Analysis of Statistical QoS in Half Duplex and Full Duplex Dense Heterogeneous Cellular Networks

Alireza Sadeghi, Student Member, IEEE, Michele Luvisotto, Student Member, IEEE, Farshad Lahouti, Senior Member, IEEE, Michele Zorzi, Fellow, IEEE.

Abstract—Statistical QoS provisioning as an important performance metric in analyzing next generation mobile cellular network, aka 5G, is investigated. In this context, by quantifying the performance in terms of the effective capacity, we introduce a lower bound for the system performance that facilitates an efficient analysis. Based on the proposed lower bound, which is mainly built on a per resource block analysis, we build a basic mathematical framework to analyze effective capacity in an ultra dense heterogeneous cellular network. We use our proposed scalable approach to give insights about the possible enhancements of the statistical QoS experienced by the end users if heterogeneous cellular networks migrate from a conventional half duplex to an imperfect full duplex mode of operation. Numerical results and analysis are provided, where the network is modeled as a Matérn point process. The results demonstrate the accuracy and computational efficiency of the proposed scheme, especially in large scale wireless systems. Moreover, the minimum level of self interference cancellation for the full duplex system to start outperforming its half duplex counterpart is investigated.

Index Terms—Full Duplex, Heterogeneous Cellular Network, Statistical QoS, 5G.

I. INTRODUCTION

The ever increasing demand for mobile data traffic continues with the advent of smart phones, tablets, mobile routers, and cellular M2M devices. This is accompanied by user behavioral changes from web browsing towards video streaming, social networking, and online gaming with distinct quality of service (QoS) requirements [2]. To handle this evolution, researchers are examining different enabling technologies for 5G, including mmWave communications for wider bandwidth, extreme densification of the network via low power base stations, large-scale antenna systems (known as massive MIMO), and wireless full duplex communications [3].

The new cellular architecture known as Heterogeneous Cellular Networks (HCNs) refers to a scenario in which a macro cell is overlaid by heterogeneous low–power base stations (BSs). Such low power BSs have small coverage areas and are characterized by their own transmit power and named accordingly as micro, pico and femto cells. They are used to increase the capacity of the network while eliminating coverage holes [5]. In terms of modeling and analyzing HCNs, stochastic geometry has been vastly used in recent years [6]. For instance Poisson point processes (PPPs) have provided tractable, simple, and accurate expressions for coverage probabilities and mean transmission rates [6], [7], [8]. However, in a real cellular network, the adoption of a simple PPP to model the location of the BSs does not capture an important characteristic of a real network deployment. Spectrum access policies, network planning, and the MAC layer impose constraints on the minimum distance between any two BSs or UEs in the network that are operating in the same resource blocks (RBs) [6]. According to these considerations, a repulsive point process with a minimum distance such as the Matérn hard core point process (HCPP), despite its higher complexity, represents a better candidate to model a HCN compared to a simple PPP [6], [9].

The Full Duplex (FD) radio technology can enhance spectrum efficiency by enabling simultaneous transmission and reception in the same frequency band at the same time. This new emerging technology has the potential to double the physical layer capacity [4] and enhance the performance even more, when higher layer protocols are redesigned accordingly [10].

Due to the hurdles of canceling self-interference (SI) in FD devices via active and passive suppression mechanisms, FD operations are more reliable in low power wireless nodes. For instance in [11] and [12] the authors have implemented a FD WiFi radio operating in an unlicensed frequency band with 20 dBm transmit power while the same trend is followed in other works like [13], where the maximum transmit power is 15 dBm. All these implementations suggest FD technology as a reasonable candidate to be employed in the low power BSs deployed within HCNs. Moreover, the increased spectral efficiency of the FD systems, combined with that of HCNs, provides another serious motivation in attempting to analyze a completely FD HCN.

From another perspective, next generation mobile networks (5G) will aim not only to increase the network capacity but also to enhance several other performance metrics, including lower latency, seamless connectivity, and increased mobility [3]. These requirements can be generally regarded as improvements in the QoS experienced by the network entities. According to a forecast by Ericsson, approximately 55 percent of all the mobile data traffic in 2020 will account for mobile video traffic while another 15 percent will account for social networking [2]. These multimedia services require a bounded...
delay. The delay requirements of time sensitive services in 5G will greatly vary from milliseconds to a second [14]. Consequently, the analysis of statistical QoS in HCNs will become extremely important in the near future.

The objective of this paper is to analyze and compare FD and HD HCNs in provisioning statistical QoS guarantees to the users in the network. In this regard, we model the HCN as a Matérn HCPP, and statistically assess the QoS in terms of Effective Capacity (EC) or the maximum throughput under a delay constraint [15]. We provide insights on possible improvements in the QoS experience of end users if the current architecture migrates from conventional HD to FD. We propose a lower bound for the EC which greatly reduces the complexity of the analysis, while tightly approximating the system performance, especially in very large scale systems. The presented analysis is validated with simulation results.

The rest of the paper is organized as follows. A brief review of FD, statistical QoS provisioning, and stochastic geometry is provided in Section II. The system model is described in Section III. Section IV presents the proposed lower bound for the system performance and the corresponding theoretical analysis, whose results are validated through simulations in Section V. Some discussion on user distributions in the network, coverage areas of the small cells, and channel models is provided in Section VI. Finally, Section VII concludes the paper.

II. PRELIMINARIES

A. In-band Full Duplex Wireless Communications

In-band full-duplex (IBFD) devices are capable of transmitting and receiving data in the same frequency band at the same time. In traditional wireless terminals, the self interference (SI) power is much larger than that of the received intended signal, making any reception infeasible while a transmission is ongoing. To overcome this issue, FD terminals are equipped with active and passive cancellation mechanisms to suppress their own SI in the received signal [4]. However, in practice, because of the many imperfections in transceiver characteristics and operations, namely phase noise, nonlinearities, I/Q imbalances, receiver quantization noise and constraints on the accuracy of the SI channel estimation, full cancellation of the SI signal is not possible. Therefore, some residual self-interference (RSI) always remains after all cancellation steps, and results in a degraded system performance [4], [11], [12], [13].

The RSI signal represents the main obstacle for a perfect FD communication and, similar to noise, is essentially uncorrelated with the original transmitted signal. Among different methods to model RSI, we consider RSI as a zero mean complex Gaussian random variable whose variance is a function of the transmitted power [16], [17]. We have

\[ \text{RSI} \sim \mathcal{CN}(0, \eta P^\kappa) \]  

where \( P \) is the transmit power, and \( \eta \) and \( \kappa \), \( 0 \leq \eta, \kappa \leq 1 \), are respectively the linear and the nonlinear SI cancellation parameters capturing the SI cancellation performance. When no SI cancellation is performed \( \eta, \kappa = 1 \), while \( \eta = 0 \) represents the ideal case of perfect SI cancellation.

B. Statistical QoS guarantees

Real time multimedia services like video streaming require bounded delays. In this context, a received packet that violates its delay bound requirement is considered useless and is discarded. Due to the wireless nature of the access links in a mobile cellular network, providing deterministic delay bound guarantees is impossible. Thus, effective capacity defined as the maximum throughput under a given delay constraint, has been used to analyze multimedia wireless systems [15]. Any active entity in a cellular network can be regarded as a queueing system: it generates packets according to an arrival process, stores them in a queue and transmits them according to a service process. For stationary and ergodic arrival and service processes, the queue length process \( Q(t) \) converges, in distribution, to a random variable \( Q(\infty) \) as follows

\[ - \lim_{Q_{th} \to \infty} \frac{\log(\Pr\{Q(\infty) > Q_{th}\})}{Q_{th}} = \theta, \]  

where \( Q_{th} \) is the queue length threshold and \( \theta \) is a constant value called QoS exponent. In fact (2) states that the probability that the queue length process violates this threshold decays exponentially fast as the threshold increases, i.e., \( \Pr\{Q(t) > Q_{th}\} \sim e^{-\theta Q_{th}} \) as \( Q_{th} \to \infty \) [18]. In addition, it was shown in [15] that for small values of \( Q_{th} \), the violation probability can be approximated more accurately as \( \Pr\{Q(t) > Q_{th}\} \approx \delta e^{-\theta Q_{th}} \), where \( \delta \) represents the probability that the buffer is not empty, and the QoS exponent, \( \theta \), characterizes the decaying rate of the QoS violation probability. A large value for \( \theta \) corresponds to a faster decaying rate, which means that the system has stringent statistical delay constraints. For instance, when \( \theta \to \infty \) the system is very sensitive to delay and generally no delay is tolerable. On the other hand, for small values of \( \theta \) the system can tolerate the delays and, when \( \theta \to 0 \) there is no sensitivity at all to delays in the system, and in this case the EC tends to the Shannon capacity [19].

Let us define the service provided by the channel until time slot \( t \) as

\[ C(0, t) = \sum_{k=1}^{t} R[k] \]  

where \( R[k] \) denotes the bits served in time slot \( k \). The effective capacity of the channel is defined as [15]

\[ EC(\theta) = -\frac{\Lambda_C(-\theta)}{\theta} \]  

where \( \Lambda_C(-\theta) = \lim_{t \to \infty} \frac{1}{t} \log E\{e^{-\theta C(0, t)}\} \) is the Gärtner - Ellis limit of the service process \( C(0, t) \). If the instantaneous service process, \( R[k] \), is independent over time, EC can be simplified to

\[ EC(\theta) = -\frac{1}{\theta} \log E\{e^{-\theta R[k]}\}. \]  

EC is defined as the maximum arrival rate a given service process can support in order to guarantee a QoS requirement specified by \( \theta \) [19]. We note that EC is defined for the service process and not for the arrival process. However, the
dual definition of EC for the arrival process, referred to as the effective bandwidth, serves this purpose. It indicates the minimum service process required by a given arrival process for which the QoS exponent $\theta$ is satisfied \cite{19}.

### III. System Model

We consider a HCN consisting of a macro cell, overlaid by different tiers of small cells, each with its own characteristics including transmit power, path loss exponent, and coverage range. Each tier is assumed to have a circular coverage area provided by an omni-directional antenna to serve any user within its coverage range. The small cells are assumed to use out of band resources like fiber optics, wire, or microwave links for backhauling. Fig. 1 depicts an example of such a HCN.

In this notation, the numerator represents the desired signal power received from a BS in the $i^{th}$ tier which serves the UE. Here, a tier is defined as the set of BSs with a given density and the same characteristics including average transmit power, supported data rate, and coverage area \cite{7}. The first and second terms in the denominator represent the interference from other BSs and UEs in the network operating in the same RBs as the desired UE. Specifically, $\Phi_k^{BS}$ and $\Phi_k^{UE}$ indicate sets containing the positions of all interfering BSs and UEs in the $k^{th}$ tier, and the summation is over all possible tiers. The third term is the RSI signal power as modeled in (1). Finally, $\sigma^2$ is the additive noise power.

This paper analyzes the system performance on a resource block (RB) basis. An illustration of a single RB in LTE-A is presented in Fig. 2. The reason for following a RB analysis is three-fold: i) the service rate, $R[k]$, can be replaced by the instantaneous capacity of the channel. This is due to the fact that the service time, $T_f$, is less than the coherence time of the channel, but still sufficiently long to support the information theoretic assumption of adaptive modulation and capacity achieving coding; ii) the service rate, $R[k]$, becomes statistically independent from one RB to another, as was assumed in deriving \cite{5}; iii) most importantly, the RB based analysis will assist to find a very tight, computationally efficient and scalable lower bound for analyzing EC in the meaningful range of values for the QoS exponent. Consequently, in this work, the unit for both the service rate, $R[k]$, and the effective capacity, $EC(\theta)$, is defined as bits per RB.

Fig. 1. System model: an HCN with one macrocell and several low-power BSs.

Fig. 2. Illustration of one resource block in LTE-A.
We recall that the number of bits delivered to a UE during an interval, $T_f$, in a given bandwidth, $BW$, if capacity achieving modulation and coding are used, can be represented as

$$R = T_f BW \log_2 (1 + \text{SINR}).$$

Therefore, the effective capacity of the desired UE based on (5) can be expressed as

$$EC(\theta) = -\frac{1}{\theta} \log (\mathbb{E}\{\exp (-\theta T_f BW \log_2 (1 + \text{SINR}))\}) = -\frac{1}{\theta} \log \left( \mathbb{E} \left( (1 + \text{SINR})^{-\theta T_f BW \log_2 e} \right) \right)$$

where the expectation is taken with respect to SINR.

In a HD scenario, a 1/2 scaling factor is needed and, also, the SINR would become

$$\text{SINR}_{\text{HD}} = \sum_k \sum_{x \in \Phi_{\text{BS}}} P_k h_{x,i} \|x_i\|^{-\alpha_i} + \sigma^2$$

where the numerator represents the desired signal power received from a BS in the $i^{th}$ tier, which serves the UE and the denominator includes noise and DL interference from interfering BSs in the $k^{th}$ tier and the summation is over all tiers.

**IV. Theoretical Analysis**

We aim at computing the QoS experienced by a generic UE, that can be placed anywhere in the coverage area of its own small cell with uniform distribution. To find the exact EC in a given topology, one needs to compute (8) either through extensive simulations or by mathematical analysis. It is worth mentioning that, if there are $M$ small cells within the macro cell communicating with $M$ FD UEs, the associated integrals (or simulation setups) would be in a $2M + 1$ dimensional parameter space. To understand why this is the case just note that the expectation in (8) is with respect to SINR which is a parameter space. To understand why this is the case just note that the average interference on the desired network entity ($\bar{I}$) is known. Therefore, the proposed bound makes this calculation scalable with the size of the network, at the possible cost of losing some precision.

In order to efficiently compute the proposed lower bound in (13), i.e., $EC_{\text{LB}}(\theta)$, one has to calculate analytically the average interference on the desired UE, $\bar{I}$. Recalling the expression of the interference from Section III, the average power from an interferer located at distance $x$ from the considered UE can be found as

$$\mathbb{E}[P h \|x\|^{-\alpha}] \stackrel{(a)}{=} P \mathbb{E}(h) \mathbb{E}[\|x\|^{-\alpha}] \stackrel{(b)}{=} P \mathbb{E}[\|x\|^{-\alpha}]$$

where (a) follows from the fact that the channel coefficient and distance between the interferer and the desired UE are independent random variables and (b) holds because we assumed the Rayleigh fading channel, $h$, has unit mean. Consequently, all our efforts will be focused on finding the average path loss from the desired UE to the interferers.

It is noteworthy that in deriving the lower bound proposed in (13) we implicitly have a constraint on the QoS exponent, $\theta$, through Lemma 1, where we assumed the constraint on $\beta$ as $0 \leq \beta \leq 1$. This imposes a constraint on $\theta$, as $0 \leq \theta \leq \frac{1}{T_f BW \log_2 e}$. While in general, the QoS exponent is defined in the range $[0, \infty)$. Due to the per RB based analysis this constraint on $\theta$ would amount to

$$0 \leq \theta \leq \frac{1}{T_f BW \log_2 e} \approx 10^{-2}.$$  \hspace{1cm} (15)

Fortunately, the range of $\theta$ in (15) lies in the meaningful range of the QoS exponent [19]. However, simulations show that our method gives a good approximation for EC even in a wider range.

**Proof.** By assuming $s$ as a constant, taking the second derivative of $g$ with respect to $I$ leads to

$$\frac{\partial^2 g(s, I)}{\partial I^2} = \frac{\beta s}{(I + a)^2} \left( \frac{1 + s}{I + a} \right)^{-(\beta + 2)} (-2(I + a) + (\beta - 1)s)$$

which is negative, for all values of $0 \leq \beta < 1 + \frac{2}{\text{SINR}}$, where SINR $\geq \frac{1}{T_f e}$. This readily shows that $g$ is a concave function of $I$ for any $0 \leq \beta \leq 1$.

The concavity of $g$ helps us find a tight lower bound for the EC with greatly decreased complexity. To this end, by exploiting Jensen’s inequality, we obtain

$$EC_{\text{LB}}(\theta) = -\frac{1}{\theta} \log \mathbb{E}_s \left( \left( \frac{1 + s}{I + a} \right)^{-\beta} \right) \leq EC(\theta) \hspace{1cm} (13)$$

Here, $I = \mathbb{E}(I)$ is the average interference experienced by the UE, and the remaining expectation only applies to the desired signal power. The advantage of this lower bound is its much smaller computational complexity. Indeed, calculating this lower bound only requires a 1-dimensional integral (or simulation) with respect to the desired signal power provided that the average interference on the desired network entity ($I$) is known. Therefore, the proposed bound makes this calculation scalable with the size of the network, at the possible cost of losing some precision.

A. Approximating EC

Let us define a generic function $g$ of $s, I, a$, and $\beta$ as follows

$$g(s, I) \triangleq \left( 1 + \frac{s}{I + a} \right)^{-\beta}.$$  \hspace{1cm} (10)

This function has the same structure of the expectation argument in (8), where $s$ models the received signal power, $I$ represents the overall interference from other BSs and UEs in the network, $a$ represents the RSI and noise, and $\beta = \theta \cdot T_f \cdot BW \cdot \log_2 e$. Based on this definition, the EC can be rewritten as

$$EC(\theta) = -\frac{1}{\theta} \log (\mathbb{E}_{s,I} g(s, I)) \hspace{1cm} (11)$$

**Lemma 1.** For $0 \leq \beta \leq 1$, $g$ is always a concave function of $I$. 

- **Proof.** By assuming $s$ as a constant, taking the second derivative of $g$ with respect to $I$ leads to

$$\frac{\partial^2 g(s, I)}{\partial I^2} = \frac{\beta s}{(I + a)^2} \left( \frac{1 + s}{I + a} \right)^{-(\beta + 2)} (-2(I + a) + (\beta - 1)s)$$

which is negative, for all values of $0 \leq \beta < 1 + \frac{2}{\text{SINR}}$, where SINR $\geq \frac{1}{T_f e}$. This readily shows that $g$ is a concave function of $I$ for any $0 \leq \beta \leq 1$. 

The concavity of $g$ helps us find a tight lower bound for the EC with greatly decreased complexity. To this end, by exploiting Jensen’s inequality, we obtain

$$EC_{\text{LB}}(\theta) = -\frac{1}{\theta} \log \mathbb{E}_s \left( \left( \frac{1 + s}{I + a} \right)^{-\beta} \right) \leq EC(\theta) \hspace{1cm} (13)$$

Here, $I = \mathbb{E}(I)$ is the average interference experienced by the UE, and the remaining expectation only applies to the desired signal power. The advantage of this lower bound is its much smaller computational complexity. Indeed, calculating this lower bound only requires a 1-dimensional integral (or simulation) with respect to the desired signal power provided that the average interference on the desired network entity ($I$) is known. Therefore, the proposed bound makes this calculation scalable with the size of the network, at the possible cost of losing some precision.

In order to efficiently compute the proposed lower bound in (13), i.e., $EC_{\text{LB}}(\theta)$, one has to calculate analytically the average interference on the desired UE, $\bar{I}$. Recalling the expression of the interference from Section III, the average power from an interferer located at distance $x$ from the considered UE can be found as

$$\mathbb{E}[P h \|x\|^{-\alpha}] \stackrel{(a)}{=} P \mathbb{E}(h) \mathbb{E}[\|x\|^{-\alpha}] \stackrel{(b)}{=} P \mathbb{E}[\|x\|^{-\alpha}] \hspace{1cm} (14)$$

where (a) follows from the fact that the channel coefficient and distance between the interferer and the desired UE are independent random variables and (b) holds because we assumed the Rayleigh fading channel, $h$, has unit mean. Consequently, all our efforts will be focused on finding the average path loss from the desired UE to the interferers.

It is noteworthy that in deriving the lower bound proposed in (13) we implicitly have a constraint on the QoS exponent, $\theta$, through Lemma 1, where we assumed the constraint on $\beta$ as $0 \leq \beta \leq 1$. This imposes a constraint on $\theta$, as $0 \leq \theta \leq \frac{1}{T_f BW \log_2 e}$, while in general, the QoS exponent is defined in the range $[0, \infty)$. Due to the per RB based analysis this constraint on $\theta$ would amount to

$$0 \leq \theta \leq \frac{1}{T_f BW \log_2 e} \approx 10^{-2}.$$  \hspace{1cm} (15)

Fortunately, the range of $\theta$ in (15) lies in the meaningful range of the QoS exponent [19]. However, simulations show that our method gives a good approximation for EC even in a wider range.
position for the interferer, i.e., fixed \((r_1, \theta_1)\). Subsequently, we compute the expectation of the resulting quantity with respect to all possible values of \((r_1, \theta_1)\).

Regarding the first step, since we assumed \((r_1, \theta_1)\) is fixed, \(c\) and \(\psi\) become constants, which facilitates the analysis

\[
\mathbb{E}_{(r_2, \theta_2)} \left[ \left( c^2 + r_2^2 + 2cr_2 \cos (\theta_2 + \psi) \right)^{-\frac{\alpha}{2}} \right] \left| (r_1, \theta_1) \right| 
\]

\[
= \int_0^{2\pi} \int_0^{2\pi} \left( c^2 + r_2^2 + 2cr_2 \cos (\theta_2 + \psi) \right)^{-\frac{\alpha}{2}} \frac{1}{\pi R_2^2} dr_2 d\theta_2 
\]

\[
= \frac{2\pi R_2}{\alpha} \int_0^{2\pi} \left( 1 + \left( \frac{r_2}{c} \right)^2 \right) + 2 \left( \frac{r_2}{c} \cos (\theta_2 + \psi) \right) \frac{1}{\pi R_2^2} dr_2 d\theta_2 
\]

\[
\leq \frac{\alpha (\alpha + 2)}{8} \left[ \left( \frac{r_2}{c} \right)^2 + 2 \left( \frac{r_2}{c} \cos (\theta_2 + \psi) \right) \right] \frac{1}{\pi R_2^2} dr_2 d\theta_2 
\]

where in (a) we used the first three terms of the Taylor series expansion of \((1 + x)^{-\alpha} = 1 - \alpha x + \frac{\alpha (\alpha + 1)}{2} x^2 + \ldots\). The Taylor approximation is legitimate if \(x < 1\) (i.e., \(c > d\)) which is already satisfied considering the repulsive point process we have assumed for the small cells, characterized by the hard core distance \(r_h \geq R_1 + R_2\).

We recall that this result holds for any fixed values of \((r_1, \theta_1)\). In particular, by setting \(r_1 \rightarrow 0\) (i.e., \(c \rightarrow d\) in (25)), we obtain the average path loss component between a randomly deployed UE and the BS. Therefore, the average interference that an external BS causes to the considered UE uniformly placed in any point within the coverage area of its small cell would become

\[
I_{BS-UE} = P_{BS} d^{-\alpha} \left[ 1 + \frac{\alpha^2}{8} \left( \frac{R_1^2}{3c^3} + \frac{R_2^2}{c^2} \right) + \frac{\alpha R_1^2}{4 3d^3} \right]. 
\]
where $P_{UE}$ is the transmit power of the interferer UE, and $c$ is the distance between the interferer and the BS which serves the UE. The expectation terms are with respect to the location of the interferer, i.e., $(r_1, \theta_1)$.

**Proof.** We compute the expectation of (25) with respect to the position of the interferer, i.e., $(r_1, \theta_1)$, which leads to (27).

It must be noted that to compute the quantity in (27), one needs to calculate only $E_{r_1, \theta_1}[c^{-\alpha}]$, since the other two expectations can be immediately obtained by replacing $\alpha$ with $\alpha + 2$ and $\alpha + 4$. We further observe that this expectation corresponds to the average path loss component between the interferer and the desired UE’s BS. This quantity can be derived from (26) by setting $P_{BS} = 1$ and changing $R_2$ to $R_1$.

The proposed relations in (26) and (27) can hence be used as a basic mathematical tool to investigate the system performance.

### C. Average Interference from a free UE

We should note that the only source of interference on a desired UE that does not follow the structure provided in Fig. 3 is the UE connected to the macro BS, that we refer to as free UE. We make here a simplifying assumption in order to derive an approximate expression for the interference caused by the free UE to the desired UE, of which we are computing the effective capacity. Specifically, let $R_1'$ be the coverage radius of the macro BS, $R_2'$ the coverage radius of the small cell where the desired UE is placed, $d'$ the distance between the macro BS and the BS of the small cell and $\theta^*$ the angle of the circular sector centered in the macro BS which fully encloses the coverage area of the small cell, as represented in Fig. 4. We assume that the free UE can be deployed anywhere in the coverage area of the macro BS except in the sector with angle $2\theta^*$, i.e., the shaded area in Fig. 4. Thus, the PDF of the free UE in the polar coordinates of the macro cell would be

$$f(r_1', \theta_1') = \begin{cases} \frac{r_1'}{(\pi-\theta^*)R_1^2} & \theta^* < \theta_1' < 2\pi - \theta^* \\ 0 & -\theta^* < \theta_1' < \theta^* \end{cases} \tag{28}$$

The next proposition presents the average interference from a free UE in this setting.

**Proposition 3.** The approximate average interference from the free UE on a desired UE can be found by calculating

$$I_{free \cdot UE \cdot UE} \approx P_{UE} \cdot \frac{(\alpha + 2)(\alpha + 3)}{6} + \frac{2}{\alpha - 2} \cdot \frac{1}{R_1'^2} \mathbb{E}\left(c^{-(\alpha - 2)}\right) - \frac{\alpha (3\alpha^3 + 11\alpha^2 - 18\alpha - 56) \sin \theta^*}{15(\pi - \theta^*) (\alpha^2 - 1)} \cdot \frac{1}{R_1'^2} \mathbb{E}\left(c^{-(\alpha - 2) \cos \psi'}\right) - \frac{\alpha (\alpha + 2)(\alpha + 4) \sin 2\theta^*}{16(\pi - \theta^*)} \cdot \frac{1}{R_1'^2} \mathbb{E}\left(c^{-(\alpha - 2) \cos \psi'}\right) \tag{29}$$

where $c'$ is the distance between the macro BS to the desired UE which may deployed in any location of the macro cell according to (28). Moreover $\mathbb{E}\left(c^{-(\alpha - 2) \cos \psi'}\right)$, and

![Fig. 4. Structure of interference caused by a free UE on the desired UE. The shaded sector is the area where we assume that the free UE can not be deployed.](image)

It is worth noting that in the simulations, we let the free UE be in any location of the macro cell except the small cell under investigation. We will then draw comparisons between simulations and the analysis.

### V. Simulations and Results

This section provides numerical simulations to analyze the performance of the FD and HD HCNs under statistical QoS constraints. In addition, results are reported to assess the scalability of our proposed approach to analyze system performance. Although a single macro cell overlaid with randomly placed small cells is considered, the analytic framework proposed in this paper readily allows the study of a scenario with multiple macro cells. In the reported results, the UE in the dashed small cell is under consideration. The exact curves are obtained via Monte-Carlo simulation. A lower bound is obtained using (13) and the average interference on the desired UE, $I$, calculated through Propositions 1, 2. During the simulations the system for the HD scenario is considered to be FDD. See Tab. 1 for the parameters adopted to carry out the simulations.

In the following, Section V-A provides results on HD and FD HCNs and studies the impact of FD technology on the QoS experienced by the end users. In Section V-B the accuracy and complexity of the proposed lower bound are investigated. Finally in Section V-C the crossover point under different QoS requirements is investigated, and the minimum level of SI
A. Performance evaluation of HD and FD HCNs

Fig. 5 reports the DL EC of a perfect FD system compared with that of a HD network for different QoS exponents. As it can be seen, a 1.95X gain from a perfect FD compared to HD in terms of EC is obtained. This is due to the fact that in completely FD systems there would be more interference in the network compared to a completely HD one since UL and DL transmissions may occur in the same RBs. It is worth noting again that when \( \theta \rightarrow 0 \) there is no sensitivity at all to delays in the system, and in this case the EC tends to the Shannon capacity.

Figs. 6 and 7 show DL EC versus linear SI cancellation parameter, \( \eta \), under sparse and dense HCN scenarios, respectively. As it can be seen, depending on the level of SI cancellation, the system performance under FD scenario exhibits various performance results. For instance, for low SI cancellation parameter, i.e., \( \eta \rightarrow 0 \) dB, the system can not provide any service to the UE. However, as the system goes towards perfect FD, i.e., \( \eta \rightarrow -\infty \) dB, the system...
provides approximately 2X service to the UE compared to the HD counterpart. In addition, there is a minimum linear SI cancellation required so that FD operation mode outperform HD one. We call this crossover point and as it can bee seen, it is approximately $-50$ dB in this case. Detailed discussions on this crossover point and its dependence on the system parameters are presented in Sec. V-C.

Moreover, in Fig. 6.b the EC for the UE is approximately $1.67X$ that observed in Fig. 7.b for the same level of SI cancellation. This is mainly due to the lower level of interference in the network. However, the area spectral efficiency is much higher in the HCN of Fig. 7. In fact, as the network becomes denser, a larger number of UEs can be served simultaneously in a single RB. Thus, despite the decreased service rate to each UE in a dense HCN, a very high area spectral efficiency can be expected.

B. Assessment of the Lower Bound: Accuracy and Complexity

As Figs. 6 and 7 demonstrate, the proposed lower bound is tight, and interestingly as the system becomes denser, the lower bound becomes tighter. Furthermore, according to Figs. 6 and 7 increasing the SI parameter, $\eta$, the gap between the lower bound and the exact value of EC decreases. This is due to the fact that as the SI parameter increases, the RSI dominates other sources of interference on the desired UE and since its power is constant as $\eta P_X$, the interference tends to a constant and Jensen’s inequality (invoked in (13)) becomes an equality.

Fig. 5.b shows the impact of QoS exponent on the accuracy of the lower bound. As it is shown, under loose QoS requirements, i.e., $\theta \rightarrow 0$, the lower bound becomes looser, however, still it is tight enough to give insight on system performance. On the other hand, for stringent QoS requirements the lower bound becomes very tight. This figure demonstrates that in a very broad range of QoS exponent the proposed approach gives accurate results on the EC that a cellular network provides in DL for a UE.

Fig. 8.a depicts the relative error of the proposed lower bound versus the number of small cells. We considered the number of small cells from the set $\{1, 2, 4, 8, 16, 22, 37, 50\}$, and averaged the relative error of our scheme in 100 different realizations of the HCN deployment for each case. As shown in this figure, by increasing the number of interferes in the network, the lower bound becomes tighter.

As discussed earlier in Section IV, the complexity of the proposed approach for analyzing the system performance is almost independent of the size of the network and significantly smaller than that of the exact analysis. To be more concrete, an exact analysis has a complexity in the order of $O(N^M)$, where $N$ is the number of function evaluations in each dimension of the integration, and $M$ is the number of small cells in the network. In comparison, the computational complexity of the proposed lower bound is $O(N + M)$, where the first term of the inner argument corresponds to integration in the desired signal’s dimension and the second term corresponds to calculating the average interference. To assess this claim for the same realizations of the network considered in Fig. 8.a, the average time elapsed to get the results is provided in Fig. 8.b as a metric of complexity. As the number of interferes in the network grows, the complexity of the exact analysis increases dramatically, while the proposed lower bound has complexity almost independent of the network size.

Given the complexity and accuracy issues discussed above, it can readily be claimed that the approach proposed in this paper is scalable. That is, the denser the network, the tighter the approximation, without any significant additional complexity.

C. Crossover Point

The SI cancellation techniques play an essential role in enabling a wireless device to operate in FD mode; see e.g., [4], [11], [12], [20], [21], [22]. Despite the abundance of many different proposed RF, active, and passive interference

2All the simulations were carried out with an Intel Core i5-2.53GHz processor and 4G RAM on a Dell Inspiron 5010.
cancellation techniques with a variety of performance results, establishing any FD wireless link requires carefully designed SI cancellation techniques specific to the required application. In this regards, a system level study helps gaining insight on the required level of SI cancellation and thus on finding the appropriate SI cancellation techniques. To this aim, in the following we investigate the crossover point, $\eta^*$, i.e., the minimum SI cancellation ($\eta$) required for the FD system to start outperforming the HD counterpart. This would require solving the equation:

$$EC_{HD}(\theta) = -\frac{1}{\theta} \log \left[ E_{s,f} (1 + \text{SINR}_{\text{HD}})^{-\theta T_f BW \log_2 e} \right] = -\frac{1}{\theta} \log \left[ E_{s,f} (1 + \text{SINR}_{\text{FD}})^{-\theta T_f BW \log_2 e} \right] = EC_{FD}(\theta)$$

(30)

where SINR$_{\text{HD}}$ and SINR$_{\text{FD}}$ are provided in (9) and (6), respectively. Eq. (30) is a complicated non-linear function of the parameter $\eta$, the topology of the network, the QoS constraint, $\theta$, and the transmit powers in the network, which makes finding any closed form expression for the trade off point not possible. However, an estimate of such a point can be found numerically.

As an example, we plotted this trade off point in Fig. 9 for a random realization of the HCN. This figure indicates the relation between the crossover point, $\eta^*$, and the QoS exponent, $\theta$. Not intuitively, this plot indicates that the more stringent the QoS constraint, the higher the SI cancellation level required in order to make FD preferable to HD. In other words, if the QoS constraint is higher, the impact is profound on a FD system compared to its HD counterpart.

Finally, Figs. 10 and 11 provide the results for the DL EC of a typical UE as a function of the transmit power of the small cell BSs. As these figures show, the transmit power for which the FD starts outperforming the HD counterpart is a function of the topology of the network and, more importantly, of the SI cancellation performance, $\eta$, $\kappa$. For instance, in Fig. 10 this occurs at 35 dBm, while for higher SI cancellation parameters, e.g., $\eta = -100$ dB, FD would outperform HD in a wider transmit power ranges as shown in Fig. 11.

VI. DISCUSSIONS

A. UE Distribution, Coverage Areas, and Channel Models

The uniform distribution of the UEs around their corresponding BSs, which is a result of the Matérn HCPP model for the system, can be relaxed. This is due to the fact that the result given in (13) is true for any distribution of the UEs and BSs in the network. In fact, we can expect to find similar results for other point processes that can be used to model the nodes. For instance, we can expect the same procedure to be applied also for Thomas point processes, where child points (UEs) are distributed with an isotropic Gaussian distribution with a given variance around the cluster heads (BSs) [6]. In fact, for any generic distribution we only need to build a computationally efficient procedure to find the average interference on the
desired network entity (UE or BS), as we have done in Section IV-B.

Finally, we assumed a pathloss dominated AWGN Rayleigh fading channel model, but, since the lower bound presented in [13] is always valid for any distribution of the interferer’s channel, it is straightforward to extend this work to other channel models, e.g., including log-normal shadowing.

B. Overlapping Small Cells

The analysis carried out in Section IV relies on the assumption that two small cells deployed in the same macro cell do not overlap. However, in very dense network deployments, small cells are not perfectly separated and may overlap in some cases, thus impairing the effectiveness of the analysis proposed in this paper.

In detail, when two small cells overlap, as depicted in Fig. 12 there can be two users deployed in the overlapping regions which are served in the same RB, since we did not make any assumption on the resource scheduling strategy. In this case, the interference caused by the neighboring UE to the free UE (whose average value is expressed by Proposition (2)) is very high, since the distance between the two UEs is small. This is particularly relevant if both UEs are FD and, hence, both are transmitting and receiving at the same time. As a consequence, the term \( E_{f} (I) = I \) may become very high in this scenario and, consequently, the lower bound proposed in [13] is much less tight.

Fortunately, in a real network scenario, the possibility that two UEs very close to each other are served in the same RB is practically very small, and can be neglected. Indeed, exploiting cell coordination and channel measurements, two close UEs deployed in different but overlapping cells will be scheduled in different RBs. Consequently, the theoretical approach proposed in this paper is of practical interest.

VII. CONCLUSIONS

In this paper we introduced a lower bound for the evaluation of EC in a wireless cellular scenario based on which we built a scalable mathematical framework to analyze the statistical QoS performance of ultra dense next generation HCNs. Our proposed scalable scheme helped us analyzing HD and imperfect FD HCNs from the EC perspective with high accuracy at only a fraction of the complexity needed for an exact analysis. Moreover, we investigated the trade off between FD and HD systems by finding the minimum amount of SI cancellation needed for FD HCN to perform similar to HD in terms of EC.

APPENDIX A

The average path loss between the free UE and the desired UE based on Fig. 4 can be found by computing

\[
E \left( \| x' \|^{-\alpha} \right) = E \left( r_1'^2 + c'^2 - 2c' r_1' \cos (\theta_1' - \psi) \right)^{-\frac{3}{2}} \quad (31)
\]

where

\[
c'^2 = d^2 + r_2'^2 + 2r_2' d' \cos \theta_2' \quad (32)
\]

To approximately calculate (31) this time we assume that \( (r_2', \theta_2') \), i.e., the position of the desired UE, is fixed, and
then take the expectation with respect to the position of the free UE and finally with respect to the desired UE.

\[
E(r'_i, \theta'_i) \left[ \left( r'_1^2 + c'^2 - 2c'r'_1 \cos (\theta'_1 - \psi') \right)^{-\frac{\alpha}{2}} \right]_{(r'_2, \theta'_2)}
\]

\[
R'_1 2\pi \theta^\star \int_{0}^{\theta^\star} \int_{0}^{r'_1} \left( r'_1^2 + c'^2 - 2c'r'_1 \cos (\theta'_1 - \psi') \right)^{-\frac{\alpha}{2}} r'_1 dr'_1 d\theta'_1
\]

\[
\approx \frac{c'^{-\alpha}}{R'_1 2\pi \theta^\star} \left( 1 + \frac{\alpha}{4} + \frac{2\alpha}{3(\pi - \theta^*)} \sin \theta^* \cos \psi' \right) + \frac{\alpha (\alpha + 2)}{8}
\]

\[
\times \left( \frac{4}{3} + \frac{1}{(\pi - \theta^*)} \left( \frac{8}{5} \sin \theta^* \cos \psi' - \frac{1}{2} \sin 2\theta^* \cos 2\psi' \right) \right)
\]  

(34)

In the second integral, again by using the same approach as in (34), we obtain

\[
I_2 = \frac{1}{R'_1 2\pi \theta^*} \times
\]

\[
\int_{c'}^{\infty} \int_{\theta^*}^{\theta'_1} \left( 1 + \frac{c'^2}{\theta'_1} \right)^{-\frac{\alpha}{2}} \left( \frac{r'_1}{R'_1} \right)^2 \left( c'^{-\alpha - 1} \right) \left( 1 + \frac{c'^2}{\theta'_1} \right)^{\frac{\alpha}{2}} \cos \left( \theta'_1 - \psi' \right) d\theta'_1 dr'_1
\]

\[
\approx \frac{1}{R'_1 2\pi \theta^*} \left\{ \frac{2}{\alpha - 2} \left( c'^{-\alpha - 1} - R'_1 c'^{-\alpha} \right) - \frac{2\alpha}{\alpha - 2} \left( c'^{-\alpha} \right) \left( c'^{-\alpha - 1} - R'_1 c'^{-\alpha} \right) \sin \theta^* \cos \psi' \right\}
\]

\[
+ \frac{\alpha (\alpha + 2)}{8} \left[ \frac{2c'^4}{\alpha + 2} \left( c'^{-\alpha} - R'_1 c'^{-\alpha} \right) + \frac{4c'^2}{\alpha} \cos \left( \theta'_1 - \psi' \right) \right]
\]

\[
+ \frac{8c'^3}{(\alpha + 1)(\pi - \theta^*)} \left( c'^{-\alpha} - R'_1 c'^{-\alpha} \right) \sin 2\theta^* \cos 2\psi'
\]  

(35)

We can simplify (35) by recalling that \( R'_1 \) is the radius of the coverage area of the macro BS, thus we have \( R'_1 \ll c'^{-a} \), \forall t > 0. By taking this into account and using (34) and (35), we obtain the approximation for the average path loss from a free UE to the desired UE as

\[
I_{\text{UE - UE}} \approx P_{\text{UE}} \left[ \left( \frac{(\alpha + 2)}{(\alpha + 3)} + \frac{2}{\alpha - 2} \right) \frac{1}{R'_1 2\pi} E \left( c'^{-\alpha - 2} \right)
\]

\[
- \frac{\alpha (3\alpha^3 + 11\alpha^2 - 18\alpha - 56)}{15(\pi - \theta^*) (\alpha^2 - 1)} \frac{1}{R'_1 2\pi} E \left( c'^{-\alpha - 2} \cos \psi' \right)
\]

\[
- \frac{(\alpha + 2)(\alpha + 4)}{16(\pi - \theta^*)} \frac{1}{R'_1 2\pi} E \left( c'^{-\alpha - 2} \cos 2\psi' \right)
\]

(36)

where \( E \left( c'^{-\alpha - 2} \right) \) is already provided in (25) and \( E \left( c'^{-\alpha - 2} \cos \psi' \right) \) and \( E \left( c'^{-\alpha - 2} \cos 2\psi' \right) \) can be calculated as follows

\[
E \left( c'^{-\alpha - 2} \cos \psi' \right) = \frac{d' E \left( c'^{-\alpha - 1} \right)}{d' E \left( c'^{-\alpha - 1} \right) + E \left( c'^{-\alpha - 1} \sin \psi' \right)} \approx \frac{d' E \left( c'^{-\alpha - 1} \right)}{d' E \left( c'^{-\alpha - 1} \right) + E \left( c'^{-\alpha - 1} \sin \psi' \right)} \approx \frac{d' E \left( c'^{-\alpha - 1} \right)}{d' E \left( c'^{-\alpha - 1} \right) + E \left( c'^{-\alpha - 1} \sin \psi' \right)}
\]

(37)
and
\[
\begin{align*}
\mathbb{E}\left(e^{-\alpha z_2} \cos^2 \psi_1\right) &= \\
\mathbb{E}\left(e^{-\alpha z_2} \left(1 - 2 \left(\frac{r_2' \sin \theta_2'}{c'}\right)^2\right)^2\right) &\approx \\
\mathbb{E}\left(e^{-\alpha z_2} - \frac{d^{\alpha - 2} R_2^2}{2} + \frac{\alpha (4\alpha + 7)}{6} \frac{d^{\alpha - 2} R_2^4}{2}ight) \\
&+ \frac{\alpha (2 + d^{\alpha - 4} R_2^6)}{32} \quad (38)
\end{align*}
\]

In deriving (37) and (38) we assumed \( R_2^2 < d' \).

REFERENCES

[1] A. Sabeghi, M. Luvisotto, F. Lahouti, S. Vitturi, and M. Zorzi, “Statistical QoS analysis of full duplex and half duplex heterogeneous cellular networks,” in 2016 IEEE International Conference on Communications (ICC), Kuala Lumpur Malaysia, May 2016.

[2] Ericsson Press Release, “Ericsson mobility report,” February 2015.

[3] J. Andrews, S. Buzzi, W. Choi, S. Hanly, A. Lozano, A. Soong, and J. Zhang, “What will 5G be?” IEEE Journal on Selected Areas in Communications, vol. 32, no. 6, pp. 1065–1082, June 2014.

[4] A. Sabharwal, P. Schniter, D. Guo, D. Bliss, S. Rangarajan, and R. Wichman, “In-band full-duplex wireless: Challenges and opportunities,” IEEE Journal on Selected Areas in Communications, vol. 32, no. 9, pp. 1637–1652, September 2014.

[5] A. Ghosh, N. Mangalvedhe, R. Ratasuk, M. Cudak, E. Visotsky, T. Thomas, J. Andrews, P. Xia, H. Jo, H. Dhillon, and T. Novlan, “Heterogeneous cellular networks: From theory to practice,” IEEE Communications Magazine, vol. 50, no. 6, pp. 54–64, June 2012.

[6] D. Bharadia, E. McMilin, and S. Katti, “Full duplex radios,” in Proc. 11th USENIX Symposium on Networked Sys. Design and Implementation, April 2014.

[7] D. Bharadia and S. Katti, “Full duplex MIMO radios,” in Proc. 11th USENIX Symposium on Networked Sys. Design and Implementation, April 2014.

[8] M. Duarte and A. Sabharwal, “Experiment-driven characterization of full-duplex wireless systems,” IEEE Transactions on Wireless Communications, vol. 11, no. 12, pp. 4296–4307, December 2012.

[9] M. Luvisotto, A. Sadeghi, F. Lahouti, S. Vitturi, and M. Zorzi, “RCFD: A novel channel access scheme for full-duplex wireless networks based on contention in time and frequency domains,” arXiv preprint arXiv:1606.01038, 2017.

[10] C.-S. Chang, “Stability, queue length, and delay of deterministic and stochastic queueing networks,” IEEE Transactions on Automatic Control, vol. 39, no. 5, pp. 913–931, May 1994.

[11] X. Zhang, J. Tang, H.-H. Chen, S. Ci, and M. Guizani, “Cross-layer-based modeling for quality of service guarantees in mobile wireless networks,” IEEE Communications Magazine, vol. 44, no. 1, pp. 100–106, January 2006.

[12] M. Duarte and A. Sabharwal, “Full-duplex wireless communications using off-the-shelf radios: Feasibility and first results,” in Conference Record of the Forty Fourth Asilomar Conference on Signals, Systems and Computers (ASILOMAR), 2010, pp. 1558–1562.

[13] J. I. Choi, M. Jain, K. Srinivasan, P. Levis, and S. Katti, “Achieving single channel, full duplex wireless communication,” in Proceedings of the Sixteenth Annual International Conference on Mobile Computing and Networking (MobiCom), New York, NY, USA, 2010.

[14] A. Sabeghi, F. Lahouti, and M. Zorzi, “Constellation shaping and LDPC coding in a bidirectional full duplex communication,” in IEEE Global Communications Conference (GLOBECOM), San Diego, USA, December 2015.