Coplanar Turbo-FSK: a Flexible and Power Efficient Modulation for the Internet-of-Things

Yoann Roth, Jean-Baptiste Doré, Laurent Ros, Vincent Berg

Coplanar Turbo-FSK: a Flexible and Power Efficient Modulation for the Internet-of-Things

Yoann Roth*†, Jean-Baptiste Doré*, Laurent Ros† and Vincent Berg*

*CEA-Leti, Minatec Campus, Grenoble, France.

†Univ. Grenoble Alpes, CNRS, Grenoble INP, GIPSA-lab, Grenoble, France.

Corresponding author: J.-B Doré, jean-baptiste.dore@cea.fr - The research leading to these results received funding from the French Agence Nationale de la Recheche (ANR), under grant agreement ANR-16-CE25-0002905 (project EPHYL).

Abstract

As the Internet-of-Things (IoT) has expanded, multiple solutions have attempted to address the issues of the Low Power Wide Area (LPWA) networks physical layer. In a previous work, we proposed the Turbo-FSK, a constant envelope modulation with orthogonal alphabet that allows the receiver to operate at very low levels of power (high sensitivity performance) and very low levels of energy per bit E_b. The scheme was demonstrated to approach Shannon's limit as close as 0.29dB. However, the scheme lacks of flexibility in terms of spectral efficiency (always lower than 10^{-1} bits/s/Hz), especially compared to the recently standardized Narrow-Band IoT (NB-IoT) solution. In this work, we propose an evolution of the initial scheme, so-called Coplanar Turbo-FSK (C-TFSK). In order to increase the spectral efficiency of the system, two new features are introduced: a modulation combining linear and orthogonal properties where only subsets of the alphabet are orthogonal and a puncturing mechanism. Several aspects of the scheme are then studied under asymptotic hypothesis, such as the influence of the linear component of the alphabet and the effects of puncturing. The high flexibility in term of spectral efficiency, the short distance to Shannon's limit and the constant envelope property make the C-TFSK a serious contender for the physical layer of the IoT.

Index Terms

Internet-of-Things, IoT, Low Power Wide Area, LPWA, Turbo, FSK, Constant envelope, Orthogonal modulations

I. Introduction

The Internet-of-Things (IoT) is in the process of global expansion but the definition of its technical features is still in progress [1]. According to the definition from the International Telecommunication Union (ITU) [2], the IoT will interconnect existing networks as well as evolving technologies. In this context, Low Power Wide Area (LPWA) networks [3], [4] are expected to represent approximately 10% of the overall IoT connections [5]. Requirements for LPWA include long range communication and low energy consumption at the device level, i.e. long battery life. Low consumption can be achieved by reducing the constraints on the power amplifier using a constant envelope waveform, or by reducing the energy per bit E_b of the considered modulation [6]. Long range is obtained by decreasing the required received power, also called the sensitivity level.

The problem of achieving a low sensitivity level can be solved by ensuring a low data rate transmission [7]. It is commonly obtained following two possible strategies: reduce the bandwidth or reduce the level of Signal-to-Noise Ratio (SNR) required for the given Quality of Service (QoS). The first strategy leads to narrow band signaling, the keystone of the Sigfox LPWA solution [8]. The second strategy is pursued by reducing the spectral efficiency η of the selected technique (expressed in bits/s/Hz), or by reducing the E_b/N_0, the SNR per bit, required for the QoS. A commonly used technique to reduce the spectral efficiency is use of the repetition code (repetition factor), used for example in the IEEE 802.15.4k standard [9]. Repetitions are recombined at the receiver side, which lowers the level of sensitivity while maintaining a required E_b/N_0 constant. The required E_b/N_0 can be reduced by using Forward Error Correction (FEC) [10]. FEC is usually associated with a loss in spectral efficiency and introduces the notion of capacity.

The maximum achievable data rate over Additive White Gaussian Noise (AWGN) for an arbitrarily small level of error, also called the capacity, was established by Shannon [11]. Many research works have focused on approaching the capacity through the use of FEC [10]. A system operating close to the capacity ensures an efficient if not optimal use of the spectral and energetic resource. Another formulation of the limit expresses the relationship between the minimum achievable E_b/N_0, $(E_b/N_0)_{\min}$, and the maximum achievable spectral efficiency $\eta_{\max}$ with [12]

$$\left(\frac{E_b}{N_0}\right)_{\min} = \frac{2^{\eta_{\max}} - 1}{\eta_{\max}}. \tag{1}$$

This expression illustrates the perpetual trade-off between spectral efficiency and energy efficiency, which is related to the required E_b/N_0. It also gives the ultimate limit in E_b/N_0, obtained when η_{max} tends toward 0 and equal to $(E_b/N_0)_{\text{lim}} \simeq -1.59$dB. Having a high spectral efficiency will result in a high E_b/N_0, and inversely operating a low E_b/N_0 requires the technique to have a low spectral efficiency. For the LPWA context, ensuring both a low spectral efficiency and a low level of required E_b/N_0 will lead to a low level of sensitivity. Moreover, an efficient use of the resource is guaranteed if the long range transmission technique approaches the capacity.

When it comes to existing solutions, the Narrow-Band IoT (NB-IoT) solution from the 3rd Generation Partnership Project (3GPP) [13], [14], [15], [16] combines the elements necessary for an efficient LPWA solution. Through the use of the powerful $[1315]$ Turbo Code (TC) [17], [18], a low E_b/N_0 and a short distance to capacity can be achieved. The additional use of a repetition factor ensures low levels of sensitivity by reducing the spectral efficiency [19]. The variable number of repetitions and the possibility to puncture parity bits generated by the TC allows for a wide range of spectral efficiencies. With a constant bandwidth, applications demanding higher data rates are supported by increasing the spectral efficiency, while reducing the range of the transmission.

An alternative strategy is to consider ultra low rate FEC schemes, such as Super Orthogonal Turbo Codes (SOTC) [20] or Turbo-Hadamard Codes (THC) [21]. The use of orthogonal codes can actually be shown to reach the capacity for an infinite size of alphabet [12]. As this leads to a spectral efficiency tending towards 0, the use of orthogonal codes alone is not practical. Nevertheless, through the use of a combination of orthogonal Hadamard codes, a convolutional code and a turbo receiver, THC have been demonstrated to get as close as 0.29dB from the $(E_b/N_0)_{\text{lim}}$ for a spectral efficiency equal to $1.06 \cdot 10^{-2}$ bits/s/Hz [22]. The joint optimization of modulation and coding scheme have also been studied in the context of deep space communication [23]. In this case Pulse Position Modulation (PPM) is considered, and a turbo-like code is optimized to bridge the gap to Shannon capacity. However, this scheme is not well optimized for the next generation of IoT communications as the size of the codeword is quite important (4000 bits). Moreover PPM is not well adapted compared to constant amplitude modulation when state of the art power amplifier is used. In a previous work, we proposed the Turbo-FSK modulation [7], [19], [24], which is inspired by the THC and can be either viewed as a bit interleaved coded modulation (BICM) or parallel concatenated trellis-coded modulation (PCTCM) [25] schemes with orthogonal signaling. Low consumption of the transmitter is ensured

by a constant envelope signal and a low complexity encoder. Turbo-FSK is adapted to the LPWA context [26] as it ensures low levels of sensitivity to be reached while approaching the channel capacity [27]. However, unlike the technologies relying on the use of a powerful FEC combined with a repetition factor, Turbo-FSK is restrained to low levels of spectral efficiency. This limits the number of possible applications for the solution.

In this paper, we propose an extension of the initial Turbo-FSK technique presented in [7]. An alphabet of modulation combining orthogonal and linear properties, the so-called Orthogonal with Coplanar Subsets (OCS) alphabet is introduced. It allows achieving higher levels of spectral efficiencies. After giving a general definition of the alphabet, its application to Frequency Shift Keying (FSK) signaling is reviewed. The use of this alphabet as a substitution for the FSK alphabet in the Turbo-FSK scheme is considered, leading to the Coplanar Turbo-FSK (C-TFSK) modulation. In addition to the more spectrally efficient alphabet, C-TFSK also includes a puncturing mechanism which trades performance for a shorter transmitted sequence. The introduction of these two new features increases the possible number of configurations for C-TFSK. Additionally, it improves the flexibility of the scheme and makes it a serious contender for IoT technology. The scheme is analyzed using the EXtrinsic Information Transfer (EXIT) chart tool [28], which has already been introduced for the Turbo-FSK [27].It should be noted that minimum distance analysis is also a powerful tool for Treillis-Coded-Modulation scheme, but is not straightforward to derive for our PCTCM-like scheme (it has been partially done in [29][24](page 38) for more basic scheme). The performance evaluation presented in this work is limited to the AWGN channel. However, this is motivated by the following reasons. (i) This type of channel is convenient to consider for the theoretical study of the structure of the code. (ii) For low throughput scenarios it can be reasonable to consider the transmission channel as non-frequency selective channel. It is particularly true for small bandwidth and terrestrial networks. It should be however mentioned that performance evaluation of C-TFSK under frequency selective channels demonstrated that the scheme is competitive versus standard Orthogonal Frequency Division Multiplexing (OFDM) signaling [30].

The paper is organized as follows. Section II is dedicated to the presentation of the OCS alphabet and Section III to the C-TFSK transmitter and the receiver's architecture. In Section IV, various features of the scheme are studied and Section V concludes the paper.

II. Orthogonal with Coplanar Subsets Modulation

The modulation mapping procedure consists in associating binary words of size m to a symbol (or codeword) taken from the alphabet of modulation of size $M = 2^m$. There are two main families of modulations: linear modulations and orthogonal modulations. For linear modulations, the alphabet is constructed from a base symbol which is multiplied by M various complex values. Since the symbols belong to the complex plane, there is a coplanar relation between the symbols of the alphabet. Linear modulations are spectrally efficient, both the spectral efficiency and the required E_b/N_0 increase with the size of alphabet, without limitation. Quadrature Amplitude Modulation (QAM) and Phase Shift Keying (PSK) modulation are examples of linear modulations. For orthogonal modulations, symbols of the alphabet are orthogonal to each other (the scalar product is equal to 0). These modulations are energy efficient as the energy efficiency increases with the size of the alphabet, but inversely the spectral efficiency decreases with an increasing size of alphabet. Theoretically, orthogonal modulations can reach the $(E_b/N_0)_{\text{lim}}$ for an infinite size of alphabet [12]. However, this is not conceivable for a practical system, as the spectral efficiency would tends toward 0. FSK and Pulse Position Modulation (PPM) are examples of orthogonal modulations.

Linear and orthogonal modulations, due to their complementary spectral efficiencies and required E_b/N_0 ranges, seem to be fundamentally opposed. We propose nevertheless the definition of the hybrid OCS modulation, which holds properties from both modulations. This alphabet is a generalization of the modulation presented in [31]. The hybridation of FSK modulation and a linear modulation was also presented in [32], [33], [34], [35], but no channel coding was considered. [36] considered the use of binary and multidimensional TCs associated with a hybrid FSK and QAM modulation.

A. Definition of the Alphabet of Modulation

The alphabet of modulation is denoted by $\boldsymbol{\mathcal{A}}$ and contains M elements. While having an orthogonal alphabet would require a total of M orthogonal dimensions, we relax the orthogonality constraint for subsets of the alphabet only. These subsets contain M_L symbols (with $M_L \geq 1$) and are denoted by $\boldsymbol{\mathcal{A}}_\delta$ with $\delta \in \{0, \ldots, M_O - 1\}$, where M_O is the number of orthogonal dimensions, with $M = M_O M_L$. Two symbols from two distinct subsets are orthogonal, but two symbols from the same subset belong to the same complex plane and are thus coplanar.

Considering two symbols $\boldsymbol{s}^i$ and $\boldsymbol{s}^{i'}$ from the alphabet (with $\{i, i'\} \in \{0, \ldots, M-1\}^2$), the definition of the alphabet leads to

$$\langle \boldsymbol{s}^i, \boldsymbol{s}^{i'} \rangle = \begin{cases} A_{i-i'} & \text{if } \exists\ \delta \in \{0,\ \ldots,\ M_O - 1\},\ \text{such that } \boldsymbol{s}^i, \boldsymbol{s}^{i'} \in \boldsymbol{\mathcal{A}}_\delta \\ 0 & \text{otherwise} \end{cases},$$

where $A_{i-i'}$ is a complex number proportional to the coplanarity relation between two symbols from the same subset and $\langle \cdot, \cdot \rangle$ is the scalar product operator[1] defined as follows:

$$\langle \boldsymbol{x}, \boldsymbol{y} \rangle = \sum_i x_i \overline{y_i}$$

with $\overline{y_i}$ denotes the conjugate of complex number y_i.

In the alphabet, each subset is equivalent to a linear modulation with M_L elements. Orthogonality holds between the subsets only.

We consider the baseband model. Through the addition of the M_L linear components, the number of orthogonal dimensions have been divided by M_L and so was the number of samples (or chips) required to represent each symbol. M_O chips are now required to represent one symbol. Symbols from the alphabet are expressed as

$$\boldsymbol{s}^i = \boldsymbol{f}^\delta z_p, \tag{2}$$

with i a function of δ and p and where $\boldsymbol{f}^\delta = \left[f_0^\delta,\ f_1^\delta,\ \ldots,\ f_{M_O-1}^\delta\right]$ is the length-M_O base vector of the subset $\boldsymbol{\mathcal{A}}_\delta$. The coefficient $z_p \in \mathbb{C}$, with $p \in \{0,\ \ldots,\ M_L - 1\}$, is taken from the complex plane and represents the linear modulation. The spectral efficiency of the modulation η, expressed in bit/s/Hz, is given by the ratio between the number of information bits contained in one symbol and the number of chips per symbols. It is given by

$$\eta = \frac{\log_2(M)}{M_O} = \frac{\log_2(M)}{M} \cdot M_L. \tag{3}$$

The second form of the expression offers a better understanding of the effect of the modulation. While keeping a constant value for M, the spectral efficiency increases with the value of M_L, which reminds the property of linear modulation. However, increasing the size of alphabet for a constant M_L (i.e. increasing M_O, the number of orthogonal dimensions) reduces the spectral efficiency, similarly to orthogonal modulations.

[1]It should be noticed that the number of orthogonal dimension is equal to the number of chips M_O and not $2M_O$ since we need to use the complex scalar product instead of the real one.

B. Likelihood of a Symbol

Probabilistic demodulation is considered. The likelihood of all the symbol from the alphabet must be evaluated. The observation $\boldsymbol{r}$ of the symbol $\boldsymbol{s}$ is given by

$$\boldsymbol{r} = \boldsymbol{s} + \boldsymbol{n},$$

where $\boldsymbol{n}$ is a zero-mean circular complex Gaussian white noise. The elements of $\boldsymbol{n}$ are i.i.d. Gaussian random variables with zero mean and variance equal to σ_{n}^2.Assuming that $\boldsymbol{s}^i$ was transmitted, the likelihood of the symbol is given by

$$\begin{aligned} p\left(\boldsymbol{r} \mid \boldsymbol{s}^i\right) &= \prod_{k=0}^{M_O-1} p\left(r_k \mid s_k^i\right) \\ &= \left(\frac{1}{2\pi\sigma_{\mathrm{n}}^2}\right)^{M_O} \exp\left\{-\frac{1}{2\sigma_{\mathrm{n}}^2}\sum_{k=0}^{M_O-1}|r_k - s_k^i|^2\right\} \\ &= \mathcal{C}\exp\left\{-\frac{1}{2\sigma_{\mathrm{n}}^2}\sum_{k=0}^{M_O-1}\left|s_k^i\right|^2 + \frac{1}{\sigma_{\mathrm{n}}^2}\sum_{k=0}^{M_O-1}\langle r_k, s_k^i\rangle\right\}, \end{aligned} \tag{4}$$

where $\mathcal{C}$ is a constant independent of i. Following the definition of the vector $\boldsymbol{s}$ given in (2), the summation of the scalar product over k is expressed

$$\begin{aligned} \sum_{k=0}^{M_O-1}\langle r_k, s_k^i\rangle &= \sum_{k=0}^{M_O-1}\mathrm{Re}\left(r_k \cdot \overline{f_k^\delta} \cdot \overline{z_p}\right) \\ &= \mathrm{Re}\left(\overline{z_p} \cdot \langle \boldsymbol{r}, \boldsymbol{f}^{\boldsymbol{\delta}}\rangle\right), \end{aligned} \tag{5}$$

where $\langle \boldsymbol{r}, \boldsymbol{f}^{\boldsymbol{\delta}}\rangle$ represents the projection of the received vector $\boldsymbol{r}$ on the subset $\boldsymbol{\mathcal{A}}_\delta$ which includes $\boldsymbol{s}^i$.

Additionally, the definition of the symbols in (2) yields

$$\sum_{k=0}^{M_O-1}\left|s_k^i\right|^2 = |z_p|^2 \cdot \left\|\boldsymbol{f}^\delta\right\|^2. \tag{6}$$

Using (5) and (6) in (4), the likelihood of the symbol $\boldsymbol{s}^i$ is expressed

$$p\left(\boldsymbol{r} \mid \boldsymbol{s}^i\right) = \mathcal{C}\exp\left\{-\frac{|z_p|^2}{2\sigma_{\mathrm{n}}^2} \cdot \left\|\boldsymbol{f}^\delta\right\|^2 + \frac{1}{\sigma_{\mathrm{n}}^2}\mathrm{Re}\left(\overline{z_p} \cdot \langle \boldsymbol{r}, \boldsymbol{f}^\delta\rangle\right)\right\}. \tag{7}$$

Maximum Likelihood (ML) decision could be performed by searching the symbol $\boldsymbol{s}^i$ which has the greatest likelihood.

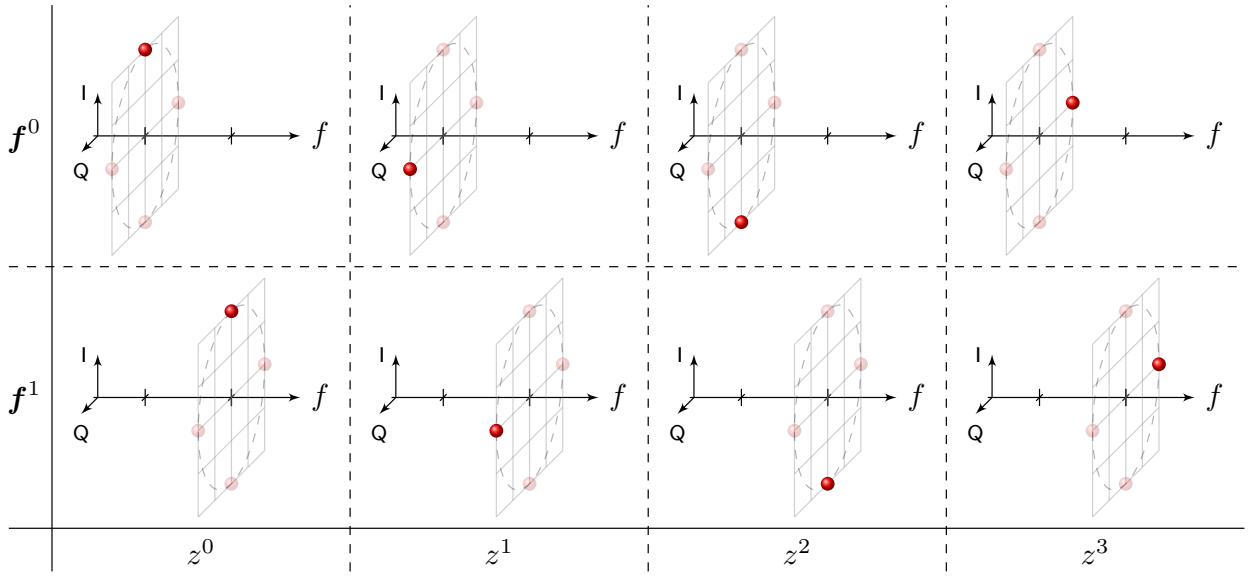


Figure 1: Alphabet of modulation for the case $M = 8$, 2-FSK and $M_L = 4$.

C. Application to FSK Signaling

In practice, common modulations can be chosen to represent both orthogonal and linear parts of the OCS modulation. For the considered application C-TFSK, we select the FSK modulation as orthogonal modulation. For this specific case, the vectors $\boldsymbol{f}^\delta$ can be picked as rows or columns from the Discrete Fourier Transform (DFT) matrix of size M_O. The coplanar value can take any complex value. However, we will consider the case where roots of the unity are selected in order for the modulation to have a constant envelope. The M_L values are given by $z_p = \exp\left\{j\frac{2\pi}{M_L}p\right\}$, which reminds of the construction of the PSK modulation alphabet. In Figure 1, the alphabet of modulation for the case $M = 8$ and $M_L = 4$ is depicted. For this example, the vectors $\boldsymbol{f}^\delta$ are of size 2. In the figure, an entire line of symbols corresponds to a subset and rows are orthogonal to each other. The spectral efficiency, given by (3), is equal to $3/2$. Due to the choice of the FSK modulation as orthogonal modulation and roots of the unity as coplanar points, this modulation has a constant envelope.

III. Coplanar Turbo-FSK

The alphabet introduced previously is now applied to the Turbo-FSK principle presented in [19]. After reviewing the operations performed at the transmitter side, the receiver is presented.

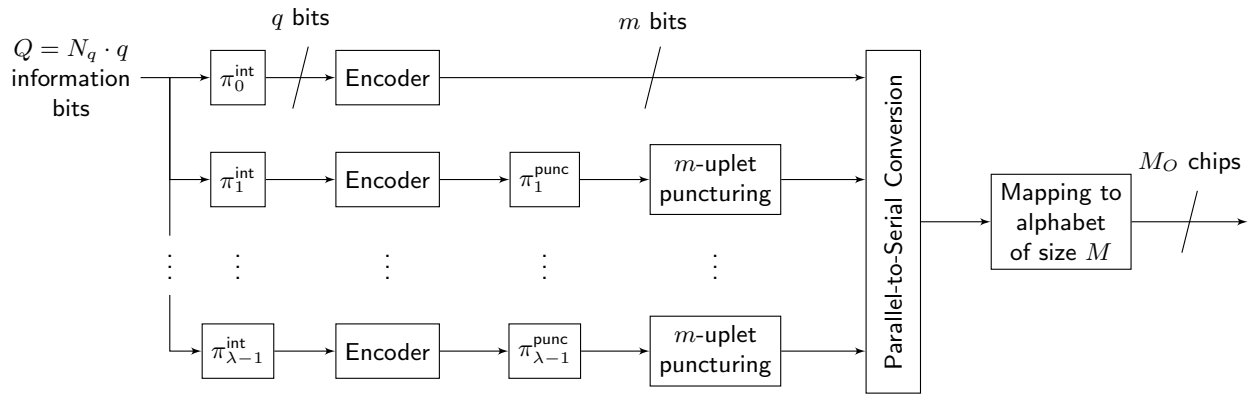


Figure 2: The C-TFSK transmitter architecture.

A. *Transmitter*

The architecture of the transmitter is presented in Figure 2. Q is the information block size, and it is divided into N_q binary words of size q (thus $Q = N_q \cdot q$). The transmitter is composed of λ similar stages and the stage $\ell \in \{0, \ldots, \lambda - 1\}$ is considered. The operations performed by one stage are presented.

1) Encoder: The information block is interleaved with the function π_ℓ^{int}, where int stands for internal, as this interleaver corresponds to the interleaver used for the turbo process. The N_q words of length q are then encoded using the encoder presented in Figure 3. The encoder in each branch adds one parity bit to each length-q information word and an extra word to force the memory state back to 0. The rate of the code (one branch) is

$$R_c = \frac{Q}{(N_q + 1)(q + 1)} \approx \frac{q}{q + 1},$$

where the approximation is valid for large sizes of Q. The size of the encoded words is denoted m, with $m = q + 1$. For each stage, the sequence of $N_q + 1$ binary words of length m is to be modulated. Each binary word of size m will be associated with one of the elements of the alphabet of modulation $\boldsymbol{\mathcal{A}}$, referred to as codewords. The encoder needs to be systematic as the codewords will be mapped to a physical waveform. Omitting the systematic bits as in the case of Turbo-Hadamard [21] cannot be considered when dealing with non-binary transmitted codewords. Contrary to usual efficient binary coding scheme, note that the systematic bits are then sent multiple times, and that more than $\lambda = 2$ stages should be considered for turbo-encoders.

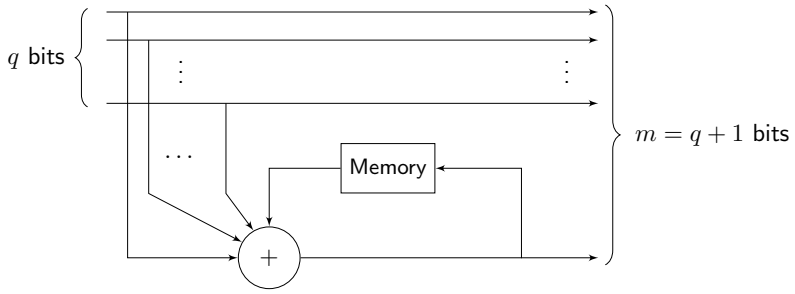


Figure 3: The Parity-Accumulator encoder

However, the performance of this type of architecture has been demonstrated previously for Turbo-Hadamard (see [29] figure 3) as well as for the Turbo-FSK scheme (see [19],[30]). In order to increase the spectral efficiency of the scheme, a puncturing step is also considered.

2) Puncturing: Puncturing commonly consists in removing some parity bits of the encoded sequence. Only the systematic bits and a reduced number of parity bits are sent. For the considered scheme, removing some parity bits would involve changing the size m to q for some specific words. However, the size of the words to be modulated should be kept to a constant value, equal to m. In order to fulfill this condition, entire words of size m are punctured. If the first stage is not punctured, then every systematic bit is sent at least once. Every other stage is punctured at a rate $R_p^\ell \geq 1$. This rate may differ depending on the stage. The overall puncturing rate is denoted by R_p, and is equal to

$$R_p = \frac{\lambda(N_q+1)}{(N_q+1)+\sum_{\ell=1}^{\lambda-1}\frac{(N_q+1)}{R_p^\ell}} = \frac{\lambda}{1+\sum_{\ell=1}^{\lambda-1}\frac{1}{R_p^\ell}},$$

with $R_p \geq 1$.

In order to avoid puncturing too many consecutive codewords and induce holes in the trellis, a block interleaver π_ℓ^{punc} of size G is used. The block is interleaved by filling a matrix of size $G \times (N_q+1)$ row by row, and the same matrix is read column by column to form the interleaved output. Considering the puncturing rate of stage ℓ to be R_p^ℓ, the actual step of puncturing consists in selecting only the first $(N_q+1)/R_p^\ell$ binary words. The use of the block interleaver scatters the punctured binary words over the whole sequence.

3) Modulation: After puncturing, the overall sequence consists of $\lambda(N_q+1)/R_p$ binary words of size m. These words are mapped to the alphabet of codewords $\boldsymbol{\mathcal{A}}$ of size M. The alphabet is

chosen as the OCS alphabet with FSK as orthogonal modulation, as described in Section II-C. The number of elements in each subset is M_L and the number of frequencies is given by the ratio M/M_L. The size of the alphabet is given by $M = 2^m$. With the spectral efficiency of the modulation as defined by Equation (3), the overall spectral efficiency of the scheme is given by

$$\eta = \frac{N_q \cdot q}{\lambda(N_q + 1)M} \cdot M_L \cdot R_p. \tag{8}$$

under the assumption of critically sampled and spaced frequency shift signals, i.e an orthogonal basis built from discrete Fourier series. Following the relation $m = q + 1$, the number of information bits per codeword is given by

$$q = \log_2(M) - 1.$$

The spectral efficiency can be approximated, for large sizes of information block Q, with

$$\eta \approx \frac{\log_2(M) - 1}{\lambda M} \cdot M_L \cdot R_p.$$

This expression includes the 4 main parameters which define the scheme: the number of stages λ, the size of alphabet M, the number of elements in each orthogonal subset M_L and the puncturing rate R_p (the number of frequencies, denoted previously with M_O, is obtained with M/M_L). The introduction of puncturing and a linear component allows for higher ranges of spectral efficiency. Increasing the size M_L or the puncturing rate R_p clearly increases the spectral efficiency. Typically, multiplying M_L by 2 with a constant size of alphabet M increases the spectral efficiency by a factor 2. By setting $M_L = 1$ and without puncturing, the scheme is strictly equivalent to the Turbo-FSK from [26], [19].

4) Trellis: A trellis is established due to the use of the accumulator in the encoding process. Each input information word is associated to one of the 4 possible transitions. As there are $M = 2^m$ possible encoded words, each transition has $M/4$ branches. The mapping performed between the binary words of length m and the codewords from the alphabet can be represented on the trellis.

An example of trellis for the case $M = 32$ and $M_L = 8$ is depicted in Figure 4. For this configuration, there are 8 branches for each transition. The mapping is then done so that each transition corresponds to a unique orthogonal subset. All the parallel branches are mapped by codewords belonging to the same subset. Since FSK is considered for the orthogonal modulation, the vectors f are lines of the DFT matrix of size 4.

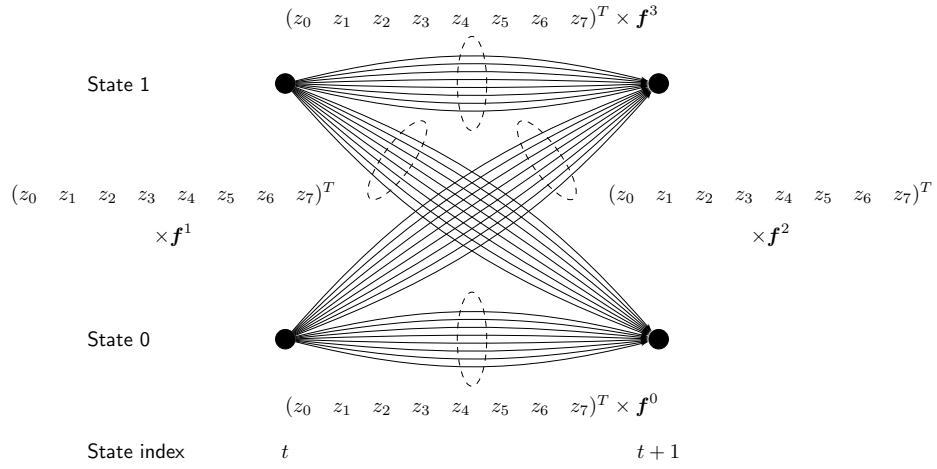


Figure 4: A section of the trellis for the configuration $M = 32$ and $M_L = 8$, with one possible mapping .

When the chosen alphabet is orthogonal without any linear component (i.e. $M_L = 1$), the choice of mapping has no influence. The distance between the branches in the trellis is a constant as all branches are in different orthogonal dimensions. However, when considering $M_L > 1$, this is no longer the case. Codewords belonging to a same subset are not orthogonal between each other. The distance between the various symbols of the alphabet is modified by the introduction of a linear component. When mapping the binary words to the codewords, consideration must be taken on the orthogonality of the trellis, i.e. on the distances between the symbols associated to the transitions.

Using the mapping presented in Figure 4, the transitions in the trellis are ensured to be further apart, as the distance between the orthogonal subsets is larger than the minimum distance in each subset. This may not be the case for other mappings or for other values of M and M_L. The influence of the choice of mapping on the performance is studied later.

B. Receiver

1) Derivation of the Receiver: First of all, we consider in this work a coherent receiver perfectly synchronized in time and frequency. As in conventional turbo receivers, the Bahl, Cocke, Jelinek and Raviv (BCJR) algorithm [37] is used in order to compute the log-*A Posteriori*

Probabilities (APP) ratios of the information bits. The log-APP are given by

$$L\left(b_{n,t} \,\middle|\, \boldsymbol{R}_1^{N_q+1}\right) = \log \frac{\Pr\left(b_{n,t}=0 \,\middle|\, \boldsymbol{R}_1^{N_q+1}\right)}{\Pr\left(b_{n,t}=1 \,\middle|\, \boldsymbol{R}_1^{N_q+1}\right)},$$

where $\boldsymbol{R}_1^{N_q+1}$ is the received codewords sequence and $b_{n,t}$ the bit at index n of the information word at time index t. The derivations of the receiver are similar to the case of Turbo-FSK presented in [7]. The use of the BCJR introduces the terms α, β and γ similarly to [17]. The final expression of the log-APP is

$$L\left(b_{n,t} \,\middle|\, \boldsymbol{R}_1^{N_q+1}\right) = \log \frac{\sum\limits_{i\in\mathcal{B}_{+1}^n} \alpha_{t-1}(s_i') \cdot p\left(\boldsymbol{r}_t \mid \boldsymbol{c}^i\right) \Pr\left(\boldsymbol{c}^i\right) \cdot \beta_t(s_i)}{\sum\limits_{i\in\mathcal{B}_{-1}^n} \alpha_{t-1}(s_i') \cdot p\left(\boldsymbol{r}_t \mid \boldsymbol{c}^i\right) \Pr\left(\boldsymbol{c}^i\right) \cdot \beta_t(s_i)}, \tag{9}$$

where $\mathcal{B}_{+1}^n$ (resp. $\mathcal{B}_{-1}^n$) is the group of codewords that encodes information words for which the bit b_n is equal to $+1$ (resp. -1). $p\left(\boldsymbol{r}_t \mid \boldsymbol{c}^i\right)$ is the likelihood of the codeword $\boldsymbol{c}^i$ and $\Pr\left(\boldsymbol{c}^i\right)$ its *a priori* probability of the codeword, given by the other decoders. Considering the chosen OCS alphabet, for which the symbols likelihood was developed in Section II-B, Equation (7), and with $\boldsymbol{c}^i = z_p \boldsymbol{f}^\delta$, the product can be expressed

$$p\left(\boldsymbol{r}_t \mid \boldsymbol{c}^i\right) \Pr\left(\boldsymbol{c}^i\right) = \mathcal{C} \exp\left\{ -\frac{|z_p|^2}{2\sigma_{\mathrm{n}}^2} \cdot \left\|\boldsymbol{f}^\delta\right\|^2 + \frac{1}{\sigma_{\mathrm{n}}^2} \mathrm{Re}\left(\overline{z_p} \cdot \langle \boldsymbol{r}_t,\ \boldsymbol{f}^\delta \rangle\right) + \frac{1}{2} \sum_{k=0}^{q-1} L_A(b_k) b_k^i \right\}. \tag{10}$$

In this expression, both the orthogonal and linear components of the codeword $\boldsymbol{c}^i$ appear.

2) Architecture: The architecture of the receiver is depicted in Figure 5. The λ stage observations are first retrieved. For each stage, the observation consists in $(N_q+1)/R_p^\ell$ codewords of length $M_O = M/M_L$. The detail of one stage is depicted in Figure 6. The detector computes both the correlation on the orthogonal dimensions and the two first terms of (10) in order to estimate a quantity related to the likelihood of the M codewords. The result is a matrix of size $(N_q+1)/R_p^\ell \times M$. Then, the depuncturing block appends $(N_q+1)(1-1/R_p^\ell)$ extra lines to the aforementioned matrix, corresponding to the same number of punctured codewords at the transmitter side. For these lines, the M likelihoods are equal, i.e. the receiver considers each codeword to be equally probable. Interleaving is then applied and scatters the appended lines over the whole sequence. This step requires the receiver to know the puncturing rate as well as the puncturing pattern for each stage. When considering the trellis representation, puncturing consists in not transmitting some sections of the trellis. The receiver then considers every branch

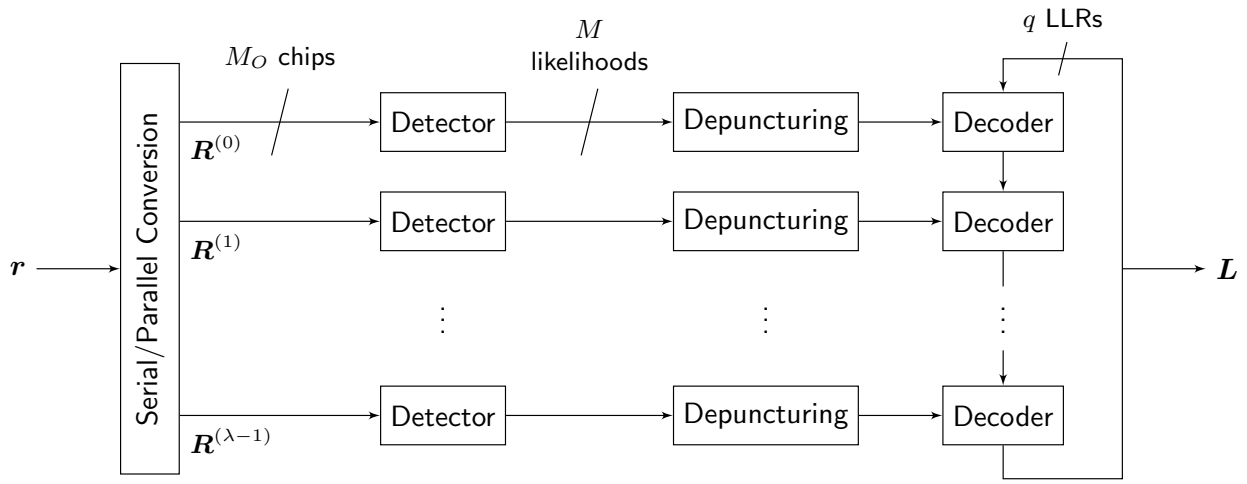


Figure 5: The C-TFSK receiver architecture.

of the punctured sections to be equally probable. This procedure is similar to the puncturing and depuncturing process widely considered with Turbo codes [38].

The detector and the puncturing management provide the $\log$ of the likelihood part of (10). It is fed to the decoder, which first computes the $\log$ of the *a priori* part (the third term) of (10) by combining the *a priori* $\log$ ratios. After summation, the $\log$ of (10) is fed to the BCJR algorithm, which estimates the $\log$-APP of the information bits using Equation (9). The extrinsic $\log$ ratio is then computed and stored for use at the following iteration. After interleaving, the $\log$-APP are fed to the next decoder. Complexity of the BCJR algorithm can be reduced by using the common $\max$-$\log$ approximation. It avoids the computation of exponential functions, which are required when computing the Maximum *A Posteriori* (MAP) algorithm. For the case where the FSK orthogonal alphabet is selected, the correlation on the orthogonal dimensions may be done using the Fast Fourier Transform (FFT) algorithm. The detector would then consist in a FFT operation followed by normalization and real part selection operations.

The initial motivation for C-TFSK was to introduce flexibility in terms of spectral efficiency. Starting from the initial Turbo-FSK scheme, a specific puncturing procedure is introduced in order to reduce the number of transmitted codewords by a factor R_p, the puncturing rate. The alphabet of codewords is also modified by introducting M_L linear components, so that the size of the codewords is divided by M_L, while the number of elements of the alphabet stays equal to M. The introduction of these linear components may also be interpreted as a reduction in the

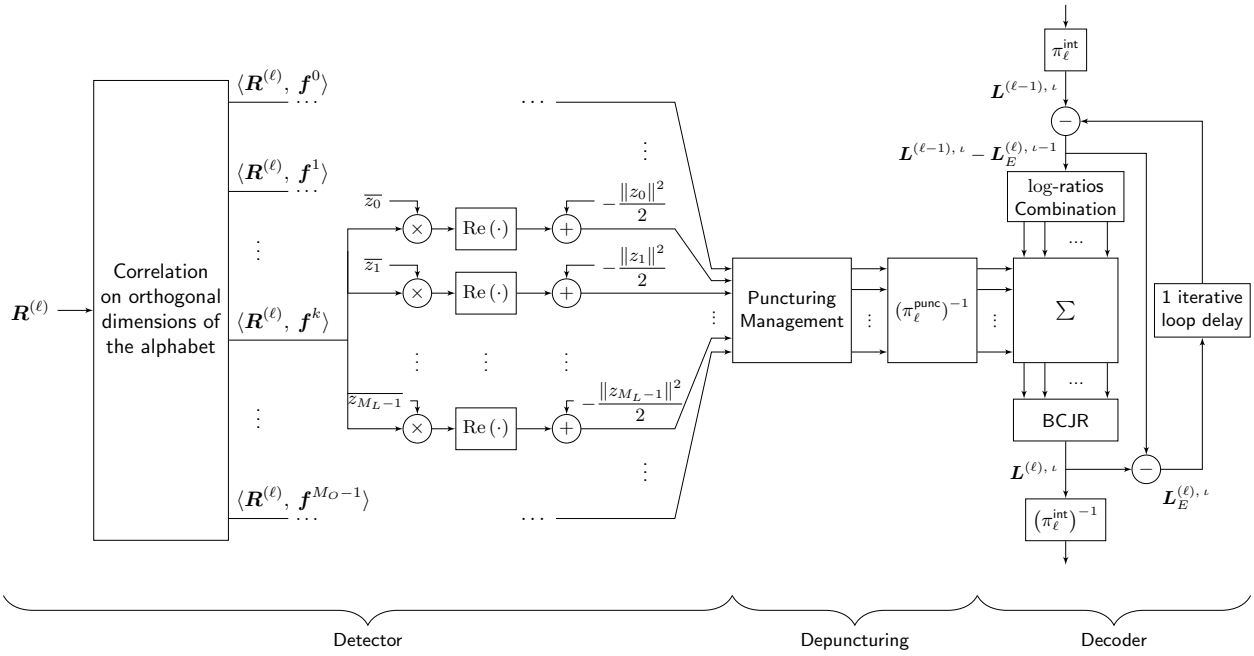


Figure 6: The C-TFSK one stage receiver's architecture.

orthogonality of the alphabet. In the next section, we propose a study of the performance of the scheme.

IV. Analysis

The design of the scheme gives 5 parameters to study: the alphabet size M, the size of the linear modulation M_L, the number of stages λ, the puncturing rate R_p and the choice of mapping on the trellis. In this section, each parameter is studied and the performance of the scheme according to this parameter is assessed. The EXIT chart for Turbo-FSK, presented in [27], was demonstrated to be able to predict the asymptotic behavior of the decoder, and that performance when the block size is shortened follow the same trend. Its use for the analysis of the C-TFSK is considered.

The impact of the choice of how to map codewords on the trellis is first studied. The puncturing mechanism, introduced for the C-TFSK scheme is then studied. It allows for gains in spectral efficiency, at the expense of some performance loss, which is evaluated for some use cases. Finally, some configurations of the scheme are compared to the maximum achievable spectral efficiency for a given E_b/N_0, and the flexibility in terms of spectral efficiency of the C-TFSK

Section	Study	Changing Parameter(s)	Constant Parameter(s)
III.A	Influence of the Choice of Trellis Mapping	M, λ, Mapping	R_p
III.B	Effects of Puncturing	R_p	M, M_L, λ, Mapping
III.C	Impact on the Spectral Efficiency	M, M_L, λ	R_p, Mapping

Table I: Organization of the section concerning the analysis of the parameters.

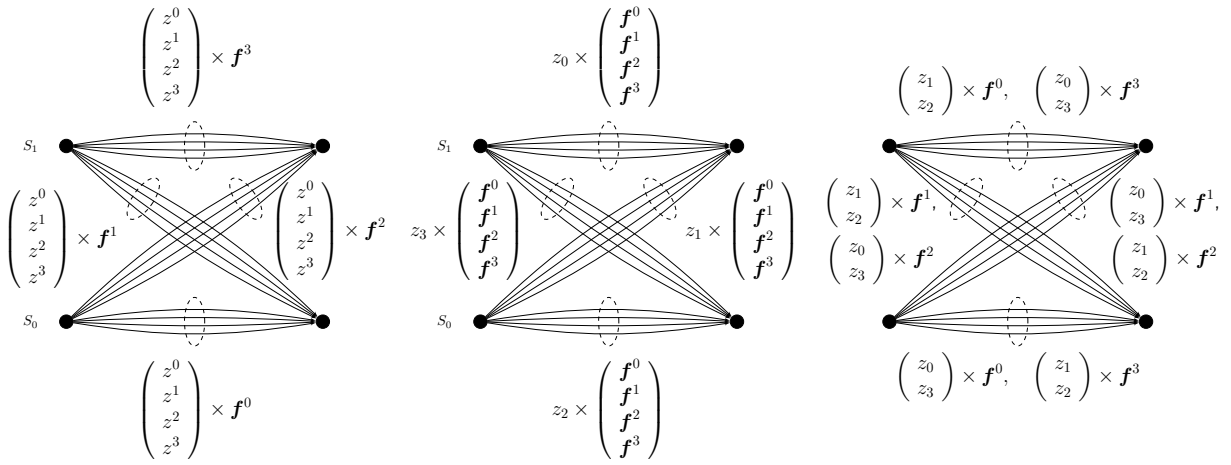
is demonstrated. The starting E_b/N_0 value of the waterfall can be estimated by computing the threshold of the decoder (also called the turbo cliff position). It is the minimum level of E_b/N_0 required for the decoder to be able to correct all the erroneous bits, regardless of the interleaver size and the number of iterations performed. It represents the asymptotic behavior of the decoder.

Table I summarizes the content of the section and mentions what are the changing or constant parameters for each subsection. For all simulations, FSK is used as orthogonal modulation and PSK as linear modulation. When Bit Error Rate (BER) performance is computed, a random interleaving function is selected.

A. Influence of the Choice of Trellis Mapping

As presented when describing the trellis, the introduction of linearities in the alphabet can also be interpreted as a loss of orthogonality. For the initial case of the Turbo-FSK, the M codewords of the alphabet are orthogonal to each other. When using the OCS modulation, only subsets (each containing M_L codewords) are orthogonal to each other. The choice of the mapping of the binary words of length m on the alphabet of size M can be expected to have an impact on the performance. In order to quantify that impact, various trellis and mappings are compared. The thresholds are estimated using the EXIT chart technique. No puncturing is applied (i.e. $R_p = 1$) and the values z_p are taken as roots of the unity.

1) Trellis Classification: Since the alphabet has M elements and there are M branches in the trellis ($M/4$ branches per transition), the total number of mapping strategies is equal to $M!$.



(a) (b) (c)

Figure 7: Example of $\perp$ mapping (a), L mapping (b) and X mapping (c) for the case $M = 16$ and $M_L = 4$.

As this value increases exponentially (for example, $16! \simeq 2.09 \cdot 10^{13}$), we focus our analysis on three types of mapping rules.

The first type, denominated $\perp$ mapping, is characterized by the group of orthogonal vectors $\mathcal{F}_{s' \to s}$, and

$$\boldsymbol{c}^i \in \mathcal{T}_{s' \to s} \quad \text{if} \quad \boldsymbol{f}^\delta \in \mathcal{F}_{s' \to s},$$

where $\boldsymbol{c}^i = z_p \boldsymbol{f}^\delta$ and $\mathcal{T}_{s' \to s}$ is the group of branches corresponding to the transition from s' to s. This means that two different transitions are mapped by codewords from different subsets. This mapping ensures orthogonality between the transitions. The example of mapping for $M = 32$ and $M_L = 8$ given in Figure 4 belongs to the $\perp$ mapping, with *e.g* $\mathcal{F}_{1 \to 1} = \{\boldsymbol{f}^3\}$. In Figure 7 (a), the case $M = 16$ and $M_L = 4$ is represented for a $\perp$ mapping.

The second type of mapping considered is the L mapping. It is characterized by the group $\mathcal{Z}_{s' \to s}$ with

$$\boldsymbol{c}^i \in \mathcal{T}_{s' \to s} \quad \text{if} \quad z_p \in \mathcal{Z}_{s' \to s}.$$

With this mapping, two different transitions are mapped by codewords with different z_p, but that may belong to the same subset. There is no orthogonality between the transitions, but there

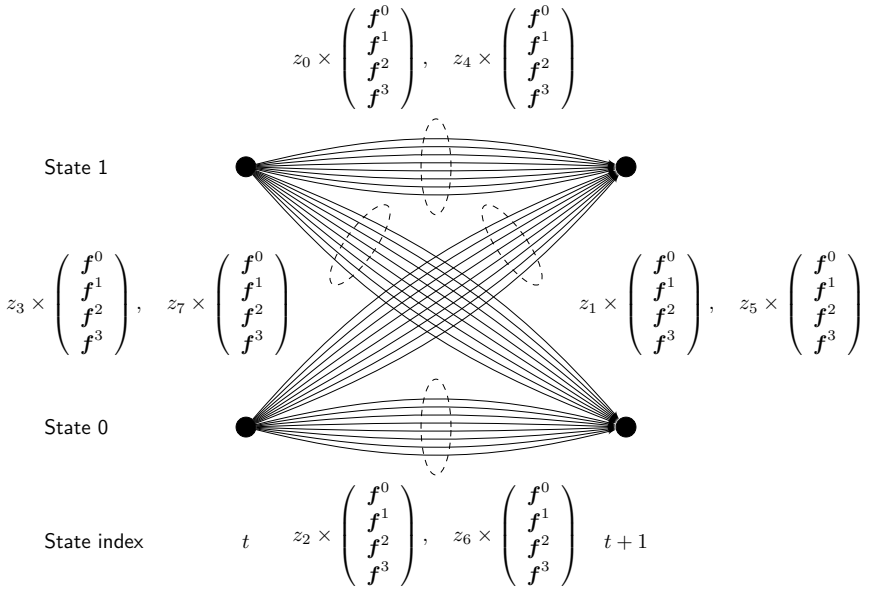


Figure 8: A section of the trellis with L mapping for the configuration $M = 32$ and $M_L = 8$.

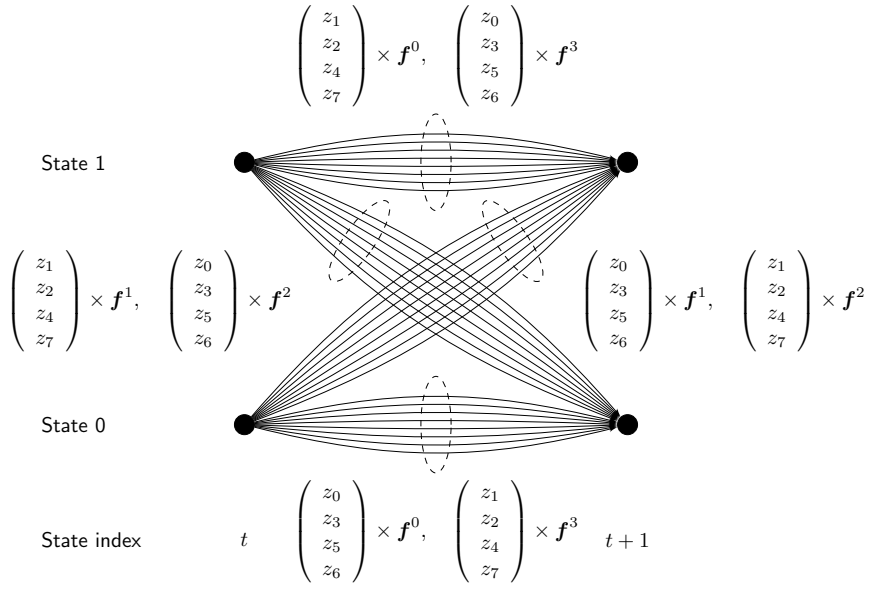


Figure 9: A section of the trellis with X mapping for the configuration $M = 32$ and $M_L = 8$.

is orthogonality between the branches of a transition. A L mapping for the case $M = 32$ and $M_L = 8$ is depicted in Figure 8, with for example $\mathcal{Z}_{1\to1} = \{z^0,\, z^4\}$. In Figure 7 (b), the case $M = 16$ and $M_L = 4$ is represented for a L mapping.

The third mapping, denominated X mapping, is a hybrid mapping. For each transitions, the $M/4$ branches are mapped in $M_O/2$ of the orthogonal dimensions, and each orthogonal dimension can be modulated by $M_L/2$ of the linear values. The mapping is selected so that pairs of orthogonal dimensions are modulated by the M_L linear values. An X mapping for the case $M = 16$ and $M_L = 4$ is depicted in Figure 7 (c), and another for the case $M = 32$ and $M_L = 8$ is depicted in Figure 9. This mapping somehow mixes the two mappings $\perp$ and L. The pair of transitions $1 \rightarrow 1$ and $0 \rightarrow 0$ is orthogonal to the pair $1 \rightarrow 0$ and $0 \rightarrow 1$, but the two transitions $1 \rightarrow 1$ and $0 \rightarrow 0$ are mapped by codewords in the same orthogonal subsets.

The three mapping types correspond to several of the $M!$ possible mappings, and two different mappings of the same type may lead to variation in performance. However, each mapping defines a general structure for the trellis, and the study will help evaluate the impact of the mapping on the performance.

2) Threshold Comparisons: In order to estimate the asymptotic performance, thresholds of the C-TFSK with various parameters and mappings are computed. The EXIT Chart technique was used to estimate the thresholds, in a similar process as the one presented in [26]. For each size of alphabet M, four trellis mappings are considered: the three mappings previously introduced and the fully orthogonal mapping, where each branch is orthogonal to all the other branches. This mapping is strictly equivalent to the Turbo-FSK ($M_L = 1$).

For the first comparison, we consider the C-TFSK where the number of orthogonal subsets is kept constant and equal to 4. The number of elements in the subsets increases with M as $M_L = M/4$. The estimated thresholds for values of M spanning from 16 to 512 and with the parameter $\lambda = 4$ are depicted in Figure 10. The trellis depicted and presented previously actually corresponds to some of the mappings used for these simulations: the case $M = 16$ with $\perp$ mapping is depicted in Figure 7 (a); the case $M = 32$ with $\perp$ mapping is depicted in Figure 4, the case $M = 16$ with L mapping is depicted in Figure 7 (b), the case $M = 32$ with L mapping is depicted in Figure 8 and the case $M = 32$ with X mapping is depicted in Figure 9.

Since one of the properties of the linear modulations is to have an energy efficiency decreasing with the order of the modulation, a loss in performance is expected when increasing the value of M_L. For all cases but $M = 16$, the thresholds of the C-TFSK is indeed higher than the threshold of the Turbo-FSK, for all types of mapping rules. However, the loss varies depending on the selected mapping. More particularly, the L mapping (which totally gives up orthogonality between the transitions) has a threshold more than 6dB higher than the Turbo-FSK case for $M =$

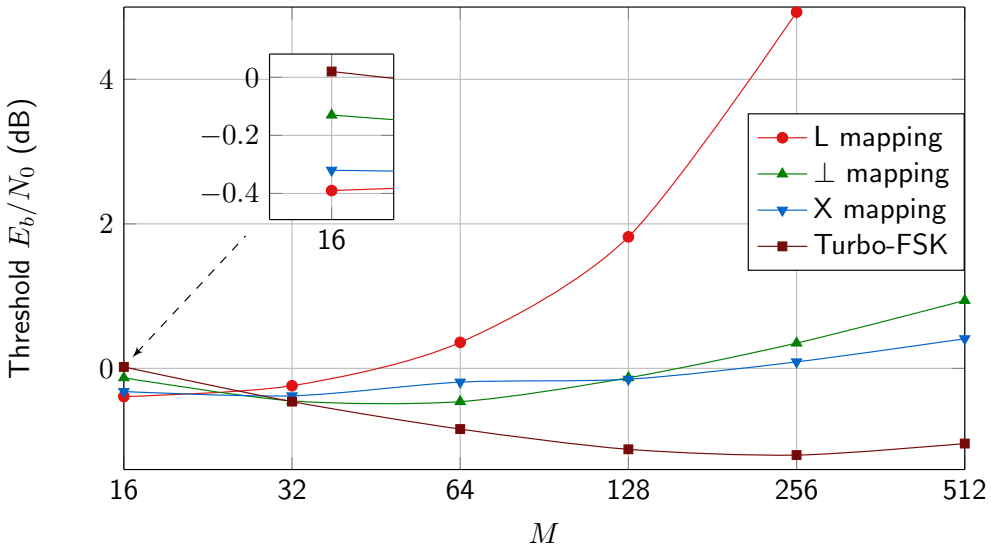


Figure 10: Thresholds for various size of alphabet M and mappings, with a constant number of orthogonal dimensions $M_O = 4$ (4-FSK) and $\lambda = 4$.

256 (i.e. $M_L = 64$), while using the X mapping reduces this gap to 1dB. This mapping performs better than the $\perp$ mapping for larger sizes of alphabet, with a gain of 0.5dB for $M = 512$. This figure emphasizes the importance of choosing mappings that maintain a certain degree of orthogonality between the transition in the trellis. However, preserving complete orthogonality between the transitions (i.e. $\perp$ mapping) reveals itself less interesting than introducing linearities between some transitions (i.e. X mapping).

The case $M = 16$ (thus $M_L = 4$) shows interesting results. Indeed, for this size of alphabet, the C-TFSK performs better than the Turbo-FSK. Also, the L mapping outperforms the two other mappings, and offers 0.4dB of gain versus the Turbo-FSK. The trellis for this mapping and configuration is depicted in Figure 7 (b).

In order to explain these differences in performance, it is necessary to study the distance between the various codewords in the alphabet. A useful representation is the projection of all codewords of the alphabet on one of the orthogonal dimensions $\boldsymbol{f}^{\delta}$. For an arbitrary size of alphabet M and $M_L = 4$, this projection is depicted in Figure 11. There are M points: the 4 coplanar values and $M - 4$ points at coordinates $(0, 0)$. These $M - 4$ points correspond to all the vectors that are orthogonal to the considered subset δ. This representation also gives an indication of the distances between the various codewords. With a normalized constellation, the

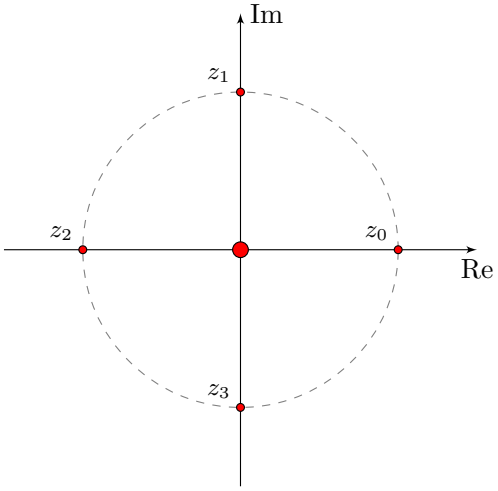


Figure 11: Projection on the δ-th orthogonal dimension with $M_L = 4$.

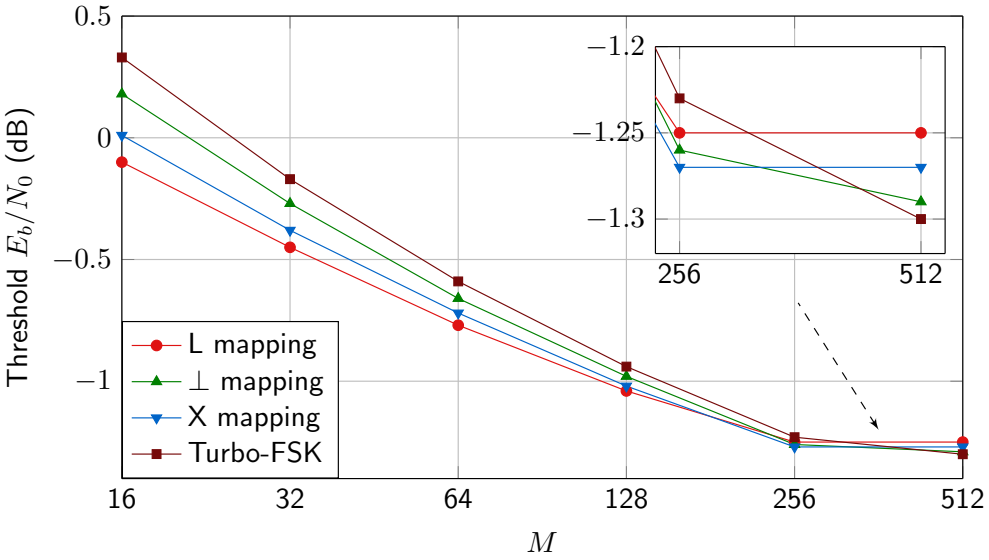


Figure 12: Thresholds for various size of alphabet M and mappings, with $M_L = 4$ (QPSK) and $\lambda = 3$.

distance between *e.g* z_0 and $(0, 0)$ is 1. But the distance between z_0 and z_1 is equal to $\sqrt{2} \simeq 1.41$ and the distance between z_0 and z_2 is equal to 2. Since the L mapping with $M_L = 4$ associates each transition to one of the four possible z_p, transitions are further apart than when simple orthogonality is considered.

With these considerations on the 4-PSK modulation, and its potential gain versus the Turbo-

FSK scheme, we wish to study this size of linear modulation versus various values of M_O. The parameter $\lambda = 3$ is selected, and alphabet sizes spanning from $M = 16$ to 512 are chosen (equivalent to sizes of FSK spanning from 4 to 128). The same 4 mappings are considered, and the estimated thresholds are depicted in Figure 12. Up to an alphabet size equal to $M = 128$, the L mapping outperforms the other mappings, with a gain of 0.42dB versus the Turbo-FSK for $M = 16$. For $M = 256$ (i.e. 64-FSK), the C-TFSK (regardless of the mapping) has a lower threshold than the Turbo-FSK, but the difference is much less significant. For $M = 512$, the thresholds of the C-TFSK are no longer lower than the Turbo-FSK. The impact of having larger distances due to the choice of linear modulation is reduced when large numbers of orthogonal subsets are considered. Considering the L mapping, the distance between the transitions is the same when $M = 16$ and when $M = 512$, but each transition has 4 orthogonal branches per transition in the first case and 128 in the second case.

As expected when first considering the use of a linear component, some performance loss is observed for large sizes of M_L. Nonetheless, the impact of the loss of orthogonality can be reduced by selecting an appropriate mapping that maintains some orthogonal properties between the transitions. Also, using the parameter $M_L = 4$ is beneficial for the performance when a relatively small number of orthogonal subsets are considered (up to $M_O = 64$).

In addition to the three parameters that are M, M_L and λ, the selected mapping must be considered for the C-TFSK, and there are $M!$ possible mappings. This makes the scheme arduous to study, and only a small proportion of the parameters have been explored. It is possible for another configuration to exceed the initial Turbo-FSK, but this configuration has not been found yet. For the rest of the study, the X mapping is considered, as it seems to offer the best compromise in terms of performance.

B. Effects of Puncturing

The puncturing procedure included in the design of C-TFSK offers additional flexibility in data rate, which may be required for potential applications. In classical coding schemes, puncturing is done at the bit level. In the case of C-TFSK, puncturing must be applied at binary words level, as they are associated to a non-binary non-systematic waveforms in the modulation process. Overall, the puncturing mechanism reduces the number of transmitted symbols. At the receiver side, entire sections of the trellis are unknown. For these sections, the likelihood of all the codewords from the alphabet is considered equal. While a gain in spectral efficiency is achieved, the receiver has

R_p	R_p^0	R_p^1	R_p^2	R_p^3	G	η
1	1	1	1	1	N/A	0.093
4/3	1	3/2	3/2	3/2	3	0.124
3/2	1	3/2	3/2	3	3	0.140
8/5	1	2	2	2	2	0.149
12/7	1	3/2	3	3	3	0.160
2	1	3	3	3	3	0.186

Table II: Parameters of the different puncturing patterns tested. R_p is the global puncturing rate and R_p^ℓ, with $\ell \in \{0, \ldots, \lambda - 1\}$, the puncturing rate of stage ℓ. G is the parameter of the block interleaver, and η is the spectral efficiency, expressed in bits·s^{-1}·Hz^{-1}. The other parameters are set to $M = 128$, $M_L = 8$ and $\lambda = 4$.

less information from the channel to work with. Using puncturing usually has a negative impact on the performance.

The case $M = 128$ and $M_L = 8$ is selected. Using the parameter $\lambda = 4$ and an X mapping, 5 different puncturing patterns are considered, for which the parameters are given in Table II. Puncturing rates spanning from $4/3$ to 2 are considered, as well as cases where the puncturing rate of the different branches is equal or varies. The BER performance with $Q = 1000$ is depicted in Figure 13. As expected, increasing the puncturing rate induces a loss in performance, at the benefit of a gain in spectral efficiency (which is given, for each case, in Table II). For the case $R_p = 2$ and $Q = 1000$, only 336 codewords are transmitted instead of 672 for the case without puncturing. The spectral efficiency is multiplied by 2 at the expense of a performance loss of 1.3dB for a BER of 10^{-5}. The cases $R_p = 3/2$ and $R_p = 8/5$ give almost the same performance, even though the puncturing rates are different. For the case $R_p = 8/5$, branches puncturing rates are all equal while the $R_p = 3/2$ uses two different puncturing rates on the branches.

In order to study the influence of the choice of puncturing rate on each branch, we select the case $R_p = 4/3$. Four combinations of branch puncturing rates are selected, and the BER performance is depicted in Figure 14. The combinations are given in the form $[R_p^0\ R_p^1\ R_p^2\ R_p^3]$. $G = 3$ for all cases except the case $[1 \quad 1 \quad 4/3 \quad 4]$ where $G = 4$. For the case $[1 \quad 1 \quad 1 \quad \infty]$, the last branch is not sent. This case corresponds to $\lambda = 3$, as the last branch does not give any information.

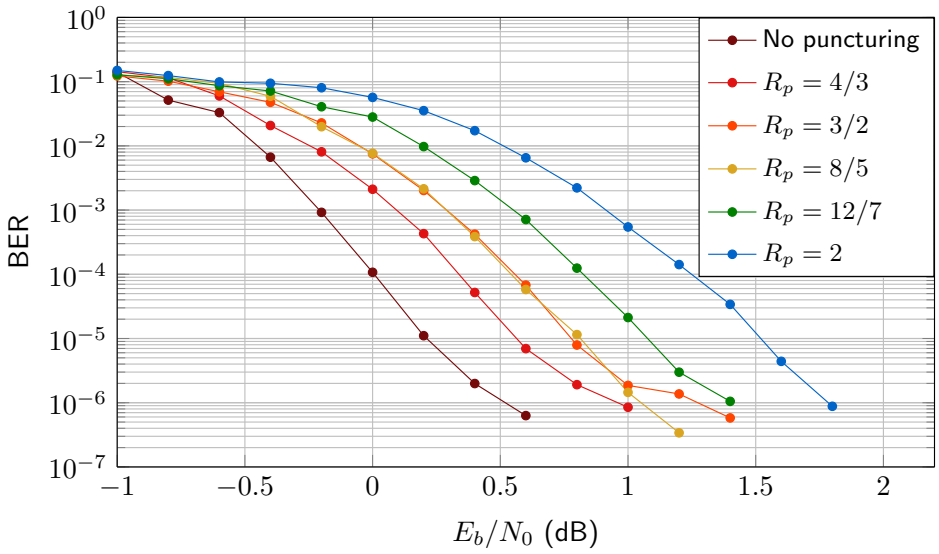


Figure 13: BER performance under the AWGN channel versus E_b/N_0 of C-TFSK with the configuration $M = 128$, $M_L = 8$, $\lambda = 4$, X mapping and various puncturing rates. The spectral efficiency without puncturing is equal to $9.32 \cdot 10^{-2}$, and for punctured cases the spectral efficiencies are given in Table II. The interleaver size is set to $Q = 1000$ and 10 decoder iterations are performed.

The choice of puncturing rate on each branch has a minimal influence on the performance, except for the case where the last branch is not sent. This last case has an increased error floor when compared to the other puncturing schemes. This shows that by considering an additional branch with only a limited number of symbols (for the case $[1 \quad 1 \quad 4/3 \quad 4]$, only 42 are sent for the last branch), the error floor can be reduced for a constant value of spectral efficiency.

In order to illustrate this effect, the Packet Error Rate (PER) performance for the case $\lambda = 3$ without puncturing and the case $\lambda = 4$ with $R_p = 4/3$ is depicted in Figure 15. These two configurations have the same spectral efficiency and are thus comparable in SNR. The PER performance for two other configurations, the case $\lambda = 4$ without puncturing and the case $\lambda = 5$ with $R_p = 5/4$ (with $R_p^\ell = 4/3$ and $G = 4$), are also depicted in the figure. These two configurations also have the same spectral efficiency, equal to $\eta = 9.319 \cdot 10^{-2}$ bits·s^{-1}·Hz^{-1}. For both cases, considering an extra branch $(\lambda + 1)$ only few symbols gives performance close to the case λ without puncturing, but the error floor is significantly reduced. For a constant spectral efficiency, puncturing can thus be used to reduce the error floor. However, complexity

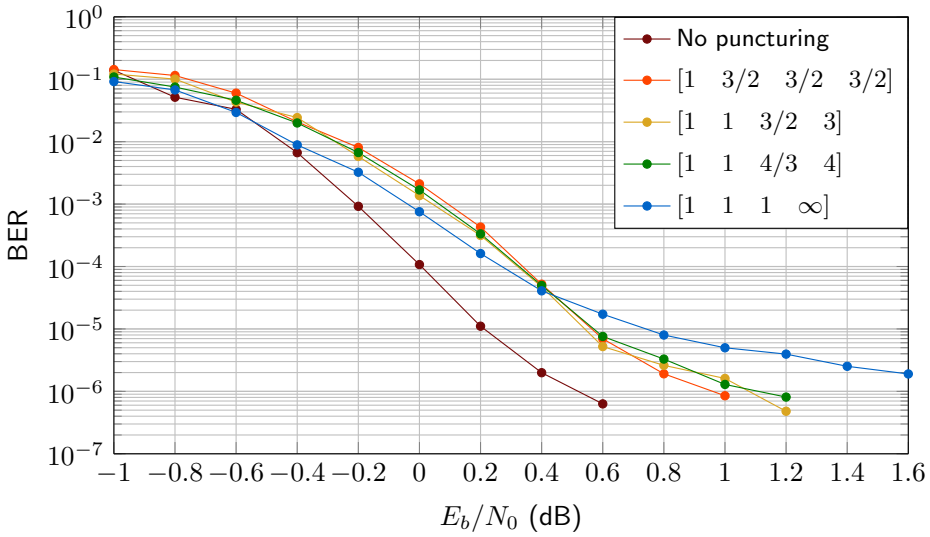


Figure 14: BER performance under the AWGN channel versus E_b/N_0 of C-TFSK with the configuration $M = 128$, $M_L = 8$, $\lambda = 4$, X mapping and a puncturing rate of $R_p = 4/3$. The spectral efficiencies are equal to $9.32 \cdot 10^{-2}$ bits·s^{-1}·Hz^{-1}(without puncturing) and $1.24 \cdot 10^{-1}$ bits·s^{-1}·Hz^{-1}(with puncturing). The interleaver size is set to $Q = 1000$ and 10 decoder iterations are performed.

is increased as when considering a receiver with λ instead of $\lambda - 1$ branches, the execution an additional trellis decoding is implied.

Puncturing is very often used to increase the spectral efficiency or the coding rate. For the C-TFSK scheme, the puncturing procedure introduced has the expected effect. It allows more flexibility and an additional trade-off between performance and spectral efficiency. Moreover, simulations demonstrated that puncturing can be used to reduce the error floor, at the expense of a complexity increase.

C. Influence of the linear component

C-TFSK introduced a linear component with the aim to improve spectral efficiency flexibility. We wish to study the impact of the parameter M_L regarding both the spectral efficiency and the performance of the scheme. During this section, the mapping and puncturing are selected constant, respectively chosen as the X mapping and $R_p = 1$. There are still numerous combinations of M, λ and M_L possible. Two cases are selected: $M = 128$ with $\lambda = 4$ and $M = 512$ with $\lambda = 3$. These combinations correspond to some of the best ones in terms of performance and according

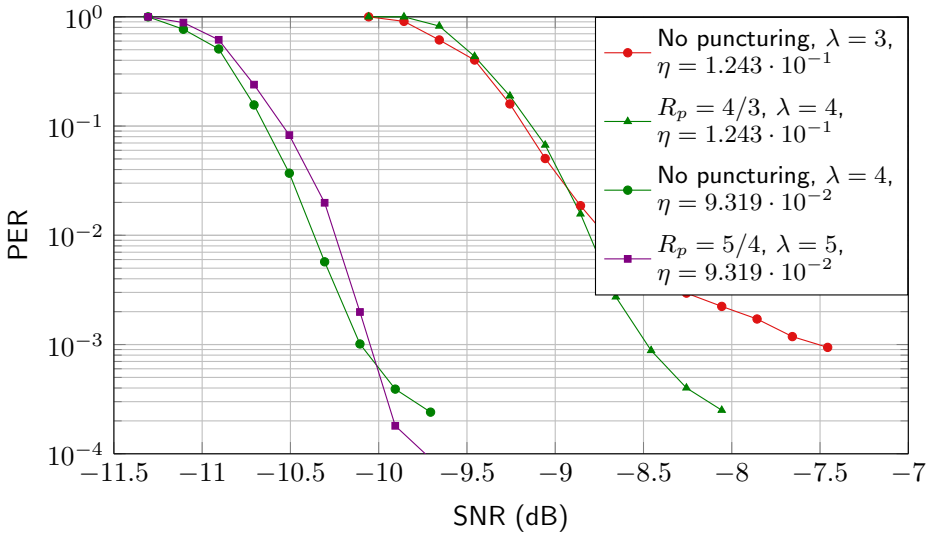


Figure 15: PER performance under the AWGN channel versus the SNR of C-TFSK with the configuration $M = 128$, $M_L = 8$ and X mapping with two different values for λ, with and without puncturing. The interleaver size is set to $Q = 1000$ and 10 decoder iterations are performed.

to [27]. For both cases, there are respectively 7 and 9 possible values for M_L, considering the condition $M_O = M/M_L \geq 2$ (i.e. a minimum of 2 orthogonal subsets). Thresholds for both cases and all possible values of M_L are computed using the EXIT chart. To each configuration corresponds a spectral efficiency, computed with Equation (8). Two elements of comparison are considered for the study: the maximum achievable spectral efficiency according to the channel capacity, computed using (1), and the threshold of the $[13\,15]$ TC associated with Binary Phase Shift Keying (BPSK) modulation. The result is depicted in Figure 16. For each point, the parameter couple $(M_O,\ M_L)$ is labeled. For example, the point labeled $(8, 64)$ for the curve $M = 512$ and $\lambda = 3$ corresponds to $M_L = 64$ with $M_O = 8$ orthogonal subsets.

For both sizes of alphabet, the effect of increasing the size of M_L (i.e. decreasing the number of orthogonal subsets) also increases the spectral efficiency. Up to a certain value of M_L, the performance loss in E_b/N_0 is limited, an effect previously observed. The minimum distance in the trellis is maintained, and the performance depends on the choice of mapping (for all configurations, the X mapping described in Section IV-A is selected). When $M_L \geq 8$, this minimum distance starts to decrease, with a negative effect on the threshold level. As mentioned previously, the Shannon limit (expressed here in terms of maximum achievable spectral

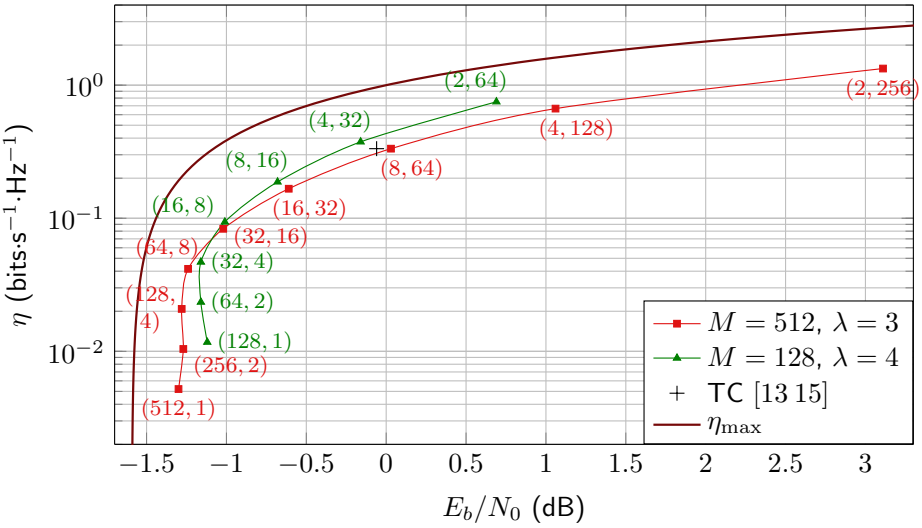


Figure 16: Spectral efficiency versus threshold for various sizes of alphabet M and parameter λ, using various combinations of $(M_O,\ M_L)$. The spectral efficiency is computed with (8). The maximum achievable spectral efficiency according to the channel capacity is computed using (1). A X mapping was selected, without any puncturing (i.e. $R_p = 1$).

efficiency) traduces the required trade-off between energy efficiency and spectral efficiency. The performance observed for C-TFSK also follows this principle, as the energy efficiency decreases when the spectral efficiency increases. However, the performance of the scheme stays efficient as the distance to the maximum achievable spectral efficiency for a given E_b/N_0 remains relatively small.

The $[13\ 15]$ TC, which threshold is depicted in Figure 16, is the channel code used in the NB-IoT approach for LPWA solutions. Associated with BPSK modulation, this scheme has a threshold of -0.06dB for a spectral efficiency of $1/3$. As shown in the figure, the performance of the TC can be closely approached using different configurations. In order to improve the range of the communication, the NB-IoT approach also considers the use of a Spreading Factor (SF). By reducing the spectral efficiency, the required level of SNR for the QoS is reduced. However, the use of SF does not improve the energy efficiency of the scheme, i.e. the level of E_b/N_0 required for the QoS will remain unchanged. Unlike the NB-IoT approach, C-TFSK offers flexibility in configuration. By selecting a configuration with a lower spectral efficiency, both the range and the energy efficiency can be improved, up to more than 1dB.

These results clearly demonstrate how the C-TFSK scheme allows for a gain in spectral

efficiency while having thresholds close to the minimum achievable E_b/N_0. Thanks to the introduction of a linear modulation, spectral efficiencies up to $4/3$ ($M_O = 2$, $M_L = 256$ and $\lambda = 3$) can be achieved. However, in agreement with the information theory's limit (see equation 1), this comes with a significant performance loss in required E_b/N_0 compared to the Turbo-FSK with the same size of alphabet (4.41dB for the same configuration). Spectral efficiency can also be increased by using puncturing (at the expense of more performance loss).

V. Conclusion

Solutions dedicated to LPWA networks always try to offer low levels of sensitivity at the expense of a low data rate. However, having a system with flexibility in data rate or spectral efficiency can lead to a larger number of potential applications. The original design, the Turbo-FSK, is a constant envelope modulation restrained to low levels of spectral efficiencies. In order to improve the flexibility of the system, two new features have been introduced: the OCS modulation, which introduces a linear component in the modulation's alphabet, and a puncturing mechanism. The scheme, so-called C-TFSK, also has a constant envelope and is defined by several parameters. Due to the very large number of possible configurations, strict optimization is cumbersome and seems hardly practicable. Instead, an asymptotic non-exhaustive study has been realized, and demonstrated interesting properties for C-TFSK. While a very large number of possible mappings of the codewords on the trellis are possible, it is more beneficial to maintain some orthogonal properties between the transition of the trellis. Puncturing allows for an increase in the spectral efficiency at the expense of some performance loss. It can also reduce the error floor for a given spectral efficiency, at the expense of an increase in complexity. Finally, the comparison to the maximum achievable spectral efficiency reveals that C-TFSK maintains an energy efficiency close to the limit, despite the fact that increasing the size of the linear component of the alphabet increases the value of the threshold. When compared to the TC used for the NB-IoT solution, C-TFSK is more flexible and can provide a better energy efficiency when low spectral efficiencies are considered.

References

[1] M. R. Palattella, M. Dohler, A. Grieco, G. Rizzo, J. Torsner, T. Engel, and L. Ladid, "Internet of Things in the 5G Era: Enablers, Architecture, and Business Models," *IEEE Journal on Selected Areas in Communications*, vol. 34, no. 3, pp. 510–527, March 2016.

[2] International Telecommunication Union (ITU-T), "Overview of the Internet-of-Things," Recommendation ITU-T Y.2060, June 2012.

[3] U. Raza, P. Kulkarni, and M. Sooriyabandara, "Low Power Wide Area Networks: An Overview," *IEEE Communications Surveys Tutorials*, vol. PP, no. 99, pp. 1–1, 2017.

[4] H. P. Enterprise, "Low Power Wide Area (LPWA) Networks Play an Important Role in Connecting a Range of Devices," *Business white paper*, November 2016.

[5] T. Rebbeck, M. Mackenzie, and N. Afonso, "Low-Powered Wireless Solutions Have the Potential to Increase the M2M Market by Over 3 Billion Connections," Analysys Mason, Sept 2014.

[6] S. Cui, A. J. Goldsmith, and A. Bahai, "Energy-constrained modulation optimization," *IEEE Transactions on Wireless Communications*, vol. 4, no. 5, pp. 2349–2360, Sept 2005.

[7] Y. Roth, J.-B. Doré, L. Ros, and V. Berg, "Turbo-FSK, a Physical Layer for Low-Power Wide-Area Networks: Analysis and Optimization," *Elsevier Comptes Rendus Physique*, vol. 18, no. 2, pp. 178 – 188, 2017.

[8] "SigFox website," http://www.sigfox.com/, accessed: July 16, 2018.

[9] "802.15.4k: Low-Rate Wireless Personal Area Networks (LR-WPANs) Amendment 5: Physical Layer Specifications for Low Energy, Critical Infrastructure Monitoring Networks." IEEE Standard for Local and metropolitan area networks, pp. 1–149, Aug 2013.

[10] J. Costello, D.J. and J. Forney, G.D., "Channel coding: The road to channel capacity," *Proceedings of the IEEE*, vol. 95, no. 6, pp. 1150–1177, June 2007.

[11] C. Shannon, "A mathematical theory of communication," *The Bell System Technical Journal*, vol. 27, no. 3, pp. 379–423, July 1948.

[12] J. Proakis, *Digital Communications 3rd Edition*, ser. Communications and signal processing. McGraw-Hill, 1995.

[13] "LTE Evolved Universal Terrestrial Radio Access (E-UTRA): Physical Channels and Modulation," 3GPP TS 36.211, V13.2.0, Release 13, Aug 2016.

[14] "Whitepaper Narrowband Internet of Things," Rohde & Schwarz, Aug 2016.

[15] R. Ratasuk, B. Vejlgaard, N. Mangalvedhe, and A. Ghosh, "NB-IoT System for M2M communication," in *2016 IEEE Wireless Communications and Networking Conference*, April 2016, pp. 1–5.

[16] Y. P. E. Wang, X. Lin, A. Adhikary, A. Grovlen, Y. Sui, Y. Blankenship, J. Bergman, and H. S. Razaghi, "A Primer on 3GPP Narrowband Internet of Things," *IEEE Communications Magazine*, vol. 55, no. 3, pp. 117–123, March 2017.

[17] C. Berrou, A. Glavieux, and P. Thitimajshima, "Near Shannon limit error-correcting coding and decoding: Turbo-codes," in *IEEE International Conference on Communications (ICC). Geneva.*, vol. 2, May 1993, pp. 1064–1070.

[18] "LTE Evolved Universal Terrestrial Radio Access (E-UTRA): Multiplexing and Channel Coding," 3GPP TS 36.212, V12.6.0, Release 12, pp. 12–15, Oct 2015.

[19] Y. Roth, J.-B. Doré, L. Ros, and V. Berg, "Turbo-FSK: A New Uplink Scheme for Low Power Wide Area Networks," in *2015 IEEE 16th International Workshop on Signal Processing Advances in Wireless Communications (SPAWC)*, Stockholm, Sweden, June 2015, pp. 81–85.

[20] P. Komulainen and K. Pehkonen, "Performance Evaluation of Superorthogonal Turbo Codes in AWGN and Flat Rayleigh Fading Channels," *IEEE Journal on Selected Areas in Communications*, vol. 16, no. 2, pp. 196–205, Feb 1998.

[21] L. Ping, W. Leung, and K. Y. Wu, "Low-rate turbo-Hadamard codes," *IEEE Transactions on Information Theory*, vol. 49, no. 12, pp. 3213–3224, Dec 2003.

[22] Y.-J. Wu and L. Ping, "On the Limiting Performance of Turbo-Hadamard Codes," *IEEE Communications Letters*, vol. 8, no. 7, pp. 449–451, July 2004.

[23] M. Peleg and S. Shamai(Shitz), "Efficient communication over the discrete-time memoryless rayleigh fading channel with turbo coding/decoding," *European Transactions on Telecommunications*, vol. 11, no. 5, pp. 475–485. [Online]. Available: https://onlinelibrary.wiley.com/doi/abs/10.1002/ett.4460110505

[24] Y. Roth, "The Physical Layer for Low Power Wide Area Networks: A Study of Combined Modulation and Coding Associated with an Iterative Receiver," PhD Thesis, Université Grenoble Alpes, Jul. 2017. [Online]. Available: https://hal.archives-ouvertes.fr/tel-01568794

[25] C. Fragouli and R. D. Wesel, "Turbo-encoder design for symbol-interleaved parallel concatenated trellis-coded modulation," *IEEE Transactions on Communications*, vol. 49, no. 3, pp. 425–435, Mar 2001.

[26] Y. Roth, J.-B. Doré, L. Ros, and V. Berg, "A Comparison of Physical Layers for Low Power Wide Area Networks," in *11th EAI International Conference on Cognitive Radio Oriented Wireless Networks (Crowncom)*, Grenoble, France, June 2016.

[27] ——, "EXIT Chart Optimization of Turbo-FSK: Application to Low Power Wide Area Networks," in *9th International Symposium on Turbo Codes & Iterative Information Processing 2016 (ISTC'16)*, Brest, France, Sep. 2016.

[28] S. ten Brink, "Convergence behavior of iteratively decoded parallel concatenated codes," *IEEE Transactions on Communications*, vol. 49, no. 10, pp. 1727–1737, Oct 2001.

[29] L. Ping, W. Leung, and K. Wu, "Low rate turbo-Hadamard codes," in *IEEE International Symposium on Information Theory*, 2001, pp. 211–.

[30] Y. Roth, J.-B. Doré, L. Ros, and V. Berg, "5G Contender Waveforms for Low Power Wide Area Networks in a 4G OFDM Framework," in *2018 25th International Conference on Telecommunications (ICT) (ICT 2018)*, Saint Malo, France, Jun. 2018.

[31] R. Padovani and J. Wolf, "Coded Phase/Frequency Modulation," *IEEE Transactions on Communications*, vol. 34, no. 5, pp. 446–453, May 1986.

[32] R. A. Khalona, G. E. Atkin, and J. L. LoCicero, "On the Performance of a Hybrid Frequency and Phase Shift Keying Modulation Technique," *IEEE Transactions on Communications*, vol. 41, no. 5, pp. 655–659, May 1993.

[33] A. Latif and N. D. Gohar, "Error Rate Performance of Hybrid QAM-FSK in OFDM Systems Exhibiting low PAPR," *Science in China Series F: Information Sciences*, vol. 52, no. 10, pp. 1875–1880, Oct 2009. [Online]. Available: https://doi.org/10.1007/s11432-009-0165-y

[34] P. Loskot, "A Generalized FSK-Based PHY Layer Design for Wireless Sensor Networks," in *7th International Conference on Communications and Networking in China*, Aug 2012, pp. 362–367.

[35] I. Veřtàt, "Hybrid M-FSK/DQPSK modulations for CubeSat picosatellites," *Radioengineering*, vol. 22, no. 1, pp. 389–393, 2013.

[36] S. Hong, M. Sagong, C. Lim, K. Cheun, and S. Cho, "FQAM : A Modulation Scheme for Beyond 4G Cellular Wireless Communication Systems," in *2013 IEEE Globecom Workshops (GC Wkshps)*, Dec 2013, pp. 25–30.

[37] L. Bahl, J. Cocke, F. Jelinek, and J. Raviv, "Optimal decoding of linear codes for minimizing symbol error rate (corresp.)," *IEEE Transactions on Information Theory*, vol. 20, no. 2, pp. 284–287, Mar 1974.

[38] O. F. Acikel and W. E. Ryan, "Punctured turbo-codes for BPSK/QPSK channels," *IEEE Transactions on Communications*, vol. 47, no. 9, pp. 1315–1323, 1999.