\frac{(l'-l-mN)(k+nM)}{MN}}, \tag{10}$$

where $k', k = 0,\ldots,M-1$, $l', l = 0,\ldots,N-1$, $\mathbf{\Phi} = \mathrm{diag}(\phi[0],\ldots,\phi[M-1])\otimes\mathbf{I}_N$, $\dot{\mathbf{\Phi}} = \mathrm{diag}(\dot{\phi}[0],\ldots,\dot{\phi}[M-1])\otimes\mathbf{I}_N$.

PN model: The oscillator PN is characterized in the frequency domain in terms of the PSD, given by [3]
$L(f_m) = \left[B^2_{\rm PLL}L_0/(B^2_{\rm PLL}+f^2_m)\right](1+f_{\rm corner}/f_m) + L_{\rm floor}$, where $B_{\rm PLL}$ is the -3 dB bandwidth of the PLL, L_0 is the in-band single-side PN level in dBc/Hz, f_m is the frequency offset from the carrier frequency, $f_{\rm corner}$ is the flicker corner frequency, and $L_{\rm floor}$ is the noise floor.

III. Proposed PN-compensating Receiver

Leveraging the structure of the derived Zak-OTFS I/O relation in the presence of PN in (9), here we develop the proposed PN-compensating receiver. Note that the PN introduces two additive terms in the DD domain, PNI-T1 and PNI-T2, which must be estimated and compensated to mitigate the degrading effect of PN. Among the two terms, our simulations have shown that the PNI-T2 term is more detrimental. Therefore, we estimate and compensate the PNI-T2 term. The proposed receiver has a hybrid architecture consisting of processing blocks that are conventional (yellow blocks) in Fig. 1(a) and learning-based (green blocks). The overall receiver processing flow is summarized in Algorithm 1.

Remark on multi-pilot structure in Fig. 1(b): The quantities $\phi[k]$ and $\dot{\phi}[k]$ in PNI-T1 and PNI-T2 terms remain constant across the Doppler index l for each delay index k. This delay-wise structure motivates the multi-pilot design along the delay axis in Fig. 1(b), and this exposes the delay-dependent PN terms and facilitates better channel learning. In the proposed receiver, we use a CNN for channel estimation and a DNN for PN derivative estimation. The frame structure is designed by balancing two requirements. The first requirement is to suppress interference from surrounding data symbols into the pilot region. For example, for a frame with $M = 14, N = 18$, a guard region of one Doppler bin above and one below the pilot region (see Fig. 1(b)) results in 50% overhead. The second requirement concerns the number of pilots and the spacing among them. A large pilot separation is preferred to achieve good CNN performance because of low interference between pilot responses. On the other hand, to achieve good DNN performance, it is preferred that the collective responses of the pilots adequately span the entire pilot region, for which a large number of closely spaced pilots are preferred. Therefore, the number of pilots and spacing have to be chosen balancing this trade-off to achieve the best overall performance.

A. CNN-based effective channel estimation

The pilot response in the embedded frame is affected by data interference, PN, and AWGN. A CNN is employed to obtain enhanced estimates of the effective channel coefficients $h_{\rm eff}[k,l]$ in the presence of these impairments. The training dataset is generated using (9) by extracting the received signal in the pilot region $\mathcal{P}$. For the system configuration in Sec. IV with $M = 14$, we found that 3 pilots at $(k_{p_1}, k_{p_2}, k_{p_3}) = (2, 7, 12)$ offer good overall performance. We describe the estimation algorithms for this pilot configuration, while noting that they are valid for other frame sizes and pilot configurations[6]. Let $\mathbf{Y}'_{\rm p1}, \mathbf{Y}'_{\rm p2}, \mathbf{Y}'_{\rm p3}$ denote the responses extracted from the three pilot subregions $\mathcal{S}_1, \mathcal{S}_2, \mathcal{S}_3$, respectively (see Fig. 1(b)). Each extracted response is embedded at the center of a separate $M \times N$ zero matrix, yielding $\mathbf{Y}_{\rm p1}, \mathbf{Y}_{\rm p2}, \mathbf{Y}_{\rm p3} \in \mathbb{C}^{M\times N}$, which are stacked to form the three-channel tensor $\mathcal{Y}^{\rm pn}_{\rm pilot} = \mathrm{stack}(\mathbf{Y}_{\rm p1}, \mathbf{Y}_{\rm p2}, \mathbf{Y}_{\rm p3})$. For the ith realization, the CNN input set is $\mathbb{X}^{(i)}_{\rm train} = \{\Re(\mathcal{Y}^{{\rm pn},(i)}_{\rm pilot}), \Im(\mathcal{Y}^{{\rm pn},(i)}_{\rm pilot})\}$. The CNN is trained[7] in a supervised manner over multiple independent realizations, where each stacked pilot tensor is paired with the corresponding ground-truth effective channel coefficients $h^{(i)}_{\rm eff}[k,l]$, $k = -\frac{M}{2},\ldots,\frac{M}{2}-1$, $l = -\frac{N}{2},\ldots,\frac{N}{2}-1$. Let $\mathbf{H}^{(i)}_{\rm grid} \in \mathbb{C}^{M\times N}$ denote the matrix representation of $h^{(i)}_{\rm eff}[k,l]$ over the DD grid. The corresponding training-target set is $\mathbb{Y}^{(i)}_{\rm train} = \{\Re(\mathbf{H}^{(i)}_{\rm grid}), \Im(\mathbf{H}^{(i)}_{\rm grid})\}$. The real and imaginary components of $\mathbb{X}^{(i)}_{\rm train}$ are processed separately using the same CNN and mapped to the corresponding components of $\mathbb{Y}^{(i)}_{\rm train}$. Denoting the resulting estimates by $\widehat{\mathbf{H}}^{(i)}_{{\rm grid},\Re}$ and $\widehat{\mathbf{H}}^{(i)}_{{\rm grid},\Im}$, the estimated effective channel is $\widehat{\mathbf{H}}^{(i)}_{\rm grid} = \widehat{\mathbf{H}}^{(i)}_{{\rm grid},\Re} + j\widehat{\mathbf{H}}^{(i)}_{{\rm grid},\Im}$. The entries of $\widehat{\mathbf{H}}_{\rm grid}$ are used to construct $\hat{\mathbf{H}}_{\rm eff}$ as per (10).
CNN (Channel Estimation): *Input → Conv(3×3,12) → ReLU → Conv(3×3,12) → ReLU → Conv(3×3,24) → ReLU → Conv(3×3,24) → ReLU → Conv(1×1,1) → Output.*
The training parameters are summarized in Table I.

B. DNN-based PN derivative estimation

Accurate estimation of the PN derivative $\dot{\phi}[k]$ in PNI-T2 term in (9) is essential for effective PN compensation.

[6]When scaling to larger frame sizes (e.g., $M = 28, N = 36$) where $K_{\max}\tau_{\rm p}B_{\rm PLL} = 1.75 < 2$, additional pilots (e.g., 6) can be incorporated with proper spacing to adequately span the pilot region.

[7]Since the CNN input size depends on the number of pilots, it needs retraining when used for a system with a different number of pilots.

TABLE I: Architectural/training parameters of the learning modules in the proposed PN-compensating receiver.

Parameter	**CNN (Enhanced channel estimation)**	**DNN ($\dot{\phi}$ estimation)**
Input / Output size	$14 \times 18 \times 3 \rightarrow 14 \times 18 \times 1$	$28 \rightarrow 1$
Hidden layers (Conv / Fully-connected (FC))	5 Conv	3 FC
Training / Model	Epochs: 100, Params: 9,493	Epochs: 200, Params: 4,146
MACs / FLOPs	$2.37 \times 10^6 / 4.74 \times 10^6$	$4.29 \times 10^3 / 8.58 \times 10^3$
SNR / Initial LR / Batch size	Training SNR (dB): $\{5, 24, 30\}$, Initial LR: 10^{-3}, Batch size: 512	
Activation: ReLU; Normalization: LayerNorm (LN) for DNNs; Loss: RMSE; Optimizer: Adam; Learning-rate decay: 0.9 every 20 epochs Training/validation split: 85%/15% of the total dataset.		
Training data: Veh-A channel ($P = 6$ paths); PN params: $B_{\text{PLL}} = 2$ kHz, $L_0 = -65$ dBc/Hz, $f_{\text{corner}} = 1$ kHz, $L_{\text{floor}} = -140$ dBc/Hz.		

Towards this end, using the channel estimates from CNN and the known pilot symbols, the received pilot sample $y_{\text{p}}[k,l]$, for $(k,l) \in \mathcal{P}$, can be expressed through the reconstructed pilot response as $y_{\text{p}}[k,l] = \big(\hat{\mathbf{H}}_{\text{eff}}\mathbf{x}_{\text{p}}\big)[k,l] - \alpha\,\dot{\phi}[k]\,\big(\mathbf{D}_\nu\hat{\mathbf{H}}_{\text{eff}}\mathbf{x}_{\text{p}}\big)[k,l] + \eta[k,l]$, where $\mathbf{x}_{\text{p}}$, $\mathbf{D}_\nu$, and $\alpha = \frac{N}{2\pi\nu_{\text{p}}}$ are known quantities corresponding to the pilot vector, the Doppler central-difference operator, and the model constant, respectively, and $\eta[k,l]$ denotes residual interference and noise. Defining the reconstruction error $r[k,l] \triangleq (\hat{\mathbf{H}}_{\text{eff}}\mathbf{x}_{\text{p}})[k,l] - y_{\text{p}}[k,l]$ and $v[k,l] \triangleq \big(\mathbf{D}_\nu\hat{\mathbf{H}}_{\text{eff}}\mathbf{x}_{\text{p}}\big)[k,l]$, the model reduces to $r[k,l] = \alpha\,\dot{\phi}[k]\,v[k,l] + \eta[k,l]$. Let N_{p} denote the number of Doppler bins in the pilot strip. Stacking $r[k,l]$ and $v[k,l]$ over these bins for each delay index k yields $\mathbf{r}_k, \mathbf{v}_k \in \mathbb{C}^{N_{\text{p}}}$, with $\mathbf{r}_k = \alpha\,\dot{\phi}[k]\,\mathbf{v}_k + \boldsymbol{\eta}_k$. A lightweight DNN is proposed to estimate $\dot{\phi}[k]$ using the feature vectors $\mathbf{r}_k$ and $\mathbf{v}_k$. The network is trained in a supervised manner over multiple independent realizations. For each delay index k, $\big[\Re\{\mathbf{r}_k^{(i)}\};\ \Im\{\mathbf{r}_k^{(i)}\};\ \Re\{\mathbf{v}_k^{(i)}\};\ \Im\{\mathbf{v}_k^{(i)}\}\big] \in \mathbb{R}^{4N_{\text{p}}}$ is provided as input to the DNN along with the corresponding ground-truth PN derivative $\dot{\phi}^{(i)}[k]$. The estimated PN derivative $\hat{\dot{\phi}}[k]$ is obtained as the DNN output.

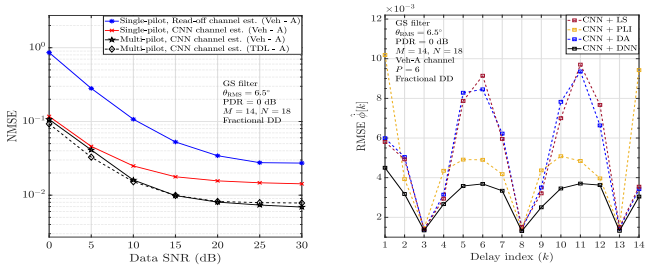
DNN ($\dot{\phi}$ Estimation): *Input $\rightarrow$ 50 $\rightarrow$ ReLU $\rightarrow$ LayerNorm $\rightarrow$ 35 $\rightarrow$ ReLU $\rightarrow$ LayerNorm $\rightarrow$ 20 $\rightarrow$ ReLU $\rightarrow$ 1 $\rightarrow$ Output.* The training parameters are summarized in Table I.

C. PN-aware MMSE data detection

Using the estimated PN derivative $\hat{\dot{\phi}}[k]$, a PN-aware effective channel is constructed by incorporating PNI-T2 into the channel model as $\tilde{\mathbf{H}}_{\text{eff}} = \hat{\mathbf{H}}_{\text{eff}} + \alpha\,\hat{\dot{\boldsymbol{\Phi}}}\,\mathbf{D}_\nu\,\hat{\mathbf{H}}_{\text{eff}}$, where $\hat{\dot{\boldsymbol{\Phi}}} = \text{diag}(\hat{\dot{\phi}}[0], \ldots, \hat{\dot{\phi}}[M-1]) \otimes \mathbf{I}_N$. Although $\tilde{\mathbf{H}}_{\text{eff}} \in \mathbb{C}^{MN \times MN}$, data detection only involves the received samples in the data and guard regions ($\mathcal{D} \cup \mathcal{G}$). Accordingly, we form the reduced effective channel $\bar{\mathbf{H}}_{\text{eff}} \in \mathbb{C}^{|\mathcal{D}\cup\mathcal{G}| \times |\mathcal{D}|}$ by retaining $|\mathcal{D} \cup \mathcal{G}|$ rows of $\tilde{\mathbf{H}}_{\text{eff}}$ (corresponding to the received-sample vector $\mathbf{y}^{\text{pn}}_{\mathcal{D}\mathcal{G}} \in \mathbb{C}^{|\mathcal{D}\cup\mathcal{G}| \times 1}$) and $|\mathcal{D}|$ columns of $\hat{\mathbf{H}}_{\text{eff}}$ (corresponding to the data-symbol vector $\mathbf{x}_{\text{d}} \in \mathbb{C}^{|\mathcal{D}| \times 1}$). Data detection is then performed using an MMSE equalizer based on the reduced effective channel, yielding $\widehat{\mathbf{x}}_{\text{d}} \in \mathbb{C}^{|\mathcal{D}| \times 1}$ as $\widehat{\mathbf{x}}_{\text{d}} = \big(\bar{\mathbf{H}}^H_{\text{eff}}\mathbf{C}^{-1}_{\bar{\mathbf{n}}}\bar{\mathbf{H}}_{\text{eff}} + \mathbf{I}_{|\mathcal{D}|}\big)^{-1}\bar{\mathbf{H}}^H_{\text{eff}}\mathbf{C}^{-1}_{\bar{\mathbf{n}}}\mathbf{y}^{\text{pn}}_{\mathcal{D}\mathcal{G}}$, where $\bar{\mathbf{n}} \in \mathbb{C}^{|\mathcal{D}\cup\mathcal{G}| \times 1}$ denotes the noise corresponding to the received samples in $\mathcal{D} \cup \mathcal{G}$, with covariance $\mathbf{C}_{\bar{\mathbf{n}}} = \mathbb{E}[\bar{\mathbf{n}}\bar{\mathbf{n}}^H]$. The MMSE equalizer requires the inversion of a $|\mathcal{D}| \times |\mathcal{D}|$ matrix and therefore has a computational complexity of $\mathcal{O}(|\mathcal{D}|^3)$.

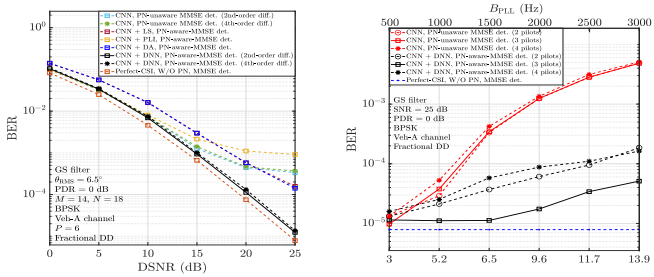
IV. Results and Discussion

In this section, we present the performance of the proposed receiver. The following system parameters are considered:



(a) (b)

Fig. 2: (a) NMSE of effective channel estimation vs data SNR. (b) RMSE of PN derivative estimate $\hat{\dot{\phi}}[k]$ vs delay index k.



(a) (b)

Fig. 3: (a) BER vs DSNR. (b) BER vs θ_{RMS} performance.

$\nu_{\text{p}} = 30$ kHz, $\tau_{\text{p}} = 1/\nu_{\text{p}} \approx 33.33\,\mu$s, $M = 14$ and $N = 18$, frame duration $T = N\tau_{\text{p}} \approx 0.6$ ms and bandwidth $B = M\nu_{\text{p}} = 420$ kHz. Vehicular-A channel with $P = 6$ fractional DD paths with maximum Doppler $\nu_{\max} = 1000$ Hz and maximum delay $\tau_{\max} = 2.51\,\mu$s, and power delay profile as given in Table 2 in [2] is considered. The Doppler shift of each path follows $\nu_i = \nu_{\max}\cos(\theta_i)$ with $\theta_i \sim \mathcal{U}[0, 2\pi]$. Gaussian-sinc (GS) pulse-shaping filter which offers the orthogonality and localization advantages of sinc and Gaussian pulses is considered [6]. The PN parameters are $B_{\text{PLL}} = 1.5$ kHz, $L_0 = -67$ dBc/Hz, $f_{\text{corner}} = 1$ kHz, and $L_{\text{floor}} = -140$ dBc/Hz, corresponding to an RMS phase metric of $\theta_{\text{RMS}} \approx 6.5^\circ$, where $\theta_{\text{RMS}} \triangleq 1/\sqrt{\text{SNR}_{\text{PN}}}$ with SNR_{PN} denoting the ratio of the desired signal power to the PN-induced interference power [3]. For these parameters, the generated oscillator phase process yields $\mathbb{E}[\phi^2(t)] = 4.24 \times 10^{-4} \ll 1$ and the 2nd-, 3rd-, and 4th-order terms contribute only 1.26%, 0.0128%, and 0.00011%, respectively, which validate the first-order approximation $e^{-j\phi(t)} \approx 1 - j\phi(t)$. Also, $K_{\max}\tau_{\text{p}}B_{\text{PLL}} = 0.85 < 2$, which validates the slow phase-variation approximation.

Figure 2(a) shows the normalized mean squared error (NMSE), defined as $\mathbb{E}[(\|\mathbf{H}_{\text{grid}} - \widehat{\mathbf{H}}_{\text{grid}}\|_F^2)/(\|\mathbf{H}_{\text{grid}}\|_F^2)]$, for effective channel estimation in the presence of PN as a function of data SNR (DSNR) at a pilot-to-data power ratio

(PDR) of 0 dB. NMSE plots for the cases of single-pilot (in the middle of the frame) and multi-pilot (3 pilots as in Fig. 1(b)) are shown. It can be seen that while the NMSE performance of the simple read-off based channel estimation in [2] is significantly compromised in the presence of PN, the proposed CNN-based channel estimation achieves significantly improved estimation accuracy. Between single-pilot and multi-pilot (keeping the total pilot power to be the same), multi-pilot offers superior performance due to better learning from multiple measurements from multiple pilots placed along the delay-strip matching the delay-wise structure of the PN process. We can also see that the Veh-A-trained CNN achieves nearly identical NMSE performance when tested on the TDL-A channel with $P = 23$ paths [7] without any retraining/fine-tuning, demonstrating good generalization across different channel models. Our simulations with this channel model mismatch also showed similar generalizability of the DNN.

In Fig. 2(b), we compare the RMSE performance of the following four PN derivative estimation methods. (i) Proposed DNN-based estimation scheme described in Sec. III-B (denoted by CNN+DNN). (ii) LS-based estimation scheme (denoted by CNN+LS). In CNN+LS, the LS estimate of $\dot{\phi}[k]$ is obtained as $\hat{\dot{\phi}}_{\rm LS}[k] = \mathbf{v}_k^H \mathbf{r}_k / (\alpha \mathbf{v}_k^H \mathbf{v}_k)$ at all the delay indices. (iii) An interpolation based scheme motivated by a pilot-aided interpolation (PLI) based PN estimation/compensation scheme in [5] (denoted by CNN+PLI). In CNN+PLI, estimates of the PN derivative are obtained as per the LS estimate only at the pilot locations and the estimates for other delay indices in between are obtained through linear interpolation as $\hat{\dot{\phi}}[k] = \text{interp}\big(\{k_{p_1}, k_{p_2}, k_{p_3}\}, \{\hat{\dot{\phi}}_{\rm LS}[k_{p_1}], \hat{\dot{\phi}}_{\rm LS}[k_{p_2}], \hat{\dot{\phi}}_{\rm LS}[k_{p_3}]\}, k\big)$. (iv) A scheme motivated by data decision-aided (DA) PN estimation/compensation in [4] (denoted by CNN+DA). In CNN+DA, estimates of the PN derivative at all delay indices are obtained as per the LS estimate which are re-estimated after removal of PN derivative reconstructed using detected data symbols. That is, the reconstructed received signal using the initial LS estimate $\hat{\dot{\mathbf{\Phi}}}^{(0)}$ and detected symbols $\hat{\mathbf{x}}^{(0)}$ is obtained as $\hat{\mathbf{y}} = \big(\hat{\mathbf{H}}_{\rm eff} + \frac{1}{2\pi}\frac{N}{\nu_p}\hat{\dot{\mathbf{\Phi}}}^{(0)}\mathbf{D}_\nu \hat{\mathbf{H}}_{\rm eff}\big)\hat{\mathbf{x}}^{(0)}$, and the PN derivative is estimated again using the signal $\mathbf{y} - \hat{\mathbf{y}}$. In Fig. 2(b), we see that the CNN+LS scheme exhibits pronounced RMSE peaks at delay indices far from the pilots (see Fig. 1(b)) where the estimates are influenced mainly by noise. The CNN+DA scheme exhibits similar RMSE performance indicating that the improvement due to data decision-aided PN derivative removal is not significant. This is attributed to the reduced data leakage because of guard/pilot regions and low side lobes in the GS pulse. The interpolation in the CNN+PLI scheme provides better estimates compared to CNN+LS and CNN+DA schemes due to reduced influence of noise. However, it suffers from larger estimation errors at the edge bins[8]. In contrast, the proposed CNN+DNN scheme achieves the lowest RMSE across all delay indices due to the DNN's effective learning[9].

In Fig. 3(a), we present the BER performance of the receivers using the four PN derivative estimation schemes considered in Fig. 2(b). The PN-unaware MMSE receiver (employing CNN-based effective channel estimation without PN compensation) exhibits a BER floor at higher DSNRs due to the uncompensated PN impairment. Among the benchmark receivers, CNN+LS and CNN+DA achieve nearly identical BER performance, consistent with their similar PN derivative estimation accuracies observed in Fig. 2(b). The CNN+PLI receiver provides a slight BER improvement over CNN+LS and CNN+DA at low DSNRs. However, at higher DSNRs, the increased estimation errors at the edge bins become dominant, resulting in BER that is worse than that of the PN-unaware receiver. In contrast, the proposed CNN+DNN receiver achieves the lowest BER, approaching the performance with perfect CSI and no PN. Figure 3(b) shows the robustness of the proposed receiver under varying oscillator PN levels, characterized by $\theta_{\rm RMS}$, for 25 dB DSNR and 0 dB PDR, comparing the 3-pilot configuration against 2-pilot ($k_p \in \{4, 10\}$) and 4-pilot ($k_p \in \{2, 5, 9, 12\}$) configurations. $B_{\rm PLL}$ is varied from 500 Hz to 3000 Hz, giving $\theta_{\rm RMS}$ from 3° to 13.9°. For $\theta_{\rm RMS} = 13.9^\circ$, $\mathbb{E}[\phi^2(t)] = 9.97 \times 10^{-4} \ll 1$, where the 2nd-, 3rd-, and 4th-order terms contribute only 1.95%, 0.031%, and 0.00043%, respectively, and $K_{\max}\tau_{\rm p}B_{\rm PLL} = 1.7 < 2$. The proposed receiver outperforms PN-unaware baselines, and the 3-pilot configuration balances pilot span and interference well, approaching closer to the no-PN, perfect-CSI performance. Similar trend was observed for increased guard region size.

V. Conclusions

We derived a closed-form expression for Zak-OTFS I/O relation in the presence of PN, using the structure of which we devised a PN-compensating receiver using CNN-based channel estimation, DNN-based PN derivative estimation, and PN-aware MMSE detection. The proposed receiver offered improved BER performance compared to PN-unaware receiver. Future works can consider other hardware impairments such as IQ imbalance and carrier frequency offset.

References

[1] S. K. Mohammed, R. Hadani, and A. Chockalingam, *OTFS Modulation – Theory and Applications*, IEEE Press-Wiley, 2024.

[2] S. K. Mohammed, R. Hadani, A. Chockalingam, and R. Calderbank, "OTFS – predictability in the delay-Doppler domain and its value to communication and radar sensing," *IEEE BITS the Inform. Theory Mag.*, vol. 3, no. 2, pp. 7-31, Jun. 2023.

[3] L. Smaini, *RF Analog Impairments Modeling for Communication Systems Simulation: Application to OFDM-Based Transceivers*, Wiley, 2012.

[4] D. Petrovic, W. Rave, and G. Fettweis, "Effects of phase noise on OFDM systems with and without PLL: Characterization and compensation," *IEEE Trans. on Commun.*, vol. 55, no. 8, pp. 1607-1616, Aug. 2007.

[5] Q. Zou, A. Tarighat, and A. H. Sayed, "Compensation of phase noise in OFDM wireless systems," *IEEE Trans. Signal Process.*, vol. 55, no. 11, pp. 5407-5424, Nov. 2007.

[6] A. Das, F. Jesbin, and A. Chockalingam, "A Gaussian-sinc pulse shaping filter for Zak-OTFS," *IEEE Trans. Veh. Tech.*, vol. 75, no. 3, pp. 4307-4322, Mar. 2026.

[7] 3GPP TR 38.901, "Study on channel model for frequencies from 0.5 to 100 GHz," version 16.1.0, Release 16, 2020-11.

[8] Since $\dot{\phi}[k]$ exhibits slow phase-variation and is estimated from the guard-protected pilot region only, higher-order PLI or additional DA iterations provide no noticeable performance gain. The corresponding $\dot{\phi}[k]$ estimation costs are 1.16×10^3, 2.86×10^2, 2.72×10^7, and 8.58×10^3 FLOPs/frame for LS, PLI, DA (single iteration), and the proposed DNN, respectively, showing that the DNN provides the lowest RMSE with some added complexity.

[9] Our simulations show that the DNN RMSE increases with the CNN NMSE; e.g., for $\theta_{\rm RMS} = 3^\circ, 13.9^\circ$, CNN NMSEs of $0.0089, 0.013$ yield DNN RMSEs of $0.0026, 0.0039$, respectively, as expected.