

Channel-Adaptive Autoencoder Architecture for Terahertz Aerospace Communications

Yicheng Huang¹, Tianqi Mao¹, *Member, IEEE*,
Dongxuan He¹, *Member, IEEE*, Qi Wang²,
Chong Han³, *Senior Member, IEEE*,
and Zhaocheng Wang¹, *Fellow, IEEE*

Abstract—The terahertz (THz) aerospace communication has garnered increasing attention due to its potential to meet the ultra-broadband transmission demand of space-air backbone networks. However, several practical issues have to be considered, which include the complex channel characteristics, antenna misalignment, hardware imperfections, etc. To tackle these issues, a dynamic autoencoder system based on channel prediction is proposed to enhance its robustness against complicated dynamic channel fading and distortions in aerospace THz communication systems. Specially, a new architecture with a set of autoencoders and fitting networks is proposed for resilient transmission against the strong Doppler shifts and hardware imperfections. Besides, a single point channel prediction model is utilized to forecast the variations of the channel and then subsequently select the suitable autoencoder. Numerical results demonstrate the superiority of the proposed dynamic autoencoder architecture under an intricate dynamic THz channel.

Index Terms—Aerospace communication, antenna misalignment, autoencoder, high mobility, terahertz communication.

I. INTRODUCTION

For future 6G communication, aerospace communication networks, constituted by low-altitude drones, high-altitude jet aircraft, stratospheric airships and satellite platforms, are anticipated to deliver broadband ubiquitous coverage to billions of user terminals, necessitating over gigabits-per-second (Gb/s) transmission within the space-air backbone network [1]. However, due to the scarcity of spectrum resources, it is challenging to enable such high-speed transmission over the traditional microwave frequencies. Therefore, terahertz (THz) band (0.1 to 10 THz) is explored to be utilized, which is capable

of achieving terabits-per-second (Tb/s) transmission with ultra-wide available bandwidth. Moreover, unlike terrestrial scenarios, suffering from severe molecular absorption loss and obstructions, aerial scenarios with sparse atmosphere can be well-suited for the line-of-sight (LoS) long-range THz communications. Additionally, the compact size of THz antennas can alleviate the payload budget of the light-weight aerospace platforms.

Nevertheless, the practical implementation of aerospace THz communication encounters several technical challenges. First, THz signals may experience non-uniform molecular absorption at different altitudes [2]. Besides, to combat the severe path loss, the aerospace THz communications can be highly directional, which poses significant challenges in beam alignment, especially for high-mobility transceivers like aircraft and satellites [3]. Furthermore, the rapid motion of aerospace vehicles results in considerable Doppler shift, hence adversely affecting channel prediction and estimation. Finally, THz communication systems suffer from inevitable hardware imperfections including in-phase/quadrature (I/Q) imbalance, phase noise, and the nonlinearity of power amplifiers (PA) [4], which pose difficulties to signal reconstruction.

To address these issues, resilient THz signal processing techniques against the sophisticated channel and front-end impairments can be necessary [5], [6]. Specifically, a single-carrier THz communication system utilizing the frequency-multiplier-last architecture was investigated in [5] under significant transmitter imperfections. Additionally, a channel estimation and tracking scheme [6] was introduced to solve the performance deterioration problem caused by the unique triple delay-beam-Doppler squint effects of aeronautical THz ultra massive MIMO channels. Nevertheless, these traditional methods failed to consider the complex distortion and fading issues in a comprehensive manner especially under aerospace THz communication scenarios. Moreover, existing schemes tend to design different transceiver modules separately, which causes high computational complexity and possible local optima.

To this end, artificial intelligence (AI) techniques, such as autoencoder [7], generative adversarial network (GAN) [8] and reinforcement learning (RL) [9], have been adopted in THz communication systems for performance enhancement. However, due to the limited computational resources of the aircraft, the high computational complexity of GAN and RL makes it difficult to support high-speed real-time data transmission. In contrast, the autoencoder has a simpler model, enabling efficient data processing. Moreover, the architecture of the autoencoder is more flexible, allowing easy integration with other modules to extend its functionality. For instance, [7] proposed an enhanced autoencoder by employing a fitting network to approximate the properties of the static THz channel. However, straightforward application of existing autoencoder algorithms to aerospace THz communications can be rather challenging due to the necessity and difficulties in adaptive training of the autoencoder and the fitting network for gradient feedback, under drastically time-variant channel in high-mobility scenarios.

Against this background, this paper presents a dynamic autoencoder architecture with strong robustness to the complicated time-varying channel fading and hardware imperfections in aerospace THz communications. Specially, a bank structure of autoencoders plus fitting networks is proposed for resilient transmission against the strong Doppler shifts and hardware imperfections, such as I/Q imbalance, PA nonlinearity and phase noise. Furthermore, a single point channel prediction model (SPCPM) [10] is utilized to forecast variations in channel

Received 23 January 2025; revised 28 June 2025; accepted 27 July 2025. Date of publication 22 August 2025; date of current version 13 February 2026. This work was supported in part by the National Natural Science Foundation of China under Grant 62401054 and Grant 62088101, in part by the Open Project Program of State Key Laboratory of CNS/ATM under Grant 2024B12, and in part by Young Elite Scientists Sponsorship Program by China Association for Science and Technology under Grant 2022QNRC001. The review of this article was coordinated by Prof. Jiguang He. (*Corresponding authors: Tianqi Mao; Dongxuan He.*)

Yicheng Huang is with the State Key Laboratory of CNS/ATM, Beijing Institute of Technology, Beijing 100081, China (e-mail: huangyc2024@bit.edu.cn).

Tianqi Mao is with the State Key Laboratory of Environment Characteristics and Effects for Near-space, Beijing Institute of Technology, Beijing 100081, China (e-mail: maotq@bit.edu.cn).

Dongxuan He is with the School of Information and Electronics, Beijing Institute of Technology, Beijing 100081, China (e-mail: dongxuan_he@bit.edu.cn).

Qi Wang is with Huawei Device Company Ltd., Shenzhen 518129, China (e-mail: steven.wangqi@huawei.com).

Chong Han is with Terahertz Wireless Communications Laboratory, Shanghai Jiao Tong University, Shanghai 200240, China, and also with Cooperative Medianet Innovation Center, School of Information Science and Electronic Engineering, Shanghai Jiao Tong University, Shanghai 200240, China (e-mail: chong.han@sjtu.edu.cn).

Zhaocheng Wang is with the Department of Electronic Engineering, Tsinghua University, Beijing 100084, China (e-mail: zewang@tsinghua.edu.cn).

Digital Object Identifier 10.1109/TVT.2025.3601728

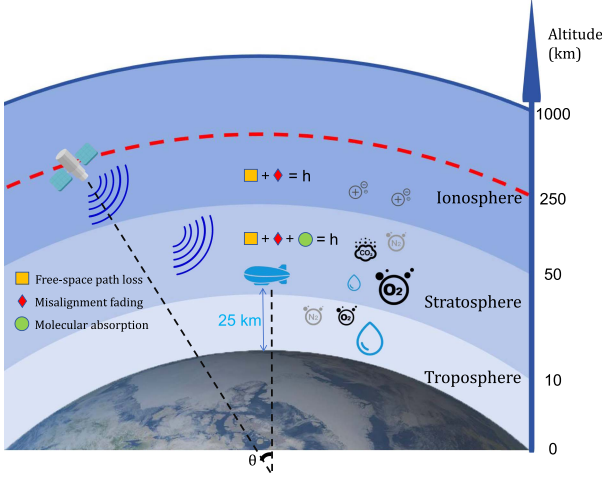


Fig. 1. The considered communication scenario, the wave propagation model and the molecular absorption model from the troposphere to the ionosphere. According to [12], the water vapor dominates the molecular absorption in the troposphere, while oxygen becomes the main factor in the stratosphere.

conditions, allowing real-time awareness of the propagation channel for efficient autoencoder selection. Simulation results demonstrate that the proposed dynamic autoencoder achieves precise channel reconstruction and resilient signal transmission under an intricate dynamic THz channel.

II. SYSTEM MODEL

A. Time-Varying Channel Model

The communication scenario is shown in Fig. 1, where a stratospheric airship is stationed at the altitude of 25 km, while a very low earth orbit (VLEO) satellite is orbiting the Earth at an altitude of 250 km. For simplicity, it is assumed that the airship is quasi-stationary with a fixed position, with antennas prone to mild jitter, and the satellite maintains a circular orbit without trajectory changes. Additionally, the line-of-sight (LoS) path is mainly considered due to the highly directional (quasi-optical) THz signal propagation [11].

According to [2], the attenuation effect of THz wave is caused by free-space path loss h_f as well as molecular absorption loss h_a in the stratosphere (ranging from 10 to 50 km in altitude). Moreover, the THz transmission is expected to be highly directional to combat the severe path loss, which is extremely fragile to wind-driven vibration of the airships, leading to beam misalignment fading h_m . Therefore, the THz channel path loss h can be expressed as

$$h = h_f h_m h_a \quad (1)$$

1) *Free-Space Path Loss*: The signal experiences free-space path loss h_f during propagation, affected by the distance d between the satellite and the airship, the signal frequency f , and the speed of light $c = 3 \times 10^8$ m/s. The path loss can be expressed as the well-known Friis equation.

2) *Misalignment Fading*: The misalignment fading h_m can be expressed as

$$h_m = \sqrt{G_T(N_T, \theta_T, \phi_T) G_R(N_R, \theta_R, \phi_R)}, \quad (2)$$

where the subscript T indicates the transmitting antenna, and R indicates the receiving antenna. Let $S \in \{T, R\}$, G_S denotes the power fading coefficient, which is influenced by the antenna array size N_S , the roll ϕ_S , and the composite angle of yaw and pitch θ_S [3].

3) *Molecular Absorption Loss*: As stated in [12], molecular absorption loss is around 10^{-3} dB/km at an altitude of 20 km, which decreases by almost two orders of magnitude every 10 kilometers as the altitude increases.

4) *Doppler Shift*: Satellites in orbits around the Earth can attain velocities of many kilometers per second, which makes Doppler frequency shift between aircraft substantial. Ignoring multipath effects, a narrow THz channel impulse response can be represented as

$$H = h e^{j\psi}, \quad (3)$$

where $\psi = 2\pi\Delta f \cdot t$ denotes the Doppler phase shift, and $\Delta f = \frac{v}{c} f \cos(\alpha)$ indicates the Doppler frequency shift. Here v and α denote the velocity of the satellite and the angle between the direction of the satellite-airship line and that of the satellite's motion, respectively.

B. Mixed Distortion

As shown in Fig. 2, the communication system encounters severe mixed distortion, including I/Q imbalance, nonlinearity of the THz PA, phase noise, and distortion caused by the time-varying channel [4].

At the transmitter, due to the non-ideal nature of the quadrature modulator, the transmitted signal $\tilde{x} = \tilde{x}_I + j\tilde{x}_Q$ suffered from I/Q imbalance can be expressed as

$$s = (\cos(\varphi_T) - j\epsilon_T \sin(\varphi_T)) \tilde{x} + (\epsilon_T \cos(\varphi_T) - j \sin(\varphi_T)) \tilde{x}^*, \quad (4)$$

where ϵ_T and φ_T are the amplitude and phase imbalance parameters at the transmitter, respectively. \tilde{x}^* denotes the complex conjugate of \tilde{x} . Additionally, signal is influenced by phase noise caused by oscillators, electromagnetic interference, and clock jitter, given by $\tilde{s} = s e^{j\delta_T}$, where δ_T represents the phase noise at the transmitter, which follows an independent and identically distributed (i.i.d.) Gaussian distribution. Besides, the THz PA saturates at high power levels, resulting in signal distortion, which is characterized by an odd-order memoryless polynomial model, given by

$$\tilde{s} = \sum_{k=1}^K \alpha_{2k-1} \tilde{s} | \tilde{s} |^{2k-2}, \quad (5)$$

where the order of nonlinearity is $2K - 1$, and α_{2k-1} are the complex parameters.

The signal arriving at the receiver is affected by the time-varying channel, expressed as $r = H\tilde{s} + w$, where H represents the channel impulse response at this instant, given in (3), and w represents the additive white Gaussian noise (AWGN) with variance σ_w^2 .

At the receiver, the quadrature demodulator imperfections also induce non-negligible I/Q imbalance, given by

$$\tilde{r} = (\cos(\varphi_R) - \epsilon_R \sin(\varphi_R)) r + (\epsilon_R \cos(\varphi_R) - j \sin(\varphi_R)) r^*, \quad (6)$$

where ϵ_R and φ_R are the amplitude and phase imbalance parameters at the receiver, respectively. Furthermore, phase noise at the receiver also affects the signal. Thus, the received signal is given by $\tilde{y} = \tilde{r} e^{j\delta_R}$, where δ_R denotes the phase noise at the receiver, following an i.i.d. Gaussian distribution. According to (1)-(6), the overall expression of the received signal is provided in (7) shown at the bottom of the next page.

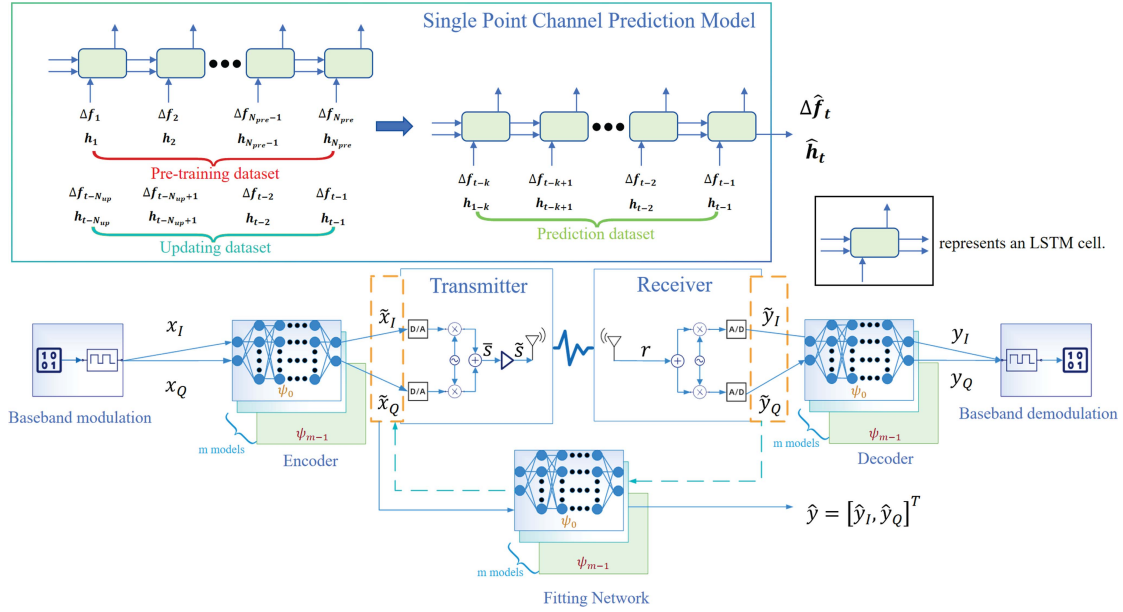


Fig. 2. Diagram of the proposed autoencoder architecture with fitting networks and SPCPM. There are m autoencoders and m fitting networks in the structure. More specifically, the slightly dark blue solid rightward arrow indicates the inference (forward) link for data transmission from the transmitter to the receiver. The cyan dotted leftward arrow represents the backward link, which is used for gradient backpropagation during the training phase of the autoencoder.

III. PROPOSED AUTOENCODER ARCHITECTURE

In this section, the autoencoder philosophy is invoked for performance enhancement of aerospace THz communications, where a bank structure of the autoencoders plus fitting networks are incorporated to accommodate dynamic propagation channel. Besides, an SPCPM network supported by long short-term memory (LSTM) model [13] is employed for real-time prediction of the channel coefficients and Doppler frequency shifts using historical data.

A. Bank Structure of Autoencoders Plus Fitting Networks

In the considered THz communication system, the complex distortions originated from the propagation channel and the RF front-ends hinder the backpropagation of gradient information for end-to-end training. To address this issue, a fitting network can be employed to replicate the sophisticated distortion effects. Empirically speaking, one single autoencoder trained offline cannot adapt to the drastic variations in Doppler phase shifts under high mobility, while online training suffers from slower convergence with limited computational resources. Therefore, we invoke a bank of autoencoders with fitting networks to accommodate different levels of Doppler phase shifts.

As shown in Fig. 2, totally m pairs of the autoencoders and fitting networks are tailored in parallel to m sub-intervals $\{\psi_i = [-\frac{\pi+2i\pi}{m}, \frac{\pi+2i\pi}{m}] \mid i = 0, 1, \dots, m-1\}$ of Doppler phase shift values. The midpoint set of the m intervals is $\mathbf{C} = \{C_i = \frac{2i\pi}{m} \mid i = 0, 1, \dots, m-1\}$. Each network pair is designated by a phase interval $\psi_i (i = 0, 1, \dots, m-1)$, indicating its capability to manage a channel exhibiting a Doppler phase shift $\psi \in \psi_i$. We train the i -th pair networks with the channel exhibiting a Doppler phase shift of C_i . Note that the design principles and training methodologies are identical for different network pairs. Hence, below we merely focus on one single autoencoder with its corresponding fitting network.

Fig. 2 illustrates that all three networks in one pair utilize feedforward neural network (FNN) structures, where the signal flows unidirectionally from the input layer to the output layer. Let e , d and f represent the encoder, decoder, and fitting network, respectively. Define $n \in \{e, d, f\}$, let A_n denote the total number of layers for the network n , including an input layer, an output layer, and several hidden layers. The output vector of the j -th layer of network n can be expressed as

$$\mathbf{o}_{n,j} = f(\mathbf{W}_{n,j}\mathbf{o}_{n,j-1} + \mathbf{b}_{n,j}), \quad (8)$$

where $\mathbf{W}_{n,j}$ and $\mathbf{b}_{n,j}$ represent the weight matrix and bias vector of the j -th layer, respectively. The activation function $f(\cdot)$ is chosen as the

$$\begin{aligned} \tilde{y} = e^{j\delta_R} \left\{ \left(\cos(\varphi_R) - j\epsilon_R \sin(\varphi_R) \right) \left\{ h_f h_m h_a e^{j2\pi \frac{v}{c} f \cos(\alpha) t} \sum_{k=1}^K \alpha_{2k-1} \left\{ e^{j\delta_T} [(\cos(\varphi_T) - j\epsilon_T \sin(\varphi_T))\tilde{x} + (\epsilon_T \cos(\varphi_T) - j\sin(\varphi_T))\tilde{x}^*] \right\} \right. \right. \\ \left. \left. \left| e^{j\delta_T} [(\cos(\varphi_T) - j\epsilon_T \sin(\varphi_T))\tilde{x} + (\epsilon_T \cos(\varphi_T) - j\sin(\varphi_T))\tilde{x}^*] \right|^{2k-2} + w \right\} \right. \\ \left. + \left(\epsilon_R \cos(\varphi_R) - j\sin(\varphi_R) \right) \left\{ h_f h_m h_a e^{j2\pi \frac{v}{c} f \cos(\alpha) t} \sum_{k=1}^K \alpha_{2k-1} \left\{ e^{j\delta_T} [(\cos(\varphi_T) - j\epsilon_T \sin(\varphi_T))\tilde{x} + (\epsilon_T \cos(\varphi_T) - j\sin(\varphi_T))\tilde{x}^*] \right\} \right. \right. \\ \left. \left. \left| e^{j\delta_T} [(\cos(\varphi_T) - j\epsilon_T \sin(\varphi_T))\tilde{x} + (\epsilon_T \cos(\varphi_T) - j\sin(\varphi_T))\tilde{x}^*] \right|^{2k-2} + w \right\}^* \right\}. \end{aligned} \quad (7)$$

hardtanh function [14], which constrains the network's output within the range $[-1, 1]$. This function can not only bound the signal output power but also simplify the gradient computation.

The overall training process of the proposed autoencoder architecture can be divided into two stages:

1) *Fitting Network Training*: As shown in Fig. 2, we use the encoder output $\tilde{\mathbf{x}} = [\tilde{\mathbf{x}}_I, \tilde{\mathbf{x}}_Q]^T$ as the training data and the decoder input $\tilde{\mathbf{y}} = [\tilde{\mathbf{y}}_I, \tilde{\mathbf{y}}_Q]^T$ as the target data. $\tilde{\mathbf{x}}$ is simultaneously input into the transmitter and the fitting network. To align the fitting network output $\hat{\mathbf{y}} = [\hat{\mathbf{y}}_I, \hat{\mathbf{y}}_Q]^T$ with the received distorted signal $\tilde{\mathbf{y}}$ corresponding to $\tilde{\mathbf{x}}$, the loss function for the fitting network is characterized as the mean squared error (MSE) between $\tilde{\mathbf{y}}$ and $\hat{\mathbf{y}}$, given by

$$L_f = \mathbf{E}_{\tilde{\mathbf{x}}}(|\tilde{\mathbf{y}} - \hat{\mathbf{y}}|^2). \quad (9)$$

The Adam algorithm [15] is employed for fitting network training. In every batch of training, the values of $\{\mathbf{W}_{f,1}, \mathbf{W}_{f,2}, \dots, \mathbf{W}_{f,A_f}\}$ and $\{\mathbf{b}_{f,1}, \mathbf{b}_{f,2}, \dots, \mathbf{b}_{f,A_f}\}$ are updated iteratively until the loss function falls below a predefined threshold.

2) *Autoencoder Training*: We use the encoder input $\mathbf{x} = [\mathbf{x}_I, \mathbf{x}_Q]^T$ as the training data. \mathbf{x} is fed into encoder, traversing the autoencoder and the actual channel. To synchronize the decoder output $\mathbf{y} = [\mathbf{y}_I, \mathbf{y}_Q]^T$ with the input data \mathbf{x} , the loss function for the autoencoder is characterized as the MSE between \mathbf{y} and \mathbf{x} , given by

$$L_a = \mathbf{E}_{\mathbf{x}}(|\mathbf{y} - \mathbf{x}|^2). \quad (10)$$

The gradient information of the loss function is first backpropagated to the decoder, to update the decoder parameters. Then, the gradient information is backpropagated through the fitting network to encoder with the fitting network remaining fixed, which only serves as a channel for gradient backpropagation. Finally, the gradient information is backpropagated to the encoder, updating its parameters.

The Adam algorithm is also employed for autoencoder training. In every batch of training, the values of $\{\mathbf{W}_{e,1}, \mathbf{W}_{e,2}, \dots, \mathbf{W}_{e,A_e}, \mathbf{W}_{d,1}, \mathbf{W}_{d,2}, \dots, \mathbf{W}_{d,A_d}\}$ and $\{\mathbf{b}_{e,1}, \mathbf{b}_{e,2}, \dots, \mathbf{b}_{e,A_e}, \mathbf{b}_{d,1}, \mathbf{b}_{d,2}, \dots, \mathbf{b}_{d,A_d}\}$ are updated iteratively until the loss function falls below a predefined threshold.

B. Dynamic Channel Prediction

To adapt to the drastic time-variations of the propagation channel, the SPCPM is introduced to predict the channel coefficient and Doppler frequency shift of the next time step. Specifically, Fig. 2 illustrates that SPCPM forecasts the channel amplitude coefficient \hat{h}_t utilizing past data $\{h_1, \dots, h_{t-1}\}$. Simultaneously, it forecasts the Doppler shift Δf_t utilizing past data $\{\Delta f_1, \dots, \Delta f_{t-1}\}$. The execution process of SPCPM is divided into three stages:

- 1) *Pre-training*: The initial N_{pre} historical channel data are used to pre-train the model.
- 2) *Updating the model*: The historical channel data from the $(t - N_{up})$ -th to the $(t - 1)$ -th time steps are used to update the model parameters.
- 3) *Prediction*: The historical channel data from the $(t - k)$ -th to the $(t - 1)$ -th time steps are sent into the model, which has been trained in the previous steps. The predicted channel data of the (t) -th time step is the output.

For subsequent time steps, the updating and prediction processes are repeated. The system computes the anticipated Doppler phase shift $\hat{\psi}_t = \Delta f_t \cdot t$ according to the prediction results and system clock. Then, the system selects the network pair corresponding to the interval $\{\psi_i | \hat{\psi}_t \in \psi_i\}$ and compensates for the channel amplitude coefficient at the receiver.

TABLE I
THE HIDDEN LAYER SIZES OF ENCODER, DECODER AND FITTING NETWORK

Network	Sizes of the hidden layers
Encoder	(16,64,64,16)
Decoder	(16,128,128,16)
Fitting Network	(40,40,40)

C. Complexity Analysis

In this subsection, we analyze the complexity of the proposed architecture and traditional method. All computational complexity analyses are conducted in terms of complex multiplications. For the bank structure of autoencoders plus fitting networks (total of m pairs), assume each pair has L layers, and the number of neurons in the l -th layer is N_l . Then, the training computational complexity per iteration is $\mathcal{O}(m \sum_{l=1}^{L-1} N_l N_{l+1})$, which scales linearly with m . During inference, only one pair is executed each time, resulting in a computational complexity of $\mathcal{O}(\sum_{l=1}^{L-1} N_l N_{l+1})$. The number of parameters (weights and biases) of all m network pairs is $m(\sum_{l=1}^{L-1} N_l N_{l+1} + \sum_{l=2}^L N_l)$, which reflects the storage complexity of the networks. For the SPCPM, assuming the hidden size of the LSTM is N_h , the computational complexity of the pre-training phase is $\mathcal{O}(N_{pre} N_h^2)$. For each time step during deployment, the parameter update incurs a computational complexity of $\mathcal{O}(N_{up} N_h^2)$, while the inference process incurs a complexity of $\mathcal{O}(k N_h^2)$. For traditional methods such as minimum mean square error (MMSE) equalization, if the received signal vector dimension is N_{mmse} , the computational complexity of MMSE equalization scales with $\mathcal{O}(N_{mmse}^3)$ due to matrix inversion.

In summary, since the network is trained in an offline manner, the additional complexity induced by the resolution enhancement of Doppler phase shifts, is totally acceptable, which will not affect the operation of autoencoder during inference process.

IV. SIMULATION RESULTS

This section addresses two facets of simulation verification. Initially, we simulate and validate the efficacy of the fitting network and SPCPM in channel fitting. Then, we assess the efficacy of our suggested method regarding BER under our proposed system model.

As shown in Fig. 1, the altitudes of the airship and the satellite are 25 km and 250 km, respectively. The airship is quasi-stationary, while the satellite moves in a circular orbit around the Earth at a velocity of 7.755 km/s. The simulation scenario involves the satellite transitioning from $\theta = 30^\circ$ to $\theta = 15^\circ$, separated into 100 time steps. The signal frequency f is 220GHz (0.22 THz), which lies within the effective window of 100-300 GHz [16]. The Doppler shift, on the order of MHz, varies rapidly with satellite motion. $N_T = N_R = 16$, 16×16 planar antenna arrays are used at the transmitter and the receiver, respectively. The beam misalignment angles θ_T and θ_R both follow a Gaussian distribution with a mean of 0° and a variance of 1° [3]. These angles are i.i.d. random variables at different time steps. At the transmitter, the amplitude and phase imbalance parameters ϵ_T and φ_T are 0.3 and -5° , respectively [7]. The nonlinearity order of the PA is 5 with the complex parameters $\alpha_1 = 1.1214 + j0.0321$, $\alpha_3 = 0.0994 - j0.2355$ and $\alpha_5 = -0.952 - j1.032$ [17]. At the receiver, the amplitude and phase imbalance parameters ϵ_R and φ_R are 0.3 and -5° , respectively [7]. The phase noise δ_T and δ_R follow a Gaussian distribution with a mean of 0° and a variance of 3° [4]. For SPCPM, the hyperparameters are summarized in Table III. For networks, $m = 40$, so the system has 40 pairs of autoencoders and fitting networks, resulting in a Doppler phase shift resolution of 9° . Both the input and output layers have a

TABLE II
THE HYPERPARAMETERS USED IN THE TRAINING

Network	Epochs	Batch size	Learning rate
Fitting Network	15	10^4	10^{-4}
Autoencoder	500	10^3	10^{-3}

TABLE III
THE HYPERPARAMETERS USED IN SPCPM

Parameter	Value	Parameter	Value
Input size	1	N_{pre}	40
Output size	1	N_{up}	20
Hidden state dimension	40	k	5
Learning rate	0.06		

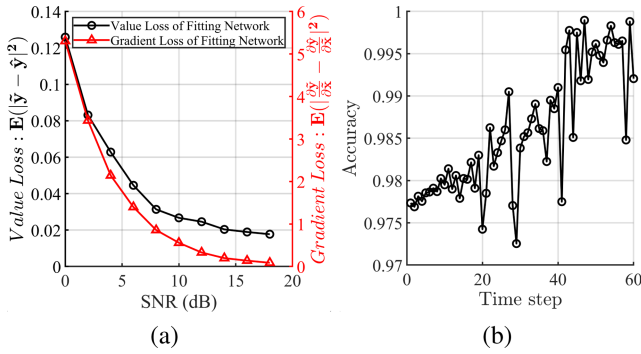


Fig. 3. (a) The value loss and gradient loss of the fitting network. Value loss can be expressed as $\mathbb{E}(|\tilde{y} - \hat{y}|^2)$ and gradient loss can be expressed as $\mathbb{E}(|\frac{\partial \tilde{y}}{\partial x} - \frac{\partial \hat{y}}{\partial x}|^2)$. (b) The accuracy of selecting the autoencoder based on SPCPM predictions over the last 60 time steps.

size of 2, and the sizes of the hidden layers are shown in Table I, where for example the fitting network has three hidden layers and each one has 40 neurons. The hyperparameters used in the training are shown in Table II. To assess the system's capability to manage fluctuations in Doppler phase shifts, 20,000 symbols are transmitted at each time step, with a symbol duration of 0.1 ns.

Fig. 3(a) illustrates the average effectiveness over the 40 fitting networks. As the signal-to-noise-ratio (SNR) increases, both value loss and gradient loss approach zero, indicating that the fitting network effectively reconstructs the actual channel characteristics. Fig. 3(b) demonstrates that the prediction accuracy of SPCPM is precise enough to select the correct autoencoder.

Fig. 4 illustrates the BER performance of our design, compared with that of MMSE equalization, the autoencoder system in [7] and the twin-phase FNN design in [18]. All of the four methods are evaluated under the dynamic channel model proposed in this article, while the autoencoder is also assessed under a static channel devoid of Doppler shift. Fig. 4 demonstrates that the autoencoder fails to handle highly dynamic channels but achieves excellent BER performance under static channel conditions. The performance achieved by our approach under dynamic channels closely resembles the results obtained by the autoencoder in static channels and outperforms twin-phase design and MMSE equalization. Our design achieves approximately 4 dB SNR gain over the twin-phase design at the BER level of 10^{-3} under QPSK modulation. The SNR gain of our design over MMSE equalization for the QPSK modulation, for instance, is around 5 dB at the BER level of 10^{-3} . The SNR difference between our design and autoencoder at the

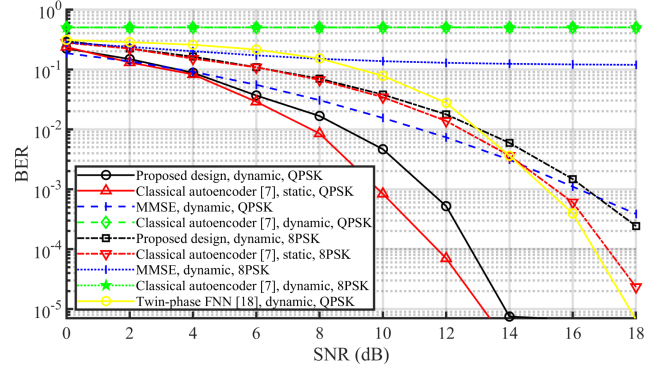


Fig. 4. BER performance comparison with classical autoencoder [7], MMSE equalization, and twin-phase FNN design [18].

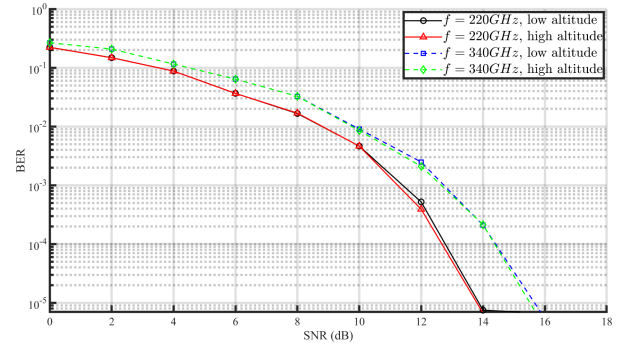


Fig. 5. BER performances for the proposed dynamic autoencoder architecture under different configurations.

BER level of 10^{-3} is minimal, approximately 2 dB for QPSK and 1 dB for 8PSK.

Fig. 5 illustrates the BER performance of our channel-adaptive autoencoder under QPSK modulation for two altitude configurations: a low-altitude scenario (airship/satellite at 25 km/250 km) and a high-altitude scenario (airship/satellite at 30 km/300 km). To reflect the more severe hardware distortions often encountered at higher frequency, the I/Q imbalance parameters are adjusted to $\epsilon_T = \epsilon_R = 0.4$, $\varphi_T = \varphi_R = -6^\circ$ at 340 GHz. It can be observed that, due to the more severe hardware impairments, the performance of the system at 340 GHz is slightly worse than that at 220 GHz. However, our proposed architecture can adapt to changes in altitude.

V. CONCLUSION

This work proposed a dynamic autoencoder system and a SPCPM to adapt to time-varying channels and hardware hybrid distortions in an aerospace THz communication scenario. A bank structure of autoencoders plus fitting networks was developed to ensure robust transmission despite varying Doppler shifts in the channel, while SPCPM was employed to anticipate fluctuations in channel conditions, enabling the system to adapt to temporal variations and select the appropriate autoencoder. A communication scenario was established between a quasi-stationary stratospheric airship and a VLEO satellite, validating that the proposed dynamic autoencoder achieved BER performance gains compared to other methods under intricate dynamic channel conditions and the performance was comparable to that of the static autoencoder functioning under more relaxed channel conditions.

There exist several open issues for our proposed autoencoder architecture. Firstly, current design needs further generalizations to support

multiple moving nodes at the base station. One feasible approach is to serve different users in a time-division manner, where the SPCPM and the autoencoder need to be retrained when switching the connected users. This approach necessitates more advanced scheduling scheme as well as lightweight design of autoencoder structures. Secondly, more intricate dynamic factors (e.g., unpredictable movement/jitter of UAVs) need careful consideration, which can be alleviated by leveraging the integrated sensing and communication techniques, where sensing capabilities can assist communication tasks, particularly for precise beam alignment/tracking and Doppler shift estimation. These strategies will be explored in our future work.

REFERENCES

- [1] T. Mao, L. Zhang, Z. Xiao, Z. Han, and X. -G. Xia, "Terahertz-band near-space communications: From a physical-layer perspective," *IEEE Commun. Mag.*, vol. 62, no. 2, pp. 110–116, Feb. 2024.
- [2] Z. Yang, W. Gao, and C. Han, "A universal attenuation model of terahertz wave in space-air-ground channel medium," *IEEE Open J. Commun. Soc.*, vol. 5, pp. 2333–2342, 2024.
- [3] M. T. Dabiri and M. Hasna, "Pointing error modeling of mmWave to THz high-directional antenna arrays," *IEEE Wireless Commun. Lett.*, vol. 11, no. 11, pp. 2435–2439, Nov. 2022.
- [4] W. Zhao, C. Han, H.-J. Song, and E. Björnson, "DNN based two-stage compensation algorithm for THz hybrid beamforming with imperfect hardware," *IEEE Trans. Wirel. Commun.*, early access, Aug. 5, 2025, doi: [10.1109/TWC.2025.3593266](https://doi.org/10.1109/TWC.2025.3593266).
- [5] T. Mao, Q. Wang, and Z. Wang, "Receiver design for the low-cost terahertz communication system with hardware impairment," in *Proc. IEEE Int. Conf. Commun.*, 2020, pp. 1–5.
- [6] A. Liao et al., "Terahertz ultra-massive MIMO-based aeronautical communications in space-air-ground integrated networks," *IEEE J. Sel. Areas Commun.*, vol. 39, no. 6, pp. 1741–1767, Jun. 2021.
- [7] Z. Huang, D. He, J. Chen, Z. Wang, and S. Chen, "Autoencoder with fitting network for terahertz wireless communications: A deep learning approach," *China Commun.*, vol. 19, no. 3, pp. 172–180, Mar. 2022.
- [8] H. Ye, L. Liang, G. Y. Li, and B. -H. Juang, "Deep learning-based end-to-end wireless communication systems with conditional GANs as unknown channels," *IEEE Trans. Wireless Commun.*, vol. 19, no. 5, pp. 3133–3143, May 2020.
- [9] T. -Y. Tung, S. Kobus, J. P. Roig, and D. Gündüz, "Effective communications: A joint learning and communication framework for multi-agent reinforcement learning over noisy channels," *IEEE J. Sel. Areas Commun.*, vol. 39, no. 8, pp. 2590–2603, Aug. 2021.
- [10] S. Wang, T. Lv, X. Zhang, Z. Lin, and P. Huang, "Learning-based multi-channel access in 5G and beyond networks with fast time-varying channels," *IEEE Trans. Veh. Technol.*, vol. 69, no. 5, pp. 5203–5218, May 2020.
- [11] H. Saeeddeen, M. -S. Alouini, and T. Y. Al-Naffouri, "An overview of signal processing techniques for terahertz communications," *Proc. IEEE*, vol. 109, no. 10, pp. 1628–1665, Oct. 2021.
- [12] C. Han, W. Gao, N. Yang, and J. M. Jornet, "Molecular absorption effect: A double-edged sword of terahertz communications," *IEEE Wireless Commun.*, vol. 30, no. 4, pp. 140–146, Aug. 2023.
- [13] Y. Yu, X. Si, C. Hu, and J. Zhang, "A review of recurrent neural networks: LSTM cells and network architectures," *Neural Computation*, vol. 31, no. 7, pp. 1235–1270, Jul. 2019.
- [14] C. Nwankpa, W. Ijomah, A. Gachagan, and S. Marshall, "Activation functions: Comparison of trends in practice and research for deep learning," 2018, *arXiv:1811.03378*.
- [15] D. P. Kingma and J. Ba, "Adam: A method for stochastic optimization," in *Proc. Int. Conf. Learn. Representations (ICLR)*, May 2015, pp. 1–15.
- [16] H. Tataria, M. Shafi, A. F. Molisch, M. Dohler, H. Sjöland, and F. Tufvesson, "6G wireless systems: Vision, requirements, challenges, insights, and opportunities," *Proc. IEEE*, vol. 109, no. 7, pp. 1166–1199, Jul. 2021.
- [17] N. Moghadam, G. Fodor, M. Bengtsson, and D. Love, "On the energy efficiency of MIMO hybrid beamforming for millimeter-wave systems with nonlinear power amplifiers," *IEEE Trans. Wireless Commun.*, vol. 17, no. 11, pp. 7208–7221, Nov. 2018.
- [18] D. He, Z. Wang, T. Q. S. Quek, S. Chen, and L. Hanzo, "Deep learning-assisted terahertz QPSK detection relying on single-bit quantization," *IEEE Trans. Commun.*, vol. 69, no. 12, pp. 8175–8187, Dec. 2021.