HAL
open science

Superimposed Frame Synchronization Optimization for Finite Blocklength Regime

Alex The Phuong Nguyen, Raphaël Le Bidan, Frédéric Guilloud

To cite this version:

Alex The Phuong Nguyen, Raphaël Le Bidan, Frédéric Guilloud. Superimposed Frame Synchronization Optimization for Finite Blocklength Regime. WCNCW 2019 : IEEE Wireless Communications and Networking Conference Workshops, Apr 2019, Marrakech, Morocco. 10.1109/WCNCW.2019.8902552 . hal-02076366

HAL Id: hal-02076366

https://imt-atlantique.hal.science/hal-02076366v1

Submitted on 22 Mar 2019

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Superimposed Frame Synchronization Optimization for Finite Blocklength Regime

Alex The Phuong Nguyen
IMT Atlantique, Lab-STICC, UBL
29238 Brest, France
thephuong.nguyen@imt-atlantique.fr

Raphaël Le Bidan
IMT Atlantique, Lab-STICC, UBL
29238 Brest, France
raphael.lebidan@imt-atlantique.fr

Frédéric Guilloud
IMT Atlantique, Lab-STICC, UBL
29238 Brest, France
frederic.guilloud@imt-atlantique.fr

***Abstract*—Considering a short frame length, which is typical in Ultra-Reliable Low-Latency and massive Machine Type Communications, a trade-off exists between improving the performance of frame synchronization (FS) and improving the performance of information throughput. In this paper, we consider the case of continuous transmission over AWGN channels where the synchronization sequence is superimposed to the data symbols, as opposed to being added as a frame header. The advantage of this superposition is that the synchronization length is as long as the frame length. On the other hand, its power has to be traded-off not to degrade the code performance. We first provide the analysis of FS error probability using an approximation of the probability distribution of the overall received signal. Numerical evaluations show the tightness of our analytic results. Then we optimize the fraction of power allocated to the superimposed synchronization sequence in order to maximize the probability of receiving a frame without synchronization errors nor decoding errors. Comparison of the theoretical model predictions to a practical setup show very close optimal power allocation policies.**

***Index Terms*—Frame synchronization, preamble design, short-packet transmission, finite blocklength, mMTC, URLLC, 5G.**

I. Introduction

Short packet transmissions are fundamental bricks to build future applications such as intelligent houses, connected vehicles, automated factories and Internet of Things. The physical layer design rules for short packet transmissions are quite different from the classic asymptotic counterpart where blocklengths are assumed asymptotically long. For example, inserting pilots for channel estimation becomes surprisingly expensive [1], capacity achieving codes like LDPC and turbo code do not perform as well as their lengths are reduced [2] and the widely-used ergodic and outage capacities are overshooting [3]. As a result, new finite blocklength (FBL) design rules are required. Recently, the bounds of maximal channel coding rate for the FBL regime, which are the peers of asymptotic capacity metrics, were derived in [4], [5]. Nevertheless, successful frame synchronization (FS) is also required prior to decoding to achieve the promised throughput. FS can be achieved by blind estimators [6]. However, for complexity issues, the use of sequences of known symbols denoted synchronization words (SW) are preferred. SW can be concatenated to the information symbols by means of a frame header [7]. When the length of frame is fixed, including a header for FS reduces both spectral efficiency and the coding length. Conversely, reducing the FS length will lower the FS performance. Hence there exists a trade-off to be found to optimize the chance of receiving a frame without errors. In the context of low latency communications and/or massive connectivity, the frame length can be reduced while keeping a maximal FS length (i.e. the length of the entire frame) by superimposing SW to data symbols [8]. Then the optimization is related to the power of the FS symbols.

FS is a well investigated topic. FS methods can be roughly divided into two groups according to two distinct transmission modes. The first one is *burst transmissions* for which FS is performed thanks to binary hypothesis testing. Binary hypothesis testing requires a threshold that is usually derived from the Neyman-Pearson lemma. Optimum metrics, in terms of minimizing FS error probability, for binary modulation over AWGN channels are studied in [9], [10] and [11]. The approximation of optimum metric for M-PSK modulation with phase offset error over AWGN channels is provided by [12]. In the finite blocklength regime, a study of FS for burst transmission was recently proposed in [13].

The second mode is *continuous transmissions* where frames are transmitted successively. The knowledge of frame length enables Maximum Likelihood (ML) synchronization: a metric to be maximized is computed for all possible SW positions. The synchronization metric to be used depends on channel models and data distributions. For example, the correlation metric minimizes the FS error probability for binary symmetric channels (BSC) [14], whereas for AWGN channels with binary modulation, the optimal metric requires an additional energy-correction term [15]. As an extension of [15], the authors of [16] present ML metrics for general M-ary phase-coherent and phase-noncoherent AWGN channels. For Rayleigh fading models, the optimum metrics for non-coherent detection is provided in [17]. Regarding the analytic performance (the probability of erroneous frame synchronization), little work has been proposed so far. The work from [7] is one of a few that we are aware of, and provides both metrics and analytic performance results for the particular case of binary modulations over AWGN channels.

In this paper, we provide an analysis of FS error probability for continuous transmissions with superimposed SW over AWGN channels. In the considered setup FS is performed once for the whole transmission rather than repeated on each

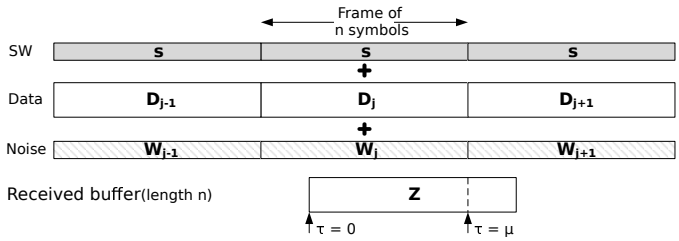

Fig. 1. System model (from the $(j-1)$-th frame to the $(j+1)$-th frame). The frame beginning is at position $\tau = \mu$ in buffer $\mathbf{Z}$.

frame. We will assume hereafter that the coded data symbols follow a spherical uniform distribution. Resorting to Gaussian approximation leads us to a simple expression that can be used to optimize the trade-off between FS overhead and FBL decoding by using the results of [5]. Numerical evaluations show the tightness of our analytic results. Finally, this theoretical model is compared to a practical setup to show that their optimal power allocations are very close. A similar issue has been tackled in [13] but with a different setup. Indeed, the authors of [13] assume *burst transmissions* thus implementing binary hypothesis testing detection. On the contrary, our setup assumes *continuous transmissions* for which the derivation of optimal thresholds is not required, leading to different and complementary calculations and results.

The paper is organized as follows. The system model and the optimization target are described in Section II. The analytic FS error probability is derived in Section III where an approximation of the received data distribution is proposed and justified. Applications to the optimization of the synchronization word power, numerical evaluations and comparisons to a practical setup are reported in Section IV. Finally, conclusions and perspectives can be found in Section V.

Throughout the paper, random variables are denoted by upper case letters whereas their realizations are written in lower case, and boldface letters denote column-wise vectors and matrices. Hence, X and x stand for a random scalar variable and its realization respectively, whereas $\mathbf{X}$ and $\mathbf{x}$ are used to denote a random vector and its deterministic version. Notation $[\mathbf{a};\mathbf{b}]$ denotes the vertical concatenation of two vectors $\mathbf{a}$ and $\mathbf{b}$. Superscripts ^{T} and ^{H} stand for transpose and Hermitian transpose, respectively. $\|.\|$ is used to denote the L_2-norm and $\mathrm{Re}\{x\}$ is the real part of complex x. We adopt conventional notations $\mathbf{0}$ and $\mathbf{I}$ for a zero matrix and an identity matrix of whatever dimension needed to make sense of the expressions being computed. To be more explicit, $\mathbf{0}_n$ and $\mathbf{I}_n$ can be used to clarify the dimensions $n \times n$ of these particular matrices. The Delta function, the Gamma function, the tail distribution function of the standard Normal distribution and the n-order modified Bessel function of the first kind are denoted by $\delta(.)$, $\Gamma(.)$, $\mathcal{Q}(.)$ and $\mathcal{I}_n(.)$, respectively.

II. System model and Target

A. Frame structure

The frame is made of coded data $\mathbf{D} \in \mathbb{C}^n$ superimposed with a fixed known synchronization word (SW) $\mathbf{s} \in \mathbb{C}^n$ where n is the number of symbol (channel-use) per frame. Frames are assumed to be transmitted continuously over an AWGN channel. The receiver is synchronized at the symbol level. By buffering n successive symbols, the frame beginning is surely included in the buffer which is denoted by $\mathbf{Z}$ (see Fig. 1). Note that although each frame includes a SW, the receiver only needs to synchronize once when starting the reception of the data stream. So the analysis hereafter is based on the probability of failure of this initial synchronization step.

The coded data $\mathbf{D}$ is assumed to follow the spherical uniform probability distribution with radius $\sqrt{n\rho}$: codewords $\mathbf{D} \in \mathbb{C}^n$ are uniformly distributed on the surface of a complex sphere and $\|\mathbf{D}\|^2 = n\rho$. This assumption follows from Shannon's achievability proof of the AWGN channel capacity theorem that optimal codes for complex AWGN channels in the asymptotic regime consist of dense packing of signal points within a sphere.

The power of noise $\mathbf{W}$ is normalized to unity and the average energy per symbol of $\mathbf{s}$, $\mathbf{D}$ and of the superposition of the two over the entire frame are denoted by ρ_s, ρ and ρ_{tot} respectively. Hence $\|\mathbf{s}\|^2 = n\rho_s$ with $\rho_{\text{tot}} = \rho_s + \rho$ and we define the *overhead* ratio between the average symbol energy of SW and the average symbol energy of the entire frame as

$$\alpha \triangleq \frac{\rho_s}{\rho_{\text{tot}}}. \tag{1}$$

Intuitively, if α is increased, the frame beginning will be easier to locate but less power for data will lower the coding performance, and vice versa.

As illustrated in Fig. 1, if frame beginning is at position μ of buffer $\mathbf{Z}$, then $\mathbf{Z} = \mathbf{s}_\mu + \mathbf{D}_z + \mathbf{W}_z$ where $\mathbf{s}_\mu$ denotes the right circular shift by μ positions of $\mathbf{s}$, $\mathbf{D}_z$ is the concatenation of a part of frame-(j)-data $\mathbf{D}_j$ and a part of frame-$(j+1)$-data $\mathbf{D}_{j+1}$, $\mathbf{W}_z$ is an i.i.d. Gaussian noise vector. Hence $\mathbf{Z}$ consists of the deterministic $\mathbf{s}_\mu$ and the random part $\mathbf{Y}_z \triangleq \mathbf{D}_z + \mathbf{W}_z$.

The frame synchronization can be modeled as the optimization problem $\hat{\tau} = \underset{0 \le \tau < n}{\operatorname{argmax}}\ f(\mathbf{z}, \tau)$ where τ is the tested position of the frame beginning within the received vector $\mathbf{z}$ and $f(\mathbf{z}, \tau)$ is a *decision metric*. If the SW is at position $0 \le \mu < n$, with $f(\mathbf{Z}, \tau)$ denoting *decision-metric random variables*, the probability P_e of erroneous synchronization can be written as $P_e = \Pr\{\hat{\tau} \neq \mu\} = \Pr\left\{\bigcup_{\tau \neq \mu} f(\mathbf{Z}, \mu) \le f(\mathbf{Z}, \tau)\right\}$ which can be upper-bounded by

$$P_e \le P_{e,u} \triangleq \sum_{\tau \neq \mu} \mathcal{P}_e(\tau), \tag{2}$$

where $\mathcal{P}_e(\tau) \triangleq \Pr\{f(\mathbf{Z}, \mu) \le f(\mathbf{Z}, \tau)\}$. In this paper, we focus on the widely-used correlation rule $f(\mathbf{z}, \tau) = \mathrm{Re}\{\mathbf{s}_\tau^H \mathbf{z}\}$ [18, Chapter 3, Page 69] where the channel phase is assumed to be well estimated and compensated.

B. Optimization target

The objective is to find the optimum value of the energy allocation ratio (the overhead) α for a given fixed total symbol energy ρ_{tot}. We optimize α so as to minimize the frame error rate P_f defined by

$$P_f \triangleq 1-(1-P_e)(1-\epsilon^\star) \tag{3}$$

where P_e is the probability of erroneous synchronization and where $\epsilon^\star$ denotes the minimal channel-coding packet error probability that one can obtain with the best channel code/decoder pair for a finite-length-n packet of k information bits at SNR ρ [4]. Rigorously speaking, P_f is an upper-bound on the true frame error rate since it disregards the possibility that, at least in principle, a frame could be correctly decoded even though it has not been correctly synchronized.

Under the spherical uniform distribution assumption, $\epsilon^\star$ can be approximated by

$$\epsilon^\star(n,k,\rho) \approx \mathcal{Q}\left(\frac{nC(\rho)-k+\frac{1}{2}\log_2(2n)}{\sqrt{nV(\rho)}}\right) \tag{4}$$

where $C(\rho)=\log_2(1+\rho)$ and $V(\rho)=\log_2^2(e)\frac{\rho(\rho+2)}{(\rho+1)^2}$ for the *complex* continuous-input AWGN channel [5]. One could alternatively use any other achievability bounds of [4], e.g. RCU and $\kappa\beta$, to bound $\epsilon^\star$ in (3).

Replacing P_e by its union bound (2) results in the following upper bound $P_{f,u}$ on frame error rate

$$P_{f,u} \triangleq 1-(1-P_{e,u})(1-\epsilon^\star). \tag{5}$$

The optimization target becomes finding the overhead α that minimizes $P_{f,u}$:

$$\hat{\alpha} = \text{argmin}_{0\le\alpha\le 1} P_{f,u}. \tag{6}$$

The evaluation of $P_{f,u}$ is the subject of the next Section.

III. FS ERROR PROBABILITY ANALYSIS

As in (2), the distributions of $\mathbf{Z}$ and $\mathbf{Y}_z$ are required for FS error probability analysis. Since the distribution of $\mathbf{Y}_z$ is quite involved, we propose hereafter an approximation of this distribution which will be used further to obtain an analytic expression of FS error probability.

In Section III-A, we first analyze the distribution of $\mathbf{Y}=\mathbf{D}+\mathbf{W}$ (i.e. assuming $\mu=0$) where $\mathbf{D}\in\mathbb{R}^n$ follows a uniform distribution on the surface of a sphere of radius $\sqrt{n\rho}$ and $\mathbf{W}\sim\mathcal{N}(\mathbf{0},\mathbf{I}_n)$. An approximation of the distribution of $\mathbf{Y}$ that simplifies the FS analysis is given in Section III-B. The choice of the real space is simply for illustrative purpose. Section III-C provides the extension to the complex space that is used to obtain final results on FS error probability in Section III-D.

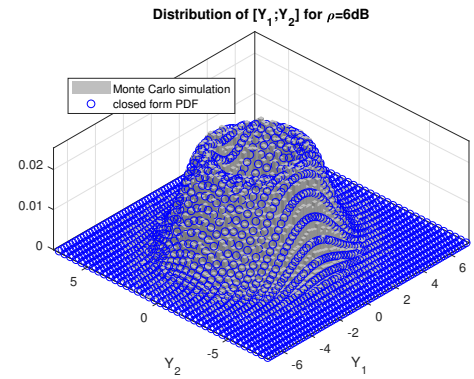

Fig. 2. PDF of $\mathbf{Y}\in\mathbb{R}^2$ at $\rho=6$dB.

A. The probability density function (PDF) of $\mathbf{Y}$

Theorem III.1 (Distribution of $\mathbf{Y}$ for real space)**.** *Let $\mathbf{D}\in\mathbb{R}^n$ follow uniform distribution on the surface of a sphere of radius $\sqrt{n\rho}$ and $\mathbf{W}\sim\mathcal{N}(\mathbf{0},\mathbf{I}_n)$, then $\mathbf{Y}=\mathbf{D}+\mathbf{W}$ has PDF*

$$\begin{aligned} p_{\mathbf{Y}}(\mathbf{y}) =& \frac{\Gamma(n/2)}{2\pi^{n/2}}(n\rho)^{1/2-n/4}\|\mathbf{y}\|^{1-n/2}e^{-(\|\mathbf{y}\|^2+n\rho)/2} \\ &\times \mathcal{I}_{n/2-1}(\|\mathbf{y}\|\sqrt{n\rho}). \end{aligned} \tag{7}$$

Proof. We have $\|\mathbf{Y}\|^2 = n\rho+\|\mathbf{W}\|^2+\mathbf{D}^H\mathbf{W}+\mathbf{W}^H\mathbf{D}$, then $p_{\mathbf{Y}}(\mathbf{y}) = p_{\mathbf{D},\mathbf{W}}(\mathbf{d},\mathbf{w})$. Because of radial symmetry, the PDF of $\|\mathbf{Y}\|^2$ is independent of the rotation applied on $\mathbf{D}$. Indeed, by applying any unitary matrix $\mathbf{v}\in\mathbb{R}^{n\times n}$ on the data $\mathbf{D}$, the corresponding received random vector has the following square norm: $\|\mathbf{vD}+\mathbf{W}\|^2 = n\rho+\|\mathbf{W}\|^2+\mathbf{D}^T\mathbf{v}^T\mathbf{W}+\mathbf{W}^T\mathbf{vD}$. Then $p_{\mathbf{D},\mathbf{W}}(\mathbf{vd},\mathbf{w}) = p_{\mathbf{D},\mathbf{W}}(\mathbf{d},\mathbf{v}^T\mathbf{w}) = p_{\mathbf{D},\mathbf{W}}(\mathbf{d},\mathbf{w})$. As a consequence, we can choose $\mathbf{d}=\mathbf{d}_0 \triangleq [\sqrt{n\rho};0;...;0]$ to show that $\|\mathbf{Y}\|^2$ is distributed according to the non-central χ^2 PDF with n degrees of freedom and a non-centrality parameter of $n\rho$, $\|\mathbf{Y}\|^2\sim\chi_n^2(n\rho)$. Hence $\|\mathbf{Y}\|$ follows a non-central Chi distribution $\chi_n(\sqrt{n\rho})$.

Again, the radial symmetry implies that $\mathbf{Y}$ is statistically invariant against rotation. Thus for a given $\|\mathbf{Y}\|=r$, vector $\mathbf{Y}$ is uniformly distributed on the n-dimensional hyper-sphere of surface $\mathcal{S}(r)=\frac{2\pi^{n/2}}{\Gamma(n/2)}r^{n-1}$.

By a change of variables from Cartesian coordinates $\mathbf{y}$ to spherical coordinates $[r;\phi_1;...;\phi_{n-1}]$ for which the determinant of the Jacobian matrix is $r^{n-1}\prod_{j=1}^{n-2}\sin^{n-j-1}(\phi_j)$, the conditional PDF $p_{\mathbf{Y}|R=r}(\mathbf{y}|r) = \frac{1}{\mathcal{S}(r)}\delta(\|\mathbf{y}\|=r)$ is obtained. Finally, the PDF of $\mathbf{Y}$ is obtained through $p_{\mathbf{Y}}(\mathbf{y}) = \int_r p_{\mathbf{Y}|R=r}(\mathbf{y}|r)p_R(r)\mathrm{d}r = \frac{1}{\mathcal{S}(\|\mathbf{y}\|)}p_{\|\mathbf{Y}\|}(\|\mathbf{y}\|)$. ■

The result of Theorem III.1 is illustrated in Fig. 2.

B. Gaussian approximation of the PDF of $\mathbf{Y}$

Although the PDF from Theorem III.1 is numerically computable, it is still very cumbersome. We propose hereafter a Gaussian approximation of $p_{\mathbf{Y}}(\mathbf{y})$.

Theorem III.2 (Gaussian approximation of $\mathbf{Y}$ for real space)**.** *Let $\mathbf{D} \in \mathbb{R}^n$ follow a uniform distribution on the surface of a sphere of radius $\sqrt{n\rho}$ and $\mathbf{W} \sim \mathcal{N}(\mathbf{0}, \mathbf{I}_n)$. Then among all the centered Gaussian distribution, the one with covariance matrix $\Sigma = (1+\rho)\mathbf{I}_n$ minimizes its Kullback-Leibler divergence to the distribution of $\mathbf{Y} = \mathbf{D} + \mathbf{W}$.*

Proof. Let $\mathbf{Q} \sim \mathcal{N}(\mathbf{0}, \Sigma)$ denote a zero-mean Gaussian random vector with a $n \times n$ covariance matrix Σ. The PDF of $\mathbf{Q}$ is given by

$$p_{\mathbf{Q}}(\mathbf{q}) = \frac{1}{(2\pi)^{n/2}\sqrt{\det(\Sigma)}} e^{-\frac{1}{2}\mathbf{q}^T\Sigma^{-1}\mathbf{q}}.$$

The covariance matrix Σ can be obtained by minimizing the Kullback-Leibler (KL) divergence between the distributions represented by $p_{\mathbf{Q}}$ and $p_{\mathbf{Y}}$:

$$\begin{aligned} D(p_{\mathbf{Y}}||p_{\mathbf{Q}}) &= \mathbb{E}\left[\log\left(\frac{p_{\mathbf{Y}}(\mathbf{Y})}{p_{\mathbf{Q}}(\mathbf{Y})}\right)\right] = \log\Gamma(n/2) \\ &+ \left(\frac{n}{2}-1\right)\log 2 + \left(\frac{1}{2}-\frac{n}{4}\right)\log(n\rho) - \frac{1}{2}n\rho \\ &+ \left(1-\frac{n}{2}\right)\mathbb{E}\left[\log\|\mathbf{Y}\|\right] + \mathbb{E}\left[\log\mathcal{I}_{n/2-1}(\|\mathbf{Y}\|\sqrt{n\rho})\right] \\ &+ \frac{1}{2}\log\det(\Sigma) + \frac{1}{2}\mathbb{E}\left[\mathbf{Y}^T(\Sigma^{-1}-\mathbf{I}_n)\mathbf{Y}\right]. \end{aligned} \tag{8}$$

Let Σ^* denote the optimal covariance matrix given by the minimization of (8), $\Sigma^* \triangleq \underset{\Sigma}{\operatorname{argmin}}\, D(p_{\mathbf{Y}}||p_{\mathbf{Q}}) = \underset{\Sigma}{\operatorname{argmin}}\, \log\det\Sigma + \mathbb{E}\left[\mathbf{Y}^T\Sigma^{-1}\mathbf{Y}\right]$. Taking the derivative of $D(p_{\mathbf{Y}}||p_{\mathbf{Q}})$ with respect to $\tilde{\Sigma} \triangleq \Sigma^{-1}$ yields:

$$\begin{aligned} \frac{\partial D(p_{\mathbf{Y}}||p_{\mathbf{Q}})}{\partial\tilde{\Sigma}} &= \frac{\partial}{\partial\tilde{\Sigma}}\left(\log\det\Sigma + \mathbb{E}\left[\mathbf{Y}^T\Sigma^{-1}\mathbf{Y}\right]\right) \\ &= \frac{\partial}{\partial\tilde{\Sigma}}\left(-\log\det\tilde{\Sigma} + \mathbb{E}\left[\mathbf{Y}^T\tilde{\Sigma}\mathbf{Y}\right]\right). \end{aligned} \tag{9}$$

Thanks to the Jacobi formula of matrix calculus which expresses the derivative of the determinant we obtain $\frac{\partial}{\partial\tilde{\Sigma}}\log\det\tilde{\Sigma} = \Sigma^T$. Also, $\frac{\partial}{\partial\tilde{\Sigma}}\mathbb{E}\left[\mathbf{Y}^T\tilde{\Sigma}\mathbf{Y}\right] = \mathbb{E}\left[\mathbf{Y}\mathbf{Y}^T\right]$. Hence the solution of $\frac{\partial D(p_{\mathbf{Y}}||p_{\mathbf{Q}})}{\partial\tilde{\Sigma}} = 0$ is given by:

$$\Sigma^* = \mathbb{E}\left[\mathbf{Y}\mathbf{Y}^T\right]. \tag{10}$$

We recall that $\mathbf{Y} = \mathbf{D} + \mathbf{W}$ where the Cartesian coordinates of Gaussian noise $\mathbf{W}$ are identically distributed uncorrelated and have zero mean by definition. The uniform spherical distribution of $\mathbf{D}$ implies the same properties because two n-dimensional random vectors $[D_1; ...; +D_l; ...; D_n]$ and $[D_1; ...; -D_l; ...; D_n]$ follow the same distributions hence they must have the same mean, i.e. $\mathbb{E}[D_l] = \mathbb{E}[-D_l]$, and the same covariance, i.e. $\mathbb{E}[D_l D_m] = \mathbb{E}[-D_l D_m]$, $\forall m \neq l$. Therefore the Cartesian coordinates of $\mathbf{D}$ are also zero-mean identically distributed uncorrelated random variables. As a consequence, the Cartesian coordinates of $\mathbf{Y}$ share the same properties and (10) becomes $\Sigma^* = \mathbb{E}\left[\mathbf{Y}\mathbf{Y}^T\right] = \sigma^2\mathbf{I}_n$ where σ^2 denotes the variance of Cartesian coordinates of $\mathbf{Y}$.

The value of σ^2 is derived as follows. By using the non-central law $\chi^2_n(n\rho)$ of $\|\mathbf{Y}\|^2$ (see proof of Theorem III.1), $\mathbb{E}\left[\|\mathbf{Y}\|^2\right] = n + n\rho$. On the other hand, since the Cartesian coordinates of $\mathbf{Y}$ are identically distributed, we have $\mathbb{E}\left[\|\mathbf{Y}\|^2\right] = n\sigma^2$. Hence $\sigma^2 = 1+\rho$. Finally, we conclude that among centered Normal distributions $\mathcal{N}(\mathbf{0}, \Sigma)$, the one minimizing the KL divergence is the i.i.d. Normal distribution $\mathcal{N}(\mathbf{0}, (1+\rho)\mathbf{I}_n)$. ■

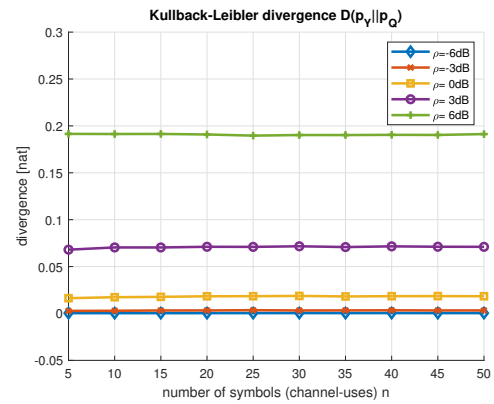

Fig. 3. KL divergences $D(p_{\mathbf{Y}}||p_{\mathbf{Q}})$ for several SNR ρ as a function of the number of symbols/channel-uses.

The KL divergence between the distribution of $\mathbf{Y}$ and its Gaussian approximation is illustrated in Fig. 3 for several SNR ρ, showing that it is almost independent of n. This means that the approximation is good even for small n. Also, it is observed that the lower the SNR, the better the approximation. Nevertheless, for a wide range of SNR, the divergence remains small enough to enable a reliable approximation of the theoretical probability of synchronization error as detailed hereafter.

C. Extension to Complex Signals

To analyze the FS error probability of the system model in Section II, we provide hereafter the complex space versions of Theorems III.1 and III.2.

Theorem III.3 (Distribution of $\mathbf{Y}$ for complex space and its Gaussian approximation)**.** *Let $\mathbf{D} \in \mathbb{C}^n$ follow uniform distribution on the surface of sphere having radius $\sqrt{n\rho}$ and $\mathbf{W} \sim \mathcal{CN}(\mathbf{0}, \mathbf{I}_n)$, then $\mathbf{Y} = \mathbf{D} + \mathbf{W}$ has PDF*

$$\begin{aligned} p_{\mathbf{Y}}(\mathbf{y}) =& \Gamma(n)\pi^{-n}(n\rho)^{(1-n)/2}\|\mathbf{y}\|^{1-n}e^{-(\|\mathbf{y}\|^2+n\rho)} \\ &\times \mathcal{I}_{n-1}(2\|\mathbf{y}\|\sqrt{n\rho}). \end{aligned} \tag{11}$$

Furthermore, among circular Normal distributions $Q = \mathcal{CN}(0, \Sigma)$, the one minimizing the KL divergence $D(p_{\mathbf{Y}}||p_{\mathbf{Q}})$ is $\mathcal{CN}(\mathbf{0}, (1+\rho)\mathbf{I}_n)$.

Proof. The proof is similar to the one of Theorem III.1 with slight differences in power normalization, and with the surface of *complex* hyper-sphere having radius r given by $\frac{2\pi^n}{\Gamma(n)}r^{2n-1}$. The optimal Gaussian approximation is obtained by the same reasoning than in Section III-B. ■

D. FS error probability

One advantage of the proposed distribution approximation is that by approximating both $(\mathbf{D}_j+\mathbf{W}_j)$ and $(\mathbf{D}_{j+1}+\mathbf{W}_{j+1})$ by $\mathcal{CN}(\mathbf{0},(1+\rho)\mathbf{I}_n)$, $\mathbf{Y}_z=(\mathbf{D}_z+\mathbf{W}_z)$ can also be approximated by $\mathcal{CN}(\mathbf{0},(1+\rho)\mathbf{I}_n)$.

If the frame starts at position μ of buffer $\mathbf{Z}$, the correlation metric yields $f(\mathbf{Z},\mu)=\|\mathbf{s}\|^2+\mathrm{Re}\{\mathbf{s}_\mu^H(\mathbf{D}_z+\mathbf{W}_z)\}$ and $f(\mathbf{Z},\tau)=\mathrm{Re}\{\mathbf{s}_\tau^H\mathbf{s}_\mu\}+\mathrm{Re}\{\mathbf{s}_\tau^H(\mathbf{D}_z+\mathbf{W}_z)\}$. Because of the Gaussian approximation of $\mathbf{Y}_z$, we obtain that random variable $(f(\mathbf{Z},\mu)-f_C(\mathbf{Z},\tau))$ can be well approximated by $\mathcal{CN}(\|\mathbf{s}\|^2-\mathrm{Re}\{\mathbf{s}_\tau^H\mathbf{s}_\mu\},\frac{1+\rho}{2}\|\mathbf{s}_\mu-\mathbf{s}_\tau\|^2)$. Finally,

$$\mathcal{P}_e(\tau)=\Pr\{f(\mathbf{Z},\mu)\leq f(\mathbf{Z},\tau)\}\approx\mathcal{Q}\left(\frac{\|\mathbf{s}\|^2-\mathrm{Re}\{\mathbf{s}_\tau^H\mathbf{s}_\mu\}}{\sqrt{\frac{1+\rho}{2}\|\mathbf{s}_\mu-\mathbf{s}_\tau\|^2}}\right). \tag{12}$$

For a given frame length n and FS power allocation α, the upper bound $P_{f,u}$ on frame error rate can be calculated by combining (12) and (4) into (5):

$$\begin{aligned}P_{f,u}\approx 1-&\left(1-\sum_{\tau\neq\mu}\mathcal{Q}\left(\frac{\|\mathbf{s}\|^2-\mathrm{Re}\{\mathbf{s}_\tau^H\mathbf{s}_\mu\}}{\sqrt{\frac{1+\rho}{2}\|\mathbf{s}_\mu-\mathbf{s}_\tau\|^2}}\right)\right)\\ &\times\left(1-\mathcal{Q}\left(\frac{nC(\rho)-k+\frac{1}{2}\log_2(2n)}{\sqrt{nV(\rho)}}\right)\right).\end{aligned} \tag{13}$$

Note that in (13), the overhead α is not explicit because $P_{f,u}$ depends on the specific structure of the SW $\mathbf{s}$. We can bring out α by using particular sequences that have ideal auto-correlation property $\mathrm{Re}\{\mathbf{s}_\tau^H\mathbf{s}_\mu\}=\delta(\tau-\mu)$ (e.g. Zadoff-Chu sequences [19], [20]). Then (12) simplifies into $\mathcal{P}_e(\tau)\approx\mathcal{Q}\left(\frac{\|\mathbf{s}\|}{\sqrt{1+\rho}}\right)=\mathcal{Q}\left(\sqrt{n\frac{\rho_s}{1+\rho}}\right)$ and (13) becomes

$$\begin{aligned}P_{f,u}\approx 1-&\left(1-(n-1)\mathcal{Q}\left(\sqrt{n\frac{\alpha\rho_{\text{tot}}}{1+(1-\alpha)\rho_{\text{tot}}}}\right)\right)\\ &\times\left(1-\mathcal{Q}\left(\frac{nC((1-\alpha)\rho_{\text{tot}})-k+\frac{1}{2}\log_2(2n)}{\sqrt{nV((1-\alpha)\rho_{\text{tot}})}}\right)\right).\end{aligned} \tag{14}$$

IV. Numerical evaluation

The upper bound on the probability of erroneous synchronization (2) is evaluated based on (12) and compared to Monte Carlo simulations. Zadoff-Chu sequences of root 1 are used as SW $\mathbf{s}$. The comparison is illustrated in Fig. 4 for frame with short lengths. Despite the Gaussian assumption and the union bound approximation, the proposed upper bounds are tight compared to their corresponding Monte-Carlo simulations, even at high SNR. It is worth emphasizing that while Monte-Carlo simulations remain manageable at SNRs such that $P_e>10^{-8}$, the proposed theoretical approximation allows much faster evaluations for every SNR.

Thanks to the distribution approximation of the erroneous FS probability, the overhead ratio α that characterizes the fraction of total power allocated to the SW can be optimized. In Fig. 5, the optimization of the frame error rate under both synchronization errors and decoding errors is illustrated for

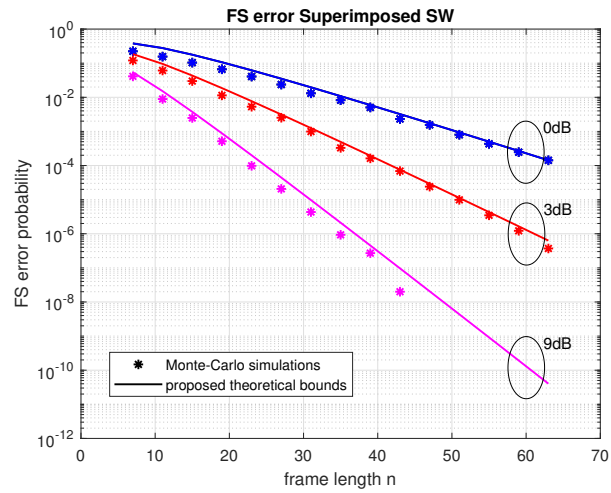

Fig. 4. FS error probability and its approximated union bounds for $\rho_{\text{tot}}=0\text{dB}, 3\text{dB}$ and 9dB for short frames. Equal power allocation for SW and data ($\alpha=0.5$).

several transmitted average symbol energy ρ_{tot} for a short frame of $n=63$ symbols (the frame length is constrained by Zadoff-Chu sequences and should therefore be odd). and $k=32$ information bits. The small mismatches at low power ratio and low SNR ($\rho_{\text{tot}}=0$dB and 1dB) are due to the fact that less power allocated to SW with respect to the noise power results in an increase of the FS error probability for which the union bound becomes useless (i.e. $P_{e,u}$ exceeds 1). However, the optimum overheads obtained by simulation and by the theoretical expression remain coherent.

As illustrated in Fig. 5, above 6dB, the Monte-Carlo simulations become *extremely costly* due to tiny frame error rates around 10^{-10}. This demonstrates the practical interest of this theoretical model as the optimal frame structure can be found very fast (minutes compared to days of Monte-Carlo simulations) by evaluating the proposed theoretical bounds, which are close to the Monte-Carlo simulations.

Finally the performance obtained with a practical setup using QPSK modulation combined with the 3GPP 5G NR Downlink Polar channel code [21] is illustrated in Fig. 6. The channel code decoder is the CRC-Aided Successive Cancellation List Decoder with list size 32. This 3GPP Polar code is powerful for short packets and includes a 24-bit CRC that helps the list-based channel decoder. The performance gap between the proposed theoretical bounds and the practical setup partly comes from the sub-optimal decoder (a larger list is required to reach ML performance), and also from the QPSK modulation which departs from the spherical uniform distribution assumption. However the main purpose of this comparison is to underline that the optimal overheads coincide: the proposed theoretical bounds help finding the optimal overheads very fast thus avoiding time-consuming simulations.

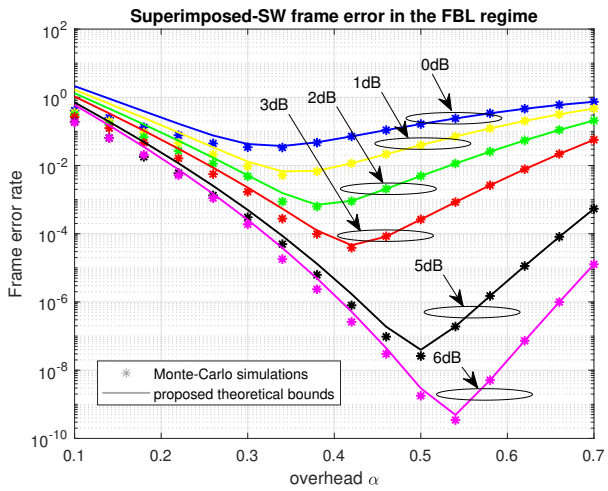

Fig. 5. Trade-off between FS and channel-coding. Zadoff-Chu sequences of root 1 are used as SW. Frame length $n = 63$ and the number of information bits $k = 32$. The total power ρ_{tot} varies from 0dB to 6dB.

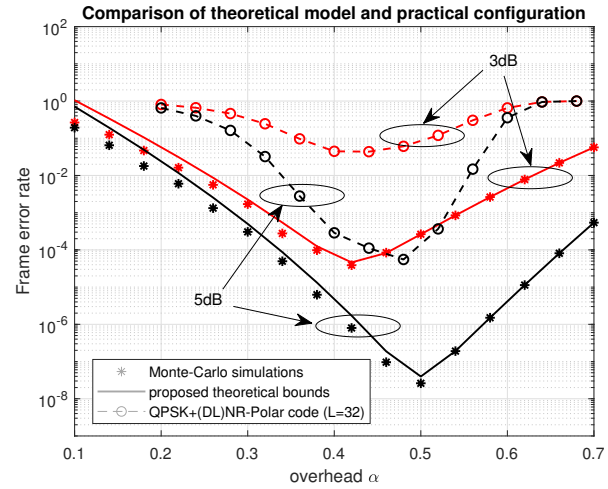

Fig. 6. Optimal frame structure comparison with QPSK and 3GPP 5G NR Downlink Polar code. Zadoff-Chu sequences of root 1 are used as SW. Frame length $n = 63$ symbols and $k = 32$ information bits. The total power $\rho_{\text{tot}} =$ 3dB and 5dB.

V. Conclusion

In this paper, we have developed an upper bound on the frame synchronization error probability in the context of continuous transmissions over AWGN channel. A synchronization word is assumed to be superimposed to the data symbols instead of being embedded in a frame header. The advantage is that the synchronization word length is as long as the frame itself. The proposed error probability is used to optimize the power overhead between the synchronization word and the data symbols for a given total transmitted average symbol energy using recent results on the finite blocklength coding performances. Comparisons with Monte-Carlo simulations of the theoretical scheme show that our results are tight enough to find the optimal power overhead. Furthermore, numerical evaluations show that the optimal overheads obtained with this theoretical model coincide with the practical 3GPP 5G NR Downlink scheme. The comparison between using a superimposed synchronization word and concatenating it to data symbols is part of ongoing work. Further perspectives also include an extension to mobile fading channels.

References

[1] J. Östman, G. Durisi, and E. G. Ström, "Finite-blocklength bounds on the maximum coding rate of Rician fading channels with applications to pilot-assisted transmission," in *2017 IEEE 18th International Workshop on Signal Processing Advances in Wireless Communications (SPAWC)*. IEEE, 2017, pp. 1–5.

[2] "Channel coding schemes for URLLC scenario," R1-1611692 TSG RAN WG1 Meeting 87, 3GPP, 2016.

[3] G. Durisi, T. Koch, J. Östman, Y. Polyanskiy, and W. Yang, "Short-packet communications over multiple-antenna Rayleigh-fading channels," *IEEE Transactions on Communications*, vol. 64, no. 2, pp. 618–629, 2016.

[4] Y. Polyanskiy, H. V. Poor, and S. Verdú, "Channel coding rate in the finite blocklength regime," *IEEE Transactions on Information Theory*, vol. 56, no. 5, pp. 2307–2359, 2010.

[5] T. Erseghe, "Coding in the finite-blocklength regime: Bounds based on laplace integrals and their asymptotic approximations," *IEEE Transactions on Information Theory*, vol. 62, no. 12, pp. 6854–6883, 2016.

[6] R. Imad, G. Sicot, and S. Houcke, "Blind frame synchronization for error correcting codes having a sparse parity check matrix," *IEEE Transactions on Communications*, vol. 57, no. 6, pp. 1574–1577, 2009.

[7] M. Chiani and M. G. Martini, "Analysis of optimum frame synchronization based on periodically embedded sync words," *IEEE Transactions on Communications*, vol. 55, no. 11, pp. 2056–2060, 2007.

[8] Y. Wang, J. Oostveen, A. Filippi, and S. Wesemann, "A novel preamble scheme for packet-based OFDM WLAN," in *Wireless Communications and Networking Conference, 2007. WCNC 2007. IEEE*. IEEE, 2007, pp. 1481–1485.

[9] M. Chiani and M. G. Martini, "Practical frame synchronization for data with unknown distribution on AWGN channels," *IEEE Communications Letters*, vol. 9, no. 5, pp. 456–458, 2005.

[10] ——, "On sequential frame synchronization in AWGN channels," *IEEE Transactions on Communications*, vol. 54, no. 2, pp. 339–348, 2006.

[11] W. Suwansantisuk, M. Chiani, and M. Z. Win, "Frame synchronization for variable-length packets," *IEEE Journal on Selected Areas in Communications*, 2008.

[12] A. Elzanaty, K. Koroleva, S. Gritsutenko, and M. Chiani, "Frame synchronization for M-ary modulation with phase offsets," in *2017 IEEE 17th International Conference on Ubiquitous Wireless Broadband (ICUWB)*. IEEE, 2017, pp. 1–6.

[13] A.-S. Bana, K. F. Trillingsgaard, P. Popovski, and E. de Carvalho, "Short packet structure for Ultra-Reliable Machine-type Communication: Tradeoff between detection and decoding," *International Conference on Acoustics, Speech and Signal Processing (ICASSP) 2018*, 2018.

[14] R. H. Barker, "Group synchronization of binary digital systems," *in Communication Theory*, pp. 273–287, 1953.

[15] J. Massey, "Optimum frame synchronization," *IEEE Transactions on Communications*, vol. 20, no. 2, pp. 115–119, 1972.

[16] G. Lui and H. Tan, "Frame synchronization for Gaussian channels," *IEEE Transactions on Communications*, vol. 35, no. 8, pp. 818–829, 1987.

[17] H. Jia and D. E. Dodds, "Frame synchronization for PSAM in AWGN and Rayleigh fading channels," in *2005. Canadian Conference on Electrical and Computer Engineering*. IEEE, 2005, pp. 44–50.

[18] F. Ling, "Signal Model and ML Detection of Known Symbol Sequence in Received Signal," in *Synchronization in Digital Communication Systems*. Cambridge University Press, 2017, ch. 3, p. 69.

[19] H.-J. Zepernick and A. Finger, *Pseudo random signal processing: theory and application*. John Wiley & Sons, 2013.

[20] *LTE; E-UTRA; Physical channels and modulation*, 3GPP Std. 36.211 v14.6.0, 2018.

[21] *LTE; E-UTRA; 5G; NR; Multiplexing and channel coding Rel.15*, 3GPP Std. 38.212 V15.2.0, 2018.