Is Synchronization a Bottleneck for Pilot-Assisted URLLC Links?

A. Oguz Kislal, Madhavi Rajiv, Giuseppe Durisi, Erik G. Ström, Urbashi Mitra

Abstract

We propose a framework to evaluate the random-coding union bound with parameter s (RCUs) on the error probability achievable in the finite-blocklength regime for a pilot-assisted transmission scheme operating over an imperfectly synchronized and memoryless block-fading channel. Unlike previously reported results, our framework does not assume perfect synchronization. Instead, we take advantage of the pilots for both synchronization and channel estimation. Additionally, we utilize the saddlepoint approximation to provide a numerically efficient method for evaluating the RCUs bound in this scenario. Our numerical experiments show that the saddlepoint approximation accurately represents the RCUs bound over a wide range of parameters. Furthermore, with the suggested synchronization algorithm, the SNR penalty of the imperfect synchronization at the target packet error probability 10^{-5} is observed to be about $0.6\,\mathrm{dB}$.

I. Introduction

The rapid advancement of electronic devices has paved the way for the rise of novel applications like remote surgery [1], [2], factory automation [3] and autonomous driving that significantly impact various aspects of people's lives. However, these applications generate substantial data traffic, posing a formidable challenge for current cellular communication technologies to address. The total mobile data traffic is projected to reach 329 exabytes per month by the end of 2028, a significant increase compared to the 93 exabytes per month recorded in 2022 [4]. Handling the different requirements of these applications is a difficult task and 5G is envisioned to support all these requirements [5]. One interesting scenario, denoted as ultra-reliable low-latency communication (URLLC), is designed for mission-critical applications targeting 99.999% reliability with end-to-end latency as low as $1\,\mathrm{ms}$ [6]. Moreover, 6G is targeting even higher reliability, with the goal of achieving $10^{-5} - 10^{-7}$ packet error rate with a more stringent latency constraint [7], [8].

A key feature of URLLC traffic is the frequent use of small information payloads accompanied by short packets which consist of a limited number of encoded symbols. To understand the need for short packets, it is important to remember that the size of a data packet is determined by the available bandwidth and the signal duration. In URLLC, the duration of the signal is constrained due to the latency requirements imposed by specific applications such as automated factory control and critical internet-of-things services. This limitation on signal duration is crucial to ensure timely and responsive communication. Additionally, the available bandwidth is often limited in URLLC scenarios, since the transmission of multiple users needs to be orthogonalized in order to mitigate multiuser interference. Such interference can have a detrimental effect on the packet error probability and overall system performance. Therefore, the allocation of limited bandwidth resources becomes a critical factor in URLLC communication. The conventional asymptotic performance metrics commonly employed in the design of communication systems, namely the ergodic and outage rates, are not suitable for the short-packet regime [9]. Therefore, there is a need for a more accurate characterization of the tradeoff between transmission rate and error probability in order to address the specific requirements of short-packet communications.

Due to the use of short packets in URLLC, the field of finite-blocklength information theory gained significant attention over the last decade, particularly following the seminal work from Polyanskiy [10]. This theory offers a precise understanding of the tradeoffs in the non-asymptotic upper (achievability) and lower (converse) bounds on the minimum error probability for a given SNR, transmission rate, and packet size. These bounds offer valuable insights into the limits of a communication system's performance when operating with finite blocklengths.

In this paper, we focus on communication over memoryless block-fading channels with imperfect synchronization. We propose a computationally efficient method based on the saddlepoint approximation technique [11], to evaluate the random coding union bound with parameter s (RCUs) proposed in [12]. This bound is particularly well-suited for communication over fading channels due to its ability to provide achievable results applicable to both the optimal noncoherent maximum-likelihood (ML) decoder and more practically significant decoders that utilize pilot-assisted transmission (PAT) [13]. One such example is using the pilots for delay and channel estimation, and then treating these estimates as perfect via the use of scaled nearest-neighbor (SNN) decoder [14]. Approximation methods such as the one proposed in this paper are necessary since evaluating the RCUs bound (and other available information-theoretic error-probability bounds) for scenarios of interest for URLLC can be extremely computationally expensive [15],

[16]. Indeed, due to this complexity, the RCUs bound is not feasible to be used directly within URLLC optimization routines such as resource-allocation [17], [18] and scheduling algorithms [19], [20].

State of the art: In the literature, the most common approach while benchmarking a URLLC system is either assuming perfect channel state information and synchronization at the receiver or assuming perfect synchronization and only dealing with imperfect channel estimation. In both cases, the random variable whose tail probability is of interest can be written as a sum of independent random variables. Then this tail probability can be approximated by a Gaussian tail probability using central-limit theorem. This approximation is typically called the *normal approximation* [10]. Typically, the normal approximation proves to be accurate when dealing with moderate error probabilities and when the transmission rate approaches the channel capacity [21]. Since this is not a viable assumption in URLLC, this approximation does not provide an accurate characterization of the RCUs bound for the error probabilities relevant for the URLLC regime [12], [22].

A much more accurate approximation (e.g., [21], [23], [24]) can be obtained through the *saddlepoint method* which consists of applying normal approximation over a tilted distribution [11], [25]. For the case of the optimal ML decoder, the saddlepoint approximation is studied in [21]. The authors of [24] have provided an analysis using saddlepoint approximation of RCUs bound for PAT transmission, SNN decoding over SISO and MIMO channels. In [13], [22] the saddlepoint approximation for PAT transmission and SNN decoding over massive MIMO is computed when only one fading block is available. This work has been extended for the case of multiple fading blocks in [16].

All of the aforementioned references assume perfect synchronization in their system model which simplifies both normal and saddlepoint approximations. When synchronization is imperfect, the underlying assumption that has been made in these approximations (i.e., received vector having independent symbols per channel use), does not hold anymore. As explained in [11, Ch. 6], it is possible to derive a saddlepoint approximation when the tail probability of interest is written as a sum of dependent random variables, and its usefulness in the general case will be investigated in this work. As we shall see, the imperfect synchronization causes the random variables to be dependent in a structured manner which we take advantage of and derive a saddlepoint approximation for the tail probability of a Markov chain, which has also been studied in [11, Ch. 9].

Contributions: In this paper, we provide a framework that takes imperfect synchronization into account to provide design guidelines for the URLLC regime. To do so, we have used the RCUs bound for the case of PAT and SNN decoding. Both the synchronization and channel estimation have been obtained imperfectly using the pilot symbols. We then show how the imperfect synchronization causes the data symbols in a fading block to be dependent. Noting that this dependence violates one of the fundamental assumptions of the previously reported saddlepoint approximations for RCUs bound, we derive a saddlepoint approximation tailored for this unique case. Our numerical analyses show that the saddlepoint approximation is accurate for a wide range of parameters. Additionally, we observed that with PAT, the impact of the imperfect synchronization diminishes as the number of diversity branches increases.

Notation: We denote random vectors and random scalars by upper-case boldface letters such as $\boldsymbol{X}$ and upper-case standard letters, such as X, respectively. Their realizations are indicated by lower-case letters of the same font. We use upper-case letters of two special fonts to denote deterministic matrices (e.g., Y). To avoid ambiguities, we use another font, such as R for rate, to denote constants that are typically capitalized in the literature. The circularly-symmetric Gaussian distribution is denoted by $\mathcal{CN}(\mu, \sigma^2)$, where μ and σ^2 denote the mean and the variance respectively. The superscripts $(\cdot)^T$ and $(\cdot)^H$ denote transposition and Hermitian transposition respectively. We write $\log(\cdot)$ to denote the natural logarithm, $\|\cdot\|$ stands for the ℓ^2-norm, $\mathbb{P}[\cdot]$ for the probability of an event, $\mathbb{E}[\cdot]$ for the expectation operator, $*$ for the convolution operation, $\mathbb{1}(\cdot)$ for the indicator function, and $Q(\cdot)$ for the Gaussian Q-function.

Organization of the paper: In Section II, we present our system model. In Section III, we present a finite-blocklength achievability bound on the error probability. We then introduce the saddlepoint approximation to efficiently evaluate this bound in the URLLC regime with imperfect synchronization. In Section IV, we provide an example on how saddlepoint approximation can be applied when BPSK constellation is used. In Section V, we discuss the accuracy of the saddlepoint approximation and the impact of the imperfect synchronization with the help of numerical examples. Concluding remarks are provided in Section VI.

II. System Model

A. Overview

We consider pilot-assisted transmission of a uniformly distributed message over a single-input single-output (SISO) block-flat-fading channel with unknown delay. The setup is illustrated by

the block diagram in Fig. 1. The encoder maps the realization of the message $\boldsymbol{W}$ to a complex-valued codeword, which in turn is split into the subcodewords as

$$\mathbf{x}_{\ell}^{(d)} = \begin{bmatrix} x_{0,\ell}^{(d)} & \cdots & x_{n_{\mathrm{s}}-1,\ell}^{(d)} \end{bmatrix}^{\mathrm{T}} \in \mathbb{C}^{n_{\mathrm{s}}}, \qquad \ell = 0, \ldots, n_{\mathrm{b}} - 1 \tag{1}$$

where n_{b} is the number of fading blocks used for transmission of the message. The subcodewords are prepended by the pilot sequence

$$\mathbf{x}^{(p)} = \begin{bmatrix} x_{0}^{(p)} & \cdots & x_{n_{\mathrm{p}}-1}^{(p)} \end{bmatrix}^{\mathrm{T}} \in \mathbb{C}^{n_{\mathrm{p}}} \tag{2}$$

to form n_{b} subpackets which are transmitted over different channel fading blocks. Hence, the ℓth subpacket, which consists of $n_{\mathrm{c}} = n_{\mathrm{p}} + n_{\mathrm{s}}$ pilot and data symbols, sees a flat-fading channel with complex channel gain H_{ℓ}, where $H_0, \ldots, N_{n_{\mathrm{b}}-1}$ are i.i.d.. However, all subpackets experience the same propagation delay D, a fact which will be used in the synchronization algorithm[1]. We assume that the transmitter and receiver are coarsely synchronized and model D as uniformly distributed over $[0, d_{\max}]$, where $d_{\max}$ is a known constant.

We assume that the data and pilot symbols are subject to the same power constraint $\mathbb{E}[\|\boldsymbol{X}_{\ell}^{(d)}\|^2] = n_{\mathrm{s}}\rho$ and $\|\mathbf{x}^{(p)}\|^2 = n_{\mathrm{p}}\rho$. As is clear from (1) and (2), the packet consists of $n_{\mathrm{b}}n_{\mathrm{c}}$ complex pilot and data symbols, and the codebook therefore contains $\lceil \exp(n_{\mathrm{b}}n_{\mathrm{c}}\mathrm{R}) \rceil$ codewords, where R is the rate in nats per complex symbol.

The receiver use knowledge of the pilot sequence to estimate the common channel delay and the individual fading block gains. The receiver picks the codeword in the codebook that, after being delayed and scaled with the delay and channel gain estimates, is at minimum Euclidean distance to received vector. If the synchronization and channel estimation is error-free, this would correspond to ML decoding (which minimizes the packet error probability when the message is uniformly distributed). We finally define the packet error probability as

$$\epsilon_{pep} = \mathbb{P}\left(\boldsymbol{W} \neq \hat{\boldsymbol{W}}\right). \tag{3}$$

where $\hat{\boldsymbol{W}}$ is the decoded message.

In the following subsections, we will define the pulse shaping and receiver processing considered in this paper.

[1]This model covers, e.g., the case when we transmit the subpackets in parallel over sufficiently spaced subcarriers.

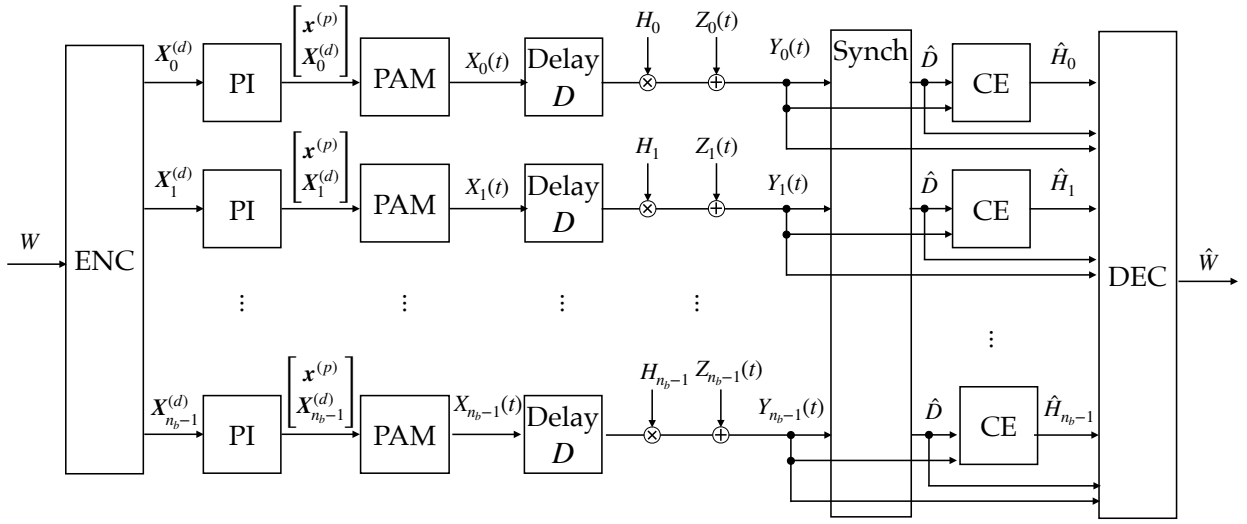


Fig. 1: Block diagram of the system model.

B. Modulation and Waveform Channel

The continuous-time pilot signal, which is the same for all fading blocks, is formed as

$$x^{(p)}(t) = \sum_{k=0}^{n_p-1} x_k^{(p)} s_{t_\mathrm{p}}(t - kt_\mathrm{p}) \tag{4}$$

where $s_\tau(t)$ is a unit-energy square pulse with duration τ defined as

$$s_\tau(t) = \begin{cases} 1/\sqrt{\tau}, & t \in [0, \tau) \\ 0, & \text{otherwise} \end{cases}. \tag{5}$$

The ℓth fading block data signal is

$$x_\ell^{(d)}(t) \;\; = \;\; \sum_{k=0}^{n_s-1} x_{k,\ell}^{(d)} s_{t_\mathrm{p}}(t - kt_\mathrm{p} - n_p t_p) \tag{6}$$

and the ℓth subpacket signal is $x^{(p)}(t) + x_\ell^{(d)}(t)$. Hence, the data symbols follows immediately after the pilot symbols.

The received signal due to the ℓth subpacket is

$$Y_\ell(t) \;\; = \;\; H_\ell\Big(x^{(p)}(t - D) + x_\ell^{(d)}(t - D)\Big) + Z_\ell(t) \tag{7}$$

where H_ℓ is the complex channel gain, D is the channel delay, and $Z_0(t), \ldots, Z_{n_\mathrm{b}-1}(t)$ are independent white complex Gaussian processes with power spectral density N_0. For convenience, we set $N_0 = 1$, and we can therefore interpret ρ as the SNR per pilot or data symbol.

C. Synchronization and Channel Estimation

The first step in the receiver is to estimate the delay and channel gains. The synchronization and channel estimation algorithm makes use of an upsampled version of the received signal. We see from (4) and (5) that we can write the pilot signal as

$$x^{(p)}(t) = \sum_{n=0}^{\mathrm{N}n_p-1} x^{(p)}_{\lfloor n/\mathrm{N}\rfloor}\sqrt{\frac{t_\mathrm{s}}{t_\mathrm{p}}}s_{t_\mathrm{s}}(t-nt_\mathrm{s}) = \sum_{n=0}^{\mathrm{N}n_p-1} x^{(p)}_{\mathrm{N},n}s_{t_\mathrm{s}}(t-nt_\mathrm{s}) \tag{8}$$

where the integer $\mathrm{N}\geq 1$ is the upsampling factor, $t_\mathrm{s}=t_\mathrm{p}/\mathrm{N}$, and

$$x^{(p)}_{\mathrm{N},n} = \begin{cases} \frac{1}{\sqrt{\mathrm{N}}}x^{(p)}_{\lfloor n/\mathrm{N}\rfloor}, & n=0,1,\ldots,\mathrm{N}n_\mathrm{p}-1 \\ 0, & \text{otherwise} \end{cases}$$

is a scaled and upsampled version of the pilot sequence. We can, of course, recover $x^{(p)}_{\mathrm{N},n}$ from $x^{(p)}(t)$ by sampling the output of a causal filter with impulse response $s_{t_\mathrm{s}}(t)$ and input $x^{(p)}(t)$. Indeed, the filter output is

$$(x^{(p)} * s_{t_\mathrm{s}})(t) = \sum_{n=0}^{\mathrm{N}n_p-1} x^{(p)}_{\mathrm{N},n}r_{t_\mathrm{s}}(t-nt_\mathrm{s}) \tag{9}$$

where $r_{t_\mathrm{s}}(t)$ is a triangular pulse with duration $2t_\mathrm{s}$ and unit peak value,

$$r_{t_\mathrm{s}}(t) = (s_{t_\mathrm{s}} * s_{t_\mathrm{s}})(t) = \begin{cases} t/t_\mathrm{s}, & 0\leq t<t_\mathrm{s} \\ 1-t/t_\mathrm{s}, & t_\mathrm{s}\leq t\leq 2t_\mathrm{s} \\ 0, & \text{otherwise} \end{cases}. \tag{10}$$

Clearly, $x^{(p)}_{\mathrm{N},m} = (x^{(p)} * s_{t_\mathrm{s}})((m+1)t_\mathrm{s})$. However, for a sampling offset $e\in[0,t_\mathrm{s})$, we have that

$$(x^{(p)} * s_{t_\mathrm{s}})((m+1)t_\mathrm{s}+e) = \left(1-\frac{e}{t_\mathrm{s}}\right)x^{(p)}_{\mathrm{N},m} + \frac{e}{t_\mathrm{s}}x^{(p)}_{\mathrm{N},m+1} \tag{11}$$

which will be used later.

To capture the entire pilot signal from the received signal due to the ℓth subpacket, we take M samples of the matched filter output

$$Y^{(p)}_{m,\ell} = \left(Y_\ell * s_{t_\mathrm{s}}\right)\left((m+1)t_\mathrm{s}\right), \qquad m=0,\ldots,\mathrm{M}-1 \tag{12}$$

where

$$\mathrm{M} = \left\lceil \frac{d_\mathrm{max}}{t_\mathrm{s}} \right\rceil + n_p\mathrm{N} \tag{13}$$

is sufficiently large, since the delay $D \in [0, d_{\max}]$. We form the received vector (for the purpose of synchronization and channel estimation) as $\boldsymbol{Y}_\ell^{(p)} = [Y_{0,\ell}, \ldots, Y_{\mathrm{M}-1,\ell}^{(p)}]^T \in \mathbb{C}^{\mathrm{M}}$. If we let $D = Qt_{\mathrm{s}} + E$, where $Q \in \mathbb{Z}$ and $E \in [0, t_{\mathrm{s}})$, then from (11) and (12) we can write

$$\boldsymbol{Y}_\ell^{(p)} = \mathbf{v}(Q, E)H_\ell + \boldsymbol{C}_\ell + \boldsymbol{Z}_\ell \tag{14}$$

where

$$\mathbf{v}(q, e) = \begin{bmatrix} \mathbf{x}_{\mathrm{N}}^{(p)}(q) & \mathbf{x}_{\mathrm{N}}^{(p)}(q+1) \end{bmatrix} \begin{bmatrix} 1 - \frac{e}{t_{\mathrm{s}}} \\ \frac{e}{t_{\mathrm{s}}} \end{bmatrix} \tag{15}$$

$$\mathbf{x}_{\mathrm{N}}^{(p)}(q) = \begin{bmatrix} \mathbf{0}_q^T & x_{\mathrm{N},0}^{(p)} & \cdots & x_{\mathrm{N},\mathrm{N}n_p-1}^{(p)} & \mathbf{0}_{\mathrm{M}-q-\mathrm{N}n_p}^T \end{bmatrix}^T \tag{16}$$

$\boldsymbol{Z}_\ell \sim \mathcal{CN}(\mathbf{0}, \mathbf{I}_{\mathrm{M}})$ is the noise vector due to $Z_\ell(t)$, and $\boldsymbol{C}_\ell$ accounts for the interference due to the data signal $x_\ell^{(d)}(t)$.

In this paper, we choose to estimate the channel gains $\hat{\boldsymbol{H}} = [\hat{H}_0, \ldots, \hat{H}_{n_{\mathrm{b}}-1}]^T$ and delay $\hat{D} = \hat{Q}t_{\mathrm{s}} + \hat{E}$ as

$$[\hat{\boldsymbol{H}}, \hat{Q}, \hat{E}] = \arg\min_{\bar{\mathbf{h}}, \bar{q}, \bar{e}} \sum_{\ell=0}^{n_{\mathrm{b}}-1} \|\boldsymbol{Y}_\ell^{(p)} - \mathbf{v}(\bar{q}, \bar{e})\bar{h}_\ell\|^2. \tag{17}$$

Hence, we see that (17) is a least squares estimator which coincides with the ML estimator of H and D if we can ignore the influence of $\boldsymbol{C}_\ell$. However, even for the case when $\boldsymbol{C}_\ell \neq \mathbf{0}$, the numerical results presented below indicate that the mean square error of the estimator in (17) approaches the Cramér-Rao bound as the SNR increases. Hence, the estimator appears to be asymptotically efficient (just as we expect the ML estimator to be).

We note that the ℓth element of $\bar{\mathbf{h}} = [\bar{h}_0, \ldots, \bar{h}_{n_{\mathrm{b}}-1}]^T$ affects only the ℓth term in the objective function in (17). Hence, we can minimize the sum in (17) with respect to $\bar{\mathbf{h}}$ by considering each element in $\bar{\mathbf{h}}$ separately. Given the observation $\boldsymbol{Y}_\ell^{(p)} = \mathbf{y}_\ell^{(p)}$, the $\bar{h}_\ell$ that minimizes the cost function in (17) can be written as a function of $\bar{q}$ and $\bar{e}$:

$$\hat{h}_\ell(\bar{q}, \bar{e}) = \arg\min_{\bar{h}_\ell} \sum_{k=0}^{n_{\mathrm{b}}-1} \|\mathbf{y}_k^{(p)} - \mathbf{v}(\bar{q}, \bar{e})\bar{h}_k\|^2 \tag{18}$$

$$= \arg\min_{\bar{h}_\ell} \|\mathbf{y}_\ell^{(p)} - \mathbf{v}(\bar{q}, \bar{e})\bar{h}_\ell\|^2 \tag{19}$$

$$= \frac{\mathbf{v}(\bar{q}, \bar{e})^H \mathbf{y}_\ell^{(p)}}{||\mathbf{v}(\bar{q}, \bar{e})||^2}. \tag{20}$$

Hence, we can estimate the delay parameters as

$$[\hat{q},\hat{e}] = \arg\min_{\bar{q},\bar{e}} \sum_{\ell=0}^{n_{\mathrm{b}}-1} \|\mathbf{y}_\ell^{(p)} - \mathbf{v}(\bar{q},\bar{e})\hat{h}_\ell(\bar{q},\bar{e})\|^2 \tag{21}$$

$$= \arg\min_{\bar{q},\bar{e}} \sum_{\ell=0}^{n_{\mathrm{b}}-1} \|\mathbf{y}_\ell^{(p)}\|^2 - 2\,\mathrm{Re}\{(\mathbf{y}_\ell^{(p)})^H\mathbf{v}(\bar{q},\bar{e})\hat{h}_\ell(\bar{q},\bar{e})\} + \|\mathbf{v}(\bar{q},\bar{e})\|^2|\hat{h}_\ell(\bar{q},\bar{e})|^2 \tag{22}$$

$$= \arg\min_{\bar{q},\bar{e}} \sum_{\ell=0}^{n_{\mathrm{b}}-1} -2\,\mathrm{Re}\left\{\frac{|\mathbf{v}(\bar{q},\bar{e})^H\mathbf{y}_\ell^{(p)}|^2}{\|\mathbf{v}(\bar{q},\bar{e})\|^2}\right\} + \frac{|\mathbf{v}(\bar{q},\bar{e})^H\mathbf{y}_\ell^{(p)}|^2}{\|\mathbf{v}(\bar{q},\bar{e})\|^2} \tag{23}$$

$$= \arg\max_{\bar{q},\bar{e}} \sum_{\ell=0}^{n_{\mathrm{b}}-1} \frac{|\mathbf{v}(\bar{q},\bar{e})^H\mathbf{y}_\ell^{(p)}|^2}{\|\mathbf{v}(\bar{q},\bar{e})\|^2} \tag{24}$$

$$= \arg\max_{\bar{q},\bar{e}} \frac{a(\bar{q},\bar{e})}{b(\bar{q},\bar{e})} \tag{25}$$

where

$$a(\bar{q},\bar{e}) = \sum_{\ell=0}^{n_{\mathrm{b}}-1} |\mathbf{v}(\bar{q},\bar{e})^H\mathbf{y}_\ell^{(p)}|^2, \tag{26}$$

$$b(\bar{q},\bar{e}) = \|\mathbf{v}(\bar{q},\bar{e})\|^2. \tag{27}$$

For a fixed $\bar{q}$ both $a(\bar{q},\bar{e})$ and $b(\bar{q},\bar{e})$ are second-degree polynomials in $\bar{e}$. We find the extreme points of $a(\bar{q},\bar{e})/b(\bar{q},\bar{e})$, for a fixed $\bar{q}$, by differentiating with respect to $\bar{e}$ and finding the roots, i.e., the solutions to

$$\frac{\partial}{\partial\bar{e}}\frac{a(\bar{q},\bar{e})}{b(\bar{q},\bar{e})} = \left.\frac{b(\bar{q},\bar{e})a'(\bar{q},\bar{e}) - a(\bar{q},\bar{e})b'(\bar{q},\bar{e})}{b(\bar{q},\bar{e})^2}\right|_{\bar{e}=\bar{e}^*} = 0 \tag{28}$$

We note that the roots are easily found, since the numerator of (28) is polynomial with degree 3 (or less) in $\bar{e}$. If (28) has a solution $\bar{e}^*$ in the range $(0, t_s)$, then $(\bar{q},\bar{e}^*)$ is a candidate for $(\hat{q},\hat{e})$. We also consider the boundary points $(\bar{q},0)$ and $(\bar{q},t_s)$ as candidates, since they might be the solution of (25) in the case when no extreme point $\bar{e}^*$ can be found in $(0,t_s)$ or when the extreme points are minima points (instead of maxima). In passing, we also note that $a(\bar{q},\bar{e})/b(\bar{q},\bar{e})$ is not necessarily differentiable at $\bar{e}=0$, but this will not be a problem for the algorithm. Finally, we find $(\hat{q},\hat{e})$ as the best element in $\mathcal{D}$.

We are now ready to summarize the delay and channel estimation algorithm. Let $\mathcal{D}$ be the set of candidates for $(\hat{q},\hat{e})$. We initialize $\mathcal{D}$ as $\mathcal{D} = \{(\bar{q},0) : \bar{q} = 0,1,\ldots,\lceil d_{\max}/t_{\mathrm{s}}\rceil\}$. Next, we find the solutions $\bar{e}^*$ to (28) for each $\bar{q} \in \{0,1,\ldots,\lceil d_{\max}/t_{\mathrm{s}}\rceil\}$. If a solution $\bar{e}^*$ is in the range $[0,t_{\mathrm{s}})$, then we add $(\bar{q},\bar{e}^*)$ to $\mathcal{D}$. We then find $(\hat{q},\hat{e})$ as the entry in $\mathcal{D}$ that maximizes the

Algorithm 1 Synchronization and Channel Estimation

Require: t_{s}, $d_{\max}$, $\mathbf{y}_0^{(p)}, \mathbf{y}_1^{(p)}, \ldots, \mathbf{y}_{n_{\mathrm{b}}-1}^{(p)}$

Ensure: A solution $\hat{q}$, $\hat{e}$, $\hat{h}_0, \hat{h}_1, \ldots, \hat{h}_{n_{\mathrm{b}}-1}$ to (17)

$\mathcal{D} \leftarrow \{(q,0) : q = 0, 1, \ldots, \lceil d_{\max}/t_{\mathrm{s}} \rceil\}$

for $\bar{q} \in \{0, 1, \ldots, \lceil d_{\max}/t_{\mathrm{s}} \rceil\}$ **do**

  for each root $\bar{e}^*$ to (28) such that $\bar{e}^* \in (0, t_s)$ **do**

    add $(\bar{q}, \bar{e}^*)$ to $\mathcal{D}$

  end for

end for

$(\hat{q}, \hat{e}) \leftarrow$ element in $\mathcal{D}$ that maximizes the objective function in (25)

for $\ell = 0, 1, \ldots, n_{\mathrm{b}} - 1$ **do**

  $\hat{h}_\ell \leftarrow \mathbf{v}(\hat{q}, \hat{e})^H \mathbf{y}_\ell^{(p)} / ||\mathbf{v}(\hat{q}, \hat{e})||^2$

end for

objective function in (25). Finally, we use $(\hat{q}, \hat{e})$ to compute $\hat{h}_0, \hat{h}_1, \ldots, \hat{h}_{n_{\mathrm{b}}-1}$ according to (20). Pseudocode for the algorithm is given in Algorithm 1.

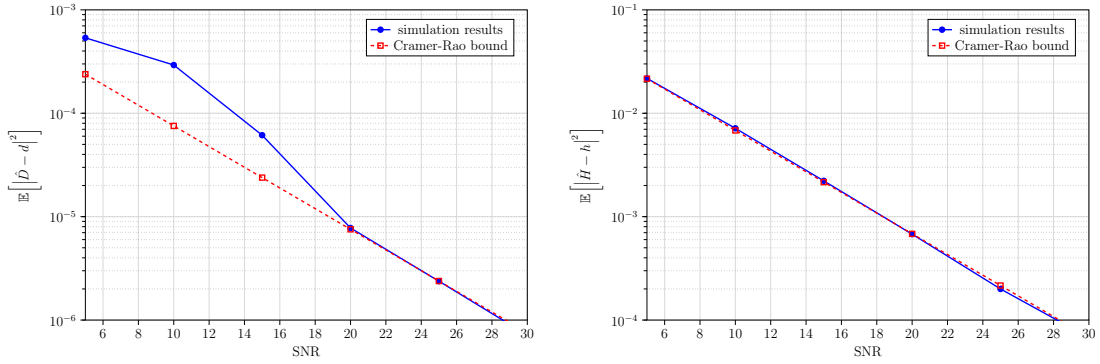
To validate our algorithm, we computed the Cramer-Rao bound (CRB) and compared it with Monte Carlo simulation results for different $D = d$ and $H = h$. In all our simulations, both the channel and delay estimation mean squared errors approach the CRBs as ρ increases. An example of one such simulation is found in Fig. 2. The details of CRB evaluation can be found in the Appendix A.

D. Codeword Decoding Phase

The codeword decoding phase makes use of the estimated delay and channel gain from the previous phase to decode the transmitted data symbols. The input-output relationship for the kth symbol in the lth block, assuming that the synchronization is not off by more than one symbol (i.e., $|\hat{D} - D| \leq t_{\mathrm{p}}$) is

$$Y_{k,\ell} = \left(Y_\ell * \tilde{s}_{t_{\mathrm{p}}}\right)(t)\big|_{t=kt_p+n_pt_p+\hat{D}} \tag{29}$$

$$= H_\ell((1-\Delta)x_{k+\Lambda,\ell} + \Delta x_{k,\ell}) + Z_{k,\ell}, \tag{30}$$



(a) Synchronization error (b) Channel estimation error

Fig. 2: Comparison of mean square error and CRB of the estimators for N = 20, $n_{\mathrm{b}} = 1$, $d_{\max} = 20$, $\mathbf{c}_\ell = \mathbf{0}$, $d = 8.34$, $h = 0.835 - 0.536i$.

where $\tilde{s}_{t_{\mathrm{p}}}(t) = s_{t_{\mathrm{p}}}(-t + t_{\mathrm{p}})$, $\Delta = 1 - (D - \hat{D})/t_{\mathrm{p}}$, $\Lambda = \mathrm{sign}(\hat{D} - D)$.[2]

In the decoding process, the receiver tries to find the codeword in the codebook closest to the received signal after scaling each part of the codeword corresponding to a different fading block by the available channel estimate. Mathematically, given the received vector and the channel estimates, the decoded codeword $\hat{\mathbf{x}} = [\hat{\mathbf{x}}_0^T, \ldots, \hat{\mathbf{x}}_{n_{\mathrm{b}}-1}^T]^T$ is determined as follows:

$$\hat{\mathbf{x}} = \underset{\bar{\mathbf{x}}=[\bar{\mathbf{x}}_0^T,\ldots,\bar{\mathbf{x}}_{n_{\mathrm{b}}-1}^T]^T \in \mathcal{C}}{\arg\min} \sum_{\ell=0}^{n_{\mathrm{b}}-1} \|\mathbf{y}_\ell - \hat{h}_\ell \bar{\mathbf{x}}_l\|^2. \tag{31}$$

where $\mathbf{y}_\ell = [y_{0,\ell}, \ldots, y_{n_{\mathrm{s}}-1,\ell}]^T$ and $\mathcal{C}$ is the codebook. This decoder, known as the *mismatched SNN decoder*, coincides with the ML decoder only when the receiver has perfect channel state information, i.e., $\hat{h}_\ell = h_\ell$ for $\ell = 0, \ldots, n_{\mathrm{b}} - 1$ and perfect synchronization, i.e. $\delta = 1$. This decoder, although not optimal, is practically relevant and also yields tractable analysis of information-theoretic bounds [13].

III. A Non-Asymptotic Upper Bound on The Error Probability

We may evaluate the error probability as

$$\epsilon_{\mathrm{pep}} = \mathbb{P}\Big(\Big|\hat{D} - D\Big| \le t_{\mathrm{p}}\Big)\,\epsilon_1 + \mathbb{P}\Big(\Big|\hat{D} - D\Big| > t_{\mathrm{p}}\Big)\,\epsilon_2 \tag{32}$$

where ϵ_1 and ϵ_2 are the probability of erroneous packet decoding when the synchronization is off by less than and more than one symbol respectively. When the synchronization is off more than

[2]We have omitted the superscript $(.)^{(d)}$ for the rest of the paper to keep the notation simple.

one symbol, there is a minuscule probability of successful decoding since our decoder operates under the assumption of perfect synchronization, By assuming the decoder cannot decode the packet when synchronization is off more than one symbol, i.e. $\epsilon_2 = 1$, we may bound the error probability as

$$\epsilon_{\text{pep}} \leq \mathbb{P}\Big(\Big|\hat{D} - D\Big| \leq t_{\text{p}}\Big)\,\epsilon_1 + \mathbb{P}\Big(\Big|\hat{D} - D\Big| > t_{\text{p}}\Big) \tag{33}$$

In the next section, we will present the RCUs bound for ϵ_1 and its corresponding saddlepoint approximation.

A. The RCUs Finite-Blocklength Bound

Like most of the achievability results in information theory, the RCUs bound which we will focus on in this paper is obtained by a random-coding argument. This means that instead of analyzing the performance of a particular code, we evaluate the average error probability averaged over a randomly constructed ensemble of codebooks. In this paper, we consider an i.i.d. discrete ensemble in which each symbol of every codeword is drawn independently (and uniformly) from a constellation set with u elements (e.g. for BPSK, $u = 2$) and power ρ. Note that for $\sigma_\ell^2 = 1$, ρ determines the SNR of the transmission. Although suboptimal, this choice of the random code leads to tractable expressions once we introduce the asymptotic expansion of the RCUs bound.

For our setup, the RCUs achievability bound ϵ_{ub} on ϵ_1 is given by

$$\epsilon_1 \leq \epsilon_{\text{ub}} = \mathbb{P}\left[\frac{\log U}{n_{\text{c}} n_{\text{b}}} + \frac{1}{n_{\text{c}} n_{\text{b}}} \sum_{\ell=0}^{n_{\text{b}}-1} \sum_{k=0}^{n_{\text{s}}-1} \imath_s(X_{k,\ell}; Y_{k,\ell}, \hat{H}_\ell) \leq \text{R}\right] \tag{34}$$

where U is a random variable that uniformly distributed on $[0, 1]$ and independent of all other quantities, and $\imath_s(x, y, \hat{h})$ is the so-called *generalized information density* which is defined as [12]

$$\imath_s(x; y, \hat{h}) = \log \frac{e^{-s\left|y - \hat{h}x\right|^2}}{\mathbb{E}_{\bar{X}}\left[e^{-s\left|y - \hat{h}\bar{X}\right|^2}\right]}, \tag{35}$$

where $\bar{X}$ is independent and has the same distribution as X and $s > 0$ is an optimization parameter that can be used to obtain a tighter bound.

There is no closed-form expression for the RCUs bound (34), in general. Therefore, we need to evaluate it using numerical methods, such as Monte-Carlo simulations. Unfortunately, such an evaluation can be extremely time consuming due to the low target error probabilities of interest in URLLC. Next, we introduce an asymptotic expansion and a saddlepoint approximation of (34) to compute it efficiently.

B. The Saddlepoint Approximation for Markov Chains

We first note that the RCUs bound in (34) can also be stated in the conditional form as

$$\epsilon_{\text{ub}} = \mathbb{E}_{\boldsymbol{H},\hat{\boldsymbol{H}},\Delta}\left[\epsilon_{\text{ub}}\left(\boldsymbol{H},\hat{\boldsymbol{H}},\Delta\right)\right] \tag{36}$$

where

$$\epsilon_{\text{ub}}\left(\mathbf{h},\hat{\mathbf{h}},\Delta\right) = \mathbb{P}\left[\frac{\log U}{n_{\text{c}}n_{\text{b}}} + \frac{1}{n_{\text{c}}n_{\text{b}}}\sum_{\ell=1}^{n_{\text{b}}}\sum_{k=1}^{n_{\text{s}}}\imath_s\left(X_{k,\ell};Y_{k,\ell},\hat{h}_\ell\right) \le R \middle| \boldsymbol{H}=\mathbf{h},\hat{\boldsymbol{H}}=\hat{\mathbf{h}},\Delta=\delta\right] \tag{37}$$

$\boldsymbol{H} = [H_0,\ldots,H_{n_{\text{b}}-1}]^T$, and U is a random variable uniformly distributed over [0,1]. Here, given $\mathbf{h}$, $\hat{\mathbf{h}}$ and δ, $X_{k,\ell}$ are conditionally i.i.d. in k, yet the same cannot be said for $Y_{k,\ell}$. As a result, $\imath_s\left(X_{k,\ell};Y_{k,\ell},\hat{h}_\ell\right)$ is also dependent in k. This is due to the imperfect synchronization as shown in (30). Note that given $\Delta=\delta$, all random quantities in (37) are conditionally independent in ℓ.

When analyzing the error probability, we assume that the interference due to the imperfect synchronization is always caused by the previous symbol, or equivalently $\hat{D} \le D$ and $\Lambda = -1$ for simplicity.

Let $b_{k,\ell}$ be the enumeration of a state with two elements $[x_{k-1,\ell}, x_{k,\ell}]$. Then,

$$\mathbb{P}(B_{k,\ell} \,|\, B_{k-1,\ell},\ldots,B_{0,\ell}) = \mathbb{P}(B_{k,\ell} \,|\, X_{k-1,\ell},X_{k-2,\ell}) \tag{38}$$

$$= \mathbb{P}(B_{k,\ell} \,|\, B_{k-1,\ell}), \tag{39}$$

which proves that $B_{k,\ell}$ forms a Markov chain. Note that $\imath_s(X_{k,\ell};Y_{k,\ell},\hat{h}_\ell)$ in (37) can be expressed as a function of $B_{k,l}$ and $Z_{k,\ell}$. Also note that for our setup, the variable $y_{k,\ell}$ given in (30) depends on both $x_{k,\ell}$ and $x_{k-1,\ell}$ which means that $y_{k,\ell}$ depends on both $B_{k,\ell}$ and $B_{k-1,\ell}$, and it does not depend on $B_{k-i,\ell}$ for $k \ge i \ge 2$.

Let $\varphi_\ell(\zeta)$ and $\kappa_\ell(\zeta)$ be moment-generating and cumulant-generating functions of $\sum_{k=1}^{n_{\text{s}}} -\imath_s\left(X_{k,\ell};Y_{k,\ell},\hat{h}_\ell\right)$ respectively; then

$$\varphi_\ell(\zeta) = \mathbb{E}\left[e^{-\zeta\sum_{k=1}^{n_{\text{s}}}\imath_s\left(X_{k,\ell};Y_{k,\ell},\hat{h}_\ell\right)}\right] \tag{40}$$

and $\kappa_\ell(\zeta) = \log(\varphi_\ell(\zeta))$. We also let

$$\mu_\ell(\zeta) = \frac{1}{n_{\text{s}}}\frac{\partial\kappa_\ell(\zeta)}{\partial\zeta} \tag{41}$$

and

$$\sigma_\ell^2(\zeta) = \frac{1}{n_{\text{s}}}\frac{\partial^2\kappa_\ell(\zeta)}{\partial\zeta^2}. \tag{42}$$

Indeed, for $\zeta = 0$, $\mu_\ell(0)$ and $\sigma^2_\ell(0)$ correspond to the mean and variance of $\sum_{k=1}^{n_\mathrm{s}} -\imath_s\Big(X_{k,\ell}; Y_{k,\ell}, \hat{h}_\ell\Big)$ respectively. We also let $\mu(\zeta) = \sum_{\ell=0}^{n_\mathrm{b}-1} \mu_\ell(\zeta)$, and $\sigma^2(\zeta) = \sum_{\ell=0}^{n_\mathrm{b}-1} \sigma^2_\ell(\zeta)$ which correspond to the mean and the variance of $\sum_{k=0}^{n_\mathrm{s}-1}\sum_{\ell=0}^{n_\mathrm{b}-1} -\imath_s\Big(X_{k,\ell}; Y_{k,\ell}, \hat{h}_\ell\Big)$ for $\zeta = 0$ respectively.

Before we introduce the saddlepoint approximation for (37), we first introduce the saddlepoint approximation for the sum of a function of dependent random variables as (see [11, Ch-6] for details): Assume for $\underline{\zeta} \le \zeta \le \overline{\zeta}$, $\kappa_\ell(\zeta)$ and its first two derivatives exist for every ℓ. If there exists a $\underline{\zeta} \le \zeta \le \overline{\zeta}$ satisfying $-\mu(\zeta)n_\mathrm{s}/(n_\mathrm{c}n_\mathrm{b}) = \mathrm{R}$, then for $\zeta \in [0, \overline{\zeta}]$

$$\mathbb{P}\left[\frac{1}{n_\mathrm{c}n_\mathrm{b}}\sum_{k=0}^{n_\mathrm{s}-1}\sum_{\ell=0}^{n_\mathrm{b}-1} \imath_s\Big(X_{k,\ell}; Y_{k,\ell}, \hat{h}_\ell\Big) \le R\right] \approx e^{\kappa(\zeta)-n_\mathrm{s}\zeta\mu(\zeta)}\left[e^{\frac{\beta_\zeta^2}{2}}Q(\beta_\zeta)\right] \tag{43}$$

where $\beta_u = u\sqrt{n_\mathrm{s}\sigma^2(\zeta)}$ and $\kappa(\zeta) = \sum_{\ell=0}^{n_\mathrm{b}-1}\kappa_\ell(\zeta)$. For $\zeta \in [\underline{\zeta}, 0]$

$$\mathbb{P}\left[\frac{1}{n_\mathrm{c}n_\mathrm{b}}\sum_{\ell=0}^{n_\mathrm{b}-1}\sum_{k=0}^{n_\mathrm{s}-1} \imath_s\Big(X_{k,\ell}; Y_{k,\ell}, \hat{h}_\ell\Big) \le R\right] \approx 1 - e^{\kappa(\zeta)-n_\mathrm{s}\zeta\mu(\zeta)}\left[e^{\frac{\beta_\zeta^2}{2}}Q(\beta_{-\zeta})\right]. \tag{44}$$

Here, compared to (37), we are missing a $\log U$ term inside the probability on the left hand side of (43) and (44). How saddlepoint approximation should be updated to regarding for the $\log U$ term is explained in [21, App. B] (also see [26, App. E]). Note that both in [21] and [26], authors consider the tail of the sum of independent random variables. However, their methods can be applied to the sum of dependent random variables as well. Thus, by following the same steps from [21, App. B] in our setup, we obtain the saddlepoint approximation of (37) as follows:

Assume for $\underline{\zeta} \le \zeta \le \overline{\zeta}$, $\kappa_\ell(\zeta)$ and its first two derivatives exist for every ℓ and let there be a ζ in $\underline{\zeta} < \zeta < \overline{\zeta}$ satisfying $-\mu(\zeta)n_\mathrm{s}/(n_\mathrm{c}n_\mathrm{b}) = \mathrm{R}$. If $\zeta \in [0, 1]$ then

$$\epsilon_{\mathrm{ub}}\Big(\mathbf{h}, \hat{\mathbf{h}}, \delta\Big) \approx e^{\kappa(\zeta)-n_\mathrm{s}\zeta\mu(\zeta)}\left[e^{\frac{\beta_\zeta^2}{2}}Q(\beta_\zeta) + e^{\frac{\beta_{1-\zeta}^2}{2}}Q(\beta_{1-\zeta})\right]. \tag{45}$$

If $\zeta > 1$, then

$$\epsilon_{\mathrm{ub}}\Big(\mathbf{h}, \hat{\mathbf{h}}, \delta\Big) \approx e^{\kappa(1)-n_\mathrm{s}\mu(\zeta)}\Big[\tilde{\Psi}_{n_\mathrm{s}}(1,1) + \tilde{\Psi}_{n_\mathrm{s}}(0,-1)\Big] \tag{46}$$

where

$$\tilde{\Psi}_{n_\mathrm{s}}(\alpha_1,\alpha_2) = e^{\alpha_1\left[-n_\mathrm{s}\mu(1)-n_\mathrm{s}\mathrm{R}+\frac{n_\mathrm{s}\sigma(1)}{2}\right]}Q\left(\alpha_1\sqrt{n_\mathrm{s}\sigma(1)} - \alpha_2\frac{n_\mathrm{s}\mu(1)+n_\mathrm{s}\mathrm{R}}{\sqrt{n_\mathrm{s}\sigma(1)}}\right). \tag{47}$$

If $\zeta < 0$, then

$$\epsilon_{\mathrm{ub}}(\mathbf{h}, \hat{\mathbf{h}}, \delta) \approx 1 - e^{\kappa(\zeta)-n_\mathrm{s}\zeta\mu(\zeta)}\left[e^{\frac{\beta_{-\zeta}^2}{2}}Q(\beta_{-\zeta}) - e^{\frac{\beta_{1-\zeta}^2}{2}}Q(\beta_{1-\zeta}) + \mathcal{O}\left(\frac{1}{\sqrt{n_\mathrm{s}}}\right)\right]. \tag{48}$$

Algorithm 2 Saddlepoint Approximation

Input : $\rho, \mathrm{R}, n_{\mathrm{s}}, n_{\mathrm{p}}, n_{\mathrm{b}}, \mathrm{N_{MC}}$

Output : ϵ_{sp}

Draw $(\mathbf{h}_i, \hat{\mathbf{h}}_i, \delta_i)$ for $i = 1, \ldots, \mathrm{N_{MC}}$

for every $(\mathbf{h}_i, \hat{\mathbf{h}}_i, \delta_i)$ **do**

    Solve $\mathrm{R} = \frac{-\kappa'(\zeta) n_{\mathrm{s}}}{n_{\mathrm{c}} n_{\mathrm{b}}}$ for ζ, denote solution by ζ^*

    Compute $\kappa(\zeta^*)$, $\kappa''(\zeta^*)$

    Compute $\epsilon_{\mathrm{sp}}(\mathbf{h}_i, \hat{\mathbf{h}}_i, \delta_i)$

end for

$\epsilon_{\mathrm{sp}} = \frac{1}{\mathrm{N_{MC}}} \sum_{i=1}^{\mathrm{N_{MC}}} \epsilon_{\mathrm{sp}}(\mathbf{h}_i, \hat{\mathbf{h}}_i, \delta_i)$

Pseudo-code for the saddlepoint approximation is given in Algorithm 2. The saddlepoint approximation requires the existence of the MGF $\varphi(\zeta)$ and its derivatives. Due to the dependent random variables we are dealing with in our model, the MGF and its derivatives cannot be factorized, which is one of the main challenges here compared to the i.i.d. case. To tackle with this problem, we next take advantage of the Markovian structure of our setup and provide a practical method to evaluate the MGF.

C. MGF of the Sum of Information Densities

For the purpose of the derivations in this section, we assume that $X_{-1,\ell}$ takes a value from the constellation. If the pilot symbols are selected from the same constellation, $X_{-1,\ell}$ would take a value from our constellation since the pilot transmission phase always takes place before the data transmission. Even when this is not the case, it should be assumed that it takes a value from the constellation, as it is crucial for the following derivations.

Let $b_{k,\ell} = i$ is equivalent to having $x_{k-1,\ell} = i_1$ $x_{k,\ell} = i_2$ and let P_ℓ be the $u^2 \times u^2$ transition probability matrix for the Markov chain with entries $\{p_{ij,\ell}\}$ as

$$p_{ij,\ell} = \mathbb{P}(B_{k+1,\ell} = j \mid B_{k,\ell} = i) \tag{49}$$

$$= \mathbb{P}(X_{k+1,\ell} = j_2, X_{k,\ell} = j_1 \mid X_{k,\ell} = i_2, X_{k-1,\ell} = i_1) \tag{50}$$

$$= \mathbb{P}(X_{k+1,\ell} = j_2, X_{k,\ell} = j_1 \mid X_{k,\ell} = i_2) \tag{51}$$

$$= \mathbb{P}(X_{k+1,\ell} = j_2)\, \mathbb{1}(j_1 = i_2) \tag{52}$$

and $\boldsymbol{\nu} = [\nu_0, \ldots, \nu_{u^2-1}]^T$ be the probability distribution of the initial state. Furthermore, we let

$$\varphi_{j,\ell}(\zeta) = \mathbb{E}_Z\left[e^{-\zeta \imath_s\left(X_{k,\ell}, Y_{k,\ell}, \hat{h}_\ell\right)} \;\middle|\; B_{k,\ell} = j\right] \tag{53}$$

$$= \mathbb{E}_Z\left[e^{-\zeta \imath_s\left(X_{k,\ell}, Y_{k,\ell}, \hat{h}_\ell\right)} \;\middle|\; X_{k,\ell} = j_2, X_{k-1,\ell} = j_1\right] \tag{54}$$

$\mathsf{P}_\ell(\zeta)$ be a $u^2 \times u^2$ matrix with entries $\{p_{ij,\ell}\varphi_{j,\ell}(\zeta)\}$, and

$$\boldsymbol{\nu}_\ell(\zeta) = [\nu_0\varphi_{0,\ell}(\zeta), \ldots, \nu_{u^2-1}\varphi_{u^2-1,\ell}(\zeta)]^T. \tag{55}$$

For our case, it is straightforward to show that Markov chain is irreducible and aperiodic, since $p_{ij,\ell}^{(n)} > 0$, where $p_{ij,\ell}^{(n)}$ are the elements of P_ℓ^n for $n > 0$. Using these definitions, the moment generating function of $\sum_{k=0}^{n_s-1} \imath_s\Big(X_{k,\ell}; Y_{k,\ell}, \hat{h}_\ell\Big)$ can be found as

$$\varphi_\ell(\zeta) = \mathbb{E}\left[e^{\zeta \sum_{k=0}^{n_s-1} \imath_s\left(X_{k,\ell}; Y_{k,\ell}, \hat{h}_\ell\right)}\right] \tag{56}$$

$$= \mathbb{E}\left[e^{\zeta \imath_s\left(X_{0,\ell}; Y_{0,\ell}, \hat{h}_\ell\right)} e^{\zeta \sum_{k=1}^{n_s-1} \imath_s\left(X_{k,\ell}; Y_{k,\ell}, \hat{h}_\ell\right)}\right] \tag{57}$$

$$= \mathbb{E}\left[\mathbb{E}\left[e^{\zeta \imath_s\left(X_{0,\ell}; Y_{0,\ell}, \hat{h}_\ell\right)} e^{\zeta \sum_{k=1}^{n_s-1} \imath_s\left(X_{k,\ell}; Y_{k,\ell}, \hat{h}_\ell\right)} \;\middle|\; B_{0,\ell} = i, B_{n_s-1,\ell} = j\right]\right] \tag{58}$$

$$= \mathbb{E}\left[\mathbb{E}\left[e^{\zeta \imath_s\left(X_{0,\ell}; Y_{0,\ell}, \hat{h}_\ell\right)} \;\middle|\; B_{0,\ell} = i\right]\right] \times \mathbb{E}\left[\mathbb{E}\left[e^{\zeta \sum_{k=1}^{n_s-1} \imath_s\left(X_{k,\ell}; Y_{k,\ell}, \hat{h}_\ell\right)} \;\middle|\; B_{0,\ell} = i, B_{n_s-1,\ell} = j\right]\right] \tag{59}$$

$$= \sum_{i=0}^{u^2-1} \mathbb{E}\left[e^{\zeta \imath_s\left(X_{0,\ell}; Y_{0,\ell}, \hat{h}_\ell\right)} \;\middle|\; B_{0,\ell} = i\right] \nu_i \times \sum_{i=0}^{u^2-1} \sum_{j=0}^{u^2-1} \mathbb{E}\left[e^{\zeta \sum_{k=1}^{n_s-1} \imath_s\left(X_{k,\ell}; Y_{k,\ell}, \hat{h}_\ell\right)} \;\middle|\; B_{0,\ell} = i, B_{n_s-1,\ell} = j\right] p_{ij,\ell}^{(n_s-1)} \tag{60}$$

$$= \sum_{i=0}^{u^2-1} \varphi_{i,\ell}(\zeta)\nu_i \sum_{j=0}^{u^2-1} \left(p(\zeta)^{(n_s-1)}\right)_{ij,\ell} \tag{61}$$

$$= \left(\boldsymbol{\nu}_\ell(\zeta)\right)^T \mathsf{P}_\ell(\zeta)^{n_s-1} e_{u^2} \tag{62}$$

where $p_{ij,\ell}^{(n_s-1)}$ and $\left(p(\zeta)^{(n_s-1)}\right)_{ij,\ell}$ are the ijth entries of matrices $\mathsf{P}_\ell^{n_s-1}$ and $\mathsf{P}_\ell(\zeta)^{n_s-1}$ respectively, $\mathbf{e}_{u^2}$ is column ones vector with length u^2, (59) follows since given $B_{0,\ell}$, $e^{\zeta \imath_s\left(X_{0,\ell}; Y_{0,\ell}, \hat{h}_\ell\right)}$ is conditionally independent of $e^{\zeta \sum_{k=1}^{n_s-1} \imath_s\left(X_{k,\ell}; Y_{k,\ell}, \hat{h}_\ell\right)}$, and (61) follows from [11, Lemma 9.1.2]

$$\left(p(\zeta)^{(n_s-1)}\right)_{ij,\ell} = \mathbb{E}\left[e^{\zeta \sum_{k=1}^{n_s-1} \imath_s(X_{k,\ell}; Y_{k,\ell}, \hat{h}_\ell)} \;\middle|\; B_{0,\ell} = i, B_{n_s-1,\ell} = j\right] p_{ij,\ell}^{(n_s-1)}. \tag{63}$$

With this, we now have an efficient way of evaluating $\varphi_\ell(\zeta)$, and as a result, we can evaluate $\kappa(\zeta)$ and its first two derivatives efficiently.

$b_{k,\ell}$	$x_{k-1,\ell}$	$x_{k,\ell}$
0	$x^{(1)}$	$x^{(1)}$
1	$x^{(1)}$	$x^{(2)}$
2	$x^{(2)}$	$x^{(1)}$
3	$x^{(2)}$	$x^{(2)}$

TABLE I: The states for the Markov chain $b_{k,\ell}$.

IV. Example: BPSK Constellation

In this section, we show how the saddlepoint approximation of the RCUs bound for the introduced system model can be efficiently evaluated when the data points $x_{k,\ell}$ are selected from a BPSK constellation. We assume the constellation points are $x^{(1)} = -\sqrt{\rho}$ and $x^{(2)} = \sqrt{\rho}$, and as a result $u = 2$. Without loss of generality, we defined the states for the Markov chain $b_{k,\ell}$ as in Table I.

We next show that $\varphi_0(\zeta) = \varphi_3(\zeta)$ as

$$\varphi_{0,\ell}(\zeta) = \mathbb{E}_Z \left[e^{\zeta s \left| Y_{k,\ell} - \hat{h}_\ell x^{(1)} \right|^2} \left(\frac{1}{2} \left(e^{-s \left| Y_\ell - \hat{h}_\ell x^{(1)} \right|^2} + e^{-s \left| Y_\ell - \hat{h}_\ell x^{(2)} \right|^2} \right) \right)^\zeta \right.$$
$$\left. \Big| X_{k,\ell} = x^{(1)}, X_{k-1,\ell} = x^{(1)} \right] \tag{64}$$
$$= \frac{1}{2^\zeta} \int e^{\zeta s \left| x^{(1)}(h_\ell - \hat{h}_\ell) + Z \right|^2} \left(e^{-s \left| x^{(1)}(h_\ell - \hat{h}_\ell) + Z \right|^2} + e^{-s \left| x^{(1)}(h_\ell + \hat{h}_\ell) + Z \right|^2} \right)^\zeta \mathrm{d}F(z) \tag{65}$$
$$= \frac{1}{2^\zeta} \int e^{\zeta s \left| -x^{(2)}(h_\ell - \hat{h}_\ell) - \tilde{Z} \right|^2} \left(e^{-s \left| -x^{(2)}(h_\ell - \hat{h}_\ell) - \tilde{Z} \right|^2} + e^{-s \left| -x^{(2)}(h_\ell + \hat{h}_\ell) - \tilde{Z} \right|^2} \right)^\zeta \mathrm{d}F(z) \tag{66}$$
$$= \frac{1}{2^\zeta} \int e^{\zeta s \left| x^{(2)}(h_\ell - \hat{h}_\ell) + \tilde{Z} \right|^2} \left(e^{-s \left| x^{(2)}(h_\ell - \hat{h}_\ell) + \tilde{Z} \right|^2} + e^{-s \left| x^{(2)}(h_\ell + \hat{h}_\ell) + \tilde{Z} \right|^2} \right)^\zeta \mathrm{d}F(z) \tag{67}$$
$$= \frac{1}{2^\zeta} \int e^{\zeta s \left| x^{(2)}(h_\ell - \hat{h}_\ell) + \tilde{Z} \right|^2} \left(e^{-s \left| x^{(2)}(h_\ell - \hat{h}_\ell) + \tilde{Z} \right|^2} + e^{-s \left| x^{(2)}(h_\ell + \hat{h}_\ell) + \tilde{Z} \right|^2} \right)^\zeta \mathrm{d}F(\tilde{z}) \tag{68}$$
$$= \frac{1}{2^\zeta} \int e^{\zeta s \left| x^{(2)}(h_\ell - \hat{h}_\ell) + \tilde{Z} \right|^2} \left(e^{-s \left| x^{(2)}(h_\ell - \hat{h}_\ell) + \tilde{Z} \right|^2} + e^{-s \left| x^{(2)}(h_\ell + \hat{h}_\ell) + \tilde{Z} \right|^2} \right)^\zeta \mathrm{d}F(z) \tag{69}$$
$$= \mathbb{E}_Z \left[e^{\zeta s \left| x^{(2)}(h_\ell - \hat{h}_\ell) + Z \right|^2} \frac{1}{2^\zeta} \left(e^{-s \left| x^{(2)}(h_\ell - \hat{h}_\ell) + Z \right|^2} + e^{-s \left| x^{(2)}(h_\ell + \hat{h}_\ell) + Z \right|^2} \right)^\zeta \right.$$
$$\left. \Big| X_{k,\ell} = x^{(2)}, X_{k-1,\ell} = x^{(2)} \right] \tag{70}$$
$$= \varphi_{3,\ell}(\zeta) \tag{71}$$

where $Z \sim \mathcal{CN}(0,1)$, $\tilde{Z} = -Z$, $F(z)$ is the CDF of Z and $F(\tilde{z})$ is the CDF of $\tilde{Z}$, which by the definition of CDF $F(\tilde{z}) = F(z)$. By following same steps, we can also show that $\varphi_{1,\ell}(\zeta) = \varphi_{2,\ell}(\zeta)$ as

$$\varphi_{1,\ell}(\zeta) = \mathbb{E}_Z \left[e^{\zeta s \left| Y_{k,\ell} - \hat{h}_\ell x^{(2)} \right|^2} \left(\frac{1}{2} \left(e^{-s \left| Y_{k,\ell} - \hat{h}_\ell x^{(1)} \right|^2} + e^{-s \left| Y_{k,\ell} - \hat{h}_\ell x^{(2)} \right|^2} \right) \right)^\zeta \right.$$
$$\left. \left| X_{k-1,\ell} = x^{(1)}, X_{k,\ell} = x^{(2)} \right. \right] \tag{72}$$

$$= \frac{1}{2^\zeta} \int e^{\zeta s \left| \bar{x}_\ell - x^{(2)} \hat{h}_\ell + Z \right|^2} \left(e^{-s \left| \bar{x}_\ell - \hat{h}_\ell x^{(1)} + Z \right|^2} + e^{-s \left| \bar{x}_\ell - \hat{h}_\ell x^{(2)} + Z \right|^2} \right)^\zeta \mathrm{d}F(z) \tag{73}$$

$$= \frac{1}{2^\zeta} \int e^{\zeta s \left| -\bar{x}_\ell + x^{(2)} \hat{h}_\ell - Z \right|^2} \left(e^{-s \left| -\bar{x}_\ell + \hat{h}_\ell x^{(1)} - Z \right|^2} + e^{-s \left| \bar{x}_\ell + \hat{h}_\ell x^{(2)} - Z^2 \right|} \right)^\zeta \mathrm{d}F(z) \tag{74}$$

$$= \frac{1}{2^\zeta} \int e^{\zeta s \left| -\bar{x}_\ell - x^{(1)} \hat{h}_\ell + \tilde{Z} \right|^2} \left(e^{-s \left| -\bar{x}_\ell - \hat{h}_\ell x^{(2)} + \tilde{Z} \right|^2} + e^{-s \left| -\bar{x}_\ell - \hat{h}_\ell x^{(2)} + \tilde{Z} \right|^2} \right)^\zeta \mathrm{d}F(\tilde{z}) \tag{75}$$

$$= \mathbb{E}_Z \left[e^{\zeta s \left| Y_{k,\ell} - \hat{h}_\ell x^{(1)} + Z \right|^2} \left(\frac{1}{2} \left(e^{-s \left| Y_{k,\ell} - \hat{h}_\ell x^{(1)} + Z \right|^2} + e^{-s \left| Y_{k,\ell} - \hat{h}_\ell x^{(2)} + Z \right|^2} \right) \right)^\zeta \right.$$
$$\left. \left| X_{k-1,\ell} = x^{(2)}, X_{k,\ell} = x^{(1)} \right. \right] \tag{76}$$

$$= \varphi_{2,\ell}(\zeta) \tag{77}$$

where

$$\bar{x}_\ell = h_\ell \big((1-\delta) x^{(1)} + \delta x^{(2)} \big) \tag{78}$$

and (76) follows from

$$-\bar{x}_\ell = h_\ell \big((1-\delta) x^{(2)} + \delta x^{(1)} \big) . \tag{79}$$

Assuming $p_{ij,\ell} = 1/2$ for all i, j and ℓ, and $\nu_i = 1/4$ for all i, $\boldsymbol{\nu}_\ell(\zeta)$ and $\mathsf{P}_\ell(\zeta)$ are given by

$$\boldsymbol{\nu}_\ell(\zeta) = \begin{bmatrix} \frac{\varphi_{0,\ell}(\zeta)}{4} & \frac{\varphi_{1,\ell}(\zeta)}{4} & \frac{\varphi_{1,\ell}(\zeta)}{4} & \frac{\varphi_{0,\ell}(\zeta)}{4} \end{bmatrix}^T \tag{80}$$

$$\mathsf{P}_\ell(\zeta) = \begin{bmatrix} \frac{\varphi_{0,\ell}(\zeta)}{2} & \frac{\varphi_{1,\ell}(\zeta)}{2} & 0 & 0 \\ 0 & 0 & \frac{\varphi_{1,\ell}(\zeta)}{2} & \frac{\varphi_{0,\ell}(\zeta)}{2} \\ \frac{\varphi_{0,\ell}(\zeta)}{2} & \frac{\varphi_{1,\ell}(\zeta)}{2} & 0 & 0 \\ 0 & 0 & \frac{\varphi_{1,\ell}(\zeta)}{2} & \frac{\varphi_{0,\ell}(\zeta)}{2} \end{bmatrix} . \tag{81}$$

Note that one of the required quantities for saddlepoint approximation is the cumulant $\kappa_\ell(\zeta)$ which requires the evaluation of $\mathsf{P}_\ell(\zeta)^{n_s}$. Considering the specific structure of $\mathsf{P}_\ell(\zeta)$, its

eigenvalue decomposition can be computed straightforwardly. This decomposition can then be employed for the efficient evaluation of $\mathsf{P}_\ell(\zeta)^{n_\mathrm{s}}$ as we show next.

Let V_ℓ be the matrix whose ith column is the eigenvector $\mathbf{v}_i$ of $\mathsf{P}_\ell(\zeta)$, and Λ_ℓ be the diagonal matrix whose diagonal elements are the corresponding eigenvalues. Then $\mathsf{P}_\ell(\zeta)^{n_\mathrm{s}} = \mathsf{V}_\ell \Lambda_\ell^{n_\mathrm{s}} \mathsf{V}_\ell^{-1}$ with

$$\Lambda_\ell^{n_\mathrm{s}} = \frac{1}{2^{n_\mathrm{s}}} \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & (\varphi_{0,\ell}(\zeta) - \varphi_{1,\ell}(\zeta))^{n_\mathrm{s}} & 0 \\ 0 & 0 & 0 & (\varphi_{0,\ell}(\zeta) + \varphi_{1,\ell}(\zeta))^{n_\mathrm{s}} \end{bmatrix} \tag{82}$$

$$\mathsf{V}_\ell = \begin{bmatrix} 0 & -\frac{\varphi_{1,\ell}(\zeta)}{\varphi_{0,\ell}(\zeta)} & -1 & 1 \\ 0 & 1 & 1 & 1 \\ -\frac{\varphi_{0,\ell}(\zeta)}{\varphi_{1,\ell}(\zeta)} & 0 & -1 & 1 \\ 1 & 0 & 1 & 1 \end{bmatrix}. \tag{83}$$

After simple mathematical operations $\kappa_\ell(\zeta)$ can be found as

$$\kappa_\ell(\zeta) = \log\Big((\boldsymbol{\nu}_\ell(\zeta))^T \, \mathsf{P}_\ell(\zeta)^{n_\mathrm{s}} \mathbf{e}_4 \Big) \tag{84}$$

$$= \log\Big((\boldsymbol{\nu}_\ell(\zeta))^T \, \mathsf{V}_\ell \Lambda_\ell^{n_\mathrm{s}} \mathsf{V}_\ell^{-1} \mathbf{e}_4 \Big) \tag{85}$$

$$= \log\big(2^{-(n_\mathrm{s}+1)} (\varphi_{0,\ell}(\zeta) + \varphi_{1,\ell}(\zeta))^{n_\mathrm{s}+1} \big) \tag{86}$$

To evaluate (44) we need $\mu_\ell(\zeta)$ in (41) and $\sigma_\ell^2(\zeta)$ in (42) which are given as

$$\mu_\ell(\zeta) = \frac{1}{n_\mathrm{s}} \frac{(1+n_\mathrm{s})\varphi'_{0,\ell}(\zeta) + \varphi'_{1,\ell}(\zeta)}{\varphi_{0,\ell}(\zeta) + \varphi_{1,\ell}(\zeta)} \tag{87}$$

$$\sigma_\ell^2(\zeta) = \frac{1}{n_\mathrm{s}} \left[-\frac{(1+n_\mathrm{s})\big(\varphi'_{0,\ell}(\zeta) + \varphi'_{1,\ell}(\zeta)\big)^2}{(\varphi_{1,\ell}(\zeta) + \varphi_{1,\ell}(\zeta))^2} + \frac{(1+n_\mathrm{s})\varphi''_{0,\ell}(\zeta) + \varphi''_{1,\ell}(\zeta)}{\varphi_{0,\ell}(\zeta) + \varphi_{1,\ell}(\zeta)} \right] \tag{88}$$

where

$$\varphi'_{j,\ell}(\zeta) = \mathbb{E}_Z \left[\frac{1}{2^\zeta} e^{s\zeta \left| Y_{k,\ell} - \hat{h}_\ell x^{(j_2)} + Z \right|^2} \Big(e^{-s\left| Y_{k,\ell} - \hat{h}_\ell x^{(1)} + Z \right|^2} + e^{-s\left| Y_{k,\ell} - \hat{h}_\ell x^{(2)} + Z \right|^2} \Big)^\zeta \Big(-\log(2) \right.$$

$$\left. + s \left| Y_{k,\ell} - \hat{h}_\ell x^{(j_2)} + Z \right|^2 + \log\Big(e^{-s\left| Y_{k,\ell} - \hat{h}_\ell x^{(1)} + Z \right|^2} + e^{-s\left| Y_{k,\ell} - \hat{h}_\ell x^{(2)} + Z \right|^2} \Big) \Big) \,\Big|\, X_{k-1,\ell} = j_1, X_{k,\ell} = j_2 \right] \tag{89}$$

and

$$\varphi''_{j,\ell}(\zeta) = \mathbb{E}_Z \Bigg[\frac{1}{2^\zeta} e^{s\zeta \left|Y_{k,\ell} - \hat{h}_\ell x^{(j_2)} + Z\right|^2} \Big(e^{-s\left|Y_{k,\ell} - \hat{h}_\ell x^{(1)} + Z\right|^2} + e^{-s\left|Y_{k,\ell} - \hat{h}_\ell x^{(2)} + Z\right|^2} \Big)^\zeta \Bigg((\log(2))^2$$
$$+ s^2 \left|Y_{k,\ell} - \hat{h}_\ell x^{(j_2)} + Z\right|^4 + 2s \left|Y_{k,\ell} - \hat{h}_\ell x^{(j_2)} + Z\right|^2 \Bigg(- \log(2)$$
$$+ \log\Big(e^{-s\left|Y_{k,\ell} - \hat{h}_\ell x^{(1)} + Z\right|^2} + e^{-s\left|Y_{k,\ell} - \hat{h}_\ell x^{(2)} + Z\right|^2} \Big) \Bigg) - 2\log(2) \log\Big(e^{-s\left|Y_{k,\ell} - \hat{h}_\ell x^{(1)} + Z\right|^2} + e^{-s\left|Y_{k,\ell} - \hat{h}_\ell x^{(2)} + Z\right|^2} \Big)$$
$$+ 2\log\Big(e^{-s\left|Y_{k,\ell} - \hat{h}_\ell x^{(1)} + Z\right|^2} + e^{-s\left|Y_{k,\ell} - \hat{h}_\ell x^{(2)} + Z\right|^2} \Big) \Bigg) \Bigg| X_{k-1,\ell} = j_1, X_{k,\ell} = j_2 \Bigg]. \tag{90}$$

We may evaluate the $\varphi_{j,\ell}(\zeta)$ and its derivatives efficiently using numerical integration methods. Also the pdf of $\imath_s(x; Y_\ell, \hat{h}_\ell)$ for the BPSK case is given in Appendix B which can be used to further simplify the numerical integration process.

V. Numerical Results and Discussion

In this section, we report numerical experiments to shed a light to the following three questions:

1) How does the upsampling rate N affects the packet error probability?
2) In the URLLC regime, is the introduced saddlepoint approximation sufficiently accurate in the presence of imperfect synchronization?
3) For a target packet error probability 10^{-5}, how much should we increase the SNR to negate the impact of imperfect synchronization? How does the number of fading blocks affect this?

To do so we have used a discrete ensemble with BPSK constellation for transmission, and m-sequences as our pilot sequence. We also consider Rayleigh fading scenario where H_ℓ for $\ell = 0 \ldots, n_{\mathrm{b}} - 1$ generated independently from $\mathcal{CN}(0, 1)$ distribution.

In Fig. 3, we report the SNR required to achieve a packet error rate $\epsilon = 10^{-5}$ for $\mathrm{R} = 0.104$ bit per channel use, $n_{\mathrm{b}} = 8$ and $n_{\mathrm{b}} n_{\mathrm{c}} = 288$ as a function of upsampling rates N. Here we see that the improvement as we increased the upsampling rate is significant until $\mathrm{N} = 5$, but after that, the performance improvement as we increase N is negligible. In this figure, we also present the results for the case in which pilot symbol sampling is not influenced by the data symbols. As expected, the difference resulting from this effect is negligible.

In Fig. 4, we report the SNR required to achieve error probability $\epsilon = 10^{-5}$ for $n_{\mathrm{b}} n_{\mathrm{c}} = 288$ and $\mathrm{R} = 0.104$ bit per channel used, as a function of the number of fading blocks n_{b} spanned by each codeword. In this figure, we compared the results for the synchronization algorithm applied independently to each fading block, the synchronization done jointly in fading

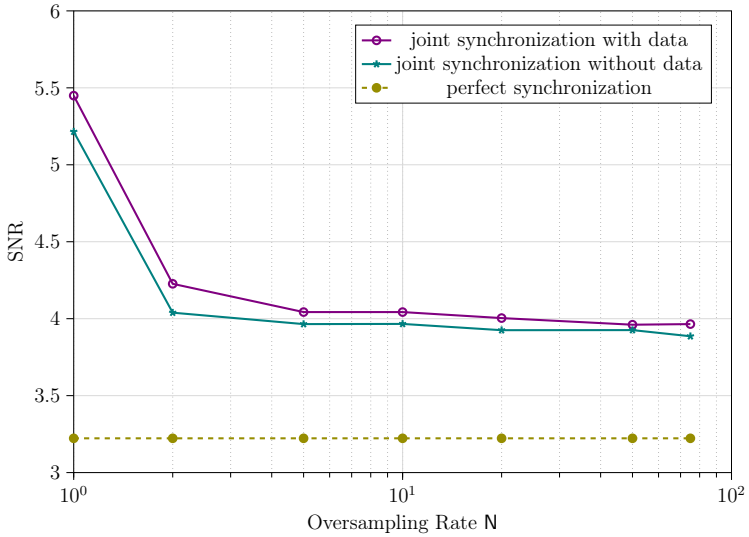


Fig. 3: Upper bound on the SNR required to achieve $\epsilon_{\text{pep}} = 10^{-5}$ as a function of N. Here, $n_{\text{b}} n_{\text{c}} = 288$, $\mathrm{R} = 0.104$ bit per channel use and $n_{\text{b}} = 8$; n_{p} and s are optimized.

blocks, the case with the perfect synchronization but imperfect channel estimation, and the case where the synchronization and the channel estimation are both perfect. As expected, performing independent synchronization over the fading blocks deteriorates the performance greatly as more than $2.5\,\mathrm{dB}$ loss is observed compared to the required SNR with joint synchronization algorithm. Another interesting observation is that even with the perfect synchronization, the required SNR to achieve $\epsilon = 10^{-5}$ only differs by $0.6\,\mathrm{dB}$ at most compared to the required SNR when joint synchronization is used for $n_{\text{b}} > 4$. For the case with perfect synchronization and channel estimation (black curve in Fig. 4), we used the number of pilot symbols n_{p} that is optimized for the cases of perfect synchronization and joint synchronization, respectively. These values are the same for this set of parameters and the optimized n_p's are provided in Table II, for a fair comparison. We observe that for $n_{\text{b}} \geq 12$, having perfect channel estimation significantly decreases the required SNR compared to perfect synchronization case. This is expected as we increase n_{b}, n_{s} decreases, and as a result, the optimal n_{p} decreases. Low n_{p} causes deterioration on channel estimation which has a significant impact on the error probability.

In Fig. 5, we report upper bounds on the packet error probability ϵ as a function of n_{p} for $\mathrm{R} = 0.104$ bit per channel use, $n_{\text{b}} = 2$, $n_{\text{c}} n_{\text{b}} = 288$, $\mathrm{N} = 20$ and $\rho = 8.45\,\mathrm{dB}$. As in the previous results, the saddlepoint approximation provide accurate results for a large range of n_{p}. This is of utmost importance because the parameter n_{p} may not be optimized specifically to minimize

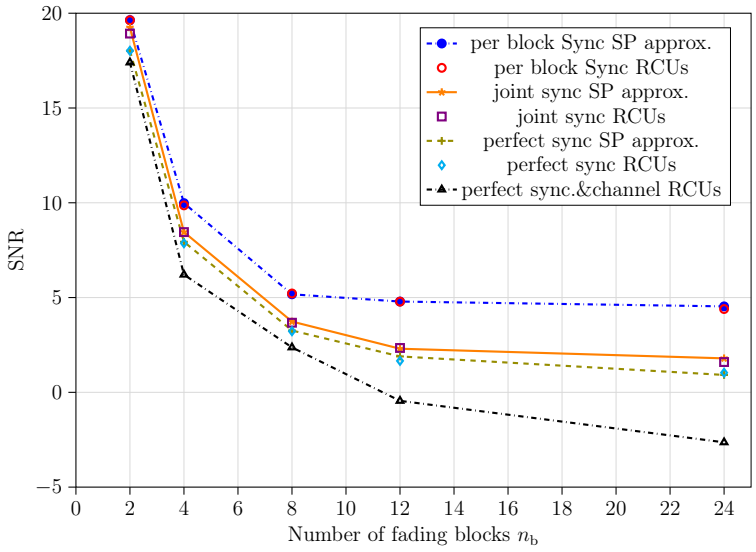


Fig. 4: Upper bound on the SNR required to achieve $\epsilon_{\mathrm{pep}} = 10^{-5}$ as a function of n_{b}. Here, $n_{\mathrm{b}} n_{\mathrm{c}} = 288$, $\mathrm{R} = 0.104$ bit per channel use, $\mathrm{N} = 20$; n_{p} and s are optimized.

n_{b}	Per block Sync n_{p}	Joint sync n_{p}	Perfect sync n_{p}
2	31	31	31
4	15	15	15
8	15	15	15
12	7	7	7
24	3	3	3

TABLE II: Optimal number of n_{p} for the data points in Fig. 4.

the packet error probability in every system. Therefore, the ability to utilize the approximations, irrespective of the chosen value for n_{p}, holds significant relevance.

VI. Conclusions

We have presented an efficient method to evaluate an upper bound on the error probability achievable over memoryless block-fading channels, with a pilot assisted-transmission for channel estimation and synchronization and nearest-neighbor decoding at the receiver. Our method is based on the saddlepoint approximation which is novel for the case with imperfect synchronization. Our numerical experiments show that the saddlepoint approximation can be safely used to benchmark practically relevant URLLC systems. We have also observed that, depending on the, synchronization method, the number of fading blocks and pilot symbols available for

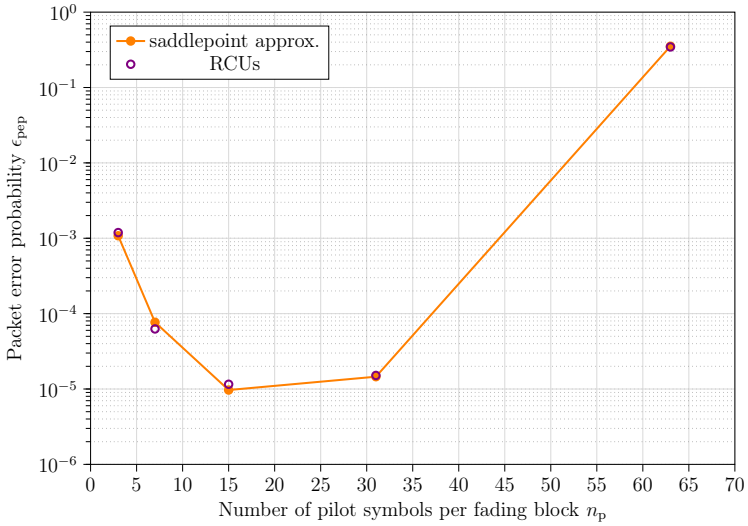


Fig. 5: Packet error probability as a function of n_{p}. Here, $\rho = 8.45\,\text{dB}$, $n_{\text{b}} = 4$, $n_{\text{b}}n_{\text{c}} = 288$, $\text{R} = 0.104$ bit per channel use, $\text{N} = 20$; s is optimized.

transmission, the SNR required to achieve packet error probability 10^{-5} non-negligibly increases. Therefore, it is advisable to utilize the suggested approximation when assessing error probabilities in URLLC optimization routines.

APPENDIX A
EVALUATION OF THE CRB OF CHANNEL AND DELAY ESTIMATE

We consider the problem of estimating the parameter vector $\boldsymbol{\theta} = [\Re\{\hat{\mathbf{h}}^T\}, \Im\{\hat{\mathbf{h}}^T\}, \epsilon]$ for the received vector given in (14). The Fisher information matrix $I(\boldsymbol{\theta})$ under the complex Gaussian additive noise with zero mean and variance σ^2 can be found in [27, Ch. 15.7] as

$$[I(\boldsymbol{\theta})]_{mn} = \frac{2}{\sigma^2}\Re\left\{\frac{\partial\phi^H(\boldsymbol{\theta})}{\partial\theta_m}\frac{\partial\phi(\boldsymbol{\theta})}{\partial\theta_n}\right\}, \quad m,n \in \{0,\dots 2n_{\text{b}}\}, \tag{91}$$

where

$$\boldsymbol{Y}^{(p)} = \left[\left(\boldsymbol{Y}_0^{(p)}\right)^T, \dots, \left(\boldsymbol{Y}_{n_{\text{b}}-1}^{(p)}\right)^T\right]^T \tag{92}$$

$$\phi(\boldsymbol{\theta}) = \mathbb{E}\left[\boldsymbol{Y}^{(p)}\right] \tag{93}$$

$$\frac{\partial\phi(\boldsymbol{\theta})}{\partial\theta_i} = \begin{cases} \left[\mathbf{0}_{\text{M}i}^T, \mathbf{l}(q,\epsilon)^T, \mathbf{0}_{\text{M}(n_{\text{b}}-i+1)}^T\right]^T, & i < n_{\text{b}} \\ \left[\mathbf{0}_{\text{M}i}^T, j\mathbf{l}(q,\epsilon)^T, \mathbf{0}_{\text{M}(n_{\text{b}}-i+1)}^T\right]^T, & n_{\text{b}} \le i < 2n_{\text{b}} \\ \left[\frac{h_0\boldsymbol{\gamma}}{t_{\text{s}}}, \dots, \frac{h_{n_{\text{b}}-1}\boldsymbol{\gamma}}{t_{\text{s}}}\right]^T, & i = 2n_{\text{b}} \end{cases} \tag{94}$$

$$\boldsymbol{\gamma} = \mathbf{x}_{p,\mathrm{N}}(q+1)^T - \mathbf{x}_{p,\mathrm{N}}(q)^T \tag{95}$$

θ_m is the mth element of $\boldsymbol{\theta}$, and M is given in (13). Note that this bound only holds over a deterministic realization of $\boldsymbol{\theta}$, meaning that the expectation in the derivation is taken only over the distribution of the AWGN $\boldsymbol{Z}_\ell$ as seen in (14).

APPENDIX B
PDF OF THE INFORMATION DENSITY

Without the loss of generality, let us assume that the same mapping in Section IV is used. For $b_{k,\ell} = [i_1, i_2]^T$, $H_\ell = h_\ell$ and $\hat{H}_\ell = \hat{h}_\ell$, Let $\imath_s(x_\ell; Y_\ell, \hat{h}_\ell)$ is shown as a function of the white Gaussian noise Z as

$$f_{i,\ell}(Z) = \log\left(\frac{2e^{-s\left(\left|c_\ell(i_1,i_2)-\hat{h}_\ell x^{(i_2)}+Z\right|^2\right)}}{e^{-s\left(\left|c_\ell(i_1,i_2)-\hat{h}_\ell x^{(1)}+Z\right|^2\right)} + e^{-s\left(\left|c_\ell(i_1,i_2)-\hat{h}_\ell x^{(2)}+Z\right|^2\right)}}\right) \tag{96}$$

where $c_\ell(i_1, i_2) = h_\ell(\tilde{\delta}x^{(i_1)} + \delta x^{(i_2)})$ and $\tilde{\delta} = 1 - \delta$. For $i = 0$ we get

$$f_{0,\ell}(Z) = \log\left(\frac{2}{1 + e^{-s\left(\left|h_\ell\sqrt{\rho}+\hat{h}_\ell\sqrt{\rho}+Z\right|^2-\left|h_\ell\sqrt{\rho}-\hat{h}_\ell\sqrt{\rho}+Z\right|^2\right)}}\right) \tag{97}$$

$$= \log\left(\frac{2}{1 + e^{-s\left(|\tilde{\alpha}_\ell+Z|^2-|\delta_\ell+Z|^2\right)}}\right) \tag{98}$$

$$= \log\left(\frac{2}{1 + e^{s\left(4\sqrt{\rho}(\mathrm{Re}\{Z\}\mathrm{Re}\{\hat{h}_\ell\}+\mathrm{Im}\{Z\}\mathrm{Im}\{\hat{h}_\ell\})+|\tilde{\alpha}_\ell|^2-|\alpha|_\ell^2\right)}}\right) \tag{99}$$

$$= \log(2) - \log\left(1 + e^{s\left(4\sqrt{\rho}\left(\mathrm{Re}\{Z\}\mathrm{Re}\{\hat{h}_\ell\}+\mathrm{Im}\{Z\}\mathrm{Im}\{\hat{h}_\ell\}\right)+|\tilde{\alpha}_\ell|^2-|\alpha|_\ell^2\right)}\right) \tag{100}$$

where $\alpha_\ell = h_\ell\sqrt{\rho} - \hat{h}_\ell\sqrt{\rho}$ and $\tilde{\alpha}_\ell = h_\ell\sqrt{\rho} + \hat{h}_\ell\sqrt{\rho}$. It is straightforward to show that $f_{3,\ell}(Z) = f_{0,\ell}(Z)$ as well. In other words, whenever the consecutive symbols are matching, they can be expressed by the same function. We next find $f_{1,\ell}(Z)$ as

$$f_{1,\ell}(Z) = \log\left(\frac{2}{1 + e^{-s\left(\left|h_\ell(\delta\sqrt{\rho}-\tilde{\delta}\sqrt{\rho})+\hat{h}_\ell\sqrt{\rho}+Z\right|^2-\left|h_\ell(\delta\sqrt{\rho}-\tilde{\delta}\sqrt{\rho})-\hat{h}_\ell\sqrt{\rho}+Z\right|^2\right)}}\right) \tag{101}$$

$$= \log\left(\frac{2}{1 + e^{-s\left(\left|\tilde{\lambda}_\ell+Z\right|^2-|\lambda_\ell+Z|^2\right)}}\right) \tag{102}$$

$$= \log\left(\frac{2}{1 + e^{s\left(4\sqrt{\rho}(\Re\{Z\}\Re\{\hat{h}_\ell\}+\Im\{Z\}\Im\{\hat{h}_\ell\})+\left|\tilde{\lambda}_\ell\right|^2-|\lambda_\ell|^2\right)}}\right) \tag{103}$$

$$= \log(2) - \log\left(1 + e^{s\left(4\sqrt{\rho}(\Re\{Z\}\Re\{\hat{h}_\ell\}+\Im\{Z\}\Im\{\hat{h}_\ell\})+\left|\tilde{\lambda}_\ell\right|^2-|\lambda_\ell|^2\right)}\right) \tag{104}$$

where $\lambda_\ell = h_\ell(\delta\sqrt{\rho} - \tilde{\delta}\sqrt{\rho}) - \hat{h}_\ell\sqrt{\rho}$ and $\tilde{\lambda}_\ell = h_\ell(\delta\sqrt{\rho} - \tilde{\delta}\sqrt{\rho}) + \hat{h}_\ell\sqrt{\rho}$. As before, $f_{2,\ell}(Z) = f_{1,\ell}(Z)$.

To find the density of (100) and (104), we let $\tilde{Z}_{\tilde{x},x} \sim \mathcal{N}(|\tilde{x}|^2 - |x|^2, 8\rho|\hat{h}_\ell|^2)$. Then $f_{0,\ell}(Z) = \tilde{f}_{0,\ell}(\tilde{Z}_{\tilde{\delta}_\ell,\delta_\ell})$ with

$$\tilde{f}_{0,\ell}(\tilde{Z}_{\tilde{\alpha}_\ell,\alpha_\ell}) = \log(2) - \log\left(1 + e^{s\tilde{Z}_{\tilde{\alpha}_\ell,\alpha_\ell}}\right) \tag{105}$$

and similarly $f_{1,\ell}(Z) = \tilde{f}_{1,\ell}(\tilde{Z}_{\tilde{\lambda}_\ell,\lambda_\ell})$ with

$$\tilde{f}_{1,\ell}(\tilde{Z}_{\tilde{\lambda}_\ell,\lambda_\ell}) = \log(2) - \log\left(1 + e^{s\tilde{Z}_{\tilde{\lambda}_\ell,\lambda_\ell}}\right). \tag{106}$$

We next find the density using change of variables as follows.

Lemma 1: Let $Z \sim \mathcal{N}(\mu, \sigma^2)$, $f(z)$ be an invertible function, $h(y) = f^{-1}(z)$ and the first derivative of $h(y)$ exists. Then, the pdf of $f(Z)$ can be found as

$$g_{f(Z)}(y) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{\frac{-(h(y)-\mu)^2}{2\sigma^2}} |h'(y)| \tag{107}$$

where $h'(y)$ is the first derivative of $h(y)$.

After applying Lemma 1, pdfs of $f_{0,\ell}(Z)$ and $f_{1,\ell}(Z)$ can be found as

$$g_{\tilde{f}_{0,\ell}(\tilde{Z}_{\tilde{\alpha}_\ell,\alpha_\ell})}(y) = \frac{1}{\sqrt{16\pi\rho\left|\hat{h_\ell}\right|^2}} \frac{1}{s(1 - e^{y-\log(2)})} e^{\frac{-\left(-\frac{\log(e^{-y+\log(2)}-1)}{s} - \left(|\tilde{\alpha}_\ell|^2 - |\alpha_\ell|^2\right)\right)^2}{16\rho|\hat{h}_\ell|^2}} \tag{108}$$

$$g_{\tilde{f}_{1,\ell}(\tilde{Z}_{\tilde{\lambda}_\ell,\lambda_\ell})}(y) = \frac{1}{\sqrt{16\pi\rho\left|\hat{h_\ell}\right|^2}} \frac{1}{s(1 - e^{y-\log(2)})} e^{\frac{-\left(-\frac{\log(e^{-y+\log(2)}-1)}{s} - \left(|\tilde{\lambda}_\ell|^2 - |\lambda_\ell|^2\right)\right)^2}{16\rho|\hat{h}_\ell|^2}} \tag{109}$$

respectively which concludes our proof.

REFERENCES

[1] G. Kolovou, S. Oteafy, and P. Chatzimisios, "A remote surgery use case for the IEEE p1918.1 tactile internet standard," in *Proc. IEEE Int. Conf. Commun. (ICC)*, Aug. 2021, pp. 1–6.

[2] E.-C. Liou and S.-C. Cheng, "A QoS benchmark system for telemedicine communication over 5G uRLLC and mMTC scenarios," in *Proc. IEEE 2nd Eurasia Conf. Biomed. Eng., Healthcare, Sustainability (ECBIOS)*, Sep. 2020, pp. 24–26.

[3] Q. Peng, H. Ren, C. Pan, N. Liu, and M. Elkashlan, "Resource allocation for uplink cell-free massive MIMO enabled URLLC in a smart factory," *IEEE Trans. Commun.*, vol. 71, no. 1, pp. 553–568, Nov. 2023.

[4] "Ericsson mobility report," Ericsson, Tech. Rep., Jun. 2023. [Online]. Available: https://www.ericsson.com/en/reports-and-papers/mobility-report/reports/june-2023

[5] B. S. Khan, S. Jangsher, A. Ahmed, and A. Al-Dweik, "URLLC and eMBB in 5G industrial iot: A survey," *IEEE Open J. Commun. Soc.*, vol. 3, pp. 1134–1163, Jul. 2022.

[6] "Study on physical layer enhancements for NR ultra-reliable and low latency case (URLLC) (release 16)," 3GPP, Tech. Rep., Mar. 2019. [Online]. Available: https://portal.3gpp.org/desktopmodules/Specifications/SpecificationDetails.aspx?specificationId=3498

[7] H. Tataria, M. Shafi, A. F. Molisch, M. Dohler, H. Sjöland, and F. Tufvesson, "6G wireless systems: Vision, requirements, challenges, insights, and opportunities," *Proc. IEEE*, vol. 109, no. 7, pp. 1166–1199, Mar. 2021.

[8] C. Yue, V. Miloslavskaya, M. Shirvanimoghaddam, B. Vucetic, and Y. Li, "Efficient decoders for short block length codes in 6G URLLC," *IEEE Commun. Mag.*, vol. 61, no. 4, pp. 84–90, 2023.

[9] G. Durisi, T. Koch, and P. Popovski, "Towards massive, ultra-reliable, and low-latency wireless communication with short packets," *Proc. IEEE*, vol. 104, no. 9, pp. 1711–1726, Sep. 2016.

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

[11] J. L. Jensen, *Saddlepoint approximations*. Oxford, U.K.: Oxford Univ. Press, 1995.

[12] A. Martinez and A. Guillén i Fàbregas, "Saddlepoint approximation of random-coding bounds," in *Proc. Inf. Theory Appl. Workshop*, San Diego, CA, USA, Feb. 2011, pp. 257–262.

[13] J. Östman, G. Durisi, E. G. Ström, M. C. Coşkun, and G. Liva, "Short packets over block-memoryless fading channels: Pilot-assisted or noncoherent transmission?" *IEEE Trans. Commun.*, vol. 67, no. 2, pp. 1521–1536, Feb. 2019.

[14] A. Lapidoth and S. Shamai (Shitz), "Fading channels: how perfect need "perfect side information" be?" *IEEE Trans. Inf. Theory*, vol. 48, no. 5, pp. 1118–1134, May 2002.

[15] A. O. Kislal, A. Lancho, G. Durisi, and E. Ström, "Efficient evaluation of the error probability for pilot-assisted finite-blocklength transmission," in *Proc. Asilomar Conf. Signals, Syst., Comput*, Pacific Grove, CA, U.S.A., Nov. 2022.

[16] A. O. Kislal, A. Lancho, G. Durisi, and E. G. Ström, "Efficient evaluation of the error probability for pilot-assisted URLLC with massive MIMO," *IEEE J. Sel. Areas Commun.*, vol. 41, no. 7, pp. 1969–1981, May 2023.

[17] A. Anand and G. de Veciana, "Resource allocation and HARQ optimization for URLLC traffic in 5G wireless networks," *IEEE J. Sel. Areas Commun.*, vol. 36, no. 11, pp. 2411–2421, Oct. 2018.

[18] W. R. Ghanem, V. Jamali, M. Schellmann, H. Cao, J. Eichinger, and R. Schober, "Codebook based two-time scale resource allocation design for irs-assisted eMBB-URLLC systems," in *2022 IEEE Globecom Workshops (GC Wkshps)*, Dec. 2022, pp. 419–425.

[19] M. Darabi, V. Jamali, L. Lampe, and R. Schober, "Hybrid puncturing and superposition scheme for joint scheduling of URLLC and eMBB traffic," *IEEE Commun. Lett.*, vol. 26, no. 5, pp. 1081–1085, Feb. 2022.

[20] Q. Chen, J. Wu, J. Wang, and H. Jiang, "Coexistence of URLLC and eMBB services in MIMO-NOMA systems," *IEEE Trans. Veh. Technol.*, vol. 72, no. 1, pp. 839–851, Sep. 2023.

[21] A. Lancho, J. Östman, G. Durisi, T. Koch, and G. Vazquez-Vilar, "Saddlepoint approximations for short-packet wireless communications," *IEEE Trans. Wireless Commun.*, vol. 19, no. 7, pp. 4831–4846, Jul. 2020.

[22] J. Östman, A. Lancho, G. Durisi, and L. Sanguinetti, "URLLC with massive MIMO: Analysis and design at finite blocklength," *IEEE Trans. Wireless Commun.*, vol. 20, no. 10, pp. 6387–6401, Oct. 2021.

[23] J. Font-Segura, G. Vazquez-Vilar, A. Martinez, A. Guillén i Fàbregas, and A. Lancho, "Saddlepoint approximations of lower and upper bounds to the error probability in channel coding," in *Proc. Conf. Inf. Sci. Sys. (CISS)*, Princeton, NJ, Mar. 2018.

[24] G. C. Ferrante, J. Ostman, G. Durisi, and K. Kittichokechai, "Pilot-assisted short-packet transmission over multiantenna fading channels: A 5G case study," in *52nd Conf. on Inf. Sci. and Sys. (CISS)*, Mar. 2018, pp. 1–6.

[25] W. Feller, *An introduction to probability theory and its applications*. Wiley, 1971, vol. 2.

[26] J. Scarlett, A. Martinez, and A. Guillén i Fàbregas, "Mismatched decoding: Error exponents, second-order rates and saddlepoint approximations," *IEEE Trans. Inf. Theory*, vol. 60, no. 5, pp. 2647–2666, May 2014.

[27] S. M. Kay, *Fundamentals of Statistical Signal Processing, Volume I: Estimation Theory*. Prentice-Hall, 1993.