Ultra-Wideband Channel Modeling on the Basis of Information-Theoretic Criteria

Ulrich G. Schuster and Helmut Bölcskei
Communication Technology Laboratory
Swiss Federal Institute of Technology (ETH), CH-8092 Zurich
email: {schuster, boelcskei}@nari.ee.ethz.ch

Giuseppe Durisi
Istituto Superiore Mario Boella
I-10138 Torino, Italy
email: durisi@ismb.it

***Abstract*—We present results of two ultra-wideband (UWB) channel measurement campaigns in the 2–5 GHz frequency band, and use Akaike's Information Criterion (AIC) to determine suitable distributions for the channel impulse response taps. Despite the large bandwidth, AIC supports the complex Gaussian tap distribution, with mean depending on the measurement setting. We estimate the empirical covariance matrix of the channel impulse response, and demonstrate that the number of corresponding significant eigenvalues scales approximately linearly with bandwidth, albeit we find that channel taps are correlated.**

I. Introduction and Overview of Results

Wireless systems operating over *ultra-wideband* (UWB) channels with several gigahertz of bandwidth promise ease of multiple access, efficient use of the radio spectrum through overlay techniques, and improved link reliability due to frequency diversity. Two key assumptions are often made in theoretical analyses of UWB systems:

A1: The taps of the equivalent baseband channel impulse response are circularly symmetric complex Gaussian distributed.

A2: The number of independent diversity branches, denoted as *stochastic degrees of freedom* in the following, scales linearly with bandwidth.

A1 can be justified by the central limit theorem: if each tap is made up of a large number of contributions from different scatterers, the resulting tap distribution can be modeled as complex Gaussian [1]. The reasoning behind A2 is that the continuous-time channel satisfies the uncorrelated scattering (US) assumption [2].

Pierce [3] and Viterbi [4] relied[1] on A1 to show that the capacity of the infinite bandwidth fading channel is equal to the capacity of the infinite bandwidth AWGN channel. Médard and Gallager [7], and Subramanian and Hajek [8], showed that under A1 and A2 the mutual information attainable with white-like signals approaches zero with increasing bandwidth.

A1 and A2 are valid for small bandwidths. However, the high temporal resolution of UWB systems might lead to a reduced number of contributing partial waves per tap [9], [10], making A1 questionable. Several UWB channel measurement results show a better fit of Nakagami [10], lognormal [11], or Weibull [12] tap amplitude distributions, while others support A1 [13], [14]. Satisfying A2 for increasing bandwidth requires arbitrarily rich scattering, which real-world environments might not provide. However, to the best of our knowledge, only a single measurement campaign has addressed A2 so far [15]. We conducted two extensive UWB channel measurement campaigns from 2 GHz to 5 GHz in an indoor public space environment to assess the validity of A1 and A2 at large bandwidths. Our main results and contributions can be summarized as follows:

- We propose to use Akaike's Information Criterion (AIC) as a tool to compare the fit of the tap amplitude distributions put forward in the literature [10]–[14], and argue that AIC is better suited for this task than the goodness-of-fit (GOF) tests often used in this context.
- We find that, depending on the type of channel, the Rayleigh or the Rice amplitude distribution is still adequate to model small-scale UWB fading. This result supports A1, despite the large bandwidth of 3 GHz. However, the alternative distributions [10]–[12] are often close to the Rayleigh or Rice distribution in the sense of AIC.
- Under the assumption that the Rice amplitude distribution results from jointly complex Gaussian distributed channel taps, we investigate the intertap correlation. Although we find that A2 holds, our results indicate that the taps are correlated.

Notation: $\mathbb{E}_X$ stands for the expectation with respect to the random variable (RV) X. Cumulative distribution functions (CDFs) are denoted by the letters F and G, corresponding probability density functions (PDFs) by f and g. If a CDF, say G, depends on some parameter vector $\mathbf{\Theta}$, we write $G_{\mathbf{\Theta}}$. Estimated quantities are indicated by a hat $\widehat{\ }$. The superscript ^{T} stands for transposition, and ^{H} for conjugate transposition. All logarithms are to the base e.

II. The Statistical Modeling Problem

As the effective channel, i.e., the physical channel in conjunction with transmit and receive filters, is always band limited, it can be described in terms of samples $h[n]$ of the continuous-time impulse response. Therefore, we consider the discrete-time complex baseband equivalent channel with input-

This work was supported in part by the Swiss *Kommission für Technologie und Innovation (KTI)* under grant 6715.2 ENS-ES, and by the Swiss *Staatssekretariat für Bildung und Forschung* as part of the European Integrated Project PULSERS under contract FP6-506897.

[1] Telatar and Tse [5], and Verdú [6] demonstrated that A1 is not necessary to obtain this result.

output relation

$$y[n] = \sum_{l=0}^{L-1} h[l] s[n-l] \quad (1)$$

where $s[n]$ denotes the transmitted signal, and $y[n]$ is the resulting output signal. Our goal is to statistically characterize the L-dimensional channel vector $\mathbf{h} = \left[h[0]\, h[1] \ldots h[L-1]\right]^T$ by selecting suitable marginal tap distributions $F_{h[l]}$ and a joint distribution $F_{\mathbf{h}}$.

We think of the channel taps $h[n]$ as being distributed according to the *operating model*, defined as the nearest representation of the true situation that can be constructed by means of a probability model [16]. The operating model is unknown; hence, the main goal in statistical channel modeling is to find a probability model for the channel taps that approximates the operating model as closely as possible. This probability model should be based on *physical insight*, be *mathematically tractable*, and lead to *consistent predictions*.

The most widely used approach to characterize the marginal tap amplitude distributions on the basis of measurement data is through GOF tests [17]: The null hypothesis H_0 is that a given parametrized function $G_{\mathbf{\Theta}}$ equals the CDF F of the operating model, that is, $H_0 : F = G_{\mathbf{\Theta}}$. The alternative hypothesis is the complementary event $H_1 : F \neq G_{\mathbf{\Theta}}$. If a suitable test statistic (i.e., a function of the measured data) exceeds a given threshold, H_0 is accepted. The probability that H_0 is rejected although it is true is called the *significance level* of the test. GOF tests are commonly applied by first estimating the parameter vector $\mathbf{\Theta}$ from measured data, and subsequently computing the corresponding test statistic on the basis of the same data that was used for parameter estimation. The test is performed for every candidate distribution, using a common significance level (chosen ad hoc), and the distribution with the highest passing percentage across several measurement locations is selected [11], [12]. If we admit that the operating model is complex and the space of possible channel tap distributions is infinite, the chance of selecting $G_{\mathbf{\Theta}}$ so that $G_{\mathbf{\Theta}} = F$ is zero. Thus, a sensible approach to model selection is to find a suitable *approximation* of the operating model. A measure of approximation between a candidate distribution and the distribution of the operating model is called a *discrepancy* [16]. A GOF test, however, does not lead to a well-defined measure of approximation quality [18]. Pioneering work in the field of model selection through discrepancy minimization was done by Akaike, whose information criterion [19] has found widespread use. One of the two main contributions of this paper is to use AIC for UWB channel modeling to investigate A1. To the best of our knowledge, only [20] suggests the use of model selection methods for wireless channel modeling, the main idea being to apply the principle of minimum description length (MDL) to determine the amplitude distribution of narrowband channels.

III. A Brief Review of Model Selection Using AIC

Our review follows the book by Linhart and Zucchini [16]. We restrict our discussion to univariate CDFs; the multivariate case is hardly feasible because of the large bandwidth and the resulting large number of taps. Denote the unknown CDF of the operating model by F, and the set of all univariate CDFs by $\mathcal{M}$. A parametric *candidate family* $\mathcal{G}^j = \{G^j_{\mathbf{\Theta}^j} \,|\, \mathbf{\Theta}^j \in \mathcal{T}^j\}$ is a subset of $\mathcal{M}$, with individual CDFs $G^j_{\mathbf{\Theta}}$ parametrized[2] by the U-dimensional vector $\mathbf{\Theta}^j \in \mathcal{T}^j$, with $\mathcal{T}^j \subset \mathbb{R}^U$. Candidate families need to be chosen *in advance* to reflect prior knowledge about the modeling problem. The set of J candidate families $\mathcal{C} = \bigcup_{j=1}^{J} \mathcal{G}^j$ constitutes the *candidate set*. A *discrepancy* is a functional $\Delta : \mathcal{M} \times \mathcal{M} \to \mathbb{R}$ satisfying $\Delta(G^j_{\mathbf{\Theta}}, F) \geq \Delta(F, F)$ for all $j \in \{1, 2, \ldots, J\}$. A consistent estimator for the discrepancy $\Delta(G^j_{\mathbf{\Theta}}, F)$ on the basis of N independent samples $\mathbf{x} = [x_1\, x_2 \ldots x_N]^T$ distributed according to F is called an *empirical discrepancy*, and will be denoted by $\Delta_N(G^j_{\mathbf{\Theta}}, F)$. The goal of the model selection procedure is to choose the distribution that minimizes the discrepancy among all members of the candidate set. The procedure consists of two steps: With the operating model unknown, we first estimate the parameter vector $\mathbf{\Theta}^j$ for each candidate family $\mathcal{G}^j$, using the *minimum discrepancy estimator*[3] $\widehat{\mathbf{\Theta}}^j(\mathbf{x}) = \arg\min_{\mathbf{\Theta} \in \mathcal{T}^j} \Delta_N(G^j_{\mathbf{\Theta}}, F)$. The resulting discrepancy $\Delta(G^j_{\widehat{\mathbf{\Theta}}}, F)$ is thus an RV. As a model should be predictive, it must be a good approximation for all realizations of the RV X, not just for the actual observations $\mathbf{x}$. The second step is thus to find j so that the expected discrepancy $\mathbb{E}_X[\Delta(G^j_{\widehat{\mathbf{\Theta}}}, F)]$ is minimized over the candidate set. As the operating model is unknown, the expected discrepancy cannot be computed, but has to be estimated.

AIC [19] is an approximately unbiased estimator of the expected Kullback-Leibler (KL) discrepancy[4] $\mathbb{E}_{\widehat{\mathbf{\Theta}}, Y}[\log g^j_{\widehat{\mathbf{\Theta}}}(Y)]$:

$$\mathrm{AIC}_j = -2 \sum_{n=1}^{N} \log g^j_{\widehat{\mathbf{\Theta}}}(x_n) + 2U. \quad (2)$$

The minimum discrepancy estimator is the maximum likelihood (ML) estimator:

$$\widehat{\mathbf{\Theta}}^j = \arg\max_{\mathbf{\Theta} \in \mathcal{T}^j} \frac{1}{N} \sum_{n=1}^{N} \log g^j_{\mathbf{\Theta}}(x_n). \quad (3)$$

AIC estimates the approximation quality of different probability models; it can be used to rank the individual candidate distributions, with the minimum AIC value indicating the best fit. We define the AIC differences $D_j = \mathrm{AIC}_j - \min_i \mathrm{AIC}_i$, where $\min_i \mathrm{AIC}_i$ is the minimum AIC value over all J candidate families. As AIC is an approximately unbiased estimator of the KL discrepancy, an estimate of the likelihood of the model m_j with CDF $G^j_{\widehat{\mathbf{\Theta}}} \in \mathcal{G}^j$, given the data $\mathbf{x}$, is obtained through the transformation $\mathcal{L}(m_j \,|\, \mathbf{x}) = \alpha e^{-D_j/2}$, with a suitably chosen constant α. The normalized likelihood

[2]In the following, we indicate that a CDF $G^j_{\mathbf{\Theta}^j}$ is parametrized by the vector $\mathbf{\Theta}^j$ by simply writing $G^j_{\mathbf{\Theta}}$.

[3]We write $\widehat{\mathbf{\Theta}}^j$ instead of $\widehat{\mathbf{\Theta}}^j(\mathbf{x})$ in the remainder of the paper.

[4]As the first term in the KL distance $D(f \,||\, g) = \mathbb{E}_Y[\log f(Y)] - \mathbb{E}_Y[\log g(Y)]$ is a function of the operating model only, it can be ignored.

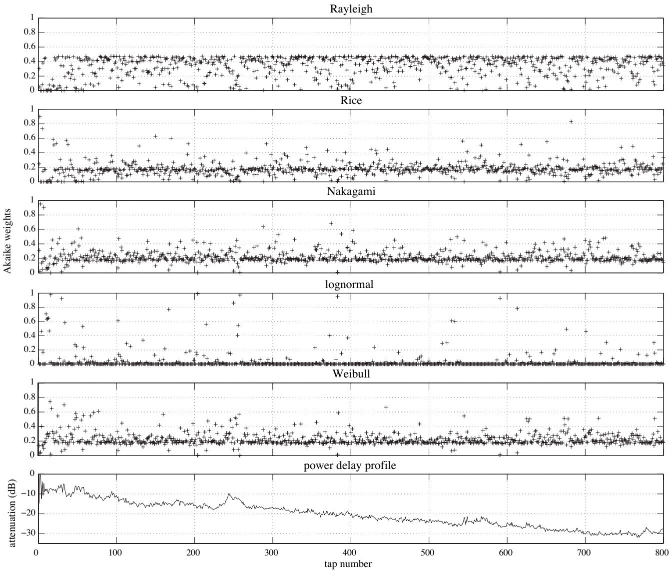

Fig. 1. PDP and Akaike weights for Measurement Campaign I

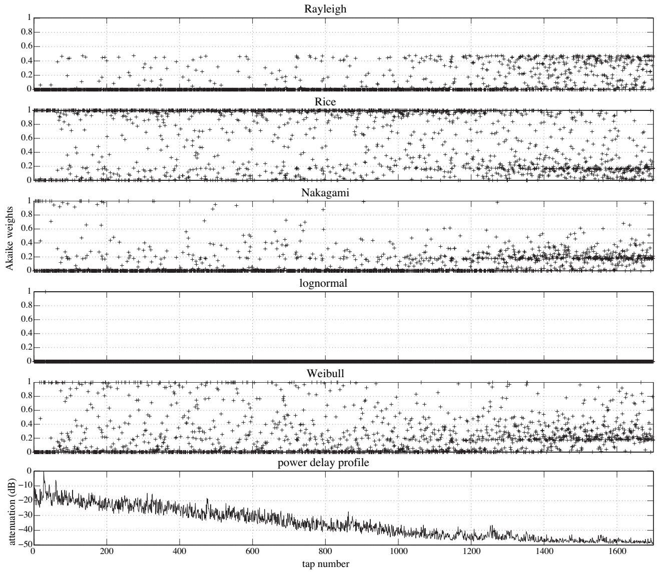

Fig. 2. PDP and Akaike weights for Measurement Campaign II

estimates yield the so-called *Akaike weights* [21]

$$w_j = \frac{e^{-\frac{1}{2}D_j}}{\sum_{i=1}^{J} e^{-\frac{1}{2}D_i}} \quad (4)$$

which satisfy $\sum_{j=1}^{J} w_j = 1$. The weight w_j can be interpreted as an estimate of the probability that the CDF $G^j_{\widehat{\Theta}}$ shows the best fit within the candidate set, given the data $\mathbf{x}$ [21]. Consequently, the Akaike weights allow us not only to select the best distribution in the candidate set, but also provide information about the relative approximation quality of each distribution.

IV. The UWB Channel Measurement Campaigns

We conducted two measurement campaigns in the entrance lobby of the ETZ building at ETH Zurich, a typical public open space environment with large windows, a tiled floor, and concrete columns. Both sets of measurements were obtained for a setting where the transmitter and the receiver were separated by approximately 20 m, and the line of sight was partially obstructed by the concrete columns.

Measurement Campaign I (MC I): We fixed the position of the receive antenna, kept the environment static, moved the transmit antenna on a rectangular 9×5 grid with 7 cm spacing in both dimensions, displaced the grid by 50 cm, and repeated the procedure. For each of the resulting $N = 90$ antenna locations, we recorded 1601 frequency points in the band from 2 GHz to 5 GHz, using an HP 8722D vector network analyzer.

Measurement Campaign II (MC II): To characterize channel variations induced by people moving in the lobby, we fixed the position of the transmit and the receive antenna, sounded the channel with a pseudonoise sequence of chip rate 10 GHz, and recorded the received signal in real time with the digital sampling oscilloscope Agilent DSO81204A at a sampling rate of 40 GHz and with an analog bandwidth of 12 GHz. Through postprocessing as described in [22], we extracted $N = 2722$ sample impulse responses in the frequency band from 2 GHz to 5 GHz, acquiring one impulse response per second.

V. Tap Statistics (A1)

We apply AIC to our measurement data to evaluate the different tap amplitude distributions put forward in the UWB literature [10]–[14]. Our candidate set $\mathcal{C}$ thus consists of the single-parameter ($U = 1$) Rayleigh family and the two-parameter ($U = 2$) Rice, Nakagami, lognormal, and Weibull families. The Rice, Nakagami, and Weibull families contain the Rayleigh family as a special case. Rayleigh, Rice, and Nakagami amplitude distributions can be derived from physical principles [1]. The Weibull [12] and lognormal [9], [11] distributions seem to lack physical support for small-scale fading.

1) MC I: Fig. 1 shows the normalized empirical power–delay profile (PDP) of the measured channel for the first 800 taps, along with the Akaike weights for each candidate family. Our findings can be summarized as follows:

1) The Rayleigh distribution shows the best fit, followed by the Rice, Nakagami, and Weibull distributions in no particular order.
2) The lognormal distribution shows a consistently bad fit, with the exception of a few isolated taps.
3) The variability of the Akaike weights is high across taps.

Interpretation of the Results: A closer look at the parameter estimates $\widehat{\Theta}^j$ shows that the variability of the Akaike weights across taps is due to the high sensitivity of w_j to $\widehat{\Theta}^j$. The parameter estimates for the Rice, Nakagami and Weibull distributions are close to the values that reduce the respective distributions to the Rayleigh distribution. The Ricean K-factor, for example, varies between 0 and 1.6 across the 800 taps shown. The average difference of the Akaike weights between the Rayleigh distribution ($U = 1$) and the Rice, Nakagami

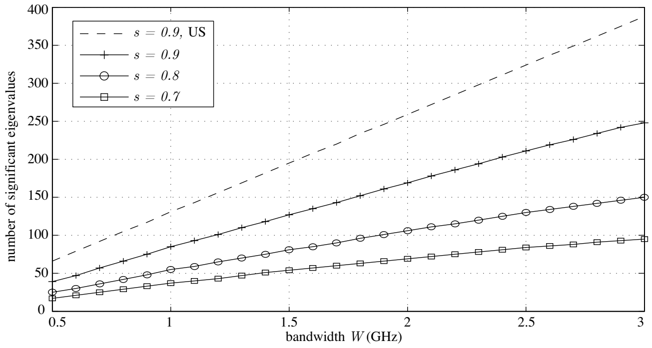

Fig. 3. Number of significant eigenvalues of $\widehat{\mathbf{K}}$ as a function of bandwidth

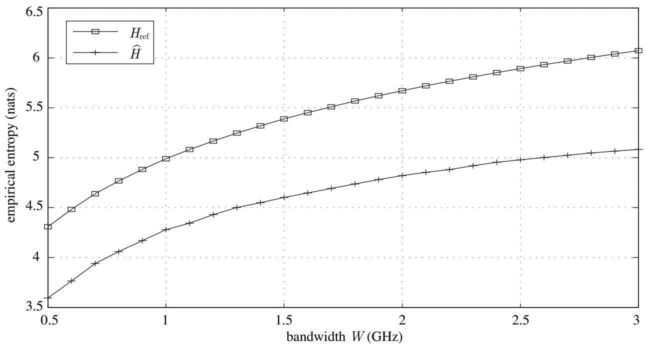

Fig. 4. $\widehat{H}$ and H_{ref} as a function of bandwidth

and Weibull distributions ($U = 2$) mainly results from the term $2U$ in (2). The reason for this behavior lies in the nature of AIC, which consists of a contribution measuring the fit of the distribution to the data, and a contribution penalizing the complexity of the distribution, as quantified by the number of free parameters U. The fit of the Nakagami, Rice, and Weibull distributions to our measured data is evidently not good enough to warrant this additional complexity. Of course, this does not preclude the existence of other distributions, potentially with more parameters, that would provide a better fit than the Rayleigh distribution.

2) MC II: Fig. 2 depicts the normalized empirical PDP and the Akaike weights obtained from the time-domain measurement campaign. Our findings are as follows:

1) The Ricean distribution shows the best fit for the first 1200 taps.
2) The Rayleigh distribution fits well for taps $\gtrsim$ 1200.
3) The lognormal distribution does not fit the measurements at all, while the Nakagami distribution is suitable only for the first few taps.
4) The Weibull distribution shows a good fit for some taps.

Interpretation of the Results: Ricean fading is often attributed to a strong mean component in the impulse response [1]. While in MC I large scattering objects, like windows and walls, move relative to the position of the antennas, the moving scatterers in MC II, i.e., people in the lobby, are much smaller. Hence, significant contributions in most channel taps are mainly due to static scatterers, consistent with the physical interpretation of the Rice model. Approximately from tap 1200 on in MC II, the Rayleigh distribution exhibits a better fit than the Ricean distribution. This can be attributed to the low measurement SNR of these taps, so that we are effectively fitting Gaussian distributed noise.

3) General Comments: Our analysis shows that even for bandwidths of up to 3 GHz the Rayleigh or the Rice distribution provides a good fit, although the differences to the Nakagami and Weibull distributions in terms of the Akaike weights are often small, especially in MC I. Consequently, the data do not provide enough evidence to unequivocally select a single distribution. However, the empirical support for Rayleigh and Rice fading, combined with the mathematical tractability of these distributions, leads us to advocate their use.

We do not apply model selection tools to determine the distribution of the phase of the complex-valued channel taps, in part because of the lack of a physically motivated alternative to the almost exclusively used uniform phase assumption, and in part because of the difficulty to obtain accurate phase estimates through measurements. The combination of the Rayleigh, respectively Rice, amplitude distribution with the assumption of a uniformly distributed phase of the zero-mean component of the taps results in the circularly symmetric complex Gaussian distribution for each individual tap.

Marginally Gaussian distributions do not imply joint Gaussianity of the channel impulse response vector $\mathbf{h}$. Unfortunately, selecting the joint PDF of $\mathbf{h}$ using AIC is a hopeless endeavor, as already discussed. As a simple heuristic test, we verify through AIC that the amplitudes of the samples of the Discrete Fourier Transform of $\mathbf{h}$ are Rayleigh, respectively Rice, distributed. Together with the uniform phase assumption in the frequency domain, this shows that at least certain linear combinations of the taps can be again modeled as complex Gaussian. Hence, with no evidence against the joint Gaussianity assumption, and again referring to the analytical tractability of the model, we advocate the use of the jointly complex Gaussian distribution for $\mathbf{h}$.

VI. The Uncorrelated Scattering Assumption (A2)

Correlation between channel taps in (1) can result from correlated scattering in the underlying continuous-time propagation channel [1], or from the effect of the antennas and the transmit and receive filters. Separating these two sources of correlation on the basis of measurements is difficult. Therefore, we consider the discrete-time effective channel only, and analyze the corresponding intertap correlation.

We decompose the random channel vector $\mathbf{h}$ according to $\mathbf{h} = \mathbf{m} + \tilde{\mathbf{h}}$, with $\mathbf{m} = \mathbb{E}[\mathbf{h}]$. Under the assumption that $\mathbf{h}$ is jointly complex Gaussian distributed with circularly symmetric $\tilde{\mathbf{h}}$, the joint distribution of $\mathbf{h}$ is specified through $\mathbf{m}$ and the covariance matrix $\mathbf{K} = \mathbb{E}\big[\tilde{\mathbf{h}}\tilde{\mathbf{h}}^H\big]$. We truncate all measured impulse responses $\mathbf{h}_n$, $n = 1, 2, \ldots, N$ after $L = 701$ taps (above the noise floor), and compute the empirical $L \times L$ covariance matrix $\widehat{\mathbf{K}} = \frac{1}{N}\sum_{n=1}^{N}(\mathbf{h}_n - \widehat{\mathbf{m}})(\mathbf{h}_n - \widehat{\mathbf{m}})^H$, where $\widehat{\mathbf{m}} = \frac{1}{N}\sum_{n=1}^{N}\mathbf{h}_n$. Accurate estimation of $\mathbf{K}$ requires many samples; hence, we use data from MC II in the following.

Note that the diversity order of a frequency-selective Ricean channel is well-defined, and given by the rank of $\mathbf{K}$ [23].

The discrete-time US assumption leads to a linear scaling of the number of "significant" eigenvalues of $\mathbf{K}$ as a function of W. We denote the kth eigenvalue of $\widehat{\mathbf{K}}$ as $\widehat{\lambda}_k$, use a normalization so that $\sum_{k=1}^{L} \widehat{\lambda}_k = 1$, and arrange the eigenvalues in decreasing order. Eigenvalues $\widehat{\lambda}_k$ with index $k \leq L_s$ are declared significant, where L_s is the largest integer satisfying $\sum_{k=1}^{L_s} \widehat{\lambda}_k \leq s$, with $0 \leq s \leq 1$. This criterion essentially measures the number of diversity branches with a branch receive SNR above a certain threshold. Fig. 3 depicts the scaling behavior for different s as a function of W. The scaling is approximately linear in all cases. For reference, we show the scaling behavior of a discrete-time US channel with PDP obtained by uniformly sampling (at rate W) the empirical PDP after removing the mean. To double-check our results, we use the model order selection criteria AIC and MDL [24], and observe the same scaling behavior. However, we note that sublinear scaling of the number of stochastic degrees of freedom was observed in [15], where a measurement setup similar to MC I was used. This indicates that there might be a fundamental difference between the spatially varying channel measured in [15] and the channel with static terminals in MC II.

Linear scaling of the number of significant eigenvalues with W is not sufficient to conclude that the discrete-time US assumption holds. We evaluate the empirical entropy $\widehat{H} = -\sum_{k=1}^{L} \widehat{\lambda}_k \log \widehat{\lambda}_k$ and compare it to $H_{\text{ref}} = -\sum_{k=1}^{L} p_k \log p_k$, where p_k is again obtained by uniformly sampling the zero-mean empirical PDP. For the discrete-time US assumption to hold, we need $H = H_{\text{ref}}$. Fig. 4 shows, however, that $H < H_{\text{ref}}$ with a gap of up to 1 nat. We can, therefore, conclude that even though the number of stochastic degrees of freedom scales linearly with bandwidth, the discrete-time US assumption is not satisfied. This finding is supported by an analysis of the correlation coefficients between individual taps [22].

VII. Conclusion

On the basis of indoor UWB channel measurements in the frequency band from 2 GHz to 5 GHz, we found that AIC supports Rayleigh or Ricean tap amplitude distributions. This is somewhat surprising, as it is often argued that for large bandwidths the number of partial waves contributing to each tap is not high enough to justify the complex Gaussian assumption by the central limit theorem. We also demonstrated that the differences between the Rayleigh distribution and the Rice, Nakagami, and Weibull distributions in MC I, and between the Rice and the Weibull distribution in MC II are often minor.

The number of significant eigenvalues of the channel impulse response covariance matrix scales approximately linearly with bandwidth. Consequently, the diversity order of the channel shows the same scaling behavior, a common assumption in information-theoretic studies of UWB systems. Nevertheless, we found that the individual channel taps are correlated, thus invalidating the discrete-time US assumption.

References

[1] R. Vaughan and J. B. Andersen, *Channels, Propagation and Antennas for Mobile Communications*. London, UK: The Institution of Electrical Engineers, 2003.

[2] P. A. Bello, "Characterization of randomly time-variant linear channels," *IEEE Trans. Commun.*, vol. 11, no. 4, pp. 360–393, Dec. 1963.

[3] J. N. Pierce, "Ultimate performance of M-ary transmission on fading channels," *IEEE Trans. Inform. Theory*, vol. 12, no. 1, pp. 2–5, Jan. 1966.

[4] A. J. Viterbi, "Performance of an M-ary orthogonal communication system using stationary stochastic signals," *IEEE Trans. Inform. Theory*, vol. 13, no. 3, pp. 414–422, July 1967.

[5] I. E. Telatar and D. N. C. Tse, "Capacity and mutual information of wideband multipath fading channels," *IEEE Trans. Inform. Theory*, vol. 46, no. 4, pp. 1384–1400, July 2000.

[6] S. Verdú, "Spectral efficiency in the wideband regime," *IEEE Trans. Inform. Theory*, vol. 48, no. 6, pp. 1319–1343, June 2002.

[7] M. Médard and R. G. Gallager, "Bandwidth scaling for fading multipath channels," *IEEE Trans. Inform. Theory*, vol. 48, no. 4, pp. 840–852, Apr. 2002.

[8] V. G. Subramanian and B. Hajek, "Broad-band fading channels: Signal burstiness and capacity," *IEEE Trans. Inform. Theory*, vol. 48, no. 4, pp. 809–827, Apr. 2002.

[9] A. F. Molisch, J. R. Foerster, and M. Pendergrass, "Channel models for ultrawideband personal area networks," *IEEE Wireless Commun. Mag.*, vol. 10, no. 6, pp. 14–21, Dec. 2003.

[10] D. Cassioli, M. Z. Win, and A. F. Molisch, "The ultra-wide bandwidth indoor channel: From statistical models to simulations," *IEEE J. Select. Areas Commun.*, vol. 20, no. 6, pp. 1247–1257, Aug. 2002.

[11] J. Keignart and N. Daniele, "Channel sounding and modelling for indoor UWB communications," in *Proc. Int. Workshop on Ultra Wideband Systems*, Oulu, Finland, June 2003.

[12] P. Pagani and P. Pajusco, "Experimental assessment of the UWB channel variability in a dynamic indoor environment," in *Proc. Int. Symp. Personal, Indoor and Mobile Radio Communications*, vol. 4, Barcelona, Spain, Sept. 2004, pp. 2973–2977.

[13] S. S. Ghassemzadeh, L. J. Greenstein, and V. Tarokh, "The ultra-wideband indoor multipath model," AT&T Labs, Florham Park, NJ, U.S.A., Tech. Rep. P802.15 02/282r1SG3a, July 2002, IEEE P802.15 SG3a contribution.

[14] J. Kunisch and J. Pamp, "Measurement results and modeling aspects for the UWB radio channel," in *IEEE Conference on Ultra Wideband Systems and Technologies, Digest of Papers*, Baltimore, MD, U.S.A., May 2002, pp. 19–23.

[15] R. Saadane, A. Menouni, R. Knopp, and D. Aboutajdine, "Empirical eigenanalysis of indoor UWB propagation channels," in *Proc. IEEE Global Telecommunications Conference (GLOBECOM)*, vol. 5, Dallas, TX, U.S.A., Nov. 2004, pp. 3215–3219.

[16] H. Linhart and W. Zucchini, *Model Selection*. New York, NY, U.S.A.: Wiley, 1986.

[17] R. B. D'Agostino and M. A. Stephens, Eds., *Goodness-of-Fit Techniques*. New York, NY, U.S.A.: Marcel Dekker, 1986.

[18] L. Guttman, "What is not what in statistics," *The Statistician*, vol. 26, no. 2, pp. 81–107, June 1977.

[19] H. Akaike, "Information theory and an extension of the maximum likelihood principle," in *Breakthroughs in Statistics*, S. Kotz and N. L. Johnson, Eds. New York, NY, U.S.A.: Springer, 1992, vol. 1, pp. 610–624, originally published in the *Proceedings of the Second International Symposium on Information Theory*, Budapest, Hungary, 1973.

[20] M. A. Taneda, J. Takada, and K. Araki, "The problem of the fading model selection," *IEICE Transactions on Communications*, vol. E84-B, no. 3, pp. 660–666, Mar. 2001.

[21] H. Akaike, "On the likelihood of a time series model," *The Statistician*, vol. 27, no. 3/4, pp. 217–235, Dec. 1978.

[22] U. G. Schuster and H. Bölcskei, "Ultra-wideband channel modeling on the basis of information-theoretic criteria," *IEEE J. Select. Areas Commun.*, 2005, submitted.

[23] R. U. Nabar, H. Bölcskei, and A. J. Paulraj, "Diversity and outage performance in Ricean MIMO channels," *IEEE Trans. Wireless Commun.*, 2005, to appear.

[24] M. Wax and T. Kailath, "Detection of signals by information theoretic criteria," *IEEE Trans. Acoust., Speech, Signal Processing*, vol. 33, no. 2, pp. 387–392, Apr. 1985.