Ultra-Reliable Wireless Communications Using a D2D Protocol for Industrial Automation

SUKANTBEHERA, M.Tech,
Associate Professor, Department of Electronics Communication Engineering
Gandhi Institute For Technology,Bhubeneswar.

Abstract—

As an indispensable use case for the 5G wireless systems on the roadmap, ultra-reliable and low latency communications (URLLC) is a crucial requirement for the coming era of wireless industrial automation. The key performance indicators for URLLC stand in sharp contrast to the requirements of enhanced mobile broadband (eMBB): lowlatency and ultra-reliability are paramount but high data rates are often not required. This paper aims to develop communication techniques for making a paradigm shift from the conventional human-type broadband communications to the emerging machine-type URLLC. One fundamental task for URLLC is to deliver short commands from a controller to a group of actuators within the stringent delay requirement and with high-reliability. Motivated by the factory automation setting in which tasks are assigned to groups of devices that work in close proximity to each other thus can form clusters of reliable device-to-device (D2D) networks, this paper proposes a novel two-phase transmission protocol for achieving URLLC. In the first phase, within the latency requirement, th e multi-antenna base station (BS) combines the messages of al l devices within each group together and multicasts them to the corresponding groups; messages for different groups ar e spatially multiplexed. In the second phase, the devices tha t have decoded the messages successfully, herein defined as th e leaders, help relay the messages to the other devices in thei r groups. Under this protocol, we design an innovative leader selection based beamforming strategy at the BS by utilizing sparse optimization technique. The proposed strategy leads to a desired sparsity pattern in user activity with at least one leader being able to decode its message in each group in the first phase, thus ensuring full utilization of the reliability enhancing D2D transmissions in the second phase. Simulation results are provided to show that the proposed two-phase transmission protocol considerably improves the reliability of the entire system within the stringent latency requirement as compare d to existing schemes for URLLC. Index Terms—Ultra-reliable and low latency communications (URLLC), 5G, industrial automation, device-to-device (D2D) communications, machine-type communications (MTC) , multicasting, beamforming, sparse optimization.

I. INTRODUCTION

Enhanced mobile broadband (eMBB), massive machine type communications (mMTC), and ultra-reliable and low Manuscript received October 14, 2017, revised January 28, 2018 and March 6, 2018, accepted April 26, 2018.

of Canada. The authors are with The Edward S. Rogers Sr. Department of Electrical and Computer Engineering, University of Toronto, Toronto, Ontario, Canada, M5S3G4, (e-mails:lianguot.liu@utoronto.ca; weiyu@comm.utoronto.ca).
latency communications (URLLC) are the three main use cases that the 5G technology must support [1]. Addressing the above requirements in 5G calls for new methods and ideas at both the component and architectural levels, including massive multiple-input multiple-output (MIMO) [2], [3], millimeter wave (mmWave) communications [4], and cloud radio access network (C-RAN) [5] for eMBB, as well as multiple access schemes to support a massive number of devices for mMTC [6]–[8]. This paper focuses on URLLC. Specifically, we aim to tackle the latency and reliability requirement motivated by industrial automation applications [9]–[13]. In a typical closed loop industrial control scenario, group s of sensors and actuators are deployed in a fixed area in a factory setting. Periodically or when triggered by externa l events, the sensors send their measurements to the central controller, which then makes decisions and sends commands to the actuators for action. Under the current technology, sensors and actuators are typically connected to the centra l controller via a wired configuration in most factories. In th e near future, under the fourth industrial revolution roadma p (known as Industry 4.0), the communication networks in the factory setting are expected to migrate from wired to wireless for the purpose of increasing the flexibility in moving machinery and also for reducing the infrastructure expenditure [9]–[13]. As factory automation systems are highly sensitive to signal delays or distortions, such a transition will impose challenging requirements in terms o f latency as well as reliability for the wireless technologies. In the current 4G cellular

network, the end-to-end latency (which includes data transmission, packet retransmission , signal processing, protocol handling, and switching and network delays) can be in the order of 30-40ms, with the physical-layer latency accounting for about 15-20ms. For mission-critical applications for industrial automation, the latency requirement of 5G physical layer is expected to be pushed down to less than 1ms, an order of magnitude shorter than 4G. Further, such low latency requirement needs to be satisfied with ultra-reliability, e.g., 99 .999% or higher. This paper aims to address the challenge of wireless factory automation by focusing on the downlink URLLC in one cell (factory) of a cellular system, where the multiantenna base station (BS) (the central controller) needs to send a small amount of information bits (command) to each user (actuator) within the latency requirement (1ms). The core question this paper tries to answer is how to achiev
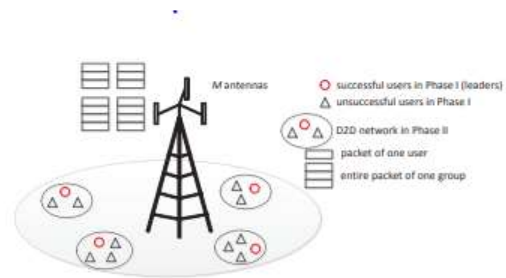

Fig. 1. Illustration of the two-phase transmission protocol: in the first phase, BS combines each group's messages together and multicasts them to the leaders in the corresponding groups; in the second phase, each leader helps relay the messages to the unsuccessful users in its group via the D2D network. the above goal with ultra-reliability in the sense that all the users can decode their messages with a very high probability. Achieving URLLC with the conventional broadcasting strategy is difficult due to the fact that a typical factory may have hundreds (or

even thousands) of actuators. Specifically, in a massive connectivity scenario where the number of devices is larger than the number of antennas at the BS, it can be difficult to transmit at an appreciable data rate to each user reliably, especially for the cell-edge users that suffer from strong inter-cell interference. Moreover, timedivision multiple-access (TDMA) may not be a feasible strategy because each user would only be allocated a fraction of the total transmission time, then the required signal-tonoise ratio (SNR) to achieve the target rate would be very high. This paper proposes a novel transmission strategy for URLLC that relies on a key observation that in practice, devices in a factory setting, e.g., robots or 3D printers, typically work in close proximity to each other and thus can potentially form a device-to-device (D2D) network for peer-to-peer communications. It is envisioned that the communication within each D2D network is significantly more reliable than that from the BS to the users due to the much stronger channels between the users in the same group. To exploit the reliable D2D networks, this paper proposes a novel D2D-based two-phase transmission protocol as shown in Fig. 1, in which the BS sends the messages to the users in the first phase, while in the second phase, the users who have already decoded the messages successfully (defined as the leaders of the groups) help relay the information to the other users in the same groups who have failed to receive their messages previously. Note that the reliability of the overall system is limited by the reliability of the cell-edge users. Our proposed protocol can opportunistically activate the cell-edge users who happen to not suffer from strong inter-cell interference due to the channel fading and let these leaders help the other cell-edge users with low signal-to-interference-plus-noise ratio (SINR) to achieve high reliability in the second phase. To enable the proposed relay strategy to work, the leaders

of each group need to receive the entire messages for all the users in their group. This paper devises a multi-group multicasting technique in the first phase [14], in which the user messages in each group are combined together as a single message and multicast to the leaders in the corresponding group, while messages for different groups are spatially multiplexed. Such a message combination strategy typically results in a manageable multicasting rate, since in most URLLC scenarios each device only requires a very small number of information bits within the latency requirement such that the total rate over all users in the same group is still reasonably small. Moreover, a substantially smaller number of users need to be activated in Phase I as compared to the broadcasting scenario since one group only needs one leader. Since the users in the same group usually belong to the same factory, the incentive mechanism and security, which are challenging issues in practical D2D networks [15], are no longer the main considerations in our investigated setup. Instead, leader selection in the first phase becomes the deciding factor of our protocol, since the groups without leaders cannot utilize the reliable D2D networks in the second phase. This paper proposes a dynamic leader selection based beamforming solution based only on the instantaneous downlink channel state information (CSI) (without needing the CSI of the D2D networks) such that at the end of the first phase each group has at least one leader with high probability. A. Prior Work The wireless inter-connection of the traditional manufacturing industries is a crucial goal for future wireless standards [9] – [13]. The current wireless techniques are not designed for the stringent reliability and latency requirements of the mission-critical

applications. As a result, designing new techniques for URLLC is considered as an increasingly important goal for 5G [16], [17], with some initial efforts already taking place. For example, [18] provides a highlevel discussion about the potential to utilize diversity, e.g., MIMO, convolutional codes, and hybrid automatic repeat request (HARQ) scheme [19] to achieve URLLC. Moreover, coordinated multi-point (CoMP) [20], deployment strategies such as adjusting the cell size [21], adaptive modulation and coding (AMC) [22], as well as reduced transmission time intervals and shorter symbol durations in orthogonal frequency-division multiplexing (OFDM) systems [23] are also investigated to improve the reliability of wireless communications. However, these works in general are built on the traditional wireless techniques that are mainly driven by the broadband communications and are often inefficient for URLLC. There is general consensus that some fundamental change in the transmission protocols is necessary to satisfy the stringent latency and reliability requirements imposed by the future wireless industrial automation [17]. A recent work [24] presents an interesting two-phase transmission protocol, named Optimizing Cooperative Communication for Ultra-reliable Protocols Yoking Control Onto Wireless (Occupy CoW), for both the uplink and downlink URLLC, that makes use of cooperative relaying to reach very high levels of reliability, while maintaining a fixed cycle time of 2ms in a network of 30 nodes. For the downlink communication, specifically, the BS combines all the users' messages together and multicasts them to the users in the first phase, while the users that can decode the messages help relay them to the other users in the second phase. However, if there are too many users in the system, such a combination of all users' messages may lead to a very high multicasting

rate, resulting in too few leaders in the first phase. Our work builds upon the Occupy CoW protocol but differs in the sense that the geographic information of the users is utilized to divide them into groups: only the messages of each group is sent to the corresponding leaders, as each leader can subsequently help its neighbors. One obvious advantage of such a grouping strategy is a lower multicasting rate for each group, instead of a higher multicasting rate across all the users in the network. It is also worth noting that the conventional approach for sending individual messages to users in the downlink is information broadcasting. In the case of a single-antenna BS, the joint power and admission control problem for such a setting is investigated in [25], [26], in which the number of users achieving their SINR targets is maximized in the event that not all of them can achieve their SINR targets (e.g., when the SINR feasibility condition for power control defined in [27] does not hold). However, the goal for URLLC is to provide reliable services to all the devices, rather than a subset of devices. As a result, our work can be interpreted as an effort to enlarge the feasible SINR regime of the conventional power control and beamforming technique [27] by a utilization of the D2D network such that URLLC can be achieved in more challenging settings. B. Main Contributions The main contributions of this paper are summarized as follows. First, this paper proposes a novel two-phase transmission protocol for URLLC based on the observation that a group of devices in close proximity to each other can form a D2D network in which reliable communication is possible. Under the proposed protocol, the BS combines each group's messages together and multicasts them to the corresponding groups in the first phase, while the users that decode the messages successfully, i.e., leaders, help relay the messages to the other users in the same

group in the second phase. We point out that since reliable communication is possible in each D2D network, the core issue under our proposed protocol is the beamforming design in the first phase that aims to successfully transmit to at least one leader in each group, while the other users can rely on the leaders in the second phase. Second, we formulate the leader selection based beamforming problem in the first phase from a sparse optimization perspective by introducing a set of auxiliary variable that indicate the gap between each user's SINR and its SINR target. Such a formulation enables us to design the beamforming at the BS and select the leaders of each group jointly rather than separately. Moreover, leader selection results in a new and non-trivial sparsity pattern for the auxiliary variables since in each group at least one auxiliary variable should be zero (which implies zero gap between the SINR and SINR target and thus a leader). To achieve this desired sparsity pattern, we introduce a novel geometricmean based penalty for the auxiliary variables of each group, which is minimized to zero when each group has at least one leader. Numerical results are provided to show that such a penalty guarantees a fair leader assignment among groups. Finally, we provide a comprehensive performance comparison between our proposed strategy and the existing ones in the literature, e.g., Occupy CoW [24] and traditional information broadcasting. For various schemes, the probability of URLLC is defined as the probability that all the users in the system receive their messages successfully within the delay requirement. It is shown by simulation that with intercell interference, our proposed scheme is able to achieve a probability of URLLC above 99.99% for a much larger rate regime as compared to all the existing URLLC schemes. C. Organization The rest of this paper is organized as follows. Section II describes the system model for

URLLC. Section III introduces the D2D-based two-phase transmission protocol for URLLC. Section IV describes the corresponding leader selection based beamforming design. Section V introduces some benchmark schemes. Section VI provides the numerical simulation results pertaining to performance comparison between our proposed scheme and benchmark schemes. Finally, Section VII concludes this paper. II. SYSTEM MODEL Consider the downlink communication in one cell (factory) consisting of one BS (controller) and K users (actuators) as shown in Fig. 1. It is assumed that the BS is equipped with M antennas, and each user is equipped with one single antenna. It is further assumed that the K users form N disjoint groups based on their geographic locations, while the users in each group are in close proximity to each other. Let Gn denote the set of users that belong to group n, and its cardinality Kn = |Gn| denote the number of users in this group, n = 1, · · · , N, respectively. Note that each user belongs to only one group, thus Gn T Gj = $\emptyset$ if n 6= j, and PN n=1 Kn = K. In practice, the BS can decide how to group the users based on its knowledge of user locations, and send this information to all the users such that each user is aware of which group it belongs to. For convenience, in this paper we assume that user grouping is already done at the network planning stage and also the user grouping information is known to the users. The downlink channel from the BS to the kth user in group n is denoted by $h_{k,n} \in \mathbb{C}^{M \times 1}$, $k = 1, \cdots, K_n$,

n = 1, · · · , N, while the channel from the ith user in group j to the kth user in group n is denoted by ˜$h_{k,n,i,j} \in$ C, $\forall$(k, n) 6= (i, j). This paper adopts a block-fading model, in which all the channels follow independent quasi-static flat-fading within a block of coherence time, where $h_{k,n}$'s and ˜$h_{k,n,i,j}$'s remain constant, but vary independently from block to

block. For convenience, it is assumed that the coherence time and bandwidth of hk,n's are the same as those of h˜ k,n,i,j 's. They are denoted by T second and B Hz, respectively. It is further assumed that the downlink channels hk,n's are perfectly known at the BS, but the channels between the users h˜ k,n,i,j 's are not known at the BS. At last, we assume that for any user k in group n, it knows its downlink channel hk,n and the channels from other users to it, i.e., h˜ k,n,i,j 's, ∀(i, j) 6= (k, n), for information decoding. For URLLC, let τ denote the delay requirement for all the users, which in general is much smaller than the channel coherence time, i.e., τ <T . Furthermore, let Ωk,n and Dk,n denote the set and the number of information bits that need to be conveyed to the kth user in group n within the delay requirement, i.e., τ, respectively. The core question for ultrareliable communications in this scenario is the following: How to design a protocol such that each of the K users can receive its messages with a very low decoding error probability within τ seconds? III. PROPOSED PROTOCOL FOR ULTRA-RELIABLE COMMUNICATIONS In this section, we propose a D2D-based two-phase transmission protocol to achieve ultra-reliable communications for the K users located in N groups. We assume that h˜ k,n,i,j is very strong if n = j since the users in the same group are close to each other, but relatively weak otherwise. As a result, the users in the same group can form a D2D network in which the communications can be made reliable. The D2D-based two-phase transmission protocol is briefly outlined as follows: in the first phase with a duration of τ1 < τ, the BS combines each group's messages together, i.e., Ω (n) = SKn k=1 Ωk,n with PKn k=1 Dk,n bits information, ∀n, and sends Ω (n) 's to the corresponding groups simultaneously via multi-group multicasting; in the second phase with a duration of τ2 = τ − τ1, the users that

decode the information successfully in the first phase can help relay the messages to the other users in the same group via the D2D network. Note that under the proposed protocol, each user not only decodes its own messages, but also receives its neighbors' messages, since the successful users in Phase I need to relay other users' messages in the

same group in Phase II. In the following, we elaborate this protocol in details. A. Phase I In the first phase with a duration of τ1 seconds, let sn denote the combined symbol intended for all the users in group n, which is modeled as a circularly symmetric

complex Gaussian (CSCG) random variable with zero-mean and unit-variance, i.e., sn~CN (0, 1), ∀n. Then, the transmit signal of the BS in Phase I is expressed as

$$\boldsymbol{x}^{(I)} = \sum_{n=1}^{N} \boldsymbol{w}_n^{(I)} s_n, \qquad (1)$$

where w (I) n ∈ CM×1 denotes the transmit beamformer for the combined user massages of group n. Note that messages of different groups are spatially multiplexed. Suppose that the BS has a transmit power constraint PBS; from (1), we thus have

$$\sum_{n=1}^{N} \|\boldsymbol{w}_n^{(I)}\|^2 \leq P_{\mathrm{BS}}. \qquad (2)$$

The received signal of the kth user in group n in Phase I is expressed as

$$
\begin{aligned}
y_{k,n}^{(I)} &= \boldsymbol{h}_{k,n}^T \boldsymbol{x}^{(I)} + z_{k,n}^{(I)} \\
&= \boldsymbol{h}_{k,n}^T \boldsymbol{w}_n^{(I)} s_n + \boldsymbol{h}_{k,n}^T \sum_{j \neq n} \boldsymbol{w}_j^{(I)} s_j + z_{k,n}^{(I)}, \quad \forall k, \ \forall n, \quad (3)
\end{aligned}
$$

where $z_{k,n}^{(I)} \in \mathcal{CN}(0, I_{k,n}^{(I)})$ denotes the superposition of the additive white Gaussian noise (AWGN) and the inter-cell interference at the kth user in group n in Phase I, with a power $I_{k,n}^{(I)}$. The SINR of the kth user in group n in Phase I is then expressed as

$$\gamma_{k,n}^{(I)} = \frac{|h_{k,n}^T w_n^{(I)}|^2}{\sum_{j \neq n} |h_{k,n}^T w_j^{(I)}|^2 + I_{k,n}^{(I)}}. \qquad (4)$$

For the first transmission phase, in total there are $\tau 1B$ symbols available for the BS to perform information multicasting. Note that although the seminal work [28] shows that in an AWGN channel, encoding over finite blocklength can result in a penalty on the channel capacity, it is recently shown in [29] that in a fading channel, the error event is dominated by the outage due to channel fading, rather than the finite blocklength

effect. As a result, in this paper we ignore the effect of finite blocklength coding, and the minimum SINR target required to convey PKn k=1 Dk,n bits messages to any user in group n using τ1B symbols is then expressed as

$$\bar{\gamma}_n^{(I)} = 2^{\frac{\sum_{k=1}^{K_n} D_{k,n}}{\tau_1 B}} - 1, \quad \forall n. \qquad (5)$$

$$\phi_{k,n}^{(I)} = \begin{cases} 1, & \text{if } \gamma_{k,n}^{(I)} \geq \bar{\gamma}_n^{(I)}, \\ 0, & \text{if } \gamma_{k,n}^{(I)} < \bar{\gamma}_n^{(I)}. \end{cases} \qquad (6)$$

Definition 1: The kth user in group n is defined as a leader of group n if it can decode the messages Ω (n) in the first transmission phase, i.e., φ (I) k,n = 1. Moreover, the set of leaders in group n is defined as Φ (I) n = {k : φ (I) k,n = 1}, n = 1, · · · , N.

B. Phase II In the second phase over the remaining τ2 seconds, the leaders in each group relay the messages to the unsuccessful users in their respective group via the local D2D networks. To ensure that all the leaders in one group transmit the same message without generating intra-group interference, we can use one of the following two strategies. In the first strategy, each leader can re-transmit the entire data packet for all the users in its group (including the data for those who already decoded their data). In this case, P k Dk,n bits are transmitted by the leaders of group n over τ2 seconds in Phase II. As an alternative strategy, if each leader is made aware of which other users have successfully decoded their packet (i.e., which other users become leaders), they can subtract the messages of all the leaders from the data packet, re-encode the rest of the messages together, and transmit the newly combined packet in the second phase. Since the leaders' messages S k∈Φ (I) n Ωk,n are not encoded, only P k /∈Φ (I) n Dk,n information bits need to be transmitted by the leaders of group n over τ2 seconds in Phase II. This paper advocates the first transmission strategy mentioned above in Phase II. Although the second strategy transmits fewer number of information bits in Phase II, it requires more overhead for the feedback of control information after Phase I so that each leader is made aware of the other leaders in its group. Specifically, each user needs to feed back one bit message at the end of Phase I to report whether it becomes a leader, then the BS needs to send control messages to the leaders to let them know about the other leaders in their groups. The overhead involved is significant and is likely to overwhelm the benefit of shorter message in Phase II. For this reason, the rest of the paper assumes the use of the first strategy above. Due to the lack of global CSI of the D2D networks, in this paper we

$$
\begin{aligned}
y_{k,n}^{(II)} &= \sum_{(i,j)\neq(k,n)} \tilde{h}_{k,n,i,j} x_{i,j}^{(II)} + z_{k,n}^{(II)} \\
&= \sum_{i\neq k} \tilde{h}_{k,n,i,n} \phi_{i,n}^{(I)} \sqrt{P} s_n + \sum_{j\neq n} \sum_{i=1}^{K_j} \tilde{h}_{k,n,i,j} \phi_{i,j}^{(I)} \sqrt{P} s_j \\
&\quad + z_{k,n}^{(II)}, \text{ if } \phi_{k,n}^{(I)} = 0, \quad (8)
\end{aligned}
$$

assume that power control is not performed among users and each leader simply transmits at its full power for relaying the messages. Suppose that all the users possess a common transmit power constraint P. The transmit signal of the kth user in group n in the second phase is thus expressed as x (II) k,n = φ (I) k,n √ P sn, k = 1, · · · , Kn, n = 1, · · · , N. (7) As a result, in Phase II, the received signal for each unsuccessful user in Phase I is expressed as

where z (II) k,n∼ CN (0, I (II) k,n) denotes the superposition of the AWGN and the inter-cell interference at the kth user in group n in Phase II, with a power I (II) k,n .

The corresponding SINR to decode sn based on y (II) k,n is thus

$$
\gamma_{k,n}^{(II)} = \frac{\left| \sum_{i\neq k} \tilde{h}_{k,n,i,n} \phi_{i,n}^{(I)} \sqrt{P} \right|^2}{\sum_{j\neq n} \left| \sum_{i=1}^{K_j} \tilde{h}_{k,n,i,j} \phi_{i,j}^{(I)} \sqrt{P} \right|^2 + I_{k,n}^{(II)}}, \text{ if } \phi_{k,n}^{(I)} = 0.
$$

(9)

Similar to (5), the minimum SINR requirement to relay $\sum_k D_{k,n}$ bits of information using $\tau_2 B$ symbols in Phase II can be expressed as

$$
\bar{\gamma}_n^{(II)} = 2^{\frac{\sum_{k=1}^{K_n} D_{k,n}}{\tau_2 B}} - 1, \quad \forall n. \quad (10)
$$

Then, we define an indicator function for each user as follows:

$$
\phi_{k,n}^{(II)} = \begin{cases} 1, & \text{if } \gamma_{k,n}^{(I)} < \bar{\gamma}_n^{(I)} \text{ and } \gamma_{k,n}^{(II)} \geq \bar{\gamma}_n^{(II)}, \\ 0, & \text{otherwise.} \end{cases} \quad (11)
$$

As a result, φ (II) k,n = 1 if an unsuccessful user in Phase I decodes the messages in Phase

II, and φ (II) k,n = 0 otherwise. Define Φ (II) n = {k : φ (II) k,n = 1} as the set of users that can decode the messages successfully in group n, n = 1, · · · , N, in Phase II. Then, the cardinalities of Φ (I) n (Definition 1) and Φ (II) n , i.e., |Φ (I) n | and |Φ (II) n |, indicate the numbers of users in group n who have decoded their messages successfully in Phase I (leaders) and Phase II, respectively. Moreover, PN n=1 |Φ (I) n | + PN n=1 |Φ (II) n | denotes the total number of successful users within the cell after τ second. We define reliable communications in the whole system as follows. Definition 2: Given a time slot of duration τ seconds, if some beamforming vectors w (I) n 's satisfying the transmit power constraints (2) can be found at the BS such that all the users can receive their messages under the proposed two-phase transmission protocol, i.e., PN n=1 |Φ (I) n | + PN n=1 |Φ (II) n | = K, then ultra-reliable communication is achieved over the τ seconds. Otherwise, we say that an outage has occurred

C. Problem Formulation To achieve ultra-reliable communication over any particular duration τ, we need to design the beamformingvectors at the BS to maximize the total number of the successful users, i.e.,

$$\underset{\{\boldsymbol{w}_n^{(\mathrm{I})}\}}{\text{maximize}} \sum_{n=1}^{N} |\Phi_n^{(\mathrm{I})}| + \sum_{n=1}^{N} |\Phi_n^{(\mathrm{II})}| \qquad (12\mathrm{a})$$

$$\text{subject to} \sum_{n=1}^{N} \|\boldsymbol{w}_n^{(\mathrm{I})}\|^2 \leq P_{\mathrm{BS}}. \qquad (12\mathrm{b})$$

If the optimal value to the above problem is K, then ultrareliable communication is achieved according to Definition 2.

IV. LEADER SELECTION BASED BEAMFORMING DESIGN Since in practice it is hard to acquire the CSI of the D2D networks, i.e., h˜ k,n,i,j 's, at the BS, in this section, we propose a reformulation of

problem (12) without assuming any knowledge of h˜ k,n,i,j 's. A. Problem Reformulation The proposed two-phase transmission protocol in Section III arises from the observation that if a group has at least one leader in the first phase, then with a very high probability, all the other users in the group would be able to decode the messages successfully over the D2D network in the second phase due to their proximity to the leaders. Motivated by this observation, this paper reformulates the problem by setting the constraint of Phase I so as to ensure that each group has at least one leader. This can be done without knowledge of h˜ k,n,i,j 's. Moreover, among all beamforming strategies w (I) n 's that yield at least one leader for each group, we choose the one that maximizes the total number of leaders in Phase I, so that fewer users need to satisfy their SINR requirements in Phase II. In this way, we formulate the following beamforming design problem, assuming no knowledge of h˜ k,n,i,j 's:

$$\underset{\{\boldsymbol{w}_n^{(\mathrm{I})}\}}{\text{maximize}} \sum_{n=1}^{N} |\Phi_n^{(\mathrm{I})}| \qquad (13\mathrm{a})$$

$$\text{subject to } |\Phi_n^{(\mathrm{I})}| \geq 1, \quad \forall n, \qquad (13\mathrm{b})$$

$$\sum_{n=1}^{N} \|\boldsymbol{w}_n^{(\mathrm{I})}\|^2 \leq P_{\mathrm{BS}}. \qquad (13\mathrm{c})$$

In problem (13), φ (I) k,n's as given in (6) are complicated and discrete functions over the beamforming vectors, which make it challenging to apply optimization technique to solve problem (13). To tackle the issue arising from the discrete φ (I) k,n's, let us define t (I) n = [t (I) 1,n, · · · , t(I) Kn,n] T ∈ C Kn×1 , ∀n. It can then be shown that problem (13) is equivalent to the following problem:

$$\underset{\{w_n^{(I)}, t_n^{(I)}\}}{\text{minimize}} \sum_{n=1}^{N} \|t_n^{(I)}\|_0 \tag{14a}$$

$$\text{subject to } \gamma_{k,n}^{(I)} + t_{k,n}^{(I)} \geq \bar{\gamma}_n^{(I)}, \ \forall k, \ \forall n, \tag{14b}$$

$$\|t_n^{(I)}\|_0 \leq K_n - 1, \ \forall n, \tag{14c}$$

$$t_{k,n}^{(I)} \geq 0, \ \forall k, \ \forall n, \tag{14d}$$

$$\sum_{n=1}^{N} \|w_n^{(I)}\|^2 \leq P_{\text{BS}}, \tag{14e}$$

where γ (I) k,n's are given in (4). The equivalence between problem (13) and problem (14) is due to the fact that the auxiliary variables t (I) k,n's characterize the gap between the SINR targets and achievable SINRs in Phase I, thus the number of zero t (I) k,n's denotes the number of leaders in Phase I, and constraint (14c) guarantees at least one leader in each group. Such an equivalent transformation based on the auxiliary variables t (I) k,n's results in a continuous problem (14). Moreover, since we optimize the number of zero elements in t (I) n 's, sparse optimization techniques can now be used to solve the problem. The difficulty to solve (14) lies in the multiple cardinality constraints (14c). To ensure that there is at least one leader in each group in Phase I without having to deal with the complicated cardinality constraints (14c), in this paper, we propose the following novel leader selection based beamforming problem:

$$\underset{\{w_n^{(I)}, t_n^{(I)}\}}{\text{minimize}} \sum_{n=1}^{N} \|t_n^{(I)}\|_1 + \sum_{n=1}^{N} \beta_n \sqrt[K_n]{\prod_{k=1}^{K_n} t_{k,n}^{(I)}} \tag{15a}$$

$$\text{subject to } (14b), (14d), (14e), \tag{15b}$$

where $\beta_n > 0$ is the corresponding penalty weight for group n, $n = 1, \cdots, N$.

Note that in the objective function of problem (15), we use the convex functions $\|t_n^{(I)}\|_1$'s to approximate non-convex functions $\|t_n^{(I)}\|_0$'s in problem (14) based on standard sparse optimization technique. In addition, we define a new penalty[1] for each group as $\beta_n \sqrt[K_n]{\prod_{k=1}^{K_n} t_{k,n}^{(I)}}$ in the objective function of problem (14). It is easily observed that for each group n, its penalty is zero if at least one element of $t_n^{(I)}$ is zero. As a result, the new penalty can lead to the desired sparsity pattern, i.e., at least one zero in $t_n^{(I)}$, $\forall n$. Moreover, since the geometric mean $f(x = [x_1, \cdots, x_N]) = \sqrt[N]{\prod_{n=1}^{N} x_n}$ is a concave function over x [30], the function $\sum_{n=1}^{N} \beta_n \sqrt[K_n]{\prod_{k=1}^{K_n} t_{k,n}^{(I)}}$ is concave over $t_{k,n}^{(I)} > 0$'s. The objective function of problem (15) is thus the difference of convex functions, and standard techniques such as the successive convex approximation method can be used to solve the problem to a local optimum. In the rest of this section, we give details on how to solve the above leader selection based beamforming design problem.

B. Beamforming Design to Problem (15) The new penalty in problem (15) enables us to bypass the complicated cardinality constraints of t (I) n 's in problem (14). However, problem (15) is a non-convex problem and it is thus difficult to find its globally optimal solution. On one hand, the penalty in the objective function is non-convex. On the other hand, the SINR constraints given in (14b) are also non-convex. In the following, we propose an algorithm based on the successive convex approximation technique that can yield a locally optimal solution to problem (15). First, we provide a convex upper bound to the objective function of problem (15). Since the function

$\sum_{n=1}^{N} \beta_n \sqrt[K_n]{\prod_{k=1}^{K_n} t_{k,n}^{(I)}}$ is concave over $t_{k,n}^{(I)} > 0$'s as

1 In this paper, we set the penalty weight as $\beta_n = 2K_n$, $\forall n$, i.e., a higher penalty is introduced to the groups with more users since a leader in such groups can help more users in the second phase

illustrated in Section IV-A, its first-order approxin
any given point $\hat{t}_n^{(I)} = [\hat{t}_{1,n}^{(I)}, \cdots, \hat{t}_{K_n,n}^{(I)}]^T$, $\forall n$:

$$f(\{t_{k,n}^{(I)}, \hat{t}_{k,n}^{(I)}\}) = \sum_{n=1}^{N} \beta_n \sqrt[K_n]{\prod_{k=1}^{K_n} \hat{t}_{k,n}^{(I)}} +$$

$$\sum_{n=1}^{N} \frac{\beta_n}{K_n} \sqrt[K_n]{\prod_{k=1}^{K_n} \hat{t}_{k,n}^{(I)}} \left[\frac{1}{\hat{t}_{1,n}^{(I)}}, \cdots, \frac{1}{\hat{t}_{K_n,n}^{(I)}}\right] (t_n^{(I)} - \hat{t}$$

serves as its upper bound. Next, we deal with the non-convex SINR constraints (14b). Well-known methods to deal with the multicasting SINR constraints include semidefinite relaxation (SDR) [31] and successive convex approximation [32]. Recently, it is shown in [33] that the successive convex approximation based algorithm in general achieves better performance in multicasting than the SDR-based algorithm when the number of antennas at the BS and the number of devices are large. As a result, in this paper, we adopt the successive convex approximation technique to deal with the non-convex SINR constraints (14b). First, the SINR constraints (14b) can be re-formulated as

$$\frac{|h_{k,n}^T w_n^{(I)}|^2}{\bar{\gamma}_n^{(I)}} + I_{k,n}^{(I)} t_{k,n}^{(I)} \geq \sum_{j \neq n} |h_{k,n}^T w_j^{(I)}|^2 + I_{k,n}^{(I)}, \forall$$

The main issue is that $|h_{k,n}^T w_n^{(I)}|^2$ is a convex function, rather than a concave function, over $w_n^{(I)}$. We use the first-order approximation to provide concave upper bounds to $|h_{k,n}^T w_n^{(I)}|^2$'s as in [32]. Specifically, define

$$a_{k,n} = \Re(h_{k,n}^T w_n^{(I)}), \quad \forall k, \forall n, \quad (18)$$

$$b_{k,n} = \Im(h_{k,n}^T w_n^{(I)}), \quad \forall k, \forall n. \quad (19)$$

Then we have $|h_{k,n}^T w_n^{(I)}|^2 = a_{k,n}^2 + b_{k,n}^2$, $\forall k, n$. Given any $\hat{w}_n^{(I)}$'s that satisfy the transmit power constraint (2), we can define $\hat{a}_{k,n}$'s and $\hat{b}_{k,n}$'s as (18) and (19). Then, the first-order approximation to $|h_{k,n}^T w_n^{(I)}|^2$ at this particular point is given by

$$g_{k,n}(a_{k,n}, b_{k,n}, \hat{a}_{k,n}, \hat{b}_{k,n}) = \hat{a}_{k,n}^2 + \hat{b}_{k,n}^2$$
$$+ 2[\hat{a}_{k,n}, \hat{b}_{k,n}] \left(\begin{bmatrix} a_{k,n} \\ b_{k,n} \end{bmatrix} - \begin{bmatrix} \hat{a}_{k,n} \\ \hat{b}_{k,n} \end{bmatrix}\right), \forall k, n. \quad (20)$$

To summarize, by introducing the auxiliary variables $a_{k,n}$'s and $b_{k,n}$'s, given any fixed $\hat{w}_n^{(I)}$'s, we can use the following convex constraints to approximate the non-convex constraints (17):

$$\frac{g_{k,n}(a_{k,n}, b_{k,n}, \hat{a}_{k,n}, \hat{b}_{k,n})}{\bar{\gamma}_n^{(I)}} + I_{k,n}^{(I)} t_{k,n}^{(I)}$$
$$\geq \sum_{j \neq n} |h_{k,n}^T w_j^{(I)}|^2 + I_{k,n}^{(I)}, \quad \forall k, \forall n. \quad (21)$$

Algorithm 1 Proposed Algorithm for Solving Problem (15)

Initialization: Set the initial values for $\hat{t}_n^{(I)}$'s and $\hat{w}_n^{(I)}$'s and set $l = 1$;

Repeat:
1) Find the optimal solution to problem (22) using CVX as $\{w_n^{(I,l)}, t_n^{(I,l)}, a_n^{(l)}, b_n^{(l)}\}$;
2) Update $\hat{w}_n^{(I)} = w_n^{(I,l)}$ (thus $\hat{a}_n = a_n^{(l)}$ and $\hat{b}_n = b_n^{(l)}$), $\hat{t}_n^{(I)} = t_n^{(I,l)}$, $\forall n$;
3) $l = l + 1$.

Until convergence

With the above approximations, given any fixed $\hat{t}_n^{(I)}$'s and $\hat{w}_n^{(I)}$'s, we can solve the following convex problem:

$$\underset{\{w_n^{(I)}, t_n^{(I)}, a_n, b_n\}}{\text{minimize}} \sum_{n=1}^{N} \|t_n^{(I)}\|_1 + f(\{t_{k,n}^{(I)}, \hat{t}_{k,n}^{(I)}\}) \quad (22a)$$

$$\text{subject to} \quad (18), (19), (21), (14d), (14e), \quad (22b)$$

where $a_n = [a_{1,n}, \cdots, a_{K_n,n}]^T$, $b_n = [b_{1,n}, \cdots, b_{K_n}$ $\forall n$, and $f(\{t_{k,n}^{(I)}, \hat{t}_{k,n}^{(I)}\})$ are as given in (16). Since pro (22) is a convex problem, it can be globally solved existing standard package such as CVX [34]. The succe convex approximation based algorithm to solve the l selection based beamforming problem, i.e., (15), proc by iteratively updating $\hat{t}_n^{(I)}$'s and $\hat{w}_n^{(I)}$'s (thus $\hat{a}_n$'s $\hat{b}_n$'s) based on the solution to problem (22). The prop algorithm is summarized in Algorithm 1. The converg behavior of Algorithm 1 is guaranteed in the follo proposition.

Proposition 1: Monotonic convergence of Algorithm guaranteed, i.e.,

$$\sum_{n=1}^{N} \|t_n^{(I,l+1)}\|_1 + \sum_{n=1}^{N} \beta_n \sqrt[K_n]{\prod_{k=1}^{K_n} t_{k,n}^{(I,l+1)}}$$
$$\leq \sum_{n=1}^{N} \|t_n^{(I,l)}\|_1 + \sum_{n=1}^{N} \beta_n \sqrt[K_n]{\prod_{k=1}^{K_n} t_{k,n}^{(I,l)}},$$

where l denotes the index of iteration of Algorith Moreover, the converged solution satisfies all the const as well as the Karush-Kuhn-Tucker (KKT) conditio problem (15).

Proof: Please refer to Appendix A.

The leader selection based beamforming design for ing problem (12) to maximize the total number of users under the proposed two-phase transmission sc in Section III and to achieve ultra-reliable commu tion as defined in Definition 2 is summarized as fol First, we solve problem (15) via Algorithm 1 and $w_n^{(I)}$'s to calculate $\phi_{k,n}^{(I)}$'s based on (6). Then, we calc $\phi_{k,n}^{(II)}$'s based on (11). Reliable communication is ach if $\sum_{n=1}^{N} |\Phi_n^{(I)}| + \sum_{n=1}^{N} |\Phi_n^{(II)}| = K$.

V. BENCHMARK SCHEMES FOR ULTRA-RELIABLE COMMUNICATIONS

In this section, we briefly introduce several other potential approaches for URLLC as benchmark schemes. Numerical simulation comparisons are provided in Section VI to show the performance gain of our proposed scheme over the benchmark schemes. A. Benchmark Scheme 1: Our Proposed Scheme but without Leader Selection First, to illustrate

the importance of leader selection in our proposed two-phase transmission scheme, in the following we consider one benchmark scheme in which the number of leaders is maximized in Phase I without encouraging at least one leader for each group. In this case, problem (15) reduces to

$$\underset{\{w_n^{(I)}, t_n^{(I)}\}}{\text{minimize}} \sum_{n=1}^{N} \|t_n^{(I)}\|_0 \tag{24a}$$
$$\text{subject to } (14b), (14d), (14e). \tag{24b}$$

In other words, the penalty for encouraging at least one leader in each group is removed. Problem (24) can be solved in a similar way as problem (15). As a remark, the strategy that maximizes the total number of leaders in Phase I is expected to activate more users in the cell-center groups that are close to the BS, because they have strong direct channels and do not suffer from inter-cell interference. As a result, it is expected that the optimized beamformers for problem (24) would result in many leaders in the cell-center groups, while having no leader in the cell-edge groups. In contrast, our proposed scheme tries to activate at least one leader in each group by solving problem (15). It promotes fairness among different groups in Phase I and thus makes the best use of the D2D network in Phase II, as later verified by numerical results in Section VI. B. Benchmark Scheme 2: Occupy CoW Protocol [24] The Occupy CoW protocol is proposed in [24] for achieving URLLC in a similar setup as in this paper. Specifically, Occupy CoW protocol is also a two-phase transmission protocol where user messages are transmitted from the BS to the users in the first phase and the successful users help relay the messages to the other users via the D2D network in the second phase. The main difference of Occupy CoW protocol as compared to our proposed protocol lies in the fact that all the users form one group, rather

than N groups based on their geographic locations. As a result, in the first phase, the BS combines all users' messages together, i.e., $\Omega = S_{k,n} \Omega_{k,n}$, and multicasts this entire message to all the users, while in the second phase, the successful users in Phase I can help all the unsuccessful users via the D2D network. Note that [24] only considers the case that the BS has one single antenna. To make a fair comparison, in the following we briefly introduce how to design the beamforming under the Occupy CoW protocol if the BS has multiple antennas. Let $s \sim CN(0, 1)$ denote the entire message intended for all the users. The received signal at each user in the first phase is expressed as

$$y_{k,n}^{(I)} = \boldsymbol{h}_{k,n}^T \boldsymbol{w}^{(I)} s + z_{k,n}^{(I)}, \quad \forall k, n, \quad (25)$$

where $\boldsymbol{w}^{(I)} \in \mathbb{C}^{M \times 1}$ denotes the multicast beamforming at the BS in Phase I. The corresponding SINR to decode s is

$$\gamma_{k,n}^{(I)} = \frac{|\boldsymbol{h}_{k,n}^T \boldsymbol{w}^{(I)}|^2}{I_{k,n}^{(I)}}, \quad \forall k, n. \quad (26)$$

Similar to (5) in Section III, to decode a message of $\sum_{n=1}^{N} \sum_{k=1}^{K_n} D_{k,n}$ bits using $B\tau_1$ symbols, the identical minimum SINR requirement for all the users is expressed as:

$$\bar{\gamma}^{(I)} = 2^{\frac{\sum_{n=1}^{N} \sum_{k=1}^{K_n} D_{k,n}}{B\tau_1}} - 1. \quad (27)$$

The indicator function of each user in Phase I then depends on whether its SINR satisfies this common SINR target or not, i.e.,

$$\phi_{k,n}^{(I)} = \begin{cases} 1, & \text{if } \gamma_{k,n}^{(I)} \geq \bar{\gamma}^{(I)}, \\ 0, & \text{if } \gamma_{k,n}^{(I)} < \bar{\gamma}^{(I)}. \end{cases} \quad (28)$$

The users that decode the messages successfully in the first phase can then relay the messages to the other users in the second phase. Similar to Section III-B, each leader only needs to relay the message $\bigcup_n \bigcup_{k \notin \Phi_n^{(I)}} \Omega_{k,n}$ with $\sum_{n=1}^{N} \sum_{k \notin \Phi_n^{(I)}} D_{k,n}$ bits information to the other users, where $\Phi_n^{(I)}$ denotes the set of leaders located in group n. The minimum SINR required to send the above message using $B\tau_2$ symbols is

$$\bar{\gamma}^{(II)} = 2^{\frac{\sum_{n=1}^{N} \sum_{k \notin \Phi_n^{(I)}} D_{k,n}}{B\tau_2}} - 1. \quad (29)$$

The key to achieve ultra-reliable communications for the Occupy CoW protocol is the assumption that full diversity gain can be achieved in the second phase even though there is no cooperation between the BS and leaders. Specifically, [24] assumes that if the kth user in group n does not decode the messages in Phase I, based on some space-time coding technique, its SINR achieved from the second phase transmission is

$$\gamma_{k,n}^{(II)} = \max_{(i,j) \neq (k,n)} \gamma_{k,n}^{(i,j)}, \quad \text{if } \phi_{k,n}^{(I)} = 0, \quad (30)$$

where

$$\gamma_{k,n}^{(i,j)} = \frac{\phi_{i,j}^{(I)} P |\tilde{h}_{k,n,i,j}|^2}{I_{k,n}^{(II)}}, \quad \forall (i,j) \neq (k,n), \quad (31)$$

denotes the individual SINRs of the orthogonal signals from the ith user in group j. Depending on whether its SINR 9 in Phase II satisfies the SINR target or not, the indicator functions of the unsuccessful users in Phase I are defined as

$$\phi_{k,n}^{(II)} = \begin{cases} 1, & \text{if } \gamma_{k,n}^{(II)} < \bar{\gamma}^{(II)} \text{ and } \gamma_{k,n}^{(II)} \geq \bar{\gamma}^{(II)}, \\ 0, & \text{otherwise.} \end{cases} \quad (32)$$

Under the Occupy Cow protocol, the objective of designing the beamforming vectors $\boldsymbol{w}^{(I)}$ is to maximize the total number of successful users in Phase I such that more leaders can help relay the messages in Phase II, i.e.,

$$\underset{\{\boldsymbol{w}^{(I)}, \boldsymbol{t}_n^{(I)}\}}{\text{minimize}} \sum_{n=1}^{N} \|\boldsymbol{t}_n^{(I)}\|_0 \quad (33a)$$

$$\text{subject to } \gamma_{k,n}^{(I)} + t_{k,n}^{(I)} \geq \bar{\gamma}^{(I)}, \quad \forall k, n, \quad (33b)$$

$$t_{k,n}^{(I)} \geq 0, \quad \forall k, n, \quad (33c)$$

$$\|\boldsymbol{w}^{(I)}\|^2 \leq P_{\text{BS}}. \quad (33d)$$

Similar to Algorithm 1, we can use sparse optimization and successive convex approximation techniques to solve problem (33), and then determine whether URLLC is achieved under the Occupy CoW protocol, i.e., whether all the users receive their messages over τ seconds. Note that in [24], all the users' channels are assumed to have the same distribution since the locations of the users are not considered in the channel model. This paper shows that if in practice the effect of user locations on user channels is considered, we should group the users in close proximity together since the gain of D2D network in Phase II mainly comes from the adjacent users. Note that in Phase I, user grouping leads to multi-group multicast, instead of singlegroup multicast under the Occupy CoW protocol in [24]. If there is only one omnidirectional antenna at the BS, Occupy CoW works well since there is no inter-group interference in Phase I if the BS does not separate users into groups in the spatial domain. If the BS is equipped with multiple antennas, however, our proposed scheme in Section III works better since the BS can adjust N beams in Phase I and utilize the spatial multiplexing gain to activate one leader in each group as long as M > N. Note that for Occupy CoW, the BS only designs one beam in Phase I to satisfy the SINR targets of all the users, which is hard. The increased degree of freedom, i.e., optimization of N beams rather than one beam, ensures that our proposed scheme is more powerful than Occupy CoW, as later verified by numerical results in Section VI

C. Benchmark Scheme 3: Modified Occupy CoW Protocol with Leader Selection

Leader selection introduced in Section IV can be applied in Occupy CoW as well to improve its performance, if the users are geographically located in separate groups.. In the first phase, although all users' messages are combined

together, we can design w(I) to activate at least one leader in each geographical group. The considered problem is thus formulated as

$$\underset{\{w^{(\mathrm{I})}, t_n^{(\mathrm{I})}\}}{\text{minimize}} \sum_{n=1}^{N} \|t_n^{(\mathrm{I})}\|_0 + \sum_{n=1}^{N} \beta_n \sqrt[\kappa_n]{\prod_{k=1}^{K_n} t_{k,n}^{(\mathrm{I})}} \quad (34a)$$

$$\text{subject to } (33b), (33c), (33d). \quad (34b)$$

This problem can be solved similarly as problem (15). For Occupy CoW, combining all users' messages together significantly increases the SINR targets of all the users as compared to our proposed scheme, thus leads to a reduction of the number of leaders in Phase I. In Phase II, an geographically isolated group without a leader nevertheless cannot rely on the far away leaders in other groups. As a result, even if leader selection is considered in Occupy CoW, its probability of reliable communications is much lower than our proposed scheme, as later verified by numerical results in Section VI.

D. Benchmark Scheme 4: One-Phase Transmission Protocol with Broadcasting

The proposed scheme in Section III and Occupy CoW proposed in [24] both advocate a two-phase transmission for achieving ultra-reliable communications by utilizing the D2D network in the second phase. In the rest of this section, we introduce possible approaches with one-phase transmission where only the downlink communication from the BS to the users is considered. First, consider the case of broadcasting. In this case, the received signal of the kth user in group n is

$$y_{k,n} = h_{k,n}^T w_{k,n} s_{k,n} + h_{k,n}^T \sum_{(i,j) \neq (k,n)} w_{i,j} s_{i,j} + z_{k,n}, \quad \forall k, n, \quad (35)$$

$$y_{k,n} = h_{k,n}^T w_{k,n} s_{k,n} + h_{k,n}^T \sum_{(i,j) \neq (k,n)} w_{i,j} s_{i,j} + z_{k,n}, \quad \forall k, n, \quad (35)$$

where $s_{k,n} \sim \mathcal{CN}(0,1)$ denotes the message for the kth user in group n, $w_{k,n} \in \mathbb{C}^{M \times 1}$ denotes the corresponding beamforming vector, and $z_{k,n} \sim \mathcal{CN}(0, I_{k,n})$ denotes the superposition of the AWGN and inter-group interference, with a power $I_{k,n}$. The SINR for decoding $s_{k,n}$ is thus

$$\gamma_{k,n} = \frac{|h_{k,n}^T w_{k,n}|^2}{\sum_{(i,j) \neq (k,n)} |h_{k,n}^T w_{i,j}|^2 + I_{k,n}}, \quad \forall k, n. \quad (36)$$

Note that the minimum SINR target to deliver $D_{k,n}$ bits information using $B\tau$ symbols is:

$$\tilde{\gamma}_{k,n} = 2^{\frac{D_{k,n}}{B\tau}} - 1, \quad \forall k, n. \quad (37)$$

Then, we can design the beamforming vectors $w_{k,n}$'s to maximize the number of users whose SINRs satisfy their SINR targets, i.e.,

$$\underset{\{w_{k,n}, t_{k,n}\}}{\text{minimize}} \sum_{n=1}^{N} \sum_{k=1}^{K_n} |t_{k,n}|_0 \quad (38a)$$

$$\text{subject to } \gamma_{k,n} + t_{k,n} \geq \tilde{\gamma}_{k,n}, \quad \forall k, n, \quad (38b)$$

$$t_{k,n}^{(I)} \geq 0, \quad \forall k, n, \quad (38c)$$

$$\sum_{n=1}^{N} \sum_{k=1}^{K_n} \|w_{k,n}\|^2 \leq P_{BS}. \quad (38d)$$

Similar to [35], we can transform the SINR constraint (38b) into the following convex form:

$$\frac{h_{k,n}^T w_{k,n}}{\tilde{\gamma}_{k,n}} + I_{k,n} t_{k,n}$$
$$\geq \sqrt{\sum_{(i,j) \neq (k,n)} |h_{k,n}^T w_{i,j}|^2 + I_{k,n}}, \quad \forall k, n. \quad (39)$$

Then, via relaxing $|t_{k,n}|_0$'s by $|t_{k,n}|_1$'s in the objective function (38a), we can design the beamforming vectors by solving the following convex problem:

$$\underset{\{w_{k,n}, t_{k,n}\}}{\text{minimize}} \sum_{n=1}^{N} \sum_{k=1}^{K_n} |t_{k,n}|_1 \quad (40a)$$

$$\text{subject to } (39), (38c), (38d). \quad (40b)$$

URLLC via broadcasting is achieved if with the obtained beamforming vectors, all the users satisfy their SINR targets.

As a remark, in most use cases for URLLC the rate requirement of each user is very low, thus even if we combine each group's messages together, the SINR requirement for each group is still reasonable. In this case, intuitively, with M antennas at the BS, approximately M leaders and thus M groups with one leader in each group can be supported under our proposed scheme. On the other hand, information broadcasting can only support approximately M users in toal, i.e., URLLC is not possible if K ≫ M. With a reliable D2D network, our proposed scheme can achieve URLLC even when K ≫ M as long as the number of groups satisfy N < M.

E. Benchmark Scheme 5: One-Phase Transmission Protocol with TDMA

Another strategy to convey Ωk,n to the kth user in group n, ∀k, n, is TDMA. Specifically, when the kth user in group n is scheduled, the BS can implement a maximal-ratio transmission (MRT) beamforming, and the received signal at the user is

$$y_{k,n} = h_{k,n}^T \frac{\sqrt{p_{k,n}} h_{k,n}^*}{\|h_{k,n}\|} s_{k,n} + z_{k,n}, \quad \forall k, n, \quad (41)$$

where $p_{k,n}$ denotes the power to transmit the message $\Omega_{k,n}$. The corresponding SINR at the user is

$$\gamma_{k,n} = \frac{p_{k,n} \|h_{k,n}\|^2}{I_{k,n}}, \quad \forall k, n. \quad (42)$$

Note that with TDMA, $D_{k,n}$ bits of information needs to be transmitted to the kth user in group n with only $B\tau/K$ symbols. The corresponding minimum SINR target for each user is:

$$\hat{\gamma}_{k,n} = 2^{\frac{KD_{k,n}}{B\tau}} - 1, \quad \forall k, n. \tag{43}$$

Then, the transmit power to satisfy each user's SINR target is

$$p_{k,n}^* = \frac{\hat{\gamma}_{k,n} I_{k,n}}{\|\boldsymbol{h}_{k,n}\|^2}, \quad \forall k, n. \tag{44}$$

If the total transmit power at the BS satisfies the transmit power constraint, i.e.,

$$\sum_{n=1}^{N} \sum_{k=1}^{K_n} p_{k,n}^* \leq P_{\text{BS}}, \tag{45}$$

then all the users can be supported via TDMA, and thus URLLC is achieved.

In a massive connectivity scenario, the number of users, i.e., K, can be large. If each user is only allocated τ/K seconds for information transmission, the resulting SINR target as shown in (43) is quite high. As a result, as later shown in Section VI, TDMA cannot achieve reliable communications in general.

F. Benchmark Scheme 6: One-Phase Transmission Protocol with Multi-Group Multicasting

Moreover, similar to the first phase in our proposed scheme, we can also apply multi-group multicasting to convey the message Ω (n) with PKn k=1 Dk,n bits to all the users in group n, ∀n, but using all the Bτ symbols in one shot. The SINR of each user is given in (4). However, since the transmission time is doubled compared to our proposed two-phase transmission scheme, the SINR target for each group is reduced. Specifically, the minimum SINR requirement for the users in group n is:

$$\hat{\gamma}_n = 2^{\frac{\sum_{k=1}^{K_n} D_{k,n}}{B\tau}} - 1, \quad \forall n. \tag{46}$$

We can proceed to solve problem (24) in Benchmark Scheme 1 with the new SINR targets γ˜n's. If one beamforming solution can be found such that all the users can decode their messages, then URLLC is achieved for this scheme.

However, similar to Benchmarks 4 and 5, the lack of utilization of the reliable D2D network makes it difficult for this one-phase scheme to guarantee the performance of cell-edge users who experience strong inter-cell interference.

VI. NUMERICAL RESULTS This section presents the numerical results to verify the effectiveness of the proposed two-phase transmission protocol in Sections III and IV for URLLC as compared to the benchmark schemes introduced in Section V.
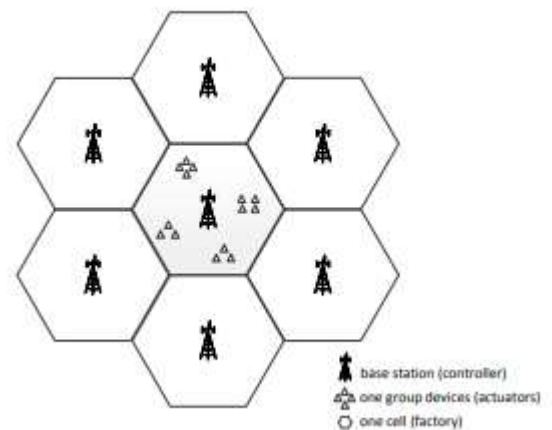

Fig. 2. The cellular-based model for industrial automation: the cell, BS, and user act as factory, controller, and actuator, respectively, in industrial automation. In this setup, the center cell is the reference cell of interests, while the other 6 cells generate inter-cell interference to the users in the reference cell.

The simulation setup is as follows. As shown in Fig. 2, we consider a network consisting of 7 cells in a wrapped around topology, in which

the center cell is the reference cell—the performance of which we are interested in, and the other 6 adjacent cells generate inter-cell interference to the users in the reference cell.[2] This cellular-based topology

is a proper model for industrial automation: each cell, BS, and device can be viewed as the factory, controller, and actuator, respectively, in which each factory consists of several groups of actuators. As a result, the simulation results shown in this section are valid in an industrial automation topology. In the simulation, the cell radius is 500m. It is assumed that there are N = 6 groups in the reference cell, while each group consists of 8 users, i.e, K = 48. The center of each group is randomly located in a doughnut shape of inner radius Rinner and outer radius Router. Moreover, each group covers an area with a radius of 20m, and its users are randomly located in the area covered by this group. The BS is assumed to be equipped with M = 8 antennas. The downlink channel from the BS to the kth user in group n is modeled as hk,n = ηk,ngk,n, where ηk,n denotes the pathloss component, and gk,n∼ CN (0, I) denotes the Rayleigh fading component. Moreover, the path-loss component is modeled as −128.1 − 36.7 log10(dk,n) in dB, where dk,n in km denotes the distance from the kth user in group n to the BS. For the communications between the users over the D2D network, if two users are in the same group, we model their channel as Rician fading; otherwise, we model their channel as Rayleigh fading. Specifically, the channel from the ith user in group n to the kth user in the same group is modeledas

$$h_{k,n,i,n} = \eta_{k,n,i,n}\left(\sqrt{\delta/(\delta+1)}\hat{g}_{k,n,i,n} + \sqrt{1/(\delta+1)}\tilde{g}_{k,n,i,n}\right).$$

, where $\eta_{k,n,i,n}$ denotes the path-loss

other 6 cells' beamformers at the BSs in Phase I and leader locations in Phase II. Then, we calculate the inter-cell interference for the reference cell and design its beamforming accordingly as shown in Section IV. In other words, only the reliability of the reference cell is considered.

component, $\hat{g}_{k,n,i,n}$ with $|\hat{g}_{k,n,i,n}|^2 = 1$ denotes the line-ofsight (LOS) deterministic component, $\tilde{g}_{k,n,i,n} \sim CN(0,1)$ denotes the Rayleigh fading component, and δ denotes the Rician factor specifying the power ratio between the LOS and fading component. In this paper, we set δ = 4 and characterize the path-loss component as −76.8 − 18.7 log10(dk,n,i,n) in dB according to the indoor channel model in [36], where dk,n,i,n denotes the distance between the two users. Moreover, the channel from the ith user in group j to the kth user in another group n is modeled as $\tilde{h}_{k,n,i,j} = \eta_{k,n,i,j}\tilde{g}_{k,n,i,j}$, where the path-loss component is modeled as −128.1 − 36.7 log10(dk,n,i,j) in dB, and $\tilde{g}_{k,n,i,j} \sim CN(0,1)$ denotes the Rayleigh fading component. The total bandwidth used is assumed to be B = 100kHz for both the downlink channels hk,n's and D2D channels $\tilde{h}_{k,n,i,j}$'s. We assume flat fading across this bandwidth B. The delay requirement for the communication is τ = 1ms. As a result, there are in total Bτ = 100 transmit symbols available for delivering the messages Ωk,n's to each user. Moreover, we set τ1 = 0.75ms in Phase I and τ2 = 0.25ms in Phase II. The motivation for allocating more time to Phase I is to reduce its SINR requirement (see (5)) such that there is at least one leader in each group. The transmit power constraint at the BS is PBS = 43dBm, and at the user is P = 23dBm. The power spectral density of the AWGN at the users is assumed to be −169dBm/Hz. For convenience, it is assumed that the message sizes of all the users are the same, i.e., Dk,n = D, ∀k, n. In the simulation, we generate 10, 000 channel realizations and

for each realization, we measure whether all the users decode their messages within $\tau = 1ms$ for each investigated scheme. The probability of URLLC for each scheme is then defined as the percentage of instances achieving reliable communication.

A. Performance Comparison between Proposed Scheme and Benchmark Schemes when D = 22

In this numerical example, we assume that each user requires a 22-bit message, i.e., D = 22. Moreover, we assume that Rinner = 250m and Router = 350m, i.e., the center of each group is randomly located in a doughnut shape of inner radius 250m and outer radius 350m. The probabilities of reliable communications and the averaged numbers of users that decode their messages successfully under the proposed scheme versus Benchmark Schemes 1 – 6 are given in Tables I and II, respectively. It is observed that our proposed scheme can achieve a probability of reliable communications above 99.99% (no outage is observed in 10, 000 channel realizations) in this setup, which is much higher than the benchmark schemes. Specifically, if leader selection is not considered in our proposed scheme, i.e., Benchmark Scheme 1, the probability of reliable communications is 0, although the averaged number of successful users is 32.241. It is worth noting that under the proposed scheme with leader selection, all the 8 groups have at least one leader after Phase I in all the 10, 000 channel realizations, while under Benchmark Scheme 1 without leader selection, on average.

TABLE I
PROBABILITY OF RELIABLE COMMUNICATIONS WHEN $D = 22$ BITS

Proposed Scheme	> 99.99%
Benchmark Scheme 1	0
Benchmark Scheme 2	0
Benchmark Scheme 3	0
Benchmark Scheme 4	11.60%
Benchmark Scheme 5	0
Benchmark Scheme 6	0

TABLE II
AVERAGED TOTAL NUMBER OF USERS THAT DECODE THEIR MESSAGES
SUCCESSFULLY WHEN $D = 22$ BITS

Proposed Scheme	48.0
Benchmark Scheme 1	32.2
Benchmark Scheme 2	0
Benchmark Scheme 3	0
Benchmark Scheme 4	45.1
Benchmark Scheme 5	0.014
Benchmark Scheme 6	11.5

only 4.2 groups have at least one leader. This verifies the effectiveness of the leader selection based beamforming design in our proposed scheme. Moreover, for the Occupy Cow protocol, no user can decode the message for both the cases without and with leader selection, i.e., Benchmark Schemes 2 and 3, since the SINR requirement for multicasting KD = 1056 bits to the users using $\tau 1B = 75$ symbols is too high (about 42dB). Finally, for the onephase transmission scheme, information broadcasting, i.e., Benchmark Scheme 4, can achieve reliable communications with a probability of 11.60%, while the probabilities of reliable communications achieved by TDMA and information multicasting, i.e., Benchmark Schemes 5 and 6, are both zero.

B How Many Bits can Be Transmitted Reliably?

In this example, we study the effect of transmission rate, which is determined by the size of message per user, i.e., D, the latency requirement 1ms and the bandwidth used, i.e., 100kHz, on the probability of reliable communications. This result is useful in practice because it indicates how much information can be reliably sent to all the users under each scheme. The simulation setup is the same as that of Tables I and II.

Fig. 3 shows the probabilities of reliable communications achieved by different schemes versus various values of D. Under our proposed scheme, when D ≤ 24 bits, outage is not observed in the 10, 000 channel realizations, while when D = 26 or 28 bits, 53 and 523 outages are observed over 10, 000 channel realizations, respectively. As a result, if the reliability requirement of the communications is 99.99%, i.e., no more than 1 outages in 10, 000 channel realizations, then we can transmit at most D = 24 bits to each user in this setup

Fig. 3.Probabilities of reliable communications versus values of D under different schemes. The center of each group is randomly located in a doughnut shape of inner radius Rinner = 250m and outer radius Router = 350m.

We remark that in Section III-B we provide an alternative transmission strategy for Phase II in which each leader is made aware of all the other leaders in its group so it can subtract all the leaders' messages and transmit a shorter packet. We ignore the overhead for the control signals and measure the reliability achieved by this scheme. It is observed that still at most D = 24 bits can be transmitted to each user if the reliability requirement of the communications is 99.99%. The reason is as follows. As observed in our simulation, when the number of information bits required by the users is D = 24 bits, there are only about 1 − 2 leaders among 8 users in each group after multi-group multicast beamforming in Phase I. In this case, the message size is not significantly reduced in Phase II even if the leaders are allowed to subtract the messages of all the leaders.

Moreover, it is observed that when D ≤ 18 bits, reliable communications can be achieved by information broadcasting as well. However, Occupy CoW Protocol cannot achieve reliable communications even when D = 12 bits. This is because when D = 12 bits, the minimum SINR requirement for multicasting KD = 576 bits to the users using τ1B = 75 symbols is 23.10dB, which is too high for the cell-edge users. In [24], it is shown that if the BS has one antenna, i.e., M = 1, many users' channels from the BS can suffer from deep fading at any time, and the two-phase Occupy Cow protocol is able to combat the fading and enhance the system reliability by increasing the diversity order in the second phase, if the rate requirement of each user is sufficiently small so that a large number of users can decode the message in the first phase.
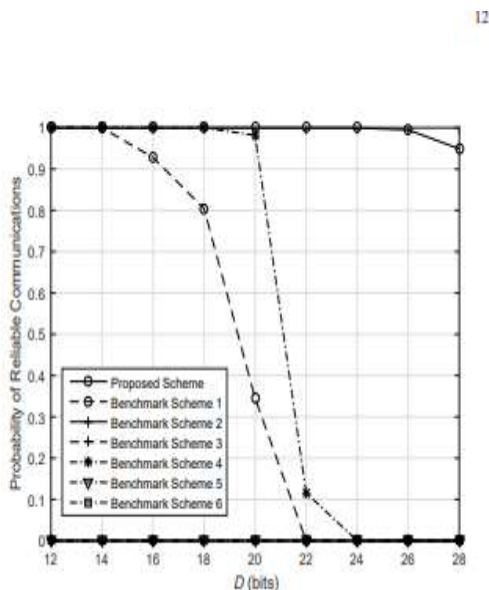
However, if the BS has multiple antennas, the effect of channel fading on the reliability of the information broadcasting is reduced thanks to the channel diversity. In this case, information broadcasting is more effective than Occupy CoW since combining all user
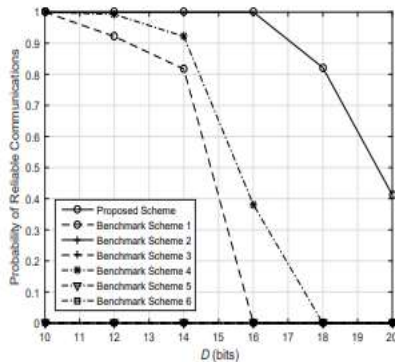

Fig. 4.Probabilities of reliable communications versus values of D under different schemes. The center of each group is randomly located in a doughnut shape of inner radius Rinner = 350m and outer radius Router = 450m.

messages together results in an infeasible rate for all the users under Occupy CoW, while broadcasting a very short message to each user would be feasible in many cases via beamforming

C Effect of User Topology on System Reliability

In this subsection, we study the effect of user topology on the reliability of our considered communication system. The simulation setup is the same as that of Fig. 3, except that the center of each group is randomly located in a doughnut shape of inner radius Rinner = 350m and outer radius Router = 450m, i.e., the users suffer from stronger intercell interference.

Fig. 4 shows the probabilities of reliable communications achieved by different

schemes versus values of D. By comparing Fig. 4 with Fig. 3, it is observed that since in this numerical example the users are located at the cell edge with much stronger inter-cell interference, the system reliability is significantly reduced. Specifically, if the reliability requirement of the communications is 99.99%, i.e., no more than 1 outages in 10, 000 channel realizations, then we can transmit at most D = 16 bits to each user under our proposed scheme, which is much lower than D = 24 bits in Fig. 3. This example shows that the system reliability heavily depends on the user distribution. Nevertheless, our scheme still outperforms the other benchmark schemes significantly in this numerical example.

VII. CONCLUSION AND FUTURE WORK

This paper proposes a novel two-phase transmission protocol to fully exploit the D2D transmission for URLLC. Under the proposed protocol, each group's messages are combined together and multicast to the leaders from the BS in the first phase, while the leaders relay the messages to the other users in their groups in the second phase. Since the D2D networks are reliable due to the strong channels between the users in

the same group, the challenge of our protocol is to select at least one leader for each group in the first phase via a proper beamforming design at the BS. Utilizing the sparse optimization technique, this paper proposes an efficient algorithm that jointly optimizes the beamforming at the BS and leader selection in each group. Simulation results show that the proposed algorithm results in a fair leader assignment among groups in the first phase, thus leading to reliable communications in all the groups in the second phase. Performance comparison to other existing schemes for URLLC is provided to show the effectiveness

of the proposed protocol with the leader selection based beamforming design. There are a number of directions along which the results of this paper can be further extended. For example, this paper assumes that the user grouping is already done at the network planning stage and known to the BS and users. Future work may study how to group the users based on their locations such that the system reliability can be maximized under our proposed downlink protocol. Moreover, this paper focuses on URLLC in the downlink. Future work may investigate efficient protocols for achieving URLLC in the uplink.

APPENDIX

A. Proof of Proposition 1

First, it can be shown that in the lth iteration of Algorithm 1, the solution obtained in the $(l-1)$th iteration is also feasible to problem (22) given $w_n^{(l)} = w_n^{(1,l-1)}$ (thus $\hat{a}_n = a_n^{(l-1)}$ and $b_n = b_n^{(l-1)}$) and $t_n^{(l)} = t_n^{(1,l-1)}$, $\forall n$. In other words, $\sum_{n=1}^N \|t_n^{(l-1)}\|_1 + f(\{t_{k,n}^{(l-1)}, \tilde{t}_{k,n}^{(l-1)}\})$ is achievable to problem (22) in the lth iteration. As a result, the optimal objective value of problem (22) in the lth iteration is no larger than that achieved by the solution $w_n^{(1,l-1)}$'s and $t_n^{(1,l-1)}$'s, i.e.,

$$\sum_{n=1}^N \|t_n^{(1,l)}\|_1 + f(\{t_{k,n}^{(1,l)}, \tilde{t}_{k,n}^{(1,l-1)}\}) \leq$$
$$\sum_{n=1}^N \|t_n^{(1,l-1)}\|_1 + f(\{t_{k,n}^{(1,l-1)}, \tilde{t}_{k,n}^{(1,l-1)}\}). \quad (47)$$

It is worth noting that $f(\{t_{k,n}^{(1,l)}, t_{k,n}^{(l-1)}\}) = \sum_{n=1}^N \beta_n \sqrt[K_n]{\prod_{k=1}^{K_n} t_{k,n}^{(1,l)}}$ according to (16). It is also worth noting that since $\sum_{n=1}^N \beta_n \sqrt[K_n]{\prod_{k=1}^{K_n} t_{k,n}^{(1)}}$ is a concave function, we have

$$f(\{t_{k,n}^{(1,l)}, \tilde{t}_{k,n}^{(1,l-1)}\}) \geq \sum_{n=1}^N \beta_n \sqrt[K_n]{\prod_{k=1}^{K_n} t_{k,n}^{(1,l)}} \quad (48)$$

Then, it follows that

$$\sum_{n=1}^N \|t_n^{(1,l)}\|_1 + \sum_{n=1}^N \beta_n \sqrt[K_n]{\prod_{k=1}^{K_n} t_{k,n}^{(1,l)}}$$
$$\leq \sum_{n=1}^N \|t_n^{(1,l)}\|_1 + f(\{t_{k,n}^{(1,l)}, \tilde{t}_{k,n}^{(1,l-1)}\})$$
$$\leq \sum_{n=1}^N \|t_n^{(1,l-1)}\|_1 + f(\{t_{k,n}^{(1,l-1)}, \tilde{t}_{k,n}^{(1,l-1)}\})$$
$$= \sum_{n=1}^N \|t_n^{(1,l-1)}\|_1 + \sum_{n=1}^N \beta_n \sqrt[K_n]{\prod_{k=1}^{K_n} t_{k,n}^{(1,l-1)}}. \quad (49)$$

REFERENCES

[1] 3GPP TR 138 913, "5G: Study on scenarios and requirements for next generation access technologies (Release 14)".

[2] T. L. Marzetta, "Noncooperative cellular wireless with unlimited numbers of base station antennas," IEEE Trans. Wireless Commun., vol. 9, no. 11, pp. 3590-3660, Nov. 2010.

[3] E. G. Larsson, F. Tufvesson, O. Edfors, and T. L. Marzetta, "Massive MIMO for next generation wireless systems," IEEE Commun. Mag., vol. 52, no. 2, pp. 186-195, Feb. 2014.

[4] W. Roh et al., "Millimeter-wave beamforming as an enabling technology for 5G cellular communications: Theoretical feasibility and prototype results," IEEE Commun. Mag., vol. 52, no. 2, pp. 106-113, Feb. 2014.

[5] O. Simeone, A. Maeder, M. Peng, O. Sahin, and W. Yu, "Cloud radio access network: Virtualizing wireless access for dense heterogeneous systems," J. Commun. and Networks, vol. 18, no. 2, pp. 135-149, Apr. 2016.

[6] L. Liu and W. Yu, "Massive connectivity with massive MIMO-Part I: Device activity detection and channel estimation," IEEE Trans. Signal Process., vol. 66, no. 11, pp. 2933-2946, Jun. 2018.

[7] L. Liu and W. Yu, "Massive connectivity with massive MIMO-Part II: Achievable Rate

Characterization," IEEE Trans. Signal Process., vol. 66, no. 11, pp. 2947-2959, Jun. 2018.

[8] E. Bj ¨ornson, E. de Carvalho, J. H. Sørensen, E. G. Larsson, and P. Popovski, "A random access protocol for pilot allocation in crowded massive MIMO systems," IEEE Trans. Wireless Commun., vol. 16, no. 4, pp. 2220-2234, Apr. 2017.

[9] A. A. Kumar Somappa, K. Øvsthus, and L. M. Kristensen, "An industrial perspective on wireless sensor networks-A survey of requirements, protocols, and challenges," IEEE Commun. Surveys Tuts., vol. 16, no. 3, pp. 1391-1412, 3rd Quart., 2014.

[10] A. Frotzscher, U. Wetzker, M. Bauer, M. Rentschler, M. Beyer, S. Elspass, H. Klessig, "Requirements and current solutions of wireless communication in industrial automation," in Proc. IEEE Int. Workshop 5G Technol. (5GWS'14), 2014.

[11] A. Varghese and D. Tandur, "Wireless requirements and challenges in Industry 4.0," in Proc. Int. Conf. ComtemporaryComput.Informat., Nov. 2015, pp. 634-638.

[12] Q. Wang and J. Jiang, "Comparative examination on architecture and protocol of industrial wireless sensor network standards," IEEE Commun. Surveys Tuts., vol. 18, no. 3, pp. 2197-2219, 3rd Quart., 2016.

[13] C. Lu, A. Saifullah, B. Li, M. Sha, H. Gonzalez, D. Gunatilaka, C. Wu, L. Nie, and Y. Chen, "Real-time wireless sensor-actuator networks for industrial cyber-physical systems," Proceedings of the IEEE, vol. 104, no. 5, pp. 1013-1024, May 2016.

[14] E. Karipidis, N. D. Sidiropoulos, and Z. Q. Luo, "Quality of service and max-min fair transmit beamforming to multiple cochannel multicast groups," IEEE Trans. Signal Process., vol. 56, no. 8, pp. 1268-1279, Mar. 2008.

[15] J. Liu, N. Kato, J. Ma, and N. Kadowaki, "Device-to-device communication in LTE-advanced networks: A survey," IEEE Commun. Surveys Tuts., vol. 17, no. 4, pp. 1923-1940, 4th Quart., 2015.

[16] P. Popovski, "Ultra-reliable communication in 5G wireless systems," in Proc. IEEE Int. Conf. 5G for Ubiquitous Connectivity (5GU), Nov. 2014, pp. 146-151.

[17] G. Durisi, T. Koch, and P. Popovski, "Toward massive, ultrareliable, and low-latency wireless communication with short packets," Proc. IEEE, vol. 104, no. 9, pp. 1711-1726, Sep. 2016.

[18] N. A. Johansson, Y.-P. E. Wang, E. Eriksson, and M. Hessler, "Radio access for ultra-reliable and low-latency 5G communications," in Proc. IEEE Int. Conf. Commun. (ICC), London, UK., Jun. 2015, pp. 1184- 1189.

[19] H. Shariatmadari, S. Iraji, and R. Jantti, "Analysis of transmission methods for ultra-reliable communications," in Proc. IEEE Pers. Indoor and Mobile Radio Commun. (PIMRC), 2015, Hong Kong, China, pp. 1126-1131.

[20] D. Ohmann, M. Simsek, and G. Fettweis, "Achieving high availability in wireless networks by an optimal number of Rayleigh-fading links," in Proc. IEEE Globecom Workshops, Dec. 2014, Austin, USA, pp. 1402-1407.

[21] N. Brahmi, O. N. C. Yilmaz, K.W. Helmersson, S. A. Ashraf, J. Torsner, "Deployment strategies for ultra-reliable and low-latency communication in factory automation," in Proc. IEEE Globecom Workshops, San Diego, USA, Dec. 2015.

[22] H. Shariatmadari, Z. Li, M. A. Uusitalo, S. Iraji, and R. J¨antti, "Link adaptation design for ultra-reliable communications," in Proc. IEEE Int. Conf. Commun. (ICC), Kuala Lumpur, Malaysia, 2016.

[23] O. N. C. Yilmaz, Y.-P. E. Wang, N. A. Johansson, N. Brahmi, S. A. Ashraf, and J. Sachs, "Analysis of ultra-reliable and low-latency 5G communication for a factory automation use case," in Proc. IEEE Int. Conf. Commun. (ICC) Workshop, London, UK., Jun. 2015, pp. 1190-1195.

[24] V. N. Swamy et al.,"Real-time cooperative communication for automation over wireless," IEEE Trans. Wireless Commun., vol. 16, no. 11, pp. 7168-7183, Nov 2017.

[25] I. Mitliagkas, N. D. Sidiropoulos, and A. Swami, "Joint power and admission control for ad-hoc and cognitive underlay networks: Convex approximation and distributed implementation," IEEE Trans. Wireless Commun., vol. 10, no. 12, pp. 4110-4121, Dec. 2011.

[26] Y.-F. Liu, Y.-H.Dai, and Z.-Q. Luo, "Joint power and admission control via linear programming deflation," IEEE Trans. Signal Process., vol. 61, no. 6, pp. 1327-1338, Mar. 2013.

[27] G. J. Foschini and Z. Miljanic, "A simple distributed autonomous power control algorithm and its convergence," IEEE Trans.

Veh. Technol., vol. 42, no. 4, pp. 641-646, Nov. 1993.

[28] Y. Polyanskiy, H. V. Poor, and S. Verd ´u, "Channel coding rate in the finite blocklength regime," IEEE Trans. Inf. Theory, vol. 56, no. 5, pp. 2307-2359, May 2010.

[29] W. Yang, G. Durisi, T. Koch, and Y. Polyanskiy, "Quasi-static multipleantenna fading channels at finite blocklength," IEEE Trans. Inf. Theory, vol. 60, no. 7, pp. 4232-4265, Jun. 2014.

[30] S. Boyd and L. Vandenberghe, Convex Optimization, Cambidge Univ. Press, 2004.

[31] N. D. Sidiropoulos, T. N. Davidson, and Z.-Q. Luo, "Transmit beamforming for physical-layer multicasting," IEEE Trans. Signal Process., vol. 54, no. 6, pp. 2239-2251, Jun. 2006.

[32] L. N. Tran, M. F. Hanif, and M. Juntti, "A conic quadratic programming approach to physical layer multicasting for large-scale antenna arrays," IEEE Signal Process. Lett., vol. 21, no. 1, pp. 114-117, Jan. 2014.

[33] C. Lu and Y.-F. Liu, "An efficient global algorithm for single-group multicast beamforming," IEEE Trans. Signal Process., vol. 65, no. 14, pp. 3761-3774, July 2017.

[34] M. Grant and S. Boyd, CVX: Matlab software for disciplined convex programming, version 1.21, http://cvxr.com/cvx/ Apr. 2011. 15

[35] A. Wiesel, Y. C. Eldar, and S. Shamai, "Linear precoding via conic optimization for fixed MIMO receivers," IEEE Trans. Signal Process. , vol. 54, no. 3, pp. 161-176, January 2006.

[36] IST-4-027756 WINNER II D1.1.2 V1.2 WINNER II Channel Models, 2008. [37] B. R. Marks and G. P. Wright, "A general inner approximation algorithm for nonconvex mathematical programs," Operations Research, vol. 26, no. 4, pp. 681-683, 1978