



Qi, Y., Shakir, M. Z., Imran, M. A., Quddus, A., and Tafazolli, R. (2016) How to Solve the Fronthaul Traffic Congestion Problem in H-CRAN? In: 2016 IEEE International Conference on Communications Workshops (ICC), Kuala Lumpur, Malaysia, 23-27 May 2016, pp. 240-245. ISBN 9781509004485.

There may be differences between this version and the published version. You are advised to consult the publisher's version if you wish to cite from it.

<http://eprints.gla.ac.uk/132998/>

Deposited on: 1 March 2017

Enlighten – Research publications by members of the University of Glasgow
<http://eprints.gla.ac.uk>

How to Solve the Fronthaul Traffic Congestion Problem in H-CRAN?

Yinan Qi¹, Muhammad Z. Shakir², Muhammad A. Imran³, Atta Quddus³, and Rahim Tafazolli³

¹Samsung Electronics R&D Institute UK, Staines, Middlesex TW18 4QE, UK

Email: naiose.k@googlemail.com

²Dept. of Systems and Computer Engineering, Carleton University, Ottawa, Canada

Email: muhammad.shakir@sce.carleton.ca

³5GIC, University of Surrey, Guildford, Surrey, United Kingdom, GU2 7XH

Email: {m.imran, a.quddus, r.tafazolli}@surrey.ac.uk

Abstract—The design of efficient wireless fronthaul connections for future heterogeneous networks incorporating emerging paradigms such as heterogeneous cloud radio access network (H-CRAN) has become a challenging task that requires the most effective utilization of fronthaul network resources. In this paper, we propose and analyze possible solutions to facilitate the fronthaul traffic congestion in the scenario of Coordinated Multi-Point (CoMP) for 5G cellular traffic which is expected to reach ZetaByte by 2017. In particular, we propose to use distributed compression to reduce the fronthaul traffic for H-CRAN. Unlike the conventional approach where each coordinating point quantizes and forwards its own observation to the processing centre, these observations are compressed before forwarding. At the processing centre, the decompression of the observations and the decoding of the user messages are conducted in a joint manner. Our results reveal that, in both dense and ultra-dense urban small cell deployment scenarios, the usage of distributed compression can efficiently reduce the required fronthaul rate by more than 50% via joint operation.

Keywords— *HetNet; H-CRAN; fronthaul; Coordinated Multi-Point; joint decompression & decoding.*

I. INTRODUCTION

Every year, mobile data traffic volume increases astronomically due to the recent widespread ubiquitous use of mobile broadband services together with the advent of many smart phones, super-phones and tablets with powerful multimedia capabilities and applications. Moreover, new data services and applications, e.g., emerging 3D multimedia applications, are making the experience of using the mobile broadband better and more exciting. Moreover, the integration of social networking and online video sharing inspires an unprecedented new market - social video which plays an important role on the interplay of social media and online videos. Social video is usually bandwidth-demanding applications, especially for HD-Video or 3D-Video. A report has predicted that 75% of all online videos will be HD by 2015, in addition to the ever-increasing length of videos [1]. The increasing trends in mobile data traffic are expected to maintain the momentum over the next decade and will be complemented by billions of connected devices involved in Machine-Type Communications (MTC). Extrapolation of growing mobile broadband traffic predicts that the future networks are required to be prepared to support up to a thousand fold increase in total mobile broadband traffic by

2020 [1]-[3]. This prediction assumes a ten-fold increase in mobile broadband subscriber and up to 100 times higher traffic per user. As a result, the current and projected growth of mobile traffic demands the development of the 5th Generation (5G) mobile communications technology.

A. Heterogeneous Networks

5G involves capacity enhancements in multiple dimensions, among which, heterogeneous network (HetNet) is emerging as a promising feature where the network is densified by deploying multiple small cells to improve the coverage and capacity. Small cells are low power, low cost user/operator deployed cells to complement the existing network. HetNet is not about just adding small cells to the networks, rather it is a paradigm shift network densification technology to optimize the user experience and add new features to significantly improve the user experience by offloading more and more users from coverage cell to small cells and freeing up resources for users on the coverage cell.

As one of the most promising potential solutions to efficiently manage hyper-dense small cell deployment, cloud radio access network (C-RAN) attracted a great deal of attention and the concept of integration of HetNet with C-RAN to form a heterogeneous C-RAN (H-CRAN) has been proposed recently [4]. The traditional C-RAN only consists of a large number of spatially separated small cells and a centralized data processing centre, where the small cells are simplified and activated whenever the demand arises and some or all other functionalities, e.g., digital baseband processing, are centralized [5]-[6]. On the contrary, macro cells are also deployed in H-CRAN, which creates a potential opportunity to decouple the control-plane and user-plane. The essence of the control-user plane splitting concept lies in the small cells handling the user traffic while other (macro) cells deal with the signaling traffic (similarly to, but not equally to software-defined networking (SDN) as the intelligence can remain in each individual network element). The whole network is divided into two layers: control layer and user layer. The control layer consists of newly deployed or legacy macro cells whose main task is to exchange control signaling with user equipments (UEs) to maintain connectivity in a large area, thus ensuring that the UEs are always connected to the network. The user layer consists of small cells overlaid in the same area to boost the traffic. This splitting provides a promising

direction that maximizes the flexibility in the potential allocations of UEs and reduces the signaling traffic between the small cell and CRAN, i.e., over the fronthaul and accordingly it can improve the offered quality of experience (QoE), the resource-usage efficiency and the energy efficiency.

B. Fronthaul Requirements

With the continued evolution in network densification into hyper-dense network, the small cells are required to bring the content and information for the users from the Internet through the fronthaul, thus the capacity requirement of the fronthaul should be in the same order of the network in order to avoid the fronthaul rate bottleneck, which is obvious in dense or hyper-dense H-CRAN where low rate wireless fronthaul solutions are preferred for small cells instead of Fiber to the home (FTTH) technology due to deployment and operations expenses. Each added small cell has a profound effect on overall network performance, both with the increased capacity it brings as well as the large amount of interferences it generates. The effect is even significant under hyper-dense H-CRAN scenario when the fronthaul requirements under such hyper-dense network configuration have become stringent to support an efficient usage of spectrum and integration of potential user-centric services and applications.

C. Coordinated Multi-Point

One of the most important ways to improve the spectral efficiency of the system and mitigate the adverse effect of the huge interference is the use of advanced Coordinated Multi-Point (CoMP) technology among the small cells in order to combat the generated interference using interference management or interference alignment-based transmission solutions [7]-[8]. Moreover, once the UEs are connected to small cells because of high traffic demand, the mobility management is of paramount importance because the small cells merely cover a very small area so that frequent handover is expected even with very low moving velocity. CoMP is envisaged to be able to tackle this problem because the moving UE is connected to multiple adjacent small cells. Thus, CoMP together with H-CRAN is considered as a potential hybrid solution to the excessive interferences and limited fronthaul under dense and hyper dense small cell networks deployments.

CoMP imposes unprecedented challenges on the fronthaul network in H-CRAN. As prescribed earlier, the data rate of future communication networks will be increased by a factor 1000, compared to the current LTE system. The fronthaul requirements will follow the same patterns. For instance, it is anticipated that coordinated multipoint enables gain factor of 3 by exchanging user data in the clusters between 3 coordinating cells at distant sites, on average. Accordingly, the factor becomes 3000 in 5G fronthaul, e.g., several 100 Gb/s per site will be needed to support the emerging features in the future networks. Therefore, in spite of increasing the wireless network capacity by employing advanced PHY techniques, high data rates might still not be achievable unless high performance efficient fronthaul are designed for the excessive data [1].

While most of the previous research on uplink CoMP assumes lossless fronthaul links with infinite capacity and the obtained results are only of theoretical interest, some

researchers have brought attention to realistic fronthaul links with capacity limit [9]-[10]. However, in these works the coordinating small cells are assumed to simply quantize their own received signals before forwarding them to the processing centre via the fronthaul links. Considering the fact that the observations of all coordinating cells are actually correlated because they are broadcasted from the same source, distributed Wyner-Ziv compression can be applied to exploit this correlation and reduce the required fronthaul link rate. Distributed Wyner-Ziv compression is proposed in [11] and further extended to multiple sources in [12]. In distributed compression, each coordinating small cell compresses its own received signal and at the processing centre the correlation between the receptions of all coordinating small cells will be exploited to decompress and reconstruct their observations which will then be used to decode the user message.

D. Our Contributions

In this paper, we are motivated to consider the H-CRAN with focus on practical small cell uplink CoMP framework incorporating capacity-limited multiple fronthaul links. The main contributions are listed as follows:

- 1) Under the proposed framework, distributed Wyner-Ziv compression is employed to exploit the joint processing capability of the processing centre and minimize the required fronthaul rate for a given target user rate.
- 2) Some typical uplink CoMP application scenarios with practical deployment and noise and interference assumptions are considered to yield insight into not only the value and efficiency of proposed schemes but also the applicability of small cell uplink CoMP.

The rest of the paper is organized as follows: In the next section, the distributed compression and joint processing are introduced and analyzed. The efficiency evaluation is carried out in section III. The final section draws conclusions.

Notation: The following notations and definitions are employed in this work: capital letters e.g., X , stand for random variables and lower case letters e.g., x , represent realization of these variables. Vectors are represented by bold letters, e.g., \mathbf{X} . \mathbf{X}^H stands for the conjugate transpose of \mathbf{X} . Equalities of vectors are element-wise, i.e., $\mathbf{X} > \mathbf{Y}$ means $X_i > Y_i$ for $\forall i$. A superscript, e.g., \mathbf{X}^n represents the vector (X_1, \dots, X_n) . Euclidean letters denote sets, e.g., $T \equiv \{1, \dots, T\}$, whose cardinality is $|T|$. A subset Σ which fulfils $\Sigma \subseteq T$, has a complementary set Σ^C where $\Sigma \cup \Sigma^C = T$ and $\Sigma \cap \Sigma^C = \emptyset$. A subset with element t removed is denoted as $\Sigma \setminus t$ and x_Σ refers to $\{x_1, \dots, x_i, \dots, x_{|\Sigma|}\}$. Mutual information and entropy are denoted as $I(\cdot)$ and $H(\cdot)$, respectively. Unless otherwise stated, all the logarithms are in base 2. $\mathbb{E}\{\cdot\}$ denotes expectation.

II. FRONTHAUL COMPRESSION

Distributed Wyner-Ziv compression is proposed to reduce the required fronthaul traffic in uplink CoMP, tailored for future cellular network architectures, such as H-CRAN, that make use of numerous wireless fronthaul links with capacity limit and thus demand effective usage of the fronthaul network

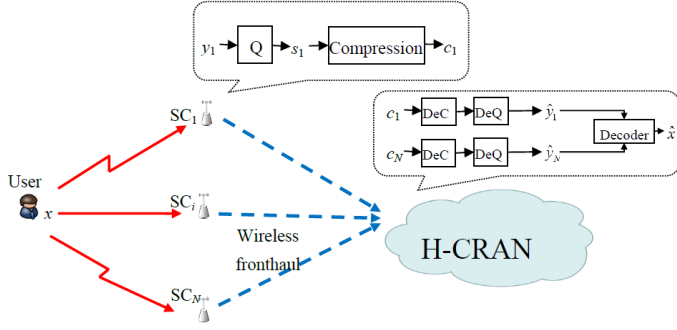


Fig. 1. Graphical illustration of compression/decompression in H-CRAN where small cells are connected with the processing centre via capacity limited wireless fronthaul links.

resources. The fronthaul rate is essential to the coordinating gain obtained. Actually, with a low fronthaul rate, introducing more coordinating cells only achieves marginal improvement. The distributed compression can be more effective than conventional schemes in terms of the obtained fronthaul traffic reduction.

A. Successive Decompression and Decoding

As shown in Fig. 1, a user encodes source message w and simultaneously transmits to N coordinating small cells which are then connected to the processing centre of H-CRAN via wireless fronthaul links. Upon reception of the signal, the small cells need to forward it to the processing centre for joint decoding. Since the received signal is continuous, it should be quantized into binary index s first and then the obtained binary index will be sent to the destination via the wireless fronthaul links. Quantization is the process of mapping a large set of input values to a (countable) smaller set. Here, the large set is actually an infinite set for continuous value. At the processing centre, the binary data s will first be used to reconstruct the received signal at each small cell. There is some difference between the original signals received by the small cells and the reconstructed ones. This difference is called quantization distortion and can be modeled by additional noise. Then, the reconstructed signals will be used for final joint decoding of the user message.

On the contrary, in distributed compression, the i th small cell compresses the received signal into the Wyner-Ziv bin index c_i using Wyner-Ziv lossy distributed source coding and forwards c_i to the processing centre. This compression can be done in two steps: 1) received continuous signal is quantized into binary index s ; 2) the binary index is then compressed into compression index c . We assume that the fronthaul transmission operates at a different frequency band from the users, the coordinating small cells can be regarded as full-duplex nodes because they can forward the compressed index of the $(n-1)$ th received frame and at the same time receive the n th frame from the user as long as the duration of compression bin index transmission is shorter than the user frame, which is normally true because of the relatively high capacity of fronthaul links. All small cells forward the Wyner-Ziv bin index to the processing centre simultaneously via the wireless fronthaul links. We can assume that the multiple fronthaul links

do not interfere with each other. This is because in some of the 5G proposals, the backhaul/fronthaul are designed to operate at very high carrier frequency [13], where the wave length is in millimeters thus the transmission is highly directional. The power leakage from one backhaul/fronthaul link to another is small enough to be negligible. At the processing centre, the decoding of the user message is conducted in a successive manner: firstly, the compression indices are decompressed to reconstruct the observations of the coordinating small cells; secondly, the reconstructed signals are coherently combined to decode the messages from the user. Similar to the quantization only case, the reconstructed signals are not exactly the same as the actual received signals and the difference is compression noise. Considering an extreme case where the fronthaul capacity approaches infinite so that the quantization, i.e., mapping from large or infinite set to another finite set can be extremely accurate, the compression noise approaches zero and the coordinating small cells can actually be equivalent to a MIMO system.

The achievable rate for such a system is a straightforward extension in [14] as

$$R \leq I(X; \hat{Y}) \quad (1)$$

subject to $\forall S \subseteq J : I(\hat{Y}_S; Y_S | \hat{Y}_{S^c}) \leq \sum_{n \in S} C_n$

where $\mathcal{S} = \{1, \dots, N\}$, \hat{Y}_n is the estimation of Y_n and C_n is the capacity of fronthaul link n . With circularly symmetric complex Gaussian (CSCG) distributed codebooks, the achievable rate is derived as

$$R = \log \left(1 + \sum_{n \in J} \frac{\gamma_n P_X}{\sigma_c + \sigma_{w,n}} \right), \quad (2)$$

subject to $f_S(\sigma_{w,J}) \leq \sum_{n \in S} C_n$

for all $S \subseteq \mathcal{S}$, where

$$f_S(\sigma_{w,J}) = \log \left(\prod_{n \in J} (\sigma_c + \sigma_{w,n}) + \sum_{n \in J} \left(\gamma_n P_X \prod_{i \in J, i \neq n} (\sigma_c + \sigma_{w,i}) \right) \right) - \log \left(\prod_{n \in S} \sigma_{w,n} \right) - \log \left(\prod_{n \in S^c} (\sigma_c + \sigma_{w,n}) + \sum_{n \in S^c} \left(\gamma_n P_X \prod_{i \in S^c, i \neq n} (\sigma_c + \sigma_{w,i}) \right) \right)$$

$E\{X\} \leq P_X$ for $1 \leq n \leq N$, σ_c and $\sigma_{w,n}$ are the variances of Gaussian noise and compression noise introduced by distributed Wyner-Ziv compression at SC_n , respectively, and $\gamma_n = |h_j|^2$. Notice that the achievable rate depends on the variances of the compression noises. The optimization objective function can be formed as,

$$\begin{aligned} &\text{maximize} \quad R(\sigma_{w,J}) = \log \left(1 + \sum_{n \in J} \frac{\gamma_n P_X}{\sigma_c + \sigma_{w,n}} \right) \\ &\text{subject to} \quad f_{S^l}(\sigma_{w,J}) \leq C_{S^l}, \quad \sigma_{w,n} > 0, \quad n \in J \end{aligned} \quad (3)$$

for $\forall S^l$, where

$$C_{S^l} = \sum_{i \in S^l} C_i, \quad \text{for } \forall S^l \subseteq J, \quad \text{and } |S^l| = l, \quad 1 \leq l \leq N.$$

The optimization can be done using Lagrangian dual optimization approach as shown in [15].

B. Joint Decompression and Decoding

The aforementioned successive decompression and joint decoding requires the received signals compressed at the coordinating small cells to be correctly decompressed first at the processing centre, followed by joint decoding of the user message. However, since the main objective of the processing centre is to decode the user message rather than successful decompression, this requirement is actually unnecessarily imposed. Even if the decompressed information has errors and the observation reconstruction is not correct, it is still possible for the processing centre to correctly decode the user message, i.e., the reception of the compressed information can tolerate errors. In this regard, the decompression and decoding procedure at the processing centre should be modified accordingly. It is pointed out that rather than a successive decoding process the decompression and joint decoding of the user message can be conducted in a joint manner and improves the fronthaul usage efficiency [16].

When only the errors of the user message are taken into consideration with joint decompression and decoding, the achievable rate is given in [17] as

$$R = \min_{S \subseteq \mathcal{I}} \left\{ \sum_{n \in S} [C_n - I(\hat{Y}_n; Y_n | X)] + I(\hat{Y}_{S^c}; X) \right\}. \quad (4)$$

With Gaussian distributed codebooks, (4) can be written as the following expression:

$$R = \min_{S \subseteq \mathcal{I}} \left\{ \sum_{n \in S} \left[C_n - \log \left(1 + \frac{\sigma_c}{\sigma_{w,n}} \right) \right] + \log \left(1 + \sum_{n \in S^c} \left(\frac{\gamma_n P_X}{\sigma_c + \sigma_{w,n}} \right) \right) \right\}. \quad (5)$$

The achievable rate still depends on the variances of the compression noises but the difference is that the constraints are no longer imposed, i.e., (5) should be optimized as

$$R = \max_{\sigma_{w,j}} \left\{ \min_{S \subseteq \mathcal{I}} \left\{ \sum_{n \in S} \left[C_n - \log \left(1 + \frac{\sigma_c}{\sigma_{w,n}} \right) \right] + \log \left(1 + \sum_{n \in S^c} \left(\frac{\gamma_n P_X}{\sigma_c + \sigma_{w,n}} \right) \right) \right\} \right\}. \quad (6)$$

Notice that (6) is not in standard concave form with respect to $\sigma_{w,j}$. However, by using an intermediate variable z_n defined as

$$z_n = \log \left(1 + \frac{\sigma_c}{\sigma_{w,n}} \right) = \log \left(1 + \frac{1}{\sigma_{w,n}} \right), \quad (7)$$

(6) can be expressed as

$$R = \max_{\sigma_{w,j}} \left\{ \min_{S \subseteq \mathcal{I}} \left\{ \sum_{n \in S} [C_n - z_n] + \log \left(1 + \sum_{n \in S^c} \left(\frac{\gamma_n P_X (1 - 2^{-z_n})}{\sigma_c} \right) \right) \right\} \right\}. \quad (8)$$

The equation inside the maximization function in (8) is a concave function and similar to the previous case, can be solved by using the subgradient method. The direction of the subgradient search is

$$g_n(z_n) = \begin{cases} -1, & \text{if } n \in S_{\min} \\ \frac{\ln(2) \gamma_n P_X 2^{-z_n}}{\sigma_c \left(1 + \sum_{j \in S^c} \left(\frac{\gamma_j P_X (1 - 2^{-z_j})}{\sigma_c} \right) \right)}, & \text{if } n \notin S_{\min} \end{cases} \quad (9)$$

where S_{\min} is the solution subset to (4) and can be obtained by exhaustive searching, i.e., comparing all subsets of \mathcal{I} at z_0 . The searching criterion is: if $g_n(z_n) \leq 0$, decrease z_n ; otherwise increase z_n . When $\sigma_{w,j}$ is in the region $(0, +\infty)$, according to (7), z_n is also in the range of $(0, +\infty)$. When the optimal $\sigma_{w,j}^*$ is found, it can be used as the initial value for the optimization using iterative algorithms proposed in [15].

III. EFFICIENCY EVALUATION OF FRONTHAUL COMPRESSION

We present some numerical results to illustrate the gain brought by the distributed compression schemes. Realistic assumptions are considered to provide insight into the practical performance of the fronthaul. As shown in Fig. 1, the small cells are connected with the processing centre via capacity limited wireless fronthaul links. The cell edge users are associated with multiple small cells and a coordinating area is formed. All fronthaul links are assumed to have the same capacity C for simplicity. The carrier frequency f_c is 2 GHz and the frequency reuse factor is 1, i.e., all small cells use the same frequency but the users within one coordinating area occupy different frequency bands based on certain scheduling mechanisms. Therefore, there is no intra-cell interference from the users within the same coordinating area but other users in adjacent cells outside the area generate inter-cell interference. We consider the worst case that there is one user occupying exactly the same frequency band in each interfering small cell, transmitting with maximal power 23 dBm. For simplicity, we assume that the interfering users are always located at the centre of the cells.

Two different scenarios are studied in this paper: 1) dense urban small cells deployment where the radius of small cells is assumed to be 200 meters and both the target user and the interfering users are assumed to be outdoor, and 2) ultra-dense urban small cells deployment where the radius is only 20 meters and the target user is assumed to be indoor but the interfering users are assumed to be outdoor. The first scenario is a more general case and the second scenario is more applicable in some specific situations, where ultra-dense small cells are deployed to cover hotspot events like exhibitions and games. Path loss and Rayleigh fast fading are taken into account for both scenarios. For the first scenario, the path loss model is given as $PL = 22.7 + 36.7 \log_{10} d + 26.0 \log_{10} f_c$ [18], where d is the distance in meter and f_c is in GHz. For the second scenario, the path loss model within the coordinating area is given as $PL = 37 + 30.0 \log_{10} d$, and the path loss model outside the coordinating area is $PL = 15.3 + 37.6 \log_{10} d + A_d$, where $A_d = 20$ dB is the penetrating loss. We only consider the interference from the users in the first tier interfering cells. Both the users and the small cells are equipped with omni-directional antennas with 0dB antenna gain. We assume that the noise power density $N_o = -171$ dBm/Hz and the bandwidth $B = 1$ MHz. The overall noise level should be the summation of thermal noise and interference.

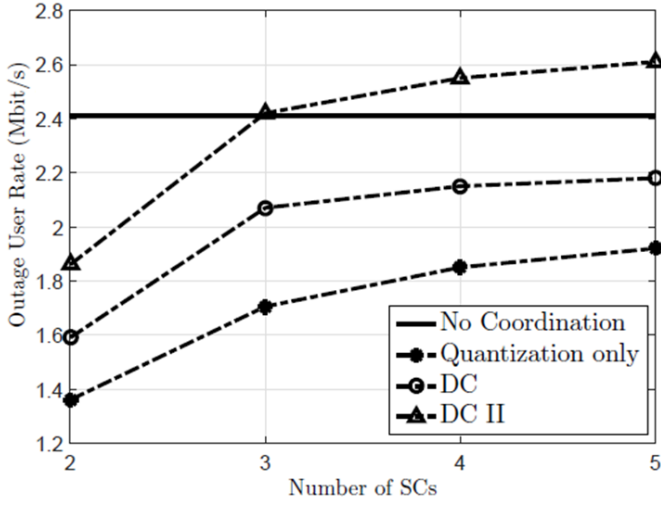


Fig. 2. Outage user rate vs number of coordinated small cells in dense deployment.

A. Outage Performance Comparison

For comparison purpose, we consider the following schemes:

- 1) Quantization only, where the small cells simply quantize the received signal and send the quantized information to the processing centre.
- 2) DC: all small cells conduct distributed compression and the decompression and decoding of message are carried out in a successive manner at the processing centre.
- 3) DC II: all small cells conduct distributed compression but the decompression and decoding of message are carried out in a joint manner at the processing centre.
- 4) Upper Limit: the capacity of the fronthaul links is assumed to be infinity.
- 5) No coordination case, where the user is still assumed to be located at the cell edge.

Fig. 2 and Fig. 3 plot the outage user rate with a low fronthaul rate $C=1$ Mbits/s for different number of small cells and the outage probability $P_{out}=10^{-2}$. It is clearly shown that distributed compression outperforms the quantization only scheme and the performance can be further enhanced by applying joint decompression and decode. Two important observations are obtained. Firstly, the gain of adding a new small cell diminishes as the size of the coordinating network grows. It is shown in Fig. 2 that the joint operation achieves almost 0.6 Mbit/s gain from 2 to 3 small cells but only less than 0.2 Mbit/s gain from 3 to 4 small cells. Secondly, if the fronthaul link capacity is low it is evident that coordination does not always outperform non-coordination. Actually, the joint operation can only outperform non-coordination when coordinating small cells are increased to 3 and 4 for dense and ultra-dense deployment, respectively. This is because with low fronthaul rate the variances of the noise introduced by quantization or compression are so large that the potential coordinating gain cannot be harvested.

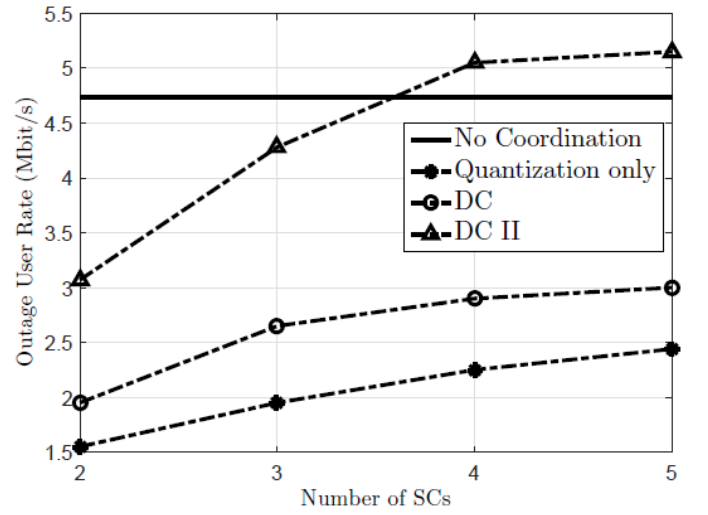


Fig. 3. Outage user rate vs number of coordinated small cells in ultra-dense deployment.

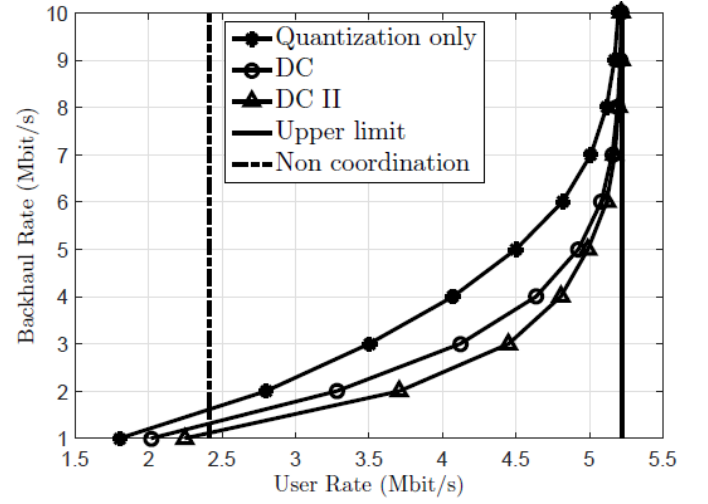


Fig. 4. Required C for dense small cell deployment.

B. Required Fronthaul Rate

Fig. 4 illustrates the required fronthaul link rate C for a given outage user rate at $P_{out}=10^{-2}$ in the dense deployment scenario. The two vertical lines are non-coordination and upper limit when C approaches infinity, respectively. The non-coordination lines intersect with the required fronthaul rate curves and only beyond these intersection points does coordination outperform non-coordination. It is clearly shown that distributed compression effectively reduces the required fronthaul rate C to achieve a target user rate. For 4 small cells, at user rate = 3.5 Mbit/s the required C reduces 0.6 Mbit/s from quantization to distributed compression and a further 0.2 Mbit/s to joint decompression & decoding. With a higher target user rate, the distributed compression and joint decompression & decoding schemes are more efficient than quantization. However, when the target user rate is further increased to approach the upper limit, the gain diminishes eventually and the required fronthaul rate C is approaching infinity.

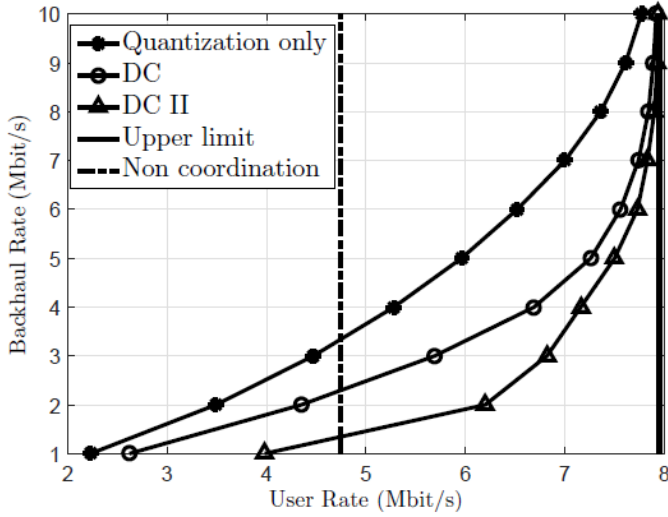


Fig. 5. Required C for ultra-dense small cell deployment.

Fig. 5 shows the ultra-dense deployment indoor scenario. Because of the relatively weak interference due to penetration loss, the user rate is more quantization/compression noise dominant and the gaps between three curves are more remarkable due to enhanced efficiency of the proposed optimization algorithms. The same trend as the outdoor dense deployment scenario is observed: the gains of distributed compression and joint decompression & decoding increase with the target user rate firstly but eventually diminishes when approaching the upper limit. As shown in Fig. 4 and 5, up to 45% and 60% fronthaul traffic reduction are observed for dense deployment and ultra-dense small cells deployment, respectively.

IV. CONCLUSIONS AND FUTURE OUTLOOK

In this paper, we investigate the problem of using distributed Wyner-Ziv compression to reduce the demanded fronthaul traffic to forward the received signals in uplink CoMP, tailored for future cellular network architectures such as H-CRAN with control-user plane split, where the small cells carry user data and make use of numerous wireless fronthaul links with capacity limit and thus demand effective usage of the fronthaul network resources. Numerical results are shown for two typical scenarios: outdoor dense deployment and indoor ultra-dense deployment. Our results reveal that, in both scenarios, the fronthaul rate is essential to the coordinating gain obtained. With a low fronthaul link rate, introducing more coordinating cells merely achieves marginal improvement. The gain of coordination can only be harvested in the existence of medium to high fronthaul capacity.

More importantly, we compare the distributed compression with the conventional quantization only scheme and find that for a given target user rate, the required fronthaul link rate can be effectively reduced by distributed compression, and with joint operation of decompression and decoding further improvement can be obtained. The achieved fronthaul traffic reduction becomes more significant with the increased target user rate firstly but eventually diminishes when approaching upper limit where the required fronthaul rate approaches

infinity. Compared with the outdoor coordination case, distributed compression yields greater reduction in the indoor case where the compression noise rather than the co-channel interference dominates the performance.

REFERENCES

- [1] CISCO, Cisco Visual Networking Index: Global Mobile Data Traffic Forecast, 2013–2018, Feb. 2014.
- [2] M. Z. Shakir, M. Imran, X. Wang, J. Wu, A. Ghosh, H. Lundqvist, and L. Liu, "Smart fronthauling and backhauling for 5G systems: from precoding to network architecture," *IEEE Wireless Commun. Mag.*, vol. 22, no. 5, pp. 10–12, Oct. 2015.
- [3] NSN white paper, 2020: Beyond 4G Radio Evolution for the Gigabit Experience, 2011.
- [4] M. Peng, Y. Li, J. Jiang, J. Li, and C. Wang, "Heterogeneous Cloud Radio Access Networks: A New Perspective for Enhancing Spectral and Energy Efficiencies," *IEEE Wireless Commun. Mag.*, vol. 21, no. 6, pp. 1536–1284, Dec. 2014.
- [5] China Mobile Research institute, C-RAN White paper V2.5, Oct 2011.
- [6] China Mobile Research Institute, "C-RAN - Road Towards Green Radio Access Network," in *Proc. of Wireless World Research Forum, Meeting 27 (WWRF 27)*, pp. 73–80, Duesseldorf, 2011.
- [7] Z. Ghebretensae, K. Laraoui, S. Dahlfort, and J. Chen, "Transmission solutions and architectures for heterogeneous network built as C-RANs," in *Proc. 7th Intl. ICST Conf. Commun. and Networking in China*, pp. 748–752, KunMing, Aug. 2012.
- [8] K. Sakaguchi, S. Sampei, H. Shimodaira, and R. Rezagah, "Cloud cooperated heterogeneous cellular networks," in *Proc. Intl. Symposium Intelligent Signal Processing and Communication Systems (ISPACS 2013)*, pp. 787–791, Naha, Nov. 2013.
- [9] P. Marsch, Coordinated multi-point under a constrained backhaul and imperfect channel knowledge, Vogt Verlag, Jun. 2010, Ph.D. thesis, ISBN 978-3-938860-35-9.
- [10] S. Wang, L. Chen, Y. Yang, and G. Wei, "Quantization in uplink multi-cell processing with fixed-order successive interference cancellation scheme under backhaul constraint," in *Proc. IEEE 79th Vehicular Technology Conf. (VTC 2014 Spring)*, Seoul, Korea, May 2014.
- [11] A. D. Wyner and J. Ziv, "The rate-distortion function for source coding with side information at the decoder," *IEEE Trans. Information Theory*, vol. 22, no. 1, pp. 1–10, Jan. 1976.
- [12] M. Gastpar, "The Wyner-Ziv problem with multiple sources," *IEEE Trans. Information Theory*, vol. 50, no. 11, pp. 2762–2768, Nov. 2004.
- [13] F. Boccardi, J. Heath, R.W., A. Lozano, T. Marzetta, and P. Popovski, "Five disruptive technology directions for 5G," *IEEE Commun. Mag.*, vol. 52, no. 2, pp. 74–80, Feb. 2014.
- [14] G. Kramer, M. Gastpar, and P. Gupta, "Cooperative strategies and capacity theorems for relay networks," *IEEE Trans. Information Theory*, vol. 51, no. 9, pp. 3037–3063, Sep. 2005.
- [15] Y. Qi, M. A. Imran, A. Qudus, and R. Tafazolli, "Achievable rate optimization for coordinated multi-point transmission (CoMP) in cloud-based RAN architecture," in *Proc. IEEE Intl. Conf. Commun., (ICC 2014)*, Sydney, Jun. 2014.
- [16] X. Wu and L. Xie, "On the optimal compressions in the compress-and-forward relay schemes," *IEEE Trans. on Information Theory*, vol. 59, no. 5, pp. 2613–2628, May 2013.
- [17] A. Sanderovich, S. Shamai, Y. Steinberg, and G. Kramer, "Communication via decentralized processing," *IEEE Trans. Information Theory*, vol. 54, no. 7, pp. 3008–3023, Jul. 2008.
- [18] H. Holtkamp (et al.), D2.2: Definition and parameterization of reference systems and scenarios, Tech. Rep. of FP7 project EARTH, Jun. 2010, available online at <https://www.ict-earth.eu>.