Leveraging D2D Communication to Maximize the Spectral Efficiency of Massive MIMO Systems

Asma Afzal[†◇], Afef Feki[*], Merouane Debbah[*◇], Syed Ali Zaidi[†], Mounir Ghogho[†‡] and Des McLernon[†]
[†] University of Leeds, United Kingdom, [◇]CentraleSupélec, Gif-sur-Yvette, France
[*]Mathematical and Algorithmic Sciences Lab, Huawei, France, [‡]International University of Rabat, Morocco
Email: {elaaf, s.a.zaidi, d.c.mclernon, m.ghogho}@leeds.ac.uk, {afef.feki, merouane.debbah}@huawei.com

***Abstract*—In this article, we investigate how the performance of Massive MIMO cellular systems can be enhanced by introducing D2D communication. We consider a scenario where the base station (BS) is equipped with large, but finite number of antennas and the total number of UEs is kept fixed. The key design question is that what fraction of users should be offloaded to D2D mode in order to maximize the aggregate cell level throughput. We demonstrate that there exists an optimal user offload fraction, which maximizes the overall capacity. This fraction is strongly coupled with the network parameters such as the number of antennas at the BS, D2D link distance and the transmit SNR at both the UE and the BS and careful tuning of the offload fraction can provide up to $5\times$ capacity gains.**[1]

I. Introduction

Massive MIMO systems are envisioned to be a key enabler for 5G cellular network because of their capability to simultaneously serve multiple user equipments (UE) on a single resource block (time/frequency) [1]. The distinct feature of massive MIMO is that the number of antennas deployed at the BS is much larger than the number of UEs to be served. This allows for significant improvements in link reliability and data rates due to increased spatial directivity. The additional degrees of freedom alleviate the need for sophisticated signal processing and simple linear processing achieves near-optimal performance [2]. In this paper, we study massive MIMO cellular network which exploits device-to-device (D2D) communication as an overlaid network to enhance aggregate throughput. D2D provisions direct communication between UEs without the intervention of the BS [3]. The short range of D2D communication improves coverage and it also reduces the burden of access on the BS and the core network.

D2D networks have been studied extensively in the context of cellular networks with BSs equipped with a single antenna. However, researchers have recently begun to analyze the potential gains of D2D communication with massive MIMO. In [4] and [5], the authors study an isolated cell with a single cellular UE and D2D pair and investigate how the excess antennas at BS can eliminate the interference at the D2D receiver. The sum capacity of an isolated cell with a fixed number of cellular UEs and a random number of D2D pairs has been studied in [6] for the case of cellular uplink (UL). Expressions for signal-to-interference-and-noise ratio (SINR) are derived for both the cellular and D2D cases for fixed spatial locations of UEs and the randomness is accounted for in simulations. The corresponding downlink (DL) analysis is conducted in [7] and the density of D2D pairs maximizing the sum capacity is explored.

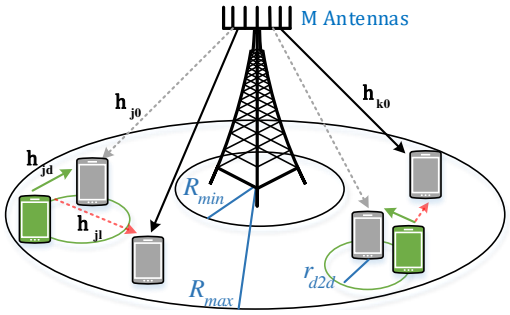


Figure 1: System Model.

The research on massive MIMO with D2D thus far does not consider dynamic mode selection for the UEs. It is only in [8] that the authors consider mode switching for a UE (between cellular and D2D) in cellular uplink (UL) for a simple network setting with a single D2D pair while assuming fixed locations of interfeing D2D pairs. The obtained results, however, cannot be directly translated to downlink (DL) and scaled for multiple randomly located D2D pairs. This is because, the interference from active D2D pairs is strongly coupled with the distance between UE-UE links and will significantly impact the findings.

Motivated by this, we study the UE offloading problem for a single cell scenario in DL, where a fixed number of UEs N is distributed uniformly around the BS. We focus on D2D in DL time slot as it is more suited for massive MIMO scenario. This is because the BS can make use of the excess degrees of freedom to interference at the D2D receivers, whereas this is not possible in the UL with single antenna UEs [4], [5], [9], [10]. While D2D communication between UEs in close proximity can provide high data rates, the transmit power of BS is much higher than a UE and it is not clear under what circumstances offloading is a better choice. There exists an inherent tradeoff as offloading UEs to operate in D2D mode will improve the SE, but at the same time, a large number of D2D UEs may induce higher interference causing

[1]This research has been supported by the ERC Starting Grant 305123 MORE (Advanced Mathematical Tools for Complex Network Engineering).

the SE to drop. The incentive of this work is to answer the following question: *Given a certain number of UEs inside a cell, what is the optimal offload fraction which maximizes the sum capacity in a massive MIMO system and how the system parameters affect this fraction?*. Our main contribution is to explore this trade off and derive closed-form expressions for the approximation of the unconditional overall capacity.

II. System Model

We consider a TDD DL transmission scenario where the BS is equipped with M antennas and $N < M$ single antenna UEs distributed uniformly in an annular region of inner radius R_{min} and outer radius R_{max} centered at the BS as shown in Fig. 1. K out of N UEs are served directly by the BS, while the remaining $(N-K)$ UEs are offloaded to D2D mode. Each of the $(N-K)$ D2D receiving UEs is associated to a unique D2D transmitter UE located randomly at the perimeter of a disk of radius r_{d2d} centered at the UEs. These transmitters can be thought of as idle UEs which can establish D2D connections with their neighboring UEs to share previously downloaded files [11].

Without any loss of generality, the set of all N UE locations can be written as $\mathcal{U} = \{\mathbf{x}_1, .., \mathbf{x}_K, \mathbf{x}_{K+1}, .., \mathbf{x}_N\}$. Assuming that the BS is located at the origin, the distance between the k^{th} UE and the BS $r_{k0} = \|\mathbf{x}_k\|$ is distributed as

$$f_{r_{k0}}(x) = \frac{2x}{R_{max}^2 - R_{min}^2}, R_{min} \leq x \leq R_{max}. \quad (1)$$

We adopt a simple power-law path loss model where the signal power attenuates according to $r^{-\alpha_b}, m = \{b, d\}$, where r is the distance separation and α_m denotes the path loss exponent in mode m where $m = b$ stands for cellular mode when the UE is served by the BS and $m = d$ stands for D2D mode. The BS-UE and UE-UE links suffer from small scale Rayleigh fading. This implies that the channel gain is independent and identically distributed (i.i.d) complex Gaussian variable with zero mean and unit variance. We further assume that the D2D pairs share the same resources as the cellular UEs and hence, both the BS-UE and UE-UE links interfere with each other. The BS is considered to have full channel state information (CSI) of the UEs and it employs zero-forcing beamforming (ZFBF) precoding. As a result, there is no signal leakage within the cellular UEs. The BS transmits a total power p_b, which is equally distributed for cellular UEs and the D2D UEs transmit a fixed power p_d, where $p_d < p_b$. The preliminary analysis for the SINR at the UEs in cellular and D2D modes is presented as follows.

A. Cellular Mode

The signal received at the k^{th} cellular UE under ZFBF can be written as

$$y_k = \sqrt{\frac{p_b r_{k0}^{-\alpha_b}}{K}} \left(\mathbf{h}_{k0}^{BS-UE}\right)^H \mathbf{w}_{k0}^{BS} s_{k0}^{BS-UE} + \sqrt{p_d} \sum_{l=1}^{N-K} \sqrt{r_{kl}^{-\alpha_d}} h_{kl}^{UE-UE} s_l^{UE} + v_k^{BS}, \quad (4)$$

where, $\mathbf{h}_{k0}^{BS-UE} \in \mathbb{C}^{M\times 1}$ is a vector of M Rayleigh fading channel gains, v_k^{BS} is the zero mean additive white Gaussian noise (AWGN) with variance σ_{BS}^2, s_{k0}^{BS-UE} is the complex scalar signal and $\mathbf{w}_{k0} \in \mathbb{C}^{M\times 1}$ is the precoding vector. To satisfy the maximum BS power constraint, s_{k0}^{BS-UE} and $\mathbf{w}_{k0}^{BS} = \mathbf{g}_{k0}^{BS}/\|\mathbf{g}_{k0}^{BS}\|^2$ are normalized such that $\mathbb{E}\left[\left\|s_{k0}^{BS-UE}\right\|^2\right] = \|\mathbf{w}_{k0}\|^2 = 1$. The un-normalized precoding vector $\mathbf{g}_{k0}^{BS}$ for ZFBF is given as $\mathbf{G}^{BS} = \mathbf{H}^{BS-UE}\left(\left(\mathbf{H}^{BS-UE}\right)^H \mathbf{H}^{BS-UE}\right)^{-1}$, where $\mathbf{H}^{BS-UE} = \left[\mathbf{h}_{10}^{BS-UE}, .., \mathbf{h}_{K0}^{BS-UE}\right]$ and $\mathbf{G}^{BS-UE} = \left[\mathbf{g}_{10}^{BS-UE}, .., \mathbf{g}_{K0}^{BS-UE}\right]$. The second term in (4) denotes the interference signal from all the $(N-K)$ active D2D transmitters to the k^{th} cellular UE, where $r_{kl} = \|\mathbf{x}_k - \mathbf{x}_l\|$ is the distance between the k^{th} UE and the l^{th} D2D transmitter and s_l^{UE} is the information symbol transmitted by the l^{th} D2D transmitter. The average SE for the k^{th} UE in cellular mode can be written as $SE_k^{BS-UE} = \mathbb{E}\left[\log_2\left(1 + SINR_k^{BS-UE}\right)\right]$, where

$$SINR_k^{BS-UE} = \frac{\gamma_b r_{k0}^{-\alpha_b} \left\|\left(\mathbf{h}_{k0}^{BS-UE}\right)^H \mathbf{w}_{k0}^{BS}\right\|^2}{\gamma_d \sum_{l=1}^{N-K} r_{kl}^{-\alpha_d} \left\|h_{kl}^{UE-UE}\right\|^2 + 1}, \quad (5)$$

where $\gamma_b = p_b/\sigma_{BS}^2$ and $\gamma_d = p_d/\sigma_{UE}^2$ are the cellular and D2D transmit signal-to-noise ratios (SNRs).

B. D2D Mode

The signal received at the j^{th} UE x_j in D2D mode from its corresponding dth D2D transmitter can be written as $y_j = \sqrt{p_d r_{d2d}^{-\alpha_d}} h_{jd}^{UE-UE} s_d^{UE} + I_j^{UE-UE} + I_j^{BS-UE} + v_d^{UE}$, where, v_d^{UE} is the zero mean additive white Gaussian noise (AWGN) with variance σ_{UE}^2, $I_j^{UE-UE} = \sqrt{p_d} \sum_{l\neq d}^{N-K} \sqrt{r_{jl}^{-\alpha_d}} h_{jl}^{UE-UE} s_l^{UE}$ is the interference signal received by the j^{th} UE in D2D mode from other active D2D transmitters and $I_j^{BS-UE} = \sqrt{\frac{p_b r_{j0}^{-\alpha_b}}{K}} \sum_{k=1}^{K} \left(\mathbf{h}_{j0}^{BS-UE}\right)^H \mathbf{w}_{k0}^{BS} s_{k0}^{BS-UE}$ is the interference from the BS. The average SE for the j^{th} UE in D2D mode can then be written as $SE_j^{UE-UE} = \mathbb{E}\left[\log_2\left(1 + SINR_j^{UE-UE}\right)\right]$, where

$$SINR_j^{UE-UE} = \frac{\gamma_d r_{d2d}^{-\alpha_d} \left\|h_{jd}^{UE-UE}\right\|^2}{\left\|I_j^{BS-UE}\right\|^2/\sigma_{UE}^2 + \left\|I_j^{UE-UE}\right\|^2/\sigma_{UE}^2 + 1}. \quad (6)$$

III. Spectral Efficiency Analysis

The goal of this work is to evaluate the optimal fraction of UEs to be offloaded in D2D mode. We define our performance metric as follows.

Definition 1. Given a fixed number of UEs N, the maximum attainable overall capacity is given by the following optimiza-

$$SE_k^{BS-UE}|r_{kl}\left(\beta_{kl}\right) \approx \frac{\log_2\left(1+\beta_{kl}\right)}{R_{max}^2} + \frac{2\sqrt{\beta_{kl}}}{\ln\left(2\right)}\tan^{-1}\left(\sqrt{\frac{1}{\beta_{kl}}}\right) \tag{2}$$

$$SE_j^{UE-UE}|r_{jl}\left(\beta_{jl1},\beta_{jl2}\right) \approx \frac{1}{R_{max}^2}\log_2\left(1+\frac{\gamma_d r_{d2d}^{-4}}{\beta_{jl1}+\gamma_b/R_{max}^4}\right) + \frac{2}{\ln\left(2\right)}\left[\sqrt{\frac{\gamma_b/R_{max}^4}{\beta_{jl2}+\gamma_d r_{d2d}^{-4}}}\tan^{-1}\left(\sqrt{\frac{\beta_{jl2}+\gamma_d r_{d2d}^{-4}}{\gamma_b/R_{max}^4}}\right) - \sqrt{\frac{\gamma_b/R_{max}^4}{\beta_{jl2}}}\tan^{-1}\left(\sqrt{\frac{\beta_{jl2}}{\gamma_b/R_{max}^4}}\right)\right] \tag{3}$$

tion problem

$$C_{tot} = \max_{K} \sum_{k=1}^{K} SE_k^{BS-UE} + \sum_{j=K+1}^{N} SE_j^{UE-UE}$$

where K is the number of UEs in cellular mode. Our aim is to evaluate $\mu^* = {}^{(N-K^*)}/_{N}$, which is the optimal offload fraction.

In the following subsections, we present our analysis pertaining to the cellular and D2D SEs.

A. Cellular Mode

The following Lemma provides the SE of a UE in cellular mode conditioned on UE locations.

Lemma 1. *Conditioned on the location of the UEs, the average SE of a UE in cellular mode can be approximated as*

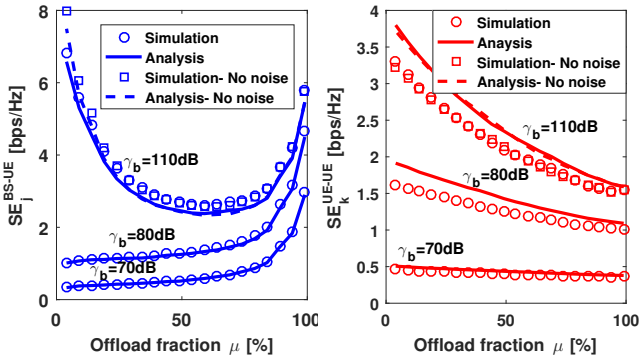
$$SE_k^{BS-UE}|r_{k0},r_{kl} \approx \log_2\left(1+\frac{\gamma_b\left(M-K\right)r_{k0}^{-\alpha_b}}{K\left(1+\gamma_d\sum_l^{N-K} r_{kl}^{-\alpha_d}\right)}\right). \tag{7}$$

Proof: Since $\log\left(1+x^{-1}\right)$ is convex in x, we employ Jensen's inequality to obtain

$$SE_k^{BS-UE}|r_{k0},r_{kl} \leq \log_2\left(1+\mathbb{E}\left[\frac{1}{SINR_k^{BS-UE}}\right]^{-1}\right). \tag{8}$$

The inverse of the desired power in (5) is an inverse chi-squared random variable such that $\frac{2}{\left\|\left(\mathbf{h}_{k0}^{BS-UE}\right)^H\mathbf{w}_{k0}^{BS}\right\|^2} \sim$ Inv-$\chi^2_{2(M-K+1)}$. This is because the isotropic M dimensional vector is projected onto $M-K+1$ dimensional beamforming space [9]. As a consequence, we have $\mathbb{E}\left[\frac{1}{\left\|\left(\mathbf{h}_{k0}^{BS-UE}\right)^H\mathbf{w}_{k0}^{BS}\right\|^2}\right] = (M-K)^{-1}$. The interference power from each D2D UE is a unit mean exponential random variable. $\left\|h_{kl}^{UE-UE}\right\|^2 \sim \exp(1)$. Exploiting the independence of these random variables and plugging in the expected power values in (8), we obtain (7). The approximation sign is used as the bound is very tight for a wide range of network parameters. ■

We de-condition $SE_k^{BS-UE}|r_{j0},r_{jl}$ in (7) with respect to distances in the following Lemma 2 and Proposition 1.



(a) Cellular mode. (b) D2D mode.

Figure 2: Effect of the offload fraction on the SE of an arbitrary UE.

Lemma 2. *The average SE of an arbitrary UE in cellular mode conditioned on the location of interfering D2D UEs can be approximated in closed-form for $\alpha_b = \alpha_d = 4$ as (2) where $\beta_{kl}\left(\psi\right) = \frac{\gamma_b(M-K)}{K(1+\gamma_d\psi)R_{max}^4}$ and $\psi = \sum_{l=1}^{N-K} r_{kl}^{-4}$.*

Proof: The proof follows by averaging (7) over r_{k0} which is distributed according to (1). ■

Proposition 1. *The bounds on the unconditional average SE of a UE in cellular mode $SE_{k,LB}^{BS-UE} \leq SE_k^{BS-UE} \leq SE_{k,UB}^{BS-UE}$ can be written in closed-form as*

$$SE_{k,UB}^{BS-UE} = SE_k^{BS-UE}|r_{kl}\left(\beta_c^{UB}\right), \tag{9}$$

where $\beta_c^{UB} = \beta_{kl}\left(\psi_c^{UB}\right)$ with $\psi_c^{UB} = (N-K)\,\mathbb{E}\left[r_{kl}\right]^{-4}$ and $\mathbb{E}\left[r_{kl}\right] = \frac{128}{4\pi}R_{max}$ for the upper bound and

$$SE_{k,LB}^{BS-UE} = SE_k^{BS-UE}|r_{kl}\left(\beta_c^{LB}\right), \tag{10}$$

where $\beta_c^{LB} = \beta_{kl}\left(\psi_c^{LB}\right)$, with $\psi_c^{LB} = (N-K)\,\mathbb{E}\left[r_{kl}^{-4}\right]$ and

$$\mathbb{E}\left[r_{kl}^{-4}\right] \approx \rho_g^{-1}\left[-\frac{3\sqrt{(4R_{max}^2-1)}}{4R_{max}^2}, + \left(1+\frac{1}{2R_{max}^2}\right)\cos^{-1}\left(\frac{1}{2R_{max}}\right)\right]$$

$\rho_g = \sqrt{(4R_{max}^2-1)}\left(\frac{2R_{max}^2+1}{8R_{max}^2}\right)$ *for the lower bound.*

Proof: The first term in (2) is of the form $\log\left(1+\left(A+B\,r_{kl}^{-4}\right)^{-1}\right)$, which is concave in r_{kl} and

Parameter	Value
BS antennas M, Total users N	200, 100
BS coverage radius (Outer R_{max}, Inner R_{min})	200 m, 2m
D2D range r_{d2d}	12 m
Path loss exponents: α_c, α_d	4, 4
Ratio of cellular and D2D SNR $(\gamma_b/\gamma_d)_{\text{dB}}$	30 dB

Table I: List of simulation parameters.

convex in r_{kl}^{-4} for $A, B > 0$. The second term is of the form $g\left(r_{kl}, A, B\right) = \left(A + B\, r_{kl}^{-4}\right)^{-1/2} \tan^{-1}\left(\left(A + B\, r_{kl}^{-4}\right)^{1/2}\right)$, which can be easily shown to be a non-decreasing quasi-concave function satisfying the necessary condition $\left(r_2 - r_1\right) g'\left(A, B, r_1\right) \geq 0$, where $r_2 > r_1$. The function $g(r_{kl}, A, B)$ transitions from convex to concave and contains a point of inflection at $r_{kl} > 0$. Simulations reveal that this inflection point takes a very small value of $r_{kl} < 10$m for various values of A and B.. Since $0 < r_{kl} < 2R_{max}$, we treat $g(r, A, B)$ as concave and make use of Jensen's inequality to shift the expectation operator inside these functions to obtain the bounds[2]. The D2D UEs are i.i.d distributed and their respective transmitters are uniformly located at a fixed distance r_{d2d}. For tractability, we assume that $R_{min} = 0$. This does not impact the result as $R_{max} \gg R_{min}$. The effective distance between the k^{th} UE and l^{th} transmitting D2D UE is then distributed according to [12]

$$f_{r_{kl}}(x) = \frac{2x}{\pi R_{max}^2}\left(2\cos^{-1}\left(\frac{x}{2R_{max}}\right) - \frac{x}{R_{max}}\sqrt{1-\left(\frac{x}{2R_{max}}\right)^2}\right),\ 0 \leq x \leq 2R_{max},$$

where $\mathbb{E}\left[r_{kl}\right] = \frac{128}{4\pi} R_{max}$. It is slightly tricky to obtain $\mathbb{E}\left[r_{kl}^{-4}\right]$. Since $f_{r_{kl}}(0) = 2/R_{max}^2 > 0$, it implies that the expectation $\mathbb{E}\left[r_{kl}^{-4}\right]$ is unbounded even when the cell size is large. To tackle this issue and avoid singularity, we introduce a minimum separation distance of 1m. Therefore, we have $\mathbb{E}\left[r_{kl}^{-4} | r_{kl} \geq 1\right] = \int_{x=1}^{2R_{max}} x f_{r_{kl}}(x | r_{kl} \geq 1)\, dx$, where $f_{r_{kl}}(x | r_{kl} \geq 1) = {}^{f_{r_{kl}}(x)}/_{\mu_g}, 1 \leq x \leq 2R_{max}$ and $\mu_g = \mathbb{P}\left[r_{kl} \geq 1\right]$. This completes the proof. ∎

B. D2D Mode

The SE of a UE in D2D mode conditioned on UE locations is given by the following Lemma.

Lemma 3. *Conditioned on the location of the UEs, the average SE of a UE in D2D mode can be approximated as*

$$SE_j^{UE-UE} | r_{j0}, r_{jl} \approx \log_2\left(1 + \frac{\gamma_d r_{d2d}^{-\alpha_d}}{1 + \gamma_d \sum_{l \neq d}^{N-K} r_{jl}^{-\alpha_d} + \gamma_b r_{j0}^{-\alpha_c}}\right). \quad (11)$$

[2]The analysis in the subsequent sections also assumes $g(r, A, B)$ to be concave.

Proof: We follow a different approach compared to the proof of Lemma 1. Since the desired power is exponentially distributed $\left\|h_{jd}^{UE-UE}\right\|^2 \sim \exp(1)$, the expected value of its inverse does not exist. We therefore exploit the concavity of $\log(1 + x)$ to obtain $\log_2\left(1 + \mathbb{E}\left[SINR_j^{UE-UE}\right]\right)$. We can write $\mathbb{E}\left[SINR_j^{UE-UE}\right] = \int_0^\infty \mathrm{e}^{-s_z} \mathbb{E}\left[\exp\left(-s_z {I_j}^{BS-UE}\right)\right] \mathbb{E}\left[\exp\left(-s_z {I_j}^{UE-UE}\right)\right] dz$, where $s_z = {}^{z\, r_{d2d}^{-\alpha_d}}/_{\gamma_d}$. Since $\left(\mathbf{h}_{j0}^{BS-UE}\right)^H$ is independent of $\mathbf{w}_k^{BS}$, $\left\|\left(\mathbf{h}_{j0}^{BS-UE}\right)^H \mathbf{w}_k^{BS}\right\|^2 \sim \exp(1)$. $\left\|{I_j}^{BS-UE}\right\|^2$ is the superposition of K independent data streams, which implies $\sum_{k=1}^{K} 2\left\|\left(\mathbf{h}_{j0}^{BS-UE}\right)^H \mathbf{w}_k^{BS}\right\|^2 \sim \chi_{2K}^2$. For the D2D interference power $\left\|{I_j}^{UE-UE}\right\|^2$, we have $\left\|h_{jl}^{UE-UE}\right\|^2 \sim \exp(1)$. To ensure tractability, we invoke Jensen's inequality once again to draw the expectation inside the exponential to obtain (11). ∎

Similar to the analysis for cellular mode, we derive the expressions for unconditional SE of a UE in D2D mode as follows.

Lemma 4. *The average SE of an arbitrary UE in D2D mode conditioned on the location of interfering D2D UEs can be approximated in closed-form for $\alpha_b = \alpha_d = 4$ as (3) where $\beta_{jl1}(\psi) = \beta_{jl2}(\psi) = 1 + \gamma_d \psi$ and $\psi = \sum_{l \neq d}^{N-K} r_{jl}^{-4}$.*

Proof: The proof follows by averaging (11) over r_{k0}. ∎

Proposition 2. *The bounds on the unconditional average SE of a UE in D2D mode $SE_{j,LB}^{UE-UE} \leq SE_j^{UE-UE} \leq SE_{UB}^{UE-UE}$ can be written in closed-form as*

$$SE_{j,UB}^{UE-UE} = SE_j^{UE-UE} | r_{jl}\left(\beta_d^{UB}, \beta_d^{LB}\right), \quad (12)$$

where $\beta_d^{UB} = \beta_{jl1}\left(\psi_d^{UB}\right)$ with $\psi_d^{UB} = (N - K - 1)\,\mathbb{E}\left[r_{kl}\right]^{-4}$ and $\beta_d^{LB} = \beta_{jl1}\left(\psi_d^{LB}\right)$ with $\psi_d^{LB} = (N - K - 1)\,\mathbb{E}\left[r_{kl}^{-4}\right]$ for the upper bound and

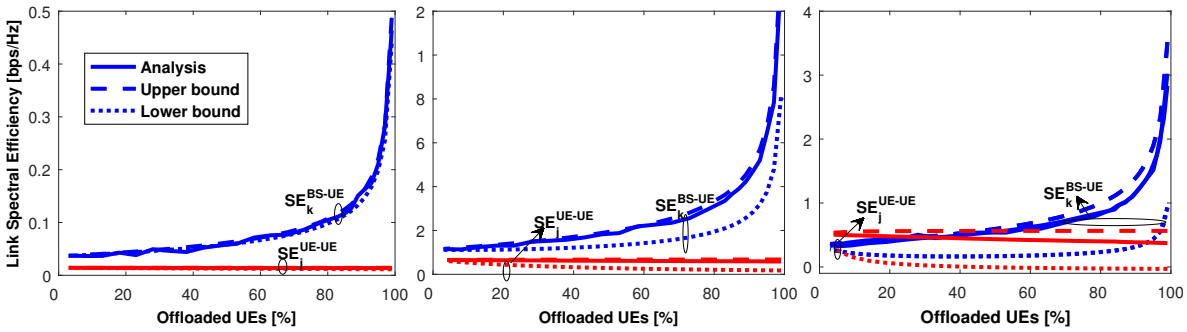
$$SE_{j,LB}^{UE-UE} = SE_j^{UE-UE} | r_{kl}\left(\beta_d^{LB}, \beta_d^{UB}\right), \quad (13)$$

for the lower bound.

Proof: The proof is similar to that of Prop. 1 with the exception that there is a negative sign inside the second term of (3). It can be easily shown that for $g\left(r_{jl}, A, B\right)$, we have $g\left(r_{jl}, A_1, B\right) \leq g\left(r_{jl}, A_2, B\right)$ for $A_1 \geq A_2$ or $A_1 - A_2 = \gamma_d r_{d2d}^{-4} > 0$. Therefore, if we re-write $SE_j^{UE-UE} | r_{jl} = T_1 + T_2$, then T_2 exhibits the opposite behavior of T_1. It is concave in r_{jl}^{-4} and convex in r_{jl}. The coefficient $(N - K - 1)$ in ψ_1 and ψ_2 denotes the number of interfering D2D pairs, excluding the one on which the performance is being measured. ∎

IV. Results and Discussion

In this section, we present the evaluation procedure adopted to assess the performance of the offloading mechanism. As a first step, we validate our analysis in (7) and (11) with the help of Monte Carlo simulations in Figs. 2a and 2b



(a) $\gamma_b = 50$dB (b) $\gamma_b = 60$dB (c) $\gamma_b = 70$dB

Figure 3: Bounds on SE_k^{BS-UE} and SE_j^{UE-UE} from Prop. 1 and 2.

respectively. The simulation parameters are listed in table I unless stated otherwise. The simulations are repeated for 10^4 network realizations and each realization, the SE is measured at an arbitrary UE operating in cellular or D2D mode. The SE obtained from (7) and (11) and the simulations is averaged over all the realizations and hence, the effect of link distances is also averaged out. We study the average SE per UE as function of the offload fraction $\mu = (N-K)/N$. Clearly, the analysis for both SE_k^{BS-UE} and SE_j^{UE-UE} is in good agreement with the simulations for various transmit SNR[3] values γ_b and γ_d. In addition, we study the SE for the case when there is no noise, i.e. $v_k^{BS}, v_d^{UE} = 0$ or alternatively $\gamma_d, \gamma_b \to \infty$. In that case, the analysis in (7) and (11) reduces to

$$\lim_{\gamma_b,\gamma_d\to\infty} SE_k^{BS-UE}|r_{k0}, r_{kl} \approx \log_2\left(1 + \frac{\gamma_b/\gamma_d\,(M-K)\,r_{k0}^{-\alpha_c}}{K\sum_l^{N-K} r_{kl}^{-\alpha_d}}\right)$$

and

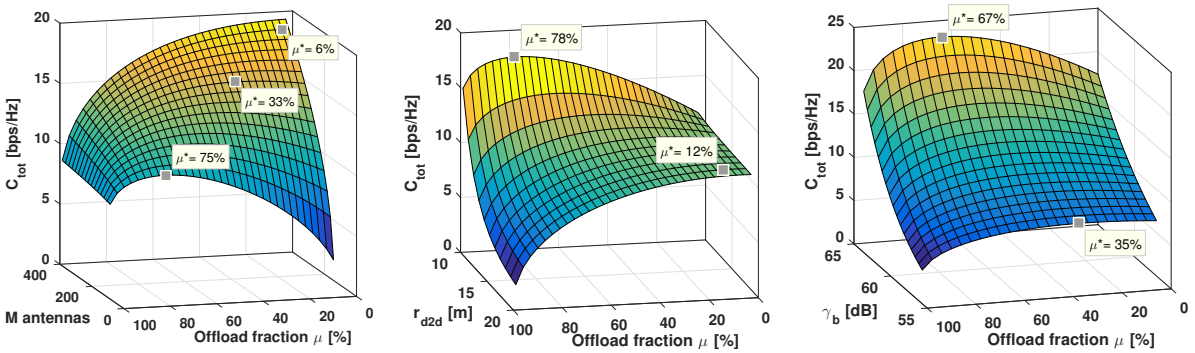
$$\lim_{\gamma_b,\gamma_d\to\infty} SE_j^{UE-UE}|r_{j0}, r_{jl} \approx \log_2\Bigg(1 + \frac{r_{d2d}^{-\alpha_d}}{\sum_{l\neq d}^{N-K} r_{jl}^{-\alpha_d} + \gamma_b/\gamma_d r_{j0}^{-\alpha_c}}\Bigg).$$

We observe that for low transmit SNR values γ_b and γ_d, SE_k^{BS-UE} increases monotonically, while there is a drop in SE_j^{UE-UE} with the increase in μ. We refer to this as the low-SNR (LS) regime. The rise in SE_k^{BS-UE} with the increase in μ is because as more UEs are offloaded to D2D mode, the number of the cellular UEs K inside the cell decreases and the power allocated to each cellular UE by the BS increases. The fall in SE_j^{UE-UE}, on the other hand, is due to the increasing interference from D2D UEs. This gap becomes more pronounced for higher values of γ_d. A different behavior is observed for SE_k^{BS-UE} in the high SNR (HS) regime; it closely resembles the case when $\gamma_d, \gamma_b \to \infty$, i.e. the system is interference-limited. The SE_k^{BS-UE} in HS regime is initially quite high when no UEs are offloaded. At offload percentage of 1%, the BS power is being distributed over $N-1$ UEs. Because of negligible D2D interference, a smaller allocated power is still sufficient to counter the $BS-UE$ link path loss in HS scheme. As more UEs are offloaded, the allocated power for cellular UE increases, but SE_k^{BS-UE} decreases steadily. This is because the increase in the BS power per UE is unable to cope with the increase in the D2D interference. After a certain fraction of UEs has been offloaded ($\mu \sim 50\%$), SE_k^{BS-UE} begins to rise again. This rise is now dominated by the increase in the allocated power per cellular UE. The value of SE_k^{BS-UE} at $\mu = 100\%$ is lower compared to that at $\mu = 1\%$ because of the adverse effects of the aggregate D2D interference power. In the rest of this paper, we will focus on the LS regime as HS regime is more suited for multi-cell environment, where inter-cell interference also plays a critical role. An interesting observation from Figs. 2a and 2b is that while SE_k^{BS-UE} monotonically increases in the LS regime and SE_j^{UE-UE} monotonically increases, there must exist an optimal offload fraction $\mu = \mu^*$ which maximizes C_{tot}.

After validation of our analysis, we study the accuracy of the bounds derived in Prop. 1 and 2. Fig. 3 shows that the bounds closely match SE_k^{BS-UE} and SE_j^{UE-UE} from (7) and (11) respectively. The bounds are fairly tight especially for low values of γ_b and γ_d. For high values of γ_b and γ_d, the bounds on SE_j^{UE-UE} begin to deviate significantly while the bounds on SE_k^{BS-UE} still remain tight. The upper bound is tighter compared to the lower bound for both SE_j^{UE-UE} and SE_k^{BS-UE}. For the rest of the discussion, we use the upper bounds $SE_{j,UB}^{UE-UE}$ and $SE_{k,UB}^{BS-UE}$ to analyze the overall capacity C_{tot}.

We study the behavior of C_{tot} with respect to μ in Figs. 4a-4c. We also explore the impact of key design parameters on the optimal offload fraction $\mu = \mu^*$and the corresponding C_{tot}. These parameters include, i) the number of antennas M at the BS, ii) D2D link distance r_{d2d}, and iii) the transmit

[3]The variation in transmit SNRs $\gamma_b = p_b/\sigma_{BS}^2$ and $\gamma_d = p_d/\sigma_{UE}^2$ is governed by several parameters including the BS and UE transmit powers p_b and p_d, the noise spectral density, etc. In this paper, we implicitly treat the effect of these parameters by varying γ_b and γ_d directly to assess the performance of our setup. To ensure a fair comparison, a fixed, positive ratio γ_b/γ_d is maintained.



(a) . (b) (c)

Figure 4: Effect of the number of antennas M, D2D link distance r_{d2d} and γ_d on C_{tot} and K^*: $\gamma_b = 60$dB

SNRs γ_b and γ_d. From (2), we see that the SE of cellular UE SE_k^{BS-UE} increases with the increase in M, while the SE of D2D UE SE_j^{UE-UE} in (3) does not depend on M. As M increases, more and more UEs can be offlaoded to D2D mode as seen from Fig. 4a. When M=N=100, it is better to offload 75% UEs in D2D mode while only 6% UEs should to be offloaded when M is increased to 400, while $N = 100$. Another important observation is that the selection of μ is crucial for smaller M. We can see that when $M = N = 100$, $C_{tot} = 2$ bps/Hz for $\mu = 3\%$, whereas $C_{tot} = 10$ bps/Hz for $\mu = 75\%$ giving 5 times better performance.

Fig. 4b shows the effect of D2D link distance r_{d2d} on C_{tot} and μ^*. The increase in r_{d2d} aggravates D2D link path loss and degrades SE_j^{UE-UE}, while SE_k^{BS-UE} is independent of r_{d2d}. We see that a high overall capacity C_{tot} can be achieved with smaller values of r_{d2d} and it is better to offload UEs in D2D mode if their respective D2D transmitter is located close by. We further notice that even a slight increase of a few meters in r_{d2d} significantly reduces gains in C_{tot} from offloading, thereby causing μ^* to drop. As a consequence, the BS has to carefully evaluate the offloading strategy based on the D2D link distances before scheduling UEs for transmission.

We also study the effect of γ_b and γ_d in Fig. 4c. We observe from (2) and (3) that as γ_b increases while γ_b/γ_d is fixed, both SE_k^{BS-UE} and SE_j^{BS-UE} increase causing C_{tot} to rise. The increase in SE_j^{BS-UE}, however, is more that the increase in SE_k^{BS-UE} as evident from Figs. 2b and 3. This implies that with higher SNR, more UEs should be offloaded to D2D mode to maximize C_{tot}. We see that for a 10dB rise in γ_b and γ_d, up to 30% more UEs can be offloaded to maximize C_{tot}.

V. Conclusion

In this paper we studied the performance gains achieved by employing network-assisted D2D communication in a massive MIMO system. We derived closed-form expressions for the spectral efficiency of an arbitrary UE in the cell in both D2D and cellular modes. Our results reveal that there exist a trade off between the number of UEs a BS can offload and the maximum achievable capacity of the cell. With careful selection of the offload fraction, the overall capacity of the cell can be improved up to 5×.

References

[1] D. Gesbert, M. Kountouris, R. W. Heath Jr, C.-B. Chae, and T. Salzer, "Shifting the MIMO paradigm," *IEEE Signal Process. Mag.*, vol. 24, no. 5, pp. 36–46, 2007.

[2] T. L. Marzetta, "Massive MIMO: an introduction," *Bell Labs Technical Journal*, vol. 20, pp. 11–22, 2015.

[3] K. Doppler, M. Rinne, C. Wijting, C. B. Ribeiro, and K. Hugl, "Device-to-device communication as an underlay to lte-advanced networks," *IEEE Commun. Mag.*, vol. 47, no. 12, pp. 42–49, 2009.

[4] Y. Ni, S. Jin, W. Xu, Y. Wang, M. Matthaiou, and H. Zhu, "Beamforming and interference cancellation for D2D communication underlaying cellular networks," *IEEE Trans. Commun.*, vol. 64, no. 2, pp. 832–846, 2016.

[5] W. Xu, L. Liang, H. Zhang, S. Jin, J. C. Li, and M. Lei, "Performance enhanced transmission in device-to-device communications: Beamforming or interference cancellation?" in *Global Communications Conference (GLOBECOM), 2012 IEEE*. IEEE, 2012, pp. 4296–4301.

[6] X. Lin, R. W. Heath, and J. G. Andrews, "The interplay between massive MIMO and underlaid D2D networking," *IEEE Trans. Wireless Commun.*, vol. 14, no. 6, pp. 3337–3351, 2015.

[7] S. Shalmashi, E. Björnson, M. Kountouris, K. W. Sung, and M. Debbah, "Energy efficiency and sum rate tradeoffs for massive mimo systems with underlaid device-to-device communications," *arXiv preprint arXiv:1506.00598*, 2015.

[8] S. Shalmashi, E. Björnson, S. B. Slimane, and M. Debbah, "Closed-form optimality characterization of network-assisted device-to-device communications," in *2014 IEEE Wireless Communications and Networking Conference (WCNC)*. IEEE, 2014, pp. 508–513.

[9] J. Hoydis, K. Hosseini, S. t. Brink, and M. Debbah, "Making smart use of excess antennas: Massive MIMO, small cells, and TDD," *Bell Labs Technical Journal*, vol. 18, no. 2, pp. 5–21, 2013.

[10] J. C. Li, M. Lei, and F. Gao, "Device-to-device (D2D) communication in mu-mimo cellular networks," in *Global Communications Conference (GLOBECOM), 2012 IEEE*. IEEE, 2012, pp. 3583–3587.

[11] A. Afzal, S. A. R. Zaidi, D. McLernon, and M. Ghogho, "On the analysis of cellular networks with caching and coordinated device-to-device communication," in *IEEE International Conference on Communications (ICC)*, 2016, pp. 1–7.

[12] D. Moltchanov, "Distance distributions in random networks," *Ad Hoc Networks*, vol. 10, no. 6, pp. 1146–1166, 2012.