ICT-248891 STP FREEDOM

# Femtocell-based network enhancement by interference management and coordination of information for seamless connectivity 

D3.2
Interference Coordination Protocols in Femto-based Networks



#### Abstract

: This document looks into coordinated radio resource allocation algorithms in deployments of huge number of femto-access points (FAP) by exploiting the message exchange at control-plane level through the wired ISP backhaul link. The algorithms present in this document are based on decentralized cooperative strategies supported by Game Theory. Decentralized resource allocation algorithms based on price exchange have been derived under different criteria: guaranteeing a minimum rate with the minimum power, maximizing the weighted sum-rate of the system, maximizing the opportunistic throughput when there are different source of randomness like random link failures and noise quantization or when the activity of macro-users is not known by FAPs (coordinated channel sensing or modelling the activity). Additionally, a genetic-based resource allocation is investigated, providing centralized and decentralized algorithms, to which the Game Theory solutions are challenged . In all cases, the techniques derived require modifications of the current LTE-A standard which are identified. Finally, some of those algorithms have been evaluated in a realistic corporate scenario elucidating the advantages of coordinated resource allocation techniques proposed.


Keyword list: Cooperative games, MIMO, Genetic optimization, Decentralized optimization, resource block optimization, multi-user communication, opportunistic throughput, backhaul limitation.

## Document Revision History

| DATE | ISSUE | AUTHOR | SUMMARY OF MAIN CHANGES |
| :--- | :--- | :--- | :--- |
| $11-07-2011$ | a | All partners | Contributions are collected and the first draft is <br> elaborated |
| $9-09-2011$ | b | UPC | 2nd draft version including the latest <br> contributions |
| $15-11-2011$ | c | UPC | 3rd draft version with the evaluation of the <br> UPC's algorithms in a sector-based scenario |
| $23-11-2011$ | d | All partners | Final contributions. Updates on INFO <br> contribution, and comparison of two algorithms <br> in a common scenario, UPC and DUNE |
| $28-11-2011$ | e | UPC | New contributions from UPC related to a LTE <br> Standard contribution and Patent Application |

## Executive Summary

The theme addressed in this work is the development of distributed interference-aware resource allocation algorithms that coordinate different FAPs by exchanging parameters at control-plane level through the backhaul link. We are assuming a dense femtocell deployment and hence a high level of interference that negatively influences the system spectral efficiency. Under the assumption that femtocells are connected through a wired ISP backhaul link, the quality of that connection impacts the exchange of information messages at the control-plane level and needs to be taken into consideration.
We propose a set of algorithms that are able to address the resource allocation of the system in a decentralized way under the following criteria:

- Power minimization. The objective is to allocate resources (power allocation) efficiently using the minimum power and guaranteeing a minimum rate per user (in section 5).
- Weighted sum-rate (WSR) maximization. Assuming that the attained bitrate is weighted by a factor that accounts for the priority of each user in a scenario where there are multiples FUEs per FAP, the resource allocation (power allocation and carrier assignment per FAP) is optimized so that the weighted sum-rate is maximum, constrained to a maximum transmitted power and a maximum rate per FAP (given by the quality of the backhaul) (in section 6).
Those criteria are adopted following a decentralized operation principle, and their performance is compared to a centralized solution based on Genetic Algorithm (in section 8). This latter solution is attractive for its simplicity, modularity and suitable for both synchronous and asynchronous scenario.
A femto-cell system in which channel access is not coordinated with the Macro Users' network is inherently subject to the performance of the new system functionalities that are introduced to implement the proposed algorithms. We follow the approach of using a statistical model for a particular aspect of the problem and build on this model to devise suitable algorithms. Namely, in Section 7, we describe algorithms based on:
- Coordinated channel sensing. FAPs/FUEs track the activity of the interference (i.e. MBS) and exchange the measures over nearby FAPs. Using that set of measures, each FAP is able to allocate the resources in a smart way when the detection and channel access parameters are jointly done. The spectrum sensing detection performance has a key role in the definition of the optimal access strategy. The performance metric to maximize is opportunistic throughput, a notion that redefines the concept of throughput taking into account the presence of the macro users' communication, which the proposed system should preserve as much as possible.
- Markovian interference model. The activity on each frequency subchannel is assumed to follow a general Markov model whose parameters are estimated on the basis of recorded observations of the interference over time.
- Random failures and noise quantization. The exchanged control messages (prices) are quantized or possibly loss in a random way.
The main benefits showed by the proposed algorithms for coordinating the generated interference are:
- Possibility of addressing a joint resource allocation in a decentralized way.
- To partially overcome the inefficiency of Nash Equilibria present in similar algorithms that avoid using exchanging messages at the control plane (like those in [FREEDOM-D3.1]).
- Messages only have to be exchanged amongst the interfering neighboring terminals, so that, algorithm are scalable with the number of FAPs.
- Finally, a simpler and better adapted LTE based pricing solution has been derived for MBSFAPs interference coordination and the required standard enhancements have been identified.


## DISCLAIMER

The work associated with this report has been carried out in accordance with the highest technical standards and the FREEDOM partners have endeavoured to achieve the degree of accuracy and reliability appropriate to the work in question. However since the partners have no control over the use to which the information contained within the report is to be put by any other party, any other such party shall be deemed to have satisfied itself as to the suitability and reliability of the information in relation to any particular use, purpose or application.

Under no circumstances will any of the partners, their servants, employees or agents accept any liability whatsoever arising out of any error or inaccuracy contained in this report (or any further consolidation, summary, publication or dissemination of the information contained within this report) and/or the connected work and disclaim all liability for any loss, damage, expenses, claims or infringement of third party rights.

## Table of Contents

1 INTRODUCTION ..... 14
2 SCENARIOS ..... 17
2.1 BUSINESS SCENARIOS ..... 17
2.2 TECHNICAL SCENARIOS ..... 17
3 SYSTEM ASSUMPTIONS ..... 19
3.1 PHY ASSUMPTIONS. ..... 19
3.2 MAC SUPPORT ..... 19
3.3 NETWORK ARCHITECTURE ..... 20
4 SIMULATION METHODOLOGY ..... 21
5 DECENTRALIZED POWER MINIMIZATION ..... 23
5.1 SISO CASE AND PRICING MECHANISMS ..... 23
5.1.1 Preliminaries ..... 23
5.1.2 Problem formulation ..... 24
5.1.3 Numerical results. ..... 27
5.2 MIMO CASE ..... 28
5.2.1 Preliminaries ..... 28
5.2.2 Description ..... 28
5.2.3 Numerical results. ..... 32
5.3 Conclusions ..... 34
6 DECENTRALIZED WEIGHTED SUM RATE MAXIMIZATION ..... 35
6.1 SISO CASE: COMPLEX VS. SIMPLE TRANSMITTERS ..... 35
6.1.1 System model ..... 36
6.1.2 Decentralized Resource Allocation ..... 38
6.1.3 Numerical results. ..... 46
6.2 MIMO CASE AND BACKHAUL RATE CONSTRAINTS ..... 49
6.2.1 Description ..... 49
6.2.2 Radio resource allocation at the $k$-th FAP ..... 51
6.2.2.1 Rate and Power allocation given a certain RB assignment. ..... 51
6.2.2.2 Rate, Power and RB optimization ..... 52
6.2.3 Generation of prices. ..... 53
6.2.4 Numerical results. ..... 53
6.3 Conclusions ..... 56
7 DECENTRALIZED COORDINATION STRATEGIES BASED ON STATISTICAL MODELLING ..... 57
7.1 PRELIMINARIES ..... 57
7.2 COORDINATED CHANNEL SENSING ..... 61
7.2.1 System and detection models ..... 61
7.2.2 Maximization of the aggregated opportunistic throughput ..... 62
7.2.3 Opportunistic throughput maximization: A game theoretic approach ..... 66
7.2.4 Numerical results. ..... 70
7.3 DYNAMIC RESOURCE ALLOCATION UNDER MARKOVIAN INTERFERENCE MODEL74
7.3.1 Preliminaries ..... 74
7.3.2 Maximum expected rage game ..... 75
7.3.2.1 Numerical results ..... 78
7.3.3 Min-power game subject to Markovian interference. ..... 80
7.4 Distributed Stochastic Pricing for Sum-Rate Maximization with Random Graph and Quantized Communications ..... 82
7.4.1 Preliminaries ..... 82
7.4.2 System model ..... 83
7.4.3 Distributed Stochastic Pricing Algorithm (DSPA) ..... 87
7.4.4 Numerical results. ..... 88
7.5 Conclusions ..... 89
8 CENTRALIZED DYNAMIC INTERFERENCE MANAGEMENT ..... 91
8.1 PRELIMINARIES ..... 91
8.2 CELL SYSTEM OPTIMIZATION BY GENETIC ALGORITHM. ..... 93
8.2.1 Distributed or centralized implementation. ..... 93
8.2.2 Working of the GA implementation ..... 94
8.2.3 GO implementation for a set of FAPs. ..... 95
8.2.4 GO formulation. ..... 96
8.3 GO FOR FLAT FADING SCENARIO ..... 100
8.3.1 Fitness function. ..... 100
8.3.2 Analysis of topologies for a single block ..... 101
8.3.2.1 Flat fading channel ..... 101
8.3.2.2 GO for frequency selective channel ..... 105
8.4 GENETIC ALGORITHM AND SCALABILITY ..... 107
8.4.1 Benefits of clustering ..... 107
8.4.2 Clustering of FAPs ..... 108
8.5 SYNCHRONIZATION ISSUES AND GO ..... 109
8.5.1 Asynchronous scenario ..... 109
8.5.2 Design principles and constraints ..... 112
8.5.3 GA tested on a fully asynchronous frame sequence ..... 112
8.5.3.1 Results of GA tested on fixed frame sequences. ..... 112
8.5.3.2 Results of GA tested on randomly changing frame sequences ..... 113
8.5.3.3 Conclusions on second tier synchronization and GA ..... 115
8.6 CONCLUSIONS OF THE GO OPTIMIZATION. ..... 115
9 SIMULATION RESULTS ..... 116
9.1 RESOURCE ALLOCATION BASED ON WEIGHTED SUM-RATE (WSR) MAXIMIZATION ..... 117
9.1.1 Small Corporate configuration. ..... 118
9.1.2 Simulated Corporate configuration. ..... 121
9.1.3 Conclusions ..... 124
9.2 RESOURCE ALLOCATION BASED ON POWER MINIMIZATION ..... 124
9.2.1 Results for a single scenario with a fixed number of FAPs. ..... 125
9.2.2 Results for several scenarios ..... 128
9.2.3 Results for varying FAP density. ..... 132
9.2.4 Conclusions ..... 135
9.3 DYNAMIC RESOURCE ALLOCATION UNDER MARKOVIAN INTERFERENCE MODEL135138
9.5 GO FOR REALISTIC FREQUENCY SELECTIVE CHANNEL - POWER CONSTRAINED1409.5.1 Macro-network UL140
9.5.2 Macro-network DL ..... 141
9.5.3 Overall results ..... 142
10 IMPLEMENTATION NOTES ..... 145
10.1 SCALABILITY ..... 145
10.2 APPLICABILITY ..... 146
10.3 COMPLEXITY......................................................................................................... 147

11 LTE-BASED RESOURCE ALLOCATION ............................................... 149
11.1 LTE-A ADAPTED PRICING MECHANISMS ............................................................ 149
11.1.1 A pricing based mechanism for MCS and bandwidth part selection................ 149
11.1.1.1 Fundamentals ............................................................................................ 150
11.1.1.2 Procedure for coordinated bandwidth part selection .................................. 153
11.1.1.2.1 Sensing and sharing the CQI degradation .............................................. 154

CQI detection in the serving cell..................................................................................... 154
CQI degradation in the serving cell............................................................................... 154
11.1.1.2.2 Decision making..................................................................................... 155
11.1.2 Exponential Effective SINR Mapping (EESM) .................................................. 156
11.1.3 Simulation results ............................................................................................... 158
11.1.4 Conclusions and recommended actions............................................................. 162
11.2 RATE MAX OR POWER MIN UNDER INTERFERENCE-POWER CONSTRAINTS.. 163
11.2.1 LTE signals and measurements ......................................................................... 163
11.3 RESOURCE BLOCK POWER ALLOCATION IN LTE FEMTOCELL NETWORKS ... 164

12 GENERAL CONCLUSIONS ....................................................................... 167
13 SUMMARY OF RESULTS TOWARDS OTHER ACTIVITIES............... 169
13.1 Towards WP4 ........................................................................................... 169
13.2 TOWARDS WP5 ........................................................................................... 169
13.3 TOWARDS WP6 ..................................................................................................... 169

14 APPENDIX........................................................................................................ 170
14.1 SIMULATION METHODOLOGY ............................................................................. 170
14.2 GEOMETRY OF THE SCENARIO............................................................................ 171
14.3 PATHLOSS MODELS............................................................................................... 172
14.3.1 ITU and 3GPP formulation ............................................................................... 172
14.3.2 Pathloss and clustering ..................................................................................... 173
14.4 MORE RESULTS OF SECTION 8.3.2.1..................................................................... 174
14.4.1 Frequency GO under power constraint.............................................................. 177
14.4.2 GO output for different FAP load FAP/FUE couples ...................................... 177

## References \& Standards

[3GPP TS 36.213]
[3GPP-TR36.814]
[Agustin11a] $\begin{aligned} & \text { A. Agustin, J. Vidal, O. Muñoz-Medina, "Interference Pricing for Self- } \\ & \text { organisation in OFDMA Femtocell Networks", European Workshop on } \\ & \text { Broadband Femtocell Networks - Future Network and Mobile Summit } \\ & \text { (FuNeMS), Warsaw, Poland, June } 2011\end{aligned}$
[Agustin11b] A. Agustin, J. Vidal, O. Muñoz-Medina, J.R. Fonollosa, "Decentralized Weighted Sum Rate Maximization in MIMO-OFDMA Femtocell Networks", IEEE Proc. the Second Globecom 2011 Workshop on Femtocell Networks (GC'11 Workshop-FEMnet), Houston,USA,Dec. 2011
[Ban03] S. Bandyopadhyay, E. J. Coyle, "An Energy Efficient Hierarchical Clustering Algorithm for Wireless Sensor Networks", IEEE INFOCOM 2003, vol. 3 pp. 1713-1723.
[Barbarossa10] S. Barbarossa, S. Sardellitti, A. Carfagna, P. Vecchiarelli, "Decentralized Interference Management in Femtocells: A Game-Theoretic Approach", Proc. CrownCom 2010, Cannes, France, June 2010.
[Barbarossa11a] S. Barbarossa, A. Carfagna, S. Sardellitti, M. Omilipo, L. Pescosolido "Optimal Radio Access in Femtocell Networks Based on Markov Modeling of Interferers' activity", in Proc. IEEE ICASSP 2011
[Barbarossa11b] S. Barbarossa, S. Sardellitti, A. Carfagna, "Decentralized Resource Allocation in Femtocell Networks Based on Game Theory and Markov Modeling", submitted for publication, available online at http://infocom.uniroma1.it/sergio/jsac-femtocells.pdf.
[Bertsekas95] D.P. Bertsekas, Nonlinear Programming. Athena Scientific, Belmont, Massachusetts, 1995.
[Bertsekas03] D.P. Bertsekas, Convex Analysis and Optimization. Athena Scientific, Belmont, Massachusetts, 2003.
[Boyd04] S. Boyd, L. Vandenberghe, Convex optimization, Cambridge University Press, 2004.
[Cendrillon05] R. Cendrillon, M. Moonen, "Iterative Spectrum balancing for digital subscriber lines, in Proc. IEEE Intl. Conf. on Communications (ICC), June 2005
[DiLorenzo 11] P. Di Lorenzo, S. Barbarossa, M. Omilipo, "Distributed sum-rate maximization in femtocell networks with random graph and quantized communications", submitted to IEEE Trans. on Signal Processing, 2011
[DOCOMO] TSG-RAN Working Group (Radio) meeting \#52 R4-093244 Shenzen, 2428 August 2009, NTT DOCOMO Downlink Interference Coordination Between eNodeB and Home eNodeB
[Facchinei03] F. Facchinei and J.-Pang, Finite-Dimensional Varational Inequalities and Complementarity Problems, New-York: Springer-Verlag, 2003.
[Facchinei2010] F. Facchinei, V. Piccialli and M. Sciandrone, "Decomposition algorithms for generalized potential games", Computational Optim. and Appl., Springer-Verlag New York Inc., May 2010.

| [FREEDOM-D21] | ICT FP7-248891 STP FREEDOM project deliverable D2.1, "Scenario, requirements and first business model analysis", June 2010. [Online]: http://www.ict-freedom.eu/index.php?option=com_spcom\&Itemid=88 |
| :---: | :---: |
| [FREEDOM-D31] | ICT FP7-248891 STP FREEDOM project deliverable D3.1, "Interference identification in the femto-based networks", July 2011. [Online]: http://www.ict-freedom.eu/index.php?option=com_spcom\&Itemid=88 |
| [FREEDOM-D42] | ICT FP7-248891 STP FREEDOM project deliverable D4.2, "Design and evaluation of effective procedures for MAC layer", Dec. 2011. [Online]: http://www.ict-freedom.eu/index.php?option=com_spcom\&Itemid= $=88$. |
| [FREEDOM-D52] | ICT FP7-248891 STP FREEDOM project deliverable D5.2, "System Interference Management Evaluation", Dec. 2011. [Online]: http://www.ict-freedom.eu/index.php?option=com_spcom\&Itemid=88 |
| [Huang06] | J. Huang, R.A. Berry, M.L. Honig, "Distributed Interference Compensation for Wireless Networks", IEEE J. on Selected Areas in Comтип., Vol. 24, pp. 1074--1084, May 2006. |
| [Ho09] | L.T.W. Ho, I.Ashraf, H.Claussen, "Evolving Femtocell Coverage Optimization Algorithms using Genetic Programming", IEEE Personal, Indoor and Mobile Radio Communic. Conference, PIMRC 2009, , pp. 2132-2136, 2009. |
| [http://www.nt.tuwien. ac.at/ltesimulator] | Institute of Communications and Radio-Frequency Engineering, Vienna University of Technology, Austria |
| [ITU1238] | ITU-R P.1238-6, "Propagation data and prediction methods for the planning of indoor radiocommunication systems and radio local area networks in the frequency range 900 MHz to $100 \mathrm{GHz"}, 2009$. |
| [Jorswieck08] | E. A. Jorswieck, E. G. Larsson, and D. Danev, "Complete Characterization of the Pareto Boundary for the MISO Interference Channel", IEEE Trans. on Signal Processing, vol. 56, no. 10, October 2008. |
| [Kum02] | R. Kumar, P.P. Parida, M. Gupta, "Topological design of communication networks using multiobjective genetic optimization", Evolutionary Computation, CEC '02 Proceedings, pp. 425-430, 2002. |
| [Lie98] | K. Lieska, E. Laitinen, J. Latheenmaki, "Radio coverage optimization with genetic algorithms", Ninth IEEE Intern. Symposium on Personal, Indoor and Mobile Radio Communications, PIMRC 1998, pp. 318-322, 1998. |
| [Meu00] | H. Meunier, E.-G. Talbi, P. Reininger, "A multiobjective genetic algorithm for radio network optimization", Proceedings of the 2000 Congress on Evolutionary Computation, pp. 317-324, 2000. |
| [MondererShapely96] | D. Monderer, and L. S. Shapley, "Potential Games", Games and Economics Behaviour, Vol. 14, pp. 124-143, 1996. |
| [Munoz11a] | O. Munoz-Medina, A. Pascual, P. Baquero, J. Vidal , "Preemption and QoS Management Algorithms for Coordinated and Uncoordinated Base Stations", SPAWC 2011, San Francisco, June 2011 |
| [Munozl 1b] | O. Muñoz, J. Vidal, A. Agustín, A. Pascual-Iserte, S. Barbarossa, "Coordinated MIMO Precoding for Power Minimization in Femtocell Systems", The Second GlobeCom 2011 Workshop on Femtocell Networks (GC'11 Workshop - FEMnet).Houston, USA, Dec. 2011 |

[Olfati-Saber07] R. Olfati-Saber, J.A. Fax, R.M. Murray, "Consensus and Cooperation in Networked Multi-Agent Systems," Proc. of the IEEE, Vol. 95, pp. 215233, Jan. 2007.
[Pang08] J.S. Pang, G. Scutari, F. Facchinei, C. Wang, "Distributed Power Allocation With Rate Constraints in Gaussian Parallel Interference Channels", IEEE Trans. on Inform. Theory., Vol. 54, no. 8, pp. 34713489, August 2008.
[Quan08] Z. Quan, S. Cui, H. V. Poor and A. H. Sayed, "Collaborative wideband sensing for cognitive radio," IEEE Signal Process. Mag., Vol. 25, pp. 60-73, November 2008.
[Quan09]
[R3-112701] R3-112701, "Way forward on Carrier based ICIC LTE", AT\&T, China Unicom, Mitsubishi Electric, NEC, Telefonica, TeliaSonera, Verizon Wireless, ZTE
[R3-112752]
[R3-112953] DAC-UPC, Telefonica, Text proposal for carrier-based ICIC TR on use cases for Macro - SC HeNB scenario, 3GPP TSG-RAN WG3 \#74, November 14-18, 2011, San Francisco, USA
[R4-071617] R4-071617, "HNB and HNB-Macro Propagation Models", Qualcomm Europe, 3GPP TSG-RAN Working Group 4 (Radio) meeting \#44bis, October 2007.
[R4-080149]
[Robbins51] H. Robbins, S. Monro, "A stochastic approximation method", The Annals of Mathematical Statistics 22, 1951, no.3, pp. 400-407.
[Scutari08a] G. Scutari, D.P. Palomar and S. Barbarossa, "Cognitive MIMO Radio", IEEE Signal Processing Mag., Vol. 25, pp. 46-59, Nov. 2008.
[Scutari08b] G. Scutari, D. P. Palomar, S. Barbarossa, "Optimal Linear Precoding strategies for wideband noncooperative systems based on Game Theory Part I: Nash Equilibria and - Part II: Algorithms", IEEE Trans. on Signal Proc., vol.56, no. 3 March 2008, pp. 1230-1267
[Seong06] K. Seong, M. Mohseni, J.M. Cioffi, "Optimal Resource allocation for Downlink Systems", in Proc. IEEE Intl. Symposium on Information Theory (ISIT), Seattle, USA, July 2006
[Sesia09] S. Sesia, I. Toufik , M. Baker, LTE-The UMTS Long Term Evolution, John Wiley \& Sons, 2009
[Shi08] C. Shi, R. Berry, and M. Honig, "Distributed Interference Pricing for OFDM Wireless Networks with Non-Separable Utilities", Proc. of Conference on Information Sciences and Systems (CISS),Princeton University, NJ, USA, March 2008
[Shi09a] C. Shi, R. A. Berry, and M. L. Honig, "Monotonic Convergence of Distributed Interference Pricing in Wireless Networks," IEEE Int. Symposium on Information Theory, ISIT 2009, Seoul, Korea, June 2009.
[Shi09b] C. Shi, et al., "Distributed Interference Pricing for the MIMO interference channel", in Proc. IEEE Intl. Conf. on Comm., Dresden, Germany, June 2009.
[Son09] K. Son, B. C. Jung, S. Chong and D. K. Sung, "Opportunistic underlay transmission in multi-carrier cognitive radio systems," Proc. IEEE WCNC 2009, Budapest, Hungary, April 2009.
[Schmidt09]
D. A. Schmidt, C. Shi, R. A. Berry, M. L. Honig, W. Utschick, "Distributed Resource Allocation Schemes: Pricing algorithms for power control and beamformer design in interference networks", IEEE Signal Processing Magazine, Sep. 2009.
[TSG-RAN 2009] TSG-RAN Working Group 4 (Radio) meeting \#52 R4-093244 Shenzhen, 24-28 August 2009, NTT DOCOMO Downlink Interference Coordination Between eNodeB and Home eNodeB.
[Vishwanath04] S. Vishwanath, N.Jindal, A.Goldsmith, "On Duality of Gaussian MultipleAccess and Broadcast channels", IEEE Tran. on Information Theory, vol. 50, no. 5, pp. 768-783, May 2004
[Yu07] W. Yu, "Multiuser Water-filling in the presence of Crosstalk", in Proc. of Information Theory and Applications Workshop (ITAW), Jan-Feb 2007
[Yu06] W. Yu, R. Lui, "Dual methods for nonconvex spectrum optimization of multicarrier systems", IEEE Trans. on Comm., vol.54, no.7, Jul. 2006
[Yu02] W. Yu, G. Ginis, and J.M. Cioffi, "Distributed Multiuser Power Control for Digital Subscriber Lines", IEEE J. of Selected Areas in Commun., Vol. 20, pp. 1105-1115, June 2002.
[Zhang10] R. Zhang, S. Cui, "Cooperative Interference Management with MISO Beamforming", IEEE Trans. on Signal Processing, vol. 58, no. 10, pp. 5450-5457, October 2010.

## List of abbreviations \& symbols

| AWGN | Additive White Gaussian Noise |
| :---: | :---: |
| ABS | Almost Blank Sub-frames |
| BS (or MBS) | Macrocell Base Station |
| CA | Carrier Aggregation |
| CDF | Cumulative Density Function |
| CQI | Channel Quality Indicator |
| CSG | Closed Subscriber Group |
| DL | Downlink |
| DSPA | Distributed Stochastic Pricing Algorithm |
| DTMC | Discrete Time Markov Chain |
| FAP | Femto Access Point |
| FAP-MS | FAP Management System |
| FAP-GW | FAP-Gateway |
| FUE | Femto User Equipment |
| GA | Genetic Algorithm |
| GNE | Generalized Nash Equilibrium |
| GPG | Generalized Potential Game |
| GT | Game Theory |
| IGPA | Iterative Gradient Projection Algorithm |
| ISP | Internet Service Provider |
| KKT | Karush-Khun Tucker |
| LTE | Long Term Evolution |
| LTE-A | Long Term Evolution Advanced |
| MADP | Modified Asynchronous Distributed Pricing |
| MCS | Modulation and Coding Scheme |
| MNO | Mobile Network Operator |
| MUE | Macro User Equipment |
| MIMO | Multiple Input Multiple Output |
| MISO | Multiple Input Single Output |
| MIWF | Multilevel Iterative Water-Filling |
| MRT | Maximum Ratio Transmission |
| MUI | Multi User Interference |
| NE | Nash Equilibrium |
| OFDM | Orthogonal Frequency Division Multiplexing |
| OFDMA | Orthogonal Frequency Division Multiple Access |
| PM | Pricing Mechanisms |
| PRB | Physical Resource Block |
| UL | Uplink |
| RM | Robbins Monro |
| RRA | Radio Resource Allocation |
| RRC | Radio Resource Control |
| RRM | Radio Resource Management |
| RS | Reference Sequence |
| VI | Variational Inequality |
| WPx | Work Package x |
| WSR | Weighted Sum Rate |
| ZF | Zero-forcing |


| Description | Femto Forum | 3GPP/LTE-A | 802.16/WiMAX | FREEDOM |
| :---: | :---: | :---: | :---: | :---: |
| Base Station of the macrocell | Macro Node B <br> (MNB) | Macro Node B <br> (MNB) | Macro Advanced BS <br> (Macro ABS) | $\begin{gathered} \text { Macro BS } \\ \text { (MBS or BS) } \end{gathered}$ |
| User attached to the macrocell | Macro User <br> Equipment (MUE) | Macro User <br> Equipment <br> (MUE) | Advanced (AMS) | Macro User <br> Equipment <br> (MUE) |
| femtocell | Femto Acces Point or Home Node B (HNB or FAP) | Home Node B (HNB) <br> $\mathrm{H}(\mathrm{e}) \mathrm{NB}$ includes HeNB and FAP | Femto ABS | FAP |
| Access Station, can be either FAP or BS |  |  |  | AS |
| User attached to the femtocell | Femto UE or Home UE (HUE or FUE) | Home User Equipment (HUE) | AMS | Femto UE <br> (FUE) |
| The network element that terminates TR-069 with the femtocell to hankdel the remote management of a large number of femtocells | Auto-configuration Server (ACS) | Home NodeB Management System (HMS) | femtocell <br> Management system <br> in Femto Network Service Provider (Femto - NSP) | Femtocell <br> Management system FAP Management System <br> (FAP-MS) |
| The network element that directly terminates the Iuh interface with the femtocell and the existing IuCS and IuPS interface with the core Networks | FAP Gateway (FAP-GW) | Home Node B gateway (FAP-GW) | Femto Access Service Network Gateway (Femto-ASN GW) | Femto gateway FAP gateway (FAP-GW) |
| Handovers femtofemto, BS-femto | handover (or handoff) | femto-macro: handout macro-femto: handin | handover (or handoff) | Handover <br> Hand-in: handover from macro to femto Hand-out: handover from femto to macro |

Table 1. Summary of terminology used in standard and within FREEDOM

## 1 INTRODUCTION

This document explores the activities undergone in 3A2 regarding the use of the wired ISP backhaul link for interference management in a cellular scenario where MBS and FUE might coexist on the same bands. On a first phase, it will be assumed that the available bandwidth and delay constraints on the backhaul link are enough to support parameter exchange at control-plane level among different network nodes. As the backhaul quality may not be assumed perfect, an insight will also be done to evaluate the impact of quantizing the exchanged information and the effect of packet losses.

Unlike the approach in [FREEDOM-D31], where no control information is assumed to be exchanged through the backhaul link, we can now coordinate the resource allocation in a distributed way, whereby each source takes decisions independently in an egoistic way for optimizing his resources (see for example [Pang08], [Scutari08b] for the SISO case or [Jorswiek08] for the MISO case). While the purely egoistic strategy for the sources has been proved to be inefficient from a global point of view (especially in high interference scenarios), we elaborate on the results presented in [Huang06], [Shi08] and [Shi09] to derive a decentralized resource allocation algorithm, where some parameters are exchanged among the sources in the neighbourhood. As a consequence, sources become more altruistic since the impact of their decisions is being considering beforehand.

In this respect, section 5 is devoted to optimize the resource allocation (transmitted power, allocated resource blocks and MIMO precoders) that minimizes the transmitted power and guarantees a minimum rate per user. SISO and MIMO cases are addressed. Section 6 tackles the decentralized resource allocation that maximizes the WSR assuming a maximum backhaul link capacity. Simple transmitters (each source assigns one carrier to at most one destination) and complex transmitters (multiple destinations per carrier using dirty-paper coding) are analysed. In all cases, each source can serve multiple destinations under an OFDMA scheme. All those solutions scale with the number FAPs and hence are suitable for a massive deployment and the presence of a MBS is considered as an additional FAP whose transmitter power. The possibility of associating different priorities allows preserving QoS to the macro user equipments (MUE). This way, prices exchange gives a natural support to different QoS grades for different users.

Section 7 investigates how the decentralized algorithms need to be adapted to combat different sources of randomness. In this respect, section 7.2 and 7.3 look into the case where macro BS (MBS) is serving MUEs in the same band employed by FAPs, but FUEs are not allowed to interfere the macro user communications. It is assumed here that MBS does not inform of the usage of resources to the FAPs. Therefore, FAPs must estimate somehow the activity of the MBS to avoid introducing interference. Two approximations are followed: section 7.2 assumes that neighboring FAPs can perform a coordinated channel sensing to improve the estimation of the MBS activity; section 7.3 models the MBSs activity by a Markov model and under such model the radio resource allocation at the FAPs is performed. In both cases, the obtained throughput is opportunistic, due to the presence of uncertainties. Finally, section 7.4 addresses the randomness due to the random link failure or to the dithered quantization of messages exchanged by FAPs.

In general, the analysed decentralized algorithms do not provide globally optimal solutions, because the resource allocation over an interference network is not a convex problem. In order to evaluate the loss of performance (the so-called price of anarchy), section 8 looks into a centralized algorithm based on Genetic Optimization that tackles the whole problem. In such a case, it is assumed that there is a central node able to collect all the required information from the different nodes of the network. This approach will be considered as a benchmark of the previous decentralized solutions.

The centralized approach is applicable straightforwardly for scenarios with a low/moderate number of nodes, although the computational load increases with the number of users. In a scenario with a high number of nodes, an efficient implementation of the algorithm requires a fragmentation of the network
in subsets, non-interfering each other or partially interfering, as happens for example in a cluster of FAPs placed in the same building. The central computing unit can optimize only the UL, only the DL, or both of the FAPs/FUEs network, based on MNO strategy. Information collected from the nodes of the network will be average SNRs of FAPs, FUEs, or both, respectively.

Section 9 is devoted to evaluate the techniques and compare the decentralized and centralized algorithms in a common scenario. In section 10 we analyse the applicability, scalability and complexity of the investigated techniques in sections $5-8$. On the other hand, section 11 presents algorithms adapted to the current LTE standard, like a pricing solution for MBS-FAPs interference coordination, an algorithm for maximizing the rate or minimize the power and a resource block power allocation.

Finally, section 12 present the conclusions obtained in this work.
A summary of achievements obtained in activity 3 A 2 follows:

- Pricing-based algorithms allow performing a joint resource allocation in a decentralized way, while maintaining a given QoS.
- We have proposed a decentralized algorithm that designs the resources in order to guarantee a minimum rate for all the users in the system, if the rate is feasible. In those cases where nonpricing algorithms also satisfy that minimum rate, the proposed algorithm defines a low complexity power reduction procedure.
- If the objective is to maximize the weighted sum-rate of the system, we have observed that pricing-based algorithms are able to improve the outage rate of the users in the system by a factor of 2-3. In case where each source serves multiple users in an OFDMA fashion, then the spectral efficiency is also improved when the resource blocks (or carriers) assignment is jointly optimized with transmitted power and spatial precoders.
- A centralised approach to resource allocation has been implemented by a Genetic Algorithm (GA), attractive for its simplicity, modularity and suitable for both synchronous and asynchronous scenarios. Its only inputs are the mean SNIRs from FAPs and FUEs and outputs the "optimal" (in the sense of the adopted metric) radio resources allocation. Its main drawbacks rely in the convergence time and in the distance from the optimal solutions. Many computational aids can help to overcome the latter, but the compliance between the convergence time and the requirements at system level cannot be insured a priori in all scenarios. The results of the simulations in an asynchronous scenario have shown to be consistent with the "physical meaning" implemented in the adopted fitness functions, on which a specific effort must be spent in order to obtain a meaningful target and functional slopes easing the GA convergence capabilities.
- We have determined the optimal access strategy when relying on spectrum sensing to decide whether a channel is left unused by Macro User using the notion of opportunistic throughput.
- We have derived a decentralized iterative water filling algorithm that incorporates a Markovian modelling of the resource block use by Macro users in an OFDMA-like time frequency frame structure. With respect to conventional IWFA algorithm, our proposed one exploits not only on the frequency domain, but also the time domain while allocating resources among FAPs. This creates a benefit in terms of the frequency with which the algorithm needs to be run to adapt to the macro users' activity.
- Assuming a decentralized resource allocation strategy based on the local exchange of interference prices, we have proposed a stochastic algorithm which takes into account the nonidealities of realistic communications, i.e. quantization noise and random link failures.
- A version of the cooperative game algorithms with a higher degree of compliance with LTE has been obtained for MBS-FAPs interference coordination. The solution and the required standard enhancements have been presented in the 3GPP meeting held in San Francisco in November 2011 [R3-112752]. A contribution for the specifications of the DL MeNB-HeNB use case within RAN3, addressing DL interference MeNB-HeNB scenario and Operational requirements was generated and submitted to 3GPP as document [R3-112953].
- An algorithm for rate maximization or power minimization under interference power constraints is patent application pending.


## 2 SCENARIOS

### 2.1 Business scenarios

The techniques investigated in this work envision the business scenarios BM1 and BM3 depicted in Figure 1 and defined in [FREEDOM-D21]. Both scenarios assume a highly dense femtocell network coexisting with a macrocell deployment. Scenario BM1 (Figure 1-left) defines a scenario with a variable FAP density and relatively large coverage area, like dense-urban or urban areas. The user density is high. On the other hand, scenario BM3 describes a residential scenario with user density corresponding to a sub-urban area. The maximum number of FUEs served by each FAP and the FAP parameter settings has been defined in section 5 of [FREEDOM-D21].


Figure 1. Network configurations for the business scenarios analyzed in this work.
Left - Corporate customer in a dense urban/semi-urban area.
Right - Residential customer in urban/sub-urban area

### 2.2 Technical scenarios

The technical scenario investigated in this work consists of a single macrocell served by one Macro base-station (MBS) coexisting with several femto access points (FAPs) as it is sketched in Figure 2. The user equipments (UEs) served by the MBS are denoted by MUEs while those UEs served by a FAP are identified as FUEs. MBS and FAPs are able to serve multiple users simultaneously, but each UE is only associated to a single source (MBS or FAPs).


Figure 2. Technical scenario addressed in this work

In contrast to the MBS which is placed by the mobile operator, FAPs will be installed by the endusers. That means that the positions become random and the generated interference has to be managed properly. In this work we exploit the fact that MBS and FAPs are connected through a backhaul link that allows the exchange of parameters at the control-plane level, in order to design the techniques to manage the interference.

The techniques investigated are developed in a framework compliant with the LTE air-interface based on Orthogonal Frequency Division Multiple Access (OFDMA) with multi-antenna terminals. Those techniques can be applied to the uplink (UL) and downlink (DL).

Three scenarios describe the type of coexistence between the macro cell and femto cells in terms of occupied bandwidth and generated interference:

1. MBS and FAPs operate in orthogonal (or non-overlapped) bands. The resources allocated to the MUEs and to the FUEs are orthogonal while the band employed by all FAPs is the same. In this case, MBS perform the resource allocation to its associated MUEs independently from the resource allocation performed by the FAPs to their associated FUEs.
2. MBS and FAPs operate in the same band with same role. In this case MBS and FAPs perform a joint resource allocation and they are mutually influenced by the decisions taken.
3. MBS and FAPs operate in the same band with different role. Here, FAPs design their resource allocation in a jointly way but taking into account their impact on MUEs in terms of achieved rate or generated interference. In contrast, MBS do not take into consideration the generated interference to the FUEs.

In must be emphasized that in all cases, FAPs must perform a joint resource allocation over the same band. In this regard, the second scenario is easily addressed by the same algorithm considered for the first scenario at the FAPs, but assuming the MBS as an additional entity with the appropriate channel models and power transmission levels. In this regard, the techniques investigated in sections 5.1, 5.2, $6.1,6.2$ and 8 apply to technical scenarios 1 and 2 . On the other hand, the work presented in sections 7.2 and 7.3 consider the technical scenario 3. Finally, the technique introduced in section 7.4 accounts for technical scenarios 1 and 2 assuming a quantified (or missing) exchanged information.

Notice that the techniques are analysed in a general framework, without imposing any constraint on neither the current standard (LTE or WiMAX) nor the UL/DL duplexing mode. However, section 10 is devoted to address how the new techniques fit in the current version of LTE and WiMAX standards and what changes would be required.

## 3 SYSTEM ASSUMPTIONS

### 3.1 PHY assumptions

We consider the typical structure of an OFDMA-based system. For the sake of providing order of magnitudes of the involved quantities, (i.e. Physical Resource Block (RB) size, frame duration, intersubcarrier spacing, etc.) we will refer to the LTE standard [3GPP-TR36.814]. However, our results will be general enough to be extended to the WiMAX case. Other assumptions are:

- The channel state information (CSI) of the different links is perfectly known at receiver and transmitter side thanks to the use of pilot sequences present in the different frames and the proper codebooks to feedback the channel state. The overhead and estimation errors are not considered in this work.
In the LTE standard, there are a set of tones that allows estimating the channel over all the frequencies considered for the transmission.
The PHY and MAC overheads introduced by the control channels are not considered. With respect the algorithm presented in section 7, a study of the impact of this overhead on the performance will be presented in [FREEDOM-D5.2].
- All power available at the terminals is devoted for the data channels.
- When the spectral efficiency and outage rate are evaluated, they are obtained as a result of the optimization based on Shannon rate formulas, but assuming a maximum bitrate as a consequence of the maximum MCS allowed in the standard.
- The channel coefficients over the different carriers of the same radio block (RB) are assumed to be constant over the scheduling period, i.e. we assume a slowly varying frequency selective fading channel. This assumption looks reasonable, especially in the context of indoor communications..


### 3.2 MAC support

The techniques investigated in this work assume the following features that need to be supported by the MAC layer:

- FAPs are synchronized at frame level with the appropriate accuracy such that the reference sequences (RS) can be found and detected. As over-the-air synchronization with MBS cannot be assumed in all cases, network-based synchronization using IEEE 1588 PTP is assumed. Nevertheless, centralized GO algorithm investigated in section 8, driven by a central unit and based on average measurements, does not require synchronization of FAPs/FUEs with the BS, nor of FAPs among them, thus relaxing constraints also on the temporal alignment of the frames structures of different nodes.
- FAPs are able to track the Physical Resource Block (RB) structure within a frame (either taking into account a RB-level synchronization capability or a low bit rate MBS-FAP channel)
- It is assumed that UEs have the capability of estimating the channels with the most disturbing interfering nodes on every RBs.
The terminals must be informed about which interfering sources are in their neighboring in order to estimate the links with those nodes.
- The algorithms investigated in section 5, 6, 7 are able to optimize the power allocation (conventional waterfilling-based solutions) over different RB and obtain different bitrates per RB, like independent messages were transmitted per RB

In the current LTE standard it is assumed that the MCS have to be the same for the messages transmitted over the RBs, and additionally, the same power is allocated to the employed RB.

- The resource allocation algorithm investigated in section 6 optimizes the RB assignment over those users associated to the same source without imposing any constraint on how RBs are distributed. Moreover, the power constraint at the source is assumed in terms of sum-power constrain.
The LTE standard adopts localized FDMA which defines that the consecutive RBs are assigned to the same user and in a given subframe they are in multiples of 2,3 and 5 for low complexity DFT implementation.


### 3.3 Network architecture

- There is a protocol able to support the exchange of control-plane information between FAPs and FAPs-MBS. In [FREEDOM-D42] some modifications are proposed for accommodating the messages to/from FAPs.
So far, the existing X2 protocol in LTE standard only allows the communication between MBS as it was described in [FREEDOM-D21]
- We assume that there is not a critical delay in the exchange of messages and there are not lost messages, except in section 7.4 we deal with the case that this situation might happen.
- For the centralized resource optimization it is assumed that a network entity exists (the central processor unit or CPU) that collects all the required information from FAPs (and possibly from MBS) and performs the optimization.


## 4 SIMULATION METHODOLOGY

The techniques investigated in this work have considered the models and parameters presented in Table 2

| Description | Key Parameters adopted |
| :---: | :---: |
| Traffic models | Since the objective of the present work is to investigate novel techniques to manage the generated interference in a coordinated way, it is assumed a full-buffer traffic model for a given set of users to be served that are defined a priori |
| Interference models | When FAPs do not know which RB are employed by the MBS when all are employing the same band, we model such activity by two-state homogeneous Discrete Time Markov Chain (DTMC). The activity on different RB is assumed to be statistically independent. <br> No inter-macrocell interference is considered in the simulations below, but this is not limiting the validity of the techniques developed |
| MAC overhead | The final results should be scaled by factors 0.745 (SISO), 0.717 (MIMO $2 \times 2$ ) or 0.677 (MIMO $4 \times 4$ ) |
| FAP access modes | The proposed techniques do not distinguish among different FAP access modes. |
| Backhaul quality model | $\{5,10,20,30,40$, Ideal ( $\infty$ ) \} Mbits/s |
| Link adaptation and packet error modelling | The techniques analyzed in this work are evaluated in term of achievable rate (mutual information) and abstract metrics. Hence, no generation of encoded packets is required. |
| System performance indicators | - Transmit Power <br> - Spectral efficiency <br> - Outage rate <br> - Number of iterations for convergence |

Table 2. System models and parameters adopted
The terminal deployments that are commonly found in the business/residential scenarios have been defined in sec. 5.2.2 of [FREEDOM-D21]. Taking into account those deployments we have considered four reference scenarios to evaluate the proposed techniques which are depicted in Figure 3, Figure 4 and Figure 5. In this regard, reference scenario I, Figure 3, assumes one FAP area [FREEDOM-D21] consisting of two buildings separated by a street, where the MBS might be active. This scenario has been considered in sections 5.1, 7.3 and 8. Additionally, reference scenario II, depicted in Figure 4, has been considered for evaluating the algorithm investigated in section 7.2 and 7.4, from which we have performed repeated trials over random FAP topologies. Finally, reference scenario III and IV, sketched in Figure 5, are considered for evaluating the techniques in a more realistic scenario taking into account all parameters defined in [FREEDOM-D21]. Both reference scenarios describe the terminal deployment for one sector of a given cell for a residential and corporate scenario. Reference scenario III has been considered in sections 5.2 and 6.2 , while reference scenario IV is tackled in section 9 .


Figure 3. Reference scenario I. (Considered in sections 5.1, 7.3 and 8)


Figure 4. Reference scenario II. (Considered in sections 7.2 and 7.4)


Figure 5. Left- Reference scenario III (Residential). Right- Reference scenario IV (Corporate)

The common scenario assumes one $120^{\circ}$ sector of a cell, featuring a MBS and a number of randomly deployed FAP, enough for the comparison pursued in this deliverable. The key-parameters have been extracted from [FREEDOM-D21].

## 5 DECENTRALIZED POWER MINIMIZATION

One of the major goals in the design and deployment of femtocell networks is the reduction of the overall transmit power, with respect to the macro counterpart, while satisfying some prescribed QoS constraint. In massive femtocell deployments, it is of interest to devise decentralized strategies able to minimize the transmit power by ensuring a desired, application and user-dependent, information rate.

A possible approach for distributed self-organizing operation is considering femtocells as selfish agents competing for the resources available in the common spectrum band. This approach comes, however, at the expenses of injecting undue interference to the whole system, lack of fairness and efficiency loss. For such a reason, we adopt an alternate method based on the exchanged of limited information among FAPs in the form of interference prices, that represent the interference cost at each receiver. This avoids performance degradation and yet gracefully scales under a massive deployment.

We focus therefore on the minimization of the total transmitted power subject to minimum user rate constraints, assuming that the different FAPs may exchange information at the control plane (pricing). Section 5.1 considers the SISO case, while section 5.2 considers multiple antennas at both transmitter and receivers, providing a close-form for the transmit covariance matrices which depends on the pricing values exchanged. Finally, section 5.3 presents the conclusions for the decentralized power minimization approach.

### 5.1 SISO case and pricing mechanisms

### 5.1.1 Preliminaries

In the case of Gaussian parallel interference channels, a game theoretic approach to the minimum power problem has been proposed in [Pang08], formulating the problem as a pure competitive game for which the generalized Nash Equilibrium (NE) point can be found through totally decentralized algorithms. A NE, however, just because of its purely competitive nature, could be Pareto inefficient. It is then of interest to check if there are strategies to modify the minimum power game in order to make its equilibrium point more efficient and improve as much as possible its performance.

To reach this goal, in this section we introduce the min-power game with pricing mechanisms where the players are the FAPs which compete against each-other by choosing the optimal power allocation subject to a rate constraint for each FAP. The introduction of pricing mechanisms implies a modification of the formulation of each FAP's strategy by incorporating a cost quantifying the "damage" that each FAP's action can induce on the other players (FAPs) strategies. In this way we incentivize each FAP to achieve more socially efficient NE points by requiring a local exchange of a few data among FAPs through the backhaul wired link. As a by-product of the proposed procedure, the aggregated interference generated towards to the other FAPs and MBSs (or MUEs) is consequently reduced. In this case, it is assumed that the receivers have access, through a spectrum sensing operation, to the current channel occupation state of the macro users.

The features of the presented technique are those described in the following, and summarized in Table 3

| Technique | Objective <br> function | Constraints | Price Exchange | Spectrum sensing |
| :---: | :---: | :---: | :---: | :---: |
| Minimum power <br> optimization with <br> rate constraint and <br> exact interference <br> knowledge | Transmit <br> power | Per user <br> information <br> rate | Deterministic: prices <br> are correctly <br> exchanged between <br> FAPs | Each user needs to sense the <br> whole set of subchannels for <br> each realization of the <br> channel use, i.e. for each <br> slot. |

Table 3. Minimum power algorithm features

### 5.1.2 Problem formulation

As first step we formulate the min-power game. We then modify it by introducing pricing mechanism. Denoting by $R_{q}^{0}$ the rate required by FAP $q, \mathbf{p}_{q}=\left[p_{1}^{q}, \ldots, p_{N}^{q}\right]^{T}$ the power vector of $q$-th FAP, with $p_{k}^{q}$ the power allocated by the $q$-th FAP on the $k$-th subchannel, the set of feasible strategies of player $q$ is

$$
\begin{equation*}
\mathcal{P}_{q}\left(\mathbf{p}_{-q}\right)=\left\{\mathbf{p}_{q} \in \mathbb{R}^{N}: R_{q}\left(\mathbf{p}_{q}, \mathbf{p}_{-q}\right) \geq R_{q}^{0}, 0 \leq p_{k}^{q} \leq p_{q}^{\max }(k), k=1, \ldots, N\right\} \tag{1}
\end{equation*}
$$

The utility of each player is the transmit power, i.e. $u_{q}\left(\mathbf{p}_{q}\right)=\sum_{k=1}^{N} p_{k}^{q}$. Hence, the min-power game is

$$
\begin{equation*}
\mathcal{G}=\left\{\Omega,\left\{\mathcal{P}_{q}\left(\mathbf{p}_{-q}\right)\right\}_{q \in \Omega},\left\{u_{q}\left(\mathbf{p}_{q}\right)\right\}_{q \in \Omega}\right\}, \tag{2}
\end{equation*}
$$

with $\Omega$ the set of players (FAPs). Each player chooses the strategy that solves the following constrained problem

$$
\left(P_{2}\right) \begin{gather*}
\min _{\mathbf{p}_{q}}  \tag{3}\\
\text { s.t. }
\end{gather*} R_{q}(\mathbf{p}) \geq R_{q}^{0} ; 0 \leq p_{k}^{q} \leq p_{q}^{\max }(k), k=1, \ldots, N .
$$

Each user, given the others' strategies optimizes on its own power, in order to find the minimum power allocation vector that assures a rate value at least equal to $R_{q}^{0}$. It is worth to point out that the feasible set of every player now depends on the strategies chosen by the other players. In other words, while the max-rate game has coupled utility functions and uncoupled constraints, the min-power game has uncoupled utilities and coupled constraints. This is clear because, in the max-rate game, the (power) constraint of FAP $q$ does not depend on the constraints of the other FAPs while, in this case, increasing or decreasing power of $q$-th FAP translates into in an increasing, or decreasing, of other FAPs' information rate, which is the constraint of the optimization problem. This makes the problem of finding an NE for the min-power game harder to solve. In this case, the possible equilibrium points of the game are called Generalized NE (GNE), to point out the coupled nature of the constraints. Anyway the GNE's of game $\mathcal{G}$ may be Pareto-inefficient, because of its purely competitive nature. Hence, it is worth asking whether it is possible to modify game $\mathcal{G}$ in order to improve its performance. This case is different from the max rate game because, even if unknowingly, every player of game $\mathcal{G}$ is already pursuing a social utility goal. In fact, game $\mathcal{G}$ is a generalized exact potential game [MondererShapely96]. We recall that a game with utility function $u_{q}\left(\mathbf{p}_{q}, \mathbf{p}_{-q}\right)$ is an exact potential game if there exists a function $U\left(\mathbf{p}_{q}, \mathbf{p}_{-q}\right)$, called the potential, such that for all $q \in \Omega$

$$
\begin{equation*}
u_{q}\left(\mathbf{x}_{q}, \mathbf{p}_{-q}\right)-u_{q}\left(\mathbf{y}_{q}, \mathbf{p}_{-q}\right)=U\left(\mathbf{x}_{q}, \mathbf{p}_{-q}\right)-U\left(\mathbf{y}_{q}, \mathbf{p}_{-q}\right), \forall \mathbf{x}_{q}, \mathbf{y}_{q} \in \mathcal{P}_{q}\left(\mathbf{p}_{-q}\right) \tag{4}
\end{equation*}
$$

It is easy to check that the potential of game $\mathcal{G}$ is simply the sum of all the powers: $U(\mathbf{p})=\sum_{q=1}^{Q} \sum_{k=1}^{N} p_{k}^{q}$.
More specifically, since the constraints of $\mathcal{G}$ are coupled, $\mathcal{G}$ is a generalized potential game (GPG) [Facchinei2010]. Hence, since in this case each player is already pursuing a social goal (minimization of the total radiated power), we may wonder whether it is still possible to improve the performance of
game $\mathcal{G}$ by incorporating some pricing mechanism, similarly to what done for max-rate game in [Shi08]. To this end, we reformulate game $\mathcal{G}$, as follows

$$
\begin{array}{rc}
\min _{\mathbf{p}} & \sum_{q=1}^{Q} u_{q}\left(\mathbf{p}_{q}\right) \\
\text { s.t. } & R_{q}(\mathbf{p}) \geq R_{q}^{0}, \quad \forall q \in \Omega  \tag{5}\\
0 \leq p_{k}^{q} \leq p_{q}^{\max }(k), k=1, \ldots, N, \forall q \in \Omega
\end{array}
$$

with $\mathbf{p}=\left[\mathbf{p}_{1}^{T}, \ldots, \mathbf{p}_{Q}^{T}\right]^{T}$ the power allocation vector of all $Q$ FAPs and $u_{q}\left(\mathbf{p}_{q}\right)$ defined as above. In principle, the solution of this problem requires the existence of a central station that has all the necessary information. Nevertheless, a limited exchange of information among nearby FAPs is sufficient to implement a decentralized solution of $(P)$, which requires only local coordination among nearby FAPs. For the optimization problem $(P)$ to be meaningful, it is necessary to check first that the feasible set is nonempty. In [Barbarossal1b] we give sufficient conditions guaranteeing that the feasible set of $(P)$ is nonempty and compact so that the problem admits at least a solution point. It can be proved that any local optimum of $(P$ ) is a regular point (a feasible point is said to be regular if the equality constraints gradients and the active inequality gradients are linearly independent [Bertsekas95]) then it must satisfy the Karush-Kuhn Tucker (KKT) necessary conditions [Bertsekas95]. In particular, the Lagrangian associated to problem (5) is

$$
\begin{equation*}
\mathcal{L}(\mathbf{p}, \lambda, \mu, v)=\sum_{q=1}^{Q} \sum_{k=1}^{N} p_{k}^{q}-\sum_{q=1}^{Q} \lambda_{q}\left(R_{q}(\mathbf{p})-R_{q}^{0}\right)-\sum_{q=1}^{Q} \sum_{k=1}^{N} \mu_{k}^{q} p_{k}^{q}+\sum_{q=1}^{Q} \sum_{k=1}^{N} v_{k}^{q}\left(p_{k}^{q}-p_{q}^{\max }(k)\right) \tag{6}
\end{equation*}
$$

and the KKT conditions are ( $\mathbf{a} \perp \mathbf{b}$ means that the vectors $\mathbf{a}$ and $\mathbf{b}$ are orthogonal):

$$
\begin{align*}
\frac{\partial \mathcal{L}(\mathbf{p}, \lambda, \mu, \nu)}{\partial p_{k}^{q}}= & -\lambda_{q} \frac{\partial R_{q}}{\partial p_{k}^{q}}-\sum_{r \neq q} \lambda_{r} \frac{\partial R_{r}}{\partial p_{k}^{q}}-\mu_{k}^{q}+v_{k}^{q}=0 \\
& 0 \leq \mu_{k}^{q} \perp p_{k}^{q} \geq 0  \tag{7}\\
& 0 \leq v_{k}^{q} \perp p_{q}^{\max }(k)-p_{k}^{q} \geq 0 \\
& 0 \leq \lambda_{q} \perp R_{q}(\mathbf{p})-R_{q}^{0} \geq 0
\end{align*}
$$

with

$$
\begin{equation*}
\frac{\partial R_{q}}{\partial p_{k}^{q}}=\frac{\left|H_{k}^{q q}\right|^{2}}{\sigma_{q k}^{2}+I_{k}^{q}+\left|H_{k}^{q q}\right|^{2} p_{k}^{q}} ; \quad \frac{\partial R_{r}}{\partial p_{k}^{q}}=-\frac{\left|H_{k}^{r r}\right|^{2}\left|H_{k}^{q r}\right|^{2} p_{k}^{r}}{\left(\sigma_{r k}^{2}+I_{k}^{r}\right)\left(\sigma_{r k}^{2}+I_{k}^{r}+\left|H_{k}^{r r}\right|^{2} p_{k}^{r}\right)} \mathbb{I}\left(r \in \mathcal{N}_{q}\right), \tag{8}
\end{equation*}
$$

where

- $\left|H_{k}^{x y}\right|^{2}$ is the channel transfer coefficient over the $k$-th subchannel between the $x$-th transmitter and the $y$-th receiver;
- $\quad I_{k}^{j}=\sum_{i \in \mathcal{N}_{j}}\left|H_{k}^{i j}\right|^{2} p_{k}^{i}$ is the interference that the $j$-th FAP receives from all its neighbours $\mathcal{N}_{j}$ on the $k$-th subchannel;
- $\mathbb{I}(X)$ is a logical function equal to one, if X is true, or zero, otherwise;

Proceeding as in [Huang06] it is useful to introduce the price coefficient for user $r$ on the $k$-th subchannel:

$$
\begin{equation*}
\pi_{k}^{r}:=-\frac{\partial R_{r}(\mathbf{p})}{\partial I_{k}^{r}} \tag{9}
\end{equation*}
$$

which is proportional to the marginal decrease of user $r$ rate because of an increase of the $q$-th transmit power, as $\frac{\partial R_{r}(\mathbf{p})}{\partial p_{k}^{q}}=-\pi_{k}^{r} \frac{\partial I_{k}^{r}}{\partial p_{k}^{q}}$. If the prices are assumed to be constant with respect to $p_{k}^{q}$, solving problem (5) with respect to $\mathbf{p}_{q}$ is equivalent to solving the following local problem (in general, the assumption of $\pi_{k}^{r}$ to be constant with respect to $p_{k}^{q}$ is only an approximation. Nevertheless, the resulting algorithm provides significant performance improvement with respect to the purely competitive game).

$$
\begin{align*}
\left(P_{2}\right) & \min _{\mathbf{p}_{q}} \quad \sum_{k=1}^{N}\left(1+\sum_{r \in \mathcal{N}_{q}} \lambda_{r} \pi_{k}^{r}\left|H_{k}^{q r}\right|^{2}\right) p_{k}^{q}  \tag{10}\\
\text { s.t. } & R_{q}(\mathbf{p}) \geq R_{q}^{0}, 0 \leq p_{k}^{q} \leq p_{q}^{\max }(k), \quad k=1, \ldots, N
\end{align*}
$$

where the local Lagrangian coefficient $\lambda_{q}$ must satisfy the equality $\lambda_{q}\left(R_{q}(\mathbf{p})-R_{q}^{0}\right)=0$, with $R_{q}(\mathbf{p}) \geq R_{q}^{0}$. This problem can be solved locally, by FAP $q$, provided that all its neighbours send the coefficients $\lambda_{r} \pi_{k}^{r}$. The local solution, given the powers used by all other FAPs, is

$$
\begin{equation*}
p_{k}^{q}=\left[\frac{\lambda_{q}}{1+b_{k}^{q}}-\frac{I_{k}^{q}+\sigma_{q k}^{2}}{\left|H_{k}^{q q}\right|^{2}}\right]_{0}^{p_{q}^{\max }(k)} \tag{11}
\end{equation*}
$$

where $[x]_{a}^{b}$ denotes the projection of x into the interval $[\mathrm{a}, \mathrm{b}]$, in our case $\left[0, p_{q}^{\max }(k)\right]$, and $b_{k}^{q}=\sum_{r \in \mathcal{N}_{q}} \lambda_{r} \pi_{k}^{r}\left|H_{k}^{q r}\right|^{2}$ (parameter that takes into account the pricing coefficients). From (11) we see that the multiplier $\lambda_{q}$ associated to the rate constraint, must be strictly greater than zero, otherwise the powers $p_{k}^{q}$ will all be equal to zero and this would contradict the inequality $R_{q}(\mathbf{p}) \geq R_{q}^{0}$. Hence, $R_{q}(\mathbf{p})=R_{q}^{0}$ and, as a consequence, $\lambda_{q}$ can be found as the coefficient that guarantees $R_{q}(\mathbf{p})=R_{q}^{0}$. Some power coefficients $p_{k}^{q}$ can be null. Let us denote by $\mathcal{D}_{q}$ the set of subcarriers where user $q$ allocates a non-null power $p_{k}^{q}<p_{q}^{\max }(k)$ and by $\mathcal{D}_{q}^{\prime}$ the set of subcarriers where user $q$ allocates a power $p_{k}^{q}=p_{q}^{\max }(k)$. After a few algebraic manipulations, we can express $\lambda_{q}$ in closed form as

$$
\begin{equation*}
\lambda_{q}=e^{\frac{1}{\left|D_{q}\right|}\left[R_{q}^{0}-\sum_{k \in D_{q}} \log \left(\frac{\mid H_{q q}^{\left.q\right|^{2}}}{\left(I_{k}^{q}+\sigma_{q k}^{2}\right)\left(1+b_{k}^{q}\right)}\right)-\sum_{k \in D_{q q}^{\prime}} \log \left(1+p_{q}^{\max }(k) \frac{\left.| | \frac{q q}{q q^{2}}\right|^{2}}{I_{k}^{q^{2}}+\sigma_{q k}^{2}}\right)\right]} \tag{12}
\end{equation*}
$$

In summary, the proposed decentralized, min-power allocation strategy for the FAPs is described by the algorithm inTable 4:

```
Algorithm: Minimum power optimization with pricing mechanisms
S.0: Choose any feasible power allocation \(\mathbf{p}^{0}=\left(\mathbf{p}_{1}^{0}, \ldots, \mathbf{p}_{Q}^{0}\right)\) and set \(\mathrm{n}=0\);
S.1: If \(\mathbf{p}_{q}(\mathrm{n})\) satisfies a suitable termination criterion then STOP,
otherwise;
S.2: Set \(\mathrm{n}=\mathrm{n}+1\) and for \(\mathrm{q}=1, \ldots, \mathrm{Q}\) compute \(p_{k}^{q}(\mathrm{n})\) from (11), using (12);
S.3: Compute \(\lambda_{q}\) and \(\pi_{k}^{q}\) and broadcast \(\lambda_{q} \pi_{k}^{q}\) to the neighbors with index
\(i \in \mathcal{N}_{q}\);
S.4: Set \(\mathbf{p}(\mathrm{n})=\left(\mathbf{p}_{1}(\mathrm{n}), \ldots, \mathbf{p}_{Q}(\mathrm{n})\right)\) and go to S.1.
```

Table 4. Minimum power algorithm with pricing.
In this context, we did not assume any MBS's activity model and we suppose that the receivers have access, through a spectrum sensing operation, to the current channel occupation state of the macro users. Furthermore we assume that each FAP broadcasts price coefficients (step S.3) to its neighbours, at the start of each subframe (or for every channel use), with no errors. This means that the IP-based backhaul link is good enough to deliver packets at control data plane, that are received with no errors or delays due to packet retransmissions

### 5.1.3 Numerical results

We will now show some numerical results for the game in (10). The next two figures (Figure 6 and Figure 7) report the sum of the radiated power expressed in dBm , and the sum of the rate expressed in bit per OFDM symbol [bps], respectively, obtained with and without pricing, vs. the number of iterations. As we can see the algorithm that takes into account the prices exchange between nearby FAPs, performs better than the one that does not consider any pricing mechanism. In fact we see in Figure 7 that in both cases rate constraint is respected, but the pricing mechanism translates into reducing the effective transmitted power for each FAP with respect to the case of purely competitive access, as is shown in Figure 6. Result refers to a scenario with ten interfering FAPs (see Figure 3). The number of subcarriers is $N=600(10 \mathrm{MHz}$ LTE-A bandwidth $)$, while $R_{q}^{0}$ is set to 0.9 Kbit per OFDM symbol, while the maximum transmit power for each FAP is set to 20 dBm .


Figure 6. Sum of the radiated power vs. iteration index for the minimum power game with and without pricing.


Figure 7. Sum rate vs. iteration index for the minimum power game with and without pricing.

### 5.2 MIMO case

### 5.2.1 Preliminaries

This section addresses the problem of coordinated radio resource allocation in the downlink for a set of femtocells operating on the same band under OFDMA access, when both transmitters and receivers are equipped with multiple antennas (MIMO).
Considering a set of multiple mutually interfering transmitters, an algorithm to compute numerically the set of beamforming vectors that maximize the sum rate utility was proposed in [Shi09a]. The beamforming design in an interference scenario was also addressed in [Jorswieck08] and [Zhang10] for the single carrier case. In [Jorswieck08], the authors designed the beamformers as linear combinations of the zero-forcing (ZF) and maximum-ratio transmission (MRT) beamformers. In [Zhang10], the authors obtained a closed-form solution for the beamformer that maximizes the transmission rate at each transmitter subject to a certain set of interference power constraints at the receivers. Less work has been done, however, for the MIMO case. In [Shi09b] distributed algorithms were proposed, considering different transmitter's beams and treating them separately.

This section extends to MIMO settings the pricing mechanisms proposed in FREEDOM for distributed resource allocation in femtocell networks. As in previous section, we focus on the minimization of the total transmission power subject to minimum rate constraints for the users attached to the femtocells in the set, assuming that the different FAPs may exchange information (pricing) at the control plane. The algorithm features are the ones depicted in Table 3. Unlike the works mentioned in previous paragraph, we provide a closed-form solution for the precoding matrices that depends on the pricing values exchanged. Both the MISO and SISO cases can be obtained as particular solutions of the proposed approach.

### 5.2.2 Description

We consider a set of $N_{F}$ femtocells being served by FAP operating on a set of $N_{C}$ common subchannels. Non-orthogonal access is assumed among the different FAPs. Therefore, each sub-channel can be used by several (even all) femtocells. The $f$-th FAP has a set $U_{f}$ of UEs receiving under

OFDMA access in the downlink. We use $C_{u}$ to represent the set of sub-channels allocated to the $u$-th user and $u(f, c)$ to denote the user attached to the $f$-th FAP receiving signal at the $c$-th sub-channel.

We assume that the FAP and the UE are provided with $M$ and $N$ antennas respectively (to simplify the notation we assume $M$ to be equal for all the FAPs and $N$ to be equal for all the UEs, although the generalization is simple). The problem consists on designing the optimum transmit covariance matrix for the $f$-th FAP at the $c$-th sub-channel, $\mathbf{S}_{f}^{c}$ under the goal of minimizing the total transmission power, and provided that every user achieves a given target rate. To further simplify notation let us define the following noise and interference matrix at the $c$-th sub-channel for the user $u(f, c)$ :

$$
\begin{equation*}
\mathbf{R}_{f}^{c}=\mathbf{I}_{N}+\sum_{f^{\prime} \neq f}^{N_{F}} \mathbf{H}_{f f^{c}}^{c} \mathbf{S}_{f f^{c}}^{c} \mathbf{H}_{f f^{\prime}}^{c H} \tag{13}
\end{equation*}
$$

where $\mathbf{I}_{N}$ is the $N \times N$ identity matrix and $\mathbf{H}_{f f^{\prime}}^{c}$ stands for the $N \times M$ MIMO channel matrix, including the path-loss and the random channel amplitude. The superindex corresponds to the sub-channel, the first subindex denotes the FAP serving the receiving user $u(f, c)$ and the second subindex corresponds to the FAP that is producing interference over the intended user. Therefore, $\mathbf{H}_{f f}^{c}$, represents the channel between the $u(f, c)$ UE and the neighbour FAP $f^{\prime}$. Without loss of generality, we consider that the channels are normalized by the noise power at each receiver.

Under the minimum power consumption criterion, the problem to solve is the following:

$$
\begin{align*}
& \min _{\left\{\mathbf{s}_{f}^{c}\right\}} \sum_{f=1}^{N_{F}} \sum_{c=1}^{N_{c}} \operatorname{trace}\left(\mathbf{S}_{f}^{c}\right)  \tag{14}\\
& \text { s.t. } \sum_{c \in C_{u}} \log _{2}\left|\mathbf{I}_{M}+\mathbf{S}_{f}^{c} \mathbf{H}_{f f}^{c H}\left(\mathbf{R}_{f}^{c}\right)^{-1} \mathbf{H}_{f f}^{c}\right| \geq R_{u}, \text { for } u \in U_{f}, f=1, \ldots, N_{F},  \tag{15}\\
& \mathbf{S}_{f}^{c} \succeq 0, \text { for } c=1, \ldots, N_{C} \text { and } f=1, \ldots, N_{F} . \tag{16}
\end{align*}
$$

where $\succeq 0$ stands for positive semidefinite.
This is a non-convex problem and, therefore, it may have multiple local optima. We focus on finding a local optimum. Furthermore, for the sake of scalability, we also focus on finding a local optimum in a distributed way. To that end, we proceed as in [Huang06] and [Shi08]: we take into account that any local optimum must fulfil the Karush-Kuhn-Tucker (KKT) conditions [Boyd04] of the global problem and we separate the KKT conditions in subsets of equations, one for each femtocell. This way, we obtain a set of $N_{F}$ distributed problems.
The solution of each problem, however, depends on other femtocells variables, which results in a highly complex coupled problem for which a closed-form solution has not been found. In order to derive a solution to the previous problem, we consider that, at a given time, the allocation of resources of a single FAP is optimized while considering that powers and precoders for the rest of FAP are fixed. Using this approach, the precoding matrix for the $f$-th femtocell at the $c$-th sub-channel is given by the solution of the following convex problem [Munoz11b], that can be solved separately for each transmitter (either centralized or decentralized):

$$
\begin{align*}
& \min _{\mathbf{s}_{f}} \sum_{c=1}^{N_{C}} \operatorname{trace}\left(\left(\mathbf{I}_{M}+\sum_{f \neq f} \mu_{u\left(f f^{\prime},\right)} \mathbf{H}_{f^{\prime} f}^{c}{ }^{H} \Pi_{f}^{c} \mathbf{H}_{f^{\prime \prime} f}^{c}\right) \mathbf{S}_{f}^{c}\right)  \tag{17}\\
& \text { s.t. } \sum_{c \in C_{u}} \log _{2}\left|\mathbf{I}_{M}+\mathbf{S}_{f}^{c} \mathbf{H}_{f f}^{c H}\left(\mathbf{R}_{f}^{c}\right)^{-1} \mathbf{H}_{f f}^{c}\right| \geq R_{u}, \text { for } u \in U_{f} \tag{18}
\end{align*}
$$

$$
\begin{equation*}
\mathbf{S}_{f}^{c} \succeq 0, \text { for } i=1, \ldots, N . \tag{19}
\end{equation*}
$$

Notice that the objective function in (17) is different from the objective function in the original problem in eq. (14). Instead of minimizing the trace of the transmit covariance matrix, $\mathbf{S}_{f}^{c}$, the goal now is to minimize the trace of the product between the so called pricing matrix, defined as follows:

$$
\begin{equation*}
\mathbf{B}_{f}^{c}=\mathbf{I}_{M}+\sum_{f^{\prime} \neq f} \mu_{u(f, f) c} \mathbf{H}_{f^{\prime} f}^{c}{ }^{H} \mathbf{\Pi}_{f}^{c} \mathbf{H}_{f^{\prime} f}^{c} \tag{20}
\end{equation*}
$$

and the transmit covariance matrix, $\mathbf{S}_{f}^{c}$.
The pricing matrix defined in eq. (20) is a full rank matrix with probability one. In the MISO case the rank of this matrix is one, and in the SISO case is a scalar value. The role of the pricing matrix is to reflect the compensation to pay for the interference generated to the users connected to other FAPs. It measures the degradation on other FAPs performance due to the interference generated by the $f$-th FAP. It is computed, for each sub-channel, from three factors:

- The Lagrange multiplier, $\mu_{u\left(f^{\prime}, c\right)}$, associated to rate constraint of the user allocated by the neighbor FAP $f^{\prime}$ at the $c$-th sub-channel. It is greater as the rate constraint is tighter, meaning that a small change in the constraint will affect greatly to the optimal value of the cost function.
- The cross channel $\mathbf{H}_{f^{\prime} f}^{c}$.
- The sensitivity to the noise and interference, in the $c$-th sub-channel, for the user allocated by the neighbor FAP $f^{\prime}$ at this sub-channel. In the general MIMO case, this sensitivity is given by matrix $\Pi_{f}^{c}$ computed as

$$
\begin{equation*}
\boldsymbol{\Pi}_{f}^{c}=-\left(\frac{\partial}{\partial \mathbf{R}_{f}^{c}} \log _{2}\left|\mathbf{I}_{N}+\mathbf{H}_{f f}^{c} \mathbf{S}_{f}^{c} \mathbf{H}_{f f}^{c H}\left(\mathbf{R}_{f}^{c}\right)^{-1}\right|\right)^{T}=\log _{2} e\left(\mathbf{R}_{f}^{c}\right)^{-1}\left(\mathbf{I}_{N}+\mathbf{X}_{f}^{c}\left(\mathbf{R}_{f}^{c}\right)^{-1}\right)^{-1} \mathbf{X}_{f}^{c}\left(\mathbf{R}_{f}^{c}\right)^{-1}, \tag{21}
\end{equation*}
$$

which for the SISO and MISO cases boils down to a scalar value, $\pi_{f^{\prime}}^{c}$, which ranges between 0 and 1.

The global procedure for computing the pricing matrix at each transmitter is as follows:
Step 1. Each UE detects the global identifier (GI) and channel state information (CSI) of surrounding FAPs and sends reports to the FAP that this UE is connected to. The UE also reports his interference sensitivity.
Step 2. The FAP that the UE is connected to, signals a DL-pricing message to potentially interfering FAPs. This message consists of the GI of the sending FAP and the identifier of the potentially interfering FAP and a field/s indicating the price for each RB.
Step 3. The FAP receiving a DL pricing message or messages from neighbor FAPs computes a global price per sub-channel according to eq. (20), and solves problem (17)-(19). Intuitively, for the SISO case, if the global price is high, the FAP will avoid allocating large power in this subchannel. In the MISO (and MIMO) case, the FAP will avoid the sub-channels and directions with a greater price.

Based on the pricing information sent by the neighbors FAPs, and its own users measurements (channel and noise plus interference correlation matrix), each transmitter computes and equivalent channel per sub-channel that depends on the pricing matrix, $\mathbf{B}_{f}^{c}$, the user channel, $\mathbf{H}_{f f}^{c}$, and the measured noise and interference, $\mathbf{R}_{f}^{c}$. By the eigenvalue decomposition theorem, this equivalent channel can be decomposed as follows:

$$
\begin{equation*}
\left(\mathbf{B}_{f}^{c}\right)^{-1 / 2} \mathbf{H}_{f f}^{c H}\left(\mathbf{R}_{f}^{c}\right)^{-1} \mathbf{H}_{f f}^{c}\left(\mathbf{B}_{f}^{c}\right)^{-1 / 2}=\mathbf{U}_{f}^{c} \boldsymbol{\Lambda}_{f}^{c} \mathbf{U}_{f}^{c H} \tag{22}
\end{equation*}
$$

where $\mathbf{U}_{f}^{c}$ is a unitary matrix, containing the eigenvectors of the equivalent channel, and $\boldsymbol{\Lambda}_{f}^{c}$ is a diagonal matrix containing the corresponding eigenvalues, $\lambda_{f}^{c, m}$, for $m=1, \ldots, M$. The optimal transmission strategy, when fixing the interference sensitivity matrices and the precoders for the rest of the transmitters, is given by the following transmit covariance matrix [Munoz11b]:

$$
\begin{equation*}
\mathbf{S}_{f}^{c}=\left(\mathbf{B}_{f}^{c}\right)^{-1 / 2} \mathbf{U}_{f}^{c} \tilde{\mathbf{S}}_{f}^{c} \mathbf{U}_{f}^{c H}\left(\mathbf{B}_{f}^{c}\right)^{-1 / 2} \tag{23}
\end{equation*}
$$

where $\tilde{\mathbf{S}}_{f}^{c}$ is a diagonal matrix, whose entries depend on the eigenvalues of the equivalent channel, $\lambda_{f}^{c, m}$, and the Lagrange multiplier, $\mu_{u(f, c)}$, which is adjusted to satisfy the user rate constraint defined in (18):

$$
\begin{equation*}
\tilde{s}_{f}^{c, m}=\left(\mu_{u(f, c)}-\frac{1}{\lambda_{f}^{c, m}}\right)^{+} \tag{24}
\end{equation*}
$$

By setting the pricing matrix equal to the identity matrix, the solution obtained boils down to the waterfilling solution [Pang08]. Intuitively, if the users connected to other FAPs are far away or the interference is not affecting them (low cross channel eigenvalues or low interference sensitivity), the pricing matrix in (20) tends to be the identity matrix, and the solution given by eq. (22)-(24) tends to be the waterfilling solution.

The solution given by eq. (22)-(24) can be particularized for the SISO and MISO case. In the MISO case, the interference observed by the UE allocated at the $c$-th sub-channel by the $f$-th FAP is a scalar value given by

$$
\begin{equation*}
I_{f}^{c}=\sum_{f^{\prime} \neq f}^{N_{F}} \mathbf{h}_{f f^{\prime}}^{c} \mathbf{S}_{f}^{c} \mathbf{h}_{f f^{\prime}}^{c H}, \tag{25}
\end{equation*}
$$

with $\mathbf{h}_{f f}^{c}$, the $1 \times M$ channel vector. The equivalent channel in (22) has only one eigenvalue different from 0 . Therefore, the optimal transmit covariance matrix is a rank one matrix which is given by

$$
\begin{equation*}
\mathbf{S}_{f}^{c}=\mathbf{w}_{f}^{c} \mathbf{w}_{f}^{c H} \tag{26}
\end{equation*}
$$

with

$$
\begin{equation*}
\mathbf{w}_{f}^{c}=\sqrt{\left(\mu_{f}^{c}-\frac{1+I_{f}^{c}}{\left\|\left(\mathbf{B}_{f}^{c}\right)^{-1 / 2} \mathbf{h}_{f f}^{c H}\right\|^{2}}\right)^{+} \frac{\left(\mathbf{B}_{f}^{c}\right)^{-1} \mathbf{h}_{f f}^{c H}}{\left\|\left(\mathbf{B}_{f}^{c}\right)^{-1 / 2} \mathbf{h}_{f f}^{c H}\right\|}} . \tag{27}
\end{equation*}
$$

Different from the work in [Jorswieck08], where the authors designed the beamformers as linear combinations of ZF and MRT beamformers, here the solution is obtained from MRT beamformer multiplied by a prewhitening matrix which steers nulls in the directions of the potentially interfered neighbors. The deep of the nulls, however, depend on the impact of the transmission on the neighbors' performance, measured through the pricing matrix.

In the SISO case, the equivalent channel is scalar and the optimal transmission power particularizes to the following expression that takes into account the own channel, normalized by the noise, the measured interference, and also the impact on the neighbors' performance.

$$
\begin{equation*}
p_{f}^{c}=\frac{\mu_{u(f, c)}}{\left(1+\left.\sum_{f^{\prime} \neq f} \mu_{u\left(f^{\prime}, c\right)} \pi_{\left.f^{\prime}\right|^{c}\left|h_{f^{\prime} f}^{c}\right|^{c}}\right|^{2}\right.}-\frac{1}{\frac{\left|h_{f f}^{c}\right|^{2}}{1+I_{f}^{c}}} \tag{28}
\end{equation*}
$$

which is the same solution provided in section 5.1.

As in the MISO case, the sensitivity factors, $\pi_{f^{\prime}}^{c}$, are now scalar values. Notice that higher power is allocated to the RBs with better signal to noise plus interference ratio and with lower price $\left(1+\sum_{f^{\prime} \neq f} \mu_{u\left(f^{\prime}, c\right)} \pi_{f^{\prime} \mid}^{c}\left|h_{f^{\prime} f}^{c}\right|^{2}\right)$.

Table 5 summarizes the control information required to support the computation of the optimal precoding matrices according the proposed algorithm, i.e., pricing and CSI information. The table also includes a proposal for the interface to exchange this information. In practical systems such as LTE, there is no interface defined between Macro Base Stations (MBs) and FAPs or between FAPs. However, a proposal [DOCOMO] exists to extend the X2 interface, which is the interface defined for connection of MBSs.

| Control information to be exchanged | Direction | Possible Interface |
| :---: | :---: | :---: |
| Each UE collects and reports CSI regarding <br> the link with the serving FAP | From UE to serving FAP | LTE Uu |
| Each UE collects and reports CSI regarding <br> the link with interfering FAPs | From UE to serving FAP | LTE Uu |
| Each UE estimates and reports how the <br> interference affects the quality of the <br> communication (interference sensitivity) | From UE to serving FAP | LTE Uu |
| From the information received from UEs, each <br> FAP computes and sends DL-cost information <br> to each member of a list of potentially <br> interfering FAPs | From each FAP to potentially <br> interfering FAPs | X2 extension |

## Table 5 Control information to be exchanged

### 5.2.3 Numerical results

For the simulation results, we consider 4 FAPs each one with 2 UEs attached. We consider 24 subchannels that are equally split among the 2 users connected to each FAP. The channels are generated in time with 4 taps of Rayleigh modulus of equal gain. The mean signal to noise ratio between a UE and the serving FAP is 20 dB , when the transmission power is 0 dBm . This means that when the FAP is transmitting with 20 mW (a typical value for FAPs) the signal to noise ratio will be 33 dB .

Likewise, for a transmission power of 0 dBm , the mean interference to noise ratio between a UE and an interfering FAP is 5 dB . In this set up, we evaluate two different schemes:

1. The FAPs do not exchange interference prices. They update their own strategies simultaneously, according to the single-transmitter waterfilling solution, but using the interference reported by its own users in the previous iteration (simultaneous iterative waterfilling, IWFA [Pang08]). Notice that this solution is obtained from the one provided in previous section, eq. (22)-(24), but using the identity matrix instead of the pricing matrix.
2. Every iteration, the FAPs update their pricing matrices, based on the information received from other FAPs, and compute the solution given by eq. (22)-(24).
Figure 8 shows the time evolution of the overall power necessary to guarantee the target rate per user. Despite we are focusing in the minimization of the total transmission power, it is clear that each FAP will have its own transmit power constraint. In order to compare the individual power required to fulfill the target rate constraints, Figure 9 depicts the cumulative density function (CDF) of the transmission power per FAP, obtained from 100 independent channel realizations. When exceeding the individual power constraint, for instance if the required power is greater than 13 dBm , a possible solution would be to reduce the target rates for the users connected to this FAP. This approach, however, is outside the scope of this paper.


Figure 8. Time evolution of the total transmitted power

For both, Figure 8 and Figure 9, the SISO 1x1, MISO 2x1, and MIMO $2 \times 2$ cases are compared with and without pricing. Two target rates are considered: $\mathrm{R}=1.5 \mathrm{bps} / \mathrm{Hz}$ (solid lines for SISO, MISO and MIMO) and $\mathrm{R}=2 \mathrm{bps} / \mathrm{Hz}$ (dashed lines for MIMO). As the problem is not convex, we cannot claim that the solution achieved is the optimal solution for the original problem, (17)-(16). However, it is a better solution than the IWFA, in terms of both total and individual power. Indeed, the use of the pricing matrix takes into account the potential degradation on the other users' performance, reducing the interference on other users, and so helping to reduce to transmission power required for the other transmitters to fulfill the rate constraints. Also, lower power is required if 2 transmit antennas are used. When no pricing information is exchanged, this reduction comes from the fact that each FAP transmits on the eigenmodes of the intended receiver channel, therefore a more efficient use of the transmitted power is achieved. When pricing information is exchanged, the solution would be the same if the potential degradation on the neighbors' performance is low.

On the contrary, if the potential degradation is high, the precoders will minimize the transmission power in the direction of those neighbors which have strongest constraints. Nevertheless, the improvement due to the exchange of pricing information is greater in the SISO case, as in this situation the spatial dimension cannot be used to minimize the interference. For the MIMO case, when the target rate is $1.5 \mathrm{bps} / \mathrm{Hz}$, the advantage using pricing seems to be marginal. However, if the user requirements increases (target rate equal to $2 \mathrm{bps} / \mathrm{Hz}$, dashed lines) the transmitter needs to increase the power. In such a case, it is clear that the exchange of pricing information helps to reduce the degradation on other users, and consequently the required individual power of other transmitters, resulting in a global reduction of the overall transmitted power within the set of FAPs.


Figure 9. CDF of the transmission power per transmitter.

### 5.3 Conclusions

In section 5 we have proposed alternative game-theoretic techniques that exploit the backhaul link among femto-access points to set up local coordination games which provide performance improvement with respect to purely competitive games. More specifically, we have proposed the minimum power game with pricing by showing the advantage of using pricing mechanism to reduce the radiated power still maintaining the same service quality (desired information rate).

For MIMO systems, a closed-form solution for the transmit covariance matrix has been obtained, assuming that the different transmitters may exchange information at the control plane. MISO and SISO cases can be obtained as particular cases of the general MIMO solution.

When the impact of the interference caused to the users connected to other FAPs is low, the performance for the proposed strategy approaches the performance of the waterfilling solution for the minimum power problem [Pang08]. However, the reduction in power increases when the requirements of the transmitters in terms of user target rates increases.

## 6 DECENTRALIZED WEIGHTED SUM RATE MAXIMIZATION

This section addresses decentralized algorithms to maximize the weighted sum rate (WSR) of a wireless cellular system based on femtocells following two research lines. First, section 6.1 analyses a single-input single-output (SISO) configuration where two sources (FAPs or MBS) are serving two user equipments (UEs) each one. Transmission is carried out using OFDM over multiple resource blocks (RBs). In that configuration, sources have a limitation in terms of maximum transmitted power. Two transmission strategies are looked into:
a) sources are equipped with complex transmitters based on dirty-paper coding (o receivers have multi-user decoding capabilities)
b) sources have simple transmitters and the access of the UEs served by each source is orthogonal, i.e. OFDMA

Section 6.2 investigates the multiple-input multiple-output (MIMO) case where each source transmits to its served UEs under OFDMA. Moreover, the analysed problem deals with a rate constraint per source that accounts for a maximum served rate that comes up due to the capacity-limited backhaul that inter-connects sources with the core network.

### 6.1 SISO case: complex vs. simple transmitters

The present work looks into efficient and distributed solutions to deal with the generated interference in a femtocell (FAP) scenario. Since such scenario resembles to conventional ad-hoc network or cognitive radio scenarios, the techniques developed there can be applied for our purposes. Usually, those techniques are based on game-theoretic formulation with non-cooperative players. For example, optimal linear precoding strategies are derived in [Scutari08b], where the competitive players participates in a game where each ones tries to maximize its transmitted rate taking into account a given spectral mask. Likewise, the authors provide sufficient conditions for ensuring the uniqueness of the Nash equilibrium. Those distributed techniques are based on the iterative waterfilling concept, [Yu02].

However, in the previous strategy all transmitting terminals (or players) act independently, reacting to the received interference power. In order to perform a joint resource allocation, the authors in [Cendrillon05] propose an iterative algorithm for maximizing the weighted sum-rate in digital subscriber lines, mitigating the crosstalk. In this regard, a modified waterfilling algorithm is proposed in [Yu07], where a parameter can be exchanged between transmitters in order to optimize their resources. When that technique is applied in OFDM systems, it tends to be the optimal for a large number of data carriers, [Yu06]. The parameter exchanged between transmitters also can be seen as some kind of pricing if the whole problem is formulated as a non-cooperative game, [Schmidt09].

Here we explore the case where each source is able to serve multiple users simultaneously, a reasonable scenario under a high FAP deployment. Sources are connected through a backhaul link which we will exploit for combating the generated interference. In order to dwell more on such configuration we limit the number of sources to two and the number of destinations to two per source, see Figure 10. Notice that is such scenario, each source is transmitting to their associated users (i.e. $\mathrm{T}_{1}$ to $\mathrm{U}_{1,1}, \mathrm{U}_{2,1}$ and $T_{2}$ to $\mathrm{U}_{1,2}, \mathrm{U}_{2,2}$ ), while at the same time each source generates interference to those users associated to other sources (i.e. $\mathrm{T}_{1}$ to $\mathrm{U}_{1,2}, \mathrm{U}_{2,2}$ ). Nevertheless, we will also indicate how the algorithm scales with the number of sources/destinations and show results including multiple FAPs. Notice that scenario sketched in Figure 10 assumes two sources that can be either both FAPs or FAP and MBS. Consequently, the users might be FUEs or/and MUEs.


Figure 10. Broadcast Interference scenario. Sources $T_{1}, T_{2}$ might represent FAPs and or FAP and MBS, while $\mathbf{U}_{1,1}, \mathbf{U}_{2,1}, \mathbf{U}_{1,2}, \mathbf{U}_{2,2}$ are FUEs or MUEs, depending on the associated source.

The user access for those destinations associated to a given source (i.e. $\mathrm{U}_{1,1}, \mathrm{U}_{2,1}$ with $\mathrm{T}_{1}$ ) can be based on orthogonal or non-orthogonal mode. In the first case, single-user transmitters and receivers are employed. However, in the non-orthogonal user-access it is assumed that either transmitter (dirty paper coding) or receivers (successive interference canceller) become complex. For both types of user access, the signal received from other neighbouring sources is considered as additive noise (i.e. signal from $T_{2}$ at $\mathrm{U}_{1,1}, \mathrm{U}_{2,1}$ ).

If we had just a single source in the scenario depicted in Figure 10 and the user-access was orthogonal, the resource optimization (power allocation and RB assignment) is a non-convex problem. Nevertheless, the best known solution is obtained by an algorithm with polynomial complexity proposed in [Seong06] where each RB is allocated to just one of the users. In contrast, under nonorthogonal user access with complex transmitter/receivers, the radio resource allocation can be written as a convex problem thanks to the BC-MAC duality [Vishwanath04].

In the present section we show that the joint resource optimization of the scenario shown in Figure 10 can be addressed by a decentralized optimization where each source optimizes the resources allocated to its associated users. The joint resource management can be seen as a non-cooperative game with pricing values exchanged through the backhaul link. Each player, consisting of a source with its associated destinations (i.e. $\mathrm{T}_{1-}\left\{\mathrm{U}_{1,1}, \mathrm{U}_{2,1}\right\}, \mathrm{T}_{2}-\left\{\mathrm{U}_{1,2}, \mathrm{U}_{2,2}\right\}$ ), receives from and generates to other sources parameters to be considered for designing its individual radio resources algorithm (i.e. how each player distributes power and assigns RB among its intended users).

The novelty aspects addressed in this work are:

- Multiple users per source.
- Study of orthogonal and non-orthogonal user access., with simple or complex transmitters, respectively


### 6.1.1 System model

In the scenario depicted in Figure 10 we assume that all terminals are equipped with a single antenna (SISO, single-input single output channels). The channels of the different links are frequency selective and OFDM is the selected transmission technology. Transmissions are done over $N$ parallel resource blocks (RB), where each RB consists of a group of 15 carriers. We define the signal-to-noise ratio (SNR) per RB in the different links as,

$$
\begin{cases}\rho_{i, j}^{(l)}=\left|h_{i, j}^{(l)}\right|^{2} \frac{P}{N_{0}} & \text { (link to from the j-th source to its i-th user) }  \tag{29}\\ \rho_{i, j, t}^{(L)}=\left|h_{i, j, t}^{(t)}\right|^{2} \frac{P}{N_{0}} & \text { (interfering link from the t-th source to the i-th user of the j-th source) }\end{cases}
$$

where $l$ defines the selected RB, $i, j$ indicates that SNR is that of the link from the $j$-th source to its $i$-th destination, $h_{i, j}^{(l)}$ is the channel coefficient of the link at the $l$-th RB, $P$ denotes the maximum transmitted power and $N_{0}$ stands for the AWGN (Additive White Gaussian Noise) power. The interfering links are denoted with sub-indexes $i, j, t$, describing the link from the $t$-th source to the $i$-th destination attached to the $j$-th source. In such a case the channel coefficient is represented by $h_{i, j, t}^{(l)}$. It should be remarked that the instantaneous SNR per RB is dependent on the fraction of allocated power. Hence, equation (29) denotes the maximum SNR for a given RB in case all power was allocated to that RB.

All sources share the same bandwidth, that is, all destinations are interfered by the active sources in their neighbourhood. Likewise, since each source is serving multiple users, the destinations might observe an additional kind of interference coming from its associated source, due to the messages intended to the other destinations and served by the same source. Notice that such interference depends on the employed user access mode. For example, under the orthogonal user access only one user is active per RB and that interference does not come up. We define the set of RB allocated to the different destinations under orthogonal user access as,

$$
\begin{equation*}
S_{u, k}=\left\{i \mid \quad U_{u, k} \text { active }\right\} \quad \text { (orthogonal user access) } \tag{30}
\end{equation*}
$$

The bitrate attained by the $u, k$-th user ( $u$-th user attached to the $k$-th source) will be denoted by,

$$
R_{u, k}=\frac{1}{N} \sum_{i \in S_{u, k}} \log _{2}\left(1+\gamma_{u, k}^{(i)} \frac{\left|h_{u, k}^{(i)}\right|^{2} P}{N_{0}\left(1+I_{u, k}^{(i)}\right)}\right)=\frac{1}{N} \sum_{i \in S_{u, k}} \log _{2}\left(1+\frac{\gamma_{u, k}^{(i)} \rho_{u, k}^{(i)}}{1+I_{u, k}^{(i)}}\right)
$$

where $\gamma_{u, k}^{(i)}$ denotes the fraction of power allocated to $u, k$-th user at the $i$ th RB and $I_{u, k}^{(i)}$ stands for the received interference normalized by the noise power. The generated interference at the UE can be defined by,
$\left\{\begin{array}{l}I_{1,1}^{(i)}=\left(\gamma_{1,2}^{(i)}+\gamma_{2,2}^{(i)}\right) \rho_{1,1,2}^{(i)}\left\{\begin{array}{lll}\gamma_{1,2}^{(i)}=0 & \text { if } & i \notin S_{1,2} \\ \gamma_{2,2}^{(i)}=0 & \text { if } & i \notin S_{2,2}\end{array}, \quad I_{2,1}^{(i)}=\left(\gamma_{1,2}^{(i)}+\gamma_{2,2}^{(i)}\right) \rho_{2,1,2}^{(i)}\left\{\begin{array}{lll}\gamma_{1,2}^{(i)}=0 & \text { if } & i \notin S_{1,2} \\ \gamma_{2,2}^{(i)}=0 & \text { if } & i \notin S_{2,2}\end{array}\right.\right. \\ I_{1,2}^{(i)}=\left(\gamma_{1,1}^{(i)}+\gamma_{2,1}^{(i)}\right) \rho_{1,2,1}^{(i)}\left\{\begin{array}{llll}\gamma_{1,1}^{(i)}=0 & \text { if } & i \notin S_{1,1} \\ \gamma_{2,1}^{(i)}=0 & \text { if } & i \notin S_{2,1}\end{array}, \quad I_{2,2}^{(i)}=\left(\gamma_{1,1}^{(i)}+\gamma_{2,1}^{(i)}\right) \rho_{2,2,1}^{(i)}\left\{\begin{array}{llll}\gamma_{1,1}^{(i)}=0 & \text { if } & i \notin S_{1,1} \\ \gamma_{2,1}^{(i)}=0 & \text { if } & i \notin S_{2,1}\end{array}\right.\right.\end{array}\right.$

On the other hand, when there is a non-orthogonal user access we assume that sources operate under superposition coding as a transmission strategy to its associated users (i.e. $\mathrm{T}_{1}\left\{\left\{\mathrm{U}_{1,1}, \mathrm{U}_{2,1}\right\}\right.$ in Figure 10). Since there is a SISO channel configuration, the channel is degraded and the associated destinations can be classified as strong or weak over different RB. The strong user is able to decode its message and the one intended to the other destination. Consequently, receivers must have successive interference canceller (SIC) capability. Notice that we are assuming that independent messages are transmitted over each RB by the same transmitter. In contrast, the weak user only is able to decode its intended message assuming the other message as additive noise. Such classification of the users per RB is summarized by the following ensembles,

$$
\begin{align*}
& S_{k}=\{i \mid  \tag{32}\\
& \bar{S}_{k}=\{i \mid \\
& \left.U_{1, k} \text { strong, } U_{2, k} \text { weak }\right\} \\
& \left.U_{1, k} \text { weak, } U_{2, k} \text { strong }\right\}
\end{align*} \quad k=\{1,2\} \quad \text { (non-orthogonal user-access) }
$$

Notice that the previous sets satisfy $S_{k} \cap \bar{S}_{k}=\varnothing, \quad S_{k} \cup \bar{S}_{k}=\{1, \ldots, N\}$. In contrast to (30), all users are active in all RBs. It turns out that a given user can be served as a weak user in some RBs, while is considered as the strong user in others. In this regard, we define the following achievable rates for a given user,

$$
\begin{equation*}
R_{u, k}=R_{u, k}^{w}+R_{u, k}^{s} \tag{33}
\end{equation*}
$$

where super-indexes $w, s$ denote weak and strong, respectively, and $R_{u, k}^{w}, R_{u, k}^{s}$ stands for the total rate attained in the data carriers where that user is identified as weak or strong user. For example for the ( 1,1 )-th user ,

$$
R_{1,1}=\frac{1}{N} \sum_{i \in S_{1}} \log _{2}\left(1+\gamma_{1,1}^{(i)} \frac{\left|h_{1,1}^{(i)}\right|^{2} P}{N_{0}\left(1+\gamma_{2,1}^{(i)} \frac{\left|h_{1,1}^{(i)}\right|^{2} P}{N_{0}}+I_{1,1}^{(i)}\right)}\right)+\frac{1}{N} \sum_{i \in S_{1}} \log _{2}\left(1+\gamma_{1,1}^{(i)} \frac{\left|h_{1,1}^{(i)}\right|^{2} P}{N_{0}\left(1+I_{1,1}^{(i)}\right)}\right)
$$

where $\gamma_{2,1}^{(i)}$ is the fraction of power allocated to second user associated the first source at the $i$ th RB. The generated interference at the different UE under non-orthogonal access is described by,

$$
\begin{cases}I_{1,1}^{(i)}=\left(\gamma_{1,2}^{(i)}+\gamma_{2,2}^{(i)}\right) \rho_{1,1,2}^{(i)}, & I_{2.1}^{(i)}=\left(\gamma_{1,2}^{(i)}+\gamma_{2,2}^{(i)}\right) \rho_{2,1,2}^{(i)}  \tag{34}\\ I_{1,2}^{(i)}=\left(\gamma_{1,1}^{(i)}+\gamma_{2,1}^{(i)}\right) \rho_{1,2,1}^{(i)}, & I_{2,2}^{(i)}=\left(\gamma_{1,1}^{(i)}+\gamma_{2,1}^{(i)}\right) \rho_{2,2,1}^{(i)}\end{cases}
$$

It must be remarked the difference between equations (31) and (34). In the first one, only the power allocated to one of users served by an interfering source contributes to the interference. This is due to the orthogonal user access per source. However equation (34) considers the power allocated to all the users served by the interfering source, because the two users can employ a given RB.

### 6.1.2 Decentralized Resource Allocation

The joint resource allocation for the scenario shown in Figure 10 is in general a non-convex problem which difficulties finding the optimal solution, thus imposing drawbacks to find efficient distributed solutions. Let us evoke that objective pursued by design a distributed resource allocation algorithm is to efficiently deal with multiple sources and possibly multiple destinations, as it is found in a scenario with a dense deployment of FAPs. In this regard we can exploit the Lagrangian formulation of the problem because it offers the appealing property of decomposing the whole dual function of problem into multiple subproblems that can be solved in parallel with the proper exchange of data.

The computation of the dual function at a given Lagrange multiplier requires minimizing the Lagrangian over the optimization variables. The general problem is solved iteratively where at each iteration all sources solves its local subproblems and exchange the proper parameters to other sources. Such formulation becomes optimal if the problem is strongly dual ([Boyd04, Sec. 5.2.3]) then there is no duality gap and solving the dual problem yields the solution of the primal one. Otherwise, the solution provided by the Lagrangian becomes a lower bound of the optimal solution if the optimization is done with the maximization function (it becomes an upper bound with the minimization function).

We will be able to provide a decentralized solution to our resource allocation problem under the assumption that the received interference at each iteration by the users is constant and do not depend on the transmitted power of the other sources. The obtained distributed RRM could be seen as a noncooperative game where each player (sources with its associated destinations) maximizes its own payoff function which takes into account pricing values. The pricing value has information about the changes of the generated interference by a given source over the non-associated destinations of its neighbourhood, [Schmidt09].

In the ensuing sections, A and B, we present how the joint resource allocation problem for the scenario introduced in Figure 10 is tackled in a decentralized way when the user-access is orthogonal with simple transmitters or non-orthogonal with complex transmitters.

## A. Orthogonal user-access with simple transmitters

The joint radio resource allocation is obtained as the solution of the following optimization problem,

$$
\begin{align*}
\left(P_{0}^{\text {Orth }}\right) & \underset{\substack{\left.\left\{\gamma_{u, k}\right\}\left\{\left\{, R_{u, k},\right\}, S_{u, k}\right\} \\
k\{1,2\}, u=1,2\right\}}}{\operatorname{minimize}}-\sum_{u=1}^{2} \sum_{k=1}^{2} \mu_{u, k} R_{u, k} \\
& \begin{cases}\left(v_{u, k}\right): \quad f_{u, k}=R_{u, k}-\frac{1}{N} \sum_{i \in S_{k}} \log _{2}\left(1+\frac{\gamma_{u, k}^{(i)} \rho_{u, k}^{(i)}}{1+I_{u, k}^{(i)}}\right) \leq 0, \quad k=\{1,2\}, u=\{1,2\} \\
\left(\lambda_{k}\right): & h_{k}=\sum_{i=1}^{N} \sum_{u=1}^{2} \gamma_{u, k}^{(i)}-1 \leq 0, \\
\quad R_{u, k} \geq 0\end{cases} \tag{35}
\end{align*}
$$

where $R_{\mathrm{u}, \mathrm{k}}$, denotes the rate allocated to the messages intended to $u, k$-th destination, $\boldsymbol{\gamma}_{u, k}$ stands for the vector of allocated powers, $S_{u, k}$ defines the set of carriers assigned to that user, (30). Lagrange multiplier $v_{u, k}$, is tied to the rate constraint (function $f_{u, k}$ ), while Lagrange multiplier $\lambda_{k}$ is associated to the power constraint per source (function $h_{k}$ ). The observed interference by each user is defined in (31).

The Lagrange function of problem $\left(P_{0}^{\text {Orth }}\right)$ in (35) is given by,

$$
\begin{equation*}
L_{\left(\left\{\gamma_{u, k}\right\}, \lambda_{1}, \lambda_{2},\left\{R_{u, k},\left\{\left\{_{u, k}\right\},\left\{\left\{_{u, k}\right\}\right)\right.\right.\right.}=-\sum_{u=1}^{2} \sum_{k=1}^{2} \mu_{u, k} R_{u, k}+\sum_{u=1}^{2} \sum_{k=1}^{2} v_{u, k} f_{u, k}+\lambda_{1} h_{1}+\lambda_{2} h_{2} \tag{36}
\end{equation*}
$$

Any local optimum with respect the power allocation variable $\gamma_{j}^{(i)}$ should satisfy

$$
\begin{equation*}
\frac{\partial}{\partial \gamma_{u, k}^{(i)}} L_{\left(\left\{v_{u, k}\right\}, \lambda_{1}, z_{2},\left\{R_{u, k}\right\},\left\{\left\{_{u, k}\right\},\left\{s_{u, k}\right\}\right)\right.}=v_{u, k} \frac{\partial}{\partial \gamma_{u, k}^{(i)}} f_{u, k}+\sum_{t=1, t, t k}^{2} \sum_{p=1}^{2} v_{p, t} \frac{\partial}{\partial \gamma_{u, k}^{(i)}} f_{p, t}+\lambda_{k}=0 \tag{37}
\end{equation*}
$$

with

$$
\begin{align*}
& \frac{\partial}{\partial \gamma_{u, k}^{(i)}} f_{u, k}=-\frac{\frac{\left|h_{u, k}^{(i)}\right|^{2} P}{N_{0}}}{1+I_{u, k}^{(i)}+\gamma_{u, k}^{(i)} \frac{\left|h_{u, k}^{(i)}\right|^{2} P}{N_{0}}} \\
& \frac{\partial}{\partial \gamma_{u, k}^{(i)}} f_{p, t}=\frac{\partial}{\partial I_{p, t}^{(i)}} f_{p, t} \times \frac{\partial}{\partial \gamma_{u, k}^{(i)}} I_{p, t}^{(i)}=\frac{\frac{\gamma_{p, t}^{(i)}\left|h_{p, t}^{(i)}\right|^{2} P}{N_{0}}}{1+I_{p, t}^{(i)}+\frac{\gamma_{p, t}^{(i)}\left|h_{p, t}^{(i)}\right|^{2} P}{N_{0}}} \frac{\rho_{u, k, t}^{(i)}}{1+I_{p, t}^{(i)}} \tag{38}
\end{align*}
$$

We define

$$
\left\{\begin{array}{l}
\pi_{p, t}^{(i)}=\frac{\gamma_{p, t}^{(i)} \frac{\left|h_{p, t}^{(i)}\right|^{2} P}{N_{0}}}{1+I_{p, t}^{(i)}+\gamma_{p, t}^{(i)} \frac{\left|h_{p, t}^{(i)}\right|^{2}}{N_{0}}} \frac{1}{1+I_{p, t}^{(i)}}  \tag{39}\\
\Omega_{p, t, k}^{(i)}=v_{p, t} \pi_{p, t}^{(i)} \rho_{p, t, k}^{(i)}
\end{array}\right.
$$

Assuming that $\Omega_{p, t, k}^{(i)}$ are fixed and a given power profiles [Shi08] the KKT conditions of problem $\left(P_{0}^{\text {Orth }}\right)$ are the same as those if each transmitter carries out the following optimization problem

$$
\begin{align*}
& \left(P_{k}^{\text {Orth }}\right) \underset{\substack{\left\{\gamma_{u, k},\left\{,\left\{R_{u}, k\right\},\left\{\left\{_{u, k}\right\} \\
u=\{1,2\}\right.\right.\right.}}{\operatorname{minimiz}}-\sum_{u=1}^{2}\left(\mu_{u, k} R_{u, k}-\sum_{i=1}^{N}\left(\sum_{t=1, t \neq k}^{2} \sum_{p=1}^{2} \Omega_{p, t, k}^{(i)}\right) \gamma_{u, k}^{(i)}\right)  \tag{40}\\
& \begin{cases}\left(v_{u, k}\right): & f_{u, k}=R_{u, k}-\frac{1}{N} \sum_{i \in S_{k}} \log _{2}\left(1+\frac{\gamma_{u, k}^{(i)} \rho_{u, k}^{(i)}}{1+I_{u, k}^{(i)}}\right) \leq 0, \quad u=\{1,2\} \\
\left(\lambda_{k}\right): & h_{k}=\sum_{i=1}^{N} \sum_{u=1}^{2} \gamma_{u, k}^{(i)}-1 \leq 0, \\
R_{u, k} \geq 0\end{cases}
\end{align*}
$$

Therefore, $\left(P_{0}^{\text {Orth }}\right)$ is solved in a decentralized way by applying the iterative procedure described in Table 6.
0. Initialize $R_{p, k}, \quad I_{p, k}^{(i)}=0, \quad \boldsymbol{\Omega}_{p, t, k}^{(i)}=0 \quad p, k, t \in\{1,2\}$

1. Every source solves problem $\mathrm{P}_{k}^{\text {Orth }}$ using eq.(40), $k \in\{1,2\}$
2. The $t$-th source generates $\boldsymbol{\Omega}_{p, t, k}^{(i)}$ using eq.(39), $p \in\{1,2\}, k, t \in\{1,2\}$
3. Price exchange through the backhaul link
4. Update transmit precoders $\gamma_{p, k}^{(i)}$ with the solution of step 1
5. The $k$-th source collects all prices generated by neigboring sources $k \in\{1,2\}$
6. Every source collects the noise plus interference, $I_{p, k}^{(i)}$
7. Go to 1

Table 6. Decentralized algorithm to solve problem $P_{0}^{\text {orth }}$ (35).

Although we have reduced the complexity of problem $\left(P_{0}^{\text {Orth }}\right)$ thanks to the decentralized solution, the problem to be solved at each transmitter $\left(P_{k}^{\text {Orth }}\right)$ is still non-convex due to the RB assignment. In this regard, we will consider the algorithm proposed in [Seong06] that provides the best known solution with a polynomial complexity.

Let us introduce the Lagrangian of problem $\left(P_{k}^{\text {Orth }}\right)$ as

$$
\begin{equation*}
G_{\left(\left\{v_{u, k}\right\}, \lambda_{k}\left\{R_{u, k}\right\},\left\{\gamma_{u, k}\right\},\left\{s_{u, k}\right\}\right)}=-\left(\sum_{u=1}^{2}\left(\mu_{u, k} R_{u, k}-\sum_{i=1}^{N} w_{k}^{(i)} \gamma_{u, k}^{(i)}\right)\right)+\sum_{u=1}^{2} v_{u, k} f_{u, k}+\lambda_{k} h_{k} \tag{41}
\end{equation*}
$$

where $w_{u, k}^{(i)}$ collects all the parameters generated by the neighboring sources and intended to the $k$-th source,

$$
\begin{equation*}
w_{k}^{(i)}=\sum_{t=1, t \neq k}^{2} \sum_{p=1}^{2} \Omega_{p, t, k}^{(i)} \tag{42}
\end{equation*}
$$

One of the conditions to by satisfied with respect the variable $R_{\mathrm{u}, \mathrm{k}}$ is,

$$
\begin{equation*}
\frac{\partial}{R_{u, k}} G_{\left(\left\{s_{u, k}\right\}, \lambda_{k}\left\{R_{u, k}\right\},\left\{v_{u, k}\right\},\left\{S_{u, k}\right\}\right)}=-\mu_{u, k}+v_{u, k}=0 \quad \Rightarrow \quad v_{u, k}=\mu_{u, k} \tag{43}
\end{equation*}
$$

Hence, the Lagrangian given in (41) turns out into,

$$
\begin{equation*}
G_{\left(\lambda_{k},\left\{\gamma_{u, k}\right\},\left\{S_{u, k}\right\}\right)}=\sum_{i=1}^{N} w_{p, t}^{(i)} \gamma_{u, k}^{(i)}-\sum_{u=1}^{2} \frac{\mu_{u, k}}{N} \sum_{i \in S_{u, k}} \log _{2}\left(1+\frac{\gamma_{u, k}^{(i)} \rho_{u, k}^{(i)}}{1+I_{u, k}^{(i)}}\right)+\lambda_{k} \sum_{i=1}^{N} \sum_{u=1}^{2} \gamma_{u, k}^{(i)}-\lambda_{k} \tag{44}
\end{equation*}
$$

Moreover, an additional condition to be satisfied by the power allocation is given by,

$$
\begin{equation*}
\frac{\partial}{\partial \gamma_{u, k}^{(i)}} G_{\left(\lambda_{k},\left\{\left\{_{u, k}\right\},\left\{\xi_{u, k}\right\}\right)\right.}=w_{t}^{(i)}-\frac{\mu_{u, k}}{N \ln 2} \frac{\frac{\rho_{u, k}^{(i)}}{1+I_{u, k}^{(i)}}}{1+\frac{\gamma_{u, k}^{(i)} \rho_{u, k}^{(i)}}{1+I_{u, k}^{(i)}}}+\lambda_{k}=0 \tag{45}
\end{equation*}
$$

Where we can get a semiclosed-form expression for the power allocation,

$$
\begin{equation*}
\gamma_{u, k}^{(i)}=\left[\frac{\mu_{u, k}}{N \ln 2} \frac{1}{\lambda_{k}+w_{k}^{(i)}}-\frac{1+I_{u, k}^{(i)}}{\rho_{u, k}^{(i)}}\right]^{+} \tag{46}
\end{equation*}
$$

Since each RB only is assigned to one user, we follow the same approach used in [Seong06] and we perform such allocation in order to minimize the Lagrangian

$$
\begin{align*}
G_{\left(\lambda_{k},\left\{S_{u, k}\right\}\right)} & =\sum_{u=1}^{2} \sum_{i=1}^{N} w_{t}^{(i)} \gamma_{u, k}^{(i)}\left(\lambda_{k}\right)-\sum_{u=1}^{2} \frac{\mu_{u, k}}{N} \sum_{i \in S_{u, k}} \log _{2}\left(1+\frac{\gamma_{u, k}^{(i)}\left(\lambda_{k}\right) \rho_{u, k}^{(i)}}{1+I_{u, k}^{(i)}}\right)+\lambda_{k} \sum_{i=1}^{N} \sum_{u=1}^{2} \gamma_{u, k}^{(i)}\left(\lambda_{k}\right)-\lambda_{k}=  \tag{47}\\
& =-\lambda_{k}+\sum_{i=1}^{N} G_{(i)}\left(\lambda_{k}\right)
\end{align*}
$$

It turns out that the minimization is obtained by searching over all the served users by a given source at each carrier, and select that one that mimimizes $G_{(i)}\left(\lambda_{k}\right)$,

$$
\begin{equation*}
G_{(i)}\left(\lambda_{k}\right)=\sum_{u=1}^{2} w_{t}^{(i)} \gamma_{u, k}^{(i)}\left(\lambda_{k}\right)-\frac{\mu_{u, k}}{N} \log _{2}\left(1+\frac{\gamma_{u, k}^{(i)}\left(\lambda_{k}\right) \rho_{u, k}^{(i)}}{1+I_{u, k}^{(i)}}\right)+\lambda_{k} \gamma_{u, k}^{(i)}\left(\lambda_{k}\right) \tag{48}
\end{equation*}
$$

Finally, problem $\left(P_{k}^{\text {Orth }}\right)$ is efficiently solved using the algorithm depicted in Table 7, where the optimization of variable $\lambda_{k}$ is carried out based on the bisection method [Boy04].

1. Initialize: $\lambda^{\text {max }}, \lambda^{\text {min }}$
2. while $\left|\lambda^{\text {max }}-\lambda^{\text {min }}\right| \leq \varepsilon$ do
3. $\lambda_{k}=\frac{1}{2}\left(\lambda^{\text {max }}+\lambda^{\text {min }}\right)$
4. Initialize: $S_{u, k}=\left[1, \ldots, N_{R B}\right] u \in[1,2]$
5. Calculate $\gamma_{u, k}^{(i)}\left(\lambda_{k}\right)$ for $u \in[1,2], i \in\left[1 \ldots N_{R B}\right] \quad$ eq.(46)
6. Assign RB to FUEs in order to minimize $\mathrm{G}_{(i)} i \in\left[1 \ldots N_{R B}\right]$ eq.(48)
7. Re-define $S_{u, k}$
8. Set $\gamma_{u, k}^{(i)}\left(\lambda_{k}\right)=0$ if $i \notin S_{u, k}, u \in[1,2], i \in\left[1 \ldots N_{R B}\right]$
9. if $h_{k}\left(\lambda_{k}\right)<0, \lambda^{\max }=\lambda_{k}$, else $\lambda^{\text {min }}=\lambda_{k}$
10. end while

Table 7. Efficient search of $\boldsymbol{\lambda}_{\boldsymbol{k}}$ in problem $P_{k}^{\text {orth }}$ (35).

## B. Non-Orthogonal user-access with complex transmitters

The radio resource algorithm aims to maximize the weighted sum-rate of the system when certain RB can be simultaneously employed by the associated UE to a given source. This is possible by means of the superposition coding strategy (or dirty paper coding). Assuming a constant received interference at each RB, each source can define its associated users as strong UE o weak UE per RB as has been introduced in section 6.1.1 in equations (32)-(34).

The power allocation and the attained rates are obtained as a solution of the following optimization problem,

$$
\begin{align*}
& \left\{\begin{array}{l}
\left(v_{1, k}^{s}\right): f_{1, k}^{s}=R_{1, k}^{s}-\frac{1}{N} \sum_{i \in S_{k}} \log _{2}\left(1+\frac{\gamma_{1, k}^{(i)}}{1+I_{1, k}^{(i)}}\right) \leq 0 \\
\left(v_{2, k}^{w}\right): f_{2, k}^{w}=R_{2, k}^{w}-\frac{1}{N} \sum_{i \in S_{k}}^{(i)} \log _{2}\left(1+\frac{\gamma_{2, k}^{(i)} \rho_{2, k}^{(i)}}{1+\gamma_{1, k}^{(i)} \rho_{2, k}^{(i)}+I_{2, k}^{(i)}}\right) \leq 0 \\
\left(v_{1, k}^{w}\right): f_{1, k}^{w}=R_{1, k}^{w}-\frac{1}{N} \sum_{i \in S_{k}} \log _{2}\left(1+\frac{\gamma_{1, k}^{(i)} \rho_{1, k}^{(i)}}{1+\gamma_{2, k}^{(i)} \rho_{1, k}^{(i)}+I_{1, k}^{(i)}}\right) \leq 0, \quad k=\{1,2\}, u=\{1,2\} \\
\left(v_{2, k}^{s}\right): f_{2, k}^{s}=R_{2, k}^{s}-\frac{1}{N} \sum_{i \in S_{k}} \log _{2}\left(1+\frac{\gamma_{2, k}^{(i)} \rho_{2, k}^{(i)}}{1+I_{2, k}^{(i)}}\right) \leq 0 \\
\left(\lambda_{k}\right): h_{k}=\sum_{i=1}^{N} \sum_{u=1}^{2} \gamma_{u, k}^{(i)}-1 \leq 0, \\
R_{u, k}^{w} \geq 0, R_{u, k}^{s} \geq 0
\end{array}\right. \tag{49}
\end{align*}
$$

where $\gamma_{u, k}$ denotes the $N$-vector with the power distribution employed by the $u, k$-th user, $R_{u, k}^{s}, R_{u, k}^{w}$ stands for the attained rate of the $u, k$-th user in those RBs where it is the strong user (ensemble $S_{k}$ in
(32)) or the weak user (ensemble $\bar{S}_{k}$ in (32)). Received interference at the $i$-th RB, $I_{u, k}^{(i)}$, is introduced in (34). The problem considers a maximum transmitted power per source. Finally, weights $\mu_{u, k}$ are used to describe the priority given to $u, k$-th destination. Notice that the total noise plus interference observed at destination may have up to three components, the additive white Gaussian noise (already included in the SNR definition (29)), the message intended to the strong user when the given destination is a weak user (i.e. component $\gamma_{1, k}^{(i)} \rho_{2, k}^{(i)}$ present in the $R_{2, k}^{w}$ in (49)) and the transmission done by other sources (for example the term $I_{2, k}^{(i)}$ in the $R_{2, k}^{w}$ in (49))

The problem $\left(P_{0}^{\text {Non-Orth }}\right)$ is non-convex due to the rate constraints given by the Broadcast Channel when superposition coding has been considered. However, such drawback can be combated by applying the BC-MAC duality [Vishwanath04]. Basically, it transforms the problem and introduces a new variables $\lambda_{1}, \lambda_{2}, \lambda_{3}, \lambda_{4}$ connected with the actual power allocation variables $\left(\gamma_{1}, \gamma_{2}, \gamma_{3}, \gamma_{4}\right)$ that satisfy the following property,

$$
\begin{equation*}
\gamma_{1, k}^{(i)}+\gamma_{2, k}^{(i)}=\varphi_{1, k}^{(i)}+\varphi_{2, k}^{(i)} \tag{50}
\end{equation*}
$$

An illustrative example of how variables $\boldsymbol{\lambda}_{1}, \boldsymbol{\lambda}_{2}$ and $\boldsymbol{\gamma}_{1}, \boldsymbol{\gamma}_{2}$ are connected is

Notice that the received interference also can be written as

$$
\begin{cases}I_{1,1}^{(i)}=\left(\varphi_{1,2}^{(i)}+\varphi_{2,2}^{(i)}\right) \rho_{1,1,2}^{(i)}, & I_{2.1}^{(i)}=\left(\varphi_{1,2}^{(i)}+\varphi_{2,2}^{(i)}\right) \rho_{2,1,2}^{(i)}  \tag{52}\\ I_{1,2}^{(i)}=\left(\varphi_{1,1}^{(i)}+\varphi_{2,1}^{(i)}\right) \rho_{1,2,1}^{(i)} & I_{2,2}^{(i)}=\left(\varphi_{1,1}^{(i)}+\varphi_{2,1}^{(i)}\right) \rho_{2,2,1}^{(i)}\end{cases}
$$

Using the previous BC-MAC duality the problem presented in (49) is transformed into a convex one,

$$
\begin{align*}
& \left(\left(\tilde{V}_{1, k}^{s}\right): \tilde{f}_{1, k}^{s}=R_{1, k}^{s}+R_{2, k}^{w}-\frac{1}{N} \sum_{i \in S_{k}} \log _{2}\left(1+\frac{\varphi_{\varphi, k}^{(i)} \rho_{1, k}^{(i)}}{1+I_{1, k}^{(i)}}+\frac{\varphi_{2, k}^{(i)} \rho_{2, k}^{(i)}}{1+I_{2, k}^{(i)}}\right) \leq 0\right. \\
& \left(\tilde{v}_{2, k}^{w}\right): \tilde{f}_{2, k}^{w}=R_{2, k}^{w}-\frac{1}{N} \sum_{i \in S_{k}} \log _{2}\left(1+\frac{\varphi_{2, k}^{(i)} \rho_{2, k}^{(i)}}{1+I_{2, k}^{(i)}}\right) \leq 0 \\
& \left\{\begin{array}{l}
\left(\tilde{1}_{1, k}^{w}\right): \tilde{f}_{1, k}^{w}=R_{1, k}^{w}-\frac{1}{N} \sum_{i \in S_{k}} \log _{2}\left(1+\frac{\varphi_{1, k}^{(i)}}{1+I_{1, k}^{(i)}}\right) \leq 0 \\
\left(\tilde{v}_{2, k}^{s}\right): \tilde{f}_{2, k}^{s}=R_{1, k}^{w}+R_{2, k}^{s}-\frac{1}{N} \sum_{i \in \mathcal{S}_{k}} \log _{2}\left(1+\frac{\varphi_{1, k}^{(i)}}{1+I_{1, k}^{(i)}}+\frac{\varphi_{2, k}^{(i)}}{1+\rho_{2, k}^{(i)}}\right) \leq 0
\end{array}\right.  \tag{53}\\
& \left(\tilde{\lambda}_{k}\right): \tilde{h}_{k}=\sum_{i=1}^{N} \sum_{u=1}^{2} \varphi_{u, k}^{(i)}-1 \leq 0, \\
& R_{u, k}^{w} \geq 0, R_{u, k}^{s} \geq 0
\end{align*}
$$

The Lagrangian is given by,

$$
L_{\left(\left\{\tilde{u}_{u, k},\right\}, \tilde{\tilde{L}}_{1}, \tilde{h}_{1},\left\{R_{u, k}^{s}\right\},\left\{\left\{\begin{array}{l}
w, k,\}  \tag{54}\\
w
\end{array}\right\}\left\{\left\{_{u, k}\right\}\right)\right.\right.}=-\sum_{u=1}^{2} \sum_{k=1}^{2} \mu_{u, k}\left(R_{u, k}^{s}+R_{u, k}^{w}\right)+\sum_{u=1}^{2} \sum_{k=1}^{2} \tilde{V}_{u, k}^{s} \tilde{f}_{u, k}^{s}+\sum_{u=1}^{2} \sum_{k=1}^{2} \tilde{V}_{u, k}^{s} \tilde{f}_{u, k}^{w}+\sum_{k=1}^{2} \tilde{\lambda}_{k} \tilde{h}_{k}
$$

Any local optimum with respect the power allocation $\varphi_{u, k}^{(i)}$ should satisfy the following equation

$$
\begin{align*}
& \frac{\partial}{\partial \varphi_{u, k}^{(i)}} L_{\left(\left\{\tilde{i}_{u, k},\right\}, \tilde{u}_{1}, \tilde{\tilde{L}}_{2},\left\{R_{u, k}^{s},\left\{\left\{R_{u, k}^{u},\left\{,\left\{\varphi_{u, k}\right\}\right)\right.\right.\right.\right.}= \\
& \quad=\sum_{z=1}^{2} \tilde{v}_{z, k}^{s} \frac{\partial}{\partial \varphi_{u, k}^{(i)}} \tilde{f}_{z, k}^{s}+\tilde{v}_{u, k}^{w} \frac{\partial}{\partial \varphi_{u, k}^{(i)}} \tilde{f}_{u, k}^{w}+\sum_{t=1, t, k k}^{2} \sum_{p=1}^{2}\left(\tilde{v}_{p, t}^{s} \frac{\partial}{\left.\partial \varphi_{u, k}^{(i)} \tilde{f}_{p, t}^{s}+\tilde{v}_{p, t}^{w} \frac{\partial}{\partial \varphi_{u, k}^{(i)}} \tilde{f}_{p, t}^{w}\right)+\lambda_{k}=0}\right. \tag{55}
\end{align*}
$$

where derivatives of the functions associated to other sources different the $k$-th one are addressed as,

$$
\left\{\begin{array}{l}
\frac{\partial}{\partial \varphi_{u, k}^{(i)}} \tilde{f}_{p, t}^{s}=\frac{\partial}{\partial I_{1, t}^{(i)} \tilde{f}_{p, t}^{s} \times \frac{\partial}{\partial \varphi_{u, k}^{(i)}} I_{1, t}^{(i)}+\frac{\partial}{\partial I_{2, t}^{(i)}} \tilde{f}_{p, t}^{s} \times \frac{\partial}{\partial \varphi_{u, k}^{(i)}} I_{2, t}^{(i)}} \\
\frac{\partial}{\partial \varphi_{u, k}^{(i)}} \tilde{f}_{p, t}^{w}
\end{array}=\left\{\begin{array}{lll}
\frac{\partial}{\partial I_{1, k}^{(i)}} \tilde{f}_{p, t}^{w} \times \frac{\partial}{\partial \varphi_{u, k}^{(i)}} I_{1, t}^{(i)} & \text { if } & p=1 \\
\frac{\partial}{\partial I_{2, k}^{(i)}} \tilde{f}_{p, t}^{w} \times \frac{\partial}{\partial \varphi_{u, k}^{(i)}} I_{2, t}^{(i)} & \text { if } & p=2
\end{array}\right.\right.
$$

Following similar steps as in (39) we define,



Likewise, the interference values received at the $k$-th source and generated at the $t$-th source are defined by,

$$
\begin{align*}
& \tilde{\Omega}_{k, t}^{(i)}= \begin{cases}\tilde{v}_{1, t}^{s} \pi_{1, t, k}^{s(i)}+\tilde{v}_{2, t}^{w} \pi_{2, t, k}^{w(i)} \quad \text { if } i \in S_{t} \\
\tilde{v}_{1, t}^{w} \pi_{1, t, k}^{w(i)}+\tilde{v}_{2, t}^{s} \pi_{2, t, k}^{s(i)} \quad \text { if } i \in \bar{S}_{t}, \quad k \neq t, \quad k=t=\{1,2\} \\
\tilde{w}_{k}^{(i)}=\sum_{t, t \neq k} \tilde{\Omega}_{k, t}^{(i)}\end{cases} \tag{56}
\end{align*}
$$

Finally, the problem ( $\left.P_{0}^{\text {Non-Orth }}\right)$ introduced in (49) is addressed in a decentralized and iterative way using the algorithm depicted in Table 8 where each source optimizes the resources allocation by minimizing the following problem at each step,

$$
\left.\begin{array}{l}
\left(P_{k}^{\text {Non-Orth }}\right) \\
\underset{\substack{\left\{\varphi_{u, k}\right\},\left\{R_{u, k, k}^{s}\right\},\left\{R_{u, k}^{w}\right\} \\
, u=\{1,2\}}}{\operatorname{minize}}-\left(\sum_{u=1}^{2} \sum_{k=1}^{2} \mu_{u, k}\left(R_{u, k}^{s}+R_{u, k}^{w}\right)-\sum_{i=1}^{N} \tilde{w}_{k}^{(i)}\left(\varphi_{1, k}^{(i)}+\varphi_{2, k}^{(i)}\right)\right)  \tag{57}\\
\left(\tilde{v}_{1, k}^{s}\right): \tilde{f}_{1, k}^{s}=R_{1, k}^{s}+R_{2, k}^{w}-\frac{1}{N} \sum_{i \in S_{k}} \log _{2}\left(1+\frac{\varphi_{1, k}^{(i)} \rho_{1, k}^{(i)}}{1+I_{1, k}^{(i)}}+\frac{\varphi_{2, k}^{(i)} \rho_{2, k}^{(i)}}{1+I_{2, k}^{(i)}}\right) \leq 0 \\
\left(\tilde{v}_{2, k}^{w}\right): \tilde{f}_{2, k}^{w}=R_{2, k}^{w}-\frac{1}{N} \sum_{i \in S_{k}} \log _{2}\left(1+\frac{\varphi_{2, k}^{(i)} \rho_{2, k}^{(i)}}{1+I_{2, k}^{(i)}}\right) \leq 0 \\
\left(\tilde{v}_{1, k}^{w}\right): \tilde{f}_{1, k}^{w}=R_{1, k}^{w}-\frac{1}{N} \sum_{i \in \bar{S}_{k}} \log _{2}\left(1+\frac{\varphi_{1, k}^{(i)} \rho_{1, k}^{(i)}}{1+I_{1, k}^{(i)}}\right) \leq 0 \\
\left(\tilde{v}_{2, k}^{s}\right): \tilde{f}_{2, k}^{s}=R_{1, k}^{w}+R_{2, k}^{s}-\frac{1}{N} \sum_{i \in \bar{S}_{k}} \log _{2}\left(1+\frac{\varphi_{1, k}^{(i)} \rho_{1, k}^{(i)}}{1+I_{1, k}^{(i)}}+\frac{\varphi_{2, k}^{(i)} \rho_{2, k}^{(i)}}{1+I_{2, k}^{(i)}}\right) \leq 0 \\
\left(\tilde{\lambda}_{k}\right): \tilde{h}_{k}=\sum_{i=1}^{N} \sum_{u=1}^{2} \varphi_{u, k}^{(i)}-1 \leq 0, \\
R_{u, k}^{w} \geq 0, R_{u, k}^{s} \geq 0
\end{array}, \quad u=\{1,2\}\right\}
$$

It has to be remarked that the previous solution is found assuming that the interference received at each step remains constant. Moreover, the actual power allocation is obtained by using equation (51) once we have obtained variables $\left\{\varphi_{u, k}\right\}$,
0. Initialize $R_{p, k}^{s}, R_{p, k}^{w}, \quad I_{p, k}^{(i)}=0, \quad \tilde{w}_{k}^{(i)}=0 \quad p, k, t \in\{1,2\}$

1. Every source solves problem $\mathrm{P}_{k}^{\text {Non-Orth }}, k \in\{1,2\}$
2. The $t$-th source generates $\tilde{\Omega}_{\mathrm{k}, \mathrm{t}}^{(i)}$ using eq. $(56), k, t \in\{1,2\}$
3. Price exchange through the backhaul link
4. Update transmit precoders $\gamma_{p, k}^{(i)}$ using $\varphi_{p, k}^{(i)}$ and eq. (51)
5. The $k$-th source collects all prices generated by neigboring sources $\tilde{w}_{k}^{(i)}$ eq.(56)
6. Every source collects the noise plus interference, $I_{p, k}^{(i)}$
7. Go to 1

Table 8. Decentralized algorithm to solve problem $P_{0}^{\text {Non-Orth }}$ (35).

### 6.1.3 Numerical results

We have considered the scenario depicted in Figure 10 with the following average SNR at the different channels,

$$
\begin{array}{ll}
\rho_{11}=14 d B & \rho_{32}=12 d B \\
\rho_{21}=12 d B & \rho_{42}=10 d B
\end{array}, \quad \rho_{12}=\rho_{22}=\rho_{31}=\rho_{41}=\operatorname{Interf} d B
$$

where the interference received by each terminal becomes the same (variable Interf) and we will tune it in order to analyse the performance of the investigated techniques.

The channel is frequency selective and presents a Rayleigh distributed with 6 symbol length and equal average power taps. All terminals are equipped with single antennas and the transmission is carried out along 16 orthogonal carriers or resource blocks (RB). We will study the sum-rate served by $\mathrm{T}_{2}\left(\mathrm{U}_{1,2}\right.$, $\mathrm{U}_{2,2}$ ), i.e. $R_{1,2}+R_{2,2}$, as a function of the sum-rate served by $\mathrm{T}_{1}\left(\mathrm{U}_{1,1}, \mathrm{U}_{2,1}\right)$, i.e. $R_{1,1}+R_{2,1}$, for four distributed resource allocation algorithms depending on the type of coordination and the user-access. They are based on a competitive user approach (Iterative Waterfilling, ITW) or in a coordinated approach (ITW + pricing). Likewise, we also take into account if we have an orthogonal user access with simple transmitters (Orth UA), see section A, or in contrast, there is a non-orthogonal user access with complex transmitters (Non-Orth UA), see section B.

We have considered different values of the priorities assigned to the uses, but limited to be,

$$
\left\{\begin{array}{l}
\mu_{2,1}=2 \mu_{1,1} \\
\mu_{2,2}=3 \mu_{1,2}
\end{array},\left\{\begin{array}{c}
\mu_{1,1}=(1-\kappa) \\
\mu_{1,2}=\kappa
\end{array}, \quad \kappa \in(0,1)\right.\right.
$$

Figure 11 depicts the achievable rate region of the four transmissions schemes when the received interference at each terminal is equal to Interf $=\{0,5,10,15\} \mathrm{dB}$, which accounts for a very low, low, medium and high interference scenarios. We can observe that in all cases, the ITW and ITW+pricing approaches tend to have a similar performance when priorities $\mu_{1,1} \approx \mu_{1,2}(\kappa=1 / 2)$, and when only one source is active. For other values of user's priorities, the pricing method allows improving the achieved rates. Such gain increases as the generated interference is higher.


Figure 11. Sum-rate served by $T_{2}$ as a function of the sum-rate served by $T_{1}$. Figure 11-\{top-left, top-right, bottom-left, bottom-right\} depict very low, low, medium and high interference cases

The achievable rate regions shown in Figure 11 tackle four rates, but we represent the sum-rate served by each source. Since the priorities of those users associated to the same source are not equal it might happen that the achievable rate region of the pricing mechanisms would be inferior to the non-pricing schemes, like in Figure 11-bottom-left, but the weighted sum-rate is in general improved by the pricing schemes.

Nevertheless, Figure 11-bottom-right shows the Non-Orthogonal user access with the pricing mechanism describes a non-convex achievable region that in some cases gets a worse performance than the one attained by the non-pricing mechanism in terms of weighted sum-rate, see Figure 12-left. These results are due to the pricing algorithm does not converge to a stable solution when the interference is high. Figure 12-right illustrates the total sum-rate evolution along the different iterations where the decentralized resource allocation algorithm is carried out. We can observe that the algorithm tends to a solution where the sum-rate oscillates. In order to avoid such drawback, we have incorporated some kind of memory capability at the pricing exchange [Agustin11a]. Basically, the exchanged pricing value weights with variable $\alpha$ the current parameter calculated in (56) with that value calculated in the previous iteration (using 1- $\alpha$ ), like in the low pass filter (LPF). This approach provides a stable solution in terms of sum-rate and gives us a convex achievable rate region, see Figure 13.


Figure 12. Left) Weighted sum rate (WSR) served by $\mathrm{Tx}_{2}$ as a function of the WSR served by $\mathbf{T x}_{1}$ in the high interference scenario, Right) Sum-rate evolution for the NonOrthogonal user-access when $\mu_{1,1}=0.41, \mu_{1,2}=0.59, \mu_{2,1}=2 \mu_{1,1}, \mu_{2,2}=3 \mu_{1,2}$. Introduction of a LPF with $\alpha=0.7$ for the pricing


Figure 13. Sum-rate served by $T_{2}$ as a function of the sum-rate served by $T_{1}$ in the high interference scenario. Introduction of a LPF with $\alpha=0.7$ for the pricing

### 6.2 MIMO case and backhaul rate constraints

In this section we look into a decentralized technique for downlink interference management in overlaid MIMO femtocell-macrocell deployments operating in the same band under OFDMA access. The carrier allocation and MIMO precoders are optimized by considering the weighted sum rate as target function. Being the overall problem non-convex, we propose to solve it in a decentralized way where each femto access point (FAP) optimizes a convex problem based on interference levels reported by the user equipments (UE) and interference prices exchanged with other FAPs and with the macro base station (MBS), exploiting the wired backhaul connection. The problem includes a rate restriction that accounts for the limited bitrate provided by the Internet Service Provider (ISP) wired backhaul. Numerical results compare strategies based on pricing exchange and on pure competition in dense FAP deployments

Present work looks into decentralized solutions for the resource allocation in downlink transmissions based on interference price exchange. Such coordination is carried out through the IP-based backhaul link present in femtocell networks. The key innovative aspects analyzed are

- Backhaul rate constraint. As the air interface might provide much larger spectral efficiency than current ISP connection, resource allocation has to consider a total rate constraint that accordingly limits the transmitted power and the number of resource blocks (RB) used.
- Resource block assignment. Each source terminal (be it macro base station, MBS, or FAP) is serving multiple users simultaneously under OFDMA. Assuming low complexity receivers, a given RB is assigned to just one UE associated to that source terminal. The optimal solution requires an exhaustive search over all RBs. However, we borrow results from [Seong06] where an efficient polynomial complexity search is presented.
- Multi-antenna configuration. We derive the transmit precoder structure and interference prices to be generated for the MIMO case. In contrast to [Shi09b], we do not consider the different transmitter beams separately.
- Preserving Quality of Service. By maximizing the weighted sum-rate of the system, we can assign different priorities to users. If the MBS becomes an additional player it may hold higher priority than FUEs to preserve quality for the macro user equipments (MUE). We will see that user priorities affect the interference prices.


### 6.2.1 Description

We consider an OFDMA-based wireless deployment of one macro base station (MBS) and $N_{F A P}$ femto access points (FAPs), all connected through an IP-based backhaul link. The $k$-th FAP serves up to $N_{\text {FUE }}^{k}$ FUEs simultaneously. We impose that a physical resource block (RB) is assigned to just one of the FUEs associated to the same FAP. All terminals are equipped with multiple antennas. The bitrate of the $u$-th FUE associated to the $k$-th FAP, denoted by $(u, k)$-th FUE in the following, is given by the Shannon capacity of the wireless MIMO link assuming the received interference as additive white Gaussian nose (AWGN) noise,
where $N_{\mathrm{RB}}$ denotes the total number of RBs, $U_{u, k}$ defines the set of RBs assigned to the $(u, k)$-th user, $\mathbf{H}_{u, k}^{(i)}$ is the MIMO channel matrix at the $i$-th RB between the $k$-th FAP and the $u$-th FUE, $\mathbf{S}_{u, k}^{(i)}$ stands for the transmit covariance matrix employed by the $(u, k)$-th FUE at the $i$-th RB, while $\mathbf{S}_{-k}^{(i)}$ denotes the transit covariance matrix used by the interfering FAPs (all except the $k$-th one) at the $i$-th RB, and finally $\mathbf{R}_{u, k}^{(i)}$ is the noise plus interference covariance matrix measured by the $(u, k)$-th FUE at the $i$-th

RB. Notice that (58) considers the interference received from the remaining FAPs with $\mathbf{H}_{u, k, t}^{(i)}$ the MIMO channel between the $(u, k)$-th FUE and the $t$-th interfering FAP. However, the maximum transmitted rate per FAP (sum of bitrates of its associated FUEs) depends also on available quality of the backhaul link. Such limitation influences on the resource allocation and the service given.

The resources to be optimized at the k-th FAP

$$
\begin{align*}
& \mathbf{r}_{k}=\left[n_{1, k}, \ldots, r_{N_{\text {FUE }}^{k}}^{k}, k\right] \\
& \boldsymbol{\Sigma}_{k}=\left[\boldsymbol{\Sigma}_{1, k}, \ldots, \boldsymbol{\Sigma}_{N_{F L E}^{k}, k}\right], \boldsymbol{\Sigma}_{u, k}=\left[\mathbf{S}_{u, k}^{(1)}, \ldots, \mathbf{S}_{u, k}^{\left(N_{R B}\right)}\right]  \tag{59}\\
& U_{k}=\left[\left\{U_{1, k}\right\}, \ldots,\left\{U_{N_{F U E}^{k}, k}, k\right]\right.
\end{align*}
$$

where $\mathbf{r}_{k}, \boldsymbol{\Sigma}_{k}$ and $U_{k}$ define the bitrate vector, the transmit covariance matrices and the set of RB assigned to the FUEs associated to the $k$-th FAP, respectively. If the $j$-th RB is not assigned to the $(u, k)$-th FUE then $\mathbf{S}_{u, k}^{(j)}=\mathbf{0}$.

The resource allocation is obtained as the solution of the following optimization problem where the weighted sum-rate (WSR) of the system is maximized,
where $\omega_{u, k}$ is the priority of the $(u, k)$-th FUE, $r_{u, k}$ denotes the optimized bitrate and $C_{k}$ is the set of constraints to be satisfied by the $k$-th FAP which are given by

$$
C_{k}=\left\{\begin{array}{l}
\left(v_{u, k}\right): f_{u, k}=r_{u, k}-R_{S}^{u, k}\left(\left\{\mathbf{S}_{u, k}^{(i)}\right\},\left\{\mathbf{S}_{-k}^{(i)}\right\}, U_{u, k}\right) \leq 0  \tag{61}\\
\left(\lambda_{k}\right): h_{k}=\sum_{u=1}^{N_{\text {FUE }}^{k}} \sum_{i \in U_{u, k}}^{U_{u, k} \mid} \operatorname{tr}\left(\mathbf{S}_{u, k}^{(i)}\right)-P_{M A X}^{k} \leq 0 \\
\left(\psi_{k}\right): \quad g_{k}=\sum_{u=1}^{N_{\text {IUE }}^{k}} r_{u, k}-R_{B}^{k} \leq 0, \\
-\operatorname{tr}\left(\mathbf{S}_{u, k}^{(i)}\right) \leq 0, \quad \mathbf{S}_{u, k}^{(i)}=\mathbf{0} \quad \text { if } \quad i \notin U_{u, k} \\
U_{u, k} \cap U_{p, k}=\varnothing, \quad \forall u \neq p \in\left[1, \ldots, N_{F U E}^{k}\right]
\end{array}\right.
$$

where $R_{B}^{k}$ stands for the maximum rate that can be served by the $k$-th FAP due to the current quality of the backhaul link, $P_{M A X}^{k}$ denotes the maximum transmitted power, $U_{u, k}$ defines the set of RBs assigned to the ( $u, k$ )-th FUE and finally, $v_{u, k}, \lambda_{k}, \psi_{k}$ stand for the Lagrange multipliers associated to the Shannon capacity constraint at the ( $u, k$ ) -th FUE (function $f_{u, k}$ ), sum-power (function $h_{k}$ ) and sumrate (function $g_{k}$ ) constraints at the $k$-th FAP. Notice that the precoding matrix of a given user is set to zero on those RBs assigned to other FUEs.

Let us define the price generated by the FUEs associated to the $t$-th FAP and intended to $k$-th FAP by $\mathbf{\Omega}_{k, t}^{(i)}$ and the total price received by the $k$-th FAP by $\mathbf{W}_{k}^{(i)}$ in the $i$-th RB,

$$
\begin{align*}
& \mathbf{\Omega}_{k, t}^{(i)}=\sum_{p=1}^{N_{t=1}^{t} v_{p, t}} v_{p, t, k}^{(i)}{ }^{H} \boldsymbol{\Pi}_{p, t}^{(i)} \mathbf{H}_{p, t, k}^{(i)}, \quad \mathbf{w}_{k}^{(i)}=\sum_{t=1, t \neq k}^{N_{E, p, p}+1} \mathbf{\Omega}_{k, t}^{(i)}  \tag{62}\\
& \boldsymbol{\Pi}_{p, t}^{(i)}=\left(\mathbf{R}_{p, t}^{(i)}\left(\mathbf{S}_{-t}^{(i)}\right)+\mathbf{H}_{p, t}^{(i)} \mathbf{S}_{p, t}^{(i)} \mathbf{H}_{p, t}^{(i) H}\right)^{-1}-\left(\mathbf{R}_{p, t}^{(i)}\left(\mathbf{S}_{-t}^{(i)}\right)\right)^{-1}
\end{align*}
$$

Following [Shi09] and assuming given fixed interference prices and power profiles, the global problem $\mathrm{P}_{0}$ is decomposed into $N_{F A P}$ subproblems, [Agustin1 1a],[Agustin1 1b] where each FAP has to solve,

$$
\begin{align*}
& \left(\mathrm{P}_{k}\right): \underset{\left\{r_{u, k}\right\},\left\{\mathbf{\Sigma}_{u, k}\right\},\left\{U_{u, k}\right\}}{\operatorname{minimize}}-\sum_{u=1}^{N_{F V E}^{k}} \omega_{u, k} r_{u, k}+\sum_{i \in U_{u, k}}^{\left|U_{u, k}\right|} \operatorname{tr}\left(\mathbf{W}_{k}^{(i)} \mathbf{S}_{u, k}^{(i)}\right)  \tag{63}\\
& \quad \text { subject to } C_{k}
\end{align*}
$$

It is important to remark the information to be exchanged also depends on Lagrange multiplier $v_{p, t}$, associated to the Shannon rate constraint for the ( $p, t$ )-th FUE.. The final values of $v_{p, t}, \Pi_{p, t}^{(i)}$ are obtained in the following sections.

Finally, problem $\mathrm{P}_{0}$ can be addressed in a decentralized way following the iterative algorithm depicted in Table 9.
0. Initialize $\mathbf{W}_{k}^{(i)}=\mathbf{0}, \quad \mathbf{R}_{p, k}^{(i)}=\sigma_{p, k}^{2} \mathbf{I}$

1. Every FAP solves problem $\mathrm{P}_{\mathrm{k}}$ using eq.(79), $k \in\left[1, \ldots, N_{F A P+1}\right]$
2. Every FAP generates the $\boldsymbol{\Omega}_{p, k}^{(i)} \quad$ eq.(80), $p \in\left[1, \ldots, N_{F U E}^{k}\right]$
3. Pricing exchange through the backhaul link
4. Update transmit precoders $\mathbf{S}_{p, k}^{(i)}$ with the solution of step 1
5. Every FAP collects all prices into $\mathbf{W}_{k}^{(i)}$,
6. Every FAP collects the noise plus interference, $\mathbf{R}_{p, k}^{(i)}$ measured at the UE eq.(55), $p \in\left[1, \ldots, N_{F U E}^{k}\right]$
7. Go to 1

## Table 9. Decentralized WSR maximization of $\mathbf{P}_{\mathbf{0}}$

### 6.2.2 Radio resource allocation at the k-th FAP

Problem $P_{k}$ in (63) is a convex problem when the bitrate and precoding matrices are the optimized variables. However, when RB assignment is an additional variable, then problem $\mathrm{P}_{\mathrm{k}}$ becomes nonconvex.

### 6.2.2.1 Rate and Power allocation given a certain RB assignment

The optimal transmit covariance when there is a RB pre-assignment over the served FUEs such as $U_{u, k} \cap U_{p, k}=\varnothing \quad \forall u \neq p$, can be obtained as a solution of,

$$
\begin{align*}
& \left(\tilde{\mathrm{P}}_{k}\right): \operatorname{minimize}_{\left\{r_{u, k},\right\},\left\{\sum_{u, k}\right\}} \sum_{u=1}^{N_{k, u \in k}^{k}} \omega_{u, k} r_{u, k}+\sum_{i \in U_{u, k}}^{\left|U_{u, k}\right|} \operatorname{tr}\left(\mathbf{W}_{k}^{(i)} \mathbf{S}_{u, k}^{(i)}\right)  \tag{64}\\
& \text { subject to } C_{k}
\end{align*}
$$

Proposition 1. The transmit covariance matrix used by the $(u, k)$-th user at the $i$-th RB $\left(i \in U_{u, k}\right)$ is defined by,

$$
\begin{align*}
& \mathbf{S}_{u, k}^{(i)}\left(\lambda_{k}, \psi_{k}\right)=\mathbf{Z}_{u, k}^{(i)} \mathbf{\Theta}_{u, k}^{(i)}\left(\lambda_{k}, \psi_{k}\right) \mathbf{Z}_{u, k}^{(i) H} \\
& {\left[\mathbf{\Theta}_{u, k}^{(i)}\left(\lambda_{k}, \psi_{k}\right)\right]_{m, m}=\left[\frac{\omega_{u, k}-\psi_{k}}{N_{R B} \ln 2}-\left[\mathbf{\Lambda}_{u, k}^{(i)}\left(\lambda_{k}\right)\right]_{m, m}^{-1}\right]^{+}} \tag{65}
\end{align*}
$$

where $\lambda_{k}$ and $\psi_{k}$ are the Lagrange multipliers associated the max sum-power and max sum-rate constraints, $\Theta_{u, k}^{(i)}$ stands for a diagonal matrix that contains the power allocation over the different modes (up to $\left.\beta_{u, k}^{(i)}\right),[\mathrm{a}]^{+}$is an operator used for $\max (a, 0)$ and remaining matrices are defined by,

$$
\left\{\begin{array}{l}
\mathbf{B}_{k}^{(i)}=\operatorname{eigenvec}\left(\mathbf{W}_{k}^{(i)}\right), \quad \mathbf{V}_{u, k}^{(i)}=\operatorname{eigenvec}\left(\overline{\mathbf{H}}_{u, k}^{(i) H} \overline{\mathbf{H}}_{u, k}^{(i)}\right)  \tag{66}\\
\overline{\mathbf{H}}_{u, k}^{(i)}=\left(\mathbf{R}_{u, k}^{(i)}\right)^{\frac{1}{2}} \mathbf{H}_{u, k}^{(i)} \mathbf{B}_{k}^{(i)}\left(\mathbf{D}_{k}^{(i)}+\lambda_{k} \mathbf{I}\right)^{\frac{1}{2}} \\
\mathbf{Z}_{u, k}^{(i)}=\mathbf{B}_{k}^{(i)}\left(\mathbf{D}_{k}^{(i)}+\lambda_{k} \mathbf{l}\right)^{\frac{-1}{2}} \mathbf{V}_{u, k}^{(i)}, v_{u, k}=\left[\omega_{u, k}-\psi_{k}\right]^{+} \\
\mathbf{D}_{k}^{(i)}=\operatorname{eigenval}\left(\mathbf{W}_{k}^{(i)}\right), \quad \beta_{u, k}^{(i)}=\operatorname{rank}\left(\overline{\mathbf{H}}_{u, k}^{(i)} \overline{\mathbf{H}}_{u, k}^{(i)}\right) \\
\boldsymbol{\Lambda}_{u, k}^{(i)}=\operatorname{eigenval}\left(\overline{\mathbf{H}}_{u, k}^{(i) H} \overline{\mathbf{H}}_{u, k}^{(i)}\right)^{(i)}
\end{array}\right.
$$

The bitrate $r_{u, k}$ is obtained by applying the Shannon capacity, (58), with the optimal transmit covariance matrix depicted in (65).

## Proof. See [Agustin 11b]

The optimal values for the Lagrange multipliers $\lambda_{k}$ and $\psi_{k}$ in Proposition 1 can be efficiently obtained using the algorithm presented in Table 10. It is based on the bisection method [Boyd04] and the subgradient concept [Yu06]. Functions $h_{k}$ and $g_{k}$ in (64) are the subgradients employed to update $\lambda_{k}$ $\psi_{k}$, respectively.

```
1. Initialize: \(\lambda^{\text {max }}, \lambda^{\text {min }}\)
2. while \(\left|\lambda^{\text {max }}-\lambda^{\text {min }}\right| \leq \varepsilon\) do
3. \(\lambda_{k}=\frac{1}{2}\left(\lambda^{\text {max }}+\lambda^{\text {min }}\right)\)
4. Initialize: \(\psi^{\text {max }}=\max \left(\omega_{u}\right), \psi^{\text {min }}=0\)
5. while \(\left|\psi^{\text {max }}-\psi^{\text {min }}\right| \leq \varepsilon\) do
6. \(\psi_{k}=\frac{1}{2}\left(\psi^{\text {max }}+\psi^{\text {min }}\right)\)
7. Calcuate power allocation \((\lambda, \psi)\) using eq.(82)
8. Get bitrates eq.(76)
9. Obtain subgradients \(h_{k}\left(\lambda_{k}, \psi_{k}\right), g_{k}\left(\lambda_{k}, \psi_{k}\right) \quad\) eq.(81)
10. if \(g_{k}\left(\lambda_{k}, \psi_{k}\right)<0, \psi^{\text {max }}=\psi_{k}\), else \(\psi^{\text {min }}=\psi_{k}\)
11. end while
12. if \(h_{k}\left(\lambda_{k}, \psi_{k}\right)<0, \lambda^{\text {max }}=\lambda_{k}\), else \(\lambda^{\text {min }}=\lambda_{k}\)
13. end while
```


## Table 10. Efficient Search of $\lambda_{k}, \psi_{k}$ in problem $P_{k}$

### 6.2.2.2 Rate, Power and RB optimization

Problem $\mathrm{P}_{\mathrm{k}}$ in (63) becomes non-convex when the RB assignment has to be optimized. Similarly for the SISO-OFDMA case, an efficient method with polynomial complexity is proposed in [Seong06] to maximize the WSR. It is based on the Lagrange dual function where each tone is taken by at most one user. We apply the same principle to our problem. where the Lagrangian at the $k$-th FAP is defined by,

$$
\begin{equation*}
G_{\left(\left\{v_{u, k}\right\}, \lambda_{k}, \psi_{k},\left\{\tilde{r}_{u, k},\left\{\left\{\sum_{u, k},\left\{U_{u, k},\right\}\right)\right.\right.\right.}=-\sum_{u=1}^{N_{F U E}^{k}} \omega_{u, k} r_{u, k}+\sum_{u=1}^{N_{\text {FUE }}^{k}} v_{u, k} f_{u, k}+\lambda_{k} h_{k}+\psi_{k} g_{k} \tag{67}
\end{equation*}
$$

where $f_{u, k}, h_{k}$, and $g_{k}$ depend on $U_{u, k}$ and are given in (61). In those RB where the $(u, k)$-th FUE is active the transmit covariance matrix satisfy Proposition 1. Hence, (67) becomes,

$$
\begin{align*}
G_{\left(\lambda_{k}, \psi_{k},\left\{U_{u, k}\right\}\right)} & =-\lambda_{k} P_{M A X}^{k}-\psi_{k} R_{B}^{k}-\sum_{u=1}^{N_{F U E}^{k}} \sum_{i \in U_{u, k}}^{\left|U_{u, k}\right|}\left(\left[\omega_{u, k}-\psi_{k}\right]^{+} R_{S}^{u, k}(i)+\lambda_{k} \operatorname{tr}\left(\mathbf{S}_{u, k}^{(i)}\right)\right)=  \tag{68}\\
& =-\lambda_{k} P_{M A X}^{k}-\psi_{k} R_{B}^{k}+\sum_{i \in U_{u, k}}^{\left|U_{u, k}\right|} G_{(i)}\left(\lambda_{k}, \psi_{k}\right)
\end{align*}
$$

where $R_{S}^{u, k}(i)$ is defined in (58). Since each RB is only assigned to one FUE in the same FAP, the minimization of the Lagrangian depicted in (68) is equivalent to search over all $N_{\text {FUE }}^{k}$ FUE assignments at each RB and assign the $i$-th RB to the FUE that minimizes $G_{(i)}\left(\lambda_{k}, \psi_{k}\right)$.
It must be emphasized that only the FUE at the $k$-th FAP that has been allocated the $i$-th RB presents $\mathbf{S}_{u, k}^{(i)} \neq 0$. The optimal Lagrange multipliers $\lambda_{k}, \psi_{k}$ can be obtained from
Table 10 replacing step 7 by the steps depicted in Table 11 , which describe RB assignment for each value of $\lambda_{k}, \psi_{k}$.

$$
\begin{array}{ll}
\text { 7.a } & \text { Initialize: } U_{u, k}=\left[1, \ldots, N_{R B}\right] u \in\left[1 \ldots N_{F U E}^{k}\right] \\
\text { 7.b } & \text { Calculate } \mathbf{S}_{u, k}^{(i)}\left(\lambda_{k}, \psi_{k}\right) \text { for } u \in\left[1 \ldots N_{F U E}^{k}\right], i \in\left[1 \ldots N_{R B}\right] \quad \text { eq.(11) } \\
\text { 7.c } & \text { Assign RB to FUEs in order to minimize } \mathrm{G}_{(i)} i \in\left[1 \ldots N_{R B}\right] \text { eq.(13) } \\
\text { 7.d } & \text { Re-define } U_{u, k} \\
\text { 7.e } & \text { Set } \mathbf{S}_{u, k}^{(i)}\left(\lambda_{k}, \psi_{k}\right)=\mathbf{0} \text { if } i \notin U_{u, k}, u \in\left[1 \ldots N_{F U E}^{k}\right], i \in\left[1 \ldots N_{R B}\right]
\end{array}
$$

Table 11. Modifications to the Algorithm given in Table 10 to address problem ( $\mathbf{P}_{\mathrm{k}}$ ), (63)

### 6.2.3 Generation of prices

After solving $\mathrm{P}_{\mathrm{k}}$, the $k$-th FAP must calculate the interference prices that must transmit to the $t$-th neighboring FAP through the IP-based backhaul, step 2 shown in Table 9.

$$
\begin{equation*}
\mathbf{\Omega}_{t, k}^{(i)}=\sum_{u=1}^{N_{F U E}^{k}}\left[\omega_{u, k}-\psi_{k}\right]^{+} \mathbf{H}_{u, k, t}^{(i)}{ }^{H} \boldsymbol{\Pi}_{u, k}^{(i)} \mathbf{H}_{u, k, t}^{(i)} \tag{69}
\end{equation*}
$$

Notice that interference price generated by the $k$-th FAP intended to the $t$-th FAP depends on the priorities of its associated FUES $\left(\omega_{u, k}\right)$. However, in the results section we will observe that individual bitrates might oscillate when pricing exchange is assumed. We deal with that drawback by including some memory in the interference price, using a low pass filter (LPF), as it was proposed in [Yu07] for the SISO-OFDM case,

$$
\begin{equation*}
\overline{\mathbf{\Omega}}_{t, k}^{(i)}(n)=(1-\tau) \overline{\mathbf{\Omega}}_{t, k}^{(i)}(n-1)+\tau \mathbf{\Omega}_{t, k}^{(i)} \tag{70}
\end{equation*}
$$

where $n$ and $n$-1 denote the current and previous iteration respectively, $\boldsymbol{\Omega}_{t, k}^{(i)}$ is calculated according to (69). Variable $\tau$ controls how the current value impacts in the final price.

### 6.2.4 Numerical results

The investigated radio resource algorithms are evaluated in a residential scenario consisting of 49 houses uniformly distributed in a sector of a macro-cell with radii $r_{0}=500 \mathrm{~m}$. The FAPs are placed inside squared-houses, while the served FUEs can be indoors or outdoors with a certain probability. Moreover, the MUEs are uniformly distributed over the sector of the cell. The channel models for the different links involved in the transmission and parameter settings are defined in [FREEDOM-D21]. MBS and FAPs perform their radio resource allocation in order to maximize the WSR of the system over 24 RB on which the total 10 MHz bandwidth is slotted. Each source (MBS or FAP) is serving 2 UEs on a given scheduling period. In order to preserve the services offered by the MBS, we have assumed all FUEs with the same priority, while MUEs have a priority 5 times higher. The IP-based
backhaul link limits the total served rate per FAP up to $20 \mathrm{Mbits} / \mathrm{s}$. Finally, MBS, FAPs and UEs are equipped with 4,2 and $\{1$ or 2$\}$ antennas, respectively.

The individual rate convergence is illustrated in Figure 14 showing the attained bitrates by one MUE and FUEs associated to a given FAP (out of 49) while the decentralized algorithm introduced in Table 9. is performed. The RBs are pre-assigned and only precoders and bitrates are optimized. We can observe that when sources are competitive (no pricing) the algorithm converges to stable values in a few iterations. But when sources coordinate their transmissions by exchanging interference prices the individual bitrates may oscillate. That drawback is avoided if adding some memory information to the price value by using (70). $\tau=0.7$ has been used below. The bitrates converge to some constant solution in less than 10 iterations in this case.


Figure 14. Individual bitrate as a function of the evolution of the decentralized resource allocation when RB are pre-assigned. Antenna conf: $4 \times 2 \times 2$ (MBS, FAP, UE). Cases: No pricing, Pricing and Price-filtering, (70)

It remains to see what is the performance in terms of WSR and how it improves due to the pricing. Figure 15 depicts the evolution of total WSR when RBs are pre-assigned (Only Pow in legend) or are optimized (RB+Pow). Thanks to the coordination of transmitters by means of pricing exchange, the total WSR is enhanced with respect the competitive approach (no-pricing). Moreover, the use of pricefiltering values in (70) does not significantly degrade the total WSR and allow stable solutions in terms of individual bitrate (see Figure 14).

Figure 16 and Figure 17 are devoted to show the system spectral efficiency attained by MUEs and FUEs under different antenna configurations and with no-price and price-filtering exchange. The resource allocation optimizes RBs and precoding matrices. One interesting result for the $4 \times 2 \times 2$ antenna configuration is that by introducing price exchange the FUE system spectral efficiency is reduced (around $3 \mathrm{bps} / \mathrm{Hz}$ at cdf 0.1 in Figure 17), but MUE system spectral efficiency improves, due to MUEs priorities which are 5 times larger than FUEs (around $3 \mathrm{bps} / \mathrm{Hz}$ at cdf 0.1 in Figure 16). On the other hand, for the $4 \times 2 \times 1$ antenna configuration the price exchange enhances the MUE and FUE spectral efficiency


Figure 15. Evolution of the total WSR when decentralized resource allocation is carried out: RBs are pre-assigned or RBs are also optimized. Antenna conf: $\mathbf{4 \times 2 \times 2}$ (MBS, FAP, UE). Cases: No pricing, Pricing and Price-filtering, (70)


Figure 16. Macro-User system spectral efficiency when no-pricing or price-filtering exchange is assumed. Antenna conf: $\mathbf{4 \times 2 \times \{ 2 , 1 \}}$


Figure 17. Femto-User system spectral efficiency when no-pricing or price-filtering exchange is assumed. Antenna conf: $4 \times 2 \times\{2,1\}$

### 6.3 Conclusions

We have investigated a radio resource allocation algorithm suitable for downlink transmissions in femtocell networks which is scalable with the number of active FAPs. It is based on the exchange of interference prices through wired backhaul link connecting FAPs and MBS. The procedure adjusts the transmitted power and served rate as a function of current backhaul link bitrate. When maximizing the WSR criterion, the priorities of all users are taken into account because they are present in the exchanged prices, in contrast to competitive strategies. Moreover, incorporating some memory term to the exchanged price is useful to reduce the individual bitrate oscillations when the iterative decentralized resource allocation is performed without degrading significantly the total WSR

We have analysed the SISO and MIMO cases when each FAP is serving multiple users under OFDMA criterion. In this regard, the investigated algorithm shows that the best results are obtained when resource blocks ( RB ) are optimized along with the transmit precoders.

## 7 DECENTRALIZED COORDINATION STRATEGIES BASED ON STATISTICAL MODELLING

In this section, we propose a set of decentralized resource allocation algorithms allowing a set of nearby FAPs to automatically adapt to the channel conditions and the macro users' activity, seen as interference. The objective of the allocation strategies can be either the sum-rate maximization, under a power constraint, or the power minimization under a per-user rate constraint. The common characteristic in the proposed strategies, is that they are based on a statistical model of a particular aspect of the system. Specifically focus on the following properties and/or assumptions:

1. FAPs, thanks to backhaul link availability, can perform coordinated channel sensing, the sensing performance are by simple parameters, and such parameters are incorporate in the optimization of the channel access.
2. The MBS's activity, on each sub-channel, is modelled as a set of Discrete Time Markov Chain. Assuming error-free communication among the FAPs through the backhaul link, gametheory based decentralized algorithms are proposed that incorporate the parameters of such a model.
3. Assuming a given, deterministic, MBS's activity, and FAPs still operating in a coordinated manner, we propose a strategy that takes into account statistical models for the quantization error inherent to the inter-FAP signalling messages and the probability of packet losses in such communications, and incorporate them in the design of a suitable resource allocation strategy.

These algorithms have been devised following an approach where a set of femtocells self-organizes, exchanging data at the control plane with the neighboring FAPs only, in order to find out the most appropriate resource allocation strategy and improve the performance that the system would have in the case of a purely competitive approach. In the practical situation of a femtocell deployment, this approach reflects the situation in which 1) the resource allocation mechanism is decentralized; and 2): there is exchange of signalling traffic among the neighboring cells. From a qualitative point of view, in terms of assumptions, what differentiates the contribution of this section from the rest of the deliverable is the presence, for each algorithm, of a "source of randomness" to which the algorithm is tailored.

In the first case randomness is explicitly taken into account when we consider the false alarm probabilities, or equivalently detection probabilities, that come out from channel sensing procedures. In the second one, it is related to the fact that the MBS's activity is described as a DTMC then it assumes a statistical description. While, in the last case, we explicitly take into account the fact that prices are quantized at the transmitter side (thus we have a quantization noise), and then received with a given probability $p$, because of the possibility of link failures between neighboring FAPs, at control data plane.

### 7.1 Preliminaries

Unlike macro networks, femtocells are typically installed by subscribers and they are deployed and maintained without global planning, with no special consideration about traffic demands or interference with other cells, either femto or macro cells. Hence, a potential massive deployment of FAPs might induce an intolerable interference from femto to macro users, as well as from femto to femto users, as in Figure 18, where the black lines represent the IP-based backhaul link (where neighboring FAPs can exchange signalling traffic at the control plane), while the red lines and blue lines represent the interference and data link, respectively.

Interference management is then arguably the major challenge to be faced. In principle, the optimal solution to interference management would require an accurate global planning. However, a centralized planning is not really a viable solution, for several reasons and a global optimization would require the exchange of a huge amount of data among the many FAPs and the macro base stations inducing an excessive signalling traffic. A more interesting approach consists in devising decentralized mechanisms able to adapt resource allocation in order to limit interference adequately and to get the advantages offered by the capillary deployment of FAPs. A fundamental tool to devise innovative decentralized resource allocation strategies is game theory, a branch of mathematics studying interactive decision problems connected to multi-objective optimization.


Figure 18. Femtocell network scenario.
The multi-objective optimization in such a case may consist, for example, in maximizing the transmission rate of a FAP, under power budget constraints, or in minimizing the FAP transmission power necessary to guarantee a desired rate (see section 5). In this kind of game, FAPs are the players, and since the wireless channel is an interference channel, the transmission strategy adopted by any player is going to affect the performance of the other players - FAPs - who are then going to react and change their strategy consequently.

Game theory provides the basic tools to study this kind of problems. A possible form of equilibrium is the celebrated Nash Equilibrium (NE), indicating the condition in which every player has no incentive to unilaterally deviate from his strategy, given the strategies of the other players. In this definition, the adverb "unilaterally" plays a key role. In fact, a NE does not consider the situation where two or more players may decide to form a coalition. From a practical viewpoint, the absence of cooperation corresponds to the absence of backhaul links among the FAPs. Dealing with multi-objective optimization, it is first of all fundamental to specify what it means to achieve an optimal solution.
A global notion of optimality in a multi-objective context is the so called Pareto optimality, defined as follows: a set of strategies is Pareto efficient, or Pareto optimal, if it is not possible to make at least some player better off without making any other player worse off. If an equilibrium point belongs to the Pareto boundary, the equilibrium is said to be efficient. In general, however, NE is not necessarily efficient and this is a possible consequence of the competitive nature of the underlying game. In spite of this inefficiency, a NE can be achieved through a purely decentralized approach. Nevertheless it's possible to achieve a more efficient solution than the purely competitive NE, when the IP-based backhaul link between neighboring FAPs is of a sufficient quality, such that it allows the exchange of local information between FAPs. It is then of interest to check if there are strategies to modify the game in order to move the equilibrium point of the modified game towards the Pareto optimal
boundary. One of the mechanisms to be used to achieve such a goal is pricing, which requires some exchange of information (prices) among players (FAPs) at control plane level. From a practical point of view, each price reflects the marginal cost of increasing interference to the other users on a particular subchannel. This possible local coordination among FAPs is made possible in femtocell networks through the, IP-based, backhaul link, which creates an underlying wired network connecting FAPs and MBSs [Barbarossa10].
In this section, we categorize our techniques considering the following issues (the same description is given for the minimum power game in Section 5.1):

1. Objective function: the description of any allocation strategy must clearly identify the objective function, which is being optimized. The objective function is then a natural performance indicator for the considered technique.
2. Constraints: typically, the optimization problem that the proposed technique attempts to solve in a distributed fashion is characterized by one or more constraints on some system parameters or indicators.
3. Price Exchange: the way in which prices are exchanged. It could be deterministic (backhaul link is good enough to allow price exchange with no errors), or stochastic (backhaul link allows price exchange with a given probability). This last case is taken into account in Section 7.4.
4. Spectrum Sensing: in this deliverable we assume that the FAPs are able to implement spectrum sensing, with a certain periodicity, over the system bandwidth. Furthermore, as exchange of local information among FAPs is possible, they could perform coordinated spectrum sensing.

The presented techniques are decentralized strategies based on the achievement of an equilibrium that, thanks to this local coordination, will be surely more efficient than the purely competitive NE. Our work assumes as the operating situation one in which the backhaul network is good enough to allow the exchange of signalling traffic between neighboring FAPs or FAPs and MBSs.

As usual, the FAPs operate as competitors over common radio resources, but when they decide for a given strategy, they take into account those price coefficients exchanged at the control plane level, performing a sort of coordination. Just an introductory example to better understand the effect of pricing mechanisms, Figure 19 shows the sum rate in bit per OFDM symbol [bps] of 20 active FAPs deployed over a given area vs. the interference coverage radius of each FAP ${ }^{1}$. Sum-rate values are those taken at convergence, when a simultaneous distributed gradient projection algorithm with pricing (red lines) and without pricing (blue lines), is run. As expected, with the increasing of interference coverage radius, we have a considerable loss in term of sum rate (interference is higher), but pricing mechanisms allow us to improve equilibrium efficiency.

[^0]

Figure 19. Pricing vs. No Pricing.

The features of the proposed techniques are described in the following, and summarized in Table 12.

| Technique | Objective <br> function | Constraints | Price <br> Exchange | Spectrum sensing |
| :---: | :---: | :---: | :---: | :---: |
| Joint optimization of <br> FAP's throughput <br> and detection <br> parameters | Opportunistic <br> throughput | Interference <br> constraint <br> towards the <br> MUEs | No price <br> exchange, IP- <br> based backhaul <br> link is exploited <br> to perform <br> coordinated <br> channel sensing | Each user needs to sense the <br> whole set of subchannels for <br> each realization of the channel <br> use, i.e. for each slot. FAPs <br> aggregate their observations to <br> identify unused subbands |
| Max rate <br> optimization with <br> power constraint and <br> statistical interference <br> knowledge. | Per user rate | Average <br> power | Deterministic: <br> prices are <br> correctly <br> exchanged <br> between FAPs | Periodical channel sensing with <br> a period larger than the <br> channels allocation time unit of <br> the macro system. |
| Minimum power <br> optimization with <br> rate constraint and <br> exact interference <br> knowledge | Transmit <br> power | Per user <br> information <br> rate | Deterministic: <br> prices are <br> correctly <br> exchanged <br> between FAPs | Each user needs to sense the <br> whole set of subchannels for <br> each realization of the channel <br> use, i.e. for each slot. |

Table 12. Decentralized resource allocation strategies classification
In Section 7.2 we propose a distributed mechanism where FAPs perform a coordinated channel sensing, and focus on how to incorporate the performance and parameter setting of the spectrum sensing phase in the resource allocation algorithm design, so as to optimize the overall network efficiency. Each FAP optimizes its own opportunistic throughput by choosing detection thresholds, sensing time and allocation power vector jointly, under a constraint on the interference to MUEs and a constraint on the total transmit power.

In Section 7.3, following the current trend in 3G systems and their evolution like WiMax and LTE, we consider the problem of allocating power optimally in the joint time-frequency domain, exploiting the use of a statistical model for the interference activity in the time-frequency plane. We propose an optimal power allocation strategy based on modelling the interferer's activity, in particular, as a two-
state Markov chain and show how to maximize the expected value of the femto-users' rate, averaged over the interference statistical model, when some pricing mechanisms are taken into account. In this case we assume that FAPs perform a spectrum sensing periodically at the start of allocation time interval, in order to acquire knowledge of the current channel occupation state of the macro users. Finally in Section 7.4 we address the problem of rate maximization when price coefficients are quantized, and then received with a given probability $p$.

It's important to remark that algorithms in Section 7.2 and 7.3 follow a game theoretic approach, while algorithm in Section 7.4 is a decentralized gradient projection algorithm aimed to maximize the network sum-rate function this algorithm does not follow any game theoretic approach..

### 7.2 Coordinated channel sensing

In a scenario where FAPs coexist with one or more MBSs, without the latter sending real-time resource allocation information to the FAPs, coordinated channel sensing by a group of nearby FAPs may be a useful tool to enhance their capability to track the activity of the MBS, and hence be able to allocate resources in a smart way. It can be showed that, within this framework, to maximize what we define the opportunistic throughput of the FAPs, which quantifies their ability to send information on a given portion of the spectrum, the choice of the detection and channel access parameters should be done in a joint manner.
Therefore, we consider the joint optimization of sensing time, detection thresholds and power allocation across multichannel links, in order to maximize the aggregated opportunistic throughput, given a probabilistic constraint on the interference generated towards the macro users. The problem will be first studied in the single FAP case and then extended to the multi FAP scenario by following a game theoretic set-up, where each FAP competes against others to maximize its own opportunistic throughput by choosing, jointly, the detection parameters, and the vector of power allocation over a multichannel link.

It is important to remark that, in this particular case, coordination among FAPs is taken into account in the detection phase. All the algorithms follow a purely competitive approach, with the only exception of algorithm shown in Table 15, where a particular step of the algorithm is performed running an average consensus, which provides a local exchange of information between neighboring FAPs

### 7.2.1 System and detection models

Let us consider a femtocell system where a group of nearby FAPs (e.g. co-located in a building or in the same area, such as a corporate compound) utilize a wideband channel composed by $N$ nonoverlapping subbands. Each FAP senses the available spectrum with the aim of identifying subbands unused by the MBS, to be used for transmission. To do this, the FAPs aggregate their sensing result and decide, for each subband, on the basis of such an aggregated data set, between the two hypotheses $\mathcal{H}_{1, k}$ and $\mathcal{H}_{0, k}$, denoting, respectively, the presence and absence of MBS's signal. Let $Q_{C}$ be the number of coordinating FAPs and $K_{S}$ the number of observations taken by each FAP on each subband. We assume that the FAPs acquire their signal samples in the same interval. At the end of the sensing interval, through a suitable signalling protocol and using the backhaul link, each FAP disseminates its results to the remaining FAPs. After that, each FAP is able to take a decision on the usage of each subband by the MBSs on the basis of the whole data set. We denote with $x_{q, k}(m)$ the $m$-th observation taken by the $q$-th FAP on the $k$-th subband with $k=1, \ldots, N, q=1, \ldots, Q_{C}, m=1, \ldots, K_{S}$ and introduce the corresponding vector of observations $\mathbf{x}_{q, k} \triangleq\left[x_{q, k}(1), \ldots, x_{q, k}\left(K_{S}\right)\right]^{T}$. The whole data set used for the final decision on the $k$-th subband is represented by the vector $\mathbf{x}_{k} \triangleq\left[\mathbf{x}_{1, k}^{T}, \ldots, \mathbf{x}_{M, k}^{T}\right]^{T}$, which has $K_{S} \times Q_{C}$ elements. For the $k$-th subband, under the two hypotheses of being unused or used by the MBS, we have

$$
\begin{align*}
& \mathcal{H}_{0, k}: x_{q, k}(m)=w_{q, k}(m)  \tag{71}\\
\mathcal{H}_{1, k} & : x_{q, k}(m)=I_{q, k}(m)+w_{q, k}(m) \quad k=1, \ldots, N,
\end{align*}
$$

for $m=1, \ldots, K_{S}$ and $q=1, \ldots, Q_{C}$, where $I_{q, k}(m)$ and $w_{q, k}(m)$ denote, respectively, the $m$-th sample of the aggregated ${ }^{2}$ signal from the MBSs and noise, over the $k$-th subband. Under the assumption that noise and aggregated interference samples, over each subband, are zero mean, statistically independent, Gaussian random variables, the optimal detector, over each subband, is the energy detector [Quan08]:

$$
\begin{equation*}
D_{k}\left(\mathbf{x}_{k}\right)=\sum_{i=1}^{\varrho_{c} K_{S}}\left|\mathbf{x}_{k}(i)\right|^{2} \stackrel{\mathcal{H}_{1, k}}{\gtrless} \gamma_{k}, \quad k=1, \ldots, N \tag{72}
\end{equation*}
$$

where $\gamma_{k}$ is the threshold associated to the $k$-th subband. If the received samples are not Gaussian, the energy detector is no longer optimal, but it is still the most popular detector. For $K_{s}$ sufficiently large, invoking the central limit theorem, $D_{k}\left(\mathbf{x}_{k}\right)$ can be well approximated by a Gaussian probability density function (pdf) [Quan08]. In particular, $D_{k}\left(\mathbf{x}_{k}\right) \sim \mathcal{N}\left(\mu_{0, k}, \sigma_{0, k}\right)$ under $\mathcal{H}_{0, k}$ and $D_{k}\left(\mathbf{x}_{k}\right) \sim \mathcal{N}\left(\mu_{1, k}, \sigma_{1, k}\right)$ under $\mathcal{H}_{1, k}$ with

$$
E\left[D_{k}\left(\mathbf{x}_{k}\right)\right]=\left\{\begin{array}{c}
\mu_{0, k}=Q_{C} K_{S} \sigma_{n}^{2}(k)  \tag{73}\\
\mu_{1, k}=Q_{C} K_{S}\left(\sigma_{n}^{2}(k)+\sigma_{I}^{2}(k)\right)
\end{array}\right.
$$

where $\sigma_{n}^{2}(k)$ and $\sigma_{I}^{2}(k)$ denote the variance of noise and aggregated MBSs, respectively, over the $k$ th subchannel. Furthermore, we have

$$
\operatorname{Var}\left[D_{k}\left(\mathbf{x}_{k}\right)\right]=\left\{\begin{array}{c}
\sigma_{0, k}^{2}=2 Q_{C} K_{S} \sigma_{n}^{4}(k)  \tag{74}\\
\sigma_{1, k}^{2}=2 Q_{C} K_{S} \sigma_{n}^{2}(k)\left(\sigma_{n}^{2}(k)+2 \sigma_{I}^{2}(k)\right)
\end{array}\right.
$$

As sensing performance metrics per subchannel, we use the probability of successfully identifying the unused band (or spectral holes), i.e. $\operatorname{Prob}\left(\widehat{\mathcal{H}}_{k}=\mathcal{H}_{0, k} \mid \mathcal{H}_{0, k}\right)$, and the probability of transmitting over bands occupied by the MBS-MUE links, i.e. $\operatorname{Prob}\left(\widetilde{\mathcal{H}}_{k}=\mathcal{H}_{0, k} \mid \mathcal{H}_{1, k}\right)$, where $\widehat{\mathcal{H}}_{k}$ is the final decision of the hypotheses test. According to these assumptions, the false alarm and detection probabilities depend on the detection thresholds and the number of observation as:

$$
\begin{equation*}
P_{f a}\left(\gamma_{k}, Q_{C} K_{S}\right)=Q\left(\frac{\gamma_{k}-\mu_{0, k}\left(Q_{C} K_{S}\right)}{\sigma_{0, k}\left(Q_{C} K_{S}\right)}\right) \tag{75}
\end{equation*}
$$

and

$$
\begin{equation*}
P_{d}\left(\gamma_{k}, Q_{C} K_{S}\right)=Q\left(\frac{\gamma_{k}-\mu_{1, k}\left(Q_{C} K_{S}\right)}{\sigma_{1, k}\left(Q_{C} K_{S}\right)}\right) \tag{76}
\end{equation*}
$$

where $Q(x)$ denotes the $Q$-function.

### 7.2.2 Maximization of the aggregated opportunistic throughput

Since the detection process inevitably affects the power allocation of the FAP over the sensed spectrum, we provide now a novel formulation of the optimal access strategy, considering jointly the detection parameters (thresholds and sensing time) and the power allocation over all the available

[^1]subbands. We assume that the FAP knows the channel towards the intended receiver. This information can be made available either by a feedback link from receiver to transmitter, or just exploiting channel reciprocity when transmitter and receiver transmit over the same band (Time Division Duplexing).
The goal of the optimization is the maximization of the so called aggregated opportunistic throughput, defined in [Quan09] as the aggregate rate effectively transmitted by the FAP over the unoccupied bands, under the constraint that the power generated towards MBSs be confined within prescribed limits. We provide now a mathematical formulation of this problem. As channel model we consider frequency selective channels with uncorrelated scattering. Furthermore we assume that the spatial distribution of the macro user receiver follows an homogeneous Poisson point process with known spatial density $\rho$ (number of MUEs $/ \mathrm{m}^{2}$ ). The rate of the $k$-th subband for each FAP is:
\[

$$
\begin{equation*}
\tilde{r}_{k}=\log _{2}\left(1+\frac{p_{k}|H(k)|^{2}}{\sigma_{k}^{2}}\right), \tag{77}
\end{equation*}
$$

\]

where $p_{k}$ is the power transmitted over the $k$-th subchannel, $H(k)$ denotes the secondary channel (complex) transfer function over the $k$-th subband, and $\sigma_{k}^{2}=\sigma_{n}^{2}(k)+\sigma_{l}^{2}(k)$ represents the total disturbance, i.e. noise plus interference generated by MBS. Since the $k$-th subcarrier is used by the FAP only when no MBS is detected over that subcarrier, the probability of using that carrier is $\operatorname{Prob}\left(\widehat{\mathcal{H}}_{k}=\mathcal{H}_{0, k} \mid \mathcal{H}_{0, k}\right)=1-P_{f a}\left(\gamma_{k}, Q_{C} K_{s}\right)$, where $\gamma_{k}$ represents the detection threshold over subcarrier $k$. Hence, denoting by $\mathbf{p}=\left[p_{1}, \ldots, p_{N}\right]$ the power vector, the aggregated opportunistic throughput for a given interference profile $\sigma_{k}^{2}$ can be written as

$$
\begin{gather*}
R\left(\mathbf{p} ; p_{f a}\left(\gamma_{1}, Q_{C} K_{s}\right), \ldots, p_{f a}\left(\gamma_{N}, Q_{C} K_{s}\right)\right)=\sum_{k=1}^{N} \operatorname{Prob}\left(\hat{\mathcal{H}}_{k}=\mathcal{H}_{0, k} \mid \mathcal{H}_{0, k}\right) \operatorname{Prob}\left(\mathcal{H}_{0, k}\right)\left(\frac{T-T_{s}}{T}\right) \tilde{r}_{k}= \\
=\sum_{k=1}^{N}\left(1-P_{f a}\left(\gamma_{k}, Q_{C} K_{s}\right)\right)\left(1-\frac{K_{s}}{N_{T}}\right) \log _{2}\left(1+\frac{p_{k}|H(k)|^{2}}{\sigma_{k}^{2}}\right) \operatorname{Prob}\left(\mathcal{H}_{0, k}\right) \tag{78}
\end{gather*}
$$

where we denote with $T$ the total frame duration, $T_{S}$ the sensing time, and thus $\left(T-T_{s}\right) / T$ denotes the portion of the frame duration available for the data transmission while $N_{T}$ represents the total number of observations in the frame. We adopt as detection rule the Neyman-Pearson criterion, so that the false alarm probability is fixed and the threshold is chosen in order to maximize the detection probability. In particular, we assume the same false alarm rate over all the subchannels, i.e., we set $P_{f a}\left(\gamma_{k}, Q_{C} K_{s}\right)=p_{f a}$. Since, in general, the disturbance level over the available subcarriers is not constant, we allow the detection thresholds to vary over the subchannels, under the constraint that the false alarm rate over all the channels is the same.

Introducing the variable $a_{k}=\frac{|H(k)|^{2}}{\sigma_{k}^{2}}$, the aggregate opportunistic throughput can be written as

$$
\begin{equation*}
R\left(\mathbf{p}, p_{f a}, K_{s}\right)=\sum_{k=1}^{N}\left(1-p_{f a}\right)\left(1-\frac{K_{s}}{N_{T}}\right) \log _{2}\left(1+p_{k} a_{k}\right) \operatorname{Prob}\left(\mathcal{H}_{0, k}\right) . \tag{79}
\end{equation*}
$$

The major constraint in a femtocell deployment scenario is to either avoid the interference towards macro-users or to keep it under a tolerable level. Interference over, let us say, the $k$-th subchannel is generated only when the FAP erroneously misses the presence of a MBS over that channel. Hence, in order to limit the femto-to-mobile interference we can consider the interference power generated by the FAP, i.e.

$$
\begin{aligned}
& \sum_{k=1}^{N} \operatorname{Prob}\left(\widehat{\mathcal{H}}_{k}=\mathcal{H}_{0, k} \mid \mathcal{H}_{1, k}\right) \operatorname{Prob}\left(\mathcal{H}_{1, k}\right) p_{k} \frac{\left|H_{F M}(k)\right|^{2}}{r^{\eta}}= \\
& =\sum_{k=1}^{N}\left(1-P_{d}\left(p_{f a}, Q_{C} K_{s}, k\right)\right) \operatorname{Prob}\left(\mathcal{H}_{1, k}\right) p_{k} \frac{\left|H_{F M}(k)\right|^{2}}{r^{\eta}}
\end{aligned}
$$

where

$$
\begin{equation*}
P_{d}\left(p_{f a}, Q_{C} K_{s}, k\right)=Q\left(Q^{-1}\left(p_{f a}\right) b_{k}+c_{k}\left(Q_{C} K_{s}\right)\right) \tag{80}
\end{equation*}
$$

with $b_{k}=\frac{\sigma_{0, k}\left(Q_{C} K_{s}\right)}{\sigma_{1, k}\left(Q_{C} K_{s}\right)}, c_{k}\left(Q_{C} K_{s}\right)=\frac{\mu_{0, k}\left(Q_{C} K_{s}\right)-\mu_{1, k}\left(Q_{C} K_{s}\right)}{\sigma_{1, k}\left(Q_{C} K_{s}\right)}$, while $\left|H_{F M}(k)\right|^{2}$ denotes the channel coefficient between the femto transmitter and the MUE receiver. In order to take into account the distance and the attenuation $\eta$ from wall penetration we have included the term $r^{\eta}$. Since the FAP transmitter has no knowledge of the channel coefficient $\left|H_{F M}(k)\right|^{2}$ we can consider its expected value, that in the case of a channel with $L$ uncorrelated paths, assumes a constant value or $E\left[\left|H_{F M}(k)\right|^{2}\right]=\sum_{l=1}^{L} \sigma_{F}^{2}(l)=K_{1}$. Furthermore since for a Poisson point the distance r between femtomacro user follows the Rayleigh distribution

$$
\begin{equation*}
f_{R}(r)=2 \pi r \rho e^{-r^{2} \pi \rho} \quad \text { for } \quad r \geq 0 \tag{81}
\end{equation*}
$$

and denoting for simplicity $d_{k}=\left(1-P_{d}\left(p_{f a}, Q_{C} K_{s}, k\right)\right) \operatorname{Prob}\left(\mathcal{H}_{1, k}\right)^{3}$, we can consider the following probabilistic constraint

$$
\begin{equation*}
\operatorname{Prob}\left\{\sum_{k=1}^{N} d_{k} \frac{p_{k}}{r^{\eta}} K_{1} \leq p_{\min }\right\} \geq P_{0} \tag{82}
\end{equation*}
$$

where $P_{0}$ is the minimum desired probability value and $p_{\text {min }}$ is the minimum received power. In words, we impose that the probability of the macro-user receiver to be far enough from the FAP so that no interference is perceived, is higher than a minimum fixed value $P_{0}$. Hence we can easily derive the probability density function of the random variable $z=r^{\eta}$ as

$$
\begin{equation*}
f_{Z}(z)=\frac{2 \pi \rho}{\eta} z^{2 / \eta-1} e^{-\pi \rho z^{2 / \eta}} \quad \text { for } \quad z \geq 0 \tag{83}
\end{equation*}
$$

Then defining $\gamma=\left(\sum_{k=1}^{N} d_{k} p_{k}\right) K_{1} / p_{\text {min }}$, we obtain

$$
\begin{equation*}
\operatorname{Prob}\{z \geq \gamma\}=e^{-\pi \rho \gamma^{2 / n}} \tag{84}
\end{equation*}
$$

so that the constraint in (82) can be written as

$$
\begin{equation*}
\gamma \leq P_{1} \quad \text { with } \quad P_{1}=\left(\frac{-\ln \left(P_{0}\right)}{\pi \rho}\right)^{\eta / 2} \frac{p_{\min }}{K_{1}} \tag{85}
\end{equation*}
$$

According to this result we can impose the following final constraint

$$
\begin{equation*}
\sum_{k=1}^{N} d_{k} p_{k} \leq P_{1} \tag{86}
\end{equation*}
$$

[^2]We are now able to formulate the constrained optimization problem. Denoting by $P_{t}$ the transmit power budget of the single FAP, the goal of this last one is the following:

$$
\begin{align*}
\left(\mathbf{p}^{*}, p_{f a}^{*}, K_{s}^{*}\right)=\operatorname{argmax} & \left\{\sum_{k=1}^{N}\left(1-p_{f a}\right)\left(1-\frac{K_{s}}{N_{T}}\right) \log _{2}\left(1+p_{k} a_{k}\right)\right\} \\
\text { subject to } \quad & \sum_{k=1}^{N}\left(1-P_{d}\left(p_{f a}, Q_{C} K_{s}, k\right)\right) p_{k} \leq P_{1}  \tag{87}\\
& \sum_{k=1}^{N} p_{k} \leq P_{t}, \quad \mathbf{p} \geq \mathbf{0} \\
& 0 \leq p_{f a} \leq 1 / 2, \quad 1 \leq K_{s} \leq N_{T}-1
\end{align*}
$$

Note that the first constraint is meaningful if we assume $P_{0}>e^{-\left(P_{t} K_{1} / \rho_{\text {min }}\right)^{2 / 7} \pi \rho}$. Unfortunately, this problem is not convex because the feasible set (the interference constraint) is not convex. Nevertheless, it is interesting to note that fixed $p_{f a}$ and $K_{s}$ the resulting optimization problem is convex and a closed-form solution can be found. Furthermore we will show next that the problem can be reduced into the exhaustive search for the maximum of a two variables function, without loss of generality. In particular, it is easy to check that, if the false alarm probability value $p_{f a}$ and the number of observations $K_{s}$ are given, the previous problem simplifies into the search of the power allocation vector $\mathbf{p}=\mathbf{p}^{*}\left(p_{f a}, K_{s}\right)$ solving the following optimization problem:

$$
\begin{gather*}
\mathbf{p}^{*}\left(p_{f a}, K_{s}\right)=\operatorname{argmax}\left\{\sum_{k=1}^{N}\left(1-p_{f a}\right)\left(1-\frac{K_{s}}{N_{T}}\right) \log _{2}\left(1+p_{k} a_{k}\right)\right\} \\
\text { subject to } \sum_{k=1}^{N} d_{k}\left(p_{f a}, Q_{C} K_{s}\right) p_{k} \leq P_{1}  \tag{88}\\
\sum_{k=1}^{N} p_{k} \leq P_{t} \\
\mathbf{p} \geq \mathbf{0}
\end{gather*}
$$

with $\left.d_{k}\left(p_{f a}, Q_{C} K_{s}\right)=1-Q\left(Q^{-1}\left(p_{f a}\right) b_{k}\left(Q_{C} K_{s}\right)+c_{k}\left(Q_{C} K_{s}\right)\right)\right)$. This problem is strictly convex and the Slater's constraints qualifications hold true, so it has a unique solution, which can be found exploring the KKT optimality conditions. The solution can be found in closed form as a function of the Lagrange multipliers $u$ and $v$, associated to the first and second constraint, and is equal to:

$$
\begin{equation*}
p_{k}^{*}(\mathbf{y})=\left[w_{k}(\mathbf{y})-\frac{1}{a_{k}}\right]^{+} \tag{89}
\end{equation*}
$$

where $\mathbf{y}=\left[p_{f a}, K_{s}\right],[a]^{+}=\max (0, a), w_{k}(\mathbf{y})=\frac{g(\mathbf{y})}{\left(v+u d_{k}(\mathbf{y})\right)}$ and $g(\mathbf{y})=\left(1-p_{f a}\right)\left(1-\frac{K_{s}}{N_{T}}\right) / \ln (2)$.The Lagrange multipliers $u, v \geq 0$ have to be chosen in order to satisfy the following complementary constraints

$$
\begin{array}{cl}
u \cdot\left(\sum_{k=1}^{N} d_{k}(\mathbf{y}) p_{k}-P_{1}\right)=0, & v \cdot\left(\sum_{k=1}^{N} p_{k}-P_{t}\right)=0,  \tag{90}\\
\sum_{k=1}^{N} d_{k}(\mathbf{y}) p_{k} \leq P_{1}, & \sum_{k=1}^{N} p_{k} \leq P_{t} .
\end{array}
$$

Several algorithms can be used to solve the nonlinear system (90). A practical algorithm can be found, e.g. in [Son09]. Once the multipliers $u$ and $v$ are found, (89) provides a closed form solution for the optimal power vector, which we denote as $\mathbf{p}^{*}\left(p_{f a}, K_{s}\right)$, to remark its dependence on $p_{f a}$ and $K_{s}$. As a consequence, the original optimization problem boils down to searching for the $p_{f a}$ and $K_{s}$ values that maximizes $R\left(\mathbf{p}^{*}\left(p_{f a}, K_{s}\right), p_{f a}, K_{s}\right)$. The solution of this problem can be obtained numerically, as shown in the following table.

```
Algorithm I: Single user optimization problem
S.0: Set the vectors \(\mathbf{p}_{f a}^{\prime}=[0: 1 / 2]\) and \(\mathbf{K}_{s}^{\prime}=\left[1: \mathrm{N}_{\mathrm{r}}-1\right]\)
S.1: Set \(R_{\text {max }}=0\)
S.2: for \(\mathrm{n}=1\) :length \(\left(\mathbf{p}_{f a}\right)\)
        \(p_{f a}=\mathbf{p}_{f a}^{\prime}(\mathrm{n})\)
            for \(\mathrm{k}=1\) : length \(\left(\mathbf{K}_{S}^{\prime}\right)\)
            \(K_{s}=\mathbf{K}_{s}^{\prime}(\mathrm{k})\);
        find \(\mathbf{p}^{*}\left(p_{f a}, K_{s}\right)\) by solving the convex problem in (88);
            if \(\mathrm{R}\left(\mathbf{p}^{*}\left(p_{f a}, K_{s}\right), p_{f a}, K_{s}\right)>\mathrm{R}_{\max }\);
            \(\mathrm{R}_{\max }=\mathrm{R}\left(\mathbf{p}^{*}\left(p_{f a}, K_{s}\right), p_{f a}, K_{s}\right)\);
            \(p_{f a}^{*}=p_{f a} ;\)
            \(K_{s}^{*}=K_{s} ;\)
            end
    end
    end
```

Table 13. Single User Optimization Algorithm.

### 7.2.3 Opportunistic throughput maximization: A game theoretic approach

In a femtocell deployment scenario, FAPs are expected to operate in an uncoordinated manner, so that they are competing over the same set of radio resources and there is no central authority assigning them the channel resources. A game theoretic formulation of the access from secondary users, under constraints imposed by the macro-users, is then well appropriate in this context, to find out decentralized solutions, as proposed in [Scutari08a]. However, in [Scutari08a] the payoff of each player did not take into account the decision process and the sensing time. But the detection process is inevitably going to affect the overall network efficiency on one side, as well as the level of undue interference towards macro-users, on the other side. For this reason, we provide in this section a novel formulation of the game, where the payoff function of each player is the opportunistic throughput, which takes into account the detection thresholds and the sensing time. Furthermore, we add a probabilistic constraint on the average interference towards the MBSs that is explicitly written as a function of the detection thresholds and the sensing time. The problem can be formulated as follows. The transmission rate over the $k$-th channel for the $q$-th FAP is

$$
\begin{equation*}
r_{k}^{q}=\log _{2}\left(1+\frac{p_{q}(k)\left|H_{q q}(k)\right|^{2}}{\sigma_{k}^{2}+\sum_{r=1, r \neq q}^{M}\left|H_{r q}(k)\right|^{2} p_{r}(k)}\right) \tag{91}
\end{equation*}
$$

where $p_{q}(k)$ is the power allocated by the $q$-th FAP user over the $k$-th subchannel, $H_{r q}(k)$ denotes the FAP channel transfer function between source $r$ and destination $q$ over the $k$-th subband, and $\sigma_{k}^{2}=\sigma_{n}^{2}(k)+\sigma_{I}^{2}(k)$ represents the total disturbance, i.e. noise plus interference generated by macrousers (for simplicity and without loss of generality we have assumed that $\sigma_{q, n}^{2}(k)$ is independent of $q$ and equal to $\left.\sigma_{n}^{2}(k)\right)$. Hence, denoting by $\mathbf{p}_{q}=\left[p_{q}(1), \ldots, p_{q}(N)\right]$ the $q$-th user power vector and assuming $\mathbf{p}=\left[\mathbf{p}_{1}, \ldots, \mathbf{p}_{M}\right]$, and with $Q_{C}^{q}$ the number of FAPs that coordinate with the $q$-th one in the sensing phase (remark that value $Q_{C}^{q}$ takes into account the $q$-th FAP too), the opportunistic throughput is

$$
\begin{align*}
& R_{q}\left(\mathbf{p} ; p_{f a}\left(\gamma_{1}^{q}, Q_{C}^{q} K_{s}^{q}\right), \ldots, p_{f a}\left(\gamma_{N}^{q}, Q_{C}^{q} K_{s}^{q}\right)\right)=\sum_{k=1}^{N} \operatorname{Prob}\left(\hat{H}_{k}=\mathcal{H}_{0, k} \mid \mathcal{H}_{0, k}\right) \operatorname{Prob}\left(\mathcal{H}_{0, k}\right)\left(\frac{T-T_{s}}{T}\right) r_{k}^{q}= \\
& =\sum_{k=1}^{N}\left(1-p_{f a}\left(\gamma_{k}^{q}, Q_{C}^{q} K_{s}^{q}\right)\right)\left(1-\frac{K_{s}^{q}}{N_{T}}\right) \operatorname{Prob}\left(\mathcal{H}_{0, k}\right) \log _{2}\left(1+\frac{p_{q}(k)\left|H_{q q}(k)\right|^{2}}{\sigma_{k}^{2}+\sum_{r \neq q}\left|H_{r q}(k)\right|^{2} p_{r}(k)}\right) \tag{92}
\end{align*}
$$

with $q=1, \ldots, Q$. The competition among FAPs can then be cast as a non-cooperative strategic game whose structure is

$$
\begin{equation*}
\mathcal{G}=\left\{\Omega,\left\{\Gamma_{q}\right\}_{q \in \Omega},\left\{R_{q}\right\}_{q \in \Omega}\right\} \tag{93}
\end{equation*}
$$

where $\Omega=\{1,2, \ldots, Q\}$ is the set of the players (the FAPs), $\Gamma_{q}$ represents the set of admissible strategies for the $q$-th user given by

$$
\begin{equation*}
\Gamma_{q}=\left\{\mathbf{p}_{q} \in \mathbb{R}_{+}^{N}, p_{f a}^{q} \in[0,1 / 2], 1 \leq K_{s}^{q} \leq N_{T}-1: \sum_{k=1}^{N} p_{q}(k) \leq P_{t}^{q}, \sum_{k=1}^{N}\left(1-P_{d}\left(p_{f a}^{q}, Q_{C}^{q} K_{s}^{q}, k\right)\right) \operatorname{Prob}\left(\mathcal{H}_{4, k}\right) p_{q}(k) \leq P_{1}^{q}\right\} \tag{94}
\end{equation*}
$$

where $P_{t}^{q}$ is the transmit power budget, while $P_{1}^{q}=\left(\frac{-\ln \left(P_{0}^{q}\right)}{\pi \rho}\right)^{\eta / 2} \frac{p_{\min }}{K_{1}}$ depends on the lower bound $P_{0}^{q}$ which we impose on the probability that the macro-user receiver perceives no interference by the FAP transmitter. Finally, $R_{q}$ is the opportunistic throughput, given in (79). In the game $\mathcal{G}$, each player aims to solve the following optimization problem

$$
\begin{array}{cl}
\underset{\mathbf{p}_{q}, p_{f a}^{q}}{\operatorname{maximize}} & R_{q}\left(\mathbf{p}_{q}, \mathbf{p}_{-q}, p_{f a}^{q}, K_{s}^{q}\right) \\
\text { subject to } & \left(\mathbf{p}_{q}, p_{f a}^{q}, K_{s}^{q}\right) \in \Gamma_{q}  \tag{95}\\
\forall q \in \Omega &
\end{array}
$$

where $\mathbf{p}_{-q} \triangleq\left\{\mathbf{p}_{j}\right\}_{j=1, j \neq q}^{Q}$. Unfortunately, the existence of a Nash equilibrium for the game $\mathcal{G}$ appears quite difficult to prove because the admissible strategies space is nonconvex (observe that the second constraint in (94) compromises the convexity of the admissible set). Nevertheless, it can be noted that given $p_{f a}^{q}$ and $K_{s}^{q}$ the feasible set $\Gamma_{q}$ is convex. This allows us to propose two alternative strategies, one aiming at optimizing the power allocation, $p_{f a}^{q}$ and $K_{s}^{q}$, for every FAP, the other forcing the same $p_{f a}^{q}$ and $K_{s}^{q}$ for all users and maximizing the sum throughput.

## Multilevel Iterative Water-Filling (MIWF)

The non-convexity of the feasible set (94) in the joint domain $\left\{\mathbf{p}_{q}, p_{f a}^{q}, K_{s}^{q}\right\}$ makes the search for a solution non trivial. However, as in the single user problem we assume that $p_{f a}^{q}$ and $K_{s}^{q}$ are somehow
given. Indeed, for every vector $\mathbf{y}^{q}=\left[p_{f a}^{q}, K_{s}^{q}\right]$, each FAP can find the optimal vector power $\mathbf{p}_{q}\left(\mathbf{y}^{q}\right)$ in closed form, as a function of $\mathbf{y}^{q}$ and of the Lagrange multipliers $u_{q}$ and $v_{q}$, associated to the power and interference constraints given in (94). The solution is:

$$
\begin{equation*}
\tilde{p}_{q}^{*}\left(\mathbf{y}^{q}, k\right)=\left[w_{k}^{q}\left(\mathbf{y}^{q}\right)-\frac{1}{a_{k}^{q}}\right]^{+} \tag{96}
\end{equation*}
$$

where:

$$
a_{k}^{q}=\frac{\left|H_{q q}(k)\right|^{2}}{\sigma_{k}^{2}+\sum_{r \neq q}\left|H_{r q}(k)\right|^{2} p_{r}(k)} ; \quad w_{k}^{q}\left(\mathbf{y}^{q}\right)=\frac{g\left(\mathbf{y}^{q}\right)}{\left(v_{q}+u_{q} d_{k}^{q}\left(\left(\mathbf{y}^{q}\right)\right)\right)} g\left(\mathbf{y}^{q}\right)=\left(1-p_{f a}^{q}\right)\left(1-\frac{K_{s}^{q}}{N_{T}}\right) / \ln (2) .
$$

The Lagrange multipliers $u_{q}, v_{q} \geq 0$ have to be chosen in order to satisfy the following complementary constraints

$$
\begin{array}{cc}
u_{q} \cdot\left(\sum_{k=1}^{N} d_{k}^{q}\left(\mathbf{y}^{q}\right) p_{q}(k)-P_{1}^{q}\right)=0, & v_{q} \cdot\left(\sum_{k=1}^{N} p_{q}(k)-P_{t}^{q}\right)=0 \\
\sum_{k=1}^{N} d_{k}^{q}\left(\mathbf{y}^{q}\right) p_{q}(k) \leq P_{1}^{q}, & \sum_{k=1}^{N} p_{q}(k) \leq P_{t}^{q} \tag{97}
\end{array},
$$

where $d_{k}^{q}\left(\mathbf{y}^{q}\right)=1-P_{d}\left(\mathbf{y}^{q}, k\right)$. Once the multipliers $u_{q}$ and $v_{q}$ are found, e.g. following the approach proposed in [Son09], (96) provides a closed form solution for the optimal power vector, that we denote as $\mathbf{p}_{q}\left(p_{f a}^{q}, K_{s}^{q}\right)$, as a function of $p_{f a}^{q}$ and $K_{s}^{q}$. Plugging this expression in (92), user $q$ is then able to express his own payoff as a function of the two unknowns $p_{f a}^{q}$ and $K_{s}^{q}$. A numerical search over $p_{f a}^{q}$ and $K_{s}^{q}$ provides then the whole set of unknowns $\left\{\mathbf{p}_{q}, p_{f a}^{q}, K_{s}^{q}\right\}$.
This algorithm can then be used as the best response of every user against the strategies of the other users. The resulting multiuser algorithm is illustrated in Algorithm II, where at the $k$-th iteration step, $\operatorname{MIWF}_{q}\left[\mathbf{p}_{1}^{(k)}, \ldots, \mathbf{p}_{q-1}^{(k)}, \mathbf{p}_{q+1}^{(k-1)}, \ldots, \mathbf{p}_{Q}^{(k-1)}, p_{f a}, K_{s}\right]$ denotes the multilevel water-filling operator (96) based on the power vectors of the users from 1 to $q-1$ updated at the $k$-th step, while the users from $q+1$ to $Q$ are still using the power vectors as computed at step $k-1$.

```
Algorithm II: Multilevel Iterative Water-Filling (MIWF)
S.0: Set the vectors \(\mathbf{p}_{f a}^{\prime}=[0: 1 / 2], \mathbf{K}_{S}^{\prime}=\left[1: \mathrm{N}_{\mathbb{T}}-1\right]\) and fix \(Q_{C}\)
S.1: Set \(k=0\), and initialize with any \(\mathbf{p}^{(0)} \geq 0\)
S.2: while \(k \leq\) num_iter
    \(\mathrm{k}=\mathrm{k}+1\)
    for \(q=1: 0\)
        Max \(=0\);
        for \(\mathrm{n}=1\) : length \(\left(\mathbf{p}_{f a}^{\prime}\right)\)
            \(p_{f a}=\mathbf{p}_{f a}^{\prime}(\mathrm{n}) ;\)
            for \(\mathrm{m}=1\) : length \(\left(\mathbf{K}_{S}^{\prime}\right)\)
                    \(K_{s}=\mathbf{K}_{S}^{\prime}(\mathrm{m}) ;\)
                        \(\mathbf{p}_{o p}^{q}=\operatorname{MIWF}_{q}\left[\mathbf{p}_{1}^{(k)}, \ldots, \mathbf{p}_{q-1}^{(k)}, \mathbf{p}_{q+1}^{(k-1)}, \ldots, \mathbf{p}_{Q}^{(k-1)}, p_{f a}, K_{s}\right]\);
                if \(\mathrm{R}_{\mathrm{q}}\left(\mathbf{p}_{o p}^{q}, \mathbf{p}_{-q}, p_{f a}, K_{s}\right)>\operatorname{Max}\)
                \(\operatorname{Max}=\mathrm{R}_{\mathrm{q}}\left(\mathbf{p}_{o p}^{q}, \mathbf{p}_{-q}, p_{f a}, K_{s}\right)\);
                \(\mathbf{p}_{q}^{(k)}=\mathbf{p}_{o p}^{q} ;\)
                end;
            end;
        end
    end
end
```

Table 14. Multilevel Iterative Water-Filling Algorithm.

## Sum-throughput decentralized maximization

The game theoretic approach illustrated in the previous section shows interesting numerical results but it is hard to substantiate in mathematical terms. To overcome this difficulty, in this section we propose an alternative approach that requires some exchange of information between nearby nodes, that is possible, thanks to backhaul link availability. The approach is sub-optimal in the sense that the false alarm rate and the number of observations used in the sensing phase are forced to be the same for all the FAPs, but the common values are still the object of the optimization. Proceeding as in Algorithm II, the search over the optimal $p_{f a}^{q}$ and $K_{s}^{q}$ are carried out numerically. For each value of the vector $\mathbf{y}=\left[p_{f a}, K_{s}\right]$, every FAP runs a two-step algorithm composed of two phases:

1. Compute the Nash equilibrium of the game $\mathcal{G}$ for the given $\mathbf{y}$ using the same steps as in Algorithm II;
2. Run an average consensus algorithm [Olfati-Saber07] to evaluate the throughput averaged over all the FAPs.

The second step requires the interaction between nearby nodes. It is known that if the FAP's network is connected, under mild assumptions on the consensus algorithm, every FAP is able to get the average throughput, for any given values of $p_{f a}$ and $K_{s}$. Repeating the previous two-steps procedure, for every $\mathbf{y}$ value, each FAP ends up with a curve of the average throughput as a function of $p_{f a}$ and $K_{s}$. This curve allows every FAP to choose the (common) $p_{f a}$ and $K_{s}$ values that maximize the average throughput. In [Scutari08b], the authors provided sufficient conditions guaranteeing the uniqueness of the Nash equilibrium $\mathbf{p}^{*}$ of the game $\mathcal{G}$, for fixed $p_{f a}$ and $K_{s}$, as well as the global convergence of
totally asynchronous algorithms based on the MIWF mapping. Under these conditions Algorithm III, described in the next table, is thus guaranteed to converge.

```
Algorithm III: Sum-throughput decentralized maximization
S.0: Set the vectors \(\mathbf{p}_{f a}^{\prime}=[0: 1 / 2], \mathbf{K}_{S}^{\prime}=\left[1: \mathrm{N}_{\mathbb{T}}-1\right]\) and fix \(Q_{C}\)
S.1: for \(\mathrm{n}=1\) :length \(\left(\mathbf{p}_{f a}^{\prime}\right)\)
        \(p_{f a}=\mathbf{p}_{f a}(\mathrm{n})\);
        for \(\mathrm{m}=1\) : length \(\left(\mathbf{K}_{S}^{\prime}\right)\)
            \(K_{s}=\mathbf{K}_{S}^{\prime}(\mathrm{m}) ;\)
            Set \(K_{s}^{q}=\mathbf{K}_{S}^{\prime}(\mathrm{m}), \quad \forall \mathrm{q}=1, \ldots, \mathrm{\ell}\);
            Compute the Nash equilibrium \(\mathbf{p}^{*}\left(p_{f a}^{q}, K_{s}^{q}\right)\) of the game \(\mathcal{G}\) keeping
            fixed \(p_{f a}^{q}=p_{f a}(\mathrm{n})\) and \(K_{s}^{q}=\mathbf{K}_{S}^{\prime}(\mathrm{m})\);
            \(\operatorname{sum}_{-} \operatorname{rate}(\mathrm{n}, \mathrm{m})=\sum_{r=1}^{Q} \mathrm{R}_{\mathrm{r}}\left(\mathbf{p}^{*}\left(p_{f a}(\mathrm{n}), \mathbf{K}_{S}^{\prime}(\mathrm{m})\right) p_{f a}(\mathrm{n}), \mathbf{K}_{S}^{\prime}(\mathrm{m})\right)\), (average
            consensus);
    end
end
```

Table 15. Sum-throughput decentralized maximization algorithm.

### 7.2.4 Numerical results

In this section we show some numerical results proving the effectiveness of the proposed approaches.

## A. Single FAP scenario

The first important property to verify is the behaviour of the objective function $R\left(\mathbf{p}^{*}\left(p_{f a}, K_{S}\right)\right)$, obtained using the optimal power allocation $\mathbf{p}^{*}\left(p_{f a}, K_{S}\right)$, for any given false alarm value $p_{f a}$ and number of observations $K_{s}$. As an example, in Figure 20, we plot the opportunistic throughput $R\left(\mathbf{p}^{*}\left(p_{f a}, K_{S}\right)\right)$ as a function of $p_{f a}$ and $K_{S}$. It can be seen that there exists a couple of values $\left(p_{f a}^{*}, K_{s}^{*}\right)$ that maximize the opportunistic throughput. Simulation results are related to a single FAP scenario with $P_{t}=10, P_{1}=0.1, \rho=1, p_{\min }=1$ and assuming that the number of coordinating FAPs during the sensing phase is $Q_{c}=5$. In Figure 21 we report $R\left(\mathbf{p}^{*}\left(p_{f a}, K_{S}^{*}\right)\right)$ vs. $p_{f a}$ for different values of the minimum probability $P_{0}$, or equivalently for several values of the parameter $P_{1}$. The behaviors reported in Figure 21 show that, for every $P_{0}$, there exists a value $p_{f a}^{*}$ of the false alarm probability that maximizes the aggregate rate. The value $p_{f a}^{*}$ increases, as the minimum value $P_{0}$ imposed to the probability to have no interference is higher, or the same as $P_{1}$ increases. In particular, the interference constraint is inactive when $P_{1} \geq P_{t}$; in such a case, the optimal solution to (88) reduces to the classical waterfilling and $p_{f a}^{*}=0$.

Title of deliverable: Interference coordination protocols in femto-based networks


Figure 20. Optimal sum rate versus the false alarm probability $p_{f a}$ and the number of observations $K_{s}$ for $N=10$.


Figure 21. $R\left(\mathbf{p}^{*}\left(p_{f a}, K_{S}^{*}\right)\right)$ vs. $\boldsymbol{p}_{f a}$ -


Figure 22. Optimal probability of missing and optimal rate vs. $Q_{C}$.
Finally in Figure 22 the optimal probability of missing detection $1-P_{d}\left(p_{f a}^{*}, K_{S}^{*}\right)$ and the optimal opportunistic rate $R\left(\mathbf{p}^{*}\left(p_{f a}^{*}, K_{S}^{*}\right)\right)$ have been plotted versus the number of cooperating FAPs $Q_{C}$. More specifically, in the upper subplot the several curves refers to the optimal probability of missing detection over different subchannels and it can be noted that by increasing $Q_{C}$ a lower probability can be achieved for some subchannels. Additionally in the lower subplot we can observe that the optimal rate improves by increasing the number of FAPs which cooperate in the sensing phase.

## B. Multi FAP scenario

Let us now consider the multi FAPs scenario in order to numerically test the proposed Algorithms II and III reported respectively, in Table 14 and Table 15. Hence the curves in Figure 23 show the throughputs versus the iteration index of four FAPs obtained with different initializations of the power vectors and assuming, for simplicity, $P_{1}^{q}=P_{1}, P_{t}^{q}=P_{t}$ for $q=1, \ldots, Q$. Interestingly, it can be seen that the algorithm converges in few iterations.

By using the Algorithm III we consider in Figure 24 the optimal sum rate sum_rate $\left(\mathbf{p}^{*}\left(\mathrm{p}_{f a}, K_{s}^{*}\right)\right)$ versus $p_{f a}$ by considering two different values of the spatial density of the MUEs, i.e. $\rho=1,4$. We can note that it exists an optimal value of the false alarm probability which maximizes the sum rate and by increasing $\rho$ maintaining fixed the values of the probability $P_{0}$ a performance loss in terms of sum rate can be observed due to the higher interference versus the nearest MUEs. Furthermore, keeping $\rho$ fixed, a lower sum rate is achieved when the interference constraint is more stringent, i.e. when the probability lower bound $P_{0}$ increases.


Figure 23. Optimal opportunistic throughput vs. iterations.


Figure 24. Optimal sum rate versus the false alarm probability

Finally in Figure 25 in order to quantify the dependence on the number of FAPs which cooperate in the sensing phase, we plot the optimal sum rate versus $Q_{C}$. It can be observed that by increasing $Q_{C}$ the optimal sum rate obtained by running the Algorithm III improves since more accurate sensing results can be achieved.


Figure 25. Optimal sum rate versus the number of cooperating FAPs.

### 7.3 Dynamic resource allocation under Markovian interference model

### 7.3.1 Preliminaries

In this section we incorporate pricing mechanism to the maximum rate and min power games proposed in Section 7.6.3 of [FREEDOM-D31], assuming that the IP-based backhaul link allows a limited local exchange of data among the FAPs. Let us first recall the idea behind the optimization problem in Section 7.6.3 of [FREEDOM-D31]. Femtocells must be fully compliant with cellular standards. Given the current evolution of 3 G systems (see, e.g. WiMax and LTE), it is of interest to look at power allocation techniques in a time-frequency frame. In typical distributed resource allocation algorithms based on game theory, the channel status in terms of channel gain and interference (possibly over multiple subchannels) are assumed to be constant over a given period of time, along which the resources are allocated according to the particular algorithm used. It is a fact, however, that while the constant channel assumption is perfectly reasonable given the low indoor mobility, the interference from macro-users may well vary during an allocation time interval of a group of FAPs. Tracking the interference variations at the level of time-slots would in fact be too cumbersome for the FAPs, as they would spend most of the time sensing the channel and running an iterative algorithm. In other words, the correct power allocation across time and frequency would require a non-causal knowledge of the interference, which of course is not available. To circumvent this inconvenient, we propose a time-frequency resource allocation based on a Markov modelling of the interference activity on each frequency subchannel. In the following, we generalize the max-rate and min-power games to the case of Markovian interference activity. From a system point of view, we have a MBS transmitting over the considered system bandwidth composed of $N$ subchannels, a resource allocation instance takes place every frame composed of $M$ time slots, i.e. on a timefrequency grid of $N M$ resources. For simplicity, the activity on each subchannel is modelled as a two state DTMC with transition probabilities defined as $\omega_{k}$ (idle to idle transition probability) and $\mu_{k}$ (busy to busy transition probability), which correspond to steady states probabilities

$$
\begin{aligned}
& \bar{\beta}_{k}=\frac{p_{1 \rightarrow 0}^{k}}{p_{0 \rightarrow 1}^{k}+p_{1 \rightarrow 0}^{k}}=\frac{\left(\mu_{k}-1\right)}{\left(\omega_{k}+\mu_{k}-2\right)}:=\text { probability to be in the idle state } \\
& \bar{\gamma}_{k}=\frac{p_{0 \rightarrow 1}^{k}}{p_{0 \rightarrow 1}^{k}+p_{1 \rightarrow 0}^{k}}=\frac{\left(\omega_{k}-1\right)}{\left(\omega_{k}+\mu_{k}-2\right)}:=\text { probability to be in the busy state }
\end{aligned}
$$

We remark that the transition probabilities between the different states depend on the application layer traffic statistic, but are not uniquely determined by it. Particularly, the one considered here are the traffic statistics of sub-channel occupation at the physical layer, therefore, they also depend on the implementation of the whole protocol stack, including multi user and multi radio-bearer scheduling and i.e. physical resource block assignment to different radio bearers within one frame. The set of resources is used by a set of FAPs, each one connected, for simplicity, to a single FUE. The FAPs perform spectrum sensing at the start of each frame to acquire knowledge on the channel occupation status from the MBS. Notice, however, that this knowledge will be limited to the first time-slot of the frame. This is the major difference with respect to the preceding section, in which the temporal rate at which the channel occupation status changes is not addressed as a problem. Since the interference level along the frame is unknown, we propose to modify maximum rate and minimum power games [Scutari08b] by substituting the utility function with the expected value of the rate $\bar{R}_{q}\left(\mathbf{p}_{q}, \mathbf{p}_{-q}\right)$ conditioned to the observation performed on the first time slot.

### 7.3.2 Maximum expected rage game

Considering an arbitrary time slot indexed with the letter $m$, we indicate with $\mathbb{P}\left\{S_{k, m}=0\right\}=\beta_{k, m}^{\left(m^{\prime}\right)}$ and $\mathbb{P}\left\{S_{k, m}=1\right\}=\gamma_{k, m}^{\left(m^{\prime}\right)}$ the probabilities that channel $k$ is idle or busy, at time $m$, conditioned on the observed state in an arbitrary preceding step $m^{\prime}<m$. Without loss of generality, we now set $m^{\prime}=0$ and use the notation $\beta_{k, m}$ and $\gamma_{k, m}$ to indicate the conditional occupancy states, dropping the index $m^{\prime}$ even from the expected rate definition. The expected rate is then

$$
\begin{equation*}
\bar{R}_{q}\left(\mathbf{p}_{q}, \mathbf{p}_{-q}\right)=\sum_{m=1}^{M} \sum_{k=1}^{N}\left[\beta_{k, m} \log \left(1+p_{k, m}^{q} a_{n}^{q}(k, m)\right)+\gamma_{k, m} \log \left(1+p_{k, m}^{q} a_{I}^{q}(k, m)\right)\right] \tag{98}
\end{equation*}
$$

where $p_{k, m}^{q}$ is the power over the $k$-th subchannel at the $m$-th time slot for the $q$-th FAP, while

$$
\begin{equation*}
a_{n}^{q}(k, m)=\frac{\left|H_{k}^{q q}\right|^{2}}{\sigma_{n, q}^{2}(k)+\sum_{r \in \mathcal{N}_{q}} p_{k, m}^{r}\left|H_{k}^{r q}\right|^{2}}, \quad a_{I}^{q}(k, m)=\frac{\left|H_{k}^{q q}\right|^{2}}{\sigma_{n, q}^{2}(k)+\sum_{r \in \mathcal{N}_{q}} p_{k, m}^{r}\left|H_{k}^{r q}\right|^{2}+\sigma_{I^{q}}^{2}(k, m)}, \tag{99}
\end{equation*}
$$

with $\sigma_{n, q}^{2}(k)$ that denotes the noise variance on the $k$-th subchannel, $\sigma_{I^{q}}^{2}(k, m)$ denotes the received interference power from the MBS at time $m$ on the subchannel $k$ - resource block (k,m) - and $\sum_{r \in \mathcal{N}_{q}} p_{k, m}^{r}\left|H_{k}^{r q}\right|^{2}$ the interference that the $q$-th FAP receives from its neighbors (with $\mathcal{N}_{q}$ the set of
these neighbors). Finally we indicate with $\left|H_{k}^{x y}\right|^{2}$ the channel transfer coefficient, over subchannel $k$, between the $x$-th transmitter and the $y$-th receiver (we do not have the channel depending on time slot $m$ because we assumed it to be time unvarying for the frame duration). Knowing the transition probabilities through preliminary estimation, the occupancy probabilities $\beta$ and $\gamma$ at any time $m$ are related to the occupancy probabilities of the preceding step by the following recursion rule:

$$
\binom{\beta_{k, m+1}}{\gamma_{k, m+1}}=\left(\begin{array}{cc}
\omega_{k} & 1-\mu_{k}  \tag{100}\\
1-\omega_{k} & \mu_{k}
\end{array}\right)\binom{\beta_{k, m}}{\gamma_{k, m}}
$$

with $m=0, \ldots, M-1, k=1, \ldots, N$, where $\omega_{k}$ and $\mu_{k}$ denote the idle-to-idle and busy-to-busy transition probabilities, respectively; in case of error free spectrum sensing we have $\beta_{k, 0}=0$ and $\gamma_{k, 0}=1$, if channel $k$ is sensed busy at time 0 , or $\beta_{k, 0}=1$ and $\gamma_{k, 0}=0$, if channel $k$ is idle at time 0 . In (98), we distinguish between the interference from the macro-users, which is assumed as given, and the interference from other FAPs, which are modeled as competitive players. In the maximum expected rate game, $\widetilde{\mathcal{G}}$, each player must solve the following local problem:

$$
\begin{array}{cc} 
& \max _{\mathbf{p}_{q}}  \tag{101}\\
(\tilde{P}) & \bar{R}_{q}\left(\mathbf{p}_{q}, \mathbf{p}_{-q}\right) \\
& \text { subject to } \\
\mathbf{p}_{q} \in \tilde{\mathcal{P}}_{q}
\end{array}
$$

where the feasible set of user $q$ is:

$$
\begin{equation*}
\tilde{\mathcal{P}}_{q}=\left\{\mathbf{p}_{q} \in \mathbb{R}^{N M \times 1}: \sum_{m=1}^{M} \sum_{k=1}^{N} p_{k, m}^{q} \leq P_{q}, 0 \leq p_{k, m}^{q} \leq p_{q}^{\max }(k), \forall k \in\{1, \ldots, N\}, m \in\{1, \ldots, M\}\right\} \tag{102}
\end{equation*}
$$

Since the objective function in (101) is strictly concave in $\mathbf{p}_{q} \in \tilde{\mathcal{P}}_{q}$, for any given $\mathbf{p}_{-q}$, and the feasible set $\tilde{\mathcal{P}}_{q}$ is compact and convex, game $\widetilde{\mathcal{G}}$ admits a non-empty solution set for any set of channels and transmit power constraints of the users. In [Barbarossa11b] we reformulated this game as a Variational Inequality (VI) [Facchinei03] and we applied the Iterative Gradient Projection Algorithm (IGPA) to solve it, deriving sufficient conditions for its convergence. Anyway we can't say anything about equilibrium efficiency and, due to the purely competitive approach of this kind of games, in general, it could be really inefficient. The problem is then how to modify the game in order to reduce as much as possible the performance loss related to a NE or, in other words, how to move the NE's of the modified game towards the Pareto optimal boundary. We will now show to introduce a pricing mechanism to this game, and then to improve overall performance (equilibrium). First of all, we have to introduce the price coefficient for the $r$-th FAP, for a given resource block $(k, m), \pi_{k, m}^{r}$, that's defined as:

$$
\begin{equation*}
\pi_{k, m}^{r}:=-\frac{\partial \bar{R}_{r}(\mathbf{p})}{\partial I_{k, m}^{r}\left(\mathbf{p}_{-r}\right)}, \tag{103}
\end{equation*}
$$

with $I_{k, m}^{r}\left(\mathbf{p}_{-r}\right):=\sum_{i \in \mathcal{N}_{r}} p_{k, m}^{i}\left|H_{k, m}^{i r}\right|^{2}$, the interference received by $r$-th FAP on the resource block $(k, m)$ from its neighboring FAPs, and $\bar{R}_{r}(\mathbf{p})$ defined as in (98). From (103) we see that price coefficients are introduced in order to incorporate a cost quantifying the damage that each player's action can induce on the other players utilities. In other words, pricing is a way to incorporate, in each player's strategy, some kind of care about a socially meaningful performance parameter, rather than being purely selfish. Now, the modified game, incorporating pricing mechanisms, is:

$$
\begin{array}{cc}
\max _{\mathbf{p}_{q}} & \bar{R}_{q}\left(\mathbf{p}_{q}, \mathbf{p}_{-q}\right)-\sum_{m=1}^{M} \sum_{k=1}^{N} p_{k, m}^{q}\left(\sum_{r \in \mathcal{N}_{q}} \pi_{k, m}^{r}\left|H_{k}^{q r}\right|^{2}\right)  \tag{104}\\
\text { s.t. } & \mathbf{p}_{q} \in \tilde{\mathcal{P}}_{q}
\end{array}
$$

with $\tilde{\mathcal{P}}_{q}$ defined as in (102). The proposed strategy channel allocation strategy for the FAPs is made up of the following steps:
a. Each FAP observes the set of channels during slot 0 (i.e. at the beginning of a frame)
b. Based on this observation that we assume, for simplicity, to be error free, the each FAP is able to compute the occupancy probabilities for all the slots.
c. Each FAP evaluates the price coefficients in (103) and broadcast them to its neighbors.
d. The (time-frequency) NM-dimensional power allocation vector is optimized in order to maximize the objective function in (104).

The best response of each FAP leads to power coefficients $p_{k, m}^{q}$ that, within the interval $\left[0, p_{q}^{\max }(k)\right)$, must satisfy the following equation

$$
\begin{equation*}
\tilde{a}^{q}(k, m)\left(p_{k, m}^{q}\right)^{2}+\tilde{b}^{q}(k, m) p_{k, m}+\tilde{c}^{q}(k, m)=0 \tag{105}
\end{equation*}
$$

where, denoting with $v_{q}$ the Lagrange multiplier, we have set

$$
\begin{align*}
& \tilde{a}^{q}(k, m)=\left(v_{q}+\sum_{r \in \mathcal{N}_{q}} \pi_{k, m}^{r}\left|H_{k}^{q r}\right|^{2}\right) a_{n}^{q}(k, m) a_{I}^{q}(k, m) ; \\
& \tilde{b}^{q}(k, m)=\left(v_{q}+\sum_{r \in \mathcal{N}_{q}} \pi_{k, m}^{r}\left|H_{k}^{q r}\right|^{2}\right)\left(a_{n}^{q}(k, m)+a_{I}^{q}(k, m)\right)-a_{n}^{q}(k, m) a_{I}^{q}(k, m) ;  \tag{106}\\
& \tilde{c}^{q}(k, m)=v_{q}+\sum_{r \in \mathcal{N}_{q}} \pi_{k, m}^{r}\left|H_{k}^{q r}\right|^{2}-\left(a_{n}^{q}(k, m) \beta_{k, m}+a_{I}^{q}(k, m) \gamma_{k, m}\right)
\end{align*}
$$

We can verify that, $\forall v_{q}>0$, we have $\tilde{b}^{q}(k, m)^{2}-4 \tilde{a}^{q}(k, m) \tilde{c}^{q}(k, m) \geq 0$, and the only solution is

$$
\begin{equation*}
\tilde{p}_{k, m}^{b}=\frac{-\tilde{b}^{q}(k, m)+\sqrt{\tilde{b}^{q}(k, m)^{2}-4 \tilde{a}^{q}(k, m) \tilde{c}^{q}(k, m)}}{2 \tilde{a}^{q}(k, m)} \tag{107}
\end{equation*}
$$

More specifically, we get

$$
p_{k, m}^{q}=\left\{\begin{array}{ccc}
0 & \text { if } & v_{q}+\sum_{r \in \mathcal{N}_{q}} \pi_{k, m}^{r}\left|H_{k}^{q r}\right|^{2} \geq a_{n}^{q}(k, m) \beta_{k, m}+a_{I}^{q}(k, m) \gamma_{k, m}  \tag{108}\\
\tilde{p}_{k, m}^{b} & \text { if } & v_{q}+\sum_{r \in \mathcal{N}_{q}} \pi_{k, m}^{r}\left|H_{k}^{q r}\right|^{2}<a_{n}^{q}(k, m) \beta_{k, m}+a_{l}^{q}(k, m) \gamma_{k, m}
\end{array}\right.
$$

and the optimal power allocation vector is $p_{k, m}^{q^{*}}=\left[\tilde{p}_{k, m}^{b}\right]_{0}^{p_{q}^{m a x}(k)}$ where the multiplier $v_{q}$ is chosen in order to satisfy the power constraint $\sum_{k=1}^{N} \sum_{m=1}^{M}\left[\tilde{p}_{k, m}^{b}\right]_{0}^{p_{a}^{\text {max }}(k)}=P_{q}$. The previous solution assumes, for each player, that the powers used by the other players are given. In practice, the game evolves with each FAP reacting to the choices of the other FAPs. It is then fundamental to prove the convergence of this iterative mechanism. In the following, we present a version of the so called Modified Asynchronous Distributed Pricing algorithm (MADP) proposed in [Shi08] adapted to our formulation. To find the user's best response, it is useful to rewrite (104) introducing a unique index $h$ so that the entries of the power vector $\mathbf{p}_{q}$ are $p_{h}^{q}$ for $h=1, \ldots, N M$. Then, defining the quantities

$$
\begin{equation*}
\operatorname{SNIR}_{h}^{\beta^{q}}:=\frac{p_{h}^{q}\left|H_{h}^{q q}\right|^{2}}{\sigma_{n, q}^{2}(h)+\sum_{r \in \mathcal{N}_{q}} p_{h}^{r}\left|H_{h}^{r q}\right|^{2}}, \quad \operatorname{SNIR}_{h}^{\gamma^{q}}:=\frac{p_{h}^{q}\left|H_{h}^{q q}\right|^{2}}{\sigma_{n, q}^{2}(h)+\sum_{r \in \mathcal{N}_{q}} p_{h}^{r}\left|H_{h}^{r q}\right|^{2}+\sigma_{I_{q}}^{2}(h)}, \tag{109}
\end{equation*}
$$

we can derive the $q$-th user best response as

$$
\begin{equation*}
p_{h}^{q}=\frac{2 c_{h}^{q}}{\sum_{r \in N_{q}} \pi_{h}^{r}\left|H_{h}^{q}\right|^{2}+v_{q}-\eta_{h}^{q}}-p_{h}^{q} \quad \forall h=1, \ldots, M N, \tag{110}
\end{equation*}
$$

where $c_{h}^{q}=\frac{\beta_{h}^{q} \text { SNIR }_{h}^{\beta^{q}}}{1+\text { SNIR }_{h}^{\beta^{q}}}+\frac{\gamma_{h}^{q} \text { SNIR } \gamma_{h}^{\gamma^{q}}}{1+S N I R_{h}^{\gamma^{q}}}$ and $v_{q}$ and $\eta_{h}^{q}$ are the Lagrange multipliers. Given this setting, the modified MADP algorithm is illustrated in Table 16.

```
Algorithm: Modified Asynchronous Distributed Pricing
    S.0: Each user q chooses an initial power profile in the set }\mp@subsup{\tilde{\mathcal{P}}}{q}{}\mathrm{ and set
    n=0;
    S.1: Each user computes its interference prices }\mp@subsup{\pi}{h}{q}(\textrm{n}),\textrm{h}=1,\ldots,\textrm{MN}\mathrm{ and sends
    them to the other users;
    S.2: At each time n, one user q is randomly selected to maximize its utility
    function }\mp@subsup{\overline{R}}{q}{}\mathrm{ and update its power profile given the other user's power
```



```
    ph}\mp@subsup{\mp@code{h}}{(\textrm{n}+1)=\mp@subsup{p}{h}{q}(\textrm{n})+\mp@subsup{\alpha}{q}{}(\textrm{n})*[ [\mp@subsup{p}{h}{\mp@subsup{q}{}{*}}-\mp@subsup{p}{h}{q}(\textrm{n})]\mathrm{ for }\textrm{h}=1,\ldots,MN, where }{\mp@subsup{p}{h}{\mp@subsup{q}{}{*}}}\mathrm{ is given by
    (110);
    S.3: Set n=n+1, go to step S.1 and repeat until convergence is reached.
```


## Table 16. MADP Algorithm.

Following similar arguments as [Shi08], we proved in [Barbarossa11b] that there exists a small enough step size values $\alpha_{q}(n)$ for which the MADP algorithm converges monotonically to a fixed point.

### 7.3.2.1 Numerical results

The following example shows the effects of pricing on the maximum expected rate game. In Figure 26 and Figure 27 we report the sum rate of ten FAPs, in the case where the macro user activity is modeled as a two-state first order DTMC, vs. the number of time slots. In particular, Figure 26 refers to the purely competitive maximum expected rate game of [Barbarossa11a], while Figure 27 refers to the modified game including pricing for a scenario composed by 10 active FAPs. The results are obtained with idle-to-idle transition probabilities $\omega_{k}=\omega=0.5$ and busy-to-busy transition probabilities $\mu_{k}=\mu=0.5$. The number of subcarriers is set to 600 , which is the LTE-A 10 MHz bandwidth system. MBS's activity is supposed to be the same for groups of 12 subchannels. Maximum FAPs' transmit power is set to 20 dBm . Rates are expressed in bit per OFDM symbol [bps].

The three different curves in each figure indicate the sum rate obtained by assuming perfect (noncausal) knowledge of the macro-user activity, no knowledge at all (future activity is the same of the present one), or statistical knowledge, i.e. knowledge of the Markov parameters. Surely Markov parameters are known thanks to a preliminary estimation; this means that there will be an initial time in which each FAP estimates both the transition probabilities $\omega_{k}$ and $\mu_{k}$, see [FREEDOM-D31 Appendix A].


Figure 26. Sum rate of ten FAPs vs. number of time slots for the maximum expected rate game without pricing.


Figure 27. Sum rate of ten FAPs vs. number of time slots for the maximum expected rate game with pricing.

Both figures show that the statistical knowledge (estimation) of the interference activity parameters (Markov transition rates) yields a performance advantage over the case with no information and brings the performance close to the ideal case of perfect non-causal knowledge of the interference activity. Moreover, if we compare Figure 26 and Figure 27 it is evident the gain achieved with the introduction
of pricing. Finally in Figure 28 it is shown the sum rate of ten interfering FAPs, when some pricing mechanism is taken into account or not, versus the number of algorithm's iterations. It's straightforward to see that this kind of algorithm converges in very few steps, moreover, as expected, pricing mechanisms makes the final equilibrium more efficient than the one that would be obtained in a purely competitive manner.


Figure 28. Sum rate of ten FAPs vs. iterations with and without pricing.

### 7.3.3 Min-power game subject to Markovian interference

Let us consider now the generalization of the min-power game to the Markovian interference case. The utility of each FAP is now the total transmit power over $N$ subchannels and the $M$ time slots:

$$
u_{q}\left(\mathbf{p}_{q}\right)=\frac{1}{M} \sum_{k=1}^{N} \sum_{m=1}^{M} p_{k, m}^{q}
$$

while the constraint is that the expected rate, conditioned to the observation on the first time slot, has to be no smaller than a given value, i.e.

$$
\bar{R}_{q}\left(\mathbf{p}_{q}, \mathbf{p}_{-q}\right)=\frac{1}{M} \sum_{m=1}^{M} \sum_{k=1}^{N}\left[\beta_{k, m} \log \left(1+p_{k, m}^{q} a_{n}^{q}(k, m)\right)+\gamma_{k, m} \log \left(1+p_{k, m}^{q} a_{I}^{q}(k, m)\right)\right]
$$

The feasible set becomes

$$
\widetilde{\mathcal{F}_{q}}\left(\mathbf{p}_{-q}\right)=\left\{\mathbf{p}_{-q} \in \mathbb{R}^{N M \times 1}: \bar{R}_{q}^{0}, 0 \leq p_{k, m}^{q} \leq p_{\max }^{q}(k), \forall k=1, \ldots, N, m=1, \ldots, M\right\}
$$

and the game is then

$$
\widetilde{\mathcal{G}}_{2}\left(\mathbf{p}_{-q}\right)=\left\{\Omega,\left\{\widetilde{\mathcal{F}}_{q}\left(\mathbf{p}_{-q}\right)\right\}_{q \in \Omega},\left\{u_{q}\left(\mathbf{p}_{q}\right)\right\}_{q \in \Omega}\right\}
$$

The optimal strategy for each FAP amounts to solving the following optimization problem

$$
\begin{array}{cc}
\min _{\mathbf{p}_{q}} & u_{q}\left(\mathbf{p}_{q}\right)  \tag{111}\\
\text { subject to } & \mathbf{p}_{q} \in \widetilde{\mathcal{F}}_{q}\left(\mathbf{p}_{-q}\right)
\end{array} \forall q \in \Omega
$$

where the set $\widetilde{\mathcal{F}}_{q}\left(\mathbf{p}_{-q}\right)$ given the power vector $\mathbf{p}_{-q}$ of the other FAPs, is a convex set. The minimization problem $\tilde{P}_{2}$ for each player q , given the strategies of the others, is then a convex optimization problem, since the objective function is a linear (then convex) function of $\mathbf{p}_{q}$ The best response of each FAP can be written in closed form $\mathbf{p}_{q}^{*}=\mathbf{g}\left(\mathbf{p}_{-q}\right)$, where

$$
\left[\mathbf{g}\left(\mathbf{p}_{-q}\right)\right]_{k, m}=\left[\frac{-b^{q}(k, m)+\sqrt{b^{q}(k, m)^{2}-4 a^{q}(k, m) c^{q}(k, m)}}{2 a^{q}(k, m)}\right]^{p_{\max }^{q}(k)}, \forall k=1, \ldots, N, m=1, \ldots, M
$$

with

$$
\begin{gathered}
a^{q}(k, m)=a_{n}^{q}(k, m) a_{I}^{q}(k, m) \\
b^{q}(k, m)=a_{n}^{q}(k, m)+a_{I}^{q}(k, m)-\lambda_{q} a_{n}^{q}(k, m) a_{I}^{q}(k, m), \\
c^{q}(k, m)=1-\lambda_{q}\left[a_{n}^{q}(k, m) \beta_{k, m}+a_{I}^{q}(k, m) \gamma_{k, m}\right]
\end{gathered}
$$

where the Lagrange multiplier $\lambda_{q}$ must satisfy the rate constraint $\bar{R}_{q}\left(\mathbf{p}_{q}, \mathbf{p}_{-q}\right)=R_{q}^{0}$ Since game $\widetilde{\mathcal{G}}_{2}$ is a Generalized Potential Game, the existence of a NE of the potential game can be proved directly by the existence of a maximum of the potential function $\Phi(\mathbf{p})=\sum_{q=1}^{Q} u_{q}\left(\mathbf{p}_{q}\right)$ on the feasible set of the game. In order to solve the game $\widetilde{\mathcal{G}}_{2}$ we have considered an iterative and distributed approach where at each step each FAP calculates its own best response in closed form given the strategies of all the other active FAPs as $\mathbf{p}_{q}^{*}=\mathbf{g}\left(\mathbf{p}_{-q}\right)$.
Similarly to what done for the max-rate game proposed in the previous section, we can now improve the efficiency of the Nash equilibrium towards the Pareto optimal boundary by incorporating some pricing mechanisms. To this end we can reformulate $\widetilde{\mathcal{G}}_{2}$ as

$$
\begin{array}{rcc}
\left.\min _{\mathbf{p}_{q}}\right) & u_{q}\left(\mathbf{p}_{q}\right)+\sum_{m=1}^{M} \sum_{k=1}^{N}\left(\sum_{r \in \mathcal{N}_{q}} \lambda_{r} \pi_{k, m}^{r}\left|H_{k}^{q r}\right|^{2}\right) p_{k, m}^{q} & \forall q \in \Omega  \tag{112}\\
\text { subject to } & \mathbf{p}_{q} \in \mathcal{F}_{q}\left(\mathbf{p}_{-q}\right) &
\end{array}
$$

where $\lambda_{r}$ is the Lagrangian multiplier associated to the rate constraint and the pricing coefficients are defined as in (103), i.e. $\pi_{k, m}^{r}:=-\frac{\partial \bar{R}_{r}(\mathbf{p})}{\partial I_{k, m}^{r}\left(\mathbf{p}_{-r}\right)}$ with $I_{k, m}^{r}\left(\mathbf{p}_{-r}\right):=\sum_{i \in \mathcal{N}_{r}} p_{k, m}^{i}\left|H_{k, m}^{i r}\right|^{2}$, the interference received by $r$-th FAP on the resource block $(k, m)$ from its neighboring FAPs. We can see that price coefficients are introduced in order to quantify the damage that each FAP's action can induce on the other FAPs utilities by including, in each player's payoff function, some kind of care about a socially meaningful performance parameter.

Then the proposed power allocation strategy for the FAPs can be resumed as follows:
a. Each FAP observes the set of channels during slot 0 and based on this observation is able to compute the occupancy probabilities for all the slots.
b. Each FAP evaluates the price coefficients $\pi_{k, m}^{q}$ and the Lagrangian multiplier $\lambda_{q}$ and broadcast them to its neighbors.
c. The (time-frequency) NM-dimensional power allocation vector is optimized by solving problem $\tilde{P}_{3}$.
The best response of each FAP leads to the optimal powers $p_{k, m}^{q}$ which must satisfy the following equation

$$
\begin{equation*}
\tilde{c}_{1}^{q}(k, m)\left(p_{k, m}^{q}\right)^{2}+\tilde{c}_{2}^{q}(k, m) p_{k, m}+\tilde{c}_{3}^{q}(k, m)=0 \tag{113}
\end{equation*}
$$

where, denoting with $v_{q}$ the Lagrange multiplier, we have set

$$
\begin{align*}
& \tilde{c}_{1}^{q}(k, m)=\left(1+\sum_{r \in \mathcal{N}_{q}} \lambda_{r} \pi_{k, m}^{r}\left|H_{k}^{q r}\right|^{2}\right) a_{n}^{q}(k, m) a_{I}^{q}(k, m) \\
& \tilde{c}_{2}^{q}(k, m)=\left(1+\sum_{r \in \mathcal{N}_{q}} \lambda_{r} \pi_{k, m}^{r}\left|H_{k}^{q r}\right|^{2}\right)\left(a_{n}^{q}(k, m)+a_{I}^{q}(k, m)\right)-v_{q} a_{n}^{q}(k, m) a_{I}^{q}(k, m) ;  \tag{114}\\
& \tilde{c}_{3}^{q}(k, m)=1+\sum_{r \in \mathcal{N}_{q}} \lambda_{r} \pi_{k, m}^{r}\left|H_{k}^{q r}\right|^{2}-v_{q}\left(a_{n}^{q}(k, m) \beta_{k, m}+a_{I}^{q}(k, m) \gamma_{k, m}\right)
\end{align*}
$$

We can verify that, $\forall v_{q}>0$, we have $\tilde{c}_{2}{ }^{q}(k, m)^{2}-4 \tilde{c}_{1}{ }^{q}(k, m) \tilde{c}_{3}{ }^{q}(k, m) \geq 0$, and the only solution is

$$
\begin{equation*}
\tilde{p}_{k, m}^{c}=\frac{-\tilde{c}_{2}^{q}(k, m)+\sqrt{\tilde{c}_{2}^{q}(k, m)^{2}-4 \tilde{c}_{1}^{q}(k, m) \tilde{c}_{3}^{q}(k, m)}}{2 \tilde{c}_{1}^{q}(k, m)} \tag{115}
\end{equation*}
$$

More specifically, we get

$$
\begin{equation*}
p_{k, m}^{q}=\max \left\{0, \tilde{p}_{k, m}^{c}\right\} \tag{116}
\end{equation*}
$$

and the optimal power allocation vector is $p_{k, m}^{q^{*}}=\left[\tilde{p}_{k, m}^{b}\right]_{0}^{p_{q}^{\max }(k)}$ where the multiplier $v_{q}$ is chosen in order to satisfy the rate constraint $\bar{R}_{q}\left(\mathbf{p}^{*}\right)=R_{q}^{0}$.

### 7.4 Distributed Stochastic Pricing for Sum-Rate Maximization with Random Graph and Quantized Communications

### 7.4.1 Preliminaries

We now consider a pricing mechanism aimed at maximizing the sum-rate of a femtocell network in a distributed manner, thanks to a limited exchange of information among neighbor femto access points (FAPs), in the particular case in which the exchange of information among FAPs is quantized and happens through a network graph (typically a sparse graph), whose links fail randomly across iterations. The backhaul link among the FAPs is an IP-based internet connection, which delivers packets in the network using a best-effort protocol. Hence, control packets sent through the backhaul might experience large delays, because of retransmissions of packets corrupted by errors. This random delay and the associated delay jitter could jeopardize the potential benefits of coordination. It is then of interest to examine a protocol that simply discards packets that are received with a delay exceeding a given threshold.

We analyze the effect of this protocol by modelling the graph describing the interaction among FAPs as a random graph, where each link is on with a probability equal to the probability that the packet is correctly delivered within the given maximum delay. Furthermore, we take into account the quantization of the information exchanged among the FAPs. These sources of randomness introduce stochastic noise in the pricing mechanism that needs to be handled to insure convergence of the distributed algorithm. Using results from stochastic approximation theory, we propose a distributed projection based Robbins-Monro (RM) [Robbins51] scheme that converges almost surely (a.s.) on a final allocation equilibrium dependent on the mean graph of the network, even in the presence of such imperfect communication scenario.
Numerical results show how the system performance is affected by link failures, in fact, reducing the probability to establish a communication link among FAPs, the system performance decreases due to the lower coordination to mitigate interference. Nevertheless, supposing to know the probability with which each link fails, we show how to counteract the effect of random links through a proper weighting of the price coefficients coming from the neighbors. The main assumption is that, even under busy backhaul conditions, FAPs are able to exchange data on the backhaul link, with a given probability.
Furthermore, in practice, given the limited transmit power and the attenuation resulting from indoor propagation, only nearby FAPs interfere with each other. Then, in this distributed pricing mechanism, each FAP needs to exchange interference prices only with few neighbors, thus remarkably reducing the signaling in the network.

### 7.4.2 System model

We consider an OFDMA (according to LTE-A) wireless system with $Q$ distinct pairs of transmitters and receivers, sharing the same physical resources, e.g., time and bandwidth. No multiplexing strategy is imposed a priori so that, in principle, each FAP interferes with each other. In practice, interference is limited only between nearby FAPs. Then, to study the interference mechanism, it is useful to introduce what we call the interference graph $G_{I}=\left(V_{I}, E_{I}\right)$, defined as the graph whose vertices are FAPs and where there is an edge between two vertices only if the relative FAPs interfere with each other.

Typically, the interference graph is a sparse graph (i.e., with a number of links much smaller than the maximum number of possible links). We denote by $\mathcal{N}_{q}^{i}$ the set of interfering neighbors of user $q$. We also introduce the notation $\mathbf{p}=\operatorname{col}\left\{\mathbf{p}_{q}\right\}_{q=1}^{Q}$, where $\mathbf{p}_{q}=\left(p_{1}^{q}, \ldots, p_{N}^{q}\right)$ denotes the power vector of FAP $q$, whose element $p_{m}^{q}$ is the power transmitted by node $q$ over the $m$-th subcarrier. Under the previous assumptions, resorting to Shannon capacity expression, the rate of FAP $q$ is given by

$$
\begin{equation*}
R_{q}(\mathbf{p})=\sum_{m=1}^{N} \log \left(1+\frac{\left|H_{m}^{q q}\right|^{2} p_{m}^{q}}{\sigma_{q m}^{2}+\sum_{r \in \mathcal{N}_{q}^{N}}\left|H_{m}^{r q}\right|^{2} p_{m}^{r}}\right), \tag{117}
\end{equation*}
$$

where $H_{m}^{x y}$ is the channel transfer function of the $x$-th subchannel between the $y$-th transmitter and the $q$-th receiver, and $\sigma_{q m}^{2}$ is the variance (power) received on the $m$-th subchannel including receiver noise and power coming from the macro users, while $\sum_{r \in \mathcal{N}_{q}^{t}}\left|H_{m}^{r q}\right|^{2} p_{m}^{r}$ is the interference received from the $q$-th FAP from its interfering neighbors ( $\mathcal{N}_{q}^{i}$ ) on the $m$-th subchannel.

We are interested in a cooperative approach where FAPs pursue a common social objective. In particular, the optimization problem we would like to solve is the maximization of the sum rate of the $Q$ FAPs under power constraints, i.e.,

$$
\begin{array}{ll}
\max _{\mathbf{p}} & \sum_{q=1}^{Q} R_{q}(\mathbf{p})  \tag{118}\\
\text { s.t. } & \mathbf{p} \in \mathcal{P}
\end{array}
$$

where $\mathcal{P}=\prod_{q=1}^{Q} \mathcal{P}_{q}$ and $\mathcal{P}_{q}=\left\{\mathbf{p}_{q} \in \mathbb{R}^{N}: \sum_{m=1}^{N} p_{m}^{q} \leq P_{q}, \quad 0 \leq p_{m}^{q} \leq p_{q}^{m a x}(m), \quad m=1, \ldots, N\right\}$, with $P_{q}$ and $p_{q}^{m a x}(m)$ denoting, respectively, the power budget of user $q$ and the mask constraint that limits the maximum transmit power over each channel. In general, the objective function in (118) is not concave in the power allocation $\mathbf{p}$ and, as a consequence, the problem may have multiple local optima. A local solution can be found in a centralized manner using standard optimization algorithms.

However, we focus on distributed solutions where it is allowed a local coordination among FAPs through a limited exchange of data. In this work, we take into account the presence of an imperfect communication scenario, considering the following assumptions on the stochastic processes affecting the algorithm.

## 1. Random link failure model:

The exchange of data is described by a communication graph $G_{c}=\left(V_{c}, E_{c}\right)$, i.e., the graph whose vertices are FAPs and where there is an edge between two vertices only if the relative FAPs exchange data with each other. An example of communication graph for 20 FAPs is shown in Figure 29, where the blue dots represent the FAPs, and the red line the IP-based backhaul link. We denote by $\mathbf{A}=\left\{\mathrm{a}_{q r}\right\}$ the adjacency matrix of graph $\mathcal{G}_{c}$, whose nonnegative $a_{q r}$ entries are either one or zero, depending on whether there is a link between nodes $q$ and $r$. In our intended application, each FAP needs to exchange data only with neighbor FAPs in order to manage the Multi User Interference (MUI) generated by the active links. In an ideal communication scenario, the set of neighbors of FAP q in the communication graph coincides with the set of interfering FAPs in the interference graph.
Nevertheless, in a realistic communication scenario, some packets may be lost at random times. Typically, erroneous packets have to be resent and this induces a further complexity to handle packet retransmission, besides increased risks of congestion. It is then of interest to examine what happens if the erroneous packets are simply dropped, without requiring retransmissions. Random packet drops may be analyzed by modeling the network as a time-varying, or switched, network, where the presence of an edge (link) depends on the packet error rate. The communication network at time $k$ is modeled as an undirected graph, $G_{c}[k]=\left(V_{c}, E_{c}[k]\right)$ and the graph adjacency matrices as a sequence of independent identically distributed (i.i.d.) matrices $\{\mathbf{A}[k]\}$. We model the graph adjacency matrices as

$$
\begin{equation*}
\mathrm{A}[k]=\overline{\mathrm{A}}+\tilde{\mathrm{A}}[k], \tag{119}
\end{equation*}
$$

where $\widetilde{\mathrm{A}}[k]$ is a zero-mean sequence of independent identically distributed matrices, and $\overline{\mathbf{A}}=\mathbb{E}[\mathbf{A}[k]]$ denotes the adjacency matrix of the expected graph.


Figure 29. Example of communication graph.

## 2. Dithered quantization:

We assume that each inter-node communication channel uses a uniform quantizer, whose input-output relation may be modelled by the quantizing function, $q(\cdot): \mathbb{R} \rightarrow \mathbb{Q}$ with quantization step $\Delta>0$. Conditioned on the input, the quantization error is deterministic, influencing the convergence of the algorithm. Now, adding to the input $y[k]$ a dither sequence $\{d[k]\}_{k \geq 0}$ of i.i.d. uniformly distributed random variables on $[-\Delta / 2, \Delta / 2$ ) independent of the input sequence, the resultant error sequence $\{e[k]\}_{k \geq 0}$ becomes

$$
\begin{equation*}
e[k]=q(y[k]+d[k])-(y[k]+d[k]) . \tag{120}
\end{equation*}
$$

The sequence $\{e[k]\}_{k \geq 0}$ is now an i.i.d. sequence of uniformly distributed random variables on $[-\Delta / 2, \Delta / 2)$, which is independent of the input sequence. In the following, we will study the effect of the random communication graph and of the price quantization on the gradient of the sum-rate. Introducing the notation $R(\mathbf{p})=\sum_{q=1}^{Q} R_{q}(\mathbf{p})$, that's the sum of the rates of all the $Q$ FAPs that are randomly placed in a given area A , the partial derivative of the sum-rate $R(\mathbf{p})$ with respect to $p_{m}^{q}$ can be written as

$$
\begin{equation*}
\frac{\partial R(\mathbf{p})}{\partial p_{m}^{q}}=\frac{\partial R_{q}(\mathbf{p})}{\partial p_{m}^{q}}+\sum_{r \in \mathcal{N}_{q}^{i}} \frac{\partial R_{r}(\mathbf{p})}{\partial p_{m}^{q}} . \tag{121}
\end{equation*}
$$

Following the same steps as in [Huang06], [Shi08], it is useful to introduce the price coefficients as:

$$
\begin{equation*}
\pi_{m}^{r}(\mathbf{p})=-\frac{\partial R_{r}(\mathbf{p})}{\partial I_{m}^{r}} \tag{122}
\end{equation*}
$$

with $I_{m}^{r}=\left.\sum_{s \in \mathcal{N}_{r}^{r}} H_{m}^{s r}\right|^{2} p_{m}^{r}$ the interference received from the $r$-th receiver on the $m$-th subchannel.

The price $\pi_{m}^{r}(\mathbf{p})$ is always nonnegative and is proportional to the marginal decrease of user $r^{\prime}$ s rate because of an increase of the $q$-th node's transmit power, as:

$$
\begin{equation*}
\frac{\partial R_{r}(\mathbf{p})}{\partial p_{m}^{q}}=\frac{\partial R_{r}(\mathbf{p})}{\partial I_{m}^{r}} \frac{\partial I_{m}^{r}}{\partial p_{m}^{q}}=-\pi_{m}^{r}(\mathbf{p}) \frac{\partial I_{m}^{r}}{\partial p_{m}^{q}}=-\pi_{m}^{r}(\mathbf{p})\left|H_{m}^{q r}\right|^{2} \tag{123}
\end{equation*}
$$

Then, substituting expression (124) in (122), we get

$$
\begin{equation*}
\frac{\partial R(\mathbf{p})}{\partial p_{m}^{q}}=\frac{\partial R_{q}(\mathbf{p})}{\partial p_{m}^{q}}-\sum_{r=1}^{Q} a_{q r}\left|H_{m}^{q r}\right|^{2} \pi_{m}^{r}(\mathbf{p}) . \tag{124}
\end{equation*}
$$

To evaluate expression (124), FAP $q$ needs to know the price vectors $\boldsymbol{\pi}^{r}(\mathbf{p})=\operatorname{col}\left\{\pi_{m}^{r}\right\}_{m=1}^{N} \in \mathbb{R}^{N}$, $\forall r \in \mathcal{N}_{q}^{i}$, which are transmitted by the neighbors through the backhaul link. The communication graph is then present in (124) through the coefficients $a_{q r}$ of the adjacency matrix A. Collecting all the contributions in (124), the gradient of the sum-rate with respect to $\mathbf{p}$ can be written in compact vector form as

$$
\begin{equation*}
\nabla_{\mathbf{p}} R(\mathbf{p})=\mathbf{r}(\mathbf{p})-\mathbf{A}_{H} \boldsymbol{\pi}(\mathbf{p}) \tag{125}
\end{equation*}
$$

where $\mathbf{r}(\mathbf{p})=\operatorname{col}\left[\nabla_{\mathbf{p}_{q}} R_{q}(\mathbf{p})\right]_{q=1}^{Q} \in \mathbb{R}^{Q N}, \boldsymbol{\pi}(\mathbf{p})=\left[\boldsymbol{\pi}_{1}(\mathbf{p}), \ldots, \boldsymbol{\pi}_{Q}(\mathbf{p})\right] \in \mathbb{R}^{Q N}$ is the vector collecting the interference prices of the entire network and $\mathbf{A}_{H} \in \mathbb{R}^{Q N} \times \mathbb{R}^{Q N}$ is a multidimensional adjacency matrix weighted by the cross channels between FAPs. Expression (125) shows the dependence of the sumrate gradient with respect to the network graph in the case of an ideal communication scenario. Considering the presence of failures and dithered quantization noise, expression (124) evaluated at time $k$, takes the form

$$
\begin{equation*}
\frac{\partial R(\mathbf{p}[k])}{\partial p_{m}^{q}}=\frac{\partial R_{q}(\mathbf{p}[k])}{\partial p_{m}^{q}}+\sum_{r=1}^{Q} a_{q r}[k]\left|H_{m}^{q r}\right|^{2}\left(\pi_{m}^{r}(\mathbf{p}[k])+v_{m}^{q r}[k]+\epsilon_{m}^{q r}[k]\right), m=1, \ldots, N, q=1, \ldots, Q, \tag{126}
\end{equation*}
$$

where $v_{m}^{q r}[k]$ and $\epsilon_{m}^{q r}[k]$ are contributions of dithered quantization noise that the $q$-th FAP receives from the $r$-th FAP on the $m$-th subchannel, at time $k$. To rewrite (126) in compact form, we introduce the random vectors $Y[k]=\operatorname{col}\left\{Y_{q}[k]\right\}_{q=1}^{Q} \in \mathbb{R}^{Q N}$ and $\Psi[k]=\operatorname{col}\left\{\Psi_{q}[k]\right\}_{q=1}^{Q} \in \mathbb{R}^{Q N}$, which are aggregated contribution of dithered quantization noise. Then, the sum-rate gradient with respect to the power allocation $\mathbf{p}[k]$, in the presence of random link failures and dithered quantization, can be written in compact form as

$$
\begin{equation*}
\nabla_{\mathbf{p}} R(\mathbf{p}[k])=\mathbf{r}(\mathbf{p}[k])-\mathbf{A}_{H}[k] \boldsymbol{\pi}(\mathbf{p}[k])+Y[k]+\Psi[k] . \tag{127}
\end{equation*}
$$

Now, expanding the multidimensional weighted adjacency matrix $\mathbf{A}_{H}[k]$ as in (119), we can write expression (127) as the sum of a deterministic function $\mathbf{F}(\mathbf{p}[k]) \in \mathbb{R}^{Q N}$ plus a random function $\Gamma(\mathbf{p}[k]) \in \mathbb{R}^{Q N}$, where

$$
\begin{gather*}
\mathbf{F}(\mathbf{p}[k])=\mathbf{r}(\mathbf{p}[k])-\overline{\mathbf{A}}_{H}[k] \boldsymbol{\pi}(\mathbf{p}[k]),  \tag{128}\\
\Gamma(\mathbf{p}[k])=-\widetilde{\mathbf{A}}_{H}[k] \boldsymbol{\pi}(\mathbf{p}[k])+Y[k]+\Psi[k] . \tag{129}
\end{gather*}
$$

The problem in (118) can then be converted in the search, inside the feasible set $\mathcal{P}$, for the zeros of a deterministic function $\mathbf{F}(\mathbf{p})$ whose value, measurable at each time instant $k$, is corrupted by an additive random disturbance $\Gamma(\mathbf{p}[k])$.

Document number: D3.2
Title of deliverable: Interference coordination protocols in femto-based networks
7.4.3 Distributed Stochastic Pricing Algorithm (DSPA)

To find a solution of the problem (118) affected by random disturbances, it is useful to introduce stochastic approximation algorithm. In the remainder of this section we introduce a stochastic approximation scheme for solving the problem in (118) in a distributed manner. In particular, we consider a projection-based Robbins-Monro (RM) stochastic approximation procedure, taking in consideration a simultaneous update of the users' power profiles and providing supporting convergence results. The problem is amenable for distributed solutions because the optimization set $\mathcal{P}=\prod_{q=1}^{Q} \mathcal{P}_{q}$ is given by the Cartesian product of sets $\mathcal{P}_{q}$, allowing the parallel computation of the algorithm. Let us now focus on a distributed stochastic simultaneous solution. We call it Distributed Stochastic Pricing Algorithm (DSPA) which is summarized as follows:

```
Algorithm: Distributed Stochastic Pricing
S.0: Each FAP q chooses an initial power profile }\mp@subsup{\mathbf{p}}{q}{}[0] satisfying th
power constraint;
S.1: Using (122), each FAP computes the interference price vector }\mp@subsup{\boldsymbol{\pi}}{q}{}\mathrm{ ,
given the current power profiles, and transmits this to the neighbor
FAPs;
S.2: At each time k, every FAP, given the neighbors' power profiles
and price vectors, simultaneously updates its power profile according
to
```

```
    \mp@subsup{\mathbf{p}}{q}{}[\textrm{k}+1]=[\mp@subsup{\mathbf{p}}{q}{}[\textrm{k}]+\alpha[\textrm{k}]\cdot\mp@subsup{\nabla}{\mp@subsup{\mathbf{p}}{q}{}}{}\textrm{R}(\mathbf{p}[\textrm{k}])\mp@subsup{]}{\mp@subsup{\mathcal{P}}{q}{}}{}=\mp@subsup{\mathbf{T}}{q}{}(\mathbf{p}[\textrm{k}]),
```

    \mp@subsup{\mathbf{p}}{q}{}[\textrm{k}+1]=[\mp@subsup{\mathbf{p}}{q}{}[\textrm{k}]+\alpha[\textrm{k}]\cdot\mp@subsup{\nabla}{\mp@subsup{\mathbf{p}}{q}{}}{}\textrm{R}(\mathbf{p}[\textrm{k}])\mp@subsup{]}{\mp@subsup{\mathcal{P}}{q}{}}{}=\mp@subsup{\mathbf{T}}{q}{}(\mathbf{p}[\textrm{k}]),
    with q=1,···,Q and k\geq0, where [ [.] 疎 denotes the projection over the
with q=1,···,Q and k\geq0, where [ [.] 疎 denotes the projection over the
feasible set }\mp@subsup{\mathcal{P}}{q}{},\boldsymbol{\alpha}[k] is an iteration-dependent step size and
feasible set }\mp@subsup{\mathcal{P}}{q}{},\boldsymbol{\alpha}[k] is an iteration-dependent step size and
\nabla}\mp@subsup{\mathbf{p}}{q}{}R(\mathbf{p}[k])\in\mp@subsup{\mathbb{R}}{}{N}\mathrm{ is the }q\mathrm{ -th vector element of (127);
\nabla}\mp@subsup{\mathbf{p}}{q}{}R(\mathbf{p}[k])\in\mp@subsup{\mathbb{R}}{}{N}\mathrm{ is the }q\mathrm{ -th vector element of (127);
S.3: Go to step 2 and repeat until a convergence criterion is
S.3: Go to step 2 and repeat until a convergence criterion is
satisfied.

```
satisfied.
```

Table 17. DSPA.

We consider two assumptions on the stochastic procedure in (130).

1. Persistence: The step size sequence $\alpha[k]$ satisfies:

$$
\begin{equation*}
\alpha[k]>0, \quad \sum_{k=0}^{\infty} \alpha[k]=\infty, \quad \sum_{k=0}^{\infty} \alpha^{2}[k]<\infty \tag{131}
\end{equation*}
$$

Condition (131) ensures that the step-size decays to zero, but not too fast. It is standard in stochastic adaptive signal processing and control;
2. Independence: The stochastic sequences $\{\widetilde{\mathbf{A}}[k]\}_{k \geq 0},\{\epsilon[k]\}_{k \geq 0}$ and $\{\nu[k]\}_{k \geq 0}$ are mutually independent.

Next, we provide the convergence result for the distributed stochastic algorithm in (130). Let $\{\mathbf{p}[k]\}$ the sequence generated by the distributed stochastic pricing algorithm in (130), with step-size satisfying the conditions in (131). Then, the sum rate sequence $\{R(\mathbf{p}[k])\}_{k \geq 0}$ converges a.s. to a finite value $R^{*}$, i.e.

$$
\begin{equation*}
\operatorname{Prob}\left[\lim _{k \rightarrow \infty} R(\mathbf{p}[k])=R^{*}\right]=1, \tag{132}
\end{equation*}
$$

where $\operatorname{Prob}[\mathcal{E}]$ denotes the probability of the event $\mathcal{E}$. Furthermore, let $\mathbf{p}^{*}$ be an accumulation point of the sequence $\{\mathbf{p}[k]\}$, as $k \rightarrow \infty$, the optimal solution $\mathbf{p}^{*}$ is a fixed point of the mapping $\mathbf{T}(\mathbf{p})=\operatorname{col}\left\{\mathbf{T}_{q}(\mathbf{p})\right\}_{q=1}^{Q}$, such that $\mathbf{p}^{*}=\mathbf{T}\left(\mathbf{p}^{*}\right)$. The proof can be found in [DiLorenzo11].

### 7.4.4 Numerical results

In this section we provide some numerical results to validate the theoretical findings and to assess the performance of the proposed algorithm. We assume the presence of 20 FAPs randomly placed over an area A, interfering (and communicating) according to the topology graph depicted in Figure 29. Because of the randomness introduced by the control channel, a link between two neighbors has a certain probability $p_{c}$ to be established correctly. The values to be exchanged are also affected by dithered quantization noise, supposing the presence of a 6 bit mapping. We consider a number of subchannels $N=12$ and Rayleigh frequency selective fading channels, with channel order $L_{C}=4$. FAP maximum transmit power is 20 dBm and rates are expressed in bit per OFDM symbol [bps]. In Figure 30 we show a numerical example of sum-rate behavior as a function of the iteration index, considering different values of probability $p_{c}$.


Figure 30. Sum rate vs. iteration index, for different probabilities of link failures.

The ideal case corresponds to $p_{c}=1$ and it is shown as a benchmark. We also report the case $p_{c}=0$, corresponding to the behavior of the classical iterative water filling algorithm (IWFA) without pricing. In particular, considering two intermediate values $p_{c}=0.6$ and $p_{c}=0.3$, we compare the ideal behaviors obtained having a mean graph dependent on these probability values and the average behaviors, averaged over 500 independent realizations, given by the algorithm in (130) in the presence
of random links and quantization noise. As we can notice, the stochastic algorithm converges to the same equilibrium of the correspondent ideal case evaluated for the expected graph. The effect of the diminishing step size in (131) is to reduce the convergence speed of the algorithm with respect to the correspondent ideal case.

As expected, link failures determine a performance degradation due to the lower coordination among FAPs to mitigate interference. However, from theoretical results, validated in the previous numerical example, we know that the final convergence value of the stochastic algorithm depends on the expected graph of the network. Hence, assuming that each FAP knows, through preliminary estimation, the probability $p_{c}$ to establish a communication over each link, it is possible to counteract the effect of the graph randomness by weighting the price coefficients coming from each link with the inverse of the probability $p_{c}$. In this way, we are "normalizing" the mean network graph in order to be coincident with the ideal graph in the case of absence of failures.
Considering the same settings of the previous simulation, in Figure 31 we show the behavior of the sum-rate as a function of the iteration index. In particular, we report the ideal case correspondent to $p_{c}=1$, as a benchmark, compared to the average behaviors, averaged over 500 independent realizations, given by the compensated stochastic algorithm in (130) in the presence of quantization noise and for different probabilities to establish a communication link. As we can notice, thanks to the compensation, the final equilibrium value of the stochastic algorithm coincides with the ideal case for every value of probability $p_{c}$. However, reducing the probability to establish a communication link, the network requires more time to reach the final equilibrium state.


Figure 31. Sum rate vs. iteration index, for different probabilities of link failures.

### 7.5 Conclusions

In this section we have considered different aspects of the resource allocation problem among a single or a set of FAPs operating in the same band as a Macro Base Station in a multi-carrier system setup. First, we addressed the problem of maximizing the opportunistic throughput while maintaining a prescribed maximum interference level towards the macro users: we studied an algorithm based on collaborative spectrum sensing and showed that the system gets a real benefit from a joint
optimization of the false alarm rate and the power allocation over the set of potentially available subbands. This approach only requires the knowledge of the channel from the FAP to its intended receiver. The proposed solution is shown to be a multi-level water-filling when we add to the classical transmit power constraint the constraint on the maximum allowed harmful interference to the MBSs. In particular, we showed that the FAP tends to allocate more powers over the subbands where not only the channel is stronger but also the probability of a correct decision is higher. We have shown how the rate gain with respect to the case where $p_{f a}$ is not optimized becomes more and more evident as the FAP is forced to limit its maximum interference level towards MBSs.

For this particular problem we have considered a single FAP, so that the only constraint is the interference towards MBSs. Furthermore, we have considered decentralized resource allocation strategies based on sensing at the local level to determine channel status and interference: we have proposed alternative game-theoretic techniques that exploit the backhaul link among FAPs to set up local coordination games which provide performance improvement with respect to purely competitive games. The most important contribution with respect to algorithms already present in the literature is the inclusion, in the optimization strategy, of a Markovian model of the interference activity while using alternative pricing mechanism to allocate power in the joint time-frequency plane. Since conventional algorithms require knowledge of the actual channel and interference status, they require to perform spectrum sensing and to run the algorithm with the same frequency with which the channel occupation of the Macro users varies.

In contrast, our approach allows to reduce such frequency, thus leaving more resources available for data transmission. The use of pricing appears as an interesting solution in femtocell networks where the exchange of information among nearby FAPs can occur through the backhaul wired link. The exchange of information entails the transmission of only a few data, the so called prices coefficients, so that no significant rate loss due to signalling occurs.
Finally, we have studied a decentralized (gradient projection) stochastic pricing algorithm aimed to allocate power in OFDMA based femtocell networks where the communications among FAPs are affected by random link failures and quantization noise. We have proposed a projection-based RM stochastic approximation scheme that converges almost surely to a final allocation dependent on the mean graph of the network. Numerical results show how the network performance is affected by link failures. Assuming to know the failure probability over each link, we have shown how to counteract the effect of random links, thus making the allocation algorithm robust to channel imperfections, whose effect is only to slow down the convergence process.

## 8 CENTRALIZED DYNAMIC INTERFERENCE MANAGEMENT

A mobile network operator can implement a policy for the allocation of the transmission parameters, chosen at central level by adopting a dedicated unit/processor linked to the RRM. The transmission power and frequency band of each unit can be assigned by adopting a strategy targeting (for instance) at minimising interference and maximising system capacity. To this purpose, here we investigate the search of sub-optimal transmission parameters in a dense deployment of FAPs, based on Genetic Optimization Algorithm (GO).

Since the optimization method is influenced by some parameters characterizing the deployment (e.g., number and position of buildings, number of floors), the algorithm has been tested in a variety of cases, in order to assess its performances and gain when compared to other methods of resources assignment, in terms of capacity improvement and convergence time.

The technical scenario presented in Figure 2 considers a deployment of several FAPs in a series of buildings, with varying number of floors (e.g., 3 to 6 ) for which was implemented a model of pathloss influenced by the internal structure of the buildings.

### 8.1 Preliminaries

The resources allocation for a generic deployment of FAPs requires specifying a set of parameters that can be measured and fed back from the single FAP to the RRM in order to tune the single allocations after a centralized optimization process.
The system can be constituted of a high number of FAPs (from several tens to hundreds) whose relative position is assumed unknown.
The proposed approach for the system optimization is based on the Global Optimization theory and Distributed Optimization techniques with major focus on Statistic Evolutionary approaches (based on Genetic Optimization and Multivariate Constrained Optimization), which are search methods well suited for finding suboptimal solutions of complex constrained problems with unknown objective function.

Before the development of the present activity, Genetic programming has been used in [Ho09] as a method to optimize the parameters maximizing coverage and minimizing pilot power signalling, for a slowly running algorithm (algorithm collecting data for 30 minutes of transmissions) and in other works to address planning of MBS [Kum02], [Meu00],[Lie98].
To our knowledge, Genetic Optimization (GO) has not been implemented so far to assign frequency and/or transmission power to manage/reduce interference among transmitters in a network of femtocells. In this activity we introduce the GO algorithm applied to a deployment of femtocells and address system performance, describing how it can be implemented by a Mobile Network Operator (MNO) in centralized manner. Currently deployed networks are managed from a central unit, limiting uncontrolled use of resources and possibly balancing traffic load depending on the type of data (e.g., voice vs. P2P).
Clusters of users can be figured out on the basis of their interference relation even in a homogenous deployment of transmitters (especially in an urban scenario, where buildings naturally break homogeneity) guaranteeing scalability for deployments over more than a single building. Clustering methods are not the purpose of this WP and will be addressed in [FREEDOM-D52]), considering for example [Ban03] for possible approaches.
Genetic Optimization has shown to be well suited for a wide class of multivariate problems, generally providing (sub)-optimal results for the minimization of a generic performance metric (objective
function), to be designed to represent the global property of the system (e.g. overall capacity and system fairness).
The adopted search method maximizes a performance metric (objective function) representing a global property of the system which, for our analysis, is expressed in terms of the single-user capacities and takes into account a principle of fairness
Simulations have been initially performed for limited (few units) sets of FAPs interfering in UL and then have been extended to a more populated scenario (e.g., some tens of FAPs). Scalability issues can be addressed by identifying clusters of FAPs and FUEs and adapting the optimization algorithms accordingly to the notion of clustering. In this study, the concept of clusters naturally emerges by the physical characteristics of transmission in a realistic environment and is mathematically reflected in the form of the pathloss matrix. Although forming groups from a large set of transmitter requires a certain level of abstraction, we outlined how such concept can be defined in some realistic deployment. That wok will be analyzed in [FREEDOM-D52].
The optimization algorithm collects the SNR (or the estimated capacity) sensed by each FUE or FAP (DL or UL, respectively) and outputs the power and RRA of each FAP or FUE that maximize the performance metrics (e.g. aggregate sum rate).


Figure 32. Scheme of the backhaul link activity for centralized RRM.

The analysis does not assume any interaction between FAPs and MBS. As depicted in Figure 32, the FAPs communicate to the RRM the sensed SNR through the backhaul link and the centralized algorithm feeds back the radio allocations (band and power, or power only) to the FAPs. Each FAP acts independently, excluding any direct FAP-FAP communication through the backhaul link. The algorithm developed so far considers the SISO scheme.
The optimization paradigm has been faced both on the design and on the simulation side. Starting by a geometric parameterization of a set of nearby buildings with several FAPs deployed inside, different propagation loss models have been implemented (i.e., no propagation loss, flat fading channel or frequency selective fading) to test the performances of the genetic algorithm (GA), by mimicking the internal structure of the buildings and its effect on a realistic scenario. The scenario and propagation properties are synthesized in the single parameter (the sensed SINR averaged over time, e.g. at least one frame duration) that each node, FAP and/or FUE, sends back to the central processor. The simulations were run for varying sets of genetic parameters (migration rate, generations, population size, iterations), finding the acceptable range of variability ensuring the sub-optimum results.

The system scalability has been investigated extending the algorithm architecture to take into account clustering/aggregation criteria to reduce the variables space, addressing some methods to outline how sub-sets of transmitters can be analyzed together reducing the computation load (as described in Section 10). The simulations have been employed both to confirm the theoretical results and to give indications about the tradeoff to be implemented.

### 8.2 Cell system optimization by Genetic Algorithm

### 8.2.1 Distributed or centralized implementation

Although presented as a centralized approach, the GO method can be implemented either in a centralized or a distributed manner by reproducing the optimization performed at the central processor unit (RRM in Figure 32) on each terminal. Such distributed solution is not suitable for scenarios for large number of users. For example, given a neighborhood of $N$ nodes, the decentralized solutions presented in sections 5, 6 and 7 derive algorithms where each node needs $(N-1)$ messages from its neighboring nodes. In contrast, the centralized implementation considered in this section requires of $N$ or $2 N$ messages ( $2 N$ if the GA optimizes both UL and DL, $N$ if only one traffic direction). In this view, $N$ is the number of nodes undergoing optimization that can be the overall number of users in the network or, in the cluster case, the number $N_{k}$ of nodes in a single cluster K. In the clustered scenario, interference among users in nearby clusters can be tamed or not, depending on optimization strategy. Also in that case, for moderate number of interfering users such algorithm can be taken into account. Centralized GO method also becomes an interesting solution when interfering nodes could be easily connected to certain a central controller, for instance those FAPs in the same building GW.

Centralized and distributed GO methods have pros and cons, depending on the business model of the MNO. When an operator deploys a large number of FAPs, interference among transmitters is strongly dumped after few hundred meters, therefore the optimization of the whole system of FAPs with a unique step would be a meaningless effort. The centralized case, coherently with the current commercial methodologies for traffic management, can be of interest for a MNO willing to maintain a centralized control of resources allocation and can provide subsets of users (namely grouped in clusters) for GO, for example on the basis of those linked to the same femto-gateway of an operator with a license in a certain area. In this case, all users have to send periodically their SNR to a central processor, giving rise to a maximum backhaul information overhead requirement per single link which can be estimated to be roughly around $3.2 \mathrm{Kbit} / \mathrm{s}$ : that is the worst case, with highest signaling rate, given by a feedback coded in 16 bits from the FAP to the RRM every frame ( 5 ms duration). In case of longer frame duration (e.g., 10 ms ) and implementation policies, such information overhead can be much less demanding. The optimization process is performed over a time scale according to the channel coherence time. The central unit performs one GO iteration and distributes optimized values to transmitters which use them until next SNR collection. The scheme is the following:
o Step 1. Every node sends its SNR to the central unit;
o Step 2. Central unit performs one step of GO (output is a population of $p$ individuals $I$ ). Every individual corresponds to a possible set of transmission parameters distributed to the nodes;
o Step 3. Every user adopts for a fixed number of frames the transmission parameters corresponding to all individuals $I$;
o Step 4. All SNRs sensed in correspondence of all realizations are fed back to the central processor;
o Step 5. Central unit generates a new population of $p$ individuals $J$ and distributes them.
In this implementation it is supposed that the central unit, nor the nodes, has no knowledge of channels among different nodes, thus it is necessary to realize all transmission with the parameters specified by the number of individuals of GO to collect their score. On the other side, if the system has implemented, by other resources, a method to be aware of channels and attenuation of signals among different nodes, this knowledge can be exploited to directly generate the new population at the core processor, thus speeding up the optimization. This could be gained, for example, by dedicating periodic and coordinated sensing of pilot tones among nodes, but is not relevant to the purposes of the present study. Indeed, such case would prevent the test of all individuals, in favour of a direct computation of all SNRs. The time necessary to get the final output is:
o $\mathrm{T}=2 * \mathrm{~T}$ (frame) x NGenerations x NIndividuals [effective realization of transmissions]
o $\mathrm{T}=$ computing time [prior knowledge of channels, depends on processor time only]
The resulting effectiveness of the GO convergence time must be benchmarked against the requirements at application level, e.g. against the most delay-sensitive applications. The users receive the assigned parameters, during the following transmission events perform again a SNR measurement and feed back again such values which are used as an input for the next iteration of the GO in the centralized implementation. What here is called a step or iteration of GO corresponds to a "generation", in GO language (see section 8.2.3). The sequence of computations proceeds until a stop criterion is fulfilled or can continue adapting to varying environment conditions due to channel variability or mobility.

In the distributed implementation, the GO must run on each transmitter. Similarly to the centralized case, each user has to collect the SNR of the nearby interfering units involved in the optimization, adopt such set of values as input for the GO computation and extract from the output the parameters values to use for the following transmissions. Each node has to collect N-1 SNRs from nearby nodes. How this information is encoded by the GO is explained in section 8.2.3. In the method is implicit how the computation should involve only transmitters in range. All processors get the same output (encoding resulting transmission parameters also for the other transmitters) provided that the seed adopted for (pseudo)-random generators is the same, which can be gained by the same initialization. The distributed option for a massive FAP density prescribes that GO is implemented on each set of first neighbors, thus leading to the emerging idea of clustering intrinsic to the selection of users whose mutual interference requires mitigation.

In both implementations the role of signaling (UL only) could in principle raise routing issues at local level for exchange of information (discussed in section 8.2.3).), in terms of overhead on the backhaul link and of maximum tolerable delay, thus in absence of optimized routing methods (to date not yet implemented) the required constraints on maximum delay for UL and DL could be a drawback.

Finally, as for other distributed approaches, GO could have a difficult implementation in case of many users interfering with two (or more) disjoint groups. The algorithm presented here has the same performances in terms of resources assignments in both cases (centralized or distributed). System scalability has been investigated by identifying clusters of FAPs (or FUEs) and by extending the algorithm architecture to account for clustering/aggregation criteria to reduce the variables space, by addressing methods to outline how to analyze sub-sets of transmitters, instead of the whole network in a bunch, for reducing the computation load. Such issues regard topics out of scope of the GO algorithm procedure presented here.

### 8.2.2 Working of the GA implementation

Given the such design for the optimisation procedure, the GA is a sub-optimal implementation, very simple, scalable and modular, with basically no computational burden (each cycle is just a reassignment of variables).

The employed terminology relies on the concept of:

- individual: one individual is the vector of the radio resources; for instance, assuming M FAPFUE couples and optimization is towards frequency assignments, one individual could be a 2 xM -elements vector with the starting frequency and bandwidth of each FAP-FUE couple;
- population: it is an ensemble of P individuals; for example, a population is a "guess" of P possible vectors of frequency assignments;
- generations: number representing how many times a population undergoes a selection, mating and reproduction process; for example, it is a measure of how many times the algorithm has to evaluate an ensemble of frequency allocations, measure a score (given by the optimisation metric) and, based on this score, operate a selection and mating to produce a new population.

From an implementation point of view, the complete procedure is described below for the generic $\mathrm{k}^{\text {th }}$ generation:

1. The population computed at generation k-1 is sent to the FAPs; this means that each FAPFUE couple is told to employ in the next P (population size) frames, P different frequency allocations; this process will indeed take $P$ frames to be completed.
2. At the end of each of the $P$ frames, each FAP and each FUE collects its mean SNR in the frame and sends it to the central processing unit; the delivery of these 2 NP values of the SNR will take (approximately) the time needed by P frames to be completed;
3. So, after P frames from the $\mathrm{k}-1$ generation, 2 N new SNR values for each one of the P individuals are available as input to the GA: the GA task now is to rank the score of the metric (solely depending on the input SNRs) against the P individuals and perform the GA operations of selection, mating, reproduction and mutation.
4. The output of the GA at the $\mathrm{k}^{\text {th }}$ generation is thus a new population and the process is repeated from step 1, until a stop criterion is met.

This scheme can be summarized as:

- The input of the GA is: 2 NP mean SNR values ( N mean SNR values from each FAP, N mean SNR values from each FUE, repeated P times to complete the whole P-dimensional population). The number N depends on the relevant scenario, whereas P is a GA processing parameter;
- The output of the GA at each generation is a population of P individuals (i.e. frequency/power allocations); when a stop criterion is met, the "best" individual of the latest generation is taken as "best" solution and the other P-1 individuals are disregarded;
- the initial guess of the P frequency allocations (population) should be as random as possible to better explore the solution space, unless some side or a-priori knowledge is available to restrict the search of the solution in a narrower space;
- the higher is the number of employed individuals P , the better the solution space is explored, but more time is needed to complete the cycle belonging to each individual;
- none insures that an (sub-)optimal solution is actually approached; when the GA internal parameters (e.g. mating function, mutation probability) and the metric to be optimised are not properly designed, the GA solution might be locked into a local (and unsatisfactory) minima;
- the GA convergence process is time consuming: assuming, for instance, a frame duration of $\mathrm{T}=1 \mathrm{~ms}$, a population $\mathrm{P}=50$ individuals and a convergence time of $\mathrm{G}=50$ generations, the time needed to output a (sub-)optimal frequency allocation is TGP $=2.5 \mathrm{~s}$. By increasing the population size P , the more diversity is given and the less generations are needed to attain a satisfactory solution, but the best balance between the two is not known a priori.


### 8.2.3 GO implementation for a set of FAPs

GO is based on the reduction of the problem to a population of individuals, each representing a realization of the system status. Individuals, selected on the basis of a fitness evaluation, reproduce by
crossover and mutation and finally terminate. Each step is characterized by a corresponding parameter (e.g., crossover order or mutation rate) whose values can be assigned or randomly chosen in an appropriate range at every step.

The problem of network optimization can be treated by GO by encoding the set of transmission parameters in terms of an individual, which is defined as a vector containing the set of parameters characterizing the system in the current status: for example, $N$ channel assignments and $N$ transmission powers. In our case, fitness is evaluated by measuring system capacity and the score is given by the average capacity per user. Thus, the population is constituted by a fixed number of individuals and each of them represents a trial for parameter assignment to every user before getting to the next generation of individuals, iterating the algorithm until termination is reached.

In the case of centralized implementation, incoming signaling towards the central unit (computing the optimization of the parameters) is given by the set of all SNRs sensed by users. Such an input is processed to get the transmission parameters to be used in the next frame and then broadcasted to the users involved by system optimization. In the distributed case, each unit undergoing optimization must receive as incoming signaling the sensed SNRs, where noise and interference are provided by the ensemble of users in range (i.e. first neighbors). The output of the computation contains the values of the transmission parameters for all users, thus each unit has to decode its own parameters and use them, giving no outgoing signaling traffic. Indeed, a unique initialization (including the seeds of the random sequences generators) allows each user to obtain the same overall GO output and therefore unique assignment of transmission parameters. In fact each user computes a step of GO (the reproduction of a set of individuals) obtaining an output population of individuals, each encoding a set of possible transmitting parameters to get values of the fitness function.

For GO evaluation, we will describe centralized processing, keeping in mind that the distributed one can be implemented with the same performances. It is assumed that each FUE or FAP (DL or UL, respectively) can deliver to the RRM the sensed SNR through the backhaul link, then the optimization feeds back to each entity the relevant radio allocations (band and power, or power only) that maximize the adopted performance metrics (e.g. aggregate sum rate). The system can be constituted of a high number of FAPs (from several tens to hundreds) whose relative locations are assumed unknown.

Nearby buildings, each with several FAPs within, are simulated by a geometric parameterization, implementing different propagation loss models (i.e., several realizations of flat fading channel) and mimicking building internal structure for a realistic scenario.

Convergence time of the algorithm depends on population size and number of generations needed to converge. Complexity of computations is very low due to the type of operations involved.

Pseudo-random parameters of GO are migration rate, generations, population size, and iterations, whose values don't affect complexity but only convergence effectiveness and time. GO for the considered scenarios involve a maximum 100 transmitters $(N)$ which is a reasonable number for simultaneous interferers. A high rate of changes within the search space ensures better GO performances and this parameter is given by the population size. GO performs satisfactorily for population size $P_{\text {size }}$ at least $N$ and number of generations $G$ at least $2 \times N$. System time is given by the overall number of realizations needed to evaluate the Genetic algorithm, given by $P_{\text {size }} \mathrm{X} G$.
As a benchmark, GO has been tested in scenarios where the optimal solution is known, giving solutions close to theoretical results.

### 8.2.4 GO formulation

Let $\vec{\theta}=\left\{\vartheta_{j}\right\}, \quad j=1, \ldots, n$ be a set of parameters $\vec{\theta} \in \Omega \subseteq \mathbb{R}^{n}$ where $\Omega$ is the space of variation for the single parameters values and be $f(\vec{\theta})$ a function of such parameters. The genetic algorithm is a search
method to find in the search space $\Sigma \subseteq \mathbb{R}^{n}$ where is defined as $f$, a set of values $\overrightarrow{\hat{\theta}}$ in the feasible region $\Phi \subseteq \Sigma$ such that the fitness function $\kappa=f(\overrightarrow{\hat{\theta}})$ is maximum. The search space $\Sigma$ is defined as a $n$-dimensional rectangle in $\mathbb{R}^{n}$ defined by the lower and upper bounds

$$
\begin{equation*}
l(i) \leq \vartheta_{i} \leq u(i), \quad 1 \leq i \leq n \tag{133}
\end{equation*}
$$

and the feasible region is defined by the set of additional constraints ( $m \geq 0$ )

$$
\begin{equation*}
g_{j}(\theta) \leq 0, \quad \text { for } j=1, \ldots, q, \quad \text { and } h_{j}(\theta)=0 \quad \text { for } j=q+1, \ldots, m \tag{134}
\end{equation*}
$$

Applied to a deployment of $N$ FAPs, the objective of the problem is to find the transmission parameters for all transmitters (optimization of UL, DL or both) maximizing the overall capacity of the system

$$
\begin{equation*}
c_{\text {tot }}=\max _{B_{k}, P_{k}}\left(\sum_{k=1}^{N} c_{k}\right) \tag{135}
\end{equation*}
$$

where the capacity of the single transmitter for a flat fading channel is

$$
\begin{equation*}
c_{k}=\frac{B_{k}}{W} \log _{2}\left(1+\gamma_{k}\left(P_{k}\right)\right) \tag{136}
\end{equation*}
$$

is measured in bits $/ \mathrm{s} / \mathrm{Hz}$ and is expressed in terms of the ratio between the dedicated band $B_{k}$ to the total bandwidth $W$ and of the signal to noise plus interference ratio $\gamma_{k}$, given by

$$
\begin{equation*}
\gamma_{k}=\frac{P_{k} / L_{k k}}{B_{k} N_{0}+\sum_{j ; B_{j} \cap B_{k} \neq \varnothing} P_{j} / L_{k j} \times\left(\left|B_{j} \cap B_{k}\right|\right) / B_{j}} \tag{137}
\end{equation*}
$$

where $L_{k k}$ is the link loss (including channel gain) of the $k$-th FAP/FUE couple, $L_{j k}$ is the signal attenuation of $j$-th interferer, $P_{k}$ the transmitted signal power, $N_{0}$ the noise power spectral level, $P_{j}$ the interfering powers ( $j=1, \ldots, N ; j \neq k$ ). The sum in (137) accounts for the interference arising from the overlap of the bands assigned to the $k$ - $t h$ and $j$-th transmitters.
For a frequency selective channel the algorithm can be implemented splitting the GA parameters search over each coherence band $b_{f}$ of the channel $(f=l, \ldots, F)$, evaluating $c_{k f}=c_{k}\left(b_{f}\right)$, so that the overall capacity is given by

$$
\begin{equation*}
c_{\text {tot }}=\sum_{f=1}^{F}\left(\max _{B_{k}, P_{k}} \sum_{k=1}^{N} c_{k}\left(b_{f}\right)\right) \tag{138}
\end{equation*}
$$

In this case, capacity for user $q$ reads as

$$
\begin{equation*}
c_{q}=\sum_{f=1}^{F} \log \left(1+\frac{P_{q}(f)\left|H_{q q}(f)\right|^{2}}{\sigma_{n q}^{2}(f)+\sum_{n \neq q} P_{n}(f)\left|H_{n q}(f)\right|^{2}}\right) \tag{138}
\end{equation*}
$$

In the case of UL, the signal is the transmission of the current $k$-th FUE and the interference is given by all other $j$-th $(j \neq k)$ FUEs. For the DL, one suffices to exchange the role of FUEs with that of FAPs.
The system analysis to maximize capacity has been designed to address different schemes that can be applied either to DL or to UL, the only difference being the nature of the local transmitters and interferers (FAPs or FUEs).

The metrics adopted as fitness function for measuring the goodness of the optimization algorithm are evaluable by the cell sum-rate and the geometric mean in terms of the obtained capacity

$$
\begin{gather*}
I_{1}=-\sum_{k=1}^{N} c_{k}  \tag{139}\\
I_{2}=-\left(\prod_{k=1}^{N} c_{k}\right)^{1 / N} \tag{140}
\end{gather*}
$$

The expression $I_{1}$ in (139) is directly based on the aggregate capacity of the system while the functional $I_{2}$ in (140) has been implemented to evaluate a fairness criterion on the single transmitters capacities.

In general, on the basis of the above introduced capacity per user, it is possible to design some ancillary fitness functions, whose solely inputs are the 2 N mean SNR values and the output is a singlevalued scalar, representing the "score".
The fitness functions selected for this test are:

- 'max_sum_rate_overall', i.e. the average sum-rate, descending directly from $I_{1}$,

$$
\begin{equation*}
f_{0}=-\frac{1}{2 N} \sum_{k=1}^{2 N} c_{k} \tag{140}
\end{equation*}
$$

aims to maximise the aggregated theoretical capacity both on the FAPs and FUEs (employing the mean SNR of both). This metric cannot insure any fairness of the system, as a there are infinite frequency allocations with the same aggregated theoretical capacity with unfair sharing of capacity.

- 'max_sum_rate_UL'

$$
\begin{equation*}
f_{1}=\left.f_{0}\right|_{F A P s} \tag{140}
\end{equation*}
$$

the same as above, but employing only the SNR available at the FAPs;

- 'max_sum_rate_DL'

$$
\begin{equation*}
f_{2}=\left.f_{0}\right|_{F U E s} \tag{140}
\end{equation*}
$$

the same as above, but employing only the SNR available at the FUEs;

- 'max_rate_fairness'

$$
\begin{equation*}
f_{3}=f_{0}+\alpha \cdot \sigma(c) \tag{140}
\end{equation*}
$$

a weighted metric trying to maximise the sum rate, but adding a penalty factor decreasing with the fairness of the solution;

- 'max_service'

$$
\begin{equation*}
f_{4}=-\sum_{k=1: N} e^{-\left(c_{k}-s_{k}\right)^{2} / 2 \beta^{2}} \tag{140}
\end{equation*}
$$

it does not try to assign all the available resources to the entities, but exploits the fact that each entity has a specific requirement (in terms of Mbps ) depending on the running application and thus tries to allocate only the bandwidth needed to meet the requirement of
each unit, decreasing the score when the given capacity is both above or below the requested capacity.
where $s_{k}$ is the mean capacity request and $\alpha, \beta$ are simulation parameters. Their values are summarized in Table 18.

| Parameters | Selected values |  |
| :---: | :---: | :---: |
| Ptx | $[16,15,17] \mathrm{dBm}$ | for the case of 3 couples |
| $s_{k}$ | $[1,3,4] \mathrm{Mb} / \mathrm{s}$ | for the case of 3 couples |
| $\alpha$ | $[1,1.3]$ |  |
| $\beta$ | $\sqrt{2} / 2$ |  |

Table 18. Values of parameters adopted for the computations.

The following approaches have been adopted for the GA optimization:

1. GO for power and frequency (see Figure 33, activity 1). Firstly the GA is applied to find (sub-)optimal power and band allocations for all the transmitting FAPs (or FUEs). The core investigation has been performed assuming that all nodes transmit UL or DL following the same scheme for the frame structure. The algorithm has been verified to converge also in the generic case, when all nodes have different clock drifts and frames have all different structures, even changing from frame to frame.
The band allocation is determined by the starting value and relative width of the current band allocated. The constraints over frequency and the power of transmission are summarized as

$$
\begin{align*}
& C_{k} \in\left(C_{k}^{\min }, C_{k}^{\max }\right), \quad C_{k}^{\min }<C_{k}^{\max }  \tag{141}\\
& 0 \leq P_{k} \leq P_{\max }
\end{align*}
$$

where $C_{k}$ is an adimensional label indicating the number of contiguous frequency channels
 $C_{k}^{\text {max }}=32$ ) giving rise to the $k t h$-users band, $P_{\max }=20 \mathrm{~dB}$ is the maximum overall power. Frequency allocations and transmission power per channel are treated as independent variables and the approach is equivalent for UL or DL phase. The number of variables in the GO is given by $3 \mathrm{x} N$, i.e., one for lower frequency value of user band, one for its bandwidth and one for power. The algorithm simultaneously optimizes frequency allocations and the relevant transmission power so that power is distributed among different chunks as an output of the GA run. Power and frequency are simultaneously treated as independent variables.
2. GO for power only (see Figure 33, activity 2). This assumes that the frequency allocations of each transmitting unit has been already performed, the GA is implemented to determine the solely (sub-)optimal transmission power, for FAPs co-existing (and thus interfering) on the same sub-band. This method is based on the clustering of units, considered at two levels of operation: the first is by separation of FAPs on the basis of the sub-band assigned for transmission, the other on a clustering based on the effect of the pathloss driving the mutual interference.

The constraint for system optimization is

$$
\begin{equation*}
0 \leq P_{k} \leq P_{\max } \tag{142}
\end{equation*}
$$

where $P_{\max }=20 \mathrm{~dB}$ and $k=1, \ldots, N$.
3. Power per frequency block. In this case, the variable of the optimization is still $P_{q}(f)$, with $q=1, \ldots, N, f=1, \ldots, F$. For LTE, $F$ refers to PRBs. The constraint for every user $q$ is

$$
\begin{equation*}
\sum_{f=1}^{F} P_{q}(f) \leq P_{q t o t} \tag{142}
\end{equation*}
$$

4. Scalability (see Figure 33, activity 3). For a deployment of $N$ FAPs a criterion to define a series of clusters has been investigated. Parallelization allows splitting computing load over several faster sub-processes and finally re-grouping the single outputs. This issue is addressed in details in section 8.4.


Figure 33. Scheme of the activities for parameters optimization by GA.

The simulations for the UL (optimization of FUEs only parameters) have been performed by gradually increasing the system complexity level, starting with the deployment of 4 couples of transmitters and then increasing the number up to some dozens, and testing the scalability of the procedure on a clustering scheme.

Three types of propagation loss situations have been considered:

- Case 1. no propagation loss
- Case 2. Propagation loss calculated by ITU/3GPP indoor-to-indoor model, flat fading channel (detailed in section 14.3).
- Case 3. Propagation loss calculated by ITU/3GPP indoor-to-indoor model, channel frequency selective fading (by chunks of frequencies), (detailed in section 8.3.2.2).


### 8.3 GO for flat fading scenario

### 8.3.1 Fitness function

The optimization process is intended to search for the (sub-)optimal "best" values of the power $P_{k}$ to be allocated to every transmitter and of the band interval of frequencies $B_{k}$. We remind that the "best" values sorted out by the GA are sub-optimal.

When considering the genetic search for best transmission values, the simulation does not need synchronization among FAPs since this would require other resources to be provided and is not part of the present analysis. If by other means one could obtain synchronization among FAPs, and therefore among FUEs and the entire network, such properties could be taken into account and included in a modification of the present scheme.
The algorithm is designed to search for the frequency value at the beginning of the current band allocated, its relative width and the power of transmission, under the constraints in (141). The metrics measuring the goodness of the optimization algorithm expressed by the fitness functions defined in (139) and (140), the first based on maximizing the overall capacity of the system, the latter the second inspired by a more fair allocation of resources at system level.

### 8.3.2 Analysis of topologies for a single block

The performance of the genetic algorithm optimization has been tested in several cases of pathloss matrix structures, providing an output coherent with the value theoretically expected for the considered deployments.
In the present simulations we assume frame duration of 5 ms and that each FAP feeds back a single SNR data on the backhaul link per frame.
The parameters search by means of the GA has been implemented in two schemes:

1. Frequency and power search. The algorithm provides either the frequency allocation (starting value and bandwidth), or the power of transmission for the single FUESs to the corresponding FAPs. This analysis has been performed for flat fading and for selective fading propagation scenarios.
2. Power search only. We consider the case that the system has exploited other resources/strategies to assign the frequency channels to the various equipments (eventually also a preliminary run of a separate GA devoted to it), thus the GA searches for (sub)-optimal values of the transmission power in order to maximize capacity and minimize interference. The strategy can depend on the topology of the network and on the requirements of the analysis.

### 8.3.2.1 Flat fading channel

We present the results obtained for four FAPs in the same building, to assess the properties of the optimization procedure. As a test case, the simulations refer to a flat fading channel, thus one pathloss matrix suffices to provide a snapshot of the transmission capabilities of the system.
Let us first analyze the simple case of four FAPs equally interfering each other with a pathloss matrix constituted of zeros only, as expressed in dB by

$$
L_{\text {loss }}=\left[\begin{array}{llll}
0 & 0 & 0 & 0  \tag{143}\\
0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0
\end{array}\right]
$$

The output of the genetic parameters search is depicted in Figure 34. The upper panel shows the evolution of the fitness function up to convergence, which stopped when the cumulative change in the fitness function is less than a fixed parameter (in our simulations equal to $1 \mathrm{e}-10$ ), without computing the predetermined number of generations ( 350 in this scheme) and without a significant dependence on the population size (here fixed at 100 elements). The algorithm quickly converges providing the values to adopt for the transmission of the FUEs registered to the FAPs as shown in the bottom panel. The twelve variables subject to optimization are the starting frequency chunk (1st to 4th), the
bandwidth associated to each unit (5th to 8th) and the power allocated to each transmitter (9th to 12th). Thus, the value of the 1 st variable represents the initial frequency value whose associated width is given by the value of the $5^{\text {th }}$ variable and the power given by the $9^{\text {th }}$ variable, and so on.


Figure 34. GA output for a set of 4 FUEs equally interfering each other. The variables in the bottom panel represent: 1-4 initial values of the frequency bands, 5-8 corresponding bandwidth, 9-12 power allocations.

Every FUE has been assigned the same power (see the last four bars in the bottom plot), the capacity has been distributed as $(50,68,35,56) \mathrm{bits} / \mathrm{s} / \mathrm{Hz}$ over the four FUEs. The resulting frequency and power allocations can be better viewed in the representation of Figure 35: the four FUEs can be assigned orthogonal frequency bands, thus minimizing the interference in UL, for an overall average capacity of $52.25 \mathrm{bits} / \mathrm{s} / \mathrm{Hz}$ for the system, in this idealized test case.


Figure 35. GA output for 4 FUEs in terms of frequency and power allocations.

Let us consider an interference scenario in which one FUE (for example FUE number 1 in (144)) is mildly interfering with the other three which, on the other side, have a much higher level of mutual interference, as expressed by a pathloss matrix as

$$
L_{\text {loss }}(d B)=\left[\begin{array}{cccc}
30 & 400 & 400 & 400  \tag{144}\\
400 & 30 & 30 & 30 \\
400 & 30 & 30 & 30 \\
400 & 30 & 30 & 30
\end{array}\right]
$$

Such deployment refers to a case similar to the one depicted in Figure 36, leaving unmodified the assumption of one FUE per FAP.


Figure 36. Possible deployment corresponding to an interference scenario such as for a pathloss matrix of the type in Eq. (144).

In this case one gets a genetic algorithm output as detailed in Figure 37. The convergence is indeed fast, in terms of number of generations needed to find (sub-)optimal values for the twelve parameters. Once again, the four transmitters are assigned the maximum power of transmission, with a more imbalanced capacity as $(142,23,136,5)$ bits $/ \mathrm{s} / \mathrm{Hz}$ per transmitter, but a higher capacity ( 76.65 bits $/ \mathrm{s} / \mathrm{Hz}$ ) for the overall system. This numbers are related to the chosen scenario: the pathloss expression and the little number of transmitters are simply adopted to validate the methodology. The band sharing is orthogonal among the three FUEs in a bunch (black, blue and green line in right panel of Figure 37) while the remaining FUE can use the whole spectrum (red line, in the same Figure).


Figure 37. GA output for 4 FUEs deployed as in Figure 36 and interference scenario such as for a pathloss matrix of the type in Eq. (144).

Let us now consider the generic case of a realistic pathloss given for example by the matrix

$$
L_{\text {loss }}(d B)=\left[\begin{array}{llll}
56 & 78 & 93 & 93  \tag{145}\\
55 & 45 & 84 & 85 \\
95 & 84 & 58 & 83 \\
87 & 84 & 66 & 61
\end{array}\right]
$$

which results in an optimization with capacity $(9,28,14,23)$ bits $/ \mathrm{s} / \mathrm{Hz}$ but with a spectrum assignment on orthogonal frequencies (overall capacity equal to $77 \mathrm{bits} / \mathrm{s} / \mathrm{Hz}$ ).


Figure 38. GA output for 4 FUEs deployed with an interference scenario such as for a pathloss matrix of the type in Eq. (145).

The same pathloss matrix (145) can give rise to a GA parameters search as displayed in Figure 39 three FUEs have been assigned orthogonal frequency chunks while the remaining one is fully interfering with one of them but with a definitely low power which, would indicate no transmission opportunity. The capacities are $(0.06,22.21,54.04,30.39) \mathrm{bits} / \mathrm{s} / \mathrm{Hz}$, respectively and the overall one $27 \mathrm{bits} / \mathrm{s} / \mathrm{Hz}$. The score of the fitness function is in both cases almost the same, with a value below -7 , irrespective of fairness.


Figure 39. GA output for 4 FUEs deployed with an interference scenario corresponding to a pathloss matrix of the type in Eq. (145).

More examples can be found in section 14.4.

In the following sections we will analyze an approach to parallelization of the genetic algorithm parameters search for a large number of FAPs by developing the concept of clusterization on the basis, for example, of the analysis of effective pathloss matrix as discussed in Section 14.3.2.

### 8.3.2.2 GO for frequency selective channel

Let us consider the FAPs deployment in an environment of propagation with a frequency selective channel: the whole bandwidth is separated in a set of frequency chunks, whose number depends on the scenario of propagation, and the system is described by a pathloss matrix for every chunk.
As an example, let us address the case of a frequency selective channel divided in five chunks, each constituted by 16 subcarriers, every subcarrier 10 kHz wide.
Let us consider two realizations.
In the first, a set of pathloss matrices is given as in Table 19. In the first band of frequencies (chunk number 1) every FUE can communicate with its own FAP without interfering with others, while interference increases for the other frequency chunks.

| Chunk 1 |  |  |  | Chunk 2 |  |  | Chunk 3 |  |  |
| :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: |
|  | 120 | 120 |  | 060 | 60 | 60 | 030 | 30 | 30 |
| 120 | 0 | 120 |  | $60 \quad 0$ | 120 | 120 | $30 \quad 0$ | 120 | 120 |
|  | 120 | 0 | 120 | 60120 | 0 | 120 | 3080 | 0 | 80 |
| 120 | 120 | 120 | 0 | 60120 | 120 | 0 | 30120 | 120 | 0 |
| Chunk 4 |  |  |  |  |  |  | Chunk 5 |  |  |
| 0 | 60 | 120 | 120 |  |  |  | 0 | $0 \quad 60$ | 60 |
|  | 0 | 60 | 60 |  |  |  |  | 060 | 06 |
| 20 | 60 | 0 | 120 |  |  |  | 60 | 00 | 060 |
|  | 60 | 120 | 0 |  |  |  |  | 060 | 0 |

Table 19. Propagation loss matrices, in dB, for the separate frequency chunks. Rows follow FUE's index, columns FAP's index.


Figure 40. Overall output of the GA for the set of pathlosses described in Table 19.

The corresponding output of the GA optimization is shown in Figure 40. The result can be read in terms of frequency chunks allocations: the algorithm assigns maximum power to FAP/FUE when
interference from other transmitters is negligible and tends to reduce frequency allocations when it increases.

In the second realization pathlosses are the outcome of a scenario in which every FAP interferes with all other transmitters. Thus the pathlosses for the various chunks summarized in Table 20, give rise to a more variegate framework. In this example, we assumed that the first two frequency chunks have the same pathloss structure, while the other three frequency chunks are characterized by a certain propagation environment. Values in bold are relevant for the interpretation of the output detailed in Figure 41.


Table 20. Propagation loss matrices in dB for the separate frequency chunks. Rows follow FUE's index, columns FAP's index.

In chunk number 1, the first FUE has not been assigned a transmission opportunity as the pathloss with its own FAP is sensibly higher than with other FAPs in the environment. In the third chunk, the first three FUEs are privileged with respect to the fourth while in the fourth chunk we note how the second FUE has the best link with its own FAP. The same can be said for the FUE number 3 in chunk number 5 .

As expected, the solution provided by the algorithm depends on the relative values of the pathloss matrix entries. Considering the increase of the unknown variables space by a factor equal to the number of chunks in the propagation emulation, convergence time of the GA is not affected by adopting several parallel GA processes.


Figure 41. Overall output of the GA for the set of pathlosses described in Table 20.

In Figure 42 we report the output of the GA applied to a system of 12 FAPs with a frequency allocation that can be distributed over 12 frequency chunks. The output of such analysis shows a tendency to allocate evenly the transmitters over orthogonal frequencies, a case more favored also by the higher availability of frequency chunks with respect to the previous case. The right panel of Figure 42 shows how the mean capacity (in bits $/ \mathrm{s} / \mathrm{Hz}$ ) is uniform over the frequency spectrum.
Depending on the propagation model this approach can be extended to any partition of the frequency band, which in turn could be determined by other resources.
In this view, the next Sections are devoted to increasing the parallelization of the power allocations extending the GA optimization from a few FAPs to a larger population.


Figure 42. Deployment of 12 FUEs with pathloss varying over 12 frequency chunks. Right panel: Overall output of the GA. Left panel: mean capacity per frequency chunk.

### 8.4 Genetic Algorithm and scalability

### 8.4.1 Benefits of clustering

The search of FAPs/MSs parameters to minimize interference among nearby transmitters has been so far applied to the whole system of users. The increase of the number of FAPs involved in the optimization affects the computation time since reliability of results increases with the population of individuals for the genetic optimization. When the algorithm has to search either for the frequency assignments or for the power, the number of variables is $3 \times N_{\text {FAPs }}$ (i.e., taking into account the starting value of each frequency assignment, its bandwidth and each power level).
Assignment of frequency bands among FUEs could be provided by other resources, for example by a preliminary run of the GO optimization for the whole set of variables, or by assigning nearby transmitters orthogonal frequency chunks on the basis of service level or priority or quality of traffic. In this case, adoption of GO for assignment of power only reduces complexity of optimization to $N_{\text {FAPs }}$ variables.

Section 14.3.2 analyzes some peculiar aspects of the pathloss relations among transmitters, providing a possible method to group users with higher mutual interference. Once split the overall system into subgroups, GO can be readily run in parallel on clusters of users ensuring faster convergence in terms of lower system time. Indeed, when the set of $N_{F A P s}$ transmitters is split into $N_{C l}$ clusters system time is reduced, on average, of a factor $1 / N_{C l}$.
In [FREEDOM-D52] we implement various methods for grouping users based on different criteria (e.g., traffic type, location, etc.) where dedicated metrics and measures will be described and implemented to evaluate their impact and benefits at system level.

For what regards GO, if the system has a resource to group users into subclusters, in section 8.4.2 a method for treating such smaller sets is presented.

In the next Section we will analyze how parallelization can be implemented for GO as by the scheme:

1. Identify and classify subsets of FAPs
2. Process in parallel the GO over the subsets
3. Aggregate the sub-solutions toward the whole system.

### 8.4.2 Clustering of FAPs

For an inhomogeneous deployment of users the optimization strategy can be differently implemented by outlining the criteria to identify subsets of users. The clustering scheme depends on the clustering measure and clustering length that one is going to implement and the system will provide benefits or shortcomings depending on the strategy adopted. Although a thorough activity is focused on this topic (see [FREEDOM-D52]), in this section outlines the basic ideas

Possible criteria to group users are based on physical or application characteristics. The former class includes geographical position and interference relation (e.g., provided by GPS localization, femtogateway or link loss), while the second class is based on the type of traffic generated by users, discriminating for example on the basis of priority (FTP vs. VoIP, web browsing vs. video streaming). The choice of implementation of one or the other criterion depends on MNO technical and commercial strategy: indeed the operator policy could be oriented to provide better service to privileged sets of users or, on the other side, to provide an average QoS to the highest number of users.
In view of large population deployments, the algorithm scalability performances have been analyzed, outlining some methods to reduce the complexity of GO from the whole set of FAPs to every single cluster.

The case of assignment to a cluster on the basis of pathloss matrix analysis realizes, for example, for a scenario with a set of FAPs deployed in two nearby buildings, or also when distributed over different floors of the same building. Such definitions to assign users to clusters often imply that the intersection among different subsets is non-empty, i.e. some users interfere with different adjacent groups, and require a dedicated strategy to proceed with parameters optimization.

The criteria for assignment to a cluster mentioned so far are based on users in competition for certain resources (such is the case of interference relation). A different, complementary approach is provided by grouping users with less interference with others, or less band requirements: this approach permits to remove those users from optimization, thus offloading computing effort and allowing different strategies to be implemented, such frequency reuse patterns or transmission power adaptation.
An optimization method based on clustering, foresees parallelization of computing: this is an advantage if the system can be split in sub-systems for which the border effects, quantified by the number of users belonging to more than one cluster is less the number of FAPs belonging to a single cluster.

Dealing with border effects is the focal point to find the efficiency of the parallelization algorithm.
For the sake of clarity, the meaning of the terms adopted for treating scalability issues agrees as:

- Blocks: geographical location of FAPs, e.g., those within the same building. Some of them can receive interference from other transmitters in other blocks.
- Cluster: a sub-set of FAPs which is processed in a single GA run, e.g., those within the same building plus some others in an adjacent one.
- Sub-set: any sub-group within the set of $N$ FAPs.

The best-case for parallelization results when the whole set of FAPs can be split in $N$ independent subsets, i.e. with empty intersection: in this event, the complexity of the optimization depends on hardware capability to deal with $N$ independent computations. Complexity grows when an increasing number of FAPs belonging to a certain sub-set overlaps (e.g., interferes) with some FAPs of another sub-set. The system has to implement a strategy to deal with the parameters provided by two (or more, depending on the interference scheme among FAPs belonging to different sub-sets) independent GA optimizations.
When the output of an optimization runs over different clusters provides two possible parameters values for FAPs lying within the intersection, the system can swap or average the obtained values before running the next step of system optimization (for GO, typically a dozen times), adopting the results as initial values for the next run.

### 8.5 Synchronization issues and GO

### 8.5.1 Asynchronous scenario

The reference scenario is made up of an asynchronous ensemble FAP-FUE couples (assuming, one FUE and one FAP). This assumption should be better viewed as the lack of assumption on synchronism among an ensemble of objects driven by timers. Their natural state is to be asynchronous, unless specific resources and assumptions are introduced to insure a certain (sufficient) degree of synchronism among them.
For instance, each FAP belonging to an ensemble of FAPs could be synchronised with the external BS, thus insuring a certain degree of synchronism among the ensemble. This implicitly introduces the assumption of a BS-FAP link not too weak (to let the FAP sense and synchronise) and not too strong (otherwise the FAP presence would be pointless). A similar consideration holds for a FAP-FAP synchronisation.
In general, LTE identifies seven possible UL/DL structures for a frame, as summarized in Table 21 and, in principle, every FAP can adopt one different from others, due to its own requirements (e.g., UL/DL traffic ratio, etc.).

```
case 0 subframe_structure = [llllllllllllll}-10111 1-1 0 1 1 1];
case 1 subframe_structure = [lllllllllllll}-10111-1-1 0 1 1-1];
case 2 subframe_structure = [-1 0 1-1 -1 -1 0 1 -1 -1] ;
case 3 subframe_structure = [\begin{array}{llllllllll}{-1}&{0}&{1}&{1}&{1}&{-1}&{-1}&{-1}&{-1}&{-1}\end{array}];
case 4 subframe_structure =[\begin{array}{llllllllll}{-1 1 -1-1-1-1-1-1];}\end{array}]
case 5 subframe_structure = [-1 0 1-1-1-1 -1 -1 -1 -1] ;
case 6 subframe_structure =[[\begin{array}{lllllllllll}{-1}&{0}&{1}&{1}&{1}&{-1}&{0}&{1}&{1}&{-1}\end{array}];
```

Table 21. UL/DL frame structures for LTE: -1 corresponds to DL, 1 to UL, $\mathbf{0}$ to temporal slots.


Figure 43. Example of clock drift for three FAPs (blues, green, red). In adimensional units, the line up corresponds to a transmission phase, line down to receiving phase. In this example the three FAPs have different UL/DL temporal structure. Arrows highlight directions of clock drifts.

Due to the clock differences among the FAP-FUE couples (e.g. 50-100 ppm), the frame structures drift each other with time. Thus, in such a scenario, the exchange of information between FAPs and FUEs is not a viable assumption and FAPs and FUEs belonging to different couples cannot exchange information.

Let us consider the simplified case depicted in Figure 44. The communication is established between FAP(1) and FUE (1) and between $\operatorname{FAP}(2)$ and $\operatorname{FUE}(2)$. The weakest assumption is that each FUE has completed the handshaking process with its own FUE and the protocol (whatever is employed) thus insures that $\operatorname{FAP}(1)$ is synchronised with $\operatorname{FUE}(1)$ and that $\mathrm{FAP}(2)$ is synchronised with FUE(2). This synchronisation within each couple is of course possible only because a) the parties of the couple are each other in radio range and b) the employed protocol spends specific resources to establish and then maintain synchronism between the $\operatorname{FAP}(\mathrm{k})$ and $\operatorname{FUE}(\mathrm{k})$. The consequence is that (within each couple) the UL and DL comb are aligned (and always will be) in time at $\operatorname{FAP}(\mathrm{k})$ and $\operatorname{FUE}(\mathrm{k})$.
Unless additional assumptions are introduced and specific resources are employed to establish and maintain synchronisation among different couples FAP-FUE, the situation is the following: each FAP/FUE couple remains synchronised, but different couples FAP-FUES are not. Thus FAP(j) and $\operatorname{FUE}(\mathrm{k})$ or $\operatorname{FAP}(\mathrm{j})$ and $\operatorname{FAP}(\mathrm{k})$ or $\operatorname{FUE}(\mathrm{j})$ with $\operatorname{FUE}(\mathrm{k})$ (with $\mathrm{k} \neq \mathrm{j}$ ) cannot, in principle, exchange any information simply because they cannot agree a common time basis and thus the UL phase of one party cannot correspond to the DL of another party. Even if a moment in which all the UL-DL combs are synchronised could realize, this condition will be soon lost after a few fractions of seconds, because of the clock drifts.


Figure 44. Example of two couples deployed nearby.

An example of how the frame structures drift each other has been purposely analyzed for the scenario with 3 FAP-FUE couples with different frame UL-DL combs and overall frame duration of 10 ms .
Once given the timing differences of the quartzes, Figure 45 reports an example of the time variation of the mean (mean in each frame) theoretical capacity (solely depending on SNR) of the 3 FAPs, along with a detail showing the "fine structure" of the mean theoretical capacity. The leftmost plot shows the continuous drift among the UL-DL frame structures. Upper, intermediate and bottom lines correspond to the three FAPs. The observed time variation is solely due to the frame drifting, because of the relevant overlap between the UL and DL phases (unknown to each entity, as well as to a "hypothetic" central unit).


Figure 45. Capacity evolution for three couples FUEs/FAPs.

It is assumed that

- Each unit in the scenario has at least the capability of estimating its own mean SNR experienced in a frame;
- At the end of each frame, this estimated SNR can be delivered to the central processing unit, linked to each FAP via the backhaul link;
- Each FAP may (or may not) know also the mean SNR (mean in a frame) of its FUE.

Summarising, in the case each FAP knows both its own mean SNR and the mean SNR of its FUE, once in a frame (at the end of) these 2 values can be delivered to the central processing unit. In such a case it can be implemented a procedure capable to jointly "optimise" the appropriate metric accounting for both UL and DL.

In the case (maybe way too restrictive but worth to be considered for completeness) each FAP doesn't know the mean SNR of its FUE, optimization of a suitable metric is only relevant to the UL. This is what could be referred as "optimisation of the UL": as the metric includes only (mean) information coming from the UL, none can insure that the "optimal" resources allocations found in this case could not jeopardise the DL performance.

### 8.5.2 Design principles and constraints

Given the above, the design is the following:

- The FAP-FUE couples in the scenario are not synchronised, so the design cannot rely on the exchange of information among them;
- Each FAP and each FUE knows its own mean SNR experienced in a frame;
- Once in a frame, the value of the mean SNR of each FAP and each FUE can be delivered (from the FAP ADSL link) to a central processing unit;
- The central processing unit collects ONLY the SNR of FAPs and FUEs and knows who is who, but (of course) has no knowledge of who is interfering with who;
In the asynchronous scenario the contributions to SINR cannot be separated, because each FAP or FUE cannot know who is interfering with it. For N FAP-FUE couples, there are 2 N entities in the scenario and thus the mean SNIR estimated by each unit is collected (including the interfering contributions, unknown to the unit, due to the other units). The assumption that each FAP can measure/estimate its own mean SNIR once in a frame and the same can do each FUE is extremely relaxed. These 2 N numbers (i.e. some history) are sent and processed by the central unit, providing as output a resource allocation (e.g. frequency assignments) optimising the metric.


### 8.5.3 GA tested on a fully asynchronous frame sequence

In order to evaluate the effect on GA optimization method of lack of synchronization as detailed above, in this Section we present the results of running the GO for frequency only, keeping fixed the average transmission power. It has to be considered, as a simulation example, the case limited to the management of 3 FAP-FUE couples, but the simulation principle applies to an arbitrary number of FAP-FUE couples.

Taking into account the continuous frame drift among the FAP-FUE couples, the simulator computes the mean SNR of each entity at each frame, implementing the asynchronous GA optimisation process giving as outputs the "best" frequency allocation.

### 8.5.3.1 Results of GA tested on fixed frame sequences

In this subsection, we present the results of GO for $f_{3}$ and $f_{4}$ metrics under the assumption the the system is asynchronous, every coupe FAP/FUE adopts a different frame structure and clocks have different time drifts un-compensated by synchronization methods.

By adopting the 'max_service' fitness function and assuming that the requests of FAP1, 2 and 3 are 1, 3 and 4 Mbps , respectively, the output of the GA allocation is the one reported in Figure 46.

It is clear that in both cases the output of the GA process actually approaches is compliant with the design of the fitness function.


Figure 46. Capacity output for GO implementing the 'max_service' metric.
Implementation of the 'max_rate_fairness' fitness function, gives an output of the GA allocation whose effect on SNR (and therefore capacity) in time is reported in Figure 47.


Figure 47. Capacity output for GO implementing the 'max_rate_fairness' metric.

### 8.5.3.2 Results of GA tested on randomly changing frame sequences

This subsection reports for $f_{0}, f_{1}, f_{2}, f_{3}$ and $f_{4}$ metrics influence on the SNR temporal evolution, testing the GA output in a peculiarly structured case. Although in principle unlikely, it has been tested the a scenario in which at every frame the sequence of UL and DL for the next frame is randomly changed for every couple. Convergence of the GA is verified and the optimization of results lead to a uniform of SNR in time, a pattern allowed by the intrinsic orthogonalization of frequency allocations.


Figure 48. Randomly changing frame schemes: GA capacity results for max sum rate and right max service metrics, left and right plots, respectively.


Figure 49. Randomly changing frame schemes: left UL, right DL.


Figure 50. Randomly changing frame schemes: maximum rate with fairness, left $\alpha=1$, right $\alpha=1.3$.

### 8.5.3.3 Conclusions on second tier synchronization and GA

When the synchronism assumption among all the FAPs and FUEs is removed, it is still possible to design a sub optimal resources allocation procedure, based on the mean SNR measured (in a frame) by each entity in the scenario and made available at a central processing unit.

The GA approach has a working principle suitable for such a kind of problem. Although it is not the unique possible algorithmic approach, it is attractive for its simplicity and modularity. End-to-end, it inputs the mean SNRs from FAPs and FUEs (as many times as established by the choice operated on the "population" size) and outputs the "optimal" (in the sense of the adopted metric) radio resources allocation.

Its main drawbacks rely in the convergence time and in the distance from the optimal solutions. Many computational aids can help to overcome the latter, but the compliance between the convergence time and the requirements at system level cannot be insured a priori in all scenarios.

The results of the simulations in an asynchronous scenario have shown to be consistent with the "physical meaning" implemented in the adopted fitness functions, on which a specific effort must be spent in order to obtain a meaningful target and functional slopes easing the GA convergence capabilities.

### 8.6 Conclusions of the GO optimization

The search for transmission parameters minimizing interference in a system of FUEs and maximizing the overall system capacity has been performed by a genetic algorithm method implemented in the network at a centralized level. The same optimization can also be implemented as a distributed algorithm as introduced in Section 8.2.1, providing the same performances.
On the basis of the SNR fed back at a regular basis, not relevant for the present part of the analysis, the RRM can determine the frequency allocation and/or the power of transmission for the UL of every FUE in order to maximize the system capacity.
The method has been tested starting from some study-cases with already known best-solutions for various types of metrics investigating different constraints, finally extending the analysis to deployments with more than 100 FAPs/FUEs. For larger numbers it is reasonable that not all FUEs interfere with every other radio source, but this relation can be limited on the basis of distance and of obstacles in between. When this property is expressed in terms of a set of system features, such for example the pathloss matrix, it has been proposed a method to scale the system, splitting the whole set of users into sub-sets, reducing the computing complexity and obtaining faster results.

## 9 SIMULATION RESULTS

The techniques investigated in this deliverable have been evaluated in two corporate-based scenarios:

- Small Corporate configuration which consists of one MBS and one FAP area (uniformly distributed over the sector). The small Corporate scenario is depicted Figure 51 where one FAP area is deployed in a macrocell sector of radii equal to 500 m . The FAP area contains two twin buildings separated by one street. We define by $\mathrm{FAP}_{\text {load }}$ as the probability that an office/apartment has a FAP and it is active, i.e. $\mathrm{FAP}_{\text {load }}=1$ means that all offices have an active FAP (there are 60 FAPs in average). Each FAP is serving up to two FUEs. The system operates over a bandwidth of 5 MHz and we consider several limiting values for the backhaul connection $(5 \mathrm{Mbps}, 10 \mathrm{Mbps}$, $20 \mathrm{Mbps}, \infty \mathrm{Mbps})$. See further details of this scenario in Table 22.


Figure 51. Small Corporate scenario

| Parameters | Values |
| :--- | :--- |
| Number of floors per building | Uniform r.v. [1,..6] |
| Number of offices per floor | $2 \times 5$ |
| Antenna configuration | 4 (MBS), 2 (FAP), 2 (MUE) |
| Number of FAPs | $\{12,24,36,48,60\}$ in average |
| Position of the FAP | Indoor. |
| Number of FUEs per FAP | 1 or 2 |
| Position of FUEs | Indoor or Outdoor |
| Prob. of an outdoor FUE | $\left(\frac{0.8}{n}\right)^{n-1}$ for a FAP in the $n$-th floor |
| Number of MUEs | 1 or 2 |

Table 22. Parameters considered for the scenario depicted in Figure 51

- Simulated Corporate configuration which consists of one MBS and 10 FAP areas. In contrast to the previous scenario where all channel parameters (pathloss, shadowing, ...) were obtained from [FREEDOM-D21], this configuration have obtained those channel parameters by means of simulating all the buildings and terminal deployment using a VOLCANO-based simulator from SIRADEL. In this scenario we will assume that each office/apartment have a probability equal to 0.16 to have a FAP. The number of inhabitants $/ \mathrm{km}^{2}$ is set to 740 for a given wireless mobile operator. Details on the FAP area configuration are introduced in Table 23. All FAPs are placed indoors and they are serving two FUEs. The served FUEs might be indoor or outdoor. In this
latter case, it is assumed that FUEs only can be at the street level. The probability that an apartment/office has a FAP has been set to 0.16 .


Figure 52. Cell-sector Corporate scenario with 10 FAP areas

| Parameters | Values |
| :--- | :--- |
| Number of floors per building | Uniform r.v. [1,..6] |
| Number of offices per floor | $2 \times 5$ |
| Antenna configuration | 4 (MBS), 2 (FAP), 2 (MUE) |
| Number of FAPs | $\{24,48,72,96\}$ in average <br> $(0.25,0.5,0.75,1)$ |
| Prob. on FAP deployed per office | 0.16 |
| Position of the FAP | Indoor. |
| Number of FUEs per FAP | 2 |
| Position of FUEs | Indoor or Outdoor |
| Prob. of an outdoor FUE | $\left(\frac{0.8}{n}\right)^{n-1}$ for a FAP in the $n$-th floor |
| Number of MUEs | 2 |

Table 23. Parameters considered for scenario shown in Figure 52

### 9.1 Resource allocation based on Weighted sum-rate (WSR) maximization

This section will evaluate the resource allocation techniques analyzed in section 6 in the scenarios shown in Figure 51 and Figure 52. The resource allocation tackles the optimization of transmit precoders and optionally RB assignment when each source is serving multiple users. The sources active in the same spectrum are the MBS and FAPs, while the users are the MUEs and FUEs. In order to elucidate the benefits of the different techniques and the impact of a high number of FAPs over one sector we look into:

- Small Corporate configuration which consists of one MBS and one FAP area (uniformly distributed over the sector). The impact of the number of active FAPs in the FAP area will be studied for different values of backhaul qualities.

| Parameters | Values |
| :--- | :--- |
| Priority of users | Inversely proportional to the served rate <br> (proportional fair criterion) <br> FUEs: $1 / \mathrm{r}_{\text {FUE }} \quad$ MUEs: $5 / \mathrm{r}_{\text {MUE }}$ |
| Number of RB | $25(5 \mathrm{MHz})$ |
| Optimization variables | Transmit precoders and RB assignment |
| Maximum Backhaul rate at FAP | $\{5,10,20, \infty\}$ Mbps |
| Resource allocation techniques | Pricing vs. non-pricing |
| Resource allocation criterion | Maximization of the weighted sum rate |

Table 24. Parameters considered when the evaluation is done over the Small Corporate configuration

- Simulated Corporate configuration which consists of one MBS and 10 FAP areas. In contrast to the previous scenario where all channel parameters (pathloss, shadowing, ...) were obtained from [FREEDOM-D21], this configuration have obtained those channel parameters by means of simulating all the buildings and terminal deployment using a VOLCANO-based simulator from SIRADEL. In this scenario we will assume that each office/apartment have a probability equal to 0.16 to have a FAP.

| Parameters | Values |
| :--- | :--- |
| Priority of users | FUEs: equal to 1 <br> MUEs: equal to 5 |
| Number of RB | $100(20 \mathrm{MHz})$ |
| Optimization variables | a) Transmit precoders and RB assignment <br> b) Transmit precoders |
| Maximum Backhaul rate | $\{10,20,40,80, \infty\}$ Mbps |
| Resource allocation techniques | Pricing vs. non-pricing |
| Resource allocation criterion | Maximization of the weighted sum rate |

Table 25. Parameters considered when the evaluation is done over the Simulated Corporate configuration

### 9.1.1 Small Corporate configuration

We would like to emphasize that the resource allocation algorithm tackles the optimization at the FAP and MBS, since both type of sources are sharing the same spectrum. In the scenario shown in Figure 51, MUEs are uniformly deployed over the sector, so in general will not be interfered by FAPs. However, MBS plays a dominant interferer for all FUEs. Therefore, the possible performance degradation at MUEs will be a consequence of decreasing the transmitted power used by the MBS.

Figure 53 depicts the spectral efficiency attained by FUEs and MUEs for different configurations of active FAPs (denoted by variable $\mathrm{FAP}_{\text {load }}$ ) and maximum backhaul rate when pricing and non-pricing techniques are considered. Under maximum backhaul rates $B=\{5,10\} \mathrm{Mbps}$, the spectral efficiency of pricing and non-pricing techniques are almost the same for all $\mathrm{FAP}_{\text {load }}$ configurations. When the $B=20$ Mbps the pricing technique improves the performance attained by the non-pricing one when $\mathrm{FAP}_{\text {load }}=1$ ( 60 FAPs in average) with $10 \mathrm{bps} / \mathrm{Hz}$. Furthermore, in case of $B=\infty \mathrm{Mbps}$, the gains become $\{6,17,29,39,51\} \mathrm{bps} / \mathrm{Hz}$ when there are $\{12,24,36,48,60\}$ FAPs in average, i.e. figures of merit $\{6 \%, 10 \%, 12 \%, 12 \%, 14 \%\}$ over the performance of non-pricing techniques.

On the other hand, the spectral efficiency of MUEs, shown in Figure 53-right, illustrate that pricing techniques impose a worse performance compared with the non-pricing ones for $B=\{5,10,20\} \mathrm{Mbps}$. Under the ideal backhaul configuration, the performance is the other way round. Notice, that in all cases FUEs improve their performance thanks to the pricing exchange, although in some cases the gain is very small. For example, $B=20 \mathrm{Mbps}$ and $\mathrm{FAP}_{\text {load }}=0.8$, FUEs can get $191 \mathrm{bps} / \mathrm{Hz}$ considering
the pricing technique, while they obtain $185 \mathrm{bps} / \mathrm{Hz}$ when non-pricing technique is envisioned. This gain comes at the cost of reducing the MUEs spectral efficiency from $7.98 \mathrm{bps} / \mathrm{Hz}$ to $7.14 \mathrm{bps} / \mathrm{Hz}$


Figure 53. Spectral efficiency attained by Left) FUEs and Right) MUEs for a max Backhaul rate $B=\{5,10,20, \infty\}$ Mbps with a $F A P_{\text {load }} l=\{0.2,04,0.6,0.8,1\}(12,24,36,48$, 60) FAPs in average for pricing ( P ) and non-pricing ( nP ) techniques.


Figure 54. 10\%-Outage Rate attained by Left) FUEs and Right) MUEs for max Backhaul rate $B=\{5,10,20, \infty\}$ Mbps with a $F A P_{\text {load }} I=\{0.2,04,0.6,0.8,1\}(12,24,36,48$, 60) FAPs in average for pricing ( P ) and non-pricing ( nP ) techniques

Figure 54 shows the performance of pricing and non-pricing techniques in terms of $10 \%$-Outage Rate, i.e. the minimum achieved rate by the $90 \%$ of the users (FUEs and MUEs). Similar to the spectral efficiency results, the benefits of pricing techniques come up when the backhaul quality is not limiting the wireless communications, for example when $B=20 \mathrm{Mbps}$ and $\mathrm{FAP}_{\text {load }}=1$ or when $B=\infty \mathrm{Mbps}$ and $\mathrm{FAP}_{\text {load }}=\{0.2, . ., 1\}$. FUEs get significant gains around $200-300 \%$ in this latter case. With respect the $10 \%$-Outage Rate attained by MUEs, see Figure 54 -right, using the pricing technique we can get bitrate different from zero in those cases where $B=\{20, \infty\}$ Mbps. This performance confirms that the pricing techniques are able to take into account the QoS (in terms of user priorities) of the different neighboring users. Under the configuration where all FAPs are active (and $B=\infty$ Mbps) it seems reasonable that in order to maximize the sum-rate of the system, the MBS should (the dominant interferer from the FUEs' point of view) shut down because there are many FUEs.

However, in such a case the priority of MUEs becomes higher (as a function of the proportional fair criterion). Since the resource allocation is based on the weighted sum-rate maximization, and thanks to the pricing exchange, that priority can be considered by all FAPs when individually optimize their resource allocation.

The outstanding gains obtained in terms of $10 \%$-outage rate become in the order of $\{7 \%, 13 \%, 20 \%, 24 \%, 28 \%\}$ when the $50 \%$-outage rate is considered, see Figure $55-$ top, what suggests that pricing techniques are suitable to improve the minimum rate of the system.
Finally, Figure 55-bottom presents $95 \%$-outage rate obtained by FUEs and MUEs, i.e. the rate attained by the $5 \%$ of the users. Comparing Figure 53-left, Figure 54-left and Figure 55 -bottom-left we can observe quite similar values of $10 \%-, 50 \%$ - and $90 \%$ - outage rate of the FUEs when $B=\{5,10,20\}$ Mbps because backhaul is limiting the wireless communication. In such a case, the transmit power is reduced in order to meet such constraint, so that the interference in that scenario is reduced. In that configuration, pricing and non-pricing techniques tend to get a similar performance.


Figure 55. 50\%-Outage Rate (Top) and 95\%-Outage Rate (Bottom) attained by Left) FUEs and Right) MUEs for a max Backhaul rate $B=\{5,10,20, \infty\}$ Mbps with a FAP ${ }_{\text {load }}$ $I=\{0.2,04,0.6,0.8,1\}(12,24,36,48,60)$ FAPs in average for pricing $(P)$ and non-pricing ( nP ) techniques
9.1.2 Simulated Corporate configuration

This section evaluates the decentralized resource allocation techniques under a configuration where 10 FAP areas are deployed in a cell sector with radii 500 m , see in Figure 52. We evaluate pricing and non-pricing based techniques when the optimization variables are the transmit precoders and RB assignment or just the transmit precoders. In the scenario shown in Figure 52 it has been assumed that each apartment/office has a probability equal to 0.16 , that means that there are 1.6 FAPs per floor (each floor has 10 offices/apartments). Given such deployed FAPs, we study how the performance varies as a function of the number of active FAPs, denoted by variable $\mathrm{FAP}_{\text {load }}, \mathrm{FAP}_{\text {load }}=1$ means that all FAPs are active, while $\mathrm{FAP}_{\text {load }}=0.75$ means that only $75 \%$ of the FAPs are active. With the parameters depicted in Table 23, the average number of FAPs deployed in the sector becomes 192.


Figure 56. Cumulative density function (cdf) of the individual throughput of Left) MUEs and Right) FUEs for a sector-based corporate scenario when pricing and non-pricing techniques are considered where transmit precoders and RB are optimized. The $\mathbf{F A P}_{\text {load }}=\{0.25,0.5,0.75,1\}$. Ideal backhaul. MUEs present a priority 5 times higher than FUEs. All FUEs have priority equal to 1.

Figure 56 and Figure 57 sketch the individual throughputs obtained by MUEs and FUEs when the decentralized resource allocation tackles the resource block (RB) optimization or not assuming an ideal backhaul link ( $B=\infty \mathrm{Mbps}$ ), respectively. In both cases, the pricing versus non-pricing techniques are compared. We can observe that the pricing-based algorithms improve the non-pricing ones in terms of outage rate. For example, Figure 56 -right, shows that the $70 \%$ of FUEs get a throughput higher than $1 \mathrm{bps} / \mathrm{Hz}$ with the pricing-based algorithms, while the non-pricing one only get $0.5 \mathrm{bps} / \mathrm{Hz}$, getting a gain up to $100 \%$. When the RBs are not optimized and those resources have been set a priori, i.e. $50 \mathrm{RB}(10 \mathrm{MHz})$ to each served FUE in a FAP (see Figure 57), then the previous values are obtained by the $90 \%$ of FUEs, but at the cost or reducing the maximum throughput that can be obtained by a FUE. Notice that when the RBs are optimized we could get a solution where all RBs are assigned to one of the FUEs served by a FAP, situation not possible when resources are static.

In the considered scenario we can observe that the individual throughput of the FUEs is almost independent of the number of active FAPs $\left(\mathrm{FAP}_{\text {load }}\right)$, while the individual throughput of MUEs reduces as $\mathrm{FAP}_{\text {load }}$ increases. In this regard, we have simulated the same scenario with a single FAP area and the same priorities and we have observed that the FUE results are quite similar. This performance means that the interference coming from other FAPs in other FAP areas is almost insignificant when the FAP deployment probability is equal to 0.16 (1.6 FAPs per floor).


Figure 57. Cumulative density function (cdf) of the individual throughput of Left) MUEs and Right) FUEs for a sector-based corporate scenario when pricing and non-pricing techniques are considered where only precoders are optimized. The $\mathrm{FAP}_{\text {load }}=\{0.25,0.5,0.75,1\}$. Ideal backhaul. MUEs present a priority 5 times higher than FUEs. All FUEs have priority equal to 1.

Figure 58 and Figure 59 illustrate the spectral efficiency obtained by the different techniques as a function of the number of active FAPs in the sector. Interestingly we can observe that using the decentralized resource allocation with the RB optimization we improve the non-pricing approach also in terms of spectral efficiency for the FUEs (see Figure 58), while when the RB optimization is not considered, the spectral efficiency obtained by FUEs with pricing and non-pricing schemes is almost the same (see Figure 59).


Figure 58. Cumulative density function (cdf) of spectral efficiency attained by Left) MUEs and Right) FUEs for a sector-based corporate scenario when pricing and non-pricing techniques are considered where transmit precoders and RB are optimized. The $\mathbf{F A P}_{\text {load }}=\{0.25,0.5,0.75,1\}$. Ideal backhaul. MUEs present a priority 5 times higher than FUEs. All FUEs have priority equal to 1.



Figure 59. Cumulative density function (cdf) of spectral efficiency attained by Left) MUEs and Right) FUEs for a sector-based corporate scenario when pricing and non-pricing techniques are considered where only transmit precoders are optimized. The $\mathbf{F A P}_{\text {load }}=\{\mathbf{0 . 2 5}, 0.5,0.75,1\}$. Ideal backhaul. MUEs present a priority 5 times higher than FUEs. All FUEs have priority equal to 1 .

Finally, Figure 60 depicts how the spectral efficiency attained by FUEs and MUEs evolves as a function of different qualities on the backhaul link and number of active FAPs. The pricing technique becomes useful from the point of view of FUEs when there is an ideal backhaul. Notice that in this simulated scenario all FUEs have the same priority. On the other hand, pricing enhances the nonpricing performance attained by MUEs in all cases. It has to be remarked that as we increase the quality of the backhaul, the FUEs improve their spectral efficiency. Therefore, its impact of the overall weighted sum rate (considering FUEs and MUEs) also increases. Since the MBS is a dominant interferer for all FUEs, the MBS has to reduce its transmitted power and the spectral efficiency attained by MUEs decrease.


Figure 60. Average Spectral Efficiency attained by Left) FUEs and Right) MUEs when $B=\{\mathbf{1 0}, 20,40,80, \infty\}$ Mbps and $\mathrm{FAP}_{\text {load }}=\{0.75,1\}$. Pricing and non-pricing techniques with transmit precoders and RB assignment optimization are considered. MUEs present a priority 5 times higher than FUEs. All FUEs have priority equal to 1

### 9.1.3 Conclusions

- Pricing techniques are suitable for interference-limited scenarios. The backhaul link influences on the maximum power transmitted, and the generated interference.
- Pricing techniques are able to improve gains in the order of $5-15 \%$ in terms of spectral efficiency
- Pricing allows controlling the quality of service, improving the minimum rate ( $10 \%$-outage rate) by a factor of 2-3 times, i.e. gains of 200-300\% over the non-pricing techniques
- The interference generated by other FAP areas to FUEs is almost insignificant because the FAPs are placed indoors and they are not transmitting at full power in all cases because their FUEs are nearby. .
- We have observed that the resource allocation techniques that optimize transmit precoders and RB assignment are able to get gains in terms of spectral efficiency and outage rate when pricing exchange is considered. However, when only transmit precoders are optimized, the obtained gains are in terms of outage rate.


### 9.2 Resource allocation based on power minimization

This section will present simulation results for the proposed approach in the FREEDOM corporate scenario shown in Figure 51.

Specific aspects of the simulations results included in this section are:

- MBS transmissions are considered in the FREEDOM small corporate scenario. While the MBS sends cost information to neighbor FAPs to prevent them to transmit at certain directions/carriers, the MBS ignores any pricing information from the FAPs.
- In the FREEDOM small corporate scenario, FAP and MBS nominal powers are respectively 20 dBm and 46 dBm . This power can never be surpassed, even in the initial iterations of the minimum power proposed algorithm. If at some iteration, a serving station (FAP or MBS) requires transmitting a higher power than this to fulfil the target rate, it will not be allowed. In this situation, we have use the QoS readjustment algorithm proposed in [Munoz11a] to reduce the target rate at this iteration in order to keep the power in the margin $(19.5,20] \mathrm{dBm}$ and $(45.5,46] \mathrm{dBm}$ for the FAPs and MBS respectively4. If the serving stations do not need to exceed the maximum power for the served UEs to achieve the target rate, the total transmission power will be minimized.

All the results presented in this section have been obtained for two target rates, $2 \mathrm{bps} / \mathrm{Hz}$ and $4 \mathrm{bps} / \mathrm{Hz}$ (the same target rate is considered for all FUEs and MUEs), and for the following cases regarding FAP antennas and UE antennas: $2 \times 2$ (MIMO), $2 \times 1$ (MISO), and $1 \times 1$ (SISO). The MBS is considered always with 4 antennas.

[^3]9.2.1 Results for a single scenario with a fixed number of FAPs

The first set of results corresponds to a single realization of the scenario depicted in Figure 51. A FAP area with 32 active FAPs has been deployed within a sector of the MBS. In the same sector, the MBS is serving 2 MUEs. Each FAP has 1 FUE connected to.

Figure 61 and Figure 62 depict the average rate received by FUEs and MUEs and transmitted power of FAPs and MBS, respectively. For the $2 \mathrm{bps} / \mathrm{Hz}$, all the schemes allow the FUEs to achieve the target rate with the serving FAP transmitting well below the maximum transmission power. In such a case, it is clear the advantage of using pricing in terms of the average FAP power (equivalently total transmission power) for all the antenna schemes: MIMO $(2 \times 2)$, MISO $(2 \times 1)$, and SISO $(2 \times 1)$.


Figure 61. (Top) Average rate per FUE for a UE target rate of $2 \mathrm{bps} / \mathrm{Hz}$ (top-left) and 4 bps/Hz (top-right). (Bottom) Average rate for the two MUEs, considering a UE target rate of $2 \mathrm{bps} / \mathrm{Hz}$ (bottom-left) and $4 \mathrm{bps} / \mathrm{Hz}$ (bottom-right). MUEs randomly deployed in the sector.

When increasing the target rate up to $4 \mathrm{bps} / \mathrm{Hz}$, the SISO scheme fails to provide the target rate with and without pricing (see in Figure 61-top-right). However, the achieved rate in the SISO case is higher when pricing is used, achieving the same value than the MISO $(2 \times 1)$ scheme without pricing. There is no power saving in the SISO case because of the pricing, due to the fact that, for the SISO scheme and target rate of $4 \mathrm{bps} / \mathrm{Hz}$, a lot of FAPs are in the limited power zone, and the limitation is the same for both pricing and no pricing cases, see Figure 62 -top-right. As for the MISO $(2 \times 1)$ case, while this scheme fails to provide the target rate without pricing, it can provide the target rate when pricing is used. Furthermore, the use of pricing allows reducing the transmission power significantly. For the MIMO ( $2 \times 2$ ) case, there is no difference regarding the achieved rate with and without pricing, as in
any case the target rate is achieved, but there is also a significant saving in the transmission power if pricing is used (see Figure 62-top-right).

The MBS always transmits with 4 antennas. The two MUEs are in this particular realization far from the FAP area and close to the MBS. Furthermore the MBS does not take into account the impact of its own transmissions on FUEs. This explains why there is no difference between pricing and no pricing for the MUEs performance and the MBS power, see Figure 61-bottom and Figure 62-bottom, respectively. Notice that the results are independent of the number of antennas in the FAPs, therefore the performance for the MISO $(2 \times 2)$ and SISO $(1 \times 1)$ case are equal, as the only change is on the number of FAP antennas.


Figure 62. (Top) Average power per FAP (\#FAPs=32) for a UE target rate of $2 \mathrm{bps} / \mathrm{Hz}$ (top-left) and $4 \mathrm{bps} / \mathrm{Hz}$ (top-right). (Bottom) MBS power for a UE target rate of 2 $\mathrm{bps} / \mathrm{Hz}$ (bottom-left) and $4 \mathrm{bps} / \mathrm{Hz}$ (bottom-right). MUEs randomly deployed in the sector.

Instead of deploying the MUEs close to the MBS and far from the FAP area, for the next set of results the MUEs have been placed randomly within the FAP area. Therefore, the conditions for the MUEs become worse and the MBS must increase the power to fulfill the rate constraint, interfering more the FUEs. On the other hand, when pricing is considered, the FAPs should allocate their resources to avoid degrading MUEs performance, which may result in some degradation for the FUEs performance, as the MBS does not act in the same way. In such a situation, the next figures depict the average rate at the FUEs and MUEs (2), and average transmitted power by FAPs and MBS in Figure 63 and Figure 64, respectively.


Figure 63. (Top) Average rate per FUE for a UE target rate of $2 \mathrm{bps} / \mathrm{Hz}$ (top-left) and 4 bps/Hz (top-right). (Bottom) Average rate for the two MUEs, considering a UE target rate of $2 \mathrm{bps} / \mathrm{Hz}$ (bottom-left) and $4 \mathrm{bps} / \mathrm{Hz}$ (bottom-right). MUEs in the FAP area.

For the FAPs and FUEs similar conclusions than from Figure 61 and Figure 62 can be extracted. The system is now more constrained for the reasons explained above. It is worthy of comment that the difference between pricing and no pricing in term of transmission power decreases when the station is transmitting with very high power (observe the SISO case in Figure 64-top-right), as the maximum value of the transmission power is limited. The major benefits from pricing in terms of power occur when the system is not too constrained (that is, when the power does not need to be close to the maximum value), but the required power is not too low (as if the required power is too low, it is likely because the experienced interference is too low, so pricing is not necessary).

Now the interference from FAPs to the MUEs is significant. Therefore, the MBS is forced to transmit at its maximum power (see Figure 64-bottom). Even doing so, the target rate cannot be achieved by the MUEs placed within the FAP area (see Figure 63-bottom). Nevertheless, the use of pricing allows the MUEs to achieve a higher rate, as the pricing will make the FAPs to transmit less power and also to allocate their resources to reduce the degradation on the MUEs performance.


Figure 64. (Top) Average power per FAP (\#FAPs=32) for a UE target rate of $2 \mathrm{bps} / \mathrm{Hz}$ (top-left) and $4 \mathrm{bps} / \mathrm{Hz}$ (top-right). (Bottom) MBS power for a UE target rate of 2 $\mathrm{bps} / \mathrm{Hz}$ (bottom-left) and $4 \mathrm{bps} / \mathrm{Hz}$ (bottom-right). MUEs in the FAP area.

### 9.2.2 Results for several scenarios

In the following we are going to show the CDF of the FUEs/MUEs rate in Figure 65 and CDF of the transmitted power by FAPs/MBS in Figure 66. To that end 50 independent scenarios are generated. In each of them, the FAP area is placed randomly within the sector. The average number of active FAPs is 30 , despite it is a random variable, and each FAP is serving 1 FUE. The MBS is serving 2 MUEs randomly placed within the sector.

Observing the CDF of the FUE rates in Figure 65 -top-left, we see that for a target rate of $2 \mathrm{bps} / \mathrm{Hz}$, all the FUEs achieve the target rate if pricing is used. However, if pricing is not used, $5 \%$ of the users in the MISO ( $2 \times 1$ ) case and $15 \%$ of the users in the SISO ( $1 \times 1$ ) case will not achieve the target rate. The percentage increases when increasing the target rate (see Figure 65 -top-right), up to $25 \%$ and near $40 \%$ for MISO and SISO without pricing. In such a case, the use of pricing allows to increase the number of satisfied users a $10 \%, 20 \%$ and $10 \%$ for the SISO, MISO and MIMO case respectively.

Also from the CDF of the FAP power in Figure 66-top, power savings can be observed when pricing is used. Taking for instance, a target rate of $4 \mathrm{bps} / \mathrm{Hz}$, for MIMO and MISO case, we observe a reduction between $5 \%$ and $15 \%$ respectively of the number of FAPs transmitting at the maximum power if pricing is used.

Finally, for the MBS, as the probability for the MUEs to be close to the FAP area is low, there is no difference between pricing and no pricing (see Figure 65-bottom and Figure 66-bottom).


Figure 65. (Top) FUE rate CDF for a UE target rate of $2 \mathrm{bps} / \mathrm{Hz}$ (top-left) and $4 \mathrm{bps} / \mathrm{Hz}$ (top-right). (Bottom) MUE rate CDF for a UE target rate of $2 \mathrm{bps} / \mathrm{Hz}$ (bottom-left) and 4 bps/Hz (bottom-right). MUEs randomly placed in the sector.


Figure 66. (Top) FAP power CDF for a UE target rate of $2 \mathrm{bps} / \mathrm{Hz}$ (top-left) and 4 bps/Hz (top-right). (Bottom) MBS power CDF for a UE target rate of $2 \mathrm{bps} / \mathrm{Hz}$ (bottomleft) and $4 \mathrm{bps} / \mathrm{Hz}$ (bottom-right). MUEs placed randomly in the sector.

However, if the MUEs are placed randomly within the FAP area, the situation is different and in this case, the use of pricing helps to both reducing the power and increasing the achieved rate (which is however below the target rate). Figure 67 correspond to the CDF of the FUEs/MUEs rate while Figure 68 stands for the FAPs and MBS power.

As expected, the FAPs are more constrained in this set up, but the same conclusions as before can be extracted, with a greater advantage in terms in the achieved rate for pricing versus no pricing.


Figure 67. (Top) FUE rate CDF for a UE target rate of $2 \mathrm{bps} / \mathrm{Hz}$ (top-left) and $4 \mathrm{bps} / \mathrm{Hz}$ (top-right). (Bottom) MUE rate CDF for a UE target rate of $2 \mathrm{bps} / \mathrm{Hz}$ (bottom-left) and 4 $\mathrm{bps} / \mathrm{Hz}$ (bottom-right). MUEs in the FAP area.


Figure 68. (Top) FAP power CDF for a UE target rate of $2 \mathrm{bps} / \mathrm{Hz}$ (top-left) and 4 bps/Hz (top-right). (Bottom) MBS power CDF for a UE target rate of $2 \mathrm{bps} / \mathrm{Hz}$ (bottomleft) and $4 \mathrm{bps} / \mathrm{Hz}$ (bottom-right). MUEs in the FAP area.

### 9.2.3 Results for varying FAP density

To conclude this section, Figure 69 shows the percentage of FUEs and MUEs not achieving the target rate, while Figure 70 shows the percentage of FAPs transmitting in the limited power zone, both figures as a function of the average number of active FAPs. Finally, Figure 71 depicts the $50 \%$-tile of the MBS transmission power versus the average number of active FAPs, including the case of no active FAPs. Each curve is presented for a target rate of $2 \mathrm{bps} / \mathrm{Hz}$ and $4 \mathrm{bps} / \mathrm{Hz}$. The MUEs are randomly placed within the sector.

As the average number of active FAPs increases, the percentage of FAPs operating in the limited power zone and the percentage of FUEs not achieving the target rate increases. However, the use of pricing allows for the same percentage for a higher number of active FAPs.

When the MUEs are randomly placed within the sector, the performance is not affected by the FAPs density.


Figure 69. (Top) Percentage of FUEs not achieving the target rate vs. average \# active FAPs for a UE target rate of $2 \mathrm{bps} / \mathrm{Hz}$ (top-left) and $4 \mathrm{bps} / \mathrm{Hz}$ (top-right). (Bottom) Percentage of MUEs not achieving the target rate vs. average \# active FAPs for a UE target rate of $2 \mathrm{bps} / \mathrm{Hz}$ (bottom-left) and $4 \mathrm{bps} / \mathrm{Hz}$ (bottom-right). MUEs randomly placed in the sector.


Figure 70. Percentage of FAPs transmitting in the limited power zone vs. average \# active FAPs for a UE target rate of $2 \mathrm{bps} / \mathrm{Hz}$ (left) and $4 \mathrm{bps} / \mathrm{Hz}$ (right). MUEs randomly placed in the sector.


Figure 71. 50\%-tile MBS transmission power vs. average \# active FAPs for a UE target rate of $2 \mathrm{bps} / \mathrm{Hz}$ (left) and $4 \mathrm{bps} / \mathrm{Hz}$ (right). MUEs randomly placed in the sector.

Finally, for the case of the MUEs placed randomly within the FAP area, Figure 72 shows the percentage of FUEs/MUEs not achieving the target rate as a function of the average number of active FAPs. Similarly, Figure 73 shows the percentage of FAPs transmitting in the limited power zone and Figure 74 depicts the $50 \%$-tile of the MBS transmission power versus the average number of active FAPs, including the case of no active FAPs. Each curve is presented for a target rate of $2 \mathrm{bps} / \mathrm{Hz}$ and 4 $\mathrm{bps} / \mathrm{Hz}$. We can observe that when the MUEs are randomly placed within the sector, their performance is clearly degraded. For $4 \mathrm{bps} / \mathrm{Hz}$, the MBS needs to transmit almost always at its maximum power (see Figure 74), whatever the FAPs density, pricing or not. The pricing however, allows reducing the percentage of MUEs not achieving the target rate due to the combined effect of:
o The pricing will make the FAPs to transmit less power
o The pricing will make the FAPs to allocate their resources to reduce the degradation on the MUEs performance.


Figure 72. (Top) Percentage of FUEs not achieving the target rate vs. average \# active FAPs for a UE target rate of $2 \mathrm{bps} / \mathrm{Hz}$ (top-left) and $4 \mathrm{bps} / \mathrm{Hz}$ (top-right). (Bottom) Percentage of MUEs not achieving the target rate vs. average \# active FAPs for a UE target rate of $2 \mathrm{bps} / \mathrm{Hz}$ (bottom-left) and $4 \mathrm{bps} / \mathrm{Hz}$ (bottom-right). MUEs in the FAP area.


Figure 73. Percentage of FAPs transmitting in the limited power zone vs. average \# active FAPs for a UE target rate of $2 \mathrm{bps} / \mathrm{Hz}$ (left) and $4 \mathrm{bps} / \mathrm{Hz}$ (right). MUEs in the FAP area.


Figure 74. 50\%-tile MBS transmission power vs. average \# active FAPs for a UE target rate of $2 \mathrm{bps} / \mathrm{Hz}$ (left) and $4 \mathrm{bps} / \mathrm{Hz}$ (right). MUEs in the FAP area.

### 9.2.4 Conclusions

The benefit due to pricing increases, either in terms of total FAP power or percentage of users achieving the target rate, when the system is more constrained. The system is more constrained due to worse interference conditions occurring either because the FAP density increases or because the target rate increases as this makes the FAPs to increase the transmission power, increasing therefore the interference.

When the MUEs are placed close to a FAP area, the use of pricing allows the MUEs to achieve a higher rate compared to the case of no pricing. This is because the pricing mechanism allows the FAPs to transmit less power with the consequent reduction in the interference caused, and also because the cost values sent by the MBS makes the FAPs to allocate their resources to reduce the degradation on the MUEs performance.

### 9.3 Dynamic resource allocation under Markovian interference model

In this section we report some simulation results for the max-rate and min-power game with Markovian activity of the MBS analysed in section 7.3. The scenario is similar to the small corporate configuration introduced in Figure 51 and Table 22, with 2 buildings separated by a street. More specifically, we have considered 25 resource blocks each one composed of 12 subcarriers for a total number of available subcarriers equal to $N=300(20 \mathrm{MHz}$ bandwidth). In Figure 75 and Figure 76 we have plotted the sum rate per FAP and per OFDM symbol versus the number of time slots $M$ assuming, respectively, a number of active FAPs equal to 12 and 8.

The MBS activity is the same over all the subchannels and we have considered both the cases of pricing (dashed line), and no pricing (continuous line) algorithms by setting the overall transmit power per FAP equal to 1 . The three curves in each figure indicate the sum rate per FAP and per OFDM symbol by assuming perfect (non-causal) knowledge of the macro-user activity, no knowledge at all or statistical knowledge of the macro activity which is modelled as a Markov chain of the first order with idle-to-idle transition probabilities $\omega_{k}=\omega=0.1$ and busy-to-busy transition probabilities $\mu_{k}=\mu=0.1$.

Both the figures show that the statistical knowledge of the interference activity brings the performance close to the ideal case of perfect non-causal knowledge of the interference activity. Nevertheless we can observe a small gain in term of sum rate by applying the pricing algorithms due to the low interference level among the FAPs.

Although with a small increase of the maximum transmit power in order to get a higher interference level among the FAPs, we can observe from Figure 77 that the proposed pricing mechanisms yield a performance improvement with respect to the no pricing algorithms.


Figure 75. Sum rate for FAP and for OFDM symbol versus the number of time slots for 12 active FAPs.


Figure 76. Sum rate for FAP and for OFDM symbol versus the number of time slots for 8 active FAPs.


Figure 77. Sum rate for FAP and for OFDM symbol versus the number of time slots for 8 active FAPs and .

Finally in Figure 78 and Figure 79 we have considered the simulation results for the min power game with Markovian interference considering both the cases of pricing (dashed line), and no pricing (continuous line) algorithms. In order to have a fair comparison between the transmitted powers we have forced the same value of $R_{0}$ for the three cases of perfect (non-causal) knowledge of the macrouser activity, no knowledge at all or statistical knowledge of the macro activity. The parameters used in the simulations are $R_{0}=750$ bits per OFDM symbol, with transition probabilities $\omega_{k}=\omega=0.1$, $\mu_{k}=\mu=0.1$ while the MBS activity is the same for groups of 12 consecutive subcarriers.

From Figure 78 and Figure 79, respectively, for 8 and 12 active FAPs, we can note the performance improvement of the proposed statistical approach with respect to the one that considers the future exactly equal to the present. Furthermore we can observe that the pricing mechanisms allow a performance gain with respect to the no pricing algorithms by reducing the radiated power still maintaining the same service quality.


Figure 78. Average transmitted power vs. number of time slots for 8 active FAPs.


Figure 79. Average transmitted power vs. number of time slots for $\mathbf{1 2}$ active FAPs.

### 9.4 Decentralized vs. Centralized resource allocation

We evaluated the performance of the proposed algorithms for resource allocation under a common scenario deployment with a Macro Base-Station. One terminal deployment has been selected where there are up to 20 FAPs, each serving a single user, see Figure 51 and Table 22.

The considered decentralised algorithms is the iterative water-filling algorithm based on statistical inference of the probability of RB occupation by the MBS, that we call Statistical Inference Driven

Iterative Water Filling Algorithm presented in Subsection 7.3.2 (SIDIWFA for brevity). The centralised approach is the Genetic Algorithm described in Section 8.

In order to fully compare the results, both approaches share and implement the same assumptions:

- macro and femto networks fully share the same bandwidth;
- the system is synchronous, allowing the separation of the DL and UL phases of the MBS and FAPs.

These assumptions entail that, when the MBS is in DL phase, also the FAPs are performing their DL; thus the interference suffered by each FUE comes from the BS and the other FAPs. On the other hand, when the macro-network is in the UL phase, also the FUEs are performing their UL; thus the interference suffered by each FAP comes from the other FUEs and from all the MUEs. It is worth remarking that:

- these assumptions do not emphasise the advantages of the decentralized algorithm incorporating the statistical characterization of the MBS band usage;
- the following evaluations do not take into account the overhead due to the implementation of all the control plane-related procedures required by the algorithms;

Considering the above, [FREEDOM-D52] will provide additional comparisons, taking into account the above points, along with the dynamic behaviour of the MBS.


Figure 80. Comparison between centralized and decentralized resource allocation.
Figure 80 shows the results obtained with the two algorithms, the genetic centralized algorithm (GA) and the Statistical Inference Driven Iterative Water Filling Algorithm (SIDIWFA). The chosen performance metric is the per-FAP average (achievable) rate, measured in bits per OFDM symbol interval. As expected, with the above discussed assumption of static MBS activity, the GA performs better, thanks to its centralized nature. We also remark that the above results were achieved assuming a static behavior of the MBS, i.e., for each frequency block, the MBS is either always active, or always idle. Particularly, we have considered a completely silent MBS, a MBS whose traffic saturates the whole bandwidth, and a case in which about a half of the frequency resource blocks are used while the other half is idle. We expect that the inclusion of the dynamic nature of the Macro users' traffic will show the advantages of including such time-dynamicity in the problem formulation at the basis of the algorithm derivation.

The indications coming from the performed comparison are however the offspring of the common assumptions, which are not tailored to fully exploit the features of the SIDIWFA. In this sense, the introduction (to be considered in [FREEDOM-D52]) of the overheads (associated with the control plane) and an uncoordinated activity of the MBS are expected to modify the indications coming from Figure 80

### 9.5 GO for realistic frequency selective channel - power constrained

The GO approach presented in section 8 has been implemented for a realistic propagation environment sketched in Figure 51, implementing the capacity in Eq. (138) with fitness function

$$
\begin{equation*}
f_{0}=-\frac{1}{2 N} \sum_{k=1}^{2 N} c_{k} \tag{145}
\end{equation*}
$$

and under the constraints

$$
\begin{equation*}
\sum_{f=1}^{F} P_{q}(f) \leq P_{q \text { tot }} \tag{145}
\end{equation*}
$$

recalled here for the sake of clarity.
Simulation has been performed for a single terminal configuration where up to $20 \mathrm{FAPs} / \mathrm{FUEs}$ pairs are deployed. The couples that are active is a variable that can take values of $\{4,6,8,10,12,14,16$, $18,20\}$.
In this Section are reported the details of results for the case of 4 couples and the overall results, leaving to the appendix (section 14.4.1) some details for the other cases. The scenario considered was implemented for the 5 MHz band, allocating 25 PRBs. For the case of N(couples), the GA optimizes PRB x N(couples) variables. The interference role of BS, FAPs and FUEs is evaluated based on the assumptions:

- Macro and femto networks fully share the same bandwidth;
- The system is synchronous, allowing the separation of the DL and UL phases of the MBS and FAPs.

These assumptions entail that, when the MBS is in DL phase, also the FAPs are performing their DL; thus the interference suffered by each FUE comes from the BS and the other FAPs. On the other hand, when the macro-network is in the UL phase, also the FUEs are performing their UL; thus the interference suffered by each FAP comes from the other FUEs and from all the MUEs.

### 9.5.1 Macro-network UL

Figure 81 reports the optimization for $P(f)$ where $f=1, . .25$, under constraint (145), considering the state of UL for the Macro network.





Figure 81. GO Power allocation for 4 couples of FAP/FUE users. Vertical axis is scaled for unitary power; overall transmission power in dB in the plots titles. Macro-network in UL.

| N. couples 4 |  |  |  |  |  |
| :---: | :---: | :---: | :---: | :---: | :---: |
| Avg. capacity <br> bit/s/OFDM | 5609.4 | 3733.8 | 2976.4 | 1669.3 |  |
| Avg. sum rate <br> bit/s/OFDM | 3497.2 |  |  |  |  |
| Total P dB | 18.9 | 19.4 | 19.3 | 20.0 |  |

Table 26. Summary of results for a system of 4 FAP/FUE couples. Macro-network in UL.

### 9.5.2 Macro-network DL

Figure 82 reports the optimization for $P(f)$ where $f=1, . .25$, under constraint (145), considering the state of UL for the Macro network.


Figure 82. GO Power allocation for 4 couples of FAP/FUE users. Vertical axis is scaled for unitary power; overall transmission power in dB in the plots titles. Macro-network in DL.

| N. couples 4 |  |  |  |  |  |
| :---: | :---: | :---: | :---: | :---: | :---: |
| Avg. capacity <br> bit//S/OFDM | 3432.5 | 593.6 | 1640.7 | 400.1 |  |
| Avg. sum rate <br> bit/s/OFDM | 1516.7 |  |  |  |  |
| Total P dB | 19.5 | 20.0 | 19.0 | 19.9 |  |

Table 27. Summary of results for a system of 4 FAP/FUE couples. Macro-network in DL.

### 9.5.3 Overall results

The optimization has been performed considering the UL and DL phase of the macro network and the overall results are presented in Figure 83. When the macro network is in UL, the interference is mainly due to other couples FAP/FUE and thus the mutual effect implies a significant impact of the increasing number of interferers. In the complementary case, when the macro network is in DL, the interference has a similar impact on all transmitters in the cell, giving a secondary and less important role to the interference component due to the presence of other FAPs/FUEs.


Figure 83. Average sum rate after GO for power allocation.

The plots in Figure 84 and Figure 85 report the total power allocation for the users after GO when the macro network is in UL or DL for three cases with 4, 10 and 20 FAPs showing how the overall power assignment has a slight variation around the maximum.



Figure 84. Average power allocation after GO during BS UL or DL. Left panel, 4 couples; right panel, 10 couples.


Figure 85. Average transmission power after GO in the case of $\mathbf{2 0}$ couples. Comparison of power allocation during BS UL or DL.

The plots reported in Figure 86 and Figure 87 report for two scenarios of interest corresponding to 10 and 20 couples, sketching the capacity of the single users, measured in bit/s/OFDM, showing how the interference of the macro BS strongly affects the performances of each transmitter.


Figure 86. Capacity per user after GO in the case of 10 couples FAP/FUE. Comparison during BS UL or DL.


Figure 87. Capacity per user after GO in the case of 20 couples FAP/FUE. Comparison during BS UL or DL.

## 10 IMPLEMENTATION NOTES

### 10.1 Scalability

The investigated resource allocation algorithms in this work are easily scalable as the number of the number of terminals increases. Algorithms proposed in sections 5, 6 and 7 are based on the message exchange at control-plane level at a local level, i.e. among neighboring FAPs and terminals. This imposes that the overhead due to the control-plane signalling does not scale with the total number of nodes. Rather, the impact of such overhead is determined by the nodes' density and the interference range (given by the propagating conditions). A study of the control traffic in the backhaul during the exchange of prices is done in D4.2, where it is shown its relatively modest impact.

The algorithms tend to reach a stable solution in less than 15-20 iterations, irrespective of the number of terminals in the system

Although the genetic optimization-based algorithm analysed in section 8 is centralized, it can also deal with scalability in a corporate scenario, where each building might have a central processor unit, where all the resources of the FAPs deployed in that building are optimized. Since FAPs are placed indoors and transmit with low power, in general the interference from one building to another is not so significant. The algorithm needs of the SNIRs of all transmitters to be sent to the central processor unit and provides a sub-optimal solution for the resource allocation. One of the main drawbacks is the convergence time that needs of the order of hundreds of subframes.
10.2 Applicability

Table 28. Applicability of the proposed techniques

| Technical contribution | Main enhancements required in the current LTE specifications to support FREEDOM solutions |  |
| :---: | :---: | :---: |
|  | PHY layer | MAC layer |
| Identification of the interference by coordinated channel sensing | Availability of spectrum sensing procedures at the nodes over the whole bandwidth. | The algorithm is applicable, in a general sense to any MAC which schedules femto users in a given way but foresees the sensing of the presence of a Macro user signal. With respect to LTE, the MAC protocol is inherently different due to the fact that power allocation across subcarriers, or block of subcarriers, is a determined by the sensing performance. |
| Decentralized resource allocation in DL based on pricing <br> and <br> Decentralized resource allocation based on game theory | - Availability of (complex) channel knowledge of active users of own and neighboring FAPs. <br> - Possibility of applying waterfilling-based power-allocation per RB. | - Knowledge of the neighboring terminals in order to estimate the links (information about the pilot sequences to be scanned). <br> - Define messages (information elements) to support the pricing exchange through the X2 interface over the players involved in the resource allocation. <br> - A message can employ different MCS over the different RB. |
| Centralized dynamic interference management | No specific modification at PHY layer. <br> The feasibility of a centralised interference management requires specific resources at network level and application level (e.g. the presence of a dedicated entity performing the computation, the allocation of dedicated resources of the backhaul messaging for the exchange of relevant information). | No specific modification at MAC layer. <br> The feasibility of a centralised interference management requires specific resources at network level and application level (e.g. the presence of a dedicated entity performing the computation, the allocation of dedicated resources of the backhaul messaging for the exchange of relevant information). |

10.3 Complexity

Table 29. Complexity of the proposed techniques

| Technical contribution | Complexity of the optimisation procedures | Complexity at the MBS | Complexity at the FAP | Complexity at the UE | Complexity at network management level |
| :---: | :---: | :---: | :---: | :---: | :---: |
| Identification of the interference by coordinated channel sensing | - A set of cooperating FAP's performs cooperative sensing to characterize the status of the network in terms of active users <br> - For each sensing interval, the procedure requires each FAP to disseminate its sensing outcome to its neighbors | - None, the MBS is not involved | - Each FAP must run a local sensing algorithm and a global one to fuse the sensing outcomes from its neighbors. | - FUE's not involved | - The exchange of sensing information exploits the backbone link |
| Decentralized resource allocation in DL based on pricing <br> (bitrate, power allocation and carrier assignment) | - Pricing values are generated per resource block and distributed among interfering FAPs <br> - Impact of CSI quantization | - The channel coefficients to the serving FAP and to the interfering FAP have to be measured at the FUE <br> - In the decentralized approach, the MBS can be seen as an additional FAP | - If there are multiple FUEs per FAP and OFDMA is assumed, the complexity of allocating the best resource block is proportional to the number of carriers times the number of users, and it is done at each FAP <br> - Under the non-orthogonal user-access per FAP the transmitters can be complex (based on dirty paper coding, DPC) or simple (using superposition coding, SC) | - When the user-access per FAP is orthogonal, then simple receivers can be used at the UE (FUE or MUE) <br> - When the user-access in nonorthogonal and transmitters are DPC-based devices, then simple receivers can be considered. However, if SC transmitters are used, then successive decoding receivers are required. In this latter case, the UE must be informed about the proper decoding order. | - The procedure to exchange the pricing values among FAPs is not standardized. Some modification of the X2 interface is needed. <br> - Meeting the conditions for convergence of gametheory based methods might require the presence of new network elements. This is under study at this time. |


| Decentralized resource allocation based on game theory | - Channel estimates (or equivalent information) and pricing values per subcarrier are calculated by each FAP and communicated to the other interfering FAPs. | - None: the MBS's are not involved in the procedure | - Channel estimation is required at the FAPs: channel estimation is required even for the interfering channels. | - Initially, a simple FAP-FUE link is considered; CSI is required at the FUE's if the algorithm is used for the uplink | - A suitable protocol to exchange the pricing values among FAPs must be devised for the backbone network. |
| :---: | :---: | :---: | :---: | :---: | :---: |
| Centralized dynamic interference management | - Transmitters undergoing optimization need to regularly evaluate their SNR and/or CSI and feed it back to the centralized processor. <br> - At application level, the system should provide a dedicated IP two-ways channel. <br> - Can be implemented either as a centralized or as a distributed algorithm. | - $\mathrm{n} / \mathrm{a}$ | - The backhaul transmission should allow establishing a dedicated channel for signalling. | - The communication link should allow establishing a dedicated channel for signalling, at least at application level. | - The optimization is based on a computation performed at Network level <br> - In the case of distributed implementation, signalling is required in DL only. |

## 11 LTE-BASED RESOURCE ALLOCATION

In contrast to previous sections, here we present different techniques that consider the constraints imposed by the current LTE-A standard.

### 11.1 LTE-A adapted pricing mechanisms

Two alternative forms of pricing mechanisms have been proposed to deal with the mutual interference among neighboring FAPs. One is focused on the maximization of the weighted sum-rate, while the other is focused on the minimization of the total transmission power while guaranteeing a minimum bit rate per user [Munozl1b]. Both mechanisms require the exchange of pricing information among neighboring FAPs through the backhaul, including interference sensitivity, Lagrange multipliers and cross channels. Also, each user needs to report to the serving FAP the signal to noise plus interference ratio (SNIR) measured at each resource block (or minimum resource allocation unit). With this information, the proposed mechanisms optimize the power allocation per carrier, according to one of the above mentioned criteria.

Despite the promising results, applying the proposed algorithms to LTE-A is not straightforward. Last version of LTE has some specific features that need to be taken into account:

- In LTE, there is no DL power control
- In LTE, there is no direct report of the Signal-to-Interference plus Noise Ratio (SNIR). Instead the maximum modulation and coding scheme (MCS) for a common group of resource blocks $(\mathrm{RB})$ is reported.
- In LTE, all the RBs allocated to one user at each sub-frame use the same MCS

Customizing the proposed pricing algorithms to LTE may therefore require changes at both LTE and the algorithms themselves:

- Some modifications in LTE are necessary in order to support the proposed algorithms. Indeed, MCS sensitivity to interference (to obtain pricing) is an intrinsic feature of the proposed algorithms, so new procedures are unavoidable, as well as new messages to exchange a quantized version of this information between FUE-FAP, FAP-FAP and FAP-MBS.
- At the same time, some simplifications can be done in the proposed algorithms to make them fit better within LTE, such as on/off transmission per RB, use of MCS instead of SNIR, etc.

In this section we will present a simpler and better LTE-adapted solution than those presented in sections 5, 6 and 7 , and the unavoidable modifications for LTE to support this simpler scheme will be described. The contents of this section have been presented in the last November meeting of 3GPP [R3-112752].

### 11.1.1 A pricing based mechanism for MCS and bandwidth part selection.

The goal of this section is to provide a coordinated mechanism to maximize the sum of the MCSs (and equivalently the sum of transmission rates) employed by a set of nearby FAPs. While the MBS may perform carrier aggregation (CA), the FAPs operate as a single carrier system, with different channel bandwidths. The proposed mechanism can be employed to select the DL operational frequency for wideband channels (e.g. 5 MHz ) and the bandwidth part for broadband channels (i.e. $10 \mathrm{MHz}, 15 \mathrm{MHz}$ or 20 MHz ). Also, given the operational frequency/bandwidth part, the same mechanism can be employed to select the sub-bands if the DL traffic is scheduled in a portion of the channel bandwidth or a portion of the previously selected bandwidth part. To describe the procedure, we will focus in the first case, i.e. selection of the operational frequency/ bandwidth part.

Figure 88 illustrates possible bandwidth partitions within the frequency channel, i.e. bandwidth part. Note that the width of a bandwidth part is approx. 5 MHz , when the total channel BW is 20 MHz . For a 20 MHz channel width, a bandwidth part has the same number of resource blocks as a 5 MHz carrier.


Figure 88. Example of bandwidth partitions

### 11.1.1.1 Fundamentals

Let us consider the following notation:

- $\quad N_{F}$ : Total number of serving stations (FAPs and, maybe, also the MBS).
- $\{S\}$ : Set of available bandwidth parts in the system
- $\left\{S_{f}\right\}$ : Set of available bandwidth parts for the $f$-th serving station.
- $M C S_{u}^{s}$ : Modulation and coding scheme reported by user $u$ in the bandwidth part $s$
- $\quad b_{f}^{s}$ : Binary variable which equals 1 if the $f$-th FAP transmits in the $s$-th bandwidth part, and 0 otherwise. Notice that $b_{f}^{s}=0$, if $s \notin\left\{S_{f}\right\} \subset\{S\}$.
- $\quad P_{f}$ and $P_{g}$ : Transmission power per resource block (RB) for the $f$-th and the $g$-th serving station respectively
- $\left|h_{u, g}^{s, r}\right|$ and $\left|h_{u, g}^{s, r}\right|$ : Cross- channel amplitudes between the user $u$ and the $f$-th and $g$-th serving station, respectively, on the $r$-th RB in the $s$-th bandwidth part.
- $\quad i_{u}^{s, r}$ : Interference measured by the user $u$ at the the $r$-th resource block in the $s$-th bandwidth part.
- $\quad \mathbf{i}_{u}^{s}$ : Vector containing the interference measured the user $u$ for the $N$ RBs in the $s$-th bandwidth part, i.e, $\mathbf{i}_{u}^{s}=\left[i_{u}^{s, r=1}, \ldots, i_{u}^{s, r=N}\right]^{T}$
- $\left\{U_{f}\right\}$ : Set of users connected to the $f$-th serving station.

To simplify notation, we consider initially that each serving station is serving a single user, and we will use $u(f)$ to denote the user connected to the $f$-th serving station, i.e., $\left\{U_{f}\right\}=\{u(f)\},\left|U_{f}\right|=1$. The signal to noise plus interference ratio measured at the $r$-th RB in the $s$-th bandwidth part by the user connected to the $f$-th serving station is given by:

$$
\begin{equation*}
\operatorname{SNIR}_{u(f)}^{s, r}=\frac{\left|h_{u(f), f}^{s, r}\right|^{2} P_{f}}{1+\sum_{g \neq f}^{N_{F}} b_{g}^{s, r}\left|h_{u(f), g}^{s, r}\right|^{2} P_{g}} \tag{146}
\end{equation*}
$$

While the channel could remain constant within each RB, it is unlikely that it remains constant on all the RBs in the considered band. When all the subcarriers of a specific user are modulated using the same modulation MCS a function that maps the set of SINRs to a single value of MCS for a given Block Error Rate (BLER) is required. For the $s$-th bandwidth part, the user $u$ computes a MCS value based on the SNIR measures for the $N$ RBs that compose this bandwidth part:

$$
M C S_{u}^{s}=\operatorname{function}\left(S N I R_{u}^{s, r=1}, \ldots, S N I R_{u}^{s, r=N}\right)
$$

The maxrate problem to solve by the $f$-th serving station is the following:

$$
\begin{array}{ll}
\max _{\left\{b_{f}^{s}\right\}} & \sum_{s \in\{S\}} b_{f}^{s} M C S_{u(f)}^{s} \\
\text { s.t. } & b_{f}^{s}=\{0,1\} \text { for } s \in S_{f}, \\
& b_{f}^{s}=0 \text { for } s \notin S_{f}, \\
& \sum_{s \in\{S\}} b_{f}^{s}=1 . \tag{150}
\end{array}
$$

Let us consider now the total rate achieved within a set of $N_{F}$ neighbour serving stations. In such a case, problem (147)-(150) becomes:

$$
\begin{array}{ll}
\max _{\left\{b_{f}^{s}\right\}_{f\left(1, \ldots N_{F}\right.}} & \sum_{f=1}^{N_{F}} \sum_{s \in\{\{ \}\}} b_{f}^{s} M C S_{u(f)}^{s} \\
\text { s.t. } & b_{f}^{s}=\{0,1\} \text { for } s \in S_{f}, f=1, \ldots, N_{F}, \\
& b_{f}^{s}=0 \text { for } s \notin S_{f}, f=1, \ldots, N_{F}, \\
& \sum_{s \in\{\{ \}\}} b_{f}^{s}=1, \text { for } f=1, \ldots, N_{F} . \tag{154}
\end{array}
$$

The interference measured at the the $r$-th resource block in the $s$-th sub-band by the user connected to the $f$-th serving station, $u(f)$, is

$$
\begin{equation*}
i_{u(f)}^{s, r}=\left.\sum_{\substack{g=1 \\ g \neq f}}^{N_{F}} b_{g}^{s} P_{g}| |_{u(f), g}^{s, r}\right|^{s} \tag{155}
\end{equation*}
$$

Due to the cross-interference, the modulation and coding scheme reported at each bandwidth part by the user connected to the $f$-th serving station will change if the other serving stations are transmitting or not in this bandwidth part. Let us use the following linear approximation for the $M C S_{u(f)}^{s}$ which depends on $\left\{b_{g}^{s}\right\}_{g \neq f}$ :

$$
\begin{align*}
M C S_{u(f)}^{s} \approx & \left.M C S_{u(f)}^{s}\right|_{\mathrm{i}_{f(0)}(0)}+\left.\sum_{r=1}^{N} \frac{\partial M C S_{f}^{s}}{\partial \partial_{u(f)}^{s, r}}\right|_{i_{i(f)}^{s}(0)} \Delta i_{u(f)}^{s, r}= \\
& \left.M C S_{u(f)}^{s}\right|_{i_{u(f)}^{s}(0)}-\sum_{r=1}^{N} \pi_{u(f)}^{s, r}\left(\sum_{\substack{g=1 \\
g \neq f}}^{N_{F}}\left(b_{g}^{s}-b_{g}^{s}(0)\right) P_{g}\left|h_{u(f), g}^{s, r}\right|^{2}\right) \tag{156}
\end{align*}
$$

where $\mathbf{i}_{u(f)}^{s}$ is the vector containing the interference measured for the $N$ RBs in the $s$-th bandwidth part, $\mathbf{i}_{u(f)}^{s}(0)$ and $b_{g}^{s}(0)$ denotes the previous value for $\mathbf{i}_{u(f)}^{s}$ and $b_{g}^{s}$ respectively, and $\pi_{u(f)}^{s, r}$ is the interference sensitivity factor defined as follows:

$$
\begin{equation*}
\pi_{u(f)}^{s, r}=-\left.\frac{\partial M C S_{u(f)}^{s}}{\partial u_{u(f)}^{s, r}}\right|_{\dot{i}_{u(f)}^{s,}(0)} \tag{157}
\end{equation*}
$$

Using the linear approximation in (156), the cost function in (151) can be rewritten as follows:

$$
\begin{equation*}
\max _{\left\{b_{f}^{s}\right\}_{f=1, \ldots, N_{F}}} \sum_{f=1}^{N_{F}} \sum_{s \in\{S\}} b_{f}^{s}\left(\left.M C S_{u(f)}^{s}\right|_{\mathbf{i}_{f}^{s}(0)}-\sum_{r=1}^{N} \pi_{u(f)}^{s, r}\left(\sum_{\substack{g=1 \\ g \neq f}}^{N_{F}}\left(b_{g}^{s}-b_{g}^{s}(0)\right) P_{g}\left|h_{u(f), g}^{s, r}\right|^{2}\right)\right) \tag{158}
\end{equation*}
$$

Notice that the terms for the $f$-th serving station depend on the other station variables, which results in a highly complex coupled problem. In order to derive a solution, we consider a set of distributed problems in which, at a given time, each serving station variables, i.e. $\left\{b_{f}^{s}\right\}$ are optimized while the variables for the rest of the transmitters are fixed. Under this assumption, the terms in the cost function (158) that depend on the $f$-th transmitter are:

$$
\begin{align*}
& \sum_{s \in\{S\}} b_{f}^{s} M C S_{u(f)}^{s}+\sum_{\substack{j=1 \\
j \neq f}}^{N_{F}} \sum_{s \in\{S\}} b_{j}^{s}\left(-\sum_{r=1}^{N} \pi_{u(j)}^{s, r}\left(b_{f}^{s} P_{f}\left|h_{u(j), f}^{s, r}\right|^{2}\right)\right)= \\
& \sum_{s \in\{S\}} b_{f}^{s}\left(M C S_{u(f)}^{s}-\sum_{\substack{j=1 \\
j \neq f}}^{N_{F}} b_{j}^{s}\left(\sum_{r=1}^{N} \pi_{u(j)}^{s, r} P_{f}\left|h_{u(j), f}^{s, r}\right|^{2}\right)\right) \tag{159}
\end{align*}
$$

Assuming that other serving stations' variables are given, the problem can be decomposed in subproblems, where each FAP optimizes its own variables, and the problem to be solved by the $f$-th serving station is therefore:

$$
\begin{array}{ll}
\max _{\left\{b_{f}^{s}\right\}} \sum_{f=1}^{N_{F}} \sum_{s \in\{S\}} b_{f}^{s}\left(M C S_{u(f)}^{s}-\sum_{\substack{g=1 \\
g \neq f}}^{N_{F}} b_{g}^{s}\left(\sum_{r=1}^{N} \pi_{u(g)}^{s, r} P_{f}\left|h_{u(g), f}^{s, r}\right|^{2}\right)\right) \\
\text { s.t. } & b_{f}^{s}=\{0,1\} \text { for } s \in S_{f}, \\
& b_{f}^{s}=0 \text { for } s \notin S_{f}, f=1, \ldots, N_{F}, \\
& \sum_{s \in\{S\}} b_{f}^{s}=1 . \tag{163}
\end{array}
$$

When comparing problem (160)-(163) with problem (147)-(150), we observe that the solution now is to select the bandwidth part, not with the highest reported MCS, but with the highest difference between the reported MCS and the pricing term

$$
\begin{equation*}
\text { price }_{f}^{s}=\sum_{\substack{g=1 \\ g \neq j}}^{N_{F}} b_{g}^{s}\left(\sum_{r=1}^{N} \pi_{u(g)}^{s, r} P_{f}\left|h_{u(g), f}^{s, r}\right|^{2}\right) \tag{164}
\end{equation*}
$$

This pricing term measures the degradation on other stations MCS due to the interference generated by the $f$-th serving station. It is computed as the sum over all neighbor serving stations of the cost that the transmission of the $f$-th serving stations has in terms of the supported MCS in the $s$-th bandwidth part:

$$
\begin{equation*}
\operatorname{cost}_{g, f}^{s}=\sum_{r=1}^{N} \pi_{u(g)}^{s, r} P_{f}\left|h_{u(g), f}^{s, r}\right|^{2} \tag{165}
\end{equation*}
$$

This cost is equal to the sum, over the $N$ resource blocks in the $s$-th sub-band, of the product of three terms:

1. The term $\pi_{u(g)}^{s, r}$, which measures the sensitivity, of the MCS reported for the $s$-th sub-band due to the interference in the $r$-th RB (see eq. (157)).
2. The transmission power of the $f$-th serving station.
3. The cross channel gain in the $r$-th RB of the $s$-th sub-band between the $f$-th serving station and the user connected to the neighbour serving station $g$.

Notice that the impact on the price term defined (164) of the cost for the $g$-th serving station due to the transmission by the $f$-th serving station in the $s$-th bandwidth part, $\operatorname{cost}_{g, f}^{s}$, is cero if $b_{g}^{s}=0$, i.e. if the $g$-th serving station does not use of the $s$-th bandwidth part.

In the proposed approach, each serving station updates its resource allocation strategy assuming that other FAPs' variables are given. In practice, several approaches are possible, i.e., the serving stations may perform synchronously or asynchronously (either at random time instances or sequentially). In order to speed up convergence, we will consider that the serving stations update the resource allocation simultaneously at each frame. When the serving stations update the resource allocation strategy simultaneously a problem that may happen is that they go back and forward on the same RBs. To avoid this undesired effect we propose to use some memory to update the pricing values. Therefore, for the $n$-th frame, the pricing values are computed as follows:

$$
\begin{equation*}
\operatorname{price}_{f}^{s}(n)=\alpha \operatorname{price}_{f}^{s}(n-1)+(1-\alpha) \sum_{\substack{g=1 \\ g \neq f}}^{N_{F}} b_{g}^{s}\left(\sum_{r=1}^{N} \pi_{u(g)}^{s, r} P_{f}\left|h_{u(g), f}^{s, r}\right|^{2}\right) \tag{166}
\end{equation*}
$$

with $\alpha$ a parameter between 0 a 1 .

### 11.1.1.2 Procedure for coordinated bandwidth part selection

Based on the above, our interference coordination approach involves the following steps:

1. A potential interference victim UE identifies the interference created by each one of the surrounding cells operating in the considered frequency band;
2. Based on this information, the UE determines which the strongest interferes are;
3. UE determines the influence of scheduling each strong interfering serving station DL transmission in one or more parts of the frequency channel. This influence, measured as MCS degradation, is the "cost" ${ }^{5}$ defined in eq. (165);
4. For each relevant subframe and full channel or bandwidth part, the "cost" information associated to each interfering serving station is sent by UE over the Uu interface to its serving station;
5. The high-level "costs" are shared, over the X2 interface, between the FAPs and MBS in the area;
6. Each FAP determines the "price" of its activity on the MCSs of other FAPs and on the MBS providing the coverage layer (see eq. (164);
7. Each FAP determines the frequency resource to be used such to better solve the trade-off between maximizing performance and minimizing the interference to other FAPs and to the
[^4]MBS. The MBS has a "privileged" status, while the FAPs will have to change their operating frequency channel or the bandwidth part for a specific DL transmission.

The full channel (wideband) reports are applied for wideband channels (e.g. 5 MHz ), while the channel part reporting is applied for broadband channels (i.e. $10 \mathrm{MHz}, 15 \mathrm{MHz}$ or 20 MHz ). The serving station (FAP or MBS) may alter the cost, if for example there are free resources in other subframes and may decide to transmit or not the resulting cost to other stations. If the decision is positive, the costs are distributed by the serving station over the X2 interface to the other stations in the area.

Based on the costs received from others serving stations (FAPs or MBS), each interfering FAP which is looking for scheduling DL traffic on a given part of the frequency channel will be able to calculate the interference "price" of re-using the specific part of the frequency channel for the affected UEs which are receiving DL traffic in that frequency channel part. Based on this information the interfering FAP will take steps for changing the operational frequency or the bandwidth part or the operational frequency channel such to create minimum interference to the population of other operational UEs.

### 11.1.1.2.1 Sensing and sharing the CQI degradation

In fact the UE is not reporting the MCS, but the channel quality indicator (CQI), which for the SISO case is identical to MCS. We will first look at the CQI detection, based on existing 3GPP specs, and after that at the evaluation of CQI degradation due to the other stations DL interference.

## CQI detection in the serving cell

We start from a point in which the system is operational and the serving station (FAP or MBS) transmits to UE on the frequency bandwidth part or the frequency channel chosen by the serving station. The UE has the capability of reporting to the serving station the CQI information, as detailed in clause 7.2 of TS 36.2130 , for the serving cell. It is defined that:
"Based on an unrestricted observation interval in time and frequency, the UE shall derive for each CQI value reported in uplink subframe $n$ the highest CQI index between 1 and 15 in Table 7.2.3-1 which satisfies the following condition, or CQI index 0 if CQI index 1 does not satisfy the condition:
A single PDSCH transport block with a combination of modulation scheme and transport block size corresponding to the CQI index, and occupying a group of downlink physical resource blocks termed the CSI reference resource, could be received with a transport block error probability not exceeding 0.1."

Periodic and aperiodic CQI reporting is possible for LTE system. The CQI is computed for a set a resource blocks. Wideband (the CQI is computed for the whole bandwidth) and UE-selected sub-band feedback are possible.
For instance, for periodic CQI reporting, the UE-selected sub-band is as follows. The total number of sub-bands N is divided into bandwidth parts. Considering a system bandwidth between 64 and 110 RBs, the sub-band size is 8 RBs grouped into 4 bandwidth parts (smaller sub-band sizes are considered for smaller system bandwidths). One CQI value is computed and reported for a single selected sub-band from each bandwidth part, along with the corresponding sub-band index [Sesia2009].

## CQI degradation in the serving cell

Let us suppose that UE is able to use the reference signals (RS) sent by other cells for assessing the interference caused by the DL transmissions of these cells. Let us also suppose that there is a suitable mapping of these RSs, such to associate a RS with a specific $N_{\mathrm{ID}}^{\text {cell }}$.

In such a case, the potential interference victim UE connected to the $g$-th FAP, $u(g)$, can evaluate the impact of the interference caused by the DL transmission of the cell $f$ on its achievable CQI in the bandwidth part $s$. Given the interference impact on CQI degradation, $\mathrm{UE}_{u(g)}$ can establish the "cost", which depends on the interference sensitivity and the interference power received in the RBs on the bandwidth part $k$ (see eq. (165)).

We assume that the can be reported over Uu interface to the serving station (FAP or MBS). The serving station may decide to transmit the cost information over X2 interface to other stations in the area. If the sensitivity to the interferer power is low, for example due to a high $\operatorname{SINR}$ at $\mathrm{UE}_{u(g)}$, there will be no need for further communicating this information to other stations.

The MBS may report much higher adjusted costs as compared with FAPs. In this way, the UEs served by MBSs will be better protected to the interference from FAPs.

We have identified a number of issues with the existing standards, which impede on the application of the proposed solution. These issues are listed below:

Issue 1: For determining the interference power from a single transmission, it is necessary to define changes to the standards for extending the protected measurements to bandwidth parts.
Issue 2: X2 should support the transmission of information covering the resource allocated for the protected measurement.
Issue 3: The Uu interface should support the transmission of the interference cost.
Issue 4: The existing standards do not support the measurement of the degradation caused by the activity of another station and the calculation of the "interference sensitivity".
Each report of cost related to a specific UE, serving station and interference source station, will be shared between the stations in the neighborhood, using the X2 interface. To limit the traffic, the bandwidth part should be chosen such to reflect the frequency resources needed for UE scheduling. The generated traffic should be relatively low, due to the fact that only the "potential victim" UEs and stations will generate it.
Issue 5: It is necessary to define the information elements for distributing the "interference cost" over X2 interface.
Issue 6: the existing splitting of the channel width in bandwidth parts (TS36.213 Table 7.2.2-2) is suitable for 20 MHz channels only; for 10 MHz and 15 MHz channel the channel part is not equal with the 5 MHz channel width.

Issue 7: In case that there is no synchronization between FAP and MBS and also between FAPs, the ICIC as defined in Release 8 and the eICIC based on synchronized ABS frames is not useful. FFR was designed for cases in which all the frequency channels have the same bandwidth. It is needed to define an additional eICIC mechanism, possibly using the bandwidth parts as main resource elements.

### 11.1.1.2.2 Decision making

A FAP looking to schedule new DL traffic will look first at bandwidth parts having a low "price".
As described in section 11.1.1.1, the "price" can be defined as a mathematical function which accounts for the degradations on the data rates of the UEs connected to surrounding station due to the transmissions of $f$-th station in a given channel bandwidth part $s$. Such a function was defined in section 11.1.1.1, eq. (164) as:

$$
\begin{equation*}
\text { price }_{f}^{s}=\sum_{\substack{g=1 \\ g \neq f}}^{N_{F}} b_{g}^{s} \operatorname{cost}_{g, f}^{s} \tag{167}
\end{equation*}
$$

with $b_{g}^{s}$ equals to 0 if the $g$-th station did not transmit a cost for the $s$-th bandwidth part (i.e., if the $g$-th station is not scheduling DL traffic in the $s$-th bandwidth part).

If the "price" is above a threshold, and the $f$-th station is actually a FAP, the $f$-th station will consider changing its operating frequency and selecting the frequency channel having a lower price.

If the $f$-th station is a MBS, the MBS will schedule the new traffic for an UE on that frequency channel and bandwidth part best suitable for its operation.

With this approach, each station will be able to select the frequency channel and the bandwidth part suitable for low interfering transmissions.

### 11.1.2 Exponential Effective SINR Mapping (EESM)

As described in section11.1.1.1, when all the subcarriers of a specific user are modulated using the same modulation and coding scheme (MCS) a compression function to map the instantaneous values of SNIRs to the corresponding BLER (Block Error Rate) value, is required. Furthermore, this function is required in order to measure the interference sensitivity factors, $\pi_{u(f)}^{s, r}$, defined in eq. (157), which are necessary to compute the cost values.

Despite there are different possibilities, the EESM (Exponential Effective SINR Mapping) has shown to yield an accurate estimation of the AWGN-equivalent SINR (usually referred to as 'effective SINR') for frequency selective channels, so we will consider this metric for the mapping function. The EESM method estimates the effective SINR using the following formula:

$$
\begin{equation*}
\operatorname{SNIR}_{e f f}=\operatorname{EESM}(\gamma, \beta)=-\beta \ln \left(\frac{1}{N} \sum_{i=1}^{N} e^{-\frac{S N I R_{i}}{\beta}}\right) \tag{168}
\end{equation*}
$$

where the $S N I R_{i}$ are the per sub-carrier SNIR values (we will user one value per resource block), and $\beta$ is the parameter to be determined for each Modulation Coding Scheme (MCS) level. This value is used to adjust to match the actual BLER and the predicted BLER from the effective SNIR in the AWGN channel.

For the simulation results we will consider the MCS available in LTE, along with the effective SNR and $\beta$ values provided in [http://www.nt.tuwien.ac.at/ltesimulator].

| MCS | Effective SNIR (dB) | $\beta$ value |
| :--- | :---: | :--- |
| 0.1523 | -6.934 | 1 |
| 0.2344 | -5.147 | 1.44 |
| 0.3770 | -3.18 | 1.4 |
| 0.6016 | -1.254 | 1.48 |
| 0.8770 | 0.761 | 1.5 |
| 1.1758 | 2.70 | 1.62 |
| 1.4766 | 4.697 | 3.10 |
| 1.9141 | 6.528 | 4.32 |
| 2.4063 | 8.576 | 5.37 |
| 2.7305 | 10.37 | 7.71 |
| 3.3223 | 12.3 | 15.5 |
| 3.9023 | 14.18 | 19.6 |
| 4.5234 | 15.89 | 24.7 |
| 5.1152 | 17.82 | 27.6 |
| 5.5547 | 19.83 | 28 |

Table 30 MCS and required effective SNIR

This table has been obtained through extensive simulations using the LTE codes. The effective SNR for each MCS value is the required effective SNR to achieve a BLER less than $10 \%$ when using this MCS value.

As show in Figure 89, the relationship between the MCS and the required effective SNR (red points) can be approximated by the following empirical SNIR-to CQI mapping function (solid blue line):

$$
\begin{equation*}
\operatorname{MCS}\left(S N R_{e f f}\right) \cong 1.2213 \cdot \ln \left(1+S N R_{e f f}\right) \tag{169}
\end{equation*}
$$



Figure 89. CQI mapping.

The approach is as follows: from the SINR values measured at a given bandwidth part, each user computes the effective SNIR for every possible $\beta$ value (each $\beta$ value is associated to one MCS). The so computed effective SNIR is compared with the effective SNR required for the MCS corresponding to the $\beta$ value. The highest MCS such that the computed effective SNIR is equal or greater than the required effective SNIR is selected.

Once the user $u$ has selected the MCS for the bandwidth part $s$, the sensibility to the interference in this sub-band will be computed using the empirical SNIR-to CQI mapping function in (169):

### 11.1.3 Simulation results

This section provides performance results of our LTE-A adapted pricing approach. We have simplified the simulation by limiting the simulations to a single 20 MHz frequency channel, covering a 20 MHz allocation band, using four bandwidth parts. In this case a bandwidth part is similar with a 5 MHz channel. The CQI and the cost are estimated per bandwidth part. In addition, we have applied the pricing policy, based on a deployment as shown in Figure 90. In this deployment there is one FAP area (FAP dual-strip zone) within the coverage area of the MBS.

Specific parameters are:

- Channel bandwidth: 20 MHz with four bandwidth parts corresponding each to 25 RBs , same as in 5 MHz ;
- Number of UEs per FAP: 1, SISO mode;
- Number of UEs served by MBS and placed in the FAP area: 2 (the MBS will allocate these 2 UEs in two separate bandwidth parts, each one of 5 MHz );
- The dual-strip deployment was considered over a number of floors varying between 1 and 6 ;
- One serving station considers itself interfered if the average SNR received from an interfering FAP is greater than the SNR of the serving station minus 15 dB .


Figure 90. Deployment scenario
Figure 91, Figure 92 and Figure 93 correspond to the simulation results for FAPs and FUEs.
In Figure 91, the operating frequency channel has 5MHz. The average FAP throughput depends on the number of active FAPs in the simulated area. In Figure 91 right, the average number of active FAPs in the area is 6 . The total FAPs throughput gain is aprox. $2 \mathrm{Mb} / \mathrm{s}$.



Figure 91. FAP throughput Left: vs. average \# of FAPs, Right: 6 (average) active FAPs, Average FAP throughput vs. iteration number

Figure 92 Left depicts the percentage of FUEs supporting the maximum MCS versus the average number of active FAPs in the area. On the other hand, Figure 92 right depicts the cumulative density function of the MCS supported by a FUE in the allocated bandwidth part, considering an average of 6 active FAPs in the simulated area. Given that the pricing-based scheme would greatly benefit from the existence of higher rate MCS, we have evaluated the CDF of MCS usage. In the considered case, the use of pricing allows $90 \%$ of the users to support the highest MCS. This value is reduced to $70 \%$ when pricing is not used.


Figure 92. FUEs supporting the maximum MCS; Left: percentage; Right: CDF of MCS for 6 (average) active FAPs

Using the maximum MCS translates (see Figure 93 Right) to a maximum throughput of approximately 21 Mbps (considering the physical overhead) in the selected bandwidth part ( 5 MHz ). If pricing is not used, $30 \%$ of the FUEs will achieve a throughput below this value, while only $10 \%$ of the FUEs will be below this throughput value if pricing is used.



Figure 93. Left: minimum throughput; Right: CDF of MCS for 6 (average) active FAPs
In Figure 93 Left, is shown the minimum throughput for the best $80 \%$ of FUEs vs. the average number of active FAPs, while in Figure 93 Right is shown the CDF of the throughput of FUEs, for an average number of active FUEs equal to 6. If we consider the $80 \%$ best cases (see Figure 93 Left), the minimum throughput guaranteed for a FUE in the simulated area is 6 Mbps better with pricing than without pricing, for an average number of 6 active FAPs. The difference in throughput increases to 10 Mbps for an average number of 18 FAPs. The difference starts to decrease as the density of FAPs grows due to the saturation of the system.

Figure 94, Figure 95 and Figure 96 correspond to the results for the MBS. Figure 94 Left shows the MBS throughput in $\mathrm{Mb} / \mathrm{s}$ ( 10 MHz ), with and without pricing, versus the average number of active FAPs in the area, including the case of no active FAP. On the right, it is shown the MBS throughput in $\mathrm{Mb} / \mathrm{s}(10 \mathrm{MHz})$ versus the iteration number, with and without pricing, for an average number of active FAPs equal to 6 . Notice that an average gain of 6 Mbps can be achieved when pricing is exchanged (see Figure 94 right). It has to be taken into account that the simulation conditions for the MUEs correspond to a worst case, as the two MUEs are deployed within the FAP area (with a probability of being indoor of 0.2 ).


Figure 94. MBS throughput vs. average \# of active FAPs, Left: MBS throughput vs. iteration number

Figure 95 Left shows the percentage of MUEs supporting the maximum MCS with and without pricing versus the average number of active FAPs in the area, including the case of no active FAP, while in Figure 95 Right is shown the CDF of the MCS supported by MUEs, with and without pricing, for an average number of active FAPs equal to 6 . In such a case, the experimental probability for a MUE to support the highest MCS is $26 \%$ (obtained through 100 independent realizations with 2

MUEs per scenario). If pricing is not used, this experimental probability is reduced to $17 \%$. This means (see Figure 96 Right) that the maximum throughput per bandwidth part (approximately 21 Mbps, considering the physical overhead) is not achieved $74 \%$ of the time with pricing, and this number increases to $83 \%$ of the time when pricing is not used.


Figure 95. Left: percentage of MUEs supporting the maximum MCS; Right: CDF of the supported MCS for MUEs

In Figure 96 Left is shown the minimum throughput for the best $80 \%$ of MUEs (meaning that $20 \%$ of the users will have a throughput less than this value) versus the average number of active FAPs in the area, including the case of no active FAP. On the right is shown the CDF of the throughput of MUEs, with and without pricing, for an average number of active FAPs equal to 6 .


Figure 96. Left, Min. throughput for the best $\mathbf{8 0 \%}$ of MUEs; Right, CDF of the throughput of MUEs

Finally, Figure 97 shows the average number of significant interfering FAPs, considering that a FAP is a significant interferer if the average signal strength received from this station is between 0 a 15 dB below the signal strength of the serving station. Notice that this average value is less than 2 for FUEs (even for a significantly high density of active FAPs), while is greater for MUEs deployed within the simulated area. This is due to the lower signal strength received from the MBS compared with FAPs.


Figure 97. Average number of dominant interferer FAPs detected per FUE and MUE.
For an average number of 6 active FAPs (see Figure 98 left), the probability for a FUE to detect more than 3 interfering FAPs is less than $4 \%$. In the worst case, a FUE will detect up to 6 interfering FAPs. This user must report 6 cost values per bandwidth part, which means, assuming 6 bits for quantization of each cost value every frame, i.e., every 10 ms , a rough value of 2.4 kbps . Notice, however, that detecting 6 interfering FAPs is a low likely case.


Figure 98. CDF of the number of dominant interferer FAPs detected per FUE (left) and per MUE (right).

In the case of the MUEs deployed within the simulated area (see Figure 98 right), the number of detected interferers increases a little, with 6 interferers for $90 \%$ of the cases, and a worst case value of 14 interferers.

### 11.1.4 Conclusions and recommended actions

Our simulations demonstrate significantly higher performance as compared with the reference case (no pricing), justifying the investments in standard enhancements. We summarize below the missing elements in standards:

Issue 1: For determining the interference power from a single transmission, it is necessary to define changes to the standards for extending the protected measurements to bandwidth parts.

Issue 2: X2 should support the transmission of information covering resource allocated for the protected measurement within the frequency channel.

Issue 3: The Uu interface should support the transmission of the interference cost.
Issue 4: The existing standards do not support the measurement of the degradation caused by the activity of another eNB and the calculation of the "interference sensitivity".

Issue 5: It is necessary to define the information elements for distributing the "interference cost" over X2 interface.

Issue 6: the existing splitting of the channel width in bandwidth parts (TS36.213 Table 7.2.2-2) is suitable for 20 MHz channels only; for 10 MHz and 15 MHz channel the channel part is not equal with the 5 MHz channel width.

Issue 7: In case that there is no synchronization between FAP and MBS and also between FAPs, the ICIC as defined in Release 8 and the eICIC based on ABS subframes are not useful. FFR was designed for cases in which all the frequency channels have the same bandwidth. eICIC requires inter-cell synchronization. It is needed to define an additional eICIC mechanism, possibly using the bandwidth parts as main resource elements.

### 11.2 Rate Max or Power Min under Interference-power constraints

In [Zhang 2010] the authors consider the downlink transmission of a cellular system in which base stations, each equipped with multiple antennas, cooperatively design their respective transmit beamforming vectors to optimize the overall system performance. Serving mobile stations are assumed to be equipped with a single antenna and only one of them can be active at any given time within each cell. The corresponding channel model is then that of the multiple-input single-output Gaussian interference-channel (MISO-IC). A method is proposed to characterize different rate-tuples on the Pareto boundary of the achievable rate region for the MISO-IC. It is shown that the Paretoboundary rate-tuple of the MISO-IC can be achieved in a decentralized manner when each of the base stations attains its own channel capacity subject to a certain set of interference-power constraints at the mobile stations of the other cells. A decentralized algorithm for implementing the cooperative downlink beamforming method is proposed.

The technique proposed in [Zhang 2010] has been extended to a femtocell based multi-cellular scenario with SISO, MISO or MIMO multi-carrier transmission in under the assumption that only one user is served using the same time-frequency resource. The proposed technique allows a simple pairwise optimization scheme between cooperating femtocells each of which can either maximize its rate or minimize its total transmission power.

## [Patent Application in Process]

### 11.2.1 LTE signals and measurements

LTE signals and measurements

- Reference Signals (RS), both cell specific (including multiple antenna up to four) and UEspecific allowing beamforming.
- Channel Quality Indicator (CQI) Feedback which can be:
o Aperiodic CQI Reporting. Wideband and possibly eNodeB-configured sub-band depending on the PDSCH transmission mode or UE-selected sub-band.
o Periodic CQI Reporting. Wideband and UE-selected sub-band.
- Cell Search signals:
o Primary Synchronization Signal (PSS) and Secondary Synchronization Signal (SSS)
o Physical Broadcast Channel ( PBCH ) decoding in the initial synchronization but not necessary for new cell identification
- LTE Measurements which include:
o LTE Reference Signal Received Power (RSRP): RSRP is defined for a specific cell as the linear average over the power contributions (in Watts) of the Resource Elements (REs) which carry cell-specific RS within the considered measurement frequency bandwidth.
o LTE Carrier Received Signal Strength Indicator (RSSI): total received wideband power observed by the UE from all sources, including co-channel serving and nonserving cells, adjacent channel interference and thermal noise within the measurement bandwidth. LTE carrier RSSI is not reported as a measurement in its own right, but is used as an input to the LTE RSRQ measurement.
o LTE Reference Signal Received Quality (RSRQ): This measurement is intended to provide a cell-specific signal quality metric. Similarly to RSRP, this metric is used mainly to rank different LTE candidate cells according to their signal quality. This measurement is used as an input for handover and cell reselection decisions, for example in scenarios for which RSRP measurements do not provide sufficient information to perform reliable mobility decisions. The RSRQ is defined as the ratio $\mathrm{N} \cdot \mathrm{RSRP} /(\mathrm{LTE}$ carrier RSSI), where N is the number of Resource Blocks (RBs) of the LTE carrier RSSI measurement bandwidth. The measurements in the numerator and denominator are made over the same set of resource blocks. While RSRP is an indicator of the wanted signal strength, RSRQ additionally takes the interference level into account due to the inclusion of RSSI. RSRQ therefore enables the combined effect of signal strength and interference to be reported in an efficient way


### 11.3 Resource block power allocation in LTE femtocell networks

The downlink of LTE is based on Orthogonal Frequency Division Multiple Access (OFDMA) and the smallest radio resource unit that the scheduler can assign to a user is a Resource Block (RB). Each RB has a time slot duration of 1 ms , corresponding to 12 OFDM symbols and a constraint in LTE downlink is that each RB must use the same modulation and coding scheme. Let us suppose that a given FAP wishes to allocate $N_{b}$ resource blocks each one composed of $N$ subcarriers with the goal of maximizing the overall transmission rate with the constraint of ensuring the same modulation over each RB. Then this optimization problem can be formulated as
[P0]

$$
\begin{aligned}
& \max _{\mathbf{p}} \quad \frac{1}{N N_{b}} \sum_{l=1}^{N_{b}} \sum_{k=1}^{N} \log _{2}\left(1+p_{k, l} a_{k, l}\right) \\
& \text { subject to } \quad \frac{1}{N N_{b}} \sum_{l=1}^{N_{b}} \sum_{k=1}^{N} p_{k, l} \leq P_{t} \\
& \quad \mathbf{p} \geq \mathbf{0} \\
& p_{1, l} a_{1, l}=p_{k, l} a_{k, l} \quad \forall k=1, \ldots, N, \quad \forall l=1, \ldots, N_{b},
\end{aligned}
$$

where $p_{k, l}$ represents the power allocated over the k-th subchannel of the 1-th resource block;
$\mathbf{p}$ is the power vector with entries $p_{k, l}$ for $\mathrm{k}=1, \ldots, \mathrm{~N}, \mathrm{l}=1, \ldots, N_{b} ; \quad a_{k, l}=\frac{\left|H_{F F}(k, l)\right|^{2}}{\sigma_{n}^{2}(k, l)}$ with $H_{F F}(k, l)$ the discrete frequency response over the k-th subband of the 1-th RB and $\sigma_{n}^{2}(k, l)$ denotes the noise variance.
Note that the third constraint in $[\mathbf{P 0}]$ ensures that over each resource block the same modulation scheme is adopted since

$$
p_{1, l} a_{1, l}=p_{k, l} a_{k, l} \Rightarrow \log _{2}\left(1+p_{1, l} a_{1, l}\right)=\log _{2}\left(1+p_{k, l} a_{k, l}\right) \quad \forall k=1, \ldots, N, \quad \forall l=1, \ldots, N_{b},
$$

so that the number of unknown powers for each block can be reduced to one.
More specifically, we choose as unknown power in the 1-th block the power $p_{1, l}$ corresponding to the best channel conditions, i.e. to $a_{1, l}=\max _{i \in\{1, \ldots, N\}} a_{i, l}$ in order to guarantee an efficient power allocation on each RB.
Furthermore the constraint $\frac{1}{N N_{b}} \sum_{l=1}^{N_{b}} \sum_{k=1}^{N} p_{k, l} \leq P_{t} \quad$ can be rewritten as

$$
\sum_{l=1}^{N_{b}} \sum_{k=1}^{N} p_{k, l}=\sum_{l=1}^{N_{b}}\left(p_{1, l}+\frac{p_{1, l} a_{1, l}}{a_{2, l}}+\ldots+\frac{p_{1, l} a_{1, l}}{a_{N, l}}\right) \leq P_{t} N N_{b}
$$

or

$$
\begin{gathered}
\sum_{l=1}^{N_{b}} p_{1, l} w_{1, l} \leq P_{t} N N_{b} \\
\text { with } w_{1, l}=1+a_{1, l} \sum_{k=2}^{N} \frac{1}{a_{k l}} .
\end{gathered}
$$

Then the optimization problem in $[\mathbf{P 0}]$ can be reformulated as

$$
\begin{array}{ll}
\max _{\mathbf{p}} & \frac{1}{N_{b}} \sum_{l=1}^{N_{b}} \log _{2}\left(1+p_{1, l} a_{1, l}\right) \\
\text { subject to } & \frac{1}{N N_{b}} \sum_{l=1}^{N_{b}} p_{1, l} w_{1, l} \leq P_{t}  \tag{171}\\
& p_{1, l} \geq 0 \quad \forall l=1, \ldots, N_{b}
\end{array}
$$

Note that this last problem is a convex optimization problem whose solution can be written found by using the KKT conditions as

$$
\begin{equation*}
p_{1, l}^{\star}=\left[\frac{\log _{2}(e) N}{\lambda w_{1, l}}-\frac{1}{a_{1, l}}\right]^{+}, \quad p_{k, l}^{\star}=p_{1, l}^{\star} \frac{a_{1, l}}{a_{k, l}} \quad \forall k=1, \ldots, N, \quad \forall l=1, \ldots, N_{b} \tag{172}
\end{equation*}
$$

where $\lambda$ is the Lagrangian multiplier which can be found as solution of the constraint (171). We can note that the solution in (172) is a multilevel water-filling with water levels

$$
\mu_{1, l}=\frac{\log _{2}(e) N}{\lambda w_{1, l}}
$$

As numerical examples we report in Figure 99 the optimal powers (172) allocated over each subchannel of the resource blocks, while from Figure 100 we can note that the same rate has been

Title of deliverable: Interference coordination protocols in femto-based networks
allocated over each resource block although it can vary from one resource block to another as imposed by the optimization strategy.


Figure 99. Optimal power over each subchannel.


Figure 100. Optimal rate over each subchannel.

## 12 GENERAL CONCLUSIONS

Activity 3A2 has focused on the investigation of decentralized approaches for designing the radio resource allocation when several femto access points (FAPs) and possibly the macro BS (MBS) coexist on the same band. We have elucidated that the system performance is improved when sources are able to exchange messages at control-plane level through the backhaul link that connects all the sources. Those messages, named pricing in this work, basically account for the sensitivity of the cost (or objective) function considered at each source as a function of the received interference. They also convey information about the priority of the served users and how the different constraints are met (target rate, max sum-rate, ...). Such information allows optimizing all radio resources jointly but in a decentralized way, by means of considering the individual constraints of other sources. This mode of operation is not possible when all sources optimize their resources in a competitive way (studied in activity 3 A 1 ).

In this regard, when we want to guarantee a minimum rate in the system, the radio resource allocation becomes a solution of the problem analysed in section 5 , where the power is minimized with the objective of obtain a minimum target rate (in SISO, MISO and MIMO configurations).

- With the proposed approach the total radiated power is reduced and the minimum target rate is satisfied, which is not always possible when pricing messages are not exchanged.

On the other hand, section 6 provides radio resource algorithms to maximize the weighted sum rate (WSR) of the system, comparing simple vs. complex transmitters (based on superposition coding), analysing the impact of a maximum rate served by the FAP due to the backhaul link quality. Since FAPs are serving multiple users, we also investigate the best way of assigning the resource blocks (RB) over those users served by the same FAP.

- When sources update its power simultaneously, the decentralized algorithm might lead to a solution where the bitrate of the users oscillates. By means of incorporating some memory term, that effect is reduced without degrading significantly the total WSR of the system.
- When RB assignment has to be optimized over the served users by the same FAP (when users are served in an orthogonal ways), we have analysed a polynomial algorithm
- Since the objective is to maximize the WSR, the decentralized algorithm gets solutions where some nodes should switch off their transmitters, or decrease the total transmitted power. This situation is not possible when pricing messages are not exchanged.
- Results presented in section 9.1 have elucidated the benefits of the pricing exchange, able to improve the minimum rate of the system ( $10 \%$-Outage rate) by a factor of $2-3$ ( $200-300 \%$ gain) while the spectral efficiency gets a $15 \%$ gain, compared with algorithms without using the pricing exchange.
- The gains in terms of spectral efficiency are obtained when transmit precoders and RB assignment are optimized, otherwise only gains in terms of minimum rate are attained.
- In case the backhaul link limits the communication, the performance of pricing and nonpricing techniques tends to coincide because the generated price is zero. In a nutshell, since the power is not limiting the communication, if you receive additional interference, you have power resources to combat the interference, so it is not required to exchange any price to the interfering terminal in order to control the generated interference.

Section 7 has investigated decentralized resource allocation algorithms when there is a source of randomness.

- In a scenario where FAPs employ the same band as the MBS, but they have to estimate the presence of a MBS in a given RB (with a false alarm probability) in order to not interfering the macro user, the solution to maximize the opportunistic throughput becomes an algorithm
that allocates power over those RB where the channel is stronger and the probability of correct decision is high. The probability of detection is improved thanks to the coordinated channel sensing.
- Similarly, if the activity of the MBS is modelled by Markovian model, we have obtained an algorithm that allocates power in the joint time-frequency plane.
- If the backhaul link is affected by random failures and quantization noise, one algorithm is proposed that converges almost surely.

Another type of algorithm investigated in this work is the Genetic algorithm optimization (Section 8). When it is implemented in a centralized way, i.e. assuming that the SNIRs of all transmitters undergoing optimization are fed at a given central processor unit, it provides a sub-optimal resources allocation. The solution aims at maximizing a target function which can be the system capacity or closely related to it (for example, implementing a fairness criterion or optimizing separately FUEs or FAPs). Its implementation doesn't require any specific modification at PHY or MAC level since it is based on data (FAPs and FUEs SNIRs) which are available to the core network at the level of network management (application or network layer). The analysis performed in this study take into account LTE timing at level of subframe and addresses the problem of radio synchronization. The algorithm is designed to be resilient to fast fluctuations of propagation conditions, as it is based on time averages, relatively long compared to the duration of a single PRB (i.e., taken over a few tens of OFDM symbols). The solution has shown to be stable with respect to short variations of interference levels, although having as a drawback a convergence time of the order of hundreds of subframes.

We have compared the centralized GO with the decentralized resource allocation, providing an interesting insight on algorithms performances. They have been considered in the same scenario, evaluating the same metrics and power constraints in a realistic propagation scenario. The attained solution by the decentralized approach is not too far from the solution attained by an ideal centralized algorithm that knows all the parameters of the system.

Regarding the time or number of frames to attain a stable solution:

- The decentralized approach analysed in section 5, 6, 7 tend to converge in less than 15-20 iterations, except in the link failures case, where the converge time might increase up to 30-50 iterations
- The GO algorithm typically converges in $P \mathrm{x} G$ frames, where $P$ (Population) and $G$ (Generations) are parameters related to the algorithm. The minimum is about 500 frames.

In section 11.1 a simpler pricing-based algorithm for interference coordination has been proposed with the goal of meeting most LTE constraints. More specifically, this algorithm considers on/off transmission at each resource block (no power control). Instead of the SNIR, UEs report the maximum MCS that can be used within the set of available resource blocks, along with a parameter (cost) that measures the MCS degradation because of neighbor FAPs transmissions. The results demonstrate significantly enhanced performance as compared to the no pricing case, justifying the required modifications in the enhancement of future releases of LTE.

## 13 SUMMARY OF RESULTS TOWARDS OTHER ACTIVITIES

### 13.1 Towards WP4

The investigated techniques in sections 5.2 and 6 at the physical layer have been inputs for the activities carried out in WP4 (4A2, MAC control procedures for femtocell and performance evaluation). Proper control plane procedures have been specified in order to accommodate the investigated algorithms into the standards.

### 13.2 Towards WP5

The investigated techniques have been considered for evaluation at system level in WP5.

### 13.3 Towards WP6

Possible useful information for WP6 (6A1 Hardware feasibility study and prototyping) is Table 29, describing the complexity of the different techniques. The objective of provide this input is to know the difficulties when devised WP3 techniques have to be implemented.

## 14 APPENDIX

### 14.1 Simulation methodology

The simulation of the interfering FAPs is based on defining the set of parameters describing the propagation environment and their reciprocal effect for the displacement of several FAPs, from few to tens of.

The characterization of the simulation scheme is represented in Figure 101. First the simulator sets the scenario and the relevant propagation parameters (Figure 102). Depending on the number of FAPs and on the capabilities implemented at RRM level, the transmitters can be either grouped in sub-sets or processed as a whole.
Different runs of the GA can provide somewhat different values of the parameters under search. The difference is intrinsic in the GA and the simulations included this intrinsic variability implementing a method for iterating every optimization using the current GA output as initial values for the next run.


Figure 101. Block scheme for the system simulation implementing GA optimization.

The scenario generation phase includes the definition of the transmission parameters related to the Standard of communication under study (i.e., LTE or WiMAX) which are summarized in the pathloss table representing the mutual interference among the FAPs finally deployed. According to Figure 102, the environment system generator is made of four blocks:

1. Setting the PHY characteristics of the Standard (carrier, number of subcarriers, channelization scheme, MCS, power of transmission;
2. Generating the geometry of the environment, in terms of FAPs density distribution, building shape, number of floors, number and type of internal/external walls, position within the apartments;
3. Definition of the channel model adopted, according to ITU or 3GPP;
4. Computation of the pathloss matrix, representing the pathloss for every FUE with respect to the corresponding FAP and to every other FAP/FUE in the system.

The scenario generation is summarized by the pathloss matrix which is the input for the GA optimization. Clustering and scalability based on the parallelization of sub-sets of FAPs is based on the analysis of the pathloss matrix and its structure. Some explicit examples are introduced in section 14.3.2.


Figure 102. Scheme for the scenario generation expressed in terms of pathloss.

### 14.2 Geometry of the scenario

The study analyzes the downlink or uplink co-existence within a dense deployment of FAPs in blocks of flats (or offices) in a typical urban scenario (see [FREEDOM-D21]). A single block of flats consists of a square building made of three floors divided as a regular square grid of 25 apartments on each floor, separated by 3 m . Each apartment is modeled as a square of 10 m side. Another set of parameters describe the details of the environment, as summarized in Table 31.

| Parameter used | Description | Value |
| :--- | :--- | :--- |
| min_separ | Minimum separation FAP/FUE | 1 m |
| Nfloor | Number of floors in the building | 3 |
| Gr | Size of the apartments grid | 5 |
| Lapt | Apartment side | 10 m |
| d_w | Minimum separation between walls in the same apartment | 3 m |
| prob_mw | Probability of presence of a main wall between two transmitters | 0.1 |
| p_fap | Probability of FAP presence in an apartment, i.e. FAP density | $0.33-1$ |

Table 31. Typical values for the parameters used for the scenario geometry and the deployment of FAPs.

On the basis of such parameters one can generate a deployment of FAPs in a block, summarized by a table of FAPs distances from FUEs. The effective parameters to model the propagation loss includes the presence of main walls and floors which, depending on the relative distances between FUEs and FAPs, are summarized in terms of an effective number of floors separating the devices.
The deployment of FAPs inside a building has been simulated for two different values of probability ( $33 \%$ and $100 \%$ ) of presence of FAPs in an apartment. The effective positions of each FAP in the apartments are calculated by selecting a random position within a realistic parameters range (e.g., distance from floor, distance from ceiling) and calculating the equivalent floors of separation in terms of the real number of floors and taking account of the number of walls in between. The flats can allocate one FUE, one FAP, both transmitters or none (see [R4-080149]).

### 14.3 Pathloss models

### 14.3.1 ITU and 3GPP formulation

The pathloss considered in the simulations are provided by the 3GPP and ITU models for indoor attenuation, which depend on some environment and propagation parameters.
According to [ITU1238], within the ITU scheme one has

$$
\begin{equation*}
L_{I T U}=20 \log f+\beta \log d+L_{f}\left(n_{f l}\right)-28 \tag{173}
\end{equation*}
$$

where the total path loss, in dB , depends on the frequency of transmission $f$ in Hz , the distance $d$ in meters, the power loss coefficient $\beta$ and the floor loss penetration factor $L_{f}\left(n_{f l}\right)$ varying with the number of floors $n_{f f}$. The possible values of the parameters in (173) are summarized in Table 32.

| ITU | Parameters value $[\mathbf{d B}]$ |  |
| :--- | :--- | :--- |
|  | $\beta$ | $L_{f}\left(n_{f l}\right)$ |
| Residential | 28 | $4 \times n_{f l}$ |
| Office | 30 | $15+4 \times\left(n_{f l}-1\right)$ |
| commercial | 22 | $6+3 \times\left(n_{f l}-1\right)$ |

Table 32. Values adopted for pathloss parameters for the ITU scheme.

According to the 3GPP modelization the pathloss is given by (see, for example, [R4-071617])
$L_{3 G P P}=20 \log \left(\frac{4 \pi f}{\mathrm{c}}\right)+20 \log d+q_{i n} W_{i n}+q_{e x} W_{e x}+L_{f} n_{f l}^{\left(\left(n_{f}+2\right) /\left(n_{n}+1\right)-0.46\right)}$
where $c$ is the speed of light, $q_{i n}$ is a random variable representing the total number of internal walls between transmitter and receiver, $W_{i n}$ is the partition loss corresponding to internal walls within and apartment (in dB ), $q_{e x}$ is a random variable representing the total number of external walls between transmitter and receiver (including main walls within a building), $W_{e x} W_{p x}$ is the partition loss corresponding to external walls (in dB ). While $W_{\mathrm{m}}^{*}$ and $W_{\mathrm{E}}^{*}$ are fixed parameters, $\boldsymbol{q}_{\mathrm{m}}$ and $\boldsymbol{q}_{\boldsymbol{e x}}$ take account of real-world variations of apartments' layouts. The total number of walls between transmitter and receiver is given as $q=q_{t \mathrm{~m}}+q_{e x}$ chosen randomly from the set $\left\{0,1, \ldots,\left\lfloor\frac{d}{d_{w}}\right\rfloor\right\}$. If transmitter and receiver are in different apartments, the parameters can be determined as $q_{e x}=\max (1,\lfloor q / k\rfloor)$, where k represents the average number of internal walls per external wall (of order $10 / d_{w}$ ) and $q$ is evaluated by consequence. The values adopted are in the following Table 33.

| Parameter | Value |
| :---: | :---: |
| $W_{t \mathrm{xm}}$ | 5 dB |
| $W_{e x}$ | 5 dB |
| $W_{e x}^{*}$ | 18.3 dB |
| $k$ | 5 |
| $f$ | $2 \times 10^{9} \mathrm{~Hz}$ |

Table 33. Values adopted for pathloss parameters in the 3GPP scheme.

After the scenario set up, the algorithm provides the set or values regarding linear distances, number of internal and external walls, number of floors between every FUE and every FAP, whose possible realization in summarized in Table 34.


Table 34. Values obtained for one simulation of the scenario geometry. Rows follow FUEs index, columns follow FAPs index.

For example, with the above values for the scenario definition, one gets the set of relative pathlosses in dB as

$$
L_{\text {loss }}=\left[\begin{array}{ccccc}
58 & 66 & 64 & 82 & 106  \tag{175}\\
74 & 32 & 83 & 66 & 75 \\
69 & 78 & 41 & 78 & 87 \\
65 & 67 & 80 & 76 & 92 \\
103 & 83 & 95 & 77 & 61
\end{array}\right]
$$

### 14.3.2 Pathloss and clustering

By increasing the number of FAPs, the simulation has analyzed the structure of the pathloss matrix taking into account the apartments deployment described above. The interference among FAPs is characterized by their relative distance and by the physical environment (walls, floors) surrounding the transmitters. As shown in Figure 103 a structure as clusters of interference appears as an effect of the effective distances: 3, 4, 5 or 6 clusters identify the groups of couples FAPs-FUEs which have a larger mutual influence. The typical maximum is along the diagonal (mostly blue squares, lowest pathloss) which characterizes the FUEs linked to their own FAP.


Figure 103. Pathloss representation in dB. Left: 27 FAPs, 3 floors building; Right: 64 FAPs, 4 floors building.

As supported by physical and geometrical considerations, FAPs scattered apart tend not to influence each other, therefore when considering a large number of FAPs one can set a criterion to determine if the problem is scalable or not. If from the pathloss matrix one can define a set of emerging clusters, the GO optimization method can be split in several sub-problems which can be run in parallel, saving computing time and decreasing the convergence time of the algorithm.
We will devote section 8.4 to the scalability of the GO based on the parallelization over several GO processes with a reduced number of variables.

### 14.4 More results of section 8.3.2.1

The optimization algorithm applies similarly in the case of a flat fading channel for an increasing number of FAPs/FUEs. The analysis for the case of twelve FAPs is shown for a pathloss matrix as visualized in Figure 104


Figure 104. Random deployment of 12 FUEs , pathloss matrix false color representation, values are in dB.

The genetic optimization with the fitness function evolution is shown in Figure 105 where the graph in the bottom represents the assignment of frequency (first 12 bars provide starting frequency value, second 12 bars provide bandwidth) and power (last 12 bars). Final allocation is visualized inFigure 106: the three parameters assigned to the k-th FAP are given by the value of the k -th bar in each of the three sets of 12 .


Figure 105. GO output for 12 FUEs deployed with an interference scenario corresponding to a pathloss matrix as in Figure 104.


Figure 106. GO output for 12 FUEs , power/bandwidth allocation, for a pathloss matrix as in Figure 104.

Those FAPs that can be considered independent of each other, in terms of pathloss, are assigned the same frequencies without risk of interference.
The algorithm can deal with an increasing number of FAPs/FUEs, although reducing the speed of computation. Considering a sparse density with value 0.33 , i.e. corresponding to a $33 \%$ probability of a FAP (and corresponding FUE) presence in a single apartment, the algorithm has been tested also for scenarios with an increasing number of FAPs/FUEs. The results for 32 FAPs deployed in a single building are shown in Figure 107 and Figure 108, while for 48 FAPs are reported in Figure 109 and Figure 110.


Figure 107. GO output for 32 FUEs deployed with an interference scenario corresponding to a pathloss matrix as in left panel of Figure 108.


Figure 108. Deployment of 32 FUEs. Left panel: pathloss matrix visualization, values in dB. Right panel: GO output power/bandwidth allocation.


Figure 109. Random deployment of 48 FUEs, pathloss matrix false color representation, values in dB.


Figure 110. Visualization of GO output for 48 FUEs, power/bandwidth allocation, for a pathloss matrix as in Figure 109.

### 14.4.1 Frequency GO under power constraint

In this Section we report a sample of outputs for the optimization for the variables $P(f)$ where $f$ $=1, . .25$, under the constraint in Eq. (145), considering the state of UL and DL for the Macro network for $\{6,8,10,12,14,16,18,20\}$ couples of FAP/FUE, corresponding to a FAP load of $\{0.3,0.4,0.5$, $0.6,0.7,0.8,0.9,1\}$, respectively.

### 14.4.2 GO output for different FAP load FAP/FUE couples



Figure 111. GO output for 8 couples of FAP/FUE users. Macro-network in UL. 25 PRBs, 200 variables optimized.


Figure 112. GO results for 8 couples of FAP/FUE users. 25 PRBs. Macro-network in UL.

| N. couples 8 - Macro UL |  |  |  |  |  |  |  |  |
| :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: |
| Avg. capacity <br> bit///OFDM | 5296.0 | 3287.2 | 2568.7 | 1172.8 | 2167.7 | 3072.8 | 3863.4 | 1830. <br> 7 |
| Avg. sum rate <br> bit///OFDM | 2907.4 |  |  |  |  |  |  |  |
| Total P dB | 19.0 | 19.6 | 18.7 | 19.4 | 19.5 | 18.3 | 18.3 | 19.0 |

Table 35. Summary of results for a system of 8 FAP/FUE couples. Macro-network in UL.

| N. couples 8 - Macro DL |  |  |  |  |  |  |  |  |  |
| :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: |
| Avg. capacity <br> bit///OFDM | 3401.5 | 490.0 | 1445.9 | 388.3 | 502.1 | 1360.8 | 3804.8 | 149.0 |  |
| Avg. sum rate <br> bit//OFDM | 1442.8 |  |  |  |  |  |  |  |  |
| Total P dB | 18.8 | 18.6 | 17.5 | 20.0 | 20.0 | 18.1 | 18.6 | 18.1 |  |

Table 36. Summary of results for a system of 8 FAP/FUE couples. Macro-network in DL.

## ICT-248891 STP

Document number: D3.2
Title of deliverable: Interference coordination protocols in femto-based networks


Figure 113. GO results for 12 couples of FAP/FUE users. 25 PRBs. Macro-network in UL.


[^0]:    ${ }^{1}$ The considered example was obtained considering a simple Gaussian channel, in which case the interference power of a transmitter versus a given receiver, can be interchanged with the notion of coverage radius, i.e. the distance above within which two transceivers are defined to interfere with each other. Following the so called "protocol model," interference is then modeled assuming that only nodes within interference distance from each other give a significant contribution to each other's interference terms.

[^1]:    ${ }^{2}$ In this context, i.e. for the purpose of deciding whether a given subband is used or not by any MBS, the interference signal is the sum of the signal coming from different MBSs.

[^2]:    ${ }^{3}$ In the following, we assume $\operatorname{Prob}\left(\mathcal{H}_{1, k}\right)=\operatorname{Prob}\left(\mathcal{H}_{0, k}\right)=1 / 2$ and, for simplicity, we drop in our formulas the dependence on the a priori probabilities.

[^3]:    ${ }^{4}$ A narrower power margin can be considered for the QoS readjustment at the expense of a higher computational load for the calculation of the afforded rates.

[^4]:    ${ }^{5}$ The degradation of UE data rate due to the degradation of its MCS caused by a specific FAP DL transmission in a specific frequency channel or frequency channel part and subframe is named "cost".

