Is Synchronization a Bottleneck for Pilot-Assisted URLLC Links?

N.B. When citing this work, cite the original published paper.

(article starts on next page)

Is Synchronization a Bottleneck for Pilot-Assisted URLLC Links?

A. Oguz Kislal, Madhavi Rajiv, Giuseppe Durisi, *Senior Member, IEEE*, Erik G. Ström, *Fellow, IEEE*, Urbashi Mitra, *Fellow, IEEE*

Abstract—We propose a framework to evaluate the so-called random-coding union bound with parameter s (RCUs) on the achievable error probability in the finite-blocklength regime for a pilot-assisted transmission scheme operating over an imperfectly synchronized and memoryless block-fading waveform channel. Unlike previous results, which disregard the effects of imperfect synchronization, our framework utilizes pilots for both synchronization and channel estimation. Specifically, we provide an algorithm to perform joint synchronization and channel estimation, and verify its accuracy by observing its tightness in comparison with the Cramer-Rao bound. Then, we develop an RCUs bound on the error probability, which applies for a receiver that treats the estimates provided by the algorithm as accurate. Additionally, we utilize the saddlepoint approximation to provide a numerically efficient method for evaluating the RCUs bound in this scenario. Our numerical experiments verify the accuracy of the proposed approximation. Moreover, when the delays are modeled as fully dependent across fading blocks, numerical results indicate that the number of pilot symbols needed to estimate the fading channel gains to the level of accuracy required in ultra-reliable low-latency communication is also sufficient to acquire sufficiently good synchronization. However, when the delays are modeled as independent across blocks, synchronization becomes the bottleneck for the system performance.

Index Terms—URLLC, pilot-assisted transmission, synchronization, channel estimation, finite-blocklength information theory

I. INTRODUCTION

Advancements in technology have allowed for the rise of applications like remote surgery [2], [3], factory automation [4], and autonomous driving that significantly impact various aspects of people's lives. These types of applications create substantial data traffic that future cellular communication technologies will need to accommodate. 5G has the goal to support the diverse requirements of these technologies.

The scenario considered in this paper, denoted as *ultra-reliable low-latency communication* (URLLC), is designed for mission-critical applications targeting 99.999% reliability with

A. Oguz Kislal, Giuseppe Durisi, and Erik. G. Ström are with the Department of Electrical Engineering, Chalmers University of Technology, Gothenburg 41296, Sweden (e-mail: {kislal,durisi,erik.strom}@chalmers.se). Madhavi Rajiv and Urbashi Mitra are with the Ming Hsieh Department of Electrical Engineering, University of Southern California, Los Angeles CA 90089, USA (email: {rajiv,ubli}@usc.edu).

end-to-end latency as low as 1 ms [5]. For 6G, the targets will be even more aggressive: $10^{-5} - 10^{-7}$ packet error rate with a more stringent latency constraint [6]–[8].

A key feature of URLLC traffic is the frequent use of small information payloads, which are transmitted in short packets consisting of a limited number of encoded symbols representing the payload data. In URLLC, there are limitations on the signal duration and available bandwidth, due to latency requirements and the need to orthogonalize multiple-user transmissions to mitigate multi-user interference. As a consequence, 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]. Because of its relevance for URLLC, the field of finite-blocklength information theory has been studied extensively, particularly following the seminal work [10], which offers a precise understanding of the tradeoff between error probability and packet size, for a given SNR and transmission rate, when operating with finite blocklengths.

In this paper, we focus on communication over memoryless block-fading channels with imperfect synchronization,[1] with the goal of understanding the impact of synchronization accuracy on the overall system performance. Benchmarking URLLC systems with such goals often requires the use of approximation techniques due to the computational expense of evaluating finite blocklength bounds exactly. Such approximation techniques are critical to enable the use of these bounds within URLLC optimization routines such as resource-allocation [11], [12] and scheduling algorithms [13], [14].

The computationally efficient approximation proposed in this paper is based on the saddlepoint method [15], which is used to evaluate the random coding union bound with parameter s (RCUs) proposed in [16]. This bound is well-suited for communication over fading channels, as it can be applied to both the optimal non-coherent maximum-likelihood (ML) decoder and more practically significant decoders that utilize pilot-assisted transmission (PAT), see, e.g., [17]. Specifically, the decoder considered in this paper is the so-called scaled nearest-neighbor (SNN) decoder [18], which minimizes the error probability when synchronization and channel estimation are perfect, but is suboptimal otherwise. The evaluation of the RCUs bound for URLLC scenarios can be computationally expensive, and, for some applications, such bounds need to be approximated by methods similar to the one proposed in [19].

---

[1] Throughout this paper, by synchronization, we mean time synchronization, or, more specifically, the estimation of the integer and fractional delays of the samples at the output of the upsampled receiver matched filter.

Fig. 1: Block diagram of the system model. In the case of fully dependent delays, $D_1 = \ldots = D_{n_\mathrm{b}}$.

Most existing approaches for benchmarking URLLC systems disregard synchronization errors. In these approaches, an upper bound on the error probability is obtained by evaluating the tail probability of a sum of independent random variables. This allows for approximations such as the *normal approximation* [10], or the more accurate *saddlepoint approximation* [15], [20]. The saddlepoint approximation has been applied to the case of an optimal ML decoder in [21], of PAT transmission and SNN decoding over single-input single-output (SISO) and multiple-input multiple-output (MIMO) channels in [22], and of PAT transmission and SNN decoding over massive MIMO for both single and multiple fading blocks in [17], [19], [23]. In [24], the problem of joint synchronization and decoding is analyzed from a finite-blocklength perspective. There, it is shown that one can achieve performance close to the perfectly synchronized case even for an exponentially large asynchronism level. However, asynchronism is modeled only at the symbol level, and imperfect synchronization results only in an incorrect estimate of the location of the transmitted codeword.

In this paper, we consider the more practically relevant case in which asynchronism is at the waveform channel, and imperfect synchronization yields intersymbol interference after matched filtering and sampling. As we shall see, this implies that the random variables that need to be analyzed to obtain upper bounds on the error probability are, in general, dependent, and the saddlepoint approximations provided in [17], [19], [22], [23] are therefore not applicable. However, under certain assumptions on the dependence between the random variables, saddlepoint approximations can still be derived as we do herein.

Contributions: We consider the problem of transmitting short packets over a SISO memoryless block-fading waveform channel with unknown delays, and analyze the impact of imperfect synchronization and channel estimation on system performance in the URLLC regime. Specifically, we focus on the extreme cases of delays being independent across all fading blocks, or delays being fully dependent across fading blocks, and design a synchronization and channel estimation algorithm that can be used in both scenarios. We verify the accuracy of the proposed algorithms by presenting a comparison with the Cramer-Rao bound (CRB). We then derive an RCUs bound for the case of PAT and SNN decoding to evaluate the error probability achievable by a decoder that treats the estimates for the delay and the channel gains provided by our algorithm as perfect. Finally, we provide a novel saddlepoint approximation for this bound based on [15, Ch. 6]. The novelty of the proposed approximation is that it accounts for the specific dependence between random variables that arises due to imperfect synchronization. Finally, we provide numerical results to illustrate the accuracy of the proposed approximation and its utility in obtaining insights into the design of URLLC links.

In this paper, we extend the results in [1] by providing a complete derivation of the upsampled output of the matched filter at the receiver. We also explain how to evaluate in closed form the moment-generating function appearing in the saddle-point approx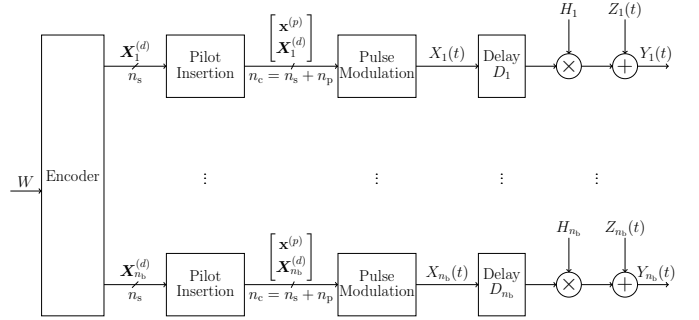imation and provide a concrete example for the case of BPSK transmission. Finally, with extensive numerical experiments, we clarify, among other aspects, the optimal number of pilot symbols to be used for synchronization and channel estimation.

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., $\mathbf{Y}$), and use the notation $[\mathbf{Y}]_{ij}$ to refer to the element on the ith row and jth column of the matrix $\mathbf{Y}$. 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, $Q(\cdot)$ for the Gaussian Q-function, $\mathbb{1}\{\cdot\}$ for the indicator function, $\mathbf{1}_u$ for the all-one column vector with a length u, and $\mathbf{I}_m$ for the $m \times m$ identity matrix, respectively. Finally, for two functions $f(n)$ and $g(n)$, the notation $f(n) = o(g(n))$ means that $\lim_{n \to \infty} f(n)/g(n) = 0$ and the notation $f(n) = \mathcal{O}(g(n))$ means that $\limsup_{n \to \infty} |f(n)/g(n)| < \infty$.

II. SYSTEM MODEL

A. Overview

We consider pilot-assisted transmission of a uniformly distributed message over a SISO block-flat-fading channel with unknown delay. We assume that the channel stays constant over the transmission of a block of n_c channel uses and that it changes independently across blocks consisting of non-overlapping channel uses. Independence may be achieved by spacing the blocks sufficiently in frequency or time. Each transmitted packet spans n_b such fading blocks. The setup is illustrated by the block diagram in Fig. 1, in which each branch corresponds to the transmission over one fading block. The encoder maps the realization of the random message W

to a complex-valued codeword of length $n_b n_s$, which, in turn, is split into n_b subcodewords $\{\mathbf{x}_\ell^{(d)}\}_{\ell=1}^{n_b}$ of length n_s

$$\mathbf{x}_\ell^{(d)} = \begin{bmatrix} x_{1,\ell}^{(d)} & \cdots & x_{n_s,\ell}^{(d)} \end{bmatrix}^{\mathrm{T}} \in \mathbb{C}^{n_s}, \tag{1}$$

where $\ell = 1, \ldots, n_b$ and n_b is the number of fading blocks used for the transmission of the message. A pilot sequence

$$\mathbf{x}^{(p)} = \begin{bmatrix} x_1^{(p)} & \cdots & x_{n_p}^{(p)} \end{bmatrix}^{\mathrm{T}} \in \mathbb{C}^{n_p} \tag{2}$$

is prepended to each codeword, to form n_b subpackets of length $n_c = n_p + n_s$, which are transmitted over different channel fading blocks and experience potentially different propagation delays. We assume that W is drawn uniformly from the set $\{1, 2, \ldots, \lceil \exp(n_b n_c \mathrm{R}) \rceil\}$ and that $\lceil \exp(n_b n_c \mathrm{R}) \rceil$ is the number of codewords in the codebook. As a consequence, R is the transmission rate in nats per symbol.

We assume that the data and pilot symbols are subject to the power constraint $\mathbb{E}[\|\boldsymbol{X}_\ell^{(d)}\|^2] = n_s\rho$ and $\|\mathbf{x}^{(p)}\|^2 = n_p\rho$, respectively.

In the following subsections, we will define the pulse shaping, introduce the propagation delays in the waveform domain, and describe the processing at the receiver considered in this paper.

B. Signal Model

The n_p pilot symbols within each subpacket are used to form the continuous-time pilot signal $x^{(p)}(t)$ defined as

$$x^{(p)}(t) = \sum_{k=1}^{n_p} x_k^{(p)} s_{t_p}(t - (k-1)t_p), \tag{3}$$

where $s_{t_p}(t)$ is a rectangular pulse[2] with normalized energy and support of size t_p, i.e.,

$$s_{t_p}(t) = \begin{cases} \frac{1}{\sqrt{t_p}}, & t \in [0, t_p) \\ 0, & \text{otherwise .} \end{cases} \tag{4}$$

The data symbols in the ℓth subpacket are sent after the pilot symbols via the continuous-time signal

$$x_\ell^{(d)}(t) = \sum_{k=1}^{n_s} x_{k,\ell}^{(d)} s_{t_p}(t - (k-1)t_p - n_p t_p). \tag{5}$$

The total continuous-time signal corresponding to the ℓth subpacket is put through a flat-fading channel to obtain the received continuous signal

$$Y_\ell(t) = H_\ell\Big(x^{(p)}(t - D_\ell) + x_\ell^{(d)}(t - D_\ell)\Big) + Z_\ell(t), \tag{6}$$

where H_ℓ denotes the scalar random fading complex channel gain for the ℓth fading block, D_ℓ is the time delay for the ℓth fading block, which we assume to be uniform in $[0, d_{\max}]$, and $Z_1(t), \ldots, Z_{n_b}(t)$ are independent white complex Gaussian processes with power spectral density N_0. For simplicity, we set $N_0 = 1$.

[2]We consider rectangular pulses for simplicity. Our analysis can be generalized to other pulse shapes by using the corresponding matched filter in (9).

We highlight that other types of asynchronism beyond the one modeled in (6), including clock and carrier asynchronism may occur in a wireless communication system. These other forms of asynchronism are not captured directly by (6) (although impairments such as carrier-phase offset can be modeled by modifying the statistics of the fading coefficients H_ℓ), and their impact on performance will not be discussed in the present paper.

Depending on the type of diversity being used, the delays D_ℓ in (6) may be dependent across ℓ. For simplicity, we focus on the two extreme cases of delays being independent across fading blocks, where $D_1, \ldots, D_{n_b}$ are i.i.d. and the case of delays being fully dependent across fading blocks, where $D_1 = \cdots = D_{n_b} = D$. In the former case, information about the pair (D_ℓ, H_ℓ) is found only in $Y_\ell(t)$, and the estimation problem decouples into n_b separate problems, one per block. However, in the case of fully dependent delays, it will be advantageous to jointly estimate $(D, H_1, \ldots, H_{n_b})$. We next introduce synchronization and channel estimation algorithms for both cases.

C. Synchronization and Channel Estimation Phase

The receiver uses the knowledge of the pilot sequence to estimate the propagation delays and the fading-block gains. Throughout, we assume that the receiver employs synchronization and channel estimation algorithms that take an upsampled version of the received signal as the input. For a given upsampling rate N, let t_s be the sampling interval; then, $t_p = t_s \mathrm{N}$ is the period of the pulses used to construct the continuous time pilot and data signals. We also let $x_{\mathrm{N},n}^{(p)} = x_{\lceil n/\mathrm{N} \rceil}^{(p)}$ be the nth element of the upsampled vector of pilot symbols, where $n = 1, \ldots, \mathrm{N}n_p$. In order to obtain an upsampled received signal, we process and sample the received signal $Y_\ell(t)$ as

$$Y_{m,\ell}^{(p)} = (Y_\ell * s_{t_s})(mt_s), \tag{7}$$

where $m = 1, \ldots, \mathrm{M}$, and the sampling endpoint is chosen to capture all pilot symbols as

$$\mathrm{M} = \lceil d_{\max}/t_s \rceil + n_p \mathrm{N}. \tag{8}$$

Let $\boldsymbol{Y}_\ell^{(p)} = [Y_{1,\ell}^{(p)}, \ldots, Y_{\mathrm{M},\ell}^{(p)}]^T$. To express $\boldsymbol{Y}_\ell^{(p)}$ compactly, we shall first isolate the contribution to $\boldsymbol{Y}_\ell^{(p)}$ of the pilot signal $x^{(p)}(t)$ in (3). In view of (7), we start by noting that the convolution between $x^{(p)}(t)$ and $s_{t_s}(t)$ results in the signal

$$(x^{(p)} * s_{t_s})(t) = \sum_{n=1}^{\mathrm{N}n_p} x_{\mathrm{N},n}^{(p)} r_{t_s}(t - (n-1)t_s), \tag{9}$$

where $r_{t_s}(t)$ is a triangular pulse with duration $2t_s$ and unit peak amplitude

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

Thus, for any $e \in [0, t_s)$

$$(x^{(p)} * s_{t_s})(mt_s + e) = \left(1 - \frac{e}{t_s}\right)x_{\mathrm{N},m}^{(p)} + \frac{e}{t_s}x_{\mathrm{N},m+1}^{(p)}. \tag{11}$$

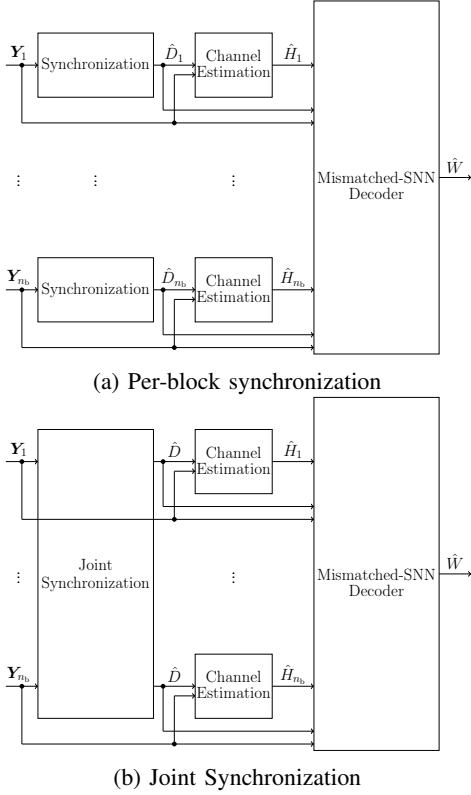(a) Per-block synchronization

(b) Joint Synchronization

Fig. 2: Block diagrams of the receiver.

Let now $Q_\ell = \lfloor D_\ell/t_s \rceil$ and $E_\ell = D_\ell - Q_\ell t_s$ so that $D_\ell = Q_\ell t_s + E_\ell$ with $Q_\ell \in \mathbb{Z}$ and $E_\ell \in [0, t_s)$. It follows from (11) that we can represent $\boldsymbol{Y}_\ell^{(p)}$ as

$$
\boldsymbol{Y}_\ell^{(p)} = H_\ell \begin{bmatrix} \mathbf{x}_N^{(p)}(Q_\ell) & \mathbf{x}_N^{(p)}(Q_\ell+1) \end{bmatrix} \begin{bmatrix} 1 - \frac{E_\ell}{t_s} \\ \frac{E_\ell}{t_s} \end{bmatrix}
$$
$$
+ \boldsymbol{Z}_\ell + \boldsymbol{C}_\ell, \tag{12}
$$

where

$$
\mathbf{x}_N^{(p)}(q_\ell) = \frac{1}{\sqrt{N}} \begin{bmatrix} \mathbf{0}_{q_\ell}^T, x_{N,1}^{(p)}, \dots, x_{N,Nn_p}^{(p)}, \mathbf{0}_{M-q_\ell-Nn_p}^T \end{bmatrix}^T, \tag{13}
$$

$\boldsymbol{Z}_\ell \sim \mathcal{CN}(\mathbf{0}_M, \mathbf{I}_M)$ is the Gaussian vector containing the sampled noise process, and $\boldsymbol{C}_\ell$ captures the potential interference from the data signal (5).

For simplicity, our algorithms are derived by considering the ML estimators of the parameters in question in the case where $\boldsymbol{C}_\ell = \mathbf{0}_M$, i.e., where no interference from the data symbols is present. As we shall verify in Section V, our assumption incurs minimal loss.

1) Per-Block Synchronization for Independent Delays: In this case, as illustrated in Fig. 2a, the receiver uses the observation of $\boldsymbol{Y}_\ell^{(p)}$ for each fading block in order to obtain the estimates $\hat{H}_\ell$, $\hat{Q}_\ell$, and $\hat{E}_\ell$. To do so, we pose the following minimization problem:

$$
[\hat{H}_\ell, \hat{Q}_\ell, \hat{E}_\ell] = \underset{\bar{h}, \bar{q}, \bar{e}}{\arg \min} \| \boldsymbol{Y}_\ell^{(p)} - \bar{h} \mathbf{v}(\bar{q}, \bar{e}) \|^2, \tag{14}
$$

where

$$
\mathbf{v}(\bar{q}, \bar{e}) = \begin{bmatrix} \mathbf{x}_N^{(p)}(\bar{q}) & \mathbf{x}_N^{(p)}(\bar{q}+1) \end{bmatrix} \begin{bmatrix} 1 - \frac{\bar{e}}{t_s} \\ \frac{\bar{e}}{t_s} \end{bmatrix}. \tag{15}
$$

For any given $\boldsymbol{Y}_\ell^{(p)} = \mathbf{y}_\ell^{(p)}$, the ML channel estimate for each fading block $\bar{h}_\ell$ can be written as a function of $\bar{q}$ and $\bar{e}$ as

$$
\hat{h}_\ell(\bar{q}, \bar{e}) = \underset{\bar{h}_\ell}{\arg \min} \sum_{k=1}^{n_b} \| \mathbf{y}_k^{(p)} - \mathbf{v}(\bar{q}, \bar{e}) \bar{h}_k \|^2 \tag{16}
$$
$$
= \underset{\bar{h}_\ell}{\arg \min} \| \mathbf{y}_\ell^{(p)} - \mathbf{v}(\bar{q}, \bar{e}) \bar{h}_\ell \|^2 \tag{17}
$$
$$
= \frac{\mathbf{v}(\bar{q}, \bar{e})^H \mathbf{y}_\ell^{(p)}}{\|\mathbf{v}(\bar{q}, \bar{e})\|^2}. \tag{18}
$$

Substituting (18) into (14), we can estimate the delay parameters as

$$
[\hat{q}_\ell, \hat{e}_\ell] = \underset{\bar{q}, \bar{e}}{\arg \min} \| \mathbf{y}_\ell^{(p)} - \mathbf{v}(\bar{q}, \bar{e}) \hat{h}_\ell(\bar{q}, \bar{e}) \|^2 \tag{19}
$$
$$
= \underset{\bar{q}, \bar{e}}{\arg \min} \| \mathbf{y}_\ell^{(p)} \|^2 - 2 \operatorname{Re} \left\{ (\mathbf{y}_\ell^{(p)})^H \mathbf{v}(\bar{q}, \bar{e}) \hat{h}_\ell(\bar{q}, \bar{e}) \right\}
$$
$$
+ \| \mathbf{v}(\bar{q}, \bar{e}) \|^2 \left| \hat{h}_\ell(\bar{q}, \bar{e}) \right|^2 \tag{20}
$$
$$
= \underset{\bar{q}, \bar{e}}{\arg \min} -2 \frac{\left| \mathbf{v}(\bar{q}, \bar{e})^H \mathbf{y}_\ell^{(p)} \right|^2}{\| \mathbf{v}(\bar{q}, \bar{e}) \|^2} + \frac{\left| \mathbf{v}(\bar{q}, \bar{e})^H \mathbf{y}_\ell^{(p)} \right|^2}{\| \mathbf{v}(\bar{q}, \bar{e}) \|^2} \tag{21}
$$
$$
= \underset{\bar{q}, \bar{e}}{\arg \max} \frac{\left| \mathbf{v}(\bar{q}, \bar{e})^H \mathbf{y}_\ell^{(p)} \right|^2}{\| \mathbf{v}(\bar{q}, \bar{e}) \|^2} \tag{22}
$$
$$
= \underset{\bar{q}, \bar{e}}{\arg \max} \frac{n(\bar{q}, \bar{e})}{d(\bar{q}, \bar{e})}, \tag{23}
$$

where we set

$$
n(\bar{q}, \bar{e}) = \left| \mathbf{v}(\bar{q}, \bar{e})^H \mathbf{y}_\ell^{(p)} \right|^2, \tag{24}
$$
$$
d(\bar{q}, \bar{e}) = \| \mathbf{v}(\bar{q}, \bar{e}) \|^2. \tag{25}
$$

For a fixed value of $\bar{q}$, both $n(\bar{q}, \bar{e})$ and $d(\bar{q}, \bar{e})$ are second-degree polynomials in $\bar{e}$, which implies that the objective function in (23) is a rational function in $\bar{e}$. For a fixed value of $\bar{q}$, we find an extreme point of this rational function by differentiating with respect to $\bar{e}$ and setting the derivative to zero as

$$
\frac{\partial}{\partial \bar{e}} \frac{n(\bar{q}, \bar{e})}{d(\bar{q}, \bar{e})} = \frac{n'(\bar{q}, \bar{e}) d(\bar{q}, \bar{e}) - n(\bar{q}, \bar{e}) d'(\bar{q}, \bar{e})}{d(\bar{q}, \bar{e})^2} = 0. \tag{26}
$$

If (26) has a solution $\bar{e}^\star$ in the range $(0, t_s)$, then $(\bar{q}, \bar{e}^\star)$ is a candidate for $(\hat{q}_\ell, \hat{e}_\ell)$. We also consider the boundary points $(\bar{q}, 0)$ and $(\bar{q}, t_s)$ as candidates, since they might be the solution of (26) in the cases where no extreme point can be found in $(0, t_s)$, or when a minimum point (instead of a maximum) is found.

We are now ready to introduce our delay and channel estimation algorithm. Let $\mathcal{D}_\ell$ be the set of candidates for $\hat{d}_\ell = \hat{q}_\ell t_s + \hat{e}_\ell$. We construct $\mathcal{D}_\ell$ as follows: for each $\bar{q} \in \{0, \dots, d_{\max}/t_s\}$ we add $\bar{q} t_s$ to $\mathcal{D}_\ell$. Next, we find the solutions $\bar{e}^\star$ for (26) for each $\bar{q}$. If $\bar{e}^\star$ is in the range $[0, t_s)$, we add $\bar{q} t_s + \bar{e}^\star$ to $\mathcal{D}_\ell$. We then find $\hat{d}_\ell$ as the entry in $\mathcal{D}_\ell$

Algorithm 1 Synchronization and Channel Estimation

Input: $t_\mathrm{s}, d_\mathrm{max}, \mathbf{y}_1^{(p)}, \mathbf{y}_2^{(p)}, \ldots, \mathbf{y}_{n_\mathrm{b}}^{(p)}$
Provide: A solution $\hat{q}, \hat{e}, \hat{h}_1, \hat{h}_2, \ldots, \hat{h}_{n_\mathrm{b}}$ to (14)
  for $\ell = 1, 2, \ldots, n_\mathrm{b}$ **do**
    $\mathcal{D}_\ell \leftarrow \{(q, 0) : q = 0, 1, \ldots, \lceil d_\mathrm{max}/t_\mathrm{s} \rceil\}$
        $\cup \{(q, t_\mathrm{s}) : q = 0, 1, \ldots, \lceil d_\mathrm{max}/t_\mathrm{s} \rceil\}$
    for $\bar{q} \in \{0, 1, \ldots, \lceil d_\mathrm{max}/t_\mathrm{s} \rceil\}$ **do**
      for each root $\bar{e}^*$ to (26) such that $\bar{e}^* \in (0, t_s)$ **do**
        add $(\bar{q}, \bar{e}^*)$ to $\mathcal{D}_\ell$
      end for
    end for
    $(\hat{q}_\ell, \hat{e}_\ell) \leftarrow$ element in $\mathcal{D}_\ell$ that maximizes (23)
    $\hat{h}_\ell \leftarrow \mathbf{v}(\hat{q}, \hat{e})^H \mathbf{y}_\ell^{(p)} / \|\mathbf{v}(\hat{q}, \hat{e})\|^2$
  end for

that maximizes the objective function in (23). This estimate is then used to evaluate the channel estimate $\hat{h}_\ell$ from (18). A pseudocode for the algorithm is given in Algorithm 1.

2) Joint Synchronization for Fully Dependent Delays: In this case, since the delay is same for all fading blocks (i.e., $D_\ell = D$ for all ℓ), the receiver can use the observations of $\{\boldsymbol{Y}_\ell^{(p)}\}_{\ell=1}^{n_\mathrm{b}}$ jointly, as illustrated in Fig. 2b. These observations are used to estimate the channel gains $\hat{\boldsymbol{H}} = [\hat{H}_1, \ldots, \hat{H}_{n_\mathrm{b}}]^T$ and delay as

$$[\hat{\boldsymbol{H}}, \hat{Q}, \hat{E}] = \arg \min_{\bar{\mathbf{h}}, \bar{q}, \bar{e}} \sum_{\ell=1}^{n_\mathrm{b}} \|\boldsymbol{Y}_\ell^{(p)} - \bar{h}_\ell \mathbf{v}(\bar{q}, \bar{e})\|^2. \quad (27)$$

Following the same steps leading to (23), we obtain

$$[\hat{q}, \hat{e}] = \arg \max_{\bar{q}, \bar{e}} \sum_{\ell=1}^{n_\mathrm{b}} \frac{|\mathbf{v}(\bar{q}, \bar{e})^H \mathbf{y}_\ell|^2}{\|\mathbf{v}(\bar{q}, \bar{e})\|^2}. \quad (28)$$

The numerator and denominator of (28) have the same polynomial structure as (23), and the same synchronization and channel estimation algorithm described in Section II-C1 can be used, but now with only one set of candidates $\mathcal{D}$.

D. Codeword Decoding Phase

The codeword decoding phase is based on a mismatch-decoding approach, where the delay and channel estimates returned by the algorithms described in Section II-C are treated as perfect. The input-output relationship for the kth symbol in the ℓth block is[3]

$$Y_{k,\ell} = \left(Y_\ell * s_{t_p}\right)(k t_p + n_p t_p + \hat{D}_\ell), \quad k = 1, 2, \ldots, n_\mathrm{s}. \quad (29)$$

In the decoding process, the receiver seeks the codeword in the codebook $\mathcal{C}$ closest to the received signal after scaling each subcodeword with the corresponding channel estimates. Hence, given the received vector and the channel estimates, the decoded codeword $\hat{\mathbf{x}} = [\hat{\mathbf{x}}_1^T, \ldots, \hat{\mathbf{x}}_{n_\mathrm{b}}^T]^T$ is determined as

$$\hat{\mathbf{x}} = \arg \min_{\substack{\bar{\mathbf{x}} = [\bar{\mathbf{x}}_1^T, \ldots, \bar{\mathbf{x}}_{n_\mathrm{b}}^T]^T \\ \bar{\mathbf{x}} \in \mathcal{C}}} \sum_{\ell=1}^{n_\mathrm{b}} \|\mathbf{y}_\ell - \hat{h}_\ell \bar{\mathbf{x}}_\ell\|^2, \quad (30)$$

[3]We will omit the superscript $^{(d)}$ in the remainder of the paper to keep the notation compact.

where $\mathbf{y}_\ell = [y_{1,\ell}, \ldots, y_{n_\mathrm{s},\ell}]^T$. This decoder, known as the *mismatched SNN decoder* [25], coincides with the ML decoder only when the receiver has perfect channel state information, i.e., $\hat{h}_\ell = h_\ell$ for $\ell = 1, \ldots, n_\mathrm{b}$, and perfect synchronization. This decoder, although not optimal, is practically relevant and the analysis of its finite-blocklength error probability is tractable [17].

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

Let $\hat{W}$ be the estimate, produced by the decoder, of the transmitted message W. We define the packet error probability as $\epsilon_\mathrm{pep} = \mathbb{P}[\hat{W} \neq W]$. We can express ϵ_pep as

$$\epsilon_\mathrm{pep} = \mathbb{P}\left[\left|\hat{D}_\ell - D_\ell\right| \leq t_\mathrm{p}\right] \epsilon_1 + \mathbb{P}\left[\left|\hat{D}_\ell - D_\ell\right| > t_\mathrm{p}\right] \epsilon_2, \quad (31)$$

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 evaluating ϵ_pep, we will assume, for simplicity, that the decoder cannot decode the packet when synchronization is off more than one symbol, i.e., we will upper-bound ϵ_pep in (31) by setting $\epsilon_2 = 1$.

In the next section, we will present an 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 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 $\mathcal{U}$ with u elements (e.g., $u = 2$ for BPSK) and power ρ. Although potentially suboptimal, this choice is practically relevant and allows us to evaluate the RCUs bound efficiently via a saddlepoint approximation.

When the synchronization is off by less than one symbol (i.e., $|\hat{D}_\ell - D_\ell| \leq t_\mathrm{p}$), the input-output relation in (29) can be stated as

$$Y_{k,\ell} = H_\ell(\Delta_\ell X_{k,\ell} + (1 - \Delta_\ell) X_{k+\Lambda_\ell,\ell}) + Z_{k,\ell}, \quad (32)$$

where we define

$$\Delta_\ell = 1 - \left|\hat{D}_\ell - D_\ell\right|/t_\mathrm{p} \quad (33)$$

$$\Lambda_\ell = \mathrm{sign}(\hat{D}_\ell - D_\ell), \quad (34)$$

and $Z_{k,\ell}$ are i.i.d. zero-mean, unit-variance, complex Gaussian random variables. Note that the lack of perfect synchronization, i.e., $\Delta_\ell \neq 1$, yields intersymbol interference. In (32), the data symbol $X_{k,\ell}$ and the intersymbol interference $X_{k+\Lambda_\ell,\ell}$ are scaled by Δ_ℓ and $1 - \Delta_\ell$, respectively, both of which are independent of Λ_ℓ. Since data symbols are also independent of each other, Λ_ℓ has an impact on the error probability only through whether the last symbol of a subcodeword i.e. $X_{n_\mathrm{s},\ell}$

is affected by intersymbol interference or not.[4] Specifically, when $\Lambda_\ell = -1$, $X_{n_s-1,\ell}$ interferes with $X_{n_s,\ell}$. Conversely, when $\Lambda_\ell = 1$, $X_{n_s,\ell}$ is free from intersymbol interference. Even if the impact of Λ_ℓ on the packet error probability is expected to be minute, we assume the worst case here and let $\Lambda_\ell = -1$ for the remainder of the paper. By choosing the worst-case scenario for our analysis, we ensure that the achievability bound we shall present next still holds as an achievability bound.

For this setup, the RCUs achievability bound ϵ_{ub} on ϵ_1 is given by [16]

$$\epsilon_1 \leq \epsilon_{\mathrm{ub}} \qquad (35)$$

with

$$\epsilon_{\mathrm{ub}} = \mathbb{E}_{\boldsymbol{H},\hat{\boldsymbol{H}},\boldsymbol{\Delta}}\Big[\epsilon_{\mathrm{ub}}\big(\boldsymbol{H},\hat{\boldsymbol{H}},\boldsymbol{\Delta}\big)\Big] \qquad (36)$$

where $\boldsymbol{H} = [H_1,\ldots,H_{n_b}]^T$, $\boldsymbol{\Delta} = [\Delta_1,\ldots,\Delta_{n_b}]^T$, and

$$\epsilon_{\mathrm{ub}}\big(\mathbf{h},\hat{\mathbf{h}},\boldsymbol{\delta}\big) = \mathbb{P}\Bigg[\frac{\log\Upsilon}{n_c n_b} + \frac{1}{n_c n_b}\sum_{\ell=1}^{n_b}\sum_{k=1}^{n_s}\imath_s\Big(X_{k,\ell};Y_{k,\ell},\hat{h}_\ell\Big)$$
$$\leq \mathrm{R}\Big|\boldsymbol{H}=\mathbf{h},\hat{\boldsymbol{H}}=\hat{\mathbf{h}},\boldsymbol{\Delta}=\boldsymbol{\delta}\Bigg]. \quad (37)$$

Here, Υ is a random variable that is 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 [16]

$$\imath_s(x;y,\hat{h}) = \log\frac{e^{-s|y-\hat{h}x|^2}}{\mathbb{E}_{\bar{X}}\Big[e^{-s|y-\hat{h}\bar{X}|^2}\Big]}, \qquad (38)$$

where $\bar{X}$ is independent of all other random variables and is drawn uniformly from the constellation and $s > 0$ is a parameter that can be optimized to obtain a tighter bound. Note that (37) can be evaluated (at least numerically) for any realization of Δ, H and $\hat{H}$, regardless of their underlying distribution. Consequently, the RCUs bound can be evaluated for all channel and delay estimation algorithms.

No closed-form expression for the RCUs bound (37) is in general available. Numerical methods to evaluate it, such as Monte-Carlo simulations, can be time consuming due to the low target error probabilities of interest in URLLC. Next, we introduce a saddlepoint approximation on (37) that allows for an efficient computation of (37).

IV. A SADDLEPOINT APPROXIMATION ON THE CONDITIONAL ERROR PROBABILITY IN (37)

The saddlepoint method is a well-established tool to obtain accurate approximations of tail probabilities involving sum of random variables. Unfortunately, none of the saddlepoint approximations for the RCUs bounds reported previously in the literature apply to the setup considered in this paper. Indeed, differently from [19], [23], in our setup, $\{Y_{k,\ell}\}_{k=1}^{n_s}$ are not conditionally independent given $\boldsymbol{H} = \mathbf{h}$, $\hat{\boldsymbol{H}} = \hat{\mathbf{h}}$, and $\boldsymbol{\Delta} = \boldsymbol{\delta}$, due to intersymbol interference (see (32)). As a result, the random variables $\{\imath_s(X_{k,\ell};Y_{k,\ell},\hat{h}_\ell)\}_{k=1}^{n_s}$ are also

not conditionally independent. This prevents us from using the saddlepoint expansion reported, e.g., in [19, Thm. 1].

In the case of independent delays, the $\{Y_{k,\ell}\}_{\ell=1}^{n_b}$ are independent for each ℓ, and since delay and channel estimation are performed per block, this implies that the random variables $\left\{\sum_{k=1}^{n_s}\imath_s\Big(X_{k,\ell};Y_{k,\ell},\hat{H}_\ell\Big)\right\}_{\ell=1}^{n_b}$ are also independent. This would in principle enable us to derive a saddlepoint expansion directly on the ϵ_{ub} in (36), similar to the one reported in [19, Thm. 2].[5] However, this approach cannot be applied to the case of fully dependent delays.

To obtain a saddlepoint approximation that applies to both the cases of fully dependent delays and independent delays, we exploit the property that, in both cases, the random variables $\left\{\sum_{k=1}^{n_s}\imath_s(X_{k,\ell};Y_{k,\ell},\hat{h}_\ell)\right\}_{\ell=1}^{n_b}$ are conditionally independent, given $\boldsymbol{H} = \mathbf{h}$, $\hat{\boldsymbol{H}} = \hat{\mathbf{h}}$, and $\boldsymbol{\Delta} = \boldsymbol{\delta}$. This allows us to perform a saddlepoint expansion of the conditional error probability $\epsilon_{\mathrm{ub}}\big(\mathbf{h},\hat{\mathbf{h}},\boldsymbol{\delta}\big)$ in (37) with respect to the number of blocks n_b. Specifically, let us fix $\boldsymbol{\delta}$, $\mathbf{h}$, and $\hat{\mathbf{h}}$, and denote for convenience

$$I_\ell = \sum_{k=1}^{n_s}\imath_s(X_{k,\ell};Y_{k,\ell},\hat{h}_\ell), \qquad (39)$$

where we let, with a slight abuse of notation, for the remainder of this section

$$\imath_s(X_{k,\ell};Y_{k,\ell},\hat{h}_\ell)$$
$$= \log\frac{e^{-s\big|h_\ell(\delta_\ell X_{k,\ell}+(1-\delta_\ell)X_{k-1,\ell})+Z_{k,\ell}-\hat{h}_\ell X_{k,\ell}\big|^2}}{\mathbb{E}_{\bar{X}}\Big[e^{-s\big|h_\ell(\delta_\ell X_{k,\ell}+(1-\delta_\ell)X_{k-1,\ell})+Z_{k,\ell}-\hat{h}_\ell\bar{X}\big|^2}\Big]}, \quad (40)$$

which is a function of the random variables $X_{k,\ell}$, $X_{k-1,\ell}$, and $Z_{k,\ell}$ (and is parameterized by h_ℓ, $\hat{h}_\ell$, and δ).

As we shall see, to obtain the desired saddlepoint expansion, we need to solve the following two issues.

- In our setup, the $\{I_\ell\}_{\ell=1}^{n_b}$ are conditionally independent, but, differently from the variables appearing in previous works, e.g., [19], [23], they are not identically distributed.
- To apply the saddlepoint expansion, we need to be able to compute efficiently (ideally, in closed-form) the moment-generating function (MGF) of each I_ℓ, which requires us to tackle the dependence between the random variables in the sum on the right-hand side of (39), caused by intersymbol interference.

A. Saddlepoint Expansion for Sum of Independent but not Identically Distributed Random Variables

We next present a general result that will allow us to obtain a saddlepoint approximation on $\epsilon_{\mathrm{ub}}\big(\mathbf{h},\hat{\mathbf{h}},\boldsymbol{\delta}\big)$ in (37). Let $\{R_\ell\}_{\ell=1}^n$ be a family of independent, but not necessarily identically distributed, nonlattice random variables.[6] Let us denote by $\varphi_\ell(\zeta) = \mathbb{E}\big[e^{-\zeta R_\ell}\big]$ the MGF of $-R_\ell$ and by

---

[4]$X_{1,\ell}$ is affected by intersymbol interference when $\Lambda_\ell = -1$ due to pilot symbols.

[5]We would, however, still need to deal with the intersymbol-interference within each block.

[6]A random variable R_ℓ is said to be lattice if it is supported on the points $b, b \pm h, b \pm 2h, \ldots$ for some $b \in \mathbb{R}$ and $h \in \mathbb{R}$. A random variable that is not lattice is said to be nonlattice.

$\kappa_\ell(\zeta) = \log \varphi_\ell(\zeta)$ its cumulant generating function (CGF). Let us also set $\kappa(\zeta) = \sum_{\ell=1}^n \kappa_\ell(\zeta)$ as well as

$$\mu(\zeta) = \frac{1}{n} \frac{\mathrm{d}\kappa(\zeta)}{\mathrm{d}\zeta} \tag{41}$$

$$\sigma^2(\zeta) = \frac{1}{n} \frac{\mathrm{d}^2\kappa(\zeta)}{\mathrm{d}\zeta^2}. \tag{42}$$

In the next theorem, we present a saddlepoint approximation for the probability that $\log \Upsilon + \sum_{\ell=1}^n R_\ell$ does not exceed $n\mathrm{R}$.

Theorem 1: Suppose that there exists a $\zeta_0 > 0$ such that

$$\sup_{|\zeta| < \zeta_0} \left| \frac{\mathrm{d}^4\varphi_\ell(\zeta)}{\mathrm{d}\zeta^4} \right| < \infty, \quad \forall \ell \in \{1, \ldots, n\} \tag{43}$$

and also positive constants $m_\mathrm{l} \leq m_\mathrm{u}$ such that

$$m_\mathrm{l} \leq \sigma^2(\zeta) \leq m_\mathrm{u} \tag{44}$$

holds for all $n \in \mathbb{N}$ and for all $|\zeta| \leq \zeta_0$. Assume that there exists a $\zeta \in [-\zeta_0, \zeta_0]$ satisfying $-\mu(\zeta) = \mathrm{R}$. If $\zeta \in [0, 1]$ then

$$\mathbb{P}\left[\log \Upsilon + \sum_{\ell=1}^n R_\ell \leq n\mathrm{R} \right]$$
$$= e^{\kappa(\zeta) - n\zeta\mu(\zeta)} \left[e^{\frac{\beta_\zeta^2}{2}} Q(\beta_\zeta) + e^{\frac{\beta_{1-\zeta}^2}{2}} Q(\beta_{1-\zeta}) + o\left(\frac{1}{\sqrt{n}}\right) \right] \tag{45}$$

where $\beta_a = a\sqrt{n\sigma^2(\zeta)}$. If $\zeta > 1$, then

$$\mathbb{P}\left[\log \Upsilon + \sum_{\ell=1}^n R_\ell \leq n\mathrm{R} \right]$$
$$= e^{\kappa(1) - n\mu(\zeta)} \left[\Psi_n(1, 1) - \Psi_n(0, -1) + \mathcal{O}\left(\frac{1}{\sqrt{n}}\right) \right] \tag{46}$$

where

$$\Psi_n(a, b) = e^{na[-\mu(1) - \mathrm{R} + \sigma^2(1)/2]}$$
$$\times Q\left(a\sqrt{n\sigma^2(1)} - b\frac{n(\mu(1) + \mathrm{R})}{\sqrt{n\sigma^2(1)}} \right). \tag{47}$$

Finally, if $\zeta < 0$, then

$$\mathbb{P}\left[\log \Upsilon + \sum_{\ell=1}^n R_\ell \leq n\mathrm{R} \right] = 1 - \left(e^{\kappa(\zeta) - n\zeta\mu(\zeta)} \right)$$
$$\times \left[e^{\frac{\beta_{-\zeta}^2}{2}} Q(\beta_{-\zeta}) - e^{\frac{\beta_{1-\zeta}^2}{2}} Q(\beta_{1-\zeta}) + \mathcal{O}\left(\frac{1}{\sqrt{n}}\right) \right]. \tag{48}$$

Proof: To prove Theorem 1 one has to combine the steps in [21, App. I] with the ones in [25, App. E]. Furthermore, one has to replace [21, Lem. 8] with [20, Thm. 1, Sec. XVI.6]. ∎

To use Theorem 1 in order to obtain a saddlepoint expansion of (37), one has to replace the $\{R_\ell\}$ with the $\{I_\ell\}$ defined in (39), n with n_b, and R by $n_\mathrm{c}\mathrm{R}$. Verifying that conditions (43) and (44) hold in our scenario is challenging. We shall return to this point in Section V. The saddlepoint approximation is then obtained by neglecting the $o(\cdot)$ and $\mathcal{O}(\cdot)$ terms in (45), (46), and (48). A pseudocode for computing ϵ_ub in (35) using this approximation is provided in Algorithm 2.

Algorithm 2 Saddlepoint Approximation

1: **Input** : $\rho, \mathrm{R}, n_\mathrm{s}, n_\mathrm{p}, n_\mathrm{b}, \mathrm{N}_\mathrm{MC}$
2: **Output** : ϵ_ub
3: Draw $(\mathbf{h}_i, \hat{\mathbf{h}}_i, \boldsymbol{\delta}_i)$ for $i = 1, \ldots, \mathrm{N}_\mathrm{MC}$
4: **for** every $(\mathbf{h}_i, \hat{\mathbf{h}}_i, \boldsymbol{\delta}_i)$ **do**
5:     Solve $\mathrm{R} = \frac{-\mu(\zeta)}{n_\mathrm{c}}$ for ζ, denote solution by ζ^*
6:     Compute $\kappa(\zeta^*)$, $\kappa''(\zeta^*)$
7:     Compute $\epsilon_\mathrm{ub}(\mathbf{h}_i, \hat{\mathbf{h}}_i, \boldsymbol{\delta}_i)$ using Theorem 1
8: **end for**
9: $\epsilon_\mathrm{ub} \leftarrow \frac{1}{\mathrm{N}_\mathrm{MC}} \sum_{i=1}^{\mathrm{N}_\mathrm{MC}} \epsilon_\mathrm{ub}(\mathbf{h}_i, \hat{\mathbf{h}}_i, \boldsymbol{\delta}_i)$

TABLE I: A possible mapping between $(X_{k-1,\ell}, X_{k,\ell})$ and $B_{k,\ell}$ for BPSK constellation; here, $x^{(1)} = -\sqrt{\rho}$ and $x^{(2)} = \sqrt{\rho}$.

$B_{k,\ell}$	$X_{k-1,\ell}$	$X_{k,\ell}$
1	$x^{(1)}$	$x^{(1)}$
2	$x^{(1)}$	$x^{(2)}$
3	$x^{(2)}$	$x^{(1)}$
4	$x^{(2)}$	$x^{(2)}$

B. Closed-Form Evaluation of the Moment Generating Function

It follows from the pseudocode provided in Algorithm 2 that a key step in the computation of the saddlepoint approximation is the evaluation of the moment-generating function $\varphi_\ell(\zeta) = \mathbb{E}\left[e^{-\zeta I_\ell} \right]$, for $\ell = 1, \ldots, n_\mathrm{b}$, which is required to determine $\kappa(\zeta)$ and its first and second derivatives. As shown in (39), each I_ℓ consists of the sum of n_s random variables. However, these random variables are not independent because of the intersymbol interference caused by errors in the estimation of the propagation delay (see (40)). Inspired by [15, Ch. 9], we shall next exploit the Markovian structure of the dependence between these random variables to provide an effective method to evaluate $\varphi_\ell(\zeta)$.

Let $\mathcal{U}$ denote the set of constellation points of cardinality u. It will turn out convenient to associate to every pair[7] $(X_{k-1,\ell}, X_{k,\ell}) \in \mathcal{U}^2$ of consecutive transmitted symbols for $k = 1, \ldots, n_\mathrm{s}$, with an index $B_{k,\ell} \in \{1, \ldots, u^2\}$. In Table I, we present one such association for the case of BPSK constellation. Note that the index $B_{1,\ell}$ depends on the pair $(X_{0,\ell}, X_{1,\ell})$, where $X_{0,\ell}$ stands for the last pilot symbol transmitted before the subcodeword. We shall assume for simplicity that, as for the data symbols, this pilot symbol is drawn uniformly and independently from the constellation $\mathcal{U}$.

We note that for all possible values of $b_{k,\ell}, b_{k-1,\ell}, \ldots, b_{1,\ell}$, we have that

$$\mathbb{P}[B_{k,\ell} = b_{k,\ell} \,|\, B_{k-1,\ell} = b_{k-1,\ell}, \ldots, B_{1,\ell} = b_{1,\ell}]$$
$$= \mathbb{P}[B_{k,\ell} = b_{k,\ell} \,|\, X_{k-1,\ell} = x_{k-1,\ell}, X_{k-2,\ell} = x_{k-2,\ell}] \tag{49}$$
$$= \mathbb{P}[B_{k,\ell} = b_{k,\ell} \,|\, B_{k-1,\ell} = b_{k-1,\ell}], \tag{50}$$

which implies that the sequence $\{B_{k,\ell}\}$ is a Markov chain in k.

---

[7]When considering pulse shapes that create more intersymbol interference, $B_{k,\ell}$ may need to be associated with a larger number of adjacent symbols, leading to a more involved state-space-evolution analysis.

Let $\{(x_{i1}, x_{i2})\}_{i=1}^{u^2}$ denote the set of all possible pairs of consecutively transmitted constellation symbols. In the remainder of the paper, we will use the convention that $B_{k,\ell} = i$ means that $X_{k-1,\ell} = x_{i1}$ and $X_{k,\ell} = x_{i2}$. Using this notation, we can express each entry $[\mathbf{P}]_{ij}$ of the $u^2 \times u^2$ transition matrix[8] $\mathbf{P}$ as

$$[\mathbf{P}]_{ij} = \mathbb{P}[B_{k+1,\ell} = j \mid B_{k,\ell} = i] \tag{51}$$

$$= \mathbb{P}[X_{k+1,\ell} = x_{j2}, X_{k,\ell} = x_{j1}$$
$$\mid X_{k,\ell} = x_{i2}, X_{k-1,\ell} = x_{i1}] \tag{52}$$

$$= \mathbb{P}[X_{k+1,\ell} = x_{j2}] \mathbb{1}\{x_{j1} = x_{i2}\} \tag{53}$$

$$= \frac{1}{u} \mathbb{1}\{x_{j1} = x_{i2}\}. \tag{54}$$

It also follows by our assumptions that the initial state $B_{1,\ell}$ is chosen uniformly at random from the set $\{1, \ldots, u^2\}$.

Let the conditional moment generating function of $-\imath_s(X_{k,\ell}; Y_{k,\ell}, \hat{h}_\ell)$, given $B_{k,\ell} = j_k$, be denoted as

$$\varphi_{j_k,\ell}(\zeta) = \mathbb{E}\left[e^{-\zeta \imath_s\left(X_{k,\ell}; Y_{k,\ell}, \hat{h}_\ell\right)} \,\Big|\, B_{k,\ell} = j_k\right]. \tag{55}$$

Note that the expectation in (55) is only with respect to the additive noise $Z_{k,\ell}$ in (32). Thus, for all $j_1, \ldots, j_{n_s} \in \{1, \ldots, u^2\}$ we have

$$\varphi_{j_k,\ell}(\zeta) = \mathbb{E}\left[e^{-\zeta \imath_s\left(X_{k,\ell}; Y_{k,\ell}, \hat{h}_\ell\right)} \,\Big|\, B_{1,\ell} = j_1, \ldots, B_{n_s,\ell} = j_{n_s}\right], \tag{56}$$

and the conditional MGF of $-I_\ell$ can therefore be stated as

$$\mathbb{E}\left[e^{-\zeta \sum_{k=1}^{n_s} \imath_s\left(X_{k,\ell}; Y_{k,\ell}, \hat{h}_\ell\right)} \,\Big|\, B_{1,\ell} = j_1, \ldots, B_{n_s,\ell} = j_{n_s}\right]$$

$$= \prod_{k=1}^{n_s} \mathbb{E}\left[e^{-\zeta \imath_s\left(X_{k,\ell}; Y_{k,\ell}, \hat{h}_\ell\right)} \,\Big|\, B_{1,\ell} = j_1, \ldots, B_{n_s,\ell} = j_{n_s}\right] \tag{57}$$

$$= \prod_{k=1}^{n_s} \varphi_{j_k,\ell}(\zeta) \tag{58}$$

where (57) follows since $\{Z_{k,\ell}\}_{k=1}^{n_s}$ are i.i.d.. We can remove the conditioning using the law of total expectation as

$$\varphi_\ell(\zeta) = \sum_{j_1=1}^{u^2} \cdots \sum_{j_{n_s}=1}^{u^2} \left(\prod_{k=1}^{n_s} \varphi_{j_k,\ell}(\zeta)\right)$$
$$\times \mathbb{P}[B_{1,\ell} = j_1, \ldots, B_{n_s,\ell} = j_{n_s}]. \tag{59}$$

Evaluating this expression requires us to sum u^{2n_s} terms, which is computationally infeasible except for very small n_s. We next exploit the Markov structure to find an efficient method to compute $\varphi_\ell(\zeta)$.

To this end, let $\mathbf{P}_\ell(\zeta)$ be the $u^2 \times u^2$ matrix with elements $[\mathbf{P}_\ell(\zeta)]_{mn} = [\mathbf{P}]_{mn}\varphi_{n,\ell}(\zeta)$ for $m, n \in \{1, \ldots, u^2\}$. The entry

in position[9] (j_1, j_{n_s}) of the matrix $\mathbf{P}_\ell(\zeta)^{n_s-1}$ can be (for $n_s \geq 3$) expressed as

$$\left[\mathbf{P}_\ell(\zeta)^{n_s-1}\right]_{j_1 j_{n_s}}$$

$$= \sum_{j_2=1}^{u^2} \cdots \sum_{j_{n_s-1}=1}^{u^2} [\mathbf{P}_\ell(\zeta)]_{j_1 j_2} \cdots [\mathbf{P}_\ell(\zeta)]_{j_{n_s-1} j_{n_s}} \tag{60}$$

$$= \sum_{j_2=1}^{u^2} \cdots \sum_{j_{n_s-1}=1}^{u^2} \left(\prod_{k=2}^{n_s} \varphi_{j_k,\ell}(\zeta)\right)$$
$$\times [\mathbf{P}]_{j_1 j_2} \cdots [\mathbf{P}]_{j_{n_s-1} j_{n_s}}. \tag{61}$$

Now, since

$$[\mathbf{P}]_{j_1 j_2} \cdots [\mathbf{P}]_{j_{n_s-1} j_{n_s}}$$

$$= \prod_{k=2}^{n_s} \mathbb{P}[B_{k,\ell} = j_k \mid B_{k-1,\ell} = j_{k-1}, \ldots, B_{1,\ell} = j_1] \tag{62}$$

$$= \mathbb{P}[B_{n_s,\ell} = j_{n_s}, \ldots, B_{2,\ell} = j_2 \mid B_{1,\ell} = j_1] \tag{63}$$

$$= \mathbb{P}[B_{n_s-1,\ell} = j_{n_s-1}, \ldots, B_{2,\ell} = j_2 \mid B_{n_s,\ell} = j_{n_s},$$
$$B_{1,\ell} = j_1] \mathbb{P}[B_{n_s,\ell} = j_{n_s} \mid B_{1,\ell} = j_1] \tag{64}$$

$$= \mathbb{P}[B_{n_s-1,\ell} = j_{n_s-1}, \ldots, B_{2,\ell} = j_2 \mid B_{n_s,\ell} = j_{n_s},$$
$$B_{1,\ell} = j_1][\mathbf{P}^{n_s-1}]_{j_1, j_{n_s}} \tag{65}$$

it follows (from the law of total expectation) that (61) simplifies to

$$\left[\mathbf{P}_\ell(\zeta)^{n_s-1}\right]_{j_1 j_{n_s}} = \mathbb{E}\left[e^{-\zeta \sum_{k=2}^{n_s} \imath_s(X_{k,\ell}; Y_{k,\ell}, \hat{h}_\ell)} \,\Big|\, B_{n_s,\ell} = j_{n_s}, B_{1,\ell} = j_1\right] [\mathbf{P}^{n_s-1}]_{j_1 j_{n_s}}. \tag{66}$$

We can now form an expression for $\varphi_\ell(\zeta)$ as

$$\varphi_\ell(\zeta) = \mathbb{E}\left[e^{-\zeta I_\ell}\right] \tag{67}$$

$$= \mathbb{E}\left[e^{-\zeta \imath_s\left(X_{1,\ell}; Y_{1,\ell}, \hat{h}_\ell\right)} e^{-\zeta \sum_{k=2}^{n_s} \imath_s\left(X_{k,\ell}; Y_{k,\ell}, \hat{h}_\ell\right)}\right] \tag{68}$$

$$= \sum_{j_1=1}^{u^2} \mathbb{E}\left[e^{-\zeta \imath_s\left(X_{1,\ell}; Y_{1,\ell}, \hat{h}_\ell\right)} e^{-\zeta \sum_{k=2}^{n_s} \imath_s\left(X_{k,\ell}; Y_{k,\ell}, \hat{h}_\ell\right)} \,\Big|\, B_{1,\ell} = j_1\right] \mathbb{P}[B_1 = j_1] \tag{69}$$

$$= \sum_{j_1=1}^{u^2} \frac{1}{u^2} \mathbb{E}\left[e^{-\zeta \imath_s\left(X_{1,\ell}; Y_{1,\ell}, \hat{h}_\ell\right)} \,\Big|\, B_{1,\ell} = j_1\right]$$
$$\times \mathbb{E}\left[e^{-\zeta \sum_{k=2}^{n_s} \imath_s\left(X_{k,\ell}; Y_{k,\ell}, \hat{h}_\ell\right)} \,\Big|\, B_{1,\ell} = j_1\right] \tag{70}$$

$$= \sum_{j_1=1}^{u^2} \frac{\varphi_{j_1,\ell}(\zeta)}{u^2}$$
$$\times \mathbb{E}\left[e^{-\zeta \sum_{k=2}^{n_s} \imath_s\left(X_{k,\ell}; Y_{k,\ell}, \hat{h}_\ell\right)} \,\Big|\, B_{1,\ell} = j_1\right] \tag{71}$$

where (69) follows from the law of total expectation and (70) follows since $e^{-\zeta \imath_s\left(X_{1,\ell}; Y_{1,\ell}, \hat{h}_\ell\right)}$ and $e^{-\zeta \sum_{k=2}^{n_s} \imath_s\left(X_{k,\ell}; Y_{k,\ell}, \hat{h}_\ell\right)}$

---

[8]The transition matrix does not depend on the index ℓ; hence, this index is omitted.

[9]The indices j_1 and j_{n_s} can be chosen arbitrarily, independent of their subscripts. We select these subscripts to keep the following derivations compact.

are independent conditioned on $B_{1,\ell}$. The last factor in (71) can be written as

$$\mathbb{E}\left[e^{-\zeta \sum_{k=2}^{n_s} \imath_s\left(X_{k,\ell};Y_{k,\ell},\hat{h}_\ell\right)} \,\Big|\, B_{1,\ell} = j_1\right]$$

$$= \sum_{j_{n_s}=1}^{u^2} \mathbb{E}\left[e^{-\zeta \sum_{k=2}^{n_s} \imath_s\left(X_{k,\ell};Y_{k,\ell},\hat{h}_\ell\right)}\Big|B_{n_s,\ell} = j_{n_s},\right.$$

$$\left. B_{1,\ell} = j_1\right] \mathbb{P}[B_{n_s} = j_{n_s} \mid B_1 = j_1] \quad (72)$$

$$= \sum_{j_{n_s}=1}^{u^2} \left[\mathbf{P}_\ell(\zeta)^{n_s-1}\right]_{j_1 j_{n_s}}. \quad (73)$$

Combining (71) and (73) yields

$$\varphi_\ell(\zeta) = \sum_{j_1=1}^{u^2} \sum_{j_{n_s}=1}^{u^2} \frac{\varphi_{j_1,\ell}(\zeta)}{u^2} \left[\mathbf{P}_\ell(\zeta)^{n_s-1}\right]_{j_1 j_{n_s}} \quad (74)$$

$$= \boldsymbol{\nu}_\ell(\zeta)^T \mathbf{P}_\ell(\zeta)^{n_s-1} \mathbf{1}_{u^2}, \quad (75)$$

where

$$\boldsymbol{\nu}_\ell(\zeta) = \left[\frac{1}{u^2}\varphi_{1,\ell}(\zeta),\dots,\frac{1}{u^2}\varphi_{u^2,\ell}(\zeta)\right]^T. \quad (76)$$

To summarize, (75) provides an efficient way to evaluate the moment generating function $\varphi_\ell(\zeta)$, provided that one can compute efficiently powers of the matrix $\mathbf{P}_\ell(\zeta)$. Whether this is possible or not depends on the constellation that is used. In the next section, we discuss in detail the practically relevant case of BPSK constellation.

C. A Case Study: BPSK Constellation

We next consider BPSK constellation and use the mapping between $(X_{k-1,\ell}, X_{k,\ell})$ and $B_{k,\ell}$ described in Table I. It follows by symmetry that $\varphi_{1,\ell}(\zeta) = \varphi_{4,\ell}(\zeta)$ and that $\varphi_{2,\ell}(\zeta) = \varphi_{3,\ell}(\zeta)$. As illustrated in Appendix A, these quantities and their derivatives, which are needed to numerically evaluate the saddlepoint expansion in Theorem 1, can be computed efficiently using numerical integration.

Using the symmetry just unveiled, we can express $\boldsymbol{\nu}_\ell(\zeta)$ and $\mathbf{P}_\ell(\zeta)$ as

$$\boldsymbol{\nu}_\ell(\zeta) = \left[\frac{\varphi_{1,\ell}(\zeta)}{4}\ \ \frac{\varphi_{2,\ell}(\zeta)}{4}\ \ \frac{\varphi_{2,\ell}(\zeta)}{4}\ \ \frac{\varphi_{1,\ell}(\zeta)}{4}\right]^T, \quad (77)$$

and

$$\mathbf{P}_\ell(\zeta) = \begin{bmatrix} \frac{\varphi_{1,\ell}(\zeta)}{2} & \frac{\varphi_{2,\ell}(\zeta)}{2} & 0 & 0 \\ 0 & 0 & \frac{\varphi_{2,\ell}(\zeta)}{2} & \frac{\varphi_{1,\ell}(\zeta)}{2} \\ \frac{\varphi_{1,\ell}(\zeta)}{2} & \frac{\varphi_{2,\ell}(\zeta)}{2} & 0 & 0 \\ 0 & 0 & \frac{\varphi_{2,\ell}(\zeta)}{2} & \frac{\varphi_{1,\ell}(\zeta)}{2} \end{bmatrix}. \quad (78)$$

To evaluate $\mathbf{P}_\ell(\zeta)^{n_s-1}$ we write $\mathbf{P}_\ell(\zeta)$ in terms of its eigenvalue decomposition as $\mathbf{P}_\ell(\zeta) = \mathbf{V}_\ell \boldsymbol{\Lambda}_\ell \mathbf{V}_\ell^{-1}$ where

$$\mathbf{V}_\ell = \begin{bmatrix} 0 & -\frac{\varphi_{2,\ell}(\zeta)}{\varphi_{1,\ell}(\zeta)} & -1 & 1 \\ 0 & 1 & 1 & 1 \\ -\frac{\varphi_{1,\ell}(\zeta)}{\varphi_{2,\ell}(\zeta)} & 0 & -1 & 1 \\ 1 & 0 & 1 & 1 \end{bmatrix} \quad (79)$$

contains the eigenvectors of $\mathbf{P}_\ell(\zeta)$ and

$$\boldsymbol{\Lambda}_\ell = \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & \frac{\varphi_{1,\ell}(\zeta)-\varphi_{2,\ell}(\zeta)}{2} & 0 \\ 0 & 0 & 0 & \frac{\varphi_{1,\ell}(\zeta)+\varphi_{2,\ell}(\zeta)}{2} \end{bmatrix}, \quad (80)$$

contains the corresponding eigenvalues. Then we use that $\mathbf{P}_\ell(\zeta)^{n_s-1} = \mathbf{V}_\ell \boldsymbol{\Lambda}_\ell^{n_s-1} \mathbf{V}_\ell^{-1}$. It then follows from (75) that

$$\varphi_\ell(\zeta) = \frac{1}{2^{n_s}}(\varphi_{1,\ell}(\zeta) + \varphi_{2,\ell}(\zeta))^{n_s}. \quad (81)$$

## V. NUMERICAL RESULTS AND DISCUSSION

In this section, we report numerical experiments illustrating the performance of our synchronization method both in terms of normalized mean square error (NMSE) and achievable packet error probability. We also verify the accuracy of the saddlepoint approximation.

Throughout, we shall consider the case of fully dependent delays; we shall also assume that a BPSK constellation is used for both pilot and data transmission, and that m-sequences [26, Ch. 8] are used as pilot sequences. We consider a Rayleigh-fading scenario, i.e., we assume that the $H_\ell, \ell = 1,\dots,n_b$, are generated independently from a $\mathcal{CN}(0,1)$ distribution. Since $N_0 = 1$, we can interpret the power constraint ρ introduced in Section II as the (average) SNR. We will analyze the performance of both the per-block and the joint synchronization algorithms proposed in Sections II-C1 and II-C2, respectively. Note that per-block synchronization, when applied to case of fully dependent delays, achieves the same performance as if the delays across blocks were independent.

To assess the performance of the estimators introduced in Section II-C1 and Section II-C2, we compare their NMSE with the CRB, whose evaluation is detailed in Appendix B. The delay NMSE is defined as $\mathbb{E}\left[(D - \hat{D})^2/t_p^2\right]$ for the case of joint synchronization, and as $\mathbb{E}\left[(D - \hat{D}_\ell)^2/t_p^2\right]$ for the case of per-block synchronization. Note that this last quantity does not depend on ℓ. The channel-gain NMSE is defined as $\mathbb{E}\left[\left|H_\ell - \hat{H}_\ell\right|^2 / |H_\ell|^2\right]$ for both algorithms. This quantity also does not depend on ℓ.

In Fig. 3, we depict the NMSE as a function of the SNR for $N = 10$, $n_p = 7$, and $n_b = 4$. The curves are obtained by assuming no interference from the data symbol, i.e., by setting $C_\ell = \mathbf{0}_M$ in (12). Under this assumption, our estimators are unbiased and the NMSE approaches the CRB as the SNR increases, as shown in Fig. 3. As expected, joint synchronization significantly outperforms per-block synchronization (which, however, is the only option if the delays are independent across blocks). Furthermore, the convergence to the CRB is more rapid.

In Fig. 4, we illustrate the impact of synchronization errors on the packet error probability, evaluated by substituting the RCUs bound ϵ_{ub} for ϵ_1 in (31). To do so, we let σ_d^2 be the mean squared error incurred when estimating D. In this analysis (and only in this analysis), we assume the channel to be known at the receiver (i.e., $\hat{h}_\ell = h_\ell$ for all ℓ), let $\hat{D} \sim \mathcal{N}(D, \sigma_d^2)$ (or
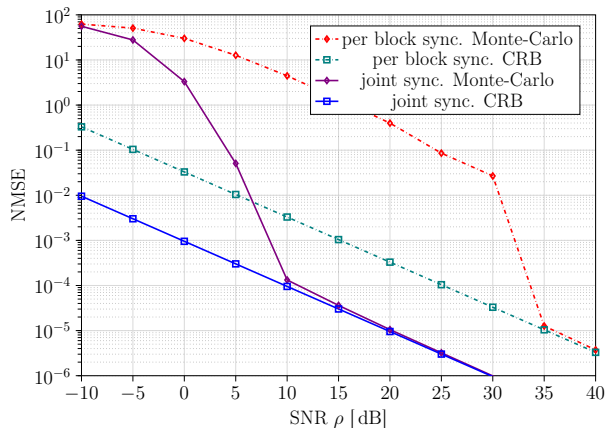
(a) Synchronization error
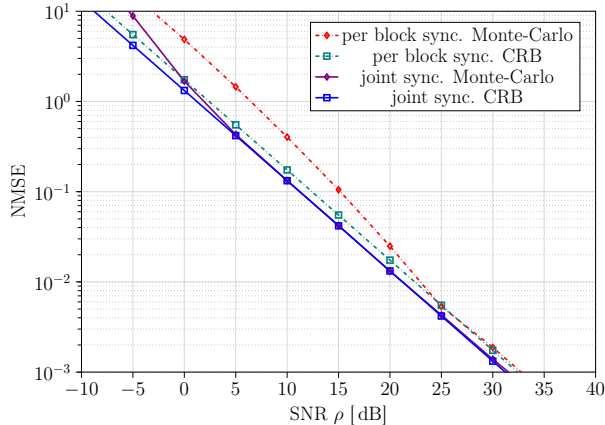
(b) Channel estimation error

Fig. 3: Comparison between NMSE of the estimators and CRB for N = 10, $n_{\mathrm{p}} = 7$, $n_{\mathrm{b}} = 4$.

Fig. 4: Achievable packet error probability evaluated using the RCUs bound as a function of the average synchronization error. Here $n = 288$, $n_{\mathrm{b}} = 8$, R = 0.104 bit per channel use; s is optimized.

Fig. 5: NMSE for the delay estimation for both joint and per-block synchronization for N = 10, $n_{\mathrm{b}} = 4$, $n_{\mathrm{p}} = 7$.

$\hat{D}_\ell \sim \mathcal{N}(D, \sigma_d^2)$ for the per-block synchronization algorithm), where we allow σ_d^2, which corresponds to the mean square error for delay estimation, to vary independently from any other system parameter, and report an upper bound on the ϵ_{pep} as a function of $\sigma_d^2/t_{\mathrm{p}}^2$ for $\rho \in \{2.5, 6.5\}$ dB, $n_{\mathrm{b}}n_{\mathrm{c}} = 288$, $n_{\mathrm{b}} = 8$, $n_{\mathrm{p}} = 0$, and R $= 30/288 = 0.104$ bit per channel use. These parameters reflect a URLLC scenario involving the transmission of compact downlink control information [22], [27].

The parameter s of the RCUs bound is optimized. Note that, since we assumed that the $\{h_\ell\}$ are known to the receiver, once we fix σ_d^2, the packet error probability achieved using joint synchronization and per-block synchronization coincide. We see from the figure that to achieve an error probability $\epsilon_{\mathrm{pep}} < 10^{-4}$ for both values of ρ, it is enough that $\sigma_d^2/t_{\mathrm{p}}^2$ is below 0.12 and that the error probability deteriorates rapidly once this value is exceeded.

We next assess the validity of the assumption that $C_\ell = \mathbf{0}_{\mathrm{M}}$ by plotting in Fig. 5 the NMSE for the delay estimation as a function of the SNR for both the case of no data interference and data interference. In the figure, we assume that $N = 10$, $n_{\mathrm{b}} = 4$, and $n_{\mathrm{p}} = 7$. As shown in the figure, the impact of data interference on the performance of our algorithm is minimal.

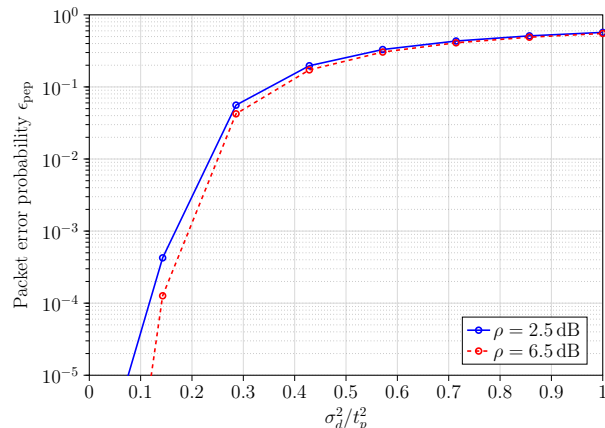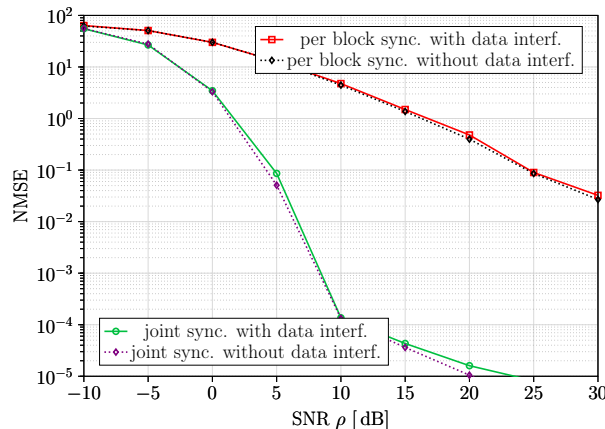Next, we report in Fig. 6 the SNR required to achieve a

packet error rate $\epsilon = 10^{-5}$ for a transmission rate R $= 0.104$ bit per channel use, $n_{\mathrm{b}} = 8$, and $n_{\mathrm{b}}n_{\mathrm{c}} = 288$ as a function of the upsampling rate N. The SNR values are obtained by performing an optimization over both s and the number of pilot symbols n_{p}. We see from the figure that a significant SNR reduction can be achieved by increasing N from 1 to 5. However, a further increase in N yields negligible performance improvements[10]. In the figure, we also assess once more the impact of interference from the data symbols and conclude again that it is negligible.

In Fig. 7, we analyze the performance achievable in the URLLC regime using the synchronization and channel estimation algorithms introduced in Section III. Specifically, we show the SNR required to achieve an error probability of 10^{-5} for $n_{\mathrm{b}}n_{\mathrm{c}} = 288$, N $= 5$, and R $= 0.104$ bit per channel, as a function of the number of fading blocks n_{b} spanned by each codeword. We obtain each value of SNR by optimizing over both the number of pilot symbols n_{p} and the s parameter in the RCUs. We consider both per-block and joint synchroniza-

---

[10]The same trend is observed for the channel estimation and synchronization errors: the corresponding NMSE either does not improve much for N > 5 or is small enough at N = 5 such that further reduction does not significantly improve the error probability.
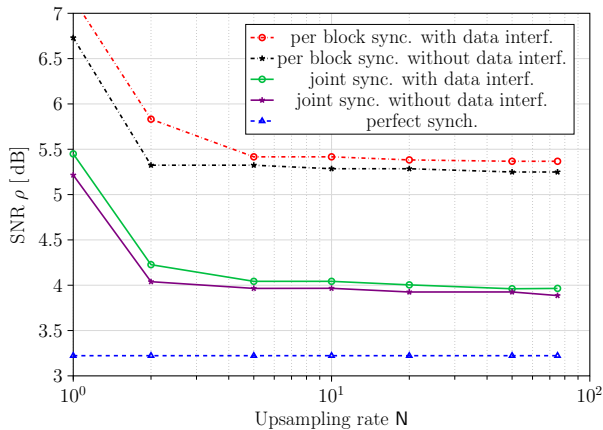
Fig. 6: Upper bound on the SNR required to achieve $\epsilon_{\mathrm{pep}} = 10^{-5}$ as a function of N. Here, $n_{\mathrm{b}}n_{\mathrm{c}} = 288$, R $= 0.104$ bit per channel use and $n_{\mathrm{b}} = 8$; n_{p} and s are optimized.

Fig. 7: Upper bound on the SNR sufficient to achieve $\epsilon_{\mathrm{pep}} = 10^{-5}$ as a function of n_{b}. Here, $n_{\mathrm{b}}n_{\mathrm{c}} = 288$, R $= 0.104$ bit per channel use, N $= 5$; n_{p} and s are optimized.

TABLE II: Optimal values of n_{p} in Fig. 7.

n_{b}	joint sync. n_{p}	perfect sync. n_{p}
2	31	31
4	15	15
8	15	15
12	7	7
24	3	3

tion, and depict for reference also the curve corresponding to perfect synchronization and channel estimation, and the one corresponding to perfect synchronization, but pilot-aided estimation of the channel gain. For each scenario, we plot both the RCUs and its saddlepoint approximation. We observe that the saddlepoint approximation provides an accurate and numerically efficient approximation of the RCUs bound for the parameters considered in the figure. This also suggests that the conditions required for the saddlepoint approximation, given in (43) and (44), hold in our setup.

We observe from the figure that per-block synchronization requires up to $3.5\,\mathrm{dB}$ higher SNR than joint synchronization to achieve the same error probability. We note that when $n_{\mathrm{b}} \geq 4$, the SNR gap between joint synchronization and perfect synchronization but pilot-aided channel estimation is no larger than $0.6\,\mathrm{dB}$. This suggests that the pilot symbols needed to estimate the fading coefficients in the perfect synchronization, pilot-aided channel estimation case are sufficient to also estimate the delay when the joint synchronization algorithm is used. To strengthen this claim, we report the optimum number of n_{p} for joint, and perfect synchronization cases in Table II. Unsurprisingly, the optimum number of n_{p} for the joint and perfect synchronization case are the same for all values of n_{b} considered in Fig. 7. It is interesting to note that the gap between the perfect synchronization, pilot-aided channel-estimation curve and the curve corresponding to perfect synchronization and channel estimation[11] increases as a function of SNR. This is because, as n_{b} is increased for a fixed $n_{\mathrm{c}}n_{\mathrm{b}}$, the number of symbols n_{c} that can be transmitted in a block decreases. As a result, the optimal value of n_{p} decreases as well, as illustrated in Table II, which causes a deterioration in the accuracy of the channel estimate and an increase of the error probability, for a fixed SNR.

The optimal number of pilots for the per-block synchronization case turns out to coincide with that for the joint-synchronization case. The reason is as follows: while increas-

ing n_{p} reduces estimation errors, it comes at the cost of a reduction of the number of data symbols n_{s}, which impacts the packet error probability. This is illustrated in Fig. 8, where we report our upper bound on the packet error probability ϵ_{pep} as a function of n_{p} for R $= 0.104$ bit per channel use, $n_{\mathrm{b}} = 4$, $n_{\mathrm{c}}n_{\mathrm{b}} = 288$, N $= 5$, and $\rho = 8.45\,\mathrm{dB}$ for both per-block and joint synchronization. We see that the value of n_{p} that minimizes our upper bound ϵ_{pep} coincides with per-block and joint synchronization. However, per-block synchronization yields a higher error probability for a fixed SNR. This implies that, as shown in Fig. 7, per-block synchronization requires a higher SNR than joint-synchronization, to achieve a given target error probability. The results in Fig. 8 illustrate also that the saddlepoint approximation provides an accurate approximation for all n_{p} values considered in the figure. This is of utmost importance because the parameter n_{p} may not be optimized specifically to minimize 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 pilot-assisted transmission for channel estimation and timing synchronization. The method is based on a novel saddlepoint approximation, which accounts for the dependence across certain random variables arising in the presence of synchronization errors.

Numerical experiments conducted using two synchronization and channel-estimation algorithms show that the proposed

---

[11]To perform a fair comparison between the two curves, the one corresponding to perfect synchronization and channel estimation is obtained by using the n_{p} values in Table II, although these pilot symbols are not used.
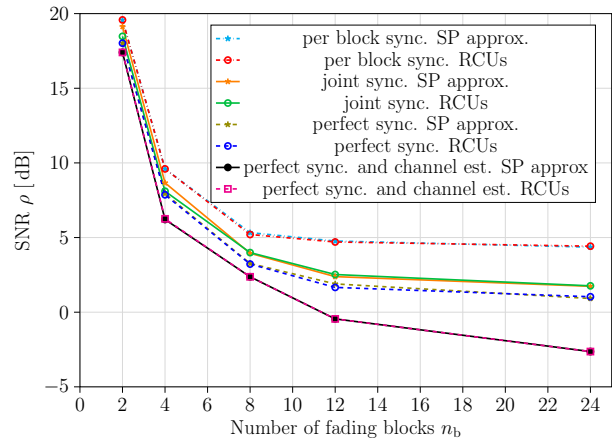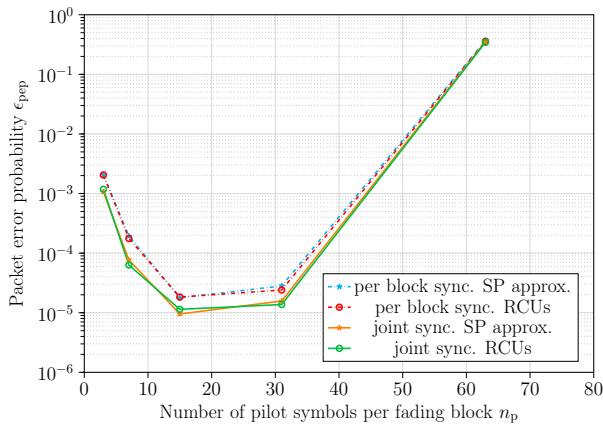
Fig. 8: Upper bound on packet error probability ϵ_{pep} as a function of n_{p}. Here, $\rho = 8.45$ dB, $n_{\text{b}} = 4$, $n_{\text{b}}n_{\text{c}} = 288$, R = 0.104 bit per channel use, N = 5; s is optimized for each value of n_{b}.

saddlepoint approximation can be safely used to benchmark URLLC systems. We show how to use our approximation to determine the synchronization level required to achieve the low error probabilities demanded in URLLC applications. Moreover, our numerical results reveal that, when the delays are fully dependent across fading blocks and synchronization is performed jointly over the blocks, the pilot symbols needed for channel estimate are sufficient to acquire synchronization. However, if we use per-block synchronization (which is unavoidable if delays are independent across blocks), the same number of pilot symbols is not enough to achieve a sufficiently good synchronization. As a result, for the case when the delays are independent, synchronization becomes the bottleneck for the system performance, and yields an SNR penalty that can be quantified using our bounds. A generalization of the analysis presented in this paper to the case of multiple-antenna systems is an interesting topic for a future work.

APPENDIX A
NUMERICAL EVALUATION OF (55) AND ITS FIRST TWO DERIVATIVES

We consider BPSK constellation and use the mapping described in Table I. For this case, (38) reduces to

$$\imath_s(x; y, \hat{h}) = -s\left|y - \hat{h}x\right|^2$$
$$- \log \frac{e^{-s|y - \hat{h}x^{(1)}|^2} + e^{-s|y - \hat{h}x^{(2)}|^2}}{2}. \quad (82)$$

For a given $B_{k,\ell} = j$, we have that $X_{k-1,\ell} = x_{j1}$ and $X_{k,\ell} = x_{j2}$. Hence, conditioned on $\Delta_\ell = \delta_\ell$, $H_\ell = h_\ell$, $\hat{H}_\ell = \hat{h}_\ell$, we have

$$Y_{k,\ell} = h_\ell(\delta x_{j2} + (1 - \delta)x_{j1}) + Z_{k,\ell} \quad (83)$$
$$= \alpha_j + Z_{k,\ell}, \quad (84)$$

where $\alpha_j = h_\ell(\delta x_{j2} + (1 - \delta)x_{j1})$. Hence,

$$\imath_s(X_{k,\ell}; Y_{k,\ell}, \hat{h}) = -s\left|\alpha_j + Z_{k,\ell} - \hat{h}x\right|^2$$
$$- \log \frac{1}{2}\left(e^{-s|\alpha_j + Z_{k,\ell} - \hat{h}x^{(1)}|^2} + e^{-s|\alpha_j + Z_{k,\ell} - \hat{h}x^{(2)}|^2}\right), \quad (85)$$

and

$$e^{-\zeta\imath_s(X_{k,\ell}; Y_{k,\ell}, \hat{h})} = e^{\zeta f_1(Z_{k,\ell})}f_2(Z_{k,\ell})^\zeta, \quad (86)$$

where

$$f_1(Z_{k,\ell}) = s\left|\alpha_j + Z_{k,\ell} - \hat{h}_\ell x_{j2}\right|^2, \quad (87)$$
$$f_2(Z_{k,\ell}) = \frac{1}{2}\left(e^{-s|\alpha_j + Z_{k,\ell} - \hat{h}_\ell x^{(1)}|^2} + e^{-s|\alpha_j + Z_{k,\ell} - \hat{h}_\ell x^{(2)}|^2}\right). \quad (88)$$

The nth derivative of $\exp(-\zeta\imath_s(X_{k,\ell}; Y_{k,\ell}, \hat{h}))$ is

$$\frac{d^n}{d\zeta^n}e^{-\zeta\imath_s(X_{k,\ell}; Y_{k,\ell}, \hat{h})}$$
$$= e^{\zeta f_1(Z_{k,\ell})}f_2(Z_{k,\ell})^\zeta(f_1(Z_{k,\ell}) + \log f_2(Z_{k,\ell}))^n. \quad (89)$$

Hence,

$$\varphi_{j,\ell}(\zeta) = \mathbb{E}\left[e^{\zeta f_1(Z_{k,\ell})}f_2(Z_{k,\ell})^\zeta\right] \quad (90)$$
$$\varphi'_{j,\ell}(\zeta) = \mathbb{E}\left[e^{\zeta f_1(Z_{k,\ell})}f_2(Z_{k,\ell})^\zeta(f_1(Z_{k,\ell}) + \log f_2(Z_{k,\ell}))\right] \quad (91)$$
$$\varphi''_{j,\ell}(\zeta) = \mathbb{E}\left[e^{\zeta f_1(Z_{k,\ell})}f_2(Z_{k,\ell})^\zeta(f_1(Z_{k,\ell}) + \log f_2(Z_{k,\ell}))^2\right]. \quad (92)$$

Since $Z_{k,\ell} \sim \mathcal{CN}(0, 1)$, the expectations in (90), (91), and (92) can be expressed as two-dimensional real integrals, which can be evaluated efficiently via numerical integration methods.

APPENDIX B
EVALUATION OF CRB

For the purpose of computing the CRB in the asynchronous case, we consider the problem of estimating the parameter vector $\boldsymbol{\theta} = [\text{Re}\{\mathbf{h}^T\}, \text{Im}\{\mathbf{h}^T\}, \mathbf{d}^T]^T \in \mathbb{R}^{3n_{\text{b}}}$ where $\mathbf{d} = [d_1, \ldots, d_{n_{\text{b}}}]^T$. We view the parameter vector as unknown and deterministic. Ignoring the the data interference part in (12), the observed data vector is

$$\boldsymbol{Y}^{(p)} = \left[h_1\mathbf{v}(d_1)^T, \ldots, h_{n_{\text{b}}}\mathbf{v}(d_{n_{\text{b}}})\right]^T + \left[\boldsymbol{Z}_1^T, \ldots, \boldsymbol{Z}_{n_{\text{b}}}^T\right]^T \quad (93)$$

where M is defined in (8),

$$\mathbf{v}(d_\ell) = \left(1 - \frac{e_\ell}{t_{\text{s}}}\right)\mathbf{x}_{\text{N}}^{(p)}(q_\ell) + \frac{e_\ell}{t_{\text{s}}}\mathbf{x}_{\text{N}}^{(p)}(q_\ell + 1) \quad (94)$$

and $q_\ell = \lfloor d_\ell/t_{\text{s}} \rfloor$ and $e_\ell = d_\ell - q_\ell t_{\text{s}}$. Hence, $\boldsymbol{Y}^{(p)}$ is a complex Gaussian random vector with covariance matrix $\mathbf{I}_{n_{\text{b}}M}$ and mean vector

$$\boldsymbol{\mu}(\boldsymbol{\theta}) = \left[h_1\mathbf{v}(d_1)^T, \ldots, h_{n_{\text{b}}}\mathbf{v}(d_{n_{\text{b}}})^T\right]^T. \quad (95)$$

The Fisher information matrix $\mathbf{J}(\boldsymbol{\theta})$ for this problem is given by [28, Ch. 15.7]

$$[\mathbf{J}(\boldsymbol{\theta})]_{mn} = 2\,\text{Re}\left\{\frac{\partial\boldsymbol{\mu}(\boldsymbol{\theta})^H}{\partial\theta_m}\frac{\partial\boldsymbol{\mu}(\boldsymbol{\theta})}{\partial\theta_n}\right\}, \quad (96)$$

where $m, n \in \{1, \ldots, 3n_{\text{b}}\}$, θ_m is the mth element of $\boldsymbol{\theta}$, and the derivatives are evaluated at the true parameter values. Hence, we need to compute the derivatives of $\boldsymbol{\mu}(\boldsymbol{\theta})$ with respect to the real and imaginary parts of $h_1, \ldots, h_{n_{\text{b}}}$, which

is trivial, and with respect to $d_1, \ldots, d_{n_{\mathrm{b}}}$, which requires a bit of care. We note that if $e_\ell \neq 0$, then

$$\frac{\partial}{\partial d_\ell} h_\ell \mathbf{v}(d_\ell) = \frac{h_\ell}{t_{\mathrm{s}}} (\mathbf{x}_{\mathrm{N}}^{(p)}(q_\ell + 1) - \mathbf{x}_{\mathrm{N}}^{(p)}(q_\ell)). \qquad (97)$$

However, if $e_\ell = 0$, then $\boldsymbol{\mu}(\boldsymbol{\theta})$ may not be differentiable with respect to d_ℓ, and the CRB cannot be computed for this case. We will therefore from now on assume[12] that $e_\ell \neq 0$ for $\ell \in \{1, 2, \ldots, n_{\mathrm{b}}\}$. It then follows from (94) and (95) that, for $\ell \in \{1, 2, \ldots, n_{\mathrm{b}}\}$,

$$\frac{\partial \boldsymbol{\mu}(\boldsymbol{\theta})}{\partial \theta_\ell} = \left[\mathbf{0}_{\mathrm{M}(\ell-1)}^T, \mathbf{v}(d_\ell)^T, \mathbf{0}_{\mathrm{M}(n_{\mathrm{b}}-\ell)}^T \right]^T \qquad (98)$$

$$\frac{\partial \boldsymbol{\mu}(\boldsymbol{\theta})}{\partial \theta_{n_{\mathrm{b}}+\ell}} = \left[\mathbf{0}_{\mathrm{M}(\ell-1)}^T, j\mathbf{v}(d_\ell)^T, \mathbf{0}_{\mathrm{M}(n_{\mathrm{b}}-\ell)}^T \right]^T \qquad (99)$$

$$\frac{\partial \boldsymbol{\mu}(\boldsymbol{\theta})}{\partial \theta_{2n_{\mathrm{b}}+\ell}} = \frac{h_\ell}{t_{\mathrm{s}}} \left[\mathbf{0}_{\mathrm{M}(\ell-1)}^T, \mathbf{x}_{\mathrm{N}}^{(p)}(q_\ell+1) - \mathbf{x}_{\mathrm{N}}^{(p)}(q_\ell), \mathbf{0}_{\mathrm{M}(n_{\mathrm{b}}-\ell)}^T \right]^T. \qquad (100)$$

Using the Fisher information matrix $\mathbf{J}(\boldsymbol{\theta})$, we can bound the estimation error variances for unbiased estimators as

$$\mathbb{E}\left[\left(\mathrm{Re}\{\hat{H}_\ell\} - \mathrm{Re}\{h_\ell\} \right)^2 \right] \geq \left[\mathbf{J}(\boldsymbol{\theta})^{-1} \right]_{\ell\ell} \qquad (101)$$

$$\mathbb{E}\left[\left(\mathrm{Im}\{\hat{H}_\ell\} - \mathrm{Im}\{h_\ell\} \right)^2 \right] \geq \left[\mathbf{J}(\boldsymbol{\theta})^{-1} \right]_{(n_{\mathrm{b}}+\ell)(n_{\mathrm{b}}+\ell)} \qquad (102)$$

$$\mathbb{E}\left[\left(\hat{D}_\ell - d_\ell \right)^2 \right] \geq \left[\mathbf{J}(\boldsymbol{\theta})^{-1} \right]_{(2n_{\mathrm{b}}+\ell)(2n_{\mathrm{b}}+\ell)} \qquad (103)$$

where $\ell \in \{1, \ldots, n_{\mathrm{b}}\}$. We may obtain the CRB for the channel estimation as

$$\mathbb{E}\left[\left| \hat{H}_\ell - h_\ell \right|^2 \right]$$

$$= \mathbb{E}\left[(\mathrm{Re}\{\hat{H}_\ell\} - \mathrm{Re}\{h_\ell\})^2 + (\mathrm{Im}\{\hat{H}_\ell\} - \mathrm{Im}\{h_\ell\})^2 \right] \qquad (104)$$

$$\geq \left[\mathbf{J}(\boldsymbol{\theta})^{-1} \right]_{\ell\ell} + \left[\mathbf{J}(\boldsymbol{\theta})^{-1} \right]_{(\ell+n_{\mathrm{b}})(\ell+n_{\mathrm{b}})}. \qquad (105)$$

For the case of fully dependent delays, there is only one delay parameter $d = d_1$. However, the above derivation still holds if we redefine the parameter vector as $\boldsymbol{\theta} = [\mathrm{Re}\{\mathbf{h}^T\}, \mathrm{Im}\{\mathbf{h}^T\}, d]^T \in \mathbb{R}^{2n_{\mathrm{b}}+1}$, set $d_\ell = d_1$ for $\ell = 2, \ldots, n_{\mathrm{b}}$, and restrict (100) and (103) to hold only for $\ell = 1$. Of course, the dimension of the Fisher information matrix is also reduced to be $(2n_{\mathrm{b}}+1) \times (2n_{\mathrm{b}}+1)$.

REFERENCES

[1] A. O. Kislal, M. Rajiv, G. Durisi, E. G. Ström, and U. Mitra, "Pilot-assisted URLLC links: Impact of synchronization error," in *Proc. IEEE Int. Conf. Commun. (ICC)*, Denver CO, USA, Jun. 2024, pp. 617–622.

[2] 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)*, Montreal, QC, Canada, Aug. 2021, pp. 1–6.

[3] 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)*, Tainan, Taiwan, Sep. 2020, pp. 24–26.

[4] 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.

[5] "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

[6] 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.

[7] 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.

[8] M. C. Coşkun, G. Durisi, T. Jerkovits, G. Liva, W. Ryan, B. Stein, and F. Steiner, "Efficient error-correcting codes in the short blocklength regime," *Physical Communication*, vol. 34, pp. 66–79, 2019.

[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] 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.

[12] 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 *IEEE Globecom Workshops (GC Wkshps)*, Rio de Janeiro, Brazil, Dec. 2022, pp. 419–425.

[13] 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.

[14] 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.

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

[16] 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.

[17] 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.

[18] 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.

[19] 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.

[20] W. Feller, *An Introduction to Probability Theory and Its Applications*. New York, NY, USA: Wiley, 1971, vol. 2.

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

[23] 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.

[24] Y. Polyanskiy, "Asynchronous communication: Exact synchronization, universality, and dispersion," *IEEE Trans. Inf. Theory*, vol. 59, no. 3, pp. 1256–1270, 2013.

[25] 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.

[26] J. G. Proakis, *Digital Communications*, 4th ed. Boston, MA, USA: McGraw-Hill, 2008.

[27] "R1-1720997: On PDCCH for ultra-reliable transmission," 3GPP, Tech. Rep., Nov. 2017. [Online]. Available: https://www.3gpp.org/ftp/TSG_RAN/WG1_RL1/TSGR1_91//docs/

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

---

[12]This is not a strong assumption, since $E_\ell = 0$ occurs with probability zero.