URLLC-eMBB Slicing to Support VR Multimodal Perceptions over Wireless Cellular Systems

Jihong Park and Mehdi Bennis

Abstract-Virtual reality (VR) enables mobile wireless users to experience multimodal perceptions in a virtual space. In this paper we investigate the problem of concurrent support of visual and haptic perceptions over wireless cellular networks, with a focus on the downlink transmission phase. While the visual perception requires moderate reliability and maximized rate, the haptic perception requires fixed rate and high reliability. Hence, the visuo-haptic VR traffic necessitates the use of two different network slices: enhanced mobile broadband (eMBB) for visual perception and ultra-reliable and low latency communication (URLLC) for haptic perception. We investigate two methods by which these two slices share the downlink resources orthogonally and non-orthogonally, respectively. We compare these methods in terms of the just-noticeable difference (JND), an established measure in psychophysics, and show that non-orthogonal slicing becomes preferable under a higher target integrated-perceptual resolution and/or a higher target rate for haptic perceptions.

Index Terms—Virtual reality (VR) multimodal perception, VR traffic slicing, URLLC-eMBB multiplexing, stochastic geometry.

I. Introduction

Virtual reality (VR) is often seen as one of the most important applications in 5G cellular systems [1]–[3]. As in real life, mobile VR users can interact in the virtual space immersively with virtual objects that may stimulate their multiple sensory organs. This multimodal VR perception happens, for example, when a VR user measures the size of a virtual object through visual and haptic senses. In this study, we consider the problem of supporting such visuo-haptic VR perceptions over wireless cellular networks, and focus on the downlink design.

The key challenge is that these two perceptions have completely different cellular service requirements. In fact, visual traffic requires high data rate and relatively low reliability with packet error rate (PER) on the order of $10^{-1} \sim 10^{-3}$ [4], [5]. This requirements can be supported mostly through enhanced mobile broadband (eMBB) links [6]. Haptic traffic, by contrast, should guarantee a fixed target rate and high reliability with PER on the order of $10^{-4} \sim 10^{-5}$ [7], [8], which can be satisfied via ultra-reliable and low latency communication (URLLC) links [9], [10].

Furthermore, in order to render a smooth multimodal experience, the PERs associated with the visuo-haptic VR perceptions should guarantee a target perceptual resolution. To be precise, the perceptual resolution is commonly measured by using the just-noticeable difference (JND) in psychophysics, a field of study that focuses on the quantitative relation between physical stimulus and perception [4], [11], [12]. Following Weber's law, JND describes the minimum detectable change amount of perceptual inputs, e.g., 3 mm for the object size measurement using visuo-haptic perceptions [11]. According to psychophysical experiments, the JND of the aggregate visuo-haptic perception is the harmonic mean of the squared

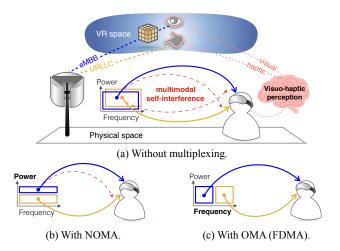


Fig. 1. An illustration of visuo-haptic VR perceptions and the required URLLC-eMBB traffic slicing: (a) without multiplexing; (b) with NOMA; and (c) with OMA.

JNDs of the individual perceptions [11], in which the JND of each perception is proportional to the PER [4].

As a result, the PERs associated with the visuo-haptic VR traffic should be adjusted so as to achieve a target JND, while abiding by the eMBB and URLLC service objectives in terms of PERs and data rates. Due to the discrepancy of the visual and haptic service requirements, it is difficult to support both perceptions through either eMBB or URLLC links. Hence, it is necessary to *slice the visuo-haptic VR traffic into eMBB and URLLC links*, leading to URLLC-eMBB multimodal transmissions. Unfortunately, such multimodal transmissions bring about *multimodal self-interference*, which is manifested through an actual wireless interference as visualized in Figs. 1-a and b, or via the necessity to share resources as shown in Figs. 1-b and c.

This critical self-interference can be alleviated by multiplexing the URLLC-eMBB multimodal transmissions over the transmit power domain with successive interference cancellation (SIC) at reception, i.e., downlink non-orthogonal multiple access (NOMA) [13], as illustrated in Fig. 1-b. Alternatively, as Fig. 1-c shows, the self-interference can be avoided via orthogonal multiple access (OMA) such as frequency division multiple access (FDMA). In this paper, using stochastic geometry, we investigate the optimal design of NOMA and OMA to support visuo-haptic VR perceptions while coping with the multimodal self-interference in a large-scale downlink system.

Related Works – The communication and computation resource management of mobile VR networks has recently been investigated in [2], [3], [14]–[16], particularly under a VR social network application [3] and a VR gaming scenario [16]. The end-to-end latency has been studied in [14] for a single-

J. Park and M. Bennis are with the Centre for Wireless Communications, University of Oulu, Oulu 90014, Finland (email: {jihong.park, mehdi.bennis}@oulu.fi).

cell scenario and in [3], [15] for a multi-cell scenario. These works focus primarily on supporting either visual or haptic perceptions. Towards supporting multimodal perceptions, suitable network architecture and coding design have been proposed in [7], [8], while not specifying the requirements on the wireless links. In an uplink single-cell system, orthogonal/nonorthogonal multiplexing of URLLC and eMBB links has been optimized by exploiting their reliability diversity in [17].

Contributions – The main contributions of this work are summarized as follows.

- To the best of our knowledge, this is the first work that combines both visual and haptic modalities in the context of mobile VR network design.
- To support visuo-haptic VR perceptions, an optimal downlink NOMA design with reliability-ordered SIC has been proposed (see Lemma 2 and Proposition 4).
- Compared to an OMA baseline (see Proposition 2), it has been observed that the proposed NOMA becomes preferable under a higher target integrated-perceptual resolution and/or a higher target rate for haptic perceptions (see Fig. 3).
- By using stochastic geometry, closed-form average rate expressions have been derived for downlink URLLCeMBB multiplexing under OMA and NOMA in a largescale cellular network (see Propositions 1 and 3).

II. SYSTEM MODEL AND PROBLEM FORMULATION

In this section, we first introduce the downlink system operation of OMA and NOMA under a single-cell scenario, and describe its extension to the operation under a large-scale network. Then, we specify visuo-haptic perceptions, followed by the problem formulation of visuo-haptic VR traffic slicing and multiplexing.

The user under study requests visuo-haptic VR perceptions that are supported through URLLC-eMBB cellular links. We use the subscript $i \in \{1, 2\}$ to indicate the URLLC link Link₁ with i = 1 and the eMBB link Link₂ with i = 2. The subscript $j \in \{0, N\}$ identifies OMA and NOMA, respectively.

A. Single-Cell Channel Model with OMA and NOMA

In a downlink scenario, we consider a single user that is associated with a single base station (BS). For both OMA and NOMA, the transmissions of the BS at a given time occupy up to the frequency bandwidth normalized to one, which is divided into the K number of miniblocks. Each miniblock is assumed to be within the frequency-time channel coherence intervals. The channel coefficients are thus constants within each miniblock, and fade independently across different miniblocks over frequency and time. The transmit power of the BS is equally divided for each miniblock, normalized to one.

1) Single-Cell OMA: A set K_i of miniblocks are allocated to Link_i, with $|\mathcal{K}_1| + |\mathcal{K}_2| = K$. Each set corresponds to a fraction $w_{i,O} = |\mathcal{K}_i|/K > 0$, with $w_{1,O} + w_{2,O} = 1$. The transmit power allocations to Link₁ and Link₂ are set as the maximum transmit power per miniblock. Denoting as $\beta_{i,O} \leq 1$ the transmit power allocation fraction to miniblock $k \in \mathcal{K}_i$, this corresponds to the allocations that equal $\beta_{1,O} = \beta_{2,O} = 1$.

The user's received signal-to-noise ratio (SNR) is determined by small-scale and large-scale fading gains. For a given user-BS association distance r, the large-scale fading gains of Link₁ and Link₂ are identically given as $r^{-\alpha}$ with the path loss exponent $\alpha > 2$. For miniblock $k \in \mathcal{K}_i$, the small-scale fading gain $g_i^{(k)}$ is an exponential random variable with unit mean, which is independent and identically distributed (i.i.d.) across different miniblocks. The user's received SNR of $Link_i$ through miniblock $k \in \mathcal{K}_i$ is then expressed as

$$\mathsf{SNR}_{i,\mathsf{O}}^{(k)} = \frac{g_i^{(k)} r^{-\alpha}}{N},\tag{1}$$

where N is the noise spectral density of a single miniblock.

2) Single-Cell NOMA: The entire bandwidth is utilized for both links in NOMA, i.e., the miniblock allocation fractions equal $w_{1,N} = w_{2,N} = 1$. This is enabled by transmitting the superposition of the signals intended for Link₁ and Link₂, with their different transmit power allocations, and then by decoding the signals with SIC at reception [13]. The transmit power allocated to Link_i has a fraction $\beta_{i,N}$ of the maximum transmit power per miniblock, with $\beta_{1,N} + \beta_{2,N} = 1$.

At reception, unless otherwise noted, we consider Link₁ is decoded prior to Link₂. This SIC order implicitly captures the low-latency guarantee of Link₁, as addressed in [17] for an uplink scenario. Furthermore, it improves the overall NOMA system performance due to the reliability diversity of Link_1 and Link₂, to be elaborated in Sect. III-B.

With the said SIC order, the signal intended for Link₁ is first decoded, while treating the signal for Link₂ as noise, i.e., multimodal self-interference. The decoded signal is then removed by applying SIC, and the remaining signal for Link₂ is finally decoded without self-interference. The user's received SNR for Link_i through miniblock $k \in \mathcal{K}_i$ is thereby obtained as

$$SNR_{1,N}^{(k)} = \frac{\beta_{1,N}g_i^{(k)}r^{-\alpha}}{\beta_{2,N}g_i^{(k)}r^{-\alpha} + N} \quad \text{and}$$

$$SNR_{2,N}^{(k)} = \frac{\beta_{2,N}g_i^{(k)}r^{-\alpha}}{N},$$
(3)

$$\mathsf{SNR}_{2,N}^{(k)} = \frac{\beta_{2,N} g_i^{(k)} r^{-\alpha}}{N},\tag{3}$$

Note that all the fading gains of Link₁ and Link₂ are identically $g_i^{(k)}r^{-\alpha}$ since their channels are identical.

B. Channel Model under a Stochastic Geometric Network

By using stochastic geometry, the aforementioned singlecell operation of OMA and NOMA is extended to a largescale multi-cell scenario as follows. The BSs under study are deployed in a two-dimensional Euclidean plane, according to a stationary Poisson point process (PPP) Φ with density λ , where the coordinates x of a BS belongs to Φ . Following the single-cell operation, each BS serves a single user through its Link₁ and Link₂.

The locations of users follow an arbitrary stationary point process. Each user associates with the nearest BS, and downloads the visuo-haptic VR traffic through the Link₁ and Link₂ of the BS. Following [18], we focus our analysis on a typical user that is located at the origin and associated with the nearest BS located at position x_o of the plane. This typical user captures the spatially-averaged performance, thanks to Slyvnyak's theorem [19] and the stationarity of Φ .

In the previous single-cell scenario, interference occurs only from the multimodal self-interference under NOMA, as shown in (2). In addition to such intra-cell self-interference, extension to the stochastic geometric network model induces inter-cell interference. As done in [5], [18], [20], inter-cell interference is treated as noise, and is assumed to be large such that the maximum noise power N is negligible. In this interferencelimited regime, channel quality is measured not by SNR but by signal-to-interference ratio (SIR), as described next.

The inter-cell interference is measured by the typical user, and comes from the set $\Phi_o = \Phi \setminus \{x_o\}$ of the BSs that are not associated with the typical user. We consider every BS always utilizes the entire bandwidth and the maximum transmit power. The average inter-cell interference per miniblock is thus identically given under both OMA and NOMA. The instantaneous inter-cell interference varies due to small-scale fading. For each miniblock, any interfering link's small-scale fading is independent of the small-scale fading of the typical user's desired Link₁ and Link₂.

Under OMA, the typical user's received SIR of Link_i through miniblock $k \in \mathcal{K}_i$ is thereby given as

$$\mathsf{SIR}_{i,\mathsf{O}}^{(k)} = \frac{g_i^{(k)} |x_o|^{-\alpha}}{\sum_{x \in \Phi_o} g_x^{(k)} |x|^{-\alpha}},\tag{4}$$

where $g_x^{(k)}$'s are exponential random variables with unit mean, which are independent of $g_i^{(k)}$ and are i.i.d. across different interfering BSs. Likewise, under NOMA, the typical user's received SIR of Link_i through miniblock $k \in \mathcal{K}_i$ is expressed as

SIR of Link_i through miniblock
$$k \in \mathcal{K}_i$$
 is expressed as
$$SIR_{1,N}^{(k)} = \frac{\beta_1 g^{(k)} |x_o|^{-\alpha}}{\beta_2 g^{(k)} |x_o|^{-\alpha} + \sum_{x \in \Phi_o} g_x^{(k)} |x|^{-\alpha}} \quad \text{and} \qquad (5)$$

$$SIR_{2,N}^{(k)} = \frac{\beta_2 g^{(k)} |x_o|^{-\alpha}}{\sum_{x \in \Phi_o} g_x^{(k)} |x|^{-\alpha}}. \qquad (6)$$

$$\mathsf{SIR}_{2,N}^{(k)} = \frac{\beta_2 g^{(k)} |x_o|^{-\alpha}}{\sum_{x \in \Phi_0} g_x^{(k)} |x|^{-\alpha}}.$$
 (6)

It is noted that all the SIRs under NOMA and OMA are identically distributed across different miniblocks. For the typical user's Link₁ and Link₂, the large-scale fading gains are identical. Their small-scale fading gains are independent under OMA, but are fully-correlated under NOMA.

C. Average Rate with Decoding Success Guarantee

In a large-scale downlink cellular system with OMA and NOMA, we derive the typical user's average rate that guarantees a target decoding success probability. Decoding becomes successful when the instantaneous downlink rate exceeds the transmitted coding rate.

To facilitate tractable analysis, we consider that the instantaneous channel information is not available at each BS. With the channel information at a BS, one can improve the average rate by adjusting the transmit power [17] and/or the coding rate [18], [20]. In addition, we assume separate coding for each miniblock, which may loose frequency diversity gain compared to the coding across multiple miniblocks [17], [21].

With these assumptions and the SIRs that are identically distributed across miniblocks, average rate is determined by the decoding success probability for any single miniblock. Therefore, we drop the superscript (k) in $SIR_{i,j}^{(k)}$ and the small-scale fading terms, and derive the average rate in the sequel.

1) OMA: The typical user can decode the signal from Link $_i$ with the decoding success probability $p_{i,O}(t_{i,O})$ that equals

$$p_{i,O} = \Pr\left(\log(1 + \mathsf{SIR}_{i,O}) \ge r_{i,O}\right) \tag{7}$$

$$= \Pr\left(\mathsf{SIR}_{i,\mathsf{O}} \ge t_{i,\mathsf{O}}\right),\tag{8}$$

where $r_{i,O}$ is the coding rate per miniblock, which is hereafter rephrased as a target SIR threshold $t_{i,0} = e^{r_{i,0}} - 1$.

For the given target decoding success probability η_i of Link_i, the average rate $R_{i,O}(\eta_i)$ of Link_i is obtained by using outage capacity [22] as

$$R_{i,0}(\eta_i) = w_{i,0}\eta_i \cdot \sup\{\log(1 + t_{i,0}) : p_{i,0}(t_{i,0}) \ge \eta_i\} \quad (9)$$

$$= w_{i,0}\eta_i \cdot \log(1 + t_{i,0}^*), \tag{10}$$

where the optimal target SIR threshold $t_{i,O}^*$ satisfies

 $p_{i,\mathrm{O}}(t_{i,\mathrm{O}}^*) = \eta_i$, and thus equals $t_{i,\mathrm{O}}^* = p_{i,\mathrm{O}}^{-1}(\eta_i)$. Note that even when the coding block length of Link₁ is short, the average rate expression in (10) still holds, since the finite-block length rate under fading channels converges to the outage capacity [23].

2) NOMA: With the SIC order that decodes Link₁ prior to Link₂, the typical user's decoding success probabilities $p_{1,N}(t_{1,N})$ and $p_{2,N}(t_{1,N},t_{2,N})$ of Link₁ and Link₂ are given as

$$p_{1,N}(t_{1,N}) = \Pr\left(\mathsf{SIR}_{1,N} \ge t_{1,N}\right) \quad \text{and}$$
 (11)

$$p_{1,N}(t_{1,N}) = \Pr\left(\mathsf{SIR}_{1,N} \ge t_{1,N}\right) \text{ and } (11)$$

$$p_{2,N}(t_{1,N}, t_{2,N}) = p_{1,N}(t_{1,N}) \cdot \Pr\left(\mathsf{SIR}_{2,N} \ge t_{2,N} \mid \mathsf{SIR}_{1,N} \ge t_{1,N}\right).$$
(12)

Following [24], our SIC do not allow to decode the Link₂ signal after the decoding failure of the Link₁ signal. With a different SIC architecture that allows such a decoding attempt, (12) is regarded as the lower bound, as done in [17].

For the given target decoding success probability η_1 of Link₁, the average rate $R_{1,N}(\eta_1)$ of Link₁ is given as

$$R_{1,N}(\eta_1) = \eta_1 \cdot \sup\{\log(1 + t_{1,N}) : p_{1,N}(t_{1,N}) \ge \eta_1\}$$
 (13)

$$= \eta_1 \cdot \log(1 + t_{1,N}^*), \tag{14}$$

where the optimal target SIR threshold equals $t_{1,N}^* = p_{1,N}^{-1}(\eta_1)$. Similarly, for the given target decoding success probability η_2 of Link₂, the average rate $R_{2,N}(\eta_1,\eta_2)$ of Link₂ is given as

$$R_{2,N}(\eta_1, \eta_2) = \eta_2 \cdot \sup\{\log(1 + t_{2,N}) : p_{2,N}(t_{1,N}^*, t_{2,N}) \ge \eta_2\}$$
 (15)

$$= \eta_2 \cdot \log(1 + t_{2.N}^*), \tag{16}$$

where $t_{2,N}^* = p_{2,N}^{-1}(\eta_1, \eta_2)$.

D. Visuo-Haptic Perceptual Resolution

The resolution of human perceptions is often measured by using JND in psychophysics. In a psychophysical experiment, the JND is calculated as the minimum stimulus variation that can be detectable during 84% of the trials [11]. For a visuohaptic perception, its integrated JND is obtained by combining the JNDs of visual and haptic perceptions.

To elaborate, when individual haptic and visual perceptions have the perceived noise variances σ_1 and σ_2 , a human brain combines these perceptions, yielding an integrated noise variance σ_{12} that satisfies $\sigma_{12}^{-2} = \sigma_1^{-2} + \sigma_2^{-2}$. This relationship was first discovered in [11] by measuring the corresponding JNDs that are proportional to the perceived noise variances. The said relationship is thus read as $\gamma_{12}^{-2}=\gamma_1^{-2}+\gamma_2^{-2}$, where γ_{12} denotes the JND measured when using both visuohapric perceptions, while γ_1 and γ_2 identify the JNDs of the individual haptic and visual perceptions, respectively.

For individual visual perceptions, it has been reported by another experiment [4] that the PER is proportional to its sole JND γ_2 due to the resulting visual frame loss. Similarly, for individual haptic perceptions, it has been observed in [12] that the PER is proportional to the elapsed time to complete a given experimental task, which increases with the corresponding JND γ_1 due to the coarse perceptions. Based on such experimental evidence, we can write that $\gamma_i = (1 - \eta_i)^2$, where $(1 - \eta_i)$ represents the PER on Link_i.

Accordingly, the JND γ_{12} of visuo-haptic perceptions is obtained from the following equation $\gamma_{12}^{-2} = (1 - \eta_1)^{-2} + (1 - \eta_2)^{-2}.$

$$\gamma_{12}^{-2} = (1 - \eta_1)^{-2} + (1 - \eta_2)^{-2}. \tag{17}$$

In the following subsection, we adjust the target decoding success probabilities η_1 and η_2 of Link₁ and Link₂, so as to guarantee a target visuo-haptic JND $\theta > 0$, i.e., $\gamma_{12} = \theta$.

E. URLLC-eMBB Multiplexing Problem Formulation

In a downlink cellular system serving visuo-haptic VR traffic, haptic and visual perceptions are supported through Link₁ and Link₂, respectively. Each link pursues different service objectives as follows. The URLLC Link₁ aims at:

- (i) Ensuring a target decoding success probability $\eta_1 < 1$;
- (ii) Ensuring a target average rate $R_1 > 0$.

In contrast, the eMBB Link₂ aims at:

- (iii) Maximizing the average rate $R_{2,j}(\eta_i) > 0$; while
- (iv) Ensuring a target decoding success probability η_2 , with

In (iv), $\hat{\eta}_2 \leq \eta_2$ follows from an experimental evidence that the quality of visual perceptions dramatically drops when PER exceeds a certain limit, e.g., 10% PER that equals $\hat{\eta}_2 = 0.9$ [4].

In addition to these individual service objectives, with Link₁ and Link₂, their aggregate JND γ_{12} should guarantee a target visuo-haptic JND θ . The said service objectives and requirements of Link₁ and Link₂ are described in the following problem formulation.

$$\begin{array}{ll} \text{(P1)} & \underset{\{W_{i,j},\;\beta_{i,j},\;t_{i,j},\;\eta_{i}\}}{\text{maximize}} & R_{2,j}(\eta_{1},\eta_{2}) & \text{(18a)} \\ & \text{s.t.} & R_{1,j}(\eta_{1}) = \hat{R}_{1} & \text{(18b)} \\ & & \gamma_{12} = \theta & \text{(18c)} \end{array}$$

s.t.
$$R_{1,i}(\eta_1) = \hat{R}_1$$
 (18b)

$$\gamma_{12} = \theta \tag{18c}$$

$$\hat{\eta}_2 \le \eta_2 < \eta_1 < 1 \tag{18d}$$

The objective functions $R_{2,j}(\eta_1,\eta_2)$ and $R_{1,j}(\eta_1)$ in the constraint (18b) are obtained from (10) for OMA and from (14) for NOMA. In the constraint (18c), γ_{12} is provided in (17). Without loss of generality, we hereafter consider a sufficiently large number K of miniblocks so that the miniblock allocation fraction $w_{i,j}$ under OMA is treated as a continuous value.

III. OPTIMAL MULTIPLEXING OF VISUO-HAPTIC VR TRAFFIC UNDER OMA AND NOMA

In this section, we optimize the multiplexing of Link₁ and Link₂ that support visuo-haptic VR traffic. With P1, for OMA, we optimize the miniblock allocation from the unit frequency block to each link. For NOMA, on the other hand, we optimize the power allocation from the unit transmit power.

A. Optimal OMA

We aim at optimizing the miniblock allocation $w_{i,O} < 1$. To this end, for given $w_{i,O}$ and η_i , we derive the average rate $R_{i,O}(\eta_i) = w_{i,O}\eta_i \log(1 + t_{i,O}^*)$ with $t_{i,O}^* = p_{i,O}^{-1}(\eta_i)$. This requires taking the inverse function of $p_{i,O}(t_{i,O})$ in (8).

The typical user's $p_{i,O}(t_{i,O}) = \Pr(SIR_{i,O} \ge t_{i,O})$ is commonly referred to as SIR coverage probability, and its closed-form expression can be derived by using stochastic geometry [18], [25]. Namely, $p_{i,O}(t_{i,O})$ is given as

$$p_{i,O}(t_{i,O}) = \mathsf{E}_{x_o,\Phi_o} \left[\Pr\left(g_i > t_{i,O} | x_o|^{\alpha} \sum_{x \in \Phi_o} g_x | x|^{-\alpha} \right) \right]$$
(19)
$$= \mathsf{E}_{x_o,\Phi_o} \left[\prod_{x \in \Phi_o} \exp\left(-t_{i,O} g_x \left(\frac{|x_o|}{|x|} \right)^{\alpha} \right) \right]$$
(20)

$$=\mathsf{E}_{x_{o},\Phi_{o}}\left[\prod_{x\in\Phi_{o}}\left(1+t_{i,\mathsf{O}}\left(\frac{|x_{o}|}{|x|}\right)^{\alpha}\right)^{-1}\right],\tag{21}$$

where (20) comes from the complementary cumulative density function (CCDF) of g, and (21) follows from the Laplace

functional of the i.i.d. exponential variables g_x 's. Applying the probability generating functional (PGFL) of a stationary PPP Φ_o [19], we obtain the integral expression

(21) =
$$\mathsf{E}_{x_o} \left[\exp \left(-2\pi \lambda \int_{v > |x_o|} \left(1 + \frac{1}{t_{i,\mathsf{O}}} \left(\frac{v}{|x_o|} \right)^{\alpha} \right) v dv \right) \right]$$
 (22)
= $1/2 F_1 \left(1, -2/\alpha; 1 - 2/\alpha; -t_{i,\mathsf{O}} \right)$. (23)

$$= 1/2\Gamma_1(1, -2/\alpha, 1-2/\alpha, -\iota_{i,0}). \tag{23}$$

The last step is derived by using the void probability of Φ , with the definition of a Gauss hypergeometric function that equals ${}_2F_1(a,b;c;z)=\sum_{n=0}^{\infty}\frac{\Gamma(a+n)\Gamma(b+n)\Gamma(c)}{\Gamma(a)\Gamma(b)\Gamma(c+n)}\frac{z^n}{n!}$ where $\Gamma(x)$ is the gamma function.

In spite of the closed-form SIR coverage probability expression in (23), due to the hypergeometric function, the inverse function of $p_{i,O}(t_{o,O})$ can only be numerically computed. We resolve this problem by exploiting a simplified SIR coverage probability bound, proposed in our previous study [26].

Lemma 1. (Closed-form SIR coverage bounds) Denoting as $\overline{\rm SIR} = g|x_o|^{-\alpha}/\sum_{x\in\Phi_o}g_x|x|^{-\alpha}$, according to Theorem 1 and Corollary 1 in [26], the coverage probability of \overline{SIR} is upper and lower bounded as

$$\Pr(\overline{\mathsf{SIR}} \ge t) = (1 + ct)^{-\frac{2}{\alpha}},\tag{24}$$

where
$$\alpha/(\alpha-2) \le c \le [2\pi/\alpha \cdot \csc(2\pi/\alpha)]^{\alpha/2}$$
.

As validated in [26], these upper and lower bounds guarantee the convergence to the exact values respectively for $t \to 0$ and $t \to \infty$, which corresponds to high and low target decoding success probabilities, respectively. We henceforth treat the bounds as the approximated coverage probability of SIR.

Applying this to $p_{i,\mathrm{O}}(t_{i,\mathrm{O}}) = \Pr(\overline{\mathsf{SIR}} \geq t_{i,\mathrm{O}})$, we obtain $p_{i,\mathrm{O}}(t_{i,\mathrm{O}}) = (1 + ct_{i,\mathrm{O}})^{-2/\alpha}$. We thereby derive the inverse function $t_{i,\mathrm{O}}^* = p_{i,\mathrm{O}}^{-1}(\eta_i)$ that equals

$$t_{i,0}^* = G(\eta_i), \tag{25}$$

where $G(\eta_i) = (\eta_i^{-\alpha/2} - 1)/c$. The specific value of c and the approximation accuracy are to be specified in Sect. IV. Applying (25) to $R_{i,O}(\eta_i)$ in (10) yields the following result.

Proposition 1. (Closed-form average rate, OMA) For a given η_i , the average rate of Link_i under OMA is given as

$$R_{i,O}(\eta_i) = w_{i,O}\eta_i \log(1 + G(\eta_i)).$$
 (26)

Finally, applying this result to P1, we obtain the optimal OMA design and its corresponding average rate as below.

Proposition 2. (Maximum average rate, OMA) The maximum average rate of Link₂ under OMA is provided as

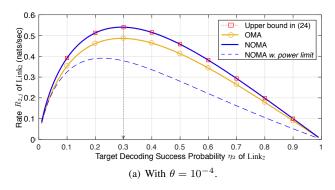
$$R_{2,0}^*(\eta_2^*) = w_{2,0}^* \eta_2^* \log(1 + G(\eta_2^*))$$
 where (27)

$$w_{2,0}^* = 1 - \frac{\hat{R}_1}{\eta_1 \log(1 + G(\eta_1^*))}, \tag{28}$$

$$\eta_1^* = 1 - \left(\frac{1}{\theta^2} - \frac{1}{(1 - \eta_2^*)^2}\right)^{-\frac{1}{2}},$$
(29)

$$\eta_2^* = \underset{\hat{\eta}_2 \le \eta_2 < \min(1, 1 - \theta\sqrt{2})}{\arg \max} R_{2,0}^*(\eta_2), \tag{30}$$

 $R_{2,\mathrm{O}}^*(\eta_2) = w_{2,\mathrm{O}}^*(\eta_2)\eta_2\log{(1+G(\eta_2))}, \text{ and } w_{2,\mathrm{O}}^*(\eta_2) \text{ is }$ obtained by replacing η_2^* with η_2 in $w_{2,0}^*$. Proof: See Appendix-A.



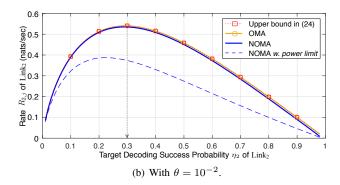


Fig. 2. Average rate $R_{2,j}^*(\eta_2)$ w.r.t. the target decoding success probability η_2 under OMA and NOMA ($\hat{R}_1 = 10^{-5}$ nats/sec, $\alpha = 4$).

Notice that $R_{2,\mathrm{O}}^*(\eta_2^*)$ still requires optimization with respect to η_2 , as shown in (30). In fact, the objective function to be maximized in (30) is concave over η_2 , as $w_{2,\mathrm{O}}^*(\eta_2)$ is a monotone increasing function of η_2 . Nevertheless, it is non-differentiable due to $w_{2,\mathrm{O}}^*(\eta_2)$, and can thus be optimized only by simulation.

In order to obtain the closed-form expression of $R_{2,\mathrm{O}}^*(\eta_2^*)$, we consider a sufficiently large $\hat{\eta}_2$ in (18d), such that the concave objective function in (30) monotonically decreases with $\eta_2 \geq \hat{\eta}_2$. In this regime, the optimum $R_{2,\mathrm{O}}^*(\eta_2^*)$ is achieved at $\eta_2^* = \hat{\eta}_2$, yielding the following corollary.

Corollary 1. (Closed-form maximum average rate, OMA) For $\hat{\eta}_2 \to 1$, the objective function in (30) is a monotone decreasing function of η_2 , and we thus obtain $R_{2,0}^*(\eta_2^*) = R_{2,0}^*(\hat{\eta}_2)$.

Note that Corollary 1 holds even under $\hat{\eta}_2$ that is much smaller than 1. For instance, as illustrated by simulation in Fig. 2, $R_{2,0}^*(\eta_2)$ monotonically decreases with $\eta_2 \geq 0.3$. This value is still much smaller than a practical value of $\hat{\eta}_2$, e.g., $\hat{\eta}_2 = 0.9$ as reported in [4], [5]. Therefore, we conclude that $R_{2,0}^*(\hat{\eta}_2)$ in Corollary 1 is the closed-form approximation. Validating its accuracy compared to Proposition 1 is deferred to Sect. IV.

B. Optimal NOMA

Our goal is to optimize the transmit power allocation fraction $\beta_{i,\mathrm{N}} \leq 1$ to Link_i . To this end, for given $\beta_{i,\mathrm{N}}$ and η_i , we derive the average rate $R_{i,\mathrm{N}}(\eta_i)$ provided in (14). The average rate is determined by the decoding order of the superposition of Link_1 and Link_2 signals. We propose the following optimal SIC decoding order.

Lemma 2. (Reliability-ordered SIC) In (P1) with $\eta_1 > \eta_2$, $R_{2,N}(\eta_2)$ is maximized when $Link_1$ is decoded prior to $Link_2$. Proof: By contradiction, we assume $Link_2$ is decoded prior to $Link_1$. We obtain the corresponding $p_{1,N}(t_{1,N},t_{2,N})$ by exchanging the subscripts 1 and 2 in (12), given as

$$p_{1,N}(t_{1,N}, t_{2,N}) = p_{2,N}(t_{2,N}) \cdot \Pr(\mathsf{SIR}_{1,N} \ge t_{1,N} \mid \mathsf{SIR}_{2,N} \ge t_{2,N}).$$

According to (15), it should guarantee $p_{1,N}(t_{1,N},t_{2,N}) \ge \eta_1$, which now reads

$$\Pr(\mathsf{SIR}_{1,N} \ge t_{1,N} \mid \mathsf{SIR}_{2,N} \ge t_{2,N}) \ge \eta_1/p_{2,N}(t_{2,N}).$$
 (32)

Since $\Pr(\mathsf{SIR}_{1,N} \geq t_{1,N} \mid \mathsf{SIR}_{2,N} \geq t_{2,N}) \leq 1$, the denominator of the RHS in (32) should satisfy $p_{2,N}(t_{2,N}) \geq \eta_1$. On Link₂,

this consequently imposes the target decoding success probability at least η_1 , which is higher than its original target η_2 . The resulting $R_{2,N}(\eta_1)$ is thus smaller than $R_{2,N}(\eta_2)$ in (14) with the opposite decoding order, finalizing the proof.

The proposed downlink SIC in order of reliability is consistent with the same SIC rule for uplink URLLC-eMBB multiplexing, proposed in [17]. Note that this reliability-ordered SIC may not always comply with a traditional SIC design that decodes the stronger signal first [13], [21]. If $\beta_{1,N} > \beta_{2,N}$, i.e., $\beta_{1,N} > 0.5$, then the reliability-ordered SIC follows the traditional power-ordered SIC. Otherwise, the reliability-ordered SIC should decode the weaker signal first, which is not allowed under the power-ordered SIC. Such restriction is examined by simulation in Sect. IV.

With the reliability-ordered SIC, as done for OMA in Sect. III-A, we exploit Lemma 1, and derive the closed-form decoding success probabilities, followed by the average rates.

First, when decoding the signal intended for Link₁, the decoding success probability $p_{1,N}(t_{1,N}) = \Pr\left(\mathsf{SIR}_{1,N} \geq t_{1,N}\right)$ is rephrased as $p_{1,N}(t_{1,N}) = \Pr\left(\overline{\mathsf{SIR}} \geq [1/t_{i,N} - \beta_{2,N}/\beta_{1,N}]^{-1}\right)$, which follows from (11) with $\mathsf{SIR}_{1,N}$ in (5). Applying Lemma 1 to this, we thereby obtain

$$p_{1,N}(t_{1,N}) = 1 - \left(1 + c\left[1/t_{1,N} - \beta_{2,N}/\beta_{1,N}\right]^{-1}\right)^{-2/\alpha}.$$
 (33)

By taking the inverse function, $t_{1,\mathrm{N}}^* = p_{1,\mathrm{N}}^{-1}(\eta_1)$ is derived as

$$t_{1,N}^* = \left[\beta_{2,N}/\beta_{1,N} + 1/G(\eta_1)\right]^{-1}$$
. (34)

Applying this to $R_{1,N}(\eta_1)$ in (14), we finally obtain the closed-form average rate, provided in Proposition 3 on the next page.

Next, when decoding the signal intended for Link₂, following from (12) with SIR_{2,N} in (6) and from $p_{1,N}(t_{1,N}^*) = \eta_1$ in (14), the decoding success probability $p_{2,N}(t_{1,N}^*, t_{2,N})$ is

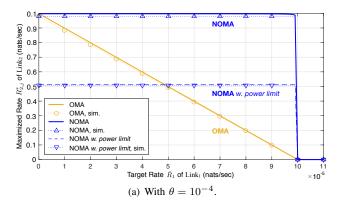
$$p_{2,N}(t_{1,N}^*, t_{2,N}) = \eta_1 \Pr\left(\mathsf{SIR}_{2,N} \ge t_{2,N} \mid \mathsf{SIR}_{1,N} \ge t_{1,N}^*\right)$$
 (35)

$$= \eta_1 \Pr\left(\overline{\mathsf{SIR}} \ge \max\{t_{2,N}/\beta_{2,N}, [1/t_{1,N}^* - \beta_{2,N}/\beta_{1,N}]^{-1}\}\right) (36)$$

$$\approx \eta_1 \Pr\left(\overline{\mathsf{SIR}} \ge t_{2,N}/\beta_{2,N}\right) = \eta_1 \Pr(\mathsf{SIR}_{2,N} \ge t_{2,N}). \tag{37}$$

In (36), we utilize the relationship $\Pr(\mathsf{SIR}_{2,\mathrm{N}} \geq t_{2,\mathrm{N}}) = \Pr(\overline{\mathsf{SIR}} \geq t_{2,\mathrm{N}}/\beta_{2,\mathrm{N}})$, as well as $\Pr\left(\mathsf{SIR}_{1,\mathrm{N}} \geq t_{1,\mathrm{N}}^*\right) = \Pr(\overline{\mathsf{SIR}} \geq [1/t_{i,\mathrm{N}}^* - \beta_{2,\mathrm{N}}/\beta_{1,\mathrm{N}}]^{-1})$. The approximation in (37) is valid for high η_1 , since $t_{1,\mathrm{N}}^*$ in (34) approaches 0 as $\eta_1 \to 1$.

is valid for high η_1 , since $t_{1,N}^*$ in (34) approaches 0 as $\eta_1 \to 1$. By comparing (37) and (35), we conclude that Link₂ signal is almost independently decoded with Link₁ decoding under high η_1 . In fact, SIR_{1,N} in (5) and SIR_{2,N} in (6) are correlated



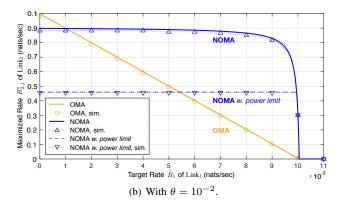


Fig. 3. Maximized average rates $R_{2,0}^*(\eta_2^*)$ and $R_{2,N}^*(\eta_1^*,\eta_2^*)$ w.r.t. a target rate \hat{R}_1 of Link₁ for different target visuo-haptic JND thresholds.

over both large-scale fading and small-scale fading, and the Link₁ decoding success may thus affect the subsequent Link₂ decoding success. With high η_1 , nevertheless, Link₁ can be decoded almost regardless of the channel quality, negligibly contributing to the Link2 decoding success.

Finally, applying Lemma 1 to (37), we obtain $t_{2.N}^* =$ $p_{2,N}^{-1}(\eta_1,\eta_2)$ that equals

$$t_{2,N}^* = \beta_{2,N} H(\eta_1, \eta_2),$$
 (38)

where $H(\eta_1, \eta_2) = [(\eta_1/\eta_2)^{\alpha/2} - 1]/c$. Plugging this in (16) yields the closed-form average rate as follows.

Proposition 3. (Closed-form average rate, NOMA) For a given η_i , the average rate of Link_i under NOMA with the reliability-ordered SIC is given as

$$R_{1,N}(\eta_1) = \eta_1 \log \left(1 + \left[\frac{\beta_{2,N}}{\beta_{1,N}} + \frac{1}{G(\eta_1)} \right]^{-1} \right)$$
 and (39)

$$R_{2,N}(\eta_1, \eta_2) = \eta_2 \log \left(1 + \beta_{2,N} H(\eta_1, \eta_2)\right).$$
 (40)

With these closed-form average rate expressions, we solve P1, and obtain the optimal NOMA design as follows.

Proposition 4. (Maximum average rate, NOMA) The maximum average rate of Link₂ under NOMA is provided as

$$R_{2,N}^*(\eta_1^*, \eta_2^*) = \eta_2 \log \left(1 + \beta_{2,N}^* H(\eta_1^*, \eta_2^*)\right)$$
 where (41)

$$\beta_{2,N}^* = \left(1 + \left[\frac{1}{\exp\left(\hat{R}_1/\eta_1^*\right) - 1} - \frac{1}{G(\eta_1^*)}\right]^{-1}\right), \quad (42)$$

$$\eta_1^* = 1 - \left(\frac{1}{\theta^2} - \frac{1}{(1 - \eta_2^*)^2}\right)^{-\frac{1}{2}},$$

$$\eta_2^* = \underset{\hat{\eta}_2 \le \eta_2 < \min(1, 1 - \theta\sqrt{2})}{\arg \max} R_{2,N}^*(\eta_1, \eta_2),$$
(43)

$$\eta_2^* = \underset{\hat{\eta}_2 \le \eta_2 < \min(1, 1 - \theta\sqrt{2})}{\arg \max} R_{2, N}^*(\eta_1, \eta_2), \tag{44}$$

 $\begin{array}{l} R_{2,\mathrm{N}}^*(\eta_1,\eta_2) = \eta_2 \log \left(1 + \beta_{2,\mathrm{N}}^*(\eta_2) H(\eta_1,\eta_2)\right)\!, \text{ and } \beta_{2,\mathrm{N}}^*(\eta_2)\\ \text{is obtained by replacing } \eta_2^* \text{ in } \beta_{2\mathrm{N}}^* \text{ with } \eta_2. \end{array}$ *Proof:* See Appendix-B.

Similar to Corollary 1 for OMA, we consider a sufficiently large $\hat{\eta}_2$ so that $R_{2,N}^*(\eta_1,\eta_2)$ in (44) monotonically decreases with η_2 , yielding the following closed-form result.

Corollary 2. (Closed-form maximum average rate, NOMA) For $\hat{\eta}_2 \rightarrow 1$, the objective function in (44) is a monotone decreasing function of η_2 , and we thus obtain $R_{2,N}^*(\eta_1^*,\eta_2^*) = R_{2,N}^*(\hat{\eta}_1,\hat{\eta}_2)$ where $\hat{\eta}_1 = 1 - [1/\theta^2 - 1/(1-\hat{\eta}_2)^2]^{-1/2}$.

As shown in Fig. 2 and discussed after Corollary 1, Corollary 2 also holds under $\hat{\eta}_2$ that is much smaller than 1. We thus consider this as the closed-form approximation, and defer its validation to Sect. IV.

IV. NUMERICAL EVALUATION

In this section, we numerically validate our analytic results on multiplexing visuo-haptic VR traffic. The default simulation parameters are: $\hat{\eta}_2 = 0.9$, $c = \alpha/(\alpha - 2)$, and $\alpha = 4$.

Fig. 3 renders the rate region of the maximized average rate $R_{2,i}^*$ of Link₂ and the target rate R_1 of Link₁, under the optimal OMA and NOMA designs. This validates that our closedform optimal rate expressions $R_{2,j}^*$'s in Corollaries 1 and 2 are only up to 1.75% less than the simulated values obtained by using Propositions 2 and 4 without the use of Lemma 1. Furthermore, in NOMA, we observe that the performance of the proposed reliability-ordered SIC is significantly degraded when the power-ordered SIC is enforced, calling for an SIC implementation that can decode the weaker signal first if it is more reliable.

Next, compared to OMA, the results in Figs. 3 shows that NOMA performs better for a higher \hat{R}_1 . In fact, NOMA is capable of achieving a higher rate, since its power-domain multiplexing logarithmically decreases the rate whereas the rate under OMA linearly decreases. In addition, as shown by comparing Figs. 3-a and b, NOMA is also preferable for a higher target visuo-haptic perceptual resolution, i.e. low θ . In this regime, due to the integrated JND relationship in (17), it leads to the higher decoding success probability η_1 that reduces the rate loss induced by SIC in NOMA, namely, the concatenated decoding at the reception of Link₂ signals.

Finally, the maximum required θ from which NOMA outperforms OMA for different \hat{R}_1 's is specified in Fig. 4. Following the same reasoning as for Fig. 3, a lower θ with NOMA results in a higher η_1 , thereby minimizing the rate loss induced by the SIC process. With OMA, on the contrary, the higher η_1 solely decreases the miniblock allocation to Link₂, reducing the rate of Link₂. For a larger \hat{R}_1 , the rate of Link₂ changes negligibly under the transmit power reduction of NOMA, yet deteriorates significantly under the linear-scale bandwidth reduction of OMA.

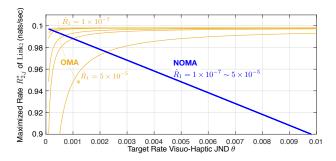


Fig. 4. Maximized average rates $R_{2,0}^*(\eta_2^*)$ and $R_{2,N}^*(\eta_1^*,\eta_2^*)$ w.r.t. a target visuo-haptic JND threshold θ for different target rates \hat{R}_1 's.

V. CONCLUSION

In this paper, we studied the multiplexing design for supporting visuo-haptic VR perceptions through downlink eMBB-URLLC links in cellular systems. Based on our closed-form average rate derivations under OMA and NOMA, we conclude that NOMA with the proposed reliability-ordered SIC outperforms OMA, for a higher target haptic data rate as well as for a higher target visuo-haptic perceptual resolution. A possible extension to this work is to optimize its uplink multiplexing design. Incorporating different types of multimodal perceptions could also be an interesting topic for further research.

ACKNOWLEDGEMENT

This paper has benefited from comments and suggestions by P. Popovski and O. Simeone.

APPENDIX - PROOFS OF PROPOSITIONS

A. Proof of Proposition 2

Now that $R_{i,O}(\eta_i)$ monotonically increases with $w_{i,j}$ while decreasing with $w_{\bar{i},j}$ for $\bar{i} \neq i$ with $\bar{i} \in \{1,2\}$, the optimal allocation $w_{1,0}^*$ for P1 is determined by the equality constraint (18b). By applying Proposition 1, (18b) reads

$$\eta_1 \log (1 + G(\eta_1)) = \hat{R}_1.$$
(45)

Solving this equation yields $w_{1,\mathrm{O}}^*$. Applying $w_{2,\mathrm{O}}^*=1-w_{1,\mathrm{O}}^*$ and (18c) to $R_{2,\mathrm{O}}(\eta_2)$ in Proposition 1 results in $R_{2,\mathrm{O}}^*(\eta_2^*)$. The feasible range of η_2 comes from (18c) and (18d).

B. Proof of Proposition 4

As $\beta_{1,N}$ grows, $R_{1,N}(\eta_1)$ monotonically increases, while $R_{2,N}(\eta_1,\eta_2)$ decreases. Therefore, the optimal allocation $\beta_{1,N}^*$ for P1 is achieved by the equality constraint (18b) that is given as

$$\eta_1 \log(1 + [\beta_{2,N}/\beta_{1,N} + 1/G(\eta_1)]^{-1}) = \hat{R}_1.$$
(46)

Solving this equation, we obtain $\beta_{1,N}^*$. Applying $\beta_{2,N}^* = 1 - \beta_{1,N}^*$ to (18c) to $R_{2,N}(\eta_1,\eta_2)$ in Proposition 1 results in $R_{2,N}^*(\eta_1^*,\eta_2^*)$. The feasible range of η_2 follows from (18c) and (18d).

REFERENCES

- ABI Research and Qualcomm, "Augmented and Virtual Reality: The First Wave of 5G Killer Apps," White Paper, Feb. 2017.
 E. Baştuğ, M. Bennis, M. Médard, and M. Debbah, "Towards Intercon-
- nected Virtual Reality: Opportunities, Challenges and Enablers," IEEE
- Commun. Mag., vol. 55, pp. 110–117, Jun. 2017.
 [3] J. Park, P. Popovski, and O. Simeone, "Minimizing latency to support vr social interactions over wireless cellular systems via bandwidth allocation," to appear in IEEE Wireless Commun. Lett. [Online]. Early access: http://ieeexplore.ieee.org/document/8332500.
- [4] Z. Shi, H. Zou, M. Rank, L. Chen, S. Hirche, and H. J. Müler, "Effects of Packet Loss and Latency on the Temporal Discrimination of Visual-Haptic Events," *IEEE Trans. Haptics*, vol. 3, pp. 28–36, Jan.-Mar. 2010. J. Park, D. Kim, P. Popovski, and S.-L. Kim, "Revisiting Frequency
- Reuse towards Supporting Ultra-Reliable Ubiquitous-Rate Communication," *Proc. IEEE WiOpt Wksp. SpaSWiN, Paris, France*, May 2017. ITU-R M.2083, IMT Vision Framework and Overall Objectives of the Future Development of IMT for 2020 and Beyond; Recommendation M.5/BL/22
- E. Steinbach, S. Hirche, M. Ernst, F. Brandi, R. Chaudhari, J. Kammerl, and I. Vittorias, "Haptic Communications," Proc. IEEE, vol. 100, pp. 937-956, Apr. 2012
- [8] Q. Zhang, J. Liu, and G. Zhao, "Towards 5G Enabled Tactile Robotic Telesurgery," [Online]. ArXiv preprint: https://arxiv.org/abs/1803.03586.
 [9] P. Popovski, J. J. Nielsen, C. Stefanovic, E. de Carvalho, E. G. Ström,
- K. F. Trillingsgaard, A. Bana, D. Kim, R. Kotaba, J. Park, and R. B. Sørensen, "Wireless Access for Ultra-Reliable Low-Latency Communication (URLLC): Principles and Building Blocks," IEEE Netw., vol. 32, pp. 16-23, Mar. 2018.
- M. Bennis, M. Debbah, and V. Poor, "Ultra-Reliable and Low-Latency Wireless Communication: Tail, Risk and Scale," [Online]. ArXiv preprint: https://arxiv.org/abs/1801.01270.
- [11] M. O. Ernst and M. S. Banks, "Humans Integrate Visual and Haptic Information in a Statistically Optimal Fashion," Nature, vol. 415, pp. 429-433, Jan. 2002
- [12] M. Rank, Z. Shi, H. J. Müller, and S. Hirche, "Predictive Communication Quality Control in Haptic Teleoperation With Time Delay and Packet Loss," IEEE Trans. Hum. Mach. Syst., vol. 46, pp. 581–592, Aug. 2016.
- 3GPP TR 36.859, Study on Downlink Multiuser Superposition Transmission (MUST) for LTE.
- A. AL-Shuwaili and O. Simeone, "Energy-Efficient Resource Allocation for Mobile Edge Computing-Based Augmented Reality Applications,' IEEE Wireless Commun. Lett., vol. 6, pp. 398-401, Apr. 2017.
- M. Chen, W. Saad, and C. Yin, "Resource Management for Wireless Virtual Reality: Machine Learning Meets Multi-Attribute Utility," Proc. IEEE GLOBÉCOM, Singapore, Dec. 2017.
- M. S. Elbamby, C. Perfecto, M. Bennis, and K. Doppler, "Toward Low-Latency and Ultra-Reliable Virtual Reality," *IEEE Netw.*, vol. 32, pp. 78– 84. Mar. 2018.
- [17] P. Popovski, K. F. Trillingsgaard, O. Simeone, and G. Durisi, "5G Wireless Network Slicing for eMBB, URLLC, and mMTC:
 A Communication-Theoretic View," [Online]. ArXiv preprint: A Communication-Theoretic https://arxiv.org/abs/1804.05057.
- J. G. Andrews, F. Baccelli, and R. K. Ganti, "A Tractable Approach to Coverage and Rate in Cellular Networks," IEEE Trans. Commun., vol. 59, pp. 3122-3134, Nov. 2011.
- M. Haenggi, Stochastic Geometry for Wireless Networks. Cambridge
- J. Park, S.-L. Kim, and J. Zander, "Tractable Resource Management with Uplink Decoupled Millimeter-Wave Overlay in Ultra-Dense Cellular Networks," IEEE Trans. Wireless Commun., vol. 15, pp. 4362-4379, Jun. 2016.
- [21] D. N. C. Tse and P. Viswanath, Fundamentals of Wireless Communications. Cambridge University Press, 2005.
- A. S. Avestimehr and D. N. C. Tse, "Outage Capacity of the Fading Relay Channel in the Low-SNR Regime," *IEEE Trans. Inf. Theory*, vol. 53, pp. 1401–1415, Apr. 2007.
- W. Yang, G. Durisi, T. Koch, and Y. Polyanski, "Quasi-Static Multiple-Antenna Fading Channels at Finite Blocklength," IEEE Trans. Inf. Theory, vol. 60, pp. 4232–4265, Jul. 2014.

 J. Blomer and N. Jindal, "Transmission Capacity of Wireless Ad Hoc
- Networks: Successive Interference Cancellation vs. Joint Detection," Proc. IEEE ICC, Dresden, Germany, Jun. 2009.
- X. Zhang and M. Haenggi, "A Stochastic Geometry Analysis of Inter-Cell Interference Coordination and Intra-Cell Diversity," 'IEEE Trans.
- Wireless Commun., vol. 13, pp. 6655–6669, Dec. 2014.

 J. Park and P. Popovski, "Coverage and Rate of Downlink Sequence Transmissions with Reliability Guarantees," IEEE Wireless Commun. Lett., vol. 6, pp. 722-725, Aug. 2017.