

Full length article

Turbo receiver design for differentially encoded PSK in bursty impulsive noise channels

Chin-Hung Chen^{a,*}, Boris Karanov^b, Wim van Houtum^{a,c}, Yan Wu^c, Alex Alvarado^a^a Information and Communication Theory (ICT) Lab, Dept. Electrical Engineering, Eindhoven University of Technology, Eindhoven, 5600 MB, The Netherlands^b Communications Engineering Lab (CEL), Karlsruhe Institute of Technology, Karlsruhe, 76187, Germany^c NXP Semiconductors, High Tech Campus 60, Eindhoven, 5656 AE, The Netherlands

ARTICLE INFO

Keywords:

Achievable information rate
Convolutional codes
Differential encoder
Hidden Markov model
Impulsive noise
Symbol detection
Turbo decoding

ABSTRACT

It has been recognized that the impulsive noise (IN) generated by power devices poses significant challenges to wireless receivers. In this paper, we comprehensively assess the achievable information rate (AIR) for the well-established Markov-Middleton IN model with a phase-shift keying (PSK) input sequence across various channel conditions, including matched and mismatched decoding scenarios. Upon determining information-theoretic bounds, we propose an optimal turbo-differentially encoded (DE)-PSK-IN receiver design based on a commonly used commercial transmission setup consisting of a convolutional encoder, bit-level interleaver, and a DE-PSK symbol mapper. We show that by incorporating the differential decoder into the maximum a-posteriori-based (MAP) IN detector, we can significantly enhance the receiver performance with a 4.5 dB gain compared to the conventional MAP-based turbo-PSK-IN receiver and a gap of around 1 dB from the theoretical bounds. We also propose a suboptimal separate receiver design that can be implemented with half the complexity of the joint design while achieving near-optimal performance. We have evaluated the performance of the proposed receiver designs through extensive simulations, demonstrating their effectiveness in real-world scenarios with limited interleaver depth and mismatched number of noise state implementation.

1. Introduction

Electromagnetic interference (EMI) generated by power devices such as switching gears, power lines, and DC-DC converters presents significant challenges to receivers traditionally designed under the additive white Gaussian noise (AWGN) assumption. Measurements have shown that electronic components in electric vehicles (EVs) [1–4] and power substations [5,6] emit EMI that can be detected by high-sensitivity receivers. The impulsive nature of this interference has been shown to strongly affect powerline communications [7,8], digital broadcasting systems [9,10], and wireless communication receivers within power substations [11–13].

Conventional clipping [14] and blanking [15] detectors for IN mitigation are widely adopted due to their simple implementation and the low computational complexity. However, their performance is far from optimal, especially in severe IN channel conditions [16]. To design robust receivers for channels impaired by IN, it is essential to obtain their statistical properties, and therefore, an appropriate IN model is necessary. Among the best-known IN models, the Middleton Class A model [17,18] provides a tractable probability density function (PDF)

based on a Poisson mixture that can be specified by only three parameters. However, it cannot capture the bursty temporal correlations observed in practice. To address this limitation, hidden Markov models (HMMs) have been adopted for IN modeling [8,19,20] to address the correlation between impulsive samples. In [19], a two-state Markov-Gaussian model was proposed, where a comprehensive achievable information rate (AIR) analysis was presented. In [20], the authors combined the HMM process with the Middleton Class A model and proposed the Markov-Middleton model. In [4], it was shown that a Markov-Middleton model with finite IN states effectively represents the EV-induced interference in practice. However, to the best of our knowledge, an information-theoretic analysis of this model has not been thoroughly investigated; therefore, one of the goals of this paper is to fill this gap by conducting a detailed AIR analysis under various channel conditions for both matched and mismatched decoding scenarios.

Optimal IN detectors exploit the IN statistics to compute maximum a posteriori (MAP) symbol decisions. In [16,20–23], several optimal MAP-based IN detectors have been derived for different IN models. To further enhance the receiver performance, powerful coding systems in combination with the IN detector have been extensively investigated to approach

* Corresponding author.

E-mail addresses: c.h.chen@tue.nl (C.-H. Chen), boris.karanov@kit.edu (B. Karanov), w.j.v.houtum@tue.nl (W. van Houtum), yan.wu_2@nxp.com (Y. Wu), a.alvarado@tue.nl (A. Alvarado).<https://doi.org/10.1016/j.phycom.2026.103224>

Received 19 September 2025; Received in revised form 3 January 2026; Accepted 16 June 2026

Available online 28 June 2026

1874-4907/© 2026 The Author(s). Published by Elsevier B.V. This is an open access article under the CC BY license (<http://creativecommons.org/licenses/by/4.0/>).

the AIR for the underlying channel model [16,19,20,24–28]. Receivers for these schemes take the soft information provided by the MAP-based IN detector and pass it to the powerful decoding algorithms, such as low-density parity-check (LDPC) codes [16,19,20,24], convolutional codes [25,26], and turbo-like codes [27,28]. However, previous studies have primarily focused on a basic binary phase-shift keying (BPSK) modulation scheme, mainly addressing the design of its corresponding IN detector and decoder. There has been limited investigation into more complex modulation schemes used in real-world transmission formats for IN receivers.

In wireless communication standards, a differential mapper, where the information is encoded in the phase difference between successive symbols, is commonly used to mitigate the effects of phase ambiguity and for noncoherent detection. For example, digital audio broadcasting, terrestrial digital multimedia broadcasting, and IEEE 802.11g WiFi employ convolutional coding with a differentially encoded (DE)-PSK symbol mapper. Although initially designed for non-coherent detection, research has demonstrated that a coherent DE-PSK demapper can be constructed, wherein the soft information output can be computed using the MAP rule [29]. Since then, turbo-DE-PSK systems, which combine a MAP DE-PSK demapper with a MAP decoder, have been explored in a broad range of literature, see [30–32] and references therein. The recursive nature of the DE-PSK modulator increases the time diversity of a serial turbo-like receiver structure. As a result, a turbo-DE-PSK system achieves significant performance gains compared to a simple turbo-PSK receiver, as reported in [33].

In this paper, we propose to combine a simple differential demapper with the MAP IN detector, forming a MAP-based joint DE-PSK-IN demapper, which can be implemented with the same complexity as a conventional PSK-IN demapper. We further construct a turbo-DE-PSK-IN receiver structure that significantly outperforms its MAP-based turbo-PSK-IN counterpart. This paper starts with an information-theoretic analysis of the Markov-Middleton channel and then performs turbo receiver designs to approach the estimated theoretical bounds. The main contributions are as follows: (a) We provide a unified HMM representation for deriving AIR, MAP detector/demapper/decoder, and iterative decoding metrics through modified branch metrics. This general representation facilitates a streamlined analysis, reveals the underlying commonalities among different algorithms, and demonstrates that they can be systematically constructed under a similar forward-backward framework. (b) We extend the AIR analysis in [19] to adapt it to a finite-state Markov-Middleton channel with a PSK input sequence. Our comprehensive analysis covers both matched and mismatched decoding scenarios, providing guidance on receiver parameter selections. (c) We propose a novel turbo-DE-PSK-IN receiver design that approaches the theoretical limits of the Markov-Middleton channel within a gap of approximately 1 dB, and show that it can achieve a significant performance gain compared to the conventional turbo-PSK-IN receiver proposed in [19,20]. The ultimate performance of the proposed receiver with perfect likelihood input is also presented. Finally, in contribution (d), we further propose a sub-optimal turbo-DE-PSK-IN receiver design that can be implemented with half the complexity of the optimal joint design with near-optimal performance. The proposed receiver with limited interleaver depth and a mismatched state implementation is also investigated and discussed. Note that in this study, we assume the IN parameters are perfectly known at the receiver to establish the achievable performance bounds of the proposed turbo-DE-PSK-IN structure. In practice, the accuracy of the estimated IN parameters can significantly influence the MAP detector's reliability. For comprehensive reviews of parameter estimation methods for Middleton Class A and Markov-Gaussian models, readers are referred to [34–37].

The paper is organized as follows: Section 2 presents the general structure and inference algorithm of HMMs. Section 3 reviews the Markov-Middleton IN channel and its HMM representation. Section 4 illustrates the computation of the AIR for the considered IN channel with a PSK input sequence. Section 5 details the proposed commu-

Table 1

List of main symbols and notation.

Symbol	Description
<i>Transmission Model Parameters</i>	
M	Modulation order
b_k, c_k, d_k	Information, coded, and interleaved coded bits
x_i, z_i	PSK-modulated and differentially encoded symbols
n_i, y_i	Noise realization and observation symbols
<i>Markov-Middleton Noise Model</i>	
W	Number of noise states
A, Λ	Impulsive index and impulsive-to-background power ratio
r	Correlation parameter between noise states
σ_0^2, σ_j^2	Background and state-dependent noise variances
P_j'	Prior state probability of a truncated noise model
P_{ij}	Transition probability of noise states
<i>Hidden State Variables</i>	
s_i	Generic hidden state
w_i	Impulsive noise state
$a_i, \hat{a}_i$	Auxiliary states for PSK-IN and DE-PSK-IN demappers
l_i	Auxiliary state of the differential decoder
g_k	Auxiliary state of the convolutional decoder

cation system and the design methodology for the optimal joint and suboptimal separate turbo-DE-PSK-IN receivers. Section 6 presents the simulation results for AIR and the proposed receiver performance, and finally, Section 7 concludes this paper.

Notation: Throughout this paper, we use lowercase letters to represent a scalar (e.g., s_i) with subscripts indicating a specific time instant and boldface letters to denote a sequence (e.g., $\mathbf{s}_1^T$), with subscripts and superscripts denoting the start and end of the sequence. The set of all states is denoted by calligraphic letters (e.g., S). For the reader's convenience, the primary symbols and parameters used throughout the paper are summarized in Table 1.

2. Hidden Markov model and its efficient inference algorithm

This section provides a general overview of the structure and statistical representation of HMMs. Additionally, we present the utilization of the forward-backward algorithm to compute joint probabilities of observations and latent variables. An HMM structure will be used to describe the Markov-Middleton IN channel in Section 3, and the inference algorithm presented in this section will later be applied to AIR computation in Section 4, while the application to symbol detection, differential demapping, and convolutional decoding based on the MAP rule will be provided in Section 5.

An HMM is a statistical model that captures the temporal behavior of a sequence of latent variables (states) $\mathbf{s}_1^T = (s_1, s_2, \dots, s_T)$ from a Markov process via a sequence of state-dependent observable variables (observations) $\mathbf{y}_1^T = (y_1, y_2, \dots, y_T)$. Here, s_i denotes the state realization at time t , which is drawn from a time-invariant finite-state set

$$s_i \in S \triangleq \{0, 1, \dots, |S| - 1\}, \quad (1)$$

where $|S|$ is the size of the state space. A general representation of the structure of an HMM is shown in Fig. 1 with the transition probability $p(s_i = j | s_{i-1} = i)$ representing the probability of moving from state i at time $t-1$ to state j at time t . The observation likelihood $p(y_i | s_i = j)$, on the other hand, describes the probability of an observation being generated from a state j at time t .

The Markovian structure of an HMM allows an efficient inference of the hidden states through the well-known Bahl-Cocke-Jelinek-Raviv (BCJR) algorithm [38]. Specifically, the inference problem aims at calculating the posterior belief of a latent state s_i at a specific time instant given the full sequence of observations as

$$p(s_i | \mathbf{y}_1^T) \propto p(s_i, \mathbf{y}_1^T) = \sum_{s_{t-1} \in S} p(s_i, s_{t-1}, \mathbf{y}_1^T). \quad (2)$$

The right-most side of (2) represents the joint distribution of the state pairs (s_i, s_{t-1}) at a specific time t and the whole received sequence $\mathbf{y}_1^T$.

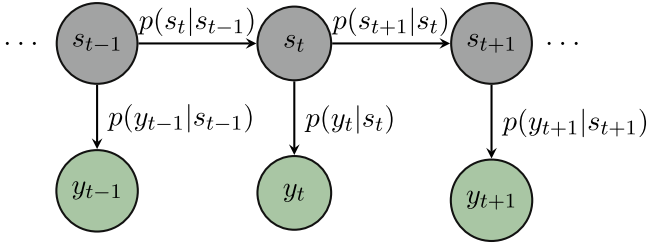


Fig. 1. Structure of a hidden Markov model, showing the factorization and conditional independencies implied by the model.

We can further factorize the joint probability as

$$p(s_t, s_{t-1}, \mathbf{y}_1^T) = \underbrace{p(\mathbf{y}_{t+1}^T | s_t)}_{\beta(s_t)} \cdot \underbrace{p(y_t, s_t | s_{t-1})}_{\gamma(s_t, s_{t-1})} \cdot \underbrace{p(s_{t-1}, \mathbf{y}_1^{t-1})}_{\alpha(s_{t-1})}. \quad (3)$$

In (3), $\alpha(s_{t-1})$, $\beta(s_t)$, and $\gamma(s_t, s_{t-1})$ represent $p(s_{t-1}, \mathbf{y}_1^{t-1})$, $p(\mathbf{y}_{t+1}^T | s_t)$, and $p(y_t, s_t | s_{t-1})$, respectively, corresponding to forward recursion, backward recursion, and branch metrics.

Using Bayes' rule and the Markovian property, the branch metric γ can be further derived as

$$\gamma(s_t, s_{t-1}) \triangleq p(y_t, s_t | s_{t-1}) = p(y_t | s_t) \cdot p(s_t | s_{t-1}) \quad (4)$$

which can be seen as the multiplication of the observation likelihood and the state transition probability. The forward recursion representation of α and the backward recursion β can be written as

$$\alpha(s_t) \triangleq p(s_t, \mathbf{y}_1^t) = \sum_{s_{t-1} \in S} \gamma(s_t, s_{t-1}) \cdot \alpha(s_{t-1}), \quad (5)$$

and

$$\beta(s_{t-1}) \triangleq p(\mathbf{y}_t^T | s_{t-1}) = \sum_{s_t \in S} \beta(s_t) \cdot \gamma(s_t, s_{t-1}). \quad (6)$$

The forward-backward computation (2)–(6) of the posterior belief comprises the conventional BCJR algorithm (see [38] for detailed derivations). In this paper, we will examine several adaptations of this algorithm, each utilizing the same forward-backward framework but with modifications to the branch metrics γ that are tailored to specific applications.

3. Markov–Middleton impulsive noise model

In this section, we summarize the Markov–Middleton model, initially developed for powerline communications [20], which captures the statistical behavior of bursty IN by combining the conventional Middleton Class A model [17] with an HMM. The integration of HMM introduces memory between noise samples, with state transition probabilities characterizing the correlation between consecutive bursts.

The Markov–Middleton model is characterized by four parameters: the background noise (BN) variance σ_0^2 , the impulsive index A , which quantifies the probability of impulsive (non-BN) events, the impulsive-to-background noise power ratio Λ , which determines the intensity of IN compared to BN, and the noise correlation parameter r , which describes the correlation between consecutive noise samples. The noise realizations generated by the Markov–Middleton model for different parameters are depicted in Fig. 2a–d, and will be further explained in the following.

The statistical properties of the IN samples n_t are completely determined by the noise state realization at time t , which is defined as

$$w_t \in \mathcal{W} \triangleq \{0, 1, \dots, |\mathcal{W}| - 1\}, \quad (7)$$

where $w_t = 0$ represents the BN state and $|\mathcal{W}| = W$ is the total number of noise states. The PDF of n_t conditioned on a specific noise state

$w_t = j$ follows a complex Gaussian distribution $\mathcal{CN}(0, \sigma_j^2)$, which can be expressed as

$$p(n_t | w_t = j) = \frac{1}{\pi \sigma_j^2} \exp\left(-\frac{|n_t|^2}{\sigma_j^2}\right), \quad (8)$$

with σ_j^2 denoting the variance of the complex-valued noise sample n_t in state $w_t = j$.

In the original Middleton Class A model [17], the number of noise states is assumed to be infinity ($W \rightarrow \infty$). Therefore, the generic PDF of the Middleton Class A noise sample can be written as the expectation over all possible states as

$$\begin{aligned} p(n_t) &= \sum_{j=0}^{\infty} P_j \cdot p(n_t | w_t = j) \\ &= \sum_{j=0}^{\infty} \frac{P_j}{\pi \sigma_j^2} \exp\left(-\frac{|n_t|^2}{\sigma_j^2}\right), \end{aligned} \quad (9)$$

where

$$P_j = p(w_t = j) = \exp(-A) A^j / j! \quad (10)$$

represents the prior probability of being at a specific noise state. From (10), we can observe that when $j = 0$, the probability of being at BN state $w_t = 0$ is $P_B = P_0 = \exp(-A)$ and the probability of being at any of the impulsive states $j > 0$ is $P_I = 1 - \exp(-A)$. In other words, a larger impulsive index A results in a larger P_I , and thus, a larger proportion of samples with large variance (σ_j^2 , $j > 0$) are generated in a noise sequence $\mathbf{n}_1^T$. As shown in Fig. 2a and b, when the impulsive index A is increased from 0.1 to 0.3, the number of samples in IN states ($j > 0$) increases.

As illustrated in (10), the probability of being in the j -th noise state decreases as j increases. Therefore, it is sufficient to use a truncated Middleton Class A model with a finite number of states W to represent the IN channel with

$$p'(n_t) = \sum_{j=0}^{W-1} \frac{P'_j}{\pi \sigma_j^2} \exp\left(-\frac{|n_t|^2}{\sigma_j^2}\right), \quad (11)$$

where the probability of being in state j is now written as

$$P'_j = \frac{P_j}{\sum_{i=0}^{W-1} P_i}. \quad (12)$$

The denominator in (12) is a normalization factor to ensure the total probability equals 1.

The noise variance in (11) is defined as

$$\sigma_j^2 = \left(1 + \frac{j\Lambda}{A}\right) \sigma_0^2, \quad (13)$$

where

$$\Lambda = \sigma_I^2 / \sigma_B^2 \quad (14)$$

is the impulsive-to-background average noise power ratio.¹ The average BN power is $\sigma_B^2 = P_0 \sigma_0^2$ and the average IN power σ_I^2 can be written as

$$\sigma_I^2 = \sum_{j=1}^{W-1} P'_j \sigma_j^2.$$

From (14), we can see that as the impulsive-to-background noise ratio Λ increases, the difference between the IN power and the BN power increases. This is clearly shown by comparing the noise realizations (black lines) in Fig. 2b and c.

To describe the burstiness (memory) of the noise observed in practice, the Markov–Middleton model [20] introduces a parameter $r \in [0, 1]$

¹ In the original paper [17], the power ratio is defined as $\Gamma = \sigma_B^2 / \sigma_I^2$.

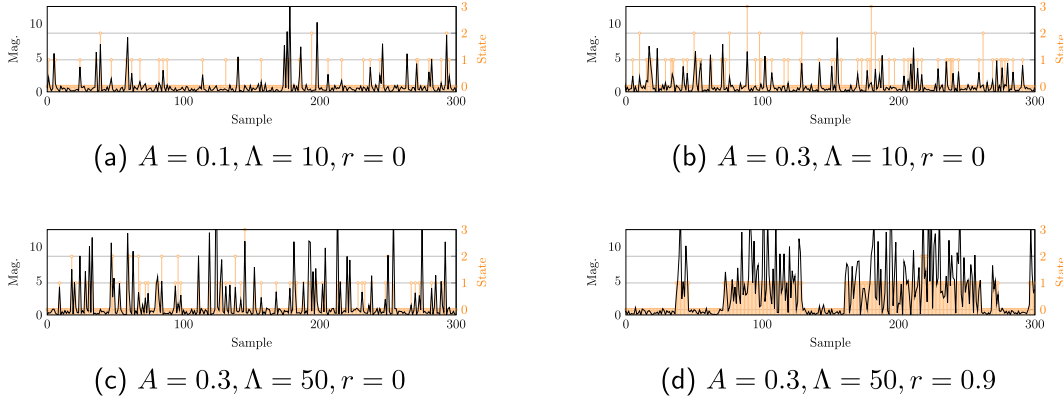


Fig. 2. Noise realization from a 4-state Markov-Middleton impulsive noise model with a background noise variance $\sigma_0^2 = 1$ for impulsive index $A = 0.1$ and $A = 0.3$, impulsive-to-background noise power ratio $\Lambda = 10$ and $\Lambda = 50$, and correlation parameters $r = 0$ and $r = 0.9$. The orange vertical bar denotes the index j of the state realizations $w_i = j$ for generating the corresponding noise samples.

to establish a correlation between consecutive noise samples. Specifically, the state transition matrix of the Markov-Middleton noise model is constructed as

$$P_{ij} \triangleq p(w_i = j \mid w_{i-1} = i) = \begin{cases} r + (1-r) \cdot P'_{ij}, & i = j \\ (1-r) \cdot P'_{ij}, & \text{otherwise.} \end{cases} \quad (15)$$

From (15), we can see that the larger the correlation parameter r , the more likely the following noise sample will stay in the same state. Namely, the bursty behavior is more prominent. In Fig. 2d, the noise samples exhibit strong correlations ($r = 0.9$), where the noise tends to remain in the same state (vertical orange bars) for consecutive samples. In contrast, in Fig. 2a-c, the noise samples are randomly chosen from $w_i \in \mathcal{W}$ with probability P'_{ij} defined in (12). Note that when $r = 0$, the Markov-Middleton model is reduced to a truncated memoryless Middleton Class A channel.

4. Achievable information rate

This section presents an information-theoretic analysis for transmission over channels with IN following the Markov-Middleton model introduced in Section 3. This is an essential reference for practical receiver design operating in specific channel conditions. We use a simulation-based method to calculate the AIR between the PSK-modulated input sequence $\mathbf{x}_1^T$ and the channel output sequence $\mathbf{y}_1^T$ with an input-output relationship defined as

$$y_i = x_i + n_i, \quad (16)$$

where

$$x_i \in \mathcal{X} = \{e^{j2\pi i/M} \mid i = 0, 1, \dots, M-1\} \quad (17)$$

denotes the transmitted M -ary PSK symbol and n_i follows the Markov-Middleton noise model described in Section 3, which is drawn from the distribution defined in (11). The computation of AIR for channels with memory is explained in detail in [39], and its application to a two-state Markov-Gaussian channel is described in [19]. The algorithm and the computation steps are briefly summarized as follows.

An AIR is the number of information bits that can be reliably transmitted and decoded, with an arbitrarily low error probability per channel use, for a specific modulation format. Specifically, the AIR we consider in this work is the mutual information $I(\mathbf{X}; \mathbf{Y})$ between the discrete input random process $\mathbf{X}$ and the continuous output random process $\mathbf{Y}$, which can be estimated via

$$I(\mathbf{X}; \mathbf{Y}) \approx \hat{I}(\mathbf{X}; \mathbf{Y}) = \frac{1}{T} \log_2 p(\mathbf{y}_1^T \mid \mathbf{x}_1^T) - \frac{1}{T} \log_2 p(\mathbf{y}_1^T) \quad (18)$$

with a very long sequence T . To calculate (18) for channels with memory, [39] introduces a simulation-based computation through a sum-product algorithm with the help of an auxiliary state process. In this paper, we define the auxiliary state realization for a PSK-modulated symbol over a Markov-Middleton IN channel as

$$a_i \triangleq f(x_i \in \mathcal{X}, w_i \in \mathcal{W}) \in \mathcal{A} \triangleq \{0, 1, \dots, |\mathcal{A}| - 1\}, \quad (19)$$

where $f: \mathcal{X} \times \mathcal{W} \rightarrow \mathcal{A}$ maps each pair (x_i, w_i) in the product space $\mathcal{X} \times \mathcal{W}$ into a unique state in $\mathcal{A}$. The size of the auxiliary state space is thus $|\mathcal{A}| = |\mathcal{X}| \times |\mathcal{W}| = M \times W$. In Fig. 3a, we show an example of the trellis section and the auxiliary state for a QPSK transmission (i.e., $M = 4$) over a 2-state (i.e., $W = 2$) Markov-Middleton channel.

The transition probability of the auxiliary states considers the prior PSK symbol probability $p(x_i)$ and the noise state transition matrix defined in (15), which can be expressed as

$$p(a_i \mid a_{i-1}) = P_{ij} \cdot p(x_i), \quad w_i = j, w_{i-1} = i. \quad (20)$$

The channel likelihood function concerning the auxiliary states (19) can be written as

$$p(y_i \mid a_i) = \frac{1}{\pi \sigma_j^2} \exp\left(-\frac{|y_i - x_i|^2}{\sigma_j^2}\right), \quad w_i = j. \quad (21)$$

After determining the auxiliary states with their statistical representation (20) and (21), the PDF of the channel output sequence $p(\mathbf{y}_1^T)$ can be calculated via

$$p(\mathbf{y}_1^T) = \sum_{a_T \in \mathcal{A}} p(a_T, \mathbf{y}_1^T) = \sum_{a_T \in \mathcal{A}} \alpha(a_T), \quad (22)$$

using forward recursion (5), with a branch metric

$$\gamma_a(a_i, a_{i-1}) \triangleq p(y_i \mid a_i) \cdot p(a_i \mid a_{i-1}). \quad (23)$$

To compute the PDF of the channel output sequence conditioned on the knowledge of the transmitted symbol sequence $p(\mathbf{y}_1^T \mid \mathbf{x}_1^T)$, we write

$$p(\mathbf{y}_1^T \mid \mathbf{x}_1^T) = \frac{p(\mathbf{y}_1^T, \mathbf{x}_1^T)}{p(\mathbf{x}_1^T)} = \frac{\sum_{a_T \in \mathcal{A}} p(a_T, \mathbf{y}_1^T, \mathbf{x}_1^T)}{p(\mathbf{x}_1^T)} \quad (24)$$

with the help of the auxiliary state. The computation of the joint probability $p(a_T, \mathbf{y}_1^T, \mathbf{x}_1^T)$ can be carried out efficiently using the forward recursion (5) with a modified branch matrix defined as

$$\begin{aligned} \gamma_{a,x}(a_i, a_{i-1}) &\triangleq p(y_i, x_i, a_i \mid a_{i-1}) \\ &= p(y_i \mid x_i, a_i, a_{i-1}) \cdot p(a_i \mid a_{i-1}, x_i) \cdot p(x_i \mid a_{i-1}) \\ &= p(y_i \mid a_i) \cdot p(a_i \mid a_{i-1}, x_i) \cdot p(x_i), \end{aligned} \quad (25)$$

where the transition probability conditioned on the knowledge of transmitted symbols is written as

$$p(a_i \mid a_{i-1}, x_i = m \in \mathcal{X})$$

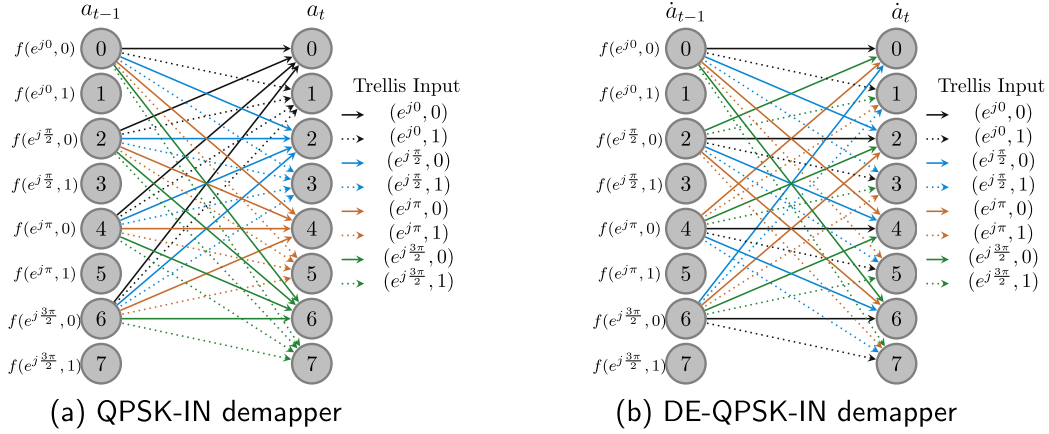


Fig. 3. Trellis sections of an impulsive noise detector with a 2-state Markov–Middleton model for (a) QPSK and (b) DE-QPSK transmissions, respectively. For clarity, we only show the state transitions from states 0, 2, 4, and 6.

$$= \begin{cases} P_{ij}, & x_t = m, w_t = j, w_{t-1} = i, \\ 0, & x_t \neq m, \end{cases} \quad (26)$$

where we use m to denote a specific PSK symbol element in $\mathcal{X}$. Given (22) and (24), the estimated AIR, represented by $\hat{I}(\mathbf{X}; \mathbf{Y})$ can be computed via (18).

5. System design

In Fig. 4, we show the overall structure and components of the transmission system over a Markov–Middleton IN channel considered in this paper. This section will first introduce the bit-interleaved convolutional-coded DE-PSK transmitter and then detail the design methodology for the corresponding turbo-DE-PSK-IN receiver.

5.1. Bit-interleaved convolutional-coded DE-PSK transmitter

The transmitter employs a serially concatenated bit-interleaved coded modulation scheme. In this scheme, a rate R convolutional encoder with an L -stage shift register takes as input the information bit sequence $\mathbf{b}_1^K$ and generates the coded bit sequence $\mathbf{c}_1^{K/R}$, where $b_k, c_k \in \{0, 1\}$, based on a certain generator constraint. A bit-level interleaver (Π) is used to permute the convolutional encoder output, where the interleaved coded bit sequence is represented as $\mathbf{d}_1^{K/R}$. The bit-interleaver is designed to transform burst errors into single-bit errors and distribute errors to improve error correction capability, which is essential for turbo decoding. After the interleaver, a symbol sequence $\mathbf{x}_1^T$ is formed by a memoryless PSK symbol mapper, which maps a number of binary coded bits to a PSK symbol with the symbol time index denoted by the letter t . Finally, a differential symbol encoder modulates the incoming PSK symbols and produces

$$z_t = x_t \cdot z_{t-1}, \quad (27)$$

with $z_0 = 1$, so the received signal is

$$y_t = x_t \cdot z_{t-1} + n_t = z_t + n_t. \quad (28)$$

This differential modulation introduces the recursive channel input-output relationship in (28).

Differential encoders are used in various wireless communication standards. One of the most significant advantages of a differential encoder is that it avoids the need for the receiver to have an exact phase reference, as the information is encoded in the phase difference of consecutive symbols. In this paper, we treat the differential encoder as a rate-one convolutional code [30,33] and thus create a serial-like turbo structure with the convolutional encoder. Its recursive structure largely enhances the time diversity in the serial turbo receiver design, leading

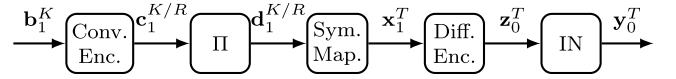
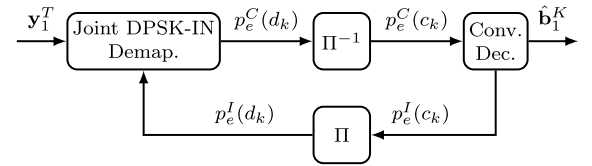
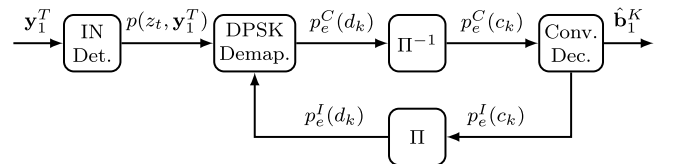


Fig. 4. Block diagram of a convolutional-coded bit-interleaved DE-PSK transmission over an impulsive noise channel.



(a) Joint receiver design



(b) Separate receiver design

Fig. 5. Block diagrams of (a) optimal joint and (b) the suboptimal separate turbo-DE-PSK-IN receiver designs.

to significant performance improvement compared to memoryless PSK modulation schemes [31,32]. This study assumes a coherent detection scenario with a known symbol phase. As a result, we will ignore the reference symbols z_0 and y_0 in the subsequent analysis.

5.2. Optimal joint turbo-DE-PSK-IN receiver

5.2.1. MAP DE-PSK-IN demapper

In the joint detector-demapper design shown in Fig. 5a, we construct a super-trellis where the structure of the Markov–Middleton IN detector and the DE-PSK demapper are considered jointly. We define the state realization of the joint super-trellis as

$$\hat{a}_t \triangleq f(z_t \in \mathcal{X}, w_t \in \mathcal{W}) \in \mathcal{A}, \quad (29)$$

which is drawn from the same state space as a_t in (19) and has a similar form of transition probability

$$p(\hat{a}_t | \hat{a}_{t-1}) = P_{ij} \cdot p(x_t), \quad w_t = j, w_{t-1} = i. \quad (30)$$

However, instead of the PSK symbol x_t , the state $\hat{a}_t$ is now associated with the differentially encoded symbol z_t with a transition driven by x_t .

The transition of the joint super-trellis section for a DE-QPSK modulation with a 2-state Markov–Middleton model is visualized in Fig. 3b. The likelihood function, on the other hand, can be written as

$$p(y_t | \hat{a}_t) = \frac{1}{\pi \sigma_j^2} \exp\left(-\frac{|y_t - z_t|^2}{\sigma_j^2}\right), \quad w_t = j. \quad (31)$$

We can now define the branch metric of the joint DE-PSK-IN detector-demapper as

$$\gamma_{\hat{a}}(\hat{a}_t, \hat{a}_{t-1}) \triangleq p(y_t | \hat{a}_t) \cdot p(\hat{a}_t | \hat{a}_{t-1}). \quad (32)$$

Given the transition probability (30) and the likelihood function (31) for computing the branch metric (32), the joint probability $p(\hat{a}_t, \hat{a}_{t-1}, \mathbf{y}_1^T)$ of the joint DE-PSK-IN detector-demapper is computed via (3)–(6). A MAP-based soft-output detection is then performed through

$$p(x_t, \mathbf{y}_1^T) = \sum_{(\hat{a}_{t-1} \rightarrow \hat{a}_t) : x_t} p(\hat{a}_t, \hat{a}_{t-1}, \mathbf{y}_1^T).$$

where $(\hat{a}_{t-1} \rightarrow \hat{a}_t) : x_t$ denotes all the state transitions $(\hat{a}_{t-1} \rightarrow \hat{a}_t)$ that are driven by the input symbol x_t .

5.2.2. Iterative turbo decoder

To perform bit deinterleaving, we need to convert the symbol probability to bit probability as

$$p(d_k, \mathbf{y}_1^T) = \Phi^{-1}[p(x_t, \mathbf{y}_1^T)] \quad (33)$$

and

$$p_e^I(d_k) = \Phi^{-1}[p_e^I(x_t)],$$

where Φ^{-1} denotes the inverse symbol mapping function that converts the probability of the PSK symbol x_t back to the probability of the corresponding bit d_k .

One essential part of the turbo decoding principle is to pass only the extrinsic information to each MAP decoder (demapper) block. Extrinsic information refers to the information provided by one decoder (demapper) about one specific bit (symbol), excluding the information directly used to derive it. In other words, it represents new information that can help the other decoder (demapper) refine its estimates. Here, we define the extrinsic likelihood information from the joint DE-PSK-IN demapper as the joint probability $p(d_k, \mathbf{y}_1^T)$ divided by the prior information regarding d_k as

$$p_e^C(d_k) \triangleq \frac{p(d_k, \mathbf{y}_1^T)}{p_e^I(d_k)}.$$

After performing the bit-deinterleaving, we can derive the extrinsic likelihood information of the coded bits as

$$p_e^C(c_k) = \Pi^{-1}[p_e^C(d_k)],$$

which is then fed to the convolutional decoder, whose state realization is defined as

$$g_k \triangleq f(b_k, b_{k-1}, \dots, b_{k-L} \in \mathcal{U}^L) \in \mathcal{G} \triangleq \{0, 1, \dots, 2^L - 1\},$$

where $f : \mathcal{U}^L \rightarrow \mathcal{G}$ maps $\mathcal{U}^L$ to distinct states in $\mathcal{G}$. The modified branch metric for the convolutional decoder is then written as

$$\gamma_g(g_k, g_{k-1}) \triangleq p_e^C(c_k) \cdot p(b_k), \quad (g_{k-1} \rightarrow g_k) : b_k. \quad (34)$$

The extrinsic likelihood information $p_e^C(c_k)$ produced by the DE-PSK demapper acts as the convolutional decoder's likelihood information, and $p(b_k) = 1/2$ is the prior bit probability. Again, we compute the joint probability $p(g_k, g_{k-1}, \mathbf{y}_1^T)$ via (3)–(6) using (34) and obtain the joint probability of the information bit as

$$p(b_k, \mathbf{y}_1^T) = \sum_{(g_{k-1} \rightarrow g_k) : b_k} p(g_k, g_{k-1}, \mathbf{y}_1^T),$$

and the joint probability of the coded bit as

$$p(c_k, \mathbf{y}_1^T) = \sum_{(g_{k-1} \rightarrow g_k) : c_k} p(g_k, g_{k-1}, \mathbf{y}_1^T).$$

To compute the extrinsic prior information produced by the convolutional decoder, we divide the joint probability $p(c_k, \mathbf{y}_1^T)$ by the likelihood function $p_e^C(c_k)$ used for deriving it as

$$p_e^I(c_k) \triangleq \frac{p(c_k, \mathbf{y}_1^T)}{p_e^C(c_k)}.$$

After the bit interleaver

$$p_e^I(d_k) = \Pi[p_e^I(c_k)],$$

we can derive the extrinsic prior information of the PSK symbol via bit-to-symbol mapping

$$p_e^I(x_t) = \Phi[p_e^I(d_k)].$$

The extrinsic symbol information $p_e^I(x_t)$ generated by the convolutional decoder is then fed back as the prior information in (30) of the joint DE-PSK-IN demapper for the next turbo iteration.

5.3. Suboptimal separate turbo-DE-PSK-IN receiver

While the joint design delivers the optimal MAP solution by iteratively updating the extrinsic information for both the IN detector and DE demapper simultaneously, the computational complexity increases significantly with the number of turbo iterations. Therefore, we introduce in this section a suboptimal design that separates the IN detector from the turbo-DE-PSK demapper-decoder to reduce the operational complexity. This separate receiver design block diagram is shown in Fig. 5b.

5.3.1. MAP IN detector

The states of the standalone MAP-based IN detector are defined as in (19), and a trellis section example is shown in Fig. 3a. Given the transition probability (20), likelihood function (21), and the branch metric (23), the joint probability $p(a_t, a_{t-1}, \mathbf{y}_1^T)$ of the IN detector is computed via (3)–(6). A MAP-based soft-output detection is then performed through

$$p(z_t, \mathbf{y}_1^T) = \sum_{(a_{t-1} \rightarrow a_t) : z_t} p(a_t, a_{t-1}, \mathbf{y}_1^T). \quad (35)$$

where $(a_{t-1} \rightarrow a_t) : z_t$ denotes all the state transitions $(a_{t-1} \rightarrow a_t)$ that are driven by the input differentially encoded symbol z_t .

5.3.2. MAP DE demapper

The joint probability (35) generated from the IN detector is used as a soft input, representing the channel information, to the differential symbol demapper. We define the state realization of the differential decoder as

$$l_t \triangleq f(z_t \in \mathcal{X}) \in \mathcal{L} \triangleq \{0, 1, \dots, M - 1\}, \quad (36)$$

where $f : \mathcal{X} \rightarrow \mathcal{L}$ maps $z_t \in \mathcal{X}$ to a specific state in $\mathcal{L}$. The trellis section of a DE-QPSK decoder is shown in Fig. 6. Note that although the state realization l_t is a mapping of the DE-PSK symbol z_t , the trellis input is the PSK modulated symbol x_t , which reveals the structure of the differential encoder in (27).

Using the state definitions (36) and the joint probability $p(z_t, \mathbf{y}_1^T)$ provided by the IN detector, we can now rewrite (4) and express the trellis branch metric for the differential decoder as

$$\gamma_p(l_t, l_{t-1}) \triangleq p(z_t, \mathbf{y}_1^T) \cdot p_e^I(x_t), \quad (l_{t-1} \rightarrow l_t) : x_t. \quad (37)$$

where the modified branch metric uses $p(z_t, \mathbf{y}_1^T)$ as the likelihood information and uses the extrinsic prior symbol information $p_e^I(x_t)$ to denote the transition probability $p(l_t | l_{t-1})$ when the state transition $(l_{t-1} \rightarrow l_t)$ is driven by the input symbol x_t . Note that we set $p_e^I(x_t) = p(x_t) = 1/M$ in the first iteration. Using the modified branch metric (37), the joint probability of the DE-PSK demapper $p(l_t, l_{t-1}, \mathbf{y}_1^T)$ is derived through (3)–(6), and the joint probability of the PSK symbol is then obtained by

$$p(x_t, \mathbf{y}_1^T) = \sum_{(l_{t-1} \rightarrow l_t) : x_t} p(l_t, l_{t-1}, \mathbf{y}_1^T).$$

Subsequently, the iterative processing is the same as in Section 5.2.2.

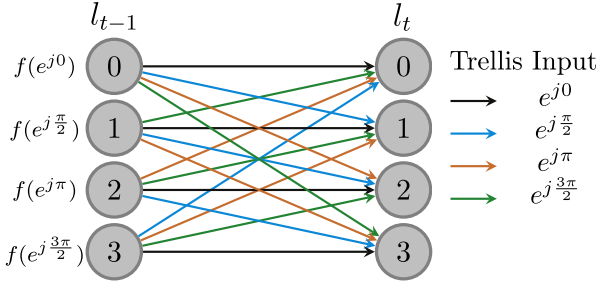


Fig. 6. Trellis section of a standalone DE-QPSK demapper.

Table 2

Number of multiplications for separate and joint turbo receiver designs with $M = 4$, $W = 4$, and $L = 2$.

	$I = 0$	$I = 1$	$I = 2$	$I = 3$	$I = 10$
C_{jo}	1024T	2048T	3072T	4096T	11264T
C_{sep}	1056T	1600T	2144T	2688T	6496T

5.4. Computational complexity analysis

This subsection offers a general comparison of the complexity between the joint and separate system designs. Since the main computational burden comes from the multiplication operation in the BCJR algorithm, we will concentrate on the complexity related to the number of multiplications used in the BCJR algorithm for two different realizations of the turbo systems.

Consider a length T sequence (T trellis sections) processed by a BCJR algorithm with N trellis states. For each state, it needs to consider transitions from all possible previous states, which results in a N^2 multiplication per trellis section. Therefore, the overall complexity of a BCJR operation is given by $2TN^2$, where the constant 2 accounts for the forward and backward operations. For transmission of a general cardinality M and W state Markov–Middleton channel, the super-trellis sections of the joint DE-PSK-IN demapper (Section 5.2) require MW states. Considering the turbo receiver operates over I iterations, we can then write the computational complexity of the joint receiver as

$$C_{\text{jo}} = 2T \cdot (I + 1)(M^2W^2 + M^{2L}).$$

The separate design (Section 5.3), on the other hand, requires an MW state MAP IN detector, an M state MAP DE-PSK demapper, and an M^L state MAP convolutional decoder. We can then write its computational complexity as

$$C_{\text{sep}} = 2T \cdot (M^2W^2 + (I + 1)(M^2 + M^{2L})).$$

In Table 2, we show the computational complexity of C_{jo} and C_{sep} for a transmission of a $L = 2$ convolutional code and a DE-QPSK ($M = 4$) modulation over a 4-state ($W = 4$) Markov–Middleton model. When operating without turbo iterations ($I = 0$), the super-trellis design in the joint system has a slightly lower complexity than a separate design. As the number of iterations increases, the computational cost of the joint design increases significantly and exceeds that of a separate design from the second turbo iteration. After 10 iterations, the suboptimal separate design has nearly half the computational complexity of the optimal receiver design.

6. Simulation results

6.1. AIR analysis

In this section, we report a detailed AIR analysis for the finite-state Markov–Middleton model with a PSK transmission scheme. Specifically, we simulated 1000 independent sequences with a length of $T = 10^6$ symbols to compute the estimated AIR, where the transmitted sequence

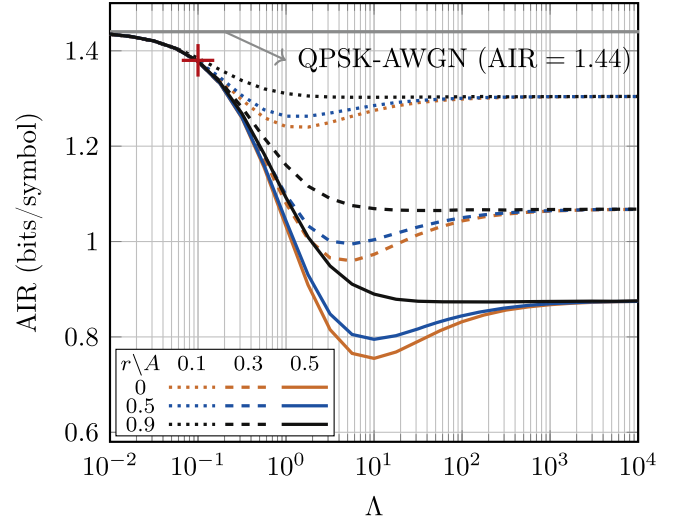


Fig. 7. AIRs versus impulsive-to-background noise ratio Λ for different impulsive indices $A = 0.1, 0.3$, and 0.5 and correlation parameters $r = 0, 0.5$, and 0.9 for a QPSK modulation scheme with a 4-state Markov–Middleton impulsive noise model with SNR = 3 dB.

$\mathbf{x}_1^T$ consists of QPSK ($M = 4$) modulated symbols with a symbol set $\mathcal{X}$ following (17). The channel output sequence $\mathbf{y}_1^T$ is obtained using (16), where the noise sequence $\mathbf{n}_1^T$ is generated by a 4-state Markov–Middleton model with noise statistics following (11)–(13), given parameters A , Λ , r , and σ_0^2 . The choice of $W = 4$ is based on both empirical evidence from our field-testing data [4] and a balanced trade-off between model accuracy and computational complexity. Note that although the mathematical derivations of this paper are in the probability domain, the actual implementation is in the log domain to prevent numerical instability. The simulated AIR values with different Markov–Middleton channel parameters are shown in Fig. 7 for an SNR of 3 dB. Here, the SNR is defined as the transmitted symbol power over the background noise power

$$\text{SNR} = E\{|\mathbf{x}_1|^2\} / \sigma_0^2.$$

To help illustrate the behavior of the AIR over impulsive-to-background noise power ratio Λ , we also plot the likelihood function for a memoryless ($r = 0$) channel with different impulsive-to-background noise power ratios ($\Lambda = 10^{-2}, 10^0, 10^4$) in Fig. 8a–c. For simplicity, we choose a two-state (background and impulsive states) channel with impulsive index $A = 0.3$ and BPSK for the likelihood function analysis, which does not affect the generality of the presentation.

Results from Fig. 7 show that a BN-limited regime can be found when $\Lambda < 10^{-1}$ (indicated by red cross), where the noise generated by an IN state is barely distinguishable from the noise generated by a BN state ($\sigma_0^2 \approx \sigma_j^2$, $j = 1, 2, 3$) as shown in the likelihood function from Fig. 8a. In this case, the performance of the Markov–Middleton channel approaches that of the AWGN channel, with an AIR of 1.44 bits/symbol, regardless of the impulsive channel parameters A and r . From Fig. 7, we can observe that as Λ increases, the channel transitions into an IN-limited regime, where A and r become crucial in determining the AIR. A non-monotonic behavior of the AIR over Λ is shown for channels with correlation parameter $r < 0.9$ in Fig. 7, where there exist inflection points where a minimum AIR occurs. Before the inflection point, the AIR quickly worsens as Λ increases. As shown in Fig. 8a, the IN state increases the uncertainty of the overall likelihood function due to its non-negligible long tail. However, after the inflection point, the impact of the impulsive state on the overall likelihood is reduced. In particular, when $\Lambda > 10^3$, the AIR for all the channels converges to a constant value. This behavior can be explained by examining the likelihood function in Fig. 8c. As the variance of the IN state increases, its likelihood

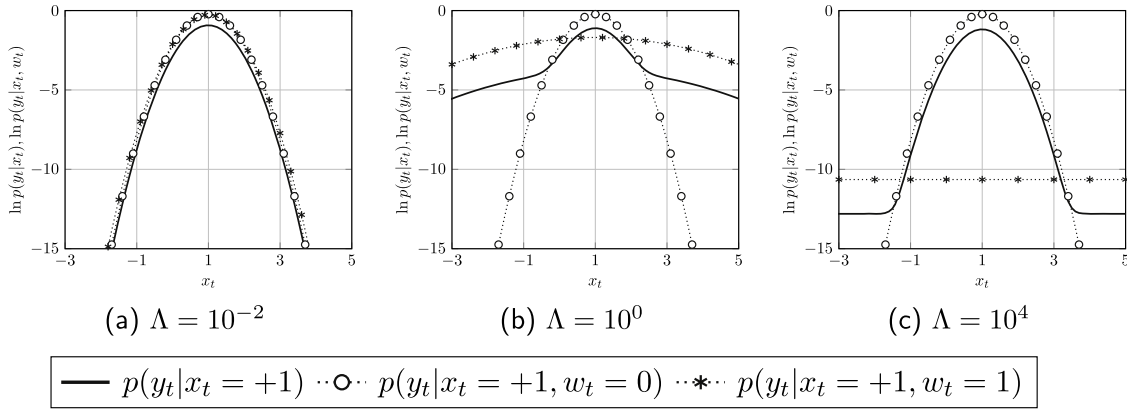


Fig. 8. The likelihood functions $p(y_t|x_t = +1)$, $p(y_t|x_t = +1, w_t = 0)$, and $p(y_t|x_t = +1, w_t = 1)$ of a 2-state memoryless ($r = 0$) Markov-Middleton model with BPSK modulation are shown for (a) $\Lambda = 10^{-2}$, (b) $\Lambda = 10^0$, and (c) $\Lambda = 10^4$.

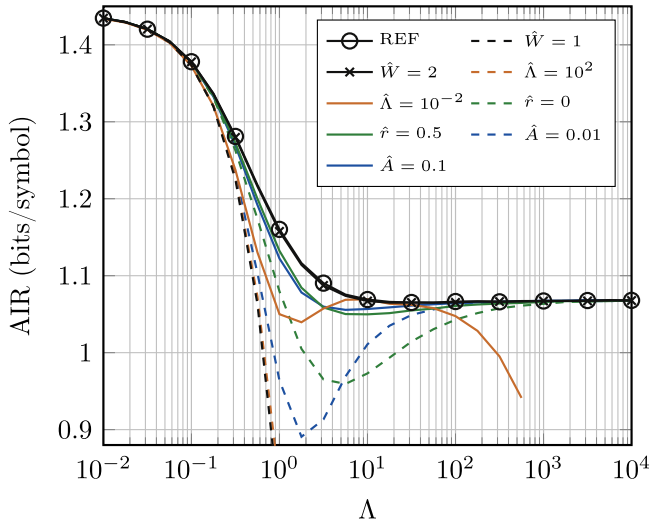


Fig. 9. AIRs versus impulsive-to-background noise ratio Λ for mismatched $\hat{A} = 0.1$ and 0.01 , $\hat{r} = 0$ and 0.5 , $\hat{\Lambda} = 10^{-2}$ and 10^2 , and $\hat{W} = 1$ and 2 , for a QPSK modulation scheme with SNR = 3 dB. The AIR with actual channel parameters $A = 0.3$, $r = 0.9$, and $W = 4$ is denoted as REF.

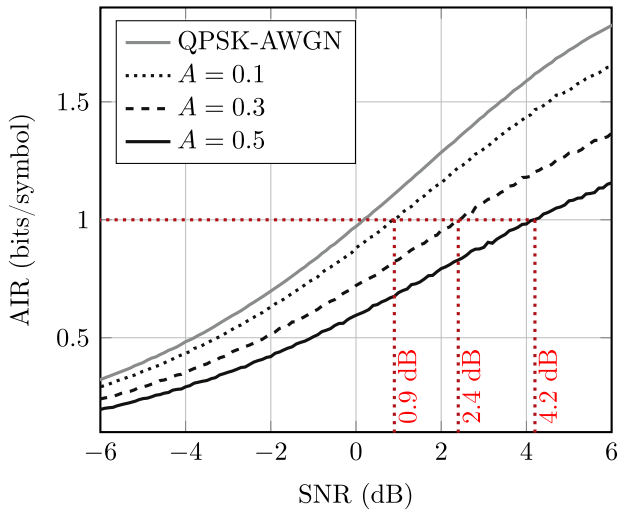


Fig. 10. AIRs versus SNR for QPSK transmission over Markov-Middleton IN channels for $A = 0.1$, 0.3 , and 0.5 with fixed $\Lambda = 10$, $r = 0.9$, and $W = 4$. The solid gray line indicates the AWGN channel.

function spreads, and the probability density around the mean decreases compared to that of the BN state. Consequently, the overall likelihood of transmitting a specific symbol x_t is almost entirely dominated by the BN state. It is also shown in Fig. 7 that the AIR deteriorates as the IN becomes more pronounced with increasing A . Besides, the inflection point shifts toward large Λ with a higher impulsive index A , which can be explained by (13). Specifically, when the IN samples are more active (larger A), the power of the IN must increase (larger Λ) to maintain the same noise power. On the other hand, the channel memory helps track the noise states through the transition matrix, reducing the uncertainties of the underlying IN channel and thus improving the AIR for any given value of A and Λ .

In Fig. 9, we evaluate the AIR performance under mismatched decoding scenarios where imperfect channel knowledge exists, providing a sensitivity analysis for parameter estimation inaccuracies. In all simulations, we assume that the Markov-Middleton channel is characterized by parameters $A = 0.3$, $r = 0.9$, and $W = 4$. This setup corresponds to a highly correlated, bursty IN environment, which is commonly observed in practice [4,6]. We first examine the extreme cases by assuming $\hat{W} = 1$ (black dashed line) and $\hat{\Lambda} = 10^{-2}$ (orange dashed line). This simplification effectively reduces the receiver to an AWGN detector. While they perform adequately in the BN-limited regime ($\Lambda < 10^{-1}$), their performance severely deteriorates as the true channel transitions into the IN-limited regime ($\Lambda > 10^{-1}$). Additionally, ignoring channel memory by assuming $r = 0$ (green dashed line) leads to an AIR loss of up to 0.1 bits/symbol compared to the memory-aware receiver under highly correlated impulsive conditions, a finding that aligns with previous observations in [19]. This highlights the need for memory-aware detection schemes or alternative techniques, such as orthogonal frequency division multiplexing (OFDM) approaches, to mitigate the impact of temporally correlated channels. Moreover, a mismatch in the impulsive parameters A and Λ imposes a notable performance loss, emphasizing the importance of accurate channel estimation in practical systems. On the other hand, the receiver is relatively robust to state mismatch: assuming a simple 2-state model (crosses) yields similar AIR performance to the true 4-state model. This suggests that the complexity of the receiver can be reduced without substantial performance loss as long as the key statistical properties of the IN are adequately captured.

As a reference for our practical receiver design, we present the AIR as a function of SNR in Fig. 10. Specifically, the SNR required to achieve a target AIR of 1 bit per QPSK symbol, which corresponds to a half-rate encoder, is 0.9 dB, 2.4 dB, and 4.2 dB for impulsive index $A = 0.1$, 0.3 , and 0.5 , respectively. These performance bounds will be used to benchmark our turbo-DE-PSK-IN receiver designs in the following subsection.

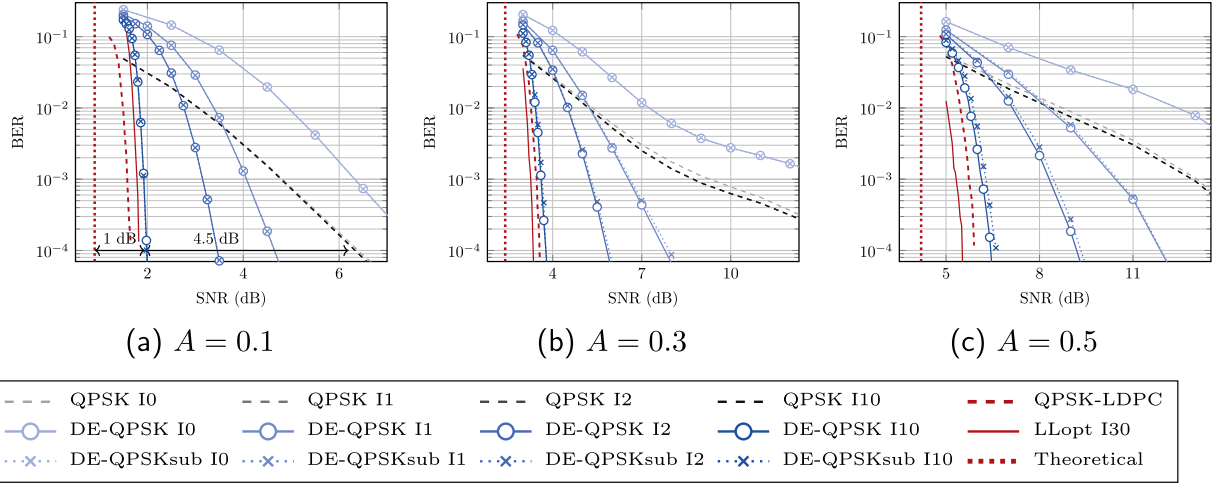


Fig. 11. BER performance over iterations for turbo-QPSK-IN (dashed lines), turbo-DE-QPSK-IN (circle solid lines), and suboptimal turbo-DE-QPSK-IN designs (cross dotted lines) over the Markov–Middleton channel with impulsive index (a) $A = 0.1$, (b) $A = 0.3$, and (c) $A = 0.5$, where $W = 4$, $\Lambda = 10$, and $r = 0.9$ are fixed for all simulations. The dotted red vertical lines indicate the simulated theoretical bound derived in Fig. 10. The dashed red lines show the QPSK-IN demapper with the LDPC code presented in [19,20]. The solid red lines indicate the ultimate performance bound of the proposed turbo-DE-PSK-IN receiver with perfect noise state information. All simulations have an interleaver depth of 64,800 bits. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

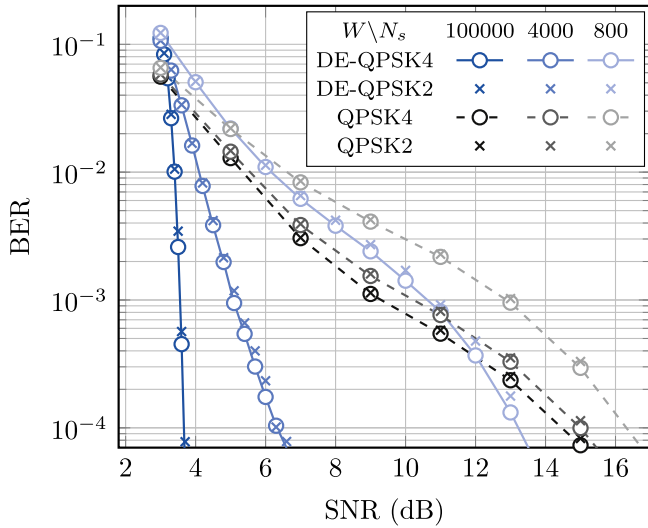


Fig. 12. BER performance for practical receiver design considerations with underestimated number of noise state $\hat{W} = 2$ and limited interleaver depths of 100,000, 4,000, and 800, respectively, for the proposed suboptimal turbo-DE-QPSK-IN receiver after 10 iterations. The IN channel is characterized by $W = 4$, $A = 0.3$, $\Lambda = 10$, and $r = 0.9$.

6.2. Proposed turbo-DE-PSK-IN receiver performance

This section reports the numerically simulated bit error rate (BER) performance of the proposed optimal and suboptimal turbo-DE-PSK-IN receivers under a 4-state Markov–Middleton IN channel. A standard half-rate convolutional code with generators $(5, 7)_8$ was used. A QPSK modulation is employed to map the interleaved coded bits to symbols, after which a differential modulator processes the symbols as in (27). The differentially modulated symbols are transmitted through a Markov–Middleton channel to generate the observed sequence y_1^T according to (28). For the conventional turbo-PSK-IN receiver, the observed symbols are generated via (16). For benchmarking, we include the performance of a state-of-the-art PSK-IN receiver employing the LDPC code, as demonstrated in [19,20]. To ensure a fair comparison with the DVB-S2

LDPC standard, we set the codeword length (interleaver depth) for all schemes to 64,800 bits.

Fig. 11a–c show the BER performance of the proposed joint turbo-DE-PSK-IN receiver (solid lines) with respect to the number of iterations for different impulsive indices $A = 0.1, 0.3$, and 0.5 . When comparing to the performance of the conventional turbo-PSK-IN receiver (dashed lines), we observe that the turbo-DE-PSK-IN receiver significantly outperforms the turbo-PSK-IN receiver in all cases after the first turbo iteration, with a performance gain of around 4.5 dB at a BER of 10^{-4} for $A = 0.1$. Notably, despite the use of a relatively simple convolutional code, the proposed receiver approaches the performance of the high-complexity DVB-S2 LDPC-based PSK-IN receiver (dashed red lines) after 10 iterations. This result highlights the efficacy of the rate-one differential structure in providing significant time diversity in impulsive environments. Furthermore, the proposed receiver maintains SNR gaps of only 1.1 to 2.1 dB relative to the theoretical bounds derived in Fig. 10. As shown in Fig. 3a and b, these gains are achieved without increasing the trellis state complexity of the joint IN detector.

The performance of the joint turbo-DE-PSK-IN receiver is also compared to its suboptimal separate design (dotted cross lines in Fig. 11a–c), where the turbo decoder and the IN detector (equalizer) are designed separately. The separate design is implemented by using the same MAP equalizer as in [16,20], which only computes the posterior symbol probability once without iterative updates. The results show that the suboptimal separate design can closely approximate the optimal joint design without much degradation, especially in a mild IN channel (Fig. 11a). A similar observation was also reported in [19]. This finding suggests that the use of a suboptimal design is preferable as it has only around half the complexity of the optimal design for 10 iterations. However, in a severely activated IN channel (Fig. 11c), the separate design shows a performance loss of around 0.2 dB compared to the joint design, which is expected to increase with a larger impulsive index A .

To gauge the ultimate performance of our proposed turbo-DE-PSK-IN scheme under the Markov–Middleton channel, we simulate the performance of the proposed turbo-DE-PSK-IN receiver with perfect noise state information (NSI) at the receiver. This is achieved by assuming that the noise state realization is perfectly known at the receiver, which allows for a perfect equalization of the received signal. The results are shown in Fig. 11a–c as solid red lines. The perfect NSI simulation serves as a benchmark for the performance of our proposed turbo system, where we

assume that a large number of iterations (30 iterations) are performed. The results show that the ultimate performance of the proposed turbo-DE-PSK-IN system is around 1 dB away from the theoretical bounds, regardless of channel conditions. For $A = 0.1$, a turbo-DE-PSK-IN receiver with 10 iterations is already close to the ultimate performance, while for $A = 0.5$, more iterations are required to achieve the best possible result.

In Fig. 12, we evaluate the performance of the separate receiver design in a setting with limited interleaver depth and mismatched channel states. The results are shown for a fixed 4-state Markov–Middleton model with $A = 0.3$, $r = 0.9$, and $\Lambda = 10$. The number of turbo iterations is set to 10. We first design a receiver that only assumes $\hat{W} = 2$, which leads to a total of $2M = 8$ trellis states in the IN detector. In this case, the receiver is designed to ignore the higher impulsive noise states. The results show that the performance of the receiver with a mismatched $\hat{W} = 2$ state assumption is very close to that of the matched receiver design with $W = 4$ states, which is also identified in Fig. 9. This is an important observation as we can design an efficient receiver with a reduced number of states without much performance loss. The results of Fig. 12 also show that the receiver is sensitive to the interleaver depth. A smaller interleaver depth of 4,000 leads to a performance loss of around 2.7 dB, while a depth of 800 results in significant performance loss of around 10 dB at a BER of 10^{-4} when compared to a system with an interleaver depth of 100,000 for the turbo-DE-PSK-IN system. The benefits of using a large interleaver depth are mainly twofold. First, it improves the MAP-based IN detector (equalizer) performance, and second, it effectively spreads the decoding errors for the two constituent decoders, which is crucial for the turbo decoding performance. We can see that the conventional turbo-PSK-IN (black dashed lines) can obtain only the detector-related benefits, while the proposed turbo-DE-PSK-IN system (blue solid lines) can benefit from both. This is because the turbo-DE-PSK-IN system can effectively spread the decoding errors thanks to the recursive nature of the differential decoder, which is crucial for turbo decoding performance.

7. Conclusions

In this paper, we propose robust turbo receiver structures that can effectively mitigate the impact of bursty impulses, modeled by a 4-state Markov–Middleton model, by incorporating a rate-one differential decoder into the receiver design. We started by investigating the AIR between the PSK-modulated symbols and the channel output sequence for various channel conditions. We found that the impulsive index significantly influences the AIR. Also, we observed a non-monotonic relationship between the estimated AIR and the impulsive-to-background noise power ratio, which can be further explained by evaluating the likelihood function for the BN and IN states. Furthermore, channel memory reduces the uncertainty of the given IN channel, thereby increasing the AIR. Mismatched decoding scenarios are also considered in the AIR analysis. In the second part of the paper, we carefully derived the proposed turbo-DE-PSK-IN receiver design, which incorporates a differential decoder into the MAP-based IN detector. Results show that the proposed turbo-DE-PSK-IN receiver significantly outperforms the conventional MAP-based turbo-PSK-IN receiver by more than 4.5 dB with an ultimate performance gap of around 1 dB from the AIR-derived performance bound. Additionally, we proposed a suboptimal separate receiver design, which only has half the complexity of the joint design but still sufficiently approximates the optimal joint receiver design. The proposed receiver design is also shown to be robust against the underestimated number of IN states.

This paper demonstrates that a simple rate-one differential encoder can effectively improve receiver performance under a serially concatenated turbo structure in the presence of bursty impulsive noise. While the present study focuses on PSK modulation and convolutional codes to clearly isolate the benefits of the proposed turbo-DE-PSK-IN architecture, the framework is inherently extensible to other modulation formats and alternative coding schemes. For instance, employing more powerful

codes such as LDPC could further enhance the quality of the extrinsic information exchanged during turbo iterations. This would potentially accelerate the convergence rate, allowing the receiver to approach theoretical information limits with fewer iterations. This remains a compelling direction for future research.

CRedit authorship contribution statement

Chin-Hung Chen: Writing – review & editing, Writing – original draft, Methodology, Investigation, Formal analysis, Conceptualization; **Boris Karanov:** Writing – review & editing, Supervision, Methodology, Investigation, Formal analysis, Conceptualization; **Wim van Houtum:** Writing – review & editing, Supervision, Project administration, Methodology, Investigation, Conceptualization; **Yan Wu:** Writing – review & editing, Investigation, Conceptualization; **Alex Alvarado:** Writing – review & editing, Supervision, Funding acquisition.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data availability

No data was used for the research described in the article.

References

- [1] A. Maouloud, M. Klingler, P. Besnier, A test setup to assess the impact of EMI produced by on-board electronics on the quality of radio reception in vehicles, *IEEE Trans. Electromagn. Compat* 63 (6) (2021) 1844–1855. <https://doi.org/10.1109/TEMC.2021.3072558>
- [2] X. Gao, D. Su, Y. Li, Study on electromagnetic interference of DC/DC converter used in the EV, in: *Asia-Pacific Symposium on Electromagnetic Compatibility (AP EMC)*, 2015, pp. 258–261. <https://doi.org/10.1109/AP EMC.2015.7175389>
- [3] K. Plakostathis, Research on EMI from modern electric vehicles and their recharging systems, in: *International Symposium on Electromagnetic Compatibility (EMC)*, 2020, pp. 1–6. <https://doi.org/10.1109/EMCEUROPE48519.2020.9245742>
- [4] C.-H. Chen, W.-H. Huang, B. Karanov, Y. Wu, A. Young, W. van Houtum, Analysis of impulsive interference in digital audio broadcasting systems in electric vehicles, in: *Symposium on Information Theory and Signal Processing in the Benelux (SITB)*, 2024, pp. 7–11.
- [5] Q. Shan, S. Bhatti, I.A. Glover, R. Atkinson, P.J. Moore, I.E. Portugues, R. Rutherford, Noise amplitude distribution of impulsive noise from measurements in a power substation, in: *International Universities Power Engineering Conference (UPEC)*, 2009, pp. 1–5.
- [6] F. Sacuto, F. Labeau, B.L. Agba, Wide band time-correlated model for wireless communications under impulsive noise within power substation, *IEEE Trans. Wireless Commun.* 13 (3) (2014) 1449–1461. <https://doi.org/10.1109/TWC.2014.011714.130685>
- [7] R. Pighi, M. Franceschini, G. Ferrari, R. Raheli, Fundamental performance limits for PLC systems impaired by impulse noise, in: *IEEE International Symposium on Power Line Communications and Its Applications (ISPLC)*, 2006, pp. 277–282. <https://doi.org/10.1109/ISPLC.2006.247475>
- [8] M. Zimmermann, K. Dostert, Analysis and modeling of impulsive noise in broadband powerline communications, *IEEE Trans. Electromagn. Compat.* 44 (1) (2002) 249–258. <https://doi.org/10.1109/15.990732>
- [9] W. van Houtum, Time-division spatial interference rejection (TDSIR)-procedure, U.S. Patent 11 (2022).
- [10] I. Landa, M.M. Vélez, A. Arrinda, R. Torre, M. Fernández, Impulsive noise characterization and its effect on digital audio quality, in: *IEEE International Symposium on Broadband Multimedia Systems and Broadcasting (BMSB)*, 2015, pp. 1–3. <https://doi.org/10.1109/BMSB.2015.7177238>
- [11] Q. Shan, Estimation of impulsive noise in an electricity substation, *IEEE Trans. Electromagn. Compat.* 53 (3) (2011) 653–663. <https://doi.org/10.1109/TEMC.2010.2092782>
- [12] S.A. Bhatti, I.A. Glover, R. Atkinson, Q. Shan, Y. Yang, J.M. R.S. Neto, Vulnerability of bluetooth to impulsive noise in electricity transmission substations, in: *IET International Conference on Wireless Sensor Network (WSN)*, 2010, pp. 53–58. <https://doi.org/10.1049/cp.2010.1027>
- [13] N.B. Sarr, A.K. Yazbek, H. Boeglen, J.P. Cances, R. Vauzelle, F. Gagnon, An impulsive noise resistant physical layer for smart grid communications, in: *IEEE International Conference on Communications (ICC)*, 2017, pp. 1–7. <https://doi.org/10.1109/ICC.2017.7996876>
- [14] G. Ndo, P. Siohan, M.H. Hamon, Adaptive noise mitigation in impulsive environment: application to power-line communications, *IEEE Trans. Power Deliv* 25 (2) (2010) 647–656. <https://doi.org/10.1109/TPWRD.2009.2035505>

- [15] H. Oh, H. Nam, Design and performance analysis of nonlinearity preprocessors in an impulsive noise environment, *IEEE Trans. Veh. Technol.* 66 (1) (2017) 364–376. <https://doi.org/10.1109/TVT.2016.2547889>
- [16] M.S. Alam, B. Selim, G. Kaddoum, B.L. Agba, Mitigation techniques for impulsive noise with memory modeled by a two state Markov-Gaussian process, *IEEE Syst. J.* 14 (3) (2020) 4079–4088. <https://doi.org/10.1109/JSYST.2020.2964980>
- [17] D. Middleton, Statistical-physical model of electromagnetic interference, *IEEE Trans. Electromagn. Compat.* 19 (3) (1977) 106–127. <https://doi.org/10.1109/TEMC.1977.303527>
- [18] D. Middleton, Non-Gaussian noise models in signal processing for telecommunications: new methods and results for class a and class b noise models, *IEEE Trans. Inf. Theory* 45 (4) (1999) 1129–1149. <https://doi.org/10.1109/18.761256>
- [19] D. Fertonani, G. Colavolpe, On reliable communications over channels impaired by bursty impulse noise, *IEEE Trans. Commun.* 57 (7) (2009) 2024–2030. <https://doi.org/10.1109/TCOMM.2009.07.070638>
- [20] G. Ndo, F. Labeau, M. Kassouf, A Markov-Middleton model for bursty impulsive noise: modeling and receiver design, *IEEE Trans. Power Del.* 28 (4) (2013) 2317–2325. <https://doi.org/10.1109/TPWRD.2013.2273942>
- [21] D. Fertonani, G. Colavolpe, Theoretical limits and practical detection schemes for channels affected by class-A impulse noise, in: *IEEE Global Telecommunications Conference (GLOBECOM)*, 2007, pp. 146–150. <https://doi.org/10.1109/GLOCOM.2007.35>
- [22] D. Fertonani, G. Colavolpe, A robust metric for soft-output detection in the presence of class-A noise, *IEEE Trans. Commun.* 57 (1) (2009) 36–40. <https://doi.org/10.1109/TCOMM.2009.0901.070041>
- [23] R. Haring, A.J. Han, Vinck, Performance bounds for optimum and suboptimum reception under class-A impulsive noise, *IEEE Trans. Commun.* 50 (7) (2002) 1130–1136. <https://doi.org/10.1109/TCOMM.2002.800806>
- [24] H. Nakagawa, D. Umehara, S. Denno, Y. Morihiro, A decoding for low density parity check codes over impulsive noise channels, in: *International Symposium on Power Line Communications and Its Applications (ISPLC)*, 2005, pp. 85–89. <https://doi.org/10.1109/ISPLC.2005.1430471>
- [25] J. Mitra, L. Lampe, Convolutionally coded transmission over Markov-Gaussian channels: analysis and decoding metrics, *IEEE Trans. Commun.* 58 (7) (2010) 1939–1949. <https://doi.org/10.1109/TCOMM.2010.07.090273>
- [26] F.G. Mengistu, D.F. Tseng, Y.S. Han, M.A. Mulatu, L.-C. Chang, A robust decoding scheme for convolutionally coded transmission through a Markov Gaussian channel, *IEEE Trans. Veh. Technol.* 63 (9) (2014) 4344–4356. <https://doi.org/10.1109/TVT.2014.2312727>
- [27] D.-F. Tseng, F.G. Mengistu, Y.S. Han, M. Mulatu, L.-C. Chang, T.R. Tsai, Robust turbo decoding in a Markov Gaussian channel, *IEEE Wireless Commun. Lett.* 3 (6) (2014) 633–636. <https://doi.org/10.1109/LWC.2014.2359464>
- [28] D. Umehara, H. Yamaguchi, Y. Morihiro, Turbo decoding in impulsive noise environment, in: *IEEE Global Telecommunications Conference (GLOBECOM)*, 2004, pp. 194–198. <https://doi.org/10.1109/GLOCOM.2004.1377938>
- [29] G. Colavolpe, Classical coherent receivers for differentially encoded M-PSK are optimal, *IEEE Commun. Lett.* 8 (4) (2004) 211–213. <https://doi.org/10.1109/LCOMM.2004.827388>
- [30] K.R. Narayanan, G.L. Stuber, A serial concatenation approach to iterative demodulation and decoding, *IEEE Trans. Commun.* 47 (7) (1999) 956–961. <https://doi.org/10.1109/26.774833>
- [31] P. Hoeher, J. Lodge, Turbo DPSK: iterative differential PSK demodulation and channel decoding, *IEEE Trans. Commun.* 47 (6) (1999) 837–843. <https://doi.org/10.1109/26.771340>
- [32] W. van Houtum, Two-dimensional block-based reception for differentially encoded OFDM systems: A study on improved reception techniques for digital audio broadcasting systems, Technical Report, Technische Universiteit Eindhoven, Eindhoven, Netherlands, 2012.
- [33] I. Lee, The effect of a precoder on serially concatenated coding systems with an ISI channel, *IEEE Trans. Commun.* 49 (7) (2001) 1168–1175. <https://doi.org/10.1109/ICC.2000.853568>
- [34] C.-H. Chen, I. Nikoloska, W. van Houtum, Y. Wu, B. Karanov, A. Alvarado, Robust blind channel estimation for bursty impulsive noise with a constrained EM approach, in: *IEEE Vehicular Technology Conference (VTC-Spring)*, 2025, pp. 1–6. <https://doi.org/10.1109/VTC2025-Spring65109.2025.11174888>
- [35] S.M. Zabin, H.V. Poor, Efficient estimation of class a noise parameters via the EM algorithm, *IEEE Trans. Inf. Theory* 37 (1) (1991) 60–72. <https://doi.org/10.1109/18.61127>
- [36] C.-Y. Chen, M.-C. Chiu, Parameter estimation of impulsive noise for channel coded communication systems, *IET Commun.* 15 (2019) 445–452. <https://doi.org/10.1049/cmu2.12077>
- [37] A. Vannucci, G. Colavolpe, R. Pecori, L. Veltri, Estimation of a Gaussian source with memory in bursty impulsive noise, in: *IEEE International Symposium on Power Line Communications and Its Applications (ISPLC)*, 2019, pp. 1–6. <https://doi.org/10.1109/ISPLC.2019.8693453>
- [38] L.R. Bahl, J. Cocke, F. Jelinek, J. Raviv, Optimal decoding of linear codes for minimizing symbol error rate, *IEEE Trans. Inform. Theory* 20 (2) (1974) 284–287. <https://doi.org/10.1109/TIT.1974.1055186>
- [39] D.M. Arnold, H.-A. Loeliger, P.O. Vontobel, A. Kavcic, W. Zeng, Simulation-based computation of information rates for channels with memory, *IEEE Trans. Inform. Theory* 52 (8) (2006) 3498–3508. <https://doi.org/10.1109/TIT.2006.878110>