This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/LWC.2014.2359464, IEEE Wireless Communications Letters

Robust Turbo Decoding in a Markov Gaussian Channel

Der-Feng Tseng, *Member, IEEE*, Fikreselam G. Mengistu, Yunghsiang S. Han, *Fellow, IEEE,* Mengistu Abera Mulatu, Li-Chung Chang, *Member, IEEE*, and Tzung-Ru Tsai

***Abstract*— Strong impulse noise is widely known to adversely affect conventional receivers designed only to consider background noise. Although sophisticated receivers offer substantial performance improvements, fully exploiting the impulse statistics, which are generally not time invariant and are difficult to model accurately, can be unrealistic. Alternatively, without making assumptions regarding the underlying impulse channel model, the authors' previously developed robust decoding schemes can achieve performance equivalent to that of their optimal counterparts in impulse noise channels. In this paper a robust turbo decoding metric is proposed to address the inherent memory in an impulse channel: a two-dimensional trellis is used to adapt for channel state transitions when statistics on the memory impulse channel model are lacking. The simulation results verified the robustness of the proposed decoder in harsh environments.**

***Index Terms*— Impulse noise, Markov Gaussian channel, turbo decoder, two-dimensional trellis.**

I. Introduction

A generic communication system is prone to substantial performance degradation in environments characterized by strong impulse noise, which is generally not time invariant and is difficult to model accurately [1], [2], causing system design to be considerably challenging. To overcome that barrier, numerous studies have used channel codes by assuming that the impulse noise can be modeled by the Class-A model [3]. In [4] a convolutional code was used and its performance bound was derived; the low density parity check code (LDPC) was employed in [5], whereas a turbo code was used and a simplified turbo decoder was suggested in [6]. Recently, studies that forgo any assumed (underlying) impulse noise model (to realistically reflect the limitations imposed on modeling) have gained substantial attention. Given a properly set clipping threshold, [7]–[10] revealed that the clipping-featured single carrier coded systems were on par with their optimal counterparts which assumed impulse noise statistics. Moreover, [11] developed a denoising method using compressive sensing in multi-carrier systems, and [12] went on to incorporate channel coding into a factor graph based algorithm, yielding a performance level near that of the optimal receiver.

Nevertheless, a memoryless impulse channel may not be adequately justified, especially when communication systems are limited by the data packet size. A first-order Markov process, the Gilbert-Elliot model [13] effectively serves as a memory impulse channel by simplistically describing the transition of two channel states. The vulnerability of decision feedback decoders [14], [15] to excessive impulse noise motivated the authors of [16] to propose a turbo coding mechanism in a binary-input binary-output Gilbert-Elliot channel: a supertrellis, comprising the encoder and channel states, was incorporated into the Bahl-Cocke-Jelinek-Raviv (BCJR) algorithm [17], and was analyzed by examining its convergence behavior, expressed as the bit error rate (BER) versus the number of iterations. A joint model-based channel parameter estimation and turbo decoding scheme that involved exploiting an underlying hidden Markov model was proposed in [18].

The Markov Gaussian channel was introduced in [19] to accommodate the bursty nature of impulse noise, and an iterative algorithm was tested in LDPC-coded system. In contrast to [19] using the Markov Gaussian channel statistics, [20] examined the BER performance of an alpha-penalty function decoder (alpha-PFD) in a convolutionally coded transmission; further, an enhanced scheme (c.f., [21]) overcame the persistent error floor typically encountered when using the alpha-PFD. The supertrellis structure justified in [16], [21] was adopted here to devise a robust and efficient turbo decoding scheme in a Markov Gaussian channel. To achieve the performance level reported in [16], the computationally intensive model-based parameter estimation [18] was requisite before the BCJR algorithm was invoked. Despite lacking channel impulse statistics, the proposed decoder induced a BER performance level markedly close to that of its sub-optimal counterpart requiring impulse channel statistics, evidencing the robustness and efficiency of the proposed decoder.

II. System Framework

For illustrative purposes, a punctured rate-$\frac{1}{2}$ parallel concatenated turbo encoder composed of two recursive systematic convolutional (RSC) codes is considered. Extension to any other rate is straightforward. Denote the codeword transmitted by $\boldsymbol{v}$ as $(v_0, v_1, \ldots, v_{2M-1})$, where M is the turbo interleaver size. Assuming that the binary phase shift keying (BPSK) modulation format is used (the symbol energy is assumed E_s), the modulated symbol is expressed as $\xi_j = (-1)^{v_j}\sqrt{E_s}$ $(0 \leq j \leq 2M-1)$; consequently, the transmitted symbol corresponding to the l-th information bit (i.e., m_l $(0 \leq l \leq M-1)$) is ξ_{2l}. To disperse burst errors, the transmitted symbol sequence $\{\xi_j\}_0^{2M-1}$ is passed through a channel interleaver.The received symbol sequence (after passing through a channel deinterleaver) is expressed as $\boldsymbol{y} = (-1)^{\boldsymbol{v}}\sqrt{E_s} + \boldsymbol{n}$, where each noise sample n_j $(0 \leq j \leq 2M-1)$ is defined by the channel state s_j, which can be either good $(s_j = G)$ or bad

The authors are with the Department of Electrical Engineering, National Taiwan University of Science and Technology, Taipei, Taiwan. (e-mail: dtseng@mail.ntust.edu.tw.) This work was supported in part by the Ministry of Science and Technology (MOST) of Taiwan under grant no. MOST 103-2221-E-011-034- and NSC 101-2221-E-011-069-MY3.

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/LWC.2014.2359464, IEEE Wireless Communications Letters

($s_j = B$). As widely assumed in relevant literature, the power of noise in the bad channel state is several times greater than that of additive white Gaussian noise (AWGN), which governs the good channel state. The single-sided power spectral density of the AWGN is assumed to be N_0. The probability density functions (PDFs), when conditioned on good and bad channel states, are respectively expressed as follows:

$$\Pr(y_j|v_j, s_j\!=\!G) = \mathcal{N}(y_j; (-1)^{v_j}\sqrt{E_s}, N_0/2) \qquad (1)$$

$$\Pr(y_j|v_j, s_j\!=\!B) = \mathcal{N}(y_j; (-1)^{v_j}\sqrt{E_s}, RN_0/2)\ , \qquad (2)$$

where $\mathcal{N}(x;\mu,\sigma^2) = e^{-\frac{(x-\mu)^2}{2\sigma^2}}/\sqrt{2\pi\sigma^2}$ is the Gaussian PDF of a random variable x with mean μ and variance σ^2; R, an index of the strength of impulse noise power, indicates the average noise power ratio between the bad and good channel states. In consideration of the memory property inherent to impulse noises, the noise state sequence is modeled based on a first-order two-state Markov process [19], where the transformation of noise sample n_j in a noise PDF is driven by four transition probabilities, namely P_{GG}, P_{BG}, P_{GB}, and P_{BB}, where $P_{s_j s_{j+1}} = \Pr(s_{j+1}|s_j)$ $(s_j, s_{j+1} \in \{G, B\})$ denotes the probability that a transition from channel state s_j to state s_{j+1} occurs. Moreover, channel state probabilities, i.e., P_G and P_B, are incorporated to complete the modeling of the Markov process. By accommodating the interleaver depth I, the transition probability matrix for the two-state stationary and irreducible Markov chain, is approximated as (c.f., [20])

$$\begin{bmatrix} P_{GG} & P_{GB} \\ P_{BG} & P_{BB} \end{bmatrix} = \begin{bmatrix} P_G & P_B \\ P_G & P_B \end{bmatrix} + \eta^I \begin{bmatrix} P_B & -P_B \\ -P_G & P_G \end{bmatrix}, \qquad (3)$$

where $\eta^I \approx 1 - I\frac{P_{BG}}{P_G}$ and P_G/P_{BG} reflects the channel memory. This work considers a persistent memory ($\eta^I > 0$).

III. Robust Turbo Decoding Metric

To accommodate the inherent impulse channel memory, the soft-in soft-out (SISO) decoder incorporates a supertrellis [16] into the BCJR algorithm.[1] First, integrate the impulse channel state with the encoder state to yield a superstate; for the information bit time instant l, denote $\tilde{\mathcal{S}}_{l-1}$ and $\tilde{\mathcal{S}}_l$ as the starting and ending states, respectively. In the illustrative example plotted in Fig. 1, superstate $\tilde{\mathcal{S}}_{l-1} \triangleq (x_{l-1}, s_{2l-1}), l \in \{0, \ldots, M\}$ is composed of the encoder state $x_{l-1} \in \{0, 1, \ldots, 2^m - 1\}$ and channel state $s_{2l-1} \in \{G, B\}$. The SISO decoder measures the log-likelihood ratio of m_l as follows:

$$L(m_l) = \log\left(\frac{\sum_{\chi^+} \alpha(\tilde{\mathcal{S}}_{l-1})\beta(\tilde{\mathcal{S}}_l)\gamma(\tilde{\mathcal{S}}_{l-1}, \tilde{\mathcal{S}}_l)}{\sum_{\chi^-} \alpha(\tilde{\mathcal{S}}_{l-1})\beta(\tilde{\mathcal{S}}_l)\gamma(\tilde{\mathcal{S}}_{l-1}, \tilde{\mathcal{S}}_l)}\right)\ , \qquad (4)$$

where χ^+ denotes the set of all allowable (two-dimensional) state transitions from the starting state $\tilde{\mathcal{S}}_{l-1}$ to ending state $\tilde{\mathcal{S}}_l$ that correspond to $m_l = 1$; similarly, χ^- is the set of the remaining state transitions (associated with $m_l = 0$). The forward and backward recursions, denoted by $\alpha(\cdot)$ and $\beta(\cdot)$, respectively, are yielded according to the following:

$$\alpha(\tilde{\mathcal{S}}_l) \;=\; \textstyle\sum_{\tilde{\mathcal{S}}_{l-1}} \alpha(\tilde{\mathcal{S}}_{l-1})\gamma(\tilde{\mathcal{S}}_{l-1}, \tilde{\mathcal{S}}_l) \qquad (5)$$

$$\beta(\tilde{\mathcal{S}}_{l-1}) \;=\; \textstyle\sum_{\tilde{\mathcal{S}}_l} \beta(\tilde{\mathcal{S}}_l)\gamma(\tilde{\mathcal{S}}_{l-1}, \tilde{\mathcal{S}}_l)\ . \qquad (6)$$

[1]The trellis of the mother code of the punctured code is implemented in the decoder, which also uses the punctured information in the decoding process. To simplify notation, the proposed decoding algorithm is applied to the equivalent trellis of the punctured code.

The appendix first details the derivation of $\gamma(\tilde{\mathcal{S}}_{l-1}, \tilde{\mathcal{S}}_l)$ (c.f., (8)) when the impulse channel model is used in the sub-optimal decoder. By contrast, unknown to any impulse channel model, the proposed decoder is subjected to estimate (2) to yield a robust metric (9). To do so, the proposed decoder opts to neutralize the excessive Euclidean bit metric (i.e., $-(y_j - (-1)^{v_j}\sqrt{E_s})^2$) whenever y_j is deemed to be corrupted by an impulse; as exemplified in Fig. 1, the Euclidean bit metrics associated with those dash-dot lines terminated at blank circles are untrustworthy. Noteworthily, the product $P_{s_{j-1},s_j=B}\Pr(y_j|v_j, s_j = B)$–a byproduct for calculating those metrics–is small for a small $P_{s_{j-1},s_j=B}$ as well as a large R, and can thus be approximated by ignoring the term $\Pr(y_j|v_j, s_j = B)$, facilitating erasure marking, as detailed in [7], to overcome the negative effect from the strong impulse. Furthermore, the derived Chernoff bound in [7] reveals that the BER performance level is crucially improved by including an offset provider (an estimate of $\Pr(y_j|v_j, s_j = B)$), which happens to be vastly smaller than one (used by erasure marking). From this perspective, [21] estimates the noise PDF in a bad channel state in the memory channel model using

$$\Pr(y_j|v_j, s_j = B) \;=\; 1/\sqrt{\pi R^{(d)} N_0}\ . \qquad (7)$$

Given a signal-to-noise ratio (SNR $\triangleq E_b/N_0$, where E_b is the information bit energy) region of interest, the offset remains markedly smaller than one when the decoder voluntarily sets the estimate $R^{(d)}$ to be large to reflect a context where strong impulses likely occur; simulation results attest the robustness of the decoder against a wide range of large $R^{(d)}$ values.

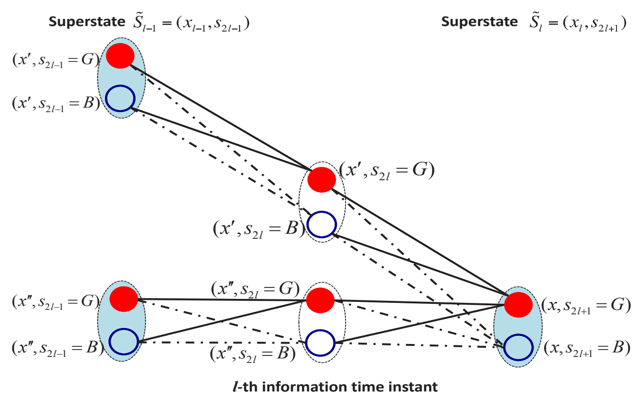

Fig. 1. Example of an expanded trellis diagram for a rate $\frac{1}{2}$ RSC code.

Other than invoking the estimation for the noise PDF (2) to yield (9), obtaining the estimates of impulse channel parameters $P_B^{(d)}(P_G^{(d)})$, and $P_{s_j s_{j+1}}^{(d)}$ $(s_j, s_{j+1} \in \{G, B\})$ (the superscript "(d)" represents the proposed decoder) seems to be indispensable. Notably, the transition matrix (3) can be properly approximated when only estimates $P_B^{(d)}$ and $P_{BG}^{(d)}$, and an interleaver depth I are available. As performed in [21], conducting a person-by-person optimization search [22] – by testing the proposed decoder in various P_{BG} values – leads to a selective estimate $P_{BG}^{(d)}$, given a SNR and a depth I; further, Section IV reveals that the BER level remains nearly identical when a variety of plausible estimates $P_B^{(d)}$ are tested.

To activate (5) ((6)), the superstates $\tilde{\mathcal{S}}_{-1} = (x_{-1}, s_{-1})$ and $\tilde{\mathcal{S}}_{M-1} = (x_{M-1}, s_{2M-1})$ must be identified, and $\alpha(\tilde{\mathcal{S}}_{-1}) \triangleq \alpha\left((x_{-1}, s_{-1})\right)$ and $\beta(\tilde{\mathcal{S}}_{M-1}) \triangleq \beta\left((x_{M-1}, s_{2M-1})\right)$ must be

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/LWC.2014.2359464, IEEE Wireless Communications Letters

set beforehand. Because state zero of the first trellis level is the initial encoder state ($x_{-1} = 0$), set $\alpha(0, s_{-1}) = 1$ and $\alpha(x, s_{-1}) = 0$ for all $x \in \{1, \ldots, 2^m - 1\}$. Moreover, the transition probability is initialized at $P^{(d)}_{s_{-1}s_0} = P^{(d)}_{s_0}$. When these parameters have been initialized and the a priori probability of m_l is available, $\alpha((\tilde{\mathcal{S}}_0))$ can be yielded after (9) and (5) have been solved. In regard to initializing $\beta(\tilde{\mathcal{S}}_{M-1})$, a subtle difference between the two SISO decoders (attributable to whether the information bit stream is passed through the turbo interleaver) is identified as follows. Set $\beta((0, s_{M-1})) = 1$ and $\beta((x, s_{M-1})) = 0$ for all $x \in \{1, \ldots, 2^m - 1\}$ if $x_{M-1} = 0$ holds; otherwise, $\beta(\tilde{\mathcal{S}}_{M-1})$ is set as $1/2^m$ if the final encoder state does not return to zero. Furthermore, the time symmetry property of this underlying Markov process enables $P^{(d)}_{s_{2M-1}s_{2M-2}} = P^{(d)}_{s_{2M-2}}$, successively yielding $\beta(\tilde{\mathcal{S}}_{M-2})$ after (9) and (6) have been invoked.

IV. Simulation Results

This section presents simulation results regarding the BER versus SNR for the turbo-coded system in a Markov Gaussian channel and the ratio between E_s and the code rate. The transmission framework comprised a punctured rate $\frac{1}{2}$ parallel concatenated encoder composed of two identical RSC codes, of which the generator matrix is [23 35]; a random interleaver with size $M = 16384$; the BPSK modulation format; and a channel interleaver with $I = 20$. The transition probability P_{BG} for the Markov Gaussian channel was 0.025, while its estimate $P^{(d)}_{BG}$ was yielded at 0.035 (refer to Section III). Despite the lack of complete statistical information on the memory impulse channel, the proposed decoder was compared with the sub-optimal decoder (developed in [16]), which however assumed the impulse channel, in terms of the BER, and all of the regarding curves were derived after eight iterations.

The effect of the impulse occurrence probability P_B on the BER performance of both the sub-optimal and proposed decoders was investigated and the results are shown in Fig. 2. The power ratio of the noise between the bad and good channel states R was set as 100 in all contexts (regarding a wide range of P_B values: 0.01, 0.02, 0.05, 0.1), and the proposed decoder assumed $R^{(d)} = 200$, and $P^{(d)}_B = P_B$ but $P^{(d)}_{BG} \neq P_{BG}$. The greater the probability P_B was, the less favorable the BER performance was. The sub-optimal decoder exhibited a performance gap of 1.3 dB SNR when the performance level at $P_B = 0.1$ (see the dash-dot line marked with "○") was compared with that at $P_B = 0.01$ (see the dash-dot line marked "□") when the BER was 10^{-5}, and the proposed decoder exhibited a slightly wider gap. The performance of the proposed decoder was similar to that of its sub-optimal counterpart requiring the impulse statistics. A mere 0.6-dB SNR loss at $P_B = 0.1$ was observed (see solid and dash-dot lines marked "○"), and SNR loss was almost indistinguishable at $P_B = 0.01$ (see solid and dash-dot lines marked "□"), evidencing the robustness of the proposed decoder. A fixed estimate of R was applied in the previous experiment – $R^{(d)} = 200$ was set in Fig. 2; however, in the subsequent experiment the degree to which the magnitude of $R^{(d)}$ affects the BER performance was examined. Fig. 3 shows that the BER performance level of the proposed decoder in the Markov Gaussian channel ($P_B = 0.05, R = 100$) did not vary according to the value of $R^{(d)}$, although a slight performance loss was observed when an excessively low estimate ($R^{(d)} = 50$) was used. Thus, $R^{(d)}$ was set as 200 in the following tests.

Compared with the conventional SISO decoder for which a decoding metric was revised according to an alpha-PFD, the proposed decoder showed a substantial SNR gap (Fig. 4, which used $\alpha = 0.5$ (suggested otherwise for convolutionally coded systems in [20])), verifying that the proposed decoding metric is effective because it accommodates unknown transition probabilities, which were also neglected in [20], causing the BER to be persistently high even at moderate SNR values; the BER did not decrease to 10^{-5} until the SNR increased to greater than 8 and 10 dB at $P_B = 0.02$ and 0.05, respectively (see dashed lines marked "□" and "○" in Fig. 4).

To validate the robustness of the proposed decoder, all of the channel parameters in the Markov Gaussian channel, $P_B = 0.1(0.05), P_{BG} = 0.025$, and $R = 100$ were assumed to be unavailable. The proposed decoder used $R^{(d)} = 200$ and $P^{(d)}_{BG} = 0.035$ and tested various $P^{(d)}_B$ values. Regardless of the assumed $P^{(d)}_B$ ($P^{(d)}_B = 0.02, 0.05, 0.1$), the BER performance level was nearly identical at $P_B = 0.05$, as shown in Fig. 5. By referring the context in which $P_B = 0.1$ to those solid lines in Fig. 5, a similar observation can be made. These results reinforced the assertion that the proposed decoder is robust against the estimate $P^{(d)}_B$ (together with $R^{(d)}$ and $P^{(d)}_{BG}$, (Fig. 3)).

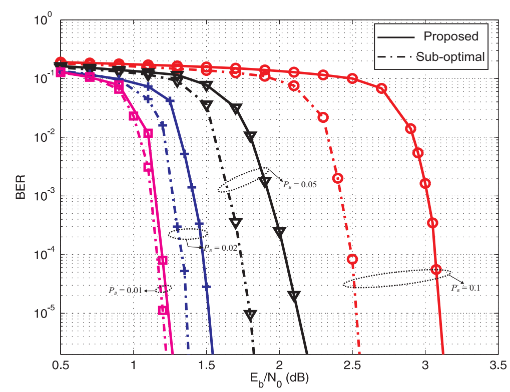

Fig. 2. BER performance for the proposed and sub-optimal turbo decoders in Markov Gaussian channel ($P_B = 0.01, 0.02, 0.05, 0.1$; $R = 100$).

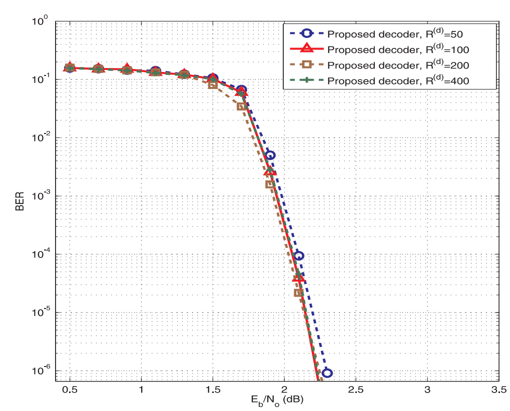

Fig. 3. BER performance for the proposed turbo decoder with various guesses of R in Markov Gaussian channel ($P_B = 0.05; R = 100$).

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/LWC.2014.2359464, IEEE Wireless Communications Letters

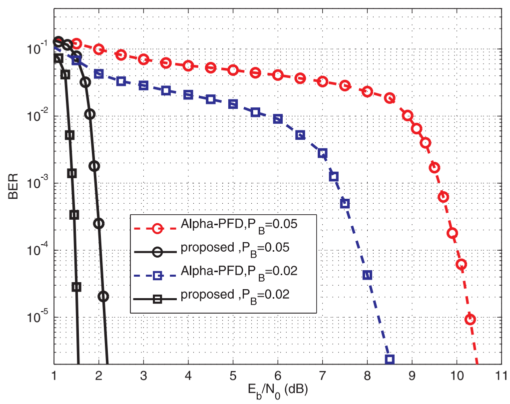

Fig. 4. BER performance comparison for the proposed decoder and alpha-PFD in Markov Gaussian channel ($P_B = 0.02, 0.05;\ R = 100$).

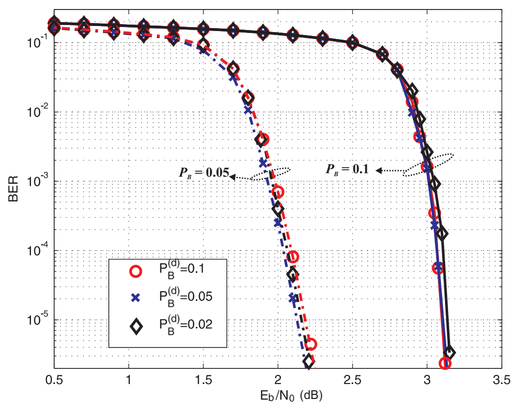

Fig. 5. BER performance for the proposed decoder assuming $R^{(d)} = 200$ and various $P_B^{(d)}$ in Markov Gaussian channel ($P_B = 0.05, 0.1; R = 100$).

V. Conclusion

This paper describes a robust turbo decoding scheme implemented in a Markov Gaussian channel. The branch metric was derived from a two-dimensional trellis, which accommodated the impulse channel state; the bad channel state is characterized by an excessively large power relative to that of the noise in the good channel state. Moreover, the channel state transition was incorporated into the decoding algorithm despite that the statistics on the underlying memory impulse channel were not assumed. Compared with conventional schemes, the proposed decoder yielded a substantial performance enhancement and a performance similar to that of its sub-optimal counterpart requiring statistical impulse knowledge in harsh scenarios, evidencing the robustness of the proposed decoder.

Appendix

Given impulse channel statistics, the sub-optimal decoder yields $\gamma(\tilde{\mathcal{S}}_{l-1}, \tilde{\mathcal{S}}_l)$ (c.f., Fig. 1) by measuring

$$\begin{aligned}&\Pr(\tilde{\mathcal{S}}_l, y_{2l}, y_{2l+1}|\tilde{\mathcal{S}}_{l-1}) = \Pr(m_l = d^{(x,x')})\\ &\cdot\textstyle\sum_{s_{2l}\in\{G,B\}} \Pr(s_{2l}, s_{2l+1}|s_{2l-1}) \Pr(y_{2l}, y_{2l+1}|\tilde{\mathcal{S}}_{l-1}, \tilde{\mathcal{S}}_l)\\ &= \Pr(m_l = d^{(x,x')}) \cdot \textstyle\sum_{s_{2l}\in\{G,B\}} \prod_{k=2l}^{2l+1} P_{s_{k-1}s_k} \Pr(y_k|\xi_k, s_k), \end{aligned} \quad (8)$$

where $\Pr(m_l = d^{(x,x')})$ is the a priori probability, $m_l = d^{(x,x')}$ is the information bit, represented by the modulated symbol ξ_{2l} that governs the transition from the encoder state x (regarding the superstate $\tilde{\mathcal{S}}_{l-1}$) to state x' (with respect to $\tilde{\mathcal{S}}_l$), and ξ_{2l+1} is the accompanied redundancy. By contrast, the proposed decoder, which lacks statistical information on the impulse channel, calculates $\gamma(\tilde{\mathcal{S}}_{l-1}, \tilde{\mathcal{S}}_l)$ using

$$\Pr(m_l = d^{(x,x')}) \cdot \max_{s_{2l}\in\{G,B\}} \Big\{ \textstyle\prod_{k=2l}^{2l+1} P^{(d)}_{s_{k-1}s_k} \Pr(y_k|\xi_k, s_k) \Big\}, \quad (9)$$

where the estimate (7) is used to replace the PDF (c.f., (2), which is used in (8)) when the state $s_k (k \in \{2l, 2l+1\})$ is assumed to be B; besides, the estimates of transition probability (i.e., $P^{(d)}_{s_{k-1}s_k}$) substitute the unknown $P_{s_{k-1}s_k}$.

References

[1] K. L. Blackard, T. S. Rappaport, and C. W. Bostian, “Measurements and models of radio frequency impulsive noise for indoor wireless communications,” *IEEE J. Sel. Areas Commun.*, vol. 11, pp. 991–1001, Sep 1993.

[2] M. Zimmermann and K. Dostert, “Analysis and modeling of impulsive noise in broad-band powerline communications,” *IEEE Trans. Electromagn. Compat.*, vol. 44, no. 1, pp. 249–258, Feb. 2002.

[3] D. Middleton, “Canonical and quasi-canonical probability models of class a interference,” *IEEE Trans. Electromagn. Compat.*, vol. EMC-25, no. 2, pp. 76–106, May 1983.

[4] J. Haring and A. H. Vinck, “Performance bounds for optimum and suboptimum reception under class-A impulsive noise,” *IEEE Trans. Commun.*, vol. 50, no. 7, pp. 1130–1136, 2002.

[5] H. Nakagawa, D. Umehara, S. Denno, and Y. Morihiro, “A decoding for low density parity check codes over impulsive noise channels,” in *Proc. IEEE ISPLC*, 2005, pp. 85 – 89.

[6] D. Umehara, H. Yamaguchi, and Y. Morihiro, “Turbo decoding in impulsive noise environment,” in *Proc. IEEE Global Telecommun. Conf.*, vol. 1, 2004, pp. 194–198.

[7] D.-F. Tseng, Y. Han, W. Mow, P.-N. Chen, J. Deng, and A. Vinck, “Robust decoding for convolutionally coded systems impaired by memoryless impulsive noise,” *IEEE Trans. Commun.*, vol. 61, no. 11, pp. 4640–4652, Sep. 2013.

[8] T. Chuah, “Robust iterative decoding of turbo codes in heavy-tailed noise,” *IEE Proceedings-Commun.*, vol. 152, no. 1, pp. 29–38, 2005.

[9] D.-F. Tseng, T.-R. Tsai, and Y. S. Han, “Robust turbo decoding in impulse noise channels,” in *Proc. IEEE ISPLC*, 2013, pp. 230–235.

[10] D. Fertonani and G. Colavolpe, “A robust metric for soft-output detection in the presence of class-A noise,” *IEEE Trans. Commun.*, pp. 36–40, Jan 2009.

[11] J. Lin, M. Nassar, and B. L. Evans, “Impulsive noise mitigation in powerline communications using sparse Bayesian learning,” *IEEE J. Sel. Areas Commun.*, pp. 1172–1183, July 2013.

[12] M. Nassar, P. Schniter, and B. L. Evans, “A factor graph approach to joint OFDM channel estimation and decoding in impulsive noise environments,” *IEEE Trans. Signal Process.*, pp. 1576–1589, Mar. 2014.

[13] M. Mushkin and I. Bar-David, “Capacity and coding for the Gilbert-Elliott channels,” *IEEE Trans. on Inf. Theory*, pp. 1277–1290, 1989.

[14] A. J. Goldsmith and P. P. Varaiya, “Capacity, mutual information, and coding for finite-state Markov channels,” *IEEE Trans. on Inf. Theory*, vol. 42, no. 3, pp. 868–886, 1996.

[15] C. Pimentel and L. Rego, “Analysis of soft decision decoding of interleaved convolutional codes over burst channels,” in *Proc. IEEE Conf. on Wireless Commun. and Networking*, 1999, pp. 1090 –1094.

[16] J. Garcia-Frias and J. D. Villasenor, “Turbo decoders for Markov channels,” *IEEE Commun. Letters*, vol. 2, no. 9, pp. 257–259, 1998.

[17] L. R. Bahl, J. Cocke, F. Jelinek, and J. Raviv, “Optimal decoding of linear codes for minimizing symbol error rate,” pp. 284–287, Mar. 1974.

[18] J. Garcia-Frias and J. D. Villasenor, “Turbo decoding of Gilbert-Elliot channels,” *IEEE Trans. Commun.*, vol. 50, no. 3, pp. 357–363, 2002.

[19] D. Fertonani and G. Colavolpe, “On reliable communications over channels impaired by bursty impulse noise,” *IEEE Trans. Commun.*, vol. 57, no. 7, pp. 2024–2030, July 2009.

[20] J. Mitra and L. Lampe, “Convolutionally coded transmission over Markov-Gaussian channels: Analysis and decoding metrics,” *IEEE Trans. Commun.*, vol. 58, no. 7, pp. 1939–1949, July 2010.

[21] F. Mengistu, D.-F. Tseng, Y. Han, M. Abera, and L.-C. Chang, “A robust decoding scheme for convolutionally coded transmission through a Markov Gaussian channel,” *to appear in IEEE Trans. Veh. Tech.*

[22] D. Bauso and R. Pesenti, “Generalized person-by-person optimization in team problems with binary decisions,” in *Proc. IEEE American Control Conf.*, June 2008, pp. 717–722.