Boyd, Christopher; Vehkalahti, Roope; Tirkkonen, Olav

Grant-Free Access in URLLC with Combinatorial Codes and Interference Cancellation

Published in:
2018 IEEE Globecom Workshops, GC Wkshps 2018 - Proceedings

DOI:
10.1109/GLOCOMW.2018.8644207

Published: 01/01/2018

Document Version
Peer-reviewed accepted author manuscript, also known as Final accepted manuscript or Post-print

Please cite the original version:
Boyd, C., Vehkalahti, R., & Tirkkonen, O. (2018). Grant-Free Access in URLLC with Combinatorial Codes and Interference Cancellation. In *2018 IEEE Globecom Workshops, GC Wkshps 2018 - Proceedings* Article 8644207 IEEE. https://doi.org/10.1109/GLOCOMW.2018.8644207

Grant-Free Access in URLLC with Combinatorial Codes and Interference Cancellation

Christopher Boyd, Roope Vehkalahti and Olav Tirkkonen
School of Electrical Engineering, Aalto University, Finland
Email: {christopher.boyd,roope.vehkalahti,olav.tirkkonen}@aalto.fi

***Abstract*—When targeting ultra-reliability for randomly accessing uplink users, an irreducible error floor exists when employing random repetition coding in synchronised frames. Combinatorial Code Designs (CCDs) that produce deterministic access patterns have been shown to overcome this performance limit in MAC layer channel models, where the physical layer has been abstracted to a collision model. In this work, we demonstrate that CCDs remain a promising option for URLLC for a properly modelled physical layer with fading and collision recovery through diversity reception. We investigate uplink users arriving according to a Poisson process, attempting to directly transmit data packets in a grant-free manner. We measure the achievable rate using a system outage capacity formulation which takes into account the access intensity, the size of the data packets transmitted, and the packet error rate, with the packet loss rate constrained to be less than a given URLLC target. We compare random and deterministic access patterns by system simulation in a factory environment. The results show that considerable gains in system throughput can be achieved by using deterministic access patterns. Unlike random patterns, patterns based on CCDs have the potential to support considerable communication rates with very high success probability. With ideal channel estimation, CCDs enable an aggregate system spectral efficiency of 2 bps/Hz in a grant-free URLLC access scenario.**

I. INTRODUCTION

Ultra-Reliable Low-Latency Communications (URLLC), characterised by low end-to-end latencies and high reliability, is one of the core novelties in 5G systems [1]–[3]. Novel mission-critical use cases, such as vehicle-to-vehicle communication, smart grid control, and factory automation are considered under this framework. A target envisioned for URLLC is 99.999% reliability with latencies less than 1ms [4]. This presents significant challenges at multiple levels, requiring new approaches to wireless system design.

URLLC challenges have been addressed from the perspective of cooperative communication [5] and routing diversity [6], as well as Hybrid ARQ [7] and control channel redesign [8]. Most of the URLLC literature assumes grant-based, i.e. scheduled, access. This involves either reserving resources semi-persistently to the users, or very reliable hand-shake protocols involving multiple round trips over the air interface within the latency window [8]. Contention-based or grant-free access is complementary to scheduled access.

In the context of URLLC, contention-based access becomes particularly challenging. Allowing uplink users to contend over channel resources seems in contradiction with the principles of URLLC; the inherent insecurity of random access seems severely counter-productive. Scheduled access is clearly beneficial from the perspective of reliability, as users can be orthogonalised to limit intra-cellular interference. Nevertheless, in [9] it was shown that by transmitting replicas of packets, in the form of multichannel slotted ALOHA, and by applying multi-user detection (MUD), URLLC targets can be met for sporadically transmitting users even without reserved channels. In [9], there was no coordination between the packet repetitions of different users. The observed reduction in collision probability holds with small access intensity, but when intensity grows, three- and multi-way collisions start to dominate performance. Methods to control such multi-way collisions have been addressed in [10], where Combinatorial Code Designs (CCDs) were discussed. In comparison to random or regular repetitions, discussed in [9], CCDs offer repetition coding that is coordinated over the whole user population, with interference cancellation (IC) and MUD applied at the MAC layer.

In this paper, we investigate grant-free access in URLLC for a realistic system simulation model. We extend the MAC layer analysis of [10], where collision channels were considered, to multi-point and multi-antenna fading channels in a multi-cell communication network. The communicating devices are synchronised to the network, and apply open-loop power control. Accordingly, the transmissions are subject to fast fading of the wireless channels, in addition to random interference from potential simultaneously accessing users. Receive diversity combining is applied on the physical layer to combat fading, and a CCD access code is employed to mitigate multi-user collisions. We compare the deterministic CCD approach with methods where packets replicas are transmitted in randomly chosen slots in a system simulation setting modeling a factory environment. We observe that for achieving URLLC targets in grant-free random access, access codes based on CCDs are superior to random repetition coding. For 4-antenna base stations (BS), CCDs achieve ultra-reliability with 5 times the access intensity when compared to the random approach.

II. RANDOM ACCESS FRAMEWORK

We consider random access in an network with N_{bs} access points (AP), and N machine-type user equipments (UE) in a closed environment, such as a factory. The UE population is fixed and under control of the network management. Accordingly, the number of UEs is known. Each AP is equipped with M_{r} antennas. At each instance, a random subset of N_{act} UEs is active, each of which transmit an uplink packet. The objective is to meet URLLC targets: packets received within a delay window of, e.g., 1 ms, and with a maximum packet loss rate of $\leq 10^{-5}$.

To limit latency, grant-free access is considered. In grant-free access, UEs sporadically transmit data on the RACH without first sending a scheduling request and obtaining a resource allocation from the AP. Traditional means of collision control on the RACH, including feedback and retransmissions, are eschewed in favour of a connectionless one-shot transmission approach. From the perspective of delay, overhead and power consumption, grant-free access is eminently preferable. The increased probability of inter-user interference however necessitates different means of collision avoidance and resolution in order to meet reliability targets.

In grant-free access, the UEs randomly access the available radio resources in an uncoordinated manner. To make low latency transmissions possible, the resources within a RACH frame of duration $\leq$ the latency target, have been divided into n access slots. The slots may be divided both in time and frequency to allow grant-free transmission of sizable packets. Here, we assume that the UEs do not know the precise channel quality. All transmission packets of all users are such that the same transmission rate R is attempted. A low user activity is considered, in that the expected number of UEs active in a frame is $N_{\mathrm{act}} << N$. Accesses follow an independent Poisson process, for which the expected number of access events per frame is denoted by intensity λ.

We assume that each UE is carrier synchronised and aware of a timing advance before transmission. The user transmissions may then be received at the APs in slot and frame synchrony, subject to conventional tolerances in OFDM networks, such as the Cyclic Prefixes (CPs). Moreover, we assume that the network coverage area is small enough (as compared to CP length) such that transmissions from all users are synchronous within the CP at all APs. Furthermore, we assume transmission power control with full compensation of average path loss, as seen at the AP with the strongest channel to the user.

Repetition coding of access packets is considered in order to provide diversity and robustness to inter-user collisions in the MAC layer, a primary source of packet loss in contention-based access. In the vein of slotted ALOHA and its evolutions, UEs transmit $k < n$ packet replicas spread across the frame according to a particular pattern. Such access patterns may be generated randomly by UEs on each access attempt, or pre-assigned from an access code. In the latter case, it is assumed that users may be allocated unique patterns from the deterministic code prior to system operation. Such a code can be understood as a collection of N length n binary vectors, where 1's indicate slots containing an access packet replica. It is further assumed that each access packet contains pointers to the locations of its replicas, enabling more robust MAC layer collision recovery and decoding.

Contemporary random access protocols such as contention resolution diversity slotted ALOHA (CRDSA) leverage successive interference cancellation (SIC) to facilitate collision recovery to improve the packet loss rate when decoding MAC frames [11]. The iterative decoding process involves the receiver parsing the n slots, decoding packets free from interference, extracting the locations of all replicas, and removing them from the frame. Upon each pass, new interference free packets may be revealed, decoded and removed in the same manner. This process continues until either all active users have been successfully received, or the remaining superposition of access patterns proves unresolvable (a stopping set) and all undecoded packets are lost (in the absence of advanced physical layer techniques, e.g. MUD).

Access protocols that employ random repetition coding and SIC decoding are subject to an error floor which limits their usability when targeting ultra-reliable communications. Recently, it has been shown that pre-distributing access patterns from combinatorial code designs to the UEs avoids this error floor, at a cost to the supportable user population [10], [12]. Here we consider a length $n = 26$, weight 5 deterministic code of size $N = 260$ based on the $S(3, 5, 26)$ Steiner system [13]. This code is ideal for the ultra-reliable grant-free access case study here, so we will consider frames of $n = 26$ times slots and a user population limited to $N = 260$ UEs.

III. PHYSICAL AND MAC LAYER MODELLING

To obtain a more complete picture of the impact of deterministic access coding on meeting URLLC reliability targets, we model both the physical layer and MAC layer operation of the network.

A. PHY Layer

We number the N UEs in the coverage area of the system by indices $i \in \{1, \dots, N\}$, and the APs by $s \in \{1, \dots, N_{\mathrm{bs}}\}$. We assume that each of the UEs have a single antenna and each of the APs have M_{r}. Communication happens inside of frames of n time slots, with each slot consisting of a modulation symbols. For simplicity, we suppress these symbols for this discussion. The procedure below is repeated a times in each frame.

If we assume that the ith UE is transmitting during the jth time slot, the sth AP will receive the signal $\mathbf{h}_{s,i} X_{i,j}$ from UE i, where $X_{i,j}$ corresponds to a length a row vector consisting of the information symbols transmitted by user i in slot j. The $M_{\mathrm{r}} \times 1$ channel between the UE and the AP is

$$\mathbf{h}_{s,i} = \begin{pmatrix} h_{1,s,i} \\ h_{2,s,i} \\ \vdots \\ h_{M_{\mathrm{r}},s,i} \end{pmatrix},$$

where the coefficients $h_{v,s,i}$ represent the fading the between vth antenna of the sth AP and ith user.

The random set of N_{act} users active in this frame is indexed by the set $S \subset \{1, \ldots, N\}$. Accordingly, in the jth time slot, the total received signal of AP s is

$$\mathbf{y}_{s,j} = \sum_{i \in S} \mathbf{h}_{s,i} X_{i,j} + N_t \ , \tag{1}$$

where N_t is a i.i.d noise vector. We assume that in the n access slots there are sufficient resources for pilot transmissions, such that the channels between the UEs and the APs can be reliably estimated. For simplicity, we shall assume perfect estimation, and that the channel is constant during a frame. Imperfect estimation would become a further impediment to be controlled in order to achieve URLLC targets. To simplify the discussion, we also assume perfect activity detection. This is not a necessity, as activity detection can be handled in a similar way as grant-free reception.

For each UE, there is a serving AP which has the strongest received signal from the UE. We assume that the channels to each active UE are estimated at each AP, but decoding is performed at the serving AP. For each detected active user i in a slot, the serving AP applies a matched filter receiver, matched to that user. The filter output then is

$$z_{i,j} = \mathbf{h}_{s,i}^{\mathrm{H}} \mathbf{y}_{s,j} = \|\mathbf{h}_{s,i}\|^2 X_{i,j} + \sum_{k \in S, k \neq i} \mathbf{h}_{s,i}^{\mathrm{H}} \mathbf{h}_{s,k} X_{k,j} \ + \ \tilde{N}_t \tag{2}$$

where we see the coherently combined wanted signal, the non-coherently combined interfering signal, and the processed noise $\tilde{N}_t$. After diversity combining, the Signal-to-Interference-plus-Noise Ratio (SINR) for receiving the signal from i at s in slot j is

$$\gamma_{i,j} = \frac{\|\mathbf{h}_{s,i}\|^4}{M_{\mathrm{r}} \sigma^2 + \sum_{k \in S, k \neq i} \left|\mathbf{h}_{s,i}^{\mathrm{H}} \mathbf{h}_{s,k}\right|^2} \ , \tag{3}$$

where σ^2 represents the power of the additive noise. For a given active user i, the SINR differs from slot to slot. In the slots where the user is not transmitting, $\gamma_{i,j} = 0$. In the slots where the user is transmitting, the SINR depends on which set of other users transmits in the slot.

For simplicity, we characterise the physical layer by a transmission efficiency function f, which maps a rate R and a SINR γ to a binary decoding "success" or "error" variable. The underlying idea is that information is gathered when attempting to decode a user with SINR γ_{ij} in the a symbols in the slot. With a being relatively large, the average SINR is a good indication of the reliability of the decoding for a transmission with a given spectral efficiency [14]. The modulation and the channel code acting on the bits in the a symbols is thus abstracted by f. In reality, such a mapping would be to a probability distribution. However, here we assume a sufficient margin such that a packet is correctly decoded when γ is greater than some threshold θ with very high probability. The rate of transmission of the users is modelled by θ. This simplification has the consequence that the PHY model does not contribute to the unreliability. We use Shannon's law with an implementation margin of $\mu = 2$ dB to represent this efficiency function, as in [15]. This margin is subtracted from the threshold when computing rate, i.e.

$$R = \log_2(1 + \theta/\mu) \ , \tag{4}$$

and the transmission spectral efficiency is chosen accordingly. Packet decoding per slot per active user is therefore modelled such that $P_{\text{error}} = P(\gamma_{i,j} < \theta)$. Accordingly, for each of the N_{act} active users, physical layer processing outputs a decoding success or failure in each of the n slots. The transmission efficiency function is

$$f(\gamma) = \frac{1}{n} \log_2(1 + \gamma/\mu) \ , \tag{5}$$

where n is the frame length.

B. MAC Layer

The N_{bs} APs are assumed to be connected, such that all information obtained at the APs is available for MAC layer processing. Consider N_{act} UEs transmitting during a frame. The transmission of a single active user i is received at all APs using a corresponding filter. A centralised decoder receives all N_{bs} frames, where each slot is as in Equation (2). The decoder parses each frame, and if a slot corresponding to the ith user contains a packet whose received SINR with respect to all interferers is larger than a predetermined threshold θ, user i is considered to have been received correctly. After parsing all available frames, the contributions of the successfully received users are removed and the process iterates until either none of the remaining users have a packet with SINR large enough, or until all users have been received. Packets from undecodable users are lost. The assumptions of AP interconnectedness and the availability of perfect channel state information are necessary to enable this multi-cellular IC.

IV. ULTRA-RELIABLE RANDOM ACCESS IN A FACTORY ENVIRONMENT

We consider random access in a network topology consisting of 4 wireless access points and $N = 260$ machine-type UEs, in a small factory environment. The APs are affixed to the ceiling of the factory, serving four equally-sized cells. The UEs move freely within the limits of the factory and are typically not in LOS of the APs. Each node is equipped with $M_r = 4$ antennas. The network operates with reuse factor 1, i.e. UEs transmitting simultaneously map potentially cause inter-user interference at all receivers. At each instance of time, the UEs are uniformly distributed inside the factory and select their cell according to the best channel. Uplink transmit power is controlled in an open loop manner based on the average received signal from the strongest access point.

We are interested in the expected aggregate amount of reliably communicated data in the grant-free access system considered. This depends on the access intensity λ, a transmission data rate R (per resource reserved for the access protocol), and a probability of packet loss P_{error}. Here, we estimate the data rate in terms of the transmission efficiency

TABLE I
NETWORK PARAMETERS [16]

Factory Size (w x l x h)	50 x 100 x 5m
Number of APs	4
Antenna Configuration ($M_{\text{UE}} \times M_{\text{AP}}$)	1 x 4
Distribution of UEs	Uniform
Carrier Frequency	1.3 GHz
Bandwidth	20 Mhz
Bandwidth Efficiency	90%
Reuse Factor	1
UE Transmit Power	20 dBm
UE Antenna Gain	0 dB
Thermal Noise Power	-174 dBm/Hz
Noise Figure	9 dB
Path Loss Exponent (NLOS, Heavy Clutter)	1.69
Shadowing s.d. (NLOS, Heavy Clutter)	6.62 dB

function $f(\gamma)$, and the SINR threshold θ required for reliable communication. The aggregate reliable communication rate is

$$C(\lambda, \theta, P_{\text{error}}) = \lambda(1 - P_{\text{error}} f(\theta)) \ , \tag{6}$$

measured in units of bps/Hz. Maximizing C over λ and θ, and constraining the reliability by $P_{\text{error}} \leq P_{\text{target}}$, we have

$$C(P_{\text{target}}) = \max_{\lambda,\theta;\ P_{\text{error}} \leq P_{\text{target}})} C(\lambda, \theta, P_{\text{error}}) \ , \tag{7}$$

which we will refer to as the *outage capacity* of the grant-free access system.

To obtain the outage capacity for a given reliability target, we simulate a packet arrival scenario in a multi-cell network based on the parameters in Table I. The 260 UEs are uniformly distributed inside the factory. In a given access frame, a subset of users, of size determined by a Poisson arrival process with intensity λ, are active. Each active UE transmits its packet replicas according to either a preallocated pattern from an access code, or randomly. We generate the SINR experienced by every active user in each slot, at every access point, and determine the respective decoding success by comparing to θ. Successfully decoded users are stripped from the frame in an iterative IC process at the MAC layer. The packet decoding error rate is obtained over 50 million instances.

Figure 1 depicts the outage probability as a function of the access intensity for a number of SINR thresholds θ. The decoding performance of random and deterministic access patterns are compared. When the SINR threshold θ for successful decoding is 20 dB or smaller, deterministic patterns provide a considerable advantage over random patterns. For example, when $\theta = 5$ dB, there is a dramatic difference in the performance of random and deterministic access patterns when the access intensity is less than 20.

One of the penalties for random patterns appears in the scenario where two active users happen to choose the same access patterns and be relatively close to each other. In such situation, both of these users may be served by the same access point. Then, all packets will collide and their received signal strengths in all slots will be roughly equal. The SINR will not be sufficient for correct decoding, and an outage will result. The probability of such collisions grows with the number of accessing users.

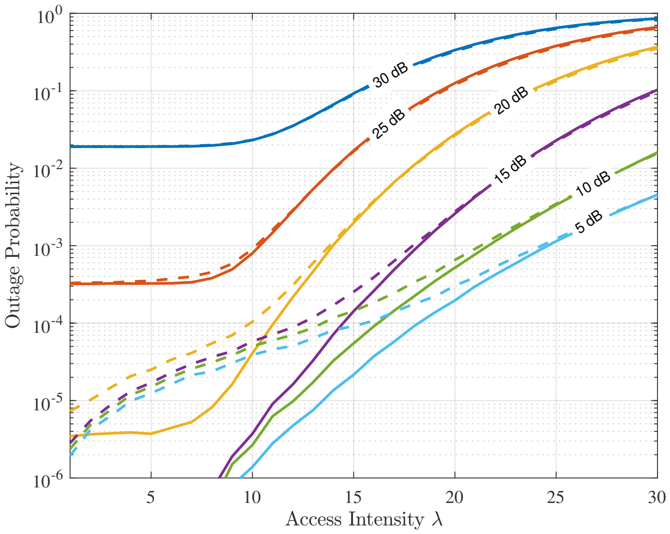

Fig. 1. Outage probability for deterministic (solid) and random (dashed) access codes, as a function of frame access intensity λ for various SINR thresholds θ, in RACH frames of $n = 26$ slots.

On the other hand, the deterministic patterns in use guarantee that up to three simultaneously active users may always be decoded, even when they are served by the same AP. We note that with a 5 dB SINR threshold, deterministic patterns can support an average of 13 active users with outage probability less than 10^{-5}, while random patterns can support only 4 users. When the access intensity is higher than 25, the difference in performance between random and deterministic access patterns diminishes. As the frame size is only 26, it is to be expected that the channel becomes clogged and coordination of access patterns become irrelevant. We observe a similar behaviour in the curves up to a SINR threshold of 20dB. However, for the larger thresholds of 25dB and 30dB, the performance of random and deterministic patterns are almost equal, and both curves suffer an error floor. When the expected number of active users tends to zero, the limiting factor is no longer interference, but noise. When the intensity is less than 7, the probability of deep fading for users without collisions in a slot dominates performance, despite the fourth order diversity reception applied.

Figure 2 details the corresponding capacity curves. The average throughput for a given intensity is considered, without regard for the reliability target. In this setting, random and deterministic repetition coding have very little differences, except that, for very high intensities, the random access patterns perform slightly better.

In Figure 3, the highest possible supportable rate (over all thresholds) both for the random and deterministic approaches is shown for various outage targets. While for outage probabilities greater than 5×10^{-3} there is little difference between the two, it can be observed that deterministic access

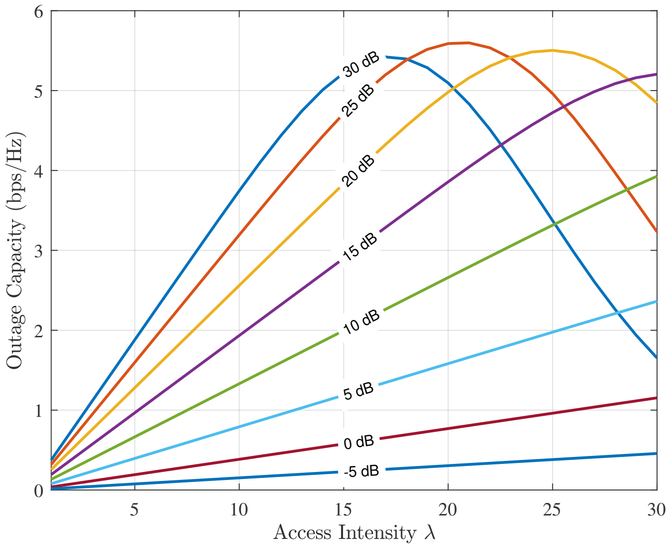

Fig. 2. Outage capacity for deterministic access code as a function of frame access intensity λ for various SINR thresholds θ, in RACH frames of $n = 26$ slots.

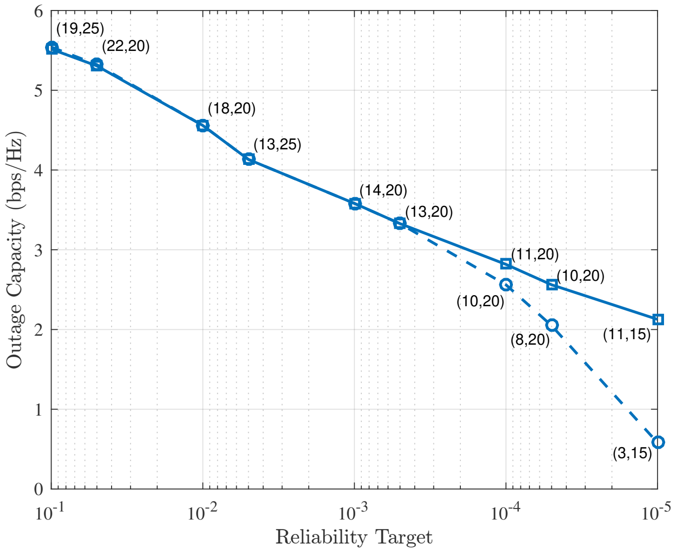

Fig. 3. Outage capacity for deterministic (solid) and random (dashed) access codes maximised by intensity threshold pair (λ, θ dB) for various reliability targets.

codes provide considerably higher rates in the ultra-reliability region.

V. CONCLUSION

We considered Combinatorial Codes Designs for URLLC in a factory environment. We observed that the gains achieved in a collision channel model were amplified when we consider a more realistic communication scenario. Moving from random to deterministic access patterns allowed up to three times more users to simultaneously access the uplink resources, while still guaranteeing data transmission to five nines of reliability. This suggests that the gains from employing deterministic access codes based on CCDs will likely be visible in scenarios where more sources of potential error are taken into account.

We conclude with comments on the simplifying assumptions made during this work, and how they can possibly be overcome. The first assumption was that packets transmitted inside a given slot can be perfectly recovered. This condition can be replaced with a guarantee on correct decoding with very high probability, which makes it possible to consider finite length blocks. Another problematic assumption was the availability of perfect channel state information (at transmitter/at receiver). This is not only crucial from the point of view of achieving high rates, but also for enabling SIC. This assumption necessitates significant training, which will further limit the achievable rates. However, at least in scenarios where fading is static within a frame, we believe that the superiority of deterministic over random access pattern, will hold under more realistic conditions as well.

ACKNOWLEDGEMENTS

This work was funded in part by the Academy of Finland (grant 299916) and EIT ICT (HII:ACTIVE).

REFERENCES

[1] A. Osseiran, J. Monserrat, and P. Marsch, eds., *5G Mobile and Wireless Communications Technology*. Cambridge University Press, 2016.

[2] G. Durisi, T. Koch, and P. Popovski, "Toward Massive, Ultrareliable, and Low-Latency Wireless Communication With Short Packets," in *Proc. IEEE*, vol. 104, no. 9, 2016, pp. 1711–1726.

[3] P. Popovski *et al.*, "Wireless Access for Ultra-Reliable Low-Latency Communication: Principles and Building Blocks," *IEEE Network*, vol. 32, no. 2, pp. 16–23, 2018.

[4] 3GPP, "Study on Scenarios and Requirements for Next Generation Access Technologies," Tech. Rep., Jun. 2017.

[5] V. N. Swamy *et al.*, "Cooperative Communication for High-Reliability Low-Latency Wireless Control," in *Proc. Int. Conf. on Comm.*, Jun. 2015, pp. 4380–4386.

[6] B. Singh, O. Tirkkonen, Z. Li, M. Uusitalo, and R. Wichman, "Selective Multi-Hop Relaying for Ultra-Reliable Communication in a Factory Environment," in *Proc. IEEE Int. Symp. on Pers., Ind. and Mobile Radio Comm.*, Sep. 2016, pp. 1–6.

[7] K. I. Pedersen, S. R. Khosravirad, G. Berardinelli, and F. Frederiksen, "Rethink Hybrid Automatic Repeat reQuest Design for 5G: Five Configurable Enhancements," *IEEE Wireless Comm.*, vol. 24, no. 6, pp. 154–160, 2017.

[8] H. Shariatmadari *et al.*, "Fifth-Generation Control Channel Design: Achieving Ultrareliable Low-Latency Communications," *IEEE Veh. Tech. Mag.*, vol. 13, no. 2, pp. 84–93, 2018.

[9] B. Singh, O. Tirkkonen, Z. Li, and M. Uusitalo, "Contention-Based Access for Ultra-Reliable Low Latency Uplink Transmissions," *IEEE Wireless Comm. Letters,*, vol. 7, no. 2, pp. 182–185, 2018.

[10] C. Boyd, R. Vehkalahti, and O. Tirkkonen, "Interference Cancelling Codes for Ultra-Reliable Random Access," *Int. J. Wireless Inf. Net.*, 2018. [Online]. Available: https://doi.org/10.1007/s10776-018-0411-6

[11] E. Casini, R. D. Gaudenzi, and O. del Rio Herrero, "Contention Resolution Diversity Slotted ALOHA (CRDSA): An Enhanced Random Access Scheme for Satellite Access Packet Networks," *IEEE Trans. Wireless Comm.*, vol. 6, no. 4, pp. 1408–1419, Apr. 2007.

[12] E. Paolini, G. Liva, and A. G. i Amat, "A Structured Irregular Repetition Slotted ALOHA Scheme with Low Error Floors," in *Proc. Int. Conf. on Comm.*, 2017, pp. 1–6.

[13] La Jolla Covering Repository. Steiner Systems. [Online]. Available: https://www.ccrwest.org/cover/steiner.html

[14] W. Yang, G. Durisi, T. Koch, and Y. Polyanskiy, "Quasi-Static Multiple-Antenna Fading Channels at Finite Blocklength," *IEEE Trans. Inf. Theory*, vol. 60, no. 7, pp. 4232–4265, 2014.

[15] P. Mogensen *et al.*, "LTE Capacity Compared to the Shannon Bound," in *Proc. IEEE Veh. Tech. Conf.*, 2007, pp. 1234–1238.

[16] T. S. Rappaport and C. D. McGillem, "Characterising the UHF Factory Radio Channel," *Electron. Lett.*, vol. 23, no. 19, pp. 1015–1017, 1987.