Non-Stationarity Characterization and Geometry-Cluster-Based Stochastic Model for High-Speed Train Radio Channels

Yan Zhang, *Member, IEEE,* Kaien Zhang, Ammar Ghazal, *Member, IEEE,* Wancheng Zhang, Zijie Ji, and Limin Xiao, *Member, IEEE*

***Abstract*—In time-variant high-speed train (HST) radio channels, the scattering environment changes rapidly with the movement of terminals, leading to a serious deterioration in communication quality. In the system- and link-level simulation of HST channels, this non-stationarity should be characterized and modeled properly. In this paper, the sizes of the quasi-stationary regions are quantified to measure the significant changes in channel statistics, namely, the average power delay profile (APDP) and correlation matrix distance (CMD), based on a measurement campaign conducted at 2.4 GHz. Furthermore, parameters of the multi-path components (MPCs) are estimated and a novel clustering-tracking-identifying algorithm is designed to separate MPCs into line-of-sight (LOS), periodic reflecting clusters (PRCs) from power supply pillars along the railway, and random scattering clusters (RSCs). Then, a non-stationary geometry-cluster-based stochastic model is proposed for viaduct and hilly terrain scenarios. Furthermore, the proposed model is verified by measured channel statistics such as the Rician K factor and the root mean square delay spread. The temporal autocorrelation function and the spatial cross-correlation function are presented. Quasi-stationary regions of the model are analyzed and compared with the measured data, the standardized IMT-Advanced (IMT-A) channel model, and a published non-stationary IMT-A channel model. The good agreement between the proposed model and the measured data demonstrates the ability of the model to characterize the non-stationary features of propagation environments in HST scenarios.**

***Index Terms*—Average power delay profile (APDP), correlation matrix distance (CMD), geometry-cluster-based stochastic model, non-stationarity, quasi-stationary region.**

I. Introduction

In recent years, the high-speed train (HST) gets rapid development and widespread deployment, resulting in growing demands for wireless connections from users on the train. The Long-Term Evolution-Railway (LTE-R) [1] and the fifth-generation communication system for railway (5G-R) [2] are recommended for future high data rate transmissions. The speed of HST can be up to 600 km/h, which brings about several challenges such as high penetration losses, fast handover, and fast travel through diverse scenarios [3]. Consequently, it leads to increasing research interests in high-mobility communication, calling for the development of new theories and technologies.

One fundamental work for the design and deployment of the HST mobile communication systems is to investigate the channel models in typical propagation scenarios. Most classic channel models, including some standardized channel models, are based on the wide-sense stationary (WSS) assumption. However, within the high-mobility scenarios, e.g., HST, vehicle-to-vehicle (V2V) [4], [5], and unmanned aerial vehicle (UAV) [6], [7], this assumption may be only valid in a small region [8]. Due to the fast movements of the transmitter (Tx), receiver (Rx), and scatterers, the channel statistics will change rapidly. In [9], it has been proved that using the WSS assumption in these high-mobility scenarios may lead to performance degradation. Therefore, in order to describe the channels in high-mobility scenarios, the non-stationarity must be measured and modeled [10]–[12].

The premise of characterizing the non-stationarity is to determine the size of the region in which the stationary assumption holds. Instead of using the WSS region in which the first- and second-order statistical moments are time-invariant, we use the term quasi-stationary region (or quasi-stationarity region) similar to the definitions in [13]–[15]. The quasi-stationary region characterizes the region in which the statistics of channel are similar enough compared to the statistics of the neighboring channel so that the channel can be approximately considered to be stationary [13]. It should be noted that the WSS assumption is stricter than the quasi-stationary condition according to their definitions. However, in practical communication systems, we are more likely to concern with channel statistics rather than the instantaneous complex channel responses or the statistical moments. Thus, using the definition of the quasi-stationary region could be useful for system design and technological development [15].

To quantify the size of the quasi-stationary region from measurements, several metrics have been proposed in the literature. According to the channel statistics, these metrics could be generally classified into: 1) metrics based on delay and power statistics, including those defined by power delay profile (PDP) [16], average power delay profile (APDP) [17],

This work was supported by the National Key Research and Development Program of China (No.2020YFB1804901) and the National Natural Science Foundation of China (No.62271051). (Corresponding author: Yan Zhang.)

Y. Zhang, K. Zhang, and W. Zhang are with the School of Information and Electronics, Beijing Institute of Technology, Beijing 100081, China (e-mail: zhangy@bit.edu.cn; 3120210829@bit.edu.cn; zhangwancheng@bit.edu.cn).

A. Ghazal is with the Faculty of Computing Engineering and Media, De Montfort University, Leicester LE1 9BH, UK (Email: ammar.ghazal@dmu.ac.uk).

Z. Ji is with the Institute of Telecommunication and Navigation Satellite, China Academy of Space Technology, Beijing 100094, China (Email: jizijie@bit.edu.cn).

L. Xiao is with the Beijing National Research Center for Information Science and Technology, Tsinghua University, Beijing 100084, China (e-mail: xiaolm@tsinghua.edu.cn).

and spectral divergence (SD) [15], [18]; 2) metrics based on spatial statistics, like correlation matrix distance (CMD) [15], [19]–[21] and channel correlation function [22], [23]; 3) metrics based on the collinearity of local scattering function [24]; and 4) metrics based on large-scale fading correlation, e.g., shadow fading (SF) [15], [21] and root mean square (RMS) delay spread [25]. The existing research works of the non-stationary analysis based on measurements were mainly for vehicular communications [15], [16], [21]–[23]. In HST communication scenarios, although there have been many measurement campaigns, e.g., those introduced in [26]–[30], studies on characterizing the non-stationarity based on measurements were still relatively few [18], [31]. For deployments of HST mobile communication systems, more results of the non-stationarity for different scenarios based on measurement are still needed. In this paper, measurement campaigns at 2.4 GHz with 40 MHz bandwidth in two typical HST scenarios, i.e., the viaduct and hilly terrain, are conducted. Based on the measured data, the non-stationarity is characterized by the APDP and CMD metrics, and the results are compared and analyzed. Moreover, the choice of metrics and decision thresholds, which are crucial in assessing the sizes of quasi-stationary regions [32] and ought to be based on the system configuration [15], are discussed in this work.

Furthermore, in order to provide suitable references for system- and link-level simulation in HST radio channels, some classic channel models have been given in existing research works, which could be categorized as: 1) statistical models directly obtained from measurements, including tap-delay-line (TDL) [33], [34] model and joint modeling of large-scale parameters [35]; 2) deterministic model represented by the ray-tracing model [33], [36], [37]; and 3) theoretical geometry-based stochastic model (GBSM) [10], [38]–[41]. The statistical models have low complexity as they directly reflect the characteristics of measured data, but usually with poor generality. The ray-tracing model [36], [37] is accurate, however, it can only be deployed in the simulated scenarios and with high complexity. By generating parameters randomly based on empirical values and distributions, the GBSMs are suitable for plenty of similar scenarios, but at the same time, the propagation environment might be oversimplified [37].

In order to make an appropriate compromise between consistency with the actual propagation environments and generality of the channel model, one type of geometry-cluster-based stochastic model has been proposed in recent years. It is a kind of measurement-based channel model, and it assumes that the scatterers are not distributed within a specific regular shape, but are given according to the characteristics of the actual propagation scenario. The geometry-cluster-based stochastic model belongs to irregular-shaped GBSM, which is established by integrating different kinds of clusters. Channel parameters of the multi-path components (MPCs) are similar inside each cluster and demonstrate regular statistical properties, corresponding to the special propagation conditions. In [42], a geometry-based random-cluster model for line-of-sight (LOS) was established under distributed antennas in HST channels. In [43], a geometry-cluster-based multiple-input multiple-output (MIMO) model for V2V considering the scattering environments was proposed. In [44], the vehicular traffic density was taken into account to characterize the non-stationarity in cluster-based V2V modeling. The existing works aimed at V2V scenarios or merely studied the properties of the LOS component in HST scenarios, and the discussion of non-stationarity was not taken into account. To fill this research gap, a non-stationary model is proposed for the viaduct and hilly terrain scenarios in HST radio channels. The proposed model is based on the particular geometric relationships and the statistical properties taken from measurements, extracted by a novel clustering-tracking-identifying (CTI) algorithm inspired by some published works [45]–[47].

The major contributions and novelties of this paper are summarized as follows.

- The results of quasi-stationary distances in viaduct and hilly terrain scenarios are presented using APDP and CMD metrics based on wideband HST channel measurements. The selection of the averaging window is discussed, and the impact of different thresholds on deciding the quasi-stationary distance is illustrated. The comparison results of the two scenarios and different metrics are shown and discussed.
- Based on the measured data, channel parameters are estimated and an automatic CTI algorithm is designed to separate MPCs into three categories, i.e., the LOS path, periodic reflecting clusters (PRCs) from power supply pillars along the railway, and random scattering clusters (RSCs).
- A non-stationary geometry-cluster-based model is implemented based on the geometric relationships and the statistical properties of the extracted clusters. The model is simulated and validated by channel statistics, including the Rician K factor and the RMS delay spread. The non-stationarity is verified by the temporal autocorrelation function (ACF) and the spatial cross-correlation function (CCF). Quasi-stationary distances defined by APDP and CMD metrics are presented. Better agreements with the measured data could be observed compared to the standardized IMT-Advanced (IMT-A) channel model and a published non-stationary IMT-A channel model.

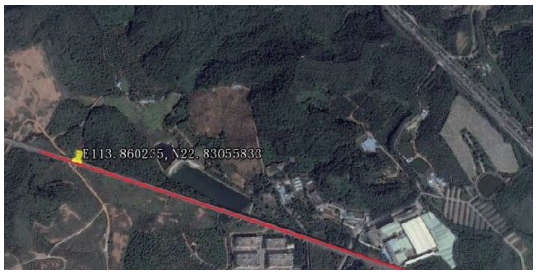
The rest of this paper is organized as follows. In Section II, the measurement setup is introduced including the descriptions of the propagation scenarios and the measurement system. Section III presents the measured results of quasi-stationary distances using two different metrics and the relevant analysis. In Section IV, the channel parameter estimation and CTI algorithm are designed and conducted. The proposed model is then implemented in detail. Model validation and result analysis are presented in Section V. Finally, the conclusions are drawn in Section VI.

II. Measurement Setup

A. Measurement Scenarios

The measurement campaign was carried out on the Guangzhou-Shenzhen Express Rail Link in Guangdong Province, China. Two typical scenarios, namely, viaduct and hilly terrain, were taken into account. Fig. 1(a) and Fig. 1(b)

(a)



(b)

Fig. 1. Top views of the measurement routes in (a) viaduct scenario and (b) hilly terrain scenario.

show the satellite top views of the measurement scenarios where the yellow marker represents the position of the base station (BS), which was regarded as the reference point of distances between Tx and Rx (T-R distance), and the red line represents the measured route. In both scenarios, the train carrying Rx traveled from west to east, and the Tx antenna on the BS was directed towards the east. The lengths of the considered routes were both about 1700 m.

In the viaduct scenario, there were only a few buildings and trees in the propagation environment. The height of the Tx antenna was 40 m, and the train was 20 m above the ground with most of the scatterers under the height of the train. In the hilly terrain scenario, the Tx antenna was 30 m higher than the track on the ground. More propagation-wise surroundings could be observed, consisting of hills, buildings, factories, and woods. In both scenarios, there were periodic power supply pillars along the railway. The horizontal distance of the Tx antenna to the track is 50 m and that is 1.5 m for the periodic power supply pillars. The power supply pillars are distributed parallel and equally spaced on both sides of the track with a 50 m spacing, which are symmetrical about the track.

During the measurement campaign, the railway was in a trial operation stage. As a result, no other trains were travelling in the same/opposite direction. Moreover, no movement of scatterers was observed in the surrounding rural or suburban areas during the measurement. As there were no passengers in the carriages, the in-cabin propagation environment kept almost stationary. A global positioning system (GPS) device connecting with the Rx was used to record the mobile speed and positions of the train. In both scenarios, the velocity of the train was 300 km/h and remained constant over the measured routes.

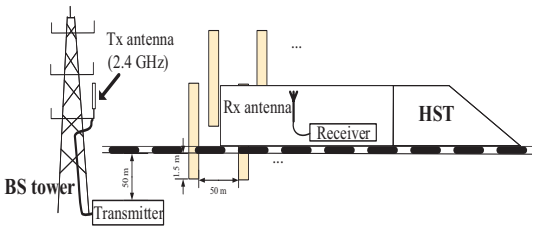


Fig. 2. An illustration of the measurement configuration.

TABLE I
SYSTEM PARAMETERS

Parameter	Value
Central frequency	2.4 GHz
Bandwidth	40 MHz
Tx signal duration	12.8 μs
Snapshot duration	25.6 μs
Transmitted power	20 dBm
Tx antenna configuration	Directional antenna with 17 dBi gain
Rx antenna configuration	Omni-directional antenna with 7 dBi gain
Train speed	300 km/h

B. Measurement System

The Tsinghua University (THU) channel sounder [48] was used to collect raw measured data. During the measurements, it worked at 2.4 GHz central frequency with 40 MHz bandwidth. A wideband linear frequency modulation signal with 12.8 μs duration was radiated at Tx, and an empty guard interval of 12.8 μs was inserted between the signals to resist the interference of propagation delays. Thus, the duration of each snapshot is 25.6 μs. The configurations of this measurement campaign are listed in Table I and the structure of the measurement system is illustrated in Fig. 2. The Tx included a signal generator, a rubidium clock reference, and a power amplifier. The test signal was fed to the Tx antenna with a low insertion loss cable, and the radiated power was 20 dBm. The directional Tx antenna with 17 dBi gain was fixed on the BS tower pointing towards the east. The beamwidth in horizontal direction of the main lobe of Tx antenna pattern is 65 degrees, and is placed symmetrically along the track. The antenna gain remains almost constant within the beamwidth and is sufficient to illuminate all scatterers present in the environment. At the Rx side, an omnidirectional antenna with 7 dBi gain was fixed on the window inside the carriage. The pictures of Tx and Rx antennas are shown in Fig. 3. The Rx owned a rubidium clock to eliminate the synchronization offset. Then, the measured data was stored online, and the data processing and calibration were carried out offline.

There are generally two wireless coverage schemes for HST [49], direct coverage (DC) and relay coverage (RC). In the DC scheme, terminals directly establish a link to the BS located near the rail track. In the RC scheme, the dedicated mobile relay stations [50] are deployed on the surface of the train to extend the coverage of the outdoor BS into train carriages. During our measurement campaign, the DC scheme was adopted.

Furthermore, due to safety restrictions, only one antenna is

(a) (b)

Fig. 3. Antenna positions of (a) Tx at BS tower and (b) Rx in the carriage.

allowed to be employed on the BS side and train side, establishing a single-input single-output (SISO) system. However, in the following analysis, we will consider CMD as one of the metrics for measuring the non-stationarity and use the space-alternating generalized expectation-maximization (SAGE) algorithm to estimate parameters of MPCs, which both require the knowledge of the spatial structure, i.e., channel responses from multiple antennas. Thus, the virtual array technique [51]–[53], which utilizes several moving antennas to achieve a high spatial resolution, is used to form virtual uniform linear arrays (ULAs) parallel to the railway. In theory, any type of array can be deployed, we use ULAs due to the limitations of measurement system. Several adjacent snapshots in the moving single-antenna measurement, where the channel is regarded as constant, are considered as equivalent elements of the multiple-antenna array. Adopting the virtual array technique offers an additional advantage, i.e., since there is no mutual coupling between neighboring elements, the antenna calibration is not required [53]. The number of equivalent array elements K, is limited by [52]

$$K < 1 + \frac{1}{2 \cdot \frac{v\Delta t}{\lambda}} \tag{1}$$

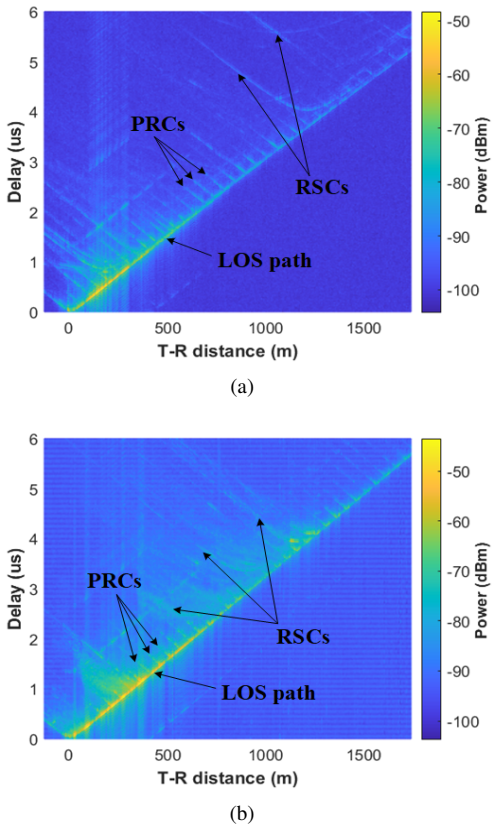
where v is the speed of the train, Δt is the time interval of adjacent snapshots, and λ is the wavelength. Substituting the values in Table I to (1), the number of virtual array elements $K < 30.3$. Thus, an equivalent ULA with a maximum of 30 elements can be constructed, and the spacing between adjacent virtual elements is equal to $v\Delta t$.

III. Channel Quasi-Stationary Distance

Since the Rx on the train travels with a constant speed v on a given route, we use the quasi-stationary distance $d_{\mathrm{qs}}(t)$ or quasi-stationary interval $T_{\mathrm{qs}}(t)$ to define the size of the quasi-stationary region in HST scenarios. Their relationship can be expressed as

$$d_{\mathrm{qs}}(t) = vT_{\mathrm{qs}}(t). \tag{2}$$

Here, both $d_{\mathrm{qs}}(t)$ and $T_{\mathrm{qs}}(t)$ are functions of time. It is noteworthy that the speed is only related with the quasi-stationary interval but is uncorrelated with the quasi-stationary distance. By selecting different channel statistics as metrics, we can obtain different definitions for quasi-stationary distances. In the following analysis, two kinds of metrics defined by APDP and CMD are taken into account.



(a)

(b)

Fig. 4. APDPs in (a) viaduct scenario and (b) hilly terrain scenario.

A. APDP Metric

After offline processing of the measured data, we can get the wideband channel impulse response (CIR) $h(i\Delta t, p\Delta\tau)$, where i is the snapshot index in the time domain, p is the delay bin index, and $\Delta\tau = 10$ ns is the delay interval between two adjacent delay bins.

To remove the impact of the small-scale fading, a sliding window is adopted for averaging over snapshots and the APDP can be calculated as

$$\overline{P_h}\left(i\Delta t, p\Delta\tau\right) = \frac{1}{W}\sum_{k=i}^{i+W-1}\left|h(k\Delta t, p\Delta\tau)\right|^2 \tag{3}$$

where $|\cdot|$ denotes the absolute value operator and W is the number of snapshots to be averaged, which is selected as 900 and the reason will be discussed in Section III-C.

The APDPs in viaduct and hilly terrain scenarios are shown in Fig. 4. The horizontal axis and vertical axis denote the T-R distance and propagation delay, respectively. It should be noted that since the Rx antenna is deployed on the north side of the carriage, the power of scattering clusters from the south side of the train may suffer a great loss due to the train body, so the detected scatterers are mainly located in the north side of the train.

As depicted in Fig. 4, examples of three main kinds of MPCs, i.e., the LOS path, the PRCs, and the RSCs, are marked by arrows. Firstly, because there is no obstructor between the BS and the train in both scenarios, the LOS path always has the

strongest power and shortest delay which decreases until the train passes the BS and increases almost linearly afterwards. Owing to the directional Tx antenna, the received signal is relatively weak at first while it becomes significantly strong after the train passes the BS. Then, the LOS signal strength decreases with the increasing T-R distance. Apart from the LOS path, the clusters go through a general birth-death process with the movement of the train. In both scenarios, it can be observed that some trajectories are approximately periodic and parallel to each other, which are close to the LOS path. The intersections between these trajectories and the LOS path have almost equivalent intervals. We define this kind of clusters as PRCs, due to the fact that they are from the reflections of the power supply pillars along the railway with identical separation between adjacent ones, which are ubiquitous in outdoor HST communication environments. The other kind of clusters, namely the RSCs, usually have strong powers and large delay spreads. They are corresponding to some large scatterers near the railway, e.g, hills and buildings. More trajectories of the RSCs can be observed in the hilly terrain scenario due to its more propagation-wise environment, which is also proven by the RMS delay spread calculated from the measured data, i.e., the mean value in the viaduct is 166.6 ns while that in the hilly terrain is 237.9 ns.

In order to calculate the quasi-stationary distance with respect to the APDPs, the correlation coefficient between two APDPs can be calculated as

$$c\left(i,j\right)=\frac{\sum\limits_{p=1}^{N_\tau}\overline{P_h}\left(i\Delta t,p\Delta\tau\right)\cdot\overline{P_h}\left(j\Delta t,p\Delta\tau\right)}{\max\left\{\sum\limits_{p=1}^{N_\tau}\overline{P_h}(i\Delta t,p\Delta\tau)^2,\sum\limits_{p=1}^{N_\tau}\overline{P_h}(j\Delta t,p\Delta\tau)^2\right\}} \tag{4}$$

where N_τ is the total number of delay bins and $\max\{\cdot,\cdot\}$ is the operator that takes the maximum value. Then, the quasi-stationary interval can be expressed as

$$T_{\mathrm{qs}}(i)=(i_{\max}-i_{\min})\cdot\Delta t \tag{5}$$

where $i_{\max}$ and $i_{\min}$ denote the time-variant maximum and minimum bounds of the quasi-stationary intervals at time $i\Delta t$, respectively, and can be expressed as

$$\begin{cases} i_{\min}=\underset{1\le j\le i-1}{\arg\max}\, c(i,j)<c_{\mathrm{th}} \\ i_{\max}=\underset{i+1\le j\le N_t-W}{\arg\min}\quad c(i,j)<c_{\mathrm{th}}. \end{cases} \tag{6}$$

Here, N_t is the total number of snapshots and c_{th} is a given threshold of the correlation coefficient between two APDPs. After choosing the appropriate c_{th}, the quasi-stationary interval $T_{\mathrm{qs}}(i)$ at time $i\Delta t$ can be determined. Then, the quasi-stationary distance can be calculated according to (2). The impact of the c_{th} on the quasi-stationary distance results will be discussed in Section III-D.

B. CMD Metric

The channel transfer function (CTF) within bandwidth B at time $i\Delta t$ can be computed by discrete Fourier transformation (DFT) of the CIR, which is expressed as

$$H(i\Delta t,q\Delta f)=\sum_{p=1}^{N_\tau}h\left(i\Delta t,p\Delta\tau\right)e^{-j2\pi q\Delta f p\Delta\tau} \tag{7}$$

where $q=1,...,N_f$ is the index of frequency bins and $\Delta f=B/N_f$ is the frequency difference. N_f here is the total number of frequency bins.

As mentioned previously in Section II, we employ the virtual array technique to form ULAs at the Rx side. With the movement of the Rx antenna, continuous samples in the time domain can be viewed as different virtual antennas. The CTF for the virtual array with K elements at time $i\Delta t$ can be expressed as

$$\begin{aligned}\mathbf{H}(i\Delta t,q\Delta f)=[&H(i\Delta t,q\Delta f),H((i+1)\Delta t,q\Delta f),\\&...,H((i+K-1)\Delta t,q\Delta f)]^{\mathrm{T}}\end{aligned} \tag{8}$$

where $[\cdot]^{\mathrm{T}}$ denotes the transpose operation.

Then, the narrowband correlation matrices at the Rx side can be expressed as

$$\mathbf{R}_{\mathrm{Rx}}(i\Delta t)=\frac{1}{N_fW}\sum_{k=i}^{i+W-1}\sum_{q=1}^{N_f}\mathbf{H}(k\Delta t,q\Delta f)\mathbf{H}(k\Delta t,q\Delta f)^{\mathrm{H}} \tag{9}$$

where $[\cdot]^{\mathrm{H}}$ denotes the Hermitian transpose operation.

The CMD is usually used as a measure to evaluate whether the spatial structure of the channel has changed to a significant extent [20]. It is defined as the inner product between the correlation matrices for $i\Delta t$ and $j\Delta t$

$$d_{\mathrm{corr}}(i,j)=1-\frac{\mathrm{tr}\{\mathbf{R}_{\mathrm{Rx}}(i\Delta t)\cdot\mathbf{R}_{\mathrm{Rx}}(j\Delta t)\}}{\|\mathbf{R}_{\mathrm{Rx}}(i\Delta t)\|_{\mathrm{F}}\cdot\|\mathbf{R}_{\mathrm{Rx}}(j\Delta t)\|_{\mathrm{F}}} \tag{10}$$

where $\mathrm{tr}\{\cdot\}$ is the trace operator and $\|\cdot\|_{\mathrm{F}}$ is the Frobenius norm. The CMD ranges from zero to one, it becomes zero if the correlation matrices are equal up to a scaling factor and one if they differ to a maximum extent.

Based on the CMD metric, we can also define the quasi-stationary interval $T'_{\mathrm{qs}}(i)$ at time $i\Delta t$ as the maximum interval over which the CMD remains below a certain threshold as

$$T'_{\mathrm{qs}}(i)=(i'_{\max}-i'_{\min})\cdot\Delta t \tag{11}$$

where the minimum and maximum time-variant bounds of the quasi-stationary intervals are

$$\begin{cases} i'_{\min}=\underset{1\le j\le i-1}{\arg\max}\, d_{\mathrm{corr}}(i,j)\ge c'_{\mathrm{th}} \\ i'_{\max}=\underset{i+1\le j\le N_t-W-L}{\arg\min}\quad d_{\mathrm{corr}}(i,j)\ge c'_{\mathrm{th}}. \end{cases} \tag{12}$$

Similarly, the quasi-stationary distance defined by CMD can also be computed according to (11), (12), and (2).

C. Discussion about the Averaging Window

Average operations are carried out in the time domain and frequency domain to estimate the correlation matrix in (9). Enough independent samples for averaging must be provided to these two operations. In the time domain, the number of independent samples is decided by the window W and the coherence time T_c, which is the time duration over which two

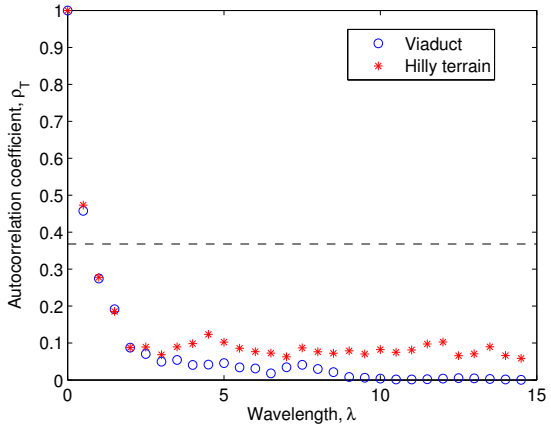


Fig. 5. The envelop autocorrelation coefficient in viaduct and hilly terrain scenarios with the threshold $1/e$.

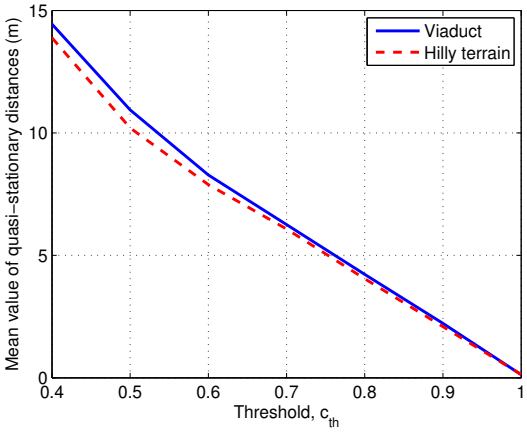


Fig. 6. Mean values of the estimated time-variant quasi-stationary distances defined by APDP with different threshold values.

received signals have a strong potential for amplitude correlation. Consequently, T_c quantifies the similarity of channel responses at different times [54], and can be defined as the time duration over which the signal autocorrelation coefficient reduces to a certain value like $1/e$ [18].

The instantaneous path gain at time $i\Delta t$ is

$$P_G(i\Delta t) = \sum_{p=1}^{N_\tau} |h(i\Delta t, p\Delta\tau)|^2. \tag{13}$$

After removing the mean value of P_G within each window, we get the normalized path gain NP_G as [15]

$$\mathrm{NP}_G(i\Delta t) = \sqrt{\frac{P_G(i\Delta t)}{\frac{1}{W}\sum_{k=i}^{i+W-1} P_G(k\Delta t)}}. \tag{14}$$

Then, the signal autocorrelation coefficient can be expressed as [15]

$$\rho_T(i\Delta t, \Delta i\Delta t) = \frac{\mathrm{E}\left[\mathrm{NP}_G(i\Delta t)\cdot \mathrm{NP}_G((i+\Delta i)\Delta t)\right]}{\sigma_{\mathrm{NP}_G}} \tag{15}$$

where $\mathrm{E}[\cdot]$ is the expectation operator and σ_{NP_G} is the standard deviation of normalized path gain.

The estimated autocorrelation coefficient values in both scenarios are shown in Fig. 5. It is demonstrated that the samples with one wavelength spacing have weak correlations when using the threshold $1/e$. Similar results are observed in all the measurements. Thus, with a window $W = 900$ equaling to 15.36λ and the coherence time about one wavelength, we can obtain at least 15 independent samples in the time domain. It is noted that the length of window W is far smaller than the size of quasi-stationary distance, and will not impact the identification of quasi-stationary regions.

Meanwhile, in the frequency domain, the coherent bandwidth is a statistical measure of the frequency range over which the channel can be considered flat, i.e., the approximate maximum bandwidth over which two frequencies of a signal are likely to experience comparable or correlated amplitude fading [54]. The coherent bandwidth can be estimated according to the RMS delay spread. For example, if the coherent bandwidth is defined as the bandwidth over which the frequency correlation coefficient is above 0.5, the coherent bandwidth can be approximately calculated as [54]

$$B_c \approx \frac{1}{5\sigma_\tau} \tag{16}$$

where σ_τ is the RMS delay spread. As aforementioned, the average RMS delay spreads in the viaduct and hilly terrain are 166.6 ns and 237.9 ns, respectively. Therefore, the corresponding coherent bandwidth values are 1.2 MHz and 0.84 MHz, respectively. Within the whole bandwidth of 40 MHz, 33 and 47 independent frequency samples can be extracted by division operations.

As a result, we can get at least $15 \times 33 = 495$ independent samples in viaduct scenario and $15 \times 47 = 705$ independent samples in hilly terrain scenario for the averaging operation, both of which are large enough for removing the impact of the small-scale fading and analyzing the quasi-stationary distance [15].

D. Analysis of Different Thresholds, Scenarios, and Metrics

According to (6) and (12), different thresholds will lead to different values of quasi-stationary distances. In Fig. 6, the mean values of the estimated time-variant quasi-stationary distances defined by APDP in both scenarios with different thresholds are illustrated. It is shown that with the decrease of the threshold c_{th}, the mean value of the quasi-stationary distance increases. Moreover, the relationship between the threshold and the quasi-stationary distance defined by CMD in both scenarios is shown in Fig. 7. As calculated in (1), the virtual array contains 30 virtual antennas with 0.5λ spacing. With an increasing threshold c'_{th}, the mean value of quasi-stationary distance increases. The opposite trend of changing with thresholds is consistent with their definitions in (4) and (10).

The values of the quasi-stationary distances are listed in Table II. The threshold is selected as 0.5 for the APDP metric

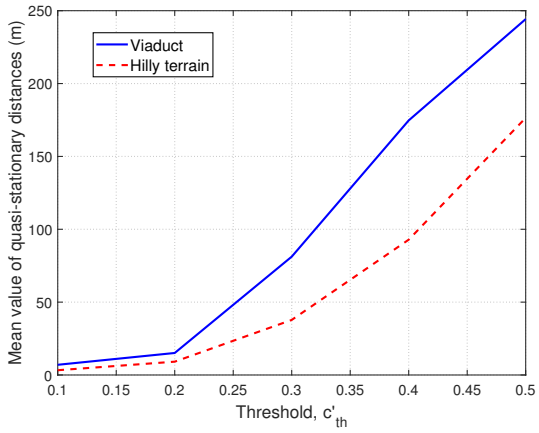


Fig. 7. Mean values of the estimated time-variant quasi-stationary distances defined by CMD with different threshold values.

[55] and 0.2 for the CMD metric [15] in both scenarios. From Table II, it can be observed that the quasi-stationary distances in the hilly terrain scenario are smaller than those in the viaduct scenario whether APDP or CMD is selected as the metric. Considering CMD-related results, the quasi-stationary distance in the hilly terrain is equal to 18.22 m for 90% while that in the viaduct scenario is 40.15 m for 90%. The mean value of the quasi-stationary distances in the hilly terrain is 7.95 m corresponding to about 64 wavelengths and that in the viaduct scenario is 13.82 m corresponding to about 111 wavelengths.

Generally, the APDP metric can reflect the dissimilarity of the statistics in the power-delay domain for a wideband channel, while the CMD metric can be used to describe whether the spatial structures change significantly. These two metrics focus on channel statistics in different domains, so the sizes of the quasi-stationary region defined by APDP and CMD may be different. Although there is no quantitative relationship between these two metrics, they show the same qualitative results, namely, the quasi-stationary distances in the hilly terrain are smaller whatever metric or threshold is chosen. It can be explained as the geometrical distribution of the scatterers in the environment and the relative positions of Tx and Rx determine the spatial structure as well as the power-delay property. In the hilly terrain, more scatterers exist and the propagation environment varies more rapidly with the movement of the train, leading to a faster birth-death process of clusters, and consequently resulting in a faster change in channel statistics.

IV. Geometry-Cluster-Based Stochastic Channel Model

According to the previous analysis, in HST radio channels, the non-stationarity leads to fast time-variance in both the power-delay domain and the spatial domain. The reason for this is the rapid change of scattering environment with different locations of Rx, which maps to time-variant channel parameters and the birth-death process of clusters in the measured data. Motivated by this, we propose a channel model to depict the non-stationarity considering actual propagation environments. The channel parameters are firstly estimated by employing the SAGE algorithm. Then, an automatic CTI algorithm is designed to classify MPCs into three aforementioned categories: the LOS path, PRCs, and RSCs. Then, the stochastic channel model is established by integrating the geometric relationships and statistical properties extracted from measured data.

A. Channel Parameter Estimation

In order to obtain the angular information, the virtual array is deployed as described in Section II. The number of equivalent array elements is chosen as 16 snapshots, which is far smaller than the quasi-stationary distances so that the environments can be regarded invariant. As long as the array number is smaller than the quasi-stationary distance, the estimation error will not be obvious. The received signal can be expressed as

$$h[\mathbf{\Phi}(t)] = \sum_{l=1}^{L} \alpha_l(t) \exp\{j2\pi\nu_l(t)t\}\delta(\tau - \tau_l(t))\mathbf{c}(\varphi_l(t)) \tag{17}$$

where $\delta(\cdot)$ is the Dirac function, L denotes the number of MPCs detected by Rx, $\mathbf{c}(\cdot)$ is the steering vector of Rx antenna, and the parameter set $\mathbf{\Phi}(t)$ contains the time-variant delay, the azimuth of arrival (AOA), Doppler frequency, and complex amplitude at moment t. The radiation pattern of Tx antenna is not included here because the antenna gain is almost constant within the beamwidth.

The high-resolution and high-precision SAGE algorithm [56], [57], which can estimate the parameters in $\mathbf{\Phi}(t)$ jointly and iteratively based on maximum likelihood, is implemented over sliding windows of virtual ULAs. L is selected to be 40, which is sufficient to separate all MPCs above the noise level in our measured data. The Doppler frequency ranges from -666.7 Hz to 666.7 Hz according to the speed of the train.

In addition, only azimuth is considered in (17) because the elevation information cannot be obtained through the ULA. (17) is designed to estimate parameters, and is not used to generate channel responses. It is noted that the proposed model in Section IV-C takes the elevation into account, but is validated only on the azimuth direction due to the restriction of measured data, which does not affect the performance of the channel model.

B. CTI Algorithm

It can be observed from the measured data that MPCs generally appear in the form of scattering clusters. Similar parameters, i.e., delays, AOAs, and amplitudes are shown inside the same cluster with dispersions. In addition, continuous trajectory of the same cluster is observed between adjacent snapshots, i.e., each cluster evolves slowly and smoothly in its survival period. The above two crucial properties offer the theoretical basis for the algorithm design.

The KPowerMeans algorithm [45], which improves the K-means algorithm by adding MPCs power weighting and using

TABLE II
VALUES OF THE QUASI-STATIONARY DISTANCES IN METER

	APDP				CMD			
	Mean	90%	50%	10%	Mean	90%	50%	10%
Viaduct	10.94	16.74	9.82	6.29	13.82	40.15	3.97	1.64
Hilly terrain	10.20	15.36	9.06	6.04	7.95	18.22	2.77	1.26

the multipath component distance (MCD) as the distance metric, has been widely used in clustering MPCs in wireless channels. However, some defects still exist in the KPowerMeans algorithm that impact the clustering performance severely. For instance, results are sensitive to the selection of initial clustering centers, the number of clusters, and some outliers in the measured data. Moreover, plenty of unpredictable missing points exist due to package loss in the measurement, arousing trouble with regard to tracking over snapshots. In order to solve these problems, as illustrated in Algorithm 1, a novel CTI algorithm is designed and divided into the following three steps.

(1) Clustering: unsupervised clustering is carried out to classify the MPCs in each snapshot. Taking the kth snapshot as an example, the noise-free $\mathbf{\Phi}_k$ contains the parameters of all the L_k MPCs and serves as the input of the density-based spatial clustering of applications with noise (DBSCAN) algorithm [58]. After pruning some outliers by setting the minimum number of inclusion points in DBSCAN, the MPCs are clustered into P_k groups, namely, $\mathbf{\Psi}_k = \left\{\mathbf{\Psi}_k^1, ..., \mathbf{\Psi}_k^r, ..., \mathbf{\Psi}_k^{P_k}\right\}$ and each of its elements $\mathbf{\Psi}_k^r = \left[\mathbf{\Phi}_k^{r,1}, ..., \mathbf{\Phi}_k^{r,s}, ..., \mathbf{\Phi}_k^{r,L_k^r}\right]$ contains the parameters of MPCs in the current group. Then, by finding the MPC with the maximum power in each cluster, we obtain the appropriate initial clustering centers and number in $\mathbf{C}_k^{\mathrm{DC}}$. The subscript (u, v) means to take the element of the uth row and vth column of the current matrix. Then, the KPowerMeans algorithm is further used to obtain clustering centers $\mathbf{C}_k^{\mathrm{KPM}}$ for the next step.

(2) Tracking: the cluster sequence is initialized as $\mathbf{CL} = \left\{\mathbf{CL}_1, ..., \mathbf{CL}_w, ..., \mathbf{CL}_{\widetilde{N_c}}\right\}$, where $\widetilde{N_c}$ represents the total number of identified clusters changing with iteration. Each of its components $\mathbf{CL}_w$ contains $\widetilde{L_w}$ clustering centers already found. In order to decide the index to which new clusters belong, we first choose the criteria for comparison $\mathbf{C}_w^{\mathrm{cmp}}$, which is calculated as the median of the last ϖ values for delay and AOA, so that the influence of abnormal values is reduced. Then, we calculate the MCD and the snapshot gap between the clustering centers and the criteria, and select the indexes that satisfy the condition of the MCD threshold c_{MCD} and the snapshot threshold c_t, chosen as 0.5 and 50 [59], respectively. By deploying this snapshot threshold, the influence of temporary obstruction in the measured data is reduced. As long as clusters meet the MCD and snapshot thresholds, they will be regarded as the same cluster. $\mathrm{ID}(\cdot)$ is the operator that takes the snapshot where the MPC belongs. If such clustering centers are found, the index is stored with the minimum MCD, otherwise a new index is established in the cluster sequence.

(3) Identifying: three components are identified according to their spatial characteristics. The LOS path always exists, the PRCs have a large angular spread and a survival time $2\mathrm{d_P} \cdot \cot\alpha$ which will be introduced in Section IV-C, and the rests are classified as the RSCs. As depicted in Fig. 8, AOAs of the LOS component have a jump point at BS and keep stable afterwards. As for PRCs, AOAs experience a change from 0 degree to about 180 degree, and delays decrease first and then increase. The visual results reveal that our proposed algorithm can identify different components in the propagation environments. Overall, the numbers of PRCs and RSCs found in the viaduct scenario are 22 and 6, while those in the hilly terrain are 51 and 11, respectively. This further confirms that the hilly terrain scenario is more propagation-wise.

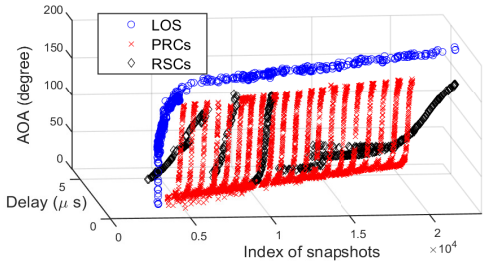


Fig. 8. Samples of LOS, PRCs, and RSCs identified by the proposed algorithm.

C. Model Implementation

In this section, a general MIMO HST channel model for the viaduct and hilly terrain scenarios is proposed. The cluster-related parameter values, which serve as the input of the channel model, are extracted by the aforementioned algorithm from measurements. The geometrical relationships of the LOS path, PRCs, and RSCs are depicted in Fig. 9, and the involved notations and their definitions are listed in Table III. Moreover, for conciseness, only the position vectors of one power supply pillar are shown in Fig. 9. A system with the number of antenna elements P at Tx and Q at Rx is established, with the indices p and q. The antenna patterns and responses can be arbitrary, which are adjusted according to real system settings. The global coordinate system (GCS) at the Tx array center and the local coordinate systems (LCSs) at the Tx and Rx array centers are constructed, aiming at calculating the antenna patterns. For more detailed information about GCS and LCS, please refer to [60]. The start time is set to t_0, when the x-coordinates of Tx and Rx are the same.

In our model, clusters are composed of rays and the whole CIR between the pth Tx and the qth Rx antenna is composed

Algorithm 1: The CTI algorithm for the measured data in HST channels

Step 1. Clustering for each snapshot
foreach $k \leq N_t$ **do**
  $\mathbf{\Phi}_k^s = [\tau_k^s, \varphi_k^s, \alpha_k^s]^{\mathrm{T}}$, $\mathbf{\Phi}_k = \left[\mathbf{\Phi}_k^1, ..., \mathbf{\Phi}_k^s, ..., \mathbf{\Phi}_k^{L_k}\right]$,
  $\mathbf{\Psi}_k = \left\{\mathbf{\Psi}_k^1, ..., \mathbf{\Psi}_k^r, ..., \mathbf{\Psi}_k^{P_k}\right\} = \mathrm{DBSCAN}\left(\mathbf{\Phi}_k\right)$, where $\mathbf{\Psi}_k^r = \left[\mathbf{\Phi}_k^{r,1}, ..., \mathbf{\Phi}_k^{r,s}, ..., \mathbf{\Phi}_k^{r,L_k^r}\right]$ and $\mathbf{\Phi}_k^{r,s} = [\tau_k^{r,s}, \varphi_k^{r,s}, \alpha_k^{r,s}]^{\mathrm{T}}$.
  foreach $r \leq P_k$ **do**
    $\zeta = \arg\max_{\xi} \mathbf{\Psi}_{k,(3,\xi)}^r$, $\mathbf{C}_{k,r}^{\mathrm{DC}} = \mathbf{\Psi}_{k,(:,\zeta)}^r$.
  end
  $\mathbf{C}_k^{\mathrm{DC}} = \left[\mathbf{C}_{k,1}^{\mathrm{DC}}, ..., \mathbf{C}_{k,r}^{\mathrm{DC}}, ..., \mathbf{C}_{k,P_k}^{\mathrm{DC}}\right]$,
  $\mathbf{C}_k^{\mathrm{KPM}} = \left[\mathbf{C}_{k,1}^{\mathrm{KPM}}, ..., \mathbf{C}_{k,r}^{\mathrm{KPM}}, ..., \mathbf{C}_{k,P_k}^{\mathrm{KPM}}\right] = \mathrm{KPowerMeans}\left(\mathbf{\Psi}_k, \mathbf{C}_k^{\mathrm{DC}}, P_k\right)$, where $\mathbf{C}_{k,r}^{\mathrm{KPM}} = [\tau_{k,r}^{\mathrm{KPM}}, \varphi_{k,r}^{\mathrm{KPM}}]^{\mathrm{T}}$.
end
Step 2. Tracking between snapshots
Initialize $\mathbf{CL} = \left\{\mathbf{CL}_1, ..., \mathbf{CL}_w, ..., \mathbf{CL}_{\widetilde{N_c}}\right\}$, where $\mathbf{CL}_w = \left[\mathbf{C}_w^1, ..., \mathbf{C}_w^r, ..., \mathbf{C}_w^{\widetilde{L_w}}\right]$.
foreach $k \leq N_t$ **do**
  foreach $r \leq P_k$ **do**
    $\mathbf{C}_w^{\mathrm{cmp}} = \left[\mathrm{ME}\left(\mathbf{CL}_{w,(1,(\widetilde{L_w}-\varpi):)}\right), \mathrm{ME}\left(\mathbf{CL}_{w,(2,(\widetilde{L_w}-\varpi):)}\right)\right]^{\mathrm{T}}$,
    $\{w_1\} = \left\{w|\mathrm{MCD}(\mathbf{C}_{k,r}^{\mathrm{KPM}}, \mathbf{C}_w^{\mathrm{cmp}}) \leq c_{\mathrm{MCD}}\right\}$, $\{w_2\} = \left\{w|k - \mathrm{ME}\left(\mathrm{ID}\left(\mathbf{CL}_{w,(:,(\widetilde{L_w}-\varpi):)}\right)\right) \leq c_t\right\}$,
    $\{w_3\} = \{w_1\} \cap \{w_2\}$.
    if $\{w_3\} \neq \emptyset$ **then**
      $w_{k,r} = \arg\min_{\{w_3\}} \mathrm{MCD}(\mathbf{C}_{k,r}^{\mathrm{KPM}}, \mathbf{C}_{\{w_3\}}^{\mathrm{cmp}})$, and store the cluster index of $\mathbf{C}_{k,r}^{\mathrm{KPM}}$ in $\mathbf{CL}_{w_{k,r}}$.
    else
      Add a new cluster index in $\mathbf{CL}$.
    end
  end
end
$\mathbf{CL} = \{\mathbf{CL}_1, ..., \mathbf{CL}_w, ..., \mathbf{CL}_{N_c}\}$, where $\mathbf{CL}_w = \left[\mathbf{C}_w^1, ..., \mathbf{C}_w^r, ..., \mathbf{C}_w^{L_w}\right]$.
Step 3. Identifying different components
foreach $w \leq N_c$ **do**
  if $L_w = N_t$ **then**
    The LOS path.
  else if $\begin{cases} \mathbf{CL}_{w,(2,L_w)} - \mathbf{CL}_{w,(2,1)} \geq 150^\circ \\ \mathrm{L_w} = 2\mathrm{d_P} \cdot \cot\alpha \end{cases}$ **then**
    The PRC.
  else
    The RSC.
  end
end

of three components, which can be expressed as

$$\begin{aligned} h_{pq}(t;\tau) = {} & h_{pq}^{\mathrm{LOS}}(t)\delta(\tau - \tau_{pq}^{\mathrm{LOS}}(t)) \\ & + \sum_{n_p=1}^{N_p} \sum_{m_{n_p}=1}^{M_{n_p}(t)} h_{pq}^{m_{n_p}}(t)\delta(\tau - \tau_{pq}^{n_p}(t) - \tau_{pq}^{m_{n_p}}(t)) \\ & + \sum_{n_r=1}^{N_r} \sum_{m_{n_r}=1}^{M_{n_r}(t)} h_{pq}^{m_{n_r}}(t)\delta(\tau - \tau_{pq}^{n_r}(t) - \tau_{pq}^{m_{n_r}}(t)). \end{aligned} \tag{18}$$

In the LOS case, the delay between the pth Tx and the qth Rx antenna $\tau_{pq}^{\mathrm{LOS}}(t)$ is calculated using $\tau_{pq}^{\mathrm{LOS}}(t) = \frac{\left\|\mathbf{\Omega}_{pq}^{\mathrm{LOS}}(t)\right\|_{\mathrm{F}}}{c}$, in which c is the speed of light. The CIR of the LOS component $h_{pq}^{\mathrm{LOS}}(t)$ can be further expressed as

$$\begin{aligned} h_{pq}^{\mathrm{LOS}}(t) = {} & \begin{bmatrix} F_{p,V}^T(\mathbf{\Omega}_{pq}^{\mathrm{LOS}}(t), \mathbf{\Psi}_p^T(t)) \\ F_{p,H}^T(\mathbf{\Omega}_{pq}^{\mathrm{LOS}}(t), \mathbf{\Psi}_p^T(t)) \end{bmatrix}^{\mathrm{T}} \\ & \cdot \begin{bmatrix} e^{j\Phi_{\mathrm{LOS}}^{VV}} & 0 \\ 0 & e^{j\Phi_{\mathrm{LOS}}^{HH}} \end{bmatrix} \begin{bmatrix} F_{q,V}^R(\mathbf{\Omega}_{pq}^{\mathrm{LOS}}(t), \mathbf{\Psi}_q^R(t)) \\ F_{q,H}^R(\mathbf{\Omega}_{pq}^{\mathrm{LOS}}(t), \mathbf{\Psi}_q^R(t)) \end{bmatrix} \\ & \cdot \sqrt{P_{\mathrm{LOS}}(t)} e^{-j2\pi \frac{\left\|\mathbf{\Omega}_{pq}^{\mathrm{LOS}}(t)\right\|_{\mathrm{F}}}{\lambda}} e^{j2\pi\nu_{pq}^{\mathrm{LOS}}(t)t}. \end{aligned} \tag{19}$$

Here, $F_{p(q),V}^{T(R)}(\cdot)$ and $F_{p(q),H}^{T(R)}(\cdot)$ are field pattern functions of the pth Tx antenna (the qth Rx antenna) for vertical and horizontal polarizations, respectively. When the antenna is single-polarized like the measurement system

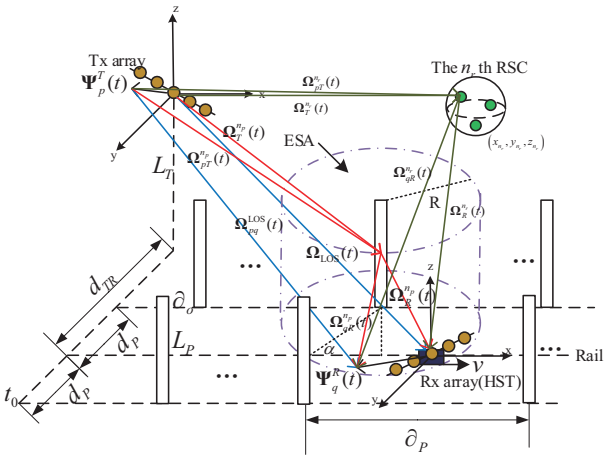


Fig. 9. Geometrical relationships of the LOS path, PRCs and RSCs.

in Section II-B, the first three terms of (19) become $\left[F_p^T(\mathbf{\Omega}_{pq}^{\mathrm{LOS}}(t),\mathbf{\Psi}_p^T(t))\right]^{\mathrm{T}} e^{j\Phi_{\mathrm{LOS}}} \left[F_q^R(\mathbf{\Omega}_{pq}^{\mathrm{LOS}}(t),\mathbf{\Psi}_q^R(t))\right]$. Position vectors $\mathbf{\Omega}_{\mathrm{LOS}}(t)$ and $\mathbf{\Omega}_{pq}^{\mathrm{LOS}}(t)$ are calculated as

$$\begin{aligned}\mathbf{\Omega}_{\mathrm{LOS}}(t) &= \hat{\mathbf{x}}\cdot v(t-t_0)+\hat{\mathbf{y}}\cdot d_{TR}-\hat{\mathbf{z}}\cdot L_T,\\ \mathbf{\Omega}_{pq}^{\mathrm{LOS}}(t) &= -\mathbf{\Psi}_p^T(t)+\mathbf{\Omega}_{\mathrm{LOS}}(t)+\mathbf{\Psi}_q^R(t)\end{aligned} \tag{20}$$

where $(\hat{\mathbf{x}},\hat{\mathbf{y}},\hat{\mathbf{z}})$ is the unit vector of GCS. Φ_{LOS}^{VV} and Φ_{LOS}^{HH} are initial phases in the corresponding direction of polarization with the uniform distribution on $[0,2\pi)$. $P_{\mathrm{LOS}}(t)$ is the received power calculated based on the log-distance model

$$P_{\mathrm{LOS}}(t) = 10^{(P_{\mathrm{Tx}}-PL_{\mathrm{LOS}}(d_{\mathrm{ref}}))/10-n_{\mathrm{LOS}}\lg(\frac{\left\|\mathbf{\Omega}_{pq}^{\mathrm{LOS}}(t)\right\|_{\mathrm{F}}}{d_{\mathrm{ref}}})} \tag{21}$$

where P_{Tx} is the transmitted power in dBW, n_{LOS} is the path loss exponent (PLE) of the LOS path, d_{ref} is the reference distance, and $PL_{\mathrm{LOS}}(d_{\mathrm{ref}})$ is the path loss at d_{ref}. The term $-2\pi\frac{\left\|\mathbf{\Omega}_{pq}^{\mathrm{LOS}}(t)\right\|_{\mathrm{F}}}{\lambda}$ is the phase deviation caused by propagation distances. At last, the Doppler frequency is given by $\nu_{pq}^{\mathrm{LOS}}(t) = \frac{1}{\lambda}\frac{\langle\mathbf{\Omega}_{pq}^{\mathrm{LOS}}(t),\mathbf{v}\rangle}{\left\|\mathbf{\Omega}_{pq}^{\mathrm{LOS}}(t)\right\|_{\mathrm{F}}}$ where $\langle\cdot,\cdot\rangle$ denotes the inner product and $\mathbf{v}$ is the speed vector of the train.

As observed in the results of the CTI algorithm, the survival time of a PRC is almost symmetric with respect to the corresponding power supply pillar. Therefore, we assume that each pillar results in a cylindrical area where Rx is effected by its reflected power, and we define this area as an effective scattering area (ESA) which merely lasts for a period of time and overlaps with adjacent ones. As such, the intersecting plane between the ESA and the railway is a circle with radius R, and the intersecting angle between the circle and the rail is $\alpha = \arcsin(d_P/R)$. The dominant ray with delay $\tau_{pq}^{n_p}(t) = (\left\|\mathbf{\Omega}_{pT}^{n_p}(t)\right\|_{\mathrm{F}} + \left\|\mathbf{\Omega}_{qR}^{n_p}(t)\right\|_{\mathrm{F}})/c$ is defined as the ray following geometrical relationships with the maximum power within each cluster, and other rays in the same cluster are generated by the distribution of relative parameter offsets fitted from measured data. We firstly draw the CDFs of the small-scale parameters from measured data in Fig. 10. The total number of rays within one cluster $M_{n_p}(t)$ is modeled as

TABLE III
DEFINITIONS OF MODEL SYMBOLS

Parameter	Definition
$\mathbf{\Psi}_p^T(t), \mathbf{\Psi}_q^R(t)$	Position vectors of the pth Tx and the qth Rx antenna
L_T	Height of Tx relative to Rx in z direction
L_P	Height of pillars relative to Rx in z direction
d_{TR}	Distance from Tx to Rx in y direction
d_P	Distance from pillars to Rx in y direction
∂_o	Distance from the first pillar to Tx in x direction
∂_P	Distance between adjacent pillars in x direction
N_p	Total number of PRCs from one side of pillars
$M_{n_p}(t)$	Total number of rays within the ESA of the n_pth PRC
N_r	Total number of RSCs
$M_{n_r}(t)$	Total number of rays within the ESA of the n_rth RSC
$\mathbf{\Omega}_{\mathrm{LOS}}(t)$	Position vector of the LOS path between Tx and Rx array center
$\mathbf{\Omega}_{pq}^{\mathrm{LOS}}(t)$	Position vector of the LOS path between the pth Tx and the qth Rx antenna
$\mathbf{\Omega}_T^{n_p}(t), \mathbf{\Omega}_R^{n_p}(t)$	Position vectors between the n_pth pillar center and the Tx (Rx) array center
$\mathbf{\Omega}_{pT}^{n_p}(t), \mathbf{\Omega}_{qR}^{n_p}(t)$	Position vectors between the n_pth pillar center and the pth Tx (the qth Rx)
$\mathbf{\Omega}_T^{m_{n_p}}(t), \mathbf{\Omega}_R^{m_{n_p}}(t)$	Position vectors between the n_pth pillar center and the Tx (Rx) array center via the m_{n_p}th ray
$\mathbf{\Omega}_{pT}^{m_{n_p}}(t), \mathbf{\Omega}_{qR}^{m_{n_p}}(t)$	Position vectors between the n_pth pillar center and the pth Tx (the qth Rx) via the m_{n_p}th ray
$\mathbf{\Omega}_T^{n_r}(t), \mathbf{\Omega}_R^{n_r}(t)$	Position vectors between the n_rth RSC and the Tx (Rx) array center
$\mathbf{\Omega}_{pT}^{n_r}(t), \mathbf{\Omega}_{qR}^{n_r}(t)$	Position vectors between the n_rth RSC and the pth Tx (the qth Rx)
$\mathbf{\Omega}_T^{m_{n_r}}(t), \mathbf{\Omega}_R^{m_{n_r}}(t)$	Position vectors between the n_rth RSC and the Tx (Rx) array center via the m_{n_r}th ray
$\mathbf{\Omega}_{pT}^{m_{n_r}}(t), \mathbf{\Omega}_{qR}^{m_{n_r}}(t)$	Position vectors between the n_rth RSC and the pth Tx (the qth Rx) via the m_{n_r}th ray
κ	Cross polarization power ratio

Poisson distribution as depicted in Fig. 10(a), with probability distribution function (PDF)

$$f(M_{n_p}(t)=k;\lambda_M) = \frac{\lambda_M^k}{k!}e^{-\lambda_M} \tag{22}$$

where λ_M is the expectation, and the minimum value of k is 1. The delay offset $\tau_{pq}^{m_{n_p}}(t)$ compared to the dominant ray within a cluster is modeled as exponential distribution as depicted in Fig. 10(b), with PDF

$$f(\tau_{pq}^{m_{n_p}}(t);\mu) = \frac{1}{\mu}e^{-\frac{\tau_{pq}^{m_{n_p}}(t)}{\mu}} \tag{23}$$

where μ is the expectation. Parameters in the distribution are

given in Table IV. The CIR of PRCs is further expressed as

$$\begin{aligned} h_{pq}^{m_{np}}(t) = &\begin{bmatrix} F_{p,V}^{T}(\mathbf{\Omega}_{pT}^{m_{np}}(t), \mathbf{\Psi}_{p}^{T}(t)) \\ F_{p,H}^{T}(\mathbf{\Omega}_{pT}^{m_{np}}(t), \mathbf{\Psi}_{p}^{T}(t)) \end{bmatrix}^{\mathrm{T}} \\ &\cdot \begin{bmatrix} e^{j\Phi_{m_{np}}^{VV}} & \sqrt{\kappa^{-1}} e^{j\Phi_{m_{np}}^{VH}} \\ \sqrt{\kappa^{-1}} e^{j\Phi_{m_{np}}^{HV}} & e^{j\Phi_{m_{np}}^{HH}} \end{bmatrix} \\ &\cdot \begin{bmatrix} F_{q,V}^{R}(\mathbf{\Omega}_{qR}^{m_{np}}(t), \mathbf{\Psi}_{q}^{R}(t)) \\ F_{q,H}^{R}(\mathbf{\Omega}_{qR}^{m_{np}}(t), \mathbf{\Psi}_{q}^{R}(t)) \end{bmatrix} \sqrt{P_{m_{np}}(t)} \\ &\cdot e^{j\left(\Phi_{m_{np}} - 2\pi \frac{\left\|\mathbf{\Omega}_{pq}^{m_{np}}(t)\right\|_{\mathrm{F}}}{\lambda}\right)} e^{j2\pi \nu_{pq}^{m_{np}}(t) t}. \end{aligned} \tag{24}$$

Here, position vectors of dominant rays within a cluster are given by

$$\begin{aligned} \mathbf{\Omega}_{T}^{n_p}(t) &= \hat{x} \cdot (\partial_o + \partial_P([(n_p > N_p) \\ &\quad \cdot \mathrm{REM}(n_p, N_p) + (n_p \le N_p) \cdot n_p] - 1)) \\ &\quad + \hat{y} \cdot (d_{TR} \pm d_P) - \hat{z} \cdot (L_T - L_P/2) \\ &= \left\|\mathbf{\Omega}_{T}^{n_p}(t)\right\|_{\mathrm{F}} \begin{bmatrix} \cos\theta_{T}^{n_p}(t) \cos\varphi_{T}^{n_p}(t) \\ \cos\theta_{T}^{n_p}(t) \sin\varphi_{T}^{n_p}(t) \\ \sin\theta_{T}^{n_p}(t) \end{bmatrix}^{\mathrm{T}}, \\ \mathbf{\Omega}_{R}^{n_p}(t) &= \mathbf{\Omega}_{\mathrm{LOS}}(t) - \mathbf{\Omega}_{T}^{n_p}(t) \\ &= \left\|\mathbf{\Omega}_{R}^{n_p}(t)\right\|_{\mathrm{F}} \begin{bmatrix} \cos\theta_{R}^{n_p}(t) \cos\varphi_{R}^{n_p}(t) \\ \cos\theta_{R}^{n_p}(t) \sin\varphi_{R}^{n_p}(t) \\ \sin\theta_{R}^{n_p}(t) \end{bmatrix}^{\mathrm{T}} \end{aligned} \tag{25}$$

where $\mathrm{REM}(\cdot,\cdot)$ is the remainder operation and $\theta_{T}^{n_p}(t)$, $\varphi_{T}^{n_p}(t)$, $\theta_{R}^{n_p}(t)$, and $\varphi_{R}^{n_p}(t)$ denote the time-variant elevation of departure (EOD), the azimuth of departure (AOD), the elevation of arrival (EOA), and AOA of the dominant ray in the n_pth PRC, respectively. Angular information of dominant rays can be obtained from (25). For instance, according to our measurement configurations, AOAs are calculated as $\varphi_{R}^{n_p}(t) = \arccos\left[\frac{\eta^2 - R\eta\cos\alpha}{\eta\sqrt{R^2+\eta^2-2R\eta\cos\alpha}}\right]$ where $\eta = v\,(t - t_0) - \partial_o - \partial_P\,(n_P - 1) + R\cos\alpha$. Within a cluster, the AOAs of other rays are calculated based on $\varphi_{R}^{m_{np}}(t) = \varphi_{R}^{n_p}(t) - \Delta\varphi_{R}^{m_{np}}(t)$, and the AOA offset $\Delta\varphi_{R}^{m_{np}}(t)$ is modeled as zero-mean Laplace distribution as depicted in Fig. 10(c), with PDF

$$f(\Delta\varphi_{R}^{m_{np}}(t); b) = \frac{1}{2b} e^{-\frac{|\Delta\varphi_{R}^{m_{np}}(t)|}{b}} \tag{26}$$

where $\sqrt{2}b$ is the standard deviation. Then, the position vectors of rays are given by

$$\begin{aligned} \mathbf{\Omega}_{T}^{m_{np}}(t) &= \left\|\mathbf{\Omega}_{T}^{n_p}(t)\right\|_{\mathrm{F}} \begin{bmatrix} \cos\theta_{T}^{m_{np}}(t) \cos\varphi_{T}^{m_{np}}(t) \\ \cos\theta_{T}^{m_{np}}(t) \sin\varphi_{T}^{m_{np}}(t) \\ \sin\theta_{T}^{m_{np}}(t) \end{bmatrix}^{\mathrm{T}}, \\ \mathbf{\Omega}_{R}^{m_{np}}(t) &= \left\|\mathbf{\Omega}_{R}^{n_p}(t)\right\|_{\mathrm{F}} \begin{bmatrix} \cos\theta_{R}^{m_{np}}(t) \cos\varphi_{R}^{m_{np}}(t) \\ \cos\theta_{R}^{m_{np}}(t) \sin\varphi_{R}^{m_{np}}(t) \\ \sin\theta_{R}^{m_{np}}(t) \end{bmatrix}^{\mathrm{T}}, \\ \mathbf{\Omega}_{pT}^{m_{np}}(t) &= -\mathbf{\Psi}_{p}^{T}(t) + \mathbf{\Omega}_{T}^{m_{np}}(t), \\ \mathbf{\Omega}_{qR}^{m_{np}}(t) &= \mathbf{\Psi}_{q}^{R}(t) + \mathbf{\Omega}_{R}^{m_{np}}(t) \end{aligned} \tag{27}$$

where $\theta_{T}^{m_{np}}(t)$, $\varphi_{T}^{m_{np}}(t)$, $\theta_{R}^{m_{np}}(t)$, and $\varphi_{R}^{m_{np}}(t)$ denote the time-variant EOD, AOD, EOA, and AOA of the m_{n_p}th ray in the n_pth PRC, respectively.

Moreover, the terms $\Phi_{m_{np}}^{VV}$, $\Phi_{m_{np}}^{VH}$, $\Phi_{m_{np}}^{HV}$, and $\Phi_{m_{np}}^{HH}$ in (24) denote initial phases in four polarizations. $\Phi_{m_{np}} - 2\pi\frac{\left\|\mathbf{\Omega}_{pq}^{m_{np}}(t)\right\|_{\mathrm{F}}}{\lambda}$ is the phase deviation caused by propagation distances. The Doppler frequency can be calculated by $\nu_{pq}^{m_{np}}(t) = \frac{1}{\lambda}\frac{\left\langle \mathbf{\Omega}_{qR}^{m_{np}}(t), \mathbf{v}\right\rangle}{\left\|\mathbf{\Omega}_{qR}^{m_{np}}(t)\right\|_{\mathrm{F}}}$. The power is given by

$$P_{m_{np}}(t) = \rho \cdot 10^{-n_{\mathrm{PRC}} \lg\left(\frac{\left\|\mathbf{\Omega}_{T}^{n_p}(t)\right\|_{\mathrm{F}} + \left\|\mathbf{\Omega}_{R}^{n_p}(t)\right\|_{\mathrm{F}}}{d_{\mathrm{ref}}}\right)}{}^{(P_{\mathrm{Tx}} - PL_{\mathrm{PRC}}(d_{\mathrm{ref}}) - \mathrm{X}_{\mathrm{PRC}})/10} \tag{28}$$

where n_{PRC} is the PLE of PRCs, $\mathrm{X}_{\mathrm{PRC}} \sim N(0, \sigma_{\mathrm{PRC}}^2)$ is a zero-mean Gaussian random variable representing the shadowing effect, and ρ is the power proportional factor defined as the ratio of each ray power to the total power within a cluster. It is modeled using Weibull distribution as depicted in Fig. 10(d), with PDF

$$f(\rho; A, B) = \frac{B}{A}\left(\frac{\rho}{A}\right)^{B-1} e^{-\left(\frac{\rho}{A}\right)^{B}} \tag{29}$$

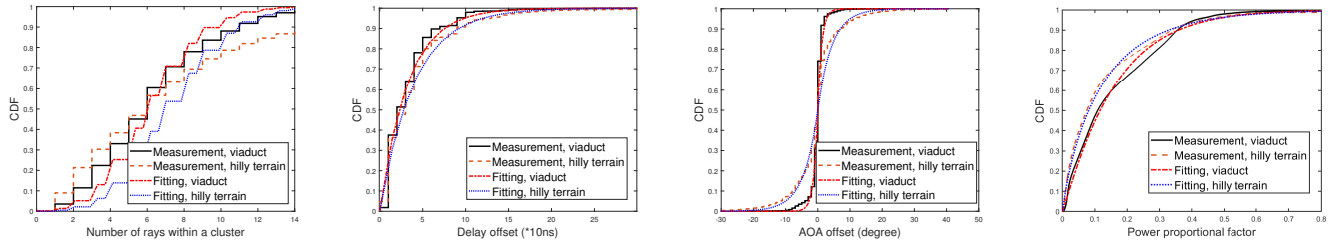
where A is the scale factor and B is the shape factor. To sum up, we characterize the inter-cluster parameters, i.e., delays and angles of dominant rays, based on the geometric relationships, and the intra-cluster parameters, i.e. number of rays, delay offsets, angular offsets, and power of rays, based on different empirical distributions. As shown in Fig. 10, good agreements can be observed between fitting distributions and the measured data.

Furthermore, as for the RSCs, single-bounce scattering is assumed in the viaduct and hilly terrain scenarios, and the scatterer center corresponding to the n_rth RSC locates at $(x_{n_r}, y_{n_r}, z_{n_r})$ in the GCS, following the uniform distribution

$$\begin{aligned} x_{n_r} &\sim U\left(0, x_{\max}^{\mathrm{bon}}\right], \\ y_{n_r} &\sim U\left[y_{\min}^{\mathrm{bon}}, d_{TR} - d_P\right), \\ z_{n_r} &\sim U\left(-L_T, z_{\max}^{\mathrm{bon}}\right] \end{aligned} \tag{30}$$

where $x_{\max}^{\mathrm{bon}}$, $y_{\min}^{\mathrm{bon}}$, and $z_{\max}^{\mathrm{bon}}$ represent the maximum boundary of x_{n_r}, the minimum boundary of y_{n_r}, and the maximum boundary of z_{n_r}, respectively. It is noted that due to the blockage of carriage, only one side of the signal can be received in the measurement. More generally, $y_{n_r} \sim U\left[y_{\min}^{\mathrm{bon}}, d_{TR} - d_P\right) \cup \left(d_{TR} + d_P, y_{\max}^{\mathrm{bon}}\right]$, where $y_{\max}^{\mathrm{bon}}$ is the maximum boundary of y_{n_r}. The delay of the dominant ray in the n_r RSC is $\tau_{pq}^{n_r}(t) = \frac{\left\|\mathbf{\Omega}_{pT}^{n_r}(t)\right\|_{\mathrm{F}} + \left\|\mathbf{\Omega}_{qR}^{n_r}(t)\right\|_{\mathrm{F}}}{c}$. The CIR of RSCs is further expressed as

$$\begin{aligned} h_{pq}^{m_{nr}}(t) = &\begin{bmatrix} F_{p,V}^{T}(\mathbf{\Omega}_{pT}^{m_{nr}}(t), \mathbf{\Psi}_{p}^{T}(t)) \\ F_{p,H}^{T}(\mathbf{\Omega}_{pT}^{m_{nr}}(t), \mathbf{\Psi}_{p}^{T}(t)) \end{bmatrix}^{\mathrm{T}} \\ &\cdot \begin{bmatrix} e^{j\Phi_{m_{nr}}^{VV}} & \sqrt{\kappa^{-1}} e^{j\Phi_{m_{nr}}^{VH}} \\ \sqrt{\kappa^{-1}} e^{j\Phi_{m_{nr}}^{HV}} & e^{j\Phi_{m_{nr}}^{HH}} \end{bmatrix} \\ &\cdot \begin{bmatrix} F_{q,V}^{R}(\mathbf{\Omega}_{qR}^{m_{nr}}(t), \mathbf{\Psi}_{q}^{R}(t)) \\ F_{q,H}^{R}(\mathbf{\Omega}_{qR}^{m_{nr}}(t), \mathbf{\Psi}_{q}^{R}(t)) \end{bmatrix} \sqrt{P_{m_{nr}}(t)} \\ &\cdot e^{j\left(\Phi_{m_{nr}} - 2\pi \frac{\left\|\mathbf{\Omega}_{pq}^{m_{nr}}(t)\right\|_{\mathrm{F}}}{\lambda}\right)} e^{j2\pi \nu_{pq}^{m_{nr}}(t) t} \end{aligned} \tag{31}$$



(a) (b) (c) (d)

Fig. 10. CDFs of intra-cluster parameters from measured data and fitting distributions for PRCs in two scenarios. (a) Number of rays with Poisson distribution; (b) delay offset with exponential distribution; (c) AOA offset with zero-mean Laplace distribution; and (d) power proportional factor with Weibull distribution.

TABLE IV
VALUES OF INTRA-CLUSTER PARAMETERS

	λ_M	μ	b	A	B
Viaduct	6.25	3.27	1.16	0.16	1.07
Hilly terrain	7.41	3.99	4.35	0.12	0.86

where the meanings of notations are the same with those in (24), except that the subscript n_r denotes the n_rth ray within a RSC. The position vectors of dominant rays are given by

$$\begin{aligned}\boldsymbol{\Omega}_T^{n_r}(t) &= \hat{x}\cdot x_{n_r} + \hat{y}\cdot y_{n_r} + \hat{z}\cdot z_{n_r} \\ &= \left\|\boldsymbol{\Omega}_T^{n_r}(t)\right\|_{\mathrm{F}} \begin{bmatrix} \cos\theta_T^{n_r}(t)\cos\varphi_T^{n_r}(t) \\ \cos\theta_T^{n_r}(t)\sin\varphi_T^{n_r}(t) \\ \sin\theta_T^{n_r}(t) \end{bmatrix}^{\mathrm{T}}, \\ \boldsymbol{\Omega}_R^{n_r}(t) &= \boldsymbol{\Omega}_{\mathrm{LOS}}(t) - \boldsymbol{\Omega}_T^{n_r}(t) \\ &= \left\|\boldsymbol{\Omega}_R^{n_r}(t)\right\|_{\mathrm{F}} \begin{bmatrix} \cos\theta_R^{n_r}(t)\cos\varphi_R^{n_r}(t) \\ \cos\theta_R^{n_r}(t)\sin\varphi_R^{n_r}(t) \\ \sin\theta_R^{n_r}(t) \end{bmatrix}^{\mathrm{T}}. \end{aligned} \tag{32}$$

Then, position vectors of rays can be calculated by (27). Also, the intra-cluster small-scale channel parameters, i.e., the number of rays within a cluster, the delay offset, the angular offset, and the power proportional factor, can be generated by (22), (23), (26), and (29), respectively.

V. MODEL VALIDATION

In this section, the proposed model is verified with channel statistics which characterize the non-stationarity of HST radio channels. Based on the channel modeling method aforementioned, simulations are conducted with the system configuration same as the measurement campaign described in Section II., and relevant parameters are set as follows if not specified otherwise. L_T is 20 m in viaduct and 30 m in hilly terrain. Values of L_P, d_P, d_{TR}, and ∂_P in both scenarios are 15 m, 1.5 m, 50 m, and 50 m, respectively. The radius of the ESA is calculated as 68 m by averaging all the detected PRCs in both scenarios. The value of ∂_o is 57.44 m in viaduct and 161.88 m in hilly terrain based on the identified results. The CIRs are calculated from (18) and discretized to the form of $h(i\Delta t, p\Delta\tau)$. Then, the APDP and the CTF are calculated using (3) and (7), respectively.

A. Statistical Properties

We validate the accuracy of the proposed model by the Rician K factor and the RMS delay spread. The Rician K factor is defined as the power ratio of the LOS path to non-line-of-sight (NLOS) paths containing the PRCs and RSCs

$$\begin{aligned}\mathrm{KF}(t) &= 10\lg\left(\frac{P_{\mathrm{LOS}}(t)}{P_{\mathrm{NLOS}}(t)}\right) \\ &= 10\lg\left(\frac{P_{\mathrm{LOS}}(t)}{\sum\limits_{m_{np}=1}^{M_{np}(t)} P_{m_{np}}(t) + \sum\limits_{m_{nr}=1}^{M_{nr}(t)} P_{m_{nr}}(t)}\right)\end{aligned} \tag{33}$$

where the power are calculated from (21) and (28). We rewrite their delays and powers as $\left\{\tau_1^{\mathrm{MPC}}, \tau_2^{\mathrm{MPC}}, ..., \tau_{L(t)}^{\mathrm{MPC}}\right\}$ and $\left\{P_1^{\mathrm{MPC}}, P_2^{\mathrm{MPC}}, ..., P_{L(t)}^{\mathrm{MPC}}\right\}$, and the RMS delay spread is defined as

$$\sigma_\tau(t) = \sqrt{\frac{\sum\limits_{l=1}^{L(t)} P_l^{\mathrm{MPC}}\cdot\left(\tau_l^{\mathrm{MPC}} - \bar{\tau}^{\mathrm{MPC}}(t)\right)^2}{\sum\limits_{l=1}^{L(t)} P_l^{\mathrm{MPC}}}} \tag{34}$$

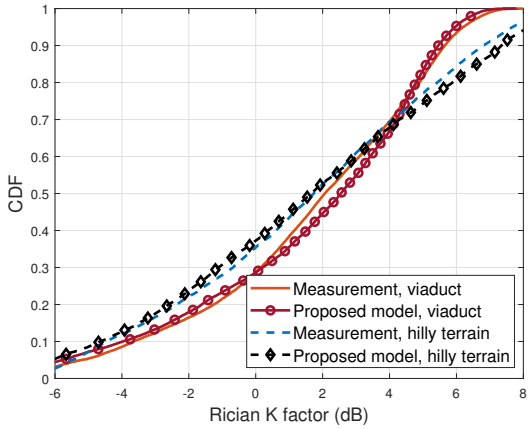
where $\bar{\tau}^{\mathrm{MPC}}(t) = \sum\limits_{l=1}^{L(t)} P_l^{\mathrm{MPC}}\cdot\tau_l^{\mathrm{MPC}} \Big/ \sum\limits_{l=1}^{L(t)} P_l^{\mathrm{MPC}}$ represents the average delay.

Moreover, in order to illustrate the non-stationarity of the model, the correlation property is characterized by the space-time correlation function (STCF)

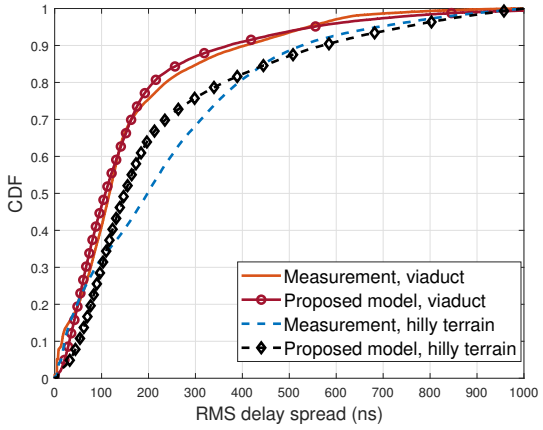
$$\begin{aligned}&\mathrm{STCF}(i\Delta t, j\Delta t, (p'-p)\Delta d_p, (q'-q)\Delta d_q) \\ &= \mathrm{E}\left[h(i\Delta t, p\Delta d_p, q\Delta d_q)\cdot h^*((i+j)\Delta t, p'\Delta d_p, q'\Delta d_q)\right]\end{aligned} \tag{35}$$

where Δd_p and Δd_q are the spacing between adjacent antennas at Tx and Rx. $h(i\Delta t, p\Delta d_p, q\Delta d_q)$ is the CIR between the pth Tx antenna and the qth Rx antenna at moment $i\Delta t$, and $h((i+j)\Delta t, p'\Delta d_p, q'\Delta d_q)$ is the CIR between the p'th Tx antenna and the q'th Rx antenna at moment $(i+j)\Delta t$. $(\cdot)^*$ is the complex conjugate operator. By setting $p = p'$ and $q = q'$, the local temporal ACF, which describes the non-stationarity in the time domain, is derived as

$$\mathrm{ACF}(i\Delta t, j\Delta t) = \mathrm{E}\left[h(i\Delta t)\cdot h^*((i+j)\Delta t)\right]. \tag{36}$$



(a)



(b)

Fig. 11. Comparison of (a) Rician K factor and (b) RMS delay spread between the measured data and the proposed model.

By setting the time difference $j\Delta t$ as zero, the spatial CCF, which describes the non-stationarity in the spatial array domain, is written as

$$\begin{aligned}&\mathrm{CCF}(i\Delta t,(p'-p)\Delta d_p,(q'-q)\Delta d_q)\\&=\mathrm{E}\left[h(i\Delta t,p\Delta d_p,q\Delta d_q)\cdot h^*(i\Delta t,p'\Delta d_p,q'\Delta d_q)\right]\end{aligned}. \quad (37)$$

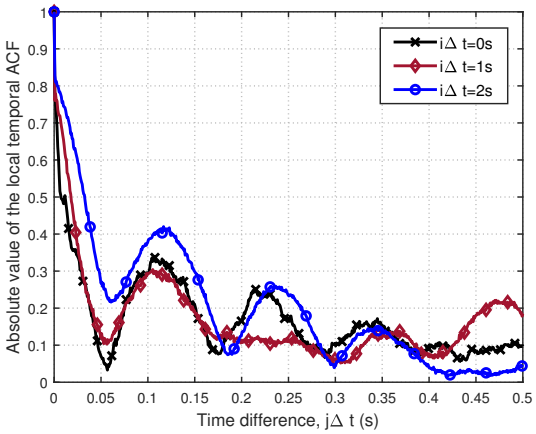
B. Results and Analyses

First, we compare the proposed model and the measured data in terms of the Rician K factor and the RMS delay spread. The comparison of Rician K factor and RMS delay spread are shown in Fig. 11, and values are given in Table V. The mean values and the empirical CDFs from the proposed model are fairly close to those from measurements, which verifies the correctness of the model.

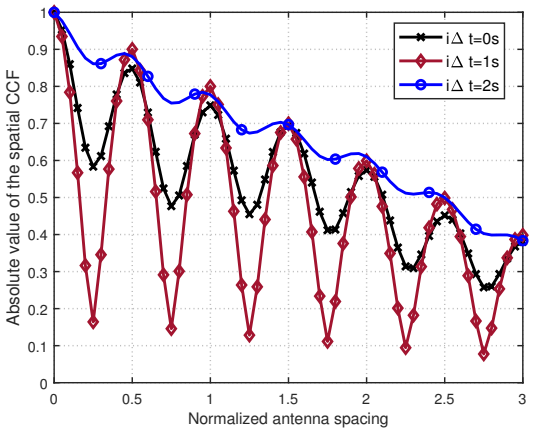
Second, the temporal ACF and the spatial CCF are simulated. Fig. 12(a) shows the absolute values of temporal ACFs at moments 0 s, 1 s, and 2 s, and Fig. 12(b) shows the absolute values of spatial CCFs at moments 0 s, 1 s, and 2 s. The array form is ULAs with antenna spacing of half wavelength. It is demonstrated that the proposed model is non-stationary in the time domain as well as in the spatial domain, for the ACF curves and CCF curves change with time.

TABLE V
COMPARISON OF RICIAN K FACTOR AND RMS DELAY SPREAD OF MEASUREMENT AND THE PROPOSED MODEL IN BOTH SCENARIOS

Viaduct (Hilly terrain)	Mean	50%
K factor, measurement (dB)	1.62 (1.51)	2.07 (1.78)
K factor, proposed model (dB)	1.63 (1.49)	2.56 (1.69)
Delay spread, measurement (ns)	166.6 (237.9)	117.1 (197.5)
Delay spread, proposed model (ns)	162.6(230.6)	108.1 (149.2)



(a)



(b)

Fig. 12. Simulated (a) temporal ACFs and (b) spatial CCFs at different moments.

Third, the quasi-stationary distances of the model are simulated and analyzed. Based on the APDP metric, Fig. 13 compares the CDFs of quasi-stationary distances of the measured data, the proposed model, the standardized IMT-A channel model [61], and the non-stationary IMT-A channel model in [62] for high-mobility scenarios. It should be noted that the standardized IMT-A channel model also introduces the concept of time-evolution to explicitly simulate the non-stationarity of the fading channels [61]. In the non-stationary IMT-A model, all the channel parameters including the delays, powers, AOAs, AODs, and number of clusters change with time. The propagation-related parameters are updated according to geometric relationships. The birth-death process described by

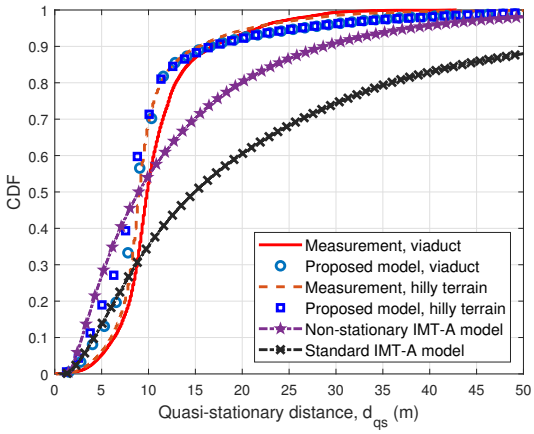


Fig. 13. Quasi-stationary distances that result from the measured data, proposed model, non-stationary IMT-A channel model, and standard IMT-A channel model based on the APDP metric.

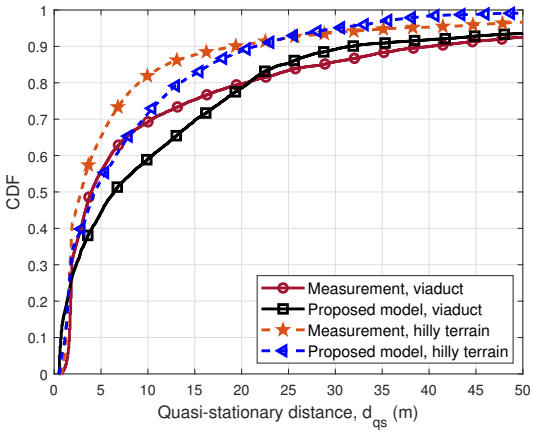


Fig. 14. Quasi-stationary distances that result from the measured data and the proposed model based on the CMD metric.

the Markov process is taken into account to model the change of the number of clusters. To generate channel responses based on the non-stationary IMT-A model, simulations are implemented according to the same parameters as in Section II, and the rural macro-cell (RMa) scenario is selected.

Here, we briefly compare the modeling mechanism of the proposed model and the non-stationary IMT-A channel model. Due to the fact that the non-stationary IMT-A model is based on the standardized IMT-A model, their generation methods of channel parameters are similar. In comparison, the proposed model is based on the assumption that the dominant ray follows the geometric relationship, which coincides with the propagation regularity summarized from the measured results. Inside a cluster, the rays are resolvable with respect to the delay, angle, and power. We deploy the distributions most consistent with the measured results. All the large-scale and small-scale channel parameters are time-variant in the proposed model. Moreover, in the non-stationary IMT-A model, clusters are assumed to be modeled in the same manner, and it is not designed specifically for high-mobility communication conditions. By contrast, inspired by the regularity of the measured data in HST channels, we separate MPCs into the LOS path, the PRCs, and the RSCs, and model them respectively with parameters directly obtained from measurements.

As for the proposed model, the mean value of the quasi-stationary distance is 10.56 m in viaduct and 10.01 m in the hilly terrain scenario, respectively, similar to those of the measurements, i.e. 10.94 m and 10.20 m, and it is 10.06 m for the non-stationary IMT-A model. When it comes to the standardized IMT-A channel model, the corresponding result is 15.74 m, which is apparently larger than the measured data. Moreover, as shown in Fig. 13, the proposed model is much closer to the measured data than the non-stationary IMT-A model. The root mean square error (RMSE) of the CDF curves between the proposed model and the measured data is 0.0595 for viaduct and 0.0568 for hilly terrain. In comparison, the RMSE values are 0.1156 and 0.1207 for the non-stationary IMT-A model in the two scenarios. Therefore, the proposed model performs better than the non-stationary IMT-A model in characterizing the non-stationarity in the power-delay domain in HST channels.

Additionally, the comparison between the proposed model and the measured data on quasi-stationary distances using the CMD metric is depicted in Fig. 14. The mean value of the quasi-stationary distance of the proposed model is 14.39 m in viaduct and 8.39 m in hilly terrain scenario, similar to those of the measurements, i.e. 13.82 m and 7.95 m, respectively. The RMSE of the CDF curves is 0.0642 in viaduct and 0.0616 in hilly terrain scenario, respectively. Therefore, the proposed model can characterize the non-stationarity in both the power-delay domain and the spatial domain.

At last, the simulation time is taken as metric to measure the complexity. The simulation time of generating CIRs for 5000 snapshots is recorded. We test on a computer with Core i7 CPU, 8GB RAM, and 512 GB hard disk drive. The operation system is Windows 10. As a result, the simulation time is 0.369 s for the standardized IMT-A model, 5.383 s for the non-stationary IMT-A model, and 8.727 s for the proposed model. It is demonstrated that the accuracy is improved and the complexity is maintained on the same order of magnitude compared with the non-stationary IMT-A model.

VI. Conclusions

Measurement campaigns of HST radio channels were conducted in viaduct and hilly terrain scenarios along the Guangzhou-Shenzhen Express Rail Link. The non-stationarity in HST radio channels was characterized and modeled based on the measured data. Two kinds of metrics including the APDP and CMD were considered to quantify the size of the quasi-stationary region. It was shown that the mean value of quasi-stationary distances in the propagation-wise hilly terrain was smaller than that in the viaduct scenario, whether APDP or CMD was selected as the metric. Parameters of MPCs were estimated using the SAGE algorithm and an automatic CTI algorithm was designed to classify MPCs into the LOS path, PRCs, and RSCs. Based on the statistical properties of clusters extracted from measurements and the geometrical

relationships, a non-stationary model for HST radio channels was proposed and implemented. The model was validated by comparing channel statistics with the measured results, and good agreements were observed. The temporal ACFs and the spatial CCFs of different times were presented. Compared with the quasi-stationary distances of the standardized IMT-A channel model and the non-stationary IMT-A channel model, the proposed model showed a better agreement with the measured data and had a good ability to characterize the non-stationary features of propagation environments in HST scenarios. For future research directions, multi-scenario, multi-band, and multi-link would be considered. In addition, big data-enabled non-stationarity characterization and modeling is also an important trend.

REFERENCES

[1] J. Bang, S. Oh, K. Kang, and Y. Cho, "A bayesian regression based LTE-R handover decision algorithm for high-speed railway systems," *IEEE Trans. Veh. Technol.*, vol. 68, no. 10, pp. 10160–10173, Oct. 2019.

[2] T. Zhou, H. Li, Y. Wang, L. Liu, and C. Tao, "Channel modeling for future high-speed railway communication systems: A survey," *IEEE Access*, vol. 7, pp. 52818–52826, 2019.

[3] G. Noh, B. Hui, and I. Kim, "High speed train communications in 5G: Design elements to mitigate the impact of very high mobility," *IEEE Wireless Commun.*, vol. 27, no. 6, pp. 98–106, Dec. 2020.

[4] Z. Huang and X. Cheng, "A 3-D non-stationary model for beyond 5G and 6G vehicle-to-vehicle mmWave massive MIMO channels," *IEEE Trans. Intell. Transp. Syst.*, vol. 23, no. 7, pp. 8260–8276, Jul. 2022.

[5] X. Cheng, Z. Huang, S. Chen, "Vehicular communication channel measurement, modelling, and application for beyond 5G and 6G," *IET Commun.*, vol. 14, no. 19, pp. 3303–3311, 2020.

[6] C. Briso, C. Calvo, Z. Cui, L. Zhang, and Y. Xu, "Propagation measurements and modeling for low altitude UAVs from 1 to 24 GHz," *IEEE Trans. Veh. Technol.*, vol. 69, no. 3, pp. 3439–3443, Mar. 2020.

[7] L. Bai, Z. Huang, X. Zhang, and X. Cheng, "A non-stationary 3D model for 6G massive MIMO mmWave UAV channels," *IEEE Trans. Wireless Commun.*, vol. 21, no. 6, pp. 4325–4339, Jun. 2022.

[8] Y. Liu, C. -X. Wang, and J. Huang, "Recent developments and future challenges in channel measurements and models for 5G and beyond high-speed train communication systems," *IEEE Commun. Mag.*, vol. 57, no. 9, pp. 50–56, Sep. 2019.

[9] J. Bian, C. -X. Wang, Y. Liu, J. Tian, J. Qiao, and X. Zheng, "3D non-stationary wideband UAV-to-ground MIMO channel models based on aeronautic random mobility model," *IEEE Trans. Veh. Technol.*, vol. 70, no. 11, pp. 11154–11168, Nov. 2021.

[10] Y. Bi, J. Zhang, Q. Zhu, W. Zhang, L. Tian, and P. Zhang, "A novel non-stationary high-speed train (HST) channel modeling and simulation method," *IEEE Trans. Veh. Technol.*, vol. 68, no. 1, pp. 82–92, Jan. 2019.

[11] Z. Huang, X. Cheng, and X. Yin, "A general 3D non-stationary 6G channel model with time-space consistency," *IEEE Trans. Commun.*, vol. 70, no. 5, pp. 3436–3450, May 2022.

[12] Z. Huang and X. Cheng, "A general 3D space-time-frequency non-stationary model for 6G channels," *IEEE Trans. Wireless Commun.*, vol. 20, no. 1, pp. 535–548, Jan. 2021.

[13] M. Paetzold and C. A. Gutierrez, "Definition and analysis of quasi-stationary intervals of mobile radio channels - Invited paper," in *Proc. IEEE 87th Veh. Technol. Conf. (VTC-Spring)*, 2018, pp. 1–6.

[14] Z. Cui, Z. Zhong, K. Guan, and D. He, "An acceleration method for ray-tracing simulation based on channel quasi-stationarity regions," in *Proc. 12th Eur. Conf. Antennas Propag. (EuCAP)*, 2018, pp. 1–5.

[15] R. He, O. Renaudin, V. Kolmonen, K. Haneda, Z. Zhong, B. Ai, and C. Oestges, "Characterization of quasi-stationarity regions for vehicle-to-vehicle radio channels," *IEEE Trans. Antennas Propag.*, vol. 63, no. 5, pp. 2237–2251, May 2015.

[16] M. Yang *et al.*, "Non-stationary vehicular channel characterization in complicated scenarios," *IEEE Trans. Veh. Technol.*, vol. 70, no. 9, pp. 8387–8400, Sep. 2021.

[17] Y. Tan, C. -X. Wang, J. O. Nielsen, G. F. Pedersen, and Q. Zhu, "A novel B5G frequency nonstationary wireless channel model," *IEEE Trans. Antennas Propag.*, vol. 69, no. 8, pp. 4846–4860, Aug. 2021.

[18] J. Yang *et al.*, "Quasi-stationarity regions analysis for channel in composite high-speed railway scenario," in *Proc. IEEE Int. Symp. Antennas Propag. (ISAP)*, 2016, pp. 1699–1700.

[19] S. Pratschner, T. Blazek, H. Groll, S. Caban, S. Schwarz, and M. Rupp, "Measured user correlation in outdoor-to-indoor massive MIMO scenarios," *IEEE Access*, vol. 8, pp. 178269–178282, 2020.

[20] V. Nikolaidis, N. Moraitis, and A. G. Kanatas, "Dual-polarized narrow-band MIMO LMS channel measurements in urban environments," *IEEE Trans. Antennas Propag.*, vol. 65, no. 2, pp. 763–774, Feb. 2017.

[21] R. He *et al.*, "Non-stationarity characterization for vehicle-to-vehicle channels using correlation matrix distance and shadow fading correlation," in *Proc. 35th Prog. Electromagn. Res. Symp. (PIERS)*, 2014, pp. 1–5.

[22] M. Yusuf *et al.*, "Experimental study on the impact of antenna characteristics on non-stationary V2I channel parameters in tunnels," *IEEE Trans. Veh. Technol.*, vol. 69, no. 11, pp. 12396–12407, Nov. 2020.

[23] M. Yusuf *et al.*, "Stationarity analysis of V2I radio channel in a suburban environment," *IEEE Trans. Veh. Technol.*, vol. 68, no. 12, pp. 11532–11542, Dec. 2019.

[24] L. Bernado, T. Zemen, F. Tufvesson, A. F. Molisch, and C. F. Mecklenbrauker, "The (in-) validity of the WSSUS assumption in vehicular radio channels," in *Proc. 23rd Inter. Symp. Pers. Indoor Mobile Radio Commun. (PIMRC)*, 2012, pp. 1757–1762.

[25] ITU-R P.1411-8, "Propagation data and prediction methods for the planning of short-range outdoor radio communication systems and radio local area networks in the frequency range 300 MHz to 100 GHz," Geneva, Switzerland, Rep. ITU-R P.1411-8, Aug. 2015.

[26] J. Kim *et al.*, "A comprehensive study on mmWave-based mobile hotspot network system for high-speed train communications," *IEEE Trans. Veh. Technol.*, vol. 68, no. 3, pp. 2087–2101, Mar. 2019.

[27] M. -S. Pan, T. -M. Lin, and W. -T. Chen, "An enhanced handover scheme for mobile relays in LTE-A high-speed rail networks," *IEEE Trans. Veh. Technol.*, vol. 64, no. 2, pp. 743–756, Feb. 2015.

[28] D. Yu *et al.*, "Empirical study on directional millimeter-wave propagation in railway communications between train and trackside," *IEEE J. Sel. Areas Commun.*, vol. 38, no. 12, pp. 2931–2945, Dec. 2020.

[29] E. P. Simon and F. Kaltenberger, "Intercarrier interference analysis for high-speed trains based on broadband channel measurements," *IEEE Antennas Wireless Propag. Lett.*, vol. 15, pp. 1502–1505, 2016.

[30] J. -J. Park, J. Lee, K. -W. Kim, and M. -D. Kim, "Large- and small-scale fading characteristics of mmWave HST propagation channel based on 28-GHz measurements," in *Proc. 15th Eur. Conf. Antennas Propag. (EuCAP)*, 2021, pp. 1–5.

[31] T. Zhou, C. Tao, and L. Liu, "LTE-assisted multi-link MIMO channel characterization for high-speed train communication systems," *IEEE Trans. Veh. Technol.*, vol. 68, no. 3, pp. 2044–2051, Mar. 2019.

[32] A. Ispas, C. Schneider, G. Ascheid, and R. Thoma, "Analysis of the local quasi-stationarity of measured dual-polarized MIMO channels," *IEEE Trans. Veh. Technol.*, vol. 64, no. 8, pp. 3481–3493, Aug. 2015.

[33] J. Yang *et al.*, "An efficient MIMO channel model for LTE-R network in high-speed train environment," *IEEE Trans. Veh. Technol.*, vol. 68, no. 4, pp. 3189–3200, Apr. 2019.

[34] H. Qiu *et al.*, "Emulation of radio technologies for railways: A tapped-delay-line channel model for tunnels," *IEEE Access*, vol. 9, pp. 1512–1523, 2021.

[35] T. Zhou, C. Tao, S. Salous, and L. Liu, "Joint channel characteristics in high-speed railway multi-link propagation scenarios: Measurement, analysis, and modeling," *IEEE Trans. Intell. Transp. Syst.*, vol. 20, no. 6, pp. 2367–2377, Jun. 2019.

[36] K. Guan *et al.*, "5G channel models for railway use cases at mmWave band and the path towards terahertz," *IEEE Intell. Transp. Syst. Mag.*, vol. 13, no. 3, pp. 146–155, Fall 2021.

[37] J. Yang *et al.*, "A geometry-based stochastic channel model for the millimeter-wave band in a 3GPP high-speed train scenario," *IEEE Trans. Veh. Technol.*, vol. 67, no. 5, pp. 3853–3865, May 2018.

[38] T. Zhou, Y. Yang, L. Liu, C. Tao, and Y. Liang, "A dynamic 3-D wide-band GBSM for cooperative massive MIMO channels in intelligent high-speed railway communication systems," *IEEE Trans. Wireless Commun.*, vol. 20, no. 4, pp. 2237–2250, Apr. 2021.

[39] Y. Liu, C. -X. Wang, C. F. Lopez, G. Goussetis, Y. Yang, and G. K. Karagiannidis, "3D non-stationary wideband tunnel channel models for 5G high-speed train wireless communications," *IEEE Trans. Intell. Transp. Syst.*, vol. 21, no. 1, pp. 259–272, Jan. 2020.

[40] T. Zhou, C. Tao, S. Salous, and L. Liu, "Geometry-based multi-link channel modeling for high-speed train communication networks," *IEEE Trans. Intell. Transp. Syst.*, vol. 21, no. 3, pp. 1229–1238, Mar. 2020.

[41] Y. Liu, C. -X. Wang, J. Huang, J. Sun, and W. Zhang, "Novel 3-D nonstationary mmWave massive MIMO channel models for 5G high-speed train wireless communications," *IEEE Trans. Veh. Technol.*, vol. 68, no. 3, pp. 2077–2086, Mar. 2019.
[42] X. Yin, X. Cai, X. Cheng, J. Chen, and M. Tian, "Empirical geometry-based random-cluster model for high-speed-train channels in UMTS networks," *IEEE Trans. Intell. Transp. Syst.*, vol. 16, no. 5, pp. 2850–2861, Oct. 2015.
[43] C. Huang *et al.*, "Geometry-cluster-based stochastic MIMO model for vehicle-to-vehicle communications in street canyon scenarios," *IEEE Trans. Wireless Commun.*, vol. 20, no. 2, pp. 755–770, Feb. 2021.
[44] L. Bai, Z. Huang, Y. Li, and X. Cheng, "A 3D cluster-based channel model for 5G and beyond vehicle-to-vehicle massive MIMO channels," *IEEE Trans. Veh. Technol.*, vol. 70, no. 9, pp. 8401–8414, Sep. 2021.
[45] N. Czink, P. Cera, J. Salo, E. Bonek, J. Nuutinen, and J. Ylitalo, "A framework for automatic clustering of parametric MIMO channel data including path powers," in *Proc. IEEE 64th Veh. Technol. Conf. (VTC-Fall)*, 2006, pp. 1–5.
[46] C. Huang *et al.*, "Artificial intelligence enabled radio propagation for communications—Part I: Channel characterization and antenna-channel optimization," *IEEE Trans. Antennas Propag.*, vol. 70, no. 6, pp. 3939–3954, Jun. 2022.
[47] G. Sun, R. He, B. Ai, C. Huang, and Z. Zhong, "Dynamic clustering of multipath components for time-varying propagation channels," *IEEE Trans. Veh. Technol.*, vol. 70, no. 12, pp. 13396–13400, Dec. 2021.
[48] Y. Zhang *et al.*, "A general coupling-based model framework for wideband MIMO channels," *IEEE Trans. Antennas Propag.*, vol. 60, no. 2, pp. 574–586, Oct. 2012.
[49] R. Ma, J. Cao, D. Feng, H. Li, and S. He, "FTGPHA: Fixed-trajectory group pre-handover authentication mechanism for mobile relays in 5G high-speed rail networks," *IEEE Trans. Veh. Technol.*, vol. 69, no. 2, pp. 2126–2140, Feb. 2020.
[50] J. Zhang, H. Du, P. Zhang, J. Cheng, and L. Yang, "Performance analysis of 5G mobile relay systems for high-speed trains," *IEEE J. Sel. Areas Commun.*, vol. 38, no. 12, pp. 2760–2772, Dec. 2020.
[51] X. Cai, X. Yin, and A. P. Yuste, "Direction-of-arrival estimation using single antenna in high-speed-train environments," in *Proc. 10th Eur. Conf. Antennas Propag. (EuCAP)*, 2016, pp. 1–4.
[52] T. Zhou, C. Tao, S. Salous, and L. Liu, "Spatial characterization for high-speed railway channels based on moving virtual array measurement scheme," *IEEE Antennas Wireless Propag. Lett.*, vol. 16, pp. 1423–1426, Aug. 2016.
[53] F. Zhang, W. Fan, J. Zhang, and G. F. Pedersen, "Virtual large-scale array beamforming analysis using measured subarray antenna patterns," *IEEE Access*, vol. 5, pp. 19812–19823, 2017.
[54] T. S. Rappaport, *Wireless Communications: Principles and Practice. 2nd ed.* NJ, USA: Prentice-Hall, 2002, pp. 163–165.
[55] A. Gehring, M. Steinbauer, I. Gaspard, and M. Grigat, "Empirical channel stationarity in urban environments," in *Proc. Eur. Pers. Mobile Commun. Conf. (EPMCC'01)*, 2001, pp. 1–6.
[56] B. H. Fleury, M. Tschudin, R. Heddergott, D. Dahlhaus, and K. Ingeman Pedersen, "Channel parameter estimation in mobile radio environments using the SAGE algorithm," *IEEE J. Sel. Areas Commun.*, vol. 17, no. 3, pp. 434–450, Mar. 1999.
[57] B. H. Fleury, P. Jourdan, and A. Stucki, "High-resolution channel parameter estimation for MIMO applications using the SAGE algorithm," in *Proc. Int. Zurich Seminar Broadband Commun. (IZSBC)*, 2002, pp. 1–9.
[58] M. Ester, H.-P. Kriegel, J. Sander, and X. Xu, "A density-based algorithm for discovering clusters in large spatial databases with noise", in *Proc. Int. Conf. Knowl. Discovery and Data Mining (KDD)*, 1996, pp. 226–231.
[59] N. Czink, C. Mecklenbrauker, and G. Del Galdo, "A novel automatic cluster tracking algorithm," in *Proc. IEEE 17th Int. Symp. Pers. Indoor Mobile Radio Commun. (PIMRC)*, 2006, pp. 1–5.
[60] S. Wu, C. -X. Wang, e. -H. M. Aggoune, M. M. Alwakeel, and X. You, "A general 3-D non-stationary 5G wireless channel model," *IEEE Trans. Commun.*, vol. 66, no. 7, pp. 3065–3078, Jul. 2018.
[61] ITU-R M.2135-1, "Guidelines for evaluation of radio interface technologies for IMT-Advanced," Geneva, Switzerland, Rep. ITU-R M.2135-1, Dec. 2009.
[62] A. Ghazal, Y. Yuan, C.-X. Wang, Y. Zhang, H. Zhou, and W. Duan, "A non-stationary IMT-Advanced MIMO channel model for high-mobility wireless communication systems," *IEEE Trans. Wireless Commun.*, vol. 16, no. 4, pp. 2057–2068, Apr. 2017.

Yan Zhang (Member, IEEE) received the B.S. degree in information engineering from the Beijing Institute of Technology, Beijing, China, in 2005 and the Ph.D. degree in information and communication engineering from Tsinghua University, Beijing, China, in 2010. From 2010 to 2013, he was with the Department of Electronic Engineering, Tsinghua University, Beijing, China, as a Post-Doctoral Researcher. From 2014 to 2015, he was a Research Assistant with the School of Engineering and Physical Sciences, Heriot-Watt University, Edinburgh, U.K. He is currently an Associate Professor with the School of Information and Electronics, Beijing Institute of Technology, Beijing, China. His main research interests include wireless channel modeling, physical security, and 5G/B5G mobile communication systems.

Kaien Zhang received the B.S. degree in electronic information engineering from the Beijing Institute of Technology, Beijing, China, in 2021, where he is currently pursuing the M.S. degree in information and communication engineering. His research interests include wireless channel characterization and modeling.

Ammar Ghazal (Member, IEEE) received the B.Sc. degree in electronics and telecommunication engineering from Damascus University, Syria, in 2006, the M.Sc. degree from Heriot-Watt University, Edinburgh, U.K., in 2010, and the Ph.D. degrees from The University of Edinburgh and Heriot-Watt University, Edinburgh, U.K., in 2015. He joined De Montfort University, Leicester, U.K., as an Early Career Academic Fellow and has been a senior Lecturer in Wireless Communications since 2020. His main research interests include wireless propagation channel characterization and modeling, nonstationary channel models, high mobility wireless propagation, and (B)5G radio communication systems.

Wancheng Zhang received the B.Eng. degree from Beijing Institute of Technology, China, in 2005, the M.Sc. degree from the University of Edinburgh, U.K., in 2006, and the Ph.D. degree from Imperial College London, U.K., in 2010, all in Electrical Engineering. Presently, he is a Lecturer at the School of Information and Electronics, Beijing Institute of Technology. His research interests center around audio and acoustic signal processing, and he has worked in particular on dereverberation, system identification and equalization, and acoustic modeling.

Zijie Ji received the B.E. degree and the Ph.D. degree in information and communication engineering from the Beijing Institute of Technology, Beijing, China, in 2016 and 2022, respectively. From 2019 to 2021, he was a visiting student with the School of Electrical and Information Engineering, the University of Sydney, NSW, Australia. Since 2022, he has been an engineer with the Institute of Telecommunication and Navigation Satellite, China Academy of Space Technology, Beijing, China. His current research interests include physical layer security, secret key generation, Internet of Things, channel modeling, and satellite communications.

Limin Xiao (Member, IEEE) received the B.S. and M.S. degrees in wireless communication engineering from the Harbin Institute of Technology, Harbin, in 1992 and 1995, respectively, and the Ph.D. degree in communication and information systems from Tsinghua University, Beijing, China, in 2000. He is currently an associate professor in Tsinghua University. His research interests are in the area of wireless digital communications, including information theory, channel coding, spread-spectrum communications, and multiple antenna systems.