Proceedings of IEEE International Symposium on Power Line Communications 2010 - Rio De Janeiro, Brasil, March 28-32, 2010
ISPLC Best Student Paper Award 2010

Analysis of the PLC Channel Statistics Using a Bottom-Up Random Simulator

Fabio Versolatto and Andrea M. Tonello
DIEGM - Università di Udine - Via delle Scienze 208 - 33100 Udine - Italy
phone: +39 0432 558042 - fax: +39 0432 558251 - e-mail: fabio.versolatto@uniud.it, tonello@uniud.it

***Abstract*—We consider a top-down statistical channel generator where the transfer function between pair of nodes is computed using transmission line theory applied to a randomly generated in-home network topology. We describe the random topology model that has been derived from the observation of regulations and common practices in real scenarios. This approach allows a strong connection with physical reality and constitutes a theoretical framework that makes it possible to derive considerations on the statistical channel characteristics. We focus on the study of the statistics of the channel, we investigate the dependency from the model parameters, and we show that the generated channel responses can be classified in terms of average capacity, or in terms of the location of the associated nodes within the topology layout.**

I. Introduction

Unlike other communication mediums, power lines suffer the substantial absence of a universally accepted model that is able to fully describe the statistical channel properties. Some models have been proposed, but most of them follow a top-down approach that is not able to take into account all the physical realities. An example is given in [1], where it is proposed to divide the set of power line channels into nine classes according to their Shannon's capacity. Each class is associated to a certain average path loss profile. Channel frequency responses of a given class are then randomly generated using a statistical model of notches and peaks. Another top-down model has been proposed in [2]. In this case the idea is to extend in a statistical way the Dostert-Zimmermann analytical model introducing a variability in the parameters. The parameters can be adjusted to obtain a partition of channels into classes following the classification in [1]. Although the analytical expression of the frequency response used in [2] is derived from the multipath propagation model, it is not fully capable of describing the physical reality since it has no strict connection to a certain topology and layout of lines, loads and outlets. Furthermore, it is not capable of describing the time variations of the channel. It has however to be recognized that the top-down channel generators enjoy a fast computer implementation and that are suitable for the analysis of signal processing and communication algorithms.

A closer connection with physical reality can be obtained following a bottom-up approach. In this case, given a certain topology the channel transfer function among two nodes can be obtained via transmission line (TL) theory. Clearly, this requires to know all the information about the topology, i.e., displacement of cables, cable types, line lengths, loads, etc. In addition, the bottom-up approach requires a fast method to compute the channel transfer function since the topology can be rather complex, especially in indoor networks as herein considered.

Now, to obtain a statistical bottom-up channel generator it is necessary to develop a random generator of topologies. To this respect, a first random model for the topology has been presented in [3]. However, the model in [3] focuses on the American indoor scenario and it is not sufficiently general and appropriate to describe the complex topologies encountered in practice. In this paper we present a novel indoor topology generator. It has been derived from the observation of wiring practices and norms in Europe. For each topology realization we compute the transfer function of any pair of outlets using the fast and computationally efficient methodology described in [4] that we refer to as voltage ratio approach.

Some preliminary results about the statistical properties of the proposed bottom-up random channel generator have been reported in [5]. In this paper we report a comprehensive analysis. We study the effect of changing the parameters and in particular the area of the topology. Furthermore, we show that the channel generator allows partitioning the channels into capacity classes or classes related to the location of the nodes. In general the developed channel generator allows having a strong connection with physical reality and constitutes a theoretical framework that makes it possible to derive considerations on the statistical channel characteristics.

The paper is organized as follows. In Section II and III we present the topology generator and the channel transfer function calculation method. In Section IV we study the statistics of the channel focusing on the RMS delay spread and average channel gain. We also deal with channel classification. The conclusions are reported in Section V.

II. The Topology Generator

The topology generator is realized from a model derived by the observation of regulations and common practices in real scenarios. As an example, we report in Fig. 1 a topology arrangement generated by the algorithm described below.

The work of this paper has been partially supported by the European Community Seventh Framework Programme FP7/2007-2013 under grant agreement n. 213311, project OMEGA-Home Gigabit Networks.

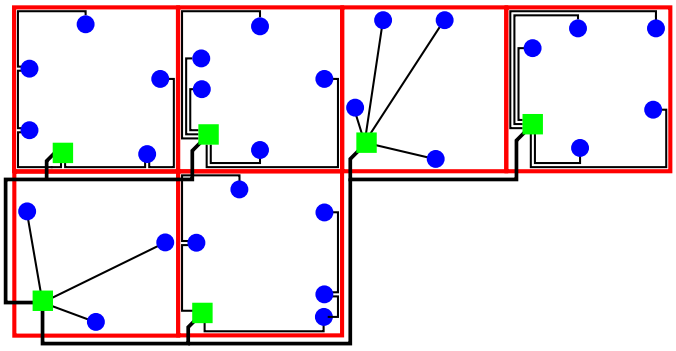

Figure 1. An example of topology arrangement generated by the algorithm. Derivation boxes and outlets are represented by the squared and dotted markers, respectively.

We have carried out an analysis of indoor scenarios that satisfy all the European rules and recommendations on that matter. In most of the cases we surprisingly found a quite regular structure where outlets of the same room are connected to the same node, referred to as "derivation box", and all the derivation boxes are connected together at a second level according to nearness and reachability rules. We have also observed that on average all the outlets fed by the same derivation box are nearby placed around the referenced derivation box in a limited area that has quite regular dimensions for all derivation boxes. Then, we have concluded that the location plan can be divided in elements that contain a derivation box and associated outlets. By the experimental observation about the existence of regular dimensions, we propose to simplify the representation using elements of the same area that are referred to as "clusters". The particularity of clusters is that even if there is no bond in their shape, in practice they can always be well represented by a rectangle with a fixed area, but variable dimension ratio. A simpler representation uses a square shaped cluster. Each cluster contains all the outlets connected to a derivation box and the derivation box itself. Different clusters are usually interconnected only through the derivation boxes. Observations of real scenarios also suggest a significant correspondence between the clusters and the room shapes. This allows an easier understanding of the network topology and of the cluster displacement starting from the location plan.

According to the experimental observations, and with the proposed partition of the topology into clusters, we have derived a statistical topology generation algorithm, where a location plan of area A_f is build up as a random displacement of neighboring clusters. For a given location plan realization, all the clusters have square shape with the same area A_c. A_c is assumed to be uniformly distributed in a realistic interval. The number of clusters N_c that build up a topology of area A_f is defined as

$$N_c = \lceil A_f / A_c \rceil \,. \tag{1}$$

Without any loss of generality, we distribute outlets only along the cluster perimeter. We do so assuming that there is no difference between an outlet placed inside the cluster and an outlet placed along its perimeter, if connection lengths are the same in both cases. We choose the number of outlets inside each cluster as a Poisson variable of intensity Λ_p, so outlets are placed according to a Poisson arrival process. This means that given their number, outlets are uniformly distributed along the cluster perimeter and this is the most plausible choice.

Derivation boxes are always placed inside their respective clusters and next to the same corner, so they are as regular spaced as we have found in practice. More in detail, to introduce some variability, each derivation box is shifted from its reference corner by a bi-dimensional offset generated as a pair of uniform distributed variables defined between 0 and half of the length of the cluster.

Outlets are connected to the associated derivation box according to the three most common practices that we have found, i.e., a star structure that satisfies the minimum distance criterion, a star topology with conductors placed along the sides only, and a bus topology. At the second layer, connections between derivation boxes fulfill the reachability and the nearness criteria with cables of section according to norms such that voltage drops are reduced. The special role played by the service panel, that is a derivation box itself, is taken also into account.

We have also studied the impact of the loads. For this purpose, we have collected from experimental measurements a set of 10 loads for the indoor scenario, such as lamps or computer transformers. These loads are randomly selected. If we define with p_v the probability that no loads are connected to a plug, the probability p_l that a given load extracted from the previous set is plugged to an outlet where an appliance is connected, reads

$$p_l = \frac{1 - p_v}{10}. \tag{2}$$

In order to simulate different scenarios the parameters p_v, as well as A_f and A_c, are appropriately adjusted.

III. Transfer Function Calculation

To compute the transfer function among any pair of outlets o_1, o_2 in a certain topology realization we use the methodology presented in [4] that exploits TL theory. More in detail, we first find the shortest signal path between the pair of outlets and we refer to it as "backbone". We then remap the entire topology such that we obtain an equivalent topology that comprises the backbone and a set of branches that depart from intermediate junctions. Then, we divide the backbone in elementary blocks called "units". A unit contains the portion of line between two intermediate nodes of the backbone, a junction to which a branch is connected to, and the branch itself. We then compute the transfer function of each unit as the voltage ratio between the two input/output ports of the unit. Finally, the overall channel transfer function $H(f)$, i.e., the insertion loss between o_1 and o_2, is obtained by the multiplication of the sub-transfer functions $H_i(f)$ of the N_u component units:

$$H(f) = \prod_{i=1}^{N_u} H_i(f) \tag{3}$$

With this methodology the channel transfer function is computed in a computationally efficient way, even for complex topologies where time-domain algorithms fail or become too complex.

IV. Statistical Analysis

We now report the analysis of the statistical properties of the channel generator. The parameters that can be adjusted are A_f, A_c, Λ_o and p_v that respectively denote the area of the topology, the area of the cluster, the intensity of the Poisson process that controls the distribution of the outlets and the probability that no loads are connected to an outlet. The parameters have been optimized to obtain results close to the ones obtained from experimental campaigns and the observations of regulations and common practices in real scenarios. Tab. I collects the parameters. As it can be noticed, we consider three topology areas A_f.

Table I
Parameter values for the setup of the bottom up simulator

Parameter	Value
A_f $[m^2]$	$\{100, 200, 300\}$
A_c $[m^2]$	$\mathcal{U}(25, 65)$
Λ_o [outlets/m^2]	0.6
p_v	0.6

The CTFs (channel transfer functions) are computed in the frequency domain in the 1-30 MHz band. The impulse response $h_i = h(iT_s)$ is obtained via the Inverse Discrete Fourier Transform with $2N$ points and sampling frequency resolution $F_s = 60/2NMHz = 100kHz = 1/T_s$.

The following analysis focuses on the root mean square delay spread and the average channel gain.

A. Delay Spread Analysis

The RMS delay spread is an important metric for the design of communication algorithms and it is defined as

$$\sigma_\tau = \sqrt{\frac{\sum_{i=0}^{2N-1} \left(iT_s\right)^2 |h_i|^2}{\sum_{i=0}^{2N-1} |h_i|^2} - \left(\frac{\sum_{i=0}^{2N-1} iT_s |h_i|^2}{\sum_{i=0}^{2N-1} |h_i|^2} \right)^2}. \quad (4)$$

In Fig. 2 we report the cumulative distribution function (CDF) of the RMS delay spread evaluated as a function of the topology area A_f that is equal to 100, 200 and 300 m^2. We have drawn 500 channel realizations for each topology size. In all cases, the CDF is well fitted by a log-normal distribution. As it can be seen, larger topologies manifest channels with larger values of RMS delay spread. In Tab. II we report the maximum, the average and the standard deviation of the RMS delay spread in all three cases.

We now investigate in more depth the log-normal behavior of the RMS delay spread. First, we merge all channel realizations for all three areas in a unique set $\mathbf{\Sigma}$. We then consider the following logarithmic normalized delay spread

$$\overline{\sigma}_\tau = \log\left(\frac{\sigma_\tau}{10^{-6}}\right), \quad \sigma_\tau \in \mathbf{\Sigma}\,. \quad (5)$$

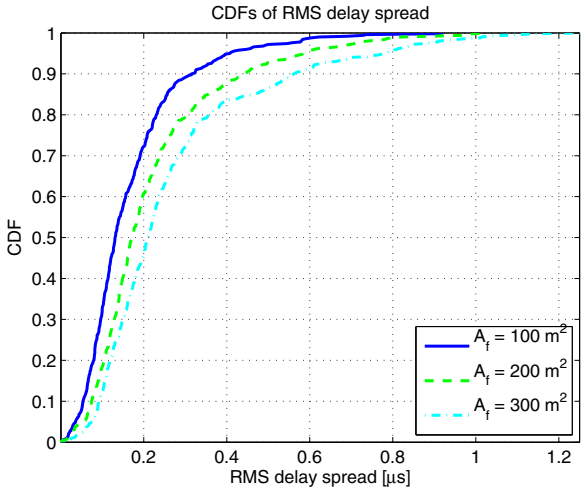

Figure 2. Cumulative distribution functions of the RMS delay spread as a function of the topology area A_f.

In Fig. 3 we report the quantile-quantile plot of $\overline{\sigma}_\tau$ versus the normal distribution. As it can be seen in Fig. 3, $\overline{\sigma}_\tau$ doesn't exhibit a linear trend for all values. Therefore, the delay spread is not strictly log-normal, and the deviations from normality of $\overline{\sigma}_\tau$ are more pronounced for lower values.

This behavior can be explained by observing that log-normality is the result of a very large number of multipath multiplicative effects [6]. If the RMS delay spread is small, it is reasonable to suppose that the associated channels are not affected by strong multipath effects, so that the log-normal behavior is no more justified, as Fig. 3 suggests.

To confirm this interpretation, we evaluate the statistics of the delay spread under the constraint that the pair of outlets is connected with a backbone of length larger than 5 m. In this way we discard the channels associated to outlets that are close and that are supposed to be less affected by multipath effects. This is because we speculate that very short backbones present less discontinuities, less reflections and therefore, less multipath propagation effects. We define with $\hat{\mathbf{\Sigma}}$ the set of RMS delay spread realizations conditioned on a backbone longer than 5 m. The logarithmic normalized conditional delay spread is denoted with $\overline{\hat{\sigma}}_\tau$. The corresponding quantile-quantile plot is shown in Fig. 3. Now, the delay spread $\overline{\hat{\sigma}}_\tau$ has a distribution closer to the normal one.

We have then performed two different tests to reject or not the null hypothesis for $\overline{\hat{\sigma}}_\tau$ samples, at a 5 % of significance level. In particular, we have performed the Jarque-Bera and

Table II
Analysis of the RMS Delay Spread

Topology Area $[m^2]$	Mean Value $[\mu s]$	Std. Deviation $[\mu s]$	Max. Value $[\mu s]$
100	0.167	0.127	0.921
200	0.222	0.168	1.027
300	0.271	0.210	1.250

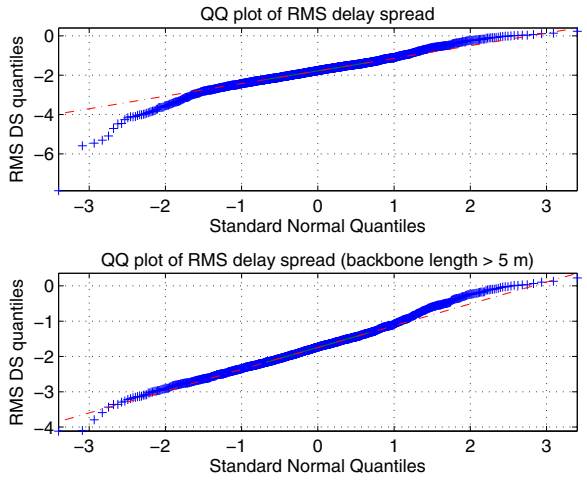

Figure 3. Quantiles of $\overline{\sigma}_\tau$ (in the top plot) and $\overline{\hat{\sigma}}_\tau$ (in the bottom plot) versus standard normal quantiles.

Lilliefors tests [7], [8] and both of them have rejected the null hypothesis. This means that even if $\overline{\hat{\sigma}}_\tau$ exhibits a quite linear trend in Fig. 3, the more accurate tests reveal that it is not normal distributed. This result appears in contrast to the one reported from the analysis of measured channels in [6] that claims a normal nature for the logarithm version of RMS delay spread. The inconsistency can be explained with the different number of realizations managed in the two cases. In fact we consider 1500 realizations, while in [6] only 120 measured channels are used.

To better compare results, we have randomly chosen 200 realizations from the set $\hat{\mathbf{\Sigma}}$. This set of channel realizations is denoted with $\hat{\mathbf{\Sigma}}_c$. We have then performed again the Jarque-Bera and Lilliefors tests for $\overline{\hat{\sigma}}_{\tau_c}$, which realizations are defined as

$$\overline{\hat{\sigma}}_{\tau_c} = \log(\sigma_\tau / 10^{-6}) \quad \sigma_\tau \in \hat{\mathbf{\Sigma}}_c \,, \tag{6}$$

The results are shown in Tab III and they say that the null hypothesis is fulfilled for $\overline{\hat{\sigma}}_{\tau_c}$. This suggests that the tests are strongly affected by the number of realizations evaluated. Therefore, we conclude that 100 realizations are not statistically sufficient to make accurate statistical considerations about the channel properties. However, the log-normal distribution remains an acceptable approximation for the delay spread distribution.

Table III
NULL HYPOTHESIS TESTS

	Jarque-Bera	Lilliefors
$\overline{\sigma}_\tau$	rejected	rejected
$\overline{\hat{\sigma}}_\tau$	rejected	rejected
$\overline{\sigma}_{\tau_c}$	rejected	rejected
$\overline{\hat{\sigma}}_{\tau_c}$	accepted (0.5)	accepted (0.38)

Finally, we have also evaluated the impact of the parameters in the statistics of the RMS delay spread. In particular, we have found that the cluster area influences the lower part of the delay spread CDF curve. The higher is A_c, the lower are the CDF values for little RMS delay spreads. A_f and Λ_o are responsible for the higher range of RMS delay spread values. We have found that when $A_f > 300\, m^2$ or $\Lambda_o > 0.8$ outlets$/m^2$ the RMS delay spread reaches high values, up to $1.5\, \mu s$.

B. Average Channel Gain Analysis

The second metric that we consider herein is the average channel gain (ACG), that is defined as

$$G = 10^{\frac{G_{dB}}{10}} = \frac{1}{N} \sum_{i=0}^{N-1} |H_i|^2 \,. \tag{7}$$

In Fig. 4 we show the quantile-quantile plot of the ACG in dB (G_{dB}) versus the standard normal. The non linear trend of G_{dB} tells us that it is not normally distributed as confirmed also by Jarque-Bera and Lilliefors tests.

To better investigate the real nature of ACG we have evaluated its behavior in relation to the topology area A_f. In the top plot of Fig. 4 we report the quantile-quantile plot of the ACG for the topologies of area 100, 200 and 300 m^2 (without any bound in the backbone length). Now, for a given topology area the ACG in dB exhibits a normal behavior. Some discrepancy from normality is exhibited only for high values.

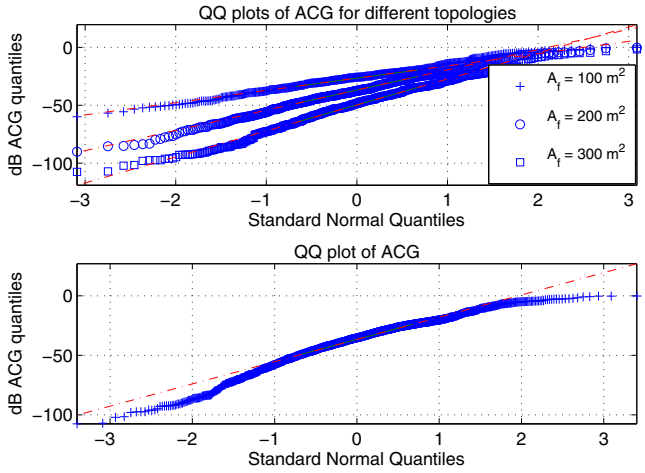

Figure 4. In the top plot: quantiles of the average channel gain in dB for different topologies areas versus the standard normal quantiles. In the bottom plot: quantiles of the "overall" average channel gain in dB versus the standard normal quantiles.

Furthermore, we have found that G_{dB} for the topologies of area $100\, m^2$ are spread down to $-60\, dB$, while for topologies of area $200\ m^2$ down to less than $-80\ dB$ and in topologies of area $300\ m^2$ down to about $-100\ dB$. It is worthwhile noting that the standard deviation is significantly different in the three cases and this could explain why the bottom plot of Fig. 4 shows that the ensemble of channels with unconstrained topology area does not have a normal behavior. In fact, if we

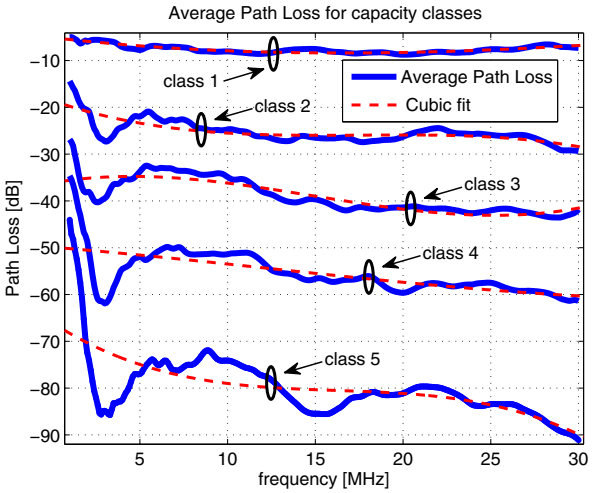

Figure 5. Average path loss profiles and their cubic fit for each class of capacity.

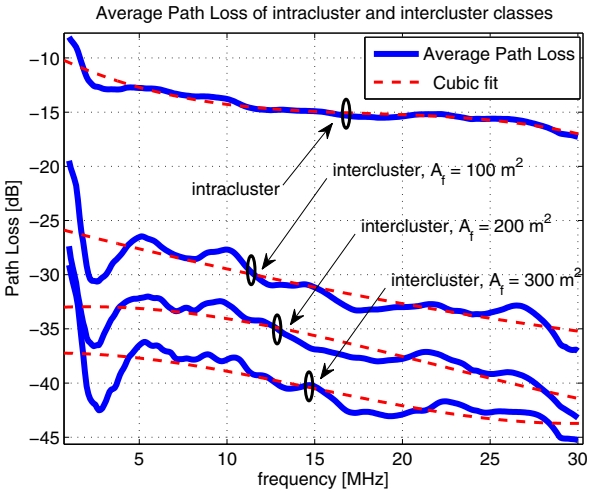

Figure 6. Average path loss profiles and their cubic fit for the intracluster and intercluster classes.

consider a random variable that is obtained merging realizations of normal variables with different standard deviations, we do not get a normal random variable.

C. Channel Classification

The bottom-up random channel generator allows generating channels and partition them into certain classes, i.e., it allows a classification.

A first classification can be made according to the topology area A_f. A second classification is based on the average capacity. We have partitioned the channel realizations into 5 classes of capacity according to Tab. IV, assuming a transmitted Power Spectral Density (PSD) of -50 dBm/Hz and additive white Gaussian noise with PSD = -140 dBm/Hz. We have divided the capacity range 0 – 900 Mbps into 5 intervals of 180 Mbps each, and we have filled each class with realizations that present a capacity in the considered interval. Tab. IV also shows the percentage of channels that fall in a certain class. This classification is similar in spirit to the one proposed in [9] and [10] where, however, 9 capacity intervals were proposed for channels in the 1-100 MHz frequency band.

Table IV
CAPACITY INTERVALS

Class	Average Class Capacity [Mbps]	Percentage of Channels [%]
1	810	6.8
2	630	20.6
3	450	29.4
4	270	27
5	90	16.2

Now, for each capacity class we have evaluated the average path loss in Fig. 5. We also report a cubic polynomial fit of the path loss which essentially shows that there is a one-to-one correspondence between average capacity and average PL profile. Worse capacity classes are characterized by ripple

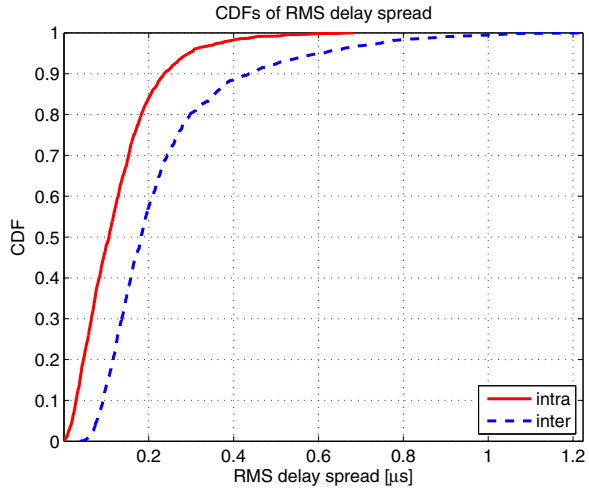

Figure 7. Cumulative distribution functions of RMS delay spread for intracluster and intercluster classes.

effects that can be traced back to particular topology practices, i.e. the star cable layout inside clusters is mostly responsible for this behavior.

The exploitation of the topology information in the bottom-up channel generator allows a third channel classification. It is based on the division of the set of channel realizations in two new sets that we define as intracluster and intercluster. The first set comprises channel realizations among pair of outlets belonging to the same cluster, while the second class comprises channels among outlets that belong to distinct clusters.

In Fig. 6 we report the average path loss and its cubic polynomial fit for the intracluster and intercluster channels. Differences between the two profiles can be traced back to the different topological nature of such channels. More in detail, the intracluster path loss profiles do not change with the topology area, but they are influenced by the cluster area.

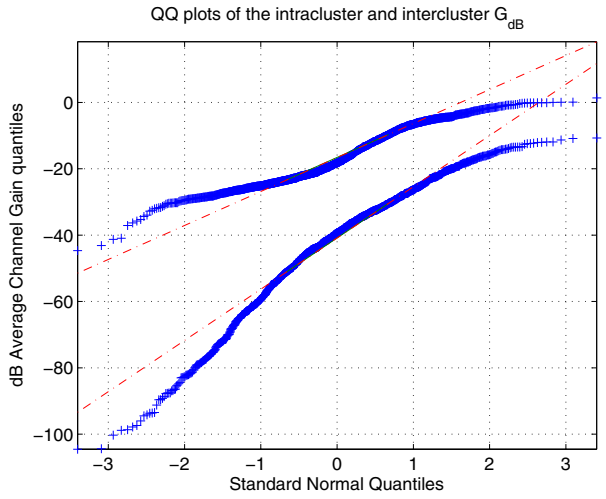

Figure 8. Quantiles of the intercluster and intracluster average channel gain in dB versus standard normal quantiles.

Conversely, it is straightforward to note that intercluster path loss profiles are a function of the topology area A_f.

In Fig. 7 we show the CDF of the RMS delay spread for the intercluster and intracluster channels, while the quantile-quantile plots of G_{dB} are shown in Fig. 8. To obtain these plots we have merged all channel realizations for the three topology areas. As the intuition suggests, the intercluster channel exhibit higher delay spreads, that can be quantified, let's say, by a factor of more than two with respect to that exhibited by the intracluster channels. Fig. 8 shows an interesting behavior too. In fact, for the intracluster channels, G_{dB} shows a lower spread trend than for the intercluster channels. This is because intracluster channels are insensitive to the topology area. Conversely intercluster channels are strongly affected by the topology area.

V. Conclusions

In this paper we have considered statistical channel modeling using TL theory on randomly generated topologies. This approach allows maintaining strong connection to physical reality and deriving a theoretical framework for the analysis of the statistical characteristics of the channel. The simulator is capable of efficiently computing the transfer function between any pair of nodes in any complex topology through a voltage ratio method that partitions the network in elementary units. We have focused on the description of the random topology model that has been derived from the observation of common practices and norms. The key idea is the division of the location plan into elementary cells that are referred to as clusters. Each cluster owns a derivation box that serves a number of outlets. Then, we have investigated the behavior of the channel as a function of the model parameters and in particular the total area and the area of the cluster. The statistical analysis of the RMS delay spread and average channel gain reveals that they are only approximately log-normal in contrast to conclusions reported from the analysis of measured channels [6]. We have shown that the discrepancy is due to the limited number of channels used in previous statistical studies as well as the inability, in those studies, of selecting channels according to the principal path distance. Even if not reported here due to space limitations, we have carried out an in-depth comparison with results from experimental campaigns in [11] which shows the good agreement between the simulated and the measured channels. In particular, we have found that excellent convergence between simulated and experimental results is achieved if the statistical analysis is done over subsets of channels that satisfy the same conditions assumed in measurement campaigns, e.g., a constraint on the minimum backbone length. In fact, measurements are typically, and deliberately, made among pair of outlets that are sufficiently far apart since they are considered as representative of locations where the modems might be installed. This is why we have herein reported a delay spread analysis for channels associated to a sufficiently large backbone length.

Finally, we have shown that the channel generator allows partitioning the channel responses into classes that are a function of certain classification metrics, e.g., the topology area, the average capacity, the backbone length among pair of nodes, the belonging of the outlets to the same cluster or to distinct clusters. In conclusion the developed channel simulator is a powerful tool that allows the generation of statistically representative channels and the in-depth understanding of their characteristics as a function of the topology layout.

References

[1] M. Tlich, A. Zeddam, A. Moulin, F. Gauthier, and G. Avril, "A broadband powerline channel generator," in *Proc. IEEE Int. Symp. Power Line Commun. and its App.*, Mar. 2007, pp. 505–510.

[2] A. M. Tonello, "Wideband impulse modulation and receiver algorithms for multiuser power line communications," *EURASIP Journal on Advances in Signal Processing*, vol. 2007, pp. 1–14, 2007.

[3] T. Esmalian, F. R. Kschischang, and P. Glenn Gulak, "In-building power lines as high-speed communication channels: channel characterization and a test channel ensemble," *International Journal of Communication Systems*, vol. 16, pp. 381–400, 2003.

[4] A. M. Tonello and T. Zheng, "Bottom-up transfer function generator for broadband plc statistical channel modeling," in *Proc. IEEE Int. Symp. Power Line Commun. and its App.*, Apr. 2009, pp. 7–12.

[5] A. M. Tonello and F. Versolatto, "New results on top-down and bottom-up statistical plc channel modeling," in *Proc. of Third Workshop on Power Line Communications and its Applications*, Oct. 2009, pp. 11–14.

[6] S. Galli, "A simplified model for the indoor power line channel," in *Proc. IEEE Int. Symp. Power Line Commun. and its App.*, Apr. 2009, pp. 13–19.

[7] H. Lilliefors, "On the Komogorov-Smirnov test for normality with mean and variance unknown," *Journal of the American Statistical Association*, no. 62, pp. 399–402, 1967.

[8] C. Jarque and A. K. Bera, "A test for normality of observations and regression residuals," *International Statistical Review*, vol. 55, no. 2, pp. 1–10, 1987.

[9] P. Pagani, M. Tlich, A. Zeddam, A. M. Tonello, F. Pecile, S. D'Alessandro, G. Mijic, and K. Kriznar, "Plc channel transfer function models for the Omega ICT project," in *Proc. of ICT Mobile Summit 2009*, Jun. 2009.

[10] FP7 Theme 3 ICT-213311 OMEGA, "Plc channel characterization and modelling," Tech. Rep. Deliverable 3.2, Dec. 2008.

[11] A. M. Tonello and F. Versolatto, "Bottom-up statistical plc channel modeling - part II: Inferring the statistics," *submitted to IEEE Trans. on Power Delivery*, Feb. 2010.