IEEE Transactions on Wireless Communications Volume: 18 , Issue: 6 , June 2019

Dynamic Transmission Policy for Multi-Pair Cooperative Device-to-Device Communication with Block-Diagonalization Precoding

Yung-Shun Wang, *Student Member, IEEE,* Y.-W. Peter Hong, *Senior Member, IEEE,* and Wen-Tsuen Chen, *Fellow, IEEE*

***Abstract*—This work proposes a dynamic precoding and power allocation policy for mutually cooperative device-to-device (D2D) transmitter-receiver pairs that underlay a cellular system in the uplink. The cooperative transmission consists of two phases: a *data-sharing phase* (i.e., phase 1) and a *joint transmission phase* (i.e, phase 2). Multicast precoders are used in phase 1 and coordinated block-diagonalization precoders are considered in phase 2. The precoders are jointly designed to maximize the long-term utility of the D2D users subject to long-term individual power and rate-gain constraints, and an instantaneous interference constraint at the base-station. The long-term objective and constraints allow cooperating users to adapt their resources more flexibly over time, but increases the complexity of the design. By adopting the Lyapunov optimization framework and by constructing virtual queues to record the temporal evolution of the system states, the long-term utility maximization problem can be decoupled into a series of short-term *weighted-rate-minus-energy-penalty (WRMEP)* optimization problems that can be solved efficiently. A low-complexity algorithm is further proposed for solving the WRMEP problem when multicasting in the data-sharing phase is performed by a spatially white input. Theoretical performance guarantees and a bound on the virtual queue backlogs are also derived.**

I. Introduction

Device-to-device (D2D) communication has been proposed as a promising technique for improving the spectrum utilization of next generation cellular systems [1], [2]. This is done by allowing nearby devices to communicate directly with each other without relaying information through the cellular base-station (BS). As the number of user devices increases, the amount of data that need to be offloaded to D2D transmissions will inevitably increase and, thus, the system must be able to accommodate a larger number of simultaneously active D2D links. However, without cooperation among these D2D pairs, the overall system performance will eventually be limited by the interference that they cause to each other as well as to the cellular system.

In the literature on D2D communications, interference management [3], resource allocation [4], power control [5], [6], and transmission mode selection [7], [8] methods were proposed to reduce the interference in D2D underlaid cellular systems. In [3]–[8], only one D2D pair was assumed to be active at a time or orthogonal resources were allocated to difference D2D pairs, and, thus, the focus was on reducing the interference between D2D and cellular transmissions. More recently, in [9]–[11], cases with multiple transmitting D2D pairs were examined with further consideration on the inter-pair interference. Specifically, in [9] and [10], both the inter-pair interference and the interference towards the cellular transmission were considered by imposing a transmit power constraint on the D2D transmitters; in [11], the issues of interference and energy-awareness were analyzed for ultra-dense D2D networks using a mean field game framework. However, these works do not exploit the advantages of cooperation. With cooperation, users will be able to better reduce interference and increase spatial spectrum utilization and, thus, allow more D2D pairs to be simultaneously active.

User cooperation was studied extensively in wireless communications since the seminal works in [12]–[14]. Many different scenarios were considered depending on the number of sources, relays, and destinations, and also on whether or not there exists dedicated relays (or sources temporarily acting as relays to help forward the information of others) [15]. Most of these works considered the use of a two-phase transmission scheme, where the source(s) first transmits information to the relay(s) in phase 1, and then the relay(s) forwards the information to the destination(s) in phase 2. The works that are most related to ours are those that consider multiple sources and multiple destinations, such as those in [16]–[24]. In [16]–[21], sources and destinations were served by dedicated relays that do not have their own data to transmit and, thus, allocate all of their resources to the forwarding of signals from the sources to the destinations. In this case, the relays can form a distributed antenna array and adopt distributed beamforming or space-time coding schemes to exploit the available spatial degrees of freedom. Moreover, in [22]–[24], sources were instead mutually cooperative and took turns acting temporarily as relays for one another. No dedicated relays were assumed to exist in these works. Space-time coding was adopted in these works to exploit the cooperative diversity gains. Furthermore, in the context of D2D networks, studies on user cooperation were examined more recently in [25]–[29], where interference against the cellular system was further taken into consideration. However, all of the above cooperative transmission schemes assume instantaneous (or short-term) power constraints, which limits the adaptability of the system over time, and do not consider fairness or performance guarantees among sources.

The authors are with the Institute of Communications Engineering, National Tsing Hua University, Hsinchu, Taiwan (Emails: {yswang@erdos.ee; ywhong@ee; wtchen@cs}.nthu.edu.tw). Y.-W. P. Hong is also with MOST Joint Research Center for AI Technology and AII Vista Healthcare.

This work was supported in part by Ministry of Science and Technology, Taiwan, under grant MOST 107-2634-F-007-005.

The main objective of this work is to propose a cooperative transmission scheme for D2D users that not only takes into consideration short-term interference temperature constraints at the BS, but also long-term power constraints and cooperative performance guarantees. Cooperation allows multiple D2D pairs to transmit simultaneously in the uplink frame of the cellular system without causing significant interference to each other as well as to the BS. Here, we adopt a two-phase cooperative transmission scheme that consists of a *data-sharing transmission* in phase 1 and a *cooperative joint transmission* in phase 2. Specifically, in phase 1, D2D transmitters (DTs) take turns broadcasting their data to other DTs using physical layer multicasting techniques. In phase 2, the DTs cooperatively transmit their data to their respective D2D receivers (DR) using coordinated multi-user precoding. Here, we assume that the DTs are located close to each other (e.g., within a certain hot spot [29]–[32]) so that the data-sharing in phase 1 can be done efficiently, and adopt the coordinated block-diagonalization (BD) precoding scheme [33]–[35] in phase 2 to ensure that inter-pair interference is eliminated.

The proposed cooperative D2D transmission scheme is formulated as a long-term precoder design problem, where the objective is to maximize the long-term utility subject to long-term individual power and rate-gain constraints as well as instantaneous constraints on the interference power at BS. This is different from most works in the literature that consider only short-term constraints on the power and the target rate. Long-term constraints allow cooperative users to allocate their resources more flexibly over time and among users. That is, some users may be favored over other users at a certain time, but not at others. In particular, the long-term individual power constraint is used to limit the transmit power of each device; the long-term rate-gain constraint is used to ensure that each D2D pair achieves a larger rate through cooperation; and the short-term interference constraint limits the interference that the D2D transmission may cause on the cellular system. By adopting the Lyapunov optimization framework [36]–[43], we are able to decouple the long-term design problem into a series of short-term subproblems, one for each time slot. To do so, we construct virtual data, rate-gain, and energy queues to record the system states and to enable separate optimization in each time slot based on these states. The effectiveness of the Lyapunov optimization framework [36], [37] in solving various issues in wireless networks was demonstrated in studies on, e.g., energy efficiency [38]–[40], downlink scheduling [41], and energy management in energy harvesting devices [42], [43]. In particular, given the queue states and the channel state information (CSI) of each time slot, we propose the maximum-weighted-rate-minus-energy-penalty (Max-WRMEP) policy to determine the BD cooperative precoders in each time slot. A low-complexity implementation is then proposed for the case where spatially white input is used for the multicasting in phase 1. Theoretical performance guarantees and a bound on the virtual queue backlogs are also derived. Our main contributions can be summarized as follows:

- the modelling of long-term optimization problems for cooperative networks, which provides better flexibility for resource-sharing over time and among users;
- the decoupling of the long-term optimization problem into practical short-term subproblems that can be solved efficiently;
- the proposition of a low-complexity scheme that utilizes spatially white input in phase 1.

The remainder of this paper is organized as follows. In Section II, we introduce the system model and the problem formulation. In Section III, we propose the Max-WRMEP policy that converts the long-term optimization problem into a series of short-term subproblems that depend only on the queue and channel states at that time. A low-complexity alternative is then proposed in Section IV. Finally, we provide numerical simulations in Section V to demonstrate the effectiveness of the proposed scheme, and conclude in Section VI.

II. System Model

Let us consider a cooperative D2D network that consists of K multi-antenna D2D transmitter and receiver pairs, denoted by the index set $\mathcal{K} = \{1, 2, \dots, K\}$, as illustrated in Fig. 1. The transmitter and the receiver of the k-th D2D pair is denoted by DT k and DR k, respectively. The number of transmit and receive antennas at each DT and DR are N_t and N_r, respectively.[1] Moreover, we assume that there exists a concurrent uplink transmission between a cellular user (CU) and the BS. The D2D and cellular transmissions are simultaneously active and, thus, may cause interference to each other. The number of CU and BS antennas are N_c and N_b, respectively.

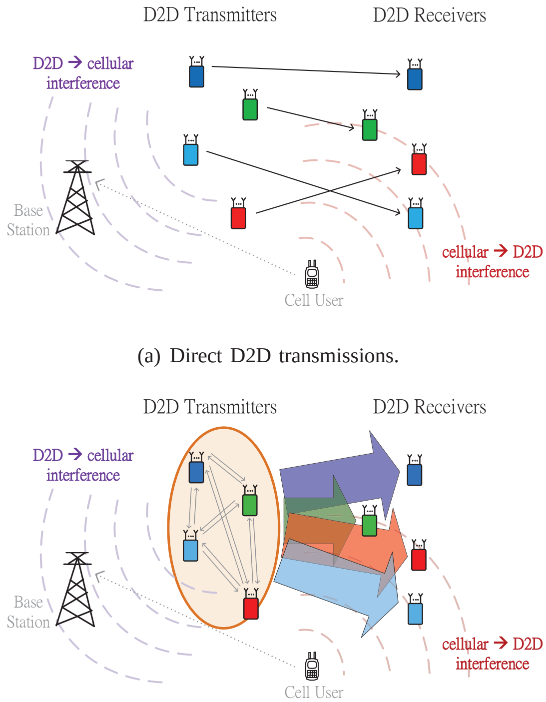
Let us consider a time-slotted system with slot duration normalized to 1. The transmission in each time slot is divided into two phases: a *data-sharing* phase (i.e., phase 1) and a *joint transmission* phase (i.e., phase 2). In phase 1, the transmitters take turns broadcasting their information to all other transmitters and, in phase 2, the transmitters together form a distributed antenna array to cooperatively transmit their data to the respective DRs. The portion of time allocated to each user in phase 1 is η_1 and that allocated to the joint transmission in phase 2 is η_2. Hence, we have $K\eta_1 + \eta_2 = 1$. Here, we assume that all D2D pairs participate in the cooperative transmission in each time slot. In particular, to ensure that cooperation is advantageous, we assume that the DTs are located close to each other, forming a cluster within a certain hot spot, so that data-sharing in phase 1 can be done efficiently. Similar settings have also been considered in [29]–[32]. The problem of determining who can or should join the cooperation (i.e., issues of admission control and partner selection) can both be studied on top of our proposed cooperative transmission scheme, but are beyond the scope of this paper.

A. Two Phase Cooperation and Constraints

Specifically, in phase 1 of time slot t, DT k transmits the signal

$$\mathbf{x}_k^{(1)}[t] = \mathbf{W}_k^{(1)}[t]\mathbf{s}_k^{(1)}[t] \tag{1}$$

[1]The number of transmit and receive antennas need not be the same for all D2D pairs, but is assumed to be the same here for ease of exposition.



(a) Direct D2D transmissions.

(b) Cooperative D2D transmissions.

Fig. 1. Illustration of D2D networks without and with cooperation.

to all other DTs for data sharing, where $\mathbf{s}_k^{(1)}[t] \in \mathbb{C}^{N_t \times 1}$ is the information-bearing signal and $\mathbf{W}_k^{(1)}[t] \in \mathbb{C}^{N_t \times N_t}$ is the multicast precoding matrix. We assume that the Gaussian codebook is used for transmission [13], [35], [41], and thus, the entries of $\mathbf{s}_k^{(1)}[t]$ are independent and identically distributed (i.i.d.) Gaussian with zero mean and unit variance, i.e., $\mathbf{s}_k^{(1)}[t] \sim \mathcal{CN}(\mathbf{0}, \mathbf{I}_{N_t})$.[2] The received signal at DT ℓ during the multicast transmission by DT k, for $k \neq \ell$, is

$$\mathbf{y}_{k,\ell}^{(1)}[t] = \mathbf{G}_{k,\ell}[t]\mathbf{W}_k^{(1)}[t]\mathbf{s}_k^{(1)}[t] + \mathbf{G}_{c,\ell}[t]\mathbf{x}_c^{(1)}[t] + \mathbf{n}_\ell^{(1)}[t], \quad (2)$$

where $\mathbf{x}_c^{(1)}[t] \sim \mathcal{CN}(\mathbf{0}, \mathbf{Q}_c[t])$ is the signal transmitted by CU during phase 1, $\mathbf{G}_{k,\ell}[t] \in \mathbb{C}^{N_t \times N_t}$ is the channel matrix between DT k and DT ℓ, $\mathbf{G}_{c,\ell}[t] \in \mathbb{C}^{N_t \times N_c}$ is the channel matrix between CU and DT ℓ, and $\mathbf{n}_\ell^{(1)}[t] \sim \mathcal{CN}(0, \sigma_n^2 \mathbf{I}_{N_t})$ is the additive white Gaussian noise (AWGN). The interference plus noise covariance matrix at DT ℓ is $\mathbf{G}_{c,\ell}[t]\mathbf{Q}_c[t]\mathbf{G}_{c,\ell}^H[t] + \sigma_n^2 \mathbf{I}_{N_t}$[3]. By choosing the noise whitening matrix as $\Upsilon_\ell^{(1)}[t] \in \mathbb{C}^{N_t \times N_t}$ such that $((\Upsilon_\ell^{(1)}[t])^H \Upsilon_\ell^{(1)}[t])^{-1} = \mathbf{G}_{c,\ell}[t]\mathbf{Q}_c[t]\mathbf{G}_{c,\ell}^H[t] + \sigma_n^2 \mathbf{I}_{N_t}$ [44], the equivalent received signal at DT ℓ after noise-whitening is

$$\tilde{\mathbf{y}}_{k,\ell}^{(1)}[t] = \Upsilon_\ell^{(1)}[t]\mathbf{y}_{k,\ell}^{(1)}[t] = \tilde{\mathbf{G}}_{k,\ell}[t]\mathbf{W}_k^{(1)}[t]\mathbf{s}_k^{(1)}[t] + \tilde{\mathbf{n}}_\ell^{(1)}[t],$$

[2]The Shannon capacity formulae in (3) and (9) are based on the use of the Gaussian codebook. The code rates achievable under this assumption can serve as an upper bound to that of practical systems where finite alphabet codebooks are often adopted.

[3]To compute the noise whitening matrix, it is necessary for DT ℓ to obtain knowledge of the channel matrix between itself and CU (and, thus, the interference plus noise covariance matrix). This channel matrix can be estimated by having DT ℓ overhear the pilot signal emitted by CU during its transmission to the BS. The noise whitening matrix at DR k can be computed similarly.

where $\tilde{\mathbf{G}}_{k,\ell}[t] \triangleq \Upsilon_\ell^{(1)}[t]\mathbf{G}_{k,\ell}[t]$ is the effective channel matrix and $\tilde{\mathbf{n}}_\ell^{(1)}[t] \triangleq \Upsilon_\ell^{(1)}[t](\mathbf{G}_{c,\ell}[t]\mathbf{x}_c^{(1)}[t] + \mathbf{n}_\ell^{(1)}[t]) \sim \mathcal{CN}(\mathbf{0}, \mathbf{I}_{N_t})$ is the effective noise. Therefore, to ensure that all other DTs can successfully decode in phase 1, the transmission rate $R_k[t]$ of DT k must satisfy

$$R_k[t] \leq \min_{\ell \neq k} \eta_1 \log_2 \left| \mathbf{I}_{N_t} + \tilde{\mathbf{G}}_{k,\ell}[t]\mathbf{Q}_k^{(1)}[t]\tilde{\mathbf{G}}_{k,\ell}^H[t] \right|, \quad (3)$$

where $\mathbf{Q}_k^{(1)}[t] \triangleq \mathbf{W}_k^{(1)}[t](\mathbf{W}_k^{(1)}[t])^H$ is the input covariance matrix of DT k in phase 1.

In phase 2, DTs transmit cooperatively to their respective DRs using joint multiuser transmit precoding. Let $\mathbf{s}_k^{(2)}[t] \in \mathbb{C}^{M \times 1}$ be the data signal intended for DR k, where M is the number of transmitted data streams due to spatial multiplexing, and let $\mathbf{W}_{\ell,k}^{(2)}[t] \in \mathbb{C}^{N_t \times M}$ be the precoding matrix employed by DT ℓ to transmit $\mathbf{s}_k^{(2)}[t]$. By collecting the precoding matrices from all DTs into a joint precoding matrix $\mathbf{W}_k^{(2)}[t] = [\mathbf{W}_{1,k}^{(2)}[t]^H, \ldots, \mathbf{W}_{K,k}^{(2)}[t]^H]^H$, $\forall k$, and by letting $\mathbf{H}_k[t] = [\mathbf{H}_{1,k}[t], \ldots, \mathbf{H}_{K,k}[t]]$, where $\mathbf{H}_{\ell,k}[t] \in \mathbb{C}^{N_r \times N_t}$ is the channel matrix between DT ℓ and DR k, the received signal at DR k can be written as

$$\mathbf{y}_k^{(2)}[t] = \mathbf{H}_k[t]\mathbf{W}_k^{(2)}[t]\mathbf{s}_k^{(2)}[t] + \sum_{\ell \neq k} \mathbf{H}_k[t]\mathbf{W}_\ell^{(2)}[t]\mathbf{s}_\ell^{(2)}[t] + \mathbf{H}_{c,k}[t]\mathbf{x}_c^{(2)}[t] + \mathbf{n}_k^{(2)}[t], \quad (4)$$

where $\mathbf{H}_{c,k}[t] \in \mathbb{C}^{N_r \times N_c}$ is the channel matrix between CU and DR k, $\mathbf{x}_c^{(2)}[t] \sim \mathcal{CN}(\mathbf{0}, \mathbf{Q}_c[t])$ is the signal transmitted by CU during phase 2, and $\mathbf{n}_k^{(2)}[t] \sim \mathcal{CN}(\mathbf{0}, \sigma_n^2 \mathbf{I}_{N_r})$ is the AWGN. Similarly, by choosing the noise whitening matrix as $\Upsilon_k^{(2)}[t] \in \mathbb{C}^{N_r \times N_r}$ such that $((\Upsilon_k^{(2)}[t])^H \Upsilon_k^{(2)}[t])^{-1} = \mathbf{H}_{c,k}[t]\mathbf{Q}_c[t]\mathbf{H}_{c,k}^H[t] + \sigma_n^2 \mathbf{I}_{N_r}$, the equivalent received signal at DR k after noise-whitening is

$$\begin{aligned} \tilde{\mathbf{y}}_k^{(2)}[t] &= \Upsilon_k^{(2)}[t]\mathbf{y}_k^{(2)}[t] \quad (5) \\ &= \tilde{\mathbf{H}}_k[t]\mathbf{W}_k^{(2)}[t]\mathbf{s}_k^{(2)}[t] + \sum_{\ell \neq k} \tilde{\mathbf{H}}_k[t]\mathbf{W}_\ell^{(2)}[t]\mathbf{s}_\ell^{(2)}[t] + \tilde{\mathbf{n}}_k^{(2)}[t] \quad (6) \end{aligned}$$

where $\tilde{\mathbf{H}}_k[t] \triangleq \Upsilon_k^{(2)}[t]\mathbf{H}_k[t]$ is the effective channel matrix and $\tilde{\mathbf{n}}_k^{(2)}[t] \triangleq \Upsilon_k^{(2)}[t](\mathbf{H}_{c,k}[t]\mathbf{x}_c^{(2)}[t] + \mathbf{n}_k^{(2)}[t]) \sim \mathcal{CN}(\mathbf{0}, \mathbf{I}_{N_r})$ is the effective noise. We assume throughout this work that all channel gains are bounded.

To eliminate interference among D2D pairs, we adopt the block diagonalization (BD) precoding technique [34], [35], where the precoding matrices $\mathbf{W}_k^{(2)}[t]$, for $k = 1, \ldots, K$, are chosen such that

$$\tilde{\mathbf{H}}_\ell[t]\mathbf{W}_k^{(2)}[t] = \mathbf{0}, \ \forall \ell \neq k. \quad (7)$$

By letting $\mathbf{Q}_k^{(2)}[t] \triangleq \mathbf{W}_k^{(2)}[t](\mathbf{W}_k^{(2)}[t])^H$ be the $KN_t \times KN_t$ joint covariance matrix for the signal intended for DR k, the constraint in (7) can be written equivalently as

$$\tilde{\mathbf{H}}_\ell[t]\mathbf{Q}_k^{(2)}[t]\tilde{\mathbf{H}}_\ell^H[t] = \mathbf{0}, \ \forall \ell \neq k. \quad (8)$$

Notice that it is necessary to have $N_t \geq N_r$ in order for the above condition to hold. To ensure that the joint transmission from the DTs to DR k in phase 2 is successful, the transmission rate $R_k[t]$ must also satisfy

$$R_k[t] \leq \eta_2 \log_2 \left| \mathbf{I}_{N_r} + \tilde{\mathbf{H}}_k[t]\mathbf{Q}_k^{(2)}[t]\tilde{\mathbf{H}}_k^H[t] \right|. \tag{9}$$

Consequently, by (3) and (9), the transmission rate $R_k[t]$ of the k-th D2D pair in time-slot t must satisfy

$$R_k[t] \leq \min \Big\{ \min_{\ell \neq k} \eta_1 \log_2 \left| \mathbf{I}_{N_t} + \tilde{\mathbf{G}}_{k,\ell}[t]\mathbf{Q}_k^{(1)}[t]\tilde{\mathbf{G}}_{k,\ell}^H[t] \right|, \eta_2 \log_2 \left| \mathbf{I}_{N_r} + \tilde{\mathbf{H}}_k[t]\mathbf{Q}_k^{(2)}[t]\tilde{\mathbf{H}}_k^H[t] \right| \Big\}. \tag{10}$$

Furthermore, to limit the interference that each D2D transmission may cause on the cellular system, we further impose an instantaneous *interference temperature (IT)* constraint on the expected interference power from each D2D pair. By assuming that the DTs know only the expected channel statistics towards BS, the IT constraints are given by [45]

$$\begin{aligned} &\mathbb{E}\big[\eta_1 \mathrm{tr}(\mathbf{G}_{k,b}[t]\mathbf{Q}_k^{(1)}[t]\mathbf{G}_{k,b}^H[t]) + \eta_2 \mathrm{tr}(\mathbf{G}_b[t]\mathbf{Q}_k^{(2)}[t]\mathbf{G}_b^H[t])\big] \\ &= \eta_1 \mathrm{tr}(\mathbf{C}_{\mathbf{G}_{k,b}}\mathbf{Q}_k^{(1)}[t]) + \eta_2 \mathrm{tr}(\mathbf{C}_{\mathbf{G}_b}\mathbf{Q}_k^{(2)}[t]) \leq \mathrm{IT}_k, \end{aligned} \tag{11}$$

for all k, where $\mathbf{G}_{k,b}[t] \in \mathbb{C}^{N_b \times N_t}$ is the channel matrix between DT k and the BS, $\mathbf{G}_b[t] \triangleq [\mathbf{G}_{1,b}[t], \ldots, \mathbf{G}_{K,b}[t]]$, and $\mathbf{C}_{G_{k,b}} \triangleq \mathbb{E}[\mathbf{G}_{k,b}^H[t]\mathbf{G}_{k,b}[t]]$ and $\mathbf{C}_{\mathbf{G}_b} \triangleq \mathbb{E}[\mathbf{G}_b^H[t]\mathbf{G}_b[t]]$ are the channel covariance matrices known at the DTs. Here, we assume that $\mathbf{C}_{\mathbf{G}_{k,b}}$ and $\mathbf{C}_{\mathbf{G}_b}$ are full rank, which occurs when $N_b \geq KN_t$ and the channel coefficients are independent with positive variances. In this case, the IT constraint in (11) imposes a constraint on the maximum transmission rate in each phase. A feasible BD cooperative precoding scheme for the k-th D2D pair refers to $\{\mathbf{Q}_k^{(1)}[t], \mathbf{Q}_k^{(2)}[t], R_k[t]\}$ that satisfies $\mathbf{Q}_k^{(1)}[t] \succeq 0$, $\mathbf{Q}_k^{(2)}[t] \succeq 0$, $R_k[t] \geq 0$ as well as (8), (10), and (11). Let us denote the set of all feasible BD cooperative precoding scheme for the k-th D2D pair by $\Gamma(\mathcal{H}[t], \mathrm{IT}_k)$, where $\mathcal{H}[t]$ is the set of all channel states at time t.

It is worthwhile to note that cooperation may not always be advantageous since additional overhead is required for DTs to share their own data with other DTs in phase 1. This holds for almost all cooperative transmission schemes [12]–[15] even in the two-user case. However, without cooperation, each D2D pair is allocated only $1/K$ portion of time for direct transmission between the corresponding DT and DR. Hence, cooperation can be beneficial if the DTs are sufficiently close to each other such that η_1 (i.e., the data-sharing overhead) can be made small, in which case, $\eta_2 = 1 - K\eta_1$ (i.e., the portion of time used to jointly serve all DRs in the cooperative case) can be greater than $1/K$. In this work, we assume that the issues of admission control and partner selection have already been resolved, and focus on the design of cooperative transmission schemes that can fully exploit the available cooperative advantages. Here, we assume that perfect CSI is available at all nodes for precoding and receiver processing, but later examine the impact of CSI acquisition and channel estimation imperfections through simulations.

B. Problem Formulation

When cooperation is employed to improve the overall system performance, D2D pairs expend their resources to aid the transmission of others (as well as serve their own transmissions). Hence, it is often difficult to ensure that all D2D pairs gain instantaneously in each time slot due to cooperation (as compared to transmitting non-cooperatively by themselves). Here, we instead determine a BD-based transmission policy (i.e., a sequence of BD cooperative precoding schemes $\{\mathbf{Q}_k^{(1)}[t], \mathbf{Q}_k^{(2)}[t], R_k[t], \forall k\}_{t=1}^{\infty}$ over time) that maximizes the long-term utility subject to long-term individual power and rate-gain constraints at all D2D pairs. In our case, the cooperative advantages are ensured in the long-term instead of in each time slot, which provides more flexibility in the resource allocation. These constraints are detailed in the following.

- *Long-Term Individual Power Constraint*: The long-term individual power constraint for DT k is given by

$$\bar{P}_k \triangleq \lim_{T \to \infty} \frac{1}{T} \sum_{t=1}^{T} \mathbb{E}[P_k[t]] \leq \mathrm{P}_{k,\mathrm{avg}}, \tag{12}$$

where

$$P_k[t] \triangleq \eta_1 \mathrm{tr}(\mathbf{Q}_k^{(1)}[t]) + \eta_2 \sum_{\ell=1}^{K} \mathrm{tr}(\mathbf{\Theta}_k \mathbf{Q}_\ell^{(2)}[t]) \tag{13}$$

is the transmit power of DT k in time slot t, and $\mathbf{\Theta}_k$ is a $KN_t \times KN_t$ diagonal matrix with $\{\mathbf{\Theta}_k\}_{j,j} = 1$, for $j = N_t(k-1)+1, \ldots, N_t k$, and $\{\mathbf{\Theta}_k\}_{j,j} = 0$, otherwise. Different from conventional per-time-slot power constraints, long-term individual power constraints provide DTs with more flexibility to distribute their power over time (e.g., allocate more power to time slots with more favorable channel conditions and vice versa). Since the slot duration is normalized to 1, the transmission power coincides with the energy consumption in each time slot.

- *Long-Term Rate-Gain Constraint*: Long-term rate-gain constraints ensure that all D2D pairs eventually achieve larger rates through cooperation. In particular, for the case without cooperation, we assume that each D2D pair is allocated $1/K$ of the total slot duration for transmission. In this case, an achievable rate for the k-th D2D pair in the non-cooperative case is given by

$$\frac{1}{K} \log_2 \left| \mathbf{I}_{N_r} + \tilde{\mathbf{H}}_{k,k}[t]\mathbf{Q}_k^{\mathrm{nc}}[t]\tilde{\mathbf{H}}_{k,k}^H[t] \right| \tag{14}$$

where $\mathbf{Q}_k^{\mathrm{nc}}[t]$ is the input covariance matrix at DT k, $\Upsilon_k^{\mathrm{nc}}[t] \in \mathbb{C}^{N_r \times N_r}$ is chosen such that $((\Upsilon_k^{\mathrm{nc}}[t])^H \Upsilon_k^{\mathrm{nc}}[t])^{-1} = \mathbf{H}_{c,k}[t]\mathbf{Q}_c[t]\mathbf{H}_{c,k}^H[t] + \sigma_n^2 \mathbf{I}_{N_r}$, and $\tilde{\mathbf{H}}_{k,k}[t] \triangleq \Upsilon_k^{\mathrm{nc}}[t]\mathbf{H}_{k,k}[t]$. Then, the maximum achievable rate for the k-th D2D pair in the non-cooperative case is given by

$$R_k^{\mathrm{nc}}[t] \triangleq \max_{\mathbf{Q}_k^{\mathrm{nc}}[t] \in \mathcal{Q}^{\mathrm{nc}}} \frac{1}{K} \log_2 \left| \mathbf{I}_{N_r} + \tilde{\mathbf{H}}_{k,k}[t]\mathbf{Q}_k^{\mathrm{nc}}[t]\tilde{\mathbf{H}}_{k,k}^H[t] \right|, \tag{15}$$

where the maximization is performed over the set of precoding matrices satisfying both power and IT constraints, i.e., $\mathcal{Q}^{\mathrm{nc}} \triangleq \{\mathbf{Q}_k^{\mathrm{nc}}[t] : \mathrm{tr}(\mathbf{Q}_k^{\mathrm{nc}}[t])/K \leq \bar{P}_k, \mathrm{tr}(\mathbf{C}_{\mathbf{G}_{k,b}}\mathbf{Q}_k^{\mathrm{nc}}[t])/K \leq \mathrm{IT}_k\}$. Note that (15) is convex and, thus, can be solved efficiently using general purpose solvers, e.g. CVX [46]. Hence, the long-term rate-gain constraint for the k-th D2D pair is given by

$$\bar{R}_k \triangleq \lim_{T\to\infty} \frac{1}{T}\sum_{t=1}^{T} \mathbb{E}[R_k[t]] \geq \lim_{T\to\infty} \frac{1}{T}\sum_{t=1}^{T} \mathbb{E}[R_k^{\mathrm{nc}}[t]] \triangleq \bar{R}_k^{\mathrm{nc}}. \quad (16)$$

Let $\Omega(\bar{\mathbf{R}})$ be the long-term utility, where $\Omega(\cdot)$ is a concave, continuous, and entry-wise non-decreasing function and $\bar{\mathbf{R}} \triangleq [\bar{R}_1, \ldots, \bar{R}_K]$ is the vector of long-term average transmission rates of the D2D pairs. For example, the utility function $\Omega(\cdot)$ can be defined as $\Omega(\bar{\mathbf{R}}) = \sum_{k=1}^{K} \log \bar{R}_k$ for proportional fairness optimization or as $\Omega(\bar{\mathbf{R}}) = \sum_{k=1}^{K} \bar{R}_k$ for sum rate maximization. Then, the problem is formulated as

$$\max_{\mathbf{Q}_k^{(1)}[t],\mathbf{Q}_k^{(2)}[t],R_k[t],\forall k,t} \quad \Omega(\bar{\mathbf{R}}) \quad (17a)$$

$$\text{subject to } \{\mathbf{Q}_k^{(1)}[t], \mathbf{Q}_k^{(2)}[t], R_k[t]\} \in \Gamma(\mathcal{H}[t], \mathrm{IT}_k), \quad (17b)$$

$$\bar{P}_k \leq \mathrm{P}_{k,\mathrm{avg}}, \ \forall k, \quad (17c)$$

$$\bar{R}_k \geq \bar{R}_k^{\mathrm{nc}}, \ \forall k, t. \quad (17d)$$

Let us denote the optimal objective value of this problem as ω_{opt}, i.e., $\omega_{\mathrm{opt}} \triangleq \Omega(\bar{\mathbf{R}}^*)$ with $\bar{\mathbf{R}}^*$ being the solution to (17). This problem is difficult to solve in practice since it generally requires noncausal CSI. However, a near-optimal policy can be obtained by adopting dynamic control and Lyapunov optimization techniques [36], [37] as we show in the following sections. It is worthwhile to note that the CSI between DTs and DRs, and also that between all D2D users and the active uplink CU are required to compute the precoders. This can be done by having DTs and DRs estimate the channel matrices locally using pilot signals emitted by DTs and CU, and forward their local estimates to the node performing the computation. This node can be either BS (in the centralized case [5]–[8]) or DTs themselves. More efficient feedback mechanisms, e.g., [47]–[49], can also be developed to reduce the overhead in practice. In this work, we assume that perfect CSI is available at all nodes, and examine the gains that can be achieved under this ideal scenario. The impact of the CSI acquisition overhead and channel estimation imperfections will be further evaluated through simulations in Section V.

Remark 1. *It is worthwhile to note that the optimization problem in* (17) *may not always be feasible due to the rate-gain and IT constraints. To ensure that these constraints can be satisfied, the DTs should be sufficiently close to each other such that the overhead required for data-sharing is tolerable and should be sufficiently far from BS so that the transmit power is not overly constrained by the IT constraint. These issues are related to studies of admission control [50], [51] and partner selection [52], [53], which are beyond the scope of this article.*

III. Dynamic BD-based Transmission Policy for Cooperative D2D Pairs

In this section, we solve the problem in (17) using the Lyapunov optimization approach [36], [37]. This approach leads to a solution in which the choice of the precoding scheme in each time slot, e.g., $\{\mathbf{Q}_k^{(1)}[t], \mathbf{Q}_k^{(2)}[t], R_k[t]\}$ for all k at time t, depends only on its current CSI and system states (and not on those of other time slots). The resulting optimization problem in each time slot can also be decoupled into K subproblems that can be solved in parallel at the DTs. The solution that is obtained, albeit suboptimal in general, can be made arbitrarily close to ω_{opt} with an appropriate choice of parameters. In particular, this is done by constructing virtual data, rate-gain, and energy queues to record the system state, and by showing the equivalence between the stability of these queues and the feasibility of the long-term constraints.

Specifically, to adopt the Lyapunov optimization technique, we introduce auxiliary variables $A_k[t]$, for all k and t, and modify the problem in (17) as

$$\max_{\mathbf{Q}_k^{(1)}[t],\mathbf{Q}_k^{(2)}[t],R_k[t],A_k[t],\forall k,t} \quad \overline{\Omega(\mathbf{A})} \quad (18a)$$

$$\text{subject to} \quad (17b), (17c), (17d), \quad (18b)$$

$$\bar{R}_k \geq \bar{A}_k, \forall k, \quad (18c)$$

$$0 \leq A_k[t] \leq A_{k,\max}, \forall k, t, \quad (18d)$$

where $\overline{\Omega(\mathbf{A})} \triangleq \lim_{T\to\infty} \frac{1}{T}\sum_{t=1}^{T} \mathbb{E}[\Omega(\mathbf{A}[t])]$ and $\mathbf{A}[t] \triangleq [A_1[t], \ldots, A_K[t]]$. The constraints of the original problem are included above and, thus, are also satisfied by the solution of the modified problem. Let $\{\mathbf{Q}_k^{(1)*}[t], \mathbf{Q}_k^{(2)*}[t], R_k^*[t], A_k^*[t]\}$, $\forall k$ and $\forall t$, be the solution of the modified problem in (18). It follows that

$$\Omega(\bar{\mathbf{R}}^*) \geq \Omega(\bar{\mathbf{A}}^*) \geq \overline{\Omega(\mathbf{A}^*)} \geq \omega_{\mathrm{opt}}(\mathbf{A}_{\max}), \quad (19)$$

where $\bar{\mathbf{A}}^* \triangleq \lim_{T\to\infty} \frac{1}{T}\sum_{t=1}^{T} \mathbb{E}[\mathbf{A}^*[t]]$, and $\omega_{\mathrm{opt}}(\mathbf{A}_{\max})$ (with $\mathbf{A}_{\max} \triangleq [A_{1,\max}, \ldots, A_{K,\max}]$) is the optimal objective value corresponding to the original problem in (17) with additional constraints $0 \leq \bar{R}_k \leq A_{k,\max}$, for all k. The first inequality follows from (18c) and the fact that $\Omega(\cdot)$ is non-decreasing; the second inequality follows from the concavity of $\Omega(\cdot)$; and the last inequality holds since, $A_k[t] = \bar{R}_k \in [0, A_{k,\max}]$, for all k and t, yields a feasible policy for (18) (see also [36, Chapter 5]). This shows that the optimal rates obtained from the modified problem can achieve at least as good an objective value as $\omega_{\mathrm{opt}}(\mathbf{A}_{\max})$.

To solve the modified problem in (18), we first transform its long-term constraints (namely, (17c), (17d), and (18c)) into queue-stability problems for virtually constructed data queues, rate-gain queues, and energy queues. The data queue records the amount of data in the transmission buffer, the rate-gain queue records the amount of service that each user should be provided in order to achieve a rate advantage over the non-cooperative case, and the energy queue records the amount of energy that each user has consumed in excess to its long-term power constraint.

Let $D_k[t]$ be the size of the virtual data queue of the k-th D2D pair at the beginning of slot t, and let $A_k[t]$ (where $0 \leq A_k[t] \leq A_{k,\max}$) and $R_k[t]$ be its arrival and departure in slot t. By letting $\mathbf{D}[t] \triangleq [D_1[t], \ldots, D_K[t]]$ and $\mathbf{R}[t] \triangleq [R_1[t], \ldots, R_K[t]]$, the evolution of the K data queues can be written as

$$\mathbf{D}[t+1] = (\mathbf{D}[t] - \mathbf{R}[t])^+ + \mathbf{A}[t], \quad (20)$$

where $(\mathbf{x})^+ = [\max\{0, x_1\}, \ldots, \max\{0, x_n\}]$ for $\mathbf{x} \triangleq [x_1, \ldots, x_n]$. Moreover, let $O_k[t]$ be the size of the rate-gain queue of DT k at the beginning of slot t. Here, the arrival in slot t is the non-cooperative rate $R_k^{\mathrm{nc}}[t]$ and the departure is the cooperative rate $R_k[t]$. The queue evolution can be written as

$$\mathbf{O}[t+1] = (\mathbf{O}[t] - \mathbf{R}[t])^+ + \mathbf{R}^{\mathrm{nc}}[t], \tag{21}$$

where $\mathbf{O}[t] \triangleq [O_1[t], \ldots, O_K[t]]$ and $\mathbf{R}^{\mathrm{nc}}[t] \triangleq [R_1^{\mathrm{nc}}[t], \ldots, R_K^{\mathrm{nc}}[t]]$. The queue size $O_k[t]$ can be viewed as the credit that the k-th D2D pair earned by sharing its resources to the cooperative transmission of other DTs' data. The k-th D2D pair is likely to take on a larger share of the cooperative resources (i.e., transmit its own data at a higher rate) when $O_k[t]$ is larger, and vice versa. Finally, let $E_k[t]$ be the size of the k-th virtual energy queue at the beginning of slot t. The arrival at time t is the transmit power $P_k[t]$, and the departure is $\mathrm{P}_{k,\mathrm{avg}}$, which is constant for all t. By letting $\mathbf{E}[t] \triangleq [E_1[t], \ldots, E_K[t]]$, $\mathbf{P}_{\mathrm{avg}} \triangleq [\mathrm{P}_{1,\mathrm{avg}}, \ldots, \mathrm{P}_{K,\mathrm{avg}}]$ and $\mathbf{P}[t] \triangleq [P_1[t], \ldots, P_K[t]]$, the queue evolution can be written as

$$\mathbf{E}[t+1] = (\mathbf{E}[t] - \mathbf{P}_{\mathrm{avg}})^+ + \mathbf{P}[t]. \tag{22}$$

Here, $E_k[t]$ can be interpreted as the amount of energy used by DT k that is in excess to its per time slot budget $\mathrm{P}_{k,\mathrm{avg}}$.

Definition 1 ([36]). *For any k, the queue $\{D_k[t]\}_{t=1}^{\infty}$ (and, similarly, for $\{O_k[t]\}_{t=1}^{\infty}$ and $\{E_k[t]\}_{t=1}^{\infty}$) is strongly stable if*

$$\limsup_{T\to\infty} \frac{1}{T} \sum_{t=1}^{T} \mathbb{E}[D_k[t]] < \infty, \tag{23}$$

and is mean-rate stable if

$$\lim_{T\to\infty} \frac{\mathbb{E}[D_k[T]]}{T} = 0. \tag{24}$$

Note that strong stability ensures that the queue-lengths are bounded, and implies mean-rate stability as long as the difference between the expected departure and arrival in each time slot is bounded [36, Theorem 2.8], which is satisfied in our case. On the other hand, mean-rate stability of the virtual queues implies that the long-term constraints in (17c), (17d), and (18c) are satisfied [36, Theorem 2.5]. Hence, a feasible solution for (18) is a BD-based transmission policy that ensures mean-rate stability of the virtual queues while having $0 \le A_k[t] \le A_{k,\max}$, for all k and t. The proposed *maximum-weighted-rate-minus-energy-penalty (Max-WRMEP)* transmission policy, to be described in the following, achieves this task.

Max-WRMEP Policy (at time slot t):

(i) **Virtual Data Arrival $\mathbf{A}[t]$:** The virtual arrival $\mathbf{A}[t]$ at time t is obtained as the solution of the following optimization problem:

$$\max_{\mathbf{A}[t]:0\le A_k[t]\le A_{k,\max},\forall k} V\Omega(\mathbf{A}[t]) - \sum_{k=1}^{K} D_k[t]A_k[t], \tag{25}$$

where $V > 0$ is some predefined constant.

(ii) **BD Cooperative Precoding Scheme $\{\mathbf{Q}_k^{(1)}[t], \mathbf{Q}_k^{(2)}[t], R_k[t], \forall k\}$:** For each k, the BD cooperative precoding scheme at time t is obtained as the solution of the following optimization problem:

$$\max_{\mathbf{Q}_k^{(1)}[t],\mathbf{Q}_k^{(2)}[t],R_k[t]} (D_k[t]+O_k[t])R_k[t] - \eta_1 \mathrm{tr}(E_k[t]\mathbf{Q}_k^{(1)}[t]) - \eta_2 \mathrm{tr}(\bar{\boldsymbol{\Theta}}[t]\mathbf{Q}_k^{(2)}[t]) \tag{26a}$$

$$\text{subject to } \{\mathbf{Q}_k^{(1)}[t], \mathbf{Q}_k^{(2)}[t], R_k[t]\} \in \Gamma(\mathcal{H}[t], \mathrm{IT}_k), \tag{26b}$$

where $\bar{\boldsymbol{\Theta}}[t] \triangleq \sum_{\ell=1}^{K} E_\ell[t]\boldsymbol{\Theta}_\ell$. This problem can be solved in parallel for the K D2D pairs.

(iii) **Updates of Queues $\mathbf{D}[t]$, $\mathbf{O}[t]$, and $\mathbf{E}[t]$:** The virtual queues are updated according to (20), (21) and (22), using the solutions obtained in the previous steps.

Notice that the original problem in (17) involves the optimization of a long-term utility function under long-term power and rate-gain constraints, which generally requires noncausal CSI and joint optimization of parameters over multiple time slots. The proposed Max-WRMEP policy instead determines the BD cooperative precoding on a slot-by-slot and user-by-user basis, which makes the problem more tractable in practice. Specifically, in (i), the virtual data arrivals are determined by exploiting the tradeoff between maximizing the objective function $\Omega(\cdot)$ and reducing the backlog $D_k[t]$. In particular, when $\Omega(\bar{\mathbf{R}})$ is chosen to be the average sum rate, i.e. $\Omega(\bar{\mathbf{R}}) \triangleq \sum_{k=1}^{K} \bar{R}_k$, the virtual data arrival is given by $A_k[t] = A_{k,\max}$, if $V \ge D_k[t]$, and $A_k[t] = 0$, otherwise; and when $\Omega(\bar{\mathbf{R}}) \triangleq \sum_{k=1}^{K} \log(\bar{R}_k)$, we have $A_k[t] = \min(V/D_k[t], A_{k,\max})$. In (ii), the input covariance matrices and rates are determined based on the queue states at that time. Here, more power is expended to increase the effective rate of the k-th D2D pair if the data queue backlog $D_k[t]$ and/or the credit accumulated through cooperation $O_k[t]$ is larger, whereas less power is used if the violation of the energy budget (i.e., $E_k[t]$) is too large. The queues are then updated in (iii).

In addition, let us consider a special class of stationary randomized policies, called $\mathcal{H}$-only transmission policies, where the transmission control actions in each time slot depend only on the channel state $\mathcal{H}[t]$ at that time. It is defined formally as follows.

Definition 2. *An $\mathcal{H}$-only transmission policy is a policy that chooses, at each time slot t, the control actions $A_k[t] \in [0, A_{k,\max}]$ and $\{\mathbf{Q}_k^{(1)}[t], \mathbf{Q}_k^{(2)}[t], R_k[t]\} \in \Gamma(\mathcal{H}[t], \mathrm{IT}_k)$, $\forall k$, based only on the set of channel states $\mathcal{H}[t]$ at that time.*

We know from [36, Theorem 4.5] that, if (17) is feasible, then there exists an $\mathcal{H}$-only transmission policy that performs arbitrarily close to the optimal solution. This policy is used to show the optimality of our proposed Max-WRMEP transmission policy and the corresponding queue-length bounds as stated in the following theorems.

Theorem 1. *Suppose that the problem in* (17) *is feasible. Then, under the proposed Max-WRMEP policy, the long-term constraints* (17c), (17d), *and* (18c) *are satisfied, and the resulting long-term average utility yields*

$$\liminf_{T\to\infty} \Omega\left(\frac{1}{T}\sum_{t=1}^{T}\mathbb{E}[\mathbf{R}[t]]\right) \ge \omega_{\mathrm{opt}}(\mathbf{A}_{\max}) - \frac{C}{V}. \tag{27}$$

where C is a positive constant specified in Appendix A.

The proof follows similar procedures as in [36], [37] and is given in Appendix A. Recall that $\omega_{\rm opt}(\mathbf{A}_{\max})$ is the optimal objective value of the problem in (17) with additional constraints $0 \le \bar{R}_k \le A_{k,\max}$, for all k. Hence, Theorem 1 implies that, for $A_{k,\max}$ sufficiently large (or, more specifically, for $A_{k,\max} \ge \bar{R}_k^{\rm opt}$, where $R_k^{\rm opt}[t]$, $\forall t$, is the rate solution to the original problem in (17)), the proposed Max-WRMEP policy can achieve a utility that is arbitrarily close to the optimal value $\omega_{\rm opt}$ by setting V to be sufficiently large. In Theorem 1, we have shown the mean-rate stability of the virtual queues and the feasibility of the proposed policy. In the following, we further show the strong stability and boundedness of the virtual queues.

Theorem 2. *Suppose that, for any $\epsilon > 0$, there exists an $\mathcal{H}$-only transmission policy (as per Definition 2) with actions $A_k(\mathcal{H}[t]) \in [0, A_{k,\max}]$, and $\{\mathbf{Q}_k^{(1)}(\mathcal{H}[t]), \mathbf{Q}_k^{(2)}(\mathcal{H}[t]), R_k(\mathcal{H}[t])\} \in \Gamma(\mathcal{H}[t], \mathrm{IT}_k)$, for all k and t, such that*

$$\mathbb{E}[A_k(\mathcal{H}[t]) - R_k(\mathcal{H}[t])] \le -\epsilon, \tag{28}$$
$$\mathbb{E}[R_k^{\rm nc}[t] - R_k(\mathcal{H}[t])] \le -\epsilon, \tag{29}$$
$$\mathbb{E}[P_k(\mathcal{H}[t]) - \mathrm{P}_{k,\rm avg}] \le -\epsilon, \tag{30}$$

for all k and t, where $P_k(\mathcal{H}[t]) \triangleq \eta_1 \mathrm{tr}(\mathbf{Q}_k^{(1)}(\mathcal{H}[t])) + \eta_2 \sum_{\ell=1}^{K} \mathrm{tr}(\mathbf{\Theta}_k \mathbf{Q}_\ell^{(2)}(\mathcal{H}[t]))$ is the transmit power. Then, the queues $\{D_k[t]\}_{t=1}^{\infty}$, $\{O_k[t]\}_{t=1}^{\infty}$ and $\{E_k[t]\}_{t=1}^{\infty}$, for all k, are strongly stable and are bounded as

$$\limsup_{T\to\infty} \frac{1}{T} \sum_{t=1}^{T} \sum_{k=1}^{K} \left(\mathbb{E}[D_k[t]] + \mathbb{E}[O_k[t]] + \mathbb{E}[E_k[t]]\right) \le \frac{C + V(\overline{\Omega(\mathbf{A})} - \omega_\epsilon)}{\epsilon}, \tag{31}$$

where $\omega_\epsilon \triangleq \mathbb{E}[\Omega(\mathbf{A}(\mathcal{H}[t]))]$ and C is the positive constant given in Theorem 1.

The proof is given in Appendix B. Theorem 2 shows that, if there exists a policy within the interior of the set of feasible policies, then the queues are strongly stable with a bound on the average queue length given in (31). This bound illustrates the dependence of the average queue length on V, ϵ and the difference between $\overline{\Omega(\mathbf{A})}$ and ω_ϵ. From the above theorems, we can see that, by choosing V to be sufficiently large, the utility value of the proposed policy can be made arbitrarily close to the optimal value, albeit at the cost of an increased average queue length. The queue length affects the time that is required for the time-averaged utility to become stable and approach its limiting value. Also, if the performance of the proposed policy, i.e., $\overline{\Omega(\mathbf{A})}$, is close to that of an $\mathcal{H}$-only policy with large ϵ, then the average queue length will be small. If the $\mathcal{H}$-only policy obtains a solution that is close to the boundary of the set of feasible policies (i.e., small ϵ), the upper bound of the average queue length becomes large and more time would be required to achieve a stable time-averaged utility value. Even though the theorems were obtained following standard procedures in Lyapunov optimization theory [36], [37], they are necessary to establish the optimality of the proposed scheme and its performance bounds.

IV. Maximum WRMEP Precoding with Spatially White Multicast Input

The Max-WRMEP policy proposed in the previous section allows the virtual data arrivals and the BD cooperative precoder in each time slot to be determined based only on the current channel and queue states. However, the precoders in both phases still require solving the optimization problem in (26) correctly. This is can be done using standard numerical optimization toolboxes, such as CVX [46], which can be inefficient when the problem size is large. In this section, we first derive the optimal structure of the BD cooperative precoding matrix, and then propose a low-complexity approximate solution that utilizes spatially white multicast input in phase 1. Further analysis on the Lagrange multipliers in the latter case leads to a more efficient iterative algorithm for solving the cooperative precoder in phase 2.

Specifically, let $\bar{\mathbf{H}}_{-k}[t] \triangleq [\tilde{\mathbf{H}}_1^H[t], \ldots, \tilde{\mathbf{H}}_{k-1}^H[t], \tilde{\mathbf{H}}_{k+1}^H[t], \ldots, \tilde{\mathbf{H}}_K^H[t]]^H$ be the collection of effective channels associated with all DRs other than DR k and let $\bar{\mathbf{H}}_{-k}[t] = \mathbf{U}_k[t][\mathbf{\Sigma}_k[t]\ \mathbf{0}][\mathbf{V}_k[t]\ \tilde{\mathbf{V}}_k[t]]^H$ be the singular value decomposition (SVD) of $\bar{\mathbf{H}}_{-k}[t]$, where $\mathbf{U}_k[t] \in \mathbb{C}^{(K-1)N_r \times (K-1)N_r}$ and $[\mathbf{V}_k[t]\ \tilde{\mathbf{V}}_k[t]] \in \mathbb{C}^{KN_t \times KN_t}$ are unitary matrices, and $\mathbf{\Sigma}_k[t] \in \mathbb{C}^{(K-1)N_r \times (K-1)N_r}$ is a diagonal matrix of singular values. Here, $\tilde{\mathbf{V}}_k[t] \in \mathbb{C}^{KN_t \times M'}$ and $\mathbf{V}_k[t] \in \mathbb{C}^{KN_t \times KN_t - M'}$, where $M' \triangleq KN_t - (K-1)N_r$. It was shown in [35] that, to satisfy the BD constraints in (8), the transmit covariance matrix $\mathbf{Q}_k^{(2)}[t]$ in phase 2 must take on the following structure

$$\mathbf{Q}_k^{(2)}[t] = \tilde{\mathbf{V}}_k[t] \tilde{\mathbf{Q}}_k^{(2)}[t] \tilde{\mathbf{V}}_k^H[t], \tag{32}$$

for all k and t, where $\tilde{\mathbf{Q}}_k^{(2)}[t] \in \mathbb{C}^{M' \times M'}$ is a positive semi-definite matrix. This implies that, to satisfy the BD constraint, the transmit signal for DR k should lie in the null space of $\bar{\mathbf{H}}_{-k}[t]$. By the optimal structure in (32), the phase-2 rate constraint in (9) for the k-th D2D pair can be written as

$$R_k[t] \le \eta_2 \log_2 \left| \mathbf{I}_{N_r} + \tilde{\mathbf{H}}_k[t] \tilde{\mathbf{V}}_k[t] \tilde{\mathbf{Q}}_k^{(2)}[t] \tilde{\mathbf{V}}_k^H[t] \tilde{\mathbf{H}}_k^H[t] \right| \tag{33}$$

where $\tilde{\mathbf{H}}_k[t]\tilde{\mathbf{V}}_k[t]$ is the projection of the effective channel matrix $\tilde{\mathbf{H}}_k[t]$ onto the null space of $\bar{\mathbf{H}}_{-k}[t]$. In this case, $\tilde{\mathbf{Q}}_k^{(2)}[t]$ is the transmit covariance matrix of the signal transmitted over the effective interference-free channel $\tilde{\mathbf{H}}_k[t]\tilde{\mathbf{V}}_k[t]$. Notice that $KN_t - (K-1)N_r$ is the maximum number of parallel data streams that can be transmitted while satisfying the BD constraint and, thus, we choose the number of transmitted data streams M equal to M'. Then, the optimization problem in (26) for finding the BD cooperative precoding scheme can be written as

$$\max_{\mathbf{Q}_k^{(1)}[t],\tilde{\mathbf{Q}}_k^{(2)}[t],R_k[t]} (D_k[t]+O_k[t])R_k[t]-\eta_1\mathrm{tr}(E_k[t]\mathbf{Q}_k^{(1)}[t]) - \eta_2\mathrm{tr}(\tilde{\mathbf{\Theta}}_k[t]\tilde{\mathbf{Q}}_k^{(2)}[t]) \tag{34a}$$

subject to

$$R_k[t]\leq\eta_1\log_2\left|\mathbf{I}_{N_t}+\tilde{\mathbf{G}}_{k,\ell}[t]\mathbf{Q}_k^{(1)}[t]\tilde{\mathbf{G}}_{k,\ell}^H[t]\right|,\forall\ell\neq k, \tag{34b}$$

$$R_k[t]\leq\eta_2\log_2\left|\mathbf{I}_{N_r}+\tilde{\mathbf{H}}_k[t]\tilde{\mathbf{V}}_k[t]\tilde{\mathbf{Q}}_k^{(2)}[t]\tilde{\mathbf{V}}_k^H[t]\tilde{\mathbf{H}}_k^H[t]\right|, \tag{34c}$$

$$\eta_1\mathrm{tr}(\mathbf{C}_{\mathbf{G}_{k,b}}\mathbf{Q}_k^{(1)}[t])+\eta_2\mathrm{tr}(\tilde{\mathbf{C}}_{\mathbf{G}_{k,b}}[t]\tilde{\mathbf{Q}}_k^{(2)}[t])\leq\mathrm{IT}_k, \tag{34d}$$

$$\mathbf{Q}_k^{(1)}[t]\succeq 0,\ \tilde{\mathbf{Q}}_k^{(2)}[t]\succeq 0,\ R_k[t]\geq 0, \tag{34e}$$

for $k=1,\ldots,K$, where $\tilde{\mathbf{\Theta}}_k[t]\triangleq\tilde{\mathbf{V}}_k^H[t]\bar{\mathbf{\Theta}}[t]\tilde{\mathbf{V}}_k[t]$, and $\tilde{\mathbf{C}}_{\mathbf{G}_{k,b}}[t]\triangleq\tilde{\mathbf{V}}_k^H[t]\mathbf{C}_{\mathbf{G}_b}\tilde{\mathbf{V}}_k[t]$. This problem is convex and can be solved by off-the-shelf solvers such as CVX [46]. However, these general purpose solvers do not further exploit the structure of this problem and the computational complexity required may increase with the number of D2D pairs.

To reduce the complexity of the design, we propose to adopt a spatially white input for multicasting in phase 1, where $\mathbf{Q}_k^{(1)}=\alpha_k\mathbf{I}_{N_t}$, for all k. This has been commonly adopted in the literature on physical layer multicasting, e.g., in [54], to reduce the complexity of the precoder design and the CSI requirement, and is known to perform well when the number of users is large or sufficiently spread out. In this case, the Lagrangian function can be written as

$$\begin{aligned}&\mathcal{L}(R_k,\alpha_k,\tilde{\mathbf{Q}}_k^{(2)},\{\lambda_{k,\ell}\}_{\ell\neq k},\mu_k,\delta_k)\\&=(D_k+O_k)R_k-\eta_1E_kN_t\alpha_k-\eta_2\mathrm{tr}\left(\tilde{\mathbf{\Theta}}_k\tilde{\mathbf{Q}}_k^{(2)}\right)\\&\quad-\sum_{\ell\neq k}\lambda_{k,\ell}\left(R_k-\eta_1\log_2\left|\mathbf{I}_{N_t}+\alpha_k\tilde{\mathbf{G}}_{k,\ell}\tilde{\mathbf{G}}_{k,\ell}^H\right|\right)\\&\quad-\mu_k\left(R_k-\eta_2\log_2\left|\mathbf{I}_{N_r}+\tilde{\mathbf{H}}_k\tilde{\mathbf{V}}_k\tilde{\mathbf{Q}}_k^{(2)}\tilde{\mathbf{V}}_k^H\tilde{\mathbf{H}}_k^H\right|\right)\\&\quad-\delta_k\left(\eta_1\alpha_k\mathrm{tr}\big(\mathbf{C}_{\mathbf{G}_{k,b}}\big)+\eta_2\mathrm{tr}\big(\tilde{\mathbf{C}}_{\mathbf{G}_{k,b}}\tilde{\mathbf{Q}}_k^{(2)}\big)-\mathrm{IT}_k\right),\end{aligned} \tag{35}$$

where $\{\lambda_{k,\ell}\}_{\ell\neq k}$, μ_k, and δ_k are the non-negative Lagrange multipliers. The Lagrange dual function is given by

$$u(\{\lambda_{k,\ell}\}_{\ell\neq k},\mu_k,\delta_k)=\max_{R_k\geq0,\alpha_k\geq0,\tilde{\mathbf{Q}}_k^{(2)}\succeq0}\mathcal{L}(R_k,\alpha_k,\tilde{\mathbf{Q}}_k^{(2)},\{\lambda_{k,\ell}\}_{\ell\neq k},\mu_k,\delta_k) \tag{36}$$

and the dual optimization problem can be written as

$$\min_{\lambda_{k,\ell}\geq0,\forall\ell\neq k,\mu_k\geq0,\delta_k\geq0}u(\{\lambda_{k,\ell}\}_{\ell\neq k},\mu_k,\delta_k). \tag{37}$$

By the stationarity condition with respect to R_k, the optimal solution must satisfy

$$D_k+O_k=\sum_{\ell\neq k}\lambda_{k,\ell}+\mu_k. \tag{38}$$

If $D_k+O_k-\mu_k=0$ (i.e., if $\lambda_{k,\ell}=0$ for all $\ell\neq k$) at the optimal point, the value of α_k that maximizes the Lagrangian function, regardless of the value of other parameters, is zero. In this case, the optimal rate R_k is 0 and, thus, the k-th D2D pair remains silent. Otherwise, if $D_k+O_k-\mu_k>0$, we can define $\lambda'_{k,\ell}\triangleq\frac{\lambda_{k,\ell}}{D_k+O_k-\mu_k}$, for all $\ell\neq k$, such that $\lambda'_{k,\ell}\geq0$, for all $\ell\neq k$, and $\sum_{\ell\neq k}\lambda'_{k,\ell}=1$. By the complementary slackness condition, we know that $\lambda_{k,\ell}>0$ (and, thus, $\lambda'_{k,\ell}>0$) only if

$$\log_2\left|\mathbf{I}_{N_t}+\alpha_k\tilde{\mathbf{G}}_{k,\ell}\tilde{\mathbf{G}}_{k,\ell}^H\right|\leq\log_2\left|\mathbf{I}_{N_t}+\alpha_k\tilde{\mathbf{G}}_{k,\ell'}\tilde{\mathbf{G}}_{k,\ell'}^H\right| \tag{39}$$

for all $\ell'\neq k$. Hence, we have

$$\sum_{\ell\neq k}\lambda_{k,\ell}\log_2\left|\mathbf{I}_{N_t}+\alpha_k\tilde{\mathbf{G}}_{k,\ell}\tilde{\mathbf{G}}_{k,\ell}^H\right|$$

$$=(D_k+O_k-\mu_k)\sum_{\ell\neq k}\lambda'_{k,\ell}\log_2\left|\mathbf{I}_{N_t}+\alpha_k\tilde{\mathbf{G}}_{k,\ell}\tilde{\mathbf{G}}_{k,\ell}^H\right| \tag{40}$$

$$=(D_k+O_k-\mu_k)\min_{\ell\neq k}\log_2\left|\mathbf{I}_{N_t}+\alpha_k\tilde{\mathbf{G}}_{k,\ell}\tilde{\mathbf{G}}_{k,\ell}^H\right|. \tag{41}$$

By (38) and (41), the dual optimization problem can be reformulated as

$$\begin{aligned}\min_{\mu_k\geq0,\delta_k\geq0}\ &\max_{\alpha_k\geq0,\tilde{\mathbf{Q}}_k^{(2)}\succeq\mathbf{0}}\\&\eta_1\Big\{(D_k+O_k-\mu_k)\min_{\ell\neq k}\log_2\left|\mathbf{I}_{N_t}+\alpha_k\tilde{\mathbf{G}}_{k,\ell}\tilde{\mathbf{G}}_{k,\ell}^H\right|\\&\quad-[E_kN_t+\delta_k\mathrm{tr}(\mathbf{C}_{\mathbf{G}_{k,b}})]\alpha_k\Big\}\\&+\eta_2\Big\{\mu_k\log_2\left|\mathbf{I}_{N_r}+\tilde{\mathbf{H}}_k\tilde{\mathbf{V}}_k\tilde{\mathbf{Q}}_k^{(2)}\tilde{\mathbf{V}}_k^H\tilde{\mathbf{H}}_k^H\right|\\&\quad-\mathrm{tr}\big[\big(\tilde{\mathbf{\Theta}}_k+\delta_k\tilde{\mathbf{C}}_{\mathbf{G}_{k,b}}\big)\tilde{\mathbf{Q}}_k^{(2)}\big]\Big\}+\delta_k\mathrm{IT}_k.\end{aligned} \tag{42}$$

It is interesting to observe that, at the optimal point, (34c) must be satisfied with equality since otherwise, we can always choose R_k to be larger or choose $\tilde{\mathbf{Q}}_k^{(2)}$ to be smaller to further increase the objective value, which contradicts with the fact that the point is optimal. By the same argument, (34b) must also be satisfied with equality for some ℓ. This implies that

$$\eta_1\min_{\ell\neq k}\log_2\left|\mathbf{I}_{N_t}+\alpha_k\tilde{\mathbf{G}}_{k,\ell}\tilde{\mathbf{G}}_{k,\ell}^H\right|-\eta_2\log_2\left|\mathbf{I}_{N_r}+\tilde{\mathbf{H}}_k\tilde{\mathbf{V}}_k\tilde{\mathbf{Q}}_k^{(2)}\tilde{\mathbf{V}}_k^H\tilde{\mathbf{H}}_k^H\right|=0 \tag{43}$$

at the optimal point. Moreover, by the complementary slackness condition on δ_k, we also know that, at the optimal point, $\delta_k>0$ only if the IT constraint is active, i.e.,

$$\eta_1\alpha_k\mathrm{tr}\big(\mathbf{C}_{\mathbf{G}_{k,b}}\big)+\eta_2\mathrm{tr}\big(\tilde{\mathbf{C}}_{\mathbf{G}_{k,b}}\tilde{\mathbf{Q}}_k^{(2)}\big)-\mathrm{IT}_k=0. \tag{44}$$

Notice that the left-hand-side of (43) is weighted by $-\mu_k$ in (42) and, thus, the maximization over α_k and $\tilde{\mathbf{Q}}_k^{(2)}$ in (42) must yield solutions that cause the left-hand-side of (43) to be non-increasing with respect to μ_k. Similarly, these solutions must also cause the left-hand-side of (44) to be non-increasing with respect to δ_k. Hence, when the IT constraint is active (i.e., when $\delta_k>0$), the optimal μ_k and δ_k can be found by employing a two-dimensional bisection search [55] aiming towards finding $\mu_k\in[0,D_k+O_k]$ and $\delta_k\in[0,\delta_{k,\mathrm{up}}]$ that satisfies (43) and (44), where $\delta_{k,\mathrm{up}}\triangleq\frac{D_k+O_k}{\mathrm{tr}(\mathbf{C}_{\mathbf{G}_{k,b}})\ln2}\max_{\ell\neq k}\mathrm{tr}(\tilde{\mathbf{G}}_{k,\ell}\tilde{\mathbf{G}}_{k,\ell}^H)-\frac{E_kN_t}{\mathrm{tr}(\mathbf{C}_{\mathbf{G}_{k,b}})}$. Notice that, when $\delta_k=\delta_{k,\mathrm{up}}$, the objective function in (42) is non-increasing with respect to $\alpha_k\geq0$ regardless of the choice of μ_k and $\tilde{\mathbf{Q}}_k^{(2)}$. On the other hand, when the IT constraint is inactive, the problem reduces to a one-dimensional bisection search over $\mu_k \in [0, D_k + O_k]$.

In each iteration of the bisection search, where μ_k and δ_k are given, the optimal transmit power α_k in phase 1 can be found by solving the following optimization problem:

$$\max_{\alpha_k \geq 0} \left\{ (D_k + O_k - \mu_k) \min_{\ell \neq k} \log_2 \left| \mathbf{I}_{N_t} + \alpha_k \tilde{\mathbf{G}}_{k,\ell} \tilde{\mathbf{G}}_{k,\ell}^H \right| - \left[E_k N_t + \delta_k \mathrm{tr}(\mathbf{C}_{\mathbf{G}_{k,b}}) \right] \alpha_k \right\}$$
$$= \max_{\alpha_k \geq 0} \left\{ (D_k + O_k - \mu_k) \min_{\ell \neq k} \sum_{n=1}^{N_t} \log_2 \left(1 + \alpha_k \varsigma_{k,\ell,n}^2 \right) - \left[E_k N_t + \delta_k \mathrm{tr}(\mathbf{C}_{\mathbf{G}_{k,b}}) \right] \alpha_k \right\}, \quad (45)$$

where $\{\varsigma_{k,\ell,n}\}_{n=1}^{N_t}$ are the singular values of $\tilde{\mathbf{G}}_{k,\ell}$. The optimal α_k can be found by simple convex optimization procedures for single parameter optimization [56]. For example, we can adopt the bisection line search over $\alpha_k \in [0, \alpha_{k,\mathrm{up}}]$, where $\alpha_{k,\mathrm{up}}$ can be chosen such that an upper bound of the derivative at this point is 0, i.e.,

$$\frac{(D_k + O_k - \mu_k) N_t}{\alpha_{k,\mathrm{up}} \ln 2} - [E_k N_t + \delta_k \mathrm{tr}(\mathbf{C}_{\mathbf{G}_{k,b}})] = 0. \quad (46)$$

In this case, we have

$$\alpha_{k,\mathrm{up}} \triangleq \frac{(D_k + O_k - \mu_k) N_t}{[E_k N_t + \delta_k \mathrm{tr}(\mathbf{C}_{\mathbf{G}_{k,b}})] \ln 2}. \quad (47)$$

Moreover, the optimal $\tilde{\mathbf{Q}}_k^{(2)}$ can be found by solving the following optimization problem:

$$\max_{\tilde{\mathbf{Q}}_k^{(2)} \succeq 0} \quad \mu_k \log_2 \left| \mathbf{I}_{N_r} + \tilde{\mathbf{H}}_k \tilde{\mathbf{V}}_k \tilde{\mathbf{Q}}_k^{(2)} \tilde{\mathbf{V}}_k^H \tilde{\mathbf{H}}_k^H \right| - \mathrm{tr}\left[(\tilde{\mathbf{\Theta}}_k + \delta_k \tilde{\mathbf{C}}_{\mathbf{G}_{k,b}}) \tilde{\mathbf{Q}}_k^{(2)} \right]. \quad (48)$$

Notice that the optimal δ_k must be positive to make $\tilde{\mathbf{\Theta}}_k + \delta_k \tilde{\mathbf{C}}_{\mathbf{G}_{k,b}}$ full rank (if $\tilde{\mathbf{\Theta}}_k$ is not) since, otherwise, (48) could become unbounded by taking $\tilde{\mathbf{Q}}_k^{(2)} = \beta_k \mathbf{v}\mathbf{v}^H$ with $\beta_k \to \infty$, where $\mathbf{v}$ is the eigenvector corresponding to the zero eigenvalue of $\tilde{\mathbf{\Theta}}_k$. By letting $\hat{\mathbf{Q}}_k^{(2)} \triangleq (\tilde{\mathbf{\Theta}}_k + \delta_k \tilde{\mathbf{C}}_{\mathbf{G}_{k,b}})^{\frac{1}{2}} \tilde{\mathbf{Q}}_k^{(2)} (\tilde{\mathbf{\Theta}}_k + \delta_k \tilde{\mathbf{C}}_{\mathbf{G}_{k,b}})^{\frac{H}{2}}$, the problem in (48) can be written as

$$\max_{\hat{\mathbf{Q}}_k^{(2)} \succeq 0} \quad \mu_k \log_2 \left| \mathbf{I}_{N_r} + \hat{\mathbf{H}}_k \hat{\mathbf{Q}}_k^{(2)} \hat{\mathbf{H}}_k^H \right| - \mathrm{tr}\left(\hat{\mathbf{Q}}_k^{(2)} \right) \quad (49)$$

where $\hat{\mathbf{H}}_k \triangleq \tilde{\mathbf{H}}_k \tilde{\mathbf{V}}_k (\tilde{\mathbf{\Theta}}_k + \delta_k \tilde{\mathbf{C}}_{\mathbf{G}_{k,b}})^{-\frac{1}{2}}$. Let $\hat{\mathbf{H}}_k = \hat{\mathbf{U}}_k \hat{\Sigma}_k \hat{\mathbf{V}}_k^H$ be the SVD of $\hat{\mathbf{H}}_k$, where $\hat{\Sigma}_k = \mathrm{diag}\{\hat{\sigma}_{k,1}, \ldots, \hat{\sigma}_{k,M'}\}$, $\hat{\mathbf{U}}_k$ is an $N_r \times M'$ semi-unitary matrix, and $\hat{\mathbf{V}}_k$ is an $M' \times M'$ unitary matrix. By Hadamard's inequality [57], the optimal $\hat{\mathbf{Q}}_k^{(2)}$ must take on the form

$$\hat{\mathbf{Q}}_k^{(2)} = \hat{\mathbf{V}}_k \mathbf{B}_k \hat{\mathbf{V}}_k^H, \quad (50)$$

where $\mathbf{B}_k = \mathrm{diag}(\beta_{k,1}, \ldots, \beta_{k,M'})$. In this case, the optimization problem further reduces to the following form:

$$\max_{\beta_{k,m} \geq 0, \forall m} \quad \mu_k \sum_{m=1}^{M'} \log_2 \left(1 + \beta_{k,m} \hat{\sigma}_{k,m}^2 \right) - \sum_{m=1}^{M'} \beta_{k,m}. \quad (51)$$

Algorithm 1 BD Cooperative Precoding with Spatially White Multicast Input

- Set $\delta_{k,\mathrm{low}} \leftarrow 0$ and $\delta_{k,\mathrm{up}} \leftarrow \frac{D_k + O_k}{\mathrm{tr}(\mathbf{C}_{\mathbf{G}_{k,b}}) \ln 2} \max_{\ell \neq k} \mathrm{tr}(\tilde{\mathbf{G}}_{k,\ell} \tilde{\mathbf{G}}_{k,\ell}^H) - \frac{E_k N_t}{\mathrm{tr}(\mathbf{C}_{\mathbf{G}_{k,b}})}$.
- If $\mathrm{rank}(\tilde{\mathbf{\Theta}}_k) = M'$, then set $\delta_k \leftarrow 0$. Else, set $\delta_k \leftarrow (\delta_{k,\mathrm{low}} + \delta_{k,\mathrm{up}})/2$.

while $\delta_{k,\mathrm{up}} - \delta_{k,\mathrm{low}} > \epsilon$
- Set $\mu_{k,\mathrm{low}} \leftarrow 0$, $\mu_{k,\mathrm{up}} \leftarrow D_k + O_k$, and $\mu_k \leftarrow (\mu_{k,\mathrm{low}} + \mu_{k,\mathrm{up}})/2$.

 while $\mu_{k,\mathrm{up}} - \mu_{k,\mathrm{low}} > \epsilon$
 - Compute the singular values $\{\hat{\sigma}_{k,m}\}_{m=1}^{M'}$ of $\hat{\mathbf{H}}_k$.
 - Compute $\beta_{k,m} \leftarrow \left(\frac{\mu_k}{\ln 2} - \frac{1}{\hat{\sigma}_{k,m}^2} \right)^+$, for all m, and $R_{k,2} \leftarrow \eta_2 \sum_{m=1}^{M'} \log_2(1 + \hat{\sigma}_{k,m}^2 \beta_{k,m})$.
 - Compute α_k by solving (45) with the bisection line search over $\alpha_k \in [0, \alpha_{k,\mathrm{up}}]$, where $\alpha_{k,\mathrm{up}}$ is defined in (47).
 - Compute $R_{k,1} \leftarrow \eta_1 \min_{\ell \neq k} \left| \mathbf{I}_{N_t} + \alpha_k \tilde{\mathbf{G}}_{k,\ell} \tilde{\mathbf{G}}_{k,\ell}^H \right|$.
 - Set $\mu_{k,\mathrm{low}} \leftarrow \mu_k$, if $R_{k,1} > R_{k,2}$, and set $\mu_{k,\mathrm{up}} \leftarrow \mu_k$, otherwise. Update $\mu_k \leftarrow \frac{\mu_{k,\mathrm{low}} + \mu_{k,\mathrm{up}}}{2}$.

 end while
- Set $\delta_{k,\mathrm{low}} \leftarrow \delta_k$, if $\eta_1 \alpha_k \mathrm{tr}(\mathbf{C}_{\mathbf{G}_{k,b}}) + \eta_2 \mathrm{tr}(\tilde{\mathbf{C}}_{\mathbf{G}_{k,b}} \tilde{\mathbf{Q}}_k^{(2)}) \geq \mathrm{IT}_k$, and set $\delta_{k,\mathrm{up}} \leftarrow \delta_k$, otherwise. Update $\delta_k \leftarrow (\delta_{k,\mathrm{low}} + \delta_{k,\mathrm{up}})/2$.

end while

The solution is given by

$$\beta_{k,m} = \left(\frac{\mu_k}{\ln 2} - \frac{1}{\hat{\sigma}_{k,m}^2} \right)^+, \quad (52)$$

for $m = 1, \ldots, M'$. It is necessary to note that the solution depends on δ_k through the values of $\hat{\sigma}_{k,m}^2$, $\forall m$, since δ_k is embedded in $\hat{\mathbf{H}}_k$. By summarizing the above, the optimal solution for $\tilde{\mathbf{Q}}_k^{(2)}$ is given by $\tilde{\mathbf{Q}}_k^{(2)} = (\tilde{\mathbf{\Theta}}_k + \delta_k \tilde{\mathbf{C}}_{\mathbf{G}_{k,b}})^{-\frac{1}{2}} \hat{\mathbf{V}}_k \mathbf{B}_k \hat{\mathbf{V}}_k^H (\tilde{\mathbf{\Theta}}_k + \delta_k \tilde{\mathbf{C}}_{\mathbf{G}_{k,b}})^{-\frac{H}{2}}$. The optimal rate is then given by

$$R_k = \min \left\{ \min_{\ell \neq k} \eta_1 \log_2 \left| \mathbf{I}_{N_t} + \alpha_k \tilde{\mathbf{G}}_{k,\ell} \tilde{\mathbf{G}}_{k,\ell}^H \right|, \eta_2 \log_2 \left| \mathbf{I}_{N_r} + \tilde{\mathbf{H}}_k \tilde{\mathbf{V}}_k \tilde{\mathbf{Q}}_k^{(2)} \tilde{\mathbf{V}}_k^H \tilde{\mathbf{H}}_k^H \right| \right\}. \quad (53)$$

The proposed BD cooperative precoding scheme with spatially white multicast input is summarized in Algorithm 1. Specifically, in Algorithm 1, we first check whether or not $\tilde{\mathbf{\Theta}}_k$, i.e., if $\mathrm{rank}(\tilde{\mathbf{\Theta}}_k) = M'$. If so, then the solution of μ_k, α_k, and $\{\beta_{k,m}\}_{m=1}^{M'}$ are computed for $\delta_k = 0$. If the solution in this case yields $\eta_1 \alpha_k \mathrm{tr}(\mathbf{C}_{\mathbf{G}_{k,b}}) + \eta_2 \mathrm{tr}(\tilde{\mathbf{C}}_{\mathbf{G}_{k,b}} \tilde{\mathbf{Q}}_k^{(2)}) < \mathrm{IT}_k$, then we are done. If not, then the algorithm continues with the bisection search over δ_k between the initial lower and upper bounds $\delta_{k,\mathrm{low}} = 0$ and $\delta_{k,\mathrm{up}} = \frac{D_k + O_k}{\mathrm{tr}(\mathbf{C}_{\mathbf{G}_{k,b}}) \ln 2} \max_{\ell \neq k} \mathrm{tr}(\tilde{\mathbf{G}}_{k,\ell} \tilde{\mathbf{G}}_{k,\ell}^H) - \frac{E_k N_t}{\mathrm{tr}(\mathbf{C}_{\mathbf{G}_{k,b}})}$.

V. Numerical Comparison

In this section, we evaluate the performance of the proposed Max-WRMEP policy through computer simulations. Notice

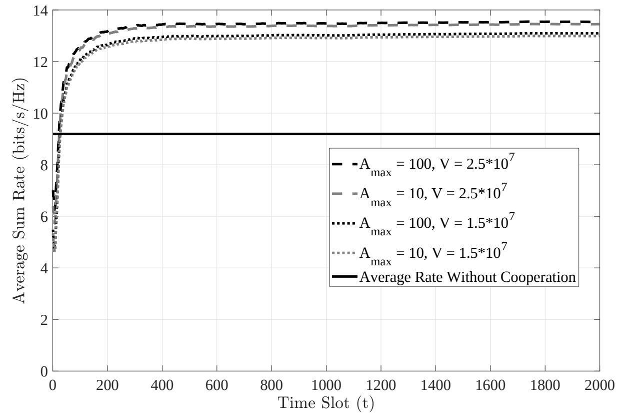


Fig. 2. Convergence of the average sum rate over time.

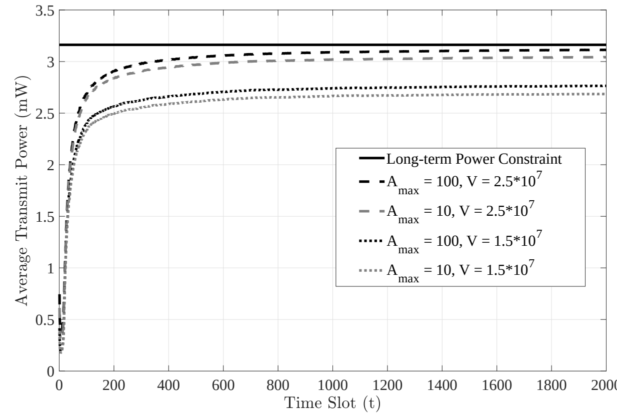


Fig. 3. Convergence of the average transmit power over time

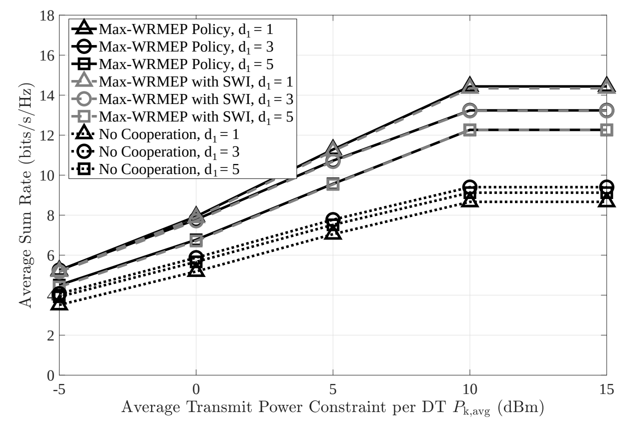


Fig. 4. Average sum rate comparison for different values of d_1 with $K = 4$, $d_2 = 30$, and $d_3 = 40$.

that cooperative D2D transmission is suitable for cases where the DTs are located close to each other, e.g., within a certain hot spot [29]–[32]. To focus on such a scenario, we assume in these experiments that the DTs are distributed according to a uniform distribution within a circular hot spot of radius d_1 meters centered at location $(-150, -150)$ and the DRs are also uniformly distributed within a ring distanced between d_2 and d_3 meters from the center of the hot spot. The BS is located at the origin and the active cell user is located randomly according to a uniform distribution within a circular area of radius $R_0 = 300$ meters. The center of the hot spot $(-150, -150)$ is only chosen so that it is more or less in the middle between the center and the edge of the cell. The number of transmit antennas at each DT and that at the CU, i.e., N_t and N_c, are both equal to 3, and the number of receive antennas at each DR and that at BS are $N_r = 3$ and $N_b = 3K$, respectively, where K is number of D2D pairs. The entries in $\mathbf{G}_{k,\ell}$ are assumed to be i.i.d. $\mathcal{CN}(0, d_{k,\ell}^{-2})$, where $d_{k,\ell}$ is the distance between nodes k and ℓ. The statistics of other channel matrices are given similarly. The results are averaged over 20 different user locations and 10^4 channel realizations per location. Moreover, $\mathbf{Q}_c[t]$ is chosen to maximize the point-to-point transmission rate between CU and BS in the absence of D2D interference under the transmit power constraint $P_{\rm CU} = {\rm tr}(\mathbf{Q}_c[t]) = 24$ dBm. The noise power is $\sigma_n^2 = -110$ dBm at all nodes. The IT constraint is set as ${\rm IT}_k = \sqrt{10} P_{\rm CU} 300^{-2}/K$, for all k, so that a target SINR of -5 dB can be achieved by cell edge users [58].

In Figs. 2 and 3, the convergence of the proposed algorithm is shown in terms of the average sum rate and the average transmit power for the case where DTs are equally spaced on the boundary of a circle with radius $1.5\,$m centered at $(-150, -150)$, DRs are $35\,$m away from the center (aligned with their corresponding DTs), and CU is placed at $(200, 0)$. The long-term average transmit power constraint $P_{k,\rm avg}$ is set as 5 dBm. Recall from (58) and (59) that the proposed Lyapunov optimization method aims to minimize an upper bound of the drift minus the V-weighted objective value. Hence, a larger utility (i.e., sum rate) is achieved with a larger choice of V. This is also shown explicitly in Theorem 1. In fact, when V is large, less emphasis is put on minimizing the drift and, thus, the average transmit power (i.e., the arrival into the virtual energy queue) will be larger. This results in a larger backlog for other queues as well, as indicated in Theorem 2. We can also see that a larger utility can be achieved by choosing $A_{\max}$ to be larger since, in this case, the additional constraint on the data arrival is loosened and, thus, the solution becomes closer to that of the original problem in (17).

In Figs. 4 and 5, the average sum rate of the proposed Max-WRMEP with spatially white input (Max-WRMEP with SWI) is compared to that of the case with no cooperation for different choices of d_1 and d_2, respectively. Specifically, in Fig. 4, we show the average sum rate versus the average transmit power constraint per DT for the case with $K = 4$, $d_2 = 30$, $d_3 = 40$, and different values of d_1. The proposed Max-WRMEP policy with optimized multicast input in problem (34), which is solved numerically by CVX, is also shown for comparison. Due to higher computational complexity, the results obtained from CVX are only averaged over 20 different user locations and 5000 channel realizations per location. We can observe that the average sum rate is larger when d_1 is smaller (i.e., when the phase 1 transmission rate can be higher). Notice that the average sum rate for the case with no cooperation is not significantly affected by the choice of d_1 since there is no need for data exchange among DTs. We can also see that the low-complexity Max-WRMEP with SWI

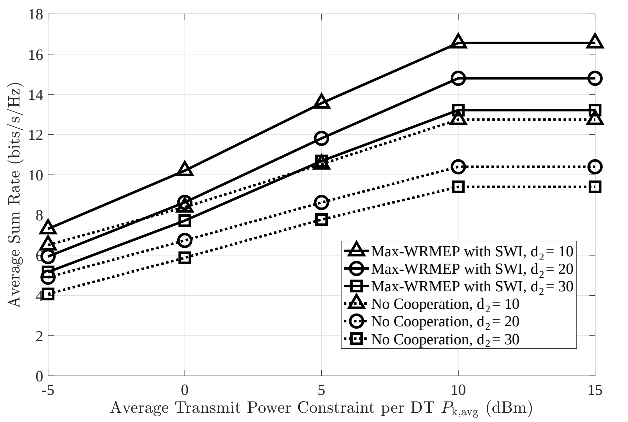


Fig. 5. Average sum rate comparison for different values of d_2 with $K = 4$, $d_1 = 3$, and $d_3 = d_2 + 10$.

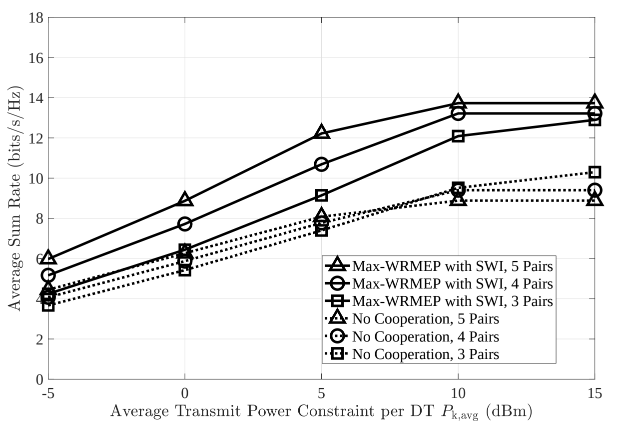


Fig. 6. Average sum rate comparison for different K, with $d_1 = 3$, $d_2 = 30$, and $d_3 = 40$.

policy performs close to the optimal Max-WRMEP policy solved by CVX. Similarly, in Fig. 5, we show the average sum rate versus the average transmit power constraint per DT for the case with $K = 4$, $d_1 = 3$, and different values of d_2 (and, thus, d_3, which is set as $d_2 + 10$). In this case, larger average sum rates are also observed for smaller values of d_2. In both figures, we can see that, even though the average sum rates increase with $P_{k,\text{avg}}$, they eventually saturate (for $P_{k,\text{avg}} \geq 10$ dBm) due to the IT constraint.

In Fig. 6, we show the average sum rate versus the average transmit power constraint per DT for the case with $d_1 = 3$, $d_2 = 30$, $d_3 = 40$, and different values of K. We can see that, for both cases with and without cooperation, the average sum rate increases monotonically with $P_{k,\text{avg}}$, but saturates earlier when K is larger since the total interference that can be allowed at the BS is fixed. For the case without cooperation, this effect causes the case with $K = 3$ to outperform the case with $K = 4$, $K = 5$ for $P_{k,\text{avg}}$ that is sufficiently large.

Up to this point, we have focused on the ideal scenario where perfect CSI is available for the computation of the precoding matrices. In particular, the CSI required for this computation can be denoted by CSI_1 to CSI_5, as defined at the bottom of Fig. 7. In practice, the channel acquisition procedure requires non-negligible overhead for pilot signalling and CSI feedback. For example, by considering the case where the

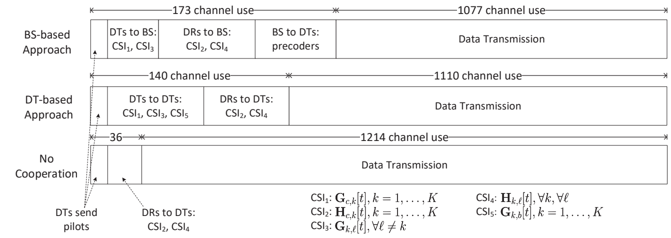


Fig. 7. Example of frame structure for channel acquisition.

computation is performed at BS (i.e., the BS-based approach), the required CSI can be acquired as follows[4]:

Step 1: DTs take turns emitting their respective pilot symbols to other DTs, DRs, and BS. (This enables DTs to estimate CSI_3, DRs to estimate CSI_4, and BS to estimate CSI_5).

Step 2: DTs take turns broadcasting their local estimates of CSI_1 and CSI_3 to BS.

Step 3: DRs take turns broadcasting their local estimates of CSI_2 and CSI_4 to BS.

Step 4: BS computes and broadcasts the precoders $\{\mathbf{W}_k^{(1)}[t]\}$ and $\{\mathbf{W}_k^{(2)}[t]\}$ to DTs.

Note that one can also consider the DT-based approach where the computation is performed directly at DTs. In this case, Step 4 can be avoided since the precoders are computed directly at the DTs. In the conventional *non-cooperative* case, CSI_1, CSI_2, part of CSI_4 (i.e., the intra-pair channels), and CSI_5 are also required at the DTs to compute their respective precoding matrices. Hence, only CSI_3 and the inter-pair channels in CSI_4 are additional requirements for cooperation. The frame structure for the abovementioned CSI acquisition procedure is illustrated in Fig. 7.

Let us consider the case with $K = 4$ D2D pairs, $N_t = N_r = N_c = 3$ antennas at each DT, DR and CU, and $N_b = KN_t = 12$ antennas at BS. We assume that each complex value is quantized into 32 bits (i.e., 16 bits each for the real and the imaginary parts). By approximating the transmission rate between nodes a and b, for $a, b \in \{\text{BS}, \text{DT}, \text{DR}\}$, as $R_{a,b} = N_t \log(1 + P_a/(d_{a,b}^2 \sigma_n^2))$, and by setting the transmit powers as $P_{\text{DT}} = P_{\text{DR}} = 24$ dBm and $P_{\text{BS}} = 40$ dBm, and the distances as $d_{\text{DR,BS}} = d_{\text{DT,BS}} + 30 = 250$ m and $d_{\text{DT,DR}} = d_{\text{DT,DT}} + 40 = 50$ m (i.e., the maximum distances between different types of nodes in our experiments), the total number of channel uses required for channel acquisition and feedback is 173 for the BS-based approach and 140 for the DT-based approach. By considering a channel with bandwidth $W = 1$ MHz and coherence time $T_c = 2.5$ ms (i.e., the case with user mobility equal to 64 km/hr [59]), the overhead occupies 13.8 and 11.2 percent of the coherence interval of 1250 channel uses respectively for the BS-based and DT-based approaches.

In Fig. 8, we show the impact of the channel acquisition overhead as well as the channel estimation imperfections for

[4]Note that DTs' estimate of CSI_1 and CSI_5 and DRs' estimate of CSI_2 can be obtained by overhearing the pilot signals emitted by CU and BS in the uplink and downlink frames. Hence, they are assumed to be available at their respective locations and, thus, are not considered explicitly in the D2D channel acquisition procedure described here.

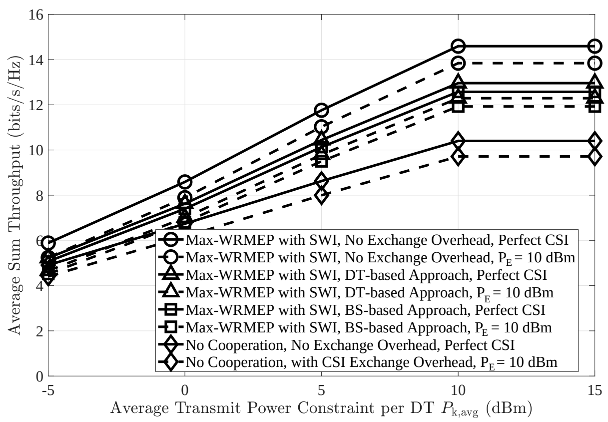


Fig. 8. Average sum rate in the presence of channel acquisition overhead and imperfect CSI for $K = 4$, $d_1 = 3$, $d_2 = 20$, and $d_3 = 30$. The MMSE estimator with pilot power $P_E = 10$ dBm is adopted here when channel estimation is considered.

the case where $K = 4$, $d_1 = 3$ m, $d_2 = 20$ m, and $d_3 = 30$ m. Here, we assume that the estimated channel matrices $\check{\mathbf{G}}_{\ell,k}[t]$ and $\check{\mathbf{H}}_{\ell,k}[t]$, for all ℓ, k, are used to compute the precoding matrices, and adopt the rate expressions in [60], [61], where the channel estimation error plus interference is treated as Gaussian noise. We can see that the proposed scheme still outperforms the non-cooperation case even when taking into account the overhead required for CSI acquisition and feedback, and the impact of imperfect CSI.

VI. Conclusion

In this paper, a dynamic cooperative transmission policy was proposed for multiple underlaying D2D pairs using Lyapunov optimization. The proposed policy employs a two-phase transmission scheme that consists of a *data-sharing transmission* in phase 1 and a *cooperative joint transmission* in phase 2. The transmit precoders in both phases were jointly designed with the goal of maximizing the sum utility of the average transmission rates subject to long-term individual power constraints, long-term rate-gain constraints, and instantaneous IT constraints. By adopting the framework of Lyapunov optimization, virtual data, rate-gain, and energy queues were constructed to record the temporal states of the system. By doing so, the long-term cooperative precoding problems were reduced to solving a series of short-term *weighted-rate-minus-energy-penalty (WRMEP)* maximization subproblems. A low-complexity cooperative precoding scheme was proposed by assuming spatially white multicast input in phase 1. Theoretical performance guarantees and bounds on the virtual queue backlogs were also derived. Finally, the effectiveness of the proposed policies was evaluated through computer simulations. In addition to the cooperative transmission scheme proposed in this work, the problem of determining who can or should join the cooperative group of D2D pairs (i.e., issues of admission control and partner selection) can also be studied on top of our proposed cooperative transmission scheme. Moreover, the framework can also be extended to cases with multiple cooperative groups of D2D pairs, in which case, the interference coordination among different cooperative groups and the group association problem should be further examined.

Appendix A
Proof of Theorem 1

Let $\mathcal{S}[t] \triangleq \{\mathbf{D}[t], \mathbf{O}[t], \mathbf{E}[t]\}$ be the state of the virtual queues at time t and let the Lyapunov function be defined as

$$L(\mathcal{S}[t]) \triangleq \frac{1}{2}\sum_{k=1}^{K} D_k^2[t] + \frac{1}{2}\sum_{k=1}^{K} O_k^2[t] + \frac{1}{2}\sum_{k=1}^{K} E_k^2[t]. \quad (54)$$

Its one-slot conditional Lyapunov drift is then given by $\Delta(\mathcal{S}[t]) \triangleq \mathbb{E}\big[L(\mathcal{S}[t+1]) - L(\mathcal{S}[t]) \mid \mathcal{S}[t]\big]$.

First, by taking the square on both sides of (20), we have

$$\begin{aligned} D_k^2[t+1] &\leq (D_k[t]-R_k[t])^2 + A_k^2[t] + 2A_k[t](D_k[t]-R_k[t])^+ \\ &\leq D_k^2[t] + R_k^2[t] - 2D_k[t]R_k[t] + A_k^2[t] + 2A_k[t]D_k[t]. \end{aligned}$$

It follows that $D_k^2[t+1] - D_k^2[t] \leq R_k^2[t] + A_k^2[t] - 2D_k[t](R_k[t] - A_k[t])$. By summing over k and by taking the condition expectation given $\mathcal{S}[t]$, we get

$$\begin{aligned} &\frac{1}{2}\mathbb{E}\Big[\sum_{k=1}^{K} D_k^2[t+1] - \sum_{k=1}^{K} D_k^2[t]\Big|\mathcal{S}[t]\Big] \\ &\leq \frac{1}{2}\sum_{k=1}^{K}\mathbb{E}\Big[R_k^2[t] + A_k^2[t]\Big|\mathcal{S}[t]\Big] \\ &\quad - \sum_{k=1}^{K} D_k[t]\mathbb{E}\Big[R_k[t] - A_k[t]\Big|\mathcal{S}[t]\Big]. \end{aligned} \quad (55)$$

Notice that $A_k[t] \leq A_{k,\max}$ and that, due to the IT constraint on the transmit power and by the boundedness assumption on the channel gains, the rates $R_k[t]$ are bounded as well. Let us denote the bound on the transmission rate by $R_{k,\max}[t]$, which may depend on the channel state at that time. Hence, the first term in (55) can be bounded as

$$\begin{aligned} &\frac{1}{2}\sum_{k=1}^{K}\mathbb{E}\Big[R_k^2[t] + A_k^2[t]\Big|\mathcal{S}[t]\Big] \\ &\leq \frac{1}{2}\sum_{k=1}^{K}\mathbb{E}\Big[R_{k,\max}^2[t]\Big] + \frac{1}{2}\sum_{k=1}^{K} A_{k,\max}^2 \triangleq C_1, \end{aligned} \quad (56)$$

where C_1 is a bounded constant since $\{R_{k,\max}[t]\}_{t=1}^{\infty}$ is stationary. It follows that

$$\begin{aligned} &\frac{1}{2}\mathbb{E}\Big[\sum_{k=1}^{K} D_k^2[t+1] - \sum_{k=1}^{K} D_k^2[t]\Big|\mathcal{S}[t]\Big] \\ &\leq C_1 - \sum_{k=1}^{K} D_k[t]\mathbb{E}\Big[R_k[t] - A_k[t]\Big|\mathcal{S}[t]\Big]. \end{aligned} \quad (57)$$

Similarly, we have $\frac{1}{2}\mathbb{E}\big[\sum_{k=1}^{K} O_k^2[t+1] - \sum_{k=1}^{K} O_k^2[t]\big|\mathcal{S}[t]\big] \leq C_2 - \sum_{k=1}^{K} O_k[t]\mathbb{E}\big[R_k[t] - R_k^{\mathrm{nc}}[t]\big|\mathcal{S}[t]\big]$ and $\frac{1}{2}\mathbb{E}\big[\sum_{k=1}^{K} E_k^2[t+1] - \sum_{k=1}^{K} E_k^2[t]\big|\mathcal{S}[t]\big] \leq C_3 - \sum_{k=1}^{K} E_k[t]\mathbb{E}\big[\mathrm{P}_{k,\mathrm{avg}} - P_k[t]\big|\mathcal{S}[t]\big]$, where C_2 and C_3 are bounded constants.

It thus follows from the definition of the Lyapunov drift and the inequalities above that

$$\begin{aligned}&\Delta(\mathcal{S}[t]) - V\mathbb{E}\big[\Omega(\mathbf{A}[t])\big|\mathcal{S}[t]\big] - C\\ \leq\ & \sum_{k=1}^{K}\mathbb{E}\Big[D_k[t]A_k[t] + O_k[t]R_k^{\rm nc}[t] - E_k[t]\mathrm{P}_{k,\rm avg}\Big|\mathcal{S}[t]\Big]\\ &-\sum_{k=1}^{K}\mathbb{E}\Big[D_k[t]R_k[t] + O_k[t]R_k[t] - E_k[t]P_k[t]\Big|\mathcal{S}[t]\Big]\\ &- V\mathbb{E}\big[\Omega(\mathbf{A}[t])\big|\mathcal{S}[t]\big], \qquad (58)\end{aligned}$$

where $C \triangleq C_1 + C_2 + C_3$. Here, an additional term $V\mathbb{E}\big[\Omega(\mathbf{A}[t])\big|\mathcal{S}[t]\big]$ is subtracted from both sides of the inequality. By substituting $P_k[t]$ with that in (13) and by rearranging the terms, the right-hand-side can be written as

$$\begin{aligned}&-\mathbb{E}\Big[V\Omega(\mathbf{A}[t]) - \sum_{k=1}^{K}D_k[t]A_k[t]\Big|\mathcal{S}[t]\Big]\\ &-\sum_{k=1}^{K}\mathbb{E}\Big[(D_k[t]+O_k[t])R_k[t] - \eta_1\mathrm{tr}(E_k[t]\mathbf{Q}_k^{(1)}[t])\\ &-\eta_2\mathrm{tr}\Big(\sum_{\ell=1}^{K}E_\ell[t]\mathbf{\Theta}_\ell\mathbf{Q}_k^{(2)}[t]\Big)\Big|\mathcal{S}[t]\Big]\\ &+\sum_{k=1}^{K}\mathbb{E}\Big[O_k[t]R_k^{\rm nc}[t] - E_k[t]\mathrm{P}_{k,\rm avg}\Big|\mathcal{S}[t]\Big], \qquad (59)\end{aligned}$$

which is minimized by (25) and (26) of the proposed Max-WRMEP policy. Hence, for any other alternative policy that yields arrival $\{A'_k[t],\forall k\}$ such that $0 \leq A'_k[t] \leq A_{k,\max}$, for all k, transmission rates $\{R'_k[t],\forall k\}$, and transmit powers $\{P'_k[t],\forall k\}$, it holds that

$$\begin{aligned}&\Delta(\mathcal{S}[t]) - V\mathbb{E}\big[\Omega(\mathbf{A}[t])\big|\mathcal{S}[t]\big] - C\\ \leq\ & \sum_{k=1}^{K}\mathbb{E}\Big[D_k[t]A'_k[t] + O_k[t]R_k^{\rm nc}[t] - E_k[t]\mathrm{P}_{k,\rm avg}\Big|\mathcal{S}[t]\Big]\\ &-\sum_{k=1}^{K}\mathbb{E}\Big[D_k[t]R'_k[t] + O_k[t]R'_k[t] - E_k[t]P'_k[t]\Big|\mathcal{S}[t]\Big]\\ &- V\mathbb{E}\Big[\Omega(\mathbf{A}'[t])|\mathcal{S}[t]\Big], \qquad (60)\end{aligned}$$

where $\mathbf{A}'[t] = [A'_1[t],\ldots,A'_K[t]]$.

Note that the feasibility of (17) in the premise of the theorem implies the feasibility of (18). Therefore, by [36, Theorem 4.5], we know that, for any $\delta > 0$ there exists an $\mathcal{H}$-only policy (as per Definition 2) with $A_k(\mathcal{H}[t]) \in [0, A_{k,\max}]$, $\forall k$, and $\{\mathbf{Q}_k^{(1)}(\mathcal{H}[t]), \mathbf{Q}_k^{(2)}(\mathcal{H}[t]), R_k(\mathcal{H}[t])\} \in \Gamma(\mathcal{H}[t], \mathrm{IT}_k)$, $\forall k, t$, such that

$$-\mathbb{E}[\Omega(\mathbf{A}'[t])] \leq -\omega_{\rm opt}(\mathbf{A}_{\max}) + \delta, \qquad (61)$$
$$\mathbb{E}[A'_k[t] - R'_k[t]] \leq \delta,\ \forall k, \qquad (62)$$
$$\mathbb{E}[R_k^{\rm nc}[t] - R'_k[t]] \leq \delta,\ \forall k, \qquad (63)$$
$$\mathbb{E}[P'_k[t] - \mathrm{P}_{k,\rm avg}] \leq \delta,\ \forall k. \qquad (64)$$

By plugging the above inequalities into (60) and by taking $\delta \to 0$, we have

$$\Delta(\mathcal{S}[t]) - V\mathbb{E}\big[\Omega(\mathbf{A}[t])\big|\mathcal{S}[t]\big] \leq C - V\omega_{\rm opt}(\mathbf{A}_{\max}). \qquad (65)$$

Then, by taking the expectation, summing over t, and rearranging the terms, we have

$$\frac{1}{T}\sum_{t=1}^{T}\mathbb{E}\big[\Omega(\mathbf{A}[t])\big] \geq \omega_{\rm opt}(\mathbf{A}_{\max}) - \frac{C}{V} - \frac{\mathbb{E}\big[L(\mathcal{S}[1])\big]}{VT}.$$

Therefore, with $L(\mathcal{S}[1]) < \infty$, it follows from Jensen's inequality that

$$\liminf_{T\to\infty}\Omega\left(\frac{1}{T}\sum_{t=1}^{T}\mathbb{E}[\mathbf{A}[t]]\right) \geq \omega_{\rm opt}(\mathbf{A}_{\max}) - \frac{C}{V}. \qquad (66)$$

Also, by rearranging (65) and by the fact that $\mathbb{E}\big[\Omega(\mathbf{A}[t])\big|\mathcal{S}[t]\big] \leq \Omega(\mathbf{A}_{\max})$, we have

$$\Delta(\mathcal{S}[t]) \leq C + V(\Omega(\mathbf{A}_{\max}) - \omega_{\rm opt}(\mathbf{A}_{\max})). \qquad (67)$$

Hence, by [36, Theorem 4.1], we know that all queues are mean rate stable which implies that the long-term constraints in (17c), (17d), and (18c) are achieved. The bound in (27) thus follows from (18c), (66), and the fact that $\Omega(\cdot)$ is continuous and entry-wise non-decreasing.

Appendix B
Proof of Theorem 2

By plugging the inequalities in (28)-(30) into (60), we have

$$\begin{aligned}&\Delta(\mathcal{S}[t]) - V\mathbb{E}\big[\Omega(\mathbf{A}[t])\big|\mathcal{S}[t]\big] - C\\ &\leq \sum_{k=1}^{K}\mathbb{E}[D_k[t](A_k(\mathcal{H}[t]) - R_k(\mathcal{H}[t]))\big|\mathcal{S}[t]]\\ &\quad+\sum_{k=1}^{K}\mathbb{E}[O_k[t](R_k^{\rm nc}[t] - R_k(\mathcal{H}[t]))\big|\mathcal{S}[t]]\\ &\quad+\sum_{k=1}^{K}\mathbb{E}[E_k[t](P_k(\mathcal{H}[t]) - \mathrm{P}_{k,\rm avg})\big|\mathcal{S}[t]] - V\mathbb{E}[\Omega(\mathbf{A}(\mathcal{H}[t]))]\\ &\leq -\epsilon\sum_{k=1}^{K}\mathbb{E}\big[(D_k[t] + O_k[t] + E_k[t])\big|\mathcal{S}[t]\big] - V\omega_\epsilon. \qquad (68)\end{aligned}$$

By taking the expectation, the time average over t, and by rearranging the terms, we get

$$\begin{aligned}&\epsilon\frac{1}{T}\sum_{t=1}^{T}\mathbb{E}\big[\sum_{k=1}^{K}(D_k[t] + O_k[t] + E_k[t])\big] - C\\ &\leq V\frac{1}{T}\sum_{t=1}^{T}\mathbb{E}\big[\Omega(\mathbf{A}[t])\big] - V\omega_\epsilon + \frac{\mathbb{E}\big[L(\mathcal{S}[1])\big]}{T}. \qquad (69)\end{aligned}$$

Then, by taking the limit on both sides and dividing by ϵ, we obtain

$$\begin{aligned}&\limsup_{T\to\infty}\frac{1}{T}\sum_{t=1}^{T}\sum_{k=1}^{K}\big(\mathbb{E}[D_k[t]] + \mathbb{E}[O_k[t]] + \mathbb{E}[E_k[t]]\big)\\ &\leq \frac{C + V(\overline{\Omega(\mathbf{A})} - \omega_\epsilon)}{\epsilon}. \qquad (70)\end{aligned}$$

It thus follows by Definition 1 that the queues $\{D_k[t]\}_{t=1}^{\infty}$, $\{O_k[t]\}_{t=1}^{\infty}$ and $\{E_k[t]\}_{t=1}^{\infty}$, for all k, are strongly stable.

REFERENCES

[1] A. Asadi, Q. Wang, and V. Mancuso, "A survey on device-to-device communication in cellular networks," *IEEE Commun. Surveys Tuts.*, vol. 16, pp. 1801–1819, Fourthquarter 2014.

[2] K. Doppler, M. Rinne, C. Wijting, C. B. Ribeiro, and K. Hugl, "Device-to-device communication as an underlay to LTE-advanced networks," *IEEE Commun. Mag.*, vol. 47, pp. 42–49, Dec. 2009.

[3] P. Janis, V. Koivunen, C. B. Ribeiro, K. Doppler, and K. Hugl, "Interference-avoiding MIMO schemes for device-to-device radio underlaying cellular networks," in *Proc. of IEEE PIMRC*, pp. 2385–2389, Sept. 2009.

[4] H. Min, J. Lee, S. Park, and D. Hong, "Capacity enhancement using an interference limited area for device-to-device uplink underlaying cellular networks," *IEEE Trans. Wireless Commun.*, vol. 10, pp. 3995–4000, Dec. 2011.

[5] C. Yu, K. Doppler, C. B. Ribeiro, and O. Tirkkonen, "Resource sharing optimization for device-to-device communication underlaying cellular networks," *IEEE Trans. Wireless Commun.*, vol. 10, pp. 2752–2763, Aug. 2011.

[6] A. Ramezani-Kebrya, M. Dong, B. Liang, G. Boudreau, and S. H. Seyedmehdi, "Joint power optimization for device-to-device communication in cellular networks with interference control," *IEEE Trans. Wireless Commun.*, vol. 16, pp. 5131–5146, Aug. 2017.

[7] G. Yu, L. Xu, D. Feng, R. Yin, G. Y. Li, and Y. Jiang, "Joint mode selection and resource allocation for device-to-device communications," *IEEE Trans. Commun.*, vol. 62, pp. 3814–3824, Nov. 2014.

[8] Y. Huang, A. A. Nasir, S. Durrani, and X. Zhou, "Mode selection, resource allocation, and power control for D2D-enabled two-tier cellular network," *IEEE Trans. Commun.*, vol. 64, pp. 3534–3547, Aug. 2016.

[9] S. Shalmashi, G. Miao, and S. B. Slimane, "Interference management for multiple device-to-device communications underlaying cellular networks," in *Proc. of IEEE PIMRC*, pp. 223–227, Sept. 2013.

[10] S. Shalmashi, G. Miao, Z. Han, and S. B. Slimane, "Interference constrained device-to-device communications," in *Proc. of IEEE ICC*, pp. 5245–5250, June 2014.

[11] C. Yang, J. Li, P. Semasinghe, E. Hossain, S. M. Perlaza, and Z. Han, "Distributed interference and energy-aware power control for ultra-dense D2D networks: a mean field game," *IEEE Trans. Wireless Commun.*, vol. 16, pp. 1205–1217, Feb. 2017.

[12] A. Sendonaris, E. Erkip, and B. Aazhang, "User cooperation diversity. part i. system description," *IEEE Trans. Commun.*, vol. 51, pp. 1927–1938, Nov. 2003.

[13] J. N. Laneman, D. N. C. Tse, and G. W. Wornell, "Cooperative diversity in wireless networks: efficient protocols and outage behavior," *IEEE Trans. Inf. Theory*, vol. 50, pp. 3062–3080, Dec. 2004.

[14] A. Scaglione and Y.-W. Hong, "Opportunistic large arrays: cooperative transmission in wireless multihop ad hoc networks to reach far distances," *IEEE Trans. Signal Process.*, vol. 51, pp. 2082–2092, Aug. 2003.

[15] Y.-W. P. Hong, W.-J. Huang, and C.-C. J. Kuo, *Cooperative communications and networking: technologies and system design*. Springer US, 2010.

[16] S. Fazeli-Dehkordy, S. Shahbazpanahi, and S. Gazor, "Multiple peer-to-peer communications using a network of relays," *IEEE Trans. Sig. Process.*, vol. 57, pp. 3053–3062, Aug. 2009.

[17] N. Bornhorst, M. Pesavento, and A. B. Gershman, "Distributed beamforming for multi-group multicasting relay networks," *IEEE Trans. Sig. Process.*, vol. 60, pp. 221–232, Jan. 2012.

[18] Y. Cheng and M. Pesavento, "Joint optimization of source power allocation and distributed relay beamforming in multiuser peer-to-peer relay networks," *IEEE Trans. Sig. Process.*, vol. 60, pp. 2962–2973, June 2012.

[19] A. Ramezani-Kebrya, M. Dong, B. Liang, G. Boudreau, and R. Casselman, "Per-relay power minimization for multi-user multi-channel cooperative relay beamforming," *IEEE Trans. Wireless Commun.*, vol. 15, pp. 3187–3198, May 2016.

[20] L. Wei, G. Wu, and R. Q. Hu, "Multi-pair device-to-device communications with space-time analog network coding," in *Proc. of IEEE WCNC*, pp. 920–925, Mar. 2015.

[21] W. Xing, C. Wang, P. Wang, and F. Liu, "Energy efficiency in multi-source network-coded device-to-device cooperative communications," in *Proc. of IEEE ICUWB*, pp. 1–4, Oct. 2016.

[22] G. S. Rajan and B. S. Rajan, "A non-orthogonal cooperative multiple access (NCMA) protocol and low ML decoding complexity codes," in *Proc. of IEEE WCNC*, pp. 885–890, Mar. 2007.

[23] A. Ribeiro, R. Wang, and G. B. Giannakis, "Multi-source cooperation with full-diversity spectral-efficiency and controllable-complexity," *IEEE J. Sel. Areas Commun.*, vol. 25, pp. 415–425, Feb. 2007.

[24] Z. Gao, H. Lai, and K. J. R. Liu, "Differential space-time network coding for multi-source cooperative communications," *IEEE Trans. Commun.*, vol. 59, pp. 3146–3157, Nov. 2011.

[25] L. Wang, T. Peng, Y. Yang, and W. Wang, "Interference constrained relay selection of D2D communication for relay purpose underlaying cellular networks," in *Proc. of WiCOM*, pp. 1–5, Sept. 2012.

[26] Y. Cao, T. Jiang, and C. Wang, "Cooperative device-to-device communications in cellular networks," *IEEE Wireless Commun.*, vol. 22, pp. 124–129, June 2015.

[27] Q. Sun, L. Tian, Y. Zhou, J. Shi, and X. Wang, "Energy efficient incentive resource allocation in D2D cooperative communications," in *Proc. of IEEE ICC*, pp. 2632–2637, June 2015.

[28] C. Karakus and S. Diggavi, "Enhancing multiuser MIMO through opportunistic D2D cooperation," *IEEE Trans. Wireless Commun.*, vol. 16, pp. 5616–5629, Sept. 2017.

[29] B. Zhou, H. Hu, S. Huang, and H. Chen, "Intracluster device-to-device relay algorithm with optimal resource utilization," *IEEE Trans. Veh. Technol.*, vol. 62, pp. 2315–2326, June 2013.

[30] A. Adhikary, H. S. Dhillon, and G. Caire, "Massive-MIMO meets HetNet: Interference coordination through spatial blanking," *IEEE J. Sel. Areas Commun.*, vol. 33, pp. 1171–1186, June 2015.

[31] I. A. Hemadeh, M. El-Hajjar, S. Won, and L. Hanzo, "Layered multi-group steered space-time shift-keying for millimeter-wave communications," *IEEE Access*, vol. 4, pp. 3708–3718, 2016.

[32] Z. Chang, Y. Hu, Y. Chen, and B. Zeng, "Cluster-oriented device-to-device multimedia communications: joint power, bandwidth, and link selection optimization," *IEEE Trans. Veh. Technol.*, vol. 67, pp. 1570–1581, Feb. 2018.

[33] D. Gesbert, S. Hanly, H. Huang, S. S. Shitz, O. Simeone, and W. Yu, "Multi-cell MIMO cooperative networks: a new look at interference," *IEEE J. Sel. Areas Commun.*, vol. 28, pp. 1380–1408, Dec. 2010.

[34] J. Zhang, R. Chen, J. G. Andrews, A. Ghosh, and R. W. Heath, "Networked MIMO with clustered linear precoding," *IEEE Trans. Wireless Commun.*, vol. 8, pp. 1910–1921, Apr. 2009.

[35] R. Zhang, "Cooperative multi-cell block diagonalization with per-base-station power constraints," *IEEE J. Sel. Areas Commun.*, vol. 28, pp. 1435–1445, Dec. 2010.

[36] M. J. Neely, *Stochastic network optimization with application to communication and queueing systems*. Morgan and Claypool, 2010.

[37] L. Georgiadis, M. J. Neely, and L. Tassiulas, "Resource allocation and cross-layer control in wireless networks," *Found. Trends Netw.*, vol. 1, pp. 1–144, Apr. 2006.

[38] M. J. Neely, "Energy optimal control for time-varying wireless networks," *IEEE Trans. Inf. Theory*, vol. 52, pp. 2915–2934, July 2006.

[39] K. Ronasi, B. Niu, V. W. S. Wong, S. Gopalakrishnan, and R. Schober, "Throughput-efficient scheduling and interference alignment for MIMO wireless systems," *IEEE Trans. Wireless Commun.*, vol. 13, pp. 1779–1789, Apr. 2014.

[40] S. Lakshminarayana, M. Assaad, and M. Debbah, "Transmit power minimization in small cell networks under time average QoS constraints," *IEEE J. Sel. Areas Commun.*, vol. 33, pp. 2087–2103, Oct. 2015.

[41] H. Shirani-Mehr, G. Caire, and M. J. Neely, "MIMO downlink scheduling with non-perfect channel state knowledge," *IEEE Trans. Commun.*, vol. 58, pp. 2055–2066, July 2010.

[42] M. Gatzianas, L. Georgiadis, and L. Tassiulas, "Control of wireless networks with rechargeable batteries [transactions papers]," *IEEE Trans. Wireless Commun.*, vol. 9, pp. 581–593, Feb. 2010.

[43] Y. Mao, J. Zhang, and K. B. Letaief, "A Lyapunov optimization approach for green cellular networks with hybrid energy supplies," *IEEE J. Sel. Areas Commun.*, vol. 33, pp. 2463–2477, Dec. 2015.

[44] S. Shim, J. S. Kwak, R. W. Heath, and J. G. Andrews, "Block diagonalization for multi-user MIMO with other-cell interference," *IEEE Trans. Wireless Commun.*, vol. 7, pp. 2671–2681, July 2008.

[45] S. Kim and G. B. Giannakis, "Optimal resource allocation for MIMO ad hoc cognitive radio networks," *IEEE Trans. Inf. Theory*, vol. 57, pp. 3117–3131, May 2011.

[46] I. CVX Research, "CVX: Matlab software for disciplined convex programming, version 2.0 beta.," 2012.

[47] H. Yoon, S. Park, and S. Choi, "Efficient feedback mechanism and rate adaptation for LTE-based D2D communication," in *Proc. of IEEE WoWMoM*, pp. 1–9, June 2017.

[48] F. Alavi, N. M. Yamchi, M. R. Javan, and K. Cumanan, "Limited feedback scheme for device-to-device communications in 5G cellular

networks with reliability and cellular secrecy outage constraints," *IEEE Trans. Veh. Technol.*, vol. 66, pp. 8072–8085, Sept. 2017.

[49] R. AliHemmati, B. Liang, M. Dong, G. Boudreau, and S. H. Seyedmehdi, "Long-term power allocation for multi-channel device-to-device communication based on limited feedback information," in *Proc. of Asilomar Conf. on Signals, Syst. and Comput.*, pp. 729–733, Nov. 2016.

[50] Y. Li, T. Jiang, M. Sheng, and Y. Zhu, "QoS-aware admission control and resource allocation in underlay device-to-device spectrum-sharing networks," *IEEE J. Sel. Areas Commun.*, vol. 34, pp. 2874–2886, Nov. 2016.

[51] X. Gong, S. A. Vorobyov, and C. Tellambura, "Joint bandwidth and power allocation with admission control in wireless multi-user networks with and without relaying," *IEEE Trans. Sig. Process.*, vol. 59, pp. 1801–1813, Apr. 2011.

[52] L. Wang and G. L. Stuber, "Pairing for resource sharing in cellular device-to-device underlays," *IEEE Network*, vol. 30, pp. 122–128, Mar. 2016.

[53] A. Nosratinia and T. E. Hunter, "Grouping and partner selection in cooperative wireless networks," *IEEE J. Sel. Areas Commun.*, vol. 25, pp. 369–378, Feb. 2007.

[54] N. Jindal and Z. Luo, "Capacity limits of multiple antenna multicast," in *Proc. of IEEE ISIT*, pp. 1841–1845, July 2006.

[55] G. R. Wood, "The bisection method in higher dimensions," *Math. Program.*, vol. 55, pp. 319–337, June 1992.

[56] S. Boyd and L. Vandenberghe, *Convex optimization*. Cambridge University Press, 2004.

[57] E. Telatar, "Capacity of multi-antenna Gaussian channels," *Eur. Trans. Telecommun.*, vol. 10, no. 6, pp. 585–595, 1999.

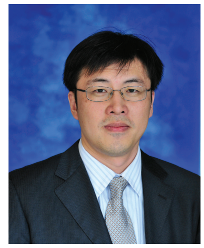
[58] A. Simonsson and A. Furuskar, "Uplink power control in LTE - overview and performance, subtitle: principles and benefits of utilizing rather than compensating for SINR variations," in *Proc. of IEEE VTC*, pp. 1–5, Sept. 2008.

[59] D. Tse and P. Viswanath, *Fundamentals of Wireless Communication*. Cambridge University Press, 2005.

[60] B. Hassibi and B. M. Hochwald, "How much training is needed in multiple-antenna wireless links?," *IEEE Trans. Inf. Theory*, vol. 49, pp. 951–963, April 2003.

[61] T. Yoo and A. Goldsmith, "Capacity and power allocation for fading MIMO channels with channel estimation error," *IEEE Trans. Inf. Theory*, vol. 52, pp. 2203–2214, May 2006.

Yung-Shun Wang received the B.S. degree in electrical engineering, and the M.S. and Ph.D. degrees in communications engineering from National Tsing Hua University, Hsinchu, Taiwan, in 2008, 2010 and 2019, respectively. He is currently a Postdoctoral Researcher with National Tsing Hua University, Hsinchu, Taiwan. His research interests include cooperative communications, signal processing for wireless networks, and machine learning for communications. He received the Best Paper Award from Asia Pacific Signal and Information Processing Association, Annual Summit and Conference in 2013.



Y.-W. Peter Hong (S'01 - M'05 - SM'13) received the B.S. degree in electrical engineering from National Taiwan University, Taipei, Taiwan, in 1999, and the Ph.D. degree in electrical engineering from Cornell University, Ithaca, NY, USA, in 2005. He joined the Institute of Communications Engineering and the Department of Electrical Engineering at National Tsing Hua University, Hsinchu, Taiwan, in 2005, where he is currently a Full Professor. He also currently serves as the Director of the Institute of Communications Engineering, and the Associate Vice President of Research and Development at National Tsing Hua University. His research interests include machine learning for wireless communications, physical layer security, distributed signal processing for IoT and sensor networks, and big data analytics.

Dr. Hong received the IEEE ComSoc Asia-Pacific Outstanding Young Researcher Award in 2010, the Y. Z. Hsu Scientific Paper Award in 2011, the National Science Council Wu Ta-You Memorial Award in 2011, the Chinese Institute of Electrical Engineering Outstanding Young Electrical Engineer Award in 2012, the Best Paper Award from the Asia-Pacific Signal and Information Processing Association Annual Summit and Conference (APSIPA ASC) in 2013, and the Ministry of Science and Technology Outstanding Research Award in 2019. He served as an Associate Editor of IEEE TRANSACTIONS ON SIGNAL PROCESSING and IEEE TRANSACTIONS ON INFORMATION FORENSICS AND SECURITY, and is currently an Editor of IEEE TRANSACTIONS ON COMMUNICATIONS. He was chair of the IEEE ComSoc Taipei Chapter in 2017-2018, and is currently co-chair of the Information Services Committee of the IEEE ComSoc Asia-Pacific Board.

Wen-Tsuen Chen (M'87 - SM'90 - F'94) received his B.S. degree in nuclear engineering from National Tsing Hua University, Taiwan, and M.S. and Ph.D. degrees in electrical engineering and computer sciences both from University of California, Berkeley, in 1970, 1973, and 1976, respectively. He has been with the Department of Computer Science of National Tsing Hua University since 1976 and served as Chairman of the Department, Dean of College of Electrical Engineering and Computer Science, and the President of National Tsing Hua University. In March 2012, he joined the Academia Sinica, Taiwan as a Distinguished Research Fellow of the Institute of Information Science until June 2018. Currently he is Sun Yun-suan Chair Professor of National Tsing Hua University. His research interests include computer networks, wireless sensor networks, mobile computing, and parallel computing.

Dr. Chen received numerous awards for his academic accomplishments in computer networking and parallel processing, including Outstanding Research Award of the National Science Council, Academic Award in Engineering from the Ministry of Education, Technical Achievement Award and Taylor L. Booth Education Award of the IEEE Computer Society, and is currently a lifelong National Chair of the Ministry of Education, Taiwan. Dr. Chen is the Founding General Chair of the IEEE International Conference on Parallel and Distributed Systems and the General Chair of the 2000 IEEE International Conference on Distributed Computing Systems among others. He is a Fellow of the Chinese Technology Management Association.