This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TCOMM.2017.2743689, IEEE Transactions on Communications

The Value of Full-duplex for Cellular Networks: A Hybrid Duplex based Study

Juan Liu, Shengqian Han, Wenjia Liu, and Chenyang Yang

***Abstract*—Recent work has demonstrated the gain of full-duplex (FD) network over half-duplex (HD) network in bidirectional sum throughput under the assumption of symmetric uplink-downlink traffic demands and perfect self-interference suppression (SIS). In this paper, we study the performance gain of FD network over HD network under asymmetric bidirectional traffic demands and non-ideal SIS. To this end, we investigate the traditional static time division duplex (TDD) transmission mode and the advanced dynamic TDD transmission mode to obtain the performance of HD network, and investigate the pure FD transmission mode and a flexible HD-FD hybrid transmission mode, namely XD mode, to obtain the performance of FD network. We use the number of users supported by a network as performance metric, which is defined as the minimum of the weighted numbers of users supported in uplink and downlink given random traffic demands of users. To maximize the number of users, we optimize the bidirectional transmit power for pure FD mode, bidirectional time slot configuration for dynamic TDD mode, and both for XD mode. Numerical results show an evident gain of pure FD mode and XD mode over static TDD mode for different levels of traffic asymmetry, but the gain over dynamic TDD mode is marginal, which cannot justify the application of FD technology in cellular systems without advanced interference control mechanisms.**

***Index Terms*—Full duplex, hybrid duplex, dynamic TDD, traffic asymmetry.**

I. Introduction

The fifth-generation (5G) wireless networks will achieve 1000-fold increase in network capacity and 10∼100-fold increase in number of connected devices [1]. To meet the challenging requirements, a host of novel techniques are being considered for future 5G systems, among which small cell network (SCN) and full duplex (FD) communication have attracted wide attention as two potential candidates.

The benefits of SCN have been well explored in the literature, which can significantly enhance spatial reuse and thus system capacity through network densification [2]. As a direct consequence of reducing cell size, the variation in traffic demands between different cells is more fluctuated [3]. Moreover, time-varying traffic asymmetry for uplink and downlink within each cell is expected due to the proliferation of diverse applications for smart wireless devices, users' mobility, and application usage behavior [4]. Therefore, technologies to deal with the spatio-temporally evolving bidirectional traffic asymmetry are needed for 5G systems [5].

This work was supported by National Natural Science Foundation of China (No. 61301084).

The authors are with the School of Electronics and Information Engineering, Beihang University, Beijing, 100191, China (e-mail: {liujuan, sqhan, liuwenjia, cyyang}@buaa.edu.cn). Part of this work was presented at IEEE ICC Workshops 2016. (Corresponding author: Shengqian Han)

Frequency division duplex (FDD) systems are generally recognized as difficult to handle bidirectional traffic asymmetry because of the equally paired frequency usage in downlink and uplink. Asymmetric FDD carrier aggregation by using more frequency bands in one direction can solve the problem to a certain extent, which however requires more transceivers at users, leading to increased cost and power consumption [5]. By contrast, time division duplex (TDD) systems have the capability to handle asymmetric bidirectional traffic demands. In traditional macro cell deployment scenarios, an asymmetric time slot configuration can be employed for uplink and downlink, which however needs to be used in all cells across the entire network in order to avoid the detrimental opposite-directional interference, e.g., the interference generated by a downlink transmitting base station (BS) to an uplink receiving BS. The synchronous time slot configuration, referred to as static TDD, is highly inefficient in SCN because of the spatio-temporal traffic asymmetry, which motivates the investigation of cell-specific dynamic TDD technology [6,7]. The mitigation of opposite-directional interference is a key challenge for dynamic TDD, and various interference control methods have been proposed in the literature, such as time slot allocation [8], power control [9], access control [10], coordinated beamforming [11], and cell clustering [3, 12]. In [13], the performance of dynamic TDD was analyzed, where a considerable gain of dynamic TDD over static TDD was demonstrated under asymmetric bidirectional traffic demands.

FD communications can be regarded as an enhanced dynamic TDD in the sense that both uplink and downlink in each cell operate simultaneously so that time slot allocation is no longer necessary, which therefore can naturally support asymmetric bidirectional traffic demands. FD communications were long believed impossible in wireless system design due to the severe self-interference within the same transceiver. However, the plausibility of FD technology was approved by recent tremendous progress in self-interference suppression (SIS), e.g., [14] and references therein, where it was reported that the self-interference can be suppressed to noise level. When applied in short-range point-to-point communications, FD system was demonstrated to achieve nearly doubled link performance over half duplex (HD) system [15, 16]. FD technology was also applied in relay systems to improve coverage and spectral efficiency, which can avoid the waste of resources in HD relay systems [17].

Although the promised benefits of FD technology have been established for point-to-point communications and relay systems, when applying to cellular network, the performance gain of FD network over traditional HD network will be

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TCOMM.2017.2743689, IEEE Transactions on Communications

impacted by new types of interference, i.e., BS-BS and user-user interference. The performance of FD network has been analyzed based on stochastic geometry model in the literature [18–24]. In [18], the performance of FD network with FD BSs and FD users was analyzed, and it is shown that FD network can provide large gain for both uplink and downlink compared to HD network when the self-interference is well cancelled. In [19], a FD network with FD BSs and HD users was considered, and the authors concluded that FD network can nearly double the downlink rate under perfect SIS but its uplink performance is susceptible to the BS-BS interference if the pathloss from adjacent transmitting BSs to a receiving BS is much smaller than that to users, e.g., in a macro-cell network with BSs deployed on the top of towers or building rooftops. The impact of different pathloss on uplink performance is alleviated in SCN due to the lower height of BS antennas. Further considering that the users and small BSs have comparable transmit power, it has been shown in [25, 26] that FD can be viable in SCN. In [20], the bidirectional sum throughput of a hybrid duplex heterogeneous network was investigated, where part of the BSs operate in FD mode communicating with FD users while other BSs operate in HD downlink mode transmitting to HD users. It was shown in [20] that the hybrid mode can achieve an evident gain over HD mode under well SIS and high BS density. Such a hybrid duplex mode was also analyzed in [21] for a wireless ad-hoc network, and it is concluded that under perfect SIS the hybrid network can achieve higher throughput than HD network but achieving the double throughput is impossible. In [22, 23], a so-called α-duplex scheme was studied and its superiority was demonstrated, with which a BS may operate in FD mode, HD uplink mode or HD downlink mode in different frequency bands. In [24], a hybrid duplex network with directional transmission and reception was investigated, where a FD BS may serve one FD user or two HD users. It was shown in [24] that serving two HD users is more beneficial than serving one FD user under imperfect SIS, which coincides with the results in [22, 23]. The aforementioned works analyzing the advantages of FD network explicitly or implicitly employed the assumption that the uplink and downlink traffic demands are symmetric, e.g., in [18, 20]. When asymmetric bidirectional traffic demands are taken into account, the system-level simulation results in [27] and [28] showed that the gain of FD network over HD network will be compromised. In our preliminary work [29], the performance gain of pure FD network over HD network was analyzed and it was shown that pure FD network has evident gain over static TDD, but the gain over dynamic TDD decreases with the increase of bidirectional traffic asymmetry.

In this paper, we investigate the performance gain of FD network over HD network under bidirectional traffic asymmetry and non-ideal SIS. The main contributions are summarized as follows.

- We develop a framework for investigating the performance of FD network and HD network based on the Poisson Point Process (PPP) topology. First, we propose a flexible HD-FD hybrid transmission mode, namely XD mode, where each frame is divided into three phases for pure FD, pure HD downlink, and pure HD uplink, respectively. The time slots allocated to the three phases and the bidirectional powers for pure FD phase are configurable. It provides a unified way to optimize and compare the performance of FD and HD networks. The resultant optimization problem and solutions are completely different from [29] where only pure FD mode was optimized. Second, to measure the performance of FD and HD networks under asymmetric bidirectional traffic demands, we introduce the number of users supported by the networks as performance metric. Considering the randomness of traffic demands of users, the number of users is defined in a probability manner. We find an approach to explicitly express the probability model, and then formulate the optimization problems to maximize the number of supported users in different operating modes.
- The formulated problem for resource configuration in XD mode is non-convex. We propose an algorithm to find its global solution, based on which the optimal resource configurations for pure HD dynamic TDD mode and pure FD mode are obtained, respectively. Numerical results show that the level of SIS has large impact on the performance of pure FD mode, which is much worse than HD network for small SIS. The gain of XD mode over static TDD mode is significant, however, the gain over dynamic TDD mode is limited especially in the typical downlink traffic dominant scenario. This implies that applying FD technology to cellular network by simply replacing current HD BSs with FD BSs is inadequate, and effective suppression of BS-BS and user-user interference is indispensable for the deployment of FD network.

II. System Model

A. Network Model and Hybrid Duplex

Consider a small cell network where the cells are grouped into clusters according to their bidirectional traffic demands, which can be formed based on existing cell clustering methods designed for dynamic TDD, e.g., [12]. Cluster-specific uplink-downlink resource configuration is employed in each cluster based on the bidirectional traffic demands, and different configurations can be used between clusters.

Consider that the BSs have FD capability, which can operate in either FD or HD mode, while the users can only operate in HD mode. Under imperfect SIS, such a three-node architecture for FD communications, i.e., one FD BS serving two HD users, has been shown to outperform the two-node architecture, i.e., one FD BS serving one FD user, in terms of either performance, complexity or energy consumption [24]. Suppose that time division multiple access (TDMA) is employed to serve multiple users within each cell. Then, within every time slot each BS will schedule one uplink or downlink user if it operates in HD mode, and will schedule one uplink user and one downlink user simultaneously if operating in FD mode.

We investigate a flexible HD-FD hybrid duplex scheme, where the BSs can operate in pure FD mode via simultaneously transmitting and receiving, pure HD uplink receiving

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TCOMM.2017.2743689, IEEE Transactions on Communications

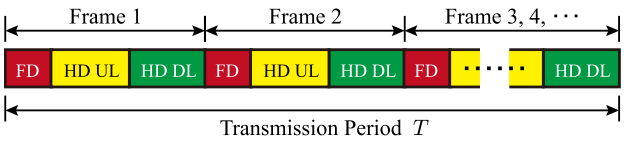

Fig. 1. Frame structure of XD mode in a transmission period.

mode, and pure HD downlink transmitting mode. We denote the hybrid duplex scheme by XD mode for notational simplicity. Let $T_{\rm f}$, $T_{\rm hu}$ and $T_{\rm hd}$ denote the fraction of time slots allocated to the three modes within each frame, respectively, $0 \le T_{\rm f}, T_{\rm hu}, T_{\rm hd} \le 1$, and $T_{\rm f} + T_{\rm hu} + T_{\rm hd} = 1$. The frame structure of the XD mode is shown in Fig. 1.

In this paper, we study the resource configuration for XD mode, including the optimization of both time slot fractions $T_{\rm f}$, $T_{\rm hu}$ and $T_{\rm hd}$ and transmit powers of BSs and users. The configuration keeps constant within a so-called transmission period T, during which the traffic demands in a cluster are stable. According to the measurement based traffic model in [30], one can compute that the transmission period T can last tens of minutes or even hours, within which the variation of traffic demands is less than 5%. Thus, we can implement the resource configuration for XD in a semi-dynamic manner.

To analyze the benefits of introducing FD communications in cellular networks, we need to optimize and compare the performance of XD mode, pure FD mode, HD dynamic TDD mode, and HD static TDD mode. Note that the latter three modes are special cases of XD mode. In particular, XD mode reduces to pure FD mode if $T_{\rm hu} = 0$ and $T_{\rm hd} = 0$, and reduces to pure HD dynamic or static TDD mode if $T_{\rm f} = 0$. Therefore, it is sufficient for us to only study the resource configuration problem for XD mode, from which the solutions to pure FD or pure HD mode can be obtained readily. Moreover, it can be expected that XD mode will never perform inferior to pure HD or FD mode, because it can adaptively select its operation mode, either FD mode, HD mode or a mixture of them, according to the bidirectional traffic demands in a cluster. This can be accomplished with the resource configuration algorithm proposed in next section.

B. Average Bidirectional Data Rates

In this subsection we analyze the average downlink and uplink data rates of a user achieved during the long-term transmission period T based on stochastic geometry model. We first obtain the complementary cumulative distribution functions (CCDF) of the downlink and uplink signal-to-interference-plus-noise ratio (SINR) of the user based on existing works [22,23,31], and then obtain its average bidirectional data rates.

The stochastic geometry model assumes that the interference is originated from an infinite network, which can be exactly applied to our problem for the scenarios where all cells have similar traffic demands so that the cluster has a very large size. When taking into account the uneven spatial distribution of traffic demands, it will be more beneficial to form smaller cell clusters. Since the resource configuration in different clusters may differ, e.g., one cluster operates in FD mode while an adjacent cluster may operate in HD downlink mode, using stochastic geometry model, which implicitly assumes all clusters have the same configuration, may lead to a certain performance bias. Nevertheless, the performance bias is not large for the considered SCN, because users and BSs have comparable transmit powers so that the difference of the interference generated from an adjacent cluster operating in different modes is generally not significant. Thus, in the following we still employ the stochastic geometry model regardless of the cluster size. The accuracy of the analytical results will be verified by simulations later.

Assume that the BSs are randomly located according to a homogeneous PPP Ω of density λ in the Euclidean plane. Each user is associated with the closest BS in both uplink and downlink. Due to the restriction that each cell schedules only one uplink user in each time slot, the locations of all scheduled uplink users in the same time slot are in fact dependent in the sense that the presence of one uplink user in a particular cell prohibits the presence of any other uplink users in the same cell. Such a location dependence is also true for all scheduled downlink users for the same reason. Moreover, the locations of uplink users and downlink users are coupled because it is implicitly assumed in FD mode that there always exists at least one uplink user located in the same cell where a downlink user presents for scheduling [22,23]. The dependence and coupling of locations of scheduled users make the analysis of network performance very complicated. For analytical tractability, as is often assumed in the literature, e.g., [22,23,31,32], we assume that the locations of the scheduled uplink or/and downlink users in each HD/FD time slot follow independent PPP Φ and Ψ, respectively, with the same density λ as the BSs. This assumption has been verified in [32] through a comparison with the real deployment of an urban 4G network. Note that this assumption does not mean that the numbers of uplink and downlink users served in the network are identical, because, for example, a downlink user with more data to transmit than an uplink user will occupy more time slots than the uplink user, leading to fewer number of downlink users served in the network. In addition, we assume standard power-law pathloss and independent and identically distributed (i.i.d.) Rayleigh fading channels as in [22, 23, 31].

We consider distance-proportional fractional power control for uplink [31]. The transmit power of an uplink user is set as $P_{\rm u} d^{\eta \alpha_{\rm bu}}$, where $\eta \in [0,1]$ is the power control factor, $\alpha_{\rm bu}$ is the pathloss exponent from the user to its serving BS, d is the distance between them, and $P_{\rm u}$ is the base power to ensure that a given percent of users, denoted by θ, in each cell can satisfy the maximal transit power constraint. Specifically, $P_{\rm u}$ is set to ensure $\mathbb{P}\{P_{\rm u} d^{\eta \alpha_{\rm bu}} \le P_{\rm u,max\,0}\} \ge \theta$, where $P_{\rm u,max\,0}$ is the maximum transmit power of users, and $\mathbb{P}\{\cdot\}$ denotes the probability. Let $P_{\rm u,max}$ denote the solution of $\mathbb{P}\{P_{\rm u} d^{\eta \alpha_{\rm bu}} \le P_{\rm u,max\,0}\} = \theta$, which can be easily found given the distribution of d. Then, we know $P_{\rm u} \le P_{\rm u,max}$ should hold. In the downlink all BSs transmit with a fixed power $P_{\rm d}$.

Assume that the users randomly move within their serving cell during the long-term transmission period T. Thus, the CCDF of the SINR of an arbitrary user is equivalent to the network coverage probability. We first consider the pure FD

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TCOMM.2017.2743689, IEEE Transactions on Communications

mode with $T_{\rm f}=1$, where each FD BS serves one uplink user and one downlink user in the same time slot. The downlink transmission experiences the interference from both other BSs and all co-scheduled uplink users. Following the derivations in [22, 23, 31], after some regular manipulations, we can obtain the CCDF of the downlink SINR of a user as

$$p_{\rm fd}(\gamma) = 2\pi\lambda \int_0^\infty r e^{-\lambda\pi r^2} e^{-\frac{r^{\alpha_{\rm bu}}\gamma N_0 W_0}{P_{\rm d}}} F_{1{\rm f}}(r,\gamma)\cdot F_{2{\rm f}}(r,\gamma,P_{\rm d},P_{\rm u}) F_{3{\rm f}}(r,\gamma,P_{\rm d},P_{\rm u}){\rm d}r, \quad (1)$$

where $F_{1{\rm f}}(r,\gamma)$, $F_{2{\rm f}}(r,\gamma,P_{\rm d},P_{\rm u})$, and $F_{3{\rm f}}(r,\gamma,P_{\rm d},P_{\rm u})$ denote the interference from inter-cell BSs, inter-cell uplink users, and intra-cell uplink user, respectively, $F_{1{\rm f}}(r,\gamma) = \exp\big(-\pi\lambda r^2\gamma^{\frac{2}{\alpha_{\rm bu}}}\int_{\gamma^{-2/\alpha_{\rm bu}}}^\infty \frac{1}{1+x^{\alpha_{\rm bu}/2}}dx\big)$, $F_{2{\rm f}}(r,\gamma,P_{\rm d},P_{\rm u}) = \exp\big(-2\pi\lambda\int_1^\infty\int_0^\infty \frac{2\pi\lambda z^3 x e^{-\pi\lambda z^2}}{1+x^{\alpha_{\rm uu}} r^{-\alpha_{\rm bu}}\gamma^{-1} z^{(\alpha_{uu}-\eta\alpha_{\rm bu})}\frac{P_{\rm d}}{P_{\rm u}}}{\rm d}z{\rm d}x\big)$, and $F_{3{\rm f}}(r,\gamma,P_{\rm d},P_{\rm u}) = \int_0^\infty\int_0^{2\pi}\frac{\lambda r_{\rm I} e^{-\pi\lambda r_{\rm I}^2}}{1+r^{\alpha_{\rm bu}}\gamma r_{\rm I}^{\eta\alpha_{\rm bu}}\bar{r}_{\rm I}^{-\alpha_{\rm uu}/2}\frac{P_{\rm u}}{P_{\rm d}}}{\rm d}\delta{\rm d}r_{\rm I}$ with $\bar{r}_{\rm I}\triangleq r_{\rm I}^2+r^2-2r_{\rm I}r\cos(\delta)$. W_0 is the system bandwidth, $\alpha_{\rm bu}$ and $\alpha_{\rm uu}$ denote the BS-user and user-user pathloss exponents, which are typically different due to the different antenna heights and propagation environments at BSs and users [22], N_0 is the noise power spectrum density, $r_{\rm I}$ denotes the distance between the intra-cell uplink user and its serving BS, and $\delta\in[0,2\pi]$ is the separation angle between the uplink and downlink users in the same cell without considering any scheduling.

The uplink transmission under pure FD mode suffers from the inter-cell interference generated by both other-cell uplink users and other BSs, and also the self-interference. As in [22], we model the attenuation power of self-interference as $\beta P_{\rm d} h_{\rm s}$, where $\beta P_{\rm d}$ is the average self-interference power after suppression, and $h_{\rm s}$ follows the exponential distribution to represent the uncertainty in SIS. The average power of the suppressed self-interference can be expressed as $10\log_{10}P_{\rm d} - 10\log_{10}(\beta P_{\rm d}) = -10\log_{10}\beta$ in units of dB, which can be more than 100 dB as reported in recent work [14]. As two extreme cases, $\beta=0$ means perfect SIS, while $\beta=1$ means no SIS.

Based on the results in [22, 23, 31], the CCDF of uplink SINR in FD mode can be obtained as

$$p_{\rm fu}(\gamma) = 2\pi\lambda \int_0^\infty r e^{-\lambda\pi r^2} e^{\frac{-r^{(1-\eta)\alpha_{\rm bu}}\gamma N_0 W_0}{P_{\rm u}}} F'_{1{\rm f}}(r,\gamma)\cdot F'_{2{\rm f}}(r,\gamma,P_{\rm d},P_{\rm u}) F'_{3{\rm f}}(r,\gamma,P_{\rm d},P_{\rm u}){\rm d}r, \quad (2)$$

where $F'_{1{\rm f}}(r,\gamma)$, $F'_{2{\rm f}}(r,\gamma,P_{\rm d},P_{\rm u})$, and $F'_{3{\rm f}}(r,\gamma,P_{\rm d},P_{\rm u})$ denote the interference from inter-cell uplink users, inter-cell BSs, and self interference, respectively, $F'_{1{\rm f}}(r,\gamma) = \exp\big(-2\pi\lambda\int_1^\infty\int_0^\infty \frac{2\pi\lambda z^3 x e^{-\pi\lambda z^2}}{1+r^{(\eta-1)\alpha_{\rm bu}}\gamma^{-1}z^{(1-\eta)\alpha_{\rm bu}}x^{\alpha_{\rm bu}}}{\rm d}z{\rm d}x\big)$, $F'_{2{\rm f}}(r,\gamma,P_{\rm d},P_{\rm u}) = \exp\big(-\pi\lambda r^{\frac{2\alpha_{\rm bu}(1-\eta)}{\alpha_{\rm bb}}}\big(\frac{P_{\rm d}}{P_{\rm u}}\gamma\big)^{2/\alpha_{\rm bb}}\int_{y>0}\frac{1}{1+y^{\alpha_{\rm bb}/2}}dy\big)$, and $F'_{3{\rm f}}(r,\gamma,P_{\rm d},P_{\rm u}) = \frac{1}{1+\frac{r^{(1-\eta)\alpha_{\rm bu}\gamma\beta P_{\rm d}}}{P_{\rm u}}}$. Herein, $\alpha_{\rm bb}$ denotes the pathloss exponent between BSs.

With the CCDF of downlink and uplink SINR, we can readily obtain the bidirectional average rates in FD mode in a transmission period by integrating over SINR, which are

$$\bar{R}_{\rm fd}(P_{\rm d},P_{\rm u}) = W_0\int_0^\infty p_{\rm fd}(e^t-1){\rm d}t, \quad (3a)$$

$$\bar{R}_{\rm fu}(P_{\rm d},P_{\rm u}) = W_0\int_0^\infty p_{\rm fu}(e^t-1){\rm d}t. \quad (3b)$$

From the average data rates under pure FD mode, we can easily obtain the average data rates under pure HD network, where $T_{\rm f}=0$ and $T_{\rm hu}+T_{\rm hd}=1$. Specifically, the average downlink rate under pure HD mode, denoted by $\bar{R}_{\rm hd}(P_{\rm d})$, can be obtained from (3a) by setting the transmit power of users as zero, i.e., $P_{\rm u}=0$. Similarly, the average uplink rate under pure HD mode, denoted by $\bar{R}_{\rm hu}(P_{\rm u})$, can be obtained from (3b) by setting the transmit power of BSs as zero, i.e., $P_{\rm d}=0$.

Under XD mode, the data of a user can be delivered in part by both FD mode and HD mode. We assume that the time slots that a BS uses to serve a user in FD and HD modes are proportional to the total time slots of FD and HD modes. Specifically, if the total time slots allocated to a downlink user is $\overline{T}$, then the BS will serve the user in FD mode with $\frac{T_{\rm f}}{T_{\rm f}+T_{\rm hd}}\overline{T}$ time slots and in HD mode with $\frac{T_{\rm hd}}{T_{\rm f}+T_{\rm hd}}\overline{T}$ time slots, respectively. Similarly, if the total time slots allocated to an uplink user is $\widetilde{T}$, then the BS will serve the user in FD mode with $\frac{T_{\rm f}}{T_{\rm f}+T_{\rm hu}}\widetilde{T}$ time slots and in HD mode with $\frac{T_{\rm hu}}{T_{\rm f}+T_{\rm hu}}\widetilde{T}$ time slots, respectively. Then, with $\bar{R}_{\rm fd}(P_{\rm d},P_{\rm u})$, $\bar{R}_{\rm fu}(P_{\rm d},P_{\rm u})$, $\bar{R}_{\rm hd}(P_{\rm d})$ and $\bar{R}_{\rm hu}(P_{\rm u})$, the average rates of a downlink user and an uplink user under XD mode within a transmission period can be obtained as

$$\bar{R}_{\rm d}(P_{\rm d},P_{\rm u},T_{\rm hd},T_{\rm f}) = \frac{T_{\rm hd}}{T_{\rm f}+T_{\rm hd}}\bar{R}_{\rm hd}(P_{\rm d}) + \frac{T_{\rm f}}{T_{\rm f}+T_{\rm hd}}\bar{R}_{\rm fd}(P_{\rm d},P_{\rm u}), \quad (4a)$$

$$\bar{R}_{\rm u}(P_{\rm d},P_{\rm u},T_{\rm hu},T_{\rm f}) = \frac{T_{\rm hu}}{T_{\rm f}+T_{\rm hu}}\bar{R}_{\rm hu}(P_{\rm u}) + \frac{T_{\rm f}}{T_{\rm f}+T_{\rm hu}}\bar{R}_{\rm fu}(P_{\rm d},P_{\rm u}). \quad (4b)$$

In this paper we focus on the analysis in the interference-limited scenario for mathematical tractability, which is relevant for SCN due to the proximity between BSs and users. We denote by $P_{\rm d\,min}$ the minimal transmit power of BSs and $P_{\rm u,min}$ the minimal base power of uplink power control, which ensure that the noise is negligible compared to the inter-cell interference. We will restrict our study to the scenarios with $P_{\rm d}\geq P_{\rm d,min}$ and $P_{\rm u}\geq P_{\rm u,min}$ for the subsequent resource configurations.

By setting $N_0=0$ in (3a) and (3b), we can find that the bidirectional average rates under pure FD mode only depend on the ratio of transmit powers of BSs and users. Defining $\kappa=\frac{P_{\rm d}}{P_{\rm u}}$, then the bidirectional average rates under pure FD mode, pure HD mode, and XD mode in the interference-limited scenario can be obtained as

$$\hat{R}_{\rm fd}(\kappa) = 2\pi W_0\lambda\int_0^\infty r e^{-\lambda\pi r^2}\int_0^\infty F_{1{\rm f}}(r,e^t-1)\cdot F_{2{\rm f}}(r,e^t-1,\kappa)F_{3{\rm f}}(r,e^t-1,\kappa){\rm d}t{\rm d}r, \quad (5a)$$

$$\hat{R}_{\rm fu}(\kappa) = 2\pi W_0\lambda\int_0^\infty r e^{-\lambda\pi r^2}\int_0^\infty F'_{1{\rm f}}(r,e^t-1)\cdot F'_{2{\rm f}}(r,e^t-1,\kappa)F'_{3{\rm f}}(r,e^t-1,\kappa){\rm d}t{\rm d}r, \quad (5b)$$

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TCOMM.2017.2743689, IEEE Transactions on Communications

$$\hat{R}_{\rm hd} = 2\pi W_0 \lambda \int_0^\infty r e^{-\lambda\pi r^2} \int_0^\infty F_{1\rm f}(r, e^t - 1) {\rm d}t {\rm d}r \triangleq \hat{R}_{\rm h0}, \tag{5c}$$

$$\hat{R}_{\rm hu} = 2\pi W_0 \lambda \int_0^\infty r e^{-\lambda\pi r^2} \int_0^\infty F'_{1\rm f}(r, e^t - 1) {\rm d}t {\rm d}r \triangleq \chi \hat{R}_{\rm h0}, \tag{5d}$$

$$\hat{R}_{\rm d}(\kappa, T_{\rm hd}, T_{\rm f}) = \frac{T_{\rm hd}}{T_{\rm f}+T_{\rm hd}} \hat{R}_{\rm h0} + \frac{T_{\rm f}}{T_{\rm f}+T_{\rm hd}} \hat{R}_{\rm fd}(\kappa), \tag{5e}$$

$$\hat{R}_{\rm u}(\kappa, T_{\rm hu}, T_{\rm f}) = \frac{T_{\rm hu}}{T_{\rm f}+T_{\rm hu}} \chi \hat{R}_{\rm h0} + \frac{T_{\rm f}}{T_{\rm f}+T_{\rm hu}} \hat{R}_{\rm fu}(\kappa), \tag{5f}$$

where we replace the term $\frac{P_{\rm d}}{P_{\rm u}}$ with κ in functions $F_{2\rm f}(\cdot)$, $F_{3\rm f}(\cdot)$, $F'_{2\rm f}(\cdot)$, and $F'_{3\rm f}(\cdot)$, and introduce a scalar χ to denote the uplink to downlink rate ratio in pure HD mode for notational simplicity.

Based on the works in [22, 23, 31], we have obtained the explicit expressions of bidirectional average rates in different modes. In next section, we will use them to formulate the resource configuration problem and solve it, which have not been addressed in the literature including [22, 23, 31].

C. Asymmetric Bidirectional Traffic

The proliferation of mobile smart devices and applications results in asymmetric average amount of traffic in uplink and downlink as reported by International Telecommunication Union (ITU) [4]. With the rapid growth of video streaming services, it is predicted by ITU that the traffic asymmetry value will be in the range of 1/7∼1/8 in favour of downlink in 2024 [4].

The traffic demands of different users are random because of diverse mobile devices and application usage behavior [4]. As measured and analyzed in [33], the traffic demands among users exhibit the "Pareto law" with less than 10% of the users creating 90% of the daily network traffic. This suggests that the traffic demand per user can be well characterized by some heavy-tailed distribution, e.g., the CDF of traffic demand per user showed in [33] can well fit the log-normal distribution.

There have been some studies regarding the mean and variance of the traffic demand per user. For example, in [34] the traffic statistics of mobile data networks around the world in second half of the year 2014 were reported. It was shown that the users in Asia-Pacific consumed the highest mobile data traffic than other regions, where the median and mean of the monthly traffic demand per user are 261.7 and 143.1 Mbyte for uplink and 298.1 and 1000 Mbyte for downlink, respectively. Based on Chebyshev's inequality, we can obtain a lower bound of the standard deviation as the absolute difference between mean and median, which is 118.6 and 701.9 Mbyte for uplink and downlink, respectively. It can be computed that the ratio of standard deviation and mean approximates 0.8 and 0.7 for uplink and downlink, respectively. As predicted by ITU, in 2020 the mean traffic demand per peak hour will approximate 2 kbps for uplink and 14 kbps for downlink.

III. Resource Configuration under Bidirectional Traffic Asymmetry

In order to compare the performance of FD network and HD network, in this section we optimize the resource configuration for different operation modes under asymmetric bidirectional traffic demands. We start by introducing the employed performance metric.

A. Performance Metric

In this paper, considering the random and asymmetric bidirectional traffic demands of users, we employ the number of users supported by a network during a long-term transmission period T as the performance metric to evaluate the serving capacity of the network. Since we are considering bidirectional communications, we define the number of users supported by a network, K, as the minimum of the weighted numbers of users supported in uplink, $K_{\rm u}$, and in downlink, $K_{\rm d}$, i.e., $K = \min(\mu K_{\rm u}, K_{\rm d})$, where $\mu > 0$ is the weight. For traditional networks only providing voice service, the serving capacity of the networks can be measured by the number of established bidirectional communication pairs, which is limited by the minimum of $K_{\rm u}$ and $K_{\rm d}$, i.e., μ can be set as 1. Unlike voice service, the uplink and downlink requests for data services do not necessarily present in pairs. For example, the services subscribed by users are automatically delivered in downlink without uplink requests, while some devices especially in the internet of things can be configured to periodically report their sensor measurements without downlink requests. According to the definition of K, we can set $\mu > 1$ if a network aims to serve more downlink requests, and set $\mu < 1$ otherwise.

Let $R_{{\rm u}i}$ and $R_{{\rm d}j}$ denote the average traffic demands of an arbitrary uplink user i and downlink user j during a transmission period T. As discussed in Section II-C, both $R_{{\rm u}i}$ and $R_{{\rm d}j}$ are random following some heavy-tailed distribution. Let $m_{\rm u}$ and $\sigma_{\rm u}^2$ denote the mean and variance of $R_{{\rm u}i}$, and $m_{\rm d}$ and $\sigma_{\rm d}^2$ denote the mean and variance of $R_{{\rm d}i}$. Given $R_{{\rm u}i}$ and $R_{{\rm d}j}$, the numbers of time slots required by uplink user i and downlink user j to accomplish their traffic demands can be computed as $\frac{R_{{\rm u}i}T}{\hat{R}_{\rm u}(\kappa, T_{\rm hu}, T_{\rm f})}$ and $\frac{R_{{\rm d}j}T}{\hat{R}_{\rm d}(\kappa, T_{\rm hd}, T_{\rm f})}$ under XD mode, respectively, where the numerators and denominators are respectively the required traffic volumes and the average data rates during the transmission period T. Thus, we obtain that $K_{\rm u}$ and $K_{\rm d}$ need to satisfy $\sum_{i=1}^{K_{\rm u}} \frac{R_{{\rm u}i}T}{\hat{R}_{\rm u}(\kappa, T_{\rm hu}, T_{\rm f})} \le (T_{\rm f}+T_{\rm hu})T$ and $\sum_{j=1}^{K_{\rm d}} \frac{R_{{\rm d}j}T}{\hat{R}_{\rm d}(\kappa, T_{\rm hd}, T_{\rm f})} \le (T_{\rm f}+T_{\rm hd})T$, i.e., the total required uplink or downlink time slots are not larger than the configured time slots. Considering the randomness of $R_{{\rm u}i}$ and $R_{{\rm d}j}$, these two conditions cannot be always ensured. Instead, we employ a probability model, where $K_{\rm u}$ and $K_{\rm d}$ are regarded feasible if the following conditions are satisfied

$$\mathbb{P}\Big\{\sum_{i=1}^{K_{\rm u}} \frac{R_{{\rm u}i}}{\hat{R}_{\rm u}(\kappa, T_{\rm hu}, T_{\rm f})} \le T_{\rm f}+T_{\rm hu}\Big\} \ge 1-\epsilon, \tag{6a}$$

$$\mathbb{P}\Big\{\sum_{j=1}^{K_{\rm d}} \frac{R_{{\rm d}j}}{\hat{R}_{\rm d}(\kappa, T_{\rm hd}, T_{\rm f})} \le T_{\rm f}+T_{\rm hd}\Big\} \ge 1-\epsilon. \tag{6b}$$

Herein, the value of ϵ is generally small, e.g., no larger than 5%. We can define in the same way the numbers of uplink and downlink users that can be supported under pure FD and HD modes as

$$\mathbb{P}\Big\{\sum_{i=1}^{K_{\rm u}} \frac{R_{{\rm u}i}}{\hat{R}_{\rm fu}(\kappa)} \le 1\Big\} \ge 1-\epsilon,\ \mathbb{P}\Big\{\sum_{j=1}^{K_{\rm d}} \frac{R_{{\rm d}j}}{\hat{R}_{\rm fd}(\kappa)} \le 1\Big\} \ge 1-\epsilon, \tag{7a}$$

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TCOMM.2017.2743689, IEEE Transactions on Communications

$$\mathbb{P}\Big\{\sum_{i=1}^{K_\mathrm{u}} \frac{R_{\mathrm{u}i}}{\hat{R}_\mathrm{hu}} \le T_\mathrm{hu}\Big\} \ge 1-\epsilon,\ \mathbb{P}\Big\{\sum_{j=1}^{K_\mathrm{d}} \frac{R_{\mathrm{d}j}}{\hat{R}_\mathrm{hd}} \le T_\mathrm{hd}\Big\} \ge 1-\epsilon. \tag{7b}$$

B. Resource Configuration under XD Mode

In this subsection we consider the flexible XD mode and study the joint optimization of the bidirectional transmit power ratio κ and the fractions T_f, T_hd, and T_hu, from which the optimal resource configuration for pure HD mode can be obtained directly. We will consider the pure FD mode in next subsections.

The joint power and time resource configuration problem under XD mode, aimed at maximizing the number of users supported by the network subject to both bidirectional transmit power constraint and total time slot constraint, can be formulated as

$$\max_{T_\mathrm{hu},T_\mathrm{hd},T_\mathrm{f},\kappa,K_\mathrm{u},K_\mathrm{d},K} \quad K=\min(\mu K_\mathrm{u}, K_\mathrm{d}) \tag{8a}$$
$$s.t.\ \mathbb{P}\Big\{\sum_{i=1}^{K_\mathrm{u}} R_{\mathrm{u}i} \le (T_\mathrm{f}+T_\mathrm{hu})\hat{R}_\mathrm{u}(\kappa, T_\mathrm{hu}, T_\mathrm{f})\Big\} \ge 1-\epsilon, \tag{8b}$$
$$\mathbb{P}\Big\{\sum_{j=1}^{K_\mathrm{d}} R_{\mathrm{d}j} \le (T_\mathrm{f}+T_\mathrm{hd})\hat{R}_\mathrm{d}(\kappa, T_\mathrm{hd}, T_\mathrm{f})\Big\} \ge 1-\epsilon, \tag{8c}$$
$$T_\mathrm{hu}+T_\mathrm{hd}+T_\mathrm{f}=1, \tag{8d}$$
$$\kappa_\mathrm{min} \le \kappa \le \kappa_\mathrm{max}, \tag{8e}$$

where constraint (8b) and (8c) come from (6), $\kappa_\mathrm{min}=\frac{P_\mathrm{d,min}}{P_\mathrm{u,max}}$ and $\kappa_\mathrm{max}=\frac{P_\mathrm{d,max}}{P_\mathrm{u,min}}$ denote the lower and upper bounds of power ratio κ, $P_\mathrm{d,max}$ is the maximal transmit powers of BSs, and $P_\mathrm{u,max}$ is the upper bound of P_u as defined in Section II-B.

We can find that $\mathbb{P}\{\cdot\}$ is a decreasing function for both K_u and K_d. Therefore, for any given resource configuration, the optimal K_u and K_d can be obtained when constraint (8b) and (8c) hold with equality.[1] As a result, problem (8) can be equivalently transformed as

$$\max_{T_\mathrm{hu},T_\mathrm{hd},T_\mathrm{f},\kappa,K_\mathrm{u},K_\mathrm{d},K} \quad K=\min(\mu K_\mathrm{u}, K_\mathrm{d}) \tag{9a}$$
$$s.t.\ \mathbb{P}\Big\{\sum_{i=1}^{K_\mathrm{u}} R_{\mathrm{u}i} \le (T_\mathrm{f}+T_\mathrm{hu})\hat{R}_\mathrm{u}(\kappa, T_\mathrm{hu}, T_\mathrm{f})\Big\} = 1-\epsilon, \tag{9b}$$
$$\mathbb{P}\Big\{\sum_{j=1}^{K_\mathrm{d}} R_{\mathrm{d}j} \le (T_\mathrm{f}+T_\mathrm{hd})\hat{R}_\mathrm{d}(\kappa, T_\mathrm{hd}, T_\mathrm{f})\Big\} = 1-\epsilon, \tag{9c}$$
$$T_\mathrm{hu}+T_\mathrm{hd}+T_\mathrm{f}=1, \tag{9d}$$
$$\kappa_\mathrm{min} \le \kappa \le \kappa_\mathrm{max}. \tag{9e}$$

To solve problem (9), we first find explicit expressions for constraint (9b) and (9c). Considering that the numbers of uplink and downlink users supported in current and future networks are generally large, we can employ the central limit theorem to simplify the constraints. Take constraint (9b) as an example, which can be transformed as

$$\begin{aligned}&\mathbb{P}\Big\{\sum_{i=1}^{K_\mathrm{u}} R_{\mathrm{u}i} \le (T_\mathrm{f}+T_\mathrm{hu})\hat{R}_\mathrm{u}(\kappa, T_\mathrm{hu}, T_\mathrm{f})\Big\}\\ \approx&\Phi\Big(\frac{(T_\mathrm{f}+T_\mathrm{hu})\hat{R}_\mathrm{u}(\kappa, T_\mathrm{hu}, T_\mathrm{f}) - K_\mathrm{u}m_\mathrm{u}}{\sigma_\mathrm{u}\sqrt{K_\mathrm{u}}}\Big) = 1-\epsilon,\end{aligned} \tag{10}$$

where the approximation follows from central limit theorem, and $\Phi(\cdot)$ is the CDF of the standard Gaussian distribution with zero mean and unit variance. Letting $\Phi^{-1}(\cdot)$ denote the inverse function of $\Phi(\cdot)$, we can rewrite constraint (9b) and (9c) explicitly as

$$m_\mathrm{u}K_\mathrm{u}+\Phi^{-1}(1-\epsilon)\sigma_\mathrm{u}\sqrt{K_\mathrm{u}} = (T_\mathrm{f}+T_\mathrm{hu})\hat{R}_\mathrm{u}(\kappa, T_\mathrm{hu}, T_\mathrm{f}), \tag{11a}$$
$$m_\mathrm{d}K_\mathrm{d}+\Phi^{-1}(1-\epsilon)\sigma_\mathrm{d}\sqrt{K_\mathrm{d}} = (T_\mathrm{f}+T_\mathrm{hd})\hat{R}_\mathrm{d}(\kappa, T_\mathrm{hd}, T_\mathrm{f}). \tag{11b}$$

Then, based on (5e) and (5f) and replacing T_f with $1-T_\mathrm{hu}-T_\mathrm{hd}$ according to (9d), we can rewrite problem (9) as

$$\max_{T_\mathrm{hu},T_\mathrm{hd},\kappa,K_\mathrm{u},K_\mathrm{d},K} \quad K=\min(\mu K_\mathrm{u}, K_\mathrm{d}) \tag{12a}$$
$$\begin{aligned}s.t.\ \ &m_\mathrm{u}K_\mathrm{u}+\Phi^{-1}(1-\epsilon)\sigma_\mathrm{u}\sqrt{K_\mathrm{u}} =\\ &\big(\chi\hat{R}_\mathrm{h0}-\hat{R}_\mathrm{fu}(\kappa)\big)T_\mathrm{hu}-\hat{R}_\mathrm{fu}(\kappa)T_\mathrm{hd}+\hat{R}_\mathrm{fu}(\kappa),\end{aligned} \tag{12b}$$
$$\begin{aligned}&m_\mathrm{d}K_\mathrm{d}+\Phi^{-1}(1-\epsilon)\sigma_\mathrm{d}\sqrt{K_\mathrm{d}} =\\ &\big(\hat{R}_\mathrm{h0}-\hat{R}_\mathrm{fd}(\kappa)\big)T_\mathrm{hd}-\hat{R}_\mathrm{fd}(\kappa)T_\mathrm{hu}+\hat{R}_\mathrm{fd}(\kappa),\end{aligned} \tag{12c}$$
$$T_\mathrm{hu} \ge 0, \tag{12d}$$
$$T_\mathrm{hd} \ge 0, \tag{12e}$$
$$T_\mathrm{hu}+T_\mathrm{hd} \le 1, \tag{12f}$$
$$\kappa_\mathrm{min} \le \kappa \le \kappa_\mathrm{max}. \tag{12g}$$

We can find that the left-hand sides of constraint (12b) and (12c) are respectively increasing functions of K_u and K_d by noting that $\Phi^{-1}(1-\epsilon) > 0$ holds for small ϵ. Further considering that the existence of bidirectional interference and self-interference under pure FD mode makes $\chi\hat{R}_\mathrm{h0} > \hat{R}_\mathrm{fu}(\kappa)$ and $\hat{R}_\mathrm{h0} > \hat{R}_\mathrm{fd}(\kappa)$ hold, we know from constraint (12b) and (12c) that for any given bidirectional transmit power ratio κ, K_u increases with T_hu and decreases with T_hd, while K_d decreases with T_hu and increases with T_hd. Therefore, in order to maximize the minimum of μK_u and K_d as in (12a), the optimal T_hu and T_hd should be selected to make $\mu K_\mathrm{u} = K_\mathrm{d} = K$ hold. Then, by replacing K_u and K_d in constraint (12b) and (12c) with $\frac{K}{\mu}$ and K, we obtain

$$\begin{aligned}&\frac{m_\mathrm{u}}{\mu}K+\frac{\Phi^{-1}(1-\epsilon)\sigma_\mathrm{u}}{\sqrt{\mu}}\sqrt{K} =\\ &\quad(\chi\hat{R}_\mathrm{h0}-\hat{R}_\mathrm{fu}(\kappa))T_\mathrm{hu}-\hat{R}_\mathrm{fu}(\kappa)T_\mathrm{hd}+\hat{R}_\mathrm{fu}(\kappa),\end{aligned} \tag{13a}$$
$$\begin{aligned}&m_\mathrm{d}K+\Phi^{-1}(1-\epsilon)\sigma_\mathrm{d}\sqrt{K} =\\ &\quad(\hat{R}_\mathrm{h0}-\hat{R}_\mathrm{fd}(\kappa))T_\mathrm{hd}-\hat{R}_\mathrm{fd}(\kappa)T_\mathrm{hu}+\hat{R}_\mathrm{fd}(\kappa).\end{aligned} \tag{13b}$$

Based on (13a) and (13b), we can solve the fractions T_hu and T_hd as

$$T_\mathrm{hu}=\frac{a'(\kappa)K+b'(\kappa)\sqrt{K}-d'(\kappa)}{c(\kappa)}, \tag{14a}$$
$$T_\mathrm{hd}=\frac{a(\kappa)K+b(\kappa)\sqrt{K}-d(\kappa)}{c(\kappa)}, \tag{14b}$$

where $a'(\kappa) = \frac{1}{\mu}m_\mathrm{u}(\hat{R}_\mathrm{h0}-\hat{R}_\mathrm{fd}(\kappa)) + m_\mathrm{d}\hat{R}_\mathrm{fu}(\kappa)$, $b'(\kappa) = \frac{1}{\sqrt{\mu}}\Phi^{-1}(1-\epsilon)\sigma_\mathrm{u}(\hat{R}_\mathrm{h0}-\hat{R}_\mathrm{fd}(\kappa)) + \Phi^{-1}(1-\epsilon)\sigma_\mathrm{d}\hat{R}_\mathrm{fu}(\kappa)$, $c(\kappa)=\chi\hat{R}_\mathrm{h0}^2-\hat{R}_\mathrm{h0}(\hat{R}_\mathrm{fu}(\kappa)+\chi\hat{R}_\mathrm{fd}(\kappa))$, $d'(\kappa) = \hat{R}_\mathrm{h0}\hat{R}_\mathrm{fu}(\kappa)$, $a(\kappa) = \frac{1}{\mu}\cdot\ m_\mathrm{u}\hat{R}_\mathrm{fd}(\kappa) + m_\mathrm{d}(\chi\hat{R}_\mathrm{h0}-\hat{R}_\mathrm{fu}(\kappa))$, $b(\kappa) = \frac{1}{\sqrt{\mu}}\Phi^{-1}(1-\epsilon)\sigma_\mathrm{u}\hat{R}_\mathrm{fd}(\kappa) + \Phi^{-1}(1-\epsilon)\sigma_\mathrm{d}(\chi\hat{R}_\mathrm{h0}-\hat{R}_\mathrm{fu}(\kappa))$, and $d(\kappa)=\chi\hat{R}_\mathrm{h0}\hat{R}_\mathrm{fd}(\kappa)$.

[1] Throughout the paper we consider that the numbers of users, K_u and K_d, are continuous numbers, and the integer constraints on K_u and K_d can be easily included by using floor operation to the obtained continuous values.

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TCOMM.2017.2743689, IEEE Transactions on Communications

With (14), we can replace $T_{\rm hu}$ and $T_{\rm hd}$ with K and reduce the variables of problem (12) to only κ and K. The resultant optimization problem can be expressed as

$$\max_{\kappa,K} \quad K \tag{15a}$$

$$s.t.\ \frac{a'(\kappa)K+b'(\kappa)\sqrt{K}-d'(\kappa)}{c(\kappa)} \geq 0, \tag{15b}$$

$$\frac{a(\kappa)K+b(\kappa)\sqrt{K}-d(\kappa)}{c(\kappa)} \geq 0, \tag{15c}$$

$$\frac{(a(\kappa)+a'(\kappa))K+(b(\kappa)+b'(\kappa))\sqrt{K}-(d(\kappa)+d'(\kappa))}{c(\kappa)} \leq 1, \tag{15d}$$

$$\kappa_{\min} \leq \kappa \leq \kappa_{\max}. \tag{15e}$$

Although we have simplified the original problem (8) into problem (15), it is still hard to directly solve it because constraint (15b), (15c) and (15d) are non-convex. To tackle this difficulty, in the following we strive to solve problem (15) in two complementary cases, based on which the solution of problem (15) can be found.

- **Case 1:** $\chi\hat{R}_{\rm h0} > \hat{R}_{\rm fu}(\kappa)+\chi\hat{R}_{\rm fd}(\kappa)$

Proposition 1: Under the constraint $\chi\hat{R}_{\rm h0} > \hat{R}_{\rm fu}(\kappa)+\chi\hat{R}_{\rm fd}(\kappa)$, κ has no impact on K in problem (15), and the optimal solution of K, denoted by K_1^*, is

$$K_1^* = \frac{\left(\sqrt{\Phi^{-2}(1-\epsilon)\sigma_{\rm ud}^2+4m_{\rm ud}\hat{R}_{\rm h0}}-\Phi^{-1}(1-\epsilon)\sigma_{\rm ud}\right)^2}{4m_{\rm ud}^2}, \tag{16}$$

where $\sigma_{\rm ud} \triangleq \frac{\sigma_{\rm u}}{\chi\sqrt{\mu}}+\sigma_{\rm d}$ and $m_{\rm ud} \triangleq \frac{m_{\rm u}}{\chi\mu}+m_{\rm d}$. With K_1^*, the optimal fractions $T_{\rm hu}^*$ and $T_{\rm hd}^*$ can be obtained from (A.3b) and (A.3c) in Appendix A.

Proof: See Appendix A. ∎

Remark 1: It can be observed from (A.3b) and (A.3c) that

$$\frac{T_{\rm hu}^*}{T_{\rm hd}^*} = \frac{m_{\rm u}\cdot\left(\sqrt{\frac{K_1^*}{\mu}}+\Phi^{-1}(1-\epsilon)\frac{\sigma_{\rm u}}{m_{\rm u}}\right)}{\chi\sqrt{\mu}m_{\rm d}\cdot\left(\sqrt{K_1^*}+\Phi^{-1}(1-\epsilon)\frac{\sigma_{\rm d}}{m_{\rm d}}\right)}. \tag{17}$$

If the number of supported users K_1^* is large so that the term $\Phi^{-1}(1-\epsilon)\frac{\sigma_{\rm u}}{m_{\rm u}}$ in (17) is ignorable, we have $\frac{T_{\rm hu}^*}{T_{\rm hd}^*} \approx \frac{m_{\rm u}}{\chi\mu m_{\rm d}}$. Under a special case with $\mu = 1$, we can obtain $\frac{T_{\rm hu}^*}{T_{\rm hd}^*} \approx \frac{m_{\rm u}/\hat{R}_{\rm hu}}{m_{\rm d}/\hat{R}_{\rm hd}}$, i.e., the fractions of time slots configured for uplink and downlink are proportional to the means of bidirectional traffic demands normalized by their corresponding average rates, which coincides with the intuition.

- **Case 2:** $\chi\hat{R}_{\rm h0} \leq \hat{R}_{\rm fu}(\kappa)+\chi\hat{R}_{\rm fd}(\kappa)$

In this case, we can find parameter $c(\kappa) \leq 0$ as defined below (14), and then constraint (15b), (15c), and (15d) can be rewritten as

$$f(\kappa,K) \leq \hat{R}_{\rm h0}, \tag{18a}$$

$$g(\kappa,K) \leq \hat{R}_{\rm h0}, \tag{18b}$$

$$h(K) \geq \hat{R}_{\rm h0}. \tag{18c}$$

Since $f(\kappa,K)$, $g(\kappa,K)$, and $h(K)$ increase with K, constraint (18a) and (18b) give two upper bounds for K, and constraint (18c) gives a lower bound for K. Considering that $\frac{\hat{R}_{\rm h0}-\hat{R}_{\rm fd}(\kappa)}{\hat{R}_{\rm fu}(\kappa)} \leq \frac{1}{\chi}$ and $\frac{\chi\hat{R}_{\rm h0}-\hat{R}_{\rm fu}(\kappa)}{\chi\hat{R}_{\rm fd}(\kappa)} \leq 1$ hold in this case, we have $f(\kappa,K) \leq h(K)$ and $g(\kappa,K) \leq h(K)$ for any given κ. This means that the maximal K can reach its upper bound specified by either (18a) or (18b) or both, and (18c) must hold and thus can be omitted. As a result, we can find the optimal κ and K in this case by enumerating the three possibilities regarding which bound is reached.

1) (18a) holds with equality: In this scenario problem (15) can be transformed as

$$\max_{\kappa,K} K \tag{19a}$$

$$s.t.\ f(\kappa,K) = \hat{R}_{\rm h0}, \tag{19b}$$

$$\frac{m_{\rm u}}{\mu}K+\frac{\Phi^{-1}(1-\epsilon)\sigma_{\rm u}}{\sqrt{\mu}}\sqrt{K} \leq \hat{R}_{\rm fu}(\kappa), \tag{19c}$$

$$\kappa_{\min} \leq \kappa \leq \kappa_{\max}, \tag{19d}$$

$$\chi\hat{R}_{\rm h0} \leq \hat{R}_{\rm fu}(\kappa)+\chi\hat{R}_{\rm fd}(\kappa), \tag{19e}$$

where(19c) is obtained from (18b) based on the equality in (19b), and (19e) is the condition of Case 2.

Proposition 2: The optimal solution of problem (15), which is the larger one of the two solutions of K obtained in Case 1 and Case 2, will not change whether or not (19e) is considered.

Proof: See Appendix B. ∎

From the expression of $f(\kappa,K)$ given in (A.2a) in Appendix A and considering the equality in (19b), we can observe that maximizing K is equivalent to minimizing the term $\frac{\hat{R}_{\rm h0}-\hat{R}_{\rm fd}(\kappa)}{\hat{R}_{\rm fu}(\kappa)}$. Further considering Proposition 1, we can convert problem (19) as

$$\max_{\kappa,K} \frac{\hat{R}_{\rm fu}(\kappa)}{\hat{R}_{\rm h0}-\hat{R}_{\rm fd}(\kappa)} \tag{20a}$$

$$s.t.\ f(\kappa,K) = \hat{R}_{\rm h0}, \tag{20b}$$

$$\frac{m_{\rm u}}{\mu}K+\frac{\Phi^{-1}(1-\epsilon)\sigma_{\rm u}}{\sqrt{\mu}}\sqrt{K} \leq \hat{R}_{\rm fu}(\kappa), \tag{20c}$$

$$\kappa_{\min} \leq \kappa \leq \kappa_{\max}. \tag{20d}$$

Problem (20) is not convex because both the objective function and the constraints in (20b) and (20c) are non-convex, making it not easy to find the optimal solution. To tackle this difficulty, we next develop a bisection based method, which obtains the optimal solution to problem (20) by solving a series of feasibility problems. The detailed algorithm is given in Appendix C for the sake of better readability of the paper.

2) (18b) holds with equality: In this scenario, problem (15) can be transformed as

$$\max_{\kappa,K} K \tag{21a}$$

$$s.t.\ m_{\rm d}K+\Phi^{-1}(1-\epsilon)\sigma_{\rm d}\sqrt{K} \leq \hat{R}_{\rm fd}(\kappa), \tag{21b}$$

$$g(\kappa,K) = \hat{R}_{\rm h0}, \tag{21c}$$

$$\kappa_{\min} \leq \kappa \leq \kappa_{\max}, \tag{21d}$$

$$\chi\hat{R}_{\rm h0} \leq \hat{R}_{\rm fu}(\kappa)+\chi\hat{R}_{\rm fd}(\kappa), \tag{21e}$$

where (21b) is obtained from (18a) based on the equality in (21c).

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TCOMM.2017.2743689, IEEE Transactions on Communications

With the equality in (21c), we can find from the expression of $g(\kappa, K)$ given in (A.2b) that maximizing K in problem (21) is equivalent to minimizing the term $\frac{\chi\hat{R}_{\rm h0}-\hat{R}_{\rm fu}(\kappa)}{\chi\hat{R}_{\rm fd}(\kappa)}$. Moreover, similar to Proposition 2, it is not difficult to show that ignoring constraint (21e) will not change the maximal value of K obtained in Case 1 and Case 2. Thus, we can recast problem (21) as

$$\min_{\kappa,K} \ \frac{\chi\hat{R}_{\rm h0}-\hat{R}_{\rm fu}(\kappa)}{\chi\hat{R}_{\rm fd}(\kappa)} \tag{22a}$$

$$s.t.\ m_{\rm d}K+\Phi^{-1}(1-\epsilon)\sigma_{\rm d}\sqrt{K}\le\hat{R}_{\rm fd}(\kappa), \tag{22b}$$

$$g(\kappa,K)=\hat{R}_{\rm h0}, \tag{22c}$$

$$\kappa_{\min}\le\kappa\le\kappa_{\max}. \tag{22d}$$

Note that the minimization of $\frac{\chi\hat{R}_{\rm h0}-\hat{R}_{\rm fu}(\kappa)}{\chi\hat{R}_{\rm fd}(\kappa)}$ is equivalent to $\max_{\kappa,K}\frac{1}{\frac{\chi\hat{R}_{\rm h0}-\hat{R}_{\rm fu}(\kappa)}{\chi\hat{R}_{\rm fd}(\kappa)}}$, where both $\frac{1}{\chi\hat{R}_{\rm h0}-\hat{R}_{\rm fu}(\kappa)}$ and $\frac{1}{\chi\hat{R}_{\rm fd}(\kappa)}$ are convex because $\chi\hat{R}_{\rm h0}-\hat{R}_{\rm fu}(\kappa)$ and $\chi\hat{R}_{\rm fd}(\kappa)$ are concave and positive for all κ constrained by (22d). Thus, problem (22) falls into the class of non-concave fractional program as problem (20). Further considering that the constraints of problem (20) and problem (22) are similar, the previously proposed algorithm in Appendix C can be directly used to solve problem (22). Let $\kappa_{\rm 2b}^*$ and $K_{\rm 2b}^*$ denote the optimal solutions to problem (22).

3) Both (18a) and (18b) hold with equality: Since (18a) and (18b) are transformed from (15b) and (15c), while (15b) and (15c) come from (12d) and (12e), when (18a) and (18b) hold with equality, we have $T_{\rm hu}=0$, $T_{\rm hd}=0$, and $T_{\rm f}=1$, i.e., the system operates in pure FD mode. Then, constraint (15b) and (15c) are simplified as

$$\frac{m_{\rm u}}{\mu}K+\frac{\Phi^{-1}(1-\epsilon)\sigma_{\rm u}}{\sqrt{\mu}}\sqrt{K}=\hat{R}_{\rm fu}(\kappa), \tag{23a}$$

$$m_{\rm d}K+\Phi^{-1}(1-\epsilon)\sigma_{\rm d}\sqrt{K}=\hat{R}_{\rm fd}(\kappa). \tag{23b}$$

Since $\hat{R}_{\rm fu}(\kappa)$ is a decreasing function of κ, $\hat{R}_{\rm fd}(\kappa)$ is an increasing function of κ, and the left-hand sides of (23a) and (23b) are both increasing functions of K, we know that there only exists a single solution of κ and K making (23a) and (23b) hold. Let κ_{2c}^* and K_{2c}^* denote the optimal solutions, which can be obtained by solving the equations in (23) with bisection method.

Now based on the obtained optimal solutions for Case 1 and the three scenarios in Case 2, we can obtain the maximal number of users supported under XD mode as

$$K^*=\max(K_1^*,K_{2a}^*,K_{2b}^*,K_{2c}^*). \tag{24}$$

The optimal κ^* is the solution of κ in the scenario where K^* is achieved. If $K^*=K_1^*$, then $\kappa^*=\frac{P_{\rm d,max}}{P_{\rm u,max}}$ because in this scenario the system operates in HD mode and both BSs and users should transmit with maximal powers. With κ^* and K^*, the optimal fractions $T_{\rm hu}^*$ and $T_{\rm hd}^*$ can be obtained from (14), and then the optimal pure FD fraction $T_{\rm f}^*=1-T_{\rm hu}^*-T_{\rm hd}^*$.

Remark 2: Through the analysis in this subsection, we can find that $T_{\rm f}^*>0$, $T_{\rm hu}^*>0$, and $T_{\rm hd}^*>0$ will not happen simultaneously. Specifically, if $K^*=K_1^*$, then $T_{\rm f}^*=0$; if $K^*=K_{2a}^*$, then $T_{\rm hu}^*=0$; if $K^*=K_{2b}^*$, then $T_{\rm hd}^*=0$; and if $K^*=K_{2c}^*$, then $T_{\rm hu}^*=T_{\rm hd}^*=0$. This can be explained by discussing the relationship between the data rates in pure FD and pure HD modes. When pure FD mode has a lower data rate than pure HD mode, i.e., $\chi\hat{R}_{\rm h0}>\hat{R}_{\rm fu}(\kappa)+\chi\hat{R}_{\rm fd}(\kappa)$ as considered in Case 1, the system will not operate in pure FD mode, leading to $T_{\rm f}^*=0$. When pure FD mode has a higher data rate as considered in Case 2, the network prefers pure FD mode. However, only operating in pure FD mode is not adequate if the bidirectional traffic demands are highly asymmetric, for example, if the downlink traffic demand is very high so that $K_{\rm d}<\mu K_{\rm u}$ during the pure FD mode. Then, the system will use the remaining time slots to operate in pure HD downlink mode to increase $K_{\rm d}$, leading to $T_{\rm hu}^*=0$.

Remark 3: The implementation of XD mode is summarized as follows. First, each cluster acquires the statistics of its bidirectional traffic demands in the upcoming transmission period, including the means and variances of uplink and downlink traffic demands. Then, given the traffic statistics, the proposed resource configuration algorithm in this subsection can be executed to obtained the fractions of time slots allocated to FD mode, HD uplink mode and HD downlink mode, as well as the transmit powers of BSs and users. The obtained resource configuration is used by all cells within the cluster, and different clusters can configure the resources in parallel and may obtain different configuration results according to their bidirectional traffic demands.

C. Resource Configuration under Pure FD Mode

In XD mode, we have shown below (12) that $\mu K_{\rm u}=K_{\rm d}$ must hold by optimizing the fractions $T_{\rm hu}$, $T_{\rm hd}$ and $T_{\rm f}$. Under pure FD mode the fractions are fixed as $T_{\rm hu}=0$, $T_{\rm hd}=0$ and $T_{\rm f}=1$ and the only configurable parameter is the power ratio κ, which simplifies the optimization problem (12) as

$$\max_{\kappa,K_{\rm u},K_{\rm d},K} \ K=\min(\mu K_{\rm u},K_{\rm d}) \tag{25a}$$

$$s.t.\ \ m_{\rm u}K_{\rm u}+\Phi^{-1}(1-\epsilon)\sigma_{\rm u}\sqrt{K_{\rm u}}=\hat{R}_{\rm fu}(\kappa), \tag{25b}$$

$$m_{\rm d}K_{\rm d}+\Phi^{-1}(1-\epsilon)\sigma_{\rm d}\sqrt{K_{\rm d}}=\hat{R}_{\rm fd}(\kappa), \tag{25c}$$

$$\kappa_{\min}\le\kappa\le\kappa_{\max}. \tag{25d}$$

We can find from (25b) that $K_{\rm u}$ decreases with κ and from (25c) that $K_{\rm d}$ increases with κ. If $\kappa\in[0,\infty]$, then there must exist a κ to make $\mu K_{\rm u}=K_{\rm d}$ hold. Yet, this may be no longer true when constraint (25d) is considered, and thus the previously proposed algorithm cannot be directly applied.

To solve problem (25), we first relax it by omitting constraint (25d). Let $\tilde{\kappa}$, $\tilde{K}_{\rm u}$, $\tilde{K}_{\rm d}$ and $\tilde{K}$ denote the solutions to the relaxed problem, which can be easily solved by bisection method. We know that $\mu\tilde{K}_{\rm u}=\tilde{K}_{\rm d}=\tilde{K}$ must hold with the relaxed $\tilde{\kappa}$.

Now we discuss the impact of constraint (25d). Recalling that $K_{\rm d}$ increases with κ and $K_{\rm u}$ decreases with κ, if $\tilde{\kappa}>\kappa_{\max}$, we know $\mu K_{\rm u}>K_{\rm d}$ for all $\kappa\in[\kappa_{\min},\kappa_{\max}]$, so that the optimal $K^*=K_{\rm d}$ according to (25a), which is obtained when $\kappa=\kappa_{\max}$. On the other hand, if $\tilde{\kappa}<\kappa_{\min}$, we have $\mu K_{\rm u}<K_{\rm d}$ for all $\kappa\in[\kappa_{\min},\kappa_{\max}]$, so that

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TCOMM.2017.2743689, IEEE Transactions on Communications

$K^* = \mu K_{\rm u}$, which is obtained when $\kappa = \kappa_{\min}$. Otherwise, if $\tilde{\kappa}^* \in [\kappa_{\min}, \kappa_{\max}]$, we have $K^* = \tilde{K}$ with the optimal $\kappa = \tilde{\kappa}$.

IV. Simulation and Numerical Results

In this section, we evaluate the performance of FD and HD networks based on the optimized resource configurations, and investigate the performance gain of FD network over traditional HD network. Unless otherwise specified, the following setups are considered throughout the simulations. The system bandwidth is 10 MHz, the noise power spectrum density is –174 dBm/Hz, and the noise figure is 9 dB, from which we can obtain the total noise power is –95 dBm [35]. The density of BSs is set as $\lambda = 6.25 \times 10^{-4}/\text{m}^2$, which corresponds to an average inter-site distance of 40 m. The pathloss for the channels between BSs, between users, and between BSs and users are based on the Hata model [36], where the heights of BSs and users are set as 10 m and 1.5 m, respectively. The maximal transmit powers of BSs and users are $P_{\rm d,max} = 30$ dBm and $P_{\rm u,max\,0} = 23$ dBm, respectively [35], considering that FD is more feasible for SCN [25, 26]. The uplink power control factor is selected as $\eta = 0$ and $\eta = 1$. In order to ensure that the network operates in interference-limited scenario, the minimal transmit power of BSs, $P_{\rm d,min}$, is selected to satisfy the condition $\frac{\hat{R}_{\rm h0} - \bar{R}_{\rm hd}(P_{\rm d,min})}{\bar{R}_{\rm hd}(P_{\rm d,min})} = 10^{-4}$, i.e., to ensure the impact of ignoring noises negligible, where $\hat{R}_{\rm h0}$ and $\bar{R}_{\rm hd}$ are the average downlink rates in the cases without and with noise as defined in Section II-B, respectively. The minimal base power of uplink power control, $P_{\rm u,min}$, can be obtained in the same way. Taking $\eta = 0$ as an example, we can obtain that $P_{\rm d,min} = -5.7$ dBm and $P_{\rm u,min} = -5.8$ dBm, resulting in the range of κ with $\kappa_{\min} = 0.0014$ and $\kappa_{\max} = 3816.8$. We consider that the uplink and downlink traffic demands of users follow log-normal distribution with means $m_{\rm u}$ and $m_{\rm d}$ for uplink and downlink, respectively, as discussed in Section II-C, where m_u is set as 2 kbps according the prediction of ITU for uplink traffic demand in a peak hour in 2020 [4], and different values of $m_{\rm d}$ will be considered.[2] Recalling that the ratios of standard derivation to mean for bidirectional traffic demands are in the range of 0.7∼0.8 as shown in Section II-C, we choose $\sigma_{\rm u} = 0.75 m_{\rm u}$ and $\sigma_{\rm d} = 0.75 m_{\rm d}$. Suppose that the network prefers to serve more downlink users, thus the weight μ is set to be larger than 1, say 5. The probability thresholds used for uplink power control and defining the number of supported users are set as $\theta = 95\%$ and $\epsilon = 5\%$, respectively. The SIS in FD mode is set as 110 dB, i.e., $\beta = 10^{-11}$. In simulations the following four transmission modes are compared.

- **XD mode:** the network operates in FD-HD hybrid mode with optimized κ, $T_{\rm hu}$, $T_{\rm hd}$, and $T_{\rm f}$.

[2]Note that we have not assumed any traffic distribution in the previous analysis and optimization, and thus choosing other traffic distributions instead of the log-normal distribution will not change the obtained conclusions. Herein, $m_{\rm u} = 2$ kbps reflects the mean traffic demands in a period T, but not the uplink transmission rate experienced by a user. For instance, in the typical scenario with $m_{\rm d} : m_{\rm u} = 7 : 1$ [4], we can compute $T_{\rm hu} = 0.03$ during which $K_{\rm u} = 269$ users are served in pure HD dynamic TDD mode, i.e., each user only occupies $T_{\rm hu} T / K_{\rm u} = 1.12 \times 10^{-4} T$ time slots, leading to the user-experienced transmission rate as $m_{\rm u} T / (1.13 \times 10^{-4} T) = 17.86$ Mbps.

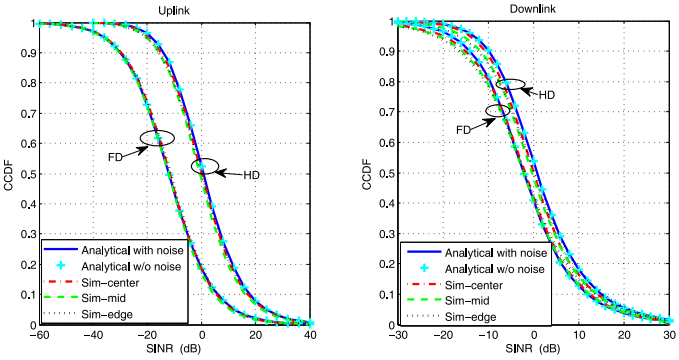

Fig. 2. CCDF of bidirectional SINRs in FD and HD modes, SIS = 110 dB, and $\eta = 0$.

- **Pure FD mode:** the network operates with optimized κ, given $T_{\rm hu} = 0$, $T_{\rm hd} = 0$ and $T_{\rm f} = 1$.
- **Pure HD dynamic TDD mode:** the network operates with optimized $T_{\rm hu}$ and $T_{\rm hd}$ given $T_{\rm f} = 0$, i.e., Case 1, where $P_{\rm d} = P_{\rm d,max}$ and $P_{\rm u} = P_{\rm u,max}$.
- **Pure HD static TDD mode:** $T_{\rm hu}$ and $T_{\rm hd}$ are fixed as $\frac{1}{2}$, $T_{\rm f} = 0$, $P_{\rm d} = P_{\rm d,max}$, and $P_{\rm u} = P_{\rm u,max}$.

A. Approximation Validation

We first evaluate the accuracy of stochastic geometry model for analyzing the performance of small-size clusters. We simulate 25 square clusters each with the side length of 200 m, where one reference cluster is surrounded by two rings of 24 interfering clusters. Recalling that the average inter-site distance is 40 m, we know that each cluster includes 25 BSs in average. In each cluster the BSs are dropped via a PPP, and a large number of users are dropped uniformly such that each BS can always select two users in its cell to serve [22]. We let every surrounding cluster randomly select its operation mode from FD, HD uplink or HD downlink with equal probability. The results are shown in Fig. 2, where legend "Analytical with noise" and "Analytical w/o noise" denote the analytical results with and without noises, respectively, based on stochastic geometry model which assume all clusters operate in the same mode as the reference cluster. To evaluate the edge effect of small-size clusters, we equally divide the reference cluster into center, middle, and edge regions and simulate their performance separately, whose results are plotted with legend "Sim-center", "Sim-mid", and "Sim-edge", respectively. It is shown that the simulation results in the three regions can well match the analytical results for both HD and FD in either uplink or downlink, indicating the appropriateness of using stochastic geometry model to analyze the performance of small-size clusters. Moreover, we can observe that ignoring noises leads to negligible impact on the analytical results.

Another approximation we used is the one based on central limit theorem to obtain the closed-form expression of the probability in (10). Figure 3 evaluates the accuracy of the approximation, where Fig. 3(a) shows the probability density function (PDF) of the term $\frac{\sum_{i=1}^{K_{\rm u}} R_{{\rm u}i}}{\sqrt{K_{\rm u}}}$. Two values of $K_{\rm u}$ corresponding to $\frac{m_{\rm d}}{m_{\rm u}} = 1$ and 7 are considered, which can be found from Fig. 3(b). We can observe that the PDFs of

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TCOMM.2017.2743689, IEEE Transactions on Communications

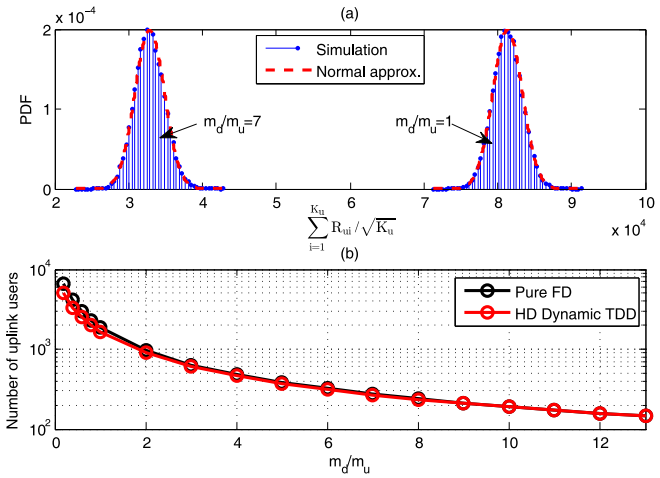

Fig. 3. Accuracy of the approximation based on central limit theorem, and the number of supported uplink users versus $\frac{m_{\rm d}}{m_{\rm u}}$.

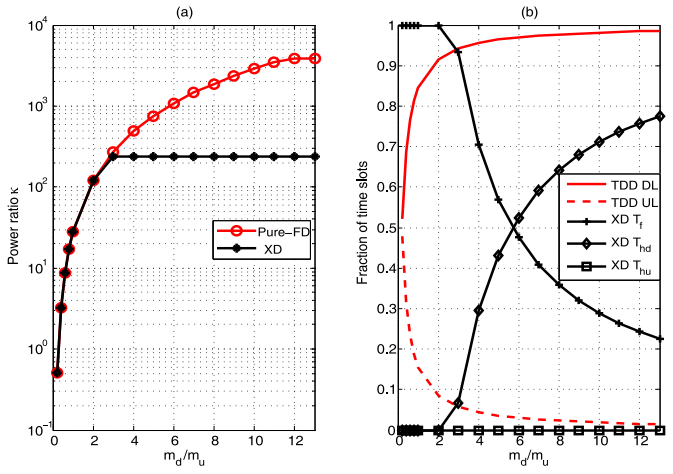

Fig. 4. Bidirectional resource configuration in pure HD dynamic TDD, pure FD, and XD modes v.s. $\frac{m_{\rm d}}{m_{\rm u}}$, $\eta = 0$, and SIS = 110 dB.

$\frac{\sum_{i=1}^{K_{\rm u}} R_{{\rm u}i}}{\sqrt{K_{\rm u}}}$ can be well fitted by Normal distributions, making the application of central limit theorem in (10) appropriate.

B. Resource Configuration Results

In this subsection we illustrate the optimized resource configurations under different $\frac{m_{\rm d}}{m_{\rm u}}$. Fig 4(a) shows the optimal downlink-to-uplink power ratio κ for XD mode and pure FD mode considering that κ only affects the performance of these two modes, and Fig 4(b) shows the optimized fractions $T_{\rm hu}$, $T_{\rm hd}$ and $T_{\rm f}$ for XD mode and pure HD dynamic TDD mode since in other modes the fractions are fixed.

We can obverse from Fig 4(a) that in pure FD mode κ increases with $\frac{m_{\rm d}}{m_{\rm u}}$ as expected until achieving its maximal value, since a larger $m_{\rm d}$ requires a higher downlink transmit power. When $\frac{m_{\rm d}}{m_{\rm u}} \leq 2$, XD mode has the same κ as pure FD mode, implying that XD mode reduces to pure FD mode in this scenario, i.e., the system operates in the scenario 3) of Case 2. This is verified by Fig 4(b), where we can find $T_{\rm hu} = 0$, $T_{\rm hd} = 0$ and $T_{\rm f} = 1$ when $\frac{m_{\rm d}}{m_{\rm u}} \leq 2$. When $\frac{m_{\rm d}}{m_{\rm u}} > 2$, in order to support the increased downlink traffic demand, we can find that the optimal strategy of XD mode is not to continuously increase downlink transmit power as shown in Fig 4(a), but to shorten the duration of FD transmission $T_{\rm f}$ and uses the saved time slots for pure HD downlink transmission, i.e., increase $T_{\rm hd}$ as shown in Fig 4(b). In this way, the numbers of supported uplink and downlink users are balanced so that $K = \min(\mu K_{\rm u}, K_{\rm d})$ is maximized. We can also observe from Fig 4(b) that $T_{\rm hu} = 0$ for all $\frac{m_{\rm d}}{m_{\rm u}}$, which agrees with Remark 1 in the sense that at least one of $T_{\rm hu}$, $T_{\rm hd}$ and $T_{\rm f}$ should be zero. For pure HD dynamic TDD mode, we can find that $T_{\rm hd}$ increases and $T_{\rm hu}$ decreases with $\frac{m_{\rm d}}{m_{\rm u}}$, respectively, as expected.

C. Gain of FD Network over HD Network

Fig. 5 depicts the performance gain of pure FD mode and XD mode over pure HD static TDD and dynamic TDD modes under different SIS with $\frac{m_{\rm d}}{m_{\rm u}} = 1$. Herein, the performance gain, for instance achieved by XD mode over static TDD mode, is defined as $\frac{K_{\rm XD} - K_{\rm Sta}}{K_{\rm Sta}}$, where $K_{\rm XD}$ and $K_{\rm Sta}$ denote the numbers of users supported by the two modes. We can observe that pure FD mode does not always outperform pure HD modes due to the existence of self-interference, e.g., when SIS < 15 dB for static TDD mode and SIS < 35 dB for dynamic TDD mode, respectively. By contrast, thanks to the flexible transmission strategy, XD mode can adaptively degenerate to pure HD dynamic TDD mode for small SIS, and degenerate to pure FD mode for large SIS, ensuring that XD mode is not inferior to both pure HD modes, as shown in Fig. 5.

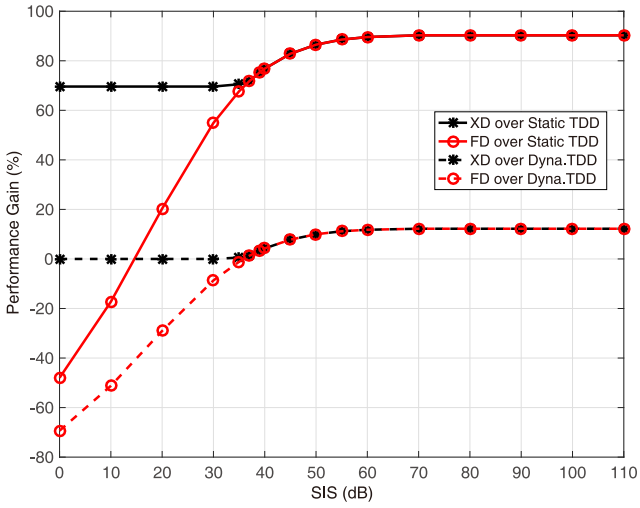

Fig. 5. Gain of pure FD and XD modes over pure HD static TDD and dynamic TDD modes v.s. SIS, $\frac{m_{\rm d}}{m_{\rm u}} = 1, \eta = 0$.

Fig. 6 depicts the performance gain of pure FD mode and XD mode over pure HD static TDD and dynamic TDD modes under different bidirectional traffic demands with SIS $=$ 110 dB. Fig. 6(a) shows the results with $\eta = 0$, i.e., when uplink power control is not considered. First, compared to static TDD mode, where identical time slots are allocated to uplink and downlink, we can observe the increasing gain of pure FD and XD modes with $\frac{m_{\rm d}}{m_{\rm u}}$. This can be explained as

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TCOMM.2017.2743689, IEEE Transactions on Communications

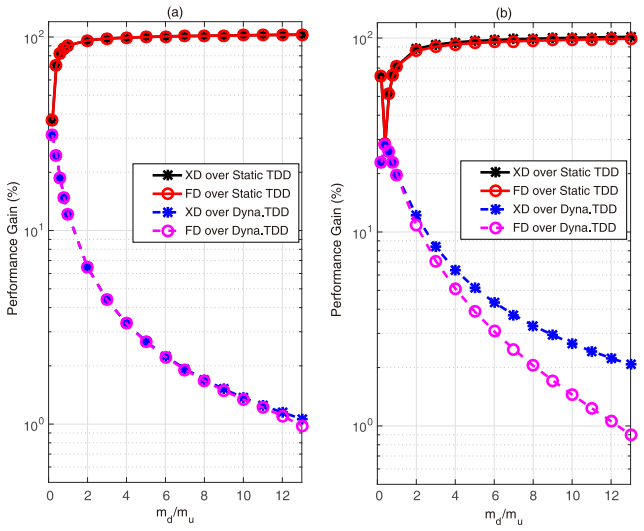

Fig. 6. Gain of pure FD and XD modes over pure HD static TDD and dynamic TDD modes v.s. $\frac{m_{\rm d}}{m_{\rm u}}$, SIS = 110 dB: (a) $\eta = 0$, (b) $\eta = 1$.

follows. According to the analysis in Remark 1, configuring identical time slots to uplink and downlink is optimal when $\frac{m_{\rm u}}{\chi\mu m_{\rm d}} \approx 1$ holds, which is equivalent to $\frac{m_{\rm d}}{m_{\rm u}} \approx 0.2$ because $\mu = 5$ and we can obtain $\chi \approx 1$ when $\eta = 0$ from (5). This leads to the minimal gain at this point (i.e., the first x-axis point). With the increase of $\frac{m_{\rm d}}{m_{\rm u}}$, the identical time slot configuration becomes farther away from the optimality, and thus results in larger performance gain. Pure FD mode performs close to XD mode under the considered large SIS. The maximal gain can be even higher than 100%, e.g., when $\frac{m_{\rm d}}{m_{\rm u}} > 4$. This is because in the simulations static TDD mode is configured with identical time slots for uplink and downlink, which largely wastes the time slotes allocated to uplink in the scenario where downlink traffic dominates.

Second, compared to dynamic TDD mode, the maximal gain of 30% is attained when $\frac{m_{\rm d}}{m_{\rm u}} = \frac{1}{\chi\mu} = 0.2$, and the gain is brought by FD transmission because $T_{\rm f} = 1$ at this point as shown in Fig. 4(b). FD transmission, allocating all time slots to both uplink and downlink, becomes inefficient with the increase of $\frac{m_{\rm d}}{m_{\rm u}}$ as expected, leading to the decrease of gain for both pure FD mode and XD mode. For typical values of $\frac{m_{\rm d}}{m_{\rm u}}$ ranging from four to seven [4], both XD mode and pure FD mode can only provide a gain less than 3.3% over dynamic TDD mode.

Similar conclusions can be drawn for the case with uplink power control, as shown in Fig. 6(b). With $\eta = 1$, we can compute from (5) that $\chi \approx 0.5$. Then, again according to Remark 1, we know that static TDD with identical uplink and downlink time slots is optimal when $\frac{m_{\rm d}}{m_{\rm u}} \approx 0.4$. This explains the minimal gain of XD mode and pure FD mode over static TDD mode at this point. For typical values of $\frac{m_{\rm d}}{m_{\rm u}}$ ranging from four to seven, the gains of both XD mode and pure FD mode over dynamic TDD mode are less than 6.5%. Indeed, such a small performance gain can hardly motivate the application of FD technology for cellular networks, and dynamic TDD mode suffices to support the asymmetric bidirectional traffic demands.

V. Conclusions

In this paper we investigated the performance gain of FD network over HD network. To this end, we analyzed the number of users supported by FD and HD networks, which is defined as the minimum of the weighted numbers of supported uplink users and downlink users given random traffic demands of users. We studied pure FD mode and a HD-FD flexible mode, namely XD mode, for FD network, and studied static TDD mode and dynamic TDD mode for HD network. We maximized the number of supported users by optimizing bidirectional transmit powers for pure FD mode, time slot configuration for pure dynamic TDD mode, and both for XD mode. Numerical results show that the performance of pure FD mode is worse than pure HD mode when self-interference is strong. If self-interference is well suppressed, both pure FD mode and XD mode exhibit evident performance gain over pure HD static TDD mode that is configured with identical time slots for uplink and downlink. However, the gain of pure FD mode and XD mode over pure HD dynamic TDD mode is marginal in typical downlink traffic dominant scenario, which cannot justify the application of FD technique to cellular networks by simply replacing current HD BSs with FD BSs. Judicious BS-BS and user-user interference control mechanisms are indispensable for FD network.

Appendix A
Proof of Proposition 1

With $\chi\hat{R}_{\rm h0} > \hat{R}_{\rm fu}(\kappa) + \chi\hat{R}_{\rm fd}(\kappa)$ and after some regular manipulations, we can rewrite constraint (15b), (15c), and (15d) as

$$f(\kappa, K) \geq \hat{R}_{\rm h0}, \tag{A.1a}$$

$$g(\kappa, K) \geq \hat{R}_{\rm h0}, \tag{A.1b}$$

$$h(K) \leq \hat{R}_{\rm h0}, \tag{A.1c}$$

where the functions $f(\kappa, K)$, $g(\kappa, K)$, and $h(K)$ are defined as follows

$$f(\kappa, K) = \frac{m_{\rm u}\frac{\hat{R}_{\rm h0}-\hat{R}_{\rm fd}(\kappa)}{\hat{R}_{\rm fu}(\kappa)} + \mu m_{\rm d}}{\mu}K + \frac{\Phi^{-1}(1-\epsilon)\sigma_{\rm u}\frac{\hat{R}_{\rm h0}-\hat{R}_{\rm fd}(\kappa)}{\hat{R}_{\rm fu}(\kappa)} + \sqrt{\mu}\Phi^{-1}(1-\epsilon)\sigma_{\rm d}}{\sqrt{\mu}}\sqrt{K}, \tag{A.2a}$$

$$g(\kappa, K) = \frac{\frac{m_{\rm u}}{\chi} + \mu m_{\rm d}\frac{\chi\hat{R}_{\rm h0}-\hat{R}_{\rm fu}(\kappa)}{\chi\hat{R}_{\rm fd}(\kappa)}}{\mu}K + \frac{\frac{\Phi^{-1}(1-\epsilon)\sigma_{\rm u}}{\chi} + \sqrt{\mu}\Phi^{-1}(1-\epsilon)\sigma_{\rm d}\frac{\chi\hat{R}_{\rm h0}-\hat{R}_{\rm fu}(\kappa)}{\chi\hat{R}_{\rm fd}(\kappa)}}{\sqrt{\mu}}\sqrt{K}, \tag{A.2b}$$

$$h(K) = \frac{\frac{m_{\rm u}}{\chi} + \mu m_{\rm d}}{\mu}K + \frac{\frac{\Phi^{-1}(1-\epsilon)\sigma_{\rm u}}{\chi} + \sqrt{\mu}\Phi^{-1}(1-\epsilon)\sigma_{\rm d}}{\sqrt{\mu}}\sqrt{K}. \tag{A.2c}$$

Considering that $\Phi^{-1}(1-\epsilon) > 0$, $\chi\hat{R}_{\rm h0} > \hat{R}_{\rm fu}(\kappa)$, and $\hat{R}_{\rm h0} > \hat{R}_{\rm fd}(\kappa)$ as discussed before, we know that $f(\kappa, K)$, $g(\kappa, K)$, and $h(K)$ are all increasing functions of K for $K \geq 0$. Thus, constraint (A.1a) and (A.1b) provide lower

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TCOMM.2017.2743689, IEEE Transactions on Communications

bounds for K, and constraint (A.1c) provides an upper bound for K. Since $\chi\hat{R}_{\rm h0} > \hat{R}_{\rm fu}(\kappa) + \chi\hat{R}_{\rm fd}(\kappa)$ in this case, we have $\frac{\hat{R}_{\rm h0}-\hat{R}_{\rm fd}(\kappa)}{\hat{R}_{\rm fu}(\kappa)} > \frac{1}{\chi}$ and $\frac{\chi\hat{R}_{\rm h0}-\hat{R}_{\rm fu}(\kappa)}{\chi\hat{R}_{\rm fd}(\kappa)} > 1$, and hence $f(\kappa,K) > h(K)$ and $g(\kappa,K) > h(K)$ as observed from (A.2). This means that the upper bound of K given by constraint (A.1c) is achievable. In order to maximize K subject to the constraints in (A.1a), (A.1b), and (A.1c), it is clear that constraint (A.1c) needs to hold with equality. Since (A.1c) is an equivalent transformation of (12f), we know that $T_{\rm hu}+T_{\rm hd}=1$ should hold, i.e., the network should operate in pure HD mode in this case.

With $T_{\rm hu}+T_{\rm hd}=1$, problem (12) can be simplified as

$$\max_{T_{\rm hu},T_{\rm hd},K} \quad K \tag{A.3a}$$

$$s.t. \quad \frac{m_{\rm u}}{\mu}K + \frac{\Phi^{-1}(1-\epsilon)\sigma_{\rm u}}{\sqrt{\mu}}\sqrt{K} = T_{\rm hu}\chi\hat{R}_{\rm h0}, \tag{A.3b}$$

$$m_{\rm d}K + \Phi^{-1}(1-\epsilon)\sigma_{\rm d}\sqrt{K} = T_{\rm hd}\hat{R}_{\rm h0}, \tag{A.3c}$$

$$T_{\rm hu} \geq 0, \tag{A.3d}$$

$$T_{\rm hd} \geq 0, \tag{A.3e}$$

$$T_{\rm hu}+T_{\rm hd}=1, \tag{A.3f}$$

which is independent from κ. Then, it is not difficult to find the optimal solution of K as (16).

Appendix B
Proof of Proposition 2

Let $\hat{\kappa}^1$ and $\hat{K}^1$ denote the optimal solutions to problem (19), and κ^1 and K^1 denote the optimal solutions to the relaxed problem (19) without constraint (19e). We know that $\hat{\kappa}^1$ must satisfy $\chi\hat{R}_{\rm h0} \leq \hat{R}_{\rm fu}(\hat{\kappa}^1) + \chi\hat{R}_{\rm fd}(\hat{\kappa}^1)$ as constrained by (19e). We next discuss the value of κ^1.

- If $\kappa^1 = \hat{\kappa}^1$ holds, then it implies that constraint (19e) does not affect the solution of problem (19) and hence can be ignored.
- If $\kappa^1 \neq \hat{\kappa}^1$, it means that κ^1 does not satisfy the ignored constraint (19e), i.e., $\chi\hat{R}_{\rm h0} > \hat{R}_{\rm fu}(\kappa^1) + \chi\hat{R}_{\rm fd}(\kappa^1)$ holds. Therefore, κ^1 is in fact a feasible solution in Case 1 where $\chi\hat{R}_{\rm h0} > \hat{R}_{\rm fu}(\kappa) + \chi\hat{R}_{\rm fd}(\kappa)$. Given κ^1, let $\tilde{K}^1$ denote the optimal solution of K in Case 1. Then, we can obtain that $f(\kappa^1,K^1) = \hat{R}_{\rm h0} \leq f(\kappa^1,\tilde{K}^1)$, where the equality is from (19b) in Case 2 and the inequality is from (A.1a) in Case 1. Since $f(\kappa,K)$ is an increasing function of K as shown in (A.2a), we obtain $K^1 \leq \tilde{K}^1$. Note that $\hat{K}^1 \leq K^1$ holds because K^1 is the solution of the relaxed problem (19), while $\tilde{K}^1$ is just a feasible solution in Case 1 given κ^1. Thus, we know that the optimal K in Case 1 is larger than $\hat{K}^1$. As a result, despite that ignoring constraint (19e) changes the solution to problem (19) in Case 2, it will not change the solution of the original problem (15), which is the larger one of the optimal solutions of K in Case 1 and Case 2.

Combining the above two scenarios, Proposition 2 is proved.

Appendix C
Algorithm to Solve Problem (20)

We propose a bisection based method to obtain the optimal solution to problem (20) by solving a series of feasibility problems. Specifically, for a given value of α, defined as $\alpha \triangleq \frac{\hat{R}_{\rm fu}(\kappa)}{\hat{R}_{\rm h0}-\hat{R}_{\rm fd}(\kappa)}$, we investigate the following feasibility problem with the objective of finding a feasible solution to satisfy all the constraint

$$\text{find } \kappa \tag{C.1a}$$

$$s.t. \frac{\frac{m_{\rm u}}{\alpha}+\mu m_{\rm d}}{\mu}K + \frac{\frac{\Phi^{-1}(1-\epsilon)\sigma_{\rm u}}{\alpha}+\sqrt{\mu}\Phi^{-1}(1-\epsilon)\sigma_{\rm d}}{\sqrt{\mu}}\sqrt{K} = \hat{R}_{\rm h0}, \tag{C.1b}$$

$$\frac{m_{\rm u}}{\mu}K + \frac{\Phi^{-1}(1-\epsilon)\sigma_{\rm u}}{\sqrt{\mu}}\sqrt{K} \leq \hat{R}_{\rm fu}(\kappa), \tag{C.1c}$$

$$\kappa_{\min} \leq \kappa \leq \kappa_{\max}, \tag{C.1d}$$

$$\frac{\hat{R}_{\rm fu}(\kappa)}{\hat{R}_{\rm h0}-\hat{R}_{\rm fd}(\kappa)} = \alpha, \tag{C.1e}$$

where (C.1b) is from (20b) by replacing the term $\frac{\hat{R}_{\rm fu}(\kappa)}{\hat{R}_{\rm h0}-\hat{R}_{\rm fd}(\kappa)}$ in the expression of $f(\kappa,K)$ given in (A.2a) with α, and (C.1e) is the definition of α .

From (C.1b) we can solve K for any given α, whose solution is denoted by K_α. Then, (C.1c) becomes $\frac{m_{\rm u}}{\mu}K_\alpha + \frac{\Phi^{-1}(1-\epsilon)\sigma_{\rm u}}{\sqrt{\mu}}\sqrt{K_\alpha} \leq \hat{R}_{\rm fu}(\kappa)$. Let $\kappa_0(K_\alpha)$ denote the value of κ making this inequality hold with equality, which can be easily computed because $\hat{R}_{\rm fu}(\kappa)$ is a decreasing function of κ as shown in (5b). Then, we know that constraint (C.1c) is equivalent to $\kappa \leq \kappa_0(K_\alpha)$, and problem (C.1) can be recast as

$$\text{find } \kappa \tag{C.2a}$$

$$s.t.\ \kappa_{\min} \leq \kappa \leq \min(\kappa_0(K_\alpha),\kappa_{\max}), \tag{C.2b}$$

$$\frac{\hat{R}_{\rm fu}(\kappa)}{\hat{R}_{\rm h0}-\hat{R}_{\rm fd}(\kappa)} = \alpha. \tag{C.2c}$$

Problem (C.2) is still difficult to solve because constraint (C.2c) is non-convex. To circumvent it, we strive to find the minimal and maximal values of $\frac{\hat{R}_{\rm fu}(\kappa)}{\hat{R}_{\rm h0}-\hat{R}_{\rm fd}(\kappa)}$ subject to constraint (C.2b). If α is between the minimal and maximal values, then there must exist a κ satisfying the constraints of problem (C.2), i.e., the given α is feasible. Otherwise, the given α is infeasible.

The maximization or minimization of $\frac{\hat{R}_{\rm fu}(\kappa)}{\hat{R}_{\rm h0}-\hat{R}_{\rm fd}(\kappa)}$ can be obtained as

$$\max_{\kappa} \ \frac{\hat{R}_{\rm fu}(\kappa)}{\hat{R}_{\rm h0}-\hat{R}_{\rm fd}(\kappa)} \quad \text{or} \quad \frac{\hat{R}_{\rm h0}-\hat{R}_{\rm fd}(\kappa)}{\hat{R}_{\rm fu}(\kappa)} \tag{C.3a}$$

$$s.t.\ \kappa_{\min} \leq \kappa \leq \min(\kappa_0(K_\alpha),\kappa_{\max}). \tag{C.3b}$$

We show in Appendix D that $\hat{R}_{\rm fu}(\kappa)$ is a convex function and $\hat{R}_{\rm fd}(\kappa)$ is a concave function. Therefore, solving problem (C.3) is equivalent to maximizing the ratio of two convex functions over a convex set, which is also known as non-concave fractional program [37]. This problem is neither convex nor quasi-convex, and thus its global optimal solution cannot be found with standard algorithms. In [37], a branch and bound based algorithm was proposed to find the global optimal solution to the non-concave fractional programming problem. We employ this method to solve problem (C.3),

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TCOMM.2017.2743689, IEEE Transactions on Communications

whose basic principle is summarized in Appendix E for the readers' convenience.

APPENDIX D
CONCAVITY OF $\hat{R}_{\mathrm{fd}}(\kappa)$

We prove the concavity of $\hat{R}_{\mathrm{fd}}(\kappa)$ in this appendix, and the extension to prove the convexity of $\hat{R}_{\mathrm{fu}}(\kappa)$ is straightforward.

Instead of directly examining the expression given in (5a), we consider an alternative expression of $\hat{R}_{\mathrm{fd}}(\kappa)$ as

$$\hat{R}_{\mathrm{fd}}(\kappa) = \int_{s>0}\int_{i_{\mathrm{BS}}>0}\int_{i_{\mathrm{UE}}>0} \tilde{R}_{\mathrm{fd}}(\kappa, s, i_{\mathrm{BS}}, i_{\mathrm{UE}}) \cdot q_{\mathrm{S}}(s) q_{\mathrm{I_{BS}}}(i_{\mathrm{BS}}) q_{\mathrm{I_{UE}}}(i_{\mathrm{UE}}) \mathrm{d}s \mathrm{d}i_{\mathrm{BS}} \mathrm{d}i_{\mathrm{UE}}, \quad \text{(D.1)}$$

where $\tilde{R}_{\mathrm{fd}}(\kappa, s, i_{\mathrm{BS}}, i_{\mathrm{UE}}) \triangleq W_0 \log\left(1 + \frac{P_{\mathrm{d}} s}{P_{\mathrm{d}} i_{\mathrm{BS}} + P_{\mathrm{u}} i_{\mathrm{UE}}}\right) = W_0 \log\left(1 + \frac{\kappa s}{\kappa i_{\mathrm{BS}} + i_{\mathrm{UE}}}\right)$ is the instantaneous rate of a downlink user with s, i_{BS} and i_{UE} denoting the instantaneous power of desired signal, interference from other BSs and interference from uplink users, and $q_{\mathrm{S}}(s)$, $q_{\mathrm{I_{BS}}}(i_{\mathrm{BS}})$ and $q_{\mathrm{I_{UE}}}(i_{\mathrm{UE}})$ denote the PDFs of s, i_{BS} and i_{UE}, respectively. Herein, the noise is omitted in the considered interference-limited scenario as discussed below (4).

Since the nonnegative weighted sum of concave functions is concave, in order to show the concavity of $\hat{R}_{\mathrm{fd}}(\kappa)$, it suffices to prove that $\tilde{R}_{\mathrm{fd}}(\kappa, s, i_{\mathrm{BS}}, i_{\mathrm{UE}})$ is concave for κ. It is easy to find that the second-order derivations of $\tilde{R}_{\mathrm{fd}}(\kappa, s, i_{\mathrm{BS}}, i_{\mathrm{UE}})$ with respect to κ is negative, which completes the proof.

APPENDIX E
ALGORITHM REVIEW FOR NON-CONCAVE FRACTIONAL PROGRAM

The branch and bound algorithm proposed in [37] is briefly reviewed in this appendix for the non-concave fractional programming problem considered in this paper

$$\max_{\kappa} \frac{l(\kappa)}{p(\kappa)} \quad s.t.\ \kappa \in \mathcal{X}_0, \quad \text{(E.1)}$$

where both $l(\kappa)$ and $p(\kappa)$ are convex functions on $\mathbb{R}$, $l(\kappa) \geq 0$ and $p(\kappa) \geq 0$ for all $\kappa \in \mathcal{X}_0$, and $\mathcal{X}_0$ is a 1-dimensional simplex or a line segment.

The branch and bound algorithm consists of three main processes including branching, lower and upper bounding, and fathoming, where branching, lower bounding and fathoming processes are standard. We next focus on the upper bounding process, where an upper bound of the objective function over a simplex $\mathcal{X}_k \subset \mathcal{X}_0$ is required in the k-th iteration.

To obtain an upper bound of $\frac{l(\kappa)}{p(\kappa)}$, [37] suggests to find a concave upper bound for $l(\kappa)$, $\hat{l}(\kappa)$, and a convex lower bound for $p(\kappa)$, $\hat{p}(\kappa)$, over $\mathcal{X}_k$. Denoting $\mathcal{X}_k = [\kappa_1, \kappa_2]$, $\hat{l}(\kappa)$ can be selected as

$$\hat{l}(\kappa) = \frac{l(\kappa_1) - l(\kappa_2)}{\kappa_1 - \kappa_2}\kappa + \frac{\kappa_1 l(\kappa_2) - \kappa_2 l(\kappa_1)}{\kappa_1 - \kappa_2}, \quad \text{(E.2)}$$

which is in fact the segment between point $(\kappa_1, l(\kappa_1))$ and point $(\kappa_2, l(\kappa_2))$. Since $l(\kappa)$ is convex, we know $\hat{l}(\kappa) \geq l(\kappa)$ for $\kappa \in \mathcal{X}$.

The convex lower bound for $p(\kappa)$ given in [37] is

$$\hat{p}(\kappa) = \max\{v_{\min}, w(\kappa)\}, \quad \text{(E.3)}$$

where $v_{\min}$ is the minimum of $p(\kappa)$ over $\mathcal{X}_0$, which is easy to obtain because $p(\kappa)$ is convex, and $w(\kappa)$ is defined as

$$w(\kappa) = p(\kappa_{\mathrm{b}}) + \nabla p(\kappa)\big|_{\kappa=\kappa_{\mathrm{b}}}(\kappa - \kappa_{\mathrm{b}}), \quad \text{(E.4)}$$

where $\nabla p(\kappa)$ denotes the gradient of $p(\kappa)$, and κ_{b} is the barycenter of $\mathcal{X}$, which is $\frac{1}{2}(\kappa_1 + \kappa_2)$ in our problem.

We can find that $w(\kappa)$ is the tangent line of $p(\kappa)$ at κ_{b}. Since $p(\kappa)$ is a convex function, we know $w(\kappa) \leq p(\kappa)$. Then, we can observe from (E.3) that $\hat{p}(\kappa)$ is a positive convex function on $\mathbb{R}$. With $\hat{l}(\kappa)$ and $\hat{p}(\kappa)$, the upper bound of $\frac{l(\kappa)}{p(\kappa)}$ can be obtained by solving the problem

$$\max_{\kappa} \frac{\hat{l}(\kappa)}{\hat{p}(\kappa)} \quad s.t.\ \kappa \in \mathcal{X}_k, \quad \text{(E.5)}$$

which is a quasi-convex problem and can be solved efficiently.

REFERENCES

[1] H. Tullberg, P. Popovski, Z. Li, M. A. Uusitalo, A. Hoglund, O. Bulakci, M. Fallgren, and J. F. Monserrat, "The METIS 5G system concept: Meeting the 5G requirements," *IEEE Commun. Mag.*, vol. 54, no. 12, pp. 132–139, Dec. 2016.

[2] D. Lpez-Prez, M. Ding, H. Claussen, and A. H. Jafari, "Towards 1 Gbps/UE in cellular systems: Understanding ultra-dense small cell deployments," *IEEE Commun. Surveys Tuts.*, vol. 17, no. 4, pp. 2078–2101, Fourth quarter 2015.

[3] Z. Shen, A. Khoryaev, E. Eriksson, and X. Pan, "Dynamic uplink-downlink configuration and interference management in TD-LTE," *IEEE Commun. Mag.*, vol. 50, no. 11, pp. 51–59, Nov. 2012.

[4] Report ITU-R M.2370, "IMT Traffic estimates for the years 2020 to 2030," 2015.

[5] S. Chen and J. Zhao, "The requirements, challenges, and technologies for 5G of terrestrial mobile telecommunication," *IEEE Commun. Mag.*, vol. 52, no. 5, pp. 36–43, May 2014.

[6] A. Dowhuszko, O. Tirkkonen, J. Karjalainen, T. Henttonen, and J. Pirskanen, "A decentralized cooperative uplink/downlink adaptation scheme for TDD small cell networks," in *Proc. IEEE PIMRC*, 2013.

[7] V. Venkatasubramanian, M. Hesse, P. Marsch, and M. Maternia, "On the performance gain of flexible UL/DL TDD with centralized and decentralized resource allocation in dense 5G deployments," in *Proc. IEEE PIMRC*, 2014.

[8] I. Sohn, K. B. Lee, and Y. S. Choi, "Comparison of decentralized time slot allocation strategies for asymmetric traffic in TDD systems," *IEEE Trans. Wireless Commun.*, vol. 8, no. 6, pp. 2990–3003, Jun. 2009.

[9] H. Lee and D.-H. Cho, "Combination of dynamic-TDD and static-TDD based on adaptive power control," in *Proc. IEEE VTC Fall*, 2008.

[10] H. Sun, M. Sheng, M. Wildemeersch, and T. Quek, "Coverage analysis for two-tier dynamic TDD heterogeneous networks," in *Proc. IEEE GLOBECOM*, 2014.

[11] P. Jayasinghe, A. Tolli, J. Kaleva, and M. Latva-aho, "Bi-directional signaling for dynamic TDD with decentralized beamforming," in *Proc. IEEE ICC Workshop*, 2015.

[12] D. Zhu and M. Lei, "Cluster-based dynamic DL/UL reconfiguration method in centralized RAN TDD with trellis exploration algorithm," in *Proc. IEEE WCNC*, 2013.

[13] M. Ding, D. L. Perez, G. Mao, and Z. Lin, "Dynamic TDD: The asynchronous case and the synchronous case," *arXiv preprint arXiv:1611.02828*, 2017.

[14] D. Korpi, M. Heino, C. Icheln, K. Haneda, and M. Valkama, "Compact inband full-duplex relays with beyond 100 dB self-interference suppression: Enabling techniques and field measurements," *IEEE Trans. Antennas and Propagation*, vol. 65, no. 2, pp. 960–965, Feb. 2017.

[15] M. Jain, J. Choi, T. Kim, D. Bharadia, S. Seth, K. Srinivasan, P. Levis, S. Katti, and P. Sinha, "Practical, real-time, full duplex wireless," in *Proc. ACM MOBICOM*, 2011.

[16] M. Duarte, C. Dick, and A. Sabharwal, "Experiment-driven characterization of full-duplex wireless systems," *IEEE Trans. Wireless Commun.*, vol. 11, no. 12, pp. 4296–4307, Dec. 2012.

[17] T. Riihonen, S. Werner, and R. Wichman, "Hybrid full-duplex/half-duplex relaying with transmit power adaptation," *IEEE Trans. Wireless Commun.*, vol. 10, no. 9, pp. 3074–3085, Sep. 2011.

[18] H. Alves, C. de Lima, P. Nardelli, R. Demo Souza, and M. Latva-aho, "On the average spectral efficiency of interference-limited full-duplex networks," in *Proc. CROWNCOM*, 2014.

[19] S. Goyal, P. Liu, S. Hua, and S. Panwar, "Analyzing a full-duplex cellular system," in *Proc. CISS*, 2013.

[20] J. Lee and T. Quek, "Hybrid full-/half-duplex system analysis in heterogeneous wireless networks," *IEEE Trans. Wireless Commun.*, vol. 14, no. 5, pp. 2883–2895, May 2015.

[21] Z. Tong and M. Haenggi, "Throughput analysis for full-duplex wireless networks with imperfect self-interference cancellation," *IEEE Trans. Commun.*, vol. 63, no. 11, pp. 4490–4500, Nov. 2015.

[22] A. Alammouri, H. Elsawy, and M. S. Alouini, "Flexible design for α-duplex communications in multi-tier cellular networks," *IEEE Trans. Commun.*, vol. 64, no. 8, pp. 3548–3562, Aug. 2016.

[23] A. Alammouri, H. Elsawy, O. Amin, and M. S. Alouini, "In-band α-duplex scheme for cellular networks: A stochastic geometry approach," *IEEE Trans. Wireless Commun.*, vol. 15, no. 10, pp. 6797–6812, Oct. 2016.

[24] C. Psomas, M. Mohammadi, I. Krikidis, and H. A. Suraweera, "Impact of directionality on interference mitigation in full-duplex cellular networks," *IEEE Trans. Wireless Commun.*, vol. 16, no. 1, pp. 487–502, Jan. 2017.

[25] R. K. Mungara and A. Lozano, "Interference surge in full-duplex wireless systems," in *Proc. Asilomar Conf. Signals, Syst. Comput.*, 2015.

[26] I. Randrianantenaina, H. ElSawy, and M.-S. Alouini, "Limits on the capacity of in-band full duplex communication in uplink cellular networks," in *Proc. IEEE GLOBECOM*, 2015.

[27] M. G. Sarret, G. Berardinelli, N. H. Mahmood, M. Fleischer, P. Mogensen, and H. Heinz, "Analyzing the potential of full duplex in 5G ultra-dense small cell networks," *EURASIP J. Wireless Commun. Networking*, vol. 2016, no. 284, 2016.

[28] M. G. Sarret, D. Catania, G. Berardinelli, and N. H. Mahmood, "Full duplex communication under traffic constraints for 5G small cells," in *Proc. IEEE VTC*, 2015.

[29] J. Liu, S. Han, W. Liu, Y. Teng, and N. Zheng, "Performance gain of full duplex over half duplex under bidirectional traffic asymmetry," in *Proc. IEEE ICC Workshops*, 2016.

[30] S. Wang, X. Zhang, J. Zhang, J. Feng, W. Wang, and K. Xin, "An approach for spatial-temporal traffic modeling in mobile cellular networks," in *Proc. International Teletraffic Congress*, 2015.

[31] S. Goyal, C. Galiotto, N. Marchetti, and S. Panwar, "Throughput and coverage for a mixed full and half duplex small cell network," in *Proc. IEEE ICC*, 2016.

[32] T. D. Novlan, H. S. Dhillon, and J. G. Andrews, "Analytical modeling of uplink cellular networks," *IEEE Trans. Wireless Commun.*, vol. 12, no. 6, pp. 2669–2679, June 2013.

[33] U. Paul, A. P. Subramanian, M. M. Buddhikot, and S. R. Das, "Understanding traffic dynamics in cellular data networks," in *Proc. IEEE INFOCOM*, 2011.

[34] Global Internet Phenomena Report: Sandvine-Intelligent Broadband Networks, www.sandvine.com, Nov. 2014.

[35] 3GPP TR 36.814, "Further advancements for E-UTRA physical layer aspects (Release 9)," 2010.

[36] A. Sadek and S. Shellhammer, "A unified path-loss model for coexistence study," 2008. [Online]. Available: https://mentor.ieee.org/802.19/file/08/19-08-0006-00-0000-unified-path-loss-model.doc

[37] H. P. Benson, "Maximizing the ratio of two convex functions over a convex set," *Naval Research Logistics*, vol. 53, no. 4, pp. 309–317, Mar. 2006.