Bernoulli-Gaussian Distribution with Memory as a Model for Power Line Communication Noise

Victor Fernandes, Weiler A. Finamore, Moisés V. Ribeiro, Ninoslav Marina, and Jovan Karamachoski

Abstract—The adoption of Additive Bernoulli-Gaussian Noise (ABGN) as an additive noise model for Power Line Communication (PLC) systems is the subject of this paper. ABGN is a model that for a fraction of time is a low power noise and for the remaining time is a high power (impulsive) noise. In this context, we investigate the usefulness of the ABGN as a model for the additive noise in PLC systems, using samples of noise registered during a measurement campaign. A general procedure to find the model parameters, namely, the noise power and the impulsiveness factor is presented. A strategy to assess the noise memory, using LDPC codes, is also examined and we came to the conclusion that ABGN with memory is a consistent model that can be used as for the evaluation of the noise effect on the digital communication (PLC) systems is the subject of this paper. ABGN is a discrete-time and memoryless model that is equivalent to the continuous-time Two-state, Middleton Class A model. Moreover, a comparison performance analysis of the digital communication system when the transmission is perturbed by the measured noise and by the noise generated according to ABGN model is discussed. Due to the limitation of ABGN model to handle memory-sensitive procedures, we propose a more sophisticated model, which takes into account the memory that is intrinsic to the nature of the PLC noise. This model is so-called ABGN with memory (ABGNM) and it is suitable for PLC systems that make use of any memory-dependent tool (e.g., error control codes).

Based on the presented analyses, we came up to the conclusion that the ABGN model is a good model for the PLC noise when the digital communication system does not use any memory-sensitive procedure. However, it does not mimic the system performance with measured noise when it has memory, which is a similar behavior of Middleton Class A model. Finally, we show that the ABGNM model can mimic the measured PLC noise, resulting in a similar performance for both measured and synthetic noises even when the digital communication system has any memory-dependent tool (e.g., error control codes).

II. COMMUNICATION SYSTEM MODEL

A digital communication system in which a modulated signal is transmitted through a medium is, in its simplest form, modeled by a channel, where the input signal at the receiver is \( y(t) = s(t) + w(t) \). This corresponds to modeling a transmission medium by the channel in which \( s(t) \) is the channel input (modulated signal), \( w(t) \) represents the noise perturbing the transmission, and \( y(t) \) is the channel output. In this section, we focus on a digital communication system with a binary phase shift keying (BPSK) modulator, transmitting a string of bits, which its \( i^{th} \) bit is represented by \( \{u_i\} (i \in \mathbb{Z} \text{ and } u_i \in \{0, 1\}) \). The modulated signal is expressed by

\[
s(t) = \sum_{i=-\infty}^{\infty} (1 - 2u_i)p(t - iT)\cos 2\pi f_c t,
\]
in which the formatting pulse is \( p(t) \) with energy equal to one, \( T \) is the signaling interval, and \( f_c \) is the modulating frequency.

A. White Gaussian Noise

The channel output (receiver input) \( y(t) = s(t) + w(t) \), representing the input signal added to the noise, is the observation that will be used by the receiver to estimate the information that has been sent. The noise perturbing the digital communication is, in many situations, modeled as a zero mean stochastic process \( W(t) \) with power spectral density \( S_w(f) = N_0/2 \), rendering a value (at the decider input) with variance \( \sigma_w^2 = N_0/2 \). This is the simple AWGN(\( \sigma_w^2 \)) model which have been extensively used to model different communication channels. The energy-per-bit \( E_b \) at the receiver input and the ratio \( E_b/N_0 \), of the energy-per-bit and the noise parameter (a parameter of interest to characterize the state of the system), can be easily calculated from (1).

In such digital communication systems, assuming perfect synchronization, the continuous signal \( y(t) \) is processed by the receiver as follows:

\[
\begin{align*}
y_o(t) &= y(t) \cos 2\pi f_c t \\
r(t) &= y_o(t) \ast p(T-t) \\
r(t) &= \{[s(t) + w(t)]\cos 2\pi f_c t \} \ast p(T-t) \\
r_n(t) &= \sum_{i=-\infty}^{\infty} (1 - 2u_i)q(t-iT) + 2w(t) \cos 2\pi f_c t \ast p(T-t) \\
x_n(t) &= \sum_{i=-\infty}^{\infty} (1 - 2u_i)\delta_{n-i} + w_n.
\end{align*}
\]

where \( * \) denotes linear convolution, \( p(t) \ast p(-t) = q(t) \) fulfill the 1st criterion of Nyquist for having free intersymbol interference, \( \delta_n \) is the impulse sequence, and \( w_n \) is the \( n \)th discrete-time noise sample. Note that \( \delta_n = 1 - 2u_n + w_n \) is mapped by a decoder into a sequence \( \delta_n = [(1 - \text{sgn}(z_n))/2] \), in which \( \text{sgn}(z_n) = -1 \), if \( z_n < 0 \) and \( \text{sgn}(z_n) = 1 \), otherwise. If the noise is absent, then the probability of having the random variables (rvs) \( U_n \) and \( U_n \), which model the sent and received bits taking the same value, would be \( P[\bar{U}_n = U_n] = 1 \). The noise component \( w_n \) is, however, always present and the performance of this digital communication system (and all digital communication systems transmitting over a transmission medium modeled as an AWGN channel, for that matter) is analyzed by examining the plot with the horizontal and vertical axis labeled by \( E_b/N_0 \) and \( P_e = P[\bar{U}_n = \bar{U}_n] \), respectively, the latter being the probability of error between the bits delivered to the destination and the transmitted ones.

If we assume that the noise component sequence \( \{w_n\} \) is modeled as a sequence of independent and identically distributed (i.i.d.) zero mean Gaussian rvs with variance \( \sigma_n^2 = N_0/2 \), the probability of error expression in this situation, as it is widely known, is

\[
P_e = Q\left(\sqrt{\frac{E_b}{\sigma_n^2}}\right),
\]

where \( Q(x) = \frac{1}{\sqrt{2\pi}} \int_{x}^{\infty} e^{-t^2/2} dt \).

B. Bernoulli-Gaussian Noise

The ABGN is a discrete-time and memoryless model that is equivalent to the continuous-time Two-state Middleton Class A model. In this model, besides the background noise power, characterized by \( N_0 \), two other parameters: \( \alpha \) (which characterizes how much stronger than the background noise the impulsive noise is) and \( p \) (which characterizes the percentage of time that only background noise is present) are used when specifying the ABGN(\( \sigma_b^2, p, \alpha \)) model.

A digital communication system in which a modulated signal is transmitted over a PLC channel is, in its simplest form, modeled as having the signal \( s(t) \), at the channel input and \( y_{PLC}(t) = s(t) + w_{PLC}(t) \) is the channel output, where \( w_{PLC}(t) \) is the PLC additive noise.

As discussed in [7], the received and filtered PLC noise can be modeled as a stochastic process \( V(t) \). At the receiver end (at the input of the decider), at a given time instant \( t \), \( V(t) \) can be a zero mean “weak” Gaussian rv with variance \( \sigma_V^2 \) and, at a distinct time instant \( t_2 \), \( V(t_2) \) can be a zero mean “strong” Gaussian rv with variance \( \sigma_V^2 = \alpha^2 \sigma_b^2 \). \( V(t) \) and \( V(t_2) \) are independent rvs. The noise component sequence \( \{V_n\} \), which can be modeled as an ABGN(\( \sigma_b^2, p, \alpha \)), is defined as follows [8]:

**Definition #1 (ABGN(\( \sigma_b^2, p, \alpha \))):** Let \( \{U_n\} \) be the \( n \)th sample of a sequence of Bernoulli rv with \( P[U_n = 1] = p \) (and, of course, \( P[U_n = 0] = 1 - p \)) and \( \{W_n\} \) be a sequence of i.i.d. Gaussian rv with zero mean and variance \( \sigma_b^2 \). The sequence of rvs, \( \{V_n\} \), with the \( n \)th sample given by

\[
V_n = U_n W_n + \alpha(1-U_n)W_n,
\]

in which \( \alpha > 1 \), is an ABGN(\( \sigma_b^2, p, \alpha \)).

Every sample \( V_n = \alpha W_n \) corresponding to \( U_n = 0 \), is called a strong variance noise component of the ABGN. These samples are, zero mean, Gaussian rvs with variance \( \alpha^2 \sigma_b^2 \). The remaining samples \( V_n = W_n \), corresponding to \( U_n = 1 \), are Gaussian rvs with zero mean and variance \( \sigma_b^2 \) referred to as background noise components. The variance of the ABGN, which models the PLC noise is thus

\[
\sigma^2 = \rho \sigma_b^2 + (1 - \rho) \alpha^2 \sigma_b^2.
\]

Let us now consider that a string of bits has been transmitted over a PLC system and the signal delivered to the receiver is a BPSK modulated signal with energy per bit equal to \( E_b \). We are interested in the performance of the power line digital digital communication system which transmits bits through a binary input channel (with input \( x_t \in \{+1, -1\} \)) perturbed by noise components modeled as \( \text{ABGN}(\sigma_b^2, \alpha, p) \). At the input of the decoder there will be a signal \( y_t = x_t + V_t \), which is modeled as the sum of \( x_t \) and a zero mean Gaussian rv \( V_t \) which, with probability \( \rho \) has variance \( \sigma_b^2 \) and, with probability \( 1 - \rho \) has variance \( \alpha^2 \sigma_b^2 \). The probability of error in the PLC system as a function of \( E_b/N_0 \), is given by the well known expression [7]

\[
P_e^{\rho, p} = P(Q\left(\sqrt{\frac{E_b}{\sigma_b^2}}\right)) + (1 - p)Q\left(\sqrt{\frac{E_b}{\alpha^2 \sigma_b^2}}\right).
\]

329
The plot of $P_e^{2.7,0.76}$ versus $E_b/N_0$ (green line), is depicted in Fig. 1 together with two other probability of error curves: $P_e$ versus $E_b/N_0$ for $p = 1$ ($\sigma = \sigma_0$) and $P_e$ versus $E_b/N_0$ for $p = 0$ ($\sigma = \alpha \sigma_0$). In the same figure we present, from results obtained in our simulations, the curve of bit error rate (BER) versus $E_b/N_0$ for a digital communication system impaired by ABGN($\sigma_0^2, 0.76; 2.7$).

Fig. 1: Probability of error, $P_e$, vs $E_b/N_0$ performance of a digital communication system which transmits a BPSK signal through (1) an AWGN($\sigma_0^2$) channel (-.), (2) an AWGN($2.7^2 \sigma_0^2$) channel (-o-), and (3) a system in which a BPSK signal is transmitted through a channel corrupted by an ABGN($\sigma_0^2, 0.76; 2.7$) (---). The BER vs $E_b/N_0$ of a simulated system perturbed by measured noise with parameters estimated to be $\hat{p} = 0.76$ and $\hat{\alpha} = 2.7$ (++) is also shown.

As remarked in [7], when the value of $E_b/N_0$ is low, both the reference model (AWGN($\sigma_0^2$)) and ABGN($\sigma_0^2, 0.76, 2.7$) exhibit a performance which are equivalent. On the other hand, when comparing the two models at large values of $E_b/N_0$, one can see that the required $E_b$ to communicate over PLC channel is quite larger than the AWGN($\sigma_0^2$) (about 8 dB, as depicted Fig. 1).

Also, Fig. 1 shows a comparison among the values of $P_e^{2.7,0.76}$ and the BER curve of the measured noise (solid magenta line with dots), which has estimated parameters ($\hat{\alpha}, \hat{p}$) = (2.7, 0.76). These curves shows that theoretically and simulation obtained results are in good agreement.

C. Bernoulli-Gaussian Noise with Memory

For digital communication systems which use error control codes, this ABGN model is too simplistic since it fails to capture the noise memory which, if we closely observe the measured PLC noise, is present. The Bernoulli-Gaussian model can be extend to include the effect of the noise memory which, as it will be seen, affects the performance of digital communication systems that use error control codes. The extension used here assumes that memory keeps the noise, for a multiple of $L$ (the memory length), in either the background state (weak noise) or in the impulsive state (strong noise). We start thus defining an ABGN with memory (ABGNM) with four parameters, $\sigma_0^2, p, \alpha$ and $L$, as follows:

**Definition #2 (ABGNM($\sigma_0^2, p, \alpha, L$)):** Let $\{U_n\}$ be the $n^{th}$ sample of a sequence of Bernoulli rv with $\mathbb{P}(U_n = 1) = 1 - \mathbb{P}(U_n = 0) = p$ and $\{W_n\}$ be a sequence of i.i.d. Gaussian rvs with zero mean and variance $\sigma_0^2$ and $L \in \mathbb{Z}_+$. Then, the sequence of rvs, $\{Z_n\}$, with the $\ell^{th}$ sample given by

$$Z_{\ell} = U_\ell W_\ell + \alpha(1 - U_\ell) W_{\ell},$$

in which $\alpha > 1$, $\ell = nL + j$, $n \in \mathbb{Z}_+$, and $j \in \mathbb{Z}_+$ ($0 \leq j \leq L - 1$), is an ABGNM($\sigma_0^2, p, \alpha, L$). ■

If we let $L = 1$, we get $\ell = n$ and we have the strict sense ABGN (i.e., memoryless) case in which

$$Z_n = U_n W_n + \alpha(1 - U_n) W_n.$$  

D. Measured Noise

Sequences of samples obtained by measuring the noise over indoor electric power grids were used to compare practical and theoretical results. The notation $\{\tilde{z}_n\}$ will be used when referring to a sequence of samples of the PLC measured noise. We then consider the channel perturbed by our measured noise $\{\tilde{z}_n\}$ and the performance of the system which transmit bits over a channel perturbed by $\{\tilde{z}_n\}$ a sequence of rv synthetically generated according to the mathematical model ABGN($\sigma_0^2, p, \alpha$) in (4).

In [9], authors compare theoretically calculated $P_e$ performance of digital communication systems (impaired by independent random noise variables) with the BER performance obtained by simulating the transmission over a channel with synthetic noise. They reached the obvious conclusion that the results are equivalent, in the same lines that we can draw the conclusion, comparing the plot of $P_e$ (theoretically calculated) and BER (obtained with synthetic noise and Monte Carlo simulation) are equivalent (of course the theoretical probability of error values and the BER results obtained by simulation are to be in good agreement).

An issue of great concern, when using a noise model, is how to find the model parameters given the samples of the measured noise. To assess how well can the ABGN model mimic the behavior of practical systems we need to find the ABGN parameters from the sequence of measured noise samples. A procedure proposed to find the ABGN parameters of a given sequence $\{\tilde{z}_n\}$ is discussed next.

III. Determination of the Bernoulli-Gaussian parameters of measured PLC noise

Performance (BER vs $E_b/N_0$) like those in Fig. 1, can be easily obtained once the parameters $\sigma_0^2$, $p$ and $\alpha$ are known. In practice, the values of the parameters for a given sequence $\{\tilde{z}_1, \ldots, \tilde{z}_L, \ldots, \tilde{z}_M\}$, of $M$ noise samples obtained by measurements are not known. The procedure described succinctly next (details of this procedure are presented in [8]), specifies how to classify each measured noise sample $\tilde{z}_n$ as either the realization of a weak rv $(1 - U_n) W_n$ or the realization of a strong rv $\alpha U_n W_n$. By doing so, we partition this sequence and find the estimates $\hat{\sigma}_0^2$, $\hat{\alpha}$ and $\hat{\beta}$.

This procedure is derived from the classical decision problem in which, upon receiving a string, drawn from known
distributions (of the rvs $U_n$ and $W_n$) with known values of $\sigma_0^2$, $\alpha$, and $p$, one has to find a threshold $\tau^*$, that minimizes the probability of taking the wrong decision.

Let us say that the decision is modeled as a rv $\hat{U}_n$ and that the decision rule declares an observed sample $\hat{z}$ to be a sample from a weak noise if $\hat{z} \leq \tau^*$ (i.e., stating that $\hat{U}_n = 0$) or, otherwise, to be a strong noise sample (i.e., stating that $\hat{U}_n = 1$). It can be shown that the probability of wrong decision (i.e., the probability $P_d = P[\hat{U}_n \neq U_n]$) for a general $\tau$ is

$$P_d = 2pQ\left(\frac{\tau}{\sigma_0}\right) - 2(1 - p)Q\left(\frac{\tau}{\alpha\sigma_0}\right) + (1 - p). \quad (9)$$

The optimal threshold $\tau^*$ that minimizes (9), the probability of making a wrong decision, is $\tau^* = \sigma_0\lambda^*$, in which

$$\lambda^* = \alpha\sqrt{\frac{2\ln(\alpha p/(1 - p))}{\alpha^2 - 1}}. \quad (10)$$

Based on (9) and (10), it is clear that $\sigma_0$, $\alpha$, and $p$ must be known or estimated.

IV. FINDING THE ABGN PARAMETERS

To find an estimate of the parameters ($\sigma_0^2, \alpha, p$), an exhaustive search algorithm has been set up to solve this problem (finding the optimal threshold). At this point, there was no attempt to make the algorithm faster or more efficient. This issue that is not crucial, is left for further investigation. To validate the procedure (before running the algorithm on measured data $\{\hat{z}_t\}$) we apply the code on synthetically generated data $\{\hat{z}_n\}$. The estimates $\hat{\alpha}$ and $\hat{\rho}$ so obtained were rather close to the chosen values used to generate the synthetic noise. This gives an indication that the developed theory as well as the numerical algorithm are a reliable tool for estimating the ABGN($\sigma_0^2, p, \alpha$) parameters for the model of noise measured in electric power grids.

The algorithm was used to obtain the PLC noise parameters by using a sequence of samples measured in electric power grids. A segment of such measured samples is exhibited in Fig. 2(a). Details of the measurement campaign can be found in [10]. Running the algorithm with the measured data lead us to the estimated values ($\hat{\alpha}, \hat{\rho}$) $\approx (2.7, 0.76)$. Also shown in this figure is a plot of synthetically generated ABGN, according to (4), with $\hat{\alpha} = 2.7$ and $\hat{\rho} = 0.76$. Corresponding histograms are also presented (just as illustrations). To assess the fitness of the ABGN model we next compare the measured data and the respective histograms, (c) and (d).

This limited analysis would lead us to the conclusion (not always true), that the ABGN($\sigma_0^2, p, \alpha$) would be a fair mathematical model for the noise over the PLC channel. We expect ABGN($\sigma_0^2, p, \alpha$) to be a fair noise model when evaluating the performance of systems that bears no memory (like for instance QAM modulation with symbol-by-symbol decisions). It is important to highlight that digital communication systems with block processing, in general, are impacted by the noise memory. We discuss next the noise memory effect and suggests a procedure to setup a model that takes in account the noise memory.

V. LDPC TO FIND PLC NOISE MEMORY

The memoryless ABGN model ($L = 1$) is too simplistic since it neglects the noise memory and it is not adequate thus for digital communication systems which use error control codes and are sensitive to the noise memory. This same sensitivity can help us to evaluate the parameter $L$ of the measured PLC noise, allowing to setup an ABGN model with memory which can be a useful tool when investigating the performance of any digital communication system.

The memory parameter can be estimated by fitting the BER performance curve of the system perturbed by the measured noise to those performance curves obtained by analytical methods or by simulating the data transmission perturbed by the synthetic noise generated by means of (7) and several values of $L$.

Figs. 3 and 4 illustrate the performance curves obtained by the digital communication system disturbed by an ABGNM model for estimate $\hat{\sigma}_0^2, \hat{\rho}, \hat{\alpha}$ (from the measured noise of Fig. 2), several values of $L$, and an irregular LDPC($n, k$) code with rate $k/n = 1/2$ ($k = 800)$ and $k/n = 1/2$ ($k = 2500$), respectively. From these plots we conclude that the digital communication system design can be conducted by modeling the power line noise as an ABGNM($\sigma_0^2, p, \alpha, L$) with
sensitive procedures. As a consequence, we could (partially
where the digital communications system does not use memory-
of the measured noise, contrary to the behavior of Fig. 1,
ABGN) does not result in a similar performance to that one
the performance of the digital communication system. In this
parameters obtained from measurements. We can see from the
experiment that the effect of the memory decreases as the
LDPC code handles longer blocks. Tuning the model with a
fitted for LDPC(800, 1600) or $L = 10$
for LDPC(2500, 5000), has to be taken into account.

VI. CONCLUSION

The use of a mathematical channel model is of paramount
importance when investigating new modulation schemes and
new error control codes for digital communication systems.
The AWGN($\sigma_0^2$) channel model, for instance, is a simple model
that has been extensively used to compare the performance of
many such modulation and coding schemes. Many models
to represent the noise impairing the transmission over PLC
systems have been proposed [2], [7] but to the best of our
knowledge, finding a noise model (the noise model parameters
consequently) that best mimic the measured data has not
received due attention. We have investigated the usefulness
of ABGNM as a model to study PLC system and came to
the conclusion that it can be used as the basis for the investiga-
tion of the PLC noise effect over digital communication.
A procedure to find the model parameters namely $\sigma_0^2$
the background noise power), $p$ (the probability that the noise is
in the background state), $\alpha$ (the impulsiveness factor) and $L$
(channel memory size) has been presented.

Although the question of which noise model to adopt has not
been entirely settled, we claim that the design of PLC systems
can be conducted by modeling the asynchronous component
of the PLC noise as a channel impaired by ABGNM($\sigma_0^2$, $p$, $\alpha$, $L$)
with the procedure proposed in this paper. This model yet
capturing the PLC noise behavior more accurately than the
plain, $L = 1$, ABGNM($\sigma_0^2$, $p$, $\alpha$), relies on some measurement
to accurately tune the model. A more sophisticated model,
in which the parameter $L$ is free to be a random variable is
currently under investigation.

REFERENCES

scheme for power line channels,” IEEE J. Sel. Areas Commun., vol. 24,
estimation based on pilots signal for OFDM modulation: A review,”
modeling power-line communication noise,” in Proc. XXXIII Simpósio
Brasileiro de Telecomunicações, Sep. 2015.
power line communication under Nakagami-like background noise,” in
Proc. Int. Conf. on Smart Grids, Green Commun. and IT Energy-aware
and outdoor electric distribution grids for PLC systems,” in Proc. PES