

# Optimum Selection of Radio Access Technologies in a 5G Heterogeneous Multicomm Environment

Markus MUECK<sup>†a)</sup>, Majdi MSALLEM<sup>†</sup>, Christian DREWES<sup>†</sup>, and Markus STINNER<sup>††</sup>, *Nonmembers*

**SUMMARY** This contribution addresses optimum wireless link selection in a 5G heterogeneous Multicomm environment in which multiple distinct Radio Access Technologies can be operated simultaneously by a given Mobile Device. The related decision making mechanisms are proposed to be part of the Mobile Device User Equipment which identifies the preferred Radio Access Technologies to be operated as well as the preferred Modulation and Coding parameters and finally selects the optimum choice either independently or through negotiation with the Network Infrastructure Equipment. In this context, the concerned Mobile Device is able to manage the bandwidth to be employed for communication — the bandwidth per Radio Access Technology is typically defined by the Network Infrastructure but the possibility for aggregating a multitude of distinct technologies introduces a new degree of freedom enabling the choice of the bandwidth selection. Simulation results demonstrate that the Mobile Device driven selection of a Multicomm configuration may either substantially reduce Mobile Device power consumption (75% in a typical example) or increase the aggregate throughput at an identical power consumption level compared to the single link case (a 3x factor in a typical example).

**key words:** 5G, heterogeneous networks, multicomm, wireless communications

## 1. Introduction

Optimum resource allocation and link selection in a heterogeneous radio context are well studied problems. Typically, a joint allocation strategy is proposed in which the network tries to identify a solution that guarantees a high level of overall resource usage efficiency subject to power constraints [1]. In practice, however, a trade-off between requirements from multiple stakeholders needs to be met: a Mobile Network Operator (MNO) is capable of finding a (near-)optimum configuration for his network. On the other side, a given Mobile Device may be able to operate links to multiple MNOs (providing access to systems such as 3GPP LTE, etc.) and to further unlicensed systems (such as WiFi, etc.) simultaneously. Indeed, current multi-SIM Mobile Device products enable End Users to operate links to a multitude of possibly competing MNOs in order to choose the best configuration from the End User perspective. In [2], such a corresponding hybrid optimization approach is discussed combining centralized and distributed decision making principles and novel mechanisms are introduced for transferring part of the decision making from the infrastruc-

ture equipment into the concerned Mobile Devices. It is assumed that the Network Infrastructure will still decide on parameters such as the overall channel bandwidth, time and frequency slots (i.e., sub-sets of the overall capacity) to be dynamically allocated to a Mobile Device, etc. However, the Mobile Device will be able to independently select the number of links to be operated simultaneous, which Modulation and Coding Scheme (MCS) should be employed and which output power levels need to be chosen for uplink communication. The expected performance gains are based on the following basic principle: from e.g. Shannon Theory, it is known, that wireless transmission modes yield an exponential behavior. For high modes, a small increase in throughput performance is associated with a large increase of power consumption. Our proposed solution avoids those high-power modes. The required capacity is obtained through bandwidth aggregation by operating multiple links simultaneously. This approach leads to a highly power efficient solution for a given Mobile Device. In order to influence this decision making process, MNOs may choose to limit the choice of available MCS modes to a subset.

In this paper, we extend the optimization approach from [2] in order to resolve an inherent sub-optimality and provide detailed analysis results based on simulation.

We employ discrete optimization techniques to identify the optimum working point for a Mobile Device, i.e. which Radio Access Technologies (RATs) to operate simultaneously, which MCS scheme to choose and which output power to set for each RAT.

This paper is organized as follows. Section 2 states the optimization procedure from a Mobile Device perspective. Section 3 proposes efficient solutions for finding the optimum solution related to the optimization problem. Section 4 analyses the performance of our approach and provides a comparison to traditional single-link and network centric optimization strategies. Finally, Sect. 5 gives a complexity analysis and Sect. 6 provides a conclusion.

## 2. Problem Statement

We consider a heterogeneous system comprising  $N_B$  BSs/APs, each of which having a distinct index  $j = [1 \dots N_B]$ . We consider one Mobile Device  $i$  among a multitude of such devices, which can simultaneously obtain different resources from any of the available BSs/APs. Denote the link between the Mobile Device  $i$  and the BS/AP  $j$  by  $i_j$ . This scheme is illustrated in Fig. 1.

Manuscript received February 11, 2015.

Manuscript revised June 29, 2015.

<sup>†</sup>The authors are with INTEL Mobile Communications, Am Campeon 10-12, 85579 Neubiberg, Germany.

<sup>††</sup>The author is with Technical University Munich, Theresienstrasse 90, 80333 Munich, Germany.

a) E-mail: markus.dominik.mueck@intel.com

DOI: 10.1587/transcom.E98.B.1949

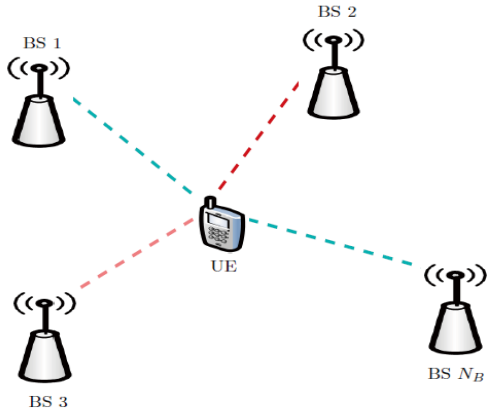


Fig. 1 Multicom communication context.

Typically, the Network Infrastructure equipment employs adaptive modulation and coding and allocates a physical configuration to each link combining a specific modulation and coding rate. These so-called *MAC Modes* relate to a specific combination of a constellation type (such as BPSK, QPSK, QAM-16, QAM-64, QAM-256, etc.) and code rate (Code rate  $R = 1/2, 2/3, 3/4$ , etc.). Those modes are referenced by index  $i_j = [1 \dots I_j]$ ; from the perspective of a Mobile Device, each MAC Mode  $i_j$  of a link to BS/AP  $j$  delivers a rate  $r_{j(i_j)}$ , requires an instantaneous transmission power level  $p_{j(i_j)}$  and uses a propagation channel of quality  $g_j$  (comprising for example the observed SINR level, etc.). The aggregate rate of all links being simultaneously operated at a given MAC mode by a Mobile Device is denoted as  $R$  and the related instantaneous power is  $P$ .

We propose a novel approach, which consists in minimizing the overall instantaneous output power of the mobile device under the condition of satisfying a minimum aggregated data rate per user. This approach satisfies a certain quality of service, i.e. the aggregate final rate is larger or equal to  $R_{\min}$  for the given Mobile Device. For the set of indexes  $I^*$ , we have

$$I^* = \arg \min_{\{i_j \in [1 \dots I_j]\}_{j=1 \dots N_B}} \sum_{j=1}^{N_B} p_j(i_j, g_j), \quad (1)$$

$$s.t. \sum_{j=1}^{N_B} r_j \geq R_{\min},$$

which is a discrete resource allocation problem. The multiple Base Stations (BSs) and/or Access Points (APs) not interfering with each other as they are operating in orthogonal spectrum bands. We refer to those stations by straightforward and trivial method is to evaluate all possible combinations and compare them in order to find the best feasible solution. This method finds ultimately the optimum solution. However, the number of tested combinations grows (quasi) exponentially for an increasing number of links  $N_{BS}$ . In the subsequent sections, different methods are outlined which are numerically more suitable and adapted to the problem.

In the pre-processing step, we consider the rate func-

tion over the SINR for each of the MAC modes for any of the single links. It is crucial to define the set of MAC modes (“points”) of each rate function which coincide with a convex envelope function. Using the following proposition, all points are discarded which are not contained in the convex envelope function.

Assume a single link with  $W$  distinct working points  $(r_i, p_i)$  with  $i \in [1, \dots, W]$ , each of which corresponding to a given MAC mode.  $r_i$  is the data rate of the concerned transmission mode and  $p_i$  is the corresponding instantaneous power consumption. Let  $i, j, k \in [1, \dots, W]$ . If  $r_i < r_j < r_k$  and  $\frac{p_k - p_i}{r_k - r_i} \geq \frac{p_k - p_j}{r_k - r_j}$ , then  $r_j$  cannot be part of the convex envelope function and must be discarded. Denote by  $Q = [1, \dots, W]$  the index set of the extreme points of the convex envelope of points and define  $\beta_d = \frac{p_d - p_{d-1}}{r_d - r_{d-1}}$  ( $\forall d \in [2, \dots, W]$ ),  $\beta_1 = -\infty$  and  $\beta_{R+1} = +\infty$  as the increasing sequence of slopes between the convex points. These slopes define the relevant Lagrangian multipliers that would lead to a different solution. This is executed separately on each link. However, for the overall solution, they need to be arranged together so that we can directly define the Lagrangian multiplier  $\lambda$  relative to the optimization problem for any target increase or decrease. The Algorithm by Shoham and Gersho [5] provides a computationally efficient approach suitable for handling such a problem statement; it ideally considers only the efficient update of  $\lambda$  as explained in the following.

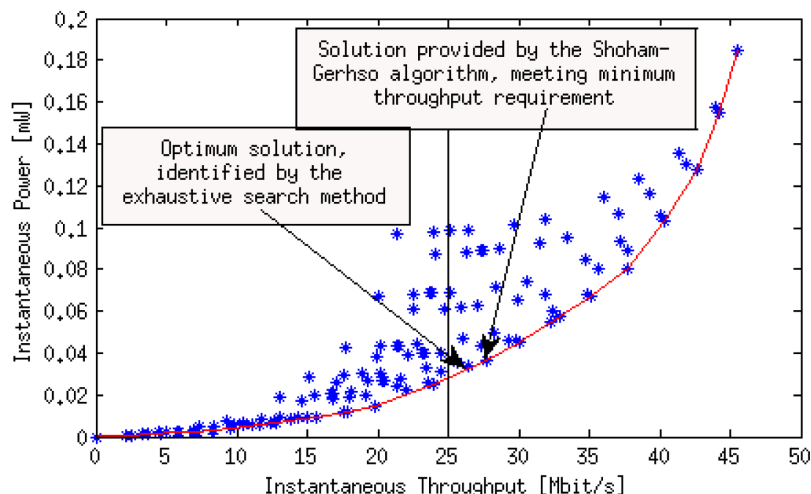
Based on the available single links, we now consider the simultaneous operation of multiple links in a given Mobile Device. For each iteration  $k$ , we compute the constraint function  $R_k$ , which corresponds to the aggregate rate available, i.e. the sum of rates  $r$  of each link being operated simultaneously:

( $R_{\min}$  is the minimum constraint that has to be fulfilled)

- If  $R_k = R_{\min}$ , the optimal allocation is found. We stop.
- If  $R_k > R_{\min}$ ,  $\lambda$  is updated to obtain a nearest value which comes closer to the constraint. We obtain a smaller decrease.
- If  $R_k < R_{\min}$ ,  $\lambda$  is updated to obtain a nearest value which comes closer to the constraint. We obtain a smaller increase.
- If  $(R_k - R_{\min})(R_{k-1} - R_{\min}) < 0$ , the best allocation on the convex envelope is found which corresponds to the one fulfilling the constraint. We stop. In this case, the algorithm changed from the unfeasible to the feasible domain (or the other way around) over the convex envelope.

Moreover, the idea by Shoham and Gersho [5] reduces the computational complexity of solving the problem if the initial  $\lambda$  is properly chosen. The closer it is to the final  $\lambda$ , the less iterations are needed. We can obtain an ‘average’ initial  $\lambda$  with low computational complexity by summing variables corresponding to average Lagrange variables of the separate slopes. The initial  $\lambda$  is, therefore, computed by

$$\lambda = \frac{1}{M} \sum_{u=1}^M \frac{\beta_{u,2} + \beta_{u,R_u}}{2}, \text{ where } \beta_{u,i} (i = [1 \dots W_u]) \text{ are the slopes}$$



**Fig. 2** Identification of the optimum working point of a Mobile Device employing the search algorithm by Shoham and Gersho along the convex envelope function (in RED). The required minimum aggregate throughput is illustrated by the vertical line at 25 Mbps.

values defined for each link  $u$ . However, this initialization method was found empirically and is not necessarily the best one, and further study of this problem may yield a better solution.

The resulting solutions correspond to points on the convex envelope as shown in Fig. 2. Each of the working points corresponds to a specific link configuration, i.e. the selection of one or multiple RATs being operated simultaneously and the respective MAC modes selected by the concerned Mobile Device. We assume that other parameters, e.g. signal bandwidth, selection of time/frequency slots, etc. are determined and pre-defined by the Network Infrastructure.

Applying the iterative Algorithm by Shoham and Gersho [5] as detailed above, each new update identifies the next working point on the convex envelope function as illustrated in Fig. 2. This method has low computational complexity and converges to the optimum solution on the convex envelope. However, if the optimum solution is above the convex envelope, it cannot be found as is illustrated in Fig. 2 where feasible working points with a lower cost function are not exactly on the convex envelope, but slightly above it.

Note that the proposed optimization approach is limited to identifying the optimum aggregate instantaneous power consumption. It is possibly different from the configuration leading to the longest Mobile Device Battery life time, since here an additional parameter needs to be identified: The time scheduling of the transmission to each of the available radio links.

In the next section, two novel approaches are presented to overcome this inherent sub-optimality.

### 3. Identification of Optimum Multicomm Configuration

To solve the previously stated discrete optimization problem, the exhaustive search approach is computational inten-

sive, in particular for  $N_{BS} > 4$ . We thus investigate alternative methods which are simple, practical and computationally more efficient, if properly implemented. Therefore, we must identify a suitable convex optimization tool. The relationship between throughput and power consumption values tends to (near) convex functions. This holds for the Shannon capacity and is also observed for experimental values.

Hence, a suitable method is based on the Lagrange Theory as illustrated in [4]. For this specific structure of problems, it is suitable to use the decomposition theory in order to solve the problem separately and to find the corresponding mode on each link efficiently. The update of the Lagrangian multiplier is independent from solving the Lagrange dual function itself and different approaches are possible. For example, the sub-gradient method relies on a step-size mechanism for each update. The dual-search procedure bases on intersection values of previous solutions [4], while the Algorithm by Shoham and Gersho as in [4], [5], [2] typically converges quickly. Using a pre-processing step, we can determine the relevant Lagrangian multipliers that lead to different solutions. The overall process is summarized in the sequel.

#### 3.1 Iterative Approach for Overcoming Sub-Optimality in the Convex Optimization

The algorithm by Shoham and Gersho identifies the optimum solution on the convex envelope with minimum power consumption and the required minimum throughput constraints. If the overall optimum working point is located above the convex envelope, it cannot be found.

To include this optimal solution in a convex envelope, we propose to change the set of variables.

We remove one working point of the available set from the previous solution. Since the problem is separable, the optimal solution is a combination of different modes on the

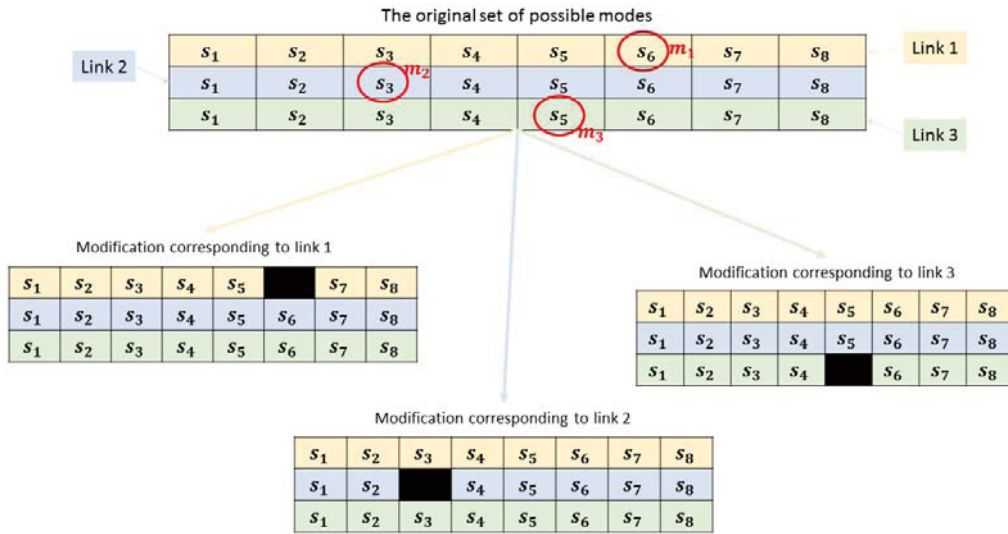


Fig. 3 Illustration of the iterative approach for overcoming sub-optimality in the convex optimization.

target links. Removing a mode of the solutions from each link eliminates many possible combinations and not just one. To eliminate a specific point, we must keep all combinations except this. Therefore, we eliminate the selected mode from each link separately.

Define  $T_j = [s_{j,1} \dots s_{j,p_j}]$  as the set of  $P_j$  possible working modes of link  $j$  (i.e., the ensemble of all possible MAC modes for which a different number can exist for each link) and  $i_j$  as the corresponding mode of the link  $j$  (i.e., the MAC mode selected by the concerned Mobile Device for link  $j$ ),  $j \in [1 \dots N]$ . For the solution by Shoham and Gersho in the first iteration, we consider the whole set  $\prod_{j \in [1 \dots N]} T_j$ . We change it into the set  $\prod_{j \neq k} T_j \times (T_k \setminus i_k)$  for each link  $k \in [1 \dots N]$  to remove one mode  $i_k$  for one specific link  $T_k$  and to combine the result with other links  $T_j \forall j \neq k$ , as illustrated in Fig. 3. The optimization algorithm obtains a new solution. This step must be applied to all links.

The first iteration requires  $N$  separate optimizations and leads to  $N$  new working points. Each of these points must be expanded again to  $N_2$  new working points after the second iteration and analogously  $Np$  new solutions after the  $p^{th}$  iteration. There is no guarantee after which iteration the true solution will be found. To reduce complexity, the number of iterations must be limited to provide a trade-off between optimization complexity and the suboptimality of the solution.

### 3.2 Sub-Block Search Based Approach for Overcoming Sub-Optimality in Convex Optimization

In this section, we combine the algorithm by Shoham and Gersho and the branch-and-bound method. This approach is different from the previous method, because it aims at finding the overall optimum solution and it only stops when all

potential solutions are examined. We use the efficient convex optimization within any specific set of points, but the management of those points is done using the domain-cut variant of the branch and bound algorithm. The algorithm by Shoham and Gersho extracts the first feasible solution and the nearest unfeasible solution. Feasible points are combinations that fulfill the throughput constraint, while unfeasible ones cannot achieve the minimum constraint. These points allow the Domain-Cut procedure to define the set of ‘non-promising’ and ‘unknown’ points, as illustrated in Fig. 4. The best feasible point defines the set of unpromising points and the first unfeasible point defines the set of ‘unfeasible’ points.

Moreover, the Domain-Cut algorithm helps dividing the remaining non-rectangular domain into a union of integer sub-boxes which can be decomposed with the Lagrangian relaxation on the revised domain.

Define  $X_k$  as the set of all integer boxes that have not been fathomed at the  $k^{th}$  iteration and let  $X_0 = X$  be the initial set of all integer boxes. In every iteration  $k$ , an integer sub-box is selected and discarded from  $X_k$  in order to be examined. The Shoham-Gersho algorithm is applied on the selected sub-box with three possible outcomes:

1. All points are feasible and we just take the minimum one and compare it to the current optimal for a potential update.
2. All points are unfeasible and discarded.
3. Definition of a feasible optimal solution  $x_k$  and a infeasible nearest point  $y_k$  along with the dual value that represents the lower bound.

In the third case, the Domain-Cut procedure uses  $x_k$  and  $y_k$  of the given integer sub-box to generate at most  $2N-1$  new integer sub-boxes of unknown points. The new sub-boxes are added to  $X_k$  to form  $X_{k+1}$ .



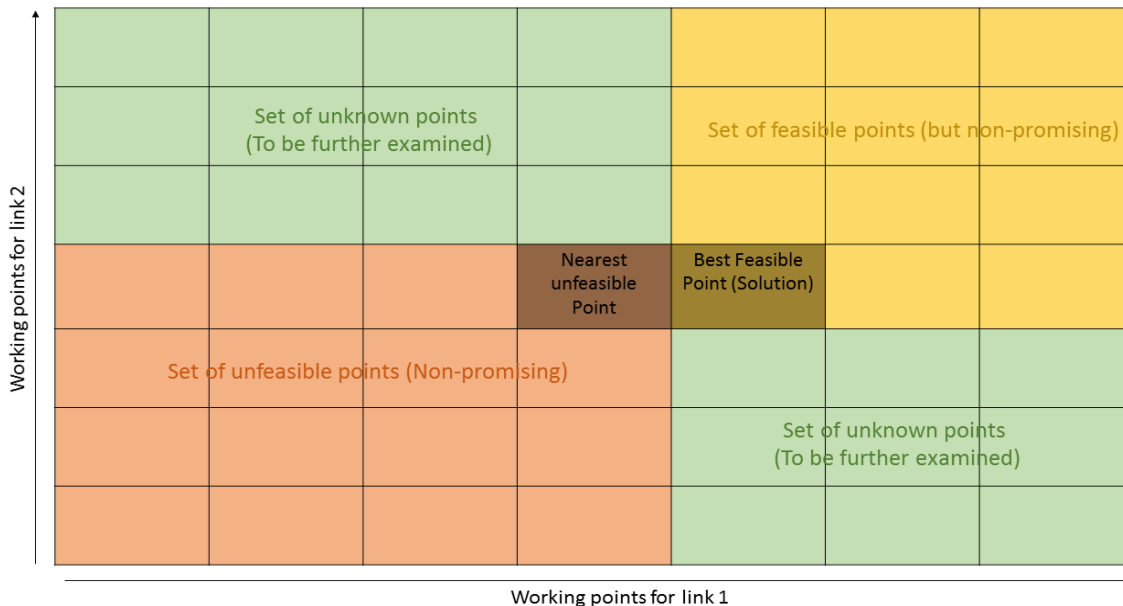


Fig. 4 Illustration of sub-block search for overcoming sub-optimality in convex optimization.

All integer sub-boxes whose lower bounds are higher than or equal to the current best feasible function value can be removed. The process terminates after  $n$  iterations when  $X_{n+1}$  is empty and the incumbent solution of  $X_n$  is the optimal solution to the primal problem.

#### 4. Simulation Results

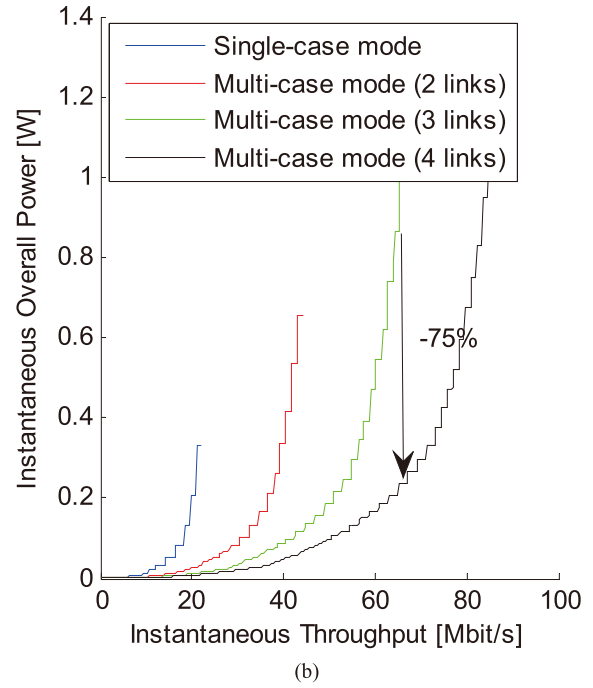
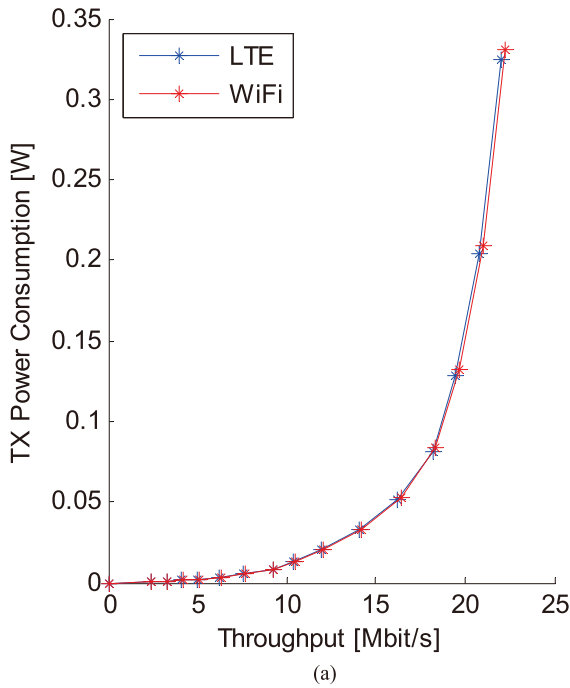
To evaluate our approach, we compare the power consumption for the same throughput constraints. In this context, we assume that the concerned Mobile Devices are able to influence resource allocation and to choose the uplink/downlink power levels each of the RATs being operated simultaneously. It is acknowledged that such a feature may not be available in current standards and is requiring future modifications. The performance depends on different parameters and on the working points of the separate links. Such diversity allows the establishment of various scenarios, combinations and results. We pick the best link for the single-link case, so that the multi-link improvement is considered for the addition of an equivalent or worse link. In Fig. 5(a) we simulate a hypothetical configuration, where the quality and the performance of each link are similar. This scenario allows the highest improvement percentage, because otherwise only the contribution of weaker links to the chosen one is possible. Comparing the power consumption values of the two-links case (red curve) with the single-link case (blue curve) in Fig. 5(b), the difference is not significant for small throughput values (0–10 Mbit/s). However, the power savings increase for large throughput values. Compared to a single optimal link, only 10.2% is needed for the multi-link configuration. In another perspective, if we fix the power values, we can compare the throughput performances between this mode and the two-links mode. Averaging all val-

ues in this configuration, we obtain 63.58% more throughput for the second mode. Note that all gains occur due to the fact that the simultaneous usage of RATs leads to a larger available bandwidth and thus lower MAC Modes can be applied leading finally to reduced output power level requirements.

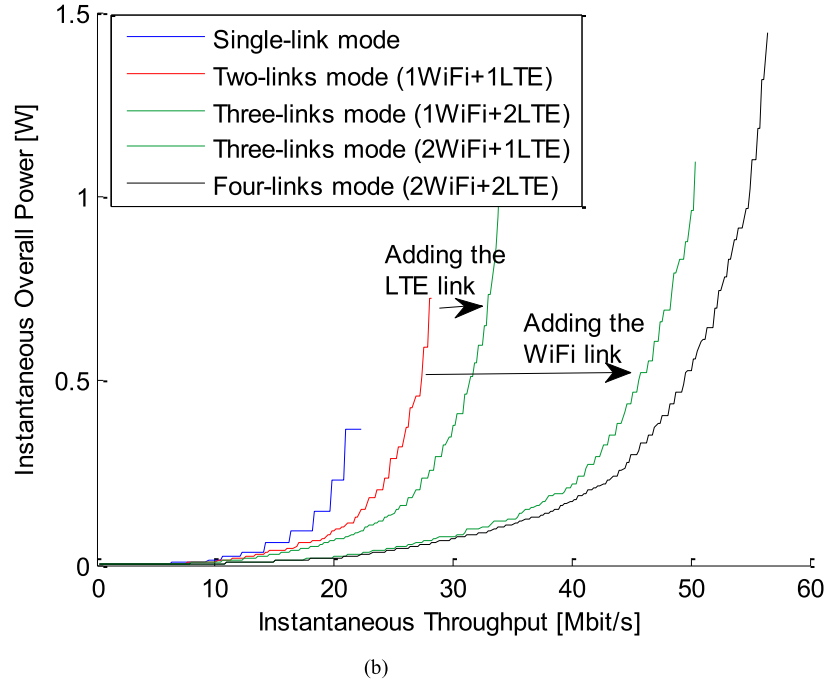
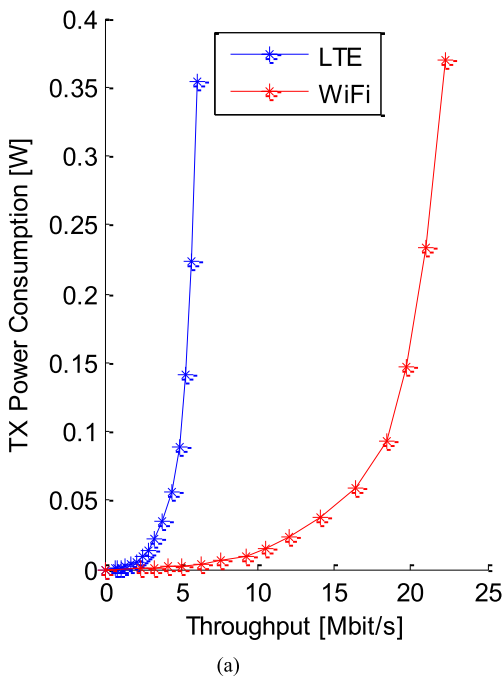
For the second hypothetical configuration, we set the LTE link to have approximately the same power values than the strong link. However, the throughput range is significantly shorter.

The working points of both links are shown in Fig. 6(a). We observe a similar behaviour to scenario 1. However, the power reduction is less drastic. In average, the two-links configuration consumes 47.4% of power needed for the single-link case. Fixing the power values and comparing the data rate values, we observe an improvement of 13.34% comparing to the 63.58% of scenario 1. Observe the improvement between adding a strong link (the WiFi in this case) and adding a weak link (the LTE link).

Taking into account our results, our first conclusion is that the power reduction increases with the quality of the second link. Within the same scenario, the power reduction is always dependant on the considered throughput. With low throughput requirements, the multi-link configuration is not particularly efficient and the improvement is only minor. However, with high requirements, and especially using aggressive MAC modes, our approach becomes attractive as well as when adding a third or fourth link. There are throughput points, where an addition a second link is attractive and the addition of a third one is not, and there are additionally particular points, where a third link is beneficial for the overall power consumption.



**Fig. 5** Results of Scenario 1 (indicating that the simultaneous usage of RATs leads to a reduced output power requirement for obtaining the target throughput due to usage of lower MAC Modes over the increased (aggregated) spectrum).



**Fig. 6** Results of Scenario 2.

### 5. Complexity Evaluation

We compare the computational cost and delivered solutions of the introduced algorithms. We discard the exhaustive search as it is inefficient for a high number of links. Already

for  $N_B = 5$ , the computational costs are infeasible. We focus on the Algorithm by Shoham and Gersho, the iterative version (second and third level) and the domain-cut variation. These methods offer a trade-off between the quality of the results and the computational costs as given in Table 1. For the numerical cost assessment, we use a model that assigns

**Table 1** Complexity Overview (The sub-block algorithm (Sect. 3.2) is used as a reference for the comparisons).

	Cases	Scenario 2 (3 links)	Scenario 2 (4 links)	Scenario 2 (5 links)	Scenario 1 (4 links)	Another scenario (4 links)
<b>Method 3 (Shoham-Gersho)</b>	Comp. Ratio	21.72%	15.19%	8.84%	9.46%	11.89%
	Power Diff.	+11.27%	+9.54%	+6.53%	+6.07%	+14.74%
<b>Method 3.1 (2 levels)</b>	Comp. Ratio	56.32%	41.56%	21.04%	23.85%	24.25%
	Power Diff.	+2.66%	+2.07%	+1.56%	+0.33%	+4.86%
<b>Method 3.1 (3 levels)</b>	Comp. Ratio	142.8%	137.4%	85.67%	69.58%	74.58%
	Power Diff.	+0.77%	+0.37%	+0.33%	+0.08	+2%

a cost to every machine operation, depending on the nature of the operation (addition, multiplication, comparison, etc.). The Sub-block algorithm (Sect. 3.2) obtains always optimal results and is used as a reference for the comparisons. The computation Ratio is the ratio of the average computational cost of one method to the reference method. The power difference is the average additional percentage of the algorithm results compared to the optimal solution results. Observe that the choice of the optimal algorithm depends on the system requirements.

If sub-optimal solutions are accepted and the computation efficiency is primordial, then the simple method by Shoham and Gersho is a good choice. If the system can handle higher computational costs and optimal transmission power is required, then the use of the Domain-Cut method is more logical. The other methods offer trade-offs between these two alternatives. It is important to estimate and compare the computational power and the transmission power consumption in order pick the most efficient method.

## 6. Conclusion

We proposed various methods for finding the optimal combination of multiple distinct Radio Access Technologies. A significant reduction in mobile device power consumption in comparison to using a single connection was observed: for optimal scenarios, 84.5% of the single-link power consumption is saved by applying this approach to achieve same throughputs. For other scenarios however, this reduction is only 38%. This depends heavily on the quality of the added link. Among the reviewed algorithms, the method by Shoham and Gersho shows lowest complexity if sub-optimal solutions are allowed. If the requirements are higher, the Domain-Cut method shows a better trade-off in terms of calculation complexity and optimality of the resulting working point.

The final implementation of the proposed methodology will, however, require an extension of existing wireless standards enabling Mobile Device to perform *MAC mode* selection as introduced in this paper.

In the next steps, the authors plan to further study the intelligent joint coding of a data stream across a multitude of heterogeneous RATs with distinct propagation characteristics. It is expected that joint coding will substantially improve the overall system performance.

## References

- [1] Y. Choi, H. Kim, S.-W. Han, and Y. Han, "Joint resource allocation for parallel multi-radio access in heterogeneous wireless networks," *IEEE Trans. Wireless Commun.*, vol.9, no.11, pp.3324–3329, 2010.
- [2] M. Msallem, Link selection optimization in a heterogeneous radio environment, Master Thesis, Technical University of Munich, Germany, 2014.
- [3] C. Gaie, Allocation de ressources à prise de décisions distribuées dans des réseaux mobiles hétérogènes, Ph.D. dissertation, Faculté des sciences d'Orsay, Université Paris-Sud 11, April, 2010.
- [4] D. Li and X. Sun, *Nonlinear Integer Programming*, Springer, Dordrecht, 2006.
- [5] Y. Shoham and A. Gersho, "Efficient bit allocation for an arbitrary set of quantizers (speech coding)," *IEEE Trans. Acoust., Speech, Signal Processing*, vol.36, no.9, pp.1445–1453, Sept. 1988.



**Markus Mueck** received the Electrical Engineering Diploma degrees of University of Stuttgart, Germany and Ecole Nationale Supérieure des Télécommunications (ENST), Paris, France in 1999 and the Doctorate Degree of ENST, Paris, France in 2006. Dr. Mueck is Senior Standardization Manager at Intel Mobile Communications, Germany, Adjunct Professor at Macquarie University, Sydney, Australia, Member of the Board at ETSI, Chairman of ETSI Reconfigurable Radio Systems Technical Body and Chairman of the IEEE Special Interest Group on Cognitive Radio in 5G.



**Majdi Msallem** received the B.S. and M.S. degrees in Electrical Engineering and Information Technology from the Technical University of Munich, Munich, Germany, in 2011 and 2014, respectively. He is now with Intel Mobile Communications working as a Systems Engineer.



**Christian Drewes** received Dipl.-Ing. and Dr.-Ing. degrees in Electrical Engineering and Information Technology from the Technische Universität München in 1995 and 1999. He's currently working at Intel on the architecture evolution of Intel's cellular modems and on end-user related cellular modem performance optimization.



**Markus Stinner** received his diploma degree in Electrical Engineering from Universität Ulm in 2011. Since 2011, he is doctoral student at the Institute of Communications Engineering where he works mainly on channel coding.