SHORT PACKET STRUCTURE FOR ULTRA-RELIABLE MACHINE-TYPE COMMUNICATION: TRADEOFF BETWEEN DETECTION AND DECODING

Alexandru-Sabin Bana, Kasper Fløe Trillingsgaard, Petar Popovski, Elisabeth de Carvalho

Department of Electronic Systems, Aalborg University, Denmark

ABSTRACT

Machine-type communication requires rethinking of the structure of short packets due to the coding limitations and the significant role of the control information. In ultra-reliable low-latency communication (URLLC), it is crucial to optimally use the limited degrees of freedom (DoFs) to send data and control information. We consider a URLLC model for short packet transmission with acknowledgement (ACK). We compare the detection/decoding performance of two short packet structures: (1) time-multiplexed detection sequence and data; and (2) structure in which both packet detection and data decoding use all DoFs. Specifically, as an instance of the second structure we use superimposed sequences for detection and data. We derive the probabilities of false alarm and misdetection for an AWGN channel and numerically minimize the packet error probability (PER), showing that for delay-constrained data and ACK exchange, there is a tradeoff between the resources spent for detection and decoding. We show that the optimal PER for the superimposed structure is achieved for higher detection overhead. For this reason, the PER is also higher than in the preamble case. However, the superimposed structure is advantageous due to its flexibility to achieve optimal operation without the need to use multiple codebooks.

Index Terms— Short packets, URLLC, detection, decoding

1. INTRODUCTION

Machine-type communication (MTC) is central to the 5G systems, where it appears in two flavors: massive MTC (mMTC), focused on serving a large number of devices and ultra-reliable low-latency communications (URLLC), focused on serving a small number of devices with stringent latency and reliability constraints. The main vehicle of MTC is transmission of short packets, especially critical for achieving low latency [1]. Short packet MTC is significantly affected by the control information [2], which can be reduced through 5G grant-free access [3], where uplink (UL) short packets are sent in specific resources, without prior device scheduling.

Packet detection is the key auxiliary procedure that needs to be carried out with very high reliability when URLLC is considered. Detection and synchronization studies date back decades in the communications context [4]. Recent information-theoretic works have investigated the achievable tradeoff between the rate of reliable communication and the asynchronism exponent [5]. It has been shown that up to an asynchronism level, reliable communication is possible by joint detection and decoding [6] in the discrete memoryless channel (DMC). Furthermore, it has been shown that it is sub-optimal to optimize the detection and decoding separately [7], and that there is a tradeoff between the probabilities of false alarm, misdetection and decoding error. Joint detection and synchronization for a DSSS

system using differential encoding is treated in [8], where the probabilities of false alarm and misdetection are derived for the AWGN channel, and further extended to a multipath Rayleigh fading channel. The false alarm is considered to be impacting only if it occurs at a time $t < t_r$ before the actual packet arrival, where t_r is the time required to recover after a false alarm event. A comparison of preamble and superimposed packet structures was done in [9] for channel frequency offset (CFO) training in OFDM WLAN, where the robustness of the superimposed case was improved through a variable data rate for the information symbols. It was shown that for the same packet length as the preamble scenario, the superimposed structure obtains similar CFO estimation performance, while reducing the data decoding error probability. In [10], the authors show that there is a tradeoff between bandwidth, latency, reliability, and rate for short packets, using the information-theoretic results on finite-blocklength regime from [11]. The paper outlines the optimal way to exploit the available spatial and frequency diversity under a latency-reliability constraint. The interaction between error-control coding and channel estimation for a short packet scenario in an AWGN channel with unknown, constant gain over a block is investigated in [12], where it is shown that for a single-antenna receiver, there is an optimal training length for which the required signal-to-noise ratio (SNR) is minimized.

These results motivate our interest in the optimal tradeoff between detection overhead and decoding of a short packet, transmitted in a point-to-point system with acknowledgement (ACK), under strict latency-reliability constraints as in URLLC. Two packet structures are analyzed. The first one is the time-multiplexed structure, in which a fraction of the Degrees of Freedom (DoFs) is used as a preamble for detection and the remaining for data transmission. The second structure uses all available DoFs for detection and data. We treat the specific instance in which the detection sequence and data are superimposed. We derive the probabilities of false alarm and misdetection, and show the tradeoff between detection and decoding by numerically determining the optimal overhead that minimizes the packet error probability (PER), making use of results from finite blocklength information theory [11]. The results show that there is a tradeoff in allocating resources between detection and decoding. The preamble case achieves the minimum PER for a smaller overhead than the superimposed structure, which implies that more resources are needed for detection in the superimposed case. Therefore, the overall packet error achieved is slightly worse for the superimposed structure. However, from a pragmatic perspective, the superimposed structure offers an enhanced adaptivity, as it can operate optimally by simply controlling the power allocation for detection, rather than changing the preamble length and coding rate.

2. SYSTEM MODEL

We consider the case of a point-to-point round-trip exchange consisting of the transmission of a short packet of b bits and reception of
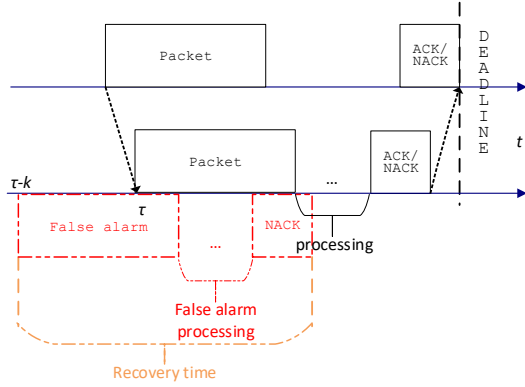
Fig. 1. The system model considered in this paper. The two time axes represent the time lines at the transmitter and receiver, respectively. The model consists of a one-shot transmission with ACK/NACK feedback within a given latency constraint. The cost of having a false alarm is also shown, which means the loss of an incoming packet within the recovery time window.

a positive or negative ACK under stringent latency-reliability constraints. The transmitter and the receiver are assumed to have established synchronization at the symbol level. The receiver does not know the precise packet arrival time τ, which must be estimated prior to attempting to decode the data. Furthermore, we assume that upon the generation of a packet, a strict deadline for reception of the acknowledgement at the device is enforced.

A transmission is made in one shot, no retransmissions, and uses all the available resources: the sum of the DoFs (channel uses) spent for the packet transmission, (N)ACK, receiver processing, and the round-trip time amount to the maximum allowed latency, see Fig. 1. In such a system there are three main sources of uncertainty: the device activity, the noise, and the channel fading coefficients. Throughout this paper, the channel gain is assumed to be constant during a packet exchange and known by both the transmitter and the receiver, such that the channel input-output relation can be modeled as an AWGN channel, $Y_j = X_j + W_j$. Here, $\{W_j\}$ are i.i.d. circularly-symmetric complex Gaussian random variables with zero mean and unit variance and X_j denotes the jth transmitted symbol, which is zero for $j \notin \{\tau, \dots, \tau+N-1\}$. Assuming a fixed ACK structure and known round-trip time, we focus on the transmitted packet structure, which contains the control information, the encoded b information bits. We assume that the b information bits include a cyclic redundancy check (CRC) which allows the receiver to detect errors. The round-trip PER can be expressed as

$$\mathcal{P}_e = 1 - (1 - \epsilon_{\text{d}})(1 - \epsilon_{\text{D}})(1 - \epsilon_{\text{ACK}}) \tag{1}$$

with ϵ_{d}, ϵ_{D} and ϵ_{ACK} being the probabilities of error for detecting the packet, decoding the data, and correct ACK reception, respectively. As we aim to investigate the interplay between the detection and decoding, we assume $\epsilon_{\text{ACK}} = 0$ and fixed number of DoFs for ACK. This leaves N channel uses for the detection sequence and the codeword that carries the data. Without loss of generality, we further abstract from the round-trip time for the remaining equations in the paper in order to improve the clarity of indexing.

The receiver operates in a sequential mode: it can either run detection procedures and buffer samples of up to one packet length N or transmit (N)ACK. Detection is performed using either time-multiplexed or superimposed Zadoff-Chu detection sequences [13], [14, Chapter 9], which have unit power per symbol and odd length. The root used to generate the sequence is chosen such that the

partial-period correlation is minimal for each sequence length. The coded information symbols, the noise, and the detection sequence symbols are further denoted by D_j, W_j, and p_j, respectively, where p_j is zero for $j \notin \{0, \dots, N_{\text{p}} - 1\}$ for the preamble case, and for $j \notin \{0, \dots, N - 1\}$ for superimposed. We sometimes let the vectors $\mathbf{D}$ and $\mathbf{p}$ denote the coded information symbols and preamble sequence, respectively, with superscripts denoting their length. The information symbols are assumed to be coded using spherical Gaussian codebooks (shell codes) for which the codewords are uniformly distributed on the shell of an N_{c}-dimensional sphere of radius $\sqrt{N_{\text{c}}}$, where N_{c} is the codeword length which is $N - N_{\text{p}}$ for the preamble case and N for the superimposed case. Such a codebook is capacity-achieving and achieves the optimal channel dispersion as well [11]. We model the codewords as random vectors defined by

$$\mathbf{D}^{(N_{\text{c}})} = \sqrt{N_{\text{c}}} \frac{\tilde{\mathbf{D}}^{(N_{\text{c}})}}{\left\| \tilde{\mathbf{D}}^{(N_{\text{c}})} \right\|} \tag{2}$$

where $\tilde{\mathbf{D}}^{(N_{\text{c}})} \sim \mathcal{N}(0, \mathbf{I}_{N_{\text{c}}})$ and $\|\cdot\|$ is the ℓ_2-norm. The transmitted signals for the preamble and superimposed case, respectively, can be expressed as

$$X_{j+\tau}^{(\text{p})} = \begin{cases} \sqrt{P} p_j, & j \in \{0, \dots, N_{\text{p}} - 1\} \\ \sqrt{P} D_{j-N_{\text{p}}}, & j \in \{N_{\text{p}}, \dots, N-1\} \\ 0, & \text{otherwise} \end{cases} \tag{3}$$

and

$$X_{j+\tau}^{(\text{SI})} = \begin{cases} \sqrt{P}\left(\sqrt{\alpha} p_j + \sqrt{1-\alpha} D_j\right), & j \in \{0, \dots, N-1\} \\ 0, & \text{otherwise.} \end{cases} \tag{4}$$

Here, α is the ratio of the powers spent for detection versus codeword for the superimposed case.

As we are interested in the regime where N is small, we make use of the asymptotic approximations for the maximum coding rate for the AWGN channel, provided in [11, Theorem 54], to analyze the decoding error probability. It has been noted in [2, footnote 1] that transmitting a length-N_{c} codeword in a complex AWGN channel is equivalent to transmitting a codeword of length $2N_{\text{c}}$ in a real-valued AWGN channel, with the same SNR. Therefore, the decoding error probability for a codeword of length N_{c}, carrying b bits, given an SNR P is well-approximated by

$$\epsilon_{\text{D}}(N_{\text{c}}, P) = Q\left(\frac{2N_{\text{c}}C(P) - b + \frac{1}{2}\log_2 2N_{\text{c}}}{\sqrt{2N_{\text{c}}V(P)}} \right) \tag{5}$$

where $C(P) = \frac{1}{2}\log_2(1 + P)$ and $V(P) = \frac{P(P+2)}{2(P+1)^2}\log_2^2 e$ denote the channel capacity and dispersion, respectively.

The detection is performed by hypothesis testing, using the log-likelihood ratio. This is optimal in an AWGN channel and the metric of comparison is known as *deterministic correlation* [15, Chapter 14]. Considering that a packet begins at time index τ, the correlation at a time index $\tau - k$ can be expressed as

$$\mathcal{R}_{Y, \tau-k} = \Re\left[\sum_{j=0}^{N_{\text{t}}-1} p_j^* Y_{j+\tau-k} \right] > \Delta \tag{6}$$

where $\Re[x]$ denotes the real part of x, p_j^* is the complex conjugate of the jth preamble symbol, $N_{\text{t}} \in \{N_{\text{p}}, N\}$ denotes the length of the preamble and superimposed detection sequence, respectively, and Δ is the detection threshold. Three error events can occur: (1) the false alarms ($\mathcal{E}_{\text{FA}}$) can occur at any time index $\tau - k$ if $\mathcal{R}_{Y, \tau-k} > \Delta$ for any offset $k \in \{1, \dots, t_r - 1\}$, where t_r denotes the *recovery time* [8] elapsed decoding, verifying the CRC bits, and transmitting NACK, during which the receiver misses any incoming packet

(Fig. 1). The receiver was assumed active for at least t_r time slots prior to a packet arrival, such that $\tau - k > 0$; (2) the misdetection event ($\mathcal{E}_{\text{MD}}$) occurs if $\mathcal{R}_{Y,\tau} \leq \Delta$; (3) erroneous decoding event ($\mathcal{E}_{\text{D}}$), signaled by the CRC, due to a bad noise realization and the receiver sends NACK. For the superimposed case, once the detection is in place, the detection sequence is subtracted from the received signal, such that it does not interfere with the decoding.

Our model for the AWGN channel and URLLC is different from the other models in the literature [5], [6], [7], which are valid for the discrete memoryless channels (DMCs) and treat an exponential level of asynchronism. This paper highlights that for URLLC in the discrete AWGN channel, it is important to model the receiver recovery time after a possible false alarm. This is especially true because for URLLC with ACK, there is an inherent cost to having false alarms, resulting in loss of the packet.

3. ANALYSIS

Using the three previously defined error events, we can formulate an upper bound on the PER as

$$\mathcal{P}_e \leq \Pr[\mathcal{E}_{\text{FA}}] + \Pr[\mathcal{E}_{\text{MD}}] + \Pr[\mathcal{E}_{\text{D}}]. \tag{7}$$

An approximation of the last term in (7) is given by $\epsilon_{\text{D}}(N - N_{\text{p}}, P)$ for the preamble case and by $\epsilon_{\text{D}}(N, (1-\alpha)P)$ for the superimposed case. The objective is to estimate the remaining two terms, namely the false alarm and misdetection probabilities and identify the parameters that influence the upper bound for both packets structures.

3.1. Time-multiplexed preamble

We first analyze the false alarm probability $\Pr[\mathcal{E}_{\text{FA}}]$. A false alarm occurs if $\mathcal{R}_{Y,\tau-k} > \Delta$ for some $k \in \mathcal{S}_{\text{FA}} = \{1, \ldots, t_r-1\}$. The correlation output when a false alarm may occur is then

$$\mathcal{R}_{Y,\tau-k} = \sqrt{P}\mathcal{R}_{\text{p}}^{\text{pre}}(k) + \Re\left[\sum_{j=0}^{N_{\text{p}}-1} p_j^* W_{j+\tau-k}\right], \tag{8}$$

where $\mathcal{R}_{\text{p}}^{\text{pre}}(k)$ is the partial-period correlation of the preamble sequence, which is deterministic and nonzero for $k \in \{1, \ldots, N_{\text{p}}-1\}$

$$\mathcal{R}_{\text{p}}^{\text{pre}}(k) = \Re\left[\sum_{j=k}^{N_{\text{p}}-1} p_j^* p_{j-k}\right]. \tag{9}$$

The second term in (8) is distributed as a Gaussian random variable. Thus, the probability of having a false alarm occurring at $\tau - k$, for any $k \in \mathcal{S}_{\text{FA}}$, is given by

$$\Pr[\mathcal{R}_{Y,\tau-k} > \Delta] = \Pr\left[\sqrt{P}\mathcal{R}_{\text{p}}^{\text{pre}}(k) + \Re\left[\sum_{j=0}^{N_{\text{p}}-1} p_j^* W_{j+\tau-k}\right] > \Delta\right]$$

$$= Q\left(\frac{\Delta - \mu_{\mathcal{R}_{Y,\text{FA}}}(k)}{\sigma_{\mathcal{R}_{Y,\text{FA}}}}\right). \tag{10}$$

Here, the mean and variance of the correlation metric are

$$\mu_{\mathcal{R}_{Y,\text{FA}}}(k) = \sqrt{P}\mathcal{R}_{\text{p}}^{\text{pre}}(k); \tag{11}$$

$$\sigma_{\mathcal{R}_{Y,\text{FA}}}^2 = \sum_{j=0}^{N_{\text{p}}-1} \text{Var}\left[\Re\left[p_j^* W_{j+\tau-k}\right]\right] = \frac{N_{\text{p}}}{2} \tag{12}$$

where (12) follows because the power of the Zadoff-Chu sequence p_j is 1 for $j \in \{0, \ldots, N_{\text{p}}-1\}$. Furthermore, the total probability of false alarms is the probability of the union of such events:

$$\Pr[\mathcal{E}_{\text{FA}}] = \Pr\left[\bigcup_{k \in \mathcal{S}_{\text{FA}}} \{\mathcal{R}_{Y,\tau-k} > \Delta\}\right] \tag{13}$$

$$\leq \sum_{k=1}^{t_r-1} Q\left(\frac{\Delta - \mu_{\mathcal{R}_{Y,\text{FA}}}(k)}{\sigma_{\mathcal{R}_{Y,\text{FA}}}}\right). \tag{14}$$

The misdetection error event ($\mathcal{E}_{\text{MD}}$) occurs if upon a packet arrival, $\mathcal{R}_{Y,\tau} \leq \Delta$. The probability of a misdetection is then

$$\Pr[\mathcal{E}_{\text{MD}}] = Q\left(\frac{\mu_{\mathcal{R}_{Y,\tau}} - \Delta}{\sigma_{\mathcal{R}_{Y,\tau}}}\right) \tag{15}$$

where

$$\mu_{\mathcal{R}_{Y,\tau}} = \text{E}\left[\sum_{j=0}^{N_{\text{p}}-1} \Re\left[\sqrt{P}p_j^* p_j\right]\right] = \sqrt{P}N_{\text{p}}; \tag{16}$$

$$\sigma_{\mathcal{R}_{Y,\tau}}^2 = \sum_{j=0}^{N_{\text{p}}-1} \text{Var}\left[\Re\left(p_j^* W_j\right)\right] = \frac{N_{\text{p}}}{2}. \tag{17}$$

3.2. Superimposed sequence

Similarly to the preamble case, we first derive the false alarm probability. For the superimposed case, the correlator output is given by

$$\mathcal{R}_{Y,\tau-k} = \sqrt{\alpha P}\mathcal{R}_{\text{p}}^{\text{SI}}(k) + \sqrt{(1-\alpha)P}\mathcal{R}_{\text{D}}(k) + \underbrace{\Re\left[\sum_{j=0}^{N-1} p_j^* W_{j+\tau-k}\right]}_{\mathcal{R}_W(k)}, \tag{18}$$

where $\mathcal{R}_{\text{D}}(k) = \Re\left[\sum_{j=k}^{N-1} p_j^* D_{j-k}\right]$ is the partial correlation between the detection sequence and the codewords. Due to the additional random variable $\mathcal{R}_{\text{D}}(k)$, we distinguish between two types of false alarms, occurring: (1) purely due to noise ($\mathcal{E}_{\text{FA1}}$) and (2) also due to partial-correlation with the preamble and codeword symbols ($\mathcal{E}_{\text{FA2}}$). The distributions of the correlation for the two types will be different. The partial period correlation of the detection sequence, $\mathcal{R}_{\text{p}}^{\text{SI}}(k)$, is defined similarly to (9) and remains deterministic.

For the purely noise-inflicted false alarms, occurring when $k \in \{N, \ldots, t_r-1\}$, both $\mathcal{R}_{\text{p}}^{\text{SI}}(k)$ and $\mathcal{R}_{\text{D}}(k)$ are equal to zero. Therefore, the noise-inflicted false alarm probability is

$$\Pr[\mathcal{E}_{\text{FA1}}] \leq (t_r-N)Q\left(\frac{\Delta}{\sigma_{\mathcal{R}_{Y,\text{FA1}}}}\right) \tag{19}$$

where $\sigma_{\mathcal{R}_{Y,\text{FA1}}}^2 = N/2$. For the partial-correlation inflicted false alarms, we shall obtain an asymptotic distribution of $\mathcal{R}_{\text{D}}(k)$ by employing Slutsky's lemma [16, Lemma 2.8]. Using the model of the codewords from (2), we obtain

$$\mathcal{R}_{\text{D}}(k) = \Re\left[\sqrt{P}\mathbf{p}^{\text{H}}\mathbf{D}^{(N)}\right] = \Re\left[\frac{\sqrt{PN}\mathbf{p}^{\text{H}}\tilde{\mathbf{D}}^{(N)}}{\|\tilde{\mathbf{D}}^{(N)}\|}\right]. \tag{20}$$

Here, $\mathbf{p}^{\text{H}}$ is the transposed conjugate of the N-dimensional vector of preamble symbols p_j. Next, we show that $\mathcal{R}_{\text{D}}(k)$ is well-approximated by Gaussian random variable for large N. To this end, let $Z = \Re\left[\sqrt{P}\mathbf{p}^{\text{H}}\tilde{\mathbf{D}}^{(N)}\right]$. Then, it follows that Z is a Gaussian random variable with zero mean and variance

$$\mu_Z = \text{E}\left[\Re\left[\sqrt{P}\mathbf{p}^{\text{H}}\tilde{\mathbf{D}}^{(N)}\right]\right] = 0; \tag{21}$$

$$\sigma_Z^2 = \text{Var}\left[\Re\left[\sqrt{P}\mathbf{p}^{\text{H}}\tilde{\mathbf{D}}^{(N)}\right]\right] = \frac{1}{2}PN. \tag{22}$$

Since the law of large numbers implies that $\|\tilde{\mathbf{D}}^{(N)}\|/\sqrt{N} \xrightarrow{\text{p}} 1$, it follows from Slutsky's lemma [16, Lemma 2.8] that

$$\mathcal{R}_{\text{D}}(k) = \Re\left[\sqrt{P}\mathbf{p}^{\text{H}}\mathbf{D}^{(N)}\right] \xrightarrow{\text{d}} Z \sim \mathcal{N}(\mu_Z, \sigma_Z^2) \tag{23}$$

where $\xrightarrow{\text{d}}$ and $\xrightarrow{\text{p}}$ denote convergence in distribution and convergence in probability, respectively. Therefore, as $N \to \infty$, $\mathcal{R}_{Y,\tau-k}$ converges to a Gaussian random variable with mean and variance

$$\mu_{\mathcal{R}_{Y,\text{FA2}}}(k) = \sqrt{\alpha P}\mathcal{R}_{\text{p}}^{\text{SI}}(k); \tag{24}$$
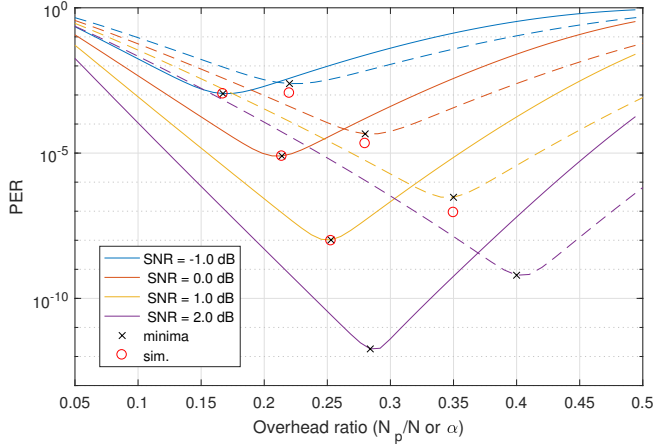
Fig. 2. Upper bounds/approximations of PER for the preamble case (solid lines) and superimposed case (dashed lines) for different SNR values. Black dots and red circles represent minima of the upper bounds/approximations and simulated PER values, respectively. A packet carrying $b = 128$ bits transmitted over $N = 257$ channel uses is considered and a recovery time of $t_r = 283$ channel uses is assumed (10% higher than N). For PER $< 10^{-8}$, computational complexity precludes accurate results.

$$\sigma^2_{\mathcal{R}_{Y,\text{FA2}}}(k) = \text{Var}\left[\mathcal{R}_W(k) + \sqrt{(1-\alpha)P}\mathcal{R}_D(k)\right] \quad (25)$$

$$= \sigma^2_{\mathcal{R}_{Y,\text{FA1}}} + (1-\alpha)\text{Var}\left[\Re\left(\sqrt{P}\sum_{j=k}^{N-1} p_j^* D_{j-k}\right)\right] \quad (26)$$

$$= \frac{N}{2} + \frac{1}{2}(1-\alpha)(N-k)P. \quad (27)$$

Here, (27) follows from (22), (23), (32), and because the covariance matrix of $\mathbf{D}^{(N)}$ is the identity matrix.[1] Therefore, a partial-correlation inflicted false alarm at time $\tau - k$ is

$$\Pr\left[\mathcal{R}_{Y,\tau-k} > \Delta\right] \approx Q\left(\frac{\Delta - \mu_{\mathcal{R}_{Y,\text{FA2}}}(k)}{\sigma_{\mathcal{R}_{Y,\text{FA2}}}(k)}\right). \quad (28)$$

Applying the union bound as in (13)-(14), the probability of having a partial-correlation inflicted false alarm for the superimposed case is upper-bounded as

$$\Pr[\mathcal{E}_{\text{FA2}}] \leq \sum_{k=1}^{N-1} Q\left(\frac{\Delta - \mu_{\mathcal{R}_{Y,\text{FA2}}}(k)}{\sigma_{\mathcal{R}_{Y,\text{FA2}}}(k)}\right). \quad (29)$$

Finally, for the misdetection event, the correlation metric can also be approximated as a Gaussian variable. Therefore, the probability of a misdetection is approximated by

$$\Pr[\mathcal{E}_{\text{MD}}] \approx Q\left(\frac{\mu_{\mathcal{R}_{Y,\tau}} - \Delta}{\sigma_{\mathcal{R}_{Y,\tau}}}\right). \quad (30)$$

Here, the mean and variance are

$$\mu_{\mathcal{R}_{Y,\tau}} = \sqrt{\alpha P}N + \sqrt{1-\alpha}\mu_Z = \sqrt{\alpha P}N \quad (31)$$

$$\sigma^2_{\mathcal{R}_{Y,\tau}} = \sigma^2_{\mathcal{R}_{Y,\text{FA1}}} + (1-\alpha)\text{Var}\left[\Re\left[\sqrt{P}\mathbf{p}^{\text{H}}\mathbf{D}^{(N)}\right]\right] \quad (32)$$

$$= \frac{N}{2} + (1-\alpha)\sigma^2_Z = \frac{N}{2} + \frac{1}{2}(1-\alpha)PN \quad (33)$$

where (33) follows because the covariance matrix of $\mathbf{D}^{(N)}$ is the identity matrix.

---

[1]Spherical symmetry implies that the vector $\mathbf{D}^{(N)}$ in (2) has zero mean with entries being uncorrelated and with entries having unit variance. The latter follows because $\sum_{i=1}^{N} \text{E}[(D_i)^2] = N\text{E}\left[\frac{\sum_{i=1}^{N}(\tilde{D}_i)^2}{||\tilde{\mathbf{D}}^{(N)}||^2}\right] = N$.
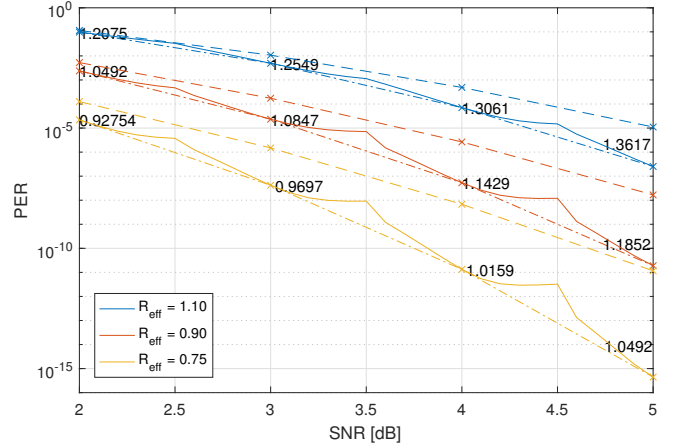

Fig. 3. PER for three effective rates (b/N, $b = 128$, N varies) for optimal preamble size (dotted and dashed lines), optimal superimposed sequence (dashed lines) and pragmatic approach of adaptive coding rate for the preamble case (solid line), where the coding rate for the data is changed every $1\,\text{dB}$ interval centered in the optimal computed rate for each integer SNR point, and labeled by the numbers next to it.

4. NUMERICAL RESULTS AND CONCLUSION

In this section, we plot our analytical bounds and approximations and compare them to simulations of the PER. In the simulations of the PER, we use the approximation (5) to compute the decoding error probability.

For the time-multiplexed preamble, we let $\overline{\mathcal{P}_e^{\text{pre}}}(\Delta, N_\text{p}, N, P)$ denote the upper bound on the PER obtained by summing (14) and (15), and $\epsilon_\text{D}(N - N_\text{p}, P)$. We are interested in solving the optimization problem

$$\underset{\substack{N_\text{p} \in \{1,\dots,N-1\} \\ \Delta \geq 0}}{\text{minimize}} \quad \overline{\mathcal{P}_e^{\text{pre}}}(\Delta, N_\text{p}, N, P). \quad (34)$$

Note that the decoding error probability depends on N_p and P and the detection error probability depends on Δ, N_p, and P.

For the superimposed sequence, we let $\overline{\mathcal{P}_e^{\text{SI}}}(\Delta, \alpha, N, P)$ be the approximation of the PER obtained by summing (19), (29), (30), and $\epsilon_\text{D}(N, (1-\alpha)P)$. Our objective is to solve the optimization problem

$$\underset{\substack{\alpha \in (0,1) \\ \Delta \geq 0}}{\text{minimize}} \quad \overline{\mathcal{P}_e^{\text{SI}}}(\Delta, \alpha, N, P). \quad (35)$$

The tradeoff between detection and decoding is shown in Fig. 2, which also depicts the solutions to the optimization problems above. For all points of the solid and dashed curves, Δ is optimized. We observe that the optimal overhead ratios depend on the SNR and the target reliability, and that the superimposed structure achieves its minimum error probability at a higher overhead ratio. Furthermore, the PER achieved with superimposed detection sequences also experiences a degradation in terms of minimum PER because a larger fraction of resources are spent on detection overhead.

The superimposed structure, however, offers enhanced adaptivity. For the system model considered, where the receiver and transmitter are aware of the SNR, the preamble and superimposed structures follow the optimal regimes with the variation of N_p and α, respectively. For the preamble case, this requires that the codewords are encoded with different rates. A pragmatic approach of what can be achieved is shown in Fig. 3 in solid lines, where a distinct codebook is required for each $1\,\text{dB}$ SNR interval. For comparison, the dashed lines and the dashed-dotted lines indicate the superimposed and the ideal preamble optimal regimes, respectively.

5. REFERENCES

[1] B. Lee, S. Park, D. J. Love, H. Ji, and B. Shim, "Packet structure and receiver design for low latency wireless communications with ultra-short packets," *IEEE Trans. Commun.*, vol. 66, no. 2, Feb. 2018.

[2] 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, Sep. 2016.

[3] P. Popovski, J. J. Nielsen, Č. Stefanović, E. de Carvalho, E. Ström, K. F. Trillingsgaard, A.-S. Bana, D. M. Kim, R. Kotaba, J. Park, and R. B. Sørensen, "Wireless access for ultra-reliable low-latency communication (URLLC): Principles and building blocks," 2017, `arXiv:1708.07862 [cs.IT]`.

[4] J. L. Massey, "Optimum frame synchronization," *IEEE Trans. Commun.*, vol. COM-20, no. 2, Apr. 1972.

[5] A. Tchamkerten, V. Chandar, and G. W. Wornell, "Communication under strong asynchronism," *IEEE Trans. Inf. Theory*, vol. 55, no. 10, Oct. 2009.

[6] Y. Polyanskiy, "Asynchronous communications: Exact synchronization, universality, and dispersion," *IEEE Trans. Inf. Theory*, vol. 59, no. 3, Mar. 2013.

[7] N. Weinberger and N. Merhav, "Codeword or noise? Exact random coding exponents for joint detection and decoding," *IEEE Trans. Inf. Theory*, vol. 60, no. 9, Sep. 2014.

[8] S. Nagaraj, S. Khan, C. Schlegel, and M. V. Burnashev, "Differential preamble detection in packet-based wireless networks," *IEEE Trans. Wireless Commun.*, vol. 8, no. 2, Feb. 2009.

[9] Y. Wang, J. Oostveen, A. Filippi, and S. Wesemann, "A novel preamble scheme for packet-based OFDM WLAN," in *Proc. IEEE Wireless Commun. and Networking Conf.*, Kowloon, China, Mar. 2007.

[10] J. Östman, G. Durisi, E. G. Ström, J. Li, H. Sahlin, and G. Liva, "Low-latency ultra-reliable 5G communications: Finite-blocklength bounds and coding schemes," in *Proc. IEEE Conf. on Sys., Commun. and Coding*, Hamburg, Germany, Feb. 2017.

[11] Y. Polyanskiy, H. V. Poor, and S. Verdú, "Channel coding rate in the finite blocklength regime," *IEEE Trans. Inf. Theory*, vol. 56, no. 5, Apr. 2010.

[12] G. Liva, G. Durisi, M. Chiani, S. S. Ullah, and S. C. Liew, "Short codes with mismatched channel state information: A case study," 2017, `arXiv:1705.05528 [cs.IT]`.

[13] D. C. Chu, "Polyphase codes with good periodic correlation properties," *IEEE Trans. Inf. Theory*, vol. 18, no. 4, Jul. 1972.

[14] F. Khan, *LTE for 4G Mobile Broadband: Air Interface Technologies and Performance.* ISBN: 9780521882217: Cambridge University Press, 2009.

[15] A. V. Oppenheim and G. C. Verghese, "6.011 Introduction to communication, control and signal processing," Massachusetts Institute of Technology: MIT OpenCourseWare. License: Creative Commons BY-NC-SA, Spring 2010. [Online]. Available: https://ocw.mit.edu

[16] A. van der Vaart, *Asymptotic Statistics.* ISBN: 978-0-521-49603-2: Cambridge University Press, 2007.