Joint Channel Coding of Consecutive Messages with Heterogeneous Decoding Deadlines in the Finite Blocklength Regime

Homa Nikbakht*, Malcolm Egan*, Jean-Marie Gorce*
*Université de Lyon, Inria, INSA Lyon, CITI, EA3720, 69621 Villeurbanne, France
{homa.nikbakht, malcom.egan, jean-marie.gorce}@inria.fr

***Abstract*—A standard assumption in the design of ultra-reliable low-latency communication systems is that the duration between message arrivals is larger than the number of channel uses before the decoding deadline. Nevertheless, this assumption fails when messages rapidly arrive and reliability constraints require that the number of channel uses exceeds the time between arrivals. In this paper, we study channel coding in this setting by jointly encoding messages as they arrive while decoding the messages separately, allowing for heterogeneous decoding deadlines. For a scheme based on power sharing, we analyze the probability of error in the finite blocklength regime. We show that significant performance improvements can be obtained for short packets by using our scheme instead of standard approaches based on time sharing.**

I. Introduction

One of the pillars of 5G is ultra-reliable low latency communications (URLLC), where the goal is to transmit typically small quantities of data with a very low probability of error and strict decoding deadlines. With applications ranging from autonomous driving to remote surgery, a number of channel coding schemes have been proposed including short LDPC and polar codes [1]–[4]. At the same time, new characterizations of fundamental tradeoffs between the size of the message set, the probability of error, and the length of the code have been obtained via achievability and converse bounds building on the work in [5].

A key assumption in existing coding schemes for URLLC and their analysis is that message arrivals and the decoding deadlines of preceeding messages are sufficiently separated. As a consequence, each packet can be encoded and decoded separately. Unfortunately, this assumption is not guaranteed to hold, particularly in industrial process control applications [1].

To give a concrete example, consider control of a conveyor belt. A key component of this system is sensor data, which is communicated to a controller. In normal operation, the sensor may send regularly timed updates of its speed, which is used in model predictive control algorithms in order to optimize actuation in order to yield a desired speed. On the other hand, when the speed requirements are varied (e.g., at start up), it may be desirable to send speed observations from the sensor more often.

In order to ensure reliability of the sensor observations, the channel uses allocated to each observation of the speed may partially overlap. It is therefore desirable to consider joint encoding of multiple sensor observations, albeit with heterogeneous decoding deadlines. That is, if the channel uses for two separate observations overlap, it is not possible to wait until the entire transmission for both sensor observations is received before decoding.

The problem of heterogeneous decoding delays has seen limited attention. The main work in this direction is in the context of broadcast communications, Shulman and Feder studied static broadcasting in [6] and [7], deriving a coding theorem for the rate region. In this model, a sender transmits a single message and multiple receivers attempt reliable decoding. Crucially, each receiver has a different decoding deadline.

Recently in [8], Langberg and Effros have also considered a variant on the network communication problem in [6]. In particular, networks consisting of multiple transmitters and receivers were studied where each receiver has different decoding deadlines for its messages of interest. A generalization of the rate region, known as the time-rate region, was introduced and an inner bound derived, which is known to not be tight.

In this paper, we derive tradeoffs between error probability, message set size, and the (finite) number of channel uses for joint channel coding of two consecutive messages with heterogeneous decoding deadlines. In contrast to the works in [6] and [8], where the messages are available at the transmitters before the transmission begins, we assume that messages arrive at different times. We focus on point-to-point Gaussian noise channels with signals subject to an average power constraint. We propose a scheme based on power sharing and analyze the probability of error. We establish a significant performance improvement of our scheme over time sharing in the finite blocklength regime for a sufficiently large transmit power.

II. Problem Setup and Proposed Coding Scheme

Consider a sensor that sends two packets, where each packet corresponds to a message in the set $\{1,\ldots,M\}$. At time $t = a_1$, transmission commences for the first packet corresponding to the message $m_1 \in \{1,\ldots,M\}$. At time $t = a_2$, transmission commences for the second packet corresponding to the message $m_2 \in \{1,\ldots,M\}$. The two messages m_1, m_2 are assumed to be drawn independently with each element in $\{1,\ldots,M\}$ occurring with probability $\frac{1}{M}$.

Each message is subject to different decoding delay constraints. In particular, at time d_1, the receiver attempts to reconstruct the message m_1. Similarly, at time $d_2 > d_1$, the receiver attempts to reconstruct the message m_2.

Given the times of arrival and decoding delay constraints, the encoder is constructed as follows. Denote the transmission window of the first and second messages by $\mathcal{W}_1$ and $\mathcal{W}_2$, respectively, where

$$\mathcal{W}_1 \triangleq \{a_1, \ldots, d_1\}, \quad \mathcal{W}_2 \triangleq \{a_2, \ldots, d_2\}. \tag{1}$$

Under the assumption $\mathcal{W}_1 \cap \mathcal{W}_2 \neq \emptyset$, the encoder outputs symbols at time $t \in \{a_1, \ldots, d_2\}$ given by

$$X_t = \begin{cases} f_t(m_1), & t \in \{a_1, \ldots, a_2 - 1\} \\ \psi_t(m_1, m_2), & t \in \{a_2, \ldots, d_1\} \\ \phi_t(m_2), & t \in \{d_1 + 1, \ldots, d_2\}, \end{cases} \tag{2}$$

where f, ψ, ϕ are the encoding functions corresponding to the channel uses where only message m_1 is arrived but not m_2, where both m_1, m_2 are present, and after m_1 has been decoded, respectively. We highlight that m_2 is not known before time $t = a_2$; i.e., encoding is causal.

For simplicity, define

$$n_1 \triangleq a_2 - a_1, \quad n_2 \triangleq d_1 - a_2 + 1, \quad \text{and} \quad n_3 \triangleq d_2 - d_1. \tag{3}$$

Given the structure of the encoding functions, receiver observations can be viewed as arising from three parallel channels: over the first channel of n_1 blocks only m_1 is transmitted; over the second channel of n_2 blocks m_1 and m_2 are *jointly* transmitted; and over the third channel of n_3 blocks only m_2 is transmitted. Our goal is to establish bounds on the decoding error probabilities of m_1 and m_2 under a Gaussian interference assumption defined precisely in Section III.

Define the following channel input vectors

$$\boldsymbol{X}_1 \triangleq \{X_{a_1}, \ldots, X_{a_2 - 1}\}, \tag{4a}$$
$$\boldsymbol{X}_2 \triangleq \{X_{a_2}, \ldots, X_{d_1}\}, \tag{4b}$$
$$\boldsymbol{X}_3 \triangleq \{X_{d_1+1}, \ldots, X_{d_2}\}. \tag{4c}$$

For the i-th channel with $i \in \{1, 2, 3\}$, the encoding functions satisfy the average block power constraint

$$\frac{1}{n_i} ||\boldsymbol{X}_i||^2 \leq P_i \tag{5}$$

almost surely.

Since the messages m_1 and m_2 are jointly transmitted over the second channel, the transmit power P_2 is divided into two parts βP_2 and $(1 - \beta) P_2$ for $\beta \in [0, 1]$. The portion βP_2 is assigned to the transmission of m_1 and the portion $(1 - \beta) P_2$ is assigned to the transmission of m_2. Thus

$$\boldsymbol{X}_2 = \boldsymbol{X}_2^{(1)} + \boldsymbol{X}_2^{(2)}, \tag{6}$$

where $||\boldsymbol{X}_2^{(1)}||^2 = n_2 \beta P_2$ and $||\boldsymbol{X}_2^{(2)}||^2 = n_2 (1 - \beta) P_2$. The corresponding outputs at the receiver similarly denoted by

$$\boldsymbol{Y}_1 \triangleq \{Y_{a_1}, \ldots, Y_{a_2 - 1}\}, \tag{7a}$$
$$\boldsymbol{Y}_2 \triangleq \{Y_{a_2}, \ldots, Y_{d_1}\}, \tag{7b}$$
$$\boldsymbol{Y}_3 \triangleq \{Y_{d_1+1}, \ldots, Y_{d_2}\}. \tag{7c}$$

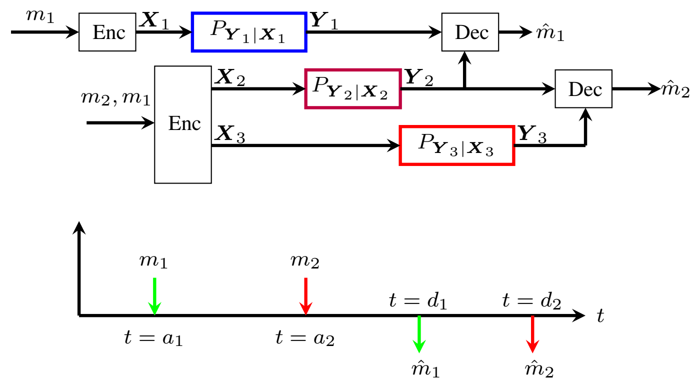

Fig. 1: System model.

Moreover, we denote the resulting three channels by $P_{\boldsymbol{Y}_1|\boldsymbol{X}_1}$, $P_{\boldsymbol{Y}_2|\boldsymbol{X}_2}$ and $P_{\boldsymbol{Y}_3|\boldsymbol{X}_3}$, respectively. We assume that each channel is additive, memoryless, stationary, and Gaussian with variance σ_i^2 with $i \in \{1, 2, 3\}$. The resulting system model is illustrated in Figure 1.

At the receiver, the decoder attempts to reconstruct the two messages m_1, m_2 based on the channel outputs via the decoding functions g_1, g_2 defined as

$$\hat{m}_1 = g_1(\boldsymbol{Y}_1, \boldsymbol{Y}_2), \tag{8}$$
$$\hat{m}_2 = g_2(\boldsymbol{Y}_2, \boldsymbol{Y}_3). \tag{9}$$

The average probability of error for each of the messages is then given by

$$\epsilon_1 = \mathbb{P}(\hat{m}_1 \neq m_1), \quad \epsilon_2 = \mathbb{P}(\hat{m}_2 \neq m_2). \tag{10}$$

The focus of this paper is to characterize the tradeoff between the size of the message set M, the error probabilities ϵ_1, ϵ_2, and the decoding deadlines d_1, d_2. Formally, we study the achievable region defined as follows.

Definition 1: Given the power constraints P_1, P_2 and P_3, a tuple $(a_1, a_2, d_1, d_2, M, \epsilon_1, \epsilon_2)$ is achievable if messages m_1, m_2 of cardinality M arriving at the a_1-th and a_2-th channel uses can be decoded by the d_1-th and d_2-th channel uses with an average probability of error not exceeding ϵ_1, ϵ_2, respectively.

III. Error Probability Analysis

In this section, we study the error probabilities of joint encoding schemes for packets with heterogeneous decoding delays. As detailed in the previous section, we consider an encoder structure which superimposes symbols corresponding to each message. As a consequence, symbols from one packet act as interference for the other.

To characterize the error probability for an optimal code, it is therefore necessary to specify the code structure. Unfortunately, the optimal code structure is not currently known. As a consequence, we first study the error probability under the assumption that the codeword for the second message is modeled as Gaussian when decoding m_1, called the *Gaussian interference approximation*. In order to verify that the Gaussian approximation is reasonable, we then consider a non-Gaussian model for the codeword of the second message, where the codeword is isotropic on the power shell.

A. Gaussian Interference Approximation

Given the set of channel uses $\{n_i\}_{i=1}^3$ and transmit powers $\{P_i\}_{i=1}^3$, and the parameter $\beta \in [0,1]$, define

$$\Omega_1 = \frac{P_1}{\sigma_1^2}, \quad \Omega_2 = \frac{\beta P_2}{(1-\beta)P_2 + \sigma_2^2}, \tag{11}$$

$$\Omega_3 = \frac{(1-\beta)P_2}{\sigma_2^2} \quad \text{and} \quad \Omega_4 = \frac{P_3}{\sigma_3^2}. \tag{12}$$

Also define

$$u_1 \sim \mathcal{X}^2\left(n_1, n_1\frac{1+\Omega_1}{\Omega_1}\right), \quad v_1 \sim \mathcal{X}^2\left(n_1, n_1\frac{1}{\Omega_1}\right), \tag{13}$$

$$u_2 \sim \mathcal{X}^2\left(n_2, n_2\frac{1+\Omega_2}{\Omega_2}\right), \quad v_2 \sim \mathcal{X}^2\left(n_2, n_2\frac{1}{\Omega_2}\right), \tag{14}$$

$$u_3 \sim \mathcal{X}^2\left(n_2, n_2\frac{1+\Omega_3}{\Omega_3}\right), \quad v_3 \sim \mathcal{X}^2\left(n_2, n_2\frac{1}{\Omega_3}\right), \tag{15}$$

$$u_4 \sim \mathcal{X}^2\left(n_3, n_3\frac{1+\Omega_4}{\Omega_4}\right), \quad v_4 \sim \mathcal{X}^2\left(n_3, n_3\frac{1}{\Omega_4}\right), \tag{16}$$

where $\mathcal{X}^2(n,s)$ denotes a non-central chi-squared random variable of order n and parameter s. Furthermore, define

$$Q_1 \triangleq \frac{v_1\Omega_1}{1+\Omega_1} + \frac{v_2\Omega_2}{1+\Omega_2}, \quad Q_2 \triangleq \frac{v_3\Omega_3}{1+\Omega_3} + \frac{v_4\Omega_4}{1+\Omega_4}, \tag{17}$$

$$Q_3 \triangleq \Omega_1 u_1 + \Omega_2 u_2, \qquad Q_4 \triangleq \Omega_3 u_3 + \Omega_4 u_4, \tag{18}$$

and for each $i \in \{1,\ldots,4\}$, define F_{Q_i} as the cumulative distribution function (CDF) of the variable Q_i.

We first establish lower bounds for ϵ_1 and ϵ_2 under the random coding scheme introduced in Section II.

Theorem 1 (Converse Bound): Under the Gaussian interference approximation, for fixed transmission rates $R_1 = \log M/(n_1+n_2)$ and $R_2 = \log M/(n_2+n_3)$ corresponding to message m_1 and m_2, respectively, the error probabilities ϵ_1 and ϵ_2 are lower bounded by

$$\epsilon_1 \geq \mathbb{P}[Q_1 > \lambda_1] = 1 - F_{Q_1}(\lambda_1) \tag{19}$$

$$\epsilon_2 \geq \mathbb{P}[Q_2 > \lambda_2] = 1 - F_{Q_2}(\lambda_2) \tag{20}$$

where λ_1 and λ_2 are set by the constraints

$$F_{Q_3}(\lambda_1) = 2^{-(n_1+n_2)R_1}, \tag{21}$$

$$F_{Q_4}(\lambda_2) = 2^{-(n_2+n_3)R_2}. \tag{22}$$

Proof: The proof follows closely the arguments in [10] and [11]. Details omitted due to space constraints, for full details see [14]. ∎

Upper bounds on the error probabilities ϵ_1 and ϵ_2 are given in the following theorem.

Theorem 2 (Achievability bound): Under the Gaussian interference approximation, for a fixed message set size M, the error probabilities ϵ_1 and ϵ_2 are upper bounded by

$$\epsilon_1 \leq 1 - F_{Q_1}(\Delta_1) + \zeta_1 + G_1(1-\zeta_1), \tag{23}$$

$$\epsilon_2 \leq 1 - F_{Q_2}(\Delta_2) + \zeta_2 + G_2(1-\zeta_2) \tag{24}$$

where

$$\begin{aligned}
\Delta_1 &\triangleq -2\ln(MG_1J_1J_2) + \frac{1}{n_1}\ln(1+\Omega_1) \\
&\quad + \frac{1}{n_2}\ln(1+\Omega_2) + n_1 + \beta n_2, \\
\Delta_2 &\triangleq -2\ln(MG_2\tilde{J}_1\tilde{J}_2) + \frac{1}{n_2}\ln(1+\Omega_3) \\
&\quad + \frac{1}{n_3}\ln(1+\Omega_4) + n_3 + (1-\beta)n_2, \\
\zeta_1 &\triangleq e^{-\kappa_1 n_1^{1/3}} + e^{-\kappa_2 n_2^{1/3}} - e^{-(\kappa_1 n_1^{1/3} + \kappa_2 n_2^{1/3})}, \\
\zeta_2 &\triangleq e^{-\tilde{\kappa}_1 n_2^{1/3}} + e^{-\tilde{\kappa}_2 n_3^{1/3}} - e^{-(\tilde{\kappa}_1 n_2^{1/3} + \tilde{\kappa}_2 n_3^{1/3})}, \\
G_1 &\triangleq \frac{1}{1-e^{-\eta}} \min\left\{\frac{L(P_2,s_2)\eta}{\sqrt{n_2\beta P_2 s_2}}, \frac{L(P_1,s_1)\eta}{\sqrt{n_1 P_1 s_1}}\right\}, \\
G_2 &\triangleq \frac{1}{1-e^{-\eta}} \min\left\{\frac{L(P_2,s_2)\eta}{\sqrt{n_2(1-\beta)P_2 s_2}}, \frac{L(P_3,s_3)\eta}{\sqrt{n_3 P_3 s_3}}\right\}, \\
s_1 &\triangleq \frac{1}{n_1}||\boldsymbol{y}_1||_2^2, \quad s_2 \triangleq \frac{1}{n_2}||\boldsymbol{y}_2||_2^2, \quad s_3 \triangleq \frac{1}{n_3}||\boldsymbol{y}_3||_2^2,
\end{aligned}$$

$$L(P,s) \triangleq \frac{(2Ps)^2}{\sqrt{2\pi}} \cdot \sqrt{\frac{1+4Ps-\sqrt{1+4Ps}}{(\sqrt{1+4Ps}-1)^5}},$$

and η, J_1, J_2, $\tilde{J}_1$, $\tilde{J}_2$, κ_1, κ_2, $\tilde{\kappa}_1$ and $\tilde{\kappa}_2$ are constants.

Proof: See Appendix A. ∎

Under the Gaussian interference assumption, the converse and achievability bounds in Theorems 1 and 2 are in agreement. This was observed for encoding of a single packet in [12] and is generalized for joint encoding with decoding delay constraints in the following corollary.

Corollary 1: By choosing the constants η, J_1, J_2, $\tilde{J}_1$, $\tilde{J}_2$, κ_1, κ_2, $\tilde{\kappa}_1$ and $\tilde{\kappa}_2$ satisfying the following conditions:

$$F_{Q_1}(\Delta_1) - F_{Q_1}(\lambda_1) = \zeta_1 + G_1(1-\zeta_1), \tag{26a}$$

$$F_{Q_2}(\Delta_2) - F_{Q_2}(\lambda_2) = \zeta_2 + G_2(1-\zeta_2), \tag{26b}$$

then, under the Gaussian interference approximation, the converse and the achievability bounds on the error probabilities ϵ_1 and ϵ_2 in Theorems 1 and 2 coincide.

B. Isotropic Interference on the Power Shell

The analysis of the probability of error in Sec. III-A relied on the assumption that when decoding m_1, the interference arising from the second message in the second channel is Gaussian. More precisely, it was assumed that $\boldsymbol{X}_2^{(2)} \sim \mathcal{N}(0, I_{n_2}(1-\beta)P_2)$ when decoding m_1. On the other hand, it is clear that for an optimal coding scheme, this assumption will not hold. Indeed, we expect that $\boldsymbol{X}_2^{(2)}$ should lie on a power shell.

In this section, we relax the Gaussian assumption on $\boldsymbol{X}_2^{(2)}$ in decoding m_1 such that it is isotropic on the power shell; i.e., $||\boldsymbol{X}_2^{(2)}||^2 = n_2(1-\beta)P_2$. A natural question is whether the resulting error probability significantly changes under this different assumption on the statistics of $\boldsymbol{X}_2^{(2)}$? In order to answer this question, we derive a lower bound on the

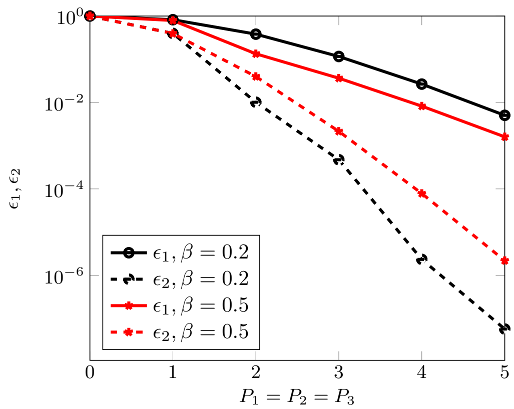

Fig. 2: Average error probabilities ϵ_1 and ϵ_2 vs the transmit power $P_1 = P_2 = P_3$ for different values of β. Here, $n_1 = 10, n_2 = 10, n_3 = 10$ and $\log M = 10$.

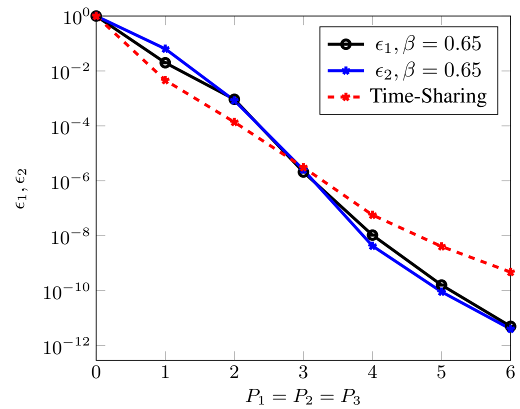

Fig. 3: Average error probabilities ϵ_1 and ϵ_2 vs the transmit power $P_1 = P_2 = P_3$. Here, $n_1 = 20, n_2 = 20, n_3 = 20$ and $\log M = 10$.

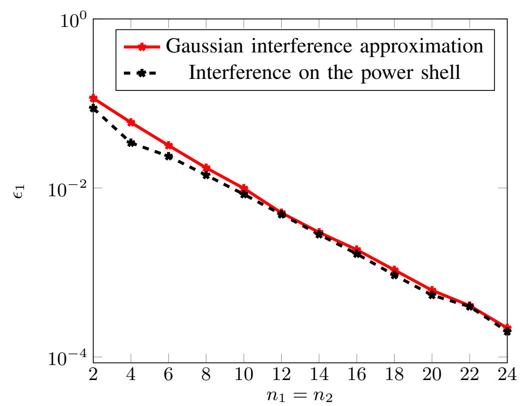

Fig. 4: Converse bounds on ϵ_1 under Gaussian interference approximation and isotropic interference on the power shell assumptions. Here, $\lambda_1 = 0.01(n_1+n_2)$, $P_1 = P_2 = 2, \beta = 0.5$ and $\log M = 10$.

probability of error and compare it to the lower bound in the previous section.

Let $Q_{\boldsymbol{Y}_2|\boldsymbol{X}_2^{(1)}}$ be the channel arising when $\boldsymbol{X}_2^{(2)}$ is isotropic on the power shell (i.e., $||\boldsymbol{X}_2^{(2)}||^2 = n_2(1-\beta)P_2$) and $P_{\boldsymbol{Y}_2|\boldsymbol{X}_2^{(1)}}$ be the channel studied in the previous section. A lower bound on the error probability under the channel $Q_{\boldsymbol{Y}_2|\boldsymbol{X}_2^{(1)}}$ can be obtained via the meta-converse argument [5].

Consider the binary hypothesis test between two distributions P and Q. Let $Z = 1$ when P is selected and $Z = 0$ when Q is selected. By the Neyman-Pearson theorem, the optimal probability of detection under a false alarm constraint $1 - \alpha$ is given by

$$\mathcal{L}_\alpha(P, Q) = \inf_{Z: P[Z=1] \geq \alpha} Q[Z = 1]. \tag{27}$$

Let ϵ be the average error probability for the channel $P_{\boldsymbol{Y}_2|\boldsymbol{X}_2^{(1)}}$ and ϵ' the average error probability for the channel $Q_{\boldsymbol{Y}_2|\boldsymbol{X}_2^{(1)}}$. Then, the meta-converse theorem [5, Theorem 26] states

$$\mathcal{L}_{1-\epsilon}(P_{\boldsymbol{Y}_2|\boldsymbol{X}_2^{(1)}}, Q_{\boldsymbol{Y}_2|\boldsymbol{X}_2^{(1)}}) \leq 1 - \epsilon'. \tag{28}$$

As a consequence, the average error probability ϵ' can be estimated from ϵ via Monte Carlo simulation. Indeed, ϵ' can be estimated by sampling from $Q_{\boldsymbol{Y}_2|\boldsymbol{X}_2^{(1)}}$ and applying the decision rule

$$Z = \mathbb{1}\left\{ \ln \frac{P_{\boldsymbol{Y}_2|\boldsymbol{X}_2^{(1)}}(\boldsymbol{y}_2|\boldsymbol{x}_2^{(1)})}{Q_{\boldsymbol{Y}_2|\boldsymbol{X}_2^{(1)}}(\boldsymbol{y}_2|\boldsymbol{x}_2^{(1)})} < \lambda_1 \right\}, \tag{29}$$

where $\mathbb{1}\{\cdot\}$ is the indicator function and λ_1 is chosen such that the probability of false alarm constraint holds.

IV. Numerical Results

In this section, we provide numerical analysis of the bounds in Theorems 1 and 2 and the performance differences between our proposed power sharing scheme and the time sharing scheme. In the time sharing scheme, messages are transmitted independently by allocating the same number of channel uses to each message. As such, the bounds on the probability of decoding error using time sharing are obtained using the results in [5]. In Fig. 2, we evaluate the bounds on ϵ_1 and ϵ_2, as a function of the transmit power for different values of the parameter β. We assume equal transmit power over the all three channels with $n_1 = n_2 = n_3 = 10$ and the Gaussian interference approximation holds. Note that utilizing Corollary 1, the upper and lower bounds are in agreement.

In the second channel recall that βP_2 is the power assigned to transmit m_1 and $(1-\beta)P_2$ is the power assigned to transmit m_2. Observe that as the power sharing parameter β increases, the error probability ϵ_1 decreases and ϵ_2 increases, as expected. When $\beta = 0.5$, i.e., when the transmit power P_2 is assigned equally to the transmission of each message, ϵ_2 is lower than ϵ_1. This is due to the fact that when decoding the first message, the transmission of the second message is considered as interference. On the other hand, when decoding the second message, the first message has already been decoded.

Fig. 3 plots the error probabilities ϵ_1 and ϵ_2 for varying power levels. The solid lines correspond to the error probabilities under our power sharing scheme and the dashed line to time sharing, where each message is allocated the same number of channel uses. Observe that for our power sharing scheme, when $n_1 = n_2 = n_3 = 20$ and $\log M = 10$, by setting β equal to 0.65, it can be seen that ϵ_1 and ϵ_2 are close. Moreover, when the transmit power is small, the error probabilities obtained under the time sharing scheme are slightly lower than the power sharing scheme. At medium and high transmit powers, however, the power sharing scheme outperform significantly the time-sharing scheme one.

Figure 4 shows the impact of relaxing the Gaussian interference approximation and assuming that the interference in the second channel when decoding m_1 is isotropic on the power shell. In particular, the lower bound on ϵ_1 is plotted for both the Gaussian interference approximation (using Theorem 1) and interference on the power shell (using the method in Sec. III-B). Observe that when $P_1 = P_2 = 2$ and the number of channel uses n_1 and n_2 are varied, the gap between the lower bounds is small. This suggests that using the Gaussian interference approximation does not significantly affect the conclusions drawn from the analysis in Theorems 1 and 2.

V. Conclusions

We derived tradeoffs between error probability, message set size, and the (finite) number of channel uses for joint channel coding of two consecutive messages with heterogeneous decoding deadlines. We considered a point-to-point communication where messages arrive at different times and are subject to heterogeneous decoding delay constraints. We proposed a joint coding scheme accounting for overlapping transmission windows in a scenario with two messages. Finally, we analyzed the probability of error in the finite block length regime and identified significant potential performance improvements over standard time sharing schemes.

VI. Acknowledgment

This work was funded by INRIA Nokia Bell Labs ADR “Network Information Theory” and by the French National Agency for Research (ANR) under grant ANR-16-CE25-0001 - ARBURST.

Appendix A
Proof of Theorem 2

In the following, we sketch the proof of Theorem 2. See [14] for full details.

For a given $\beta \in [0,1]$, define the following random coding distributions:

$$P_{\boldsymbol{X}_1}(\boldsymbol{x}) \triangleq \frac{\delta(||\boldsymbol{x}_1||_2^2 - n_1P_1)}{S_{n_1}(\sqrt{n_1P_1})}, \tag{30}$$

$$P_{\boldsymbol{X}_2}^{(1)}(\boldsymbol{x}) \triangleq \frac{\delta(||\boldsymbol{x}_2^{(1)}||_2^2 - n_2\beta P_2)}{S_{n_2}(\sqrt{n_2\beta P_2})}, \tag{31}$$

$$P_{\boldsymbol{X}_2}^{(2)}(\boldsymbol{x}) \triangleq \frac{\delta(||\boldsymbol{x}_2^{(2)}||_2^2 - n_2(1-\beta)P_2)}{S_{n_2}(\sqrt{n_2(1-\beta)P_2})}, \tag{32}$$

$$P_{\boldsymbol{X}_3}(\boldsymbol{x}) \triangleq \frac{\delta(||\boldsymbol{x}_3||_2^2 - n_3P_3)}{S_{n_3}(\sqrt{n_3P_3})}, \tag{33}$$

where $\delta(\cdot)$ is the Dirac delta and $S_n(r)$ is the surface area of a radius-r sphere in $\mathbb{R}^n$. Over the first and second channels, we sample M length-n_1+n_2 codewords independently from $P_{\boldsymbol{X}_1}(\boldsymbol{x}) \times P_{\boldsymbol{X}_2}^{(1)}(\boldsymbol{x})$ to encode m_1. Over the second and third channels, we sample M length-n_2+n_3 codewords independently from $P_{\boldsymbol{X}_2}^{(2)}(\boldsymbol{x}) \times P_{\boldsymbol{X}_3}(\boldsymbol{x})$ to encode m_2.

We start by upper bounding ϵ_1. Given $\boldsymbol{Y}_1$ and $\boldsymbol{Y}_2$, the decoder selects the message m_1 satisfying

$$q(\boldsymbol{x}(m_1), \boldsymbol{y}_1, \boldsymbol{y}_2) > \max_{\tilde{m} \in \{1,\dots,M\}\setminus m_1} q(\boldsymbol{x}(\tilde{m}), \boldsymbol{y}_1, \boldsymbol{y}_2) \tag{34}$$

where

$$q(\boldsymbol{x}(m_1), \boldsymbol{y}_1, \boldsymbol{y}_2) \triangleq \ln \frac{P_{\boldsymbol{Y}_1|\boldsymbol{X}_1}(\boldsymbol{y}_1|\boldsymbol{x}_1) \times P_{\boldsymbol{Y}_2|\boldsymbol{X}_2^{(1)}}(\boldsymbol{y}_2|\boldsymbol{x}_2^{(1)})}{P_{\boldsymbol{Y}_1}(\boldsymbol{y}_1) \times P_{\boldsymbol{Y}_2}(\boldsymbol{y}_2)}. \tag{35}$$

Definition 2 (Random coding union bound): There exists an $(n_1+n_2, M, \epsilon_1, P_1, P_2)$-code satisfying

$$\epsilon_1 \le \mathbb{E}\Big[\min\{1, M\mathbb{P}(q(\bar{\boldsymbol{X}}; \boldsymbol{Y}_1, \boldsymbol{Y}_2) \ge q(\boldsymbol{X}; \boldsymbol{Y}_1, \boldsymbol{Y}_2)|\boldsymbol{X}, \boldsymbol{Y}_1, \boldsymbol{Y}_2)\}\Big]. \tag{36}$$

where the random variables $(\bar{\boldsymbol{X}}, \boldsymbol{X}, \boldsymbol{Y}_1, \boldsymbol{Y}_2)$ are distributed as $P_{\boldsymbol{X}_1}(\bar{\boldsymbol{x}}) \times P_{\boldsymbol{X}_2^{(1)}}(\bar{\boldsymbol{x}}) \times P_{\boldsymbol{X}_1}(\boldsymbol{x}) \times P_{\boldsymbol{X}_2^{(1)}}(\boldsymbol{x}) \times P_{\boldsymbol{Y}_1|\boldsymbol{X}_1}(\boldsymbol{y}_1|\boldsymbol{x}_1) \times P_{\boldsymbol{Y}_2|\boldsymbol{X}_2^{(1)}}(\boldsymbol{y}_2|\boldsymbol{x}_2^{(1)})$.

To ease the calculation of the above expectation, we first bound the probability $\mathbb{P}(q(\bar{\boldsymbol{X}}, \boldsymbol{Y}_1, \boldsymbol{Y}_2) \ge t|\boldsymbol{Y}_1 = \boldsymbol{y}_1, \boldsymbol{Y}_2 = \boldsymbol{y}_2)$ for a constant $t \in \mathbb{R}$. For simplicity, define

$$g(t, \boldsymbol{y}_1, \boldsymbol{y}_2) \triangleq \mathbb{P}(q(\bar{\boldsymbol{X}}, \boldsymbol{Y}_1, \boldsymbol{Y}_2) \ge t|\mathcal{E}_1), \tag{37}$$

where $\mathcal{E}_1 \triangleq \{\boldsymbol{Y}_1 = \boldsymbol{y}_1, \boldsymbol{Y}_2 = \boldsymbol{y}_2\}$. By Bayes rule, we have $P_{\boldsymbol{X}_1|\boldsymbol{Y}_1}(\boldsymbol{x}|\boldsymbol{y}_1) \times P_{\boldsymbol{Y}_1}(\boldsymbol{y}_1) = P_{\boldsymbol{X}_1}(\boldsymbol{x}) \times P_{\boldsymbol{Y}_1|\boldsymbol{X}_1}(\boldsymbol{y}_1|\boldsymbol{x})$ and $P_{\boldsymbol{X}_2^{(1)}|\boldsymbol{Y}_2}(\boldsymbol{x}|\boldsymbol{y}_2) \times P_{\boldsymbol{Y}_2}(\boldsymbol{y}_2) = P_{\boldsymbol{X}_2^{(1)}}(\boldsymbol{x}) \times P_{\boldsymbol{Y}_2|\boldsymbol{X}_2^{(1)}}(\boldsymbol{y}_2|\boldsymbol{x})$ and as a result:

$$\begin{aligned} &P_{\boldsymbol{X}_1}(\bar{\boldsymbol{x}}) P_{\boldsymbol{X}_2^{(1)}}(\bar{\boldsymbol{x}}) \\ &= P_{\boldsymbol{X}_1|\boldsymbol{Y}_1}(\bar{\boldsymbol{x}}|\boldsymbol{y}_1) P_{\boldsymbol{X}_2^{(1)}|\boldsymbol{Y}_2}(\bar{\boldsymbol{x}}|\boldsymbol{y}_2) \exp(-q(\bar{\boldsymbol{x}}, \boldsymbol{y}_1, \boldsymbol{y}_2)). \end{aligned} \tag{38}$$

Thus

$$\begin{aligned} &g(t, \boldsymbol{y}_1, \boldsymbol{y}_2) \\ &= \int_{\bar{\boldsymbol{x}}} \mathbb{1}\{q(\bar{\boldsymbol{x}}, \boldsymbol{y}_1, \boldsymbol{y}_2) \ge t\} P_{\boldsymbol{X}_1}(\bar{\boldsymbol{x}}) P_{\boldsymbol{X}_2^{(1)}}(\boldsymbol{x}) d\bar{\boldsymbol{x}} \\ &= \mathbb{E}\left[\exp(-q(\boldsymbol{X}, \boldsymbol{Y}_1, \boldsymbol{Y}_2)) \mathbb{1}\{q(\boldsymbol{X}, \boldsymbol{Y}_1, \boldsymbol{Y}_2) \ge t\} | \mathcal{E}_1\right]. \end{aligned} \tag{39}$$

Following the similar steps as in [12, Section IV] with an extension to the parallel channels, we can prove that for fixed sequences $\boldsymbol{y}_1 \in \mathbb{R}^{n_1}$ and $\boldsymbol{y}_2 \in \mathbb{R}^{n_2}$ and a constant $t \in \mathbb{R}$:

$$g(t, \boldsymbol{y}_1, \boldsymbol{y}_2) \le G_1 \cdot e^{-t} \tag{40}$$

where

$$G_1 \triangleq \frac{1}{1-\exp(-\eta)} \min\left\{\frac{L(P_2, s_2)\eta}{\sqrt{n_2\beta P_2 s_2}}, \frac{L(P_1, s_1)\eta}{\sqrt{n_1 P_1 s_1}}\right\}. \tag{41}$$

In (41), $\eta > 0$ is a constant, $s_1 = \frac{1}{n_1}||\boldsymbol{y}_1||_2^2$, $s_2 = \frac{1}{n_2}||\boldsymbol{y}_2||_2^2$ and

$$L(P, s) \triangleq \frac{(2Ps)^2}{\sqrt{2\pi}} \cdot \sqrt{\frac{1 + 4Ps - \sqrt{1+4Ps}}{(\sqrt{1+4Ps} - 1)^5}}. \tag{42}$$

Given the Gaussian random noise vectors $\boldsymbol{Z}_1 \sim \mathcal{N}(0, I_{n_1}\sigma_1^2)$ and $\boldsymbol{Z}_2 \sim \mathcal{N}(0, I_{n_2}\sigma_2^2)$, we introduce the following sets of "typical" channel outputs:

$$\mathcal{F}_1 \triangleq \{\boldsymbol{y}_1 \in \mathbb{R}^{n_1} : \frac{1}{n_1}||\boldsymbol{y}_1||_2^2 \in [P_1 + \sigma_1^2 - \delta_1, P_1 + \sigma_1^2 + \delta_1]\}$$
$$\mathcal{F}_2 \triangleq \{\boldsymbol{y}_2 \in \mathbb{R}^{n_2} : \frac{1}{n_2}||\boldsymbol{y}_2||_2^2 \in [P_2 + \sigma_2^2 - \delta_2, P_2 + \sigma_2^2 + \delta_2]\},$$

where δ_1 and δ_2 are constants. By defining the following event,

$$\mathcal{E}_2 \triangleq \{\boldsymbol{Y}_1 \in \mathcal{F}_1, \boldsymbol{Y}_2 \in \mathcal{F}_2\}, \tag{43}$$

then we can rewrite the RCU bound in (36) as:

$$\begin{aligned}
\epsilon_1 &\leq \mathbb{E}\left[\min\{1, Mg(q(\boldsymbol{X};\boldsymbol{Y}_1,\boldsymbol{Y}_2),\boldsymbol{Y}_1,\boldsymbol{Y}_2)\}\right] \qquad (44)\\
&\leq \mathbb{P}\{\mathcal{E}_2^c\}\\
&+\mathbb{E}\left[\min\{1, Mg(q(\boldsymbol{X};\boldsymbol{Y}_1,\boldsymbol{Y}_2),\boldsymbol{Y}_1,\boldsymbol{Y}_2)\}\middle|\mathcal{E}_2\right]\mathbb{P}\{\mathcal{E}_2\}\\
&= \mathbb{P}\{\mathcal{E}_2^c\} + \mathbb{E}\left[\min\{1, MG_1 e^{-q(\boldsymbol{X};\boldsymbol{Y}_1,\boldsymbol{Y}_2)}\}\middle|\mathcal{E}_2\right]\mathbb{P}\{\mathcal{E}_2\}\\
&\leq \mathbb{P}\{\mathcal{E}_2^c\} + \mathbb{P}\{\mathcal{E}_2\}\Big(\mathbb{P}\left(q(\boldsymbol{X},\boldsymbol{Y}_1,\boldsymbol{Y}_2) \leq \ln(MG_1)\middle|\mathcal{E}_2\right)\\
&+MG_1\cdot\mathbb{E}\left[\mathbb{1}\{q(\boldsymbol{X},\boldsymbol{Y}_1,\boldsymbol{Y}_2) > \ln(MG_1)\}e^{-q(\boldsymbol{X},\boldsymbol{Y}_1,\boldsymbol{Y}_2)}\middle|\mathcal{E}_2\right]\Big)\\
&\leq \mathbb{P}\{\mathcal{E}_2^c\} + \mathbb{P}\{\mathcal{E}_2\}\left(\mathbb{P}\left(q(\boldsymbol{X},\boldsymbol{Y}_1,\boldsymbol{Y}_2) \leq \ln(MG_1)\middle|\mathcal{E}_2\right) + G_1\right).
\end{aligned}$$

To calculate the following probability

$$\mathbb{P}\left(q(\boldsymbol{X},\boldsymbol{Y}_1,\boldsymbol{Y}_2) \leq \ln(MG_1)\middle|\mathcal{E}_2\right), \tag{45}$$

let $P^*_{Y_1}(y_1) = \mathcal{N}(y_1; 0, P_1+\sigma_1^2)$ and $P^*_{Y_2}(y_2) = \mathcal{N}(y_2; 0, P_2+\sigma_2^2)$ be the capacity-achieving output distributions over the first channel and the second channel, respectively. Then as shown in [5, Lemma 6], given that $\boldsymbol{y}_1 \in \mathcal{F}_1$ and $\boldsymbol{y}_2 \in \mathcal{F}_2$, we have

$$\sup_{\boldsymbol{y}_1\in\mathcal{F}_1}\frac{P_{\boldsymbol{Y}_1}(\boldsymbol{y}_1)}{P^*_{\boldsymbol{Y}_1}(\boldsymbol{y}_1)} \leq J_1 \quad \text{and} \quad \sup_{\boldsymbol{y}_2\in\mathcal{F}_2}\frac{P_{\boldsymbol{Y}_2}(\boldsymbol{y}_2)}{P^*_{\boldsymbol{Y}_2}(\boldsymbol{y}_2)} \leq J_2, \tag{46}$$

where J_1 and J_2 are finite constants. Thus

$$\begin{aligned}
&\mathbb{P}\left(q(\boldsymbol{X},\boldsymbol{Y}_1,\boldsymbol{Y}_2) \leq \ln(MG_1)\middle|\mathcal{E}_2\right)\mathbb{P}\{\mathcal{E}_2\}\\
&\leq \mathbb{P}\left(q\left(\boldsymbol{X},\boldsymbol{Y}_1^*,\boldsymbol{Y}_2^*\right) \leq \ln(MG_1J_1J_2)\right)\\
&= \mathbb{P}\Bigg[\frac{\Omega_1}{(1+\Omega_1)}||\frac{\boldsymbol{y}_1^*}{\sigma_1} - \frac{(1+\Omega_1)}{\sigma_1\Omega_1}\boldsymbol{x}_1||^2\\
&+\frac{\Omega_2}{(1+\Omega_2)}||\frac{\boldsymbol{y}_2^*}{\sqrt{(1+\Omega_3)}\sigma_2} - \frac{(1+\Omega_2)}{\sigma_2\Omega_2\sqrt{(1+\Omega_3)}}\boldsymbol{x}_2^{(1)}||^2 > \Delta_1\Bigg]\\
&\overset{(a)}{=} \mathbb{P}\left[\frac{\Omega_1}{1+\Omega_1}v_1 + \frac{\Omega_2}{1+\Omega_2}v_2 > \Delta_1\right]\\
&\overset{(b)}{=} 1 - F_{Q_1}(\Delta_1), \qquad (47)
\end{aligned}$$

where in (a),

$$\Delta_1 \triangleq -2\ln(MG_1J_1J_2) + \frac{1}{n_1}\ln(1+\Omega_1) + \frac{1}{n_2}\ln(1+\Omega_2) + n_1 + \beta n_2. \tag{48}$$

and

$$v_1 \triangleq ||\frac{\boldsymbol{y}_1}{\sigma_1} - \frac{1+\Omega_1}{\sigma_1\Omega_1}\boldsymbol{x}_1||^2, \tag{49}$$

$$v_2 \triangleq ||\frac{\boldsymbol{y}_2}{\sigma_2\sqrt{(1+\Omega_3)}} - \frac{(1+\Omega_2)}{\Omega_2\sigma_2\sqrt{1+\Omega_3}}\boldsymbol{x}_2^{(1)}||^2. \tag{50}$$

Since v_1 and v_2 follow non-central chi-square distributions, i.e., $v_1 \sim \mathcal{X}^2(n_1, \frac{n_1}{\Omega_1})$ and $v_2 \sim \mathcal{X}^2(n_2, \frac{n_2}{\Omega_2})$, then in step (b) we define

$$Q_1 \triangleq \frac{\Omega_1}{1+\Omega_1}v_1 + \frac{\Omega_2}{1+\Omega_2}v_2 \tag{51}$$

and F_{Q_1} as the CDF of Q_1. To calculate this CDF, we use the analysis on the distribution of a linear combination of independent non-central chi-square variables provided in [13, Section 3] that is based on an expansion of Laguerre function.

Finally, to calculate the probability $\mathbb{P}\{\mathcal{E}_2\}$, we use *Cramer's theorem* in [9] and obtain

$$\mathbb{P}\{\mathcal{E}_2\} \geq (1 - \exp(-\kappa_1 n_1 \delta_1^2))(1 - \exp(-\kappa_2 n_2 \delta_2^2)) \tag{52}$$

for some constants κ_1 and κ_2. By setting $\delta_1 = n_1^{-1/3}$ and $\delta_2 = n_2^{-1/3}$, so

$$\mathbb{P}\{\boldsymbol{Y}_1 \in \mathcal{F}_1, \boldsymbol{Y}_2 \in \mathcal{F}_2\} \geq 1 - \zeta_1, \tag{53}$$

where by defining $\zeta_1 \triangleq \exp(-\kappa_1 n_1^{1/3}) + \exp(-\kappa_2 n_2^{1/3}) - \exp(-\kappa_1 n_1^{1/3})\exp(-\kappa_2 n_2^{1/3})$, the bound in (23) is proved. Similarly one can prove the bound in (24).

REFERENCES

[1] P. Popovski, Č. Stefanović, J. J. Nielsen, E. de Carvalho, M. Angjelichinoski, K. F. Trillingsgaard and A. Bana, "Wireless access in ultra-reliable low-latency communication (URLLC)," in *IEEE Trans. on Communications*, vol. 67, no. 8, pp. 5783–5801, Aug. 2019.

[2] X. Ge, "Ultra-reliable low-latency communications in autonomous vehicular networks," in *IEEE Trans. on Vehicular Technology*, vol. 68, no. 5, pp. 5005–5016, May 2019.

[3] M. Shirvanimoghaddam et al., "Short block-length codes for ultra-reliable low latency communications," in *IEEE Communications Magazine*, vol. 57, no. 2, pp. 130–137, Feb. 2019.

[4] X. Zhang, J. Wang and H. V. Poor, "Statistical QoS-driven energy-efficiency optimization for URLLC over 5G mobile wireless networks in the finite blocklength regime," in *Info. Sciences and Systems,* pp. 1–6, 2019.

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

[6] N. Shulman and M. Feder. "Static broadcasting", *IEEE International Symposium on Info. Theory (ISIT)*, 2000.

[7] N. Shulman, "Communication over an unknown channel via common broadcasting". PhD thesis. *Tel-Aviv University*, 2003.

[8] M. Langberg and M. Effros," Beyond capacity: The joint time-rate region", *arXiv:2101.12236v1*, Jan. 2021.

[9] I. Csiszàr and J. Körner, "Information theory", *New York: Academic Press*, 1981.

[10] J. Scarlett, "On the dispersions of the Gel'fand - Pinsker channel and dirty paper coding," in *IEEE Trans. on Info Theory*, vol. 61, no. 9, pp. 4569-4586, Sept. 2015.

[11] T. Erseghe, "Coding in the finite-blocklength regime: Bounds based on Laplace integrals and their asymptotic approximations," in *IEEE Trans. on Info. Theory*, vol. 62, no. 12, pp. 6854–6883, Dec. 2016.

[12] V. Y. F. Tan and M. Tomamichel, "The third-order term in the normal approximation for the AWGN channel," in *IEEE Trans. on Info. Theory*, vol. 61, no. 5, pp. 2430–2438, May 2015.

[13] A. Castaño-Martinez and F. López-Blázquez, "Distribution of a sum of weighted non-central chi-square variables", *TEST*, vol. 14, no. 2, pp. 397–415, 2005.

[14] H. Nikbakht, M. Egan, J-M. Gorce, " Joint channel coding of consecutive packets with heterogeneous decoding deadlines in the finite blocklength regime", [Research Report] Inria - Research Centre Grenoble – Rhône-Alpes. 2021. Available: https://hal.inria.fr/hal-03360567