IEEE Access
Multidisciplinary | Rapid Review | Open Access Journal

Received November 22, 2016, accepted December 21, 2016, date of publication January 16, 2017,
date of current version March 13, 2017.

Digital Object Identifier 10.1109/ACCESS.2017.2652983

Measurement-Based Massive MIMO Channel Modeling for Outdoor LoS and NLoS Environments

JIAJING CHEN[1], (Student Member, IEEE), XUEFENG YIN[1], (Member, IEEE), XUESONG CAI[1], (Student Member, IEEE), AND STEPHEN WANG[2], (Senior Member, IEEE)

[1]College of Electronics and Information Engineering, Tongji University, Shanghai 201804, China
[2]Telecommunications Research Laboratory, Toshiba Research Europe Ltd., Bristol BS1 4ND, U.K.

Corresponding author: X. Yin (yinxuefeng@tongji.edu.cn)

This work was supported in part by the Natural Science Foundation of China under Grant 61471268, in part by the Hong Kong, Macao, and Taiwan Science Technology Cooperation Program of China, in part by the National Science and Technology Major Project of China under Grant 2016ZX03001015, and in part by the Toshiba Research Europe Limited in the Project Massive MIMO Channel Modeling and Applications.

ABSTRACT In this paper, a measurement campaign for massive multiple-input multiple-output (MIMO) channel characterization in both line-of-sight (LoS) and non-LoS outdoor environments is introduced. The measurements are conducted at the center frequency of 15 GHz with a bandwidth of 4 GHz. A virtual 40×40 planar antenna array formed by stepping a vertically-polarized bi-conical omni-directional antenna (ODA) along regularly-spaced grids is used in the receiver (Rx). The transmitter is equipped with a single ODA. To investigate channel variations over the Rx array, this 1600-element Rx array is split into multiple 7×7 sub-arrays, and a maximum-likelihood parameter estimation algorithm implemented using the space-alternating generalized expectation-maximization principle is applied to extracting multipath components (MPCs) from sub-array outputs. The spatial variability of K-factor, composite channel spreads in delay, azimuth, and elevation of arrival are investigated. Based on the estimated MPCs' parameters, multipath clusters are identified and associated across the array to find the so-called spatial-stationary (SS) clusters. From several hundreds of SS-clusters extracted, we establish a stochastic model for their life distances in horizontal and vertical directions, two-dimensional (2-D) life region, and variations of cluster spreads. These findings are important for massive MIMO channel modeling in the cases, where 2-D large-scale arrays are considered.

INDEX TERMS Massive MIMO channel measurement and characterization, spatial stationarity, composite channel parameters, clusters of multipath, cluster life distance, and cluster life region.

I. INTRODUCTION

Massive multiple-input multiple-output (MIMO) transmission by using large-scale antenna arrays is a key technology in the fifth generation (5G) wireless communications [1]–[6]. Massive MIMO techniques have been shown to achieve the benefits of spectral efficiencies, transmit diversity by using space-time coding, and high-data-rate through spatial multiplexing [7]–[10]. Characterization of massive MIMO channels becomes highly demanded especially for high-speed-railway scenarios [11], [12], crowded scenarios [13] and local area scenarios [14] where multiuser transmission is needed. In addition, the 5G communications require a wide bandwidth ranging from 500 MHz up to 2 GHz. Higher frequency bands (HFBs) beyond 6 GHz become preferable for signal transmissions due to the abundant spectrum resources available there [15]–[18]. Channel measurement campaigns in HFBs, such as 15 GHz [19]–[21], 28 GHz [22], [23], 60 GHz [24], [25], and 70-73 GHz [26], [27] have been conducted. Recently, the NTT DoCoMo has announced their experimental results, which show that the 15 GHz band can be adopted to reach the requirement of the transmitting data rate in the 5G systems [28]. Furthermore, air and rain attenuation at 15 GHz is relatively small compared to the millimeter-wave frequency bands [19]. These works reveal the promising feature of adopting centimeter-wave frequency bands for wireless communications in a variety of 5G applications.

Many researches conducted recently demonstrate that massive MIMO techniques are appropriate solutions for meeting the wideband communication demands of 5G in numerous scenarios. Similar with the conventional MIMO systems,

communications using massive MIMO are performed via antenna selection for hybrid maximum-ratio-combining and equal-gain-combining [29], [30]. Considering the large number of antennas available and flexible deployments of the array, realistic stochastic channel models are important for generating channel realizations in massive MIMO scenarios where large-scale antenna arrays are used [31].

Massive MIMO channels are different with the conventional MIMO channels due to the fact that the antennas in massive MIMO are widely distributed in a larger spatial region in such a way that the small-scale-characteristic (SSC) assumption [32] does not apply. As a result, the parameters of propagation paths observed through different antennas fluctuate due to the spatial displacement of these antennas, and the channel exhibits a significant spatial non-stationarity. Conventionally, the temporal non-stationarity of a channel represented by e.g. the birth-and-death processes of multipath clusters has been investigated in [33]–[35]. By analogy, spatial non-stationarity is referred to the birth-and-death behavior of clusters across the antenna array, and the closer the two antennas are set, the more common scatterers they share [36]. Thus, the observed clusters can be divided into two categories: wholly visible clusters that exist in the channels observed from the entire array, and partially visible clusters that appear only to a part of the array [37]. The results reported in [37] show that the non-stationarity from the cluster partial visibility increases the channel capacity, which motivates the research of the spatial non-stationarity in massive MIMO scenarios.

Conventional measurement-based stochastic channel models, such as the 3rd Generation Partner Project (3GPP) spatial channel model (SCM) [38], Wireless World Initiative New Radio (WINNER) II SCM-enhanced (SCME) model [39], and European Cooperation in Science and Technology (COST) family of models [40]–[43] are all constructed for concentrated MIMO scenarios where the SSC assumption applies. These models were established based on the parameter estimates of multipath components (MPCs) extracted using the high-resolution parameter estimation algorithms, such as the maximum-likelihood estimation utilized in practise through the Space-Alternating Generalized Expectation-maximization (SAGE) procedure [44] and the Richter's MAXimum likelihood estimation (RiMAX) technique [45]. The modeling process does not take into account the channel spatial-variability across the antenna array as the constant channel is assumed regardless of exact transmitter (Tx) and receiver (Rx) antennas. Furthermore, since it is impossible to extend the existing SCMs and SCMEs to the massive MIMO scenarios with satisfactory fidelity to the reality [46], [47], channel sounding dedicated for massive MIMO channel characterization is needed although it takes tremendous efforts in both on-field measurements and data post-processing.

Recently, channel non-stationarity properties observed with large-scale arrays considering massive MIMO scenarios have been reported in literature. For example, in [48]–[51], measurements at 2.6 GHz frequency bands by using both a virtual linear array with 128 antenna elements and a physical cylindrical array in outdoor scenarios demonstrated that the angles of arrival of MPCs and the clusters' powers evolve along with horizontal locations of antennas. The birth-and-death process of multipath clusters are observed in angular domains with respect to antenna locations in the large-scale array [49], [50]. These findings reveal the degrees of freedom involved in the model frameworks for non-stationary channels in massive MIMO scenarios. Many open issues are raised when conducting massive MIMO channel characterization, such as how to select the types of large-scale arrays, how to effectively model non-stationary channels, and what are the parameters of interest in the model. To find clues or solutions to these problems, more measurement campaigns and results of massive MIMO channel characterizations are required, especially in the scenarios where complex-structured large-scale arrays, such as two-dimensional (2-D) planar array, or three-dimensional (3-D) arrays with hundreds of antennas are implemented.

In this work, a measurement campaign is introduced for analyzing the channels observed via a large-scale antenna array in both line-of-sight (LoS) and non-LoS cases for a roof-top environment. A virtual planar antenna array with 40×40 positions regularly distributed on a plane perpendicular to the horizontal ground is used in the receiver. The selected environment is analogous to that observed in an open hall where a large-scale antenna array is installed in its center, and metallic or concrete structures exist in the premise. A vector network analyzer (VNA) measures the channels within 13 to 17 GHz. To investigate channel variations over the apertures of the Rx array, the 40×40 planar array is divided into multiple sub-arrays, and the maximum likelihood estimation implemented using the SAGE procedure is applied to extracting MPCs observed from individual sub-arrays. The spatial variability of composite channel parameters, such as K-factor, channel delay spread, azimuth of arrival (AoA) spread and elevation of arrival (EoA) spread, is analyzed. Moreover, we introduce the so-called spatial-stationary (SS) MPC clusters and extract them from the measurement data. The statistics of SS-cluster life distance (CLD), and 2-D life region (LR), are investigated.[1] The main objective of those studies is to verify the applicability of the methods adopted for massive MIMO channel characterization with the roof-top measurement campaign as an example.

The rest of the paper is organized as follows. Section II describes the measurement equipment and environment. Section III elaborates the variations of channel parameters

[1]It is worth mentioning that the cluster LR defined in this work refers to a surface of a large-scale 2-D antenna array where an MPC cluster is observable in the channel. Such an MPC cluster is different from the cluster of physical scatterers. Compared to the scatterer cluster, an MPC cluster may be visible in a channel in a short distance that the observer moves, due to the destructive superposition of MPCs. The results reported in this work demonstrate that most of the MPC clusters, particularly those belonging to non-LoS category, have their LRs significantly less than the overall aperture of the 40×40 antenna array.

and cluster number identified from sub-arrays across the large-scale Rx array. In Section IV, the definition and the identification procedure of SS clusters are explained, and the statistics extracted for the parameters characterizing these clusters are elaborated. Finally conclusive remarks are addressed in Section V.

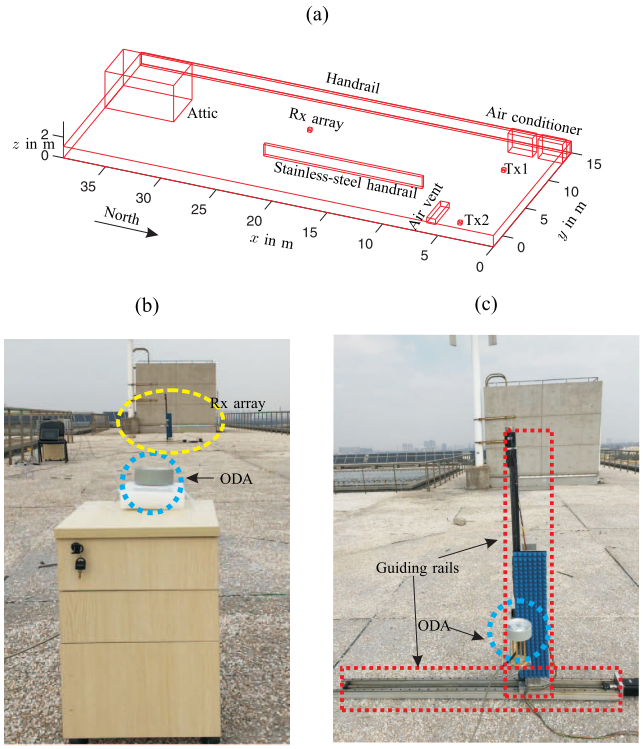


FIGURE 1. Measurement environment and the photographs of the Tx1 in a LoS scenario and Rx antenna installed on a positioner. (a) Measurement environment on the roof-top of a building. (b) Tx antenna in the location "Tx1" in Fig. 1(a). (c) Rx antenna installed in a positioner.

II. MEASUREMENT EQUIPMENT, ENVIRONMENT AND DATA PROCESSING METHOD

The measurement campaign was conducted in an outdoor environment on top of Telecom building, Tongji University. Fig. 1 (a) depicts the environment where the measurements were carried out. The objects in the premises include a concrete attic, external units of air-conditioners, metal handrails, and air vent fans. A VNA of type "Agilent 5233" configured in a single-Tx and single-Rx mode was applied in the measurements. The Tx was allocated at a LoS position and a NLoS position marked with "Tx1" and "Tx2" in two separate measurements respectively. The environment was kept stationary without humans moving during the measurements. The measurement system was triggered by remote control.

A vertically-polarized bi-conical omni-directional antenna (ODA) was installed in the Tx side at a height of 0.6 m above the ground, as illustrated in Fig. 1 (b). Fig. 1 (c) shows a virtual Rx antenna planar array mounted in a 40×40 grids on a vertical plane. The lowest position in the virtual array was 0.5 m above the ground. The neighboring positions in the array were separated by 0.5λ where λ represents the wavelength calculated with the highest frequency, i.e. 17 GHz considered in the measurements. An ODA similar to the Tx antenna was attached to a mast connected with a movable handle on the positioner. The positioning error of the system is less than 2.5 μm which coincides with the minimum displacement along the guide rail obtained when the stepping motor rotates a complete circle.

After the positioner reaches a desired sounding location, the measurement starts after additional 5 seconds in order for residual vibration to dissipate completely before data acquisition starts. Two phase-stable cables were used to connect the Tx and Rx ports from the VNA panel to the Tx and Rx antenna respectively. The cables were connected to the antennas through the feeding ports at bottom and drooped naturally to the ground. The VNA is adopted to measure the channel transfer functions with total bandwidth of 4 GHz at center frequency of 15 GHz for $40 \times 40 = 1600$ Rx positions. The channel impulse responses (CIRs) are obtained by devolving the responses of the cables and connectors measured in a calibration campaign. It took about 5 hours to complete the measurements of all 1600 spatial channels once. It is worth mentioning that when the cable is bent due to the motion of the antenna in the Rx array, the fluctuation of the phases of the cable responses is found to remain less than 1°, which introduces some errors in the MPC parameter extraction by using the high-resolution parameter estimation method. Table 1 lists the important parameter settings applied in the measurements.

The 2-D planar array is specifically considered here, because the deployment of such an array is convenient on walls in indoor hot-spot scenarios, such as railway stations, shopping malls, and auditorium in stadiums. A 3-D array can have more flexible choices for antenna locations which are hard to emulate in our measurements. Furthermore, it is important to emphasize that in our measurements, both the Tx and the Rx antennas are vertically polarized, which leads to a limitation of our results that only the vertical-to-vertical polarized components in a channel can be acquired and analyzed.

A. MPCs EXTRACTION

To extract MPCs from the observed CIRs, a square window of 7×7 grids is slid across the 40×40 Rx array, which results in totally $34 \times 34 = 1156$ sub-arrays. The outputs of each 49-element virtual sub-array are used as the observations for channel parameter estimation [52]. Under the assumption that the environment is static, the baseband representation of received signals at the output of the antenna at the nth position in the sub-array is written as

$$y_n(t) = \sum_{\ell=1}^{L} \alpha_\ell c_n(\mathbf{\Omega}_\ell) u(t-\tau_\ell) + w_n(t), \quad t \in [0, T) \quad (1)$$

TABLE 1. Specifications of the underlying antenna, the settings of antenna and of the VNA used in the measurements.

Parameters	Values
Specifications of the bi-conical omni-directional antenna	
Frequency range	2-30 GHz
Gain	4 dBi
3-dB beamwidth in azimuth	360°
3-dB beamwidth in elevation	10°
Polarization	Vertical
Settings of the virtual array	
Type of array	2-D vertical planar
Position lattice	40×40
Number of elements	1600
Array aperture	34.4 cm × 34.4 cm
Element spacing	0.88 cm
Positioner error	≤ 2.5 μm
Settings of the VNA	
Frequency span being measured	13-17 GHz
Intermediate frequency bandwidth	3 KHz
Transmission power	15 dBm
Number of sweep points	3201

where T represents the sounding duration, L denotes the total number of specular propagation paths, $c_n(\mathbf{\Omega})$ represents the antenna response when the underlying ODA is located at the nth position in the sub-array, $u(t)$ is the transmitted signal, α_ℓ, τ_ℓ, and $\mathbf{\Omega}_\ell$ denote respectively the complex attenuation, the delay and the DoA of the ℓth path, and $w_n(t)$ represents circularly symmetric white Gaussian noise components. The direction vector $\mathbf{\Omega}$ is a unite vector $\boldsymbol{e}(\phi,\theta)$ uniquely determined by its spherical coordinates $(\phi,\theta)\in[-\pi,\pi)\times[0,\pi]$ according to the relation [53]

$$\mathbf{\Omega}=\boldsymbol{e}(\phi,\theta)=[\sin(\theta)\cos(\phi)\quad\sin(\theta)\sin(\phi)\quad\cos(\theta)]^{\mathrm{T}} \tag{2}$$

with $[\cdot]^{\mathrm{T}}$ denoting the transpose operation. Notice that for simplicity, the assumption that the MPCs' parameters are constant in the bandwidth considered is applied here. We denote the unknown parameters in the signal model (1) as

$$\mathbf{\Theta}=[\alpha_\ell,\tau_\ell,\phi_\ell,\theta_\ell;\ell=1,\ldots,L]. \tag{3}$$

The maximum likelihood (ML) estimator $\hat{\mathbf{\Theta}}_{\mathrm{ML}}(\boldsymbol{y})$ of $\mathbf{\Theta}$ given the observation $\boldsymbol{y}=[y_n(t);n=1,\ldots,N]$ can be derived based on (1). Since the estimates of complex attenuations can be expressed as a function of other parameter estimates, the calculation of $\hat{\mathbf{\Theta}}_{\mathrm{ML}}(\boldsymbol{y})$ requires solving a $3L$-dimensional optimization problem, which is computationally prohibitive in practice. As an alternative, the parameter estimates can be updated iteratively by using the SAGE principle [44]. Typical examples of such algorithms for channel parameter estimation are the Improved Searching and Initialization (ISI)-SAGE introduced in [32] and the RiMAX in [45]. The performance of these algorithms is subject to the influences of incomplete and inaccurate signal models [54]. The readers are referred to these works for detailed explanation of algorithm derivation, implementation procedures, and performance assessment.

A problem when implementing the SAGE algorithm is that converging to the global maximum of the likelihood function is hard to guarantee. In our case, two approaches are adopted to assist localizing the global maximum: *i)* when initializing parameter estimates, the Bartlett beamforming technique [55] is used, and the local maxima of the delay-DoA power spectrum are identified to determine the estimates of path parameters; *ii)* the iteration number is set to 30. Furthermore, it is worth mentioning that in our case, the 7×7 sub-array has horizontal and vertical dimensions equal to 5.3 cm. Since the inverse signal bandwidth multiplied with the speed of light equals 7.5 cm, the narrow-band assumption that the array aperture should be much less than the speed of light multiplied with the inverse signal bandwidth does not hold strictly. As a consequence, the estimation accuracy of parameter estimation method derived based on the signal model (1) is reduced to a certain extent, especially when estimating the parameters of paths which own incident directions close to the plane where the 2-D array is located.

Fig. 2(a) illustrates an example the scatter-plot of estimated 100 MPCs for a sub-array in the LoS scenario, where the color of spots codes path power in dB. In the SAGE algorithm implemented, the searching step is set to 0.025 ns in delay and 1° in the angular domains respectively. It can be observed from Fig. 2(a) that the path with the strongest power has the AoA of 90°, EoA of 0°, and delay of 58 ns. These parameters are consistent with those of the LoS path calculated with the real Tx and Rx locations shown in Fig. 1(a). Figs. 2(a) to 2(d) depict the AoA-delay power spectra (PSs) of, respectively, original CIRs, reconstructed CIRs calculated based on the SAGE estimation result, and residuals of CIRs obtained by subtracting the reconstructed CIRs from the original ones. It is obvious that the PS of the reconstructed CIRs is close to that of the original CIRs, and the PS for the residuals contain components which have power 30 dB less than the maximum spectral height of the PS of the original CIRs.

In general, we found that when the SAGE algorithm is applied to estimate 100 and 300 paths in the LoS and NLoS scenarios respectively, the maximum power of the residuals computed by subtracting the CIRs reconstructed based on estimated MPCs from the original CIRs is decreased by at least 30 dB compared with the maximum power of the original CIRs. It is obvious that such numbers of paths are significantly larger than that considered when analyzing channels below 6 GHz [56]. We postulate that this is due to the following reasons: *i)* more paths are resolved due to higher delay resolution of 0.25 ns in our measurements; *ii)* the richness of scattering is more significant in 13-17 GHz frequency bands than in below 6 GHz; *iii)* the assumption that the antenna radiation pattern is frequency independent is not strictly valid; *iv)* the narrow-band assumption that the time span for a planar wave propagating across the sub-array is much less than the inverse of the sounding signal bandwidth does not hold strictly.

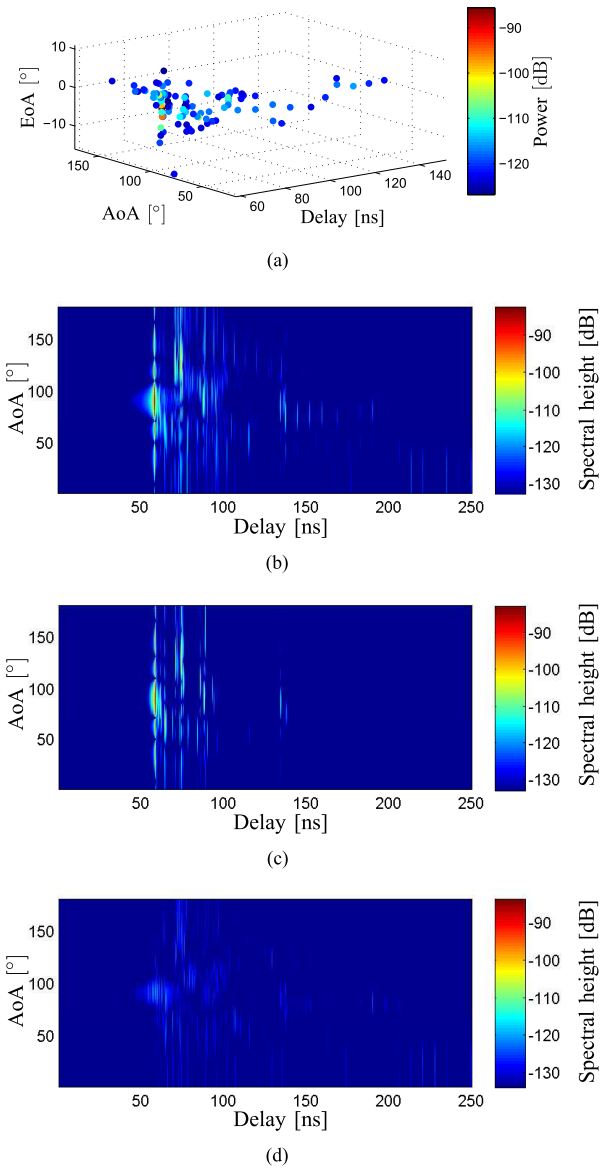


FIGURE 2. Comparison among the SAGE estimation results, and the power spectra of original CIRs, the reconstructed CIRs and the residuals. (a) MPCs estimated by using SAGE. (b) AoA-delay PS of the original CIRs. (c) AoA-delay PS of the reconstructed CIRs. (d) AoA-delay PS of the residuals.

TABLE 2. Parameter settings for the *K* PowerMeans-based clustering algorithm used in the study.

Parameter	Value
Number of clustering iterations	10
Minimum number of clusters	2
Maximum number of clusters	$4\sqrt{L}$

B. CLUSTER-INTERACTING-OBJECT MAPPING

The estimated MPCs are further grouped as clusters by using the *K* PowerMeans method introduced in [57], with parameter settings reported in Table 2. It is worth mentioning that some of the identified clusters include only a single path. This is because the channel is static, and as a result, only a single snapshot of channel is obtained in the measurement. With one realization of path constellation, the *K* PowerMeans method considers the paths isolated significantly as single-path clusters.

In order to verify whether the MPCs estimated and the clusters identified are reasonable, a cluster-interacting-object mapping procedure is conducted. The propagation paths are reconstructed based on the clusters' center of gravity (CoG) parameters by using a simple ray-tracing approach introduced in [58]. In the SIMO case considered here, path reconstruction starts by launching a ray from the center of an Rx sub-array along the direction determined by the mean cluster AoA and EoA. If an object in the environment exists along the ray, we then connect the Rx antenna with the object, and further with the Tx. If the path length is approximately equal to the average delay of the cluster multiplied with the speed of light, the identified interacting objects are considered to be the last-hops of paths within the cluster.

Notice that since a planar array consisting of ODAs is used in our measurements, there exists ambiguity that the estimated AoA, say $\hat{\phi}$ within the range $[0°, 180°]$, is in fact $-\hat{\phi}$ within the range $[-180°, 0°]$. Here the normal direction of the array plane has azimuth of 90° and elevation 90°. This creates problems when mapping clusters to interacting objects. A practical solution used in our case is that for an estimated DoA with spherical coordinate $(\hat{\phi}, \hat{\theta})$, two rays are launched from the array center, pointing towards a forward direction $\boldsymbol{e}(\hat{\phi}, \hat{\theta})$ and a backward direction $\boldsymbol{e}(-\hat{\phi}, \hat{\theta})$ respectively. The objects existing along these two rays are used to construct single-bounce paths. The path that provides similar delay to the estimated value is selected, and the other is discarded. By using this method, we found that in most cases, only one cluster can be mapped to the attic in the backward side of the Rx array, and the other clusters are mapped to the objects in the forward side of the array.

Fig. 3 illustrates as an example, a cluster-interacting-object mapping result obtained for the channel observed through a 7×7 sub-array by using the aforementioned method. It can be observed that the locations of most identified clusters are close to the objects in the environment, such as the air conditioner, stainless-steel handrail, and air vent. This implies that most clusters are generated via single-bounce paths, and moreover, the parameter estimation algorithm and the clustering technique can bring reasonable results in the case considered.

III. SPATIAL VARIATIONS OF CHANNEL PARAMETERS

In this section, we investigate the variations of K-factors, the composite channel root-mean-square (rms) delay, AoA and EoA spreads, as well as the number of clusters, calculated based on the MPCs' parameter estimates extracted from the CIRs observed by individual Rx sub-arrays across the large-scale 2-D array aperture. In order to find the analytical representation of the parameters' distributions, the Kolmogorov-Smirnov (KS) testing method is adopted to assess the "goodness-of-fit" between an empirical

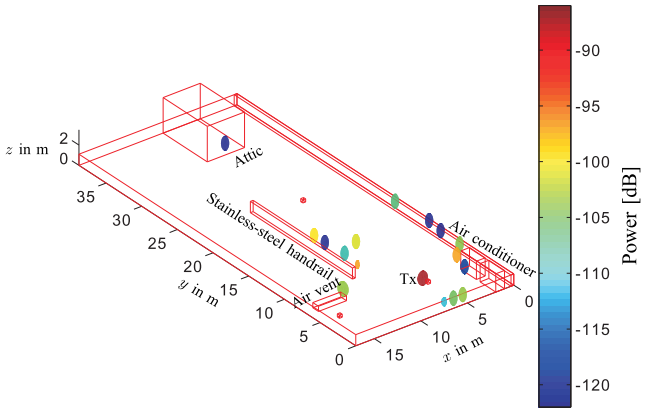


FIGURE 3. The clusters obtained in a 7 × 7 sub-array by using a ray-tracing-alike method in a LoS scenario.

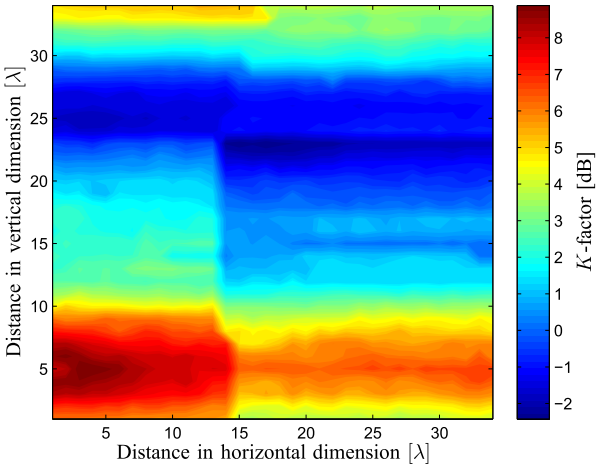


FIGURE 4. *K*-factor $K^{[i,j]}$ of channels observed at sub-arrays across the 2-D aperture of Rx array for LoS scenarios.

cumulative density function (cdf) graph and the cdf of an existing distribution with known expressions. Without being specifically mentioned, the KS-test shows that the null-hypothesis is non-rejectable for analytical cdfs and the empirical graphs shown in this section. This implies that the fitted probability density function (pdf) with selected parameters is an appropriate expression of the empirical distributions.

A. K-FACTOR

As a measure of the fading severity for narrowband channels [59], K-factor is defined to be the power ratio of LoS components to NLoS components. For notational convenience, we use $P_{i,j}(\tau)$ to denote a channel power delay profile, where $i, j \in [1, \ldots, 34]$ represent the 2-D indices of the sub-array, and τ is the delay of the path. The power of LoS component is represented with $P_{i,j}^{\text{LoS}}$ which is the maximum among all estimated MPCs, and the power of the NLoS component $P_{i,j}^{\text{NLoS}}(\tau)$ is calculated by accumulating the powers of the other paths. The K-factor $K^{[i,j]}$ in the linear scale for the channel observed by the $[i, j]$th antenna can be calculated as $K^{[i,j]} = P_{i,j}^{LoS} / P_{i,j}^{NLoS}$.

Fig. 4 depicts $K^{[i,j]}$, $i, j = [1, \ldots, 34]$ represented in dB calculated for the LoS scenario. The ticks along the x- and y-axes denote respectively the horizontal and vertical locations of the centers of the sub-arrays in wavelength. It can be observed that the K-factor varies dramatically within the range $[-2, 8]$ dB, and it gradually changes across the large-scale Rx array aperture. Specifically, the K-factor first decreases when the antenna's height increases from the bottom of the array, and then rises again when the antenna height is above 14λ. The largest K-factor can be found at the height of 3λ above the lowest position. An explanation for observing largest K-factor at this specific height is that the antenna radiation pattern has a narrow beam-width in elevation, and due to this reason, when the sub-array is 3λ above the bottom of the virtual array, the Tx and Rx antennas are aligned. Consequently, the power of the LoS component increases significantly, and a larger K-factor is obtained.

B. COMPOSITE RMS DELAY SPREAD $\sigma_\tau^{[i,j]}$

Fig. 5 depicts the contour plots of the rms delay spreads $\sigma_\tau^{[i,j]}$ of channels for both LoS and NLoS scenarios. The method applied to calculate the rms delay spread is the same as adopted for WINNER II SCME modeling [39]. It can be observed from Fig. 5 that $\sigma_\tau^{[i,j]}$, $i, j \in [1, \ldots, 34]$ in the NLoS scenario are significantly larger than those in the LoS scenario. Furthermore, it is apparent from Fig. 5(a) that the distribution of $\sigma_\tau^{[i,j]}$ over the array aperture exhibits block-like patterns which have clear boundaries in both horizontal and vertical directions for the LoS scenario. In the areas close to the boundaries, $\sigma_\tau^{[i,j]}$ changes more abruptly with a difference up to 2 ns. In the NLoS scenario, $\sigma_\tau^{[i,j]}$ is randomly distributed in most parts of the array aperture except for a rectangular region of 10λ × 5λ to the bottom where $\sigma_\tau^{[i,j]}$ fluctuates dramatically. These observations of rms delay spread varying significantly can be due to several reasons: *i)* the paths generated by scatterers experience birth and death processes from a sub-array to the next [49]; *ii)* shadow fading occurs across the array in such a way that the paths experience significant attenuation when being observed at a different location.

Fig. 6 illustrates the empirical cdfs $P(\sigma_\tau^{[i,j]} < \text{abscissa})$ of the composite rms delay spread $\sigma_\tau^{[i,j]}$ in LoS and NLoS scenarios. It can be observed by comparing these cdfs that in general, the delay spreads obtained in the LoS scenario are about one order of magnitude less than those observed in the NLoS scenario. Both cdfs can be well fitted by using normal distributions with appropriately chosen parameters reported in the figure legend, where $\mathcal{N}(\mu, \sigma)$ denotes the normal distribution with μ and σ being mean and standard deviation respectively.

C. COMPOSITE AoA SPREAD $\sigma_\phi^{[i,j]}$

Figs. 7(a) and 7(b) depict $\sigma_\phi^{[i,j]}$ of the channels observed for LoS and NLoS scenarios respectively. It is evident that $\sigma_\phi^{[i,j]}$ in the NLoS scenario is significantly larger than that observed in the LoS scenario. Furthermore, the variation pattern of

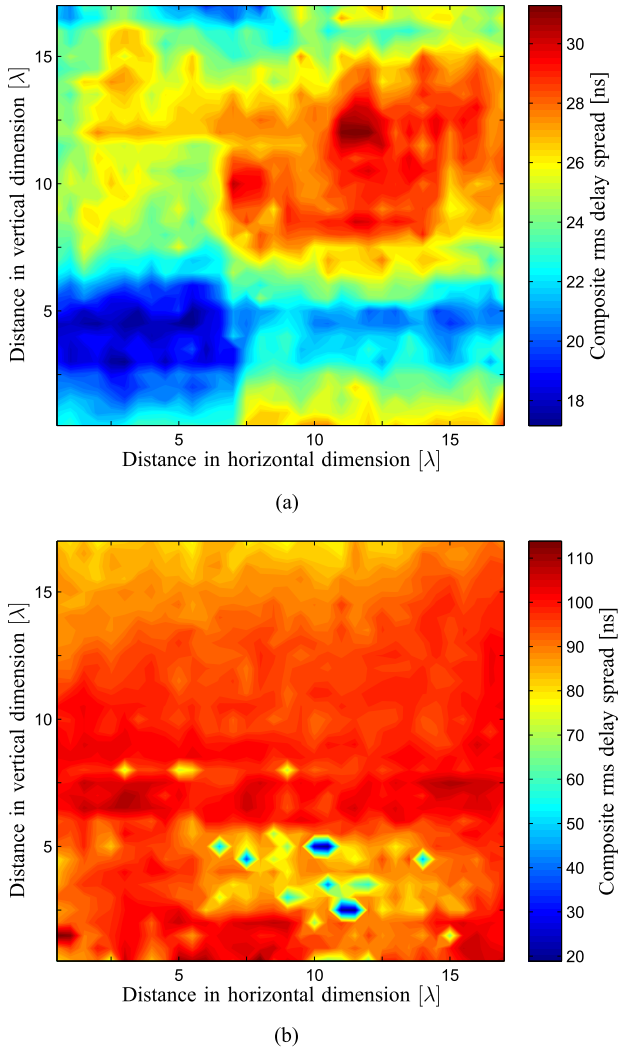


FIGURE 5. Composite delay spreads $\sigma_\tau^{[i,j]}$ of channels observed by sub-arrays across the 2-D aperture of array for LoS and NLoS scenarios. (a) Rms delay spread in LoS scenario. (b) Rms delay spread in NLoS scenario.

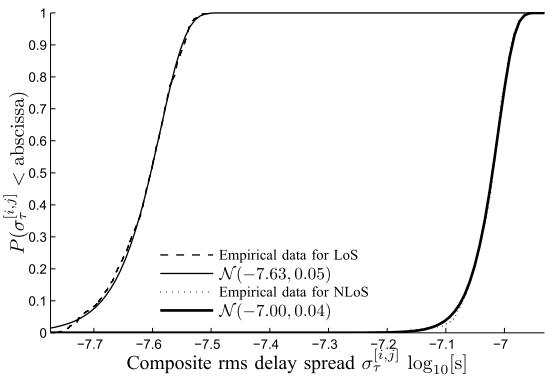


FIGURE 6. Empirical cdfs $P(\sigma_\tau^{[i,j]} <$ abscissa) for channels observed by sub-arrays across the 2-D aperture of array for LoS and NLoS scenarios.

$\sigma_\phi^{[i,j]}$ in the LoS scenario looks the same as for $\sigma_\tau^{[i,j]}$. Even the blocks' borders are identical. Fig. 8 demonstrates the empirical cdfs $P(\sigma_\phi^{[i,j]} <$ abscissa) in LoS and NLoS scenarios. The fitted curves obtained by using lognormal

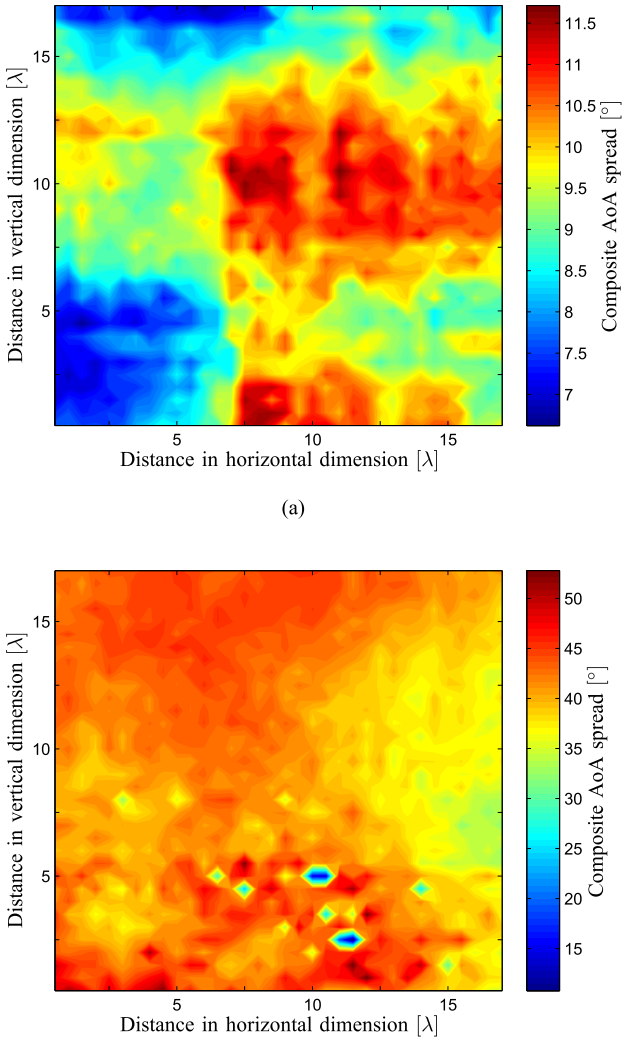


FIGURE 7. Composite AoA spreads of channels observed by sub-arrays across the 2-D aperture of array for LoS and NLoS scenarios. (a) Composite AoA spread in LoS scenario. (b) Composite AoA spread in NLoS scenario.

distributions with selected parameters reported in the legends are also illustrated in Fig. 8. It can be observed that the distributions of the composite AoA spreads are completely separated along the abscissa. The mean $\sigma_\phi^{[i,j]}$ represented in $\log_{10}([^\circ])$ equals 0.97, i.e. $10^{0.97} \approx 9.3^\circ$ and 1.61, i.e. $10^{1.61} \approx 40.74^\circ$ for the LoS and NLoS scenarios respectively. The standard deviation of $\sigma_\phi^{[i,j]}$ in $\log_{10}([^\circ])$ is observed to be small, i.e. equal to 0.06 ($\approx 1.2^\circ$) and 0.03 ($\approx 1.1^\circ$) for LoS and NLoS scenarios respectively, which implies that although the values of $\sigma_\phi^{[i,j]}$ exhibit block-like pattern as illustrated in Fig. 7, the composite AoA spread can be considered approximately identical across the array.

D. COMPOSITE EoA SPREAD $\sigma_\theta^{[i,j]}$

Fig. 9 depicts the contour plots of the composite EoA spreads $\sigma_\theta^{[i,j]}$ calculated for LoS and NLoS scenarios. It is observed

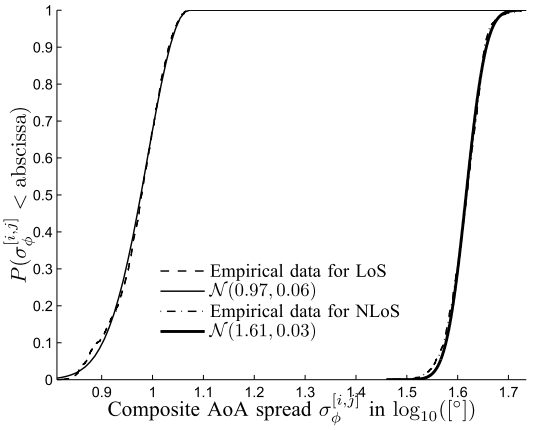
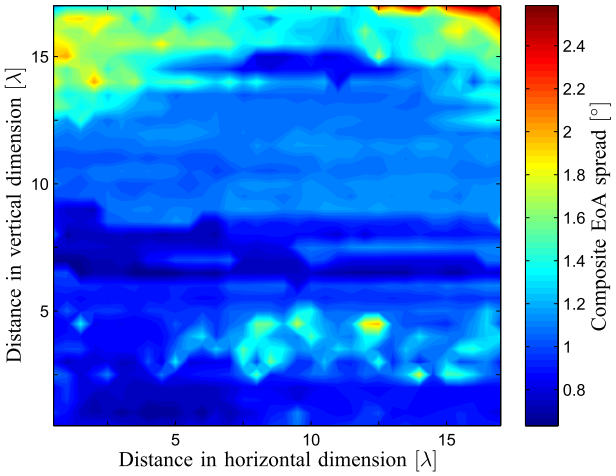
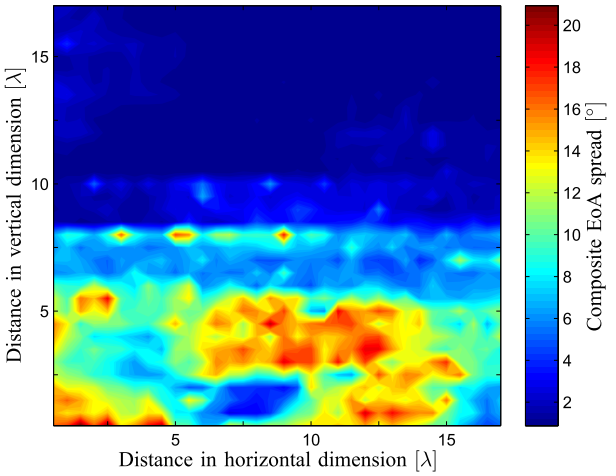


FIGURE 8. Empirical cdfs of composite AoA spreads $\sigma_{\phi}^{[i,j]}$ of channels observed by sub-arrays across the 2-D aperture of array for LoS and NLoS scenarios.



(a)



(b)

FIGURE 9. Composite EoA spreads $\sigma_{\theta}^{[i,j]}$ of channels observed by sub-arrays across the 2-D aperture of array for LoS and NLoS scenarios. (a) Composite EoA spreads $\sigma_{\theta}^{[i,j]}$ in LoS scenario. (b) Composite EoA spreads $\sigma_{\theta}^{[i,j]}$ in NLoS scenario.

that in general, $\sigma_{\theta}^{[i,j]}$ is larger in NLoS than in LoS scenarios. We postulate that this is due to the significant richness of MPCs in the NLoS case. However, by comparing the variations behavior of $\sigma_{\theta}^{[i,j]}$ versus the height of sub-array centers, we observe that in the LoS case, $\sigma_{\theta}^{[i,j]}$ increases, and in the NLoS case, $\sigma_{\theta}^{[i,j]}$ decreases as the sub-array height increases. Specifically, the average $\sigma_{\theta}^{[i,j]}$ reduces from 15° to 2° with clearly separated stages observed at the height of 6λ, 8λ and 10λ, respectively. Our conjecture is that the birth-and-death processes of MPCs occur more regularly in the vertical dimension than in the horizontal dimension. In the NLoS scenario, the values of the EoA spreads are either quite small or very large, which results in the standard deviation of the NLoS scenario being higher than that in the LoS scenarios.

Fig. 10 illustrates the empirical cdfs $P(\sigma_{\theta}^{[i,j]} < \text{abscissa})$ in both scenarios. It can be observed that $\sigma_{\theta}^{[i,j]}$'s are concentrated within [0°, 2°] in the LoS scenario, and are widely spread within [0°, 20°] in the NLoS scenario. Furthermore, as illustrated in Fig. 10, the distribution of $\sigma_{\theta}^{[i,j]}$ is fitted approximately by using a normal cdf in the LoS scenario. For the NLoS scenario, the empirical pdf exhibits multi-modal behavior, and cannot be fitted by a single distribution with a uni-modal pdf. This effect is consistent with the observation of stage-like variations of $\sigma_{\theta}^{[i,j]}$ in the NLoS scenario illustrated in Fig. 9(b), implying that $\sigma_{\theta}^{[i,j]}$ varies deterministically rather than randomly in the vertical dimension in our case.

By comparing the variation patterns of K-factor, delay spread, AoA spread, and EoA spread in the LoS scenario, as illustrated in Figs. 4, 5(a), 7(a), and 9(a) respectively, it is obvious that the variations of K-factor, delay spread, and AoA spread are highly correlated: when the K-factor increases in a certain area, the delay spread and the AoA spread in the same area are observed to decrease. However, the variations of the EoA spread is independent of the variations of the others. We postulate that these observations are due to the fact that the scatterers are distributed more horizontally in the open roof-top environment than in the vertical direction.

E. NUMBER OF CLUSTERS IN CHANNEL OBSERVED BY INDIVIDUAL SUB-ARRAYS

From the MPCs estimated by using the outputs of individual sub-arrays, the KPowerMeans clustering algorithm described in [57] and [60] is applied to extract multiple clusters of MPCs. For massive MIMO scenarios, the variations of the cluster characteristics across the array is important for understanding the evolving behavior of small-scale channel characteristics observed by using a large-scale antenna array.

Figs. 11(a) and 11(b) depict the distribution of the total number of clusters identified from individual sub-array across the large-scale Rx array for the LoS and NLoS scenarios, respectively. It can be observed from Fig. 11(a) that for the LoS scenario, the number of clusters existing in the channel observed by a 7×7 sub-array is less than 30 for most sub-arrays. It is interesting to observe that the number of clusters

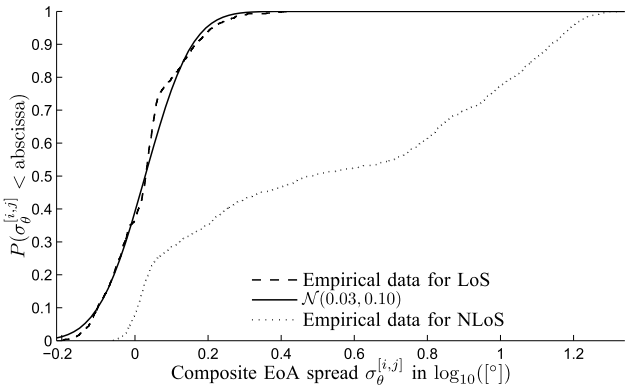


FIGURE 10. Empirical cdfs $P(\sigma_\theta^{[i,j]} < \text{abscissa})$ of composite EoA spreads $\sigma_\theta^{[i,j]}$ of channels observed by sub-arrays across the 2-D aperture of array for LoS and NLoS scenarios.

TABLE 3. Empirical statistics of channel parameters.

Channel parameters\Scenarios		LoS	NLoS
K-factor (dB)	Mean	2.33	/
	Std.	4.43	/
Composite delay spread $[\log_{10}(\text{s})]$	Mean	-7.61	-7.03
	Std.	0.05	0.06
Composite AoA spread $[\log_{10}(^\circ)]$	Mean	0.97	1.61
	Std.	0.05	0.03
Composite EoA spread $[\log_{10}(^\circ)]$	Mean	0.03	0.54
	Std.	0.10	0.45
Number of clusters in channel observed by individual sub-arrays	Mean	28	25
	Std.	3.75	18.98

maintains the same level about 30 across the array with small variance in the LoS scenario. This phenomenon is caused by the fact that many MPCs estimated with strong powers remain observable across the array. Such a high consistency in propagation constellations leads to similar cluster distribution and constant number of clusters. For the NLoS scenario illustrated in Fig. 11(b), the number of clusters is observed to fluctuate randomly with a larger standard deviation across the array than in the LoS scenario. The rapid fluctuation of cluster number is due to the frequent birth and death of paths across the array aperture. We postulate that in HFBs, scattering effects can be more significant than in the band below 6 GHz, and many objects not involved in the propagation below 6 GHz began to interact with waves. Furthermore, for the large bandwidth adopted, hundreds of paths can be resolved in delay with high resolution, which increases the randomness of the observations across the array to a certain extent. From Fig. 11, it can be observed that the number of clusters observed by individual sub-arrays can be modeled as a random process across the array in both LoS and NLoS scenarios with different standard deviations.

In summary, Table 3 reports the empirical statistics of the composite channel parameters. The observation that composite delay, AoA and EoA spreads vary in both scenarios implies that the massive MIMO transmission technique needs to be adaptive when different sub-arrays are utilized. Furthermore, the antenna selection techniques which gained research

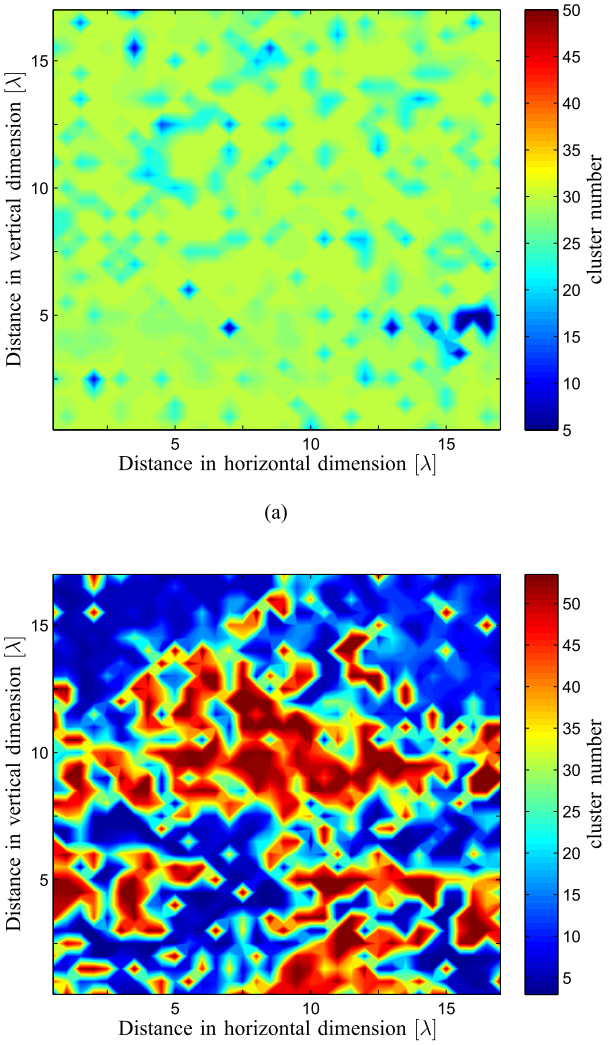


FIGURE 11. Number of clusters in LoS and NLoS scenarios. (a) Number of clusters in the LoS scenario. (b) Number of clusters in the NLoS scenario.

attentions for massive MIMO applications [31], should take into account the exact variations of channels across the array aperture.

IV. SPATIAL-STATIONARY CLUSTER CHARACTERIZATION

Inspired by the COST 2100 channel models where the visibility of cluster is probabilistic across snapshots [42], we introduce the so-called "spatial-stationary (SS) clusters" to characterize the channels in massive MIMO scenarios. An SS-cluster refers to a cluster which exists in the channels observed via more than one sub-arrays. Such a cluster can be generated by the same interacting object(s) in the environment. However, due to the spatial displacement of the sub-arrays, the cluster's parameters may slightly change with respect to individual sub-arrays, particularly in the case where interacting objects are close to the array. These changes are sufficiently small, so the cluster's parameters can be viewed as wide sense stationary processes in spatial domains. Furthermore, an SS-cluster exists in a so-called

cluster LR, referring to a 2-D surface in the array plane. Only when an antenna is within a LR for a certain cluster, the channels obtained through these antenna locations contain realizations of the cluster. It is important to note that the cluster LR defined here differs from the visibility region (VR) defined in COST 2100 channel models: the VR specifies the probability for a cluster being visible by a user equipment, and the LR is a region in the Rx array aperture where a cluster exists in the channels observed. With slight abuse of notations, the parameters characterizing SS-cluster may include *i)* the "CLD", i.e. the horizontal and vertical extension of the LR; *ii)* the area of "cluster LR" represented in λ^2, and *iii)* the statistics of rms spreads of an SS cluster in its LR.

The SS-clusters can be identified using a simple approach based on the extracted multipath clusters for channels observed with individual sub-arrays. First, a criterion for determining the stationarity of a cluster is specified. The parameters of any cluster, say the ith cluster, are organized as a vector, $\boldsymbol{\zeta}_m = [\bar{\tau}_{\mathrm{c},m}, \bar{\phi}_{\mathrm{c},m}, \bar{\theta}_{\mathrm{c},m}, \sigma_{\tau,\mathrm{c},m}, \sigma_{\phi,\mathrm{c},m}, \sigma_{\theta,\mathrm{c},m}]$, which contains the CoG of the cluster, represented by mean delay $\bar{\tau}_{\mathrm{c},m}$, mean AoA $\bar{\phi}_{\mathrm{c},m}$, mean EoA $\bar{\theta}_{\mathrm{c},m}$ of the ith cluster, and rms delay spread $\sigma_{\tau,\mathrm{c},m}$, AoA spread $\sigma_{\phi,\mathrm{c},m}$ and EoA spread $\sigma_{\theta,\mathrm{c},m}$ of the cluster. To verify whether the ith and jth clusters belong to the same group, the difference vector $\Delta\boldsymbol{\zeta}_{ab} = \boldsymbol{\zeta}_a - \boldsymbol{\zeta}_b$ is calculated. If and only if the absolute values of all entries in $\Delta\boldsymbol{\zeta}_{ab}$ are less than their respective pre-determined thresholds, we may consider these clusters with the same wide sense stationarity. The exact procedure for searching clusters with same stationarity is as follows. We select the cluster with maximum power from the set of clusters as a reference cluster. Then the other clusters with the same stationarity as the reference cluster are identified. It is possible that these clusters, including the reference cluster, may not be found from neighboring sub-arrays. In such a case, these clusters are gathered into multiple groups which are separated in the array aperture. Although these separated clusters still share the same wide sense stationarity, in order to maintain tractable modeling complexity we decide to view them as different SS-clusters. The number of SS-clusters exhibiting the same stationarity may be considered as a random variable, and its statistics extracted can be included into the model established. Those clusters assigned to the SS-clusters already identified are removed from the set of clusters valid for SS-cluster association. Then a new SS-cluster is extracted from the valid cluster set updated. This operation iterates until all clusters observed in individual sub-arrays are assigned to certain SS-clusters.

It is obvious that by changing the threshold values applied to verify stationarity, SS-clusters may have different extensions and quantities. Practically, we set the thresholds as follows. First, for mean cluster delay and rms cluster delay spread, a same threshold τ_{th} is applied, and for mean angles (i.e. azimuth and elevation) and angular spreads, another threshold ψ_{th} is specified. These thresholds are determined by the criterion that most of the identified SS-clusters should

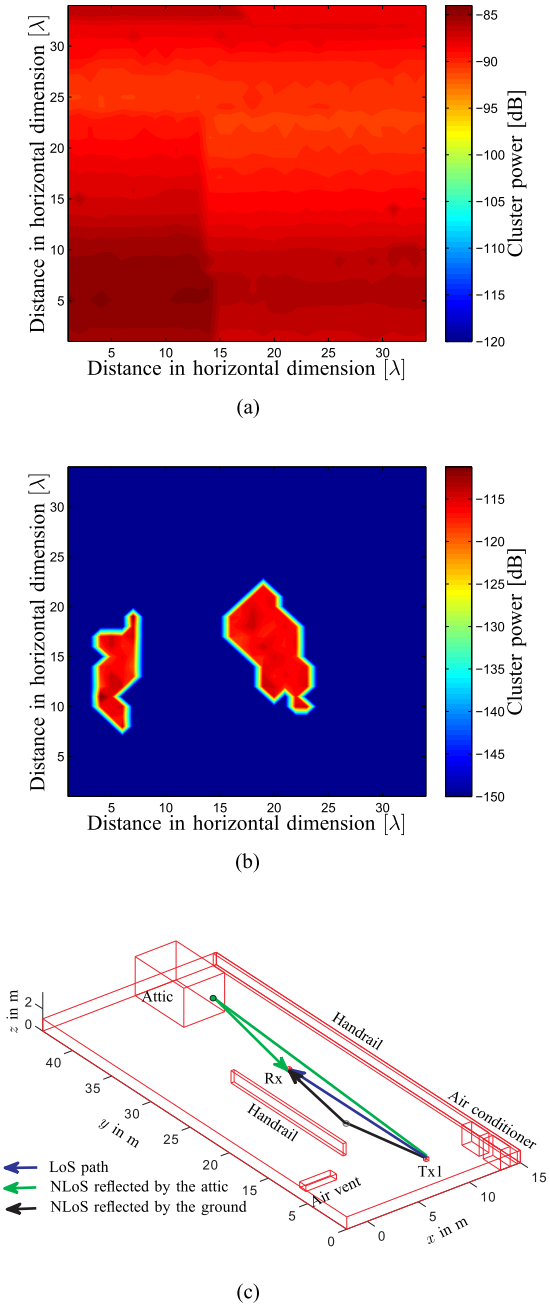


FIGURE 12. An example of extracted SS-clusters and the objects which may generate these clusters in the environment. (a) LR of an SS-cluster generated by LoS path. (b) LR of two SS-clusters generated by reflection from the ground and attic, respectively. (c) The propagation paths reconstructed and objects identified which jointly generate the clusters shown in Fig. 12(a) and Fig. 12(b).

exist in a region with certain spatial extension across the array. By applying this criterion, we found $\tau_{\mathrm{th}} = 2/B$ with B representing the total signal bandwidth and $\psi_{\mathrm{th}} = 2°$. Fig. 12(a) and Fig. 12(b) illustrate three examples of SS-clusters identified based on the aforementioned method with $\tau_{\mathrm{th}} = 2/B$ and $\psi_{\mathrm{th}} = 2°$ in the LoS scenario. It can be observed that these SS-clusters exhibit different extensions across the large-scale array. The SS-cluster shown in Fig. 12(a) is observable through the whole 34×34 sub-arrays

across the Rx array. The other two SS-clusters in Fig. 12(b) exhibit less power and smaller LRs than the aforementioned one. It can be observed that the LR of the extracted LoS SS-cluster seems to have a similar varying pattern with that of the K-factors illustrated in Fig. 4. We postulate that this is due to the following reasons: In the LoS scenario, the power of the LoS cluster is much higher than that of the NLoS clusters, and the power of the NLoS clusters observed across the subarrays is approximately equal. Thus, the pattern of the K-factors is determined by the power of the LoS components.

Fig. 12(c) illustrates the reconstructed paths for three SS-clusters whose LRs are shown in Fig. 12(a) and Fig. 12(b). It can be observed that the SS-cluster which can be observed across the whole array was generated by the LoS path, and the other two NLoS SS-clusters are generated by the reflections on the ground and on the surface of the attic to the left respectively. Notice that the interaction locations marked as ''o'' on the objects in Fig. 12(c) may change slightly when an SS-cluster is observed at different sub-arrays. In general, we may consider the extension of an SS-cluster as illustrated in Fig. 12(b) as the LR of the SS-cluster confined in a vertical planar array.

By using the SS-cluster identification method proposed, totally 138 and 331 SS-clusters are extracted for the LoS and NLoS scenarios respectively. These numbers appear to be much larger than the cluster number per channel, usually in the level of dozens, specified in the 3GPP spatial channel models. This is reasonable since in the case considered here, we have taken into account the channels observed from a total of 1156 sub-arrays. After merging the clusters that appear repetitively in channel observations from different sub-arrays into SS-clusters, it is still possible to observe hundreds of clusters through multiple portions of a large-scale array. More specifically, in the LoS scenario, about 5 SS-clusters exhibit LRs covering the whole Rx array, which according to the analysis by using the ray-tracing method, are generated by either the LoS path, or strong reflections on the ground close to the Tx. There also exists the case where the LR of a SS-cluster contains only a small number of sub-arrays, with an area less than λ^2. Practically, these SS-clusters exhibit power more than 30 dB less than the maximum cluster power. Therefore, the contributions of these clusters to the overall channel are small and thus negligible for channel modeling.

To characterize a massive MIMO channel, it is important to analyze the properties of LR for the identified SS-clusters and examine the dependence of these properties on e.g. LoS and NLoS scenarios under the influences of environments. In the following, parameters characterizing the SS-clusters are calculated, and their statistics extracted based on the measurement data are presented.

A. CLD IN VERTICAL AND HORIZONTAL DIRECTIONS - d_h AND d_v

The CLD in the vertical direction refers to the LR extension which is calculated to be $(N-1)\lambda/2$ with N being the number of consecutive sub-arrays in the vertical direction,

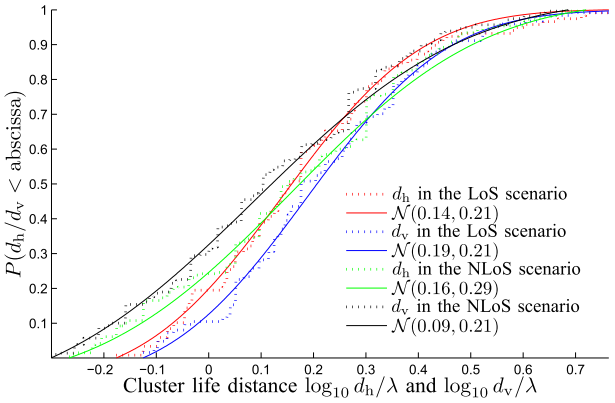


FIGURE 13. Empirical cdfs $P(d_h <$ abscissa$)$ and $P(d_v <$ abscissa$)$ of cluster life distance d_h and d_v respectively, for LoS and NLoS scenarios.

and similar definition applied to the CLD in the horizontal direction. In the case considered here, we use d_h and d_v to denote the CLD of an SS-cluster in horizontal and vertical dimensions respectively. Fig. 13 depicts the empirical cdfs $P(d_h <$ abscissa$)$ and $P(d_v <$ abscissa$)$ calculated based on 138 and 331 SS-clusters for LoS and NLoS scenarios respectively. It is worth mentioning that we found about 4% of the SS-clusters in the LoS and NLoS scenarios to have LRs covering the whole aperture of Rx array. Since these SS-clusters' CLDs are limited by the aperture of the underlying large-scale array, the cdfs shown in Fig. 13 were calculated without considering these clusters. It can be observed from Fig. 13 that d_h and d_v follow approximately lognormal distributions, within the support from 0.63λ (calculated from $10^{-0.2}\lambda$ according to Fig. 13) to 5λ (calculated with $10^{0.7}\lambda$). The parameters of the fitted pdfs are reported in the legend of Fig. 13. The CLDs' cdfs in the LoS scenario are slightly shifted towards the larger value abscissa compared with cdfs in the NLoS scenario, indicating that the longer CLD is observed with higher probabilities in LoS than in NLoS. The average of the CLD d_v is larger than that of d_h in average by 0.5 dB in LoS scenario, and the contrary can be observed for the NLoS scenario. In average, the SS-clusters maintain 1.2 to 1.5λ CLD in average in both horizontal and vertical directions. We also observed from Fig. 13 that 90% of CLDs are less than 5λ (calculated from $10^{0.7}\lambda$). In the case where the carrier frequency equals 15 GHz, this implies that most of the mean extension of LR in vertical or horizontal dimensions is less than 10 cm in the antenna array aperture.

B. SS-CLUSTER LR

The channels observed through the Rx antennas confined in an SS-cluster LR contain the realizations of the same SS-cluster. In other words, a cluster can be considered to be ''alive'' only in this 2-D LR. As illustrated in Fig. 12(b), a realistic SS can have diverse shapes which are difficult to be fitted with regular shapes. Therefore, we use the area s_{LR} of the 2-D LR as a metric to quantify the ''alive'' area of an SS-cluster in the large-scale array. It is worth mentioning that in our case, the LoS cluster is observed to exhibit a very

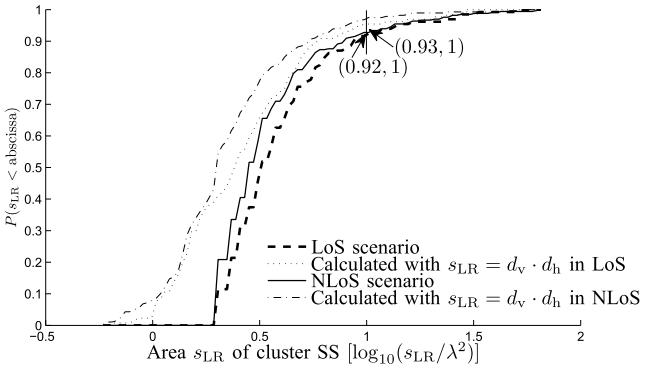


FIGURE 14. Empirical cdfs $P(s_{\rm LR} <$ abscissa) for LoS and NLoS scenarios.

large LR which covers all antennas in the large-scale array as illustrated in Fig. 12(b), thus, it is necessary to investigate the distributions of the LR of NLoS SS-clusters. Fig. 14 depicts the empirical cdfs $P(s_{\rm LR} <$ abscissa) for the NLoS clusters identified in LoS and NLoS scenarios. The abscissa represents $s_{\rm LR}$ in $\log_{10}(s_{\rm LR}/\lambda^2)$. It can be observed from Fig. 14 that the largest $s_{\rm LR}$ is up to $63\lambda^2$, which is approximately equal to 1.8 represented in $\log_{10}(s_{\rm LR}/\lambda^2)$ in the LoS scenario. The median value of $s_{\rm LR}$ is observed to be 0.45 and 0.50, i.e. $2.8\lambda^2$ in the NLoS scenario and $3.2\lambda^2$ in the LoS scenario respectively. This indicates that half of the NLoS SS-clusters are only visible within an area less than a 4×4 planar sub-array with neighboring antenna elements spaced by $\lambda/2$. Furthermore, only 8% of the SS-cluster have $s_{\rm LR}$ larger $10\lambda^2$ in both LoS and NLoS scenarios, which indicates that the majority of SS-clusters are visible within the region less than the area of 6×6 planar array. These results clearly illustrate that the channel observed across a large-scale array in the outdoor scenarios experiences significant changes in small-scale. Except for about 4% of the SS-clusters that exhibit LRs coincident with the whole aperture of the Rx array, other SS-clusters are only alive in small portions of the large-scale array.

Fig. 14 also illustrates the cdfs of an estimate $\hat{s}_{\rm LR}$ of the SS-cluster LR, which is calculated as $\hat{s}_{\rm LR} = d_{\rm v} \cdot d_{\rm h}$ for both LoS and NLoS scenarios. This estimate may closely approximate the true $s_{\rm LR}$ in the case where the distributions of the LR along vertical and horizontal axes are independent, i.e. the LR exhibits square or rectangular shapes without being tilted. It can be observed from Fig. 14 that compared with $P(s_{\rm LR} <$ abscissa), the cdf of $\hat{s}_{\rm LR}$ is more shifted to the left of abscissa in both LoS and NLoS scenarios. This indicates that the distributions of LR in vertical and horizontal dimensions are correlated, and as a result, the LRs exhibit certain tilted shapes, leading to the effect that $s_{\rm LR} \geq \hat{s}_{\rm LR}$.

C. CLUSTER-LEVEL SPREAD FOR SS-CLUSTERS

The intra-cluster path distribution in parameter space is necessary when generating paths randomly for individual clusters. In our case, the statistics of the spreads of each SS-cluster in the delay, AoA and EoA domains, as well as the dependence among the spreads in these domains are of interest for establishing GBSCMs in massive MIMO scenarios. In individual SS-clusters identified, many clusters are included which may have slightly different spreads. In such a case, we calculate the average $\mu_{\sigma_{(a)}}$ and the standard deviation $\delta_{\sigma_{(a)}}$ of the spreads of SS-cluster observed at multiple points through its LR, where a can be replaced by τ, ϕ and θ. With slight abuse of notations, these parameters are called as "SS-cluster spread" and "SS-cluster spread deviation", and used to characterize the spread of an SS-cluster when it is alive in the LR. Table 4 reports the mean and standard deviation of $\mu_{\sigma_{(a)}}$ and $\delta_{\sigma_{(a)}}$ for a being τ, ϕ and θ for the LoS and NLoS scenarios. It can be observed from Table 4 that the SS-clusters are more concentrated and the spreads are less variant in LoS scenarios than in NLoS scenarios.

TABLE 4. Statistical parameters for SS-clusters identified.

Parameters of SS-cluster spreads		LoS	NLoS
SS-cluster delay spread μ_{σ_τ} $[\log_{10}(\text{s})]$	Mean	-9.1	-8.7
	Std.	0.46	0.81
SS-cluster delay spread deviation δ_{σ_τ} $[\log_{10}(\text{s})]$	Mean	-9.45	-9.6
	Std.	0.10	0.71
SS-cluster AoA spread μ_{σ_ϕ} [°]	Mean	3.96	5.97
	Std.	3.01	2.00
SS-cluster AoA spread deviation δ_{σ_ϕ} [°]	Mean	1.97	1.99
	Std.	0.69	0.62
SS-cluster EoA spread μ_{σ_θ} [°]	Mean	0.71	4.70
	Std.	0.31	2.73
SS-cluster EoA spread deviation δ_{σ_θ} [°]	Mean	1.13	1.61
	Std.	0.72	1.11

D. DEPENDENCE OF CLUSTER SPREADS IN MULTIPLE DOMAINS FOR INDIVIDUAL SS-CLUSTERS

The dependence of cluster spreads among delay, AoA and EoA domains can be represented by the cross-correlation coefficients of the spreads, which is valuable for understanding the variations of an SS-cluster in its LR. By introducing vector $\boldsymbol{\zeta}_\sigma = [\sigma_{\text{c},\tau}, \sigma_{\text{c},\phi}, \sigma_{\text{c},\theta}]^{\rm T}$ with $(\cdot)^{\rm T}$ being the transpose operation to denote the spread vector of a cluster, the cross-correlation coefficients among the spreads for a specific SS-cluster in its LR are obtained by calculating the covariance matrix of $\boldsymbol{\zeta}_\sigma$ based on the observations of $\boldsymbol{\zeta}_\sigma$ across the LR. Fig. 15 illustrates the cdfs of cross-correlation coefficients $\rho_{\alpha\beta}$, with $\alpha \neq \beta, \alpha, \beta \in \boldsymbol{\zeta}_\sigma$ for 138 and 331 SS-clusters in LoS and NLoS scenarios respectively. It can be observed that the mean $\rho_{\alpha\beta}$'s are close to 0 in both scenarios, and the spreads of $\rho_{\alpha\beta}$ are wider in NLoS than in LoS scenario. Since the lager the absolute value of $\rho_{\alpha\beta}$, the higher is the dependence among the cluster's spreads across the LR, the observation of wider spread of $\rho_{\alpha\beta}$ in the NLoS scenario indicates that the variations of the spreads of an SS-cluster across its LR can have more complex modes in NLoS than in LoS. Specifically, with $\rho_{\alpha\beta} > 0$, an SS-cluster may shrink or expand in a and b domains at the same time with higher probablility; for $\rho_{\alpha\beta} = 0$, the spreads of an SS-cluster change independently, and with $\rho_{\alpha\beta} < 0$, the cluster may have an increasing spread in a when its spread in

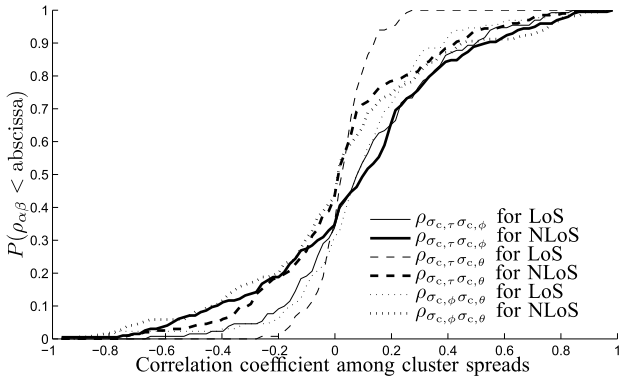


FIGURE 15. The empirical cdfs of the correlation coefficients among cluster-level delay spread, AoA spread and EoA spread in LoS and NLoS scenarios.

b decreases.

V. CONCLUSIONS

In this contribution, we introduced a recently conducted measurement campaign for characterizing channels observed with a 40×40 planar receiver (Rx) array at frequencies from 13 to 17 GHz in outdoor environments. The variations of channel parameters calculated by using sliding 7×7 planar sub-arrays were investigated. The results demonstrated that the K-factor, composite delay spread, azimuth of arrival and elevation of arrival spreads of the channel, as well as the number of multipath clusters vary in blocks with clear boundaries, which is more evident in the line-of-sight (LoS) than in non-line-of-sight (NLoS) scenarios. We introduced the concept of ''spatial-stationary (SS)-cluster'' to represent a cluster visible to multiple elements across the underlying large-scale Rx array. Statistics extracted based on hundreds of SS-clusters demonstrated that the average life distance of SS-cluster is about 1.5 wavelengths in both vertical and horizontal dimensions, and the average area of life region (LR) of SS-cluster is about 3 wavelengths squared in LoS and NLoS scenarios. The variations of SS-clusters in LR is more complicated in NLoS than LoS scenarios. From these observations, we may conclude that *i)* the spatial variations of channel characteristics are significant in massive channel modelling for 2-D array structures, and *ii)* introducing SS clusters in massive MIMO channel modeling is necessary to characterize the spatial non-stationarity of channels across antennas arrays with reduced model complexity. The analysis approaches and the model structure for massive MIMO channels presented here can also be applied to the channel measurements in other scenarios.

REFERENCES

[1] J. Hoydis, S. T. Brink, and M. Debbah, ''Massive MIMO in the UL/DL of cellular networks: How many antennas do we need?'' *IEEE J. Sel. Areas Commun.*, vol. 31, no. 2, pp. 160–171, Feb. 2013.

[2] S. K. Mohammed, A. Zaki, A. Chockalingam, and B. S. Rajan, ''High-rate spaceŰtime coded large-MIMO systems: Low-complexity detection and channel estimation,'' *IEEE J. Sel. Topics Signal Process.*, vol. 3, no. 6, pp. 958–974, Dec. 2009.

[3] K. V. Vardhan, S. K. Mohammed, A. Chockalingam, and B. S. Rajan, ''A low-complexity detector for large MIMO systems and multicarrier CDMA systems,'' *IEEE J. Sel. Areas Commun.*, vol. 26, no. 3, pp. 473–485, Apr. 2008.

[4] F. Rusek *et al.*, ''Scaling up MIMO: Opportunities and challenges with very large arrays,'' *IEEE Signal Process. Mag.*, vol. 30, no. 1, pp. 40–60, Jan. 2013.

[5] J. Chen, S. Wang, and X. Yin, ''A spherical-wavefront-based scatterer localization algorithm using large-scale antenna arrays,'' *IEEE Commun. Lett.*, vol. 20, no. 9, pp. 1796–1799, Sep. 2016.

[6] X. Yin and X. Cheng, *Propagation Channel Characterization, Parameter Estimation and Modeling for Wireless Communications*. Hoboken, NJ, USA: Wiley, 2016.

[7] H. Q. Ngo, E. G. Larsson, and T. L. Marzetta, ''Energy and spectral efficiency of very large multiuser MIMO systems,'' *IEEE Trans. Commun.*, vol. 61, no. 4, pp. 1436–1449, Apr. 2013.

[8] H. Jafarkhani, *Space-Time Coding: Theory and Practice*. Cambridge, U.K.: Cambridge Univ. Press, 2005.

[9] J. Nam, J. Y. Ahn, A. Adhikary, and G. Caire, ''Joint spatial division and multiplexing: Realizing massive MIMO gains with limited channel state information,'' in *Proc. 46th Annu. Conf. Inf. Sci. Syst.*, Mar. 2012, pp. 1–6.

[10] E. G. Larsson, O. Edfors, F. Tufvesson, and T. L. Marzetta, ''Massive MIMO for next generation wireless systems,'' *IEEE Commun. Mag.*, vol. 52, no. 2, pp. 186–195, Feb. 2014.

[11] B. Ai *et al.*, ''Future railway services-oriented mobile communications network,'' *IEEE Commun. Mag.*, vol. 53, no. 10, pp. 78–85, Oct. 2015.

[12] B. Ai *et al.*, ''Challenges toward wireless communications for high-speed railway,'' *IEEE Trans. Intell. Transp. Syst.*, vol. 15, no. 5, pp. 2143–2158, Oct. 2014.

[13] J. H. Sørensen, E. de Carvalho, and P. Popovski, ''Massive MIMO for crowd scenarios: A solution based on random access,'' in *Proc. IEEE Globecom Workshops*, Dec. 2014, pp. 352–357.

[14] S. Dierks, W. Zirwas, M. Jager, B. Panzner, and G. Kramer, ''MIMO and massive MIMO—Analysis for a local area scenario,'' in *Proc. 23rd IEEE Eur. Signal Process. Conf. (EUSIPCO)*, Aug. 2015, pp. 2451–2455.

[15] T. Jämsä and K. Kusume, ''Deliverable D1.2 initial channel models based on measurements,'' Project Name: Scenarios, Requirements and KPIs for 5G Mobile and Wireless System (METIS) document ICT-317669-METIS/D1.2, 2014.[Online]. Available: www.metis2020.com

[16] *IMT2020-5G-PG White Paper on 5G Vision and Requirements, V1.0*, Standard IMT-2020 Standard Promotion Group, China, 2014. [Online]. Available: http://www.imt-2020.cn/zh/documents/download/15

[17] J. G. Andrews *et al.*, ''What will 5G be?'' *IEEE J. Sel. Areas Commun.*, vol. 32, no. 6, pp. 1065–1082, Jun. 2014.

[18] S. Salous, V. Degli-Esposti, M. Nekovee, and S. Hur, ''Review of millimeter-wave propagation characterization and modelling towards 5G systems,'' in *Proc. Eur. Conf. Antenna Propag. (EuCAP)*, Lisbon, portgal, Apr. 2015

[19] W. C. Cheng, T. H. Liu, M. L. Hsu, and Z. M. Tsai, ''15 GHz propagation channel measurement at a university campus for the 5G spectrum,'' in *Proc. Asia–Pacific Microw. Conf. (APMC)*, Dec. 2015, pp. 1–3.

[20] J. Chen, X. Yin, and S. Wang, ''Measurement-based massive MIMO channel modeling in 13–17 GHz for indoor hall scenarios,'' in *Proc. IEEE Int. Conf. Commun. (ICC)*, May 2016, pp. 1–5.

[21] X. Yin, L. Ouyang, and H. Wang, ''Performance comparison of SAGE and MUSIC for channel estimation in direction-scan measurements,'' *IEEE Access*, vol. 4, pp. 1163–1174, 2016.

[22] X. Yin, C. Ling, and M. Kim, ''Experimental multipath-cluster characteristics of 28-GHz propagation channel,'' *IEEE Access*, vol. 3, pp. 3138–3150, 2015.

[23] Y. Azar *et al.*, ''28 GHz propagation measurements for outdoor cellular communications using steerable beam antennas in new york city,'' in *Proc. IEEE Int. Conf. Commun. (ICC)*, Jun. 2013, pp. 5143–5147.

[24] P. F. M. Smulders, ''Statistical characterization of 60-GHz indoor radio channels,'' *IEEE Trans. Antennas Propag.*, vol. 57, no. 10, pp. 2820–2829, Oct. 2009.

[25] H. Sawada, H. Nakase, S. Kato, and M. Umehira, ''Impulse response model and parameters for indoor channel modeling at 60 GHz,'' in *Proc. IEEE 71st Veh. Technol. Conf. (VTC)*, May 2010, pp. 1–5.

[26] N. Zhang, X. Yin, S. X. Lu, X. Cai, and M. Du, ''Measurement-based angular characterization for 72 GHz propagation channels in indoor environments,'' in *Proc. IEEE Globecom Workshops*, Dec. 2015, pp. 370–376.

[27] A. Karttunen, K. Haneda, J. Jarvelainen, and J. Putkonen, ''Polarisation characteristics of propagation paths in indoor 70 GHz channels,'' in *Proc. IEEE Eur. Conf. Antennas Propag. (EuCAP)*, Apr. 2015, pp. 1–4.

[28] (Dec. 2014). *NTT DoCoMo Preps for 5G With Mutielement Antenna*. [Online]. Available: http://www.fiercewireless.com/tech/story/ntt-docomo-preps-5g-multi-element-antenna/

[29] A. F. Molisch and M. Z. Win, "MIMO systems with antenna selection," *IEEE Commun. Mag.*, vol. 5, no. 1, pp. 46–56, Mar. 2004.

[30] S. Sanayei and A. Nosratinia, "Antenna selection in MIMO systems," *IEEE Commun. Mag.*, vol. 42, no. 10, pp. 68–73, Oct. 2004.

[31] X. Gao, O. Edfors, J. Liu, and F. Tufvesson, "Antenna selection in measured massive MIMO channels using convex optimization," in *Proc. IEEE Globecom Workshops Emerg. Technol. LTE-Adv. Beyond-4G*, Dec. 2013, pp. 129–134.

[32] B. H. Fleury, X. Yin, K. G. Rohbrandt, P. Jourdan, and A. Stucki, "Performance of a high-resolution scheme for joint estimation of delay and bidirection dispersion in the radio channel," in *Proc. IEEE 55th Veh. Technol. Conf. (VTC) Spring*, May 2002, pp. 522–526.

[33] T. Zwick, C. Fischer, and W. Wiesbeck, "A stochastic multipath channel model including path directions for indoor environments," *IEEE J. Sel. Areas Commun.*, vol. 20, no. 6, pp. 1178–1192, Aug. 2002.

[34] D. S. Baum, J. Hansen, and J. Salo, "An interim channel model for beyond-3G systems: Extending the 3GPP spatial channel model (SCM)," in *Proc. IEEE 61st Veh. Technol. Conf. (VTC) Spring*, vol. 5. May 2005, pp. 3132–3136.

[35] H. Xiao, A. Burr, and L. Song, "A time-variant wideband spatial channel model based on the 3GPP model," in *Proc. IEEE 64th Veh. Technol. Conf. (VTC) Fall*, Sep. 2006, pp. 1–5.

[36] Y. Xie, B. Li, X. Zuo, M. Yang, and Z. Yan, "A 3D geometry-based stochastic model for 5G massive MIMO channels," in *Proc. IEEE 11th Int. Conf. Heterogeneous Netw. Quality, Rel., Secur. Robustness (QSHINE)*, Aug. 2015, pp. 216–222.

[37] X. Li, S. Zhou, and E. Bjornson, and J. Wang, "Capacity analysis for spatially non-wide sense stationary uplink massive MIMO systems," *IEEE Trans. Wireless Commun.*, vol. 14, no. 12, pp. 7044–7056, Dec. 2015.

[38] *Universal Mobile Telecommun. System (UMTS);Spacial Channel Model for Multiple Input Multiple Output (MIMO) Simulations 3GPP TR 25.996 version 8.0.0 Release 8)*, 2003.

[39] *WINNER II interim channel models*, standard IST-4-027756 WINNER D1.1.1 Std.

[40] H. Asplund, A. Glazunov, A. Molisch, K. Pedersen, and M. Steinbauer, "The COST 259 directional channel model—Part II: Macrocells," *IEEE Trans. Wireless Commun.*, vol. 5, no. 12, pp. 3434–3450, Dec. 2006.

[41] N. Czink and C. Oestges, "The COST 273 MIMO channel model: Three kinds of clusters," in *Proc. IEEE 10th Int. Symp. Spread Spectrum Techn. Appl. (ISSSTA)*, Aug. 2008, pp. 282–286.

[42] L. Liu *et al.*, "The COST 2100 MIMO channel model," *IEEE Wireless Commun.*, vol. 19, no. 6, pp. 92–99, Dec. 2012.

[43] M. Zhu, G. Eriksson, and F. Tufvesson, "The COST 2100 channel model: Parameterization and validation based on outdoor MIMO measurements at 300 MHz," *IEEE Trans. Wireless Commun.*, vol. 12, no. 2, pp. 888–897, Feb. 2013.

[44] 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.

[45] A. Richter, M. Landmann, and R. S. Thoma, "Maximum likelihood channel parameter estimation from multidimensional channel sounding measurements," in *Proc. IEEE 57th Veh. Technol. Conf. (VTC)*, vol. 2, Aug. 2003, pp. 1056–1060.

[46] J. Hoydis, C. Hoek, T. Wild, and S. T. Brink, "Channel measurements for large antenna arrays," in *Proc. IEEE Int. Symp. Wireless Commun. Syst. (ISWCS)*, Aug. 2012, pp. 811–815.

[47] X. Fang, S. Fang, N. Ying, H. Cao, and C. Liu, "The performance of massive MIMO systems under correlated channel," in *Proc. 19th IEEE Int. Conf. Netw. (ICON)*, Dec. 2013, pp. 1–4.

[48] X. Gao, F. Tufvesson, O. Edfors, and F. Rusek, "Measured propagation characteristics for very-large MIMO at 2.6 GHz," in *Proc. IEEE Conf. Rec. 46th Asilomar Conf. Signals, Syst. Comput. (ASILOMAR)*, Nov. 2012, pp. 295–299.

[49] X. Gao, F. Tufvesson, and O. Edfors, "Massive MIMO channels; measurements and models," in *Proc. IEEE Asilomar Conf. Signals, Syst. Comput.*, Nov. 2013, pp. 280–284.

[50] X. Gao, O. Edfors, F. Rusek, and F. Tufvesson, "Massive MIMO performance evaluation based on measured propagation data," *IEEE Trans. Wireless Commun.*, vol. 14, no. 7, pp. 3899–3911, Jul. 2015.

[51] S. Payami and F. Tufvesson, "Channel measurements and analysis for very large array systems at 2.6 GHz," in *Proc. IEEE 6th Eur. Conf. Antennas Propag. (EUCAP)*, Mar. 2012, pp. 433–437.

[52] A. Manikas and C. Proukakis, "Modeling and estimation of ambiguities in linear arrays," *IEEE Trans. Signal Process.*, vol. 46, no. 8, pp. 2166–2179, Aug. 1998.

[53] B. H. Fleury, "First–and second-order characterization of direction dispersion and space selectivity in the radio channel," *IEEE Trans. Inf. Theory*, vol. 46, no. 6, pp. 2027–2044, Sep. 2000.

[54] M. Landmann, M. Kaske, and R. S. Thoma, "Impact of incomplete and inaccurate data models on high resolution parameter estimation in multi-dimensional channel sounding," *IEEE Trans. Antennas Propag.*, vol. 60, no. 2, pp. 557–573, Feb. 2012.

[55] M. S. Babtlett, "Smoothing periodograms from time-series with continuous spectra," *Nature*, vol. 161, pp. 686–687, May 1948.

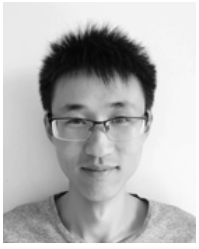
[56] B. H. Fleury, P. Jourdan, and A. Stucki, "High-resolution channel parameter estimation for MIMO applications using the SAGE algorithm," in *Proc. Int. Zurich Seminar Broadband Commun.*, vol. 30. Feb. 2002, pp. 1–9.

[57] N. Czink, P. Cera, J. Salo, E. Bonek, J.-P. 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)*, Sep. 2006, pp. 1–5.

[58] F. Fuschini, E. M. Vitucci, M. Barbiroli, G. Falciasecca, and V. Degli-Esposti, "Ray tracing propagation modeling for future small-cell and indoor applications: A review of current techniques," *Radio Sci.*, vol. 50, no. 6, pp. 469–485, Jun. 2015.

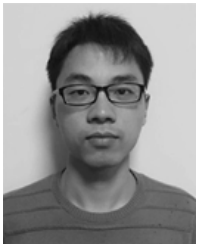
[59] A. Goldsmith, *Wireless Communications*. Cambridge, U.K.: Cambridge Univ. Press, 2005.

[60] N. Czink, P. Cera, J. Salo, E. Bonek, J.-P. Nuutinen, and J. Ylitalo, "Improving clustering performance using multipath component distance," *Electron. Lett.*, vol. 42, no. 1, pp. 33–35, Jan. 2006.



JIAJING CHEN (S'14) received the B.S. degree in electronics science and technology from Tongji University, Shanghai, China, in 2013, where he is currently pursuing the Ph.D. degree with the College of Electronics and Information. He has been a Visiting Ph.D. Student with the Department of Electronics and Information Systems, University of Bologna, Bologna, Italy, since 2016. His research interests include wideband channel measurement, and modeling.

XUEFENG YIN (S'01–M'06) received the Bachelor degree in optoelectronics engineering from the Huazhong University of Science and Technology, Wuhan, China, in 1995, and the M.Sc. degree in digital communications in 2002 and the Ph.D. degree in wireless communications from Aalborg University, Denmark, in 2006. From 2006 to 2008, he was an Assistant Professor with Aalborg University. In 2008, he joined as an Associate Professor with the College Of Electronics And Information Engineering, Tongji University, Shanghai, China, where he is currently a Vice Dean. Since 2017, he was promoted as a Full Professor. His research interests include high-resolution parameter estimation for propagation channels, channel characterization and stochastic modeling for 5G wireless communications, radar signal processing, and target recognition.



XUESONG CAI (S'14) received the B.S. degree in electronics science and technology from Tongji University, Shanghai, China, in 2013, where he is currently pursuing the Ph.D. degree with the College of Electronics and Information. His research interests are passive channel sounding, time-variant channel characterization, and modeling.

STEPHEN WANG (S'04–M'09–SM'14) received the M.Sc. degree (Hons.) in advanced photonics and communications from the University of Warwick, Warwick, U.K., in 2005, and the Ph.D. degree in cognitive radio and radar systems from the University of Bristol, Bristol, U.K., in 2009.

He served as a Wireless Engineer with China Telecom and Nokia, respectively. He then joined the Shanghai Research Center for Wireless Communications as a Senior Researcher Engineer and Team Leader in a collaborative research project with Nokia/NSN, coordinate related R&D activities in 3GPP LTE-HeNB standardization activities. From 2010 to 2016, he was with Toshiba Research Europe Ltd., Telecommunications Research Laboratory, as a Senior Research Engineer, leading several research projects in mobile body-area networks, 5G massive MIMO. From 2017, as an IT Team Leader with Ocado Technology, he is currently leading wireless IoT technology and business for Ocado's IoT-enabled fulfilment center and logistics. He holds ten patents. His research interests include industrial IoT, water resource/meter management and optimization, cognitive radio, energy-efficient resource optimization, channel propagation and modeling, and wearable heath care applications.

Dr. Wang is a U.K. Chartered Engineer. He is a Regular Reviewer of IEEE journals, a Guest Editor and the Session Chair of flagship conferences, a General Co-chair of Wireless Internet: Eighth International Conference in 2014.