On modeling power-line communication noise

Victor Fernandes, Sanja Angelova, Weiler A. Finamore and, Moisés V. Ribeiro

Abstract— **This article introduces an algorithm to estimate parameters of measured power line noise modeled as the Bernoulli-Gaussian noise. These are important parameters when designing power line communications systems. Tests with several samples of noise registered during a measure campaign in the city of Juiz de Fora, MG, were performed and the parameters of these measured noise are obtained. Tests with synthetic noise generated according to the Bernoulli model for power line noise are a good evidence of the algorithm effectiveness.**

Keywords— **Power line communication, Power line noise, Noise parameter estimation.**

I. INTRODUCTION

The characteristic of the noise perturbing Power Line Communication (PLC)—hereof called *power line noise*—has been the subject studies of many authors [1]–[3] and many models have been proposed. Power line noise is considered to be an instance of what is called *Impulsive Noise*. PLC noise is a Stochastic Process which for a given percentage of the time is in a severe state ("strong" noise) and for the remaining time is on a mild state ("weak" noise). A model which considers that the noise severity level can be classified in an infinite number of states is known as the *Middleton Class A* model [4]. As shown in [5] the *Two-state, Middleton Class A* model can be as accurate as the general *Middleton Class A*. The PLC systems to be addressed in this paper are digital communication systems and, for this reason, the noise model known as Bernoulli-Gaussian Model will be considered — the Bernoulli-Gaussian (BG) noise is the discrete-time equivalent to the continuous-time *Two-state, Middleton Class A* model. The simple Bernoulli-Gaussian Model [6] considers that the noise samples are random variables which are classified as either "mild" (low variance) or "strong" (i.e., drawn from a Gaussian distribution high variance). At a given instant of time, either a small power noise (also commonly known as the background noise component) due to thermal noise, or a large power noise (the impulsive noise component) due to natural phenomenon (atmospheric disturbance, etc) or man-made is added to the signal. The Bernoulli-Gaussian Model will be the focus of the present study. The noise affecting the PLC can be considered,

The paper has been organized as follows: in Section I-A the Bernoulli-Gauss model is presented and its relation to the the *Two-state Middleton Class A* is briefly discussed. In Section II a model similar to the Bernoulli Model (or, for that matter, to the *Two-state, Middleton Class A*) is discussed. PLC noise modeled as an additive impulsive Gaussian noise (AIGN) has been synthetically generated by using this model, with the purpose of validating the proposed algorithm. The algorithm to find the BG parameters of measured PLC noise is introduced in Section III. Results of the algorithm are presented in Section IV, and conclusions are then presented in the final section, Section V.

Victor Fernandes¸ Weiler A. Finamore and Moisés V. Ribeiro, are with Departamento de Engenharia Elétrica, Universidade Federal de Juiz de Fora, Juiz de Fora-MG, Brazil. Sanja Angelova, is with the University for Information Sciences and Technology, St Paul Apostol, Ohrid, Macedonia. E-mails: victor.fernandes@engenharia.ufjf.br, finamore@ieee.org, mribeiro@engenharia.ufjf.br, sanja.angelova@uist.edu.mk.

A. *Bernoulli-Gaussian Noise*

As discussed in [6] the PLC digital noise (or, synonymously, PLC discrete-time noise) can be modeled as a Stochastic Process $\nu(t_i)$ which, at a given time instant t_i, is either a "weak" Gaussian random variable (r.v.) — with variance $\sigma_\nu^2 = \sigma^2$ — or, "strong," i.e., a Gaussian r.v. with variance $\sigma_\nu^2 = \alpha^2\sigma^2$, ($\alpha > 1$). At any two distinct time instants t_i and t_j the r.v.'s $\nu(t_i)$ and $\nu(t_j)$ are independent. Further, given that N is the number of PLC noise samples being collected and stored, we have then that the number of weak noise samples is equal to $\lceil pN \rceil$ $(1/2 < p < 1)$ and, of course, the remaining $\lfloor (1-p)N \rfloor$ are strong noise samples. A noise with the behavior above described, known as a Bernoulli-Gaussian noise, can be defined as follows.

Definition 1: Let $\{U_i\}$ be a sequence of Bernoulli r.v. with $Pr[U_i = 1] = p$ (and, of course, $Pr[U_i = 0] = 1 - p$) and $\{W_i\}$ be a sequence of independent identically distributed Gaussian r.v. with zero mean and variance σ^2. The sequence of r.v. $\{\nu_i\}$ with every component r.v., ν_i, given by

$$\nu_i = U_i W_i + \alpha(1 - U_i)W_i, \tag{1}$$

in which $\alpha > 1$, is a Bernoulli-Gaussian noise. ■

Every sample $\nu_i = \alpha W_i$, corresponding to a Bernoulli sample such that $U_i = 0$, will be designate by *strong variance noise component* of the AIGN. Theses samples are, zero mean, Gaussian r.v. with variance $\alpha^2\sigma^2$. The remaining samples, $\nu_i = W_i$, corresponding to a Bernoulli sample such that $U_i = 1$, are Gaussian r.v. with zero mean and variance σ^2 — these are known as *background noise components*. The variance of the *impulsive noise*, used to model the PLC noise, is, thus,

$$\sigma_\nu^2 = p\sigma^2 + (1-p)\alpha^2\sigma^2 \tag{2}$$

The *Two-state, Middleton Class A* gives an statistical characterization for the continuous-time PLC noise $\nu(t)$ (see [6]) which considers that, for any time window, the noise behaves like background noise (is "weak") for a fraction $(1 - A)$ of the time $(0 < A < 1)$ — meaning that, for any instant t within this fraction, the observed noise is modeled as a zero mean, variance σ^2 Gaussian r.v. $\nu(t)$ — and, for the remaining fraction of time A, there will be a strong noise which is modeled as a Gaussian r.v. $\nu(t)$ which is zero mean and has variance $\sigma^2 + \sigma^2/A\Gamma$, $0 < \Gamma < 1$. During two distinct time instants the two defined random variables are independent.

The *Two-state, Middleton Class A* model depends on the parameters (σ^2, Γ, A), similarly, the Bernoulli-Gaussian discrete-time model depends on the parameters (σ^2, α, p). Since σ^2 represents the same variance in both models, we can easily find

$$p = 1 - A \quad (3)$$
$$\alpha = \sqrt{1 + 1/A\Gamma} \quad (4)$$

Let us now consider that a PLC system which delivers to the receiver a BPSK modulated signal, with energy per bit equal to E_b, has been transmitted trough a PLC channel (transmission over an AIGN channel). The performance of power line digital communication system which transmits bits over a binary channel with inputs $x_i \in \{+1, -1\}$ and a digital noise with components modeled as a Gaussian zero mean r.v. ν_i which, with probability p has variance σ^2 (with a flat power spectrum characterized by $N_0 = 2\sigma$) and, with probability $1-p$, has variance $\alpha^2\sigma^2$ is given by the curve of P_e (the probability of error in such a system) versus $\frac{E_b}{N_0}$, can be plotted by using the well known expression [6]

$$P_e = pQ\left(\sqrt{\frac{E_b}{\sigma^2}}\right) + (1-p)Q\left(\sqrt{\frac{E_b}{\alpha^2\sigma^2}}\right). \quad (5)$$

in which $Q(x) \triangleq \frac{1}{\sqrt{2\pi}}\int_x^\infty e^{-t^2/2}dt$.

This curve is plotted in Fig. 1 for $(\alpha, p) = (30, 0.99)$. Also

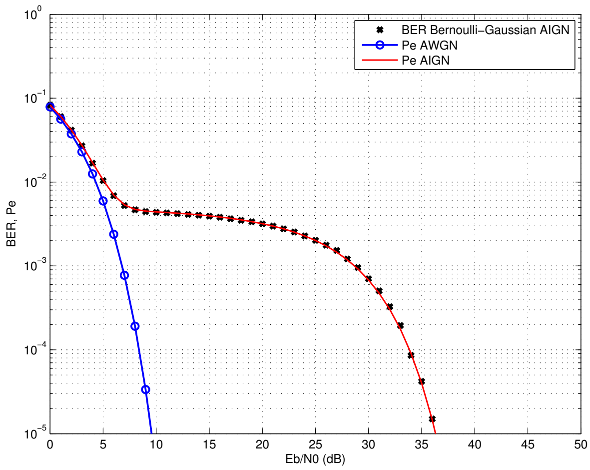

Fig. 1. Plots of P_e vs E_b/N_0 performance of two communication systems: (1) system which transmit a BPSK signal through an AWGN channel — lower curve (blue) and, (2) system which transmit a BPSK signal through an AIGN channel (PLC corrupted by BG noise with parameters $(\alpha, p) = (30, 0.99)$) — upper curve (red). Also shown, is the BER vs E_b/N_0 obtained by computer simulation — black curve (overlapped with the red curve).

plotted in this figure is the curve of P_e versus $\frac{E_b}{N_0}$ for a system (let us call this the reference system) where the transmission takes place over a BI-AWGN channel. As remarked in [6], when the value of $\frac{E_b}{N_0}$ is low, both the reference system and the PLC system exhibit a performance which are equivalent. When the values of $\frac{E_b}{N_0}$ are large, however, the required value of E_b for the communication over the PLC system to achieve the same probability of error might be quite large (over 25 dB when compared to the reference system — for the case pictured in Fig. 1). It should be noticed that BER versus $\frac{E_b}{N_0}$ curve where BER is the estimated value of P_e, obtained by simulating the transmission over a PLC channel using MATLAB to generate the AIGN is also shown in the figure. If we consider that the real life PLC noise is well modeled by the probabilistic model (the Bernoulli-Gaussian noise) and that the simulation results is a computer model that closely reflects the actual transmission over PLC (transmission over an AIGN channel) we can come up to the conclusion that the theory and practice are in good agreement.

This is, as a matter of fact, our main objective, i.e., to evaluate how well can the model (Bernoulli-Gaussian in this case) predict the behavior of practical systems (real life systems). A first step in this direction is the algorithm presented in the following section.

An issue of concern, when using a noise model is how to find the noise model parameters of the measured noise. Many papers have discussed noise models but no method to raise the values of the parameters of PLC measured noise has been found in the literature. In this direction a method to come up with the parameters of real life PLC noise is presented in the following section.

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

BER vs E_b/N_0 performance like those in Fig. 1 can be easily plotted once the parameters σ^2, α and p, are known. In practice when a given sequence of N noise samples, say $(y_1, \ldots, y_i \ldots, y_N)$, is obtained from measurements, the values of these parameters are not known. The task is then to observe these values and classify each measured noise sample y_i (considered to be zero mean) as either weak (a realization of r.v. UW) or strong (a realization of r.v. $\alpha(1-U)W$) and, by doing so, partitioning this sequence and finding the estimates $\widehat{\sigma}^2$, $\widehat{\alpha}$, and $\widehat{p}$.

To solve to this problem we will use the well known solution to the classical decision problem in which, upon receiving a string of random variables $\{\nu\}$ — drawn from known distribution for U and W (in other words, with known value of σ^2, α, and p) — find a threshold ℓ, that minimizes the probability of taking the wrong decision. The decision rule declares that an observed sample y_i is weak if $y_i \leq \ell$ or, otherwise, to be a strong noise sample. It can be shown that the probability of wrong decision is

$$P_d = 2pQ\left(\frac{\ell}{\sigma}\right) - 2(1-p)Q\left(\frac{\ell}{\alpha\sigma}\right) + (1-p). \quad (6)$$

The threshold ℓ_{OPT} that minimizes (6), the probability of taking a wrong decision is

$$\ell_{OPT} = \sigma\lambda, \tag{7}$$

in which

$$\lambda = \alpha\sqrt{\frac{2\ln(\alpha p/(1-p))}{\alpha^2-1}}. \tag{8}$$

It is easy to show, when the triple (σ^2, α, p) is assumed to be known (the development is presented in the Apendix), that $q_\ell = Pr[\nu_i \geq \ell]$ and $M_\ell = \mathbb{E}[|\nu|]$ are

$$q_\ell = 2(pQ(\lambda) + (1-p)Q(\lambda/\alpha)), \tag{9}$$

$$M_\ell = \frac{2\sigma}{\sqrt{2\pi}}(p + \alpha(1-p)). \tag{10}$$

In practice when a given sequence of noise samples are obtained, the three values are not known and an algorithm to estimate these parameters, to the authors knowledge, has never been published. In the next section we present an algorithm to solve this problem by using basically an exhaustive search (at this point, there was no attempt to make the algorithm faster or more efficient, this issue has been left for further research).

III. An algorithm to estimate the Bernoulli-Gaussian parameters of measured PLC noise

The algorithm to estimate the Bernoulli-Gaussian parameters of which models a measured PLC noise is now presented. The target is to search for a pair of consistent values $(p', \alpha') \in (1/2, 1) \times (1, \infty)$ which make the values of the error between the values of q'_ℓ and M'_ℓ obtained by taking the values (p', α') into (9) and (10) and the values of $\widehat{q}_\ell$ and $\widehat{M}_\ell$ obtained from ν. These values of error are defined as

$$E_M = \widehat{M}_\ell - M'_\ell, \tag{11}$$

$$E_q = \widehat{q}_\ell - q'_\ell. \tag{12}$$

The procedure to solve, numerically, the problem of finding the best pair (α', p') is presented in Algorithm 1 (see Appendix II). Several pairs of consistent values of (α, p) and its corresponding values of E_M and E_q are examined by the procedure. A pair (α, p) is considered a consistent pair if $E_M(p_i, \alpha_i) = E_q(p_i, \alpha_i)$. For those pairs which are consistent values, the pair such that $E_M(p_i, \alpha_i)$ is closest to $E_M = 0$ are the best choice.

IV. Results

Fig. 2 is a plot of Bernoulli-Gaussian impulsive noise samples of $\nu(t)$ generated, according to (1), with $\alpha = 10$ and $p = 0.9$.

Fig. 3 displays plots of E_M versus α for several values of p (upper set of curves) and also plots of E_q versus α for several values of p (lower set of curves) parameterized with the values of p. The noise samples used by the algorithm are synthetic generated noise. The intersection of E_M and E_q curves near the abscissas axis are marked with black dots. From these plots and intersections we can tell that $(\widehat{\alpha}, \widehat{p}) = (9.99, 0.99)$, very close to the true values. This result gives a clear indication that the theory developed as well as the numerical algorithm are a reliable tool Bernoulli-Gaussian modeling measured PLC noise (noise measured over power lines).

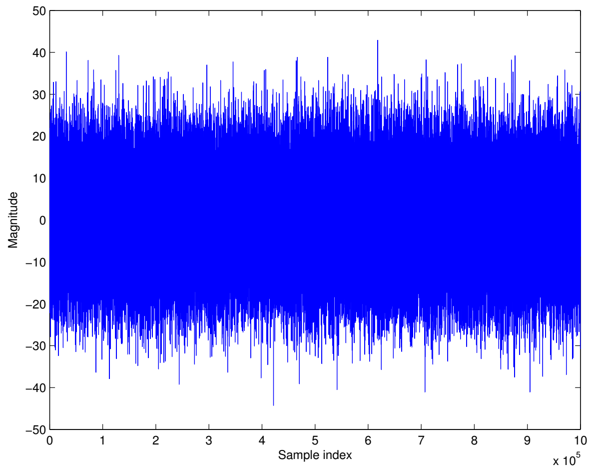

Fig. 2. Synthetically generated PLC noise with parameters $\alpha = 10$ and $p = 0.9$.

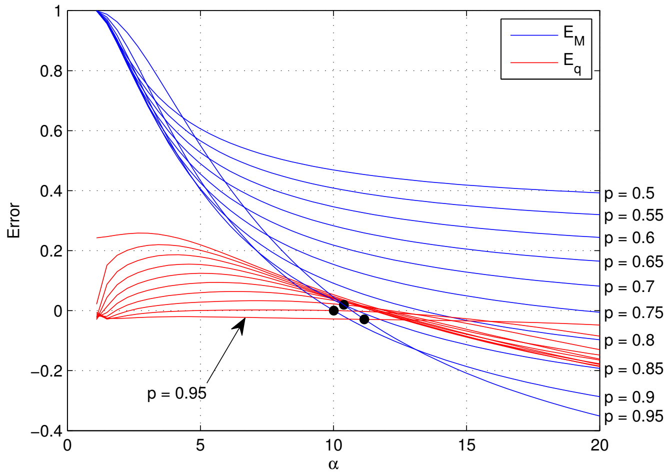

Fig. 3. Estimation of the parameters of the synthetically generated sequence of samples $\nu[k]$. The pairs (α, p) = (9.99, 0.99), (10.2, 0.85) and (11.1, 0.95) with corresponding values $E_M = 0$, 0.02 and -0.02 are indicated by the black dots. The best choice is $(\widehat{\alpha}, \widehat{p})$ = (9.99, 0.99).

To further illustrate the power of the developed tool, the parameters of a sequence with 100000 samples of the noise obtained by measurements over power lines were examined (this will be represented by $n(t)$). A segment of such samples are exhibited in Fig. 4. Details of the measurement campaign can be found in [7].

The plots of curves E_M and E_q versus α for several values of p are displayed in Fig. 5. These plots (with the intersection of lines with the same value of p marked with black dots allows us to tell that $(\widehat{\alpha}, \widehat{p}) \approx (2.8, 0.73)$ are the best estimate for the parameters of the Bernoulli-Gaussian model which better

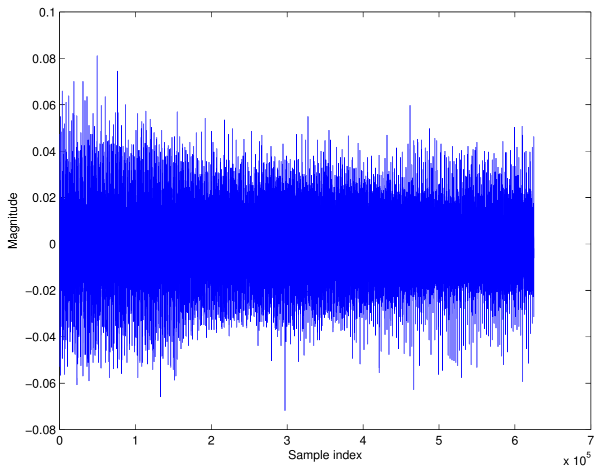

Fig. 4. Measured PLC noise samples $n[k]$.

represent the measured (real life) PLC noise.

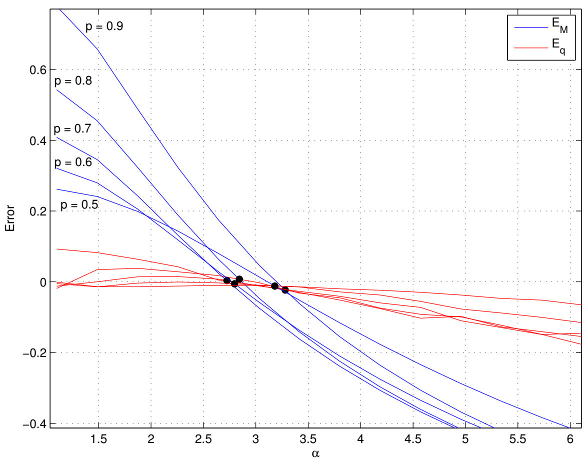

Fig. 5. Estimation of the parameters of the measured PLC noise. Using the described procedure, the best choice, examining this plot, is $(\widehat{\alpha}, \widehat{p})$ = (2.8, 0.73).

As we can see, in Fig. 6, the theoretical error probability (equation (5)) computed with the parameters obtained by the algorithm $(\widehat{\alpha}, \widehat{p}) \approx (2.8, 0.73)$, red curve, is close to the simulated BER of the same measured PLC noise (green curve). Also, we can see, in blue, the reference curve.

V. Conclusion

We have presented an algorithm to find the parameters of a Bernoulli-Gaussian noise used to model measured PLC noise. These parameters σ^2, α and p are introduced in Definition 1.

To assess the quality of the algorithm, computer generated synthetic noise with known values of σ^2, α and p has been produced according to the Bernoulli-Gaussian model. This synthetic noise when used as the input data to the algorithm had its parameters estimated with quite good accuracy (the true parameters were $\sigma^2, \alpha, p = (1, 10, 0.9)$ and the estimated parameters were $\widehat{\sigma}^2, \widehat{\alpha}, \widehat{p} = (0.98, 9.99, 0.9)$). As it has been discussed, in Section IV, good results have been reached at estimating quite accurately values of these parameters (at this point in time we can not envision a better procedure to estimate the parameters). With such a good tool the estimation of the Bernoulli-Gauss parameters of measured PLC noise can be obtained with good accuracy in the sense that Probability of Error performance of the system perturbed by the measured noise and by the synthetic noise are very close. Some interesting results, describing models, are presented in [6]—but no attempt to find the noise model parameters that best mimic the measured data has been done. With this goal in mind—how to estimate the BG parameters that characterizes the measured PLC noise—an algorithm has been proposed in the current paper.

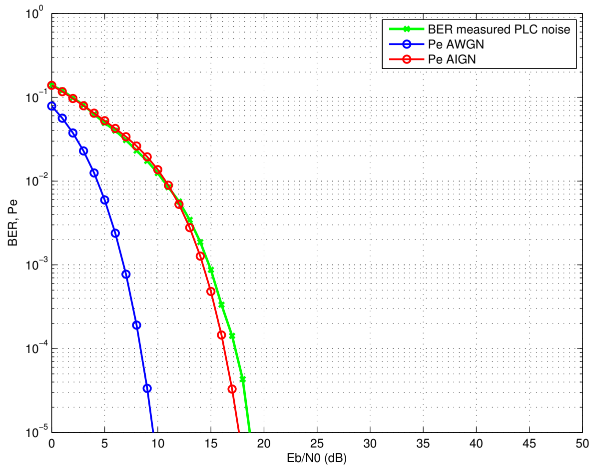

Fig. 6. Plots of P_e vs E_b/N_0 performance of two communication systems: (1) system which transmit a BPSK signal through an AWGN channel — lower curve (blue) and, (2) system which transmit a BPSK signal through an AIGN channel (PLC corrupted by BG noise with parameters $(\widehat{\alpha}, \widehat{p}) \approx$ (2.8, 0.73)) — red curve. Also shown, is the BER vs E_b/N_0 obtained by computer simulation — green curve (almost overlapped with the red curve).

Appendix I: Comments on equations (6) to (10).

Our development for equations (6) to (10) follows the rationale described next.

The probability of a wrong decision, which can be obtained by calculating the probability of the event $\widehat{U} \neq U$, in which $\{\widehat{U}\}$ is the sequence of decisions, is expressed by

$$\begin{aligned} P_d &= P[\{|\nu| > \ell; U = 0\} \cup \{|\nu| \leq \ell; U = 1\}] \\ &= P[\{|\nu| > \ell \mid U = 0\}]P[U = 0] + \\ &\quad P[\{|\nu| \leq \ell \mid U = 1\}]P[U = 1] \end{aligned} \tag{13}$$

Let us now use

$$f_W w = \mathcal{N}(w, \mu, \sigma^2) = \frac{1}{\sigma\sqrt{2\pi}} \exp\left(-\frac{(w-\mu)^2}{2\sigma_2}\right)$$

as the notation to represent the probability density functions (p.d.f.) of a Gaussian r.v. with zero expected value and variance σ_2. We have then for $\{|\nu| \mid U = 0\}$ and $\{|\nu| \mid U = 1\}$ which are Gaussian with $\sigma_0^2 = \alpha^2\sigma^2$ and $\sigma_1^2 = \sigma^2$, respectively, the conditional probability density functions (p.d.f.)

$$f_{\nu|U=u}(w) = \mathcal{N}(w, 0, \sigma^2).$$

Gathering these considerations to calculate the probability in (13) one get P_d in (6).

The value given in equation (9) is obtained by equating to zero the derivative of P_d in (6), with respect to $\lambda = \frac{\ell}{\sigma}$.

The probability $q_\ell = Pr[\nu > \ell]$ in (9), in which ν, the sum in (1), is a r.v., with p.d.f. simply given by the weighted sum of two Gaussian p.d.f.'s,

$$f_\nu(w) = p\mathcal{N}(w, 0, \sigma^2) + (1-p)\mathcal{N}(w, 0, \alpha^2\sigma^2). \qquad (14)$$

The expected in (10) follows trivially from the expression of the p.d.f. in (14).

APPENDIX II: FULL SEARCH ALGORITHM.

The Algorithm 1 describe the calculation needed to find the BG model of any stochastic process. As described the algorithm considers the samples of synthetically generated BG noise. Of course when N samples of a measured noise $n[k]$ are given the parameter of the model for this process is obtained.

Algorithm 1: Algorithm to find the BG parameters of measured PLC noise

initialization: Given the input signal $\nu[k]$ (with N samples):
$\sigma_\nu^2 = \frac{1}{N}\sum_{i=1}^{N} \nu[i]^2$;
$\widehat{M}_\ell$ = 0;
N_q = 0;
set:
p_{min};
p_{max};
α_{min};
α_{max};
begin
  for $p_{min} \le p \le p_{max}$ **do**
    for $\alpha_{min} \le \alpha \le \alpha_{min}$ **do**
      $\sigma^2 = \sigma_\nu^2/(p + \alpha^2(1-p))$;
      $\lambda = \alpha\sqrt{\frac{2\ln(\alpha p/1-p)}{\alpha^2-1}}$;
      $\ell_{OPT} = \lambda\ \sigma$;
      $M_\ell = \frac{2\sigma}{\sqrt{2\pi}}(p + \alpha(1-p))$;
      $q_\ell = 2(pQ(\lambda) + (1-p)Q(\lambda/\alpha))$;
      for $i = 1\ to\ N$ **do**
        $\widehat{M}_\ell = \widehat{M}_\ell + |\nu[k]|$;
        if $|\nu[k]| \ge \ell$ **then**
          $N_q = N_q + 1$;
        end
      end
      $\widehat{M}_\ell = \widehat{M}_\ell/N$;
      $\widehat{q}_\ell = N_q/N$;
      $E_M = \widehat{M}_\ell - M_\ell$;
      $E_q = \widehat{q}_\ell - q_\ell$;
    end
  end
end
For all the values (p, α) examined, the pair with $E_M(p, \alpha)$ and $E_q(p, \alpha)$ closest to zero is selected.
.

ACKNOWLEDGEMENTS

This work was funded in part by CNPq, INERGE, and Smarti9 Ltda.

REFERENCES

[1] M. Zimmermann and K. Dostert, "Analysis and modeling of impulsive noise in broad-band powerline communications," *IEEE Transactions on Electromagnetic Compatibility*, vol. 44, no. 1, pp. 249–258, Feb. 2002.

[2] R. Hormis, I. Berenguer, and X. Wang, "A simple baseband transmission scheme for power line channels," *IEEE Journal on Selected Areas in Communications*, vol. 24, no. 7, pp. 1351-1363, Jul., 2006.

[3] H. C. Ferreira, L. Lampe, J. Newbury, and T. G. Swart, "Power Line Communications: Theory and Applications for Narrowband and Broadband Communications Over Power Lines," Chichester, England: John Wiley and Sons, 2010.

[4] D. Middleton, *An introduction to statistical communication theory*. McGraw-Hill New York, 1960, vol. 960.

[5] K. Vastola, "Threshold detection in narrow-band non-Gaussian noise," *IEEE Transactions on Communications*, vol. 32, no. 2, pp. 134-139, Feb. 1984.

[6] T. Shongwe, A. J. Han Vinck, H. C. Ferreira, "On Impulse Noise and its Models," *18th IEEE International Symposium on Power Line Communications and Its Applications*, Glasgow, Scotland, UK. pp. 12-17, Mar. 30-Apr. 02, 2014.

[7] G. R. Colen, et alli "Setup de Medição para Caracterização de Redes de Distribuição de Energia Elétrica Externas para Sistemas PLC", (in portuguese) *XXI Seminário Nacional de Distribuição de Energia Elétrica*, Santos, SP, Brazil, 08-13 Nov., 2014.