Unsupervised Log-Likelihood Ratio Estimation for Short Packets in Impulsive Noise

Yasser Mestrah, Dadja Anade, Anne Savard, Alban Goupil, Malcolm Egan, Philippe Mary, Jean-Marie S Gorce, Laurent Clavier

Unsupervised Log-Likelihood Ratio Estimation for Short Packets in Impulsive Noise

Yasser Mestrah*, Dadja Anade†, Anne Savard‡§, Alban Goupil¶, Malcolm Egan†,
Philippe Mary*, Jean-Marie Gorce† and Laurent Clavier‡§
* Univ. Rennes, INSA Rennes, IETR, CNRS Rennes, France
† Univ. Lyon, INSA Lyon, CITI, INRIA, Lyon, France
‡ IMT Lille Douai, Institut Mines Télécom, Centre Digital Systems, F-59653 Villeneuve d'Ascq, France
§ Univ. Lille, CNRS, Centrale Lille, Univ. Polytechnique Hauts-de-France, UMR 8520 - IEMN, F-59000 Lille, France
¶ CReSTIC, Univ. Reims Champagne Ardenne, Reims, France
Email: †{dadja-toussaint.anade-apko, malcom.egan, jean-marie.gorce}@insa-lyon.fr,
‡{anne.savard, laurent.clavier}@imt-lille-douai.fr, ¶alban.goupil@univ-reims.fr, *{yasser.mestrah, philippe.mary}@insa-rennes.fr

***Abstract*—Impulsive noise, where large amplitudes arise with a relatively high probability, arises in many communication systems including interference in Low Power Wide Area Networks. A challenge in coping with impulsive noise, particularly alpha-stable models, is that tractable expressions for the log-likelihood ratio (LLR) are not available, which has a large impact on soft-input decoding schemes, e.g., low-density parity-check (LDPC) packets. On the other hand, constraints on packet length also mean that pilot signals are not available resulting in non-trivial approximation and parameter estimation problems for the LLR. In this paper, a new unsupervised parameter estimation algorithm is proposed for LLR approximation. In terms of the frame error rate (FER), this algorithm is shown to significantly outperform existing unsupervised estimation methods for short LDPC packets (on the order of 500 symbols), with nearly the same performance as when the parameters are perfectly known. The performance is also compared with an upper bound on the information-theoretic limit for the FER, which suggests that in impulsive noise further improvements require the use of an alternative code structure other than LDPC.**

***Index Terms*—Impulsive noise, BP-decoders, LDPC, short packets, LLR approximation, unsupervised estimation**

I. Introduction

The Internet of Things (IoT) has driven the emergence of Low Power Wide Area Networks (LPWANs), in which devices (e.g., sensors) with limited transmit power communicate with an access point. Key examples are LoRa, SigFox and NB-IoT. In addition to limited transmit power, these devices have restricted computational capabilities while the access point is capable of higher complexity computation and, at least in the 868 MHz ISM bands, constraints on duty cycles, as well as small quantities of data to transmit.

As a consequence: (a) codes cannot be spread over multiple packets and must be short; (b) simple modulation schemes are utilized (differential binary phase shift keying (DBPSK) for SigFox or quadrature phase shift keying (QPSK) for NB-IoT); (c) complex decoding algorithms can be implemented in the access point; and (d) training sequences to estimate channel parameters are undesirable due to short length of the code.

Despite the limitations imposed by LPWAN requirements, channel coding is mandatory to ensure robust radio links. Moreover, many soft-input decoding schemes (e.g., belief propagation or turbo decoders) rely on the log-likelihood ratio (LLR), which in turn requires an accurate characterization of the noise probability density function. In contrast to thermal noise-limited communication systems, where the noise is well-modeled as Gaussian, there is increasing evidence that non-Gaussian models are required for noise induced by interference. Indeed, theoretical analysis based on stochastic geometry models for device locations suggests that the noise distribution is impulsive or heavy-tailed in the form of α-stable models [1], [2]. Recently, this analysis has been shown to be in agreement with real experiments in the 868 MHz band [3].

A difficulty with α-stable and many other heavy-tailed models is that the LLR can be difficult to implement, requiring infinite series, lookup tables, or Fourier transforms. One approach to remedy this issue is to exploit closed-form parametric LLR approximations [4]–[6]. However, when pilot sequences are not available, it remains non-trivial to estimate the parameters of the LLR approximation due to the fact that channel inputs are not known to the access point.

An algorithm to estimate the parameters of LLR approximations in an unsupervised (USD) fashion has recently been proposed in [6]. The algorithm was shown to perform well when the code length was long; however, as the code length is reduced, it suffers from a severe performance degradation relative to an estimation scheme exploiting knowledge of the channel inputs. As such, the algorithm in [6] cannot be utilized for realistic IoT systems that exploit short packets.

In this paper, we propose a new unsupervised (NUSD) estimation algorithm for approximate LLR parameters, which overcomes the performance degradation when the code length is small (on the order of 500 symbols per packet).

Our main contributions are the following:

- The FER degrades severely when the packet size decreases. A study of the origin of the degradation is tackled, which reveals that the absence of certain low probability samples have an important impact on the performance.

- We propose and derive an analytical tool to quantify and assess the risk of failure of LLR estimation.
- We propose an *ad hoc* modification of the estimation process that makes the unsupervised receiver robust enough to behave very close to the optimal one. We therefore introduce a combination of sampling and regularization techniques to ensure that rare samples are properly accounted for.
- We also compare the FER achieved by a LDPC code with the achievable rate [7] for code length $n = 408$ and $S\alpha S$ noise. This reveals a significant gap, which suggests that further improvements can be made by modifying the LDPC code or adopting a different structure, such as turbo or polar codes.

To validate the proposed algorithm, we study the performance in the presence of a memoryless symmetric α-stable ($S\alpha S$) noise channel with the LDPC code[1] of length $n = 408$. The results can be extended to other types of codes and noises, encompassing both impulsive and Gaussian models [8]. In particular, we show via Monte Carlo simulation that our unsupervised estimation algorithm yields a FER with near optimal performance. As such, limited further performance can be gained by adapting the LDPC decoder.

II. System model and problem formulation

A. System Setup

Consider a transmission between a single device and an access point. The channel output for the i-th symbol is given by $Y_i = X_i + Z_i,\ i = 1, \ldots, n$, where the channel input is the binary sequence $X^n \in \{-1, 1\}^n$ encoded via a LDPC code and $Z^n \in \mathbb{R}^n$ is the noise sequence, independent of X^n. In particular, we assume that Z^n is an independent sequence of $S\alpha S$ random variables; that is, each element $Z_i,\ i = 1, \ldots, n$ admits a characteristic function

$$\Phi_{Z_i}(t) = \exp\left(-\gamma^\alpha |t|^\alpha\right), \tag{1}$$

where $0 < \alpha \leq 2$ is the exponent and $\gamma \geq 0$ is the scale parameter. In the case $\alpha = 2$, Z_i is Gaussian; however, for $0 < \alpha < 2$, Z_i has infinite variance and for $0 < \alpha \leq 1$, infinite mean. Except for when $\alpha = 1$ and $\alpha = 2$, $S\alpha S$ probability density functions, which we denote by $g_\alpha(\cdot; \gamma)$, do not have a closed-form representation.

At the access point, the received sequence Y^n is decoded using the belief propagation (BP) algorithm, which requires the LLRs

$$L(Y_i) = \log \frac{p_{Y|X}(Y_i|X_i = +1)}{p_{Y|X}(Y_i|X_i = -1)}, \tag{2}$$

where $p_{Y_i|X_i}(y_i|x_i) = p_{Z_i}(y_i - x_i),\ y_i \in \mathbb{R}, x_i \in \{-1, 1\}$. As the probability density function of Z_i is not available in closed form, efficient computation of (2) is not straightforward.

[1]The proposed algorithm can be extended to other types of soft-input decoders. In this paper, the selection of the best codes in the short packet length regime , is out of scope. However, we expect that polar codes with cyclic redundancy check concatenation or convolutional polar codes may outperform LDPC codes.

Instead, a parametric LLR approximation $L_\theta(\cdot)$ can be used, where θ corresponds to the vector of parameters. Several LLR approximations have been proposed [5], [6], [9]–[11]. All of them exhibit good performance in terms of bit error rate (BER) and FER with a drastic reduction in the computational complexity compared to a numerical evaluation of the true LLR. We focus on the LLR approximation in [9], defined by

$$L_\theta(y) = \mathrm{sign}(y) \min\left(a|y|, \frac{b}{|y|}\right),\ y \in \mathbb{R}, \tag{3}$$

with $\theta = (a, b) \in \mathbb{R}_+^2$. It is worth mentioning that, the demapper can be parameterized directly by a and b instead of α and γ. where, $a = \sqrt{2}/\gamma$ and $b = 2(\alpha + 1)$,

B. LLR Parameter Estimation

While a number of methods have been introduced for LLR parameter estimation, a popular one is based on minimization of the conditional entropy [12], [13]

$$H(Y|X) = \mathbb{E}_{X,Y}\left[\log_2\left(1 + e^{-XL_\theta(Y)}\right)\right]. \tag{4}$$

The parameters of the LLR approximation can then in principle be obtained by solving

$$\theta^* = \arg\min_{\theta \in \mathbb{R}_+^2} \mathbb{E}_{X,Y}\left[\log_2\left(1 + e^{-L_\theta(XY)}\right)\right]. \tag{5}$$

As a stochastic optimization problem, assuming a pilot sequence is available to yield $X_i,\ i = 1, \ldots, n$, θ^* may be estimated via sample average approximation, given by

$$\hat{\theta}^* = \arg\min_{\theta \in \mathbb{R}_+^2} \frac{1}{n} f(\theta; \{\Psi_i\}_{i=1}^n) \tag{6}$$

where $\Psi_i = X_i Y_i$, n is the packet length, and

$$f(\theta; \{\Psi_i\}_{i=1}^n) = \sum_{i=1}^n \log_2\left(1 + e^{-L_\theta(\Psi_i)}\right). \tag{7}$$

However, for short packets, pilot sequences are not available. As a consequence, an unsupervised approach is required. In the remainder of this paper, we develop a new unsupervised algorithm for the parameters θ, which has good performance even for short packets (corresponding to approximately 500 symbols).

III. Unsupervised LLR Parameter Estimation for Short Packets

In this section, we introduce our new scheme. Before presenting the proposed unsupervised estimation algorithm, we first recall the approach in [6], which will form a baseline.

A. Baseline Algorithm

The key difficulty in estimating θ directly via (6) is the lack of knowledge of $X_i,\ i = 1, \ldots, n$. In the USD algorithm proposed in [6], this problem was overcome by estimating the noise samples via $\widetilde{Z}_i = Y_i - \mathrm{sign}(Y_i)$ for each channel output $Y_i,\ i = 1, \ldots, n$. A new sequence of samples $\widetilde{Y}_i,\ i = 1, \ldots, n$ is then constructed via $\widetilde{Y}_i = 1 + \widetilde{Z}_i$ by the symmetry of the

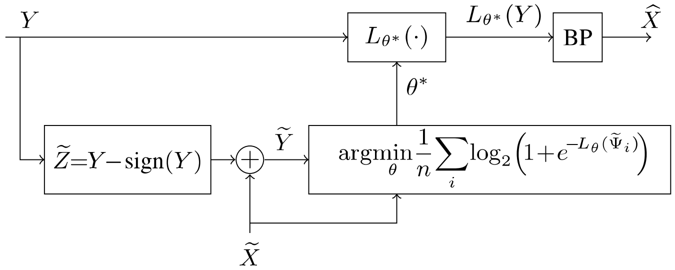

Figure 1. Unsupervised decoder (USD) of [6]

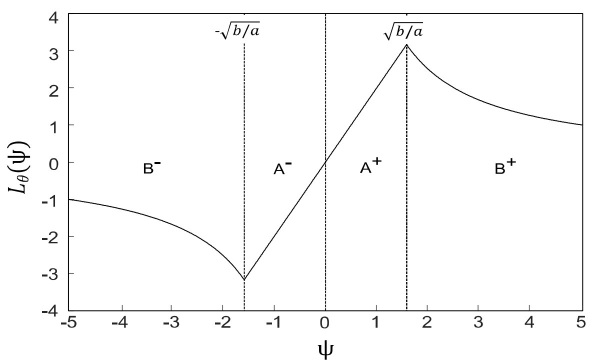

Figure 2. Illustration of the LLR approximation with the four regions influencing the shape of $L_\theta(\Psi)$.

noise distribution $p_{Z_i}(\cdot)$. In the following, a genie-aided decoder (GAD) denotes that the knowledge of $X_i,\ i=1,\dots,n$ is available at the receiver.

To obtain the parameter estimate θ^* in (6) the samples $(1,\widetilde{Y}_i),\ i=1,\dots,n$, corresponding to $\tilde{\Psi}_i=\widetilde{X}_i\widetilde{Y}_i$, are used instead of (X_i,Y_i). Which leads to better understanding without affecting performance. The complete algorithm is summarized in Fig. 1.

B. Quantifying the Risk of Poor Estimation

The USD algorithm in [6] works well for large n but, as will be illustrated in Sec. IV, poorly when n is small (on the order of 500 symbols). In order to understand the reason and to improve the algorithm in [6], we first analyze when poor estimation occurs.

To begin, the domain of the LLR approximation in (3) can be partitioned in four regions:

$$B^-=\left[-\infty,-\sqrt{b/a}\right],A^-=\left[-\sqrt{b/a},0\right],$$
$$A^+=\left[0,\sqrt{b/a}\right],B^+=\left[\sqrt{b/a},+\infty\right].$$

Indeed, the LLR approximation has the form illustrated in Fig. 2.

The objective in (7) can therefore be written as

$$f(\theta;\{\tilde{\Psi}_i\}_{i=1}^n)\propto\sum_{i:\tilde{\Psi}_i\in B^-\cup B^+}\log_2(1+e^{-\frac{b}{\tilde{\Psi}_i}}) + \sum_{i:\tilde{\Psi}_i\in A^-\cup A^+}\log_2(1+e^{-a\tilde{\Psi}_i}). \quad (8)$$

Observe that when $\tilde{\Psi}_i$ is in A^+ or B^+, the exponent inside the logarithm is negative. On the other hand, when $\tilde{\Psi}_i$ is in A^- or B^-, the exponent is positive. As a consequence, samples of $\tilde{\Psi}_i$ in B^+ tend to increase the optimized values of b, while samples in B^- tend to decrease the optimized values of b. A similar observation also holds for the a parameter. As a consequence, if only a small number of samples falls in any of these regions, the resulting LLR approximation in (3) may be poor.

It is therefore important to understand for which regions the GAD algorithm produces very few samples. Recall that the noise sample estimates are obtained via $\tilde{Z}_i=Y_i-\mathrm{sign}(Y_i)$ with $\tilde{Y}_i=1+\tilde{Z}_i$ and $\tilde{X}_i=1$. As a consequence, $\tilde{\Psi}_i=\tilde{X}_i\tilde{Y}_i<0$ when $Y_i<-2$. Hence, samples $\tilde{\Psi}_i\in A^+\cup B^+$ will occur with very high probability.

However, samples $\tilde{\Psi}_i\in A^-\cup B^-$ will occur more rarely. While this is less problematic for long packets, it greatly impacts the estimation of L_θ for short packets. In particular:

1) No samples of $\tilde{\Psi}_i$ in B^- result in $b^*\to\infty$ and consequently the threshold $\sqrt{b/a}$ tends to infinity. This is not problematic for Gaussian noise where the optimal receiver is linear (i.e., $b^*=\infty$), but causes a problem if the noise is impulsive.
2) No samples of $\tilde{\Psi}_i$ in A^- result in $a^*\to\infty$, leading to $\sqrt{b/a}\to 0$. The LLR approximation around $\tilde{\Psi}_i\approx 0$ is therefore poor. As a consequence, an error floor can occur for small values of γ.

Due to the importance of obtaining samples $\tilde{\Psi}_i$ in $A^-\cup B^-$, it is desirable for the LLR parameter estimation algorithms to yield samples in these regions. To quantify the probability such samples do indeed arise, we define the following notion of degeneration risk.

Definition 1 *The degeneration risk for the GAD and USD algorithms is the probability that the sequence Ψ_i or $\tilde{\Psi}_i,\ i=1,\dots,n$ has only positive elements. That is,*

$$\eta_n^{GAD}=\mathbb{P}(\Psi_i>0,\ \forall i\in\{1,\dots,n\})$$
$$\eta_n^{USD}=\mathbb{P}(\tilde{\Psi}_i>0,\ \forall i\in\{1,\dots,n\}). \quad (9)$$

The following proposition characterizes the degeneration risk for the GAD and USD algorithms.

Proposition 1 *The degeneration risk of the GAD algorithm and USD algorithm in [6] are given by*

$$\eta_n^{GAD}=\left[1-Q_\alpha\left(\frac{1}{\gamma}\right)\right]^n$$
$$\eta_n^{USD}=\left[1-\frac{1}{2}Q_\alpha\left(\frac{3}{\gamma}\right)-\frac{1}{2}Q_\alpha\left(\frac{1}{\gamma}\right)\right]^n, \quad (10)$$

where

$$Q_\alpha(x)=\int_x^{+\infty}g_\alpha(u;1)\,\mathrm{d}u \quad (11)$$

with $g_\alpha(\cdot;1)$ denoting the probability density function of a standard $S\alpha S$ random variable.

Proof: Let us consider the two decoders separately. Under GAD, $X_i \in \{-1,1\}$ and $\Psi_i = X_iY_i$, so that $\Psi_i = X_iY_i$ is negative if ($X_i = +1$ and $Z_i < -1$) or if ($X_i = -1$ and $Z_i > 1$). Thus,

$$\eta_1^{GAD} = 1 - \left(\frac{1}{2}\mathbb{P}(Z_1 < -1 | X_1 = +1) + \frac{1}{2}\mathbb{P}(Z_1 > 1 | X_1 = -1)\right)$$
$$= 1 - Q_\alpha\left(\frac{1}{\gamma}\right). \tag{12}$$

Under USD, the training input sequence only consists of $\widetilde{X}_i = +1$, thus $\widetilde{Y}_i = 1 + \widetilde{Z}_i$. Hence, the probability of negative $\widetilde{\Psi}_i$ is $\mathbb{P}\left(\widetilde{\Psi}_i < 0\right) = \mathbb{P}\left(\widetilde{Z}_i < -1\right) = \mathbb{P}\left(Y_i - \text{sign}\left(Y_i\right) < -1\right)$. For the last inequality to be true, sign(Y_i) is necessarily -1. Thus,

$$\begin{aligned}\eta_1^{USD} &= 1 - \mathbb{P}\left(\widetilde{\Psi}_1 < 0\right)\\ &= 1 - \mathbb{P}(X_1 + Z_1 < -2)\\ &= 1 - \frac{1}{2}\mathbb{P}(Z_1 < -3 | X_1 = 1)\\ &\quad - \frac{1}{2}\mathbb{P}(Z_1 < -1 | X_1 = -1)\\ &= 1 - \frac{1}{2}\left[Q_\alpha\left(\frac{3}{\gamma}\right) + Q_\alpha\left(\frac{1}{\gamma}\right)\right]. \end{aligned}\tag{13}$$

■

Fig. 3 compares the degeneration risk for the GAD and USD algorithms as a function of the noise scale parameter γ for $n = 1$ and $\alpha = (1.4; 1.8)$. Observe that there is a significant increase in the degeneration risk for the USD algorithm, which suggests that the lack of samples $\widetilde{\Psi}_i \in A^- \cup B^-$ may be a key factor leading to performance losses when packets are short.

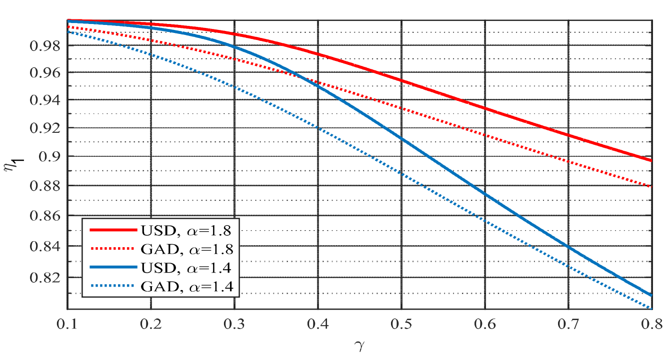

Figure 3. Comparison of the degeneration risk for $n = 1$ between GAD and USD [6] algorithms, as a function of γ, under SαS noise with $\alpha = 1.8$ and $\alpha = 1.4$.

C. New Unsupervised Estimation Method (NUSD)

Motivated by the observations in the previous section, we propose a new unsupervised estimation (NUSD) algorithm for the LLR approximation in (3). In particular, we introduce two techniques to reduce the degeneration risk: a regularization term to limit the growth of a; and an improved sampling method to obtain a better estimate of b.

1) Estimation of a: Our first goal is to improve estimation of the parameter a when no samples $\widetilde{\Psi}_i$ lie in A^-, which leads to a very large value of a^*. A simple means of avoiding this problem is to introduce the effect of a small negative sample in a modification of the objective in (7), given by

$$\tilde{f}(\theta; \{\tilde{\Psi}_i\}_{i=1}^n) = f(\theta; \{\tilde{\Psi}_i\}_{i=1}^n) + \log_2(1 + e^{a\epsilon}). \tag{14}$$

2) Estimation of b: The optimization parameter b impacts the way large amplitude samples are treated. Improving its convergence is trickier than improving the one of parameter a. Indeed, under Gaussian noise, the optimal value is $b^* = \infty$ leading to a fully linear receiver, but under SαS noise with $\alpha < 2$, a too large value of b can lead to a poorly estimated LLR. Hence, adding a regularization term in the objective function is not efficient as it would cause severe performance loss under Gaussian noise. Instead, we propose to increase the spread in the extracted noise sequence, which in turn increases the spread in $\widetilde{Y}$ and, as a consequence, reduces the degeneration risk. The proposed solution is described in Algorithm 1.

First, we double the learning sequence length by concatenating the extracted noise $\widetilde{Z}^n$ with its opposite version ($-\widetilde{Z}^n$), yielding a symmetric sequence $\widetilde{\mathbb{Z}}^c$ that contains an increased number of negative samples compared to $\widetilde{Z}^n$, note that $c = 2n$. Then, we generate the training sequence $\widetilde{\Psi}^c = \widetilde{Y}^c$. Since poor LLR approximation arises when the region B^- is empty, we ensure to populate this region by creating a negative sample when $\widetilde{\mathbb{Z}}_i > 1$. Indeed, in order for a sample to belong to B^-, the noise sample must be of opposite sign than $\widetilde{X}_i = 1$ and such that $|\widetilde{\mathbb{Z}}_i| \geq \sqrt{b/a} + 1$.

Algorithm 1 Generating the training sequence $\widetilde{\Psi}$

Input: Channel output Y^n.
Output: Training sequence $\widetilde{\Psi}^c$.

```
Compute Z̃^n = Y^n − sign(Y^n).
Concatenate Z̃^n and −Z̃^n into ℤ̃^c.
for i=1 to length(ℤ̃^c) do
    if ℤ̃_i ≥ 1 then
        Ψ̃_i = 1 − ℤ̃_i.
    else
        Ψ̃_i = 1 + ℤ̃_i.
    end if
end for
```

To summarize the above modifications of the USD algorithm, the NUSD algorithm is detailed in Algorithm 2. A simplex method based on Nelder Mead-algorithm [14] is used to solve the optimization problem defined in Algorithm 2 line 3, for more details, please refer to [6].

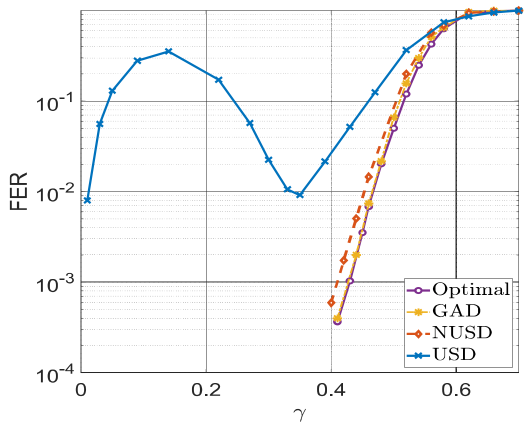

Figure 4. FER comparison between GAD, unsupervised decoder (USD) of [6], our new proposed USD (NUSD) and the optimal decoder, as a function of the scale parameter γ of a SαS noise with $\alpha = 1.8$ for the regular (3,6) LDPC code of length 408. For each FER value, we draw 500000 realizations.

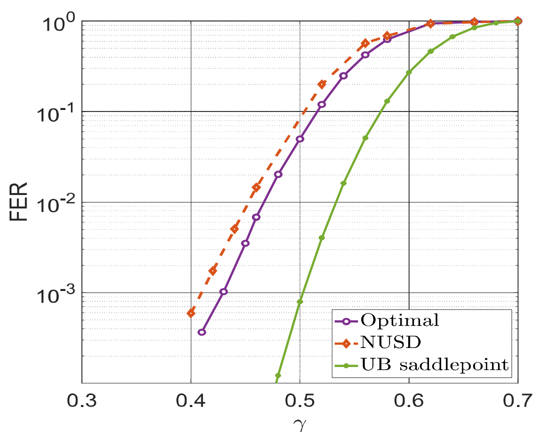

Figure 5. FER comparison between NUSD, the optimal decoder and the upper bound of the dependence testing (DT) bound (UB saddlepoint), as a function of the scale parameter γ of a SαS noise with $\alpha = 1.8$.

Algorithm 2 New Unsupervised Estimation Method

Input: Channel output Y^n.
Output: Estimated symbols $\widehat{X}^n$.
1: **for** each received packet Y^n **do**
2: Obtain $\widetilde{\Psi}^c$ from Algorithm 1 with input Y^n.
3: $\hat{\theta}^* = \arg\min_{\theta\in\mathbb{R}^2_+} \sum_{i=1}^{c} \log_2\left(1+e^{-L_\theta(\Psi_i)}\right)$.
4: Compute LLR for the sequence Y^n: $L_{\widehat{\theta}^*}(Y^n)$.
5: Compute $\hat{X}^n$ via the BP decoder with input $L_{\widehat{\theta}^*}(Y^n)$.
6: **end for**

IV. Numerical Results

A. Performance of the NUSD Algorithm

Fig. 4 plots the FER for the LLR approximation parameter estimation algorithms for varying $S\alpha S$ noise with scale parameter γ and $\alpha = 1.8$. The FER performance is tested using a regular $(3, 6)$ LDPC of packet length $n = 408$ under different decoding strategies. The curve labeled 'Optimal' is achieved with the true numerically computed LLRs, whereas the one labeled 'USD' is achieved with the previously proposed unsupervised LLR approximation of [6]. Since GAD almost achieves the FER obtained with the true LLRs, the performance loss of the USD is solely due to the unsupervised optimization and not to the LLR approximation itself.

The FER curve obtained with our proposed training sequence design and the adding of the regularization term with $\epsilon = 0.1$ is also provided in Figure 4 "NUSD". First, observe that the FER achieved by the NUSD is monotonic and does not exhibit any bump as the one obtained under the USD of [6], which appears to be due to the lack of samples $\widetilde{\Psi}_i$ lying in A^-. Furthermore, the performance of NUSD is close to that of GAD. Hence, our proposed method improves the performance of LLR parameter estimation for short packets without requiring the noise model knowledge at the receiver.

B. Limits of Short LDPC Packets

The 'UB saddlepoint' curve in Fig. 5 is an achievable FER. This bound follows from an upper-bound based on the dependence testing (DT) bound [15, Theorem 17], which is computed using the recently proposed method of [7, Theorem 5]. Since the FER achieved by our NUSD (and indeed the 'Optimal' decoder with perfect knowledge of LLR parameters) and the one achieved by the saddlepoint technique do not match, we can conclude that there is still room for efficient LDPC code design under impulsive noise and short blocklength.

V. Conclusion

In this paper, we focused on an unsupervised estimation of the LLR in the short block length regime. This estimation is performed by designing a simulated transmission at the receiver side, where the noise is extracted from the received data using a sign detector. We showed that if the simulated transmission is designed by simply adding the estimated extracted noise to the all codewords, the FER is non-monotonic and exhibits a severe performance loss compared to a genie-aided approach. We first analytically derived a metric to measure the probability of these performance loss to occur. Based on this analysis, we proposed a new design for the simulated transmission as well as the addition of a regularization term in the optimization problem so that the unsupervised LLR approximation achieves almost the same performance as the genie-aided decoder even under short codewords for a wide range of noise types, from non impulsive to highly impulsive, while exhibiting a low implementation complexity. Furthermore, the results of our simulation study show that the construction of new codes for impulsive environments remains an open problem due to a large gap in the performance with an achievable error rate.

Acknowledgment

This work has been partially supported by COST ACTION CA 15104 - IRACON, IRCICA USR CNRS 3380 and the French National Agency for Research (ANR) under grant ANR-16-CE25-0001- ARBURST.

REFERENCES

[1] M. Z. Win, P. C. Pinto, and L. A. Shepp, "A mathematical theory of network interference and its applications," *Proceedings of the IEEE*, vol. 97, no. 2, pp. 205–230, Feb 2009.

[2] M. Egan, L. Clavier, C. Zheng, M. de Freitas, and J.-M. Gorce, "Dynamic interference for uplink SCMA in large-scale wireless networks without coordination," *EURASIP Journal on Wireless Communications and Networking*, vol. 1, p. 213, 2018.

[3] L. Clavier, T. Pedersen, I. Rodriguez, M. Lauridsen, and M. Egan, "Experimental Evidence for Heavy Tailed Interference in the IoT," 2020, working paper or preprint.

[4] Y. Mestrah, A. Savard, A. Goupil, L. Clavier, and G. Gellé, "Blind estimation of an approximated likelihood ratio in impulsive environment," *IEEE PIMRC*, 2018.

[5] Y. Mestrah, A. Savard, A. Goupil, G. Gellé, and L. Clavier, "Robust and simple log-likelihood approximation for receiver design," *IEEE WCNC*, 2019.

[6] Y. Mestrah, A. Savard, A. Goupil, G. Gellé, and L. Clavier, "An unsupervised LLR estimation with unknown noise distribution," *EURASIP JWCN*, vol. 26, pp. 1–11, 2020.

[7] D. Anade, J.-M. Gorce, P. Mary, and S. M. Perlaza, "An upper bound on the error induced by saddlepoint approximations - applications to information theory," *Entropy*, vol. 22, no. 6, p. 690, Jun. 2020.

[8] Y. Mestrah, "Robust communication systems in unknown environments," Ph.D. dissertation, Computer, automatic and signal processing engineering, 2019.

[9] V. Dimanche, A. Goupil, L. Clavier, and G. Gelle, "On detection method for soft iterative decoding in the presence of impulsive interference," *IEEE Communications Letters*, vol. 18, no. 6, pp. 945–948, June 2014.

[10] Y. Hou, R. Liu, and L. Zhao, "A non-linear LLR approximation for LDPC decoding over impulsive noise channels," *IEEE/CIC ICCC*, pp. 86–90, Oct. 2004.

[11] C. Gao, R. Liu, and B. Dai, "Gradient descent bit-flipping based on penalty factor for decoding LDPC codes over symmetric alpha-stable noise channels," *IEEE/CIC ICCC*, pp. 1–4, Oct. 2017.

[12] R. Yazdani and M. Ardakani, "Linear LLR approximation for iterative decoding on wireless channels," *IEEE Transactions on Communications*, vol. 57, no. 11, pp. 3278–3287, Nov. 2009.

[13] T. J. Richardson and R. Urbkane, *Modern Coding Theory*. Cambridge University Press, 2008.

[14] J. A. Nelder and R. Mead, "A Simplex Method for Function Minimization," *The Computer Journal*, vol. 7, no. 4, pp. 308–313, 01 1965.

[15] Y. Polyanskiy, H. Poor, and S. Verdu, "Channel coding rate in the finiteblocklength regime," *IEEE Transactions on Information Theory*, vol. 56, pp. 2307–2359, Dec. 2010.