Received: 20 January 2016 | Revised: 3 October 2016 | Accepted: 2 November 2016

DOI: 10.1002/dac.3247

WILEY

RESEARCH ARTICLE

Resource allocation in heterogeneous cooperative cognitive radio networks

Babatunde S. Awoyemi[1] | **Bodhaswar T. Maharaj**[1] | **Attahiru S. Alfa**[1,2]

[1]Faculty of Engineering Built Environment and IT, Electrical, Electronic and Computer Engineering, University of Pretoria, Pretoria, Gauteng, South Africa

[2]Electrical and Computer Engineering, University of Manitoba, Winnipeg, Manitoba, Canada

Correspondence

Babatunde S. Awoyemi, Faculty of Engineering Built Environment and IT, Electrical, Electronic and Computer Engineering, University of Pretoria, Pretoria, Gauteng, South Africa.

Email: awoyemibabatunde@gmail.com

Funding information

Broadband Wireless Multimedia Communications

Summary

In cognitive radio networks (CRNs), resources available for use are usually very limited. This is generally because of the tight constraints by which the CRN operate. Of all the constraints, the most critical one is the level of permissible interference to the primary users. Attempts to mitigate the limiting effects of this constraint, thus achieving higher productivity, are a current research focus, and in this work cooperative diversity is investigated as a promising solution. Cooperative diversity has the capability to achieve diversity gain for wireless networks. In the work, therefore, the possibility of and mechanism for achieving greater utility for the CRN when cooperative diversity is incorporated are studied. To accomplish this, a resource allocation model is developed and analyzed for the heterogeneous, cooperative CRN. In the model, during cooperation, a best relay is selected to assist the secondary users that have poor channel conditions. Overall, the cooperation makes it feasible for virtually all the secondary users to improve their transmission rates while still causing minimal harm to the primary users. The results show a marked improvement in the resource allocation performance of the CRN when cooperation is used in contrast to when the CRN operates only by direct communication.

KEYWORDS

cognitive radio networks, cooperative diversity, heterogeneous users, heuristic, linear and nonlinear programming, resource allocation

1 | INTRODUCTION

Recent measurements performed in various locations on the use of the allotted spectrum spaces by the networks to which they have been assigned have shown a rather high level of inefficiency in spectrum use.[1,2] Since the discovery of this underuse of the less-freely available radio-frequency spectrum with the current static allocation arrangement, efforts have been geared at devising a newer, better, and more efficient approach to sharing and using the spectrum. One such outstanding and bold attempt has been the introduction, description, experimentation, and gradual implementation of cognitive radio networks (CRNs).[3,4] Generally, with CRN, unlicensed cognitive or secondary users (SUs) are made to access and use the same spectrum space that has been preallocated to some licensed, original or primary users (PUs) of the spectrum, provided certain preconditions (such as the amount of permissible interference) already agreed upon that are not violated by the SUs.[5,6] This kind of arrangement promises higher spectrum efficiency and system throughput. With this huge promise, the CRN therefore seems to be a likely major game-changer for spectrum use in the telecommunications space. Therefore, much work is being dedicated to this wireless communication paradigm.

As the CRN develops, several issues with its design, implementation, application, and eventual productivity emerge. Researchers in academia and the industry would generally want to know whether it would be worth their while to invest in such a supposedly huge deal. One main issue has been the possibility of very low throughput that the SU network can actualize, especially when the conditions of the PU network are rather stringent. An example is the underlay CRN arrangement where the SUs are made to transmit over the entire PU spectrum, but at such low power that the PUs are not in any

https://doi.org/10.1002/dac.3247

way adversely affected by the SUs' transmission.[7] In such a case, it becomes really difficult for the CRN to achieve excellent results, especially if the PU network has high sensitivity and a low tolerance temperature. In such situations where the permissible interference to PUs is very low or where there are other very strict conditions under which the SUs must operate, the throughput or system capacity of the CRN becomes very limited.[8] It might therefore be arguable whether the CRN is in fact a worthwhile investment, unless such issues are adequately addressed. Several research projects on how to make the CRN achieve its ends even within such tight constraints are currently being undertaken.

In some earlier works of the authors of this paper, it has been shown that to achieve an optimal or near-optimal productivity level for the distribution or allocation of the rather scarce and/or limited resources of the CRN, it is best to allocate low data rates (or none at all) to subchannels where the interference gains to PUs are quite high.[9,10] This is understandable, as allocating high data rates to such subchannels would imply high transmit power by the SUs and consequently, high interference to the PUs because of the high interference gain. This smart move by the allocating algorithms of the SUs greatly increases the throughput of the CRN. However, the throughput achieved is usually still very limited, as there are some subchannels that, because of their high interference channel gains to PUs, are either completely unallocated or only allocated to transmit very low data rates. If the throughput of the CRN is to be improved further, it therefore becomes imperative to investigate how better channelling can be actualized so as to make higher data rates possible for virtually all the available subchannels, without causing too much interference to the PU network.

To address the problem raised above, in this work cooperative diversity is investigated as a promising solution. Cooperative diversity is a recent but comprehensive proposition on how to achieve better wireless channel conditioning by providing diversity gains among spatially dispersed users.[11] This is usually actualized by the cooperating users (called nodes or relays) forming a "virtual multiple input, multiple output" arrangement. These cooperating users use their antennas, as in conventional multiple input, multiple output systems, to assist each other in transmitting (or retransmitting) their data to a given destination user. Overall, a significant increase in the reliability and capability of the system is realized. Several cooperative diversity strategies have been developed and studied; some of which are store-and-forward, amplify-and-forward, decode-and-forward, and coded cooperation.[12] Similarly, cooperative diversity has been classified in the number of cooperators selected, or whether or not the cooperation happens opportunistically or incrementally.[13] The important thing is that at the destination, a much better signal quality is achieved and network capacity is significantly improved. This work therefore investigates and develops how cooperative diversity can be used in improving the effective capacity of the CRN. Essentially, by bringing cooperative diversity into the CRN, the work reveals how the limiting effects of the interference constraint in its resource allocation (RA) problems can be adequately mitigated, thus achieving much better productivity for the network.

2 | RELATED LITERATURE

The concept of RA in CRN is no longer entirely new. Several research projects have been undertaken in this regard, and a review of relevant ones is performed in this section. Resource allocation in CRN actually deals with devising and describing mechanisms for assigning resources (frequency spectrum, transmit power, bandwidth, time slot, modulation scheme, etc) fairly and optimally to all users so that the highest possible productivity level is achieved. A number of RA problems for underlay CRN have been identified, and attempts at solving them (both optimally and sub-optimally) have been investigated. In the study of Wang et al,[14] for example, an approach for obtaining optimal solutions for RA problems in an underlay CRN is developed. A centralized algorithm, which makes use of the Lagrangian duality, is first used to solve the problem. Thereafter, a distributed algorithm that uses dual decomposition is developed to solve the same problem. Both algorithms produce near-optimal solutions. Other similar works that have developed RA models for underlay CRN can be found in several studies.[15–17] The major challenge with underlay, as can be seen from the above works and similar ones, is always the low level of utility that is achievable in its network owing to the stringent conditions of the PU and the power limitations of the secondary network.

Resource allocation problems in overlay CRN have been studied in a number of studies.[18–21] In overlay CRN generally, the SUs use free and/or available spectrum (spectrum holes) of the PUs for transmission. Both subchannel and power control were jointly considered in the study of Guo and Huang[18] with the intent of maximizing the throughput of the CRN. Expanding on this, the works of Xie et al[19] and Wang et al[20] extended the work to make room for possible imperfections in the CRN sensing capabilities of PU and developed models with some robustness to accommodate such imperfections. In the work of Chen et al,[21] the problem is studied even further to include quality-of-service provisioning. As an observation, a major problem with overlay networks, as can be identified in these and similar works, is PU interference and possible disruption in service delivery of the SU in the CRN.

As a means of addressing some of the limitations of the underlay and overlay arrangements, recent attempts at introducing user cooperation into RA in CRN have been made. A number of studies[5,22–25] have all developed models that describe possible cooperation between SUs in a CRN to help achieve a higher utility level. In two studies[5,22] relays using decode-and-forward protocol are made to assist the SUs of the CRN. For the optimization problem that has been developed to be solvable, the subchannels are first assigned to the SUs on the basis of their channel gains and possible interference to PUs. Thereafter, power is allocated to each subchannel. A similar model is developed in the work of Adian and Aghaeinia,[23] where a decode-and-forward cooperative relay network is used to assist the SUs, thereby improving throughput. The nonconvex optimization problem that was developed is solved by first dualizing, then decomposing into relay assignment and power allocation. A primary decomposition method is also used in the work of Du et al,[24] after the power allocation problem in the developed model has been formulated. The sum rate of both PUs and SUs is jointly maximized in the work of Pischella and Le Ruyet,[25] while the SUs cooperate to transmit their signals. To achieve a result close to optimal, subchannels are first allocated to the SUs; thereafter, power is assigned to each SU and PU iteratively. While the above reviewed publications have incorporated some kind of cooperation, this work differs from them all in that the cooperative diversity approach developed in this work is targeted directly at addressing the problem of PU interference. Thus, the interference problem is first taken care of by the cooperation model even before the RA to SUs is performed.

More specifically, in this work, through SU cooperation, the impact of the interference to PUs is mitigated, thus achieving greater throughput for the heterogeneous CRN. The heterogeneity in the CRN has been approached from 2 perspectives. Firstly, the channels are assumed to be heterogeneous, meaning that the available channels for the CRN do not all have the same characteristics. To capture the differing effects of channel heterogeneity, the network has been developed using an orthogonal frequency division multiple access (OFDMA) platform. With the OFDMA, the system can dynamically and

TABLE 1 List of terms and symbol notations

$K;k$	Total number of heterogeneous secondary users; k is used to identify a particular user
K_1	Number of category 1 secondary users; number of category 2 users is thus $K - K_1$
$N;n$	Number of available OFDMA subchannels; n is used to identify a particular subchannel
L	Number of primary users
SU;*SUs*	Secondary user; Secondary users
PU;*PUs*	Primary user; Primary users
SUBS	Secondary user base station
SSU	Source secondary user
CSU	Cooperative secondary user
D	Destination terminal
OFDMA	Orthogonal frequency division multiple access
BPSK	Binary phase shift keying
QAM	Quadrature amplitude modulation
$H_{k,n}^{c}$	Channel gain between SUBS and SU at the kth SU over the nth subchannel
$H_{k,n}^{s}$	Channel gain from SSU to CSU at the kth SU over the nth subchannel
$H_{k,n}^{r}$	Channel gain from CSU to D at the kth SU over the nth subchannel
$P_{k,n}^{s}$	Transmit power from SSU to CSU at the kth SU over the nth subchannel
$P_{k,n}^{r}$	Power from CSU to D at the kth SU over the nth subchannel
$c_{k,n}$	Data rate at the kth SU over the nth subchannel
$c_{k,n}^{s}$	Data rate from SSU to CSU at the kth SU over the nth subchannel
$c_{k,n}^{r}$	Data rate from CSU to D at the kth SU over the nth subchannel
$c_{k,n,D}$	Data rate at the kth SU over the nth subchannel for direct transmission
$c_{k,n,C}$	Data rate at the kth SU over the nth subchannel when cooperation is used
$P_{k,n,D}$	Transmit power at the kth SU over the nth subchannel for direct transmission
$P_{k,n,C}$	Transmit power at the kth SU over the nth subchannel when cooperation is used
$P_{\max}$	Total transmit power at SUBS
$\boldsymbol{x}$	Bit allocation vector
$\boldsymbol{x_I};\boldsymbol{x_{II}}$	Bit allocation vector for category 1 SU; bit allocation vector for category 2 SU
$\boldsymbol{b}$	Modulation order vector
$\boldsymbol{b_I};\boldsymbol{b_{II}}$	Modulation order vector for category 1 SU; modulation order vector for category 2 U
$\boldsymbol{p}$	Power transmission vector
$\boldsymbol{p_D}$	Power transmission vector for direct communication
$\boldsymbol{p_C}$	Power transmission vector for cooperative communication
R_K	Minimum rate demand of category 1 SU
γ_K	Proportional rate constraint for category 2 SU

optimally use different portions of the spectrum (heterogeneous channels) for different users at the same time. Secondly, the SUs in the network are heterogeneous. This means that the users do have different priorities, requirements, or demands, thus necessitating that the SUs be categorized. Users in each category are then serviced on the basis of their priority and/or their varying demands. During cooperation, the selection scheme used is the single-best-relay selection scheme used alongside the store-and-forward cooperative diversity technique. With this scheme, a best relay among the SUs is selected as the cooperator, which, during cooperation receives data from the source user and transmits to the destination. Overall, the heterogeneous cooperative CRN model, as developed and studied, reveals that much greater productivity is achievable by the CRN when its users cooperate. For the purpose of clarity, the most important contributions of this work are highlighted:

- Investigating the use of cooperative diversity as a means of mitigating the limiting effects of interference to PUs in the RA problem of the CRN, thereby making much better productivity possible for the CRN.
- Developing and analyzing methods for obtaining solutions (optimal and suboptimal) to the RA problem in heterogeneous, cooperative CRN. The solutions are obtained through a thorough study of the structure of the problem.

The remainder of this work is organized as follows: Section 3 describes the system model, Section 4 deals with the problem formulation and reformulation to obtain optimal solutions, Section 5 presents the heuristic developed to reduce the computational complexity, Section 6 presents and discusses the results, and finally, Section 7 gives the concluding remarks.

Table 1 contains the various notations that have been applied in the work.

3 | SYSTEM MODEL

The underlay, heterogeneous, cooperative CRN model developed consists of K heterogeneous SUs and L PUs, all located within the coverage range of the secondary user base station (SUBS). N OFDMA heterogeneous subchannels are available for the SUs, to which any of them can be assigned. The K heterogeneous SUs have different demands and priorities. These SUs are thus categorized as K_1: SUs with minimum rate guarantee and $(K - K_1)$: SUs with best effort service. Users in category 1 have a minimum rate requirement, and their demands are met first. They therefore have the higher priority. Users in category 2 are best effort users; hence, the remaining resources are proportionally shared among them (using a proportional rate constraint). All subchannels are also modelled to be in slow fading. During transmission, the network decides whether to use direct or cooperative communication on the basis of its prevalent condition. When cooperation occurs, the direct link is ignored in the model because of the high interference to PUs that will be introduced if the direct link is used, which would potentially limit the entire CRN resourcefulness.

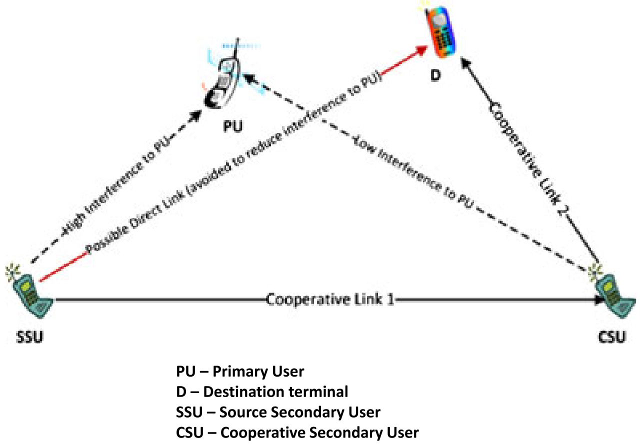

FIGURE 1 System model of the cooperative cognitive radio network

Figure 1 shows the network when cooperation is to be used. The cooperative scheme used in this work is the incremental, single-best-relay selection cooperative diversity scheme. The scheme being "incremental" means that cooperation is strictly restricted to only when it is needed and single-best-relay selective cooperation means that only one best relay is selected in such instances. The reason for this cooperative diversity choice is to ensure that the model is feasible, as well as to minimize overhead. The SU that requires cooperation, as it intends to communicate with a destination terminal (D), is referred to as the source secondary user (SSU). This SSU has a potentially high interference channel gain to the PU on the direct link and would therefore either not have been allocated subchannels at all or have been given only a few subchannels to transmit at low data rates if direct communication alone had been considered. To help mitigate this limitation, the SSU, during cooperation, selects a cooperating secondary user (CSU) with good channel quality (SSU to CSU as well as CSU to D) and poor interference channel gain to PUs. Through this cooperation, the effects of the poor channel condition are mitigated. Next, a description of how the best relay (i.e., the CSU) is selected from among the other users is presented.

In the model developed, the CRN operates a centralized control system with the SUBS as its communicating hub. Communication between the SU and the SUBS is assumed to be error free. All the SUs estimate and communicate their channel conditions as well as their interference gains to PUs to the SUBS. Any of the SUs can be the potential CSU for any other one. The SUBS determines which of the SUs needs a cooperator and chooses and contacts the best of the other SUs, which is then assigned as its CSU. It is assumed that at the moment of cooperation, the SU used as the cooperating SU is idle and has no data of its own to transmit. This information, along with the estimated channel condition and PU interference, is relayed on the control channel to the SUBS by each SU. The choice of a CSU is usually based on the SU with the best channel condition and the least interference gain to the PU. The chosen CSU is thereafter relayed by the SUBS to the SSU, while the other SUs, which have not been contacted, carry on with their normal transmission (or simply maintain their idle state, as the case may be). The SSU transmits to CSU, which then forwards the transmitted data to the destination terminal D over the assigned subchannels. The transmission is in 2 time slots. In the first time slot the SSU transmits to the CSU, and in the second time slot the CSU transmits to D. The combined channel condition of the SSU and the CSU is obtained as follows:

Denote $H^s_{k,n}$ as the channel gain between the SSU and the kth SU, used as the CSU, at the nth subchannel and $H^r_{k,n}$ as the channel gain between the CSU and the destination terminal D over the nth subchannel. The SSU transmits signals to the kth relay on the nth subchannel with power $P^s_{k,n}$ in the first slot, while the kth relay (CSU) transmits signals to D on the nth subchannel with power $P^r_{k,n}$ in the second slot. Thus, the data rate of each transmission slot is given in the work of Ge and Wang[5] as

$$\begin{aligned} c^s_{k,n} &= \log\left(1+\frac{P^s_{k,n}|H^s_{k,n}|^2}{\sigma_r^2+\sum_{l=1}^{L}J^l_{k,n}}\right),\\ c^r_{k,n} &= \log\left(1+\frac{P^r_{k,n}|H^r_{k,n}|^2}{\sigma^2+\sum_{l=1}^{L}J^l_{n}}\right), \end{aligned} \tag{1}$$

where σ_r^2 and σ^2 are the variances of the noise at the kth relay (CSU) and D, respectively. Similarly, the interference to the kth relay and that to D on the nth subchannel by the lth PU's signal is denoted by $J^l_{k,n}$ and J^l_n. This is regarded as noise and measured by the receivers of the CSU and D. It is important to state that the effective data rate during cooperative transmission is actually limited by the minimum of the 2 hops:

$$c_{k,n,C} = \min\left(c^s_{k,n}, c^r_{k,n}\right). \tag{2}$$

If no cooperation is needed, transmission is directly from the SU to D over the assigned subchannels, and the data rate is simply $c_{k,n,D}$. This data rate c for each subchannel, using either direct or cooperative transmission, is dependent on the modulation scheme to which the subchannel has been assigned. In this work, 4 modulation schemes, which are binary phase shift keying (BPSK), 4–quadrature amplitude modulation (QAM), 16-QAM, and 64-QAM, are considered. The modulation schemes transmit $c = 1,2,4$, and 6 bits per OFDMA symbol, respectively. For a given bit error rate ρ value to be met at the receiver end of communication, the minimum amount of power required over any given subchannel is dependent on the modulation scheme used. The minimum power for BPSK modulation is given as $P(c,\rho) = N_\phi[c \times erfc^{-1}(2\rho)]^2$ (where $c = 1$), while for the M-ary QAM, the minimum power is given as $P(c,\rho) = \frac{2(2^c-1)N_\phi}{3}\left[erfc^{-1}\left(\frac{c\rho\sqrt{2^c}}{2\left(\sqrt{2^c-1}\right)}\right)\right]^2$ ($c = 2,4,$ or 6 for 4-QAM, 16-QAM, and 64-QAM, respectively) where $erfc(x) = \left(\frac{1}{\sqrt{2\pi}}\right)\int_x^\infty e^{\frac{-t^2}{2}}dt$ is the complementary error function, $\pi = (22/7)$, and N_ϕ is the single-sided noise power spectral density. In this work, N_ϕ assumes the same value for all subchannels. For a given value of ρ, as the number of bits assigned to a subchannel increases, the transmit power also increases, albeit nonlinearly. The minimum power $P_{k,n}(c_{k,n},\rho)$ required at the kth SU over the nth subchannel to transmit $c_{k,n}$ bits is obtained by dividing the power $P(c_{k,n},\rho)$ of that user k on the nth subchannel by the channel gain $H^c_{k,n}$ between the SUBS and the user k over that subchannel n. This is thus given as

$$P_{k,n}(c_{k,n},\rho) = \frac{P(c_{k,n},\rho)}{H^c_{k,n}}. \tag{3}$$

4 | RESOURCE ALLOCATION PROBLEM FORMULATION

Let the minimum data rate assigned to user k in category 1 be R_k and the normalized proportional fairness factor for each SU in category 2 be γ_k with data rate R_i indicating the rate for the element i. The total power on the nth subchannel is represented as $\Phi_n = \Sigma_{k=1}^{K}P_{k,n}$ with $P_{k,n}$ being the transmit power of user k over the nth subchannel ($P_{k,n,C}$ is the power used when cooperation is used and $P_{k,n,D}$ is the transmit power for direct transmission). Also let the interference power gain matrix between the SUBS and the available PU be represented as $\boldsymbol{H^p}\in R^{L\times N}$. The vector $\boldsymbol{H^p}_{l,n}$ therefore denotes the subchannel interference power gain between the SUBS and PU l over subchannel n ($\boldsymbol{H^p}_{l,n,D}$ is the gain matrix when direct transmission is used and $\boldsymbol{H^p}_{l,n,C}$ is the gain matrix when cooperation is used). The maximum permissible level of interference to the lth PU from all the transmitting SUs is represented as ε_l while $P_{\max}$ denotes the maximum power available for transmission at the SUBS. Also let $X_{k,n,D}$ be a binary (0,1) variable used to limit each subchannel to either direct communication or cooperative communication (since each subchannel can only transmit using either of the 2, but not both). The RA problem for the heterogeneous cooperative CRN is therefore formulated as $z = \max$

$$\begin{aligned} &\sum_{n=1}^{N}\left(\sum_{k=1}^{K_1}\left[X_{k,n,D}c_{k,n,D}+(1-X_{k,n,D})c_{k,n,C}\right]\right.\\ &\qquad\left.+\sum_{k=K_1+1}^{K}\left[X_{k,n,D}c_{k,n,D}+(1-X_{k,n,D})c_{k,n,C}\right]\right);\\ &c_{k,n,D}, c_{k,n,C} \in \{0,1,2,4,6\} \end{aligned} \tag{4}$$

subject to

$$\sum_{n=1}^{N}(c_{k,n,D}+c_{k,n,C}) \geqslant R_k;\ k = 1,2,\ldots,K_1 \tag{5}$$

$$\frac{R_k}{\sum\limits_{i=K_1+1}^{K} R_i} = \gamma_k;\ k = K_1+1, K_1+2, \dots, K \tag{6}$$

$$\sum_{n=1}^{N}\left(\sum_{k=1}^{K}\left[X_{k,n,D}P_{k,n,D} + (1-X_{k,n,D})P_{k,n,C}\right]\right) \leqslant P_{\max} \tag{7}$$

$$\sum_{n=1}^{N}\Phi_n \boldsymbol{H^p}_{l,n,D} \leqslant \varepsilon_l;\ l = 1, 2, \dots, L \tag{8}$$

$$\sum_{n=1}^{N}\Phi_n \boldsymbol{H^p}_{l,n,C} \leqslant \varepsilon_l;\ l = 1, 2, \dots, L \tag{9}$$

$$\begin{gathered} c_{k,n,D} = 0 \ \mathit{if}\ c_{k',n,D} \neq 0,\ c_{k,n,C} = 0 \ \mathit{if}\ c_{k',n,C} \neq 0, \\ \forall k' \neq k;\ k = 1, 2, \dots, K \end{gathered} \tag{10}$$

$$\begin{aligned} X_{k,n,D} \in \{0,1\},\ X_{k,n,D} &= 1 \ \mathit{if}\ c_{k,n,D} \neq 0 \\ X_{k,n,D} &= 0 \ \mathit{otherwise}. \end{aligned} \tag{11}$$

The objective function 4 captures the throughput or total data rate that all the SUs in the network for both direct and cooperative transmission can realize. Constraint 5 is the minimum data rate constraint, indicating that for each SU in category 1, their minimum data rate requirement must be met. Constraint 6 is the best effort service constraint where, to determine how the remaining resources are to be assigned to each user in category 2, a proportional fairness factor is being used. Constraint 7 is used to limit the total transmit power of all the SUs during direct and cooperative transmission to the SUBS's maximum transmit power available. Constraint 8 shows that when the SUs are transmitting, the amount of interference permissible to each PU during direct transmission must not be greater than the predetermined threshold limit. Constraint 9 is similar to constraint 8, but captures the interference constraint during cooperative transmission. To restrict the allocation of subchannels to only one user per subchannel for each user, the mutually exclusive constraint in Equation 10 is given. The constraint shows that once subchannel n has been assigned to a user $k^{'} \neq k$, the data rate for subchannel n must therefore be 0 for any other user k. The equation in constraint 6 can equally be expressed as

$$R_k = \gamma_k \times \sum_{i=K_1+1}^{K} R_i,$$

where $\sum\limits_{i=K_1+1}^{K} R_i$ is the sum value of all the data rates for all SUs in category 2. Representing $\gamma_k \times \sum\limits_{i=K_1+1}^{K} R_i$ by $\tilde{\gamma}_k$, Equation 6 becomes

$$R_{K_1+1} : R_{K_1+2} : \dots : R_K = \tilde{\gamma}_{K_1+1} : \tilde{\gamma}_{K_1+2} : \dots : \tilde{\gamma}_K. \tag{12}$$

The above formulation of the RA problem in Equations 4 to 11 is not a linear programming problem because the power constraint in Equation 7 is not linear. However, by careful consideration of the problem structure, the problem has been reformulated as an integer linear programming (ILP) problem. The reformulated problem can easily be solved by using any of the classical optimization techniques. The branch-and-bound (BnB) approach is used in this work. The reformulation is achieved as described below:

Define $\boldsymbol{x_I}$ as the bit allocation vector for all subchannels assigned to all users in category 1 (both for direct and cooperative transmission, hence, $\boldsymbol{x_I} = (\boldsymbol{x_{I,D}} + \boldsymbol{x_{I,C}})$), and also define $\boldsymbol{x_{II}}$ as the bit allocation vector for all subchannels assigned to all users in category 2 (for direct and cooperative transmission so that $\boldsymbol{x_{II}} = (\boldsymbol{x_{II,D}} + \boldsymbol{x_{II,C}})$). $\boldsymbol{x_I}$ and $\boldsymbol{x_{II}}$ are given as

$$\boldsymbol{x_I} = \left[\left(\boldsymbol{x}^1_{I,N}\right)^T \left(\boldsymbol{x}^2_{I,N}\right)^T \ \dots\ \left(\boldsymbol{x}^N_{I,N}\right)^T\right]^T \in \{0,1\}^{NK_1C\times 1} \tag{13}$$

$$\boldsymbol{x_{II}} = \left[\left(\boldsymbol{x}^1_{II,N}\right)^T \left(\boldsymbol{x}^2_{II,N}\right)^T \ \dots\ \left(\boldsymbol{x}^N_{II,N}\right)^T\right]^T \in \{0,1\}^{N(K-K_1)C\times 1}, \tag{14}$$

where $\boldsymbol{x}^n_{I,N} = [x^T_{I,1,n}\ x^T_{I,2,n}\ \dots\ x^T_{I,K,n}]^T \in \{0,1\}^{KC\times 1}$ indicates that subchannel n has been assigned to a category 1 SU with $\boldsymbol{x}_{I,k,n} = [x_{k,n,1}\ x_{k,n,2}\ \dots\ x_{k,n,M}\ ]^T \in \{0,1\}^{C\times 1}$; $n = 1, \dots, N$; $k = 1, \dots, K$; M indicates the overall number of modulation schemes being used (for this work, $M = 4$). The implication is that $\boldsymbol{x}_{I,k,n} = [x_{k,n,1}x_{k,n,2}x_{k,n,3}x_{k,n,4}]^T$. Similar explanations apply to $\boldsymbol{x_{II}}$. The combined bit allocation vector $\boldsymbol{x} = \boldsymbol{x_I} + \boldsymbol{x_{II}}$. Because of the mutually exclusive constraint, $\boldsymbol{x}^n_{I,N}$ and $\boldsymbol{x}^n_{II,N}$ can be any of the vectors $\{[00\cdots0]^T,[10\cdots0]^T,[01\cdots0]^T,\cdots,[00\cdots1]^T\}$. Hence, only one component in $\boldsymbol{x}^n_{I,N}$ is 1, while the other components are all 0s (same applies for $\boldsymbol{x}^n_{II,N}$). If $x_{k,n,c}$ is 1, it means that subchannel n has been assigned to user k to transmit c bits per symbol. If $\boldsymbol{x}^n_{I,N}$ (or $\boldsymbol{x}^n_{II,N}$) has all its components as 0s, subchannel n is not being assigned to any user.

For the 2 user categories, define the modulation order vectors $\boldsymbol{b_I}$ and $\boldsymbol{b_{II}}$ as

$$\boldsymbol{b_I} = \left[\left(\boldsymbol{b}^1_{I,N}\right)^T \left(\boldsymbol{b}^2_{I,N}\right)^T \cdots \left(\boldsymbol{b}^N_{I,N}\right)^T\right]^T \in \mathbb{Z}^{NK_1C\times 1} \tag{15}$$

$$\boldsymbol{b_{II}} = \left[\left(\boldsymbol{b}^1_{II,N}\right)^T \left(\boldsymbol{b}^2_{II,N}\right)^T \cdots \left(\boldsymbol{b}^N_{II,N}\right)^T\right]^T \in \mathbb{Z}^{N(K-K_1)C\times 1}, \tag{16}$$

where $\boldsymbol{b}^n_{I,N} = \left[b^T_{I,1,n}\ b^T_{I,2,n}\ \dots\ b^T_{I,K,n}\right]^T \in \mathbb{Z}^{K_1C\times 1}$ and $\boldsymbol{b}_{I,k,n} = [b_{k,n,1}\ b_{k,n,2}\ \dots\ b_{k,n,C}\ ]^T \in \mathbb{Z}^{C\times 1}$. Similar explanations also apply to $\boldsymbol{b_{II}}$. Having considered only 4 modulation schemes (i.e., BPSK, 4-QAM, 16-QAM, and 64-QAM), $\boldsymbol{b}_{1,k,n}$=[1234]^T (the same applies to $\boldsymbol{b}^n_{II,N}$). For the 2 categories of SUs, data rate matrices $\boldsymbol{B_i} \in \mathbb{Z}^{K_1 \times NK_1C}$ and $\boldsymbol{B_j} \in \mathbb{Z}^{(K-K_1)\times N(K-K_1)C}$ are defined respectively as

$$\boldsymbol{B_i} = \begin{bmatrix} b_1 & b_1 & \cdots & b_1 \\ b_2 & b_2 & \cdots & b_2 \\ \vdots & \vdots & \ddots & \vdots \\ b_{K_1} & b_{K_1} & \cdots & b_{K_1} \end{bmatrix}, \ \boldsymbol{B_i} \in \mathbb{Z}^{K_1 \times NK_1C}$$
$$\left\{ \begin{array}{l} b_1 = [b^T \ \ 0_C^T \ \cdots \ 0_C^T] \ \in \mathbb{Z}^{1\times K_1C} \\ b_2 = [0_C^T \ \ b^T \ \cdots \ 0_C^T] \ \in \mathbb{Z}^{1\times K_1C} \\ \quad \vdots \qquad\quad \vdots \quad \ddots \qquad\quad \vdots \\ b_{K_1} = [0_C^T \ 0_C^T \ \cdots \ b^T] \ \in \mathbb{Z}^{1\times K_1C} \end{array} \right\} \tag{17}$$

$$\boldsymbol{B_j} = \begin{bmatrix} b_{K_1+1} & b_{K_1+1} & \cdots & b_{K_1+1} \\ b_{K_1+2} & b_{K_1+2} & \cdots & b_{K_1+2} \\ \vdots & \vdots & \ddots & \vdots \\ b_K & b_K & \cdots & b_K \end{bmatrix}, \ \boldsymbol{B_j} \in \mathbb{Z}^{(K-K_1)\times N(K-K_1)C}$$
$$\left\{ \begin{array}{l} b_{K_1+1} = [b^T \ \ 0_C^T \ \cdots \ 0_C^T] \ \in \mathbb{Z}^{1\times (K-K_1)C} \\ b_{K_1+2} = [0_C^T \ \ b^T \ \cdots \ 0_C^T] \ \in \mathbb{Z}^{1\times (K-K_1)C} \\ \quad \vdots \qquad\qquad \vdots \quad \ddots \qquad\qquad \vdots \\ b_K = [0_C^T \ \ 0_C^T \ \cdots \ b^T] \ \in \mathbb{Z}^{1\times (K-K_1)C} \end{array} \right\}. \tag{18}$$

Equation 4, which gives the total data rate achievable by the network, can thus be written as $\max_x \left[(\boldsymbol{b_I})^T \boldsymbol{x_I} + (\boldsymbol{b_{II}})^T \boldsymbol{x_{II}}\right]$. Define $\boldsymbol{R}_k \triangleq [R_1 \ R_2 \ \ldots \ R_{K_1}]^T \in \mathbb{R}^{K_1 \times 1}$ and $\tilde{\boldsymbol{\gamma}}_k \triangleq [\tilde{\gamma}_{K_1+1} \ \tilde{\gamma}_{K_1+2} \ \ldots \ \tilde{\gamma}_K]^T \in \mathbb{R}^{(K-K_1)\times 1}$, the constraint of Equation 5, which explains the data rate per user for category 1 SUs, can be written as $\boldsymbol{B_i x_I} \geqslant \boldsymbol{R}_k$, while the data rate constraint for category 2 SU given in Equation 6 can be written as $\boldsymbol{B_j x_{II}} = \tilde{\boldsymbol{\gamma}}_k$.

Next, a power transmission vector $\boldsymbol{p}$ is defined as

$$\boldsymbol{p} = \left[\left(\boldsymbol{p}_N^1\right)^T \left(\boldsymbol{p}_N^2\right)^T \ \ldots \ \left(\boldsymbol{p}_N^N\right)^T\right]^T \in \mathbb{R}^{NKC\times 1}, \tag{19}$$

where $\boldsymbol{p}_N^n = \left[\boldsymbol{p}_{1,n}^T \, \boldsymbol{p}_{2,n}^T \ \ldots \ \boldsymbol{p}_{K,n}^T\right]^T \in \mathbb{R}^{KC\times 1}$ and $\boldsymbol{p}_{k,n} = [p_{k,n,1} \ p_{k,n,2} \ \ldots \ p_{k,n,C} \ ]^T \in \mathbb{R}^{C\times 1}$; $p_{k,n,c}$ is the power required to transmit c bits over subchannel n for user k. Equation 7, which describes the power constraint, can now be written as $\boldsymbol{p}^T \boldsymbol{x} \leqslant P_{\max}$. Given that the transmit power is the summation of the power used for both direct and cooperation transmission, $\boldsymbol{p} = \boldsymbol{p_D} + \boldsymbol{p_C}$, where $\boldsymbol{p_D}$ and $\boldsymbol{p_C}$ are the transmit power vectors during direct and cooperation transmission, respectively. The power constraint therefore becomes $(\boldsymbol{p_D} + \boldsymbol{p_C})^T \boldsymbol{x} \leqslant P_{\max}$.

To write Equation 8, the interference power constraint (which is also applicable to Equation 9), of the bit allocation vector $\boldsymbol{x}$, define a matrix $\boldsymbol{A}\in\{0,1\}$xx-xx$^{N \times NKC}$ as below

$$\boldsymbol{A} = \begin{bmatrix} 1_{KC}^T & 0_{KC}^T & \cdots & 0_{KC}^T \\ 0_{KC}^T & 1_{KC}^T & \cdots & 0_{KC}^T \\ \vdots & \vdots & \ddots & \vdots \\ 0_{KC}^T & 0_{KC}^T & \cdots & 1_{KC}^T \end{bmatrix}, \ \boldsymbol{A} \in \{0,1\}^{N\times NKC} \tag{20}$$

$$1_{KC} = \begin{bmatrix} 1 \\ 1 \\ \vdots \\ 1 \end{bmatrix} \in \{1\}^{KC\times 1}, \qquad 0_{KC} = \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 0 \end{bmatrix} \in \{0\}^{KC\times 1}.$$

Let $\boldsymbol{p} \odot \boldsymbol{x}$ be the Schur-Hadamard (or entry-wise) product of $\boldsymbol{p}$ and $\boldsymbol{x}$, $\boldsymbol{A}(\boldsymbol{p} \odot \boldsymbol{x})$ will therefore be that $N \times 1$ vector whose nth element gives the total power the nth subchannel uses while transmitting. By defining $\boldsymbol{\varepsilon}_l \triangleq [\varepsilon_1 \ \varepsilon_2 \ \ldots \ \varepsilon_L]^T \in \mathbb{R}^{L\times 1}$, Equation 8, which describes the interference power constraint for the direct transmission, can then be written as

$$\left[\boldsymbol{H}^p_{l,n,D}(\boldsymbol{A}(\boldsymbol{P_D} \odot \boldsymbol{x}))\right] \leqslant \boldsymbol{\varepsilon}_l. \tag{21}$$

Likewise, the constraint in Equation 9, which describes the interference power for the cooperative transmission, can be written as

$$[\boldsymbol{H}^p_{l,n,C}(\boldsymbol{A}(\boldsymbol{P_C} \odot \boldsymbol{x}))] \leqslant \boldsymbol{\varepsilon}_l. \tag{22}$$

Thus, the RA problem for the modelled heterogeneous cooperative CRN described in Equations 4 to 11 can be described in the ILP form as given by the following equations:

$$z^* = \max_x \left[(\boldsymbol{b_I})^T \boldsymbol{x_I} + (\boldsymbol{b_{II}})^T \boldsymbol{x_{II}}\right] \tag{23}$$

subject to

$$\boldsymbol{B_i x_I} \geqslant \boldsymbol{R}_k; \ k = 1, 2, \ \ldots \ , K_1 \tag{24}$$

$$\boldsymbol{B_j x_{II}} = \tilde{\boldsymbol{\gamma}}_k; \ k = K_1 + 1, K_1 + 2, \ \ldots \ , K \tag{25}$$

$$(\boldsymbol{p_D} + \boldsymbol{p_C})^T \boldsymbol{x} \leqslant P_{\max} \tag{26}$$

$$\left[\boldsymbol{H}^p_{l,n,D}(\boldsymbol{A}(\boldsymbol{p_D} \odot \boldsymbol{x}))\right] \leqslant \boldsymbol{\varepsilon}_l \tag{27}$$

$$\left[\boldsymbol{H}^p_{l,n,C}(\boldsymbol{A}(\boldsymbol{p_C} \odot \boldsymbol{x}))\right] \leqslant \boldsymbol{\varepsilon}_l \tag{28}$$

$$\boldsymbol{0}_N \leqslant \boldsymbol{Ax} \leqslant \boldsymbol{1}_N \tag{29}$$

$$\boldsymbol{x_I}, \boldsymbol{x_{II}}, \boldsymbol{x} \in \{0, 1\}. \tag{30}$$

The formulation above is an ILP problem in which, in this work, the BnB approach has been used to obtain solutions. Branch-and-bound optimization is a very useful and well-developed technique for solving such problems. However, although the BnB approach may yield optimal solutions, it can be very poor at finding such solutions timeously, especially in large networks. It is imperative to investigate a much faster approach to achieve a near-optimal solution. This can be achieved by developing a heuristic for solving the problem, as provided in the next section.

5 | ITERATIVE-BASED HEURISTIC

In this section, a fast, iterative-based heuristic is developed to solve the ILP problem. Even though the results obtained are only near-optimal, the heuristic gives a good trade-off between optimality and complexity, especially for large systems. The approach used in the heuristic is an extension of the work presented by Rahulamathavan et al.[26] The algorithm involves 2 steps:

- subchannel allocation
- iterative bit and power allocation.

5.1 | Subchannel allocation

In performing the subchannel allocation for the different categories of SUs, the constraint $\boldsymbol{x} \in [0,1]$ is integer-relaxed such that the constraint becomes

$$0 \leqslant \boldsymbol{x} \leqslant 1. \tag{31}$$

In other words, $\boldsymbol{x}$ is allowed to take any value from 0 to 1 and is not necessarily restricted to either 0 or 1. All the other parts of the formulation are unchanged. By solving this integer-relaxed formulation at the first iteration, the values of $\boldsymbol{x}$ are obtained, which imply that the subchannels have been allocated to the various users. The data rate for the kth SU at the nth subchannel becomes $\left(\boldsymbol{b}_{k,n}^T \boldsymbol{x}_{k,n}\right)$. It is not impossible that there is a user $m \neq k$ whose data rate $\left(\boldsymbol{b}_{m,n}^T \boldsymbol{x}_{m,n}\right)$ on subchannel n could be larger than user k's data rate on n. It would therefore be more appropriate to give subchannel n to user m rather than k. Hence, subchannel n is only allocated to user k after ascertaining that $\left(\boldsymbol{b}_{k,n}^T \boldsymbol{x}_{k,n}\right) \geqslant \left(\boldsymbol{b}_{m,n}^T \boldsymbol{x}_{m,n}\right) \forall m \neq k$. Clearly then, each subchannel is allocated to the SU that has the highest achievable data rate over that subchannel. It is important to realize too that once the suchannels have been allocated to the different SUs using the above criterion at the first iteration, the dimension of $\boldsymbol{x}$ reduces from its initial value of $\boldsymbol{x} \in [0,1]^{KNC \times 1}$ to the smaller value of $\boldsymbol{x} \in [0,1]^{NC \times 1}$.

5.2 | Iterative bit and power allocation

Once the subchannels have been assigned to the various SUs, it then remains to determine how many bits (and by extension, what modulation scheme) and what power can be associated with each subchannel. This is performed in an iterative manner. The algorithm starts by assigning a possible number of bits (rather unambitiously) to each user, and then it determines the power used, checks that other constraints are not violated, determines if there is some excess power left, and increases the bits gradually where possible. Then it checks the power again, and the whole iterative process is repeated until no further improvement on bit allocation is possible.

The whole optimization process occurs in a number of iterations, say y. In general, therefore, the following optimization problem has to be solved at the yth iteration step:

$$\max_{\boldsymbol{x}^y} \left[\left(\boldsymbol{b}_I^y\right)^T \boldsymbol{x}_I^y + \left(\boldsymbol{b}_{II}^y\right)^T \boldsymbol{x}_{II}^y \right] \tag{32}$$

subject to

$$\boldsymbol{B}_i \boldsymbol{x}_{\boldsymbol{I}}^{\boldsymbol{y}} \geqslant [\boldsymbol{R}_k - \boldsymbol{f}^{(y-1)}]^+; \; k = 1, 2, \ldots, K_1 \tag{33}$$

$$\boldsymbol{B}_j \boldsymbol{x}_{\boldsymbol{II}}^{\boldsymbol{y}} = [\tilde{\boldsymbol{\gamma}}_k - \boldsymbol{g}^{(y-1)}]^+; \; k = K_1 + 1, K_1 + 2, \ldots, K \tag{34}$$

$$(\boldsymbol{p}^{(y-1)})^T \boldsymbol{x}^y \leqslant P_{\max} - \|\boldsymbol{u}^{(y-1)}\|_1 \tag{35}$$

$$\boldsymbol{H}^{\boldsymbol{p}} [\boldsymbol{A}(\boldsymbol{p}^{(y-1)} \odot \boldsymbol{x}^y)] \leqslant \boldsymbol{\varepsilon}_l - \boldsymbol{H}^{\boldsymbol{p}} \boldsymbol{u}^{(y-1)} \tag{36}$$

$$\boldsymbol{0}_N \leqslant \boldsymbol{A}\boldsymbol{x}^y \leqslant \boldsymbol{1}_N \tag{37}$$

$$\boldsymbol{0}_{KNC} \leqslant \boldsymbol{x}^y \leqslant \boldsymbol{1}_{KNC}, \tag{38}$$

where $\boldsymbol{f}^{(y-1)}$ and $\boldsymbol{g}^{(y-1)}$ are the allocated bits for category 1 and category 2 users at the yth iteration, respectively, and $\boldsymbol{u}^{(y-1)}$ is the allocated power at the yth iteration.

Here, a detailed explanation on the iteration process is given. Recall that the bit allocation to the nth subchannel assigned to a category 1 SU $\boldsymbol{b}_{I,n} = \left[b_{1,n}^T \; \ldots \; b_{K_1,n}^T \right]^T$ is a vector of size $K_1 C \times 1$ with possible entries 1, 2, 4, and 6. Assume that during the subchannel allocation performed in the last subsection, the first subchannel has been allocated to the second user, which happens to be a category 1 SU. Then, $\boldsymbol{b}_{I,1} = [0000,1246,0000,0000]$ for users in category 1 (assuming there are 4 users). If the third subchannel had been allocated to the first user, which happens to be a category 2 SU, then $\boldsymbol{b}_{II,3} = [1246,0000,0000,0000]$ (assuming there are also 4 users in this category) and so on. Once this has been done and certain elements of $\boldsymbol{b}_{\boldsymbol{I}}$ and $\boldsymbol{b}_{\boldsymbol{II}}$ are zeros according to the subchannel allocation, the vectors $\boldsymbol{b}_{\boldsymbol{I}}$ and $\boldsymbol{b}_{\boldsymbol{II}}$ are renamed $\boldsymbol{b}_{\boldsymbol{I}}^{\boldsymbol{1}}$ and $\boldsymbol{b}_{\boldsymbol{II}}^{\boldsymbol{1}}$, respectively. Consequently, at the first iteration (i.e., when $y = 1$), the following optimization problem is solved:

$$\max_{\boldsymbol{x}^1} \left[\left(\boldsymbol{b}_{\boldsymbol{I}}^{\boldsymbol{1}}\right)^T \boldsymbol{x}_{\boldsymbol{I}}^{\boldsymbol{1}} + \left(\boldsymbol{b}_{\boldsymbol{II}}^{\boldsymbol{1}}\right)^T \boldsymbol{x}_{\boldsymbol{II}}^{\boldsymbol{1}} \right] \tag{39}$$

subject to

$$\boldsymbol{B}_i \boldsymbol{x}_{\boldsymbol{I}}^{\boldsymbol{1}} \geqslant \boldsymbol{R}_k; \; k = 1, 2, \ldots, K_1 \tag{40}$$

$$\boldsymbol{B}_j \boldsymbol{x}_{\boldsymbol{II}}^{\boldsymbol{1}} = \tilde{\boldsymbol{\gamma}}_k; \; k = K_1 + 1, K_1 + 2, \ldots, K \tag{41}$$

$$\boldsymbol{p}^T \boldsymbol{x}^1 \leqslant P_{\max} \tag{42}$$

$$\left[\boldsymbol{H}_{l,n,D}^{\boldsymbol{p}} (\boldsymbol{A}(\boldsymbol{p}_{\boldsymbol{D}} \odot \boldsymbol{x}^1) \right] \leqslant \boldsymbol{\varepsilon}_l \tag{43}$$

$$\left[\boldsymbol{H}_{l,n,C}^{\boldsymbol{p}} (\boldsymbol{A}(\boldsymbol{p}_{\boldsymbol{C}} \odot \boldsymbol{x}^1) \right] \leqslant \boldsymbol{\varepsilon}_l \tag{44}$$

$$\boldsymbol{0}_N \leqslant \boldsymbol{A}\boldsymbol{x}^1 \leqslant \boldsymbol{1}_N \tag{45}$$

$$\boldsymbol{0}_{KNC,1} \leqslant \boldsymbol{x}^1 \leqslant \boldsymbol{1}_{KNC,1}. \tag{46}$$

$\boldsymbol{f}^{(0)}$, $\boldsymbol{g}^{(0)}$, and $\boldsymbol{u}^{(0)}$ are all going to be 0 at the first iteration; hence, they are not reflected in the formulation above. The rates $\boldsymbol{B}_i \boldsymbol{x}_{\boldsymbol{I}}^{\boldsymbol{1}}$ and $\boldsymbol{B}_j \boldsymbol{x}_{\boldsymbol{II}}^{\boldsymbol{1}}$ and power $\boldsymbol{p}^T \boldsymbol{x}^1$ obtained at the first iteration are passed on as f^1, g^1, and $\boldsymbol{u}^{(1)}$, respectively, for the second iteration. Vector $\boldsymbol{x}^1$ is used along with the power vector $\boldsymbol{p}$ to determine the initial modulation scheme (invariably, the number of bits) for each SU at various subchannels. From the explanation given earlier, the first subchannel, say, has been allocated to the second SU. Hence, all entries of $\boldsymbol{x}_I^1$ are 0s except the elements in $x^1_{2,1}$. The total power allocated to the first subchannel can then be calculated as $\left(p^T_{2,1}x^1_{2,1}\right)$. To generalize, if the nth subchannel is allocated to the kth SU, the total power allocated to it is calculated as $\left(p^T_{k,n}x^1_{k,n}\right)$. The modulation scheme q (with bits c_q) that can be used without exceeding the power $p^T_{k,n}x^1_{k,n}$ can be obtained as

$$q = arg\max_q \left\{ q \in [0,1,2,3,4] \,:\, p_{k,n,q} \leqslant p^T_{k,n}x^1_{k,n} \right\}. \qquad (47)$$

The value q answers the question, "what is the largest possible modulation scheme with which the subchannel can transmit, such that its power will not exceed the power allocated to this subchannel?" Because the modulation sizes and their corresponding powers are finite and predetermined, the set of power levels that $p_{k,n,q}$ can take will be finite. Once the bits corresponding to this $p_{k,n,q}$ are determined, the total power used up to that point will still be less than $P_{\max}$. The interference leakage to PU will also still be less than ε. As a result, it is most likely that there will be some residual power available for use (thus implying that further iterations can still be performed to increase the number of bits already allocated to each subchannel). Hence, $y = 2$ (i.e., the second iteration) becomes feasible. Since (from the subchannel allocation) the first subchannel has been allocated to the second user, which happens to be in category 1, to transmit 2 bits (4-QAM modulation), then, $\boldsymbol{b}_{I,2,1}$ can be modified as $\boldsymbol{b}^2_{I,2,1} = [0\ 0\ (4-2)\ (6-2)]^T = [0\ 0\ 2\ 4]^T$. To have allocated 2 bits to this subchannel, the power $\boldsymbol{p}_{2,1,2}$ must have been used. With the realization of excess power available for use, the allocation might then be upgraded to, say, a 16-QAM (to transmit 2 more bits) or 64-QAM (to transmit 4 more bits). For this to take place would require additional power of $(p_{2,1,3} - p_{2,1,2})$ (for 16-QAM) or $(p_{2,1,4} - p_{2,1,2})$ (for 64-QAM) respectively. Hence, the new power vector at the second iteration $\boldsymbol{p}^1_{2,1} = [p_{2,1,1}\ p_{2,1,2}\ (p_{2,1,3} - p_{2,1,2})\ (p_{2,1,4} - p_{2,1,2})]^T$. The values of the vector $\boldsymbol{p}^1$ are thus determined. If u^1_n denotes the power that was allocated to the nth subchannel in the first iteration, then $\boldsymbol{u}^1 \triangleq \left[u^1_1\ \ldots\ u^1_N\right]^T$. It therefore implies that $P_{\max} - \sum_{n=1}^{N} u^1_n$, which can rather be written as $P_{\max} - \|\boldsymbol{u}^1\|_1$, is the residual power available for the second iteration step. After this second iteration, the amount of power allocated to the nth subchannel is the sum of the power allocated at the first iteration and that allocated at the second iteration. This total power is given as

$$v^2_n = u^1_n + \left(\boldsymbol{p}^1_{k,n}\right)^T \boldsymbol{x}^2_{k,n}.$$

This new power is used to decide to what the modulation scheme q of the nth subchannel should be upgraded.

TABLE 2 Pseudo-code for the proposed iterative-based heuristic

```
Pseudo-code for the subchannel allocation
solve for $\boldsymbol{x}$ using Equations 23 - 29 and 31
set subchannel index $n = 0$
repeat
    $n \leftarrow n + 1$
    if $(\boldsymbol{b}^T_{k,n}\boldsymbol{x}_{k,n}) \geqslant (\boldsymbol{b}^T_{m,n}\boldsymbol{x}_{m,n}) \forall m \neq k$
        $n$th subchannel is allocated to user $k$
    end if
until $n < N + 1$
Pseudo-code for the bit and power allocation (i.e., at $y = 1,2,3,\ldots$)
set $n = 0, y = 0, \boldsymbol{u}^{(0)} = \boldsymbol{0}_N, \boldsymbol{p}^{(0)} = \boldsymbol{p}$
repeat
    $y \leftarrow y + 1$
    set $\boldsymbol{f}^y = \boldsymbol{0}_K, \boldsymbol{g}^y = \boldsymbol{0}_K, \boldsymbol{v}^y = \boldsymbol{0}_N$
    solve the problem (32) - (38)
    repeat
        $N \leftarrow n + 1$
        $v^y_n = u^{y-1}_n + (\boldsymbol{p}^{y-1}_{k,n})^T \boldsymbol{x}^y_{k,n}$
        if $q = arg\max_q \left\{ q \in [0,1,2,3,4] \,:\, p_{k,n,q} \leqslant v^y_n \right\}$ then
            use modulation scheme $q$ (i.e., with $c_q$ bits) on $n$th subchannel
            set $u^y_{k,n} = p_{k,n,l};\ f^y_k = f^y_k + c_q;\ \ g^y_k = g^y_k + c_q$
            set $p^y_{k,n,m} = p_{k,n,m} - p_{k,n,l},\ \ \forall m > l$
            set $b^{y+1}_{k,n,m} = b_{k,n,l} - c_q,\ \ \forall m > l$
            set $b^{y+1}_{k,n,m} = 0,\ \ \forall m \leqslant l$
        end if
    until $n < N + 1$
until no further improvement on total data rate (Equation 49)
the vectors $\boldsymbol{f}^{y+1}$ and $\boldsymbol{g}^{y+1}$ contain the bits allocated for each subchannel in category 1 and 2 respectively
the vector $\boldsymbol{u}^{y+1}$ contains the power allocated for each subchannel
```

$$q = arg\max_q \left\{ q \in [0,1,2,3,4] \,:\, p_{k,n,q} \leqslant v^2_n \right\}. \qquad (48)$$

Similarly, the interference to PUs as a result of the power allocated in the first iteration step is given as $\boldsymbol{H^p u^1}$. The remaining interference permissible must be less than $(\boldsymbol{\varepsilon}_l - \boldsymbol{H^p u^1})$ for the second iteration. Since, at this second iteration, f^1_k already becomes the data rate allocated to the kth SU in category 1 during the first iteration and g^1_k becomes the data rate already allocated to the kth SU in category 2 during the first iteration, $\boldsymbol{f}^1$ and $\boldsymbol{g}^1$ are defined as $\boldsymbol{f^1} \triangleq \left[f^1_1\ \ldots\ f^k_1\right]^T$ and $\boldsymbol{g^1} \triangleq \left[g^1_1\ \ldots\ g^k_1\right]^T$, respectively. Hence, the data rate requirement at the second iteration for category 1 users would be $(\boldsymbol{R}_k - \boldsymbol{f}^1)$, while the available data rate for category 2 users at the second iteration would be $(\tilde{\boldsymbol{\gamma}}_k - \boldsymbol{g}^1)$. The constraints on data rate then become $\boldsymbol{B_i x^2_I} \geqslant [\boldsymbol{R}_k - \boldsymbol{f}^1]^+$ for category 1 users and $\boldsymbol{B_j x^2_{II}} = [\tilde{\boldsymbol{\gamma}}_k - \boldsymbol{g}^1]^+$ for category 2 users.

This whole iteration process is repeated continuously and only stopped when no further improvement can be achieved on the total achievable data rate for each user (invariably, the throughput of the system cannot be improved any further). The stopping criterion is thus given as

$$\left[\left(\boldsymbol{b}_I^y\right)^T \boldsymbol{x}_I^y + \left(\boldsymbol{b}_{II}^y\right)^T \boldsymbol{x}_{II}^y\right] - \left[\left(\boldsymbol{b}_I^{y-1}\right)^T \boldsymbol{x}_I^{y-1} + \left(\boldsymbol{b}_{II}^{y-1}\right)^T \boldsymbol{x}_{II}^{y-1}\right] = \varsigma, \tag{49}$$

where ς is a predetermined (very small) value. After the yth iteration, the vectors $\boldsymbol{f}^{(y+1)}$ and $\boldsymbol{g}^{(y+1)}$ will contain the allocated bits for each subchannel assigned to category 1 and category 2 users, respectively. The vector $\boldsymbol{u}^{(y+1)}$ will contain the power allocated to each subchannel. The pseudo-code given in Table 2 summarizes both the subchannel allocation and the iterative bit and power allocation that form the heuristic.

6 | RESULTS AND DISCUSSION

The RA model for the underlay, heterogeneous, cooperative CRN is simulated in MATLAB. For the simulation, the parameters used are given as OFDMA subchannels $N = 64$, PUs $L = 4$, SUs $= 8$ in all, with category 1 SUs $K_1 = 2$, category 2 SUs $(K - K_1) = 2$, and SUs as possible cooperators from which the best relay (CSU) is selected $= 4$. The minimum data rate requirement for category 1 SUs is 64 bits per user while the remaining resources are proportionately distributed among the category 2 SUs with a normalized proportional rate constant γ_k summed to unity. The bit error rate requirement for all SUs $\rho = 0.01$. The choice of the number of PUs, SUs, and other parameters used in the simulation is informed by the need to compare results obtained with similar works in the literature. Hence, to validate the results presented in this section, comparative results of the works of Awoyemi et al[9] and Rahulamathavan et al[17] have been reproduced, albeit at a heterogeneous scale (involving more SUs in 2 different categories). The other results simply build on the validated ones presented.

Figure 2 shows the interference channel gain patterns for the various PUs and the consequent bit allocation to each SU when only direct communication is considered. These results are similar to and validated by the ones obtained in two studies.[9,17] The bit allocation is done with careful consideration of the interference gains to the PUs. At high interference gain, the subchannels are allocated low data rates to reduce the adverse effect of high power and/or interference gain on the PUs and vice versa. Examples of this can be seen in subchannels 2, 3, 9, 57, 63, and 64 of Figure 2B where a high data rate has been allocated. The combined interference to the PUs on those subchannels (see Figure 2A) is lower than the combined interference on the other subchannels. On subchannels 14 to 27 and 39 to 52, the combined interference to PUs is quite high, and the subchannels have been allocated low data rates to transmit. This is the fundamental principle by which the bit allocation is performed to obtain optimal results on the overall throughput of the network.

Figure 3 gives the average data rate (bits per OFDMA symbol) for each category of SUs against the maximum interference power to the PUs for both direct communication and cooperative communication. Cases when the SUBS maximum transmit power is at 20 and 40 dBm are considered in Figure 3A,B, respectively. From the results obtained, it is obvious that for the developed RA problem to have feasible solutions, the minimum rate constraint of the high priority category 1 users has to be met at all times. Again, it can be seen that at a higher SUBS power (40 dBm), the average data rate is better than at a lower SUBS power (20 dBm). Importantly, the result shows that a marked improvement in performance of the network is achieved during cooperation, compared to when direct communication alone is used. The reason for this is the improved interference gain to PUs that the cooperative network achieves, which means that the subchannels can transmit at a higher rate than they would ordinarily have been allocated by direct communication. It is also worth noting that in Figure 3A, the average data rate during cooperation eventually converges to nearly that of direct communication.

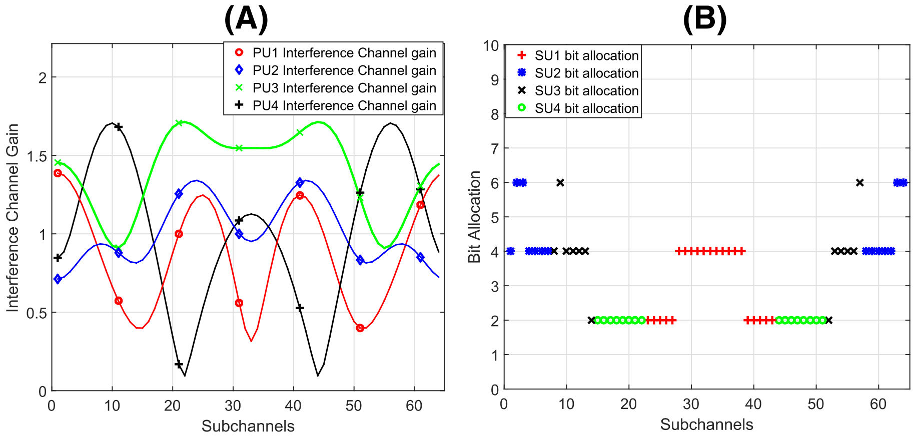

FIGURE 2 A, Interference gain of PUs. B, Subchannel allocation to SUs. Results obtained are comparable to those presented in the work of Rahulamathavan et al.[17] PU indicates primary user; SU, secondary user

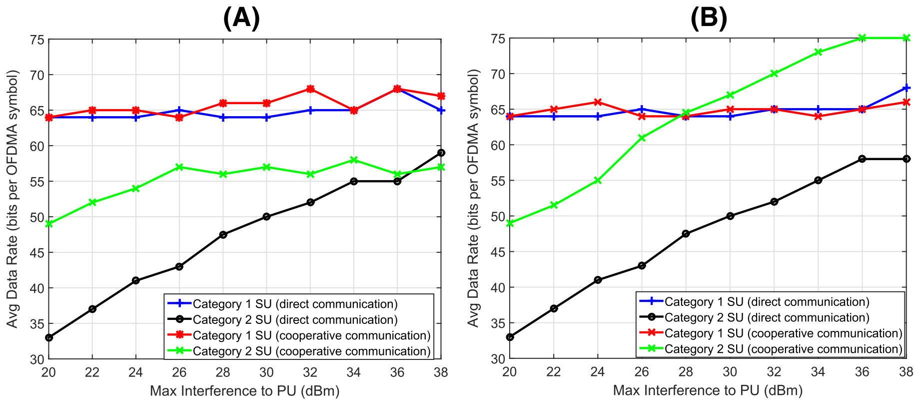

FIGURE 3 Average data rate for different categories of SUs, with both direct and cooperative communication considered. A is at 20 dBm maximum SUBS power; B is at 40 dBm maximum SUBS power. OFDMA indicates orthogonal frequency division multiple access; PU, primary user; SU, secondary user; SUBS, secondary user base station

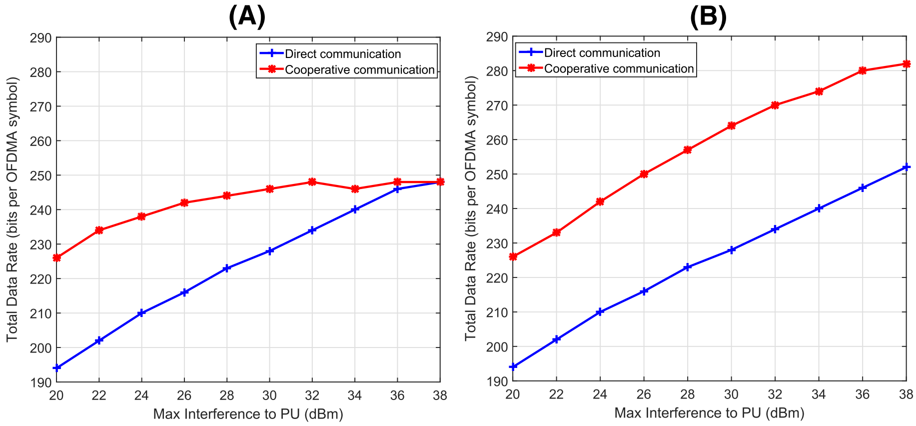

FIGURE 4 Total data rate for different categories of SUs, with both direct and cooperative communication considered. A is at 20 dBm maximum SUBS power; B is at 40 dBm maximum SUBS power. OFDMA indicates orthogonal frequency division multiple access; PU, primary user; SU, secondary user; SUBS, secondary user base station

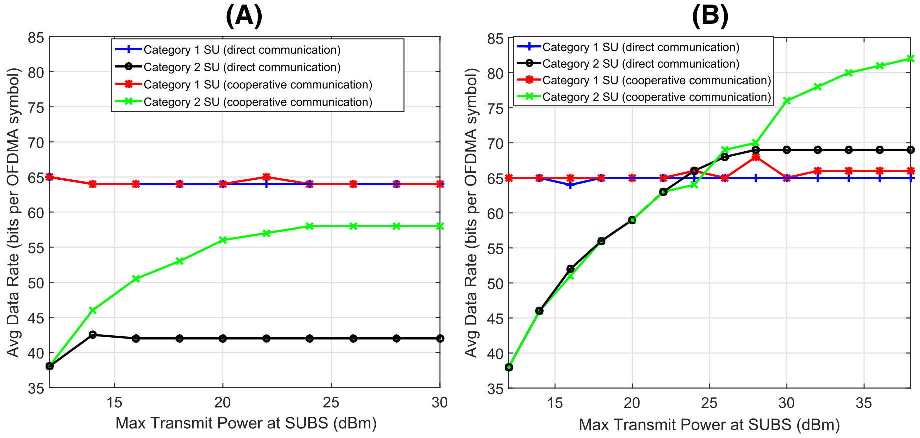

FIGURE 5 Average data rate for different categories of SUs, with both direct and cooperative communication considered. A is at 25 dBm maximum interference to PUs; B is at 45 dBm maximum interference to PUs. OFDMA indicates orthogonal frequency division multiple access; PU, primary user; SU, secondary user; SUBS, secondary user base station

This shows that as the permissible interference level to PUs increases, the need for and/or effect of cooperation diminishes. It would be better to transmit directly if the PUs are robust to the SUs' interference than to transmit using cooperation, as cooperation generally requires much more signalling overhead than direct communication.

In Figure 4, the total data rate for each category of SUs and the maximum interference power to the PUs for both direct communication and cooperative communication are compared. Similar to Figure 3, cases when the SUBS maximum transmit power is at 20 and 40 dBm are considered in Figure 4A,B, respectively. The explanations given for Figure 3 are also applicable here, as the total data rate during cooperation generally outperforms that of its direct communication counterpart. Similar reasoning and deductions about the throughput and the better performance of cooperation compared with direct communication can also be made.

Figure 5 describes the average data rate performance for increasing SUBS power. The 2 categories of SUs are covered, and both direct and cooperative communications are considered. In Figure 5A, a maximum interference power to PUs of 25 dBm is used, while the maximum interference power to PUs in Figure 5B has been increased to 45 dBm. At all times, the minimum rate guarantee of the category 1 SUs must be met for the problem to have feasible solutions. Again, as the SUBS power is increased, the average data rate improves, particularly for category 2 SUs. After a while though, no further improvement can be observed, irrespective of whether or not the SUBS power is increased. The reason for this is that the other constraints also come into play, thus making it impossible for the data rate to keep increasing indefinitely with increasing SUBS power. It is significant to note the improvement that cooperative communication achieves over direct communication. This improvement can be seen when the interference limit is at 25 dBm (Figure 5A) and at 45 dBm (Figure 5B). These results clearly show that the network would rather transmit using direct communication when the SUBS power is limited so as to maximize the power use and reduce signalling overhead. At higher power, however, cooperative communication is preferred, as the overall capacity is remarkably better.

Figure 6 presents the total data rate as realized by the network for all the categories of SUs, with both direct and cooperative communications. Also similar to Figure 5, the cases presented are when the maximum interference to PUs is at 25 and 45 dBm, as seen in Figure 6A,B, respectively. The explanations given for Figure 5 are also very appropriate for describing the performance of the network. The total data

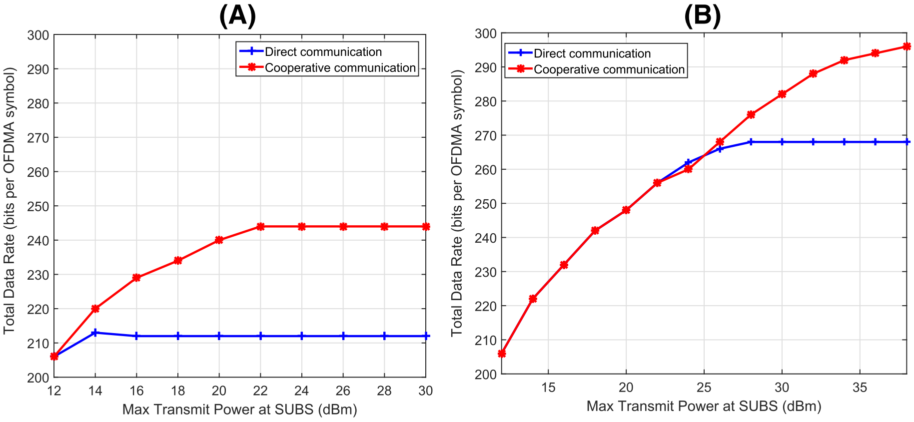

FIGURE 6 Total data rate for different categories of SUs, with both direct and cooperative communication considered. A is at 25 dBm maximum interference to PUs; B is at 45 dBm maximum interference to PUs. OFDMA indicates orthogonal frequency division multiple access; PU, primary user; SU, secondary user; SUBS, secondary user base station

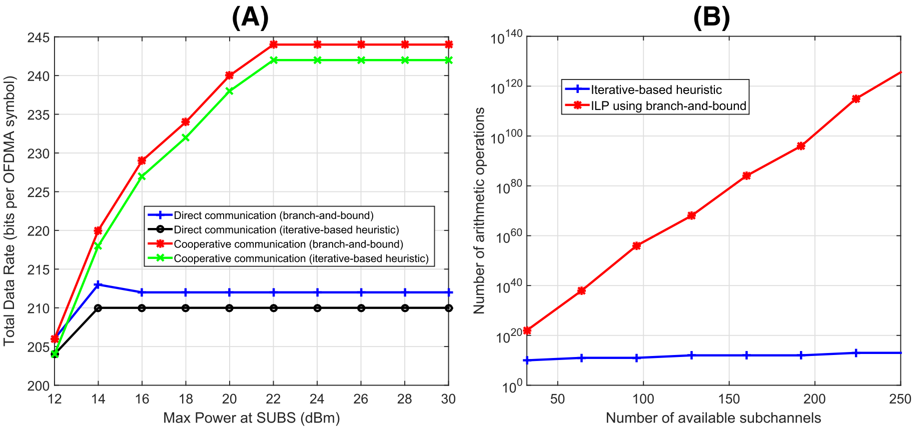

FIGURE 7 Performance of the branch-and-bound method compared with the iterative-based heuristic approach. A, Total data rate for an increasing SUBS power (max interference to PUs at 25 dBm). B, Computational complexity for different number of subchannels. Results comparable to the work of Rahulamathavan et al.[26] ILP indicates integer linear programming; OFDMA, orthogonal frequency division multiple access; PU, primary user; SUBS, secondary user base station

rates during cooperation generally outperform rates achieved during direct communication. The reasoning and inferences given in Figure 5 are also applicable to understanding the results of Figure 6

In Figure 7, the optimality of the network and the complexity are compared for the ILP using BnB and the iterative-based heuristic. The computational complexity is obtained from the number of arithmetic operations that the network undergoes before arriving at the solution.[26] For the heuristic, the total complexity is the sum of the complexities of the 2 parts (subchannel allocation and the iterative bit and power allocation). The results presented show that while the heuristic performs very close to optimality in its total data rate for the network, the complexity is significantly less, especially as the network gets larger. For such large CRN systems therefore, developing appropriate heuristic(s) to solve them, thereby providing both feasible and timeous solutions with lower complexities, is recommended.

7 | CONCLUSION

In CRN, RA models that can yield outstanding productivity even with very stringent constraints are critical. This work develops such a model whereby, in a heterogeneous CRN environment, cooperative diversity is used in mitigating the limiting effects of the interference to the PUs of the network. To make the model feasible and close to practical, only one single best relay is selected from the available ones as the cooperating relay. Also, cooperation is only used by users that have subchannels with a high interference gain to the PU. The problem that has been developed is first solved by a careful reformulation of the non-deterministic polynomial-time-hard problem into an ILP problem, and optimal solutions are obtained using the BnB method for solving ILP problems. To reduce computational complexity, an iterative-based heuristic is developed to solve the problem in a much reduced time frame. The results presented compare the average data rates and total data rates for the different categories of SUs when direct and cooperative communications are used. The optimality and computational complexity of the developed heuristic are compared with those obtained using ILP as well. The improvement in the performance of the network when cooperation is used is quite remarkable, as the results have shown.

ACKNOWLEDGMENTS

This work was supported by the SENTECH Chair in Broadband Wireless Multimedia Communications, a telecommunications research group in the Department of Electrical, Electronics and Computer Engineering at the University of Pretoria, South Africa. The authors would like to thank Dr Y. Rahulamathavan for his helpful contributions and timely support that proved vital in the course of this research work.

REFERENCES

1. F. C. Commission. Report of spectrum efficiency working group. Spectrum policy task force, Washington, DC, USA; 2002.
2. F. C. Commission. Cognitive radio technologies proceeding. Rep. ET Docket, no. 03-108; 2003.
3. Doyle LE. *Essentials of Cognitive Radio*, The Cambridge Wireless Essentials Series. New York, USA: Cambridge University Press; 2009.
4. Pretz K. Overcoming spectrum scarcity—cognitive radio networks might be one answer; 2012. Available at http://theinstitute.ieee.org/technology-focus/technology-topic/overcoming-spectrum-scarcity. Accessed August 4, 2014
5. Ge M, Wang S. On the resource allocation for multi-relay cognitive radio systems. *2014 IEEE ICC - Cognitive Radio and Networks Symposium*, Sydney, NSW; June 2014:1591–1595.
6. Li J, Luo T, Yue G. Resource allocation scheme based on weighted power control in cognitive radio systems. *Communications, Circuits and Systems (ICCCAS), 2013 International Conference on*, IEEE, Chengdu, China, vol. 1; November 2013:178–182.
7. Oh J, Choi W. A hybrid cognitive radio system: a combination of underlay and overlay approaches. *Vehicular Technology Conference Fall (VTC 2010-Fall), 2010 IEEE 72nd*, Ottawa, ON; September 2010:1–5.
8. Hao N, Yoo S-J. Interference avoidance throughput optimization in cognitive radio ad hoc networks. *EURASIP J Wirel Commun Networking*. 2012;2012(1):295–313.
9. Awoyemi BS, Maharaj BT, Alfa AS. Resource allocation for heterogeneous cognitive radio networks. *2015 IEEE Wireless Communications and Networking Conference (WCNC)*, New Orleans, LA; March 2015:1759–1763.
10. Awoyemi BS, Maharaj BT, Alfa AS. QoS provisioning in heterogeneous cognitive radio networks through dynamic resource allocation. *AFRICON, 2015*, Addis Ababa, Ethiopia; September 2015:1–6.
11. Awoyemi B. *Performance Analysis of Cooperative Diversity in Land Mobile Satellite Systems*. Durban: University of KwaZulu-Natal; 2013. Available at http://books.google.co.za/books?id=TGn6oAEACAAJ.
12. Laneman J, Tse D, Wornell GW. Cooperative diversity in wireless networks: efficient protocols and outage behavior. *IEEE Trans Inf Theory*. 2004;50(12): 3062–3080.
13. Awoyemi B, Walingo T, Takawira F. Relay selection cooperative diversity in land mobile satellite systems. *AFRICON, 2013*: IEEE, Pointe-Aux-Piments, Mauritius; 2013:1–6.
14. Wang L, Xu W, He Z, Lin J. Algorithms for optimal resource allocation in heterogeneous cognitive radio networks. *Power Electronics and Intelligent Transportation System (PEITS), 2009 2nd International Conference on*, IEEE, Shenzhen, China, vol. 2; 2009:396–400.
15. Nguyen D, Krunz M. Heterogeneous spectrum sharing with rate demands in cognitive mimo networks. *Global Communications Conference (GLOBECOM), 2013*: IEEE, Atlanta, GA; 2013:3054–3059.
16. Zheng L, Tan CW. Cognitive radio network duality and algorithms for utility maximization. *IEEE J Sel Areas Commun*. 2013;31(3):500–513.
17. Rahulamathavan Y, Cumanan K, Musavian L, Lambotharan S. Optimal subcarrier and bit allocation techniques for cognitive radio networks using integer linear programming. *Statistical Signal Processing, 2009. SSP '09. IEEE/SP 15th Workshop on*, Cardiff, Wales; 2009:293–296.
18. Guo W, Huang X. Maximizing throughput for overlaid cognitive radio networks. *Military Communications Conference, 2009. MILCOM 2009. IEEE*, Boston, MA; October 2009:1–7.
19. Xie R, Yu F, Ji H. Dynamic resource allocation for heterogeneous services in cognitive radio networks with imperfect channel sensing. *IEEE Trans Veh Technol*. 2012;61(2):770–780.
20. Wang S, Zhou Z-H, Ge M, Wang C. Resource allocation for heterogeneous cognitive radio networks with imperfect spectrum sensing. *IEEE Sel Areas Commun*. 2013;31(3):464–475.
21. Chen F, Xu W, Guo Y, Lin J, Chen M. Resource allocation in ofdm-based heterogeneous cognitive radio networks with imperfect spectrum sensing and guaranteed QoS. *Communications and Networking in China (CHINACOM), 2013 8th International ICST Conference on*, Guilin, China; August 2013: 46–51.

22. Wang S, Ge M, Wang C. Efficient resource allocation for cognitive radio networks with cooperative relays. *IEEE J Sel Areas Commun*. 2013;31(11): 2432–2441.

23. Adian M, Aghaeinia H. Optimal resource allocation for opportunistic spectrum access in multiple-input multiple-output-orthogonal frequency division multiplexing based cooperative cognitive radio networks. *IET Signal Proc*. September 2013;7(7):549–557.

24. Du S, Huang F, Wang S. Power allocation for orthogonal frequency division multiplexing-based cognitive radio networks with cooperative relays. *IET Commun*. 2014;8(6):921–929.

25. Pischella M, Le Ruyet D. Cooperative allocation for underlay cognitive radio systems. *Signal Processing Advances in Wireless Communications (SPAWC), 2013 IEEE 14th Workshop on*, Darmstadt, Germany; June 2013:245–249.

26. Rahulamathavan Y, Lambotharan S, Toker C, Gershman A. Suboptimal recursive optimisation framework for adaptive resource allocation in spectrum-sharing networks. *IET Signal Proc*. 2012;6(1):27–33.

How to cite this article: Awoyemi BS, Maharaj BT, Alfa AS. Resource allocation in heterogeneous cooperative cognitive radio networks. *Int J Commun Syst*. 2017;30:e3247. https://doi.org/10.1002/dac.3247