2019 1st Global Power, Energy and Communication Conference (IEEE GPECOM2019), June 12-15, 2019, Cappadocia, Turkey

Estimation of Gaussian Processes in Markov-Middleton Impulsive Noise

Anoush Mirbadin, Ehsan Kiani, Armando Vannucci, and Giulio Colavolpe
Department of Engineering and Architecture, University of Parma, 43124 Parma, Italy
(email: anoush.mirbadin@studenti.unipr.it)

***Abstract*—This work addresses the estimation of Gaussian signals over power line channels which are impaired by impulsive noise. The Markov-Middleton model is used to describe the memory and the multi-interferer nature of the impulsive noise. The estimation of Gaussian samples has been obtained by using a message passing algorithm. The message passing approach involves estimation of the channel states, approximation of the Gaussian mixtures and estimation of the correlated Gaussian samples. Correlation of channel states and correlation of input samples results in a loopy factor graph. To implement message passing on a loopy factor graph, we divide the graph in two main parts that exchange their messages by using a parallel iterative schedule. The lower part detects the channel states using the Bahl-Cocke-Jelinek-Raviv (BCJR) algorithm and the upper part estimates the signal samples using a Kalman smoother. The proposed approach extensively reduces the complexity of the overall estimation process.**

***Index Terms*—power line communications, impulsive noise, Markov-Middleton model, Bayesian estimation, message passing algorithms**

I. Introduction

Power Line Communications (PLC) enable broadband data transmission over existing electrical power infrastructures. One of the main limitations of the PLC systems is the channel noise which is not a simple Additive White Gaussian Noise (AWGN). Besides a stationary background noise component, the dominant noise in PLC is the impulsive noise [1], [2]. Quick changes in the state of the electrical devices and sudden connection and disconnection from power supplies are the main sources of the impulsive noise. Moreover, multicarrier modulation schemes are widely used in PLC [3], [4]. Transmission of data symbols in a multicarrier modulation system can simply be modeled by the transmission of analog Gaussian signals. Therefore, a meaningful investigation of a PLC system can be obtained through an accurate modeling of the impulsive noise and by devising an optimal estimation strategy, which is the main objective of this work. Previous attempts to model impulsive noise can be summarized as follows. The simplest model was the Bernoulli-Gaussian, in which it was assumed that impulsive noise has a Gaussian distribution with variance larger than that of background noise [5]. The model did not consider noise memory. In [6], Fertonani and Colavolpe proposed the Markov-Gaussian model which reflects the memory of the channel by considering the Markovianity of noise samples. The first report on Middleton class A noise was [7]. The model states that channel noise is a combination of independent and zero-mean Gaussian distributions. Ndo *et al.* [8] proposed a Hidden Markov Model (HMM) to describe the Middleton class A noise and the Markovianity of noise samples. Previous works on estimation of Gaussian signals in the presence of impulsive noise use the above models. In [9], Banelli derived the Minimum Mean Square Error (MMSE) Optimal Bayesian Estimation (OBE) to estimate Gaussian signals in the Middleton class A noise. Alam *et al.* [10] proposed the joint MMSE OBE of Gaussian symbols and detection of correlated channel states in the Markov-Gaussian impulsive noise. Vannucci *et al.* [11] proposed a novel factor graph based approach for joint estimation of correlated Gaussian symbols and detection of correlated channel states in a Markov-Gaussian scenario.

The main subject of this work is to investigate data transmission over PLC links, impaired by impulsive noise. To this end, we adopt the HMM proposed in [8] to model the impulsive noise and the same estimation approach as in [11] to estimate correlated Gaussian samples in the presence of Markov-Middleton class A impulsive noise, thus extending the results in [11], that are limited to a binary Markov-Gaussian channel, to a more realistic impulsive noise scenario. We focus our attention from a minimum of 4 to a maximum of 16 noise states, i.e., 3 to 15 possible sources of impulsive noise. It should be noted that the 2-state noise model is fully analyzed in [11]. Moreover, it is expected that the results would not considerably change for higher noise states than 16-state noise model. Approximate inference algorithms can be applied to the factor graph that models the overall system, in order to devise an optimal receiver. Dealing with a loopy factor graph, it is convenient to split the graph in two subgraphs and to establish a parallel iterative schedule between the two halves. One subgraph detects the channel states, through the Bahl-Cocke-Jelinek-Raviv (BCJR) algorithm, and the other subgraph estimates the correlated signal samples, through a Kalman smoother.

II. Channel Model

A sequence of correlated Gaussian samples $\{s_k\}_{k=0}^{K-1}$ are obtained by passing a white noise process through a single pole Infinite Impulse Response (IIR) digital filter, which forms an autoregressive model of order one (AR(1))

$$s_k = a_1 s_{k-1} + \omega_k \tag{1}$$

where $\{\omega_k\}$ is a sequence of i.i.d. Gaussian noise samples, $\omega_k \sim N(\eta_\omega, \sigma_\omega^2)$, and a_1 determines the pole of the filter. The

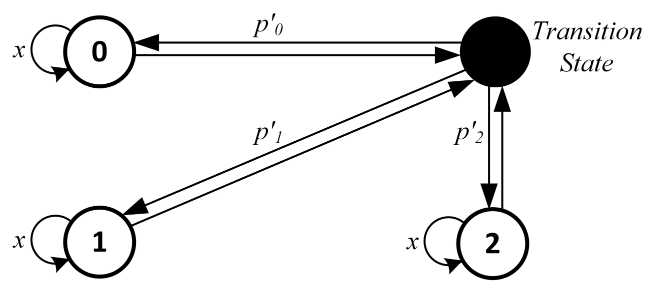

Fig. 1. Three-state Markov-Middleton noise model.

mean and variance of a signal sample s_k can consequently be calculated as $\eta_s = \eta_\omega/(1-a_1)$ and $\sigma_s^2 = \sigma_\omega^2/(1-a_1^2)$, respectively. The correlated Gaussian samples are transmitted over a Middleton Class A noise channel [7], so that the received samples can be described by

$$y_k = s_k + n_k \quad (k = 0, 1, \cdots, K-1) \tag{2}$$

in which n_k is a real-valued noise random variable whose probability density function (pdf) can be written as

$$p(n_k) = \sum_{i=0}^{\infty} \frac{p_i}{\sqrt{2\pi\sigma_i^2}} \exp\left\{-\frac{\mathrm{n_k^2}}{2\sigma_\mathrm{i}^2}\right\} \tag{3}$$

In (3), p_i represents the probability of being in the i-th noise state and has a Poisson distribution,

$$p_i = \frac{e^{-A}A^i}{i!} \tag{4}$$

in which A is the impulsive index and equals the average number of active impulsive sources per unit time; σ_i^2 is the variance of the overall impulsive noise component and can be calculated as

$$\sigma_i^2 = \left(1 + \frac{i}{A\Gamma}\right)\sigma_0^2 \tag{5}$$

which implies that the total variance of noise is equal to the variance of the background Gaussian noise, σ_0^2, plus the variance of the impulsive noise, $\frac{i}{A\Gamma}\sigma_0^2$. The physical meaning of Γ stems from the expectation of (5),

$$\sigma^2 = E[\sigma_i^2] = \sum_{i=0}^{\infty}\sigma_i^2 p_i = \left(1 + \frac{1}{\Gamma}\right)\sigma_0^2 \tag{6}$$

Equation (6) states that the average power of the noise is equal to the average power of background Gaussian noise plus the average power of impulsive noise. Therefore, Γ is the ratio of the average power of background noise to the average power of impulsive noise. Suppose that i impulsive sources are simultaneously active at a given time. In this case, the Middleton class A noise can be seen by the receiver as a single Gaussian distributed sequence with larger variance, σ_i^2.

Now we may consider the channel memory. To this end, we use the proposed HMM by Ndo *et al.* [8] which consists of the Middleton class A noise model parameters, A, Γ,σ^2, and a correlation parameter, which reflects the channel memory. The pdf of the Middleton class A noise has an infinite number of terms. An approximation can be made by considering a maximum number of impulsive sources $M-1$. Truncating (3), the pdf of the M-state Middleton class A noise can be written as

$$p(n_k) = \sum_{i=0}^{M-1} \frac{p_i'}{\sqrt{2\pi\sigma_i^2}} \exp\left\{-\frac{\mathrm{n_k^2}}{2\sigma_\mathrm{i}^2}\right\} \tag{7}$$

where, consistent with [8]

$$p_i' = \frac{p_i}{\sum_{i=0}^{M-1} p_i} \tag{8}$$

is set for normalization. The corresponding HMM state diagram is shown in Fig. 1 for $M = 3$. The correlation parameter x can be interpreted as the probability of remaining in the same state. The transition matrix $\boldsymbol{P}$ for the corresponding HMM can be written as

$$\boldsymbol{P} = \begin{bmatrix} x+(1-x)p_0' & (1-x)p_1' & (1-x)p_2' \\ (1-x)p_0' & x+(1-x)p_1' & (1-x)p_2' \\ (1-x)p_0' & (1-x)p_1' & x+(1-x)p_2' \end{bmatrix} \tag{9}$$

The average number of consecutive samples staying in a given noise state i can be derived from (9) as

$$T_i = \frac{1}{(1-x)(1-p_i')} \tag{10}$$

which thus represents the average duration of an impulsive noise event due to i active impulsive sources.

III. Estimation Strategy

A. Drawing the Factor Graph

In this section, we describe our strategy to estimate correlated Gaussian samples in the presence of impulsive noise with memory, as modeled by the above Markov-Middleton model. The proposed estimation method by Vannucci *et al.* in [11] is employed in this work. A traditional approach to signal estimation can be performed by first detecting the noise state at every time slot. On the contrary, for an optimal signal estimation, joint estimation of the correlated Gaussian samples and detection of the correlated channel states is demanded. Factorizing the joint pdf of signal samples $\underline{s} = \{s_k\}$ and channel states $\underline{i} = \{i_k\}$ conditioned on the observation samples $\underline{y} = \{y_k\}$ yields

$$\begin{aligned} p(\underline{s}, \underline{i} \mid \underline{y}) &\propto p(\underline{s}, \underline{i}, \underline{y}) = p(\underline{y} \mid \underline{s}, \underline{i}) p(\underline{s}) P(\underline{i}) \\ &= \left[\prod_{k=1}^{K-1} p(y_k \mid s_k, i_k) p(s_k \mid s_{k-1}) P(i_k \mid i_{k-1})\right] \\ &\quad p(y_0 \mid s_0, i_0) p(s_0) P(i_0) \end{aligned} \tag{11}$$

Considering the Markovianity of noise states, $P(i_k \mid i_{k-1})$ can be computed from matrix $\boldsymbol{P}$ in (9) as

$$P(i_k \mid i_{k-1}) = \boldsymbol{P}(i_{k-1}+1, i_k+1) \tag{12}$$

and its initial value is set by the distribution

$$P(i_0) = \sum_{n=0}^{M-1} p_n' \delta(i_0 - n) \tag{13}$$

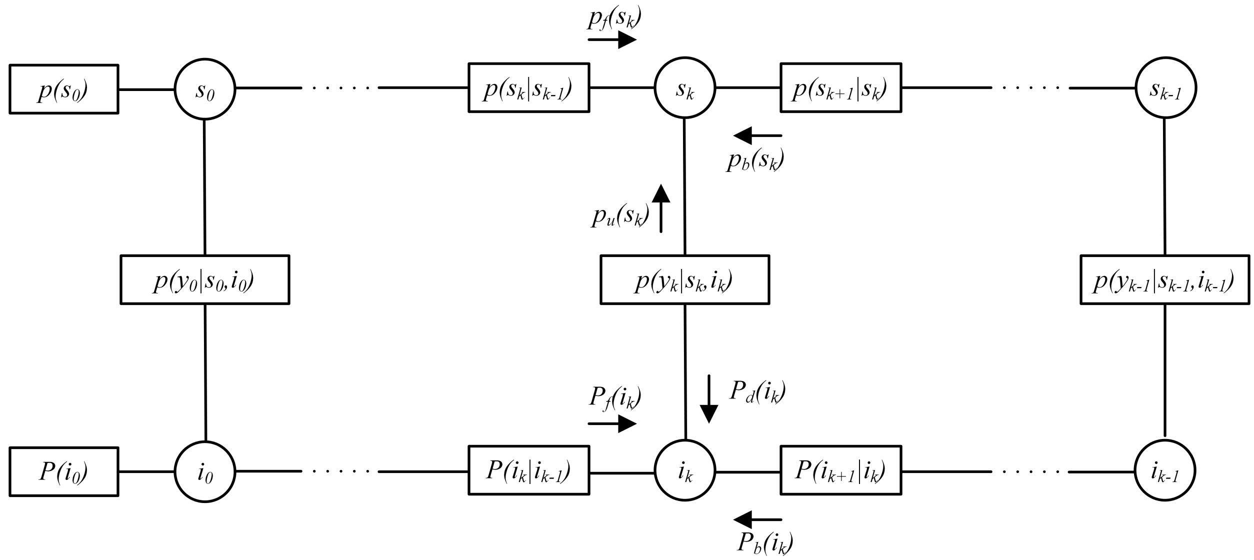

Fig. 2. Schematic of the factor graph for joint probability distribution $p(\underline{s}, \underline{i} \mid \underline{y})$. i_k and s_k represent the noise state and the signal sample at time instant k, respectively. At the k-th stage, the messages sent from factor nodes to variables nodes are indicated in the graph.

where M is the number of noise states. From the AR(1) process (1), $p(s_k \mid s_{k-1})$ can be written as

$$p(s_k \mid s_{k-1}) = g(s_k - (a_1 s_{k-1} + \eta_\omega), \sigma_\omega^2) \tag{14}$$

in which $g(x - \eta, \sigma^2)$ denotes the Gaussian pdf with mean η and variance σ^2. Substituting $\eta_\omega = \eta_s(1 - a_1)$ and $\sigma_\omega^2 = \sigma_s^2(1 - a_1^2)$ in (14) results in

$$p(s_k \mid s_{k-1}) = g((s_k - \eta_s) - a_1(s_{k-1} - \eta_s), \sigma_s^2(1 - a_1^2)) \tag{15}$$

so that the initial distribution of signal samples can be obtained by simply setting a_1 to zero

$$p(s_0) = g(s_0 - \eta_s, \sigma_s^2) \tag{16}$$

which is the prior distribution of samples s_k. From (2), the conditional pdf of the observed sample at time k is

$$p(y_k \mid s_k, i_k) = g(y_k - s_k, \sigma_{i,k}^2) \tag{17}$$

in which $\sigma_{i,k}^2$ is determined by the noise state at time k and can be computed from (5) as a function of i_k.

The factor graph corresponding to the joint pdf in (11) is depicted in Fig. 2. As extensively discussed in [11], such a factor graph with loops and mixed discrete (i_k) and continuous (s_k) random variables makes the message passing procedure intractable. One way to overcome these difficulties is to split the factor graph in its lower and upper parts and let them operate separately, while approximating the messages that these two subgraphs exchange along the vertical edges Fig. 2. Three main operations are thus implemented in the graph. Throughout the bottom line of the graph, channel states are detected using the BCJR algorithm. Throughout the top line of the graph, signal samples are estimated using a Kalman smoother. The reason for referring these two well assessed algorithms is that, as can be easily demonstrated [11], the two halves of the factor graph indeed reduce to the implementation of BCJR (lower half) and Kalman smoother (upper half). On each vertical edge of the graph, a hard decision on i_k is used to perform the approximation of Gaussian mixtures, that naturally arise as messages [11], which remarkably reduces the overall complexity of the computation. The communication procedure among the algorithmic blocks is known as *parallel iterative schedule* [11]. In the following sections, the algorithms will be briefly discussed.

B. BCJR

Suppose that the upper part of the factor graph provides the estimation of Gaussian signal samples at a given iteration, i.e., that node s_k provides its own estimated pdf, with mean $\widehat{\eta}_k$ and variance $\widehat{\sigma}_k^2$. Thus, the message sent from the variable node s_k to the factor node $p(y_k \mid s_k, i_k)$, named as $p_d(s_k)$, can be written by

$$p_d(s_k) = g(s_k - \widehat{\eta}_k, \widehat{\sigma}_k^2) \tag{18}$$

The message received by the variable node i_k from the factor node $p(y_k \mid s_k, i_k)$ can be computed with a convolution of Gaussian pdfs as

$$\begin{aligned} P_d(i_k) &= \int p_d(s_k) p(y_k \mid s_k, i_k) ds_k \\ &= \int g(s_k - \widehat{\eta}_k, \widehat{\sigma}_k^2) g(y_k - s_k, \sigma_{i,k}^2) ds_k \\ &= g(y_k - \widehat{\eta}_k, \widehat{\sigma}_k^2 + \sigma_{i,k}^2) = p(y_k \mid i_k) \end{aligned} \tag{19}$$

It should be noted that $\sigma_{i,k}^2$ can take values from a set of M different variances, each corresponds to one particular channel state. Having $P_d(i_k)$ for all channel states, enables us to compute the forward and the backward messages on the bottom line of the factor graph.

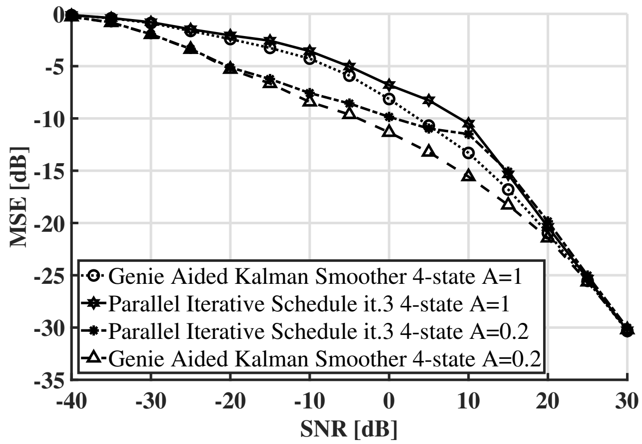

Fig. 3. MSE vs. SNR curves, 4-state noise model. The signal parameters are $a_1 = 0.9, \eta_s = 0$ and $\sigma_s^2 = 1$. The noise parameters are $x = 0.9$, $A = 0.2, 1$ and $\Gamma = 0.01$. 100 frames of 1000 samples have been transmitted. The genie aided Kalman smoother is a lower bound for the estimation. The Parallel Iterative Schedule is a suboptimal estimator in which the noise variances are provided by the hard decision unit.

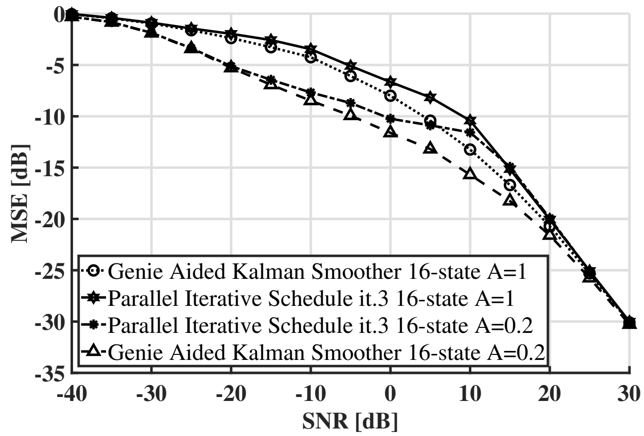

Fig. 4. MSE versus SNR curves, 16-state noise model. Same parameters as in Fig. 3. 100 frames of 1000 samples have been transmitted.

$$P_f(i_k) = \sum_{i_{k-1}} P(i_k \mid i_{k-1}) P_f(i_{k-1}) P_d(i_{k-1}) \tag{20}$$

$$P_b(i_k) = \sum_{i_{k+1}} P(i_{k+1} \mid i_k) P_b(i_{k+1}) P_d(i_{k+1}) \tag{21}$$

in which $k = 1, \cdots, K-1$ for $P_f(i_k)$ and $k = K-2, \cdots, 0$ for $P_b(i_k)$. The initial values for forward and backward equations are $P_f(i_0) = P(i_0)$ in (13) and $P_b(i_{K-1}) = 1$, respectively. It can be shown that the the above equations implement a celebrated MAP symbol detection algorithm, known as BCJR [12]. Once a forward (filtering) and backwards (smoothing) message-passing iteration has been completed, the completion step of the BCJR algorithm can be obtained by multiplying all of the incoming messages to variable node i_k:

$$\widetilde{P}(i_k) = P_f(i_k) P_b(i_k) P_d(i_k) \tag{22}$$

The message that goes upward from the variable node i_k to the factor node $p(y_k \mid s_k, i_k)$ can now be calculated as

$$P_u(i_k) = P_f(i_k) P_b(i_k) = \widetilde{P}(i_k) P_d^{-1}(i_k) \tag{23}$$

C. Kalman Smoother

Suppose that the lower part of the factor graph uses the above estimation (22) on i_k to provide $p(y_k \mid s_k, i_k)$ at the next iteration. Assume that this probability is a single Gaussian distribution with mean y_k and variance $\sigma_{n,k}^2$ at each time instant k. Therefore, the message sent from the factor node $p(y_k \mid s_k, i_k)$ to the variable node s_k, can be written in a way similar to (17), as

$$p_u(s_k) = g(y_k - s_k, \sigma_{n,k}^2) \tag{24}$$

Having $p_u(s_k)$ for all signal samples, and by considering that the noise variance changes from sample to sample, which implies that noise is not stationary, the upper part of the factor graph coincides with a Kalman smoother (see, e.g., [13]). Assuming that $p_f(s_k) = g(s_k - \eta_{f,k}, \sigma_{f,k}^2)$ and $p_b(s_k) = g(s_k - \eta_{b,k}, \sigma_{b,k}^2)$, the estimated Gaussian samples from the Kalman smoother can be described by their means and variances. After completing the message passing procedure, results

$$\widehat{\eta}_k = (\sigma_{f,k}^{-2} + \sigma_{n,k}^{-2} + \sigma_{b,k}^{-2})^{-1} [\frac{\eta_{f,k}}{\sigma_{f,k}^2} + \frac{y_k}{\sigma_{n,k}^2} + \frac{\eta_{b,k}}{\sigma_{b,k}^2}] \tag{25}$$

$$\widehat{\sigma}_k^2 = (\sigma_{f,k}^{-2} + \sigma_{n,k}^{-2} + \sigma_{b,k}^{-2})^{-1} \tag{26}$$

The estimates in (25) and (26) will be fed back to the lower half of the graph, so that BCJR and Kalman smoother can iterate until convergence.

D. Hard Decision

The Gaussianity of message $p_u(s_k)$ is not trivial. Let us describe the crucial role of the hard decision unit and provide a description of the parallel iterative schedule. According to (19), the message from the factor node $p(y_k \mid s_k, i_k)$ to the downward variable node i_k, $P_d(i_k)$, is a single Gaussian distribution. However, the upward message from the factor node $p(y_k \mid s_k, i_k)$ to the variable node s_k, $p_u(s_k)$, is a mixture of Gaussian distributions. The reason is that $P_u(i_k)$ is the estimation of the probability mass function (pmf) of the discrete variable i_k, with M different probability masses, such that the resulting observation y_k has to account for M possible noise variances (through the theorem of total probabilities). Consequently, $p_u(s_k)$ consists of M different terms each associated to one possible noise state.

$$p_u(s_k) = \sum_{j=0}^{M-1} P_u(i_k = j) g(y_k - s_k, \sigma_{j,k}^2) \tag{27}$$

Therefore, there is a mixture of Gaussian distributions at each vertical edge of the factor graph, which dramatically increases the overall estimation complexity. Our approach to overcome this problem is to approximate $p_u(s_k)$ by its n-th Gaussian term which has the largest mass in the pmf. To this end, according to (23), a hard decision is taken on $\widetilde{P}(i_k)$, as provided by the BCJR algorithm, to find out the most likely state $i_k = n$ and to compute $p_u(s_k)$ as a single Gaussian distribution.

$$p_u(s_k) \approx g(y_k - s_k, \sigma^2_{n,k}) \tag{28}$$

The parallel iterative schedule can now be described explicitly. Kalman smoother and BCJR work in parallel and they exchange their messages at every iteration. At the first iteration, Kalman smoother works with initial values of $P(i_k) = P(i_0)$ for all channel states and BCJR starts with initial values of $p(s_k) = p(s_0)$ for all signal samples. At every iteration, hard decisions for i_k are taken to compute $p_u(s_k)$ for the next iteration.

IV. Simulation Results and Discussion

The Mean Squared Error (MSE) versus Signal to Noise Ratio (SNR) curves for estimation of correlated Gaussian samples in the correlated 4-state and 16-state Middleton class A noise are depicted in Fig. 3 and Fig. 4, respectively. The simulation parameters were set as follows. We considered a correlation parameter $x = 0.9$, characterizing the Markovian memory of the channel, that tends to remain in the same state with probability 0.9. Hence, once an impulsive noise event is started due to the presence of i electromagnetic interferers, such a number is expected to change with probability 0.1. The impulsive index A, that is the average number of active interferers, was considered to be either 0.2 or 1 and we set $\Gamma = 0.01$, so that the average power of impulsive noise is 100 times larger than that of the background noise. We assumed that the transmitted signal samples s_k are strongly correlated, with $a_1 = 0.9$ being the pole of the IIR digital filter in (1). The opposite situation of independent Gaussian transmitted samples, i.e., $a_1 = 0$, is the one analyzed in [10], for which a much simpler cycle-free factor graph is obtained. We assumed zero mean correlated signal samples, so that $\eta_s = 0$, simply because a nonzero mean would result in a trivial, biased extension of the same system model. We set a normalized value $\sigma^2_s = 1$ for the variance of signal samples, which represents the asymptotic values for the MSE curves, $(\sigma^2_s)_{dB} = 0$ dB, on the left of Figs. 3 and 4, when noise dominates over signal power (so that the signal estimates degenerates onto the prior signal average). We transmitted 100 frames of 1000 samples each (for a total number of 10^5 samples, and the average MSE is reported for SNR values ranging from -40 dB (close to a purely noisy received signal) to $+30$ dB (close to a noiseless signal estimation). The curves labeled "Genie Aided Kalman Smoother" address the estimation process in which the Kalman smoother has an exact knowledge of the noise variance at any time instant. Since, with such a side information, Kalman smoother is known to be the optimal estimation strategy in additive (possibly non-stationary) Gaussian noise, then these curves represent a lower bound for the performance. The curves labeled "Parallel Iterative Schedule" address the estimation process in which the hard decision unit approximates the Gaussian mixtures and provides the noise variances for the Kalman smoother. Therefore, this is a suboptimal estimator. The final estimation is obtained after three iterations, denoted by "it.3" in the figures. Since as we verified, the iterative message passing procedure has reached convergence and results would not improve appreciably, for further iterations.

Since the 2-state scenario is fully analyzed in [11], we considered more noise states, i.e., from 4 up to 16 noise states, meaning that impulsive noise can be due to 3 to 15 possible sources of electromagnetic interference. As it can be seen in Figs. 3 and 4, increasing the value of impulsive index A results in a reduction of the system performance, as expected. This is a logical result, since the impulsive index determines the average number of active interferers per unit time. A better insight is obtained by checking the numerical values of MSE at different SNRs. For instance, in 4-state scenario, the MSE is -10.96 dB for $A = 0.2$ and is -8.265 dB for $A = 1$ at $\mathrm{SNR} = 5$ dB. These values change to -10.88 dB and -8.158 dB in the 16-state case, revealing that the maximum number of interferers has little impact, once the SNR is fixed, as seen by comparing the curves with equal A parameter, in Figs. 3 and 4. It can be concluded that for the fixed frame length of 1000 samples, increasing the number of active interferes beyond 4 noise states does not change the estimation performance considerably. The reason is that the higher order interferers, which are extra sources of noise, have smaller prior probabilities and are rarely observed in a simulated frame. A much larger number of noise states, i.e., the presence of many more than 15 sources of impulsive noise, can change results significantly but is anyhow left to further investigation since, for $M \gg 16$, simulations become computationally demanding.

V. Conclusion

In this paper, a message passing approach is proposed to estimate correlated Gaussian signals in the presence of the Markov-Middleton class A noise. The noise was modeled by considering up to 15 sources of impulsive noise. Results show that, for a given overall SNR level, the presence of many (weaker) sources of impulsive noise induces a degradation that is similar to that of fewer (stronger) ones. It has been shown that the hard approximation of Gaussian mixtures results in a suboptimal estimation strategy while leading to a considerable reduction of estimation complexity.

References

[1] M. Singh, A. Amruthakala, M. L. Sudheer, and A. S. Murthy, "Analysis of Noise in Broadband Powerline Communications (B-PLC) in Frequency Range of 150kHz–30MHz," *2017 4th IEEE Uttar Pradesh Section International Conference on Electrical, Computer and Electronics (UPCON)*, Mathura, Oct. 2017, pp. 101-106.

[2] A. Mathur, M. R. Bhatnagar, and B. K. Panigrahi, "Performance Evaluation of PLC Under the Combined Effect of Background and Impulsive Noises," *IEEE Comm. Lett.*, vol. 19, no. 7, pp. 1117-1120, July 2015.

[3] X. Lv and Y. Li, "Joint Channel Estimation and Impulsive Noise Mitigation For Power Line Communications," *2018 24th Asia-Pacific Conference on Communications (APCC)*, Ningbo, China, Nov. 2018, pp. 576-579.

[4] P. Torio and M. G. Sanchez, “Method to Cancel Impulsive Noise From Power-Line Communication Systems by Processing the Information in the Idle Carriers,” *IEEE Trans. Power Del.*, vol. 27, no. 4, pp. 2421-2422, Oct. 2012.

[5] M. Ghosh, “Analysis of the Effect of Impulse Noise on Multicarrier and Single Carrier QAM Systems,” *IEEE Trans. Comm.*, vol. 44, pp. 145-147, Feb. 1996.

[6] D. Fertonani and G. Colavolpe, “On Reliable Communications over Channels Impaired by Bursty Impulse Noise,” *IEEE Trans. Comm.*, vol. 57, no. 7, pp. 2024-2030, July 2009.

[7] D. Middleton, “Non-Gaussian Noise Models in Signal Processing for Telecommunications: New Methods an Results for Class A and Class B Noise Models,” *IEEE Trans. Inf. Th.*, vol. 45, no. 4, pp. 1129-1149, May 1999.

[8] G. Ndo, F. Labeau, and M. Kassouf, “A Markov-Middleton Model for Bursty Impulsive Noise: Modeling and Receiver Design,” *IEEE Trans. Power Del.*, vol. 28, no. 4, pp. 2317-2325, Oct. 2013.

[9] P. Banelli, “Bayesian Estimation of a Gaussian Source in Middleton’s Class-A Impulsive Noise,” *IEEE Signal Proc. Lett.*, vol. 20, n.10, pp. 956-959, oct. 2013.

[10] M. S. Alam, G. Kaddoum, and B. L. Agba, “Bayesian MMSE Estimation of a Gaussian Source in the Presence of Bursty Impulsive Noise,” *IEEE Comm. Lett.*, vol. 22, no. 9, pp. 1846-1849, Sept. 2018.

[11] A. Vannucci, G. Colavolpe, R. Pecori, and L. Veltri, “Estimation of a Gaussian Source with Memory in Bursty Impulsive Noise,” *2019 IEEE International Symposium on Power Line Communications and its Applications (ISPLC 2019)*, Prague, Czech Republic, Apr. 2019, pp. 1-6.

[12] L. R. Bahl, J. Cocke, F. Jelinek, and J. Raviv, “Optimal Decoding of Linear Codes for Minimizing Symbol Error Rate,” *IEEE Trans. Inf. Th.*, pp.284-287, Mar. 1974.

[13] J. R. Movellan, “Discrete Time Kalman Filters and Smoothers,” *MPLab Tutorials*, Univ. of California, San Diego (USA), 2011.