Lancho, A., Östman, J., Durisi, G. & Sanguinetti, L. (14-23 June 2021). *A finite-blocklength analysis for URLLC with Massive MIMO* [proceedings]. IEEE International Conference on Communications, Montreal, Canada.

DOI: 10.1109/ICC42927.2021.9500881

A Finite-Blocklength Analysis for URLLC with Massive MIMO

Alejandro Lancho*, Johan Östman*, Giuseppe Durisi* and Luca Sanguinetti†
*Chalmers University of Technology, Gothenburg, Sweden
†University of Pisa, Pisa, Italy
Emails: {lanchoa,johanos,durisi}@chalmers.se, luca.sanguinetti@unipi.it

***Abstract*—This paper presents a rigorous finite-blocklength framework for the characterization and the numerical evaluation of the packet error probability achievable in the uplink and downlink of Massive MIMO for ultra-reliable low-latency communications (URLLC). The framework encompasses imperfect channel-state information, pilot contamination, spatially correlated channels, and arbitrary linear signal processing. For a practical URLLC network setup involving base stations with $M = 100$ antennas, we show by means of numerical results that a target error probability of 10^{-5} can be achieved with MMSE channel estimation and multicell MMSE signal processing, uniformly over each cell, only if orthogonal pilot sequences are assigned to all the users in the network. For the same setting, an alternative solution with lower computational complexity, based on least-squares channel estimation and regularized zero-forcing signal processing, does not suffice unless M is increased significantly.**

I. Introduction

Ultra-reliable low-latency communications (URLLC) is one of the most challenging use cases that will be supported by the next generation of wireless communications systems. One example is factory automation [2], where small payloads on the order of 100 bits must be delivered within hundreds of microseconds and with a reliability no smaller than 99.999%. Exploiting diversity is crucial to achieve this high reliability requirements. The spatial diversity offered by multiple antennas may thus play a key role. The latest instantiation of multiple antenna technologies is the so-called Massive MIMO [3]. Despite the intense research performed during the last years, Massive MIMO results have mainly been established in the *ergodic regime*, where the propagation channel evolves according to a block-fading model, and each codeword spans an increasingly large number of independent fading realizations as the codeword length goes to infinity (infinite-blocklength regime). These assumptions are highly questionable in URLLC scenarios [4].

A different approach was followed in [5], [6] where it is assumed that the fading channel stays constant during the transmission of a codeword (so-called quasi-static fading scenario) and outage capacity [7] is used as asymptotic performance metric. Although the quasi-static fading scenario is relevant for URLLC, the infinite blocklength assumption may yield incorrect estimates of the error probability. A theoretically satisfying framework can be obtained by assuming the use of a mismatch receiver that treats the channel estimate, obtained using a fixed number of pilot symbols, as perfect. One difficulty is that a fundamental result commonly used in the ergodic case to bound the mutual information by treating the channel estimation error as noise (see, e.g., [8, Lemma B.0.1]), does not apply to the outage case.

The limitation of both ergodic and outage setups can be resolved by performing a nonasymptotic analysis of the error probability based on the finite-blocklength information-theoretic bounds introduced in [9] and extended to fading channels in [10]–[12]. This approach has been pursued in the literature. However, most of the available results [13], [14] rely on the so called *normal approximation* [9, Eq. (291)], whose applicability for URLLC is questionable [1].

To verify if the design guidelines developed for Massive MIMO in the context of non-delay limited, large-throughput, communication links apply also to the URLLC setup, we present a rigorous nonasymptotic characterization of the packet error probability achievable in Massive MIMO. Specifically, we provide a firm upper bound on the error probability, which is obtained by adapting the random-coding union bound with parameter s (RCUs) introduced in [15] to the case of Massive MIMO communications. The resulting bound applies to Gaussian codebooks, and holds for any linear processing scheme and any pilot-based channel estimation scheme.

We then use the bound to evaluate the error probability in the uplink (UL) and downlink (DL) of a Massive MIMO network, with imperfect channel state information, pilot contamination, and spatially correlated channels. Numerical results are used to evaluate the *network availability*, which we define as the fraction of user equipment (UE) placements for which the per-link error probability, averaged over the small-scale fading and the additive noise, is below a given target. In our numerical analysis, we consider a realistic URLLC multicell multiuser setup where each base station (BS) has $M = 100$ antennas and the target per-link error probability is 10^{-5}. For this setup, we show that a network availability above 90% in both UL and DL can be achieved with minimum mean-square error (MMSE) channel estimation and multicell MMSE (M-MMSE) signal processing, only when pilot contamination is avoided. If the number of BS antennas is doubled, i.e., $M = 200$, comparable

This work was partly supported by the Swedish Research Council under grant 2016-03293 and by the Wallenberg AI, autonomous systems, and software program. The extended version of this paper, which includes all the proofs and additional design insights based on the analytical analyses and simulations, is available online [1].

performance can be achieved with regularized zero-forcing (RZF) spatial processing. For maximum ratio (MR) spatial processing to achieve a 90% network availability, the number of antennas must be increased further.

Notation: Lower-case bold letters are used for vectors and upper-case bold letters are used for matrices. The circularly-symmetric Gaussian distribution is denoted by $\mathcal{CN}(0,\sigma^2)$, where σ^2 denotes the variance. We use $\mathbb{E}[\cdot]$ to indicate the expectation operator, and $\mathbb{P}[\cdot]$ for the probability of a set. The natural logarithm is denoted by $\log(\cdot)$.

II. A Finite-Blocklength Upper-Bound on the Error Probability

We start by considering a simple channel model which will constitute the building block for the analysis of the error probability achievable in the Massive MIMO network considered in Section III. Specifically, we focus on the discrete-time additive white Gaussian noise (AWGN) channel given by

$$v[k] = gq[k] + z[k], \quad k = 1,\dots,n. \tag{1}$$

Here, $q[k] \in \mathbb{C}$ and $v[k] \in \mathbb{C}$ are the input and output over channel use k, respectively, and n is the codeword length. Furthermore, $g \in \mathbb{C}$ is the channel gain, which is assumed to remain constant during transmission of the n-length codeword. The additive noise variables $\{z[k] \in \mathbb{C}, k = 1,\dots,n\}$, are independent and identically distributed (i.i.d.), $\mathcal{CN}(0,\sigma^2)$ random variables. In what follows, we assume that:

1) The receiver *does not know* the channel gain g but has an estimate $\widehat{g}$ of g that is treated as perfect.
2) To determine the transmitted codeword $\mathbf{q} = [q[1],\dots,q[n]]^{\mathrm{T}} \in \mathcal{C}$, the receiver seeks the codeword $\widetilde{\mathbf{q}}$ from the codebook $\mathcal{C}$ of size $|\mathcal{C}| = m$ with elements in $\mathbb{C}^n$ that, once scaled by $\widehat{g}$, is the closest to the received vector $\mathbf{v} = [v[1],\dots,v[n]]^{\mathrm{T}} \in \mathbb{C}^n$ in Euclidean distance. Mathematically, the estimated codeword $\widehat{\mathbf{q}}$ is obtained as

$$\widehat{\mathbf{q}} = \arg\min_{\widetilde{\mathbf{q}}\in\mathcal{C}} \|\mathbf{v} - \widehat{g}\widetilde{\mathbf{q}}\|^2. \tag{2}$$

A receiver operating according to (2) is known as mismatched scaled nearest-neighbor (SNN) decoder [8]. Note that it coincides with the optimal maximum likelihood decoder *if and only if* $\widehat{g} = g$.

We are interested in deriving an upper bound on the error probability $\epsilon = \mathbb{P}[\widehat{\mathbf{q}} \neq \mathbf{q}]$ achieved by the SNN decoding rule (2). To do so, we follow a standard practice in information theory and use a random-coding approach. Specifically, we consider a Gaussian random code ensemble, where the elements of each codeword are drawn independently from a $\mathcal{CN}(0,\rho)$ distribution.[1] Here, ρ can be thought of as the average transmit power. We consider the cases where the channel gain g in (1) can be modelled as a constant or as a random variable. This latter case is commonly referred to in the literature as quasi-static fading setting.

[1]Note that this ensemble is not optimal at finite blocklength, not even if $\widehat{g} = g$. However, it is commonly used to obtain tractable expressions and insights into the performance of communication systems. Our analysis can be extended to other ensembles.

Theorem 1: Assume that $g \in \mathbb{C}$ and $\widehat{g} \in \mathbb{C}$ in (1) are deterministic. There exists a coding scheme with m codewords of length n operating according to the mismatched SNN decoding rule (2), whose error probability ϵ is upper-bounded by

$$\epsilon \le \mathbb{P}\left[\sum_{k=1}^{n} \imath_s(q[k],v[k]) \le \log\left(\frac{m-1}{u}\right)\right] \tag{3}$$

for all $s > 0$. Here, u is a random variable that is uniformly distributed over the interval $[0,1]$ and $\imath_s(q[k],v[k])$ is the generalized information density, given by

$$\imath_s(q[k],v[k]) = -s\,|v[k] - \widehat{g}q[k]|^2 + \frac{s|v[k]|^2}{1+s\rho|\widehat{g}|^2} + \log\bigl(1 + s\rho|\widehat{g}|^2\bigr)\,. \tag{4}$$

Assume now that $g \in \mathbb{C}$ and $\widehat{g} \in \mathbb{C}$ in (1) are random variables drawn according to an arbitrary joint distribution. Then, for all $s > 0$, the error probability ϵ is upper-bounded by

$$\epsilon \le \mathbb{E}\left[\mathbb{P}\left[\sum_{k=1}^{n} \imath_s(q[k],v[k]) \le \log\left(\frac{m-1}{u}\right) \middle| g,\widehat{g}\right]\right] \tag{5}$$

where the average is taken over the joint distribution of g and $\widehat{g}$. If $g \in \mathbb{C}$ is a random variable and $\widehat{g} \in \mathbb{C}$ is a constant,[2] the average in (5) is only taken over the distribution of g.

Proof: The proof for the case of g and $\widehat{g}$ being deterministic follows by particularizing the RCUs bound introduced in [15, Th. 1] to the considered setup. The upper bound for random g and $\widehat{g}$ readily follows by taking an expectation over the joint distribution of g and $\widehat{g}$. See [1, App. A] for further details. ■

The upper bounds in (3) and (5) involve the evaluation of a tail probability that is not known in closed form and needs to be evaluated numerically. Furthermore, they can be tightened by performing an optimization over the parameter $s > 0$, which also needs to be performed numerically. All this is computational demanding, especially when one targets the low error probabilities required in URLLC applications. We can alleviate this issue by evaluating instead *saddlepoint approximations* of (3) and (5). The resulting approximations, which can be found in [1, Th. 2], turn out to be easy to compute and accurate for a large range of system parameters of practical interest. On the contrary, normal approximations of (3) and (5) resulting from the use of the central limit theorem (such as the ones used in [13]), turn out to be often inaccurate.

III. Massive MIMO Network

We consider a Massive MIMO network with L cells, each comprising a BS with M antennas and K UEs, operating according to a time-division duplex (TDD) protocol that utilizes channel reciprocity. We denote by $\mathbf{h}_{li}^{j} \in \mathbb{C}^M$ the channel vector between UE i in cell l and the BS in cell j. We use a correlated Rayleigh fading model where $\mathbf{h}_{li}^{j} \sim \mathcal{CN}(\mathbf{0}_M, \mathbf{R}_{li}^{j})$ remains constant for the duration of a UL-DL transmission round. The normalized trace $\beta_{li}^{j} = \mathrm{tr}(\mathbf{R}_{li}^{j})/M$ determines

[2]This case will turn out important to analyze the DL of Massive MIMO networks.

the average large-scale fading between UE i in cell l and the BS in cell j, while the eigenstructure of $\mathbf{R}_{li}^{j}$ describes its spatial channel correlation [16, Sec. 2.2]. The $n_{\rm p}$-length pilot sequence of UE i in cell l is denoted by the vector $\boldsymbol{\phi}_{li} \in \mathbb{C}^{n_{\rm p}}$ and satisfies $\|\boldsymbol{\phi}_{li}\|^2 = n_{\rm p}$. We assume that the K UEs in a cell use mutually orthogonal pilot sequences and these pilot sequences are reused in a fraction $1/f$ of the L cells with $n_{\rm p} = Kf$. The channel vectors are estimated using either a least-squares (LS) estimator or an MMSE estimator. For the latter estimator, we assume that the channel statistics are known at the BS.

A. Uplink pilot transmission

We consider the standard TDD Massive MIMO protocol, where the UL and DL transmissions are assigned n channel uses in total, divided in $n_{\rm p}$ channel uses for UL pilots, $n_{\rm ul}$ channel uses for UL data, and $n_{\rm dl} = n - n_{\rm p} - n_{\rm ul}$ channel uses for DL data. We assume that the $n_{\rm p}$-length pilot sequence $\boldsymbol{\phi}_{li} \in \mathbb{C}^{n_{\rm p}}$ with $\boldsymbol{\phi}_{li}^{\rm H}\boldsymbol{\phi}_{li} = n_{\rm p}$ is used by UE i in cell l for channel estimation. The elements of $\boldsymbol{\phi}_{li}$ are scaled by the square-root of the pilot power $\sqrt{\rho^{\rm ul}}$ and transmitted over $n_{\rm p}$ channel uses. When the UEs in cell j transmit their pilot sequences, the received pilot signal at cell j, $\mathbf{Y}_j^{\rm pilot} \in \mathbb{C}^{M\times n_{\rm p}}$, is

$$\mathbf{Y}_j^{\rm pilot} = \sum_{i=1}^{K} \sqrt{\rho^{\rm ul}}\mathbf{h}_{ji}^{j}\boldsymbol{\phi}_{ji}^{\rm T} + \sum_{l=1,l\neq j}^{L}\sum_{i=1}^{K}\mathbf{h}_{li}^{j}\boldsymbol{\phi}_{li}^{\rm T} + \mathbf{Z}_j^{\rm pilot} \quad (6)$$

where $\mathbf{Z}_j^{\rm pilot} \in \mathbb{C}^{M\times n_{\rm p}}$ is additive noise with i.i.d. elements distributed as $\mathcal{CN}(0,\sigma_{\rm ul}^2)$. The LS estimate of $\mathbf{h}_{li}^{j}$ is [16, Sec. 3.4]

$$\widehat{\mathbf{h}}_{li}^{j} = \frac{1}{\sqrt{\rho^{\rm ul}}n_{\rm p}}\mathbf{Y}_j^{\rm pilot}\boldsymbol{\phi}_{li}. \quad (7)$$

Assuming that $\mathbf{R}_{li}^{j}$ is known at the BS, the MMSE estimate of $\mathbf{h}_{li}^{j}$ is [16, Sec. 3.2]

$$\widehat{\mathbf{h}}_{li}^{j} = \sqrt{\rho^{\rm ul}n_{\rm p}}\mathbf{R}_{li}^{j}\mathbf{Q}_{li}^{j}\left(\mathbf{Y}_j^{\rm pilot}\boldsymbol{\phi}_{li}\right) \quad (8)$$

with

$$\mathbf{Q}_{li}^{j} = \left(\sum_{l'=1}^{L}\sum_{i'=1}^{K}\rho^{\rm ul}\mathbf{R}_{l'i'}^{j}\boldsymbol{\phi}_{l'i'}^{\rm H}\boldsymbol{\phi}_{li} + \sigma_{\rm ul}^2\mathbf{I}_M\right)^{-1}. \quad (9)$$

The MMSE estimate $\widehat{\mathbf{h}}_{li}^{j}$ and the estimation error $\widetilde{\mathbf{h}}_{li}^{j} = \mathbf{h}_{li}^{j} - \widehat{\mathbf{h}}_{li}^{j}$ are independent random vectors, distributed as $\widehat{\mathbf{h}}_{li}^{j} \sim \mathcal{CN}(\mathbf{0}_M, \boldsymbol{\Phi}_{li}^{j})$ and $\widetilde{\mathbf{h}}_{li}^{j} \sim \mathcal{CN}(\mathbf{0}_M, \mathbf{R}_{li}^{j} - \boldsymbol{\Phi}_{li}^{j})$, respectively, with $\boldsymbol{\Phi}_{li}^{j} = \rho^{\rm ul}n_{\rm p}\mathbf{R}_{li}^{j}\mathbf{Q}_{li}^{j}\mathbf{R}_{li}^{j}$.

B. Uplink data transmission

During UL data transmission, the received complex baseband signal $\mathbf{y}_j^{\rm ul}[k] \in \mathbb{C}^M$ in cell j over an arbitrary channel use k, where $k = 1,\ldots,n_{\rm ul}$, is given by

$$\mathbf{y}_j^{\rm ul}[k] = \mathbf{h}_{ji}^{j}x_{ji}^{\rm ul}[k] + \sum_{i'=1,i'\neq i}^{K}\mathbf{h}_{ji'}^{j}x_{ji'}^{\rm ul}[k] + \sum_{l=1,l\neq j}^{L}\sum_{i'=1}^{K}\mathbf{h}_{li'}^{j}x_{li'}^{\rm ul}[k] + \mathbf{z}_j^{\rm ul}[k] \quad (10)$$

where $x_{ji}^{\rm ul}[k] \sim \mathcal{CN}(0,\rho^{\rm ul})$ is the information bearing signal[3] transmitted by UE i in cell j with $\rho^{\rm ul}$ being the average UL transmit power, and $\mathbf{z}_j^{\rm ul}[k] \sim \mathcal{CN}(\mathbf{0}_M, \sigma_{\rm ul}^2\mathbf{I}_M)$ is the independent additive noise. To detect $x_{ji}[k]$, BS j selects the receive combining vector $\mathbf{v}_{ji} \in \mathbb{C}^M$, which is multiplied with the received signal $\mathbf{y}_j[k]$ to obtain the combined output

$$r_{ji}^{\rm ul}[k] = \mathbf{v}_{ji}^{\rm H}\mathbf{y}_j^{\rm ul}[k], \quad k = 1,\ldots,n_{\rm ul}. \quad (11)$$

Note that (11) can be expressed in the same form as (1) if we set $v[k] = r_{ji}^{\rm ul}[k]$, $q[k] = x_{ji}^{\rm ul}[k]$, $g = \mathbf{v}_{ji}^{\rm H}\mathbf{h}_{ji}^{j}$, $\widehat{g} = \mathbf{v}_{ji}^{\rm H}\widehat{\mathbf{h}}_{ji}^{j}$, and $z[k] = \sum_{i'=1,i'\neq i}^{K}\mathbf{v}_{ji}^{\rm H}\mathbf{h}_{ji'}^{j}x_{ji'}^{\rm ul}[k] + \sum_{l=1,l\neq j}^{L}\sum_{i'=1}^{K}\mathbf{v}_{ji}^{\rm H}\mathbf{h}_{li'}^{j}x_{li'}^{\rm ul}[k] + \mathbf{v}_{ji}^{\rm H}\mathbf{z}_j[k]$. Given all channels and combining vectors, the random variables $\{z[k] : k = 1,\ldots,n_{\rm ul}\}$ are conditionally i.i.d. and $z[k] \sim \mathcal{CN}(0,\sigma^2)$ with $\sigma^2 = \sigma_{\rm ul}^2\|\mathbf{v}_{ji}\|^2 + \rho^{\rm ul}\sum_{i'=1,i'\neq i}^{K}|\mathbf{v}_{ji}^{\rm H}\mathbf{h}_{ji'}^{j}|^2 + \rho^{\rm ul}\sum_{l=1,l\neq j}^{L}\sum_{i'=1}^{K}|\mathbf{v}_{ji}^{\rm H}\mathbf{h}_{li'}^{j}|^2$.

We assume that the BS treats the acquired (noisy) channel estimate $\widehat{\mathbf{h}}_{ji}^{j}$ as perfect. This implies that, to recover the transmitted codeword, which we assume to be drawn from a codebook $\mathcal{C}^{\rm ul}$, it performs mismatched SNN decoding with $\widehat{g} = \mathbf{v}_{ji}^{\rm H}\widehat{\mathbf{h}}_{ji}^{j}$. Specifically, the estimated codeword $\widehat{\mathbf{x}}_{ji}^{\rm ul}$ is obtained as

$$\widehat{\mathbf{x}}_{ji}^{\rm ul} = \arg\min_{\widetilde{\mathbf{x}}_{ji}^{\rm ul}\in\mathcal{C}^{\rm ul}}\|\mathbf{r}_j^{\rm ul} - (\mathbf{v}_{ji}^{\rm H}\widehat{\mathbf{h}}_{ji}^{j})\widetilde{\mathbf{x}}_{ji}^{\rm ul}\|^2 \quad (12)$$

with $\mathbf{r}_j^{\rm ul} = [r_j^{\rm ul}[1],\ldots,r_j^{\rm ul}[n_{\rm ul}]]^{\rm T}$ and $\widetilde{\mathbf{x}}_{ji}^{\rm ul} = [\widetilde{x}_{ji}^{\rm ul}[1],\ldots,\widetilde{x}_{ji}^{\rm ul}[n_{\rm ul}]]^{\rm T}$. It thus follows that (3) provides a bound on the conditional error probability for UE i in cell j given g and $\widehat{g}$. To obtain the average error probability, we need to take an expectation over $g = \mathbf{v}_{ji}^{\rm H}\mathbf{h}_{ji}^{j}$, $\widehat{g} = \mathbf{v}_{ji}^{\rm H}\widehat{\mathbf{h}}_{ji}^{j}$, and $\sigma^2 = \sigma_{\rm ul}^2\|\mathbf{v}_{ji}\|^2 + \rho^{\rm ul}\sum_{i'=1,i'\neq i}^{K}|\mathbf{v}_{ji}^{\rm H}\mathbf{h}_{ji'}^{j}|^2 + \rho^{\rm ul}\sum_{l=1,l\neq j}^{L}\sum_{i'=1}^{K}|\mathbf{v}_{ji}^{\rm H}\mathbf{h}_{li'}^{j}|^2$, which results in

$$\epsilon_{ji}^{\rm ul} \leq \mathbb{E}\left[\mathbb{P}\left[\sum_{k=1}^{n_{\rm ul}}\imath_s(r_{ji}^{\rm ul}[k], x_{ji}^{\rm ul}[k]) \leq \log\left(\frac{m-1}{u}\right)\middle| g,\widehat{g},\sigma^2\right]\right]. \quad (13)$$

The combining vector $\mathbf{v}_{ji}$ is selected at the BS based on the channel estimates $\widehat{\mathbf{h}}_{li}^{j}$. As shown in [16], the M-MMSE combining

$$\mathbf{v}_{ji}^{\rm M-MMSE} = \left(\sum_{l=1}^{L}\sum_{i'=1}^{K}\widehat{\mathbf{h}}_{li'}^{j}\left(\widehat{\mathbf{h}}_{li'}^{j}\right)^{\rm H} + \mathbf{Z}_j\right)^{-1}\widehat{\mathbf{h}}_{ji}^{j} \quad (14)$$

with $\mathbf{Z}_j = \sum_{l=1}^{L}\sum_{i'=1}^{K}\boldsymbol{\Phi}_{li'}^{j} + \frac{\sigma_{\rm ul}^2}{\rho^{\rm ul}}\mathbf{I}_M$ yields often satisfactory performance at the cost of high computational complexity. A combiner with lower computational complexity is RZF:

$$\mathbf{v}_{ji}^{\rm RZF} = \left(\sum_{i'=1}^{K}\widehat{\mathbf{h}}_{ji'}^{j}\left(\widehat{\mathbf{h}}_{ji'}^{j}\right)^{\rm H} + \frac{\sigma_{\rm ul}^2}{\rho^{\rm ul}}\mathbf{I}_M\right)^{-1}\widehat{\mathbf{h}}_{ji}^{j}. \quad (15)$$

The combiner with the lowest computational complexity is

[3]As detailed in Section II, we will evaluate the error probability for a Gaussian random code ensemble, where the elements of each codeword are drawn independently from a $\mathcal{CN}(0,\rho^{\rm ul})$ distribution.

MR, which results in $\mathbf{v}_{ji}^{\mathrm{MR}} = \widehat{\mathbf{h}}_{ji}^{j}/M$. For further information about combiners and when they are expected to provide good performance, see [16, Sec. 4.1].

C. Downlink data transmission

Assume that, to transmit to UE i in cell j, the BS in cell j uses the precoding vector $\mathbf{w}_{ji} \in \mathbb{C}^M$, which determines the spatial directivity of the transmission and satisfies the normalization $\mathbb{E}\left[\|\mathbf{w}_{ji}\|^2\right] = 1$. During DL data transmission, the received signal $y_{ji}^{\mathrm{dl}}[k] \in \mathbb{C}$ at UE i in cell j over channel use k, where $k = 1, \ldots, n_{\mathrm{dl}}$, is given by

$$y_{ji}^{\mathrm{dl}}[k] = (\mathbf{h}_{ji}^{j})^{\mathrm{H}} \mathbf{w}_{ji} x_{ji}^{\mathrm{dl}}[k] + \sum_{i'=1, i' \neq i}^{K} (\mathbf{h}_{ji}^{j})^{\mathrm{H}} \mathbf{w}_{ji'} x_{ji'}^{\mathrm{dl}}[k] + \sum_{l=1, l \neq j}^{L} \sum_{i'=1}^{K} (\mathbf{h}_{ji}^{l})^{\mathrm{H}} \mathbf{w}_{li'} x_{li'}^{\mathrm{dl}}[k] + z_{ji}^{\mathrm{dl}}[k] \tag{16}$$

where $x_{ji}^{\mathrm{dl}}[k] \sim \mathcal{CN}(0, \rho^{\mathrm{dl}})$ is the data signal intended for UE i in cell j and $z_{ji}^{\mathrm{dl}}[k] \sim \mathcal{CN}(0, \sigma_{\mathrm{dl}}^2)$ is the receiver noise at UE i in cell j. Again, we can put (16) in the same form as (1) by setting $v[k] = y_{ji}^{\mathrm{dl}}[k]$, $q[k] = x_{ji}^{\mathrm{dl}}[k]$, $g = (\mathbf{h}_{ji}^{j})^{\mathrm{H}} \mathbf{w}_{ji}$, $\widehat{g} = \mathbb{E}[(\mathbf{h}_{ji}^{j})^{\mathrm{H}} \mathbf{w}_{ji}]$ and $z[k] = \sum_{i'=1, i \neq i}^{K} (\mathbf{h}_{ji}^{j})^{\mathrm{H}} \mathbf{w}_{ji'} x_{ji'}^{\mathrm{dl}}[k] + \sum_{l=1, l \neq j}^{L} \sum_{i'=1}^{K} (\mathbf{h}_{ji}^{l})^{\mathrm{H}} \mathbf{w}_{li'} x_{li'}^{\mathrm{dl}}[k] + z_{ji}^{\mathrm{dl}}[k]$. Note that, given all channels and precoding vectors, the random variables $\{z[k] : k = 1, \ldots, n_{\mathrm{dl}}\}$ are conditionally i.i.d. and $z[k] \sim \mathcal{CN}(0, \sigma^2)$ with $\sigma^2 = \sigma_{\mathrm{dl}}^2 + \rho^{\mathrm{dl}} \sum_{i'=1, i \neq i'}^{K} |(\mathbf{h}_{ji}^{j})^{\mathrm{H}} \mathbf{w}_{ji'}|^2 + \rho^{\mathrm{dl}} \sum_{l=1, l \neq j}^{L} \sum_{i'=1}^{K} |(\mathbf{h}_{ji}^{l})^{\mathrm{H}} \mathbf{w}_{li'}|^2$.

Since no pilots are transmitted in the DL, the UE does not know the precoded channel $g = (\mathbf{h}_{ji}^{j})^{\mathrm{H}} \mathbf{w}_{ji}$ in (16). Instead, we assume that the UE has access to its expected value $\mathbb{E}[(\mathbf{h}_{ji}^{j})^{\mathrm{H}} \mathbf{w}_{ji}]$ and uses this quantity to perform mismatched SNN decoding. Specifically, we have that $\widehat{g} = \mathbb{E}[(\mathbf{h}_{ji}^{j})^{\mathrm{H}} \mathbf{w}_{ji}]$ and

$$\widehat{\mathbf{x}}_{ji}^{\mathrm{dl}} = \arg \min_{\widetilde{\mathbf{x}}_{ji}^{\mathrm{dl}} \in \mathcal{C}^{\mathrm{dl}}} \|\mathbf{y}_{ji}^{\mathrm{dl}} - \mathbb{E}[(\mathbf{h}_{ji}^{j})^{\mathrm{H}} \mathbf{w}_{ji}] \widetilde{\mathbf{x}}_{ji}^{\mathrm{dl}}\|^2 \tag{17}$$

with $\mathbf{y}_{ji}^{\mathrm{dl}} = [y_{ji}^{\mathrm{dl}}[1], \ldots, y_{ji}^{\mathrm{dl}}[n_{\mathrm{dl}}]]^{\mathrm{T}}$ and $\widetilde{\mathbf{x}}_{ji}^{\mathrm{dl}} = [\widetilde{x}_{ji}^{\mathrm{dl}}[1], \ldots, \widetilde{x}_{ji}^{\mathrm{dl}}[n_{\mathrm{dl}}]]^{\mathrm{T}}$. Obviously, channel hardening is critical for this choice to result in good performance [16, Sec. 2.5.1]. Since $\widehat{g} = \mathbb{E}[(\mathbf{h}_{ji}^{j})^{\mathrm{H}} \mathbf{w}_{ji}]$ is deterministic, the error probability at UE i in cell j in the DL can be evaluated as follows:

$$\epsilon_{ji}^{\mathrm{dl}} \leq \mathbb{E}\left[\mathbb{P}\left[\sum_{k=1}^{n_{\mathrm{dl}}} \imath_s(y_{ji}^{\mathrm{dl}}[k], x_{ji}^{\mathrm{dl}}[k]) \leq \log\left(\frac{m-1}{u}\right) \middle| g, \sigma^2\right]\right]. \tag{18}$$

Similar to the UL, the upper bound (18) holds for any precoding vector that is selected on the basis of the channel estimates available at the BS. Different precoders yield different tradeoffs between the error probability achievable at the UEs. A common heuristic comes from UL-DL duality [16, Sec. 4.3.2], which suggests to choose the precoding vectors $\mathbf{w}_{ji}$ as the following function of the combining vectors: $\mathbf{w}_{ji} = \mathbf{v}_{ji}/\sqrt{\mathbb{E}[\|\mathbf{v}_{ji}\|^2]}$. By selecting $\mathbf{v}_{ji}$ as one of the uplink combining schemes described earlier, the corresponding precoding scheme is obtained, for example, $\mathbf{v}_{ji} = \mathbf{v}_{ji}^{\mathrm{M-MMSE}}$ yields M-MMSE precoding.

D. Numerical Analysis

The simulation setup consists of $L = 4$ square cells, each of size 75 m × 75 m, containing $K = 10$ UEs each, independently and uniformly distributed within the cell, at a distance of at least 5 m from the BS. We consider a horizontal uniform linear array with $M = 100$ antennas and half-wavelength spacing. The antennas and the UEs are located in the same horizontal plane, thus the azimuth angle is sufficient to determine the directivity. We assume that the scatterers are uniformly distributed in the angular interval $[\varphi_{li} - \Delta, \varphi_{li} + \Delta]$, where φ_{li} is the nominal angle-of-arrival of UE i in cell l and Δ is the angular spread. Hence, the (m_1, m_2)th element of $\mathbf{R}_{li}^{j}$ is equal to [16, Sec. 2.6]

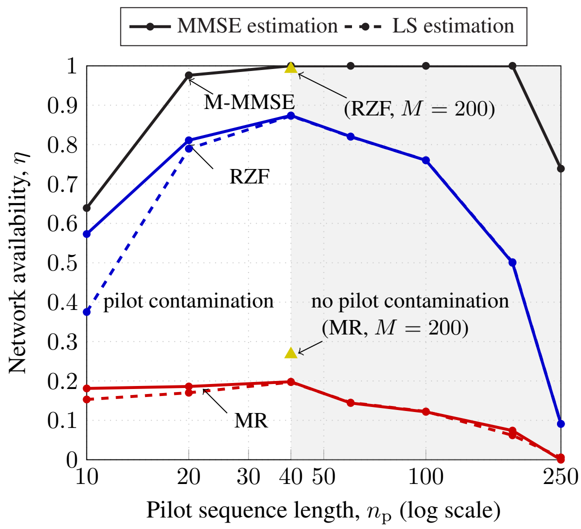
$$\left[\mathbf{R}_{li}^{j}\right]_{m_1, m_2} = \frac{\beta_{li}^{j}}{2\Delta} \int_{-\Delta}^{\Delta} e^{\mathrm{j}\pi(m_1 - m_2)\sin(\varphi_{li} + \bar{\varphi})} d\bar{\varphi}. \tag{19}$$

We assume $\Delta = 25°$ and let the large-scale fading coefficient, measured in dB, be $\beta_{li}^{j} = -35.3 - 37.6 \log_{10}(\frac{d_{li}^{j}}{1\,\mathrm{m}})$ where d_{li}^{j} is the distance between the BS in cell j and UE i in cell l. The communication takes place over a 20 MHz bandwidth with a total receiver noise power of $\sigma_{\mathrm{ul}}^2 = \sigma_{\mathrm{dl}}^2 = -94\,\mathrm{dBm}$ (consisting of thermal noise and a noise figure of 7 dB in the receiver hardware) at both the BS and UEs. The UL and DL transmit powers are equal and given by $\rho^{\mathrm{ul}} = \rho^{\mathrm{dl}} = 10$ dBm. Furthermore, we employ a wrap-around topology as in [16, Sec. 4.1.3]. We assume $n = 300$, $n_{\mathrm{ul}} = n_{\mathrm{dl}} = (n - n_{\mathrm{p}})/2$ and $\log_2 m = 160$ information bits, where m is the size of the UL and DL codebooks $\mathcal{C}^{\mathrm{ul}}$ and $\mathcal{C}^{\mathrm{dl}}$, respectively.

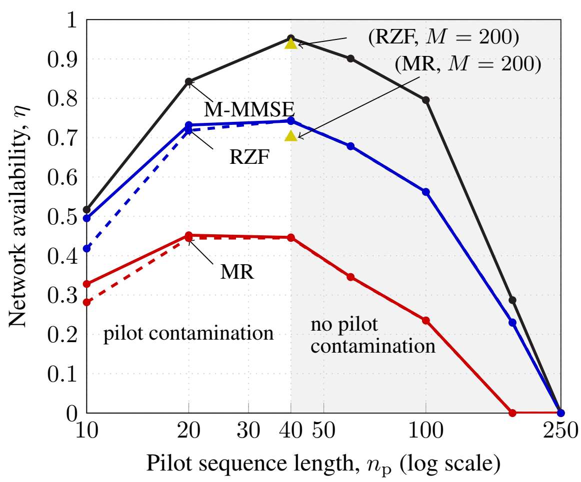
The average error probabilities $\epsilon_{ji}^{\mathrm{ul}}$ and $\epsilon_{ji}^{\mathrm{dl}}$ are computed for fixed UEs positions and averaged over the small-scale fading and the additive noise. The network availability is defined with respect to the random positions of the UEs as

$$\eta = \mathbb{P}[\epsilon \leq \epsilon_{\mathrm{target}}]. \tag{20}$$

In Fig. 1, we plot the network availability η for a fixed $\epsilon_{\mathrm{target}} = 10^{-5}$ versus the number of pilot symbols $n_{\mathrm{p}} = fK$ (recall that f is the pilot reuse factor) for MMSE (solid lines) and LS (dashed lines) channel estimation, and M-MMSE, RZF and MR combining/precoding. As shown in the figure, pilot contamination must be avoided and MMSE channel estimation with M-MMSE combining/precoding is necessary to achieve URLLC requirements for the chosen system parameters. Indeed, the largest network availability (above 90%) in both UL and DL is obtained by setting the pilot reuse factor f to 4 so that $n_{\mathrm{p}} = fK = 40$. This is the minimum value of n_{p} that results in no pilot contamination in a network with $L = 4$ cells and $K = 10$ users per cell. We can see the same behaviour when using MR and RZF, which however achieve a lower network availability (below 50% for MR in both UL and DL even when pilot contamination is avoided). Note that, in all cases, LS channel estimation achieves performance similar to MMSE channel estimation. If we consider $M = 200$ and $n_{\mathrm{p}} = 40$, the performance achieved with RZF is comparable to that of



(a) UL.



(b) DL.

Fig. 1: Network availability for $\epsilon_{\mathrm{target}} = 10^{-5}$. Here, $L = 4$, $K = 10$, $\Delta = 25°$, the cell size is 75×75 m, $\rho^{\mathrm{ul}} = \rho^{\mathrm{dl}} = 10\,\mathrm{dBm}$, $M = 100$ or 200, $\log_2 m = 160$, and $n = 300$.

M-MMSE with $M = 100$. For MR to achieve the network availability of M-MMSE, the number of antennas must be increased even further. For both MR and RZF with $M = 200$ and $n_\mathrm{p} = 40$, using LS and MMSE channel estimation yields indistinguishable performance. This is illustrated by the yellow markers in Fig. 1.

Note also that increasing n_p beyond 40 has a deleterious effect on the network availability, especially in the DL. Indeed, the corresponding reduction in the number of channel uses $n_{\mathrm{dl}} = (300 - n_\mathrm{p})/2$ available for data transmission in the DL overcomes the benefits of a more accurate channel state information (CSI). The UL achieves better performance compared to the DL since, differently from the BS, the UE has no CSI and performs mismatched decoding by relying on channel hardening.

IV. Conclusions

We have presented guidelines on the design of Massive MIMO systems supporting the transmission of short information packets under the high reliability targets demanded in URLLC. Specifically, we have shown that, for a BS equipped with 100 antennas, it is imperative to avoid pilot contamination and to use M-MMSE signal processing in order to achieve a network availability above 90%. We have also shown that lower complexity signal processing schemes such as LS estimation and RZF combining, achieves performance comparable to MMSE estimation and M-MMSE combining, provided that the BS employs a larger number of antennas. Our guidelines are based on a firm nonasymptotic bound on the error probability, which is based on recent results in finite-blocklength information theory, and applies to a realistic Massive MIMO network with imperfect channel state information, pilot contamination, spatially correlated channels, arbitrary linear signal processing, and randomly positioned UEs.

References

[1] J. Östman, A. Lancho, G. Durisi, and L. Sanguinetti, "URLLC with Massive MIMO: Analysis and design at finite blocklength," Sep. 2020. [Online]. Available: https://arxiv.org/abs/2009.10550

[2] 3GPP, "Service requirements for cyber-physical control applications in vertical domains," 3rd Generation Partnership Project (3GPP), Technical Specification (TS) 22.104, 12 2019, version 17.2.0.

[3] T. L. Marzetta, "Noncooperative cellular wireless with unlimited numbers of base station antennas," *IEEE Trans. Wireless Commun.*, vol. 9, no. 11, pp. 3590–3600, Nov. 2010.

[4] 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.

[5] M. Karlsson, E. Björnsson, and E. G. Larsson, "Performance of in-band transmission of system information in massive MIMO systems," *IEEE Trans. Wireless Commun.*, vol. 17, no. 3, pp. 1700–1712, Mar. 2018.

[6] A. Bana, G. Xu, E. D. Carvalho, and P. Popovski, "Ultra reliable low latency communications in massive multi-antenna systems," in *Proc. Asilomar Conf. Signals, Syst., Comput.*, Pacific Grove, CA, USA, Oct. 2018, pp. 188–192.

[7] L. H. Ozarow, S. Shamai (Shitz), and A. D. Wyner, "Information theoretic considerations for cellular mobile radio," *IEEE Trans. Veh. Technol.*, vol. 43, no. 2, pp. 359–378, May 1994.

[8] 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.

[9] 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.

[10] W. Yang, G. Durisi, T. Koch, and Y. Polyanskiy, "Quasi-static multiple-antenna fading channels at finite blocklength," *IEEE Trans. Inf. Theory*, vol. 60, no. 7, pp. 4232–4265, Jul. 2014.

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

[12] J. Östman, G. Durisi, E. G. Ström, M. C. Coskun, 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.

[13] J. Zeng, T. Lv, R. P. Liu, X. Su, Y. J. Guo, and N. C. Beaulieu, "Enabling ultra-reliable and low-latency communications under shadow fading by massive MU-MIMO," *IEEE Internet of Things J.*, vol. 7, no. 1, pp. 234–246, Jan. 2020.

[14] H. Ren, C. Pan, Y. Deng, M. Elkashlan, and A. Nallanathan, "Joint pilot and payload power allocation for massive-MIMO-enabled URLLC IIoT networks," *IEEE J. Sel. Areas Commun.*, vol. 38, no. 5, pp. 816–830, May 2020.

[15] A. Martinez and A. Guillén i Fàbregas, "Saddlepoint approximation of random–coding bounds," in *Proc. Inf. Theory Applicat. Workshop (ITA)*, San Diego, CA, USA, Feb. 2011.

[16] E. Björnson, J. Hoydis, and L. Sanguinetti, "Massive MIMO Networks: Spectral, Energy, and Hardware Efficiency," *Foundations and Trends® in Signal Processing*, vol. 11, no. 3-4, pp. 154–655, Nov. 2017.