

Achieving High Availability in Heterogeneous Cellular Networks via Spectrum Aggregation

Jie Jia¹, Member, IEEE, Yansha Deng², Member, IEEE, Jian Chen³, Member, IEEE, Abdol Hamid Aghvami⁴, Fellow, IEEE, and Arumugam Nallanathan⁵, Fellow, IEEE

Abstract—The exponential growth in data traffic and dramatic capacity demand in fifth generation (5G) have inspired the move from traditional single-tier cellular networks toward heterogeneous cellular networks (HCNs). To face the coming trend in 5G, the high availability requirement in new applications needs to be satisfied to achieve low latency service. Usually, these applications require an availability of six nines or even higher. In this paper, we present a tractable multitier multiband availability model for spectrum aggregation-based HCNs. We first derive a closed-form expression for the availability of spectrum aggregation-based HCNs using the signal-to-interference-plus-noise model. By doing so, we formulate two optimization problems, one is to maximize the average availability, and the other one is to minimize the average power consumption. These two optimization problems are both nonconvex problems, which are challenging to solve. To cope with them, we propose to apply genetic algorithm for the joint user equipment (UE) association, subcarrier assignment, and power allocation problem. Our results show that the average availability in spectrum aggregation-based HCNs improves with decreasing number of UEs, as well as increasing power budget ratio. We also show that increasing the maximum number of aggregated subcarriers decreases the average power consumption, but cannot guarantee the substantial improvement of average availability.

Index Terms—Genetic algorithm, heterogeneous cellular network, high availability, power consumption, spectrum aggregation.

I. INTRODUCTION

IN THE past, the target of wireless technologies has mainly focused on achieving higher data rates and data volumes. However, high average rate and high total data are not the only performance indicators that guarantee the ubiquitous connectivity in next generation wireless networks. According to ABI Research (Allied Business Intelligence Inc.), more than 30 billion devices will be wirelessly connected to the Internet by 2020 [2]. The target of next generation wireless networks has extended to realize high availability and low latency, in order to support the upcoming new applications under the context of Internet of Things (IoT), such as haptic communication [3], cloud computing [4], smart energy grids [5], vehicular communication [6], or industrial automation [7]. The availability requirement of these applications is six nines or higher. A detailed analysis on future application as well as high availability requirement can be found in [8].

The rapid growth of wireless data traffic, fueled by an ever increasing availability requirement of smart mobile computing devices, imposes a huge challenge on current cellular networks. Deploying more macro base stations (BSs) is no longer a sustainable solution to handle the traffic load. Whereas, deploying inexpensive, small-scale, low-power nodes in conventional macrocells becomes a cost-effective solution, which is the so called heterogeneous cellular networks (HCNs) [9]. These low power nodes could be pico or femto BSs. However, due to the heterogeneous deployments of those low power nodes, the interference management among tiers becomes very challenging and extremely important. In [10], [11], the ambient interference from BSs have been utilized for energy transfer to improve the energy efficiency of HetNets. With the irresistible demand to support the aforementioned new applications in HCNs, the modeling, characterization and optimization of availability in HCNs becomes extremely important.

According to the reliability theory [12], generally, there are two feasible methods to achieve high availability in a system. The first method is to substitute or improve some unreliable sub-components to make the system more reliable. The other method is to incorporate redundancy in order to improve the system reliability, through utilizing multiple sub-components in parallel. With multiple less reliable links connected to BSs in parallel boost equivalent availability as that a single more reliable link with higher transmit power or more robust coding.

Data transmission availability can be bootstrapped from physical layer technology. For instance, Spectrum aggregation

Manuscript received January 10, 2017; revised May 16, 2017 and July 24, 2017; accepted September 12, 2017. Date of publication September 26, 2017; date of current version November 10, 2017. This work was supported in part by the National Natural Science Foundation of China under Grants 61772126, 61402096, 61173153, and 61572123, in part by the Fundamental Research Funds for the Central Universities under Grant N150404006, in part by the National Science Foundation for Distinguished Young Scholars of China under Grants 61225012 and 71325002, in part by the Specialized Research Fund of the Doctoral Program of Higher Education for the Priority Development Areas under Grant 20120042130003. The review of this paper was coordinated by Prof. Y.-B. Lin. This paper was presented at the IEEE Global Communications Conference, Washington, DC, USA, December 2016. (Corresponding author: Yansha Deng.)

J. Jia and J. Chen are with the Key Laboratory of Medical Image Computing, Northeastern University, Ministry of Education, Shenyang 110819, China, and with the School of Computer Science and Engineering, Northeastern University, Shenyang 110819, China, and also with the Department of Informatics, King's College London, London WC2R 2LS, U.K. (e-mail: jiajie@mail.neu.edu.cn; chenjian@mail.neu.edu.cn).

Y. Deng and A. H. Aghvami are with the Department of Informatics, King's College London, London WC2R 2LS, U.K. (e-mail: yansha.deng@kcl.ac.uk; hamid.aghvami@kcl.ac.uk).

A. Nallanathan is with the Department of Informatics, King's College London, London WC2R 2LS, U.K., and also with the Queen Mary University of London, London E1 4NS, U.K. (e-mail: a.nallanathan@qmul.ac.uk).

Color versions of one or more of the figures in this paper are available online at <http://ieeexplore.ieee.org>.

Digital Object Identifier 10.1109/TVT.2017.2755504

(carrier aggregation) [13] is a well-known technique that enables multiple less reliable links in parallel to boost availability. As specified by 3GPP in [14], spectrum aggregation, which enables the concurrent utilization of multiple component carriers (CCs) in the physical layer, was originally proposed to increase bit rates and capacity. With spectrum aggregation, the aggregated bandwidth as large as 100 MHz can be obtained by aggregating 5 20 MHz CCs, and the propagation characteristics of different component carriers may also vary significantly. e.g., a CC in the 800 MHz has very different propagation characteristic from a CC in the 2.4 GHz. Recently, spectrum aggregation has been regarded as the primary feature deployed by operators with commercial LTE-Advanced service [15]. In [16], the spectrum aggregation was proposed to improve peak data rate in multi-band HCNs.

The spectrum aggregation has recently been applied to enhance the availability. In [17], the spectrum aggregation was applied to guarantee high availability by a joint transmission over multiple links over different carrier frequencies. However, their work was limited to Rayleigh-fading channel. The work in [17] was extended to [18] by including selection combining and maximal ratio combining over Nakagami- m fading. It is revealed in [17] and [18] that it is more beneficial in terms of power to utilize multiple links in parallel rather than boosting the power of a single link. In [19], combined macro- and microdiverse uplink connections and composite correlated distributions of Nakagami fading and log-normal shadowing was investigated. More recently, an analytical model for availability in multi-connectivity systems utilizing macro- and microdiversity was studied in [20]. Nevertheless, all of the aforementioned works have neglected path loss in the availability model or interference in each carrier.

In order to provide the availability for emergency calls, the priority based schemes has been designed, where network resources are occupied only by these emergency services [21]. Different from emergency services, IoT applications coexist with traditional data-centric applications, and share the network resources with each other. Due to the different achievable capacity of each link and cumulative interference caused by all the simultaneously transmitting nodes, nearby or faraway, simply considering the received power from the desired transmitter may not accurately capture the availability characteristics. A more appropriate model taking into account the interference statistics is the signal-to-interference-plus-noise ratio (SINR) model, which is also the main element determining the Shannon capacity. The SINR model can be widely found in solving the optimization problem in spectrum allocation [22], power control [23], load balancing [24] and UE association [25]. Assuming the shadowing fading as a random variable, [26] studied the high availability in wireless networks with different transmit power at the BS based on SINR model. However, modeling and analyzing the availability in HCNs based on SINR model can be computationally and analytically challenging.

Resource allocation has been proposed to solve power consumption problem in [27]–[30]. In [27], a power optimization scheme was proposed for interference-limited wireless communications. In [28], the energy-efficient spectrum sharing problem

was studied in cognitive radio femtocell networks. In [29], the BS sleep-mode strategies in HCNs with the small cell deployment were proposed to minimize the power consumption. In [30], the resource allocation and UE association was jointly investigated to find the near optimal solution for the minimum total energy consumption of the cellular system using iterative algorithm. However, most of existed resource allocation algorithms consider continuous transmit power allocation, which can not be directly applied to in systems supporting discrete transmit power allocation. For instance, the 3GPP LTE cellular networks only support discrete power allocation in the downlink with a use-specific data-to-pilot-power offset parameters [31]. Compared with the continuous power control, the discrete power control offers two main benefits [32]: (i) the transmitter is simplified, and more importantly, (ii) the overhead of information exchange among networks is significantly reduced. Nevertheless, using simple discretization on the solution obtained by existed continuous power control is not an effective approach. Discrete power allocation for cellular networks has been proposed in [32], [33]. In [32], two discrete power control algorithms were proposed to maximize the weighted system capacity. In [33], a discrete power control was proposed for multi-cell networks aiming at energy efficiency. However, to the best of our knowledge, there is no work dealing with the discrete power control for availability optimization.

Unlike existing works, the aim of this work is to propose a joint UE association, subcarrier assignment and discrete power allocation technique to optimize the availability and power consumption via genetic algorithms (GAs) [34] in HCNs. Due to the advantages in versatility, scalability, and computational simplicity, GAs have become increasingly popular method of solving combinatorial optimization problems in wireless networks [35]–[43]. GAs are proposed to solve the problem of antenna selection for MIMO networks [35], subcarrier pairing and power allocation for cognitive relay networks [36], channel assignment for wireless mesh networks [37], [43], channel and bandwidth allocation for mobile cellular networks [38], [39], energy saving for LTE networks [40], cell deployment for 5G networks [41], and routing and traffic scheduling for multi-hop cellular networks [42]. The main contributions of this paper are summarized as follows:

- 1) We present an analytical model for availability in HCNs based on SINR model. Unlike [44] and [45], where a UE connects to one BS offering the highest instantaneous SINR, we assume each UE connects to multiple BSs with arbitrary SINR values simultaneously. This results in a novel approach to model and analyze availability with multiple connections.
- 2) We derive an exact closed-form expression for the availability of a random UE in HCNs, which is verified by Monte Carlo simulation. Its numerical results reveal the importance of the UE association, the subcarrier assignment and the power allocation in achieving high availability.
- 3) We formulate two optimization problems with the aims of maximizing the average availability under the power budget constraint, and minimizing the average power

TABLE I
NOTATIONS

Symbol	Definition
$\mathcal{K}$	Set of BS tiers
$\mathcal{N}$	Set of all the UEs in the network
$\mathcal{B}$	Set of all the BSs in the network
$\mathcal{B}_k$	Set of BSs in tier k
$\mathcal{N}_s$	Set of UEs associated with the s th BS
$\mathcal{M}$	Set of subcarriers at each BS
ρ	Maximum number of subcarriers that can be aggregated for each UE due to the hardware constraints
$v_{s,n}^m$	Binary variable indicates if the m th subcarrier of the s th BS is allocated to the n th UE or not
$P_s^{\max}$	Maximum transmit power of the s th BS
$P_{s,m}^{\max}$	Maximum transmit power at the m th subcarrier of the s th BS
L	Maximum integer level of transmit power
$l_{s,m}$	Power allocation level at the m th subcarrier of the s th BS
δ	Power budget ratio at any BS
$H_{s,n}$	Channel power gain between the s th BS and the n th UE
$d_{s,n}$	Distance between the s th BS and the n th UE
N_0	Noise power
α_q	Path loss exponent of the q th band
C_q	Path loss constant of the q th band
μ_q	Wavelength of the q th band
τ	Predefined SINR threshold

consumption while satisfying the availability requirement. Due to the complex topology of HCNs, these two optimization problems are NP-hard in nature.

- 4) We propose to apply GA for the joint UE association, subcarrier assignment and power allocation problem. The average availability in spectrum aggregation-based HCNs improves with decreasing the number of UEs, and increasing the maximum number of aggregated subcarriers allowed for each UE. The average power consumption decreases with increasing the maximum number of aggregated subcarriers, and decreasing the number of UEs. To the best of our knowledge, this is the first work of the availability optimization in spectrum aggregation-based HCNs using GA.

The remainder of this paper is organized as follows. In Section II, we present the multi-tier multi-band availability model. Next, in Section III, we formulate the availability maximization problem and the power consumption minimization problem. Section IV applies GA for the joint UE association, subcarrier assignment and power allocation problem. Section V presents the numerical results and Section VI highlights our conclusions.

II. SYSTEM MODEL AND AVAILABILITY CHARACTERIZATION

A. System Model

We consider HCNs with $\mathcal{K} = \{1, \dots, K\}$ denoting the set of K tiers which may include macrocells, picocells, femtocells, and further radiating elements. In this paper, we focus on the downlink transmission and assume open access for all the small cells. We list all the notations in Table I.

We denote the set of UEs as $\mathcal{N} = \{1, 2, \dots, N\}$ and the set of BSs as $\mathcal{B} = \mathcal{B}_1 \cup \mathcal{B}_2 \cup \dots \cup \mathcal{B}_K = \{1, 2, \dots, S\}$, where $\mathcal{B}_k$

represents the set of BSs in tier k . To achieve high availability via multiple link connections, each UE is allowed to be connected with multiple BSs simultaneously. We assume the massive non-continuous carrier aggregation [46] is applied, where UEs can aggregate a large number of (up to 32) continuous and non-continuous subcarriers from heterogeneous spectrum bands. We denote the set of UEs associated with the s th BS as $\mathcal{N}_s$, and thus $\mathcal{N} = \mathcal{N}_1 \cup \mathcal{N}_2 \cup \dots \cup \mathcal{N}_S$. We assume that each BS has maximum Q available bands (e.g., 800 MHz, 2.5 GHz, ...), and each band contains F subcarriers. We denote the set of bands in each BS as $\mathcal{Q} = \{1, 2, \dots, Q\}$, and the set of subcarriers at each BS as $\mathcal{M} = \underbrace{\{1, \dots, F\}}_{\text{band}_1}, \dots, \underbrace{\{(Q-1)F+1, \dots, QF\}}_{\text{band}_Q}$.

We assume that the maximum subcarrier transmit power at the m th subcarrier of the s th BS is $P_{s,m}^{\max}$, and the maximum transmit power of the s th BS is $P_s^{\max}$. We consider the discrete power allocation at the m th subcarrier of the s th BS with integer level $l_{s,m}$, where

$$l_{s,m} \begin{cases} \in [1, L] & \text{If UE occupied } m\text{th subcarrier of } s\text{th BS} \\ = 0 & \text{If no UE occupied } m\text{th subcarrier of } s\text{th BS,} \end{cases} \quad (1)$$

and L is the maximum integer level. Thus, the transmit power allocated to each subcarrier of a BS belongs to the set $[0, \frac{1}{L}P_s^{\max}, \frac{2}{L}P_s^{\max}, \dots, \frac{l_{s,m}}{L}P_s^{\max}, \dots, P_s^{\max}]$.

To specify the UE association and the subcarrier assignment, we denote $v_{s,n}^m$ as the resource-allocation indicator, which is a binary variable. If $v_{s,n}^m = 1$, it indicates that the m th subcarrier of the s th BS ($s \in \mathcal{B}$) is allocated to the n th UE ($n \in \mathcal{N}$), and $v_{s,n}^m = 0$ ($m \in \mathcal{M}$) if otherwise.

We assume the following resource assignment constraint, subcarrier aggregation constraint, and per-BS power constraint need to be satisfied:

- 1) The variable $v_{s,n}^m$ must satisfy that each subcarrier for a BS can only be occupied by at most one UE.
- 2) The total number of aggregated subcarriers for each UE should be at most ρ due to hardware constraints.
- 3) The total power consumption at each BS over all its subcarriers $\sum_{m \in \mathcal{M}} \frac{l_{s,m}}{L} P_{s,m}^{\max}$ should not exceed a power budget $\delta P_s^{\max}$ with the power budget ratio δ .

We use different path loss exponents for different bands to capture the possible large differences in propagation characteristics associated with each band's carrier frequency. We formulate the SINR of the n th UE associated with the m th subcarrier of the s th BS as

$$\text{SINR}_{s,n}^m = \frac{\frac{l_{s,m}}{L} P_{s,m}^{\max} H_{s,n} C_q d_{s,n}^{-\alpha_q} v_{s,n}^m}{\underbrace{\sum_{i \in \mathcal{B} \setminus s} \frac{l_{i,m}}{L} P_{i,m}^{\max} H_{i,n} C_q d_{i,n}^{-\alpha_q}}_{I_{s,n}^m} + N_0}, \quad (2)$$

where $q = \lceil m/F \rceil$, and $\lceil \cdot \rceil$ is the ceiling function. For instance, if $m = 15$, $F = 10$, we have $q = 2$. In (2), $I_{s,n}^m$ is the aggregate interference at the n th UE from all the other BSs over the m th subcarrier, $H_{s,n}$ is the channel power gain between the s th BS and the n th UE, $d_{s,n}$ is the distance between the s th BS and the n th UE, N_0 is the noise power, α_q is the path loss exponent of the

q th band, and C_q is the path loss constant depending strongly on carrier frequency with $C_q = (\frac{\mu_q}{4\pi})^2$ for the wavelength μ_q . Similar as [16], [47]–[49], we ignore shadowing and only consider independent quasistatic Rayleigh fading with $H_{i,n} \sim \exp(1)$ for simplicity. The extension to take the shadowing into account or Rician fading can be incorporated in the availability analysis in Section II via some mathematical manipulations, remind that the GAs proposed in this work will be still valid.

B. Availability Analysis

The signal cannot be successfully received if the SINR value $SINR_{s,n}^m$ is below a certain threshold τ . Therefore, the outage probability of the n th UE associated with the m th subcarrier of the s th BS is characterized as

$$O_{s,n}^m = \mathcal{P}(SINR_{s,n}^m \leq \tau). \quad (3)$$

Thus the availability of the n th UE associated with the m th subcarrier of the s th BS can be derived as

$$\begin{aligned} A_{s,n}^m &= 1 - O_{s,n}^m = 1 - \mathcal{P}(SINR_{s,n}^m \leq \tau) \\ &= \mathcal{P}(SINR_{s,n}^m > \tau). \end{aligned} \quad (4)$$

Generally, $A_{s,n}^m$ denotes the availability of a single connection between UE n with an arbitrary BS s over subcarrier m , and $A_{s,n}^m$ is given in the form of $1 - 10^{-x}$, where x indicates the number of nines. Considering that UE n may connect multiple BSs over multiple connections, its availability is defined by the combination of multiple connection availabilities, which is derived in the following theorem.

Theorem 1: The availability of the n th UE connected to multiple BSs in HCNs is derived as

$$A_n = 1 - \prod_{s \in \mathcal{B}, m \in \mathcal{M}} (1 - A_{s,n}^m), \quad \forall n \in \mathcal{N}, \quad (5)$$

where the availability of the n th UE associated with the m th subcarrier of the s th BS is given by

$$A_{s,n}^m = \begin{cases} 0 & \text{if } v_{s,n}^m = 0 \\ \exp(-\Theta_s \tau N_0) & \text{if } v_{s,n}^m = 1, I_{s,n}^m = 0 \\ \prod_{i=1}^S \Theta_i \sum_{j=1, j \neq s}^S \frac{\exp(-\Theta_s \tau N_0)}{\Theta_s (\Theta_j + \Theta_s \tau) \prod_{k=1, k \neq s, j}^S (\Theta_k - \Theta_j)} & \text{if } v_{s,n}^m = 1, I_{s,n}^m \neq 0, \end{cases} \quad (6)$$

where

$$\Theta_\xi = L / (l_{\xi,m} P_{\xi,m}^{\max} C_q d_{\xi,n}^{-\alpha_q}), \quad (7)$$

and ξ can be s, i, j , and k .

Proof: For $v_{s,n}^m = 0$, we can directly obtain $A_{s,n}^m = 0$.

For $v_{s,n}^m = 1$ with no interference ($I_{s,n}^m = 0$), we present $A_{s,n}^m$ as

$$\begin{aligned} A_{s,n}^m &= \mathcal{P}(SINR_{s,n}^m > \tau) \\ &= 1 - \mathcal{P}\left(\frac{l_{s,m}}{L} P_{s,m}^{\max} H_{s,n} C_q d_{s,n}^{-\alpha_q} \leq \tau N_0\right) \\ &\stackrel{(a)}{=} \exp(-\Theta_s \tau N_0), \end{aligned} \quad (8)$$

where Θ_s is given by (7), and (a) is performed based on $H_{s,n} \sim \exp(1)$.

For $v_{s,n}^m = 1$ and $I_{s,n}^m \neq 0$, we employ the change of variables $X = I_{s,n}^m + N_0$, $Y = \frac{l_{s,m}}{L} P_{s,m}^{\max} H_{s,n} C_q d_{s,n}^{-\alpha_q}$, and $Z = Y/X$ to obtain

$$\begin{aligned} A_{s,n}^m &= \mathcal{P}(Z > \tau) \\ &= \int_{\tau}^{\infty} f_Z(z) dz \\ &= \int_{\tau}^{\infty} \int_0^{\infty} x f_X(x) f_Y(xz) dx dz. \end{aligned} \quad (9)$$

We have

$$f_Y(xz) = \Theta_s \exp(-\Theta_s xz), \quad (10)$$

where Θ_s is given by (7).

Next, we focus on computing $f_X(x)$ with $X = I_{s,n}^m + N_0$ and

$$I_{s,n}^m = \sum_{i \in \mathcal{B} \setminus s} \Omega_i, \quad (11)$$

where

$$\Omega_i = \frac{l_{i,m}}{L} P_{i,m}^{\max} H_{i,n} C_q d_{i,n}^{-\alpha_q}. \quad (12)$$

According to the distribution of channel power gain, we derive

$$f_{\Omega_i}(x) = \Theta_i \exp(-\Theta_i x), \quad (13)$$

where Θ_i is given by (7).

In order to obtain the probability density function (PDF) of the sum of independent exponential random variables $\sum_{i \in \mathcal{B} \setminus s} \Omega_i$, we apply the following lemma [50].

Lemma 1: Let $(W_i)_{i=1 \dots n}$, $n \geq 2$, be the independent exponential random variables with pairwise distinct respective parameters Θ_i , the PDF of their sum is given as

$$f_{W_1+W_2+\dots+W_n}(w) = \left[\prod_{i=1}^n \Theta_i \right] \sum_{j=1}^n \frac{e^{-\Theta_j w}}{\prod_{k=1, k \neq j}^n (\Theta_k - \Theta_j)}. \quad (14)$$

Based on Lemma 1, we derive the PDF of X as

$$\begin{aligned} f_X(x) &= f_{I_{s,n}^m}(x - N_0) \\ &= \prod_{i=1, i \neq s}^S \Theta_i \sum_{j=1, j \neq s}^S \frac{e^{\Theta_j N_0}}{\prod_{k=1, k \neq s, j}^S (\Theta_k - \Theta_j)} e^{-\Theta_j x}, \end{aligned} \quad (15)$$

where Θ_i , Θ_j , and Θ_k can be obtained by using (7).

Substituting (10) and (15) into (9), we obtain

$$\begin{aligned}
 A_{s,n}^m &= \int_{\tau}^{\infty} \int_{N_0}^{\infty} x \prod_{i=1, i \neq s}^S \Theta_i \sum_{j=1, j \neq s}^S \frac{e^{\Theta_j N_0} e^{-\Theta_j x} \Theta_s e^{-\Theta_s x z}}{\prod_{k=1, k \neq s, j}^S (\Theta_k - \Theta_j)} dx dz \\
 &= \prod_{i=1}^S \Theta_i \sum_{j=1, j \neq s}^S \frac{e^{\Theta_j N_0}}{\prod_{k=1, k \neq s, j}^S (\Theta_k - \Theta_j)} Z(\tau, N_0, \Theta_j + \Theta_s z).
 \end{aligned} \quad (16)$$

To solve (16), we derive $Z(\tau, N_0, \Theta_j + \Theta_s z)$ as

$$\begin{aligned}
 Z(\tau, N_0, \Theta_j + \Theta_s z) &= \int_{\tau}^{\infty} \int_{N_0}^{\infty} x e^{-(\Theta_j + \Theta_s z)x} dx dz \\
 &= \int_{\tau}^{\infty} \left(\frac{N_0 e^{-N_0(\Theta_j + \Theta_s z)}}{(\Theta_j + \Theta_s z)} + \frac{e^{-N_0(\Theta_j + \Theta_s z)}}{(\Theta_j + \lambda_s z)^2} \right) dz. \\
 &\stackrel{(b)}{=} \int_{N_0(\Theta_j + \Theta_s \tau)}^{\infty} \left(\frac{N_0}{\Theta_s} \frac{e^{-u}}{u} + \frac{N_0}{\Theta_s} \frac{e^{-u}}{u^2} \right) du \\
 &= \frac{e^{-N_0(\Theta_j + \Theta_s \tau)}}{\Theta_s (\Theta_j + \Theta_s \tau)},
 \end{aligned} \quad (17)$$

where (b) is performed by using $u = N_0 (\Theta_j + \Theta_s z)$.

Substituting (17) into (16), we finally derive $A_{s,n}^m$ as

$$A_{s,n}^m = \prod_{i=1}^S \Theta_i \sum_{j=1, j \neq s}^S \frac{e^{-N_0 \Theta_s \tau}}{\Theta_s (\Theta_j + \Theta_s \tau) \prod_{k=1, k \neq s, j}^S (\Theta_k - \Theta_j)}, \quad (18)$$

where Θ_s , Θ_i , Θ_j , and Θ_k can be obtained by using (7). ■

Note that the derived availability of an arbitrary UE in spectrum aggregation-based HCNs is a easy-to-evaluate closed-form expression. Based on this expression, each UE connects to several BSs, which enables the optimal solution of the proposed availability optimization and power consumption optimization problem. In other words, the connection between each UE and the BSs is decided to achieve the optimal overall network performance.

It should be observed that the availability defined in (5) is different from that of reliability. According to [51], reliability refers to the probability to guarantee a required function/performance under stated conditions within a given time latency, and the specific reliability requirements differ for various types of services and applications. While availability is a transport-agnostic definition from the applications point, and showcases the presence or absence of reliability [52].

Due to the fact that wireless communication systems are typically not designed to provide a reliable level at all times and in every reception scenario, this would harm the acceptance of ultra reliable communication (URC) services and restrict their usage. Our availability measurement is also different from traditional methods, where the availability can be calculated by measuring the ping non-responses and interpolating differences in time between down link alert and uplink alert during months [53]. With the help of availability definition and evaluation in

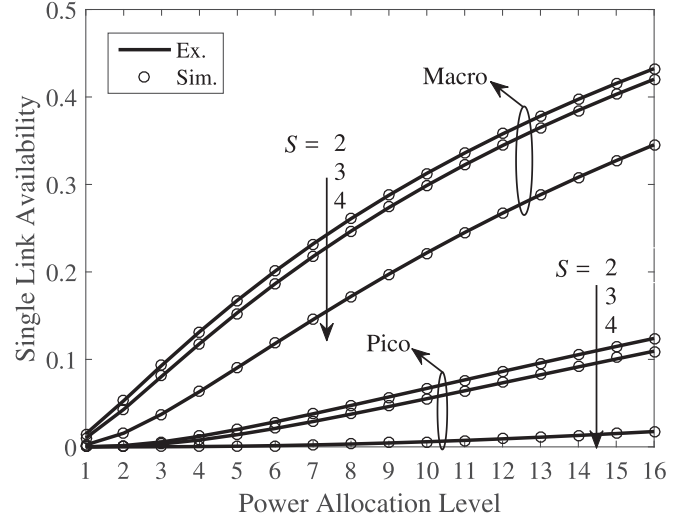


Fig. 1. Single link availability.

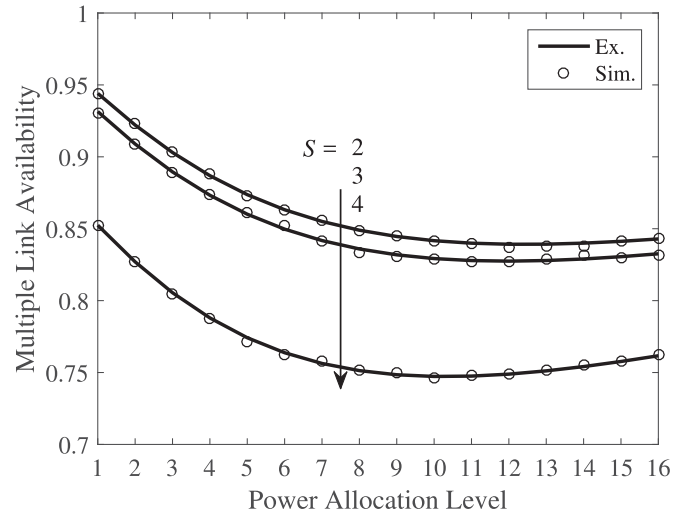


Fig. 2. Multiple links availability.

(5), we can quickly evaluate the availability under given conditions, and find those factors influencing current availability. Thus, the URC services can be quickly deployed in a wide range of scenarios by just considering whether the obtained availability meets its requirement [52].

C. Availability Validation

To verify the derived analytical results for the availability, we plot the analytical curves for the single link availability and the multiple link availability using (18) and (5) with the simulation curves using Monte Carlo simulation in Figs. 1 and 2, respectively. In these two figures, we assume the macro BS with $P_{1,m}^{\max} = 43$ dBm and all the pico BSs with $P_{j,m}^{\max} = 30$ dBm for any subcarrier ($j \neq 1$) for two-tier HCNs, where the distance between the UE and the s th BS is randomly generated. Both figures showcase that the derived analytical results match well with the simulation, which proves the accuracy of our derived results.

Fig. 1 plots the single link availability versus the power allocation level $l_{1,m}$ at the macro BS in two-tier HCNs. We set $C_q = (\frac{0.375}{4\pi})^2$ for all band q . The power allocation level $l_{j,m}$ at the pico BSs is equal to L , which indicates the full power allocation at each pico BSs. As expected, the single link availability of UE connected to the macro BS increases with increasing the transmit power of macro BS. Increasing the number of BSs in HCNs increases the interference, which degrades the single link availability. Importantly, the single link availability is very low, and can hardly achieve the availability with six nines, which reveals the potential of improving the availability via multiple links.

In Fig. 2, we assume that the number of subcarriers at each BS is $M = 2$ with $C_{q_1} = (\frac{0.375}{4\pi})^2$, and $C_{q_2} = (\frac{0.125}{4\pi})^2$, respectively. By comparing Figs. 2 with 1, we see that the availability of a UE connected with multiple links substantially outperforms that with single link, which reveals the need to apply spectrum aggregation technique. We can also see that the multiple link availability decreases with increasing the transmit power, and the highest availability of a UE achieved for the lowest power allocation level $L = 1$ and the minimum number of BSs $S = 2$ reveals the importance of joint optimization on power allocation, UE association and subcarrier assignment in multi-tier multi-band HCNs.

III. PROBLEM FORMULATION

Next, We formulate two optimization problems to achieve the maximum average availability, and to achieve minimum power consumption in spectrum aggregation-based HCNs, respectively.

Availability Maximization Problem: Network aggregate utility is conventionally regarded as a measure for evaluating the performance of resource management protocols [54]–[56]. Based on this criterion, the objective of this problem is to maximize the average availability over all the UEs. Here, the average availability is the sum of availability of all UEs averaging over the total number of UEs as shown in (19). This can be achieved by searching the optimal UE association, subcarrier assignment, and discrete power allocation under the total power consumption constraint. This availability maximization problem is formulated as

$$\max \quad \frac{\sum_{n \in \mathcal{N}} A_n}{N} \quad (19)$$

$$\text{s.t.} \quad \sum_{n \in \mathcal{N}} v_{s,n}^m \leq 1, \forall s \in \mathcal{B}, \forall m \in \mathcal{M}, \quad (20)$$

$$\sum_{s \in \mathcal{B}} \sum_{m \in \mathcal{M}} v_{s,n}^m \leq \rho, \forall n \in \mathcal{N}, \quad (21)$$

$$l_{s,m} \leq L, \forall s \in \mathcal{B}, \forall m \in \mathcal{M}, \quad (22)$$

$$\sum_{m \in \mathcal{M}} l_{s,m} \frac{P_{s,m}^{\max}}{L} \leq \delta P_s^{\max}, \forall s \in \mathcal{B}. \quad (23)$$

The constraints in (20)–(23) are named as the UE association and subcarrier assignment constraint in (20), the subcarrier aggregation constraint in (21), the power level constraint in (22)

and the per-BS power constraint in (23). The *subcarrier assignment and UE association constraint* in (20) represents that each subcarrier of each BS can be allocated to at most one UE. The *subcarrier aggregation constraint* in (21) implies that the maximum number of aggregated subcarriers must satisfy the hardware constraints. The *power level constraint* in (22) represents that the maximum discrete transmit power level of each subcarrier is L . The *per-BS power constraint* in (23) represents that the maximum transmit power at each BS is limited by its total power budget.

Power Consumption Minimization Problem: The objective of the problem is to minimize the average power consumption while satisfying each UE's availability requirement, which is formulated as

$$\min \quad \frac{\sum_{s \in \mathcal{B}} \sum_{m \in \mathcal{M}} \frac{l_{s,m}}{L} P_{s,m}^{\max}}{N} \quad (24)$$

$$\text{s.t.} \quad \sum_{n \in \mathcal{N}} v_{s,n}^m \leq 1, \forall s \in \mathcal{B}, \forall m \in \mathcal{M}, \quad (25)$$

$$\sum_{s \in \mathcal{B}} \sum_{m \in \mathcal{M}} v_{s,n}^m \leq \rho, \forall n \in \mathcal{N}, \quad (26)$$

$$l_{s,m} \leq L, \forall s \in \mathcal{B}, \forall m \in \mathcal{M}, \quad (27)$$

$$\sum_{m \in \mathcal{M}} l_{s,m} \frac{P_{s,m}^{\max}}{L} \leq \delta P_s^{\max}, \forall s \in \mathcal{B}, \quad (28)$$

$$1 - \prod_{s \in \mathcal{B}, m \in \mathcal{M}} (1 - A_{s,n}^m) \geq A_{th}, \forall n \in \mathcal{N}. \quad (29)$$

Note that the constraints of (25)–(28) are the same as (20)–(23) in the availability maximization problem, while the *per-UE availability requirement* in (29) represents that the availability requirement for each UE should be satisfied.

Instinctively, both of these two optimization problems are in the form of mixed integer non-linear programming (MINLP) problem, which are generally NP-hard and cannot be solved by traditional optimization methods [30]. In the next section, we will develop the bio-inspired GA to solve these two optimization problems.

IV. GENETIC ALGORITHM APPROACH

For these above MINLP problems, a straightforward solution is to conduct an exhaustive search by testing all feasible channel and power allocation vectors $v_{s,n}^m$ and $l_{s,m}$. This approach, however, is infeasible for networks with larger number of BSs and UEs. Some other algorithms, such as those in [30], [57], are based on decomposition. In their algorithm, the near-optimal subcarrier assignment and UE association is determined first via heuristic algorithm under fixed power allocation, and the optimal or near-optimal power allocation is obtained via Lagrangian dual based method or iterative heuristic approach with the predetermined optimal subcarrier assignment. However, their approach may be suboptimal due to the fact that the subcarrier assignment and power allocation are interacting with each other, and the subcarrier assignment and power allocation should be optimized in a compact form [58]. Therefore, we

apply GA to integrate these two steps to achieve the interaction between the subcarrier assignment and power allocation.

By simulating the process of evolution in the natural system, GA can be considered as an adaptive heuristic search algorithms, and is very suitable to provide a robust, near optimal solution for many real world NP-hard problems, such as BS placement optimization for LTE heterogeneous networks [59], channel assignment for wireless mesh networks [43]. GA is inherently an evolutionary process that involves individual encoding, selection, crossover, mutation, and replacement operations [34].

A. Individual Encoding

GA cannot deal with the solutions of the optimization problem directly. The solutions needs to be represented as chromosomes in terms of data structure. In our optimization problems, an integer-based encoding scheme reflecting the UE association, the subcarrier assignment, and the power allocation, is proposed to represent the potential solutions.

We first generate an initial population $\mathcal{R}$ with R individuals, and each individual consists of two integer-based matrices, which are the potential solutions of the considered optimization problem. These matrices are generated according to **Algorithm 1** in order to satisfy the UE association and subcarrier assignment constraint, the subcarrier aggregation constraint, the power level constraint, and the per-BS power constraint during initialization to accelerate the convergence process. We represent the two integer-based matrices in the r th individual as follows ($1 \leq r \leq R$):

1) UE association and subcarrier assignment matrix Γ^r is

$$\Gamma^r = \begin{bmatrix} \gamma_{1,1}^r & \cdots & \gamma_{1,M}^r \\ \gamma_{2,1}^r & \cdots & \gamma_{2,M}^r \\ \vdots & \vdots & \vdots \\ \gamma_{S,1}^r & \cdots & \gamma_{S,M}^r \end{bmatrix}, \quad (30)$$

where the matrix element $\gamma_{s,m}^r$ ($1 \leq s \leq S, 1 \leq m \leq M$) indicates the $\gamma_{s,m}^r$ th UE associated with the m th subcarrier of the s th BS. For instance, $\gamma_{s,m}^r = n$ indicates the n th UE associated with the m th subcarrier of the s th BS, thus $v_{s,n}^m = 1$; $\gamma_{s,m}^r = 0$ indicates no UE associated with the m th subcarrier of the s th BS, thus $\sum_{n \in \mathcal{N}} v_{s,n}^m = 0$.

Note that this matrix always satisfies the subcarrier assignment and UE association constraint. According to the population initialization in **Algorithm 1**, we count the number of subcarriers assigned to the n th UE c_n to ensure that c_n is no larger than the subcarrier aggregation constraint ρ . If $c_n > \rho$, the n th UE will become infeasible and be excluded from the set of feasible UEs $\mathcal{N}_{feasible}$.

2) Power allocation matrix L^r is

$$L^r = \begin{bmatrix} l_{1,1}^r & \cdots & l_{1,M}^r \\ l_{2,1}^r & \cdots & l_{2,M}^r \\ \vdots & \vdots & \vdots \\ l_{S,1}^r & \cdots & l_{S,M}^r \end{bmatrix}, \quad (31)$$

Algorithm 1: Population initialization.

```

set  $r = 1, c_n = 0, p_s^{assign} = 0, p_s^{remain} = \delta P_s^{\max},$ 
 $\mathcal{N}_{feasible} = \mathcal{N}$ 
while  $r \leq R$  do
  for  $BS\ s = 1$  to  $S$  do
    for subcarrier  $m = 1$  to  $M$  do
      if  $\mathcal{N}_{feasible} \neq \Phi$  then
        randomly select a UE  $n \in \mathcal{N}$ 
         $\gamma_{s,m}^r = n$ 
         $c_n = c_n + 1$ 
        if  $c_n \geq \rho$  then
           $\mathcal{N}_{feasible} = \mathcal{N}_{feasible} \setminus n$ 
        end
         $p_s^{assign} = p_s^{assign} + \frac{l_{s,m}^r}{L} P_s^{\max}$ 
         $p_s^{remain} = \delta P_s^{\max} - p_s^{assign}$ 
        if  $p_s^{remain} \geq P_s^{\max}$  then
           $l_{s,m}^r = \text{randi}(L)$ 
        else
           $l_{s,m}^r = \text{randi}(\lceil \frac{L}{P_s^{\max}} p_s^{remain} \rceil)$ 
        end
      else
         $\gamma_{s,m}^r = 0$ 
         $l_{s,m}^r = 0$ 
      end
    end
  end
  end
   $r = r + 1$ 
   $c_n = 0$ 
   $p_s^{assign} = 0$ 
   $p_s^{remain} = \delta P_s^{\max}$ 
   $\mathcal{N}_{feasible} = \mathcal{N}$ 
end

```

where $l_{s,m}^r$ represents the power level allocated to the m th subcarrier of the s th BS.

To satisfy the per-BS power constraint, the matrix element $l_{s,m}^r$ is initialized in sequence with increasing m . According to **Algorithm 1**, we compare the maximum subcarrier transmit power $P_s^{\max}$ with the remaining power p_s^{remain} at each BS, where $p_s^{remain} = \delta P_s^{\max} - p_s^{assign}$, with p_s^{assign} representing the power allocated for the s th BS. If $p_s^{remain} \geq P_s^{\max}$, the transmit power allocated to the m th subcarrier can be randomly selected from $[1, L]$, thus $l_{s,m}^r = \text{randi}(L)$. Otherwise we set $l_{s,m}^r = \text{randi}(\lceil \frac{L}{P_s^{\max}} p_s^{remain} \rceil)$, to guarantee that the assigned power cannot be larger than the maximum transmit power at each BS, where $\lceil \cdot \rceil$ is the ceiling function.

One example of encoding scheme is illustrated in Fig. 3 with 4 BSs and 6 UEs deployed in HCNs, where each BS has 3 subcarriers and each UE can associate at most 2 subcarriers. We set the maximum transmit power at each BS $P_s^{\max} = 40$ W, the maximum transmit power at each subcarrier $P_{s,m}^{\max} = 16$ W, and the maximum power level $L = 16$. For instance, $\gamma_{3,1} = 5$ and $l_{3,1} = 9$ indicates that the power level allocated by the 1st BS at the 3rd subcarrier to the 5th UE is 9. It can be also observed that this encoding scheme meets all the constraints except

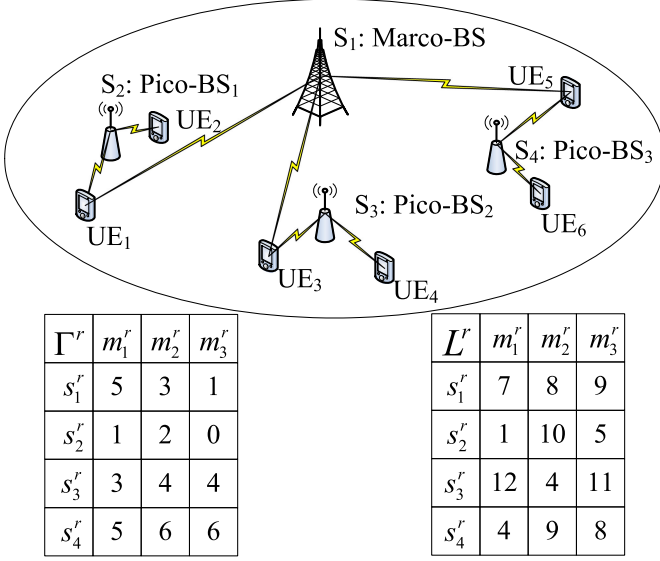


Fig. 3. Individual encoding scheme.

the per-UE availability requirement of the power consumption minimization, which will be satisfied in the following selection process.

B. Fitness Functions and Natural Selection

In GA, selection operation is applied to choose individuals to participate in reproduction, which has a significant impact on driving the search towards a promising trend and finding optimal solutions in a short time. We adopt the famous roulette wheel selection method to select the individual based on its selection probability, which is proportional to its fitness function. The selection probability of the r th individual is defined as

$$q_r = \frac{f(r)}{\sum_{r \in \mathcal{R}} f(r)}, \quad (32)$$

where $f(r)$ is the fitness function of individual r . The quality of the individual is judged by this fitness function.

For the availability maximization problem, since all the constraints are satisfied during initialization, we directly take the objective function as the fitness function, which is given by

$$f_I(r) = \frac{\sum_{n \in \mathcal{N}} A_n}{N}. \quad (33)$$

For the power consumption minimization problem, the fitness function is defined by taking the average network power consumption and a penalty function determined by the relative degree of infeasibility. To provide an efficient search and ensure that the final best solution is feasible, the penalty method [60] is adopted to deal with the availability constraint. The fitness

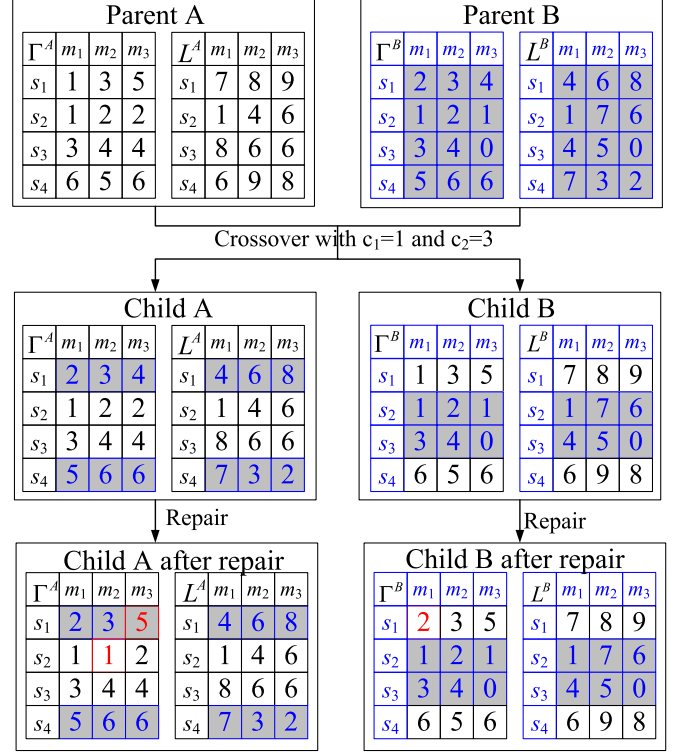


Fig. 4. Two-point crossover and individual repair.

function is expressed as

$$f_{II}(r) = - \left[\frac{\sum_{s \in \mathcal{B}} \sum_{m \in \mathcal{M}} \frac{l_{s,m}}{L} P_{s,m}^{\max}}{N} + \sum_{n \in \mathcal{N}} \alpha_n \max(A_{th} - A_n, 0) \right], \quad (34)$$

where α_n represents the penalty coefficient determined by the per-UE availability requirement. This transforms the power consumption minimization problem to a maximization problem.

C. Crossover and Mutation

The crossover operation is used to mix between the individuals to increase their fitness. In this paper, two-point crossover is performed to produce new solutions. In order to avoid violating the per-BS power constraint, we limit the crossover operation between arbitrary row of the matrices of one individual and that of another individual. Every elements between the two points are switched between two parent individuals to produce two child individuals. The subcarrier aggregation constraint may be violated after crossover operation, thus some elements of UE association and subcarrier assignment matrix need to be repaired by allocating to other UEs.

We illustrate an example of two-point crossover and individual repair operation in Fig. 4, the parameters setting of which is the same as that of Fig. 3, and the randomly generated two crossover points are $c_1 = 1$ and $c_2 = 3$. The crossover between parent A and parent B is performed by switching the rows of the 1th BS and the 4th BS in both matrices of parent A with

that of parent B . After crossover, the assigned subcarriers for the 2th UE and the 4th UE violate the subcarrier aggregation constraint $\rho = 2$ in child A . As such, we repair $\gamma_{1,3}$ and $\gamma_{2,2}$ in child A using randomly generated number 5 and 1 to obtain a repaired child A .

In the mutation operation, the elements in both matrices of each individual are randomly altered to diversify the population after the crossover operation, which will pave the way towards global optima. 1) For the mutation occurring at the arbitrary element of the UE association and subcarrier assignment matrix, repair operation may be required to satisfy the subcarrier aggregation constraint to speed up the convergence; 2) For the mutation occurring at the arbitrary element $l_{s,m}^r$ of the power allocation matrix, mutation operation will be performed using

$$l_{s,m}^r = \text{randi} \left(\left\lceil \min \left(P_s^{\max} - \sum_{i=1, i \neq m}^M l_{s,i} \frac{P_{s,i}^{\max}}{L}, P_{s,m}^{\max} \right) \frac{L}{P_{s,m}^{\max}} \right\rceil \right), \quad (35)$$

where $\lceil \cdot \rceil$ is the ceiling function.

D. Replacement

After generating a new population through the crossover and mutation operators, an elitist model based replacement is employed to update a certain number of individuals in the old population with the new generated individuals. The low quality individuals with the low fitness values in the parental population are replaced by their children in the next generation.

Now, we have designed the key components of the GA operation, which are the individual encoding, population initialization, selection, crossover, mutation, and replacement operation. The joint optimization of UE association, subcarrier assignment and power allocation based on GA is depicted in **Algorithm 2**, where G is the given number of generations, R is the population size, q_c is the crossover probability, and q_m is the mutation probability.

In the proposed GA-based optimization, the computational complexity is dominated by the complexity in evaluating the objective function in (33) or (34), which has to be evaluated R times in each iteration. For the availability maximization problem, with the number of subcarriers as M and the number of UEs as N , the time complexity in calculating the fitness function of the average availability in (33) is $O(MNR)$ within a iteration. For the power consumption minimization problem, with the number of subcarriers as M , the number of UEs as N , and the number of BSs as S , the time complexity in calculating the fitness function of the power consumption in (34) is $O(R(MS + MN))$ within a iteration.

Apart from this, a GA-based approach also depends on other factors, which are difficult to clearly enumerate, such as strategies to generate new population, and the tolerance allowable for cumulative changes in fitness values [61]. Excluding these parameters, the total complexity of our algorithm in solving the availability maximization problem and the power consumption

Algorithm 2: Joint optimization based on GA.

```

set  $g = 1$ 
Generate initiation population  $\mathcal{R}$  using Algorithm 1
Calculate fitness value for each individual in  $\mathcal{R}$ 
while  $g \leq G$  do
  Set  $\mathcal{R}' = \Phi$ 
  for  $i = 1$  to  $R/2$  do
    Select two parents  $p_1$  and  $p_2$  from  $\mathcal{R}$  using
    roulette wheel selection method
     $r_{2*i-1} = p_1$  and  $r_{2*i} = p_2$ 
    Cross  $r_{2*i-1}$  and  $r_{2*i}$  using two-point crossover
    strategy with probability  $q_c$ , and produce two
    children  $r'_{2*i-1}$  and  $r'_{2*i}$ 
    Repair elements in UE association and subcarrier
    assignment matrix if needed
    Mutate  $r'_{2*i-1}$  and  $r'_{2*i}$  using mutation strategy
    with probability  $q_m$ 
    Repair elements in UE association and subcarrier
    assignment matrix if needed
     $\mathcal{R}' = \mathcal{R}' \cup \{r'_{2*i-1}, r'_{2*i}\}$ 
    Calculate fitness value for each individual in  $\mathcal{R}'$ 
  end
   $g = g + 1$ 
  Replace the individuals with low fitness values in
  population  $\mathcal{R}$  with the children in offspring  $\mathcal{R}'$ 
end
Return the fittest individual in  $\mathcal{R}$ 

```

TABLE II
SIMULATION PARAMETERS

Parameter	Value
The number of macro BS	1
The number of pico BS	9
The number of UEs N	2 ~ 20
Maximum transmit power of macro BS	46 dBm (40 W)
Maximum transmit power of pico BS	30 dBm (1 W)
Maximum aggregated subcarriers per UE	1 ~ 10
The availability threshold A_{th}	$1 - 10^{-6}$ (six nines)
800MHz band's wavelength μ_1	0.375 m
2.5GHz band's wavelength μ_2	0.125 m
800MHz band's path loss exponent α_1	3
2.5GHz band's path loss exponent α_2	4
The number of subcarriers in each band	10
Maximum integer power level L	1 ~ 32
Maximum subcarrier transmit power of macro BS	(40/10) W
Maximum subcarrier transmit power of pico BS	(1/10) W
Noise PSD	-174 dBm
SINR threshold τ	1
Population size	20
Crossover probability	0.95
Mutation probability	0.005
Maximum generation	2000

minimization problem are $O(G(MNR + R^2))$ and $O(G(MSR + MNR + R^2))$, respectively.

V. NUMERICAL RESULTS

In this section, we provide numerical results to illustrate the performance of our proposed algorithm. We consider spectrum

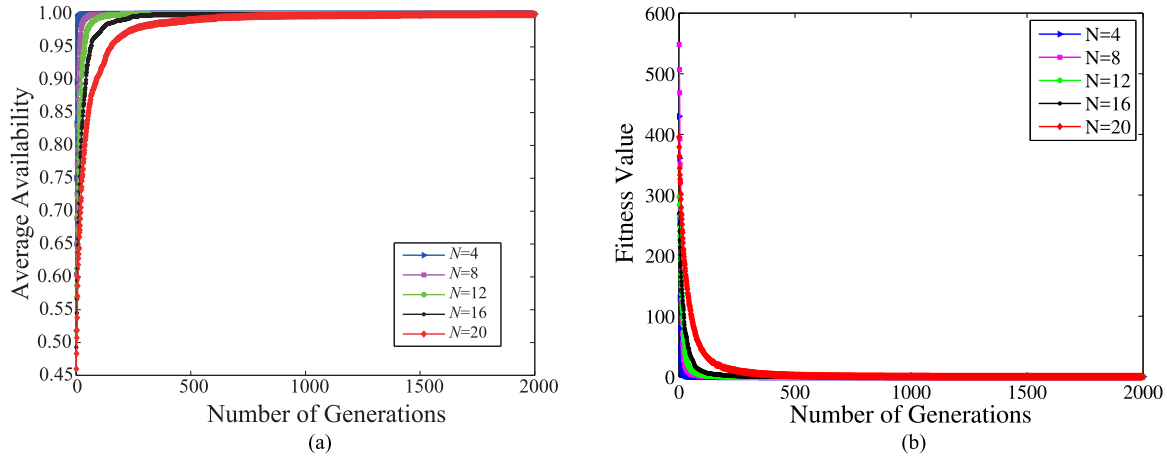


Fig. 5. (a) Convergence behavior of the availability maximization problem. (b) Convergence behavior of the power consumption minimization problem.

aggregation-based HCNs consisting of 2 tiers (macro and pico) with 2 bands (800 MHz and 2.5 GHz). The set-up is a circle area with size $(\pi 500^2) \text{ m}^2$, where the macro BS is located at the center, the pico BSs and UEs are randomly distributed in this circle area. The details of parameters are summarized in Table II unless otherwise specified. The corresponding simulations are implemented in Matlab 7 in a laptop with Intel (i5-4300) CPU. All the results are obtained by averaging 100 simulations.

A. Convergence Behavior

In GA, the convergence behavior is affected by many control parameters, such as the initial population, mutation probability, crossover mechanism, etc.. To the best of our knowledge, the conditions for GAs to converge have been proved only for the binary encoding with Markov chain models [62]. However, for the GA algorithm with integer or real encoding, the convergence is still an open problem [39]. In this paper, instead of using an analytical approach, extensive simulations are employed to look at the convergence issue. In our simulations, we set the maximum number of generation as 2000. Actually, the number of generations depends on the number of size of individuals. For instance, more generations are needed for a larger number of UEs or number of subcarriers.

Fig. 5(a) plots the convergence behavior of the availability maximization problem with the maximum number of aggregated subcarriers $\rho = 5$, and the power budget ratio $\delta = 1$. Fig. 5(b) plots the convergence behavior of the power consumption minimization problem with the availability threshold of 6 nines ($A_{th} = 1 - 10^{-6}$), and $\rho = 5$. From Fig. 5(a) and (b), we can observe that the algorithm converge after approximately 500 number of generations for various number of UEs. It takes 20 seconds to converge for $N = 10$ HCNs. This is sufficient for many applications. If we use a more powerful computer, it is expected that it can converge much faster.

For the availability maximization problem, the average availability with random allocation at the initialization is 0.564944, while the final average availability after optimization with GA is 0.999859, which showcase that the GA achieves nearly 50% more average availability compared with that of the random

TABLE III
OPTIMIZED AVERAGE AVAILABILITY VALUE

N	4	8	12	16	20
Availability by GA	10 nines	7 nines	5 nines	3 nines	3 nines
Optima	11 nines	7 nines	5 nines	3 nines	3 nines

TABLE IV
OPTIMIZED POWER CONSUMPTION VALUE

N	4	8	12	16	20
Fitness	0.009	0.023	0.138	0.771	3.448
Availability	6 nines	6 nines	5 nines	3 nines	3 nines
APC (W)	0.009	0.023	0.042	0.087	0.214
OPC (W)	0.009	0.021	0.038	0.079	0.193

resource allocation. For the power consumption minimization problem, the GA achieves a huge decrease of fitness value during evolution, this can be explained by the fact that the random resource allocation cannot satisfy the *per-UE availability requirement*, thus a large penalty value is introduced in the fitness function in (34). Additionally, it is revealed that the converge speed can be substantially increased with reduced number of UEs in HCNs.

We then present the optimized average availability, and the optimized power consumption with corresponding achieved average availability for various number of UEs in Tables III and IV, where APC means the average power consumption. Additionally, we present the optimal availability and power consumption that based on brute force approach, where OPC means the optimal power consumption. In both Tables, we see that the availability of 6 nines can be achieved when the number of UEs is less than 8. In Table IV, due to the availability of 6 nines requirement is satisfied for $N = 4$ and $N = 8$, no penalty value is introduced to the fitness value, and results in equal value as the power consumption. However, the violation of *per-UE availability requirement* (6 nines) for $N = 12, 16$, and 20 results in the added penalty values as shown in the fitness values.

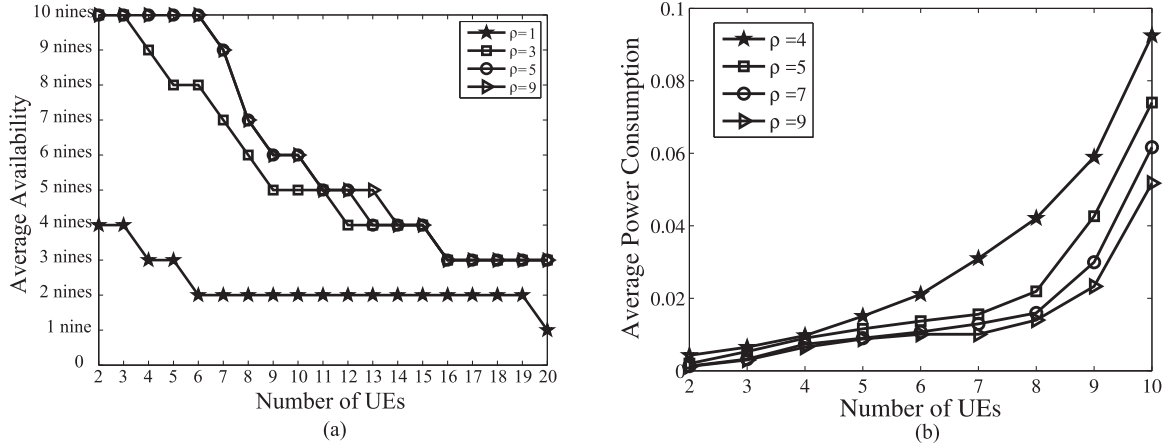


Fig. 6. (a) Average availability versus the number of UEs. (b) Average power consumption versus the number of UEs.

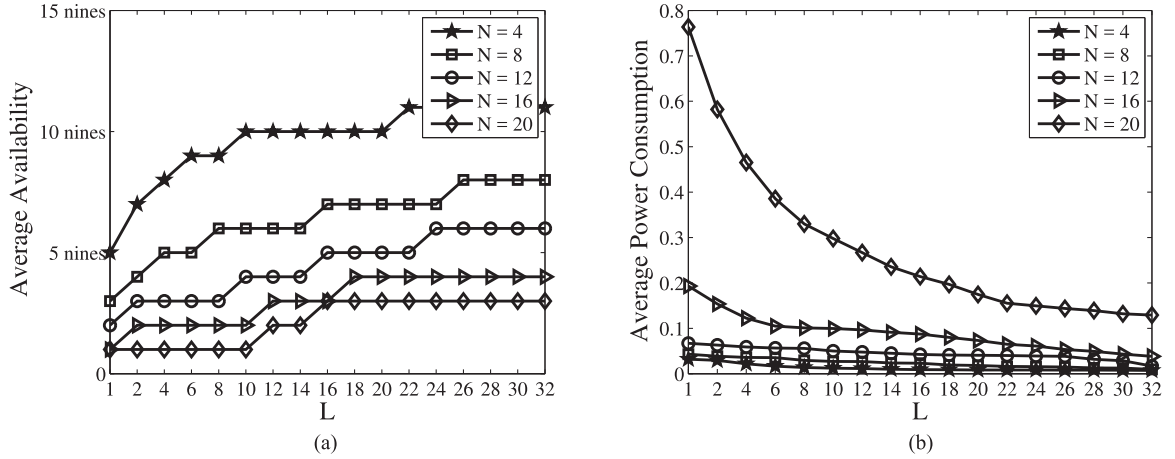


Fig. 7. (a) Average availability versus different power levels. (b) Average power consumption versus different power levels.

We also observe that the optimized value based on GA closely approaches the optima obtained by brute force approach, which showcases the effective of GA for availability maximization or power consumption minimization.

B. Impact of the Number of UEs and the Subcarrier Aggregation Constraint

Fig. 6(a) plots the average availability versus the number of UEs for various subcarrier aggregation constraint ρ . We observe that the average availability decreases with increasing the number of UEs. This can be explained by the fact that the transmit power allocated to the UE decreases and the interference from the same subcarrier at other BSs increases with increasing the number of UEs. More importantly, the average availability can be improved by relaxing the maximum number of aggregated subcarriers. For the availability maximization problem, we can observe that the substantial improvement of average availability is achieved from single subcarrier constraint to three aggregated subcarriers constraint, however further increasing the maximum number of aggregated subcarriers can not achieve much improvement. This indicates that increasing the maximum

number of aggregated subcarriers may not guarantee substantial improvement of average availability.

Fig. 6(b) plots the optimized average power consumption versus the number of UEs for various subcarrier aggregation constraint ρ . Due to the increased per-subcarrier interference with increasing the number of UEs, the average power consumption increases with increasing the number of UEs. Another important observation is that utilizing multiple connections can be an efficient way to save power and improve availability. For instance, for HCNs with 9 UEs fulfilling the availability requirement, the average power consumption with $\rho = 4$ is around 0.059 W, whereas that with $\rho = 9$ is around 0.023 W.

C. Impact of the Maximum Power Levels and Power Budget Ratio

Fig. 7(a) plots the average availability versus the maximum power levels for various number of UEs. It is shown that the average availability increases with increasing the maximum power levels for the same number of UEs. And the achieved availability is much larger than that with on power control ($L = 1$), which showcases the importance of discrete power control. However,

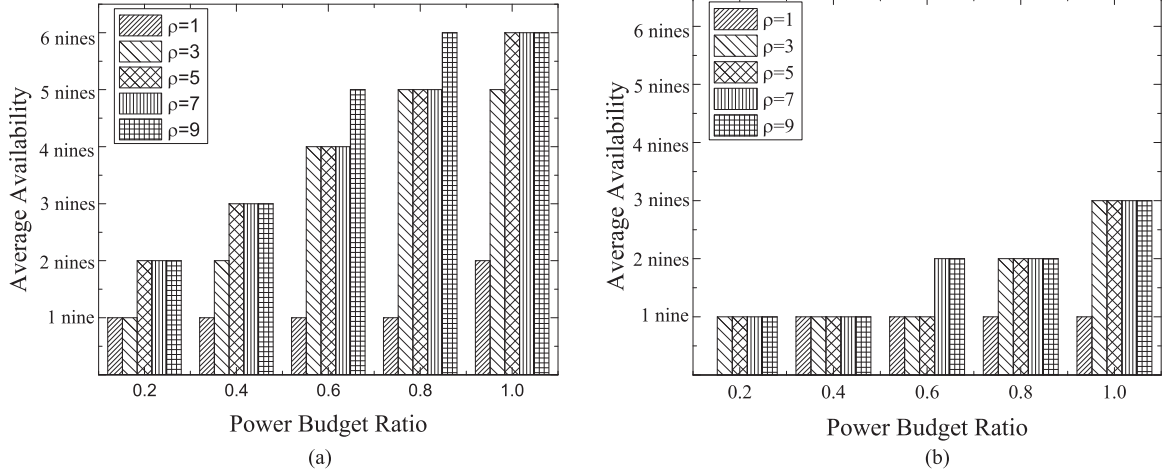


Fig. 8. (a) Average availability of 10 UEs versus different power budget ratios. (b) Average availability of 20 UEs versus different power budget ratios.

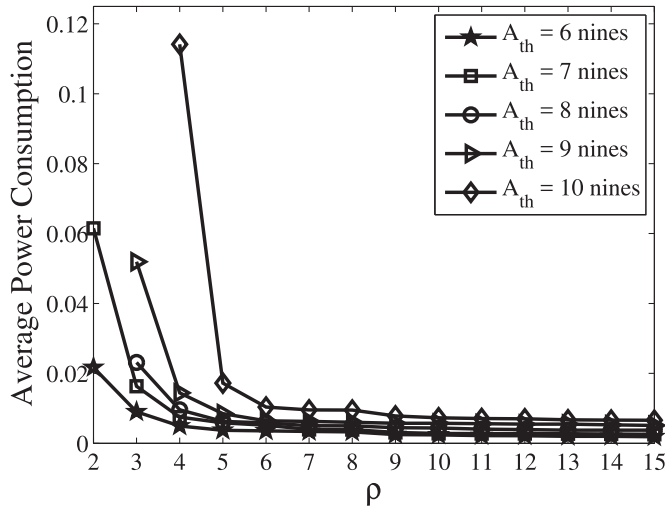


Fig. 9. Average power consumption of 4 UEs.

the average availability of 6 nines is not achievable in HCNs with $N=16$ or 20 UEs even with $L=32$, which means that increasing L can not guarantee substantial improvement in the average availability. Fig 7(b) plots the power consumption versus different L for different number of users N . We see that average power consumption decreases with increasing L , especially for N is larger. However, when N is small, increasing L can not guarantee substantial improvement in minimizing average power consumption.

Fig. 8(a) and (b) plot the average availability versus different power budget ratio δ for different ρ . It is shown that the average availability increases with increasing δ for the same ρ , which results from the increased received power. The six nines of average availability can be achieved for HCNs with 10 UEs for $\rho=5 \sim 9$ and $\delta=1$, these availability values are sufficient for the requirement of many real-time applications. However, the average availability of 6 nines is not achievable in HCNs with 20 UEs even with $\delta=1$ and $\rho=9$. Similar as the observation in Fig. 6(b), increasing the maximum number of aggregated

subcarriers can not guarantee substantial improvement in the average availability.

D. Impact of the Maximum Number of Aggregated Subcarriers

Fig. 9 plots the average power consumption versus different maximum number of aggregated subcarriers ρ for various availability threshold A_{th} with $N=4$ UEs. We see that in order to achieve higher per-UE availability requirement, more number of allowed aggregated subcarriers is needed. It is revealed that the average power consumption decreases with increasing the maximum number of aggregated subcarriers. The higher per-UE availability requirement results in higher average power consumption.

VI. CONCLUSION

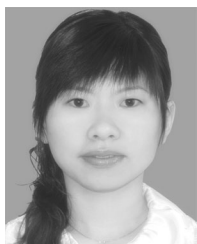
In this paper, we have presented the theoretical model and optimization algorithm to achieve high availability in spectrum aggregation-based HCNs. We have developed a novel availability model under the SINR model. We have also derived a closed-form expression for the availability in spectrum aggregation-based HCNs. We have formulated two optimization problems to maximize the average availability and minimize the average power consumption. To solve the non-convex optimization problems, we have proposed an efficient GA-based algorithm for the joint optimization of the UE association, the subcarrier assignment, and the power allocation. The average availability in spectrum aggregation-based HCNs can be improved by decreasing the number of UEs as well as increasing the power budget ratio. Increasing the maximum number of aggregated subcarriers decreases the average power consumption, but can not guarantee the substantial improvement of average availability.

REFERENCES

- [1] J. Jia, Y. Deng, S. Ping, H. Aghvami, and A. Nallanathan, "High availability optimization in heterogeneous cellular networks," in *Proc. IEEE Global Telecommun. Conf.*, Dec. 2016, pp. 1–6.

- [2] Cisco, "More than 30 billion devices will wirelessly connect to the internet of everything in 2020," May 2013, Cisco, San Jose, CA, USA. [Online]. Available: <https://www.abiresearch.com/press/more-than-30-billion-devices-will-wirelessly-connect/>
- [3] E. Steinbach *et al.*, "Haptic communications," *Proc. IEEE*, vol. 100, no. 4, pp. 937–956, Apr. 2012.
- [4] M. Jammal, A. Kanso, and A. Shami, "High availability-aware optimization digest for applications deployment in cloud," in *Proc. IEEE Int. Conf. Commun.*, Jun. 2015, pp. 6822–6828.
- [5] Z. M. Fadlullah, M. M. Fouda, N. Kato, A. Takeuchi, N. Iwasaki, and Y. Nozaki, "Toward intelligent machine-to-machine communications in smart grid," *IEEE Commun. Mag.*, vol. 49, no. 4, pp. 60–65, Apr. 2011.
- [6] W. Viriyasitavat, M. Boban, H.-M. Tsai, and A. Vasilakos, "Vehicular communications: Survey and challenges of channel and propagation models," *IEEE Veh. Technol. Mag.*, vol. 10, no. 2, pp. 55–66, May 2015.
- [7] A. Frotzschner *et al.*, "Requirements and current solutions of wireless communication in industrial automation," in *Proc. IEEE Int. Conf. Commun. Workshops*, Jun. 2014, pp. 67–72.
- [8] G. P. Fettweis, "The tactile internet: Applications and challenges," *IEEE Veh. Technol. Mag.*, vol. 9, no. 1, pp. 64–70, Mar. 2014.
- [9] D. Lopez-Perez, S. Güvenç, G. De la Roche, M. Kountouris, T. Q. S. Quek, and J. Zhang, "Enhanced intercell interference coordination challenges in heterogeneous networks," *IEEE Wireless Commun. Mag.*, vol. 18, no. 3, pp. 22–30, Jun. 2011.
- [10] Y. Deng, L. Wang, M. Elkashlan, M. Drenzo, and J. Yuan, "Modeling and analysis of wireless power transfer in heterogeneous cellular networks," *IEEE Trans. Commun.*, vol. 64, no. 12, pp. 5290–5303, Dec. 2016.
- [11] S. Akbar, Y. Deng, A. Nallanathan, M. Elkashlan, and A. H. Aghvami, "Simultaneous wireless information and power transfer in K-tier heterogeneous cellular networks," *IEEE Trans. Wireless Commun.*, vol. 15, no. 8, pp. 5804–5818, Aug. 2016.
- [12] P. Pukite and J. Pukite, *Modeling for Reliability Analysis: Markov Modeling for Reliability, Maintainability, Safety, and Supportability*. Hoboken, NJ, USA: Wiley, 1998.
- [13] H. Lee, S. Vahid, and K. Moessner, "A survey of radio resource management for spectrum aggregation in LTE-advanced," *IEEE Commun. Surveys Tuts.*, vol. 16, no. 2, pp. 745–760, May 2014.
- [14] 3GPP, "Carrier Aggregation Base Station (BS) radio transmission and reception," Sophia Antipolis Cedex, France, Tech. Rep. 36.808, 2012.
- [15] 4G Americas, "LTE aggregation and unlicensed spectrum—4G Americas," *4G Americas*, May 2015. [Online]. Available: <http://www.5gamericas.org/files/1214/4648/2397/>
- [16] X. Lin, J. G. Andrews, and A. Ghosh, "Modeling, analysis and design for carrier aggregation in heterogeneous cellular networks," *IEEE Trans. Commun.*, vol. 61, no. 9, pp. 4002–4015, Sep. 2013.
- [17] D. Ohmann, M. Simsek, and G. P. Fettweis, "Achieving high availability in wireless networks by an optimal number of Rayleigh-fading links," in *Proc. IEEE Global Telecommun. Conf. Workshops*, Dec. 2014, pp. 1402–1407.
- [18] D. Ohmann, W. Teske, and G. P. Fettweis, "Combining Nakagami-m fading links for high availability in wireless networks," in *Proc. IEEE Veh. Technol. Conf.*, May 2015, pp. 1–6.
- [19] Z. Rui, W. Jibo, and V. Leung, "Outage probability of composite microscopic and macroscopic diversity over correlated shadowed fading channels," *China Commun.*, vol. 10, no. 11, pp. 129–142, Nov. 2013.
- [20] F. Kirsten, D. Ohmann, M. Simsek, and G. P. Fettweis, "On the utility of macro- and microdiversity for achieving high availability in wireless networks," in *Proc. IEEE 24th Int. Symp. Pers., Indoor, Mobile Radio Commun.*, Sep. 2015, pp. 1723–1728.
- [21] X. Meng, P. Zerfos, V. Samanta, S. H. Y. Wong, and S. Lu, "Analysis of the reliability of a nationwide short message service," in *Proc. 26th IEEE Int. Conf. Comput. Commun.*, 2007, pp. 1811–1819.
- [22] S. Wang, M. Ge, and W. Zhao, "Energy-efficient resource allocation for OFDM-based cognitive radio networks," *IEEE Trans. Commun.*, vol. 61, no. 8, pp. 3181–3191, Aug. 2013.
- [23] Y. Kwon, T. Hwang, and X. Wang, "Energy-efficient transmit power control for multi-tier MIMO HetNets," *IEEE J. Sel. Areas Commun.*, vol. 33, no. 10, pp. 2070–2086, Oct. 2015.
- [24] Q. Ye, B. Rong, Y. Chen, M. Al-Shalash, C. Caramanis, and J. G. Andrews, "User association for load balancing in heterogeneous cellular networks," *IEEE Trans. Wireless Commun.*, vol. 12, no. 6, pp. 2706–2716, Jun. 2013.
- [25] H. Pervaiz, L. Musavian, and Q. Ni, "Joint user association and energy-efficient resource allocation with minimum-rate constraints in two-tier HetNets," in *Proc. 24th Annu. Int. Symp. Pers., Indoor, Mobile Radio Commun.*, Sep. 2013, pp. 1634–1639.
- [26] D. Oehmann, A. Awada, I. Vierung, M. Simsek, and G. Fettweis, "SINR model with best server association for high availability studies of wireless networks," *IEEE Wireless Commun. Lett.*, vol. 5, no. 1, pp. 60–63, Feb. 2016.
- [27] G. Miao, N. Himayat, G. Y. Li, and S. Talwar, "Distributed interference-aware energy-efficient power optimization," *IEEE Trans. Wireless Commun.*, vol. 10, no. 4, pp. 1323–1333, Apr. 2011.
- [28] R. Xie, F. R. Yu, and H. Ji, "Energy-efficient spectrum sharing and power allocation in cognitive radio femtocell networks," in *Proc. IEEE INFOCOM*, Mar. 2012, pp. 1665–1673.
- [29] Y. S. Soh, T. Q. S. Quek, M. Kountouris, and H. Shin, "Energy efficient heterogeneous cellular networks," *IEEE J. Sel. Areas Commun.*, vol. 31, no. 5, pp. 840–850, May 2013.
- [30] X. Sun and S. Wang, "Resource allocation scheme for energy saving in heterogeneous networks," *IEEE Trans. Wireless Commun.*, vol. 14, no. 8, pp. 4407–4416, Aug. 2015.
- [31] 3GPP, "Evolved universal terrestrial radio access (E-UTRA): Physical layer procedure," 3GPP TS 36.213 version 8.4.0 Release 8 Std., 2008.
- [32] H. Zhang, L. Venturino, N. Prasad, P. Li, S. Rangarajan, and X. Wang, "Weighted sum-rate maximization in multi-cell networks via coordinated scheduling and discrete power control," *IEEE J. Sel. Areas Commun.*, vol. 29, no. 6, pp. 1214–1224, Jun. 2011.
- [33] H. Nguyen and W. Hwang, "Distributed scheduling and discrete power control for energy efficiency in multi-cell networks," *IEEE Commun. Lett.*, vol. 19, no. 12, pp. 2198–2201, Dec. 2015.
- [34] D. E. Goldberg, *Genetic Algorithms*. Noida, India: Pearson Educ. India, 2006.
- [35] B. Makki, A. Ide, T. Svensson, T. Eriksson, and M.-S. Alouini, "A genetic algorithm-based antenna selection approach for large-but-finite MIMO networks," *IEEE Trans. Veh. Commun.*, vol. 66, no. 7, pp. 6591–6595, Jul. 2017.
- [36] H.-S. Lang, S.-C. Lin, and W.-H. Fang, "Subcarrier pairing and power allocation with interference management in cognitive relay networks based on genetic algorithms," *IEEE Trans. Veh. Commun.*, vol. 65, no. 9, pp. 7051–7063, Sep. 2016.
- [37] T.-Y. Lin, K.-C. Hsieh, and H.-C. Huang, "Applying genetic algorithms for multiradio wireless mesh network planning," *IEEE Trans. Veh. Technol.*, vol. 61, no. 5, pp. 2256–2270, Jun. 2012.
- [38] S. S. M. Patra, K. Roy, S. Banerjee, and D. P. Vidyarthi, "Improved genetic algorithm for channel allocation with channel borrowing in mobile computing," *IEEE Trans. Mobile Comput.*, vol. 5, no. 7, pp. 884–892, Jul. 2006.
- [39] W.-H. Sheen, S.-J. Lin, and C.-C. Huang, "Downlink optimization and performance of relay-assisted cellular networks in multicell environments," *IEEE Trans. Veh. Technol.*, vol. 59, no. 5, pp. 2529–2542, Jun. 2010.
- [40] H. Ghazzai, E. Yaacoub, M.-S. Alouini, and A. Abu-Dayya, "Optimized smart grid energy procurement for LTE networks using evolutionary algorithms," *IEEE Trans. Veh. Technol.*, vol. 63, no. 9, pp. 4508–4519, Nov. 2014.
- [41] C.-W. Tsai, H.-H. Cho, T. Shih, J.-S. Pan, and J. Rodrigues, "Metaheuristics for the deployment of 5G," *IEEE Wireless Commun.*, vol. 22, no. 6, pp. 40–46, Dec. 2015.
- [42] B. Lorenzo and S. Glisic, "Optimal routing and traffic scheduling for multihop cellular networks using genetic algorithm," *IEEE Trans. Mobile Comput.*, vol. 12, no. 11, pp. 2274–2288, Nov. 2013.
- [43] Y. Ding, Y. Huang, G. Zeng, and L. Xiao, "Using partially overlapping channels to improve throughput in wireless mesh networks," *IEEE Trans. Mobile Comput.*, vol. 11, no. 11, pp. 1720–1733, Nov. 2012.
- [44] J. G. Andrews, F. Baccelli, and R. K. Ganti, "A tractable approach to coverage and rate in cellular networks," *IEEE Trans. Commun.*, vol. 59, no. 11, pp. 3122–3134, Nov. 2011.
- [45] H. S. Dhillon, R. K. Ganti, F. Baccelli, and J. G. Andrews, "Modeling and analysis of K-tier downlink heterogeneous cellular networks," *IEEE J. Sel. Areas Commun.*, vol. 30, no. 3, pp. 550–560, Apr. 2012.
- [46] S. Stefanatos and F. Foukalas, "A filter-bank transceiver architecture for massive non-contiguous carrier aggregation," *IEEE J. Sel. Areas Commun.*, vol. 35, no. 1, pp. 215–227, Jan. 2017.
- [47] Y. Deng, L. Wang, M. Elkashlan, M. Di Renzo, and J. Yuan, "Modeling and analysis of wireless power transfer in heterogeneous cellular networks," *IEEE Trans. Commun.*, vol. 64, no. 12, pp. 5290–5303, Dec. 2016.

- [48] Y. Deng, L. Wang, S. A. R. Zaidi, J. Yuan, and M. ElKashlan, "Artificial-noise aided secure transmission in large scale spectrum sharing networks," *IEEE Trans. Commun.*, vol. 64, no. 5, pp. 2116–2129, May 2016.
- [49] S. Akbar, Y. Deng, A. Nallanathan, M. ElKashlan, and A.-H. Aghvami, "Simultaneous wireless information and power transfer in k -tier heterogeneous cellular networks," *IEEE Trans. Wireless Commun.*, vol. 15, no. 8, pp. 5804–5818, Aug. 2016.
- [50] M. Balázs, "Sum of independent exponentials," 2005. [Online]. Available: <http://www.math.bme.hu/balazs/sumexp.pdf>
- [51] ITU-T Telecommunication Standardization Sector of ITU, "Series E: Overall network operation, telephone service, service operation and human factors quality of telecommunication services: Concepts, models, objectives and dependability planning—use of quality of service objectives for planning of telecommunication networks," E.800, Sep. 2008.
- [52] H. D. Schotten, R. Sattiraju, D. G. Serrano, Z. Ren, and P. Fertl, "Availability indication as key enabler for ultra-reliable communication in 5G," in *Proc. Eur. Conf. Netw. Commun.*, 2014, pp. 1–5.
- [53] W. Greene and B. Lancaster, "Carrier-grade: Five nines, the myth and the reality," *Pipeline Mag.*, vol. 3, no. 11, Apr. 2007.
- [54] S. Maghsudi and S. Stańczak, "Hybrid centralized–distributed resource allocation for device-to-device communication underlying cellular networks," *IEEE Trans. Veh. Commun.*, vol. 65, no. 4, pp. 2481–2495, Apr. 2016.
- [55] A. Ferragut and F. Paganini, "Network resource allocation for users with multiple connections: Fairness and stability," *IEEE/ACM Trans. Netw.*, vol. 22, no. 2, pp. 349–362, Apr. 2014.
- [56] M. H. Cheung, H. Mohsenian-Rad, V. W. S. Wong, and R. Schober, "Utility-optimal random access for wireless multimedia networks," *IEEE Wireless Commun. Lett.*, vol. 1, no. 4, pp. 340–343, Aug. 2012.
- [57] X. Li, C. Li, and Y. Jin, "Dynamic resource allocation for transmit power minimization in OFDM-based noma systems," *IEEE Commun. Lett.*, vol. 20, no. 12, pp. 2558–2561, Dec. 2016.
- [58] T. Wang and L. Vandendorpe, "Iterative resource allocation for maximizing weighted sum min-rate in downlink cellular OFDMA systems," *IEEE Trans. Signal Process.*, vol. 59, no. 1, pp. 223–234, Jan. 2011.
- [59] S. Lee, S. K. Lee, K. Kim, and Y. H. Kim, "Base station placement algorithm for large-scale LTE heterogeneous networks," *PLoS One*, vol. 10, no. 10, Oct. 2015, Art. no. 0139190.
- [60] A. Homaifar, C. X. Qi, and S. H. Lai, "Constrained optimization via genetic algorithms," *Simulation*, vol. 62, no. 4, pp. 242–253, Apr. 1994.
- [61] D. T. Ngo, C. Tellambura, and H. H. Nguyen, "Efficient resource allocation for OFDMA multicast systems with spectrum-sharing control," *IEEE Trans. Veh. Technol.*, vol. 58, no. 9, pp. 4878–4889, Nov. 2009.
- [62] J. Suzuki, "A Markov chain analysis on simple genetic algorithms," *IEEE Trans. Syst., Man, Cybern.*, vol. 25, no. 4, pp. 655–659, Apr. 1995.



Jie Jia received the Ph.D. degree in computer science and technology in 2009 from the Northeastern University, Shenyang, China, where she is currently an Associate Professor. In 2016, she was a Visiting Research Associate at the King's College London. She is a member of various international societies, such as the IEEE and China Computer Federation. She has published more than 100 technical papers on various aspects of wireless networks. Her current research mainly includes HetNets, IoT, and cognitive radio networks.



Yansha Deng (S'13–M'17) received the Ph.D. degree in electrical engineering from the Queen Mary University of London, London, U.K., in 2015. From 2015 to 2017, she was a Postdoctoral Research Fellow in the Department of Informatics, King's College London, London, U.K., where she is currently a Lecturer. Her research interests include massive MIMO, HetNets, molecular communication, cognitive radio, cooperative networks, and physical layer security. She received the Best Paper Award in IEEE International Conference on Communications 2016.

She is currently an Editor of the IEEE COMMUNICATIONS LETTERS. She has also served as a TPC member for many IEEE conferences, such as IEEE GLOBE-COM and IEEE ICC.



Jian Chen received the Ph.D. degree in computer science and technology in 2010 from the Northeastern University, Shenyang, China, where he is currently an Associate Professor. He is also a Senior Software Engineer at the Neusoft Corporation, Shenyang, China. In 2016, he was a Visiting Research Associate at King's College London. His research interests include D2D communication, location technology, network management, and signal and image processing.



Abdol Hamid Aghvami (M'89–SM'91–F'05) is currently a Professor of telecommunications engineering at King's College London, London, U.K. He joined the academic staff at King's College London in 1984. In 1989, he was promoted to a Reader, and in 1993 was promoted to a Professor in telecommunications engineering. He is the founder of the Centre for Telecommunications Research at King's College London. He was the Director of the centre from 1994 to 2014.

He carries out consulting work on digital radio communications systems for British and international companies. He has published more than 560 technical journal and conference papers, filed 30 patents, and given invited talks and courses the world over on various aspects of mobile radio communications. He was a Visiting Professor at the NTT Radio Communication Systems Laboratories in 1990, a Senior Research Fellow at the BT Laboratories during 1998–1999, and an Executive Advisor to the Wireless Facilities Inc., USA, during 1996–2002. He is the Chairman of Advanced Wireless Technology Group Ltd. He is also the Managing Director of the Wireless Multimedia Communications Ltd, London, U.K., his own consultancy company. He is also the Founder of the International Symposium on Personal Indoor and Mobile Radio Communications, a major yearly conference attracting some 1000 attendees.

Prof. Aghvami received the IEEE Technical Committee on Personal Communications Recognition Award in 2005 for his outstanding technical contributions to the communications field, and for his service to the scientific and engineering communities. He is a Fellow of the Royal Academy of Engineering and a Fellow of the IET. In 2009, he received a fellowship from the Wireless World Research Forum in recognition of his personal contributions to the wireless world, and for his research achievements as the Director of the Centre for Telecommunications Research, King's College London.



Arumugam Nallanathan (S'97–M'00–SM'05–F'17) has been a Professor of wireless communications in the School of Electronic Engineering and Computer Science, Queen Mary University of London, London, U.K., since September 2017. From December 2007 to August 2017, he was with the Department of Informatics, King's College London, where he was a Professor of wireless communications from April 2013 to August 2017. He was an Assistant Professor in the Department of Electrical and Computer Engineering, National University of Singapore, from

August 2000 to December 2007. He has published more than 350 technical papers in scientific journals and international conferences. His research interests include 5G wireless networks, Internet of Things, and molecular communications. He is a corecipient of the Best Paper Award presented at the IEEE International Conference on Communications 2016 and the IEEE International Conference on Ultra-Wideband 2007. He is an IEEE Distinguished Lecturer. He has been selected as a Web of Science (ISI) Highly Cited Researcher in 2016. He is an Editor of the IEEE TRANSACTIONS ON COMMUNICATIONS and IEEE TRANSACTIONS ON VEHICULAR TECHNOLOGY. He was an Editor of the IEEE TRANSACTIONS ON WIRELESS COMMUNICATIONS (2006–2011), the IEEE WIRELESS COMMUNICATIONS LETTERS, and the IEEE SIGNAL PROCESSING LETTERS.