



City Research Online

City, University of London Institutional Repository

Citation: Arnold, P., Rakocevic, V. & Habermann, J. (2019). Hybrid Radio Resource Management with Co-scheduling for Relay Extended OFDMA Networks. *Wireless Personal Communications*, 109(2), pp. 1133-1160. doi: 10.1007/s11277-019-06607-4

This is the published version of the paper.

This version of the publication may differ from the final published version.

Permanent repository link: <https://openaccess.city.ac.uk/id/eprint/22437/>

Link to published version: <https://doi.org/10.1007/s11277-019-06607-4>

Copyright: City Research Online aims to make research outputs of City, University of London available to a wider audience. Copyright and Moral Rights remain with the author(s) and/or copyright holders. URLs from City Research Online may be freely distributed and linked to.

Reuse: Copies of full items can be used for personal research or study, educational, or not-for-profit purposes without prior permission or charge. Provided that the authors, title and full bibliographic details are credited, a hyperlink and/or URL is given for the original metadata page and the content is not changed in any way.

City Research Online:

<http://openaccess.city.ac.uk/>

publications@city.ac.uk



Hybrid Radio Resource Management with Co-scheduling for Relay Extended OFDMA Networks

Paul Arnold¹ · Veselin Rakocevic¹ · Joachim Habermann²

© The Author(s) 2019

Abstract

In orthogonal frequency division multiple access networks buffer aided non-transparent in-band half duplex decode and forward relay nodes aim to improve coverage and capacity under fairness considerations. Existing centralized radio resource management and inter cell interference coordination schemes achieve these goals at the cost of heavy signalling overhead. Especially for frequency division duplex downlink transmission this is a critical issue. Fully decentralized schemes often focus on different types of frequency reuse schemes with less amount of necessary feedback. Here, it is often overseen that in a practical deployment, the backhaul link quality is the bottleneck of the two hop transmission and needs to be taken into account. Moreover, it is often modelled way too optimistic and necessary co-scheduling with single hop UE further limits the possible data rate. In order to minimize the required overhead this work proposes a hybrid radio resource management (RRM) scheme. The RRM includes synchronous adapted two-hop proportional frequency selective resource scheduling as the decentralized part. Asynchronous subband power allocation scheme with very limited feedback is proposed to maximize the wireless backhaul link quality with no loss for single hop UE. Comprehensive system level simulation results show stable fairness and throughput when minimizing the required feedback and improvements for the backhaul links based on the centralized adapted power allocation including no losses in the overall system. In addition possible energy savings for the shared channel are presented when applying the proposed scheme.

Keywords Radio resource management · Relay · Meta-heuristic · Co-scheduling · CQI feedback · LTE advanced pro · 5G

✉ Paul Arnold
paularnold@mail.de
Veselin Rakocevic
veselin.rakocevic.1@city.ac.uk
Joachim Habermann
joachim.habermann@iem.thm.de

¹ School of Mathematics, Computer Science and Engineering, City, University of London, London, UK

² University of Applied Sciences Mittelhessen, Friedberg, Germany

1 Introduction

During the last decade academia and the industry have paid a lot of attention on the improvement of the system capacity of mobile networks. One possibility to satisfy the ever growing data demand and increase capacity is to densify the network, using different kind of small cells also known as heterogeneous networks [1]. Besides pico and femto cells, relay node (RN)s have been introduced to mobile networks, to improve coverage and capacity. The difference with RNs is that they are backhauled by a wireless link and therefore they might be an attractive alternative to wired backhauled pico cells [2] for operators to deploy because of reasonable capital expenditure (CAPEX) and operational expenditures (OPEX). Different kinds of RNs have been designed and standardized in recent years [3, 4] in third generation partnership project (3GPP) and will be further investigated in an upcoming study item defined as integrated access and backhaul (IAB) [5]. In general, such node types introduce a challenging research topic, primarily how to handle and allocate the available radio resources in the most efficient way and thus, how to gain the most from additional nodes in the mobile network. Those relay node (RN)s operating in half duplex mode suffer from a loss in transmission time. Sophisticated means of resource allocation have to be designed to overcome this problem and therefore, to maximize the spectral efficiency of the two hop links. Therefore, different targets for the radio resource management approaches have been identified, such as the maximization of spectral efficiency, fairness, interference mitigation, energy awareness, queueing and quality of service (QoS). Furthermore, the main challenge in designing RRM algorithms for relay extended systems is that typically an optimal solution requires prohibitively exhaustive solution search, with complexity of $O((MK)^N)$, where M, K and N represent the number of RNs, macro base station (MBS)s, and subchannels, respectively. Various RRM strategies related to the aforementioned targets for all kind of RN types have been intensively studied in the past years. Excellent surveys can be found in the literature summarizing the current state of the art and reveal unsolved challenges [6–9].

Different types of relay nodes, such as amplify and forward, decode and forward or self-backhauling RNs require different design principles of RRM schemes. It is of paramount importance for the design of RRM algorithms, to consider the underlying system assumptions.

For instance, Lee et al. conclude in [9], after an extensive literature search that, to date it remains a challenging task to design resource allocation schemes for LTE/LTE-A multihop networks with low complexity, while simultaneously excelling in aspects such as interference mitigation, resource utilization and global fairness. A comprehensive classification of RRM schemes for heterogeneous networks is done by the authors. An excellent overview of research problems in heterogeneous networks is given. Most of the considered schemes comes with the burden of heavily increased channel state information (CSI) feedback exchange, which introduces increased control traffic exchange among MBSs and additional signalling on the RN wireless backhaul (BH) links for the RNs [9]. By now, existing work fails to consider all constraints of a practical RN extended system, where only low complexity schemes can be applied not exceeding necessary CSI feedback exchange, especially for FDD downlink transmissions.

For instance, the authors in [10] consider a joint BH and access optimization for a centralized radio access network (C-RAN) time division duplex (TDD) system. They assume to have a system-wide channel knowledge at the central C-RAN control entity. Therefore the MBSs need to send huge amount of collected CSI information to the C-RAN central

entity. Besides that, it is assumed that the MBSs only serve RNs, while no MBS user equipment (UE)s are considered in an ultra dense network. Especially in the half duplex mode only UEs with bad supportable rate of the direct link might gain from the two hop links compared to single hop connections. In [11] a cell centralized scheme is proposed which adjusts the time ratio between BH and access links (AL)s to counteract on the different channel states of both. This improves fairness and system throughput. However, they only consider one macro cell and thus do not consider inter MBS interference. They assume full flexibility of switching between transmit and receive time slots which is impractical. Further, they do not adapt the transmission power for BH improvement. Liu et al. propose a joint resource and power allocation optimization for downlink in [12]. They assume that there is no direct transmission between MBS and MSs. Further, they additionally take into account only slowly varying radio channels to receive accurate feedback to be able to calculate a near optimal power and subcarrier allocation for each slot, which is an impractical solution. In [13] a simple resource allocation algorithm is suggested to maximize cell average throughput and cell edge performance. However, no fast fading and therefore perfect channel knowledge is considered. Furthermore, no fairness or energy efficient transmission is considered. Jeon et al. [14] propose a distributed non-cooperative game resource allocation strategy to provide high cell throughput and fulfil minimum data rate. Proportional fair and explicit RN wireless BH improvement is not considered. In [15] the authors analyse practical resource allocation strategies for non-transparent inband relays in an FDD system. The MBS is able to serve the MBS UE and RNs in the same time slots. By varying the number of BH subframes, they show the influence on fairness and downlink throughput. However, no additional interference mitigation or energy efficient strategy is applied. Also no further strategy is introduced to improve the BH spectral efficiency. Wang et al. [16] study fairness and throughput downlink performance based on a designed three step resource allocation scheme. They introduce buffers at the RNs to avoid limitations of the performance because of the quality imbalance of BH and ALs, when no buffers would be applied buffers and therefore, the same transmission rates of BH and RN ALs have to be reached to prevent traffic congestion at the RN. No additional technique in terms of energy efficient or interference mitigation strategy is considered. In [17] low complexity resource partitioning schemes with two-hop proportional fairness are considered. The results show improved RN UE cell edge performance as well as increased average MBS UE performance under fairness improvement. However, the BH and AL match rate is assumed and no additional buffer to further improve the system is considered. Furthermore only the average MBS UE throughput is presented, without information about the CDF of the macro UEs. No adapted power allocation is applied to reduce interference or improve energy consumption. The goal of this work is to develop an RRM scheme which improves the overall performance of a half duplex relay extended OFDMA network, under consideration of all practical limitations of an DL FDD network, by improving the RN BH spectral efficiency (SE) while protecting the performance of the UEs, directly served by MBSs. The considered relay type is a non-transparent relay in half duplex and non cooperative mode, which makes the RN a separated small cell besides the existing MBS cells in the network. The RN has its own scheduling functionality and appears at the donor MBS as an UE. Furthermore, the RN has the ability to queue received data from its donor MBS. The serving station has suboptimal frequency selective channel knowledge based on UE channel quality information (CQI) feedback as defined in [18]. A UE has exclusively orthogonal access to a subchannel within a cell (MBS/RN), which results in no intra cell interference, while full reuse of subchannels among cells is assumed. This introduces inter cell interference but provides full access to the radio resources for each cell.

The target of this contribution is to design a practical low complexity solution which aims to satisfy a combination of several research challenges identified during the past years. On the one hand, in dense RN FDD networks the BH link is often identified as the bottleneck for the downlink. Therefore the main question answered in this work is, if an adapted subband transmission power pattern can improve the BH link spectral efficiency without losing macro UE performance and introducing additional non practical high amount of periodic CSI feedback. On the other hand, fairness needs to be obtained as well as queue-awareness and energy-efficiency needs to be considered. Therefore, the proposed practical low complexity algorithm results in the design of a hierarchical RRM scheme, which consists of the asynchronous centralized part as well as the synchronous decentralized part. The algorithm aims to reduce the interference and used energy for the transmission in the downlink shared channel, while improving the wireless BH rates under fairness and overhead constraints, and therefore resulting in an improved UE throughput.

The paper is organized as follows. In Sect. 2 a practical hierarchical RRM approach is defined, followed by the considered system model in Sect. 3. Simulation results are discussed in Sect. 3.2 and concluded in Sect. 4.

2 Radio Resource Management Approach

Two different types of time transmission interval (TTI)s occur in the considered half duplex relay extended networks. The first one, further defined as backhaul subframe (BHSF) serves the RN, while they are in receiving mode, and UEs directly attached to the donor MBS. Within this TTI interference is caused only by MBSs. The possible radio link types which occur in a BHSF are illustrated in Fig. 1. The second subframe, further declared as access link subframe (ALSF), serves the directly attached MBS UEs, as well as the two hop (2H) UEs assigned to the RNs, while the RNs are in the sending mode. Here, RN interference occurs in addition. Possible link types within a ALSF are depicted in Fig. 2. The intended RRM approach aims to improve the RN BH links in the first and the directly attached UEs in the second TTI. The possible throughput of the UEs connected to the RNs will be limited by the BH transmission for each RN during the BHSF. The RN

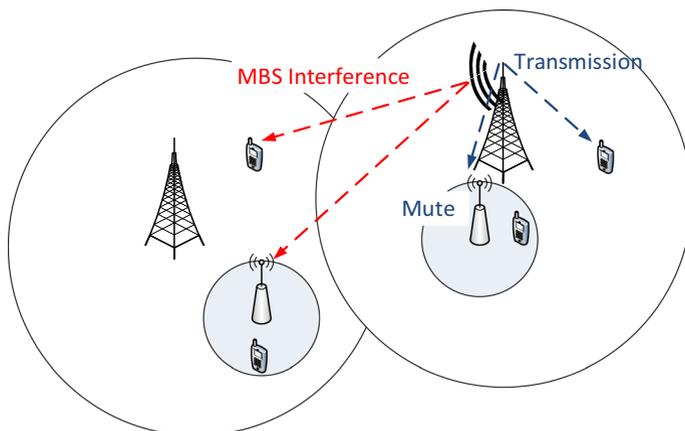


Fig. 1 Possible serving and interfering links during BH transmission time

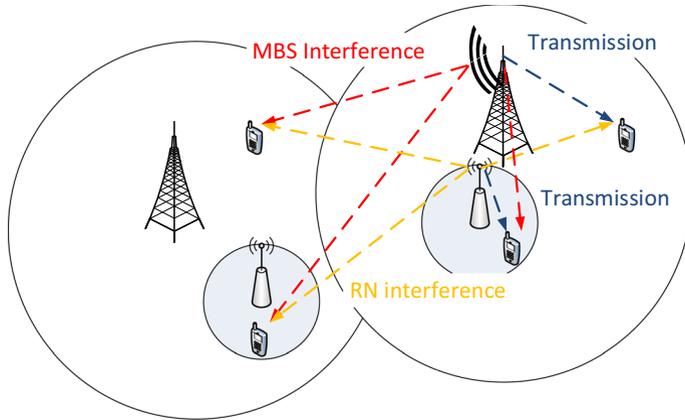


Fig. 2 Possible serving and interfering links during RN access transmission time

UE throughput will only improve if the BH links in the first subframe will have increased SE and more data per RN UE will be transmitted and buffered at the RN. The major challenge is to improve the RN BH links without a loss of the directly attached UEs, which share the capacity of the BHSF with the RNs. The ALSF, which is exclusively used by the directly linked UEs at the MBS is additionally interfered by the RN to UE transmissions, when RNs are in sending mode. Both subframes are interdependent to each other. When the RNs get too much resources the MBS UEs experience a loss in the BHSF in terms of instantaneous throughput. Thus, the resulting loss, needs to be compensated within the BHSF and ALSF by increasing the instantaneous throughput based on the improved spectral efficiency. An excellent result of the proposed RRM scheme would be to keep the overall direct link UE performance stable, while improving the BH link quality of the RNs and thus, improve the overall throughput performance of the two hop UEs. To minimize the additional interference in the ALSF, caused by the RN transmission, the quality of the RN AL should be improved as well, to provide a stable MBS UEs link quality. Further, fairness needs still to be guaranteed as well as possible energy savings need to be considered under limited amount of additional signalling overhead.

In Figs. 1 and 2 all possible link types which can occur are illustrated.

The proposed RRM scheme which aims to improve the system capacity is separated into two parts. The first step is based on a centralized asynchronous approach, which aims to adapt the transmission powers of all serving stations, MBS as well as RNs, in the network. Both types of TTIs are taken into account, while interdependencies are considered. The average received powers of serving and interfering links need to be measured by the UEs and RNs and fed back to a central optimization entity. As an additional required information, the optimizer needs to know; 1. if the measurement received was performed by an MBS, RN UE or an RN and 2. in which subframe type the measurement was carried out. Once the centralized unit has collected all feedbacks, the optimization can be initiated. The final outcome of this asynchronous procedure consists of a transmission power adaptation for each serving station as well as the total number of physical resource block (PRB)s, where the adapted transmission power patterns will be applied. The process might be repeated in a larger time scale (e.g. hundreds of ms) to prevent additional feedback overhead, caused by the necessary UE measurement reports for the calculation. Once the

improved power pattern is derived, no further action is needed by the centralized entity. It is obvious that this might result in suboptimal solution due to the nature of channel variations. However it can be easily applied in real networks due to its low amount of feedback to the centralized unit.

The second part of the process is based on a synchronous adapted scheduling procedure, where in each TTI the available PRBs are allocated to UEs and RNs to serve. This is done based on an adapted two hop proportional fair metric, which takes into account CQI reporting of RN and UEs, past decisions as well as resource allocation control rules, how to use the optimized subbands. Besides subband CQI reporting which needs to be fed back by the RN to its donor MBS for all served UEs, also relaxed different type of reports are considered to ease the amount of signalling overhead, further defined in Sect. 2.2.

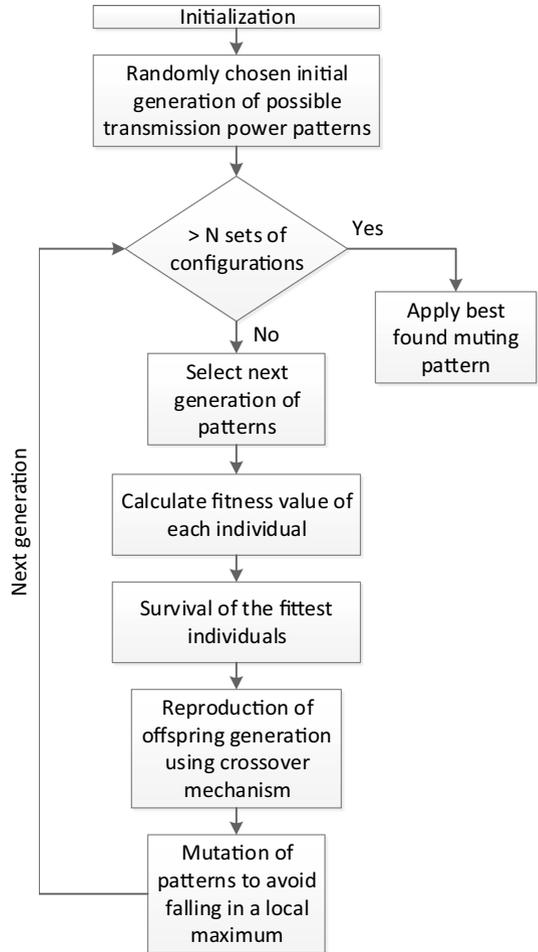
2.1 Centralized Asynchronous RRM

To find the theoretically optimal power distribution pattern would be quite challenging, due to the huge number of existing combinations. The combinatorial problem cannot be optimally solved during runtime of the system, even if only average received power values are considered. To this end, a heuristic needs to be applied to find a nearly optimal solution. As an example, it is assumed to have two MBSs with additional two RNs in each coverage area. To reduce the number of combinations a limited subset of transmission power reduction values can be defined, which will be applied for AL and BH subframes, respectively. If for instance a set of 10 possible power reduction values are assumed, it already results in a large number of possible combinations, with already 10^2 possibilities when RNs are in receiving mode multiplied with 10^4 possible options which already results in total in $1 \cdot 10^6$ combinations, according to Sect. 1. If a more realistic scenario is assumed with, e.g. 21 MBS plus 42 RNs, it results in an impractical solution space which is non-deterministic polynomial (NP) hard to solve. A near optimal result also cannot be found by an exhaustive search (deterministic approach) and thus, a heuristic approach needs to be applied, which is defined in following. For the asynchronous solution search a genetic algorithm is precisely adapted and applied to the optimization problem. In Fig. 3 the flow chart of the defined genetic algorithm (GA) is shown. The procedure first starts with the initialization of the algorithm. Several parameters, such as the choice of fitness function, mutation rate, number of generations (optimization steps), etc. are set. As an example, a parameter set applied for the conducted simulations can be found in Table 1. After the initialization is done, a randomly chosen first generation is defined. This generation includes a number of individuals which consists of a set of possible transmission power adaptation value (PAV)s for each MBS in the BH SF, the ALSF as well as a PAV for each RN in the ALSF. Based on the measurement reports received from RNs and UEs, as explained in the previous section. Based on the collected information the fitness values (FV)s for all individuals are calculated, as described in the following.

2.1.1 Calculation of Serving and Interfering Links with Power Adaptation Values

Based on the link specific average power calculation taken from [19], Eq. 1 defines the effective received average power of the serving signal \tilde{s}_0 of UE n for each individual i in the considered subframe type s . The θ is defined as the MBS m or the RN r specific PAV, as a scalar of the vector P_{pav} , which consists of the defined possible PAVs. Accordingly, Eqs. 2 and 3 define the sum of the total number of effective interfering links, either for interfering MBSs

Fig. 3 Flow chart of the applied genetic algorithm



\tilde{i}_{MBS} or RNs \tilde{i}_{RN} , respectively. The sum of the effective interference caused by the RNs (\tilde{i}_{RN}) only needs to be defined for the ALSF, due to reception mode of RNs in the BHSP.

$$\tilde{s}_0(n, i, s) = \Theta_{m,i,s} \cdot s_0(n, i), \quad \Theta_{m,i,s} \in P_{pav} \tag{1}$$

$$\tilde{i}_{MBS}(n, i, s) = \sum_{m=1}^{M_{BS}} \Theta_{m,i,s} \cdot i_{MBS}(n, i), \quad \Theta_{m,i,s} \in P_{pav} \tag{2}$$

$$\tilde{i}_{RN,AL}(n, i) = \sum_{r=1}^{R_{RN}} \Theta_{r,i} \cdot i_r(n, i), \quad \Theta_{r,i} \in P_{pav} \tag{3}$$

2.1.2 Calculation of Adapted Wideband SINR for Access and BH Subframe

To determine the FV of each individual (i) the resulting effective wideband Signal to Interference plus noise ratio (SINR)s (\tilde{g}) needs to be calculated for all UEs n , based on [20]. Equation 4 defines the wSINRs for all UEs in the ALSF ($s = 1$) including the interference ($\tilde{i}_{RN,AL}$) caused by the RNs in sending mode. Those are either UEs served by the MBSs ($N_{UE_{DL}}$) or UEs attached to the RNs ($N_{UE_{2H}}$). The additive average gaussian white noise is considered as σ .

$$\tilde{g}_{AL}(n, i) = 10\log_{10}\left(\frac{\tilde{s}_0(n, i, 1)}{\tilde{i}_{MBS}(n, i, 1) + \tilde{i}_{RN}(n, i) + \sigma}\right), \quad n \in \{N_{UE_{DL}}, N_{UE_{2H}}\} \tag{4}$$

Equation 5 defines the effective wSINR in BHSFs (\tilde{g}_{BH}) with RNs (N_{RN}) in the reception mode and for UEs ($N_{UE_{DL}}$) served by the MBS ($s = 2$). Here, no interference caused by the RNs is needed to be considered.

$$\tilde{g}_{BH}(n, i) = 10\log_{10}\left(\frac{\tilde{s}_0(n, i, 2)}{\tilde{i}_{MBS}(n, i, 2) + \sigma}\right), \quad n \in \{N_{RN}, N_{UE_{DL}}\} \tag{5}$$

2.1.3 Calculation of the Average Spectral Efficiency for Direct and Two Hop Connections

Finally, to determine the effective supportable rate (\tilde{r}) for all effective wSINRs an adapted Shannon bound curve is used based on in Eq. 6, which α set to 0.6 as a factor considering the system overhead [21]. Additionally, two boundaries are defined in Eq. 7. The minimum supported wSINR of the system g_{min} as well as a maximum reachable rate r_{max} . The minimum required SINR g_{min} is dependent on the design and robustness of the control channel of the system and r_{max} depends on the supported types of modulation and coding scheme (MCS) and spatial layers.

$$\tilde{r} = \alpha * \log_2\left(1 + 10^{\frac{\tilde{g}}{10}}\right), \tag{6}$$

$$\tilde{r}(\tilde{g} < g_{min}) = 0, \quad \tilde{r}(\tilde{r} > r_{max}) = r_{max} \tag{7}$$

For all UEs directly served by the MBS, the two existing rates need to be weighted with the percentage amount of time they can be used for transmission in principle, based on [10]. For RNs as receiver t_2 is set to zero since they can only receive data during a BHSF.

$$\tilde{r}_{DL}(n, i) = \frac{\tilde{r}_{BH}(n, i) \cdot t_1 + \tilde{r}_{AL}(n, i) \cdot t_2}{t_1 + t_2}, \quad n \in \{N_{RN}, N_{UE_{DL}}\} \tag{8}$$

For the two hop UEs $N_{UE_{2H}}$, the possible rates on the BH link of their serving RNs (M_{sRN}) need to be taken into account, as given in Eq. 9. Here, the loss in time for the two hop transmissions is already considered based on [22].

$$\tilde{r}_{2H}(n, i) = \frac{\tilde{r}_{BH}(m, i) \cdot \tilde{r}_{AL}(n, i)}{\tilde{r}_{BH}(m, i) + \tilde{r}_{AL}(n, i)}, \quad m \in M_{sRN}, \quad n \in N_{UE_{2H}} \tag{9}$$

Equation 10, finally defines the difference of the possible rates ($\Delta\tilde{r}$) compared to the reference rate (r_{ref}) without applying the current PAVs pattern for the considered UE n of the individual i in the total set of the generation.

$$\Delta\tilde{r}(n, i) = \tilde{r}(n, i) - r_{ref}(n, i), \quad n \in \{N_{RN}, N_{UE_{DL}}, N_{UE_{2H}}\} \quad (10)$$

Based on the derived possible UE specific effective rates for each individual the actual fitness values (FV)s can be derived and the best configurations of one generation can be determined. To calculate the FV two alternative optimization functions are proposed in the next section.

2.1.4 Description of Fitness Value Calculation

Equation 11 defines the first possible fitness function FV_{o1} to use. The equation consists of a sum of sums over all MBSs. At first, the possible deltas of rates for the BHSFs are considered ($\Delta\tilde{r}_{BH}$). As described in Sect.2, the algorithm aims to increase the rates of the RNs, while protecting the rates of the MBS UEs. Therefore, the delta of the RN BH n is weighted by a scalar ω_n , which represents the number of two hop UE to be served by an individual RN. The higher ω_n will be, the more important $\Delta rate$ becomes for the considered RN BH, because the RN needs to serve more two hop UEs. Second, the sum of the deltas of the rates for all MBS UEs during the BHSF is taken into account. It is additionally weighted by the factor $\frac{N_{DL_m}}{N_{2H_m}}$ which represents the importance of the sum. The more MBS UEs are attached to the MBS directly, the more important it will be also to serve these UEs and thus, the rates of them are more important. On the other hand if the RNs attached to the MBSs, N_{RN_m} needs to serve more UEs N_{2H_m} , the sum will be less prioritized. The next sum represents the value of delta rates for the UEs attached directly to the MBS and possibly served during the ALSF. Here an additional weight takes into account the importance of the sum. If a high number of two hop UEs are needed to be served by the RN attached to the donor MBS the more important the sum of the delta rates in the ALSF becomes to counteract on losses for the directly attached UE in the BHSF. Last but not least, the sum of the RN ALSF is considered as well. This aims not to introduce additional interference caused by the RNs when the AL rate is low and therefore, more PRBs would be needed for transmissions. This becomes the more relevant, the more two hop UE need to be served by RNs. Because of that, the additional weight $\frac{N_{2H_m}}{N_{DL_m}}$ gives the sum more importance when more two hop UEs have to be served. Furthermore, the term gets less important if a relatively higher number of MBS UEs are attached and need better conditions to compensate possible losses in the BHSF.

$$\begin{aligned} FV_{o1}(g, i) = & \sum_{m=1}^{MBS} \left(\sum_{n=1}^{N_{RN_m}} \omega_{m,n} \cdot \Delta\tilde{r}_{BH}(m, n) + \frac{N_{DL_m}}{N_{2H_m}} \cdot \sum_{u=1}^{U_{UE_m}} \Delta\tilde{r}_{BH}(m, u) \right. \\ & \left. + \frac{N_{2H_m}}{N_{DL_m}} \cdot \left(\sum_{u=1}^{U_{UE_m}} \Delta\tilde{r}_{AL}(m, u) + \sum_{r=1}^{R_{UE_m}} \Delta\tilde{r}_{2H_m}(m, r) \right) \right) \quad (11) \end{aligned}$$

The second fitness function FV_{o2} which is defined by Eq. 12, is more aggressive to find a good individual. It only focuses on the improvement of the RN BH link and the improvement of the direct UEs during the ALSF to protect them against losses in the BHSF.

Basically the weighted sum for the delta rates of RN BH links is used, as already explained for Eq. 11. Second sum consists of the deltas for the rates of the UE directly attached to MBSs for the ALSF, when the RNs are in sending mode.

$$FV_{o2}(g, i) = \sum_{m=1}^{M_{BS}} \left(\sum_{n=1}^{N_{RN}} \omega_{m,n} \cdot \tilde{\Delta r}_{BH}(m, n) + \sum_{u=1}^{U_{UE}} \tilde{\Delta r}_{AL}(m, u) \right) \tag{12}$$

In addition, the number of links which have a positive $\Delta rate$ are counted for ALSF and BHSF separately. Based on that, a percentage can be derived how many subbands the final individual should be applied in both subframe types to improve the system performance.

Besides that, prohibitive possible outages of UE which might be introduced by the adapted power pattern is taken into account. Therefore, the algorithm checks whether MBS UEs might end up in outage in both possible transmission subframe types. If this is the case, the UE specific $\Delta rate$ is set to -100 , which represents the negative impact on the single UE. Therefore, the total FV of the considered individual is decreased and down prioritized in the parent selection process. The same procedure is applied for RNs in the BHSF. Here, RNs which might result in outage will influence even more the FV of the individual, due to an additional weight, which represents the number of two hop UEs attached to the RN.

2.1.5 Generation of the Offspring

After the calculation of the total FVs for the considered the generation, the offspring needs to be created for the next optimization step. Therefore, half of the individuals with the highest FVs are chosen as potential parents of an offspring. The other half with lower FVs are discarded. For the selection process a fitness proportionate scheme based on Eq. 13 is used, where a probability p_i based on each individuals FV is calculated to act as a parent [23]. The exponential factor α has influence on the probability distribution of the individuals. The higher α is set, the higher the probability the best individuals are chosen as parents. Based on the outcome of a sensitivity test, α was set to 4. This procedure increases the selection probability towards individuals with higher FVs. As an example, Fig. 4 presents the procedure as an illustration of a wheel of fortune. Each parent is chosen in a single process to finally create the target number of parents needed.

$$p_i = \frac{fv_i^\alpha}{\sum_{n=1}^N fv_n^\alpha} \tag{13}$$

Therefore, a random uniformly distributed number between zero and one is determined and compared with the cumulated sum p_{cs} of each individuals' probability, as defined in Eq. 14, for the total number of potentially survived individuals N .

$$p_{cs}(i) = \sum_{n=1}^i p_i(n), \quad i = 1, 2, \dots, N \tag{14}$$

Finally, choose a parent based on Eq. 15, if:

$$p_{cs}(i) \leq x < p_{cs}(i+1), \quad x \in \{\mathbb{R} \mid x \leq 1\} \tag{15}$$

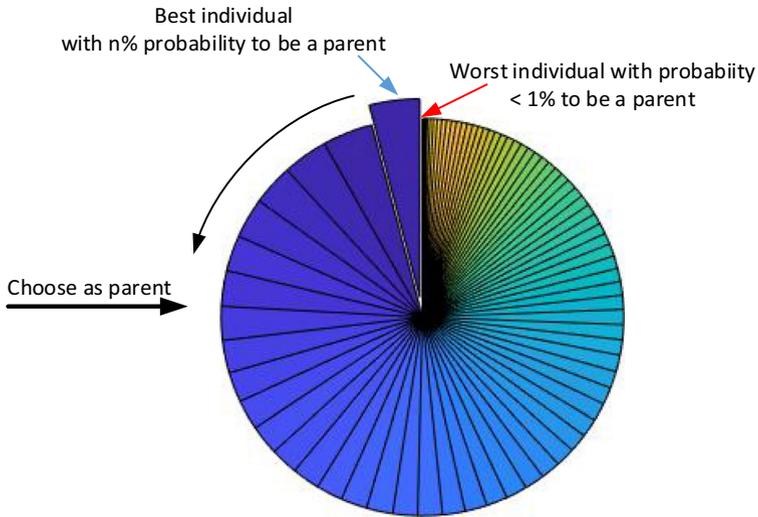


Fig. 4 Fitness proportionate selection to create couples of parents

Once, the parents are determined an one point crossover mechanism is used to create the offspring, as depicted in Fig. 5. Here, two random integer values x, y are chosen, if the number of MBS does not equal the number of RNs in the system. Otherwise one integer would be sufficient. The first value $x \in M_{MBS}$ is taken out of the total number of the MBS in the system, which equals the number of possible crossover points for the individual part

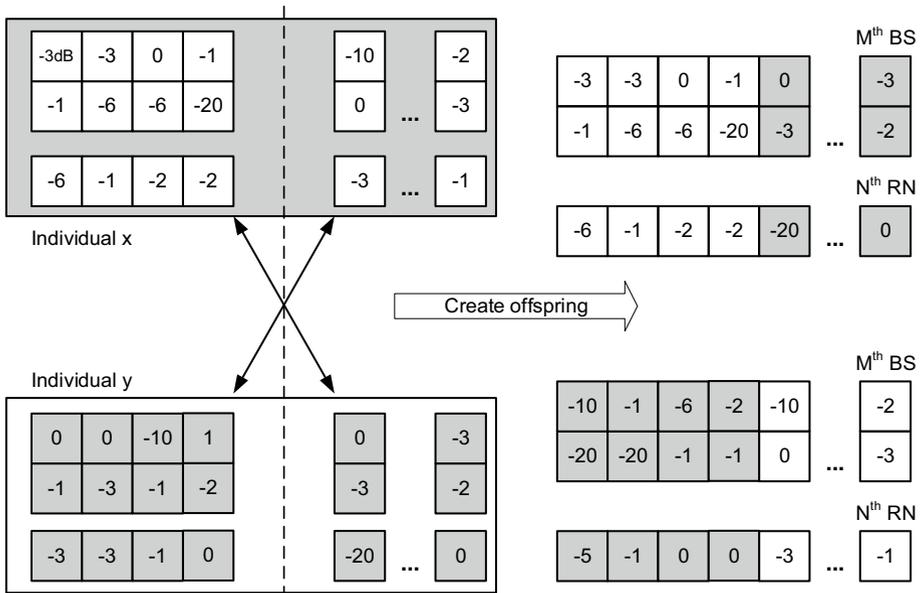


Fig. 5 One point crossover mechanism to create offspring

of the MBSs. Second value $y \in N_{RN}$ is used to find a crossover-points for the PAVs of the RNs. The couple of individuals reproduce themselves as depicted in Fig. 5. For the next couples the process is repeated and different crossover points are determined until the total set of offspring PAVs is created. Finally the offspring needs to be slightly changed to not stuck in a local optimum of the search space, when just reproducing the new set of PAVs based on the legacy. Therefore, a small number of values have to be randomly changed, which is defined as the mutation process. The final outcome of the genetic algorithm is very sensitive to the mutation rate. Therefore, a sensitivity analysis has been carried out to adjust the mutation rate to result in the best possible solution. Figure 6 gives the results of the performed sensitivity analysis. The illustrated adjustment of the mutation rate was done based on Eq. 12 with different percentage of mutation rate. For each generation only the maximum FVs are depicted. It can be observed, that the best found mutation rate was 0.05 (dotted black curve), which means 5% of the total number of PAVs of the generation were mutated. All other tested mutation rates between 1 and 10% didn't result in a higher maximum FV.

Finally, the best occurring individual with the highest FV is determined based on Eq. 16 and saved. It will be only replaced if an offspring generation includes an individual with a higher FV.

$$pav_{\max} = \text{PAV}(\max [FV(g, i)]); \quad (16)$$

2.1.6 Description of the Final Outcome of the Optimization

Finally a promising optimized PAV pattern is found which improves the total weighted sum of rates for a certain percentage of the UEs and RNs. As an numerical example of the optimization result, Fig. 7 shows a configuration found by the GA for two BS and two RNs, based on the final PAVs (individual) with the maximum FV. RN1 transmits with no adapted power, while RN2 has to decrease it by 3 dB in the ALSF. MBS1,

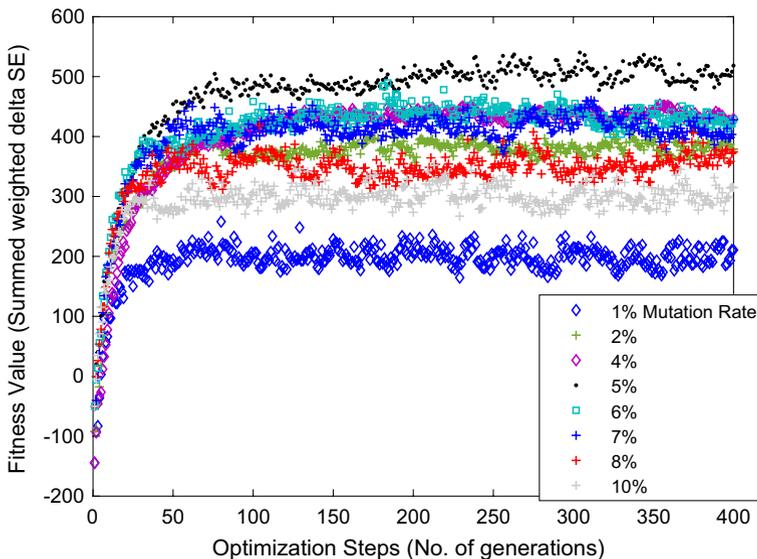


Fig. 6 Sensitivity analysis with different mutation rates

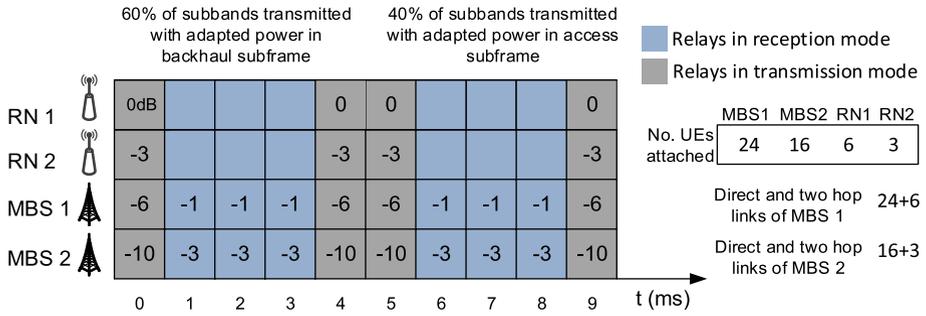


Fig. 7 Numerical example of an optimized transmission power pattern

MBS2 have an decreased transmission power by 6, 10 db respectively, in the ALSF, while decreased power of 1 dB, receptively 3 dB, has to be applied during the BHSF.

Furthermore, Fig. 8 gives an impression how a configuration of the optimized subbands could look like in time domain. As already explained in Sect. 2.1.4 the GA derives a percentage of subbands, based on the number of UEs and RNs which gain from the configuration. Figure 9 illustrates how a configuration could look like in frequency domain. Here, as an example, half of the total system bandwidth is used with the derived optimized power pattern, while the other half is transmitted with the default power settings of each serving node, to schedule UEs which have decreased rates through the applied power pattern.

After the final configuration pattern is found, it is applied and the synchronous adapted proportional fair scheduler takes the newly set decisions into account dependent on defined scheduling strategies as explained in the following section.

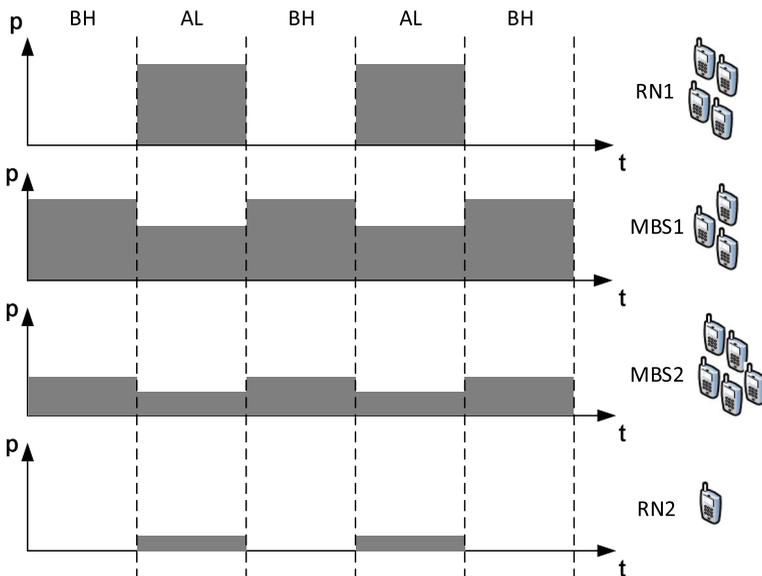


Fig. 8 Example of a possible optimized transmission power configuration over time

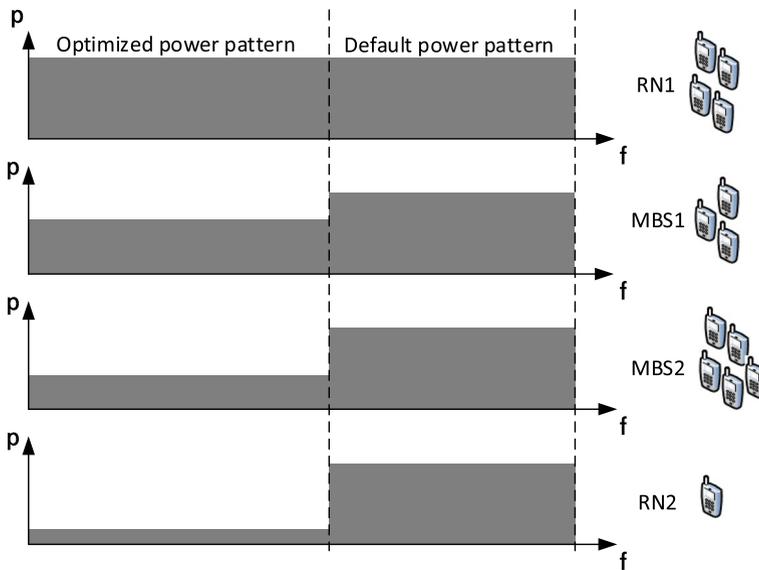


Fig. 9 Example of a possible optimized transmission power subband configuration over frequency

2.2 Decentralized Synchronous Adapted Scheduling

While the optimization process in the previous section aims to improve the BH link of the RNs and thus, the overall network performance by adjusting the transmission power on a subset of subbands for each MBS or RN respectively on a larger time scale, the TTI based scheduling strategy, which will be described in this section, aims to react on short term effects such as small scale fading as well as good performance of the link adaptation to meet the systems target block error rate (BLER). In addition to this a proportionally fair distribution of the radio resources among UE and RNs needs to be fulfilled. Due to the introduced asynchronous optimization, the short term scheduling strategy needs to take into account the imbalanced power levels and thus, larger variations in terms of SINR on the resources. The asynchronous optimization keeps the newly adjusted relation of power levels stable over a period of time, so that the CQI reports from the UEs are not outdated and the appropriate link adaptation can be performed. However, there are different strategies possible, how to make use of the optimized subbands. In the following the applied short term scheduling metric is described and different strategies are defined to allocate resources of the optimized subbands. Further different types of two hop CQI feedback reports are compared which mainly increases the signalling overhead, since the information needs to be collected and fed back from RN to MBS to distribute the radio resources in a proportional fair manner. Finally those are compared by means of system level simulation (SLS) presented in Sect. 3.2.

Figure 10 shows the flow chart of the TTI based adapted two hop proportional fair scheduler. In the first step it is determined if either a BH SF or an ALSF is present. In the second step, either an RN or an MBS is considered in the following scheduling procedure. For simplicity of the figure, the procedure to allocate resources for possible retransmissions is not shown here. In the case of an ALSF and a MBS the transport block size (TBS) of each schedulable UE is determined based on a conventional PF metric, as defined in

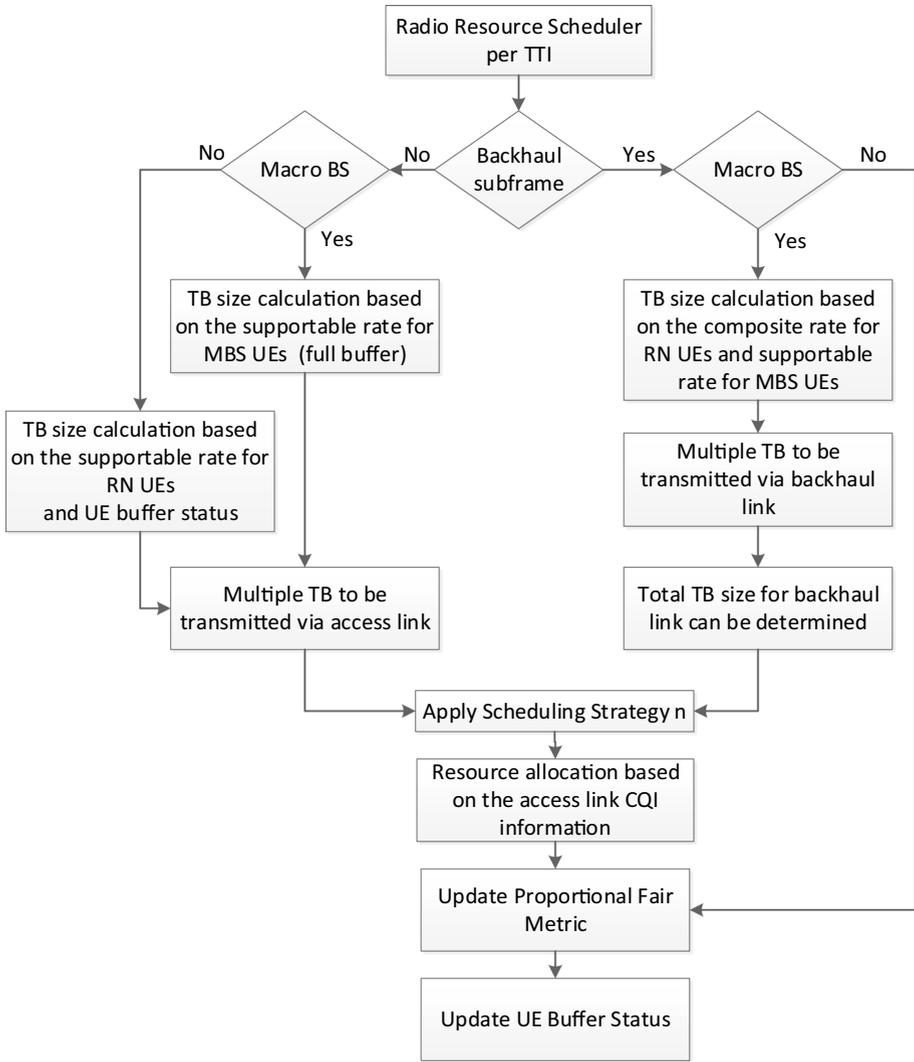


Fig. 10 Flow chart of adapted two hop proportional fair scheduling

Eqs. 17 and 18, based on [24]. All priorities P for each one hop UE k on every PRB n in TTI i are calculated.

$$P_{k,n}(i) = \frac{R_{k,n}(i)}{T_k^\alpha} \tag{17}$$

R is defined as the instantaneous supportable rate, depending on the latest received UE CQI reports on an m_{th} subband. The parameter α is an exponential scaling factor which is set to 1. The average throughput T is recursively updated, as defined in Eq. 18.

$$T_{k,i+1} = \begin{cases} \beta \cdot T_{k,i} & \text{non-scheduled UE} \\ \beta \cdot T_{k,i} + (1 - \beta) \cdot R_{k,i} & \text{scheduled UE} \end{cases} \tag{18}$$

The factor T of the unscheduled UEs/RNs is multiplied with a forgetting factor β . Scheduled UEs' T is updated by multiplying with β and adding the instantaneous data rate of the current TTI with a weight of $1 - \beta$. Based on the calculated priority matrix the PRB allocation is done by taking the UE U with the maximum priority P for each individual PRB, defined in Eq. 19.

$$U_n(i) = \max_k(P_{k,n}(i)) \tag{19}$$

While the MBS to UE transmission is done with infinite full buffer size, in the case of a RN to UE transmission in an ALSF only the data is transmitted delivered via MBS-RN BH transmission (most left case in Fig. 10). Therefore the resource allocation is done under each UEs buffer limitation. Equation 19 is only done if the buffer size of the UE is larger than or equals the total TBS of the UE. Otherwise the considered UE is excluded by setting the UE specific priorities on the left PRBs to $-\infty$. If the buffers of the left schedulable UEs are smaller as well, not all PRBs are used for transmission.

In case of a BHSF transmission the available PRBs at the MBS need to be shared among one hop UEs and RNs.

$$CR_{j,n}(i) = \left(\sum_{l=1}^L \frac{1}{R_{l,j,n}(j)} \right)^{-1} \tag{20}$$

Therefore, in Eq. 20 the composite rate (CS) replaces the instantaneous rate R of Eq. 17 for all two hop UE j , served by the considered RN, based on [22]. The specific two hop connection takes into account both, the instantaneous rates R of the BHSF and the ALSF. The harmonic mean of both hops (L) considers the loss in time of the transmission, as well. Finally, two priorities are calculated with the throughput T of the AL, T_{AL} and the BH link T_{BH} . The maximum is used as a priority of the two hop UEs. In most of the cases the priority of the AL might be used, because the UE individual past throughput might be equal or less then the total past throughput of the BH link.

$$P_{j,n}(i) = \max \left(\frac{CR_{j,n}(i)}{T_{j,AL}^\alpha}, \frac{CR_{j,n}(i)}{T_{j,BH}^\alpha} \right) \tag{21}$$

Finally, the calculated priorities for one and two hop UEs are compared and allocated to the UEs with the maximum priority on each PRB, as in Eq. 22.

$$P_{k,n}(i) = \max (P_{k,n}(i), P_{j,n}(i)) \tag{22}$$

Once, the PRBs are allocated either to one or two hop UE, the TBS can be calculated for each UE. After that, a resource reallocation needs to be done for the MBS-RN transmission with one hop feedback to get accurate frequency selective scheduling decision. Therefore the total amount of TBSs of the considered two hop UEs are summed up and the needed PRBs for one large BH TBS is derived. Furthermore, the buffer status of the two hop UEs as well as the proportional fair metric is updated afterwards.

2.2.1 Overhead Consideration and Resource Allocation for Different Types of Channel Quality Feedback

When frequency selective scheduling is applied, typically subband CQI reports are used to gain knowledge of the channel states and derive a resource allocation decision

with better resource utilization. Typically, in conventional networks the UE takes measurements and send back the report to the MBS where the MAC scheduler uses the reports to take a decision. In RN extended networks, it is additionally necessary to feedback the reports from RN to MBS for all UEs attached to the RN. This can heavily increase the additional signalling overhead in the uplink (UL) and influence the performance. Besides that, the control signalling overhead is already increased, as different subframe types occur and the MBS needs accurate information of both shadow fading (SF) types. For instance, if periodic CQI reporting is assumed, the UE needs to generate at least two CQI reports per frame, one for each occurring SF type, where different interference situations appear and additionally different optimized power patterns are applied. The reports consist of a defined number of subband reports including MCS recommendation among other information. The reports of all two-hop UEs need to be collected and sent back to the MBS. To reduce this potential large amount of overhead, different types of reports with reduced amount of overhead are compared in this study. The reference case assumes a full subband report for frequency selective decisions as input for Eq. 20. Alternative RN to MBS feedbacks with a potential overhead reduction of are used and compared in terms of throughput, SINR and fairness. Instead of frequency selective subband CQI reports either the recommended maximum, average supportable rate derived out of the CQI reports for the RN access link is fed back to the MBS. This reduces the amount of signalling by $\frac{1}{N}$, where N is the defined number of subband reports per UE. For further clarification it should be kept in mind, that the actual scheduling decision at the RN is still done in a frequency selective manner based on RN UEs' CQI feedback to the RNs.

2.2.2 Scheduling Strategies for Optimized Subbands

As already mentioned, different scheduling strategies can be applied, how the optimized subbands can be allocated to UE. Advantages and drawbacks of different strategies are discussed in the following and compared by means of SLS in Sect. 3.

Scheduling strategy 1:

Only UE with previously calculated chance to gain from optimized subband are allowed to use such resources, based on UE specific individual positive delta of the rates. If no UE with potential benefit is attached to MBS or RNs the resources will be left unallocated. This results in a better SINR in neighbouring cells on the one hand, but higher loss in terms of unallocated radio resources on the other.

Scheduling strategy 2:

UE with negative Δ rates can potentially also make use of the optimized subband. Dependent on the adapted two hop proportional fair, those UE will be less prioritized on those PRBs. For a single UE it is possible to allocate the total amount of PRBs independent if transmitted with optimized or default power. This results in a higher variation of SINR values for a single UE transmission and thus in potentially less efficient link adaptation.

Scheduling strategy 3:

UE with less rate can also make use of the optimized subband. Dependent on the adapted two hop proportional fair, those UE will be less prioritized. For a single UE it is only possible to allocate the optimized or the default subband.

3 System Level Simulations

In this section, the performance analysis of the hybrid RRM scheme is presented. First, the most important simulation assumptions are summarized, followed by the results.

3.1 Assumptions

In each sector 4 RNs are placed randomly with a minimum distance of 40 m to each other. Each cell sector is equipped with a 3D directional antenna with a 3 dB horizontal beamwidth of 70° and a vertical 3 dB HPBW of 10° . A sub-urban scenario with the commonly used macro inter-site distance of 1732 m is used. Transmitting antennas of the RN and the receiving antenna of the RNs/UEs are omni directional. No additional antenna gains for transmission and reception at the RNs is assumed which makes the solution more attractive for operators due to potential cost savings not deploying additional antennas at the RN sites. The RNs have a transmission power of 30 dBm. A hotspot UE distribution with 30 UEs in average per cell is considered. A hotspot occurs in the vicinity of the RNs. As an example, Fig. 11 shows one UE and RN drop of the simulated RN extended LTE-A network with 19 hexagonal cell sites and 3 sectors per site. The cell selection of the UEs is based on the comparison of one and two hop default rates based on Eqs. 8 and 9, which equals the end-to-end optimal routing strategy proposed in [25]. To prevent border effects, the wrap around technique is applied. WINNER+ sub urban macro (SUMa) channel model is used for the MBS-UE direct links (DL) and the MBS-RN BH links (BH). For

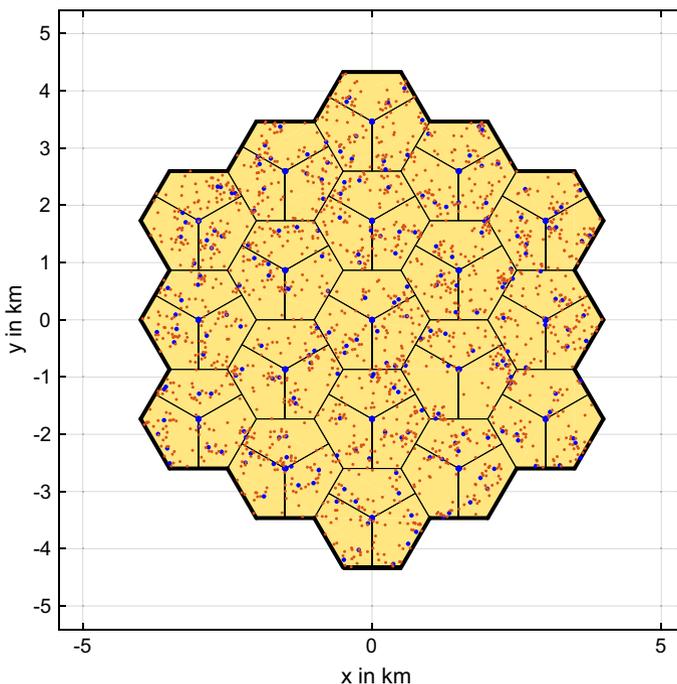


Fig. 11 Network layout, 19 MBSs, 3 sectors/site, e.g. 4 RNs per sector

the RN-UE AL urban micro (UMi) channel model is used. LoS probability of the MBS-RN BH link is only slightly increased an assumed deployment height of 5 m for the RNs. Furthermore, 256 quadrature amplitude modulation (QAM) transmission is supported based on [18]. Additional control channel overhead is assumed using 3 OFDM symbols (1 DL R-PDCCH, 1Tx-Rx, 1Rx-Tx switching per SF). It has to be kept in mind, that the chosen scenario is a very challenging, but one of the most realistic scenario for RN deployments. On the one hand RNs deployed at UE hotspots give a very high probability to provide excellent AL channel quality. On the other hand, due to the random placement of the UE hotspots and the RNs in its vicinity it can happen that the hotspot occurs in the main beam direction of the MBS antenna. In such scenarios it is quite unlikely that RNs might outperform the direct link performance, even if in such scenarios the BH link quality will be excellent as well.

Comparable results as a reference can be found in [26] based on the agreed assumptions in the 3GPP [27]. Typical performance gains by introducing relays with decentralized scheduling schemes are approximately an average user throughput of around 20% when 4 RNs per sector are deployed in an urban or sub-urban scenario [28]. However, one of the major contributions of this work is to improve the backhaul link quality of the deployed RNs without the deployment of an additional receiving antenna to save additional deployment costs and planning effort. This assumption makes the scenario more challenging and also difficult to compare with other existing results in the literature. In addition, the simulation results are quite sensitive and dependent on their underlying assumptions. In a realistic urban scenario an additional directed receiving antenna requires LoS connection to the donor MBS. Otherwise, the SINR might be decreased in an unwanted manner and possible theoretical gains are not possible due to a received signal with a high delay spread [29, 30]. This might be difficult to achieve in a realistic scenario due to limited possible RN site locations in the vicinity of a user hotspot, when RNs are deployed to increase the capacity of the network. Furthermore, possible deployment heights are limited as well in realistic scenarios. This typically decreases the LoS probability as well in the applied PL models and is often overseen in existing literature. For instance in [31] an additional RN antenna gain of 13 dBi and an RN antenna height of 15 m is assumed. In [13] an additional directional receive antenna with 7 dBi antenna gain is assumed. Nearly comparable results can be found in [15, 17, 32], with slightly deviating assumptions in terms of the deployed scenario and the considered number of users. In principle the centralized part of the proposed hybrid scheme could be applied in the considered scenarios and an additional gain could be achieved when using the solutions in a combined manner. In Fig. 15 the final result is shown of the proposed solution. In the first step a macro cell network is compared with a RN extended system. As stated before an additional gain of 18% could be reached in the average overall user throughput under fairness constraints comparable with the achieved gains in [26, 28]. When the centralized part of the hybrid RRM scheme is enabled an additional average gain of 10% could be observed. This makes it an attractive solution as no directed antennas have been applied in receiving or sending direction at the RNs. In a nutshell the chosen scenario in this work is one of the most realistic but also one of the most challenging ones to outperform the conventional MBS network and increase the performance by the hybrid RRM approach in RN extended networks. Furthermore, realistic non prohibitive low amount of feedback overhead is assumed, which limits the possible gains of the approach but shows trustworthy results. This is also often assumed way too optimistic in the existing literature, as summarized in Sect. 1.

Within the applied 2x2 MIMO system a brute force algorithm is used to create the CSI report of each UE. Based on the best derived SINR of all possible transmission modes and

codebook based precoding matrices the CQI, RI and PMI indices are chosen and the CSI report is created. A feedback delay of 5ms is assumed until the report can be taken into account at the serving node (MBS or RN). UEs attached to MBS are instructed to generate 2 feedback types, during RN reception and transmission. UEs served by RNs generate 2 feedback types as well, for TTIs in which the RNs are muted or not. This gives more accurate information due to different interference power levels. Considering the proposed GA defined in Sect. 2.1, the number of generations and individuals, crossover type, mutation rate and applied fitness functions are set, as listed in Table 1.

A comparison of the resulting UE throughput, SINR, fairness and potential energy savings on the DL physical shared channel is described in the following.

3.2 Results

In the reference simulation a conventional MBS network is considered. In the next step, the scheduling strategy is based on the decentralized adapted two hop proportional fair metric, defined in Sect. 2.2, while the centralized heuristic algorithm is disabled. To reduce RN signalling to the MBS a comparison is carried out for the different aforementioned signalling information. Finally the centralized algorithm is additionally enabled based on the defined fitness functions based on Eqs. 11 and 12, described in detail in Sect. 2.1 and further defined as o1 and o2.

3.2.1 Comparison with Different Two Hop UE Feedbacks Forwarded from RN to MBS

As described in Sect. 2.2 the forwarded additional signalling is defined as the full CQI report which includes frequency selective information (FSS), while the reduced signalling is either the maximum (MAX) or average (AVG) recommended MCS of the individual two hop UE. The figures in 12 show clearly the higher performance for the heterogeneous network compared to the reference. An average gain of approx. 18% can be observed at the 60 percentile of the CDF. The performance is kept stable, when the lower signalling overhead is used. Additionally it can be observed, that the in case of the results with AVG feedback slightly improves the performance below 60 percentile, while decreasing a bit above it. This may come from the fact that RN UEs are slightly higher prioritized, when the MBS assigns the resources to either direct UE or RN. This is also confirmed due to the higher RN BH throughput illustrated as the black curves. In addition it can be observed in Fig. 12 that the RN UEs experience a higher throughput in case of AVG or MAX feedback while

Table 1 Summary GA parameters

Parameter	Value
Fitness function type	Equation 12
Number of generations	400
Number of individuals	160
Parental selection	Fitness proportionate
Crossover inheritance type	Random one point
Mutation rate	0.05
TX power adaptation values (PAV)	[- 46, - 40, - 30, - 20, - 10, - 6, - 5, - 4, - 3, - 2, - 1, 0] dB

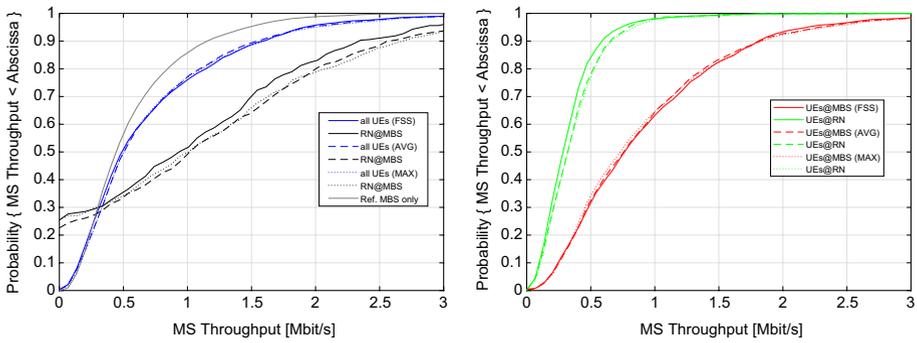


Fig. 12 Throughput of UE groups and RN BH with different feedback of RN AL CQI

the MBS UEs' throughput is slightly decreased. Furthermore, the RN UEs throughput is well below then for the MBS UEs since only the worst UEs are served by the RNs, keeping the applied end-to-end optimal routing strategy in mind (Fig. 13).

3.2.2 Throughput Comparison with Different Fitness Functions and Scheduling Policies

In the next step the centralized heuristic is applied. In the first part fitness function o1 is used. Unfortunately, the optimization approach doesn't clearly outperform the previous explained results. Due to the design of fitness function o1, there are too many counteraction terms in the equation, which wipe out possible gains. In this case, the different scheduling strategies based on Sect. 2.2.2 does not show significant influence in the performance. It can be concluded that, the centralized heuristic did not derive a proper power reduction pattern. However, the performance is kept stable and at least no increased throughput is observed.

As an alternative, fitness function o2 is now enabled and presented. As already described in Sect. 2.1 the design of the fitness function is much more aggressive to find a promising subband power reduction pattern. Only RN BH improvement in the BH subframe and possible MBS UE improvements in the RN AL are in focus without consideration of possible losses of MBS and RN AL in both subframes, respectively.

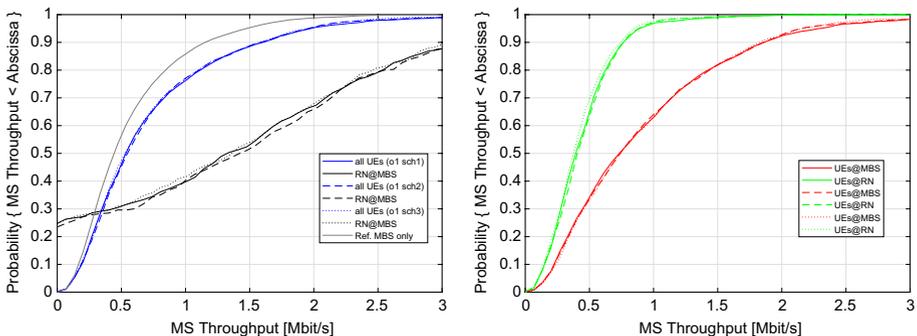


Fig. 13 Throughput of UE groups and RN BH with fitness function o1 and scheduling policy 1, 2 and 3

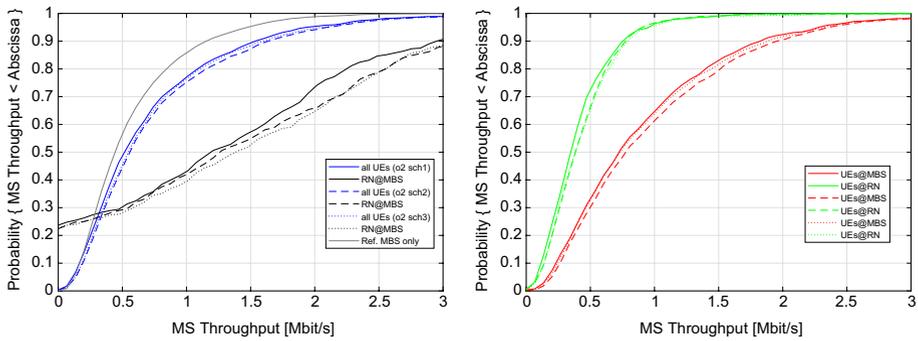


Fig. 14 Throughput of all UE groups and RN BH with fitness function o2 and scheduling policy 1, 2 and 3

Figure 14 shows the performance comparison when fitness function o2 is applied in combination with the proposed scheduling strategies. It can be observed the best performance when scheduling strategy 2 is used. To remind the reader strategy one leaves the radio resources free, when no UE individual Δ rate promises possible gains in the considered cell. On the one hand this might improve the SINR in the surrounding cells but decreases the number of available resources to be scheduled to UEs. Scheduling strategy 2 provides access to all available resources. However a single UE can only be scheduled either in the power adapted or default power resource region. In contrast to this scheduling strategy 3 gives the opportunity to allocate resource to a single UE from both resource regions. As can be observed scheduling strategy 2 outperforms 1 and 3. A closer look discloses, that scheduling strategy 3 slightly outperforms one. The difference between 2 and 3 is explained by the higher SINR variance in case of strategy 3 for a single UE. The link adaptation targets a BLER of 10 percent to provide best performance of the network. In case of strategy 3 the target BLER was narrowly missed. In Fig. 14 it can be observed, that especially in case of the MBS UE performance strategy 2 outperforms 1 and 3. For RN UEs the difference between strategy 2 and 3 can be neglected, as the RN UEs are buffer limited and do not use both bandwidth parts simultaneously due to smaller TBSs.

3.2.3 Final Comparison for Best Feedback, Power Reduction Pattern and Scheduling Settings

Figure 15 show the final results of the comparison. The UE throughput for all distinguished UE groups is shown. It can be observed, that a much more data is transmitted to the RNs when the GA was applied (black solid vs. dashed curves). A clear throughput gain can be observed for all UE when the hybrid approach with fitness function o2 and scheduling strategy 2 is applied. At 60 percentile which represents the average throughput, approx. 10 percent gain is reached, without any loss in other regions. From 0 to 100 percentile the CDF runs on the right side compared to the decentralized approach. This results gives a clear impression that the genetic algorithm optimized the BH link quality (improved throughput of RN UEs) and additionally compensated possible losses for the macro UE in the second subframe type when RNs are receiving. Even for MBS UEs the performance is increased.

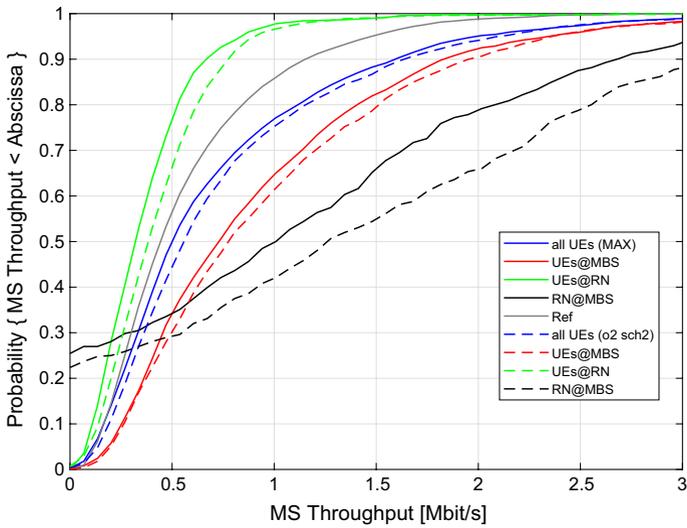


Fig. 15 Final comparison of the UE and RN throughput for reduced feedback based on max value and fitness function o2 with scheduling policy 2

3.2.4 SINR Comparison for Best Settings

To give some more insides how the system behaves the SINR values for the best derived hybrid settings (MAX feedback, fitness function o2 with scheduling strategy 2) are analysed. The SINR values are calculated on 3 subcarriers per PRB depending on the applied power per PRB, used transmission mode (spatial multiplexing or single UE beamforming), channel variations in time and frequency and codebook based precoding vectors (PMI) used for each individual transmission. Here, the CDF shows the individual subcarrier related SINR per link type. Figure 16 shows the observed SINR values for the distinguished UE groups. A clear gain of approx. 2.2 dB in average (60 percentile) for the

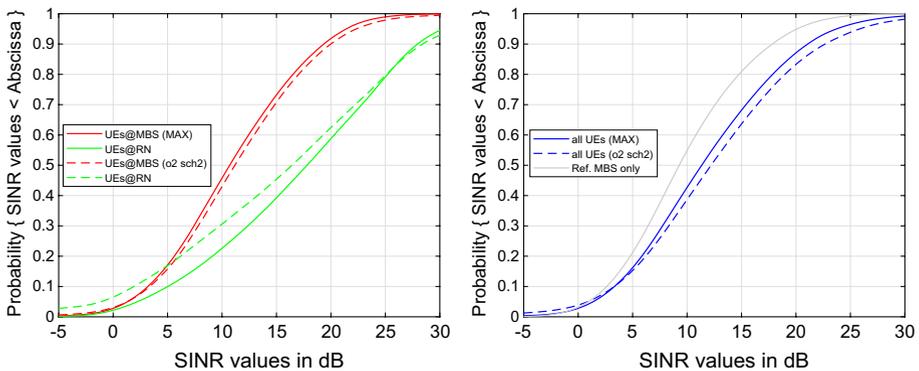


Fig. 16 Comparison of the UE SINR values for reduced feedback based on MAX feedback values and fitness function o2 with scheduling policy 2

BH can be observed. The dashed dark blue CDF (fitness function 2, scheduling strategy 2) clearly experience a better improvement than the solid curves, which represents the decentralized approach. As required no degradation for the MBS UEs (red) can be seen. Even a slightly SINR improvement of roughly 0.5 dB is observed as mean value (approx. 60 percentile) for the GA with scheduling strategy 2. For the RN UEs decreased SINR values are observed below 80 percentile. This comes from the defined fitness function, which tries to compensate the potential loss in the BH subframe for macro UEs. Above 80 percentile, the SINR it is slightly improved. However, the loss in the RN access link will not result in a huge performance degradation, due to the applied reuse one scheme, which means that all RNs have access to the total system bandwidth to serve their UEs. Since the BH subframe, when MBS UEs and RNs have to share the resources it is most likely the case, that only a limited amount of radio resources are allocated by the RN UE. Thus, the lower SINR will be compensated with higher amount of resources in use. In addition an increased SINR is observed compared to the reference MBS network. This is also confirmed considering Fig. 15. Due to the improved SINR on the BH link, the RN is higher prioritized to be scheduled and thus, the RN UEs profit from that in terms of increased throughput. Only below 5 percentile the SINR is decreased due to the aforementioned effect on the RN UE links, without any loss in the throughput performance (Fig. 17).

3.2.5 Fairness Comparison and Energy Savings for Best Settings

The fairness index according to the 3GPP criteria is defined in [27]. It is defined as the normalized UE throughput CDF with respect to the average. If it proceeds on the right side of the identity function the system performance fulfils a fair throughput distribution among UE. The limit defines a linear relation between the UE throughput and the probability to experience a certain throughput. For instance 50% of the time the UE should perceive 50% or less of the

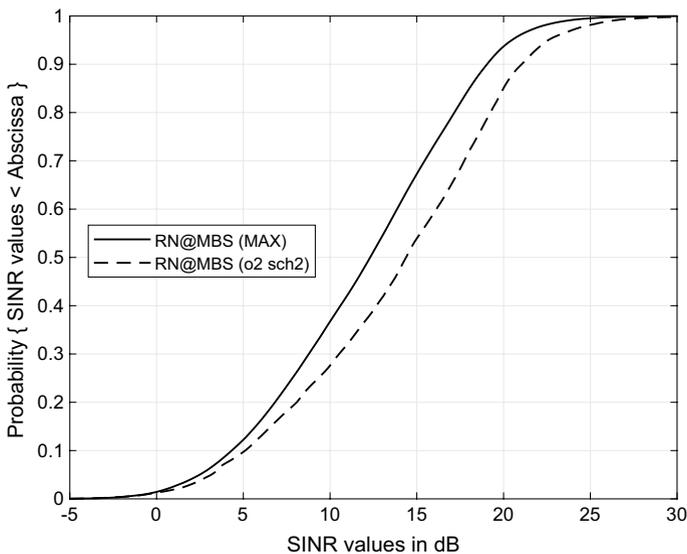


Fig. 17 Comparison of the RN BH SINR values for reduced feedback based on MAX feedback values and fitness function o2 with scheduling policy 2

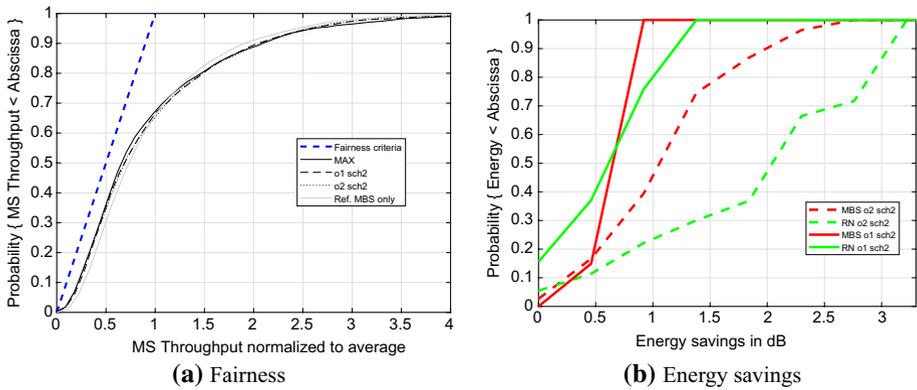


Fig. 18 Fairness comparison and energy savings

average UE throughput. In Fig 18a the fairness evaluation of the investigated system is presented. Compared to the reference network (light grey) the RN extended network has a slightly higher variance of throughput values, since it is a bit flatter. All example RN network settings shown here fulfil the fairness requirement of 3GPP.

Due to the asynchronous optimization as defined in Sect. 2.1 smaller amount of energy is used for transmission on a number of PRBs in the downlink data channel. Dependent on the MBS or RN specific power adaptation possible energy savings can be reached. Figure 18b shows the CDF with the probability of potential power savings for MBS, RNs respectively. Here, the best found setting (fitness function o2 with scheduling policy 2) provides also the highest power savings. Approx. up to 2.7 dB for MBS as well as 3.2 dB for RN transmissions could be saved.

4 Conclusions

Existing resource allocation schemes for RN extended networks fail to combine multiple targets. The hybrid scheme is separated in the decentralized part to allocate the resources in a two hop proportional frequency selective manner with co-scheduling of UEs and RNs. In the centralized part adapted subband power allocation is improved by reducing the transmission power. The scheme unites multiple objectives under detailed modelling of the wireless BH link. While improving the spectral efficiency of the BH link under fairness constraints and the limitation of possible co-scheduled single hop transmissions, it reduces the interference in the system by only reducing the transmission power on the shared channel. For the centralized part only very limited feedback is required, while for the decentralized part the necessary feedback on two hop access link is minimized. Simulation results show performance gains with stable fairness and savings in energy consumption.

Open Access This article is distributed under the terms of the Creative Commons Attribution 4.0 International License (<http://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution,

and reproduction in any medium, provided you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made.

References

- Osseiran, A., Boccardi, F., Braun, V., Kusume, K., Marsch, P., Maternia, M., et al. (2014). Scenarios for 5G mobile and wireless communications: The vision of the METIS project. *IEEE Communications Magazine*, 52, 26–35.
- Paolini, M. (2011). White paper crucial economics for mobile data backhaul. <https://cbln.com/sites/all/files/userfiles/files/CB-002070-DCLATEST.pdf>.
- Hoymann, C., Chen, W., Montojo, J., Golitschek, A., Koutsimanis, C., & Shen, X. (2012). Relaying operation in 3GPP LTE: Challenges and solutions. *IEEE Communications Magazine*, 50, 156–162.
- 3GPP. (2010). TR 36.806 v9.0.0, Relay architectures for E-UTRA (LTE-Advanced), Tech. Rep. Release 9.
- 3GPP. (2018). 3GPP TR 38.874 V0.2.1 (2018-05) Study on integrated access and backhaul; (Release 15).
- Li, G., & Liu, H. (2006). Resource allocation for OFDMA relay networks with fairness constraints. *IEEE Journal on Selected Areas in Communications*, 24, 2061–2069.
- Nam, W., Chang, W., Chung, S.-Y., & Lee, Y. H. (2007). Transmit optimization for relay-based cellular OFDMA systems. In *IEEE international conference on communications* (pp. 5714–5719). IEEE.
- Salem, M., Adinoyi, A., Rahman, M., Yanikomeroglu, H., Falconer, D., Kim, Y.-D., et al. (2010). An overview of radio resource management in relay-enhanced OFDMA-based networks. *IEEE Communications Surveys & Tutorials*, 12(3), 422–438.
- Lee, Y. L., Chuah, T. C., Loo, J., & Vinel, A. (2014). Recent advances in radio resource management for heterogeneous LTE/LTE-A networks. *IEEE Communications Surveys & Tutorials*, 16(4), 2142–2180.
- Shariat, M., Pateromichelakis, E., Quddus, A. U., & Tafazolli, R. (2015). Joint TDD backhaul and access optimization in dense small-cell networks. *IEEE Transactions on Vehicular Technology*, 64, 5288–5299.
- Han, J. A., & Jeon, W. S. (2009). Proportional fair scheduling combined with adjustment of two hop transmission time for relay-aided OFDMA systems. In *International conference on wireless communications & signal processing* (pp. 1–5). IEEE.
- Liu, C., Qin, X., Zhang, S., & Zhou, W. (2011). Proportional-fair downlink resource allocation in OFDMA-based relay networks. *Journal of Communications and Networks*, 13, 633–638.
- Zhao, Z., Wang, J., Redana, S., & Raaf, B. (2012). Downlink resource allocation for LTE-advanced networks with type1 relay nodes. In *IEEE vehicular technology conference* (p. 1).
- Jeon, W. S., Han, J. A., & Jeong, D. G. (2014). Distributed resource allocation for multi-cell relay-aided OFDMA systems. *IEEE Transactions on Mobile Computing*, 13, 2003–2015.
- Liebl, G., de Moraes, T. M., Soysal, A., & Seidel, E. (2011). Fair resource allocation for inband relaying in LTE-advanced. In *8th International workshop on multi-carrier systems & solutions* (Vol. 1, pp. 1–5). IEEE.
- Wang, L., Du, Q., Ren, P., Sun, L., & Wang, Y. (2014). Buffering-aided resource allocation for Type I relay in LTE-advanced cellular networks. In *IEEE global communications conference* (Vol. 20110201120014, pp. 4484–4489). IEEE.
- Sun, C., Wang, W., Zhang, Y., & Wang, X. (2016). Distributed two-hop proportional fair resource allocation in long term evolution advanced networks. *Wireless Communications and Mobile Computing*, 16, 264–278.
- 3GPP. (2013). 3GPP TR 36.872 Small cell enhancements for E-UTRA and E-UTRAN-Physical layer aspects (Release 12) Tech. rep., 3GPP.
- Arnold, P., Rakocevic, V., & Habermann, J. (2016). Simulation study on 2 hop-proportional fair scheduling and eCIC in relay type 1 extended LTE-A networks. In *IEEE 82nd vehicular technology conference, VTC Fall 2015—Proceedings* (pp. 0–4).
- Bulakci, O., Redana, S., Raaf, B., & Hamalainen, J. (2011). Impact of power control optimization on the system performance of relay based LTE-advanced heterogeneous networks. *Journal of Communications and Networks*, 13, 345–359.
- Mogensen, P., Na, W., Kovács, I. Z., Frederiksen, F., Pokhariyal, A., Pedersen, K. I., Kolding, T., Hugl, K., & Kuusela, M. (2007). LTE capacity compared to the Shannon bound (Vol. 1, pp. 1234–1238).
- Ahn, W.-G., & Kim, H.-M. (2008). Proportional fair scheduling in relay enhanced cellular OFDMA systems. In *IEEE 19th international symposium on personal, indoor and mobile radio communications* (Vol. 2, pp. 1–4). IEEE.
- Thierens, D., & Goldberg, D. (1994). Convergence models of genetic algorithm selection schemes. In Y. Davidor, H. P. Schwefel, & R. Männer (Eds.), *Parallel problem solving from nature — PPSN III. PPSN 1994. Lecture Notes in Computer Science* (Vol. 866). Berlin: Springer.
- Motorola. (2006). 3GPP R1-060877 frequency domain selective scheduling for E-UTRA, Tech. rep.

25. Park, W.-H., & Bahk, S. (2009). Resource management policies for fixed relays in cellular networks. *Computer Communications*, 32, 703–711.
26. Yuan, Y. (2013). *LTE-advanced relay technology and standardization. Signals and communication technology*. Berlin: Springer.
27. 3GPP. (2010). TR 36.814 v9.0.0, Further advancements for E-UTRA physical layer aspects.
28. 3GPP, P. 3GPP, and 3GPP. (2010). R1-101273-Downlink relay performance evaluation.
29. Coletti, C., Mogensen, P., & Irmer, R. (2011). Performance analysis of relays in LTE for a realistic suburban deployment scenario. In *IEEE vehicular technology conference*.
30. Coletti, C., Mogensen, P., & Irmer, R. (2011) Deployment of LTEin-band relay and micro base stations in a realistic metropolitan scenario. In *IEEE vehicular technology conference* (pp. 0–4).
31. Reddy, M. V. V., Vivier, E., & Kaddour, F. Z. (2013). Joint benefits of fractional frequency reuse and relays in LTE networks. In *IEEE online conference on green communications (OnlineGreenComm)* (pp. 127–131). IEEE.
32. Liebl, G., De Moraes, T. M., Soysal, A., & Seidel, E. (2012). Fair resource allocation for the relay backhaul link in LTE-advanced. In *IEEE wireless communications and networking conference, WCNC* (pp. 1196–1201). IEEE.

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Paul Arnold works as a research engineer at Deutsche Telekom AG, Darmstadt, Germany. He holds a Dipl.-Ing degree in Information and Communication Techniques from the University of Applied Sciences Mittelhessen, Friedberg, Germany and is a Ph.D. candidate in Telecommunications from the University of London. He has been working at Deutsche Telekom since 2007, where he has been involved in several Horizon 2020 EU funded projects such as E3, METISII, 5GNORMA and 5GMonArch. His research interests include wireless networks, interference coordination techniques, radio resource scheduling, simulation processes and protocols for radio access networks.



Veselin Rakocevic SMIEEE works as Reader in Electronic Engineering at City, University of London, United Kingdom. He holds a Dipl.-Ing degree in EEE from the University of Belgrade, Serbia, and a Ph.D. in Telecommunications from the University of London. He has been working at City, University of London since 2002, where he has been a Head of the Department of Electrical and Electronic Engineering in the period 2015–2017. His research interests include wireless networks, network optimisation, simulation processes and protocols for networks of moving objects. He participated in numerous research projects during his career and published widely.



Joachim Habermann (Senior Member IEEE) received both Diploma and Dr.-Ing. degree in Electrical Engineering from Technical University of Darmstadt, Germany. He then joined ABB research Centre in Baden Switzerland, where he worked in the Telecommunications department in the area of mobile communications. He is now a professor at the Technische Hochschule Mittelhessen (Fachhochschule), University of Applied Sciences.