ACADEMIC PRESS

Available at www.ComputerScienceWeb.com
POWERED BY SCIENCE DIRECT®

Digital Signal Processing 13 (2003) 284–300

Digital Signal Processing

www.elsevier.com/locate/dsp

Channel order and RMS delay spread estimation with application to AC power line communications ☆

Hongbin Li,[a,*] Duixian Liu,[b] Jian Li,[c] and Petre Stoica [d]

[a] *Department of Electrical and Computer Engineering, Stevens Institute of Technology, Hoboken, NJ 07030, USA*
[b] *Thomas J. Watson Research Center, International Business Machine Corporation, Yorktown Heights, NY 10598, USA*
[c] *Department of Electrical and Computer Engineering, University of Florida, Gainesville, FL 32611, USA*
[d] *Department of Systems and Control, Uppsala University, P.O. Box 27, SE-751 03, Uppsala, Sweden*

Abstract

AC power lines have been considered as a convenient and low-cost medium for intra-building automation systems. In this paper, we investigate the problem of estimating the channel order and root mean squared (RMS) delay spread associated with the power lines, which are channel parameters that provide important information for determining the data transmission rate and designing appropriate equalization techniques for power lines communications (PLC). We start by showing that the key to the RMS delay spread estimation problem is the determination of the channel order, i.e., the effective duration of the channel impulse response. We next discuss various ways to estimate the impulse response length from a noise-corrupted channel estimate. In particular, four different methods, namely a signal energy estimation (SEE) technique, a generalized Akaike information criterion (GAIC) based test, a generalized likelihood ratio test (GLRT), and a modified GLRT, are derived for determining the effective length of a signal contaminated by noise. These methods are compared with one another using both simulated and experimentally measured power line data. The experimental data was collected for power line characterization in frequencies between 1 and 60 MHz.

Keywords: Channel order estimation; RMS delay spread estimation; Estimation theory; Generalized likelihood ratio test; Power line communications

☆ This work was supported in part by the New Jersey Commission on Science and Technology, the Stevens Wireless Network Security (WINSEC), the National Science Foundation Grant MIP-9457388, the Office of Naval Research Grant N00014-96-0817, and the Swedish Foundation for Strategic Research through a Senior Individual Grant.

* Corresponding author.
E-mail address: hli@stevens-tech.edu (H. Li).

1051-2004/03/$ – see front matter
doi:10.1016/S1051-2004(02)00030-1

1. Introduction

Traditional communication carriers for home and building automation systems have been twisted pairs, coaxial cables, or optical fibers. However, the infrastructures of using such media are usually expensive to deploy. In contrast, AC power lines, which have already been installed in nearly every home and building and can be easily accessed via wall socket plugs, provide a convenient carrier for communications in home and building automation systems. The cost of such systems may be drastically reduced if power lines can be utilized in place of conventional data transmission media. In addition, power line communication (PLC) systems are subject to few regulatory issues. For these reasons, PLC has received much interest recently [1].

In spite of the above stated advantages, power lines do not represent a friendly environment for data transmission since they were not designed for such a purpose. Variable attenuation and impedance, impedance modulation, impulse noise, and continuous-wave jamming have been the major obstacles to reliable communications in PLC [1]. The recent applications of spread spectrum and forward error correction techniques to PLC have been quite successful in removing or alleviating the noise impediment [1–5]. Additionally, the consumer electronic bus (CEBus) has also been widely adopted, which provides a standard communication network protocol for the PLC industry [1,2,6]. As a result of these efforts, various PLC devices are on the market today.

Like any other communication systems, the performance of PLC is determined by a number of channel parameters, such as impedance, noise, signal attenuation, and root mean squared (RMS) delay spread. Extensive characterizations of the noise, impedance, and signal attenuation on power lines at frequencies up to 500 KHz are available in the literature (see, e.g., [7–9]). Most PLC devices currently on the market are also designed for this frequency range. These devices are usually good for applications with data rates lower than 100 Kbps. For video transmission and other similar applications, however, the data rates are in the range of Mbps, which corresponds to a much wider frequency band, typically in the range of 1–60 MHz. It is known that channel order and RMS delay spread are important parameters that affect the data transmission rate over the channel. A careful characterization of these parameters is therefore of particular importance for reliable high-speed data transmissions over power lines [10,11].

Following the general assumption for most communications channels, we herein assume that the PLC channel is a finite impulse response (FIR) filter [10,12,13]. The purpose of this study is to investigate the problem of channel order and RMS delay spread estimation for PLC. In practice, the PLC channel impulse response can be measured using some channel sounding technique, such as the impulse channel sounding described in Section 4.2. The measured channel impulse response can be considered as the true impulse response contaminated by measurement noise which does not vanish with time. With this observation, one needs to determine from the noise-contaminated channel estimate the *effective* duration of the impulse response, outside which the measurements are primarily attributed to the noise.

In this paper, we present a number of methods to solve the above effective signal length estimation problem. The first method under consideration is a signal energy estimation (SEE) based test. The idea of this method is to estimate and accumulate the signal energy

from each data sample until when the contribution from the remaining data measurements to the total signal energy becomes insignificant. This specific point in time is then identified as the terminating point of the signal. The second test makes use of the generalized Akaike information criterion (GAIC) [14], a statistical criterion that has been extensively used for model structure selection in system identification [15] and many other applications (see, e.g., [16]). As we will see, both SEE and GAIC involve parameters of user choice which may affect the performance of these tests, and yet it is unknown how to make such choices precisely in order to have the desired performance. To circumvent this difficulty, we derive a generalized likelihood ratio test (GLRT) to deal with the effective signal length estimation problem. Although GLRT has a more predictable performance, it is a valid test only when a large number of data samples are available. With limited data length, the performance of GLRT is also limited. The final method we consider for signal length estimation is a modified GLRT. The modified GLRT overcomes the drawbacks of the other three tests but retains their advantages. It is found that the modified GLRT works well in both simulated and experimental examples and may be the preferred method in practice.

In addition to PLC, the proposed algorithms can be used in other applications as well. For example, in orthogonal frequency division multiplexing (OFDM) communications systems, each OFDM symbol contains a *cyclic prefix* (CP) to combat channel dispersion [12,17]. In order to effectively remove the inter-symbol interference (ISI) at the receiver, the duration of the CP has to be longer than the impulse response of the channel. On the other hand, using the CP introduces a bandwidth and energy loss which are proportional to the length of the CP, implying that the length of the CP should be made as small as possible. Therefore, an accurate estimate of the channel order plays an important role in determining the length of the CP.

The remainder of this paper is organized as follows. Section 2 formulates the problem of interest. In Section 3 we derive the four proposed tests for effective signal length estimation. Numerical examples that compare the four tests are presented in Section 4. Finally, we summarize the study in Section 5.

2. Problem formulation

In the study, the impulse response of an AC power line $s(n)$ is assumed to have a finite duration M [10–13]. This assumption also allows the case where $s(n)$ decays fast enough such that beyond a certain point in time $s(n)$ becomes insignificant, i.e., $s(n) \approx 0$ for $n > M$. Given $s(n)$ and the (effective) signal length M, the associated mean delay and RMS delay spread normalized with respect to the sampling interval can be determined as

$$\mu = \frac{\sum_{n=1}^{M} n s^2(n)}{\sum_{n=1}^{M} s^2(n)}, \tag{1}$$

and, respectively,

$$\sigma_{\mathrm{RMS}} = \sqrt{\frac{\sum_{n=1}^{M} (n-\mu)^2 s^2(n)}{\sum_{n=1}^{M} s^2(n)}}. \tag{2}$$

In practice, the power line impulse response $s(n)$ is seldom known exactly. A typical way to estimate $s(n)$ is to measure the frequency response of the power line channel using some channel sounding technique and then apply the inverse Fourier transform to the estimated frequency response (see Section 4). The resulting signal $x(n)$ can be considered as an estimate of $s(n)$ contaminated by noise:

$$x(n) = s(n) + e(n), \quad n = 1, 2, \ldots, N, \tag{3}$$

where $e(n)$ denotes the estimation error which is modeled as a zero-mean white Gaussian noise with unknown variance σ_e^2 and is assumed to be independent of $s(n)$ [10,11], and N is chosen such that $N \gg M$. The problem of interest here is to estimate the channel order M and the RMS delay spread σ_{RMS} from the measurements $\{x(n)\}_{n=1}^{N}$.

Supposing first that M is known, we can estimate $\{s(n)\}_{n=1}^{M}$ by the maximum likelihood technique. Specifically, the negative log-likelihood function of $\{x(n)\}$ is (see, e.g., [18])

$$V_M = \frac{N}{2}\ln\sigma_e^2 + \frac{1}{2\sigma_e^2}\left\{\sum_{n=1}^{M}\big[x(n) - s(n)\big]^2 + \sum_{n=M+1}^{N} x^2(n)\right\} + \text{constant}. \tag{4}$$

The maximum likelihood estimates of $\{s(n)\}$ and σ_e^2 are obtained by minimizing the above cost function with respect to the unknown parameters, which yields

$$\hat{s}(n) = x(n), \quad n = 1, 2, \ldots, M, \tag{5}$$

and

$$\hat{\sigma}_e^2 = \frac{1}{N}\sum_{n=M+1}^{N} x^2(n). \tag{6}$$

Using (5) and (6) in (4) gives

$$\min_{\{s(n)\},\,\sigma_e^2} V_M = \frac{N}{2}\ln\hat{\sigma}_e^2 + \text{constant}. \tag{7}$$

If M is known, we can replace $s(n)$ in (1) and (2) by $\hat{s}(n)$ to obtain an estimate of the RMS delay spread. The remaining question is how to estimate the signal length M, which is discussed next.

3. Signal length estimation

In this section we derive four different tests to estimate the effective signal length M, namely the signal energy estimation (SEE) based test, the generalized Akaike information criterion (GAIC) based test, the generalized likelihood ratio test (GLRT), and the modified GLRT.

3.1. SEE based test

The SEE based test consists of choosing a sufficiently large L satisfying $M \leqslant L \leqslant N$ so that the noise variance can be estimated by using $\{x(n)\}_{n=L}^{N}$, calculating the total

signal energy based on the noise variance estimate, and using the signal energy estimate to determine the signal length. Specifically, let E_{LN} denote the total average energy of $\{x(n)\}_{n=L}^{N}$, that is,

$$E_{LN} \triangleq \sum_{n=L}^{N} E\big[x^2(n)\big] = (N-L+1)\sigma_e^2, \tag{8}$$

where $E[\cdot]$ denotes the expectation. The total noise average energy E_e is

$$E_e \triangleq \sum_{n=1}^{N} E\big[e^2(n)\big] = \frac{N}{N-L+1} E_{LN}. \tag{9}$$

Let E_x denote the total average energy of $x(n)$, i.e., $E_x \triangleq \sum_{n=1}^{N} E[x^2(n)]$. The total deterministic signal energy is obtained as $E_s \triangleq \sum_{n=1}^{N} s^2(n) = E_x - E_e$. In practice, E_x and E_{LN} can be estimated as $\widehat{E}_x = \sum_{n=1}^{N} x^2(n)$ and $\widehat{E}_{LN} = \sum_{n=L}^{N} x^2(n)$, respectively. It follows that an estimate of E_s is

$$\widehat{E}_s = \widehat{E}_x - \frac{N}{N-L+1}\widehat{E}_{LN}. \tag{10}$$

The proposed SEE test calculates the signal energy contribution at each data sample, which is $x^2(n) - \widehat{E}_e/N$. The signal energy estimate at each individual data sample is then accumulated and compared with the $\widehat{E}_s$ in (10). In more exact terms, the test consists of the following steps:

Step 1. Calculate $\widehat{E}_s$ using (10).

Step 2. Set $\breve{M} = 1$ and $\widehat{E}_s' = 0$.

Step 3. Compute $\widehat{E}_s' = \widehat{E}_s' + x^2(\breve{M}) - \widehat{E}_e/N$. Here, the updated $\widehat{E}_s'$ is the estimated total deterministic signal energy up to time index $\breve{M}$.

Step 4. If $\widehat{E}_s' \geqslant \kappa \widehat{E}_s$ or $\breve{M} = L$, then the signal length estimate $\widehat{M}_{\mathrm{SEE}}$ is equal to $\breve{M}$ and the test stops; otherwise, set $\breve{M} = \breve{M} + 1$ and go to Step 3. Here, κ is a parameter of user choice, typically $0.9 \leqslant \kappa \leqslant 0.99$.

It should be noted that the choice of L should be made with a trade-off in mind. Specifically, a small L may look appealing at a first sight since it implies that the noise variance estimate will be generated using more data samples and thus will have a better statistical property. On the other hand, a too small L may violate the condition $L \geqslant M$ and produce inaccurate noise and signal energy estimates, which will ultimately affect the signal length estimation accuracy.

As one can see, the SEE test is a method based on intuitive calculations of the signal and noise energies. It is simple but with a somewhat limited capability for detecting signal boundaries (see Section 4 for details). As such it is necessary to derive more sophisticated techniques that may yield better performance for signal length estimation.

3.2. GAIC based test

The generalized Akaike information criterion (GAIC) has been a popular statistical criterion for model structure selection in system identification [14,15]. We describe here how to adopt this criterion to determine the effective signal length M.

The GAIC cost function has the form [14,15]

$$\mathrm{GAIC}_{\breve{M}} = V_{\breve{M}} + \gamma \ln(\ln N)(\breve{M} + 1), \tag{11}$$

where (see (6) and (7))

$$V_{\breve{M}} = \frac{N}{2} \ln\left(\frac{1}{N} \sum_{n=\breve{M}+1}^{N} x^2(n) \right). \tag{12}$$

Here, $\breve{M}$ is assumed to be the signal length ($(\breve{M}+1)$ is thus the total number of unknown parameters for the data model in (3)), and γ is a parameter of user choice. The GAIC estimate $\widehat{M}_{\mathrm{GAIC}}$ of the true signal length M is obtained by minimizing the above cost function with respect to $\breve{M}$. It should be noted that the double logarithm $\ln(\ln N)$ in (11) gives the slowest rate at which the second term in the right-hand side of (11) should increase with N to guarantee the consistency of the signal length estimate obtained by minimizing (11) [19].

The proposed GAIC based test determines $\widehat{M}_{\mathrm{GAIC}}$ by the following steps:

Step 1. Choose a sufficiently large L so that $M \leqslant L \leqslant N$.

Step 2. Calculate the cost function $\mathrm{GAIC}_{\breve{M}}$ for $\breve{M} = 1, 2, \ldots, L$.

Step 3. The GAIC estimate of M is obtained as

$$\widehat{M}_{\mathrm{GAIC}} = \arg\min_{\breve{M}} \mathrm{GAIC}_{\breve{M}}, \quad \breve{M} = 1, 2, \ldots, L. \tag{13}$$

Remark 1. As one may have noticed, using either the SEE or GAIC based test for determining M involves user parameters, viz κ in SEE and the γ in GAIC, which may affect the accuracy of the signal length estimate, but whose choice is not easy. Specifically, making a choice of these parameters to achieve a certain probability of detection (or missing) is not really possible. It would be desirable to derive methods that can somehow control the risk of making a wrong decision. Such methods should be of greater interest in real applications.

3.3. GLRT

The generalized likelihood ratio for testing $M = \breve{M}$ against $M = \breve{M} + K$ (for some $K \geqslant 1$) is given by (see (6) and (7)) [15]

$$\begin{aligned} \Lambda &= N \ln\left[\frac{1}{N} \sum_{n=\breve{M}+1}^{N} x^2(n) \right] - N \ln\left[\frac{1}{N} \sum_{n=\breve{M}+K+1}^{N} x^2(n) \right] \\ &= N \ln\left[\frac{\sum_{n=\breve{M}+1}^{N} x^2(n)}{\sum_{n=\breve{M}+K+1}^{N} x^2(n)} \right]. \end{aligned} \tag{14}$$

For $N \gg 1$ and under the hypothesis

$$H_0: \quad \breve{M} \geqslant M, \tag{15}$$

it can be shown that Λ is χ^2 distributed with K degrees of freedom, denoted by

$$\Lambda \sim \chi^2(K). \tag{16}$$

To see this, we rewrite Λ as follows:

$$\Lambda = N \ln\left[1 + \frac{\sum_{n=\breve{M}+1}^{\breve{M}+K} x^2(n)}{\sum_{n=\breve{M}+K+1}^{N} x^2(n)}\right] \overset{N \gg \breve{M}+K}{\approx} N \frac{\sum_{n=\breve{M}+1}^{\breve{M}+K} x^2(n)}{\sum_{n=\breve{M}+K+1}^{N} x^2(n)}. \tag{17}$$

Let

$$\hat{\sigma}_e^2 = \frac{1}{N} \sum_{n=\breve{M}+K+1}^{N} x^2(n). \tag{18}$$

Here, $\hat{\sigma}_e^2$ is an estimate of σ_e^2. Note that for $N \gg 1$, $\hat{\sigma}_e^2 \approx \sigma_e^2$. In view of this observation, we have (under H_0)

$$\Lambda \approx \frac{1}{\sigma_e^2} \sum_{n=\breve{M}+1}^{\breve{M}+K} x^2(n) \sim \chi^2(K), \tag{19}$$

which proves (16).

The GLRT for determining the signal length estimate $\widehat{M}_{\mathrm{GLRT}}$ is summarized below:

Step 1. Choose a threshold λ from a table of the χ^2 distribution such that

$$\Pr\{y \leqslant \lambda \mid y \sim \chi^2(K)\} = \alpha, \tag{20}$$

where $0.9 \leqslant \alpha \leqslant 0.99$ (see the discussions below).

Step 2. Set $\breve{M} = 1$.

Step 3. Calculate Λ according to (14).

Step 4. If $\Lambda \leqslant \lambda$ at $\breve{M}$ and also $\Lambda \leqslant \lambda$ is true in more than $90\alpha\%$ of the cases corresponding to $\breve{M}+1, \breve{M}+2, \ldots, L-K$, then $\widehat{M}_{\mathrm{GLRT}} = \breve{M}$ and stop; otherwise, set $\breve{M} = \breve{M}+1$ and go to Step 3.

A brief explanation of Step 4 is as follows. For the reason discussed in Section 3.4, K is a small integer, typically $K \leqslant 10$. However, a small K may be a bad choice for signals that are small over some intervals within the signal duration, such as the sinc-like test signal used in Section 4. When $\breve{M}$ happens to be in one of those intervals and also K is too small to include any significant signal energy in the denominator of (17), it is very likely that the inequality $\Lambda \leqslant \lambda$ will be true. Hence, to find out the real signal boundary one has to check the inequality $\Lambda \leqslant \lambda$ not only at $\breve{M}$ but at the rest data samples as well. We shall keep in mind that even if the boundary sample has been hit, $\Lambda \leqslant \lambda$ may not be true for all of the rest data samples due to the random nature of the noise. Nevertheless, the inequality should hold true for the majority (e.g., $90\alpha\%$) of the rest data samples beyond the signal boundary.

Observe that the risk of rejecting H_0 when H_0 holds (the probability of false alarm) equals $1-\alpha$. In general, the risk of accepting H_0 when it is not true cannot be determined for the statistics introduced previously unless one restricts considerably the class of alternative hypotheses against which H_0 is tested. Thus, in applications the value of α

or, equivalently, the test threshold λ is chosen by considering only the probability of false alarm. Doing so, we shall keep in mind that as α increases, the probability of false alarm decreases, but the other type of risk increases. Typically, α is chosen between 0.9 and 0.99 [15].

Remark 2. It should be noted that the above GLRT is a valid test only when $N \to \infty$. Additionally, $\hat{\tilde{\sigma}}_e^2$ is a poor estimate of σ_e^2 if N is not large enough, particularly so if $\breve{M} + K + 1 < M$. It would be of interest to modify the GLRT somehow such that the above problems are avoided. Such a modified GLRT indeed exists, as discussed next.

3.4. Modified GLRT

As mentioned in the previous section, $\hat{\tilde{\sigma}}_e^2$ in (18) is usually not a good estimate of σ_e^2. A better estimate is

$$\hat{\tilde{\sigma}}_e^2 = \frac{1}{N-L} \sum_{n=L+1}^{N} x^2(n), \tag{21}$$

where $M \leqslant L \leqslant N$. We now replace the $\hat{\tilde{\sigma}}_e^2$ in (19) by the above $\hat{\tilde{\sigma}}_e^2$ and define

$$\Delta \triangleq \frac{N-L}{K} \cdot \frac{\sum_{n=\breve{M}+1}^{\breve{M}+K} x^2(n)}{\sum_{n=L+1}^{N} x^2(n)} \triangleq \frac{N-L}{K} \cdot \frac{\rho_1}{\rho_2}. \tag{22}$$

Under the hypothesis H_0, we have $\rho_1/\sigma_e^2 \sim \chi^2(K)$ and $\rho_2/\sigma_e^2 \sim \chi^2(N-L)$, respectively. Moreover, if $\breve{M} + K \leqslant L$, ρ_1 and ρ_2 are independent of one another. It follows that under the above conditions, Δ is F distributed with K and $N-L$ degrees of freedom [15], written as

$$\Delta \sim F(K, N-L). \tag{23}$$

Observe that (23) holds in finite samples, whereas most other tests, including the original GLRT, require $N \to \infty$.

The choice of K should be made carefully. For $\breve{M} \geqslant M$, this is perhaps not very important since any $K \geqslant 1$ will lead to a similar performance. For $\breve{M} < M$, however, the choice becomes more critical. To reduce the risk of underestimating M, a small K is recommended. To see this, let us assume that K is very large such that $K \gg M$. Then underestimating M by 1 or 2 will not increase ρ_1 too much (particularly so if M is small), and hence the risk of underestimating M may be large. As a result, a smaller K in this case should be used. However, as mentioned in Section 3.3, a small K is a bad choice for signals that are small over certain intervals within the signal duration. In a way similar to what we adopted there, we recommend choosing $K \leqslant 10$, computing Δ for $\breve{M} = 1, 2, \ldots, L-K$, and, finally, determining $\widehat{M}$ as the $\breve{M}$ at and beyond which the inequality $\Delta \leqslant \delta$ holds in more than $90\alpha\%$ of the remaining cases corresponding to $\breve{M}+1, \breve{M}+2, \ldots, L-K$. Here, δ is a threshold determined such that

$$\Pr\{z \leqslant \delta \mid z \sim F(K, N-L)\} = \alpha, \tag{24}$$

and α is between 0.9 and 0.99.

To sum up, the modified GLRT determines $\widehat{M}_{\mathrm{mGLRT}}$ in the following steps:

Step 1. Choose $K \leqslant 10$ and a threshold δ from a table of the F distribution so that (24) is satisfied.

Step 2. Calculate Δ for $\breve{M} = 1, 2, \ldots, L - K$.

Step 3. The signal length estimate $\widehat{M}_{\mathrm{mGLRT}}$ is the smallest $\breve{M}$ at which $\Delta \leqslant \delta$ is true and for which the inequality is also true in more than $90\alpha\%$ of the cases corresponding to $\breve{M} + 1, \breve{M} + 2, \ldots, L - K$.

4. Numerical results

We compare the performance of the four tests described in the previous section for effective signal length estimation. Both simulated and experimental data are used for comparison. In the following, we use $\kappa = 0.96$ for the SEE test, $\gamma = 2$ for GAIC, and $\alpha = 0.99$ for GLRT and the modified GLRT (referred to as mGLRT henceforth).

4.1. Simulated examples

The simulated data consists of a pulse having a certain shape corrupted by a zero-mean white Gaussian noise with variance σ_e^2. We consider both a rectangular pulse ($M = 40$) and a sinc-like pulse having a raised cosine spectrum. The roll-off factor for the latter is 1. The sinc-like pulse is shifted and truncated to have a duration of 80 samples. Figure 1 shows a realization of the test data corresponding to the two different pulses when $\sigma_e^2 = 0.05$, where dashdot lines represent noise-free signals and solid lines denote noise-contaminated signals, respectively. The results shown below are obtained using 200 Monte Carlo trials. For each individual trial, a total number of $N = 450$ samples are used and $L = 200$.

In the first example, we investigate the effect of the noise variance on the performance of the proposed tests. Based on our discussion in the previous section, the parameter K used in GLRT and mGLRT is suggested to be a small number ($K \leqslant 10$). Here we choose $K = 4$. (In the next example, we demonstrate numerically how different values of K affect the performance.) The results are shown in Fig. 2 for the rectangular pulse and Figs. 3a–3c for the sinc-like pulse.

For the rectangular pulse, the signal length is $M = 40$ and the associated RMS delay can be calculated using (1) and (2) as $\sigma_{\mathrm{RMS}} = 11.54$. Figure 2 shows the root mean squared errors (RMSE) of the RMS delay spread estimates as a function of the noise variance σ_e^2. In our simulation, we have observed that a number of signal length estimates obtained by GLRT are equal to L. Such an estimate is called an *outlier*. We did not weed out the outliers in the calculation of the empirical statistics of the parameter estimates since the other methods do not suffer from this problem. Due to the outliers, we can see that the performance of GLRT is considerably affected. The results in Fig. 2 also suggest that the other three tests perform similarly for a σ_e^2 up to 0.01. As the signal becomes more noisy, the performance of GAIC and mGLRT remains relatively unaffected, whereas SEE degrades substantially. We have also calculated the empirical mean and standard deviation of the signal length estimates for the rectangular pulse. The results demonstrate a similar pattern as in Fig. 2 and thus are not included here.

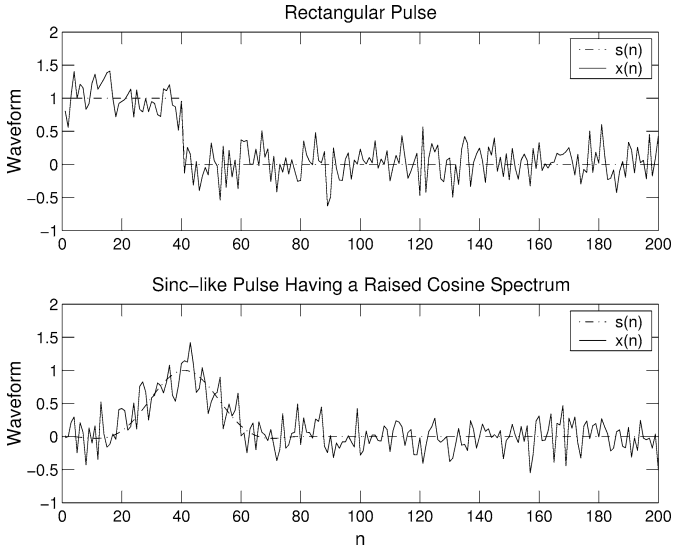


Fig. 1. Tests signals used in the simulated examples.

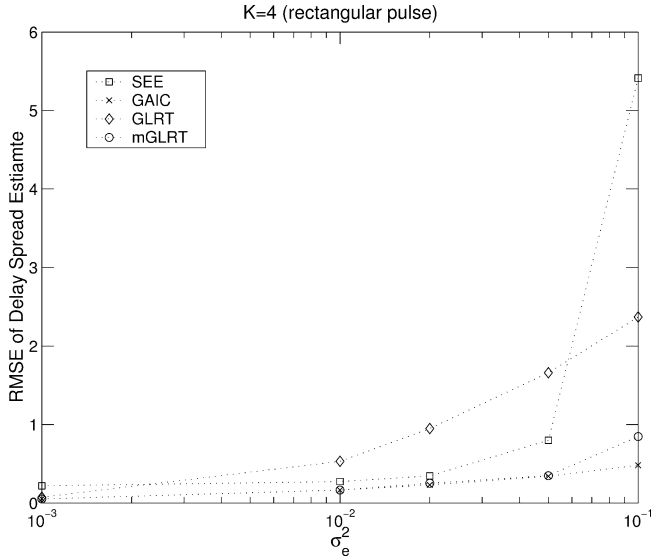
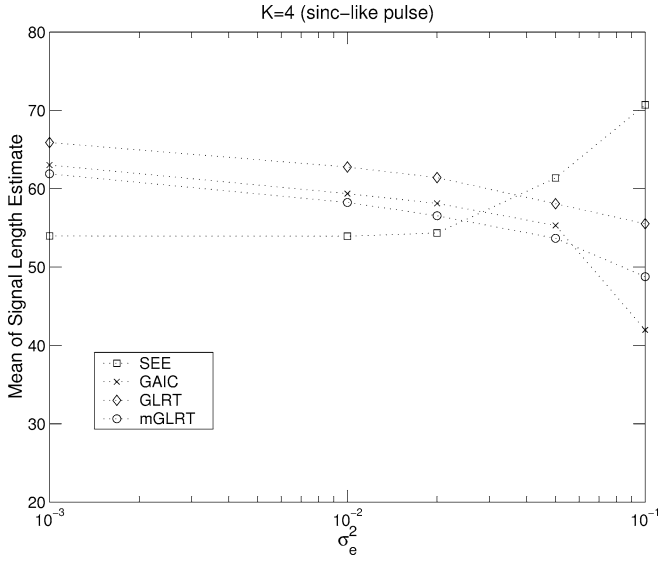
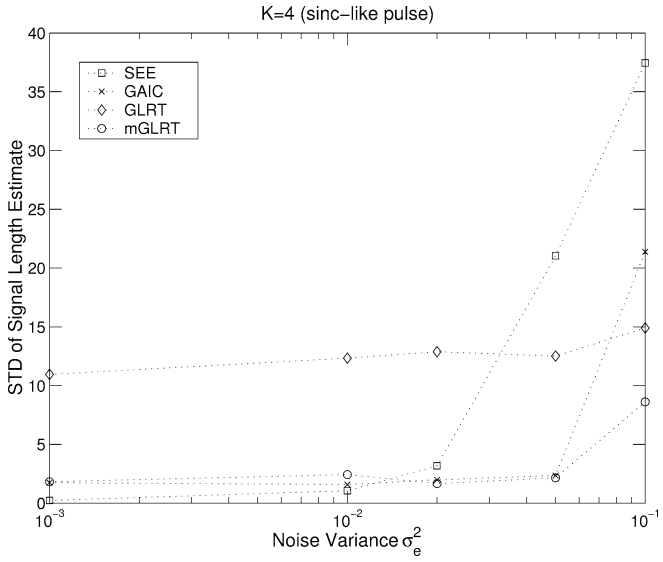


Fig. 2. RMSE of $\hat{\sigma}_{\mathrm{RMS}}$ for the rectangular pulse versus σ_e^2 when $N = 450$, $L = 200$, and $K = 4$.

For the sinc-like pulse, the situation is more subtle since the magnitude of the pulse decreases rapidly as n increases. Specifically, the magnitude for $n \geqslant 60$ is more than 15 dB smaller than the peak magnitude. Hence, it may not be appropriate to determine the *effective* length to be $M = 80$ even though the signal is truncated to zero for $n \geqslant 80$. Nevertheless, the RMS delay spread for this signal is easily calculated to be $\sigma_{\mathrm{RMS}} = 7.21$.



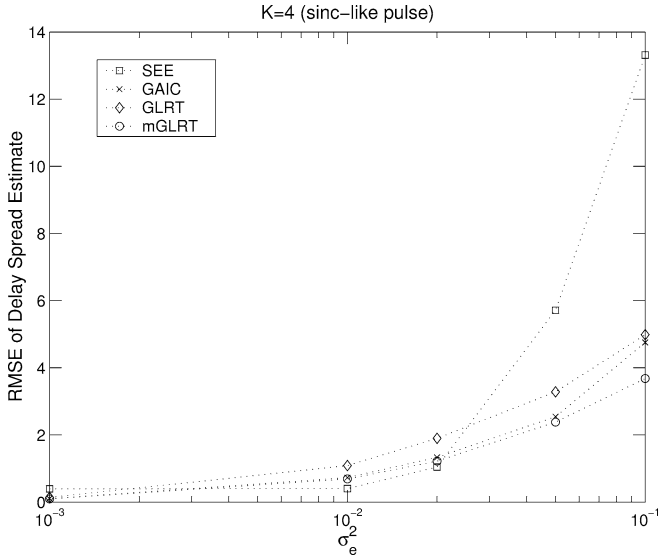
(a)



(b)

Fig. 3. Empirical statistics of the parameter estimates for the sinc-like pulse versus σ_e^2 when $N = 450$, $L = 200$, and $K = 4$. (a) Mean of $\widehat{M}$. (b) Standard deviation of $\widehat{M}$. (c) RMSE of $\hat{\sigma}_{\mathrm{RMS}}$.

Figs. 3a–3c show the empirical mean and standard deviation of the signal length estimates and the RMSE's of the RMS delay spread estimates for the sinc-like pulse. We can see that SEE in the current case appears to underestimate the signal length when σ_e^2 is small but



(c)

Fig. 3. (Continued.)

overestimate the signal length when σ_e^2 is large. The mean and standard deviation of the signal length estimates and the RMSE of the delay spread estimates obtained by GLRT are larger than those by GAIC and mGLRT, which is due to GLRT outliers. It is observed that GAIC and mGLRT in general yield similar performance. A noticeable difference between the two tests occurs at $\sigma_e^2 = 0.1$, for which the latter outperforms the former slightly.

Next, we present a numerical example to show the effect of choosing different K on the performance of GLRT and mGLRT. The setting is similar to the previous example except that σ_e^2 is fixed at some value and K is varied from 1 to 20. We consider both a moderate and a noisy scenarios, corresponding to $\sigma_e^2 = 0.01$ and $\sigma_e^2 = 0.1$, respectively. To reduce the number of figures, we only consider the sinc-like pulse here. (Similar conclusions can be drawn for the rectangular waveform.) The RMSE's of the RMS delay estimates versus K are shown in Fig. 4. We observe that the choice of $K = 1$ obtains the best performance for both tests when $\sigma_e^2 = 0.01$. However, for noisy situation corresponding to $\sigma_e^2 = 0.1$, a larger K is recommended ($K \geqslant 4$). It is also seen that GLRT in general degrades significantly as K increases, while the performance of mGLRT remains relatively unchanged for a wider range of K. Hence, the choice of K for mGLRT is easier to make than for GLRT. We do not recommend a large K (such as $K > 10$) for mGLRT since increasing K leads to an increased computational complexity of mGLRT (see (22)).

4.2. Experimental example

We now consider an experimental example. We first briefly describe the PLC channel sounding system used to obtain the measurement data. For more details of the system and measurement process, we refer the interested readers to [10]. Figure 5 shows a block

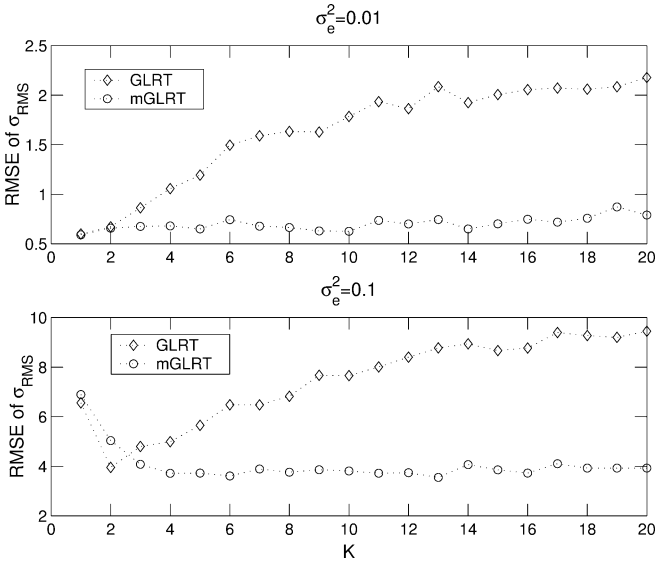


Fig. 4. RMSE of $\hat{\sigma}_{\mathrm{RMS}}$ versus K when $N = 450$ and $L = 200$ (upper figure corresponds to $\sigma_e^2 = 0.01$ and lower one corresponds to $\sigma_e^2 = 0.1$).

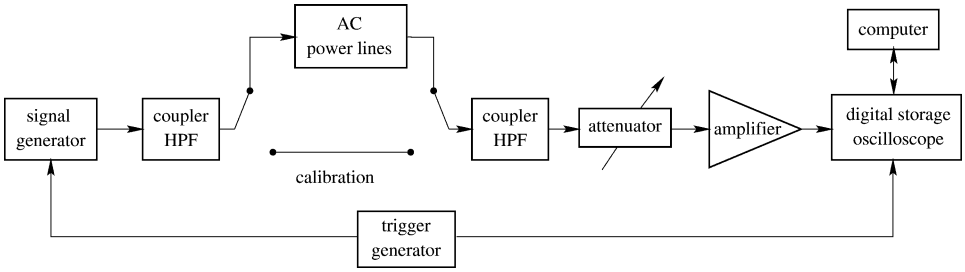


Fig. 5. Power line channel measurement system.

diagram that uses impulse channel sounding to measure the impulse response of the AC power line channel. The coupler box plugging into the AC wall outlet (the top path in Fig. 5) behaves like a highpass filter, shown in Fig. 6, with the 3 dB cutoff at 1 MHz. The probing signal passes through the coupler and the AC power line network and exits through a similar coupler plugged in a different outlet. A direct coupler to coupler connection is used to calibrate the test setup (the bottom path in Fig. 6). A low-noise amplifier (LNA) with at least 54 dB gain is used in front of the digital storage oscilloscope (DSO) to reduce the noise figure and increase the sensitivity of the system. The LNA has a built-in lowpass filter with the 3 dB cutoff frequency at 70 MHz in the front stage. Additionally, a high-precision adjustable (0–80 dB) attenuator is placed after the receiving coupler, making it possible to center the dynamic range of the LNA/DSO combination for the signal level of each outlet pair. This allows the system to capture noise spikes and temporal noise

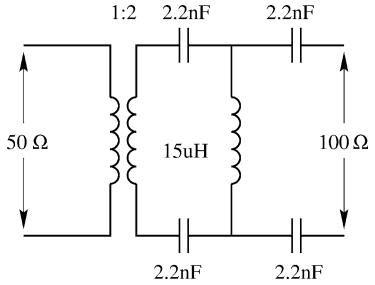


Fig. 6. Coupler circuit.

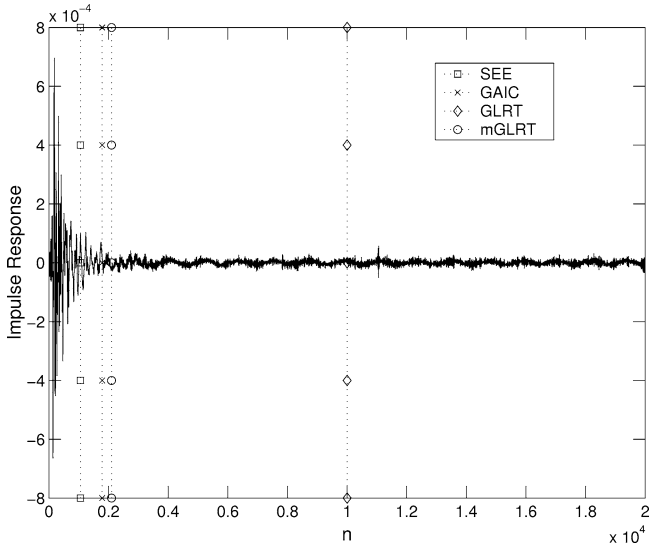


Fig. 7. Impulse response of the power line channel (1–60 MHz) and the corresponding effective signal length estimates.

fluctuations. The DSO has a bandwidth of 500 MHz, implying a high resolution, and the capability for long time captures.

The probing impulse used is a specially truncated sinc waveform, with a duration of 17 ns and a flat frequency characteristics from 0.85 to 63.6 MHz. The highpass characteristics of the couplers and the lowpass filter in the LNA limit the receiving sensitivity of the system to the 1 to 60 MHz frequency band. The sampling frequency is 1 GHz and the total number of data samples is $N = 20000$. The measurements were performed at two residential houses by averaging over 100 to 1000 scope sweeps depending on the noise situation. Figure 7 shows the impulse response of a specific power line channel corresponding to the frequency band 1–60 MHz. For channel order and RMS delay spread estimation, we choose $L = N/2$. The effective signal length estimates obtained by the four tests under study are also shown in the figure. We notice that GLRT fails again since the GLRT estimate is an outlier, with a value equal to L. It is also seen that SEE

obviously underestimates the effective signal length. On the other hand, the estimates given by GAIC and mGLRT appear to be more accurate. After obtaining the effective signal length estimate, we can use (1) and (2) to calculate the mean delay and RMS delay spread. Specifically, the RMS delay spread estimates for the 1–60 MHz frequency band obtained by SEE, GAIC, and mGLRT are 0.19, 0.27, and 0.28 μs, respectively. With no equalization, the maximum transmission rate is inversely proportional to the RMS delay spread:

$$\text{maximum transmission rate} \approx \frac{1}{2\sigma_{\text{RMS}}}. \tag{25}$$

It follows that the maximum data transmission rate is approximately 2.63 Mbps. The above calculation is somewhat optimistic since other factors, such as attenuation and noise characteristics of the PLC channel, which are important in determining the transmission rate, were not counted. Additionally, the impulse responses were obtained using one specific set of measurements. It is our experience that the RMS delay spread could vary significantly depending on the loads and environment of the power lines networks.

5. Conclusions

In this paper, we have examined the problem of channel order and RMS delay spread estimation for power line communications (PLC). It has been shown that the critical parameter is the channel order since the delay spread is readily calculated once the former is available. We have presented four different methods, namely the signal energy estimation (SEE) based test, the generalized Akaike information criterion (GAIC) based test, the generalized likelihood ratio test (GLRT), and the modified GLRT, to solve this problem. The performance of these tests has been compared using both simulated and experimental data. The experimental data has been collected to characterize the power line channel in the frequency range of 1–60 MHz. Our results have shown that the GAIC based test and the modified GLRT in general perform better than the other two tests; furthermore, the modified GLRT may be preferred to the GAIC based test because the former has more control over the performance in terms of the probability of detection, and is thus more convenient to use.

References

[1] D. Radford, Spread-spectrum data leap through ac power wiring, IEEE Spectrum (1996) 48–53.
[2] M.H. Shwehdi, A.Z. Khan, A power line data communication interface using spread spectrum technology in home automation, IEEE Trans. Power Delivery 11 (3) (1996) 1232–1237.
[3] D.E. Dodds, J.S. Hanson, Power line data communication using timed transmission, in: Proceedings of the Canadian Conference on Electrical and Computer Engineering, Vol. 1, 1984, pp. 280–283.
[4] R. Gershon, D. Propp, M. Propp, A token passing network for power line communications, IEEE Trans. Consumer Electron. 37 (1991) 129–134.
[5] R.A. Plety, Intrabuilding data transmission using power line wiring, Hewlett–Packard J. (1987).
[6] P. House, CEBus for the masses, Circuit Cellar INK (1995).
[7] J.R. Nicholson, J.A. Malack, RF impedance of power lines and line impedance stabilization networks in conducted interference measurements, IEEE Trans. Electromagnet. Compatibility 15 (1973) 84–86.

[8] R.M. Vines, H.J. Trussell, L.J. Gale, J.B. O'Neal Jr., Noise on residential power distribution circuits, IEEE Trans. Electromagnet. Compatibility 26 (4) (1984) 161–168.
[9] M.H.L. Chan, R.W. Donaldson, Attenuation of communication signals on residential and commercial intrabuilding power-distribution circuits, IEEE Trans. Electromagnet. Compatibility 28 (4) (1986) 220–230.
[10] D. Liu, E. Flint, B. Gaucher, Y. Kwark, Wide band AC power line characterization, IEEE Trans. Consumer Electron. (1999).
[11] D. Liu, E. Flint, B. Gaucher, Y. Kwark, High-speed communications system performance using the home AC power line, Technical report No. RC21604(97420), IBM T.J. Watson Research Center, 1999.
[12] J.G. Proakis, Digital Communications, 3rd ed., McGraw–Hill, 1995.
[13] H. Philipps, Performance measurements of powerline channels at high frequencies, in: Proceedings of 1998 International Symposium on Power-Line Communications and Its Applications, Tokyo, Japan, 1998.
[14] P. Stoica, P. Eykhoff, P. Janssen, T. Söderström, Model structure selection by cross-validation, Internat. J. Control 43 (6) (1986) 1841–1878.
[15] T. Söderström, P. Stoica, System Identification, Prentice Hall International, London, 1989.
[16] J. Li, P. Stoica, Efficient mixed-spectrum estimation with applications to target feature extraction, IEEE Trans. Signal Process. 44 (2) (1996) 281–295.
[17] T. Peled, A. Ruiz, Frequency domain data transmission using reduced computational complexity algorithms, in: Proceedings of the International Conference on Acoustics, Speech, and Signal Processing, Denver, CO, 1980, pp. 964–967.
[18] P. Stoica, R.L. Moses, Introduction to Spectral Analysis, Prentice Hall, Upper Saddle River, NJ, 1997.
[19] E.J. Hannan, Estimating the dimension of a linear system, J. Multivariate Anal. 11 (1981) 459–473.

Hongbin Li received the B.S. and M.S. degrees from the University of Electronic Science and Technology of China (UESTC), Chengdu, in 1991 and 1994, respectively, and the Ph.D. degree from the University of Florida, Gainesville, FL, in 1999, all in electrical engineering. From July 1996 to May 1999, he was a Research Assistant in the Department of Electrical and Computer Engineering at the University of Florida. Since July 1999, he has been an Assistant Professor in the Department of Electrical and Computer Engineering, Stevens Institute of Technology, Hoboken, NJ. His current research interests include stochastic signal processing, sensor array processing, wireless communications, and radar imaging. Dr. H. Li is a member of the IEEE, Tau Beta Pi, and Phi Kappa Phi. He received the Jess H. Davis Memorial Award for excellence in research from Stevens Institute of Technology in 2001, and the Sigma Xi Graduate Research Award from the University of Florida in 1999.

Duixian Liu received his B.S. degree in electrical engineering from XiDian University in Xi'an, China, in 1982, M.S. and Ph.D. degrees in electrical engineering from the Ohio State University in 1986 and 1990, respectively. From 1990 to 1996, he worked for Valor Enterprises Inc. in Piqua, OH, as the chief engineer. He has been with IBM at Thomas J. Watson Research Center since April 1996. His research interests are antenna design, electromagnetic theory and scattering, and digital signal processing.

Jian Li received the M.Sc. and Ph.D. degrees in electrical engineering from The Ohio State University, Columbus, in 1987 and 1991, respectively. From April 1991 to June 1991, she was an Adjunct Assistant Professor with the Department of Electrical Engineering, The Ohio State University, Columbus. From July 1991 to June 1993, she was an Assistant Professor with the Department of Electrical Engineering, University of Kentucky, Lexington. Since August 1993, she has been with the Department of Electrical and Computer Engineering, University of Florida, Gainesville, where she is currently a Professor. Her current research interests include spectral estimation, array signal processing, and signal processing for wireless communications and radar. Dr. J. Li is a member of Sigma Xi and Phi Kappa Phi. She received the 1994 National Science

Foundation Young Investigator Award and the 1996 Office of Naval Research Young Investigator Award. She is currently an Associate Editor for IEEE Transactions on Signal Processing and a Guest Editor for Multidimensional Systems and Signal Processing. She is presently a member of the Signal Processing Theory and Methods (SPTM) Technical Committee of the IEEE Signal Processing Society and an Executive Committee Member of the 2002 International Conference on Acoustics, Speech, and Signal Processing, Orlando, FL, May 2002.

Petre Stoica is a Professor of system modeling in the Department of Systems and Control at Uppsala University (UU) in Sweden. Previously he was a Professor of signal processing at the Bucharest Polytechnic Institute (BPI). He received a D.Sc. degree from the BPI in 1979 and a Honorary Doctorate from UU in 1993. He held longer visiting positions with the Eindhoven University of Technology, Chalmers Institute of Technology, Uppsala University, University of Florida, and Stanford University. His main scientific interests are in the areas of system identification, time series analysis and prediction, statistical signal and array processing, spectral analysis, wireless communications, and radar signal processing. He has published 7 books, 8 book chapters and some 370 papers in archival journals and conference records on these topics. His most recent book is "Introduction to Spectral Analysis" (with R. Moses), Prentice Hall, 1997. He was co-recipient of the 1989 ASSP Senior Award and recipient of the 1996 Technical Achievement Award of the IEEE Signal Processing Society. He is on the Editorial Boards of five journals in the field: Signal Processing; Journal of Forecasting; Circuits, Systems and Signal Processing; Multidimensional Systems and Signal Processing; and Digital Signal Processing: A Review Journal. He was guest co-editor for several special issues on system identification, spectral analysis and radar. He is a Honorary Member of the Romanian Academy, a Fellow of the Royal Statistical Society, and a Fellow of IEEE.