Modeling Noise as a Bernoulli-Gaussian Process

Weiler A. Finamore, Marcelo da S. Pinho, Manish Sharma, Moisés V. Ribeiro

***Abstract*—The transmission medium is always disturbed by noise with a random nature which can be characterized by taking a sequence of noise samples and, after analyzing the sequence, attributing a probabilistic model to represent the randomness of the noise. If thermal noise (receiver generated) is the only noise impairing the transmission (our focus is digital transmission only), the memoryless stationary discrete-time Gaussian process is the best model to represent the noise probabilistically. The mathematical representation of the transmission medium in such a situation yields the well-known Gaussian Channel. As Information Theory points out, for a fixed noise power, the Gaussian channel is the worst channel to send information through. If thermal noise is not the only noise impairing the transmission (as in sonar communication and power line communication), finding the probabilistic model other than the single-parameter Gaussian process that best matches the noise can much improve the communication system design. The Bernoulli-Gaussian process, a three parameters model, is a commonly considered option. Actually, this is a popular alternative to model the communication using power lines, which is modeled as Middleton Class-A oftenly. Finding the three parameters of the Bernoulli-Gaussian model (from known noise samples) is a formidable task that can be made simpler by considering the (original) results presented in the current paper. The Bernoulli-Gaussian model can be characterized, analytically, by using the noise power and two additional quantities: the expectation of the absolute value of the noise process plus the expected value of the third power of the absolute value. In practice, the parameters would be calculated using estimates of the mentioned expected values. In this work, it is also shown that the rate harvested when modeling the medium as a Bernoulli-Gaussian channel is increased when compared to modeling the medium with the easily obtained Gaussian channel.**

***Index Terms*—Bernoulli-Gaussian parameters, Noise modeling, Non-Gaussian stochastic process, Power Line Communication.**

I. Introduction

TRANSPORTING discrete-time information through a transmission medium from one point to another can be mathematically modeled in many ways. A common assumption is to mathematically describe the system as transmitting a sequence of samples $\{x_i\}$ (the index i is an integer usually identifying a time instant) through a transmission medium perturbed by additive noise. To each channel input x_i belonging to the set of real numbers $\mathbb{R}$, there corresponds thus a channel output $y_i = x_i + z_i$ such that a sequence of samples $\{y_i\}$, which is the addition of the signal value x_i to a noise component $z_i \in \mathbb{R}$, is delivered at the destination.

The noise is a random phenomenon. The simplest way to probabilistically model the noise[1] is to consider that the behavior of a sequence of noise samples is described by a Gaussian stochastic process, viz. a sequence of discrete-time, stochastic process $\{Z_i\}$ in which all Z_i are independent and identically distributed (i.i.d.), zero mean, Gaussian random variables (r.v.s), all with variance σ_Z^2. Thermal noise, generated at the receiver, is long known to be well-modeled as a Gaussian stochastic process. The power of the process, σ_Z^2, is well estimated by a value $\hat{\sigma}_0^2$, usually available[2]. Yet simple this model corresponds to the description of the hardest to cope with noise. Being $\hat{\sigma}_0^2$ the only history that can be taken into account to characterize the noise, adopting the Gaussian Channel (G-channel) to model the transmission medium would be the only choice but it would be the most conservative pick since, as it is well know [1], under equal conditions—same input and output alphabet and signal-to-noise ratio—the G-channel is, among all choices, the worst channel to transmit information through.

One can measure/sample noise and easily estimate the variance of the thermal noise, say $\hat{\sigma}_0^2$ but, if by any chance, the behavior of the noise under question does not match a Gaussian process (the noise is not due only to thermal causes but to man-made causes as well), the capacity of the medium would be underestimated.

In the current investigation, we tackle, however, the not so easy problem in which the sequence of samples $\{z_i\}$ is provided, and we seek to model the noise by a Bernoulli-Gaussian (BG) process or, explicitly, to model the noise by the stochastic process expressed as

$$\{Z_i\} \quad = \quad \sigma_0\{R_i^{[0]}\} + \sigma_1\{S_i R_i^{[1]}\}. \qquad (1)$$

The first component in this expression is a product of σ_0, a strictly positive real constant, and a sequence in which all $R_i^{[0]}$ are i.i.d. Gaussian r.v.s with zero mean and unity variance. We will usually refer to the noise which is well modeled as a Gaussian Process as Gaussian Noise. In the same vein noise which is well modeled by a BG-process, is said to behave as a BG process or simply to be a BG noise. The second component in (1) is considered to correspond to an intermittent stochastic process $\sigma_1\{S_i R_i^{[1]}\}$ in which S_i is a Bernoulli r.v.

W. A. Finamore (finamore@ieee.org, ORCID: 0000-0002-8963-7117), M. S. Pinho (mpinho@ieee.org, ORCID: 0000-0002-9393-8473), and M. Sharma (manish@ita.br, ORCID: 0000-0002-4792-2526) are with the Instituto Tecnológico de Aeronáutica, S. José dos Campos, SP, 12228-900, Brazil.

M. V. Ribeiro (mribeiro@engenharia.ufjf.br, ORCID: 0000-0002-6272-4675) is with the Universidade Federal de Juiz de Fora, Juiz de Fora, MG, 36036-900, Brazil.

This work was supported in part by Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES) under grant 001, Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq) under grants 404068/2020-0 and 314741/2020-8, Fundação de Amparo à Pesquisa do Estado de Minas Gerais (FAPEMIG) under grants APQ-03609-17, TEC-PPM 00787-18, and APQ-04623-22, Instituto Nacional de Energia Elétrica (INERGE), and Fundação de Amparo à Pesquisa do Estado de São Paulo (FAPESP) under grant 2022/14386-7.

Digital Object Identifier: 14209/jcis.2023.20

[1]Band-limited noise.

[2]Usually this is provided by the maker of the receiver.

such that for all integer i, $\mathsf{Prob}\{S_i = 1\} = p_1$, $0 < p_1 \leq 1$ and $\mathsf{Prob}\{S_i = 0\} = p_0 = 1 - p_1$. Furthermore, $R_i^{[1]}$ are i.i.d. Gaussian r.v.s with zero mean and unity variance. The quantity σ_1 is a real positive constant. Finding the model is, in summary, to take the available measured noise samples $\{z_i\}$ and obtain the parameters $\hat{\sigma}_0, \hat{\sigma}_1$, and $\hat{p}_1$ which best specify the process.

As highlighted in [2], there are many modern communication systems that deal with intermittent noise, such as power line communications (PLC) [3], wireless sensor networks [4], sonar communications [5], and vehicular communications [6]. In such cases, the transmitted signal is disturbed by noise which is better modeled by a non-Gaussian stochastic process. A mathematical model as the one in (1)—encompassing more sophistication than the one parameter Gaussian model—would be a good candidate (in the sense that the randomness of the physical phenomenon would be well captured by the theoretical model). An extra motivation for trying a more sophisticated model relies on the fact that non-Gaussian channels render channels less stringent than the more severe Gaussian channel (more severe in the sense that, for the same signal-to-noise power ratio, the Gaussian channel renders a smaller capacity).

Regarding communication systems with impulsive noise, we see that simplified mathematical noise models are based on BG-process [7], [8], Middleton Class-A [9]–[11], Gaussian mixtures [12], [13], and α-stable [14] random processes. Among them, Middleton Class-A and BG models can be considered the simplest because they depend on a reduced number of parameters. Even though the Middleton model contains a large number of states, the probabilities of these states follow a Poisson distribution, and thus its complexity is similar to that of the BG-process [9].

A model with a large number of parameters can yield, certainly, a probabilistic description better matched to the random behavior of the noise disturbing the transmission. The *Generalized Bernoulli-Gaussian* (gBG), a well-known stochastic process briefly discussed in Appendix A, is one such process. Finding a large number of parameters would be, however, a cumbersome task. In search for an easy to characterize model, we analyze, in this paper, the BG-process and present a closed formula solution for the determination of the parameters of the BG model when the expected values $\mathbb{E}[|Z|]$, $\mathbb{E}[Z^2]$, and $\mathbb{E}[|Z|^3]$ are known (in practice, a finite-length sequence of noise samples $\{z_i\}$, obtained from measurements, are available and one can only hope to find a good estimate of these expected values).

Of course, the higher the number of parameters, the larger the freedom to adjust the model to improve the probabilistic characterization of the noise. Yet, a simple model with a small number of parameters, but still acceptable (in the sense that it still reasonably represents the noise), is always more convenient. We conjecture that the current analysis opens up the way to study the gBG model, with a number of parameters larger than three, as well as other processes (processes that include memory, for instance).

Although the Middleton Class-A and the BG-process are both well suited to model[3] the noise in different applications with intermittent noise, the most important aspect of the modeling problem is to know how to find the parameters of the model from a sequence of measured noise samples. There are many methods that can be used to find the parameters of a process. In [15], [16] a numerical algorithm has been proposed to find the parameters of the Middleton Class-A. In the current work, an analytical solution to determine the three parameters of the BG model, when three moments of the stochastic process are known, is presented (for the first time to the authors knowledge). This is the main contribution of this paper. Yet simple the analytical solution is a very powerful tool for designing a communication system disturbed by non-Gaussian noise as we briefly discuss.

This paper presents in Section II the description of the Bernoulli-Gaussian (BG) stochastic process and introduces some theorems leading to the calculation of the parameters of the BG-process when three moments of the process are known. Proofs of theorems are placed in appendices. The discussion in Section III presents a fair comparison of the capacities of BG channels and Gaussian channels revealing the losses that are incurred when modeling noise with a BG behavior as Gaussian. Section IV is a short section describing the procedure to estimate parameters of a BG-process from a finite-length sequence of noise samples. Some results obtained by using computer generated samples of noise meant to behave as BG stochastic process are presented in Section V in order to illustrate the use of the procedure. Samples of noise over power lines obtained by measurements were also used to enhance the illustration. Section VI discusses the performance (obtained by computer simulation) of communication systems when data, encoded using Low-Density Parity Check (LDPC) code, is transmitted over the BG channel, which models the transmission medium disturbed by BG noise.

II. The Bernoulli-Gaussian Stochastic Process

The block diagram of a communication channel in which the disturbing noise is a BG-process is schematically shown in Fig. 1. Our assumption is that information samples x_i are fed to the channel input and the corresponding values observed at the output of the channel (available to the decoder) are the addition of signal plus noise $y_i = x_i + z_i$. We consider that the disturbing noise is expressed mathematically as a sequence $\{z_i\} = \sigma_0\{r_i^{[0]}\} + s_i\sigma_1\{r_i^{[1]}\}$ given by the sum of two noise components with $r_i^{[0]}$ and $r_i^{[1]}$ corresponding to variables with values belonging to the set $\mathbb{R}$ of real numbers. A particular noise sample z_i can be viewed as the manifestation of a perturbation that is in either the state $s_i = 0$ or state $s_i = 1$ (the impulsive state that occurs with probability p_1). In the first case $z_i = \sigma_0 r_i^{[0]}$ is the ubiquitous thermal noise modeled as a Gaussian r.v. $Z_i = \sigma_0 R_i^{[0]}$ in which $R_i^{[0]}$ are zero mean, Gaussian r.v.s, all with unity variance, and, when the noise is in state $s_i = 1$, the sample values are expressed as $z_i = \sigma_0 r_i^{[0]} + \sigma_1 r_i^{[1]}$, the addition of the always present background noise plus an intermittent component $\sigma_1 r_i^{[1]}$.

[3]It should be noticed that only extensive noise measurement can allow to figure out which mathematical model better represent the true noise.

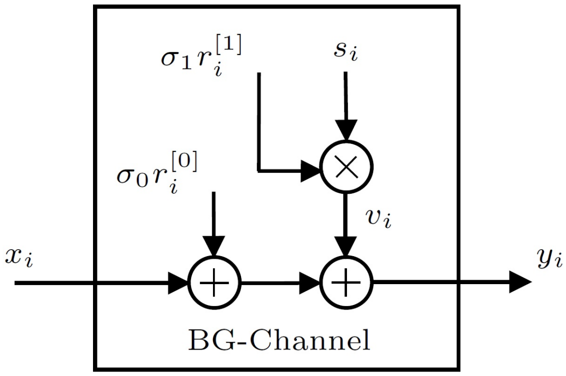


Fig. 1: Block diagram of a Bernoulli-Gaussian communication channel. The value $y_i = x_i + z_i$ observed at the output of the channel, is either x_i added to $z_i = \sigma_0 r_i^{[0]}$ or, to $z_i = \sigma_0 r_i^{[0]} + \sigma_1 r_i^{[1]}$.

We now, to make the mathematical manipulation easier, introduce some auxiliary r.v.s. We let $W_i = \sigma_0 R_i^{[0]} + \sigma_1 R_i^{[1]}$. We introduce, furthermore, the non-Gaussian i.i.d. r.v.s $V_i = S_i R_i^{[1]}$ (all S_i have the same probability mass function). The r.v.s $\sigma_0 R_i^{[0]} + \sigma_1 V_i$ are non-Gaussian. The i.i.d. assumption allows us to write $\mathsf{f}_{Z_i}(\zeta) = \mathsf{f}_Z(\zeta)$ in which

$$f_Z(\xi) = \frac{1-p_1}{\sigma_0\sqrt{2\pi}}\, e^{-\frac{\xi^2}{2\sigma_0^2}} + \frac{p_1}{\sigma_W\sqrt{2\pi}}\, e^{-\frac{\xi^2}{2\sigma_W^2}} \tag{2}$$

is not dependent on the index i. Since all pairs of distinct r.v.s are pairs of independent r.v.s we have that the sequences $\sigma_0\{R_i^{[0]}\}$ and $\sigma_1\{S_i R_i^{[1]}\}$ have power σ_0^2 and $p_1\sigma_1^2$, respectively. As for the sequence $\{Z_i\}$ its power $\sigma_Z^2 = \mathbb{E}\left[Z^2\right]$ can be written, it can be demonstrated, as

$$\sigma_Z^2 \;=\; \sigma_0^2 + p_1\sigma_1^2 \tag{3}$$

or yet, more conveniently, in terms of $\sigma_W^2 = \sigma_0^2 + \sigma_1^2$, as

$$\sigma_Z^2 \;=\; (1-p_1)\sigma_0^2 + p_1(\sigma_0^2 + \sigma_1^2). \tag{4}$$

Observe that $V_i = S_i R_i^{[1]}$ are i.i.d. r.v.s all having the same variance $p_1\sigma_1^2$ and the same probability density function (p.d.f.) $\mathsf{f}_{V_i}(\zeta) = \mathsf{f}_V(\zeta)$. The three parameters $(\sigma_0, \sigma_1, p_1)$ completely specify the BG-process $\{Z_i\}$. For convenience, we introduce the ratios

$$\alpha_{1Z}^2 \;=\; \sigma_1^2/\sigma_Z^2, \text{ and} \tag{5}$$

$$\alpha_{10}^2 \;=\; \sigma_1^2/\sigma_0^2. \tag{6}$$

Another quantity of interest, which we will call the "*process impulsivity*" is defined as

$$\Lambda \triangleq p_1\sigma_1^2/\sigma_0^2 \tag{7}$$

Notice that if $\sigma_1^2 = 0$, which is equivalent to saying that $p_1 = 0$, the impulsivity is zero (and the process is a purely Gaussian process).

The main issue when choosing a model to represent noise with known samples $\{z_i\}$ is to find the parameters of that model. The models used in this work are those with a description as in (1) and the parameter finding procedure is based on three absolute moments (which in practice can only be estimated). As will be shown, the parameters can be computed by analytically solving a system of non-linear equations. We next present some results which related the absolute moments of the BG-process to the parameters and are useful when solving the system of non-linear equations.

A. Some useful theorems

Let us consider that the process $\{Z_i\} = \sigma_0\{R_i^{[0]}\} + \sigma_1\{S_i R_i^{[1]}\}$ is stationary and that the expected values $\mathbb{E}[|Z_i|]$, $\mathbb{E}[|Z_i|^2]$, and $\mathbb{E}[|Z_i|^3]$ which, since they are not dependent on the index i, are all equal to $\mathbb{E}[|Z|]$, $\mathbb{E}[|Z|^2]$, and $\mathbb{E}[|Z|^3]$ respectively. We can write thus [17]

$$\mathbb{E}[|Z|] \;=\; \underbrace{\left(p_0 + p_1\sqrt{\alpha_{10}^2+1}\right)\sigma_0}_{\mu_1}\sqrt{\frac{2}{\pi}}, \tag{8}$$

$$\mathbb{E}[|Z|^2] \;=\; \underbrace{\left(p_0 + p_1\left(\alpha_{10}^2+1\right)\right)\sigma_0^2}_{\mu_2}, \tag{9}$$

$$\mathbb{E}[|Z|^3] \;=\; \underbrace{\left(p_0 + p_1\left(\sqrt{\alpha_{10}^2+1}\right)^3\right)\sigma_0^3}_{\mu_3}\sqrt{\frac{8}{\pi}}. \tag{10}$$

For convenience, we introduce the notations

$$\mu_1 \;=\; \mathbb{E}[|Z|]\sqrt{\frac{\pi}{2}}, \tag{11}$$

$$\mu_3 \;=\; \mathbb{E}[|Z|^3]\sqrt{\frac{\pi}{8}}. \tag{12}$$

Next theorems show important relationships among these moments.

Theorem 1 (Feasible region)**.** *The set of values $\mathbb{E}[|Z|]$, $\mathbb{E}[|Z|^2]$, and $\mathbb{E}[|Z|^3]$ of a BG-process are restricted to a region of $\mathbb{R}^3$ which we call feasible region. The feasible values of $\mathbb{E}[|Z|]$, $\mathbb{E}[|Z|^2]$ and $\mathbb{E}[|Z|^3]$ are such that the following inequalities hold:*

$$\mathbb{E}[|Z|] \;\leq\; \sqrt{\frac{2}{\pi}\mathbb{E}[Z^2]}, \tag{13}$$

$$\mathbb{E}[|Z|^3] \;\geq\; 2\mathbb{E}[|Z|]\mathbb{E}[Z^2], \tag{14}$$

$$\mathbb{E}[|Z|^3] \;\geq\; \frac{4}{\pi}\frac{\left(\mathbb{E}[Z^2]\right)^2}{\mathbb{E}[|Z|]}, \tag{15}$$

(relations will be equalities if, and only if, the r.v. Z is Gaussian). □

Proof: The proof is trivial. (See Appendix B) ■

For BG-processes with $p_1 \in (0,1)$ (i.e., for non-gaussian processes), define, the parameter

$$\kappa \triangleq \frac{\left(\mu_3 - \mu_1\mu_2\right)^2}{\left(\mu_1\mu_3 - \mu_2^2\right)\left(\mu_2 - \mu_1^2\right)}. \tag{16}$$

Next theorem deals with the just defined parameter κ.

Theorem 2 (Range of κ)**.** *Let $\sigma_0 > 0$ and $p_1 \in (0,1)$. Let $Z = \sigma_0 R^{[0]} + \sigma_1 S R^{[1]}$ denote a r.v. with associated parameter κ, defined as in (16). We have then the inequality $\kappa > 4$.* □

Proof: The proof is trivial. (See Appendix B) ■

Parameter κ, for reasons to become clear, is related to the impulsivity Λ of the BG-process. We can see that while the process impulsivity is near zero (i.e., $\Lambda \approx 0$), and if σ_0^2 is near $\mathbb{E}|Z^2|$, then process impulsivity $\Lambda = \frac{\mathbb{E}[|Z^2|]}{\sigma_0^2} - 1$ grows indefinitely as σ_0^2 goes to zero.

Results presented next are central to the calculation of the main parameters of the process.

B. Finding the parameters $(\sigma_0, \sigma_1, p_1)$ of a BG-process from expected values.

Next theorem is of paramount importance when calculating the parameters of a BG-process.

Theorem 3 (Parameters from moments)**.** *Recall that $R_i^{[0]}$ and $R_i^{[1]}$ are two independent, zero mean Gaussian r.v.s with, unity variance, and S_i is a binary r.v. Let $p_0 = \mathsf{Prob}\{S_i = 0\}$ and consider, furthermore, that any pair formed with these r.v.s is a pair of independent r.v.s. Consider the Bernoulli-Gaussian stochastic process $\{Z_i\}$ with parameters $(\sigma_0, \sigma_1, p_1)$ which can be written as $\{Z_i\} = \sigma_0\{R_i^{[0]}\} + \sigma_1\{S_i R_i^{[1]}\}$. Consider also that*

$$\xi = \frac{1}{2}\left(\kappa - 2 + \sqrt{\kappa(\kappa - 2)}\right) \tag{17}$$

is a root of the equation $\widetilde{\alpha}^2 - (\kappa - 2)\widetilde{\alpha} + 1 = 0$. If $\mathbb{E}[|Z_i|]$, $\mathbb{E}[|Z_i|^2]$, and $\mathbb{E}[|Z_i|^3]$, the first three moments of the r.v.s $|Z_i|$, are known then the three parameters can be computed by

$$\sigma_0 = \frac{\mu_3\mu_2 - \mu_1\mu_2^2}{(\mu_2^2 - \mu_1^2)(\xi + 1)} \tag{18}$$

$$\sigma_1 = \sigma_0\sqrt{\xi^2 - 1}, \text{ and} \tag{19}$$

$$p_1 = \frac{\sigma_Z^2 - \sigma_0^2}{\sigma_1^2}. \tag{20}$$

□

The proof of Theorem 3 is presented in Appendix B.

III. Modeling communication over noisy transmission media

The analysis and understanding of communication systems usually starts with measurement in order to obtain samples of the disturbing noise and, from so obtained samples $\{z_i\}$, the specification of a suitable mathematical model to probabilistically model this noise is sought. Estimating the moments of the model from the measured samples is the first step. Upon having the estimated moments the parameters of the model can be calculated. Each set of parameters specifies a given channel that can have its capacity obtained.

We proceed by plotting and examining the capacity of channels with binary input whose output are the samples of the signal added to samples of the noise. The graphics in Fig. 2 display the capacities of several channels plotted against the signal-to-noise ratio $\text{SNR} \triangleq A^2/\sigma_Z^2$ (ratio of the signal power when the input signal is restricted to binary samples belonging to the set $\{-A, A\}$, voltage unities, to the variance of process $\{Z_i\}$ chosen to model the noise). The lower curve displays the capacity of the biG-channel (Gaussian-channel with input restricted to binary values) given by [18]

$$C_{\text{biG}} = g_0\left(\frac{A}{\sigma_Z}\right), \tag{21}$$

in which

$$g_0(v) = \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\infty} e^{-\frac{(\gamma - v)^2}{2}} \log_2 \frac{2}{1 + e^{-2\gamma v}}\, d\gamma. \tag{22}$$

Also plotted in terms of SNR, with assorted values of parameters (α_{10}, p_1), are the capacities C_{biBG} of BG-channels with binary input (middle curves). The analysis presented in [18] to the biG-channel can be extended to the biBG-channel, and its capacity can be expressed as

$$C_{\text{biBG}} = g_1\left(\frac{A}{\sigma_Z}, p_1, \alpha_{10}\right) \tag{23}$$

in which

$$g_1(v, p_1, \alpha_{10}) = \\ = \int_{-\infty}^{\infty} g_2(v, \gamma, p_1, \alpha_{10}) \log_2 \frac{2}{1 + \frac{g_2(-v,\gamma,p_1,\alpha_{10})}{g_2(v,\gamma,p_1,\alpha_{10})}} d\gamma \tag{24}$$

and, besides,

$$g_2(v, \gamma, p_1, \alpha_{10}) = \frac{1 - p_1}{\sqrt{2\pi}}\, e^{-\frac{1}{2}\left(\gamma - v\sqrt{1+\alpha_{10}^2}\right)^2} + \\ + \frac{p_1}{\sqrt{2\pi(1 + \alpha_{10}^2)}}\, e^{-\frac{1}{2(1+\alpha_{10}^2)}\left(\gamma - v\sqrt{1+\alpha_{10}^2}\right)^2}. \tag{25}$$

It is also interesting to notice that the parameter κ (which is a function of the values $p_1 = 1 - p_0$ and of $\alpha_{10} = \sigma_1/\sigma_0$) determines, in fact, the impulsivity of the process — and the larger its value is, the larger is the discrepancy between the capacity of the biBG-channel and that of the biG-channel.

Another interesting aspect (not mentioned in similar published results) is that at moderate and small values of SNR, the "non-harvested rate" when modeling the noisy communication system with a purely Gaussian channel is much more noticeable than at larger SNR. At large values of SNR (above 10 dB), a region in which the Gaussian component with largest variance is predominant, corresponding to the impulsive noise, there is not much harvested rate loss.

Finding the parameter of a Gaussian process chosen to model any given noise is straightforward— the usual, simple and efficient approach is to sample the noise and use the arithmetic average of the squared sample values as the noise variance estimate. As for the parameters of a non-Gaussian process the calculation, at first sight, might not be so easy [21]. The main contribution of this work is to present a simple technique which allows to easily find the parameters of the BG-process that best[4] fit the noise, given a set of noise samples.

[4]It should be noticed that the noise might not have a random behavior that is well described by the selected BG stochastic process. The expression "non-harvested rate" is used to reveal that the designer throws away available rate by choosing a channel model that do not fully match the random behavior of the noise.

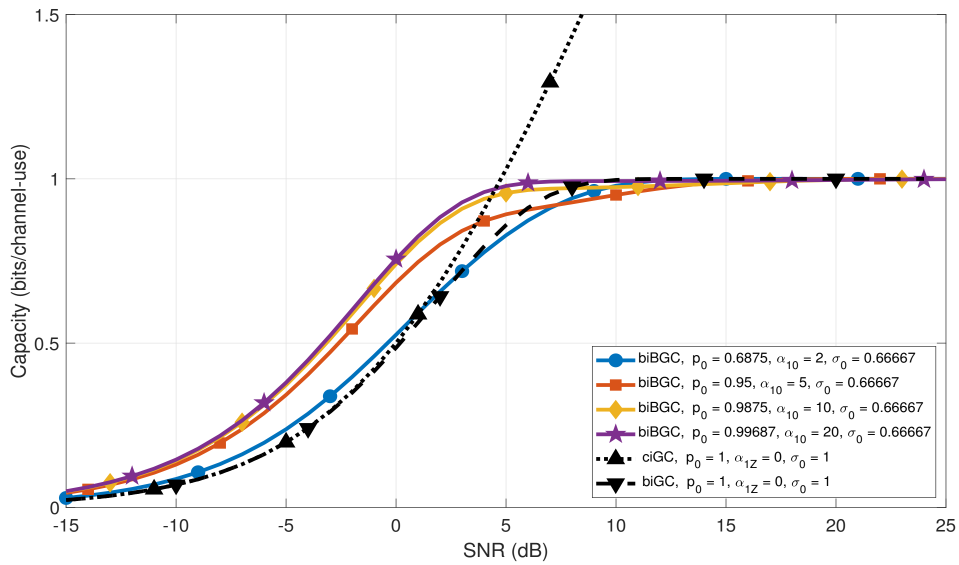


Fig. 2: Graphs illustrating the capacity C_{biG} (dashed line) of the Gaussian-channel with input restricted to binary values and the capacity C_{biBG} of several Bernoulli-Gaussian channels also with input restricted to a binary alphabet. The capacity curve C_{ciG} of a continuous input Gaussian-channel is also plotted (dashed line). The variance of the process, $\mathbb{E}[Z^2]$, as well as the background noise power, σ_0^2 are held constant (fixing the impulsivity to $\Lambda = p_1\alpha_{10}^2 = 1.25$) in all cases (except for the Gaussian process, for which $\Lambda = 0$ and the variance is $\mathbb{E}[Z^2] = \sigma_0^2$.)

IV. Simple Procedure to Obtain the Parameters of a Bernoulli-Gaussian Process from Noise Samples

When a sequence $\{z_i\}_{i=1}^N$ of values sampled from a "truly BG-noise" is available, a simple procedure — based on Theorem 3 — to find the parameters of the BG model that matches the sampled noise can be obtained by just calculating, for $\ell = \{1, 2, 3\}$, the quantities

$$M_\ell = \frac{1}{N}\sum_{i=1}^{N} |z_i|^\ell. \tag{26}$$

Quantities M_1, M_2, and M_3 are *estimates* of $\mathbb{E}[|Z|]$, $\mathbb{E}[|Z|^2]$, and $\mathbb{E}[|Z|^3]$ (the first, the second, and the third moments, respectively, of the r.v. $|Z|$.) With the values in (26) in hand and Theorem 3, the BG model parameters can be obtained (This procedure is based on the assumption that the noise is truly BG. If this is not the case, a mismatch will become evident).

To make the calculation more manageable we choose to work with normalized values. In other words, we use the notation $\overline{\sigma}_0 = \sigma_0/\sigma_Z$ and $\overline{\sigma}_1 = \sigma_1/\sigma_Z$ and introduce the normalized variable $\overline{Z}_i = Z_i/\sigma_Z$ to construct the normalized stochastic process

$$\{\overline{Z}_i\} = \overline{\sigma}_0\{R_i^{[0]}\} + \overline{\sigma}_1\{S_i R_i^{[1]}\} \in \mathbb{R}. \tag{27}$$

This modification enforces the condition $\mathbb{E}[\overline{Z}_i^2] = 1$ and introduces no loss of generality since, once $(\overline{\sigma}_0, \overline{\sigma}_1, p_1)$ is obtained, the triplet $(\sigma_0, \sigma_1, p_1)$ is easily obtainable. It is worth noticing that $\overline{\sigma}_0$ is strictly positive, i.e. $\overline{\sigma}_0 > 0$. Also, from the ratio defined in (6), we can state that $\alpha_{10} = \overline{\sigma}_1/\overline{\sigma}_0$.

With some manipulation, we rewrite equations (8), (9), and (10), and get

$$\overline{\mu}_1 = \mathbb{E}\left[|\overline{Z}|\right]\sqrt{\frac{\pi}{2}}, \tag{28}$$

$$\overline{\mu}_2 = 1, \tag{29}$$

$$\overline{\mu}_3 = \mathbb{E}\left[|\overline{Z}|^3\right]\sqrt{\frac{\pi}{8}}. \tag{30}$$

With the introduced changes, Theorem 1 is restated next showing how to get the parameters $(\overline{\sigma}_0, \overline{\sigma}_1, p_1)$ by using only the two expected values in (28) and (30).

Theorem 4 (Normalized parameters from moments)**.** *Let $\{\overline{Z}_i\} = \{\overline{\sigma}_0 R_i^{[0]} + \overline{\sigma}_1 S_i R_i^{[1]}\}$ be a normalized (unity power) Bernoulli-Gaussian process characterized by the parameters $(\overline{\sigma}_0, \overline{\sigma}_1, p_1)$ such that $p_0 = \mathrm{Prob}\{S_i = 0\}$ and $S_i R_i^{[1]}$ is the product of independent r.v.s S_i and $R_i^{[1]}$.*

If the two moments $\mathbb{E}[|\overline{Z}_i|]$, and $\mathbb{E}[|\overline{Z}_i|^3]$ of the r.v.s $|\overline{Z}_i|$, are known then the parameter $\overline{\sigma}_0$ can be obtained by first finding the roots of the equation

$$\widetilde{\alpha}^2 + (2-\kappa)\widetilde{\alpha} + 1 = 0, \tag{31}$$

in which

$$\kappa = \frac{(\overline{\mu}_3 - \overline{\mu}_1)^2}{(\overline{\mu}_1\overline{\mu}_3 - 1)\left(1 - \overline{\mu}_1^2\right)}. \tag{32}$$

And then, if ξ is a non-negative real root of (31), we have

$$\overline{\sigma}_0 = \frac{\overline{\mu}_3 - \overline{\mu}_1}{\left(1 - \overline{\mu}_1^2\right)(\xi + 1)}, \tag{33}$$

$$\overline{\sigma}_1 = \overline{\sigma}_0\sqrt{\xi^2 - 1}, \tag{34}$$

and

$$p_1 = \frac{1 - \overline{\sigma}_0^2}{\overline{\sigma}_1^2}. \tag{35}$$

$\square$

The parameters $(\sigma_0, \sigma_1, p_1)$ of the stochastic process $\{Z_i\}$ are finally obtained by taking $\sigma_0 = \overline{\sigma}_0\, \sigma_Z$ and $\sigma_1 = \overline{\sigma}_1\, \sigma_Z$.

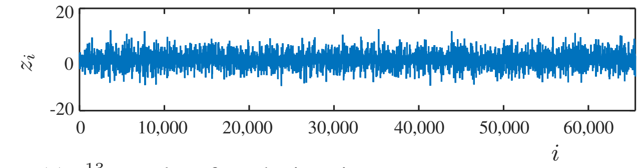
V. Comparing Achievability and Commenting some Experimental Results

Plots in Fig. 2 exhibit the code rate limits for several values of the parameters $(\sigma_0, \sigma_1, p_1)$. As it is well known [1] if symbols x generated by a source (modeled by a r.v. X) are sent to a destination (by using a rate $R_c = k/n$ ideal code to transmit through a channel) the corresponding symbols $\hat{x}$ (modeled by a r.v. $\hat{X}$) delivered to the destination are subject to negligible error probability (Prob$\{\hat{X} \neq X\}$) if $C_{\mathrm{biBG}}(\mathrm{SNR}) > R_c$.

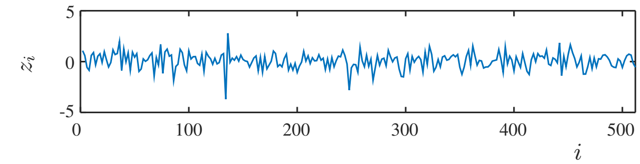
To transmit with negligible probability of error, at a rate of $R_c = 1/2$ bits-per-channel-use, for instance, when the transmission medium noise is modeled as a BG-process with parameters $p_1 = 0.05$ and $\alpha_{10} = \sigma_1/\sigma_0 = 10$, the SNR is required to be larger than about -3.0 dB. In contrast, if the transmission medium is modeled as a biG-channel (with $\sigma_Z^2 = \sigma_0^2 = 1$) the SNR requirement surpasses 0 dB. Another way to evaluate the loss of modeling the BG-process by a Gaussian stochastic process is by using the Kullback-Leibler divergence [1]. Computing numerically the divergence between a Gaussian r.v. with unitary variance for the BG-process with the given parameters, a value of 0.3 bits has been obtained.

A good communication system design benefits from the choice of efficient error control codes and the implementation of the corresponding receiver rewards from the selection of a representative channel to model the transmission medium. To assess the benefit we have generated samples belonging to a BG-process (using MATLAB) and for the sequence of samples so generated the corresponding parameters were obtained. Many synthetic noise sequences (pseudo-random generated noise samples), were examined. Synthetic noise generated with the parameters $\sigma_0 = 0.8165$, $\sigma_1 = 4.0825$, and $p_1 = 0.05$ ($\alpha_{10} = 5$), is illustrated by the plots in Fig. 3. The expected values of the BG-process having such parameters—calculated with (8), (9), and (10)—are respectively $\mathbb{E}[|Z_i|] = 0.6420$, $\mathbb{E}[|Z_i|^2] = 1$, and $\mathbb{E}[|Z_i|^3] = 3.4475$. The estimated moments, empirically calculated, are respectively $M_1 = 0.6427$, $M_2 = 0.9999$, and $M_3 = 3.5357$. The parameter estimation procedure used the estimated moments (M_1, M_2, and M_3) and produced the estimated values $\hat{\sigma}_0 = 0.8162$, $\hat{\sigma}_1 = 4.0811$, and $\hat{p}_1 = 0.0500$, ($\hat{\alpha}_{10} = 4.9981$), — which are satisfactorily close to the true parameters.

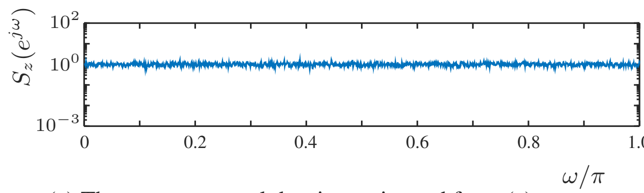
Finding the model is one question, how well does the BG-process model the random phenomenon is a separate question. Not surprisingly, when opting to model noise with an impulsive nature as a BG stochastic process rather than Gaussian, one incurs in a smaller mismatch between the mathematical results and empirical results (since the simple biG-channel is a particular instance of the more sophisticated biBG-channel, the parameter estimation procedure would deliver a Gaussian model if the estimated moments are compatible with a Gaussian model).



(a) 2^{13} samples of synthetic noise



(b) The first 2^8 samples of (a)



(c) The power spectral density, estimated from (a)

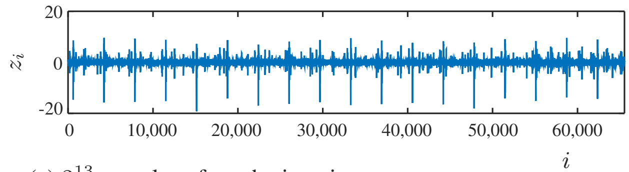
Fig. 3: Plots exhibiting segments of a pseudo random sequence of a BG-process $\{Z_i\}$ with $p_1 = 0.05$, $\sigma_0^2 = 0.6667$, and $\sigma_1^2 = 16.6667$.

The capacity versus SNR curves, pictured in Fig. 2, show that at around SNR $= 0$ dB, the disregarded capacity runs in the order of 0.25 bits-per-channel-use (0.5 instead of the 0.75 available). In other words, all theoretical analysis based on the choice of the biG-channel to model the transmission medium instead of the biBG-channel, would be, at the onset, neglecting the available channel capacity.

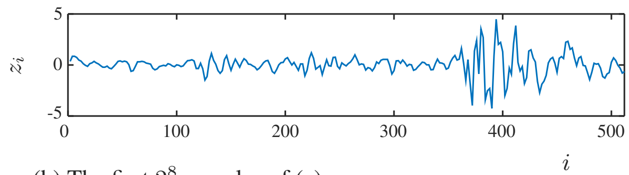
Many sequences of noise (BG and others) were analyzed and the corresponding BG parameters were estimated. As an illustration, the parameters of a $65,536$ samples long sequence of measured noise are presented in Fig. 4.

The measured noise estimated moments (normalized to enforce unitary power) namely $M_1 = 0.5932$, $M_2 = 1.0000$, and $M_3 = 4.1865$, yielded estimated parameters $\hat{\sigma}_0 = 0.5878$, and $\hat{\sigma}_1 = 3.5677$, and $\hat{p}_1 = 0.0514$, (i.e., $\hat{\alpha}_{10} = 6.0700$). This noise, as can be perceived from the Power Spectral Density (PSD) (see Fig. 4(c)), is far from a BG noise. One can surely infer that the transmission is clearly also impaired by interference[5] and not disturbed by only noise. In spite of this fact, the procedure manage to find the parameters and inform that if one's choice is to model this "strange" noise as a BG random process $\{Z_i\} = \{\hat{\sigma}_0 S_i + \hat{\sigma}_1 S_i Z_{1,i}\}$, the parameters to be used are given by the triple values $(\hat{\sigma}_0, \hat{\sigma}_1, \hat{p}_1)$ just found. From another perspective: the procedure informs that the best BG-process (built from the estimated moments) that models this "strange" noise is characterized by the obtained parameters. Even though the estimator indicates that the best model for the measured noise is not the Gaussian one, since

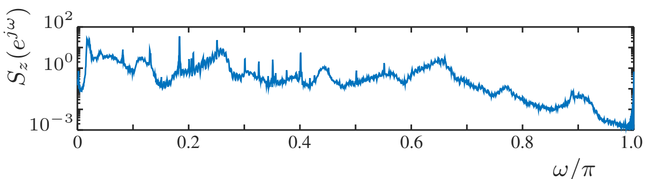
[5]The disturbing effect of interference and other impairments can, of course, be mitigated by using specific schemes. The current investigation is not aimed at combating these impairments and, for this reason, all that is not signal were, in the described experiment, treated as plain noise.



(a) 2^{13} samples of synthetic noise



(b) The first 2^8 samples of (a)



(c) The power spectral density, estimated from (a)

Fig. 4: Graphs of noise measured over power line.

the parameters obtained indicate a value of α_{10} significantly greater than zero, it is interesting to evaluate the gain of using the BG model. The log-likelihood ratio is a tool that allows a simple evaluation of this gain. Using the measured noise, $\{z_i\}_{i=1}^N$, illustrated in Fig. 4, and the p.d.f.s: f_{BG} (for BG i.i.d. r.v.s), and f_{G} (for Gaussian model), the results show that

$$\frac{1}{N}\log_2\frac{f_{\text{BG}}(z_1,\ldots,z_N)}{f_{\text{G}}(z_1,\ldots,z_N)} = 0.3925 \tag{36}$$

It is important to mention that for a long sequence of BG r.v.s, $\{Z_i\}_{i=1}^N$, the Asymptotic Equipartition Property (AEP) states that the log-likelihood ratio is approximately equal to the Kullback-Leibler divergence, 0.4685 bit (numerically computed). Therefore, the result presented in (36) shows a significant gain of the model.

In general, tests run with BG pseudo-random noise having large values of α_{10} yielded good parameters calculation. If pseudo-random samples are taken from a BG-process with very low values of α_{10} ($\alpha_{10} < 0.1$) the estimated moments fall outside the feasible region and the results are not quite as accurate as those obtained when α_{10} is large. But, it is important to emphasize, the benefit when using the BG-model over the G-model is larger when α_{10} has large values. For low values of α_{10} the gains are meager (and modeling the samples as BG or G would not lead to a large mismatch).

VI. Assessing the Utility of the Model

To assess the utility of the proposed model, the performance of communication systems in three distinct scenarios has been investigated and compared.

The performance of a system disturbed by Gaussian noise (computer generated pseudo-Gaussian noise) designed by considering that the noise that disturbs the transmission is modeled as a Gaussian channel was analyzed for reference purposes (case 1). This means, in other words, that the channel is Gaussian and the receiver is matched to Gaussian noise (the model is good and the design is matched). In the second approach, the situation is such that the channel is, in fact, a BG channel (in the sense that the synthetically generated noise is a sequence of pseudo-random BG r.v.s) but since the receiver does not have access to enough knowledge (knows only the power of the noise) the channel is modeled as Gaussian. In other words, in case 2, the channel is BG and the receiver is matched to Gaussian noise (the model is bad and the design is mismatched). In the third approach (i.e., case 3), the noise behaves as BG (a sequence of pseudo-random BG r.v.s) and the receiver knows the parameters α_{10} and p_1 to perform its task. In other words, the channel is BG and the receiver is matched to BG noise (the model is good and the design is matched). In all cases, $X \in \{\pm 1\}$ and $\mathsf{Prob}\{X = -1\} = \mathsf{Prob}\{X = +1\} = 1/2$. Bit value 0 is associated with the symbol value $A = +1$. The SNR is changed by varying σ_Z^2, the noise variance.

The impact of using the correct channel model at the receiver is analyzed by using the same channel code in all cases. A length $n = 4096$, rate 1/2 LDPC code was designed using EXIT (extrinsic information transfer) charts considering transmission over a biG channel. The Progressive Edge Growth Algorithm was used to design the encoder. The factor-graph based LDPC decoder uses soft *a priori* messages from the channel to perform decoding. The three cases contemplated build the receiver (the LDPC decoding algorithm) based on the log-likelihood function metric, named log likelihood rate (LLR), and given by

$$\text{LLR}(y) = \log\left(\frac{\mathsf{Prob}\{X = 1 \mid Y = y\}}{\mathsf{Prob}\{X = -1 \mid Y = y\}}\right) \tag{37}$$

In the case #1, z is modeled by a Gaussian r.v., while in the cases #2 and #3 it is modeled by the BG distribution given in (2).

Under the assumption that the noise is Gaussian with zero mean and variance σ_Z^2, the corresponding LLR value [22] is

$$\text{LLR}_G(y) = \frac{2y}{\sigma_Z^2}. \tag{38}$$

The second LLR is expressed by taking into consideration the p.d.f. of the BG-process r.v.s, leading to the LLR soft messages with values that depend on channel output y. Using the notation $p_X(\pm 1 \mid y) = \mathsf{Prob}\{X = \pm 1 \mid Y = y\}$ we have

$$\begin{aligned}\text{LLR}(y) &= \log\left(\frac{p_X(1 \mid y)}{p_X(-1 \mid y)}\right)\\ &= \log\left(\frac{\mathsf{f}_{Y|X=1}(y)}{\mathsf{f}_{Y|X=-1}(y)}\right),\end{aligned} \tag{39}$$

where $\mathsf{f}_{Y|X=\pm 1}(y)$ represents the p.d.f. of the output Y, calculated at point y, given that the input is equal to 1 or to -1.

When the channel is BG, and the receiver has knowledge of α_{10}, and p_1, it may use the probabilistic relationship between transmitted symbol x and received value y to generate the appropriate LLR value as:

$$\mathrm{LLR}_{BG}(y) = \log \frac{(1-p_1)\, e^{\frac{-(y-1)^2}{2\sigma_0^2}} + p_1\, e^{\frac{-(y-1)^2}{2(1+\alpha_{10}^2)\sigma_0^2}}}{(1-p_1)\, e^{\frac{-(y+1)^2}{2\sigma_0^2}} + p_1\, e^{\frac{-(y+1)^2}{2(1+\alpha_{10}^2)\sigma_0^2}}}. \quad (40)$$

The decoder in cases #1 and #2 would employ (38), while a decoder in case #3 would employ (40). The difference in performance between cases #2 and #3 would be due to knowledge of the BG model. The effect of this difference may be observed in Fig. 5 which exhibits the performance of a few communication systems perturbed by BG noise. The system which incorporates a decoder designed by considering that the noise is modeled as a BG-process renders a better performance (upper-triangles curve). The curve marked with upper-triangles shows that a probability of error of 10^{-5} can be achieved at a $\mathrm{SNR} \approx -1.2$ dB indicating a gain of about 4.5 dB over the system designed by considering that the noise is purely Gaussian. The performance (marked by circles) of the system which wrongfully considers the noise to be Gaussian, and therefore uses a Gaussian metric, is penalized even when compared to the performance (marked by squares, in the same figure) of a system that transmits over a Gaussian channel (and uses a decoder designed under a Gaussian channel assumption). Several combinations of parameters values have been examined and the mismatch loss observed is large if the process impulsivity, $p_1\alpha_{10}^2$, is high. For low values of impulsivity, the observed mismatch loss is small.

For illustration purpose, the LLRs for three scenarios are displayed in Fig. 6. The larger slope line is the LLR meant to be used by a decoder which faces a noise that is modeled as Gaussian having variance σ_1^2. The smaller slope line, on the other hand, is the LLR for a decoder facing noise modeled as Gaussian with variance σ_0^2.

The sinusoidal shaped line, on is turn corresponds to the LLR used by a decoder designed to cope with a Bernoulli Gaussian noise with parameters $(\sigma_0, \sigma_1, p_1) = (0.6667, 3.3333, 0.05)$. The curious (and desirable) behavior of the decoder is that it naturally switches the weighing of the decoder income information (the log-likelihood which is calculated with the value y observed at the channel output). In one hand, if a small likelihood value is obtained, the decoder operates as if under good channel conditions. On the other hand, if a large likelihood value is observed, the decoder operates as if under an impulsive channel condition. In between the two situations, it is noticed a soft transition between good and impulsive states.

VII. Conclusion

The increased flexibility brought when modeling the transmission medium by a probabilistic model with a large number of parameters leads to a model that can better reflect the behavior of the true noise—the observed medium noise. The two states BG-channel with three parameters $(\sigma_0, \sigma_1, p_1)$ is a model more flexible than the purely Gaussian channel with a single parameter, say σ_0'. Using a number of parameters higher than three can render, of course, a mathematical model even better adjusted to the real phenomenon—but at the expense of higher modeling complexity and, we conjecture, with most the harvested with the BG model.

This paper main contribution is a closed form calculation which helps to obtain the three parameters $(\sigma_0, \sigma_1, p_1)$ of a Bernoulli-Gaussian stationary stochastic process $\{Z_i\}$, when the three absolute moments $E[|Z_i|]$, $E[|Z_i|^2]$ and $E[|Z_i|^3]$ of the process are known. In practice, when trying to find the three parameters of the BG-channel modeling the true noise (the noise that actually disturbs the transmission) only an estimation of the three moments is available. The disparity, due to the fact that the estimates are not exact, is not of paramount concern. The most important disparity might come from the fact that the noise behavior is not BG (this mismatch can only be spelled by practical measurements).

Modeling the effect of noise disturbing the transmission with a BG-channel brings to light that extra rate can be harvested when compared to using the G-Channel (the latter, according to a well-known result from Information Theory, is the most severe assumption — which can be true or not). If the true noise is ingrained with memory a model that does not contemplate memory will not expose the available capacity. Better rate harvesting will be achieved, as it is well known [19], by using a model which tries to capture the random phenomenon memory. Modifying the BG-model to create a model that incorporates memory is a subject for further investigation.

Our investigation immediate target was to find the BG parameters. The large target, though, is to reach reliable communication through the use of error control codes for transmission over a medium disturbed by only additive noise. The capacity of the channel and the theoretical limits for the transmission rate were discussed in this paper. Studies engaged in analyzing impairments other than noise (like for instance fading, interference, and so on — impairments which are usually overcome with techniques other than error control codes) can also benefit from the current discussion.

Other existing estimation algorithms (log-likelihood estimation, etc.) will lead to a pair of estimates M_1, M_3, in a small neighborhood of the values obtained by the approach we have presented, which can *better match* the data to the theoretical model. This line of investigation—which relies on an explicit definition of the meaning of “better match”—was not pursued in the present article. It might prove fruitful but we conjecture that will not render significantly higher gains. We also conjecture that exploiting models which contemplate stochastic process embedded with memory will not significantly carve higher gains.

The PSD of a BG-process is flat. No attempt is made to shape the PSD, yet, if producing a process with a PSD exhibiting a specific shape is required, we envision that a filtered BG-process can produce the specified shape.

A question that naturally arises is: is the BG-model better than the Middleton model? This question was not addressed directly in the current paper — although we conjecture that BG might be, in general, as good as the Middleton model. Without significant distinction, it might happen that one model prevails in some scenarios and not in others. Investigation considering measured noise is needed to properly answer this question.

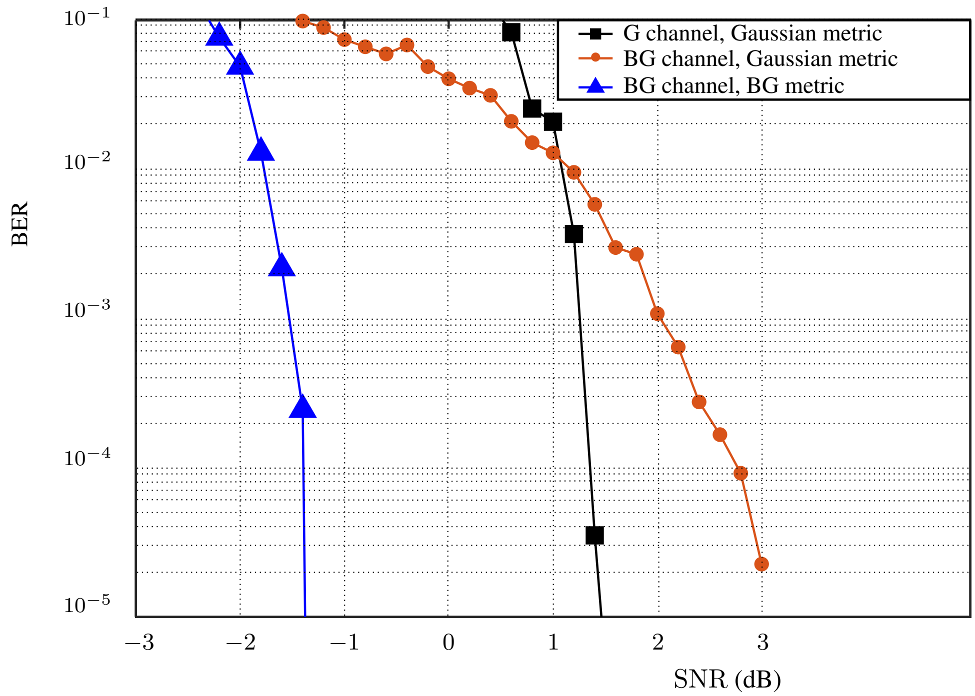


Fig. 5: Performance comparison of transmissions of data encoded with an LDPC (rate $R_c = k/n = 1/2$, $n = 4096$ over a BG-channel is shown. With the decoder designed under the assumption that the noise is modeled by a BG stochastic process would require signal-to-noise ratio $\sigma_X^2/\sigma_Z^2 = -1.1$ dB to achieve a bit-error rate $\mathrm{BER} = 10^{-5}$. If the receiver is designed by modeling the noise as a Gaussian process (wrong assumption) an extra 4.5 dB would be required to achieve the same performance. Also shown is the BER versus SNR curve of data encoded by using the same LDPC encoder and transmitted over a medium perturbed by purely Gaussian noise—with the receiver designed accordingly.

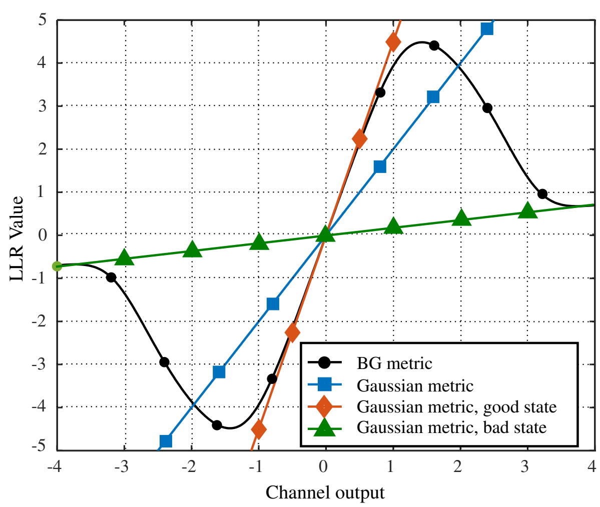


Fig. 6: Graphic of the likelihood function $\mathrm{LLR}(y)$ versus channel output y for several scenarios, $(\sigma_0, \sigma_1, p_1) = (0.6667, 3.3333, 0.05)$, and with $\mathrm{SNR} = 0$.

The issue just settled is not which model is better but how easy it is to find the BG parameters and how much can be gained when choosing to model non-Gaussian noise with a stochastic process other than the Gaussian process.

Appendix A
(A note on modeling with more than two states)

The two states BG model just discussed might be extended to contemplate more than two states. Let us consider the set $\Theta = \{0, 1, \ldots, M\}$, in which M is an integer, to be the set of states. We will need the *set indicator function* $\mathbb{I}\{S_i = \theta\}$ which is defined to be 1 if, for a given θ, the statement $S_i = \theta$ is true and, otherwise, this function of θ is defined to be zero. The noise samples perturbing the transmission can thus be expressed as $z_i = \sigma_\theta\, z_i^{[\theta]}\, \mathbb{I}\{S_i = \theta\}$ and we would be modeling this noise as a generalized-Bernoulli-Gaussian[6] (gBG) stochastic process $\{\breve{Z}_i\}$, namely,

$$\breve{Z}_i \quad = \quad \sigma_0\, R_i^{[0]} + \sum_{\theta=1}^{M} \sigma_\theta\, R_i^{[\theta]}\, \mathbb{I}\{S_i = \theta\}. \qquad (41)$$

In (41) the r.v.s $R_i^{[\theta]}$ have associated zero mean, Gaussian p.d.f.s with variance σ_θ^2. If $s_i = \theta$ the process is said to be in state θ (if $s_i = 0$ the process is in the background state). The process is specified if all the probabilities

$$\mathsf{Prob}\{S_i = \theta\} = p_\theta \qquad (42)$$

as well as all the parameters σ_θ are known. Of course it is

[6]As many readers might recognize the gBG is also know as Mixture of Gaussian process.

required to impose the condition

$$0 < \sum_{\theta=1}^{M} p_\theta < 1. \tag{43}$$

When $\theta \neq \theta'$ we have two r.v.s $R_i^{[\theta]}$ and $R_i^{[\theta']}$ which are considered as i.i.d. In the same vein, independents are the r.v.s $R_i^{[\theta]}$ and $R_j^{[\theta']}$ if $i \neq j$. With these considerations in mind, we thus conclude that any pair of distinct r.v.s $\breve{Z}_i$ and $\breve{Z}_j$ are i.i.d.

If $M = 0$ the sequence $\{S_i\}$ is, with probability 1, the sequence of all zeros, and the gBG becomes the Gaussian process. If $M = 1$ the gBG-process becomes the plain, two states, BG-process. The sequence $\{s_i\}$, probabilistically modeled by the sequence of i.i.d. r.v.s $\{S_i\}$, informs the state of the channel at time index i.

Choosing to model the noise with a gBG-process, as specified in (41), by taking $M > 1$ leads, of course, to a potentially better mathematical representation (better than that reached by selecting a BG model—which is a special instance of the gBG model with $M = 1$). But, since finding the parameters of this even more sophisticated model is a hard task, we aimed at the still sophisticated but simpler model. The procedure presented allows to find, with easiness, the BG model parameters $(\sigma_0, \sigma_1, p_1)$ when a sequence of samples of the noise which perturbs the transmission is known. Yet simple the three-parameters model enlarges the insights brought by theoretical studies of a transmission media perturbed by noise which is not well modeled as a Gaussian noise (such as noise perturbing transmission over power lines). The examination of the capacities, as evaluated by using the two models (Gaussian and Bernoulli-Gaussian), showed that under the same conditions, a non-negligible amount of capacity might be overlooked.

Examining Fig. 2 one can see that there is extra room for improvement by choosing to model the noise using the more flexible (parameters) BG-process rather than picking the more restricted (single parameter) G-channel. Error control codes design relying on the BG-channel model can benefit from the choice of a better model. A number of examined examples (one of which is discussed in Section VI) corroborate this fact.

A well-known probabilistic model for impulsive noise, proposed by Middleton [20], is an instance of the generalized BG stochastic process in (41). It is a three parameter model that can be written as

$$\breve{Z}_i = \sum_{\theta=0}^{\infty} \sigma_\theta \, R_i^{[\theta]} \, \mathbb{I}\{S_i = \theta\}, \tag{44}$$

with the channel states running over an infinite set, namely, $\theta \in \{0, 1, \ldots, \infty\}$. Middleton model [21] imposes also that Poisson be the probability law which is to rule the probability $p_\theta = \mathsf{Prob}\{S_i = \theta\}$ or, specifically,

$$p_\theta = \frac{A^\theta}{\theta!} e^{-A}. \tag{45}$$

In this case, it should be noticed that the probability of observing a noise sample in the background state is $p_0 = e^{-A}$. The power of the sequence of samples in state θ, $(\theta > 0)$ can be expressed, following the current paper notation, through

$$\sigma_\theta = \sqrt{\frac{1 + \theta(\sigma_{\breve{Z}}^2 - \sigma_0^2)}{A}}. \tag{46}$$

The p.d.f. of the r.v. $\breve{Z}$ is,

$$\mathsf{f}_{\breve{Z}}(\zeta) = \sum_{\theta=0}^{\infty} p_\theta \left(\frac{1}{\sigma_\theta \sqrt{2\pi}} e^{-\zeta^2/2\sigma_\theta^2} \right). \tag{47}$$

Since $\sigma_{\breve{Z}}^2 = \sum_{\theta=0}^{\infty} p_\theta \sigma_\theta^2$ one needs thus, to specify the process, just to know the three parameters, A, σ_0 and $\sigma_{\breve{Z}}^2$. Techniques to estimate the parameters of the Middleton Model have been presented in [9], [16], [20].

The Middleton model yet admitting that the noise goes through a very large number of states (as does the gBG) is, however, a model which restricts the states to obey a Poisson distribution, and depends on only three parameters. In scenarios in which the noise is non-Gaussian one can thus, potentially, provide equivalent estimates of the process p.d.f. by using either the BG model or the Middleton model (only practice can tell apart which model fits better — the authors conjecture that the models using one model or the other would have close performance).

APPENDIX B
(PROOFS OF THEOREMS)

Proof of Theorem 1: From the definition of μ_1, μ_2, and μ_3, and considering $\tilde{\alpha} = \sqrt{\alpha_{10}^2 + 1}$ it can be observed that

$$\mu_1 = (p_0 + p_1 \tilde{\alpha}) \, \sigma_0 \tag{48}$$

$$\mu_2 = \left(p_0 + p_1 \tilde{\alpha}^2\right) \sigma_0^2 \tag{49}$$

$$\mu_3 = \left(p_0 + p_1 \tilde{\alpha}^3\right) \sigma_0^3 \tag{50}$$

Therefore, $\mu_3 - \mu_1\mu_2$, $\mu_1\mu_3 - \mu_2^2$, and $\mu_2 - \mu_1^2$ can be factored as

$$\mu_3 - \mu_1\mu_2 = \sigma_0^3 \, p_0 \, p_1 \, (\tilde{\alpha} - 1)^2 \, (\tilde{\alpha} + 1), \tag{51}$$

$$\mu_1\mu_3 - \mu_2^2 = \sigma_0^4 \, p_0 \, p_1 \tilde{\alpha} \, (\tilde{\alpha} - 1)^2, \tag{52}$$

$$\mu_2 - \mu_1^2 = \sigma_0^2 \, p_0 \, p_1 \, (\tilde{\alpha} - 1)^2. \tag{53}$$

Since $p_0 \in (0, 1]$ and $\tilde{\alpha} \geq 1$, the theorem is proved. ∎

Proof of Theorem 2: From (51), (52) and (53), we get

$$\begin{aligned} \kappa &= \frac{(\mu_3 - \mu_1\mu_2)^2}{(\mu_1\mu_3 - \mu_2^2)(\mu_2 - \mu_1^2)} \\ &= \frac{(\tilde{\alpha} + 1)^2}{\tilde{\alpha}} \end{aligned} \tag{54}$$

which rewritten becomes

$$\kappa = \tilde{\alpha} + \frac{1}{\tilde{\alpha}} + 2.$$

Since

$$\tilde{\alpha} = \sqrt{\left(\frac{\sigma_1}{\sigma_0}\right)^2 + 1} \geq 1$$

and $\frac{\sigma_1}{\sigma_0} > 0$, we have that $(\tilde{\alpha} - 1)^2 > 0$ and, thus,

$$\tilde{\alpha}^2 - 2\tilde{\alpha} + 1 > 0. \tag{55}$$

This is equivalent to say that

$$\tilde{\alpha} + \frac{1}{\tilde{\alpha}} > 2,$$

and, therefore, $\kappa > 4$. ■

Proof of Theorem 3: From (54), it can be seen that

$$\tilde{\alpha}^2 - (\kappa - 2)\tilde{\alpha} + 1 = 0, \tag{56}$$

has two roots, namely, $\xi = \frac{1}{2}\left(\kappa - 2 \pm \sqrt{(\kappa-2)^2 - 4}\right)$. However, since

$$\begin{aligned}\sqrt{(\kappa-2)^2 - 4} &= \sqrt{\kappa(\kappa-4)} \\ &> \kappa - 4\end{aligned}$$

one of the roots is lower than one. Since $\tilde{\alpha} > 1$, this root has no meaning within the probabilistic model. Therefore

$$\xi = \frac{1}{2}\left(\kappa - 2 + \sqrt{\kappa(\kappa-4)}\right) \tag{57}$$

is the root of interest, justifying equality (17). The parameters of the model can be easily found once it is known that $\tilde{\alpha} = \xi$ is the root. The value of σ_0, established in (18), comes from the division of (51) by (53). The computation of p_0, posted in (20), comes directly from (49). ■

APPENDIX C
(NOTES ON THE IMPLEMENTATION OF THE PROCEDURE)

If the expected values $\mathbb{E}[|Z_i|]$, $\mathbb{E}\left[|Z_i|^2\right]$ and $\mathbb{E}\left[|Z_i|^3\right]$ of the stochastic process $\{Z_i\} = \{\sigma_0 R_i^{[0]} + \sigma_1 S_i R_i^{[0]}\}$ are known, the values $(\sigma_0, \sigma_1, p_1)$ of the parameters can be obtained right away by using Theorem 3. In practice, this never happens and estimated moments have to be used.

In our implementation of the procedure to find the BG parameters we work with normalized absolute moments (as if all the samples had been divided by $\hat{\sigma}_Z$) or, in other words, the estimated moments used are $\overline{M}_1 = M_1/\sqrt{M_2}$ and $\overline{M}_3 = M_3/(\sqrt{M_2})^3$ with, obviously, setting $\overline{M}_2 = 1$.

When attempting to find the parameters $(\overline{\sigma}_0, \overline{\sigma}_1, p_1)$ by using normalized estimates $\overline{M}_1$ and $\overline{M}_3$ (obtained from a given set of samples $\{z_i\}_{i=1}^N$) of the moments $\mathbb{E}\left[|\overline{Z}_i|\right]$ and $\mathbb{E}\left[|\overline{Z}_i|^3\right]$ an estimation error is inevitable. Usually, the moment estimates are quite precise. If the imprecisions of the estimated values $\overline{M}_1$ and $\overline{M}_3$ drive the algorithm to solutions falling in the feasible region there is no immediate action to be taken and the obtained parameters $(\hat{\sigma}_0, \hat{\sigma}_1, \hat{p}_1)$ are taken as the estimated parameters. If, on the other hand, the solution does not fall in the feasible region, actions have to be provided to circumvent the imprecision of the estimated parameters. Or, maybe, admit an inconsistency of the model since finding a point far outside the feasible region can be taken as a sign that the true noise definitely does not have a behavior that can be described by a BG-process. In the current investigation, no attempt was made to obtain the *best estimation* (a concept which needs a definition and justification). A way around this difficulty (if the estimated moments fall outside the feasible region) is to alter the values $(\overline{M}_1, \overline{M}_3)$ and find a near point $(\overline{M}_1', \overline{M}_3')$ falling in the feasible region.

A plot of $\hat{\sigma}_{|Z|^2}$ versus $\hat{\sigma}_{|Z|}$, using the inequalities in (13), (14), and (15), exhibit, on Fig. 7, the feasible region.

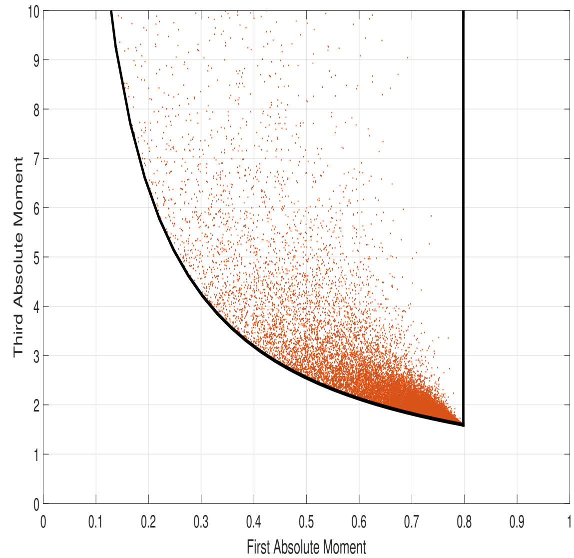


Fig. 7: The cloud of red points indicates the feasible region. Dark lines are the frontiers of the region. The feasible region corresponds to all points $(\overline{M}_1, \overline{M}_3)$ to which points $(\overline{\sigma}_0, p_1)$ in the region and $(0, 1]\times, [0, 1]$ are mapped to.

REFERENCES

[1] T. M. Cover and J. A. Thomas, *Elements of Information Theory.* Hoboken, NJ, USA: John Wiley & Sons, Inc., 2nd ed. 2006. ISBN-13 978-0471241959.

[2] L. Clavier, G. W. Peters, F. Septier, and I. Nevat, "Impulsive noise modeling and robust receiver design," *J. Wireless Com. Network*, no. 13, 2021. doi: 10.1186/s13638-020-01868-1.

[3] M. Zimmermann and K. Dostert, "Analysis and modeling of impulsive noise in broad-band power-line communications," *IEEE Trans. Electromag. Compat.*, vol. 44, no. 1, pp. 249–258, Feb. 2002. doi: 10.1109/15.990732.

[4] I. Landa, A. Blázquez, M. Vélez, and A. Arrinda, "Indoor measurements of IoT wireless systems interfered by impulsive noise from fluorescent lamps," *11th Eur. Conf. on Antennas and Propagation,* pp. 2080–2083, 2017. doi: 10.23919/EuCAP.2017.7928172.

[5] M. Stojanovic and J. Preisig, "Underwater acoustic communication channels: propagation models and statistical characterization," *IEEE Commun. Mag.*, vol. 17, no. 1, pp. 84-89, Jan. 2009. doi: 10.1109/MCOM.2009.4752682.

[6] S. Liu, F. Yang, W. Ding, and J. Song, "Double kill: compressive-sensing-based narrow-band interference and impulsive noise mitigation for vehicular communications," *IEEE Trans. Veh. Technol.*, vol. 65, no. 7, pp. 5099–5109,Jul. 2016. doi: 10.1109/TVT.2015.2459060.

[7] D. Fertonani and G. Colavolpe, "On reliable communications over channels impaired by bursty impulse noise," *IEEE Trans. Commun.*, vol. 57, no. 7, pp. 2024–2030, Jul. 2009. doi: 10.1109/TCOMM.2009.07.070638.

[8] S. P. Herath, N. H. Tran, and T. Le-Ngoc, "Optimal signaling scheme and capacity limit of PLC under Bernoulli-Gaussian impulsive noise," *IEEE Power Deliv.*, vol. 30, no. 1, pp. 97-105, Feb. 2015. doi: 10.1109/TPWRD.2014.2330197.

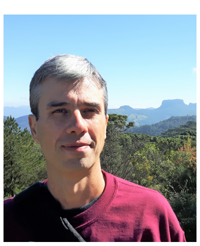
[9] D. Middleton, "Non-Gaussian noise models in signal processing for telecommunications: new methods and results for class A and class B noise models", *IEEE Trans. Inform. Theory*, vol. 45, no. 4, pp. 1129-1149, May 1999. doi: 10.1109/18.761256.

[10] F. Rouissi, A. J. Han Vinck, H. Gassara, and, A. Ghazel, "Improved impulse noise modeling for indoor narrow-band power line communication," *AEU Int. J. Electron. and Commun.*, vol. 103, pp. 74-81, May 2019. doi: 10.1109/ISPLC.2017.7897119.

[11] J. A. Cortes, A. Sanz, P. Estopiñan, and J. I. García "On the suitability of the Middleton class A noise model for narrowband PLC," *IEEE Int. Symp. Power Line Communications and its Applications*, Mar. 2016. doi: 10.1109/ISPLC.2016.7476256.

[12] M. Nassar, K. Gulati, Y. Mortazavi, and B. L. Evans “Statistical modeling of asynchronous impulsive noise in powerline communication networks,” *IEEE Global Telecom. Conf.*, Dec. 2011. doi: 10.1109/GLOCOM.2011.6134477.
[13] S. Liu, F. Yang, and J. Song, “An optimal interleaving scheme with maximum time-frequency diversity for PLC systems,” *IEEE Power Deliv.*, vol. 31, no. 3, pp. 1007-1014, Jun. 2016. doi: 10.1109/TPWRD.2014.2365579.
[14] G. Laguna-Sanchez and M. Lopez-Guerrero, “On the use of alpha-stable distributions in noise modeling for PLC,” *IEEE Power Deliv.*, vol. 30, no. 4, pp. 1863-1870, Aug. 2015. doi: 10.1109/TPWRD.2015.2390134.
[15] S. M. Zabin and H. V. Poor, “Recursive algorithms for identification of impulse noise channels,” *IEEE Trans. Inform. Theory*, vol. 36, no. 3, pp. 559-578, May 1990. doi: 10.1109/18.54878.
[16] S. M. Zabin and H. V. Poor, “Efficient estimation of class A noise parameters via the EM algorithm,” *IEEE Trans. Inform. Theory*, vol. 37, no. 1, pp. 60-72, Jan. 1991. doi: 10.1109/18.61127.
[17] A. Papoulis and S. U. Pillai, *Probability, Random Variables and Stochastic Processes.* New York, NY, USA: McGraw-Hill Eduaction, 4th ed. 2002. ISBN-13 978-0073660110 .
[18] J. G. Proakis and M. Salehi, *Digital Communications.* New York, NY, USA: McGraw-Hill Education, 5th ed. 2007. ISBN-13 978-0072957167.
[19] M. Mushkin and I. Bar-David, “Capacity and coding for the Gilbert-Elliott channels,” *IEEE Trans. Inform. Theory*, vol. 35, no. 6, pp. 1277-1290, Nov. 1989. doi: 10.1109/18.45284.
[20] D. Middleton, “Canonical non-Gaussian noise models: Their implications for measurement and for prediction of receiver performance,” *IEEE Trans. Electromag. Compat.*, vol. EMC-21, no. 3, pp. 209-220, Aug. 1979. doi: 10.1109/TEMC.1979.303732.
[21] T. Bai *et al.*, “Fifty years of noise modeling and mitigation in power-line communications,” *IEEE Commun. Surv. Tutor.*, vol. 23, no. 1, pp. 41-69, Firstquarter 2021. doi: 10.1109/COMST.2020.3033748.
[22] A. Ashikhmin, G. Kramer, and S. Ten Brink, “Extrinsic information transfer functions: Model and erasure channel properties,” *IEEE Trans. Inform. Theory*, vol. 50, no. 11, pp. 2657-2673, Nov. 2004. doi: 10.1109/TIT.2004.836693.

Weiler A. Finamore received the Ph.D. degree in electrical engineering from the University of Wisconsin, Madison, WI, USA, in 1978. He is currently with the Federal University of Juiz de Fora (UFJF), Juiz de Fora, Minas Gerais, Brazil, and the Instituto Tecnológico de Aeronáutica (ITA), as a Visiting Research. From 1979 until his retirement in 2012, he was on the faculty of Electrical Engineering with the Catholic University of Rio de Janeiro, Rio de Janeiro, Brazil. At UFJF, he was supported by the CNPq (Brazilian Science Foundation) as a Senior Visiting Researcher from October 2013 to September 2015 and by the FAPEMIG (Minas Gerais State Science Foundation) from April 2012 to September 2013. At ITA, he was supported by the FAPESP as a Visiting Research from March to July, 2023. He has also worked with the Federal University of Para, Belem, Pará, Brazil, and with the IBM Rio Scientific Center, Rio de Janeiro. Among his publications is Probabilidades, Variaveis Aleatorias e Processos Estocasticos (in Portuguese; Interciencia, 2^{nd} ed., 2017). His research interests include the areas of information theory and coding and its applications in communications systems.



Marcelo S. Pinho received the B.Sc., M.Sc. and D.Sc. degrees in electrical engineering from Pontifícia Universidade Católica do Rio de Janeiro, in 1994, 1996 and 2000, respectively. In 1998, he was a visiting scholar in the Center for Image Processing Research (CIPR) of the Rensselaer Polytechnic Institute (RPI), NY, USA. From 2000 to 2002, he worked in the Department of Electrical Engineering of the Universidade Estadual Paulista at Guaratinguetá, SP. Since April of 2002, he has been with the Department of Electronic Engineering at Instituto Tecnológico de Aeronáutica (ITA). His research interests include information theory and digital communication.

Manish Sharma received his Electrical Engineering degree in 2003, his Master’s degree in Electrical Engineering in 2006, and his Doctorate in Electrical Engineering in 2010, all from the University of Campinas (Unicamp). Since 2009 he is a professor at the Aeronautics Institute of Technology. His main research interests are multiuser communication systems and channel coding.

Moisés V. Ribeiro received a B.S. degree in Electrical Engineering from the Federal University of Juiz de Fora (UFJF), MG, Brazil, and an M.Sc. and D.Sc. Degrees in Electrical Engineering from the University of Campinas, SP, Brazil, in 1999, 2001, and 2005. He was a Visiting Scholar at the University of California in Santa Barbara, CA, USA, in 2004, a Visiting Professor (2005-2007), and an Assistant Professor (2007-2015) at UFJF. Since 2015, he has been an Associate Professor at UFJF. He is also the Director of the Brazilian National Institute of Science and Technology for Electricity (INERGE). He co-founded Smarti9 LTD. and Wari LTD. in 2012 and 2015, respectively. He had served as the Secretary of the IEEE ComSoc TC-PLC.

Dr. Ribeiro was the recipient of a Fulbright Visiting Professorship at Stanford University, Stanford, CA, USA, in 2011, and at Princeton University, Princeton, NJ, USA, in 2012. He was the General Chair of the 2010 IEEE ISPLC, 2013 IWSGC, SBrT 2015, and a Guest Co-Editor for Special Issues in the EURASIP Journal on Advances in Signal Processing and EURASIP Journal of Electrical and Computer Engineering. He was awarded Student Awards from 2001 IEEE IECON and 2003 IEEE ISIE, Winner of the 2014 I2P Global Competition, Honorable Mention in 2014 Global Venture Labs Investment Competition, 3rd Place Prêmio Mineiro de Inovação 2014, Engie Brazil Innovation Award 2016, Unicamp Inventor Awards in 2017 and 2018, and 2021 TC-PLC Outstanding Service Award from the IEEE ComSoc.

Dr. Ribeiro’s research interests include signal processing, power line communication, wireless communication, computational intelligence, and the Internet of Everything. He has advised 45 graduate students in these fields and authored over 252 peer-reviewed papers and nine book chapters. He holds 13 issued/pending patents.