Coded LoRa Frame Error Rate Analysis

Orion Afisiadis*, Andreas Burg*, and Alexios Balatsoukas-Stimming†

*Telecommunication Circuits Laboratory, École polytechnique fédérale de Lausanne, Switzerland
†Department of Electrical Engineering, Eindhoven University of Technology, Netherlands

***Abstract*—In this work, we study the coded frame error rate (FER) of LoRa under additive white Gaussian noise (AWGN) and under carrier frequency offset (CFO). To this end, we use existing approximations for the bit error rate (BER) of the LoRa modulation under AWGN and we present a FER analysis that includes the channel coding, interleaving, and Gray mapping of the LoRa physical layer. We also derive the LoRa BER under carrier frequency offset and we present a corresponding FER analysis. We compare the derived frame error rate expressions to Monte Carlo simulations to verify their accuracy.**

I. Introduction

LoRaWAN is a very popular communications protocol for the Internet of things (IoT). Its physical layer (PHY), which is called LoRa, is based on a proprietary spread spectrum modulation scheme that uses chirp modulation as its basis [1]. Along with the chirp modulation, the LoRa PHY chain includes whitening, channel coding, interleaving and Gray mapping [2]. LoRa is able to work in a wide range of operational signal-to-noise ratios (SNRs), due to the support of multiple spreading factors (SF) and code rates. Some of the details of the LoRa PHY have been revealed in patents [3], but also through several reverse-engineering efforts [4]–[7].

There are several works in the literature that derive approximate formulas for the bit error rate (BER) and the symbol error rate (SER) of the LoRa modulation under both additive white Gaussian noise (AWGN) [8], [9] and under Rayleigh fading channels [9]. Moreover, the work of [10] is the first to give a low-complexity SER approximation under AWGN and interference from a chip-aligned LoRa interferer with the same spreading factor (same-SF interference) as the user of interest. This study was extended to the more realistic non-chip-aligned and non-phase-aligned interference scenario in [11], [12]. The work in [12] also contains an approximation of the FER under AWGN and same-SF interference, which can be particularly useful for LoRa network simulators (e.g., [13]–[17]) but also for theoretical analyses (e.g., [18], [19]) instead of hard reception thresholds, which have been experimentally shown to be too simplistic [20]. However, the FER expression in [12] does not take into account the channel coding, interleaving, and Gray mapping present in a LoRa transceiver physical layer chain, which have a great impact on the error rate. The FER expression in [12] also does not take into account any potential PHY impairments, which are very important in the context of low-power radios. Finally, the work in [21] shows and numerically evaluates the bit error rate expressions of coded LoRa for both coherent and non-coherent detection.

Contributions: In this work, we analyze the channel coding and interleaving mechanisms of the LoRa PHY in order to first derive the codeword error rate (CWER) for LoRa. We then use the CWER to derive two low-complexity approximations for the coded FER of a LoRa system under AWGN. Moreover, we derive an approximation for the coded FER of a LoRa system under AWGN and residual carrier frequency offset (CFO), which is an important impairment that can significantly affect the performance of LoRa. Finally, we corroborate the accuracy of our approximations through Monte Carlo simulations.

Notation: We denote the probability density function (PDF) and the cumulative density function (CDF) of the Rayleigh and Rice distributions by $f_{\text{Ra}}(y;\sigma)$, $f_{\text{Ri}}(y;v,\sigma)$ and $F_{\text{Ra}}(y;\sigma)$, $F_{\text{Ri}}(y;v,\sigma)$, respectively, where σ and v are the *scale* and *location* parameters. Moreover, bold lowercase letters (e.g., $\mathbf{a}$) denote vectors and bold uppercase letters (e.g., $\mathbf{A}$) denote matrices. Bold calligraphic letters (e.g., $\boldsymbol{\mathcal{A}}$) denote a vector containing the frequency-domain representation of $\mathbf{a}$, i.e., $\boldsymbol{\mathcal{A}} = \text{DFT}(\mathbf{a})$, while regular calligraphic letters denote sets (e.g., $\mathcal{A}$). The i-th element of a vector $\mathbf{v}$ is denoted by $\mathbf{v}_i$.

II. LoRa PHY System Model

In this section, we provide some background on the LoRa modulation and demodulation procedures, as well as on the structure of a packet used in a LoRa data transmission.

A. LoRa Modulation and Demodulation

LoRa is a spread-spectrum frequency modulation that uses a bandwidth B and $N = 2^{\text{SF}}$ chips per symbol, where SF is called the *spreading factor* and $\text{SF} \in \{7,\dots,12\}$. Each LoRa symbol carries SF bits of information and Gray mapping is used to convert the bits to symbols. A baseband symbol $s \in \mathcal{S}$, where $\mathcal{S} = \{0,\dots,N-1\}$, begins at frequency $(\frac{sB}{N} - \frac{B}{2})$, and its frequency increases by $\frac{B}{N}$ for each chip until the Nyquist frequency $\frac{B}{2}$ is reached, at which point a frequency fold to $-\frac{B}{2}$ occurs. The general discrete-time baseband equivalent description of a LoRa symbol s, for the common case where the sampling frequency f_s is equal to B, is [2], [22][1]

$$x[n] = e^{j2\pi\left(\frac{n^2}{2N} + \left(\frac{s}{N} - \frac{1}{2}\right)n\right)}, \quad n \in \mathcal{S}. \tag{1}$$

[1]The recent work of [22] showed that the continuous-time LoRa chirp in fact occupies a bandwidth that is slightly larger than B, so that a sampling rate $f_s = B$ introduces aliasing. Nevertheless, (1) is a very good approximation of the actual samples and is analytically tractable.

978-1-7281-5089-5/20/$31.00 ©2020 IEEE

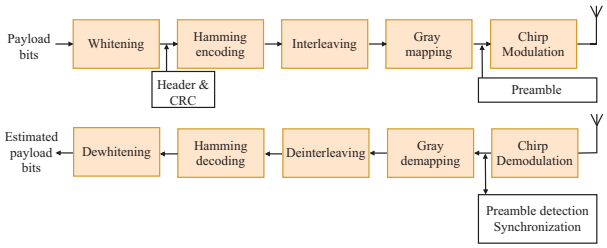

Fig. 1. Illustration of LoRa PHY Tx and Rx chains.

When transmission takes place over an AWGN channel, the received LoRa symbol is given by

$$y[n] = x[n] + z[n], \quad n \in \mathcal{S}, \tag{2}$$

where $z[n] \sim \mathcal{CN}(0, \sigma^2)$ is complex AWGN with variance $\sigma^2 = \frac{N_0}{N}$ and singled-sided noise power spectral density N_0. The SNR is defined as SNR $= \frac{1}{N_0}$. The receiver first *dechirps* the received signal through an element-wise multiplication of $y[n]$ with the complex conjugate of an *upchirp* reference signal $x_{\text{ref}}[n]$ (i.e., the LoRa symbol for $s = 0$). As a next step, the receiver computes the discrete Fourier transform (DFT) of the dechirped signal and obtains $\boldsymbol{\mathcal{Y}} = \text{DFT}\left(\mathbf{y} \odot \mathbf{x}_{\text{ref}}^*\right)$, where $\odot$ denotes the Hadamard product, $\mathbf{y} = \begin{bmatrix} y[0] & \dots & y[N-1] \end{bmatrix}$, and $\mathbf{x}_{\text{ref}}^* = \begin{bmatrix} x_{\text{ref}}^*[0] & \dots & x_{\text{ref}}^*[N-1] \end{bmatrix}$. An estimate $\hat{s}$ of the transmitted LoRa symbol is finally obtained as

$$\hat{s} = \arg\max_{k \in \mathcal{S}} \left(|\boldsymbol{\mathcal{Y}}_k| \right). \tag{3}$$

The SF-bit label of $\hat{s}$, which is defined in the LoRa standard, corresponds to an estimate of the SF transmitted bits.

B. LoRa Transceiver Chain

As shown in Fig. 1 the transmitter chain of LoRa includes a whitening block, a Hamming encoder, an interleaver, and a Gray-mapping block prior to the chirp modulation [7]. After the demodulation the receiver contains a Gray demapper, a deinterleaver, a Hamming decoder, and a dewhitening block.

LoRa uses (k_c, n_c) Hamming codes with $k_c = 4$ and $n_c \in \{5, 6, 7, 8\}$, where k_c denotes the data word length and n_c denotes the codeword length. In this work, we focus on single-bit errors, as these are the only errors that can always be corrected by the $(4, 7)$ and $(4, 8)$ Hamming codes and we assume that, when the $(4, 8)$ Hamming code detects two bit-errors, the decoder declares a failure. Since the $(4, 5)$ and $(4, 6)$ Hamming codes are effectively uncoded with respect to the error rate, the FER analysis of [12] holds. The bits included in a block of SF codewords are interleaved using a diagonal interleaver. The combination of the Hamming code and the interleaving has a great effect on the error rate and is discussed in more detail in Section III. The whitening, on the other hand, has no effect on the error rate and can thus be ignored. In the presence of only AWGN, all symbol errors are equally likely, meaning that the effect of Gray mapping can also be ignored when deriving the FER. However, in the presence of AWGN and CFO the pair-wise error probability for symbol s is non-uniform and adjacent-symbol errors are significantly more likely [2], in which case the Gray mapping plays a very important role and can thus not be ignored.

C. LoRa Packet Structure

A LoRa packet begins with a preamble, which consists of a variable number N_{pr} of upchirps, i.e., N_{pr} consecutive $\mathbf{x}_{\text{ref}}$ symbols. After the preamble, the packet contains two network identifier symbols and $2^1/_4$ frequency synchronization symbols [5]. The packet continues with an optional PHY header, which contains information about the length of the packet, the code rate, the presence of a cyclic redundancy check (CRC), and a checksum. The last part of the packet is the payload, which has a variable length with a maximum of 255 bytes, along with an optional 16-bit CRC of the payload bits [1]. We focus on the FER for the payload part of the packet, which contains the actual transmitted data.

III. LoRa Coded Frame Error Rate Under AWGN

Each LoRa payload consists of multiple interleaved Hamming codewords. Thus, in this section we first derive the codeword error rate (CWER) using existing results for the SER and then we use the CWER to derive the FER.

A. Codeword Error Rate

The uncoded symbol error probability P_s is defined as $P_s \triangleq P(\hat{s} \neq s)$. An approximation that can be used to efficiently evaluate the aforementioned probability under AWGN was derived in [9]. The work in [22] showed that [9] uses a slightly different description for the LoRa symbol than (1), but this difference does not affect the derived error rate approximation of (1). With our definition of the SNR, the symbol error probability can be calculated as

$$P_s \approx Q\left(\frac{\sqrt{\text{SNR}} - \left((H_{N-1})^2 - \frac{\pi^2}{12} \right)^{1/4}}{\sqrt{H_{N-1} - \sqrt{(H_{N-1})^2 - \frac{\pi^2}{12}} + 0.5}} \right), \tag{4}$$

where $Q(\cdot)$ denotes the Q-function and $H_n = \sum_{k=1}^{n} \frac{1}{k}$ denotes the n-th harmonic number. The term $(H_{N-1})^2$ in the above equation stems from the fact that, in the estimation of $\hat{s}$, $N-1$ symbols can be mistakenly chosen instead of s. When a symbol error happens due to AWGN only, on average half of the SF symbol bits will be erroneous, independently of whether Gray mapping is used or not [9]. Thus, for the uncoded bit error probability P_{b} we have

$$P_b = 0.5 \cdot P_s. \tag{5}$$

The codeword error probability P_{cw} is defined as the probability that the Hamming decoder output decision $\hat{\mathbf{c}}$ does not equal the Hamming encoder output codeword $\mathbf{c}$, i.e., $P_{\text{cw}} \triangleq P(\hat{\mathbf{c}} \neq \mathbf{c})$. The Hamming distance between two vectors $\mathbf{w}_1$ and $\mathbf{w}_2$, denoted by $d(\mathbf{w}_1, \mathbf{w}_2)$, is defined as the number of locations where $\mathbf{w}_1$ and $\mathbf{w}_2$ differ. Thus, if $\mathbf{v}$ is the Hamming decoder input vector, then P_{cw} for the $(4, 7)$ and $(4, 8)$ Hamming codes assuming that the decoder declares a failure when more than one error is detected can

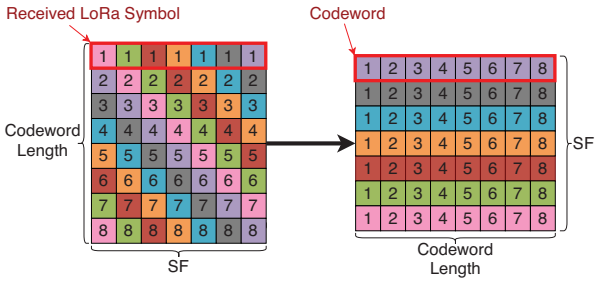

Fig. 2. LoRa deinterleaving for SF $= 7$ and a $(4, 8)$ Hamming code.

be equivalently defined as the probability of the event that $\mathbf{v}$ has *at least* two erroneous bits as

$$P_{\text{cw}} = P(\{d(\mathbf{v}, \mathbf{c}) \geqslant 2\}). \tag{6}$$

Fig. 2 illustrates LoRa deinterleaving process. Each row of the matrix before deinterleaving corresponds to the SF bits of a demodulated LoRa symbol $\hat{\mathbf{s}}_i$, $i \in \{1, \ldots, n_c\}$, while each row of the matrix after deinterleaving corresponds to the n_c bits of $\mathbf{v}_j$, $j \in \{1, \ldots, \text{SF}\}$, at the Hamming decoder input.

When a symbol error happens, e.g., $\hat{\mathbf{s}}_1 \neq \mathbf{s}_1$, on average half of the bits of row 1 of the matrix before deinterleaving will be erroneous. These bit errors are generally correlated because they come from the same symbol error, but the deinterleaver effectively removes this correlation by distributing each bit error of $\hat{\mathbf{s}}_1$ to distinct codewords, as seen in Fig. 2. If another symbol error occurs, e.g., $\hat{\mathbf{s}}_2 \neq \mathbf{s}_2$, the bit errors due to $\hat{\mathbf{s}}_2$ are again generally correlated but they are independent from the bit errors of $\hat{\mathbf{s}}_1$. As a result, the bit errors after the deinterleaver in every input codeword of the Hamming decoder $\mathbf{v}_j$ are independent and identically distributed (iid) with probability of bit error P_b given by (5). Thus, the CWER in (6) becomes

$$P_{\text{cw}} = 1 - \left((1-P_b)^{n_c} + \binom{n_c}{1} P_b (1-P_b)^{n_c-1} \right). \tag{7}$$

B. Frame Error Rate

We denote the number of payload symbols in a LoRa packet by N_{pl}. By construction, N_{pl} is an integer multiple of the codeword length n_c. The number of codewords in the payload is $N_{\text{cw}} = \frac{N_{\text{pl}}\text{SF}}{n_c}$, where each codeword has a CWER P_{cw} given by (7). Let $\mathbf{C}_{\text{bl}} \in \{0,1\}^{\text{SF}\times n_c}$ be a matrix containing a block of SF transmitted codewords $\mathbf{c}_i$, $i \in \{1, \ldots, \text{SF}\}$, that are input to one block of the interleaver. Let also $\mathbf{V}_{\text{bl}} \in \{0,1\}^{\text{SF}\times n_c}$ be a matrix containing a block of SF codewords that are the output of one deinterleaving block (and also the input to the Hamming decoder). Finally, let $\hat{\mathbf{C}}_{\text{bl}} \in \{0,1\}^{\text{SF}\times n_c}$ be a matrix containing a block of SF estimated codewords $\hat{\mathbf{c}}_i$, $i \in \{1, \ldots, \text{SF}\}$, after the Hamming decoder. The probability $P\left(\hat{\mathbf{C}}_{\text{bl}} = \mathbf{C}_{\text{bl}}\right)$ corresponds to the probability that all the SF decoded codewords of one deinterleaver block are correctly decoded, which is given by

$$P\left(\hat{\mathbf{C}}_{\text{bl}} = \mathbf{C}_{\text{bl}}\right) = \prod_{i=1}^{\text{SF}} P\left(\hat{\mathbf{c}}_i = \mathbf{c}_i | \hat{\mathbf{c}}_1 = \mathbf{c}_1, \ldots, \hat{\mathbf{c}}_{i-1} = \mathbf{c}_{i-1}\right). \tag{8}$$

Let $P_{\text{cw}}^{(i)}$ denote the conditional codeword error probability given that all previous $i-1$ codewords in the block were decoded correctly, i.e.,

$$P_{\text{cw}}^{(i)} = P(\hat{\mathbf{c}}_i \neq \mathbf{c}_i | \hat{\mathbf{c}}_1 = \mathbf{c}_1, \ldots, \hat{\mathbf{c}}_{i-1} = \mathbf{c}_{i-1}). \tag{9}$$

Using (9), we can re-write (8) as

$$P\left(\hat{\mathbf{C}}_{\text{bl}} = \mathbf{C}_{\text{bl}}\right) = \prod_{i=1}^{\text{SF}} \left(1 - P_{\text{cw}}^{(i)}\right). \tag{11}$$

Let $\mathbf{C}_{\text{pl}} \in \{0,1\}^{N_{\text{cw}}\times n_c}$ be a matrix containing all the transmitted codewords in the payload and $\hat{\mathbf{C}}_{\text{pl}}$ be a matrix containing all the decoded payload codewords. Since the payload contains N_{pl}/n_c deinterleaving blocks which are independent from each other, using (11) we can obtain the overall FER as

$$P\left(\hat{\mathbf{C}}_{\text{pl}} \neq \mathbf{C}_{\text{pl}}\right) = 1 - \left(\prod_{i=1}^{\text{SF}} \left(1 - P_{\text{cw}}^{(i)}\right) \right)^{\frac{N_{\text{pl}}}{n_c}}. \tag{12}$$

As explained in Section III-A, a correctly decoded codeword $\hat{\mathbf{c}}_i$ may come from a decoder input $\mathbf{v}_i$ that has up to one bit error. This single non-catastrophic bit-error can be in any of the n_c positions of $\mathbf{v}_i$. As a result, the computation of the product in (12), becomes a cumbersome task, since the number of all possible non-catastrophic bit-error patterns inside a deinterleaver block is a large combinatorial quantity. For this reason, in the remainder of this section we derive two low-complexity approximations for (12).

Approximation 1: A simple approximation of (12) can be obtained by ignoring the conditioning in (9) to obtain

$$P\left(\hat{\mathbf{C}}_{\text{pl}} \neq \mathbf{C}_{\text{pl}}\right) \approx 1 - (1 - P_{\text{cw}})^{\frac{N_{\text{pl}}\text{SF}}{n_c}}. \tag{13}$$

Since dependence of decoding errors between codewords of the same block intuitively decreases the probability that the i-th codeword will be erroneous, given that all previous $i-1$ codewords were decoded correctly, we expect that Approximation 1, which assumes independence, overestimates the FER.

Approximation 2: Let us now define a more elaborate approximation for (12), which has slightly higher computational complexity but gives more accurate results. To this end, in the conditional codeword error probability $P_{\text{cw}}^{(i)}$ of (9), we approximate the event that all previous $i-1$ codewords in the block were decoded correctly (i.e., $\{\hat{\mathbf{c}}_1 = \mathbf{c}_1, \ldots, \hat{\mathbf{c}}_{i-1} = \mathbf{c}_{i-1}\}$) with the event that all previous $i-1$ codewords in the block did not contain any error at all (i.e., $\{\hat{\mathbf{v}}_1 = \mathbf{c}_1, \ldots, \hat{\mathbf{v}}_{i-1} = \mathbf{c}_{i-1}\}$). Therefore, for all the previously decoded $i-1$ codewords in the block, we ignore single-bit errors. The conditional codeword error rate in (9) can then be written as

$$P_{\text{cw}}^{(i)} \approx P(\hat{\mathbf{c}}_i \neq \mathbf{c}_i | \mathbf{v}_1 = \mathbf{c}_1, \ldots, \mathbf{v}_{i-1} = \mathbf{c}_{i-1}). \tag{14}$$

Each conditional codeword error probability $P_{\text{cw}}^{(i)}$ in (14) can be interpreted as the codeword error probability of a decoder that has the additional side-information that $(i-1)$ out of the n_c bit positions in $\mathbf{v}_i$ are guaranteed to be error-free. As such, we expect this approximation to result in slightly lower

$$P\left(\hat{\mathbf{C}}_{\text{pl}} \neq \mathbf{C}_{\text{pl}}\right) = 1 - \left(\prod_{i=1}^{\text{SF}} \left(1 - P_b^{(i)}\right)^{n_c} + \binom{n_c}{1} P_b^{(i)} \left(1 - P_b^{(i)}\right)^{n_c - 1}\right)^{\frac{N_{\text{pl}}}{n_c}} \tag{10}$$

FER values compared to (12). However, we need to find an expression to calculate the conditional probability of (14).

Let $P_s^{(i)}$ be the conditional symbol error probability

$$P_s^{(i)} = P(\hat{s} \neq s | \hat{b}_1 = b_1, \ldots, \hat{b}_{i-1} = b_{i-1}), \tag{15}$$

where $\hat{b}_i = b_i$ denotes that the i-th bit of symbol s is estimated correctly. Then the conditional bit error probability is $P_b^{(i)} = 0.5 P_s^{(i)}$, since errors are again due to AWGN. The additional information included in the conditional symbol error probability of (15), compared to the unconditional one, is that a potential demodulation error can only be in $2^{\text{SF}-i-1} - 1$ DFT bins instead of the $N-1$ DFT bins with possible errors in (4). As such, the conditional symbol error probability $P_s^{(i)}$ can be calculated by adapting the harmonic number in (4) as follows

$$P_s^{(i)} \approx Q\left(\frac{\sqrt{\text{SNR}} - \left((H_{\frac{N}{2^{i-1}}-1})^2 - \frac{\pi^2}{12}\right)^{1/4}}{\sqrt{H_{\frac{N}{2^{i-1}}-1} - \sqrt{(H_{\frac{N}{2^{i-1}}-1})^2 - \frac{\pi^2}{12}} + 0.5}}\right). \tag{16}$$

Therefore, each $P_{\text{cw}}^{(i)}$ in (12) can be calculated as

$$P_{\text{cw}}^{(i)} \approx 1 - \left(\left(1 - P_b^{(i)}\right)^{n_c} + \binom{n_c}{1} P_b^{(i)} \left(1 - P_b^{(i)}\right)^{n_c - 1}\right). \tag{17}$$

Finally, our second approximation for the FER can be obtained by replacing (17) in (12), as shown in (10).

IV. LoRa Coded Frame Error Rate Under AWGN and Carrier Frequency Offset

In this section, we first derive an expression for the coded LoRa BER under AWGN and carrier frequency offset (CFO). We then use this initial result to explain how the coded FER can be calculated in the presence of both AWGN and CFO.

A. Distribution of the Decision Metric

In order to derive the coded FER under AWGN in Section III, we have used the symbol and bit error probability expressions from [9]. In order to derive the coded FER under CFO in this section, we first need an expression for the symbol and bit error probabilities of LoRa under residual CFO. Since such expressions do not exist in the literature, in the current subsection we introduce the necessary modeling that will allow us to find the uncoded symbol error probability as well as the coded bit error probability in Section IV-B.

Let f_{c_1} and f_{c_2} be the carrier frequencies that are used during up- and down-conversion, respectively. The carrier frequency offset is the difference $\Delta f_c = f_{c_1} - f_{c_2}$. Thus, in the presence of CFO, the signal model becomes

$$y[n] = c[n]x[n] + z[n], \quad n \in \mathcal{S}, \tag{18}$$

where $x[n]$ is the signal of interest, $c[n] = e^{j2\pi(n+(m-1)N)\frac{\Delta f_c}{f_s}}$ is the CFO term affecting the m-th symbol in the packet, and $z[n] \sim \mathcal{N}(0, \sigma^2)$ is AWGN. The demodulation of $y[n]$ yields

$$\boldsymbol{\mathcal{Y}} = \text{DFT}\left(\mathbf{c} \odot \mathbf{x} \odot \mathbf{x}_{\text{ref}}^*\right) + \text{DFT}\left(\mathbf{z} \odot \mathbf{x}_{\text{ref}}^*\right), \tag{19}$$

where $\mathbf{c} = \begin{bmatrix} c[0] & \ldots & c[N-1] \end{bmatrix}$. We call $\text{DFT}\left(\mathbf{c} \odot \mathbf{x} \odot \mathbf{x}_{\text{ref}}^*\right)$ the *CFO pattern*. The CFO pattern depends on the symbol value s and on the CFO value Δf_c. As explained in [2], [7], the CFO results in a frequency shift which can be modeled with an integer part L and a fractional part λ, where $L + \lambda = \frac{\Delta f_c N}{f_s}$. The fractional part λ is the fraction of the CFO relative to the distance between two adjacent DFT bins. In this work, we consider *residual CFO* after the estimation and correction procedures and we limit our study to the case where $L = 0$ and $-0.5 \leq \lambda \leq 0.5$. If the residual CFO is so large that $L \neq 0$, then all symbols will anyway be erroneously demodulated.

Let $\boldsymbol{\mathcal{R}}_k$ denote the value of the CFO pattern at frequency bin k, i.e., $\boldsymbol{\mathcal{R}}_k = \text{DFT}\left(\mathbf{c} \odot \mathbf{x} \odot \mathbf{x}_{\text{ref}}^*\right)[k], \ k \in \mathcal{S}$. Using the definition of the Fourier transform and arguments that are similar to [12], we have

$$|\boldsymbol{\mathcal{R}}_k| = \frac{\sin\left(\frac{\pi}{N}(s-k+\lambda)N\right)}{\sin\left(\frac{\pi}{N}(s-k+\lambda)\right)}. \tag{21}$$

From (21), it can be seen that, in the absence of CFO (i.e., $\lambda = 0$) we have $|\boldsymbol{\mathcal{R}}_k| = 0$, for any $k \in \mathcal{S}/s$, and $|\boldsymbol{\mathcal{R}}_s| = N$. On the other hand, if $\lambda \neq 0$, a part of the energy of bin s is spread across all N bins. We define $Y_k' = \frac{Y_k}{\sigma}$ for $k \in \mathcal{S}$, which can be used in (3) instead of Y_k. For a specific λ and symbol s, combining the CFO pattern with the AWGN leads to the demodulation metric $|Y_k'|$ used in (3) with [12]

$$|Y_k'| \sim f_{\text{Ri}}\left(y; \frac{|\boldsymbol{\mathcal{R}}_k|}{\sigma}, 1\right), \quad k \in \mathcal{S}. \tag{22}$$

B. Uncoded Symbol Error Rate

A symbol error occurs if and only if any of the $|Y_k'|$ values for $k \in \mathcal{S}/s$ exceeds the value of $|Y_s'|$, i.e.,

$$P\left(\hat{s} \neq s | s, \lambda\right) = P\left(\bigcup_{k \in \mathcal{S}/s} \{|Y_k'| > |Y_s'|\} | s, \lambda\right). \tag{23}$$

As can be deduced from (21), the DFT bins adjacent to bin s have considerably higher energy due to the fractional offset λ compared to the remaining bins, therefore, when CFO and AWGN are present the adjacent bins are generally more prone to causing symbol errors. However, in the case where the SNR is very low, all the bins in the set $\mathcal{S}/s$ still have comparable symbol error probabilities. In order to model the symbol error rate for a large range of SNRs, we rewrite the symbol error probability in (23) in the equivalent form shown in (20), which separates the set $\mathcal{D} = \{s-1, s+1\}$ of the two adjacent bins,

$$P\left(\hat{s}\neq s|s,\lambda\right)=P\left(\bigcup_{i\in\mathcal{D}}\{|Y_i'|>|Y_s'|\}|s,\lambda\right)+P\left(\bigcup_{j\in\mathcal{R}}\{|Y_j'|>|Y_s'|\}|s,\lambda\right)-P\left(\{\bigcup_{i\in\mathcal{D}}\{|Y_i'|>|Y_s'|\}\cap\bigcup_{j\in\mathcal{R}}\{|Y_j'|>|Y_s'|\}\}|s,\lambda\right) \quad (20)$$

from the rest of the bins in the set $\mathcal{R}=\mathcal{S}/\{s-1,s,s+1\}$. The third term of (20) is typically small and can be ignored.

Due to the Gray mapping, a symbol error that mistakes s for one of the symbols in $\mathcal{D}$ only causes one bit error. On the other hand, when a symbol error happens that mistakes s for one of the symbols in $\mathcal{R}$, on average half of the bits are wrong. Therefore, the bit error probability can be written as

$$P\left(\hat{b}\neq b|s,\lambda\right)\approx\frac{1}{\mathrm{SF}}P\left(\{|Y'_{\max,\mathcal{D}}|>|Y_s'|\}|s,\lambda\right)+\frac{1}{2}P\left(\{|Y'_{\max,\mathcal{R}}|>|Y_s'|\}|s,\lambda\right), \quad (24)$$

where $|Y'_{\max,\mathcal{D}}|=\max_{i\in\mathcal{D}}|Y_i'|$ and $|Y'_{\max,\mathcal{R}}|=\max_{j\in\mathcal{R}}|Y_j'|$. The first probability term in (24) is given by

$$P\left(\{|Y'_{\max,\mathcal{D}}|>|Y_s'|\}|s,\lambda\right)=1-\int_{y=0}^{+\infty}f_{\mathrm{Ri}}\left(y;v_s,1\right)F_{|Y'_{\max,\mathcal{D}}|}(y)dy, \quad (25)$$

where $v_s=|R_s|/\sigma$ is the location parameter for bin s. The second term in (24) can be obtained by replacing $\mathcal{D}$ with $\mathcal{R}$. The CDF of the n-th order statistic (i.e., the CDF of the maximum) is known to be $F_n(x)=P(X_1<x)P(X_2<x)\dots P(X_n<x)$. Due to the conditioning on λ, each $|Y_m'|$ for $m\in\mathcal{R}$ is independent from any other $|Y_n'|$ for $n\in\mathcal{D}/m$. Thus, the CDF of the maximum bin for the set $\mathcal{D}$ is $F_{|Y'_{\max,\mathcal{D}}|}(y)=\prod_{j\in\mathcal{D}}F_{\mathrm{Ri}}(y;v_j,1)$, where $v_j=|R_j|/\sigma$. The CDF $F_{|Y'_{\max,\mathcal{R}}|}$ can be obtained accordingly. Finally, s is uniformly distributed in $\mathcal{S}$, so that

$$P_{b|\lambda}=P\left(\hat{b}\neq b|\lambda\right)=\frac{1}{N}\sum_{s=0}^{N-1}P\left(\hat{b}\neq b|s,\lambda\right). \quad (26)$$

C. Complexity Reduction

We can see from (21) that all CFO patterns for all values of $s\in\mathcal{S}$ contain the same set of frequency bin magnitudes $|\boldsymbol{\mathcal{R}}_k|$, $k\in\mathcal{S}$, but are circularly shifted. This circular shift does not change the distribution of $|Y'_{\max}|$, thus the probability of $|Y'_{\max,\mathcal{D}}|>|Y_s'|$ as well as the probability of $|Y'_{\max,\mathcal{R}}|>|Y_s'|$ are not affected. Moreover, the errors in the adjacent bins always give one-bit errors due to the Gray mapping, independently of the circular shift. Therefore, the CFO patterns for all $s\in\mathcal{S}$ result in exactly the same bit error probability $P(\hat{b}\neq b|s,\lambda)$. This means that it is sufficient to compute the CFO pattern for any single symbol s for the evaluation of the bit error probability $P_{b|\lambda}$ given in (26), thus reducing the complexity of evaluating (26) by a factor of N.

D. Frame Error Rate under AWGN and CFO

We follow a similar approach to the AWGN case in order to approximate the coded FER under CFO. First, we approximate the CWER under CFO as

$$P_{\mathrm{cw,cfo}}=1-\left((1-P_{b|\lambda})^{n_c}+\binom{n_c}{1}P_{b|\lambda}(1-P_{b|\lambda})^{n_c-1}\right). \quad (27)$$

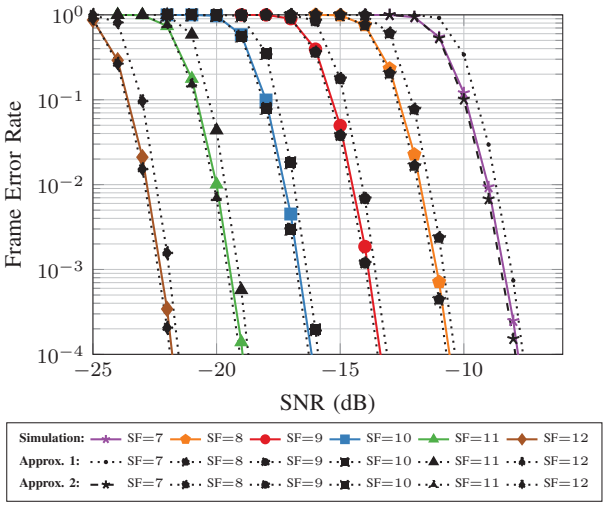

Fig. 3. Frame error rate of the LoRa modulation under AWGN for all supported spreading factors SF $\in\{7,\dots,12\}$ and $N_{\mathrm{pl}}=32$ LoRa symbols.

Following the same reasoning as Approximation 1 of Section III-B, the FER under CFO can then be written as

$$P\left(\hat{\mathbf{C}}_{\mathrm{pl,cfo}}\neq\mathbf{C}_{\mathrm{pl,cfo}}\right)\approx1-\left(1-P_{\mathrm{cw,cfo}}\right)^{\frac{N_{\mathrm{pl}}\mathrm{SF}}{n_c}}. \quad (28)$$

V. Numerical Results

For consistency of the results, we choose a common payload length of $N_{\mathrm{pl}}=32$ LoRa symbols and a common codeword length $n_c=8$, but we note that the derived approximations are also accurate for $n_c=7$, and for any frame length.

In Fig. 3, we first compare the results of a Monte Carlo simulation for the LoRa FER for all possible SF $\in\{7,\dots,12\}$ with the FER results obtained by using Approximation 1 and Approximation 2 under AWGN. We observe that both approximations are quite accurate, especially at low FER values, but Approximation 2 is visibly better than Approximation 1. However, Approximation 1 only requires a single Q-function evaluation per SNR point, while Approximation 2 requires SF Q-function evaluations per SNR point. Moreover, we observe that Approximation 1 slightly overestimates and Approximation 2 slightly underestimates the error rate, as expected from the discussion in Section III.

In Fig. 4, we show the FER for a coded LoRa system with SF $=7$ for three different values of the fractional CFO ($\lambda\in\{0.2,\ 0.3,\ 0.4\}$), as well as the case without any CFO. We observe that the impact of residual CFO in the performance is significant, especially for large λ and at low error rates. Moreover, we can see that our derived approximation is quite accurate for all values of λ.

In Fig. 5, we compare the results of a Monte Carlo simulation for the LoRa FER for all possible SF $\in\{7,\dots,12\}$ with the FER results obtained by using our derived approximation

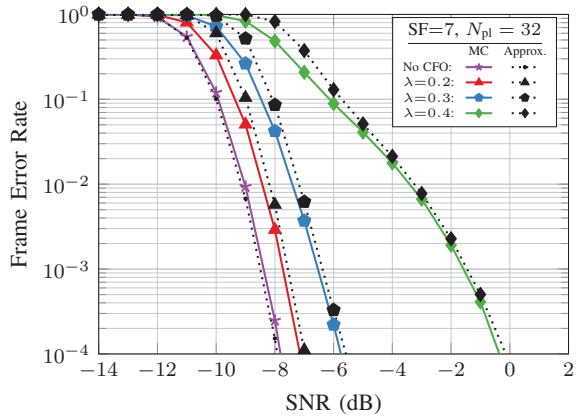

Fig. 4. LoRa FER under three different values of CFO for SF $=$ 7 and $N_{\text{pl}} = 32$ LoRa symbols..

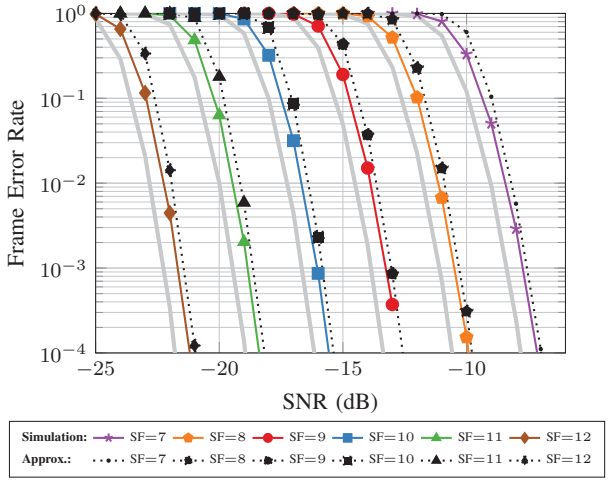

Fig. 5. Frame error rate of the LoRa modulation under AWGN and CFO ($\lambda = 0.2$) for all supported spreading factors SF $\in \{7, \ldots, 12\}$ and $N_{\text{pl}} = 32$. The thick gray lines show the FER when only AWGN is present for comparison.

for $\lambda = 0.2$. The FER under only AWGN is also included in the figure with thick gray lines (taken from Fig. 3). We note that the choice of a common fractional offset λ for all spreading factors corresponds to different Δf_c values for each SF. We observe that, for all spreading factors, a common fractional offset value leads to similar performance degradation and that our derived approximation is accurate.

VI. Conclusion

In this work, we analyzed and quantified the coded frame error rate performance of LoRa under AWGN and CFO. We first derived two low-complexity approximations for the coded LoRa FER under AWGN and we showed that the second approximation is within 0.2 dB of a Monte Carlo simulation for all LoRa spreading factors and FERs down to 10^{-5}. Moreover, we derived the FER under the impact of CFO and we showed that, thanks to the combination of Gray mapping, interleaving, and coding, LoRa is robust to small values of CFO. Finally, we derived a low-complexity approximation for the LoRa FER under CFO which is no more than 0.5 dB away from the corresponding Monte Carlo simulation.

Acknowledgment

The authors would like to thank Joachim Tapparel for useful discussions on implementation details of the LoRa PHY.

References

[1] Semtech Corporation, "SX1272/73–860 MHz to 1020 MHz low power long range transceiver." [Online]. Available: https://www.semtech.com/uploads/documents/sx1272.pdf

[2] R. Ghanaatian, O. Afisiadis, M. Cotting, and A. Burg, "LoRa digital receiver analysis and implementation," in *IEEE Int. Conf. on Acoustics, Speech and Signal Proc. (ICASSP)*, May 2019.

[3] O. B. Seller and N. Sornin, "Low power long range transmitter," US Patent 9,252,834, Feb., 2016.

[4] M. Knight and B. Seeber, "Decoding LoRa: Realizing a modern LPWAN with SDR," in *GNU Radio Conf.*, Sep. 2016.

[5] P. Robyns, P. Quax, W. Lamotte, and W. Thenaers, "A multi-channel software decoder for the LoRa modulation scheme," in *Int. Conf. on Internet of Things, Big Data and Security (IoTBDS)*, Mar. 2018.

[6] A. Marquet, N. Montavont, and G. Z. Papadopoulos, "Investigating theoretical performance and demodulation techniques for LoRa," in *IEEE Int. Symp. on A World of Wireless, Mobile and Multimedia Networks*, June 2019, pp. 1–6.

[7] J. Tapparel, O. Afisiadis, P. Mayoraz, A. Burg, and A. Balatsoukas-Stimming, "An open-source LoRa physical layer prototype on GNU radio," *ArXiv e-prints*, Feb. 2020, https://arxiv.org/abs/2002.08208.

[8] B. Reynders and S. Pollin, "Chirp spread spectrum as a modulation technique for long range communication," in *Symp. Commun. and Vehicular Technologies*, Nov. 2016, pp. 1–5.

[9] T. Elshabrawy and J. Robert, "Closed-form approximation of LoRa modulation BER performance," *IEEE Commun. Lett.*, vol. 22, no. 9, pp. 1778–1781, Sep. 2018.

[10] T. Elshabrawy and J. Robert, "Analysis of BER and coverage performance of LoRa modulation under same spreading factor interference," in *IEEE Int. Symp. on Pers., Indoor, and Mobile Radio Commun.*, Sep. 2018.

[11] O. Afisiadis, M. Cotting, A. Burg, and A. Balatsoukas-Stimming, "LoRa symbol error rate under non-aligned interference," in *Asilomar Conf. on Signals, Systems, and Computers*, Nov. 2019.

[12] O. Afisiadis, M. Cotting, A. Burg, and A. Balatsoukas-Stimming, "On the error rate of the LoRa modulation with interference," *IEEE Trans. Wireless Commun.*, vol. 19, no. 2, pp. 1292–1304, Nov. 2019.

[13] J. Markkula, K. Mikhaylov, and J. Haapola, "Simulating LoRaWAN: On importance of inter spreading factor interference and collision effect," in *IEEE Int. Conf. Commun.*, May 2019, pp. 1–7.

[14] B. Reynders, Q. Wang, and S. Pollin, "A LoRaWAN module for ns-3: Implementation and evaluation," in *Workshop on ns-3*, 2018, pp. 61–68.

[15] A. Pop, U. Raza, P. Kulkarni, and M. Sooriyabandara, "Does bidirectional traffic do more harm than good in LoRaWAN based LPWA networks?" in *IEEE Global Commun. Conf.*, Dec. 2017, pp. 1–6.

[16] F. Van den Abeele, J. Haxhibeqiri, I. Moerman, and J. Hoebeke, "Scalability analysis of large-scale LoRaWAN networks in ns-3," *IEEE Internet of Things Journal*, vol. 4, no. 6, pp. 2186–2198, Dec. 2017.

[17] M. C. Bor, U. Roedig, T. Voigt, and J. M. Alonso, "Do LoRa low-power wide-area networks scale?" in *ACM Int. Conf. on Modeling, Analysis and Simulation of Wireless and Mobile Systems*, 2016, pp. 59–67.

[18] R. B. Sorensen, N. Razmi, J. J. Nielsen, and P. Popovski, "Analysis of LoRaWAN uplink with multiple demodulating paths and capture effect," in *IEEE Int. Conf. Commun.*, May 2019, pp. 1–6.

[19] L. Amichi, M. Kaneko, N. E. Rachkidy, and A. Guitton, "Spreading factor allocation strategy for LoRa networks under imperfect orthogonality," in *IEEE Int. Conf. Commun.*, May 2019, pp. 1–7.

[20] R. Fernandes, R. Oliveira, M. Luís, and S. Sargento, "On the real capacity of LoRa networks: the impact of non-destructive communications," *IEEE Comm. Letters*, Sep. 2019, (Early Access).

[21] G. Baruffa, L. Rugini, V. Mecarelli, L. Germani, and F. Frescura, "Coded LoRa performance in wireless channels," in *IEEE Int. Symp. on Personal, Indoor, and Mobile Radio Commun.*, Sep. 2019, pp. 1–6.

[22] M. Chiani and A. Elzanaty, "On the LoRa modulation for IoT: Waveform properties and spectral analysis," *IEEE Internet of Things J.*, May 2019 (early access).