IEEE Access
Multidisciplinary | Rapid Review | Open Access Journal

Received October 31, 2017, accepted December 23, 2017, date of publication January 24, 2018, date of current version March 28, 2018.

Digital Object Identifier 10.1109/ACCESS.2018.2797418

Cross-Correlation of Large-Scale Parameters in Multi-Link Systems: Analysis Using the Box-Cox Transformation

GHASSAN DAHMAN[1], **JOSE FLORDELIS**[2], **FREDRIK TUFVESSON**[2]
[1]École de Technologie Supérieure, Montreal, QC H3C 1K3, Canada
[2]Department of Electrical and Information Technology, Lund University, 221 00 Lund, Sweden
Corresponding author: Ghassan Dahman (ghassan.dahman@lacime.etsmtl.ca)

ABSTRACT Spatially distributed transmission points connected to the same source, known as distributed antenna systems, can improve system performance compared with single-link traditional systems. However, the anticipated gain depends heavily on the cross-correlation properties of the large-scale parameters (LSPs) of the different links. Usually, measured LSPs—except the large-scale fading—have non-Gaussian distributions. Therefore, in order to study their multi-link cross-correlation properties, scenario- and parameter-specific ad-hoc transformations are applied, such that the LSPs have Gaussian distributions in the transform domain [1], [2]. In this paper, we propose using the Box-Cox transformation as a general framework for homogenizing this conversion step. The Box-Cox transformation is, by nature, not distribution specific; therefore, it can be used regardless of the empirical distributions of the studied LSPs. We demonstrate the applicability of the proposed framework by studying multi-link fully-coherent propagation measurements of four base stations and one mobile station in a suburban microcell environment at 2.6 GHz. The inter- and intra-link cross-correlation of four LSPs—the large-scale fading, the delay, azimuth, and elevation spreads—are analyzed and their distributions are modeled. Based on our analysis, it is found that for the investigated environment: 1) the LSPs of the different links can be modeled using unimodal and bimodal Gaussian distributions; and 2) the inter- and intra-link cross-correlation coefficients of the different studied LSPs can be modeled using the Truncated Gaussian distribution. The proposed models are validated using the Kolmogorov–Smirnov test, and their parameters are provided.

INDEX TERMS Distributed antenna systems, inter-link cross-correlation, intra-link cross-correlation, large-scale parameters, multi-link systems.

I. INTRODUCTION

Multi-link communications are at the core of many recently introduced concepts aiming at improving the performance of future communication systems. Using multiple links enables the exploitation of the spatial distribution of the communication nodes over the geographic area of interest. This in turn allows the system to process the signals of the multiple links jointly and to improve the quality of communications, where the performance of multi-link systems is conditioned upon the parameters of the individual links as well as their joint statistical properties [3], [4]. Modeling the multi-link cross-correlation characteristics is a key element for understanding the joint behavior of the parameters of the different links and it is therefore essential for the development and evaluation of new wireless communication systems.

A. RELATED WORK

The inter-link and intra-link correlations of the LSPs have been investigated in several studies. It was found in [5]–[8] that the inter-link correlation of the large-scale fading (LSF) may vary anywhere from very small negligible values to significant positive values. On the other hand, the inter-link correlation of the delay spread was reported to be insignificant [7]. It was also found that there is negative correlation between the LSF and the angular spread of the same link [9], [10]. In [10] and [11], it was reported that, in general, the inter-link LSP correlation decreases with increasing base station (BS) separation angle and/or relative distance. In [12], the effect of BS antenna heights and distance ranges on the cross-correlation properties of different links have been reported. In our previous work [8], the cross-correlation

properties of the LSF among different links were studied and a preliminary model was proposed.

Generating channel realizations that have specific correlation characteristics is implemented using different approaches. In the COST 2100 channel model [13], the concept of common clusters is introduced in order to model the correlation properties among different links. Where common clusters—representing the effect of the common interacting objects among the different links in the propagation environment—contribute to more than one link and may cause these links to exhibit a certain correlation [4]. In the WINNER channel model [1], [2], the correlation of the LSPs is simulated by generating independent Gaussian random processes representing the LSPs and then establishing their dependence through correlating them. Since, almost all measured/modeled LSPs (except the LSF, as will be clarified later) have non-Gaussian distributions, the WINNER model proposes, after generating the correlated Gaussian realizations, to transform each LSP using a specific scenario- and parameter-specific transformation. All the results and methods of this work are proposed having in mind the WINNER's model approach of simulating the correlation among the different links, i.e., generating correlated random processes to model the dependence of the LSPs of the different links.

B. CONTRIBUTIONS

In our previous work [8], the cross-correlation properties of the LSF among different links were studied and a preliminary model was proposed. The LSF has a Gaussian distribution (in the log domain) [14]. This property makes modeling its multi-link cross-correlation properties a simple and a straightforward task.[1] However, in [8], we failed to propose models characterizing the distributions and the cross-correlation of the other measured LSPs (namely, the delay, azimuth, and elevation spreads) due to the fact that their estimates have unknown distributions. This work extends the results of [8] in two main aspects.

Firstly, we provide a general framework that simplifies the modeling of the distributions as well as the inter-link and intra-link cross-correlation properties of the LSPs. The proposed framework uses the Box-Cox transformation as a tool for measurement-based statistical modeling of LSPs. The motivation behind our framework is the simplicity of modeling such properties when the studied parameters have Gaussian distributions. Therefore, we propose to apply the Box-Cox transformation in order to transform the measured non-normally distributed LSPs to another domain where they have approximately Gaussian distributions. Then, in the transform domain, we study the distributions and the cross-correlation properties of the LSPs of interest. In the WINNER model [1], these steps are implemented based on using scenario- and parameter-specific transformations where: 1) to map the LSPs realizations, the inverse of the cumulative distribution function (CDF) of the Gaussian random variable is applied to the Log-Gumble, Log-Logistic and Log-Normal distributions, and 2) in cases when the required mapping is unknown, this transformation step is implemented based on applying the Q-function to the empirical CDF of the modeled LSP. The framework we propose in this paper is simpler and more general as the Box-Cox transformation is by nature not distribution specific.

Secondly, based on exhaustive multi-site fully-coherent wideband propagation measurements with four BSs, we provide: 1) Gaussian models (unimodal and bimodal) to capture the distributions of four different LSPs (LSF, delay, azimuth, and elevation spreads), and 2) Truncated Gaussian models that capture the inter-link and intra-link cross-correlation behavior of the four studied LSPs. All proposed models are verified using the Kolmogorov-Smirnov (KS) test and their parameters are provided. We emphasize that, although the proposed modeling framework is general, the extracted parameters from our measurements can only be used to model scenarios that have similar propagation characteristics, i.e., suburban microcell environments as detailed in section III. On the other hand, some of our findings —in particular, the effect of interacting objects located close to BS, which give rise to a bimodal behavior on the LSPs distributions—carry over to general multi-link scenarios with full or partial blockage of the signal.

C. ORGANIZATION

The remainder of this paper is organized as follows. In section II, the formulation of the Box-Cox transformation is introduced. In section III, the measurement setup and the measurement scenario are described. In section IV, the methods used to estimate and prepare the LSPs for further analysis are detailed. Results on modeling the distributions of the LSPs of the different links, the inter-link and intra-link cross-correlation properties of the LSPs are reported in section V. Finally, the main findings are summarized in section VI.

II. USING THE BOX-COX TRANSFORMATION

To simplify the study and the modeling of the distributions and the cross-correlations of the different LSPs, we would like to first find a simple way to transform them to a domain where they have Gaussian distributions. Then, in the transform domain, we estimate the parameters of the corresponding Gaussian distributions and the cross-correlation coefficients of the transformed LSPs. In this work we suggest to utilize the Box-Cox transformation in order to simplify these modeling steps. The Box-Cox transformation belongs to the family of power transformations and is defined by [15]:

$$g(x)^{(\Lambda)} = \begin{cases} \dfrac{x^{\Lambda} - 1}{\Lambda} & \text{if } \Lambda \neq 0 \\ \log(x) & \text{if } \Lambda = 0 \end{cases} \quad (1)$$

where x is the LSP under transformation, Λ is the power parameter of the Box-Cox transformation indicating the

[1]Modeling the LSF as a log-normal distribution is a special case of applying the general transformation we are proposing, i.e., the Box-Cox transformation as can be seen in (1).

power to which all data should be raised. $g(x)$ is the corresponding value of x in the transform domain with approximately Gaussian distribution.

It is important to mention that the Box-Cox transformation does not guarantee successful transformation to normality. Instead, it searches for the best value of the power parameter (Λ) such that the transformed data has the highest likelihood of being normally distributed. In this work, successful transformation to normality is verified by using the KS test [16], a statistical nonparametric test to check the goodness-of-fit of the transformed data against unimodal as well as bimodal Gaussian distributions. The KS statistic quantifies a distance (D_{LSP}) between the empirical CDF of the sample under test and the CDF of the suggested model. More explicitly,

$$D_{\mathrm{LSP}} = \max_x |F(x) - Z_{\mathrm{LSP}}(x)| \tag{2}$$

where, $F(x)$ is the CDF of the proposed model, and $Z_{\mathrm{LSP}}(x)$ is the empirical CDF of the LSP of interest.

If D_{LSP} is found to be less than a certain defined critical value that depends on a predefined significance level of the KS test, then $F(x)$ is considered to successfully pass the KS test at the specified significance level. In this work, we perform the KS test at significance level $\alpha = 0.05$. The corresponding critical values for D_{LSP} are $1.36/\sqrt{N}$, where N is the number of the independent sets of the sample under test.

III. MULTI-LINK PROPAGATION MEASUREMENTS

The measurement campaigns were carried out with the RUSK LUND channel sounder [17] at 2.6 GHz and a measurement bandwidth of 40 MHz. The measurements took place at the campus of the faculty of Engineering, LTH, Lund University, Lund, Sweden. The measurement environment can be best characterized as a suburban microcell environment. The chosen setup consists of four transmit BSs, each equipped with a single vertically polarized antenna element.

The sounding signal is brought to each of the remote BS locations using radio-over-fiber transceivers. After having been amplified to a power of 30 dBm, the signal is broadcasted from each BS antenna element. The signal broadcasted by the BSs is received by a mobile station (MS) equipped with 64 dual-polarized antenna elements arranged in a cylindrical configuration. The 512 (4 BSs $\times$ 128 MS antenna elements) transmit-receive channels are sounded in a time-multiplexed fashion. The data resulting from this operation is referred to as a snapshot. Please refer to [18] for more details about the equipment used. The transmit antennas were located outside the windows at the second and third floors of four different buildings, which correspond to 5 to 12 m above the ground level (10 to 20 m below the rooftop of surrounding buildings). The distances among the different antenna sites are between 60 to 200 m. The area in the middle of the selected buildings is an open area with a small lake surrounded by trees, as shown in Fig. 1. The measurements took place along a predefined route circulating the lake, with a total length of about 490 m,

FIGURE 1. Aerial photo of the measurement area. Base station locations are indicated by labels BS-E (Latitude: 55.711131, Longitude: 13.210225), BS-S, BS-F and BS-M. The measurement route is plotted in blue color.

TABLE 1. Differences between the two measurement setups.

	Campaign 1	Campaign 2
No. of Tx elements at BSs [2]	1	4
Sounding signal duration	12.8 μs	3.2 μs
Trigger rate[3]	$1/\lambda_c$	$1/2\lambda_c$
Total snapshots	4200	8455

at a very low walking speed (<0.5 m/s). Depending on the propagation campaign, the sounder was wheel triggered at one snapshot per either one or half a wavelength as detailed in Table 1. In this work, given that we use a distance-depending triggering mechanism, the collected snapshots are used to study the spatial variability of the characteristics of the LSPs as the MS moves on its trajectory and no discussion on the temporal behavior of the channel is addressed.

The propagation data was collected in two measurement campaigns. All the reported results and proposed models in this work are based on analyzing the data of the two campaigns collectively. Both campaigns took place in the same area, with the same BS locations, and the same MS measurement route. The main differences between the setup of the two campaigns are listed in Table 1. The propagation conditions between the BSs and the MS can be described as obstructed line-of-sight (OLOS), or non-line-of-sight (NLOS) due to the trees in the measurement area. It should be noted that the intensity of the tree blockage varies from one MS position to another. Line-of-sight (LOS) propagation condition may occur when the MS is sufficiently close to one of the BSs. When the MS is far from the BSs, there were still possibilities of getting the optical LOS cleared between the MS and the

[2]In Campaign 2, only one Tx element is considered for the analysis in this work.

[3]Distance-based trigger. λ_c is the wavelength at the center frequency.

different BSs. However, by detailed inspection, it was found that the first Fresnel zone at these locations was not cleared. The BSs are named BS-E, BS-S, BS-F and BS-M (see Fig. 1); and their corresponding links with the MS are named E-Link, S-Link, F-Link, and M-Link, respectively.

IV. DATA ANALYSIS

A. PREPROCESSING

The raw data obtained from the measurements consist of the single-input multiple-output (SIMO) transfer functions of the different links' channels. After compensating for the differences in gain among the different links, the *links impulse response estimates* (Link-IREs) are obtained. To mitigate the effect of noise, thresholding based on [19] is applied, where multipath echoes are declared valid in a specific delay bin with a probability of false-alarm of one per 5000 snapshots per link. Also, the Link-IREs are subjected to a delay-gating filter, which is implemented using a 700 m delay-window. This filter eliminates all multipath components that are 700 m in excess of the Tx-Rx separation. Following the recommendations of [20], Link-IREs with dynamic ranges less than 18 dB are discarded from any further processing, and the rest are declared as *valid* Link-IREs. Then, Space-Alternating Generalized Expectation-Maximization (SAGE) algorithm [21], [22], is applied to each valid Link-IRE in order to extract the angle-of-arrival (both azimuth and elevation), complex amplitude, and delay of each multipath component.

B. ESTIMATION OF THE LSPs FOR EACH LINK

The multipath components resulting from SAGE are used to estimate the LSPs of the different links using these steps:

1) The measurement route is divided into consecutive non-overlapping *segments*. In order to extract and study the LSPs, we have first to eliminate the effect of small-scale fading on the extracted parameters. This is done by averaging the channels' power profiles over spatial windows (i.e., segments) each with a length of few/several wavelengths, where the appropriate length of the averaging window depends on the characteristics of the measured propagation environment. Based on the analysis detailed in [23], we found that, for our measured environment, choosing the segments length to be 5 wavelengths at the carrier frequency (thus, each segment corresponds to 5 snapshots of Campaign 1, or 10 snapshots of Campaign 2, see Table 1) is appropriate to estimate the LSPs of each link. At each segment, the group of Link-IREs that belong to a specific link are called a *Link-Segment*.
2) For each segment, if all the Link-IREs pertaining to the same Link-Segment are valid (i.e., each has a dynamic range greater than 18 dB), then it is defined as a *valid Link-Segment*. For each link, the total number of measured Link-Segments is 1685. The total number of the *valid* Link-Segments for the E-Link, S-Link, F-Link, and M-Link are 1331, 1232, 1605, and 985, respectively.
3) For each valid Link-Segment, the SAGE multipath components that belong to each link are used to obtain one *set of average power profiles*, each consisting of: an Average Power Delay Profile (APDP), an Average Power Azimuth Profile (APAP), and an Average Power Elevation Profile (APEP). This step results in 1331, 1232, 1605, and 985 sets of average power profiles for the E-Link, S-Link, F-Link, and M-Link, respectively. Then, each set of average power profiles is used to estimate one set of the LSPs consisting of: LSF, rms delay spread (τ_{rms}), rms azimuth spread (ϕ_{rms}), and rms elevation spread (θ_{rms}). The procedure of extracting the LSPs is detailed in the next paragraphs.

The LSF is defined as the power fluctuation over a large area where the small-scale fading is averaged out. To extract the LSF for each link, the following steps are followed. First, for each Link-Segment, its average received power is estimated by integrating the power of its APDP over all the delay bins. Then, the obtained average received power values are plotted against the distances from the MS to the corresponding BS in double logarithmic scale. Then, for each link, and for each measurement campaign, a linear regression is performed according to

$$P_{\text{r}}(d)_{\text{dB}} = P_{\text{r}}(d_0)_{\text{dB}} - n10\log_{10}(d/d_0), \tag{3}$$

where $P_{\text{r}}(d)_{\text{dB}}$ is the distance-dependent mean received power at distance d , d_0 is a reference distance, and $P_{\text{r}}(d_0)_{\text{dB}}$ is the received power at d_0.

The power decay exponent that minimizes the distance between the data samples and the line in the least squares sense is $n = 3.89$. Due to the height difference among the different BSs, a different offset (i.e., different value for the term $P_{\text{r}}(d_0)_{\text{dB}}$) is calculated for each link in order to find the best fit according to (3). Fig. 2 depicts the received power estimates from Campaign 1 and their best linear fit for each link. Please notice that in Fig. 2, because the measurement route has a round shape, there are two black curves representing the received power of each link instead of a single curve (i.e., for each BS, there are two MS positions within the route that have the same BS-MS distance).

Finally, the LSF is estimated as the difference, in dBs, between the received power (with the small-scale fading averaged out) and $P_{\text{r}}(d)_{\text{dB}}$, (3). Fig. 3 illustrates the obtained LSF values from Campaign 1 for the four different links throughout the whole measurement route plotted in color scale. τ_{rms} (ns), ϕ_{rms} (deg.), and θ_{rms} (deg.) are calculated as the second-order central moments of their corresponding average power profiles [24, p. 113 and 122]. This step results in 1331, 1232, 1605, and 985 sets of LSPs for the E-Link, S-Link, F-Link, and M-Link, respectively.

V. RESULTS

The results of this paper are organized as follows. In section V-A, the values of the power parameter (Λ) of the

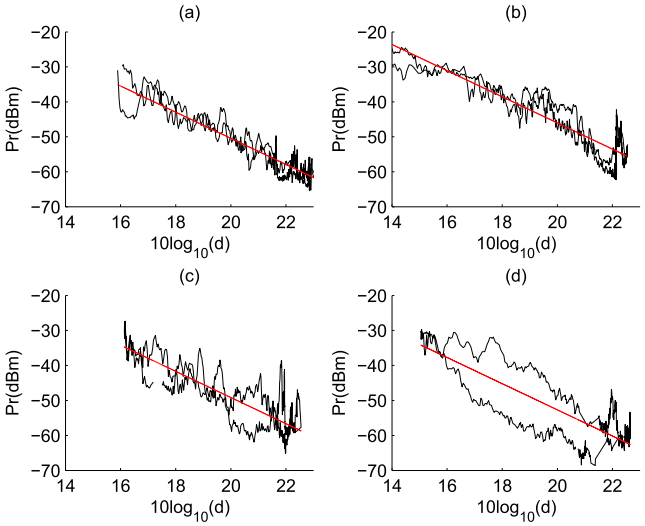

FIGURE 2. Received power as a function of distance for the four different BS links extracted from Campaign 1. (a) E-link, (b) S-Link, (c) F-Link, and (d) M-Link. Received power (black), and linear regression line (red).

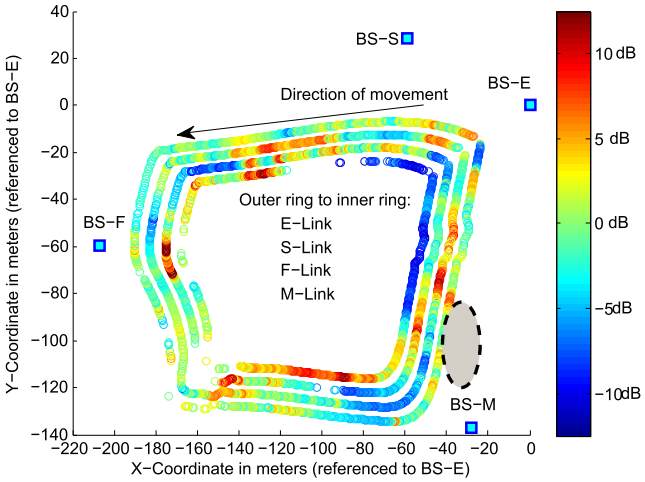

FIGURE 3. LSF estimates for the four different links throughout the whole measurement route extracted from Campaign 1. Each ring represents the LSF estimates of a single link. Missing points correspond to the invalid Link-Segments. Base station locations are indicated by labels BS-E, BS-S, BS-F and BS-M. The gray ellipse represents a group of big trees that affects the M-Link behavior as will be discussed in section V-A.

Box-Cox transformation that is applied to the different LSPs are estimated. Then, the empirical CDFs of the transformed LSPs of the different links are modeled. In section V-B, models for the inter-link and intra-link cross-correlation of the LSPs are proposed. In section V-C, the steps suggested to use the proposed models are listed.

A. MODELING THE EMPIRICAL CDFs OF THE LSPs OF INDIVIDUAL LINKS

The estimation procedures in the previous section result in 1331, 1232, 1605, and 985 sets of LSPs estimates for the E-Link, S-Link, F-Link, and M-Link, respectively. Each set of LSPs includes estimates of the LSF, τ_{rms}, ϕ_{rms}, and θ_{rms} for

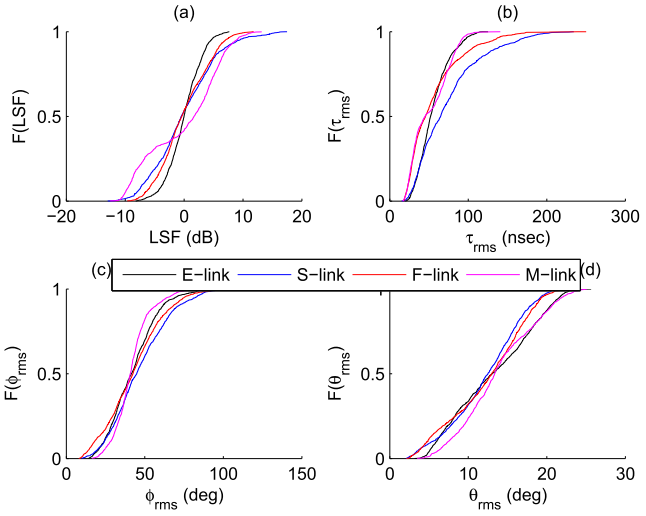

FIGURE 4. Empirical CDFs of the estimated LSPs of the different links. (a) LSF, (b) τ_{rms}, (c) ϕ_{rms}, and (d) θ_{rms}.

a particular valid Link-Segment. Fig. 4 depicts the empirical CDFs of the LSPs of the different links.

It is clear from Fig. 4 that the distributions of any LSP vary significantly from one link to another. However, for simplicity it is widely practiced to collect all the estimates of a specific LSP from all different links and model its distribution using one set of parameters. Our aim is to find appropriate models that describe the empirical CDFs of the different LSPs following two approaches.

- ***Approach 1***: modeling the empirical CDF of the LSP of interest pertaining to all links, where the estimates of the LSP of interest from the four different links are gathered and modeled collectively.
- ***Approach 2***: modeling the empirical CDF of the LSP of interest of each individual link separately.

It is obvious that Approach 1 reduces the number of modeling parameters; however, Approach 2 can capture the differences among the different links, if any. The Box-Cox-transformed delay, azimuth, and elevation spreads are denoted $g(\tau_{\text{rms}})$, $g(\phi_{\text{rms}})$, and $g(\theta_{\text{rms}})$, respectively. The values of their corresponding Box-Cox power parameters are denoted Λ_τ, Λ_ϕ, and Λ_θ. See Table 2. Please notice that there is no need to perform the Box-Cox transformation for the LSF because calculating the LSF (following the steps detailed in section IV-B) is a special case of applying the Box-Cox transformation with $\Lambda_{\text{LSF}} = 0$.

Figs. 5–8 illustrate the empirical CDFs of the transformed LSPs: LSF, $g(\tau_{\text{rms}})$, $g(\phi_{\text{rms}})$, and $g(\theta_{\text{rms}})$, along with their unimodal and bimodal Gaussian fits. Subfigures (a) in Figs. 5–8 correspond to Approach 1, where the estimates of the LSPs collected from all links are modeled together using one Gaussian distribution. The rest of the subfigures in Figs. 5–8 correspond to Approach 2, where the LSPs of each link are modeled individually. With both approaches, the empirical CDF of each LSP is modeled as follows. First, the transformed LSP is modeled using a unimodal

TABLE 2. The values of the power parameter (Λ) used in the Box-Cox transformation.

	Λ_τ	Λ_ϕ	Λ_θ
Approach 1 (all links)	-0.14	0.50	0.90
Approach 2 (E-Link)	0.07	0.35	0.65
Approach 2 (S-Link)	-0.07	0.47	1.10
Approach 2 (F-Link)	-0.34	0.60	0.65
Approach 2 (M-Link)	0.11	0.20	0.65
Approach 2 (min. value among the 4 links)	-0.34	0.20	0.65
Approach 2 (max. value among the 4 links)	0.11	0.60	1.10

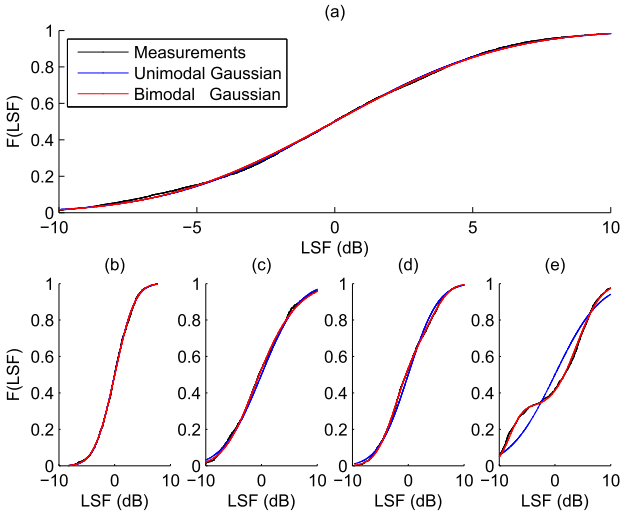

FIGURE 5. Empirical CDFs of the LSF and their unimodal and bimodal Gaussian fits. (a) the four links together, (b) E-link, (c) S-Link, (d) F-Link, and (e) M-Link.

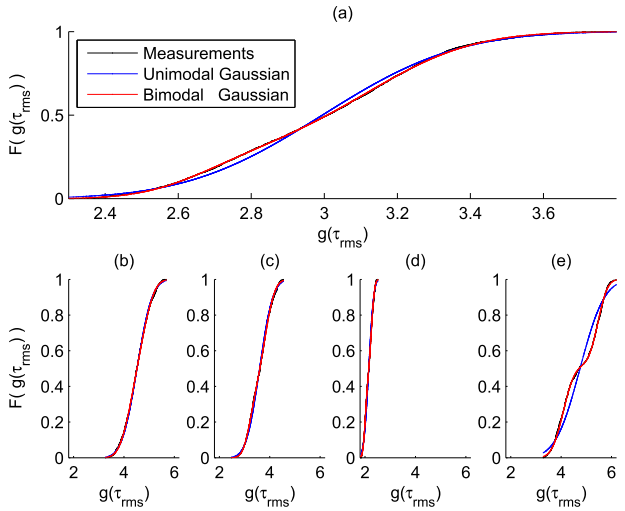

FIGURE 6. Empirical CDFs of $g(\tau_{\mathbf{rms}})$ and their unimodal and bimodal Gaussian fits. (a) the four links together, (b) E-link, (c) S-Link, (d) F-Link, and (e) M-Link.

Gaussian distribution, and then the KS test is performed. If the unimodal Gaussian distribution passes the KS test, then the unimodal Gaussian distribution is adopted as a suitable model. Otherwise, the transformed LSP is modeled using

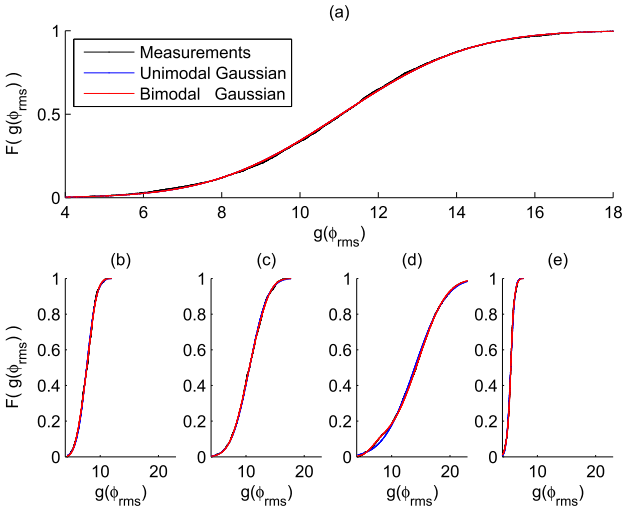

FIGURE 7. Empirical CDFs of $g(\phi_{\mathbf{rms}})$ and their unimodal and bimodal Gaussian fits. (a) the four links together, (b) E-link, (c) S-Link, (d) F-Link, and (e) M-Link.

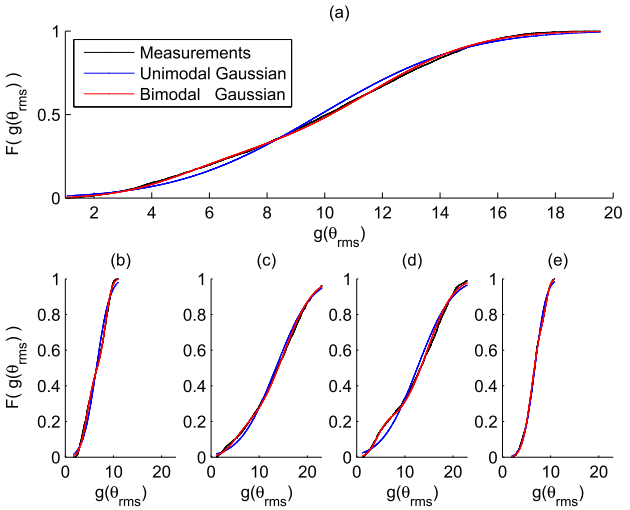

FIGURE 8. Empirical CDFs of $g(\theta_{\mathbf{rms}})$ and their unimodal and bimodal Gaussian fits. (a) the four links together, (b) E-link, (c) S-Link, (d) F-Link, and (e) M-Link.

the bimodal Gaussian distribution,[4] and again the KS test is performed. It is found that, in all cases, the bimodal Gaussian distribution passes the KS test. Throughout this work, we perform the KS test at significance level $\alpha = 0.05$. The corresponding critical values for D_{LSP} are $1.36/\sqrt{N}$, where N is the number of the sets of LSPs of each link which are assumed to be independent since N represents the number of the disjoint segments. Consequently, the values of D_{LSP} are: 0.0373, 0.0387, 0.0339, and 0.0433 for the E-Link, S-Link, F-Link, and M-Link, respectively. The detailed parameters of the unimodal and bimodal Gaussian models are listed in Tables 3-4.

[4]In this case, the LSP is modeled as a mixture of two normally distributed random variables, where: the probability of selecting the first (second) random variable is p_1 (1-p_1), and the mean and standard deviation of the first (second) random variable are μ_1 (μ_2), and σ_1 (σ_2), respectively. See Tables 3-4

TABLE 3. The parameters of the unimodal and bimodal gaussian models for approach 1 (all links together).

	Unimodal		Bimodal Model				
	μ_1	σ_1	p_1	μ_1	σ_1	μ_2	σ_2
LSF	0	4.73	-	-	-	-	-
$g(\tau_{\text{rms}})$	-	-	0.74	3.12	0.23	2.66	0.14
$g(\phi_{\text{rms}})$	11.06	2.60	-	-	-	-	-
$g(\theta_{\text{rms}})$	-	-	0.75	11.48	2.98	4.95	1.89

TABLE 4. The parameters of the unimodal and bimodal gaussian models for approach 2 (individual links).

		Unimodal		Bimodal Model				
		μ_1	σ_1	p_1	μ_1	σ_1	μ_2	σ_2
E-Link	LSF	0	2.85	-	-	-	-	-
	$g(\tau_{\text{rms}})$	4.51	0.47	-	-	-	-	-
	$g(\phi_{\text{rms}})$	7.66	1.35	-	-	-	-	-
	$g(\theta_{\text{rms}})$	-	-	0.49	8.38	1.04	4.62	1.31
S-Link	LSF	0	5.39	-	-	-	-	-
	$g(\tau_{\text{rms}})$	-	-	0.64	3.82	0.30	3.20	0.23
	$g(\phi_{\text{rms}})$	10.61	2.47	-	-	-	-	-
	$g(\theta_{\text{rms}})$	13.46	5.87	-	-	-	-	-
F-Link	LSF	-	-	0.66	-2.28	2.86	4.42	2.58
	$g(\tau_{\text{rms}})$	-	-	0.57	2.25	0.10	2.00	0.07
	$g(\phi_{\text{rms}})$	13.95	4.26	-	-	-	-	-
	$g(\theta_{\text{rms}})$	-	-	0.19	4.14	1.57	14.62	4.35
M-Link	LSF	-	-	0.67	3.91	3.45	-8.09	1.81
	$g(\tau_{\text{rms}})$	-	-	0.48	5.46	0.30	4.08	0.35
	$g(\phi_{\text{rms}})$	-	-	0.27	5.47	0.23	5.41	0.65
	$g(\theta_{\text{rms}})$	-	-	0.81	6.16	1.61	9.24	0.66

By studying the plots in Figs 5–8 and the models' parameters in Tables 3-4, the following can be concluded:

1) It is clear from subfigures (a) in Figs. 5–8 that if Approach 1 is used, i.e., collecting the estimates of the transformed LSP of interest from all links and modeling those collected values together, then: a) the LSF, as expected, is modeled as a zero mean log-normal distribution with a standard deviation equals to 4.7, b) $g(\phi_{\text{rms}})$ is modeled as a unimodal Gaussian distribution, c) both $g(\tau_{\text{rms}})$ and $g(\theta_{\text{rms}})$ are modeled as bimodal Gaussian distributions. The parameters of these unimodal and bimodal models are listed in Tables 3.
2) It is clear from subfigures (b) to (e) in Figs. 5–8 that the same LSP of the different links might have empirical CDFs with widely different behaviors. These differences suggest to model the LSP of each link individually, which is expected to provide better description of their behavior.

Of course the different values of the power parameter of the Box-Cox transformation (i.e., Λ_τ, Λ_ϕ, and Λ_θ) have a direct impact on the model parameters. However, in the following two examples, the same value of the power parameter of the Box-Cox transformation is used, yet the behaviors of the measured links are found to be totally different, which indicate that these links have different propagation characteristics. Example 1: from Table 2, it is found that $\Lambda_\theta = 0.65$ is used to get $g(\theta_{\text{rms}})$ for the E-Link, F-Link, and M-Link, yet their models have totally different parameters (see Fig. 8 (b), (d), and (e), as well as Table 4). Example 2: the LSF estimates can be looked at as a special case of applying the Box-Cox transformation to all links with $\Lambda_{\text{LSF}} = 0$. LSF is usually modeled as a unimodal log-normal distribution; however, from Fig. 5 (e) and Table 4, it is clear that the LSF of the M-Link should be modeled as a bimodal log-normal distribution (the same behavior exists, although less pronounced, in the F-link). When investigating the propagation conditions of the M-Link, the bimodal distribution of its LSF values can be explained as follows. At the vicinity of the BS-M there is a group of big trees (represented by the gray ellipse in Fig. 3). These trees are about 35 m northwest of BS-M and they cause severe attenuation of the transmitted signal for about 40% of the traveled distance. The Link-Segments that experience this attenuation correspond to the points below the regression line in Fig. 2(d), and also correspond to the Link-Segments of the inner ring in Fig. 3 that are either plotted in dark blue or identified as invalid Link-Segments. This heavy obstruction affecting a significant portion of the M-Link gives rise to the bimodal behavior to all of its LSPs. See last four rows of Table 4. A similar behavior can be expected in situations where there is partial obstruction/shadowing in a geographical area by large objects.

B. MODELING THE CROSS-CORRELATION OF THE LSPs

Simulations of wireless channels are usually performed as series of segments. Each segment represents a quasi-stationary period during which all LSPs are constant and its size can be at most a few meters [2]. In this section, we utilize our measurements in order to extract models that describe the cross-correlation of the LSPs for inter- and intra-link scenarios. To achieve this goal, the following steps are performed. First, the whole measurement route is divided into equal-length and disjoint local areas of 11.5 m (thus, each local area is 20 segment long). At each local area and for each link, time series of 20 estimates for each considered LSP are collected. Then, the inter-link and intra-link cross-correlation coefficients at a specific local area are estimated by calculating the correlation coefficient between the two concerned time series.

1) INTER-LINK CROSS-CORRELATION OF THE LSPs

For all considered LSPs and for all link pairs, it is found that the inter-link cross-correlation estimates can be modeled as a Truncated Gaussian (TG) distribution that is bounded by ± 1. If the random variable X has a Gaussian distribution with mean μ, and standard deviation σ, then X conditional on $a < X < b$ has a Truncated Gaussian distribution defined as:

$$\text{TG}(x;\mu,\sigma,a,b) = \begin{cases} \dfrac{\frac{1}{\sigma}\varphi\left(\dfrac{x-\mu}{\sigma}\right)}{\Phi\left(\frac{b-\mu}{\sigma}\right) - \Phi\left(\frac{a-\mu}{\sigma}\right)}, & \text{if } a < x < b \\ 0, & \text{otherwise} \end{cases} \tag{4}$$

where $\Phi(\cdot)$ and $\varphi(\cdot)$ denote the cumulative distribution function and density function of the standard Gaussian distribution, respectively. To simplify notations, the limits of the cross-correlation coefficient values (i.e., $a = -1$ and $b = 1$) will be dropped and the truncated distribution of interest will be denoted as $\text{TG}(\mu, \sigma)$.

TABLE 5. Model parameters of the inter-link cross-correlation.[5,6]

Link-pairs	ES	EF	EM	SF	SM	FM
μ_{LSF}	0.17	0.06	0.26	0.03	0.11	-0.18
σ_{LSF}	0.50	0.60	0.38	0.43	0.32	0.35
$\mu_{g(\tau_{\text{rms}})}$	0.12	-0.05	0.21	-0.04	0.13	-0.16
$\sigma_{g(\tau_{\text{rms}})}$	0.51	0.48	0.37	0.58	0.37	0.42
$\mu_{g(\phi_{\text{rms}})}$	0.20	-0.06	0.09	0.05	0.04	-0.06
$\sigma_{g(\phi_{\text{rms}})}$	0.43	0.57	0.45	0.40	0.34	0.38
$\mu_{g(\theta_{\text{rms}})}$	0.25	-0.07	0.13	-0.07	0.03	0.04
$\sigma_{g(\theta_{\text{rms}})}$	0.27	0.36	0.30	0.27	0.30	0.33

TABLE 6. Model parameters of the intra-link cross-correlation.[5,6]

Links	E-Link	S-Link	F-Link	M-Link	Min.	Max.
$\mu_{\text{LSF},g(\tau_{\text{rms}})}$	-1.03	-2.50	-1.10	-0.83	-2.50	-0.83
$\sigma_{\text{LSF},g(\tau_{\text{rms}})}$	0.18	0.61	0.35	0.15	0.15	0.61
$\mu_{\text{LSF},g(\phi_{\text{rms}})}$	-1.09	-1.06	-2.50	-1.36	-2.50	-1.06
$\sigma_{\text{LSF},g(\phi_{\text{rms}})}$	0.23	0.29	0.75	0.52	0.23	0.75
$\mu_{\text{LSF},g(\theta_{\text{rms}})}$	-1.48	-1.01	-1.85	-1.27	-1.85	-1.01
$\sigma_{\text{LSF},g(\theta_{\text{rms}})}$	0.80	0.45	0.80	0.80	0.45	0.80
$\mu_{g(\tau_{\text{rms}}),g(\phi_{\text{rms}})}$	1.55	1.07	0.98	2.49	0.98	2.49
$\sigma_{g(\tau_{\text{rms}}),g(\phi_{\text{rms}})}$	0.39	0.29	0.33	0.87	0.29	0.87
$\mu_{g(\tau_{\text{rms}}),g(\theta_{\text{rms}})}$	1.14	1.03	0.63	1.03	0.63	1.14
$\sigma_{g(\tau_{\text{rms}}),g(\theta_{\text{rms}})}$	0.74	0.54	0.39	0.80	0.39	0.80
$\mu_{g(\phi_{\text{rms}}),g(\theta_{\text{rms}})}$	1.82	0.82	2.09	1.55	0.82	2.09
$\sigma_{g(\phi_{\text{rms}}),g(\theta_{\text{rms}})}$	0.80	0.24	0.80	0.73	0.24	0.80

In all cases, the suggested model passes the KS test at 0.05 significance level. Table 5 lists the parameters for the suggested model for all considered LSPs and link pairs, where in the measurement setup, we have four BS links denoted as: BS-E, BS-S, BS-F, and BS-M, resulting in 6 link pairs: ES, EF, EM, SF, SM, and FM. The results reported in Table 5 can be roughly approximated by selecting μ and σ of the suggested models uniformly from the intervals $[-0.25, 0.25]$, and $[0.25, 0.60]$, respectively.

2) INTRA-LINK CROSS-CORRELATION OF THE LSPs

The same approach, i.e., using the Truncated Gaussian distribution, is used to model the intra-link cross-correlation of the LSPs. The parameters of the proposed model are listed in Table 6, where it is clear that the LSF has negative correlation with the rest of the considered LSPs, and any two of $g(\tau_{\text{rms}})$, $g(\phi_{\text{rms}})$, and $g(\theta_{\text{rms}})$ have positive correlation. Thus, our results are in line with the findings of [9] and [10], and extend them to other LSPs.

3) SHAPES OF THE DISTRIBUTIONS MODELING THE INTER-LINK AND INTRA-LINK CROSS-CORRELATION OF THE LSPs

It is important to notice that in order to simplify the usage of the herein reported results for simulation purposes, we report

[5]The reported values in these tables are the means and the standard deviations of the corresponding Gaussian distribution, before truncation.

[6]Please see section V-C for the details on the steps suggested to utilize the provided models, where after generating instances based on a Gaussian model using the herein reported parameters, truncation should be applied.

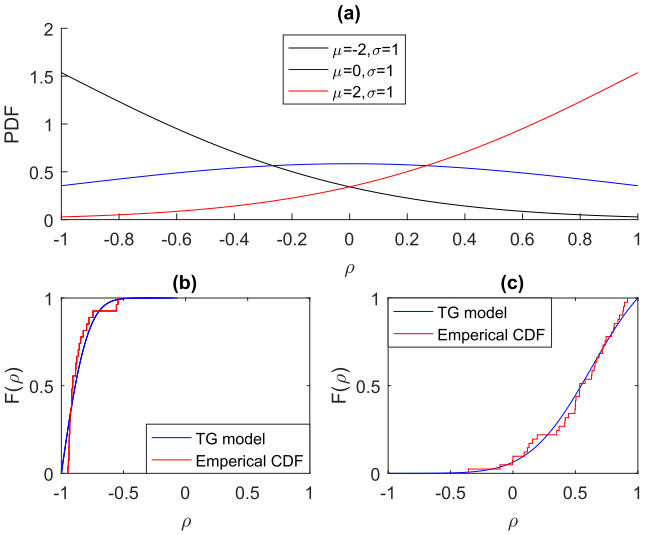

FIGURE 9. Modeling the inter- and intra-link cross-correlation of the LSPs using the Truncated Gaussian distribution. (a) Examples of the different shapes which the Truncated Gaussian distribution takes depending on the values of μ and σ. (b) and (c) Examples of the empirical CDFs of the inter- and intra-link cross-correlation and their model.

in Tables 5 and 6 the means and the standard deviations of the corresponding Gaussian distribution, before truncation. The results of Table 6 indicate that the intra-link cross-correlation of the LSPs have different behavior depending on the considered LSPs. Despite the fact that the parameters in Table 6 correspond to a Gaussian model, depending on the selected pair of parameters (i.e, μ, and σ), the actual resulted distributions (i.e., after truncation) that describe the generated cross-correlation values have different shapes. For example, selecting $(\mu = -2, \sigma = 1)$, $(\mu = 0, \sigma = 1)$, and $(\mu = 2, \sigma = 1)$[7] will result in probability distribution functions (PDFs) that look like exponential PDF, truncated bell shape PDF, and reverse exponential PDF, respectively, as depicted in Fig. 9(a). Fig. 9(b) and (c) illustrate examples of the empirical CDFs of the LSPs estimates and the CDFs of their proposed Truncated Gaussian model. As mentioned earlier, in all cases, the proposed model passes the KS test at 0.05 significance level and Fig. 9(b) illustrates the case with the largest D_{LSP}, see equation (2), among all studied cases

Selecting values for μ and σ within the aforementioned intervals produces values of the cross-correlation coefficients that take into consideration observed variations across links. For example (following the procedure detailed in section V-C), if $\mu = -0.25$ and $\sigma = 0.25$ are used to model the inter-link cross-correlation of one link pair, then the corresponding probability of having $\rho > 0.5$ is almost zero. However, for another simulated link pair, if $\mu = 0.25$ and $\sigma = 0.60$ are assumed, then the probability of having $\rho > 0.5$ will be 0.26.

[7]Please note that even though the values of μ might be outside the interval $[-1, 1]$; however, after truncation is applied, the remaining correlation coefficient values have absolute values not exceeding 1. See Fig. 9.(a).

C. USING THE PROPOSED MODELS

Let's assume that we want to generate two correlated time series to simulate S successive segments, each consisting of N LSP data points (i.e., realizations). Therefore each generated time series should have a length of SN data points. The proposed models can be used as follows[8,9]:

Step 1: *Generating instances of the cross-correlation coefficients (characterizing the correlation between the two targeted time series over individual segments)*

- Depending on the simulated LSPs, use Tables 5 or 6 and select values for μ and σ to model their cross-correlation properties. Alternatively, the results of Table 5 can be simplified by drawing μ and σ uniformly from the intervals $[-0.25, 0.25]$ and $[0.25, 0.60]$, as suggested in section V-B. Also, the results of Table 6 can be simplified by drawing μ and σ uniformly from the intervals defined by the minimum and the maximum values reported at each raw in the table.
- Generate cross-correlation instances according to the distribution $\rho \sim G(\mu, \sigma)$, where $G(\mu, \sigma)$ is the Gaussian distribution with mean, μ, and standard deviation, σ. After this step, some values of ρ might lie outside the $[-1, 1]$ interval.
- Omit the cross-correlation instances that have absolute values exceeding 1. The remaining cross-correlation instances follow the targeted distribution, i.e., $\mathrm{TG}(\mu, \sigma)$.

The above three steps should be used to generate S values each of which represents the cross-correlation between the corresponding time series during one segment.

Step 2: *Generating correlated time series of the LSPs (in the Box-Cox transform domain)*
Generate two independent Gaussian time series for the simulated LSPs each with SN data points. For each segment of N consecutive data points, use the cross-correlation instances resulted from Step 1 to generate correlated time series [25]. The resulting correlated time series describe the LSPs of interest in the transformed domain (i.e., LSF, $g(\tau_{\mathrm{rms}})$, $g(\phi_{\mathrm{rms}})$, $g(\theta_{\mathrm{rms}})$).

Step 3: *Applying inverse Box-Cox transformation*
Use the values of Λ reported in Table 2 in order to calculate the inverse Box-Cox transformation of the $g(\tau_{\mathrm{rms}})$, $g(\phi_{\mathrm{rms}})$, and $g(\theta_{\mathrm{rms}})$ generated time series (no inverse transformation is needed for the LSF). Alternatively, the results of Table 2 can be simplified by drawing Λ uniformly from the intervals defined by the minimum and the maximum values reported at each column in the table.
The result of this step is the targeted correlated LSPs time series.

To simulate more than two correlated time series, only the cross-correlation coefficient instances (generated in Step 1) that result in positive semi-definite correlation matrices are considered to be valid. Then, the sought correlations among the simulated time series are realized via generating coloring matrices (in Step 2) by applying the Cholesky decomposition to these valid correlation matrices [25].

[8] A detailed discussion on the integration of the cross-correlation instances resulting from the proposed model within a complete multiple radio links simulation framework can be found in [2, Chapter 3].

[9] The proposed procedure in this section is general and applicable to any multi-link scenario; however, the parameters provided in the tables herein are limited to scenarios with similar propagation characteristics, i.e., suburban microcell environments as detailed in section III.

VI. SUMMARY AND CONCLUSION

In this paper, the statistical properties of four large-scale parameters (LSF, as well as the delay, azimuth, and elevation spreads) are studied based on propagation measurements performed in a suburban microcell environment at 2.6 GHz with four BSs and one MS. The Box-Cox transformation is proposed as a general framework allowing to transform the measured large-scale parameters to a domain where they can be modeled as Gaussian distributions. Then, in the transform domain, the distributions and cross-correlation properties of the studied large-scale parameters are modeled. The main results of this paper can be summarized as follows. 1) Despite the fact that the overall propagation conditions of the different links are similar (e.g., comparable antenna heights, and same propagation environment), the unique distribution of the interacting objects in the environment results in different behavior for each link. Furthermore, some of the interacting objects that are close to the BSs might give rise to the bimodal behavior for the distribution of the large-scale parameters. 2) The Truncated Gaussian distribution is found to be a suitable model to describe both the inter-link and intra-link cross-correlation coefficients of the different links. 3) Finally, modeling parameters characterizing the distribution and the cross-correlation properties of the studied large-scale parameters are provided. All provided models are verified using the Kolmogorov-Smirnov test at significance level $\alpha = 0.05$.

REFERENCES

[1] D. Baum *et al.*, "WINNER I final report on link level and system level channel models," Tech. Rep. D5.4, 2005. [Online]. Available: http://www.ist-winner.org/DeliverableDocuments/D5.4.pdf

[2] P. Kyosti *et al.*, "WINNER II channel models," Tech. Rep. D1.1.2 V.1.2, 2008. [Online]. Available: http://www.ist-winner.org/deliverables.html

[3] Z. Ni and D. Li, "Effect of fading correlation on capacity of distributed MIMO," in *Proc. IEEE 15th Int. Symp. Pers., Indoor Mobile Radio Commun.*, Barcelona, Spain, Sep. 2004, pp. 1637–1641.

[4] J. Poutanen, F. Tufvesson, K. Haneda, V. Kolmonen, and P. Vainikainen, "Multi-link MIMO channel modeling using geometry-based approach," *IEEE Trans. Antennas Propag.*, vol. 60, no. 2, pp. 587–596, Feb. 2012.

[5] A. Mawira, "Models for the spatial correlation functions of the (log)-normal component of the variability of VHF/UHF field strength in urban environment," in *Proc. 3rd IEEE Int. Symp. Pers., Indoor Mobile Radio Commun.*, Oct. 1992, pp. 436–440.

[6] Z. Meifang, F. Tufvesson, and J. Medbo, "Correlation properties of large scale parameters from 2.66 GHz multi-site macro cell measurements," in *Proc. IEEE 73rd Veh. Technol. Conf.*, Budapest, Hungary, May 2011, pp. 1–5.

[7] S. Jaeckel *et al.*, "Correlation properties of large and small-scale parameters from multicell channel measurements," in *Proc. 3rd Eur. Conf. Antennas Propag.*, Berlin, Germany, Mar. 2009, pp. 3406–3410.
[8] G. Dahman, J. Flordelis, and F. Tufvesson, "On the cross-correlation properties of large-scale fading in distributed antenna systems," in *Proc. IEEE Wireless Commun. Netw. Conf.*, Istanbul, Turkey, Apr. 2014, pp. 160–165.
[9] A. Algans, K. I. Pedersen, and P. E. Mogensen, "Experimental analysis of the joint statistical properties of azimuth spread, delay spread, and shadow fading," *IEEE J. Sel. Areas Commun.*, vol. 20, no. 3, pp. 523–531, Apr. 2002.
[10] N. Jaldén, P. Zetterberg, B. Ottersten, and L. Garcia, "Inter- and intrasite correlations of large-scale parameters from macrocellular measurements at 1800 MHz," *EURASIP J. Wireless Commun. Netw.*, vol. 2007, p. 025757, Dec. 2007.
[11] J. Park, M. Kim, H. Kwon, H. Chung, X. Yin, and Y. Fu, "Measurement-based stochastic cross-correlation models of a multilink channel in cooperative communication environments," *ETRI J.*, vol. 34, pp. 858–868, Dec. 2012.
[12] A. Bottcher, P. Vary, C. Schneider, and R. S. Thoma, "Cross correlation characteristics of large scale parameters in urban macro cell," in *Proc. IEEE Veh. Technol. Conf.*, San Francisco, CA, USA, Sep. 2011, pp. 1–5.
[13] L. Liu *et al.*, "The COST 2100 MIMO channel model," *IEEE Wireless Commun.*, vol. 19, no. 6, pp. 92–99, Dec. 2012.
[14] J. Salo, L. Vuokko, H. M. El-Sallabi, and P. Vainikainen, "An additive model as a physical basis for shadow fading," *IEEE Trans. Veh. Technol.*, vol. 56, no. 1, pp. 13–26, Jan. 2007.
[15] R. M. Sakia, "The box-cox transformation technique: A review," *J. Roy. Stat. Soc. Ser. D (Stat.)*, vol. 41, no. 2, pp. 169–178, 1992.
[16] F. J. Massey, Jr., "The Kolmogorov-Smirnov test for goodness of fit," *J. Amer. Statist. Assoc.*, vol. 46, no. 253, pp. 68–78, 1951.
[17] R. S. Thoma *et al.*, "Identification of time-variant directional mobile radio channels," *IEEE Trans. Instrum. Meas.*, vol. 49, no. 2, pp. 357–364, Apr. 2000.
[18] J. Flordelis, G. Dahman, and F. Tufvesson, "Measurements of large-scale parameters of a distributed MIMO antenna system in a microcell environment at 2.6 GHz," in *Proc. 7th Eur. Conf. Antennas Propag.*, Gothenburg, Sweden, Apr. 2013, pp. 3026–3030.
[19] E. S. Sousa, V. M. Jovanovic, and C. Daigneault, "Delay spread measurements for the digital cellular channel in Toronto," *IEEE Trans. Veh. Technol.*, vol. 43, no. 4, pp. 837–847, Nov. 1994.
[20] *Multipath Propagation and Parameterization of its Characteristics*, document Recom. ITU-R P.1407-6, Jun. 2017.
[21] B. H. Fleury, M. Tschudin, R. Heddergott, D. Dahlhaus, and K. I. 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.
[22] 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 Veh. Technol. Conf.*, vol. 1. May 2002, pp. 522–526.
[23] G. Dahman, J. Flordelis, and F. Tufvesson, "Estimating the cross-correlation properties of large-scale parameters in multilink distributed antenna systems: Synchronous measurements versus repeated measurements," *IEEE Trans. Veh. Technol.*, vol. 66, no. 9, pp. 7633–7642, Sep. 2017.
[24] A. F. Molisch, *Wireless Communications*. Hoboken, NJ, USA: Wiley, 2010.
[25] R. H. Shumway and D. S. Stoffer, *Time Series Analysis and its Applications With R Examples*. New York, NY, USA: Springer, 2006.

GHASSAN DAHMAN received the Ph.D. degree from Carleton University, Ottawa, ON, Canada, in 2010. From 2010 and 2012, he was an Assistant Professor at Umm Al-Qura University, Makkah, Saudi Arabia. From 2012 to 2016, he was a Postdoctoral Researcher with Lund University, Lund, Sweden. He is currently a Researcher with the NSERC-Ultra Electronics TCS Industrial Chair, École de Technologie Suprieure, Montreal, QC, Canada. His main research interests are radio propagation and channel modeling, including massive MIMO, distributed antenna systems, overwater point-to-point communications, and anomalous propagation of microwaves in the troposphere.

JOSE FLORDELIS received the M.S. degrees in telecommunication engineering from the Universitat Politècnica de València, Spain, and Högskolan i Gävle, Sweden, in 2002. He is currently pursuing the Ph.D. degree with the Department of Electrical and Information Technology, Lund University, Sweden. From 2002 to 2012, he was with Ericsson, Sweden, where he contributed to the development of 3GPP LTE- and Bluetooth-based solutions. His main research interests are measurement, characterization, and modeling of wireless channels, especially in the areas of massive MIMO, and distributed antenna systems.

FREDRIK TUFVESSON received the Ph.D. degree from Lund University, Sweden, in 2000. After two years at a startup company, he joined the Department of Electrical and Information Technology, Lund University, where he is currently a Professor of radio systems. He has authored around 60 journal papers and 120 conference papers His main research interests are channel modeling, measurements, and characterization for wireless communication, with applications in various areas, such as massive MIMO, UWB, mm wave communication, distributed antenna systems, radio-based positioning, and vehicular communication. Dr. Tufvesson received the Neal Shepherd Memorial Award for the Best Propagation Paper in the IEEE TRANSACTIONS ON VEHICULAR TECHNOLOGY.