Reviewers Comments & Authors Replies

Manuscript No. Paper T-SP-18051-2014, submitted to *"IEEE Transactions on Signal Processing"*

Title "Traffic Aware Resource Allocation Schemes for Multi-Cell MIMO-OFDM Systems"

Authors Ganesh Venkatraman, Antti Tölli, Markku Juntti, and Le-Nam Tran

The authors would like to thank the associate editor and the reviewers for their valuable comments on the manuscript of the paper, which have been greatly helpful to improve the paper quality. Based on the comments, we have made several major revisions to the paper~~, following the suggestions of the reviewers~~. In what follows, the comments are listed, each followed immediately by the corresponding reply from the authors. The reviewers questions in the revised manuscript are highlighted in blue color and the authors responses are presented in black. ~~Following is the summary of the revision made on the manuscript in accordance with the reviewers comments~~.

1. We have shortened the discussions on the existing *backpressure algorithm* in Section III-A.
2. Initialization of the successive convex approximation (SCA) operating point is included in the revised manuscript
3. Section III-D is updated to include the guidelines involved in selecting the sub-channel order.
4. Strong convexity of the objective function is emphasized in Appendix A-B.
5. Strict monotonicity of the algorithm is discussed in a detailed manner in Appendix A-C.
6. We have provided additional details on the distributed algorithm convergence as suggested by the reviewers

In what follows, the comments are listed, each followed by the corresponding reply from the authors. Unless otherwise stated, all the numbered items (figures, equations, references, citations, etc) in this response letter refer to the revised manuscript.

Response to Reviewer - 1's Comments

Comments: The response to the reviewer's concerns are generally satisfying, except the convergence proof.

Reply: We thank the reviewer for providing valuable and insightful comments.

For the resubmitted manuscript, the reviewer still has the following concerns

1. Considering the length of the manuscript, it would be better to shorten some parts that are not new in this manuscript, e.g. III.A. More space can be left for convergence proof, which is very important.

Reply: After a careful check of the manuscript, we have removed couple of paragraphs from Section ~~III.A~~ and shortened the discussions on ~~the simulation section~~ to provide additional details for the convergence proof. We believe the removed parts don't affect the readability of the paper and the added text regarding convergence proof certainly improves the quality.

2. In convergence proof (48), why does the 2nd inequality hold? In fact, to prove the feasibility of $m_{k+1}^{(i)}, w_*^{(i-1)}; m_k^{(i)}$, the part between 2nd and 3rd inequality is not necessary, ≤ 0 directly follows the 2nd inequality since the solution $m_{k+1}^{(i)}, \gamma_{k+1}^{(i)}$ is the optimal solution, and therefore feasible.

Reply: We thank the reviewer for the critical comment. It is not possible to define the inequality with the previous optimal point. Since it is not possible to comment on the inclusion of the previous constraint set in the current update (except the earlier optimal point), the inequality is not guaranteed. Based on the comment from the reviewer, we have removed the inequality specifying the previous operating point from (49) in the revised manuscript, since it cannot be guaranteed.

3. The solutions SCA iterations $\mathbf{m}_k^{(i)}$ does not necessarily converge. In fact $\mathbf{m}$ has compact feasible region, and thus $\mathbf{m}_k^{(i)}$ has limit points for any specific i. However $m_*^{(i)}$ does not necessarily exist(the whole sequence $\mathbf{m}_k^{(i)}$ may be not convergent). Similar problem happens to $\mathbf{w}_k^{(i)}$.

Reply: We understand the reviewer concern. Even though the objective of the SCA iteration converges, it is not guaranteed that the iterates involved in the iterative algorithm, namely, $\mathbf{m}_k^{(i)}$ and $\mathbf{w}_k^{(i)}$ to converge. It is true that the iterates need not converge when the objective function is convex. We have modified the discussion by using the strong convexity of the objective function to impose the uniqueness of the iterates in each SCA update as well. By regularizing the objective function with a strongly convex term like $\|\mathbf{m} - \mathbf{m}_k^{(i)}\|^2$, we can guarantee the uniqueness of the iterates upon the SCA convergence, *i.e.*, the limit point of the SCA update is unique. We thank the reviewer for citing this issue on the sequence convergence and the uniqueness of the minimizer. We have updated the manuscript to include this information in the convergence proof in Appendix A-B around (46) and in Appendix A-C.

4. Strict monotonicity with respect to the objective function f should be rigorously proved. Note that to guarantee the uniqueness of the beamformer iterates, (52) instead of the objective function is used.

Reply: We thank the reviewer for the pointing out the issue involving the strict monotonicity. As suggested by the reviewer, we have included the strict monotonicity of the objective sequence in Appendix A-C following (52).

Note that the objective sequence $\{f(\mathbf{x}_k)\}$ generated by an iterative SCA algorithm is monotonic, *i.e*, $f(\mathbf{x}_k) \leq f(\mathbf{x}_{k-1})$ and holds with equality at the limit point of the iterative procedure. However, upon convergence, due to the convexity of the objective, we may have multiple limit points, say, $\mathbf{y}_*$ with the same objective as $f(\mathbf{y}_*) = f(\mathbf{x}_*)$. Therefore, strict monotonicity cannot be guaranteed for the objective sequence. However, when the objective function is strongly convex, as discussed in Appendix A-B, the minimizer for the subproblem in each SCA iteration k is unique as

$$f(\mathbf{x}_k) - f(\mathbf{x}_{k+1}) \geq \nabla f(\mathbf{x}_{k+1})^{\mathrm{T}}(\mathbf{x}_k - \mathbf{x}_{k+1}) + c\|\mathbf{x}_k - \mathbf{x}_{k+1}\|^2 \tag{1}$$

where $\mathbf{x}_{k+1}$ is the optimal solution for the k^{th} SCA subproblem and $\mathbf{x}_k$ is the unique minimizer obtained in the previous $k-1^{\text{th}}$ SCA step, which is a feasible point for the current problem. Since $\mathbf{x}_{k+1}$ is the solution in the k^{th} SCA problem, following conditions hold

$$\nabla f(\mathbf{x}_{k+1})^{\mathrm{T}}(\mathbf{x}_k - \mathbf{x}_{k+1}) \geq 0 \tag{2a}$$

$$f(\mathbf{x}_k) - f(\mathbf{x}_{k+1}) \geq c\|\mathbf{x}_k - \mathbf{x}_{k+1}\|^2 \tag{2b}$$

where (2a) is due to the lack of any descent direction in the feasible set and (2b) follows from (1) using (2a). As $k \to \infty$, $\|\mathbf{x}_{k+1} - \mathbf{x}_k\| \to 0$, and therefore the iterative algorithm converges to a unique minimizer, say, $\mathbf{x}_*$, which is the limit point of the algorithm. Now by using the monotonicity of the SCA updates and the uniqueness of the minimizer in each step, we can guarantee the strict monotonicity of the objective sequence generated by the iterative problem.

5. Note that the conclusions [32, Thm 2] and [26, Thm 10] have lots of assumptions. To invoke these reference, explicit exposition should be provided to show that these conclusions can be applied to our problem. The same questions occur to the proof in Appendix B, where conclusions in [11] [36] and [37] are used. Too many details are omitted to make the proof convincing and clear.

Reply: We thank the reviewer for raising the concern. We have updated the manuscript to include the details regarding the stationary point discussion in Appendix A-E and the convergence proof analysis for the primal and the alternating directions method of multipliers (ADMM) algorithms in Appendix B. Additional detail includes the conditions required for showing the stationarity of the limit point and we have included additional details to show the distributed approaches convergence convincing.

Response to Reviewer - 2's Comments

The authors have addressed many of my previous comments. However, there are still several major issues that need further clarification.

Reply: We thank the reviewer for providing useful comments. The comments are constructive and helped us to improve the manuscript better.

1. The revised paper did not address my previous comment about how to select the sub-channel ordering. I understand that finding the best sub-channel ordering requires exhaustive search which has extremely high complexity. But it is important to provide a guidance on what would be a good choice of sub-channel ordering. For example, can we achieve a good performance by using a low complexity ordering algorithm such as a greedy sub-channel ordering algorithm?

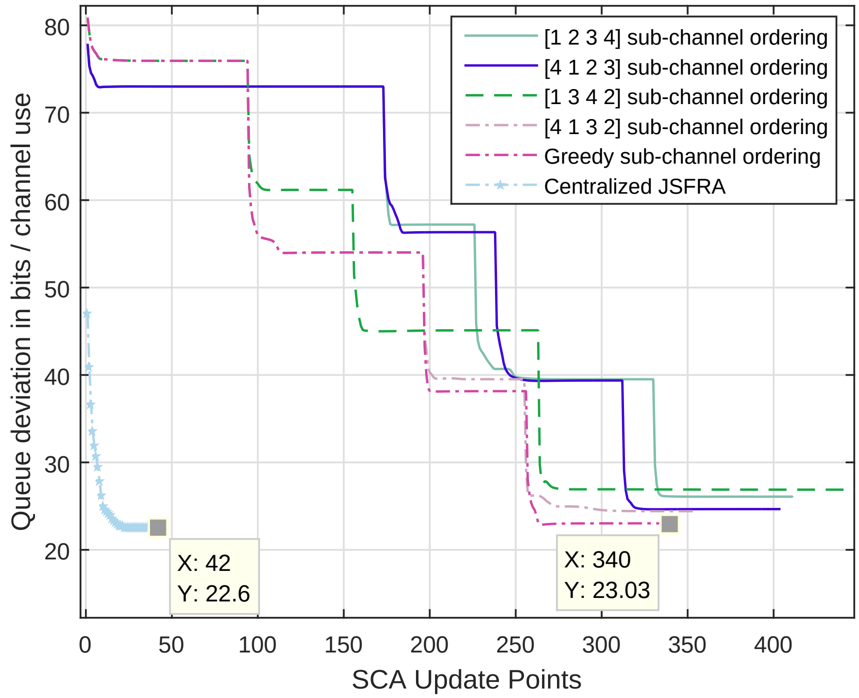
Reply: We apologize for not discussing the sub-channel ordering in a detailed manner. We have included the guidelines for selecting the sub-channel ordering based on channel gains, *i.e*, greedy selection. In random ordering scheme, after finding the precoders for a current sub-channel, we can choose any previously unselected sub-channels as the next candidate sub-channel for which the precoders are identified using the updated backlogged packets. As suggested by the reviewer, the greedy sub-channel ordering is based on sorting the best channel from each sub-channel, which is obtained by finding the highest channel norm between the users from the respective serving base station (BS). ~~Note that the users channel used in the ordering procedure is the channel seen by the users with the corresponding serving BSs only.~~

However, it may not be optimal for all system configurations, since the ordering scheme should also consider the number of backlogged packets associated with each user. We have emphasized this issue by comparing various ordering schemes for the same system model with two different set of backlogged packets associated with each user. Even though we are not including it in the paper, we have included in the response letter to answer the reviewer's question. We considered a system with $N = 4$ sub-channels, $N_B = 2$ BSs with $N_T = 4$ transmit antennas and $K = 12$ single antenna users. The path loss (PL) is distributed uniformly over $[0, -3]$ dB. The number of backlogged packets assumed for each user is provided in the corresponding captions in Fig. 1.

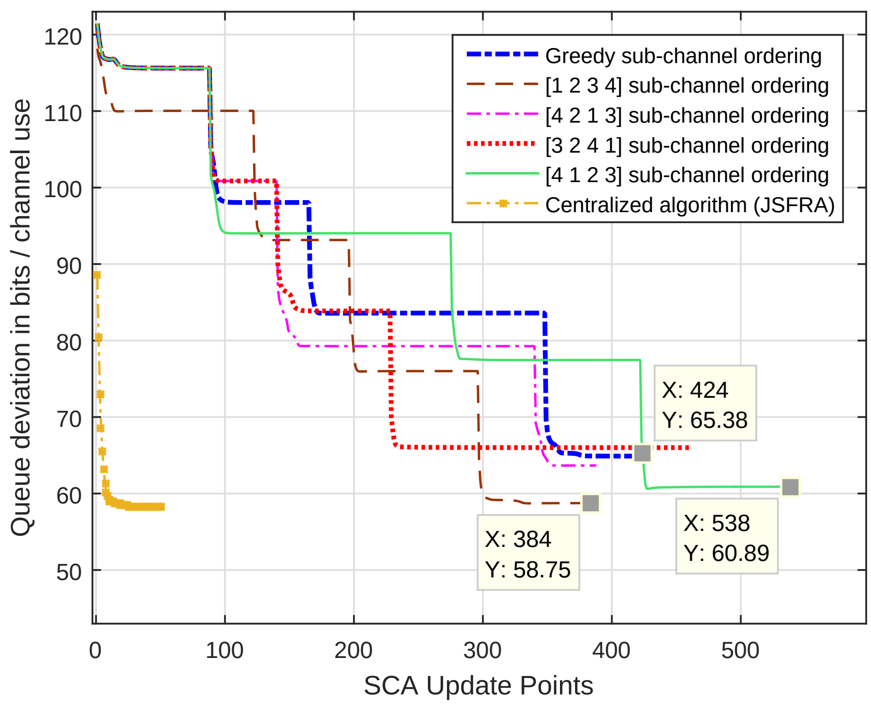
Fig. 1 compares different ordering of sub-channels with the total number of backlogged packets remaining in the system ~~as the metric~~. As can be seen from Fig. 1a, the greedy sub-channel ordering provides a favorable way of choosing the sub-channels to minimize the total number of backlogged packets. However in Fig. 1b, the greedy sub-channel ordering is not an efficient strategy in determining the order in which the sub-channels are to be chosen. We have included the discussions regarding the greedy sub-channel ordering in the revised manuscript as a heuristic method under Section III-D final paragraph. Note that as the number of users in the system increases, all channel ordering provides favorable order to minimize the total number of backlogged packets in the system.

2. The authors mentioned that the signaling overhead of the distributed algorithm can be reduced by using a smaller number of iterations $J_{\max}$. But still, you didn't answer my question about whether the signaling overhead of the distributed algorithm is smaller than the centralized algorithm. You should first analyze the signaling overhead of the distributed algorithm for fixed $J_{\max}$ and the signaling overhead of the centralized algorithm. Then you should point out under what $J_{\max}$ the distributed algorithm will have less signaling overhead than the centralized algorithm. Is it possible that the distributed algorithm always has more signaling overhead than the centralized algorithm even when $J_{\max} = 1$? Finally, there is a trade-off between performance and signaling overhead ($J_{\max}$) for the distributed algorithm. For the same signaling overhead (we can control $J_{\max}$ to make the signaling overhead of the distributed algorithm approximately equal to that of the centralized algorithm), does the distributed algorithm achieve better performance than the centralized algorithm?

Reply: We thank the reviewer for the insightful comment and we apologize for the lack of clarity in explaining this information in our earlier manuscript.



(a) Number of backlogged packets for each user in bits $Q_k = [11, 8, 14, 6, 6, 2, 10, 10, 5, 6, 9, 5]$



(b) Number of backlogged packets for each user in bits $Q_k = [8, 9, 12, 8, 12, 5, 4, 10, 8, 5, 7, 9]$

Figure 1: Convergence of the algorithms for $\{N, N_B, K, N_T, N_R\} = \{4, 2, 12, 4, 1\}$ using ℓ_1 norm

(a) The amount of signaling overhead of the distributed algorithm and the centralized one depends on the system model of consideration. For example, let us consider a model with $N = 1$ sub-channels ~~,$N_B = 2$ BSs, and~~ $K = 4$ users ~~in total~~, each BS serving 2 users. Let $N_T = 4$ be the number of transmit antennas and $N_R = 1$ be the number of receive antenna at each user. In this scenario, the amount of information exchange to perform a centralized algorithm by a common controller requires the knowledge of complete channel matrices interlinked in the system, *i.e*, the number of users times the BSs.

   In order to quantify the total number of bits required to be exchanged, let us assume that each complex channel for a single-input single-output requires 10 bits, *i.e*, 4 bits for amplitude and 6 bits for phase (assuming phase is important) or it can be a equal share of 5 bits for both amplitude and phase. Using this assumption, the total number of channel information in bits to be exchanged via backhaul requires $10 \times K \times N_B \times N_R \times N_T = 320$ bits. On the other hand, for the distributed case, let us consider 6 bits are required to quantize the scalar interference in the consensus vectors. Consequently, the proposed distributed solutions require $6 \times 2 \times 2$ bits to be exchanged in each iteration.

   In this example, for the same amount of signaling overhead as in centralized method, we can performance only up to 6 SCA updates for $J_{\max} = 2$ (*i.e.*, two updates for ADMM part). This may not be sufficient for the distributed algorithms to attain the same performance as the centralized method. However, as the number of sub-channels, users and/or the antenna element increases, it may not be a feasible option to feedback the channel state information (CSI) across the coordinating BSs to the centralized controller. In addition to comparing the signaling overhead, we also need to consider the effects of the quantization of the CSI on the performance of a centralized algorithm, which is beyond the scope of our paper. Generally, the performance is significantly degraded if the CSI is quantized [1]. Moreover, in the centralized algorithm, resulting transmit precoders need to be exchanged with the corresponding BSs before the actual transmission, involving huge overhead in the backhaul.

(b) Using the above discussion, we can say that when the system size is huge, it would be favorable to consider the distributed algorithm over the centralized approach due to the huge signaling overhead involved in exchanging the channels.

(c) In particular, ADMM and primal approach requires significant overhead compared to the centralized algorithm for a small system, since by exchanging the quantized channels, each BS can perform the centralized algorithm independently until convergence. However, it depends on the channel quantizer, which is likely to be based on the channel density function (eg. Lloyd quantizer). For a system involving more coordinating BSs, users and antenna elements, it would be beneficial to use distributed algorithm with $J_{\max} > 1$ to have a strictly monotonic decrease in the objective. If we set $J_{\max} = 1$, the proposed algorithms still converge (but monotonicity of SCA is not guaranteed). It can be easily verified from Fig. 2, which has same configuration as that of Fig. 1a. Note that Fig. 2 plots the queue deviation at the SCA update points only. The number of backlogged packets in each ADMM iteration is not shown for clarity reason.

We have included the above discussions in the first and last paragraphs of Section IV-C, and the paragraph before Section VI (i.e., the conclusion section). In reality, since the channel is time-correlated, it is enough to update the precoders once per radio frame. Thus, it is not necessary for the distributed algorithm to converge until the end. Instead, the decentralized parts only need to follow the fading process when $J_{\max} > 1$.

The performance of the distributed algorithm based on dual decomposition scheme was discussed for the time-correlated fading in Section C of [13], which shows that it is enough for the distributed precoder design to follow the fading process to provide desired performance. The distributed algorithm for the time correlated case is beyond the scope of our paper and thus is not considered in the current manuscript. However, we take this opportunity to show a plot demonstrating this behavior for the Karush-Kuhn-Tucker (KKT) based algorithm presented in Section IV-C of the manuscript.

Fig. 3 compares the performance of the distributed KKT approach presented in Section IV-C for different iteration. The signaling requirements are outlined in Algorithm 3 and the overhead involved

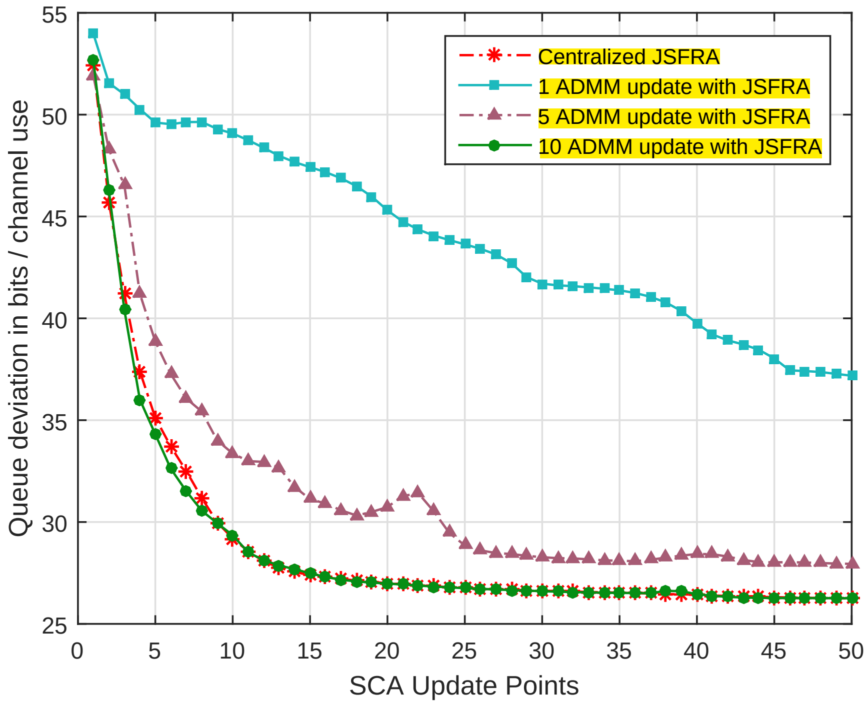


Figure 2: Convergence of the algorithms for $\{N, N_B, K, N_T, N_R\} = \{4, 2, 12, 4, 1\}$ using ℓ_1 norm

in the signaling is penalized in the achievable rate of the users. We considered that the channel is coherent over $N_S = 100$ symbols and the precoder update is performed by exchanging the equivalent channel for $J_{\max} = 3, 5, 10$ ~~number of iterations~~. The overhead is considered as $\tilde{t}_{l,k,n} = (1 - \frac{J_{\max}}{N_S}) \times t_{l,k,n}$, where $\tilde{t}_{l,k,n}$ is the rate seen by the user and the factor $(1 - \frac{J_{\max}}{N_S})$ is considered as a penalty involved due to the precoder exchange. The average number of backlogged packets after each transmission slot is evaluated as

$$\chi = \sum_{k=1}^{K} [Q_k - \tilde{t}_k]^+ \tag{3}$$

Unlike the distributed algorithm, the centralized scheme presented in Fig. 3 has no penalty term and it is used as a benchmark ~~for the other figures~~. In order to improve the performance of the distributed scheme, the operating point involved in the SCA algorithm is considered from the earlier frame instead of starting initializing randomly. Since we use KKT approach, we can either use all users in the system for the precoder design or we can utilize single-cell MU-MIMO user selection presented in the literature to limit the number of users for which the precoders are designed, which leads to the faster convergence. As we can see from Fig. 3, as the arrival rate per user increases, the performance of KKT schemes with $J_{\max} = 3, 5, 10$ converges since the number of backlogged packets are significantly large, therefore, the same set of users will be served by the algorithm with better precoders by utilizing the memory.

In spite of using memory and prior scheduling in the KKT approach, isolated single BS processing performs much better than the distributed scheme due to the limited number of iterations allowed in the algorithm. Note that the precoders are not updated for the desired users until convergence, after the limited number of iterations. However, if we perform the single cell precoder design by considering the neighboring precoders as fixed after the recent exchange as discussed in [28], we can improve the performance significantly for Algorithm 3 as shown by red curves in Fig. 3. In this approach, in between each exchange across the coordinating BSs, each BS will perform $J_{\max} = 20$ with the neighboring precoders as fixed. Once the iterations are performed to update the precoders, it is then exchanged across the coordinating BSs to perform the same procedure as mentioned earlier.

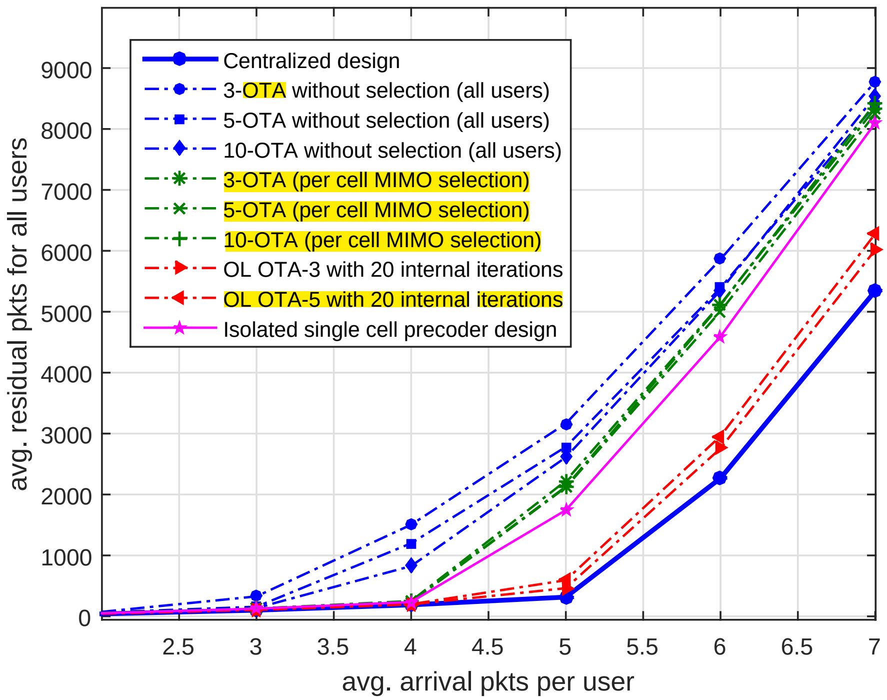


Figure 3: Average number backlogged packets after each transmission for system $\{N, N_B, K, N_T, N_R\} = \{4, 2, 16, 4, 2\}$ evaluated for 500 slots

3. If the authors can't prove the convergence of the ADMM algorithm (or the decomposition approach via KKT conditions) in Section IV.B, then at least, you should discuss the property of the fixed point of the algorithm. For example, does there exist a fixed point of the algorithm? If so, is the fixed point of the algorithm unique? Is any fixed point of the algorithm also the optimal solution of the original problem in (20)? Assuming that the ADMM algorithm converges to a fixed point, will the interference vector in (39) converges to the actual interference in the network? These questions must be clarified in the paper. Otherwise, it is not clear how the ADMM algorithm is related to the original problem in (20). Similar questions should also be answered for the decomposition approach via KKT conditions.

Reply: We thank the reviewer for raising the important concern regarding the convergence issue of the distributed algorithm with limited number of iterations. In view of this, we have updated the manuscript to include the discussion on the convergence of the distributed algorithms with a limited number of iterations in the third paragraph of Appendix B ~~after (57)~~. Since the distributed algorithm cannot be guaranteed to be monotonic in each iteration, it is not possible to prove the convergence of the algorithm. However, if the algorithm is allowed to converge or iterated to guarantee the monotonicity of the objective, it is possible to prove the convergence based on the discussions provided in the manuscript.

- If there exists a fixed point or a set of fixed points for the original nonconvex problem in (16), then the proposed centralized algorithm in Section III-B and III-C finds at least one such point if iterated until convergence, which is provided in Appendix A. In order to find a unique fixed point, we regularize the objective function in (16) with a quadratic penalty term as discussed in Appendix A to transform the objective as a strongly convex function.
- If the distributed algorithm is carried out for a limited number of iterations, it is not guaranteed to achieve a fixed point even if the outer SCA update is performed for a large number of iterations. By this approach, the distributed approaches are not guaranteed to converge to a stationary point.

In all our simulations on the primal and the ADMM approach, we have set $J_{\max} = 20$ in order to guarantee the monotonicity of the objective.

- Unless the objective function is regularized with a strongly convex term as in Appendix A-C, the uniqueness of the iterates is not guaranteed.
- In each SCA update, if the distributed algorithms are allowed to converge to the centralized solution, then the overall convergence will be a stationary point of the original nonconvex problem by following the same argument as that of the centralized algorithm.
- Since the coupling between the distributed precoder designs is the interference between the BSs and the users, in the ADMM approach, the interference is treated as a local variable, which is then included in the precoder design problem for each coordinating BS. This is treated as a local variable for individual BSs. Note that the local variable is an assumption made by the BS on the actual interference caused by the neighboring BSs. Since the actual interference caused is different, the consensus has to be made between the local interference variable maintained at each BS with the global consensus interference variable, which is nothing but the average between the corresponding BSs interference. These discussions have been made in the revised manuscript in Section IV-B. For reference purpose, we have also referred the interested reader to [11], which discusses exclusively about the ADMM approach. Upon convergence of the ADMM approach, the interference vector is equal to the actual interference seen in the network.

References

[1] M. Hanif, L.-N. Tran, A. Tolli, and M. Juntti, "Computationally Efficient Robust Beamforming for SINR Balancing in Multicell Downlink With Applications to Large Antenna Array Systems," *IEEE Trans. Commun.*, vol. 62, no. 6, pp. 1908–1920, June 2014.

Response to Reviewer - 3's Comments

The authors have introduced changes in the manuscript that improved the paper's quality. Additionally, the authors have taken into account the reviewers' comments giving clarifications and modifying the content when required. More specifically, the following aspects have been treated:

Reply: We thank the reviewer for recognizing the changes made on our earlier manuscript. We also thank the reviewer for providing constructive comments to help us to improve our manuscript.

1. Convexity of problem (16). The paragraphs surrounding (16) allow a better understanding of the usage of the additional variables, i.e. gamma and beta, to remove the equality constraint in (2). For the reviewer remains however unclear, the procedure/criterion to determine the operating point for the parameter $\tilde{\beta}$ required in (19) and used in the convex subproblems (20) and (21).

Reply: We thank the reviewer for the comment. We have included additional information on how to choose the operating point while initializing the iterative algorithm. It is included in the fourth paragraph following (23). We have also provided a detailed discussion on how to update the operating point after each iteration in the paragraph following (23). Note that the objective function is convex, and therefore we can update operating point for the current step with the solution obtained in the previous iteration as

$$\tilde{\mathbf{m}}_j^{(i)} = \mathbf{m}_j^{(i)} \tag{4}$$

where $\tilde{\mathbf{m}}_j^{(i)}$ denotes the operating point for the $j+1^{\text{th}}$ SCA iteration in the i^{th} alternating optimization (AO) step and $\mathbf{m}_j^{(i)}$ is the solution obtained by solving (20) in the j^{th} SCA iteration.

2. Proof of convergence. The proof of convergence introduced by the authors in the Appendix seems correct and enhances the content of the manuscript.

Reply: We thank the reviewer for the comment. We have included several materials to make the proof of convergence more rigorous.

Additional Comments -

(a) - The reviewer considers that closing statements regarding the applicability of the proposed schemes are missing. Since the results are quite similar (when not identical), which formulation is preferable between the centralized schemes? Which one for the distributed solutions?

Reply: We have updated the manuscript to include a discussion on the possible conditions for selecting an algorithm for a specific implementation. The choice of considering a centralized approach is equally good when the number of receive antenna is greater than one, *i.e.*, $N_R > 1$. However, in case of a single antenna receiver, the joint space-frequency resource allocation (JSFRA) formulation outlined in Section III-B is more efficient than the mean squared error (MSE) reformulation, as there is no need for the receiver update, thereby reducing the complexity significantly.

As far as the distributed approaches are concerned, the primal and the ADMM schemes are equally favorable when $N_R = 1$ since the signaling required to be exchanged between the coordinating BSs involves only the scalar interference values. However, when $N_R > 1$, then the KKT based distributed approach is more efficient than the primal or the ADMM schemes, since it has less signaling overhead ~~for a throughput improvement~~. The choice of selecting ℓ_q norm is discussed in Section II-B last paragraph. The above information is included in the paragraph before the simulations section in Section IV.

However, if the reviewer is interested to understand the performance of the KKT based scheme in a fading environment, please refer to the response provided for the reviewer 2, question 2 in this document for more details.

(b) - The last discussion in section IV-C could benefit from restructuring. The information on how to obtain a practical distributed precoder design and to avoid backhaul exchange is too condensed and difficult to understand.

Reply: It has been rewritten to improve the readability as suggested by the reviewer. We have included the reference [28] for more illustrative discussions on the backhaul signaling.

(c) - For the simulation results, why not to unify configurations when possible? Having to read a different configuration for each graph is cumbersome and no additional comparisons are possible between figures. E.g. PL uniformly distributed between [0,-6] dB in Fig. 2 and [0,-3] dB in Fig. 3.

Reply: We thank the reviewer for the comment. The main purpose of considering several path loss models is simply to show that the proposed algorithms works reasonably well on various scenarios. For benchmarking, we have also included the centralized algorithm to draw the difference between the other schemes. We are conducting more numerical experiments to have a more complete comparison of the schemes. Unfortunately those results cannot be included in the current manuscript due to space limitation. We aim to make them available online as a companion technical report whey they are ready.

(d) - In Fig. 1, the description of the system model does not agree with the statement of N = 3 subchannels.

Reply: We apologize for the error which has been fixed in the revised manuscript.

(e) - In Fig. 4(b), the performance of Q-WSRME seems to be (in average) worse than Q-WSRM. However, that should not be the case, since Q-WSRME is taking into account the over allocation. Any reason for this?

Reply: Indeed the performance of the Q-WSRME is better than Q-WSRM in Fig. 4(b). The misleading observation was probably due to the visualization effect where the colors representing the performance of sum arrivals and Q-WSRME are quite similar. We have updated Fig. 4 (and all other figures) in the revised manuscript to avoid such visualization problems.

(f) - p 6, col 2, row 49: it should be $t_{l,k,n}$ instead of $t_{l,n,k}$

Reply: This issue is corrected in the revised manuscript.

(g) - p 10, col 1, row 28: is it λ a dual variable?

Reply: We thank the reviewer for pointing out the mistake. In fact, it should be the dual variable $\sigma_{l,k,n}$ and not $\lambda_{l,k,n}$. We have fixed this problem in the revised manuscript.

(h) - p 10, col 2, row 59: typo wHith

Reply: We have fixed this error.

(i) - In general, a grammar check is recommended, several mistakes with respect to singular and plural nouns have been observed, e.g. -p 6, col 2, row 56 "... for A fixed receiverS" is not correct.

Reply: We thank the reviewer for the comment and we have proofread the manuscript carefully to eliminate all grammar errors.

Response to Reviewer - 4's Comments

This reviewer's concerns have been addressed, and this manuscript is now deemed fit for publication.

Reply: We thank the reviewer for the constructive comments and recommending the revised manuscript for the publication.