

Nonconcave Utility Maximization in the OFDMA Downlink

Johannes Brehmer and Wolfgang Utschick
Associate Institute for Signal Processing
Technische Universität München
{brehmer, utschick}@tum.de

Abstract—The problem of utility maximization in an OFDMA downlink is considered. Based on measurements of a scalable video codec, concave and nonconcave utility functions are proposed. Two approaches to solving the utility maximization problem in a modular fashion are considered, one based on Lagrange duality, the other based on a projected gradient algorithm. Both methods yield the optimum solution for concave utilities. In the nonconcave case, simulation results show almost identical performance for both methods under the assumption that the dual solution is found. In contrast, the dual approach fails if the dual subproblems can only be solved to local optimality.

I. INTRODUCTION

Utility maximization is widely recognized as a powerful framework for optimizing system parameters in wireless networks, see, e.g., [1], [2], [3]. A utility function usually represents a model of how system performance depends on the allocated resources, or, more generally, on the system's parameters. From an optimization point of view, convexity is a very desirable property of the utility maximization problem. As a result, the utility function is usually assumed to be concave. While a concave utility seems to be a good model for best-effort traffic, it may not be the most appropriate model for other traffic classes, such as voice and real-time video [4]. In this work, the scalable extension of H.264/AVC is considered as an exemplary application, which, due to its scalability with respect to source rate, fits well into a utility framework. Given a set of video quality measurements, different possibilities exist in the formulation of the utility model, in particular, one may either choose a convex or a nonconvex model.¹ The choice of the model's properties has an impact on both the accuracy of the solution as well as on the methods available for solving the utility maximization problem.

¹We use the term “convex utility model” for a model with concave utility function, in order to emphasize the fact that the resulting optimization problem is convex.

On the physical layer, multicarrier techniques represent a promising candidate for next-generation wireless systems. In this work, the downlink of an infrastructure wireless network is considered, where the users share the common resources transmit power and subcarriers, and each subcarrier is allocated to a single user only (OFDMA). It is assumed that resources are allocated by a central unit at the transmitter. Optimum power and subcarrier allocation in the OFDMA downlink has recently attracted significant interest, where optimality criteria range from weighted-sum rate maximization [5] to more general utility models [6], [7], [8]. This work differs from previous work in two aspects: First, previous works on resource allocation in the OFDMA downlink usually assumed a concave utility function. Second, a modular approach to solving the utility maximization problem is taken, i.e., given a generic utility model, the utility maximization problem is not solved from scratch, but by employing available efficient solutions for maximizing the weighted-sum rate in the OFDMA downlink. Following [9], two such modular approaches are considered, the well-known dual decomposition, and a method which operates on the boundary of the rate region.

In the context of network utility maximization (NUM), the case of nonconcave utility functions has been investigated in [4], [10]. Both [4], [10] consider a slightly different setup, with a more complex network topology, but simpler models concerning the link capacities. While [4] proposes a distributed (heuristic) solution to the nonconvex NUM problem, [10] develops a centralized method for computing tight upper bounds on the optimum utility value. In contrast, this work investigates centralized methods for computing good (but potentially suboptimum) power and subcarrier allocations in a single-hop OFDMA downlink.

II. PROBLEM SETUP

An OFDMA downlink with K users and N subcarriers is considered. The usual model of OFDMA as a set of parallel channels, with each channel allocated to a single user, is employed. The received signal on subchannel n allocated to user k is given by

$$y_{n,k} = \sqrt{p_{n,k}} h_{n,k} x_{n,k} + \eta_{n,k},$$

where $p_{n,k}$ denotes the transmit power allocated to this subchannel, $h_{n,k}$ is the channel coefficient, $x_{n,k}$ the transmitted unit power signal and $\eta_{n,k}$ additive white Gaussian noise with variance $\sigma_{n,k}^2$. Accordingly, the capacity of subchannel n is given by

$$r_{n,k}(p_{n,k}) = \log_2 \left(1 + \sigma_{n,k}^{-2} |h_{n,k}|^2 p_{n,k} \right).$$

Let $\boldsymbol{\pi} \in \{1, \dots, K\}^N$ denote the subchannel allocation, i.e., if subchannel n is allocated to user k , then $\pi_n = k$. Thus, for a given power and subchannel allocation $(\boldsymbol{p}, \boldsymbol{\pi})$, an achievable rate of user k is given by

$$r_k(\boldsymbol{p}, \boldsymbol{\pi}) = \sum_{n:\pi_n=k} r_{n,k}(p_{n,k}).$$

Power allocation is subject to a sum power constraint, i.e., the set of feasible power allocations is given by

$$\mathcal{P} = \{\boldsymbol{p} : \boldsymbol{p} \geq \mathbf{0}, \|\boldsymbol{p}\|_1 \leq P_{\text{tr}}\}.$$

Moreover, let $\Pi = \{1, \dots, K\}^N$ denote the set of feasible subchannel allocations. The rate region achievable by certain choices of the parameters $(\boldsymbol{p}, \boldsymbol{\pi})$ is given by

$$\mathcal{R} = \bigcup_{\boldsymbol{\pi} \in \Pi} \{\boldsymbol{r}(\boldsymbol{p}, \boldsymbol{\pi}) : \boldsymbol{p} \in \mathcal{P}\}, \quad (1)$$

with $\boldsymbol{r} = (r_1, \dots, r_K)$. With the usual time-sharing argument, the capacity region \mathcal{C} is defined as the convex hull of \mathcal{R} :

$$\mathcal{C} = \text{Co}(\mathcal{R}). \quad (2)$$

According to Eqs. (1) and (2), every $\boldsymbol{r} \in \mathcal{C}$ can be written as

$$\boldsymbol{r} = \sum_{w=1}^W \alpha_w \boldsymbol{r}(\boldsymbol{p}_w, \boldsymbol{\pi}_w),$$

with $\alpha_w \geq 0$ and $\sum_{w=1}^W \alpha_w = 1$. Note that the provision of a time-sharing mode may be undesirable in a practical system. This issue will be further discussed in later sections.

The properties of upper layers are modeled by a utility function $u : \mathbb{R}_{0,+}^K \rightarrow \mathbb{R}$, which maps a rate vector \boldsymbol{r} into a utility value $u(\boldsymbol{r})$. The utility function is assumed

to be strictly monotonically increasing (i.e., larger rates yield better performance) and differentiable. In contrast to the usual setup, u is not required to be concave. More details on the utility model used in this work are provided in Section III. Given a utility function u , the goal is twofold: First, find the rate vector \boldsymbol{r}^* that maximizes utility, i.e., solves

$$\max_{\boldsymbol{r} \in \mathcal{C}} u(\boldsymbol{r}). \quad (3)$$

Second, find the physical layer parameters $(\alpha_w, \boldsymbol{p}_w, \boldsymbol{\pi}_w)_{w=1}^W$ that implement the optimum rate vector. The second step is required because the optimum parameter setup is required at transmission time, i.e., the system parameters have to be adjusted such that \boldsymbol{r}^* is achievable. Problem (3) is convex if the utility function u is concave. Due to convex hull operation, however, the set \mathcal{C} is not directly accessible to optimization methods, as no explicit parametrization of \mathcal{C} is available.

A special case is given by

$$u(\boldsymbol{r}) = \boldsymbol{\lambda}^T \boldsymbol{r},$$

i.e., *weighted sum rate maximization* (WSRmax). For WSRmax,

$$\max_{\boldsymbol{r} \in \mathcal{C}} \boldsymbol{\lambda}^T \boldsymbol{r} = \max_{\boldsymbol{r} \in \mathcal{R}} \boldsymbol{\lambda}^T \boldsymbol{r} \quad (4)$$

$$= \max_{\boldsymbol{p} \in \mathcal{P}, \boldsymbol{\pi} \in \Pi} \boldsymbol{\lambda}^T \boldsymbol{r}(\boldsymbol{p}, \boldsymbol{\pi}). \quad (5)$$

Recently, efficient algorithms for (approximately) solving problem (5) have been proposed, e.g., [5]. Note that problem (5) is nonconvex, due to the discrete nature of the subcarrier allocation. In the following, however, it is assumed that the WSRmax problem can be solved to optimality.

The availability of efficient algorithms for solving the WSRmax problem in the OFDMA downlink motivates a modular approach to the utility maximization problem: Instead of solving the utility maximization problem in the OFDMA downlink from scratch, two approaches are presented which transform the original problem into a sequence of WSRmax problems.

III. UTILITY MODEL

Most works on utility maximization assume a concave utility function u . This work is motivated by the observation that a real-world application may be modeled more precisely by a nonconcave utility function.

The utility model used in this work is based on QoS measurements of a state-of-the-art scalable video codec

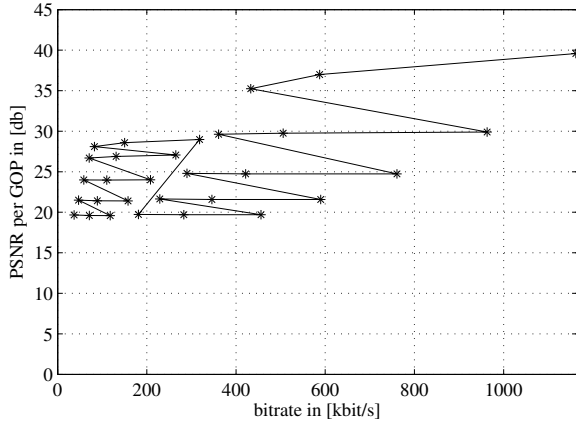


Fig. 1. Operation points of scalable video codec

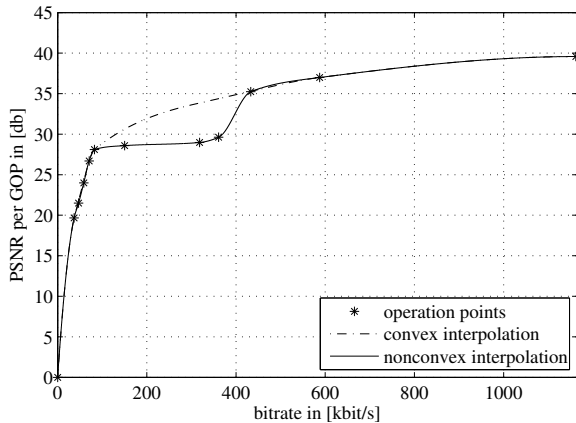


Fig. 2. Interpolated utility models

[11]. In our measurements, we used the reference implementation of the H.264/AVC scalable video extension [12]. For different operation modes, required source rate and resulting *peak signal-to-noise ratio* (PSNR) were recorded, resulting in a set of rate/PSNR pairs. Such an exemplary set of operation points is shown in Figure 1, where each star corresponds to an operation point (Note that points are connected by lines for a better visual impression).

In the chosen measurement setup, some of the operation points yield lower PSNR at higher source rate, therefore can be excluded from the set of relevant operation points. For the example under consideration, the set of relevant operation points is shown in Figure 2.

One of the basic assumptions is that the utility function u is continuous. Consequently, a utility model is generated by a continuous interpolation between the set of operation points. Two interpolation methods are

investigated: a nonconvex interpolation, which attempts to follow the operation points as closely as possible, and a convex interpolation, which further drops some points in order to obtain a concave utility function. For the example under consideration, the two resulting utility curves are also shown in Figure 2.

The details of the measurement setup are outside the scope of the paper. Instead of focusing on how the set of operation points is generated, the idea is to take the set of operation points as given and to investigate how to best deal with the properties of this set.

So far, utility models for a single video sequence have been considered. In a multi-user system, each user is assumed to receive a video sequence. How to best map the user's utilities u_k into an overall utility function u represents an important problem. In the following, the sum of users' utilities is considered:

$$u(\mathbf{r}) = \sum_{k=1}^K u_k(r_k).$$

IV. DUAL DECOMPOSITION

The dual decomposition is a well-known method to decompose the utility maximization problem into a set of coupled subproblems, with each subproblem corresponding to a layer (vertical decomposition) [2]. If the dual decomposition is applied, it is usually assumed that strong duality holds – which basically requires problem (3) to be convex.

A. Basic Algorithm

Following [2], problem (3) is first modified by introducing additional variables:

$$\max_{\mathbf{r}, \mathbf{s}} u(\mathbf{s}) \quad \text{s.t.} \quad \mathbf{s} \leq \mathbf{r}, \mathbf{r} \in \mathcal{C}. \quad (6)$$

The corresponding dual function is given by

$$g(\boldsymbol{\lambda}) = g_A(\boldsymbol{\lambda}) + g_P(\boldsymbol{\lambda}),$$

and

$$g_A(\boldsymbol{\lambda}) = \max_{\mathbf{s}} u(\mathbf{s}) - \boldsymbol{\lambda}^T \mathbf{s}, \quad \text{and} \quad (7)$$

$$g_P(\boldsymbol{\lambda}) = \max_{\mathbf{r} \in \mathcal{C}} \boldsymbol{\lambda}^T \mathbf{r}. \quad (8)$$

For a fixed $\boldsymbol{\lambda}$, the optimization is decomposed into two subproblems (7) and (8). Note that (8) is a WSRmax problem, which, based on our assumption, can be solved to optimality. Let \mathbf{s}_λ and \mathbf{r}_λ denote maximizers of problems (7) and (8), respectively. Then $\mathbf{r}_\lambda - \mathbf{s}_\lambda$ is a subgradient of g at $\boldsymbol{\lambda}$ [13]. There exist $\boldsymbol{\lambda}$ for which the maximizers are not unique, as a consequence, the

dual function is nondifferentiable. In this work, a cutting plane-based outer linearization method is used to find the optimum dual variable [13]. The outer linearization method iteratively refines a polyhedral approximation of the dual function based on subgradients of g at iterates $\lambda^{(n)}$.

Clearly, the goal is to obtain the optimum rate vector \mathbf{r}^* , as well as the optimum power and subcarrier allocation that implement \mathbf{r}^* . In other words, a strategy for recovering a primal solution \mathbf{r}^* and parameters $(\alpha_w^*, \mathbf{p}_w^*, \boldsymbol{\pi}_w^*)_{w=1}^W$ from λ^* is required.

B. Convex Case

If the utility function u is concave, problem (6) is convex and satisfies Slater's condition, thus strong duality holds. Due to strong duality, $\mathbf{r}^* \in \operatorname{argmax}_{\mathbf{r} \in \mathcal{C}} (\boldsymbol{\lambda}^*)^T \mathbf{r}$ [13]. If $\operatorname{argmax}_{\mathbf{r} \in \mathcal{C}} (\boldsymbol{\lambda}^*)^T \mathbf{r}$ is a singleton set, $W = 1$ and the solution is determined by $(\mathbf{p}^*, \boldsymbol{\pi}^*)$ that solve problem (5).

If $\operatorname{argmax}_{\mathbf{r} \in \mathcal{C}} (\boldsymbol{\lambda}^*)^T \mathbf{r}$ is *not* a singleton set, i.e., \mathbf{r}^* lies in a time-sharing region, primal recovery is more involved. To obtain the optimum solution, it would be required to construct time-sharing solutions, an issue that is further discussed in the context of a MIMO downlink in [14]. In the OFDMA downlink, a suboptimum, but simple approach is to assume that

$$u(\mathbf{r}^*) \approx u(\mathbf{r}), \forall \mathbf{r} \in \operatorname{argmax}_{\mathbf{r} \in \mathcal{R}} (\boldsymbol{\lambda}^*)^T \mathbf{r},$$

i.e., it is assumed that the corner points of the time-sharing regions are a good-enough approximation of the optimum solution. Based on this approximation, set $W = 1$ and $\hat{\mathbf{r}} = \mathbf{r}(\hat{\mathbf{p}}, \hat{\boldsymbol{\pi}})$, where $(\hat{\mathbf{p}}, \hat{\boldsymbol{\pi}})$ are the physical layer parameters returned by the WSRmax algorithm, given λ^* . Due to the fact that the optimum solution is approximated by one of the corner points of the time-sharing region, the system no longer needs to provide a time-sharing mode.

The approximation of \mathbf{r}^* by $\hat{\mathbf{r}}$ is based on the intuition that for a large enough number of subcarriers, the time-sharing regions are relatively small – and, implicitly, that the utility function is sufficiently flat around \mathbf{r}^* . While a rigorous justification remains an open problem, the approach works sufficiently well in practice [8].

C. Nonconvex Case

For nonconcave utility u , subproblem (7) is nonconvex. If it can be solved to optimality at each iteration of the cutting plane method (i.e., for each $\lambda^{(n)}$, a global maximizer $\mathbf{s}_{\lambda^{(n)}}$ is determined), the cutting plane method (or any other subgradient-based method) converges to

the optimum dual variable λ^* , as the dual function is always convex. Finding a global maximizer is possible if the maximizers can be computed analytically. On the other hand, if the function u is such that a maximizer of (7) has to be computed numerically, it may not be guaranteed that the numerically obtained maximizer is a global maximizer. If a local maximizer is obtained, then $\mathbf{r}_{\lambda} - \mathbf{s}_{\lambda}$ no longer represents a subgradient of g . As a result, a local maximizer leads to an erroneous polyhedral approximation of g , and the cutting plane method is likely to not find the optimum dual variable.

For nonconcave utility functions, strong duality can no longer be assumed and it may be the case that $\mathbf{r}^* \notin \operatorname{argmax}_{\mathbf{r} \in \mathcal{C}} (\boldsymbol{\lambda}^*)^T \mathbf{r}$. By weak duality, $g(\boldsymbol{\lambda}^*)$ provides an upper bound on the maximum utility value. The main interest, however, lies in obtaining a good rate vector $\hat{\mathbf{r}}$ and the corresponding physical layer parameters. Let $\hat{\boldsymbol{\lambda}}$ denote the dual variable returned by the cutting plane algorithm. Each $\mathbf{r} \in \operatorname{argmax}_{\mathbf{r} \in \mathcal{C}} (\hat{\boldsymbol{\lambda}})^T \mathbf{r}$ represents a feasible, though not necessarily optimum, solution to the original primal problem (3). Thus, $\hat{\mathbf{r}}$ is chosen from $\operatorname{argmax}_{\mathbf{r} \in \mathcal{R}} (\boldsymbol{\lambda}^*)^T \mathbf{r}$.

V. EFFICIENT SET APPROACH

A *primal method* [15] for solving the utility maximization problem was proposed in [9]. The method, denoted as *iterative efficient set approximation* (IEA), basically corresponds to a projected gradient algorithm on the boundary of the rate region \mathcal{C} . As an ascent method, the projected gradient method provides convergence to stationary points [15].

A. Basic Algorithm

Define the Pareto efficient set as follows:

$$\mathcal{E} = \{\mathbf{r} \in \mathcal{C} : \nexists \mathbf{r}' \in \mathcal{C}, \mathbf{r}' > \mathbf{r}\}. \quad (9)$$

The efficient set \mathcal{E} is a subset of the boundary of \mathcal{C} and contains the rate vectors which are largest under the partial Pareto order. Due to the monotonicity of the utility function, the optimum rate vector \mathbf{r}^* is element of \mathcal{E} . This motivates the use of a gradient ascent method on the efficient boundary.

The algorithm proceeds as follows: Starting at $\mathbf{r}^{(k)}$, the gradient of the utility function at $\mathbf{r}^{(k)}$ and an orthonormal basis $\mathbf{B}(\mathbf{r}^{(k)})$ of the tangent space of \mathcal{C} at $\mathbf{r}^{(k)}$ are computed. The utility gradient is projected onto the tangent space, and this projected gradient is used to give the ascent direction in the tangent space. With a stepsize t , a new point $\tilde{\mathbf{r}}^{(k+1)}$ results:

$$\tilde{\mathbf{r}}^{(k+1)} = \mathbf{r}^{(k)} + t \mathbf{B}(\mathbf{r}^{(k)}) \mathbf{B}(\mathbf{r}^{(k)})^T \nabla u(\mathbf{r}^{(k)}).$$

In general, $\tilde{\mathbf{r}}^{(k+1)} \notin \mathcal{E}$, therefore a second projection is required, which projects $\tilde{\mathbf{r}}^{(k+1)}$ onto \mathcal{E} to yield the next iterate:

$$\mathbf{r}^{(k+1)} = \text{P}_{\mathcal{E}}(\tilde{\mathbf{r}}^{(k+1)}).$$

The stepsize t is determined via a backtracking line-search.

Due to the structure of \mathcal{E} , the projection step is the most complex step in each iteration. A projection orthogonal to the tangent space is employed. Let \mathbf{n} denote the unit-norm vector that is orthogonal to $\text{span}(\mathbf{B}(\mathbf{r}^{(k)}))$ and points away from \mathcal{C} . To project $\tilde{\mathbf{r}}$ on \mathcal{E} , the following problem is solved:

$$\max_{x, \mathbf{r}} x \quad \text{s.t.} \quad \tilde{\mathbf{r}} + x\mathbf{n} \leq \mathbf{r}, \quad \mathbf{r} \in \mathcal{C}. \quad (10)$$

The solution is again obtained by Lagrange duality. As shown in [9], the optimum dual variable is found by solving

$$\min_{\boldsymbol{\lambda} \geq 0} \boldsymbol{\lambda}^T (\mathbf{r}^*(\boldsymbol{\lambda}) - \tilde{\mathbf{r}}) \quad \text{s.t.} \quad \boldsymbol{\lambda}^T \mathbf{n} = 1, \quad (11)$$

where $\mathbf{r}^*(\boldsymbol{\lambda}) \in \text{argmax}_{\mathbf{r} \in \mathcal{C}} \boldsymbol{\lambda}^T \mathbf{r}$. Again, the solution allows to reuse existing WSRmax algorithms. Given the optimum dual variable $\boldsymbol{\lambda}^*$ the same strategy for primal approximation as in case of the dual decomposition is applied.

Due to the fact that the projection has to be carried out at each outer iteration, the IEA method generally requires more calls to the WSRmax algorithm than the dual decomposition. In contrast, the dual decomposition requires the solution of an application layer problem (7) at each iteration, while IEA only requires evaluation of the gradient of the utility function.

B. Convex Case

Under convexity, any local optimum is globally optimum, therefore the gradient ascent on \mathcal{E} with backtracking line search converges to the optimum solution \mathbf{r}^* . Due to the applied strategy for primal approximation, solutions in time-sharing regions are again approximated by a corner point of the time sharing region, implying the same assumptions already discussed in case of the dual decomposition.

C. Nonconvex Case

In contrast to the dual decomposition, the efficient set approach provides convergence to a local optimum. Moreover, even if convergence to the global optimum cannot be guaranteed, starting from an initial point, the efficient set method always provides an improvement

in utility (assuming the initial point is not a stationary point). The most complex step, the projection on \mathcal{E} , is completely independent of the properties of the utility function.

VI. SIMULATION RESULTS

In this section, simulation results are used to compare the performance of the two methods discussed in the previous section. A scenario with $K = 2$ users is considered. As a reference, the utility value achieved by the rate vector that maximizes sum-rate is also shown. The sum-rate point does not take into account the properties of the utility function and serves as a starting point for the IEA algorithm.

An OFDMA system with 512 subcarriers is simulated, using the WSRmax algorithm from [5]. The channel coefficients $h_{n,k}$ are iid. complex Gaussian with zero mean and variance σ_k^2 . In the simulations, $\sigma_2^2 = 10\sigma_1^2$, i.e., user 2 has a higher channel gain on average. Results are averaged over 500 channel realizations.

Both users have the same utility function u_k , which, depending on the scenario, either correspond to the convex or the nonconvex interpolation shown in Figure 2.

Figure 3 shows the average sum utility $u = u_1 + u_2$ for different values of maximum transmit power P_{tr} that is achieved by the IEA method (IEA), by the dual decomposition (DD-GA), and by choosing the sum rate maximizing rate point, for concave u . A gradient ascent method is used to solve the application layer subproblem (7) in the dual decomposition.

As expected, for a convex problem, both IEA and dual decomposition provide identical performance. Moreover, optimizing power and subcarrier allocation with respect to the properties of the upper layers offers a significant gain in utility compared to a physical layer setup that maximizes sum rate.

Figure 4 shows the average sum utility $u = u_1 + u_2$ for different values of maximum transmit power P_{tr} in the case of nonconcave utility. For the dual decomposition, two graphs are shown. As before, the graph labelled DD-GA corresponds to the dual decomposition with the application layer subproblem solved numerically with a gradient method. In contrast, for DD-ES the subproblem is solved via an exhaustive search over sampled intervals. While the exhaustive search can always find a globally optimum solution, the gradient method in some cases only returns local maximizers. The simulation results show that DD-GA often fails in providing a good solution. On the other hand, IEA and DD-ES show almost identical performance and again provide a significant

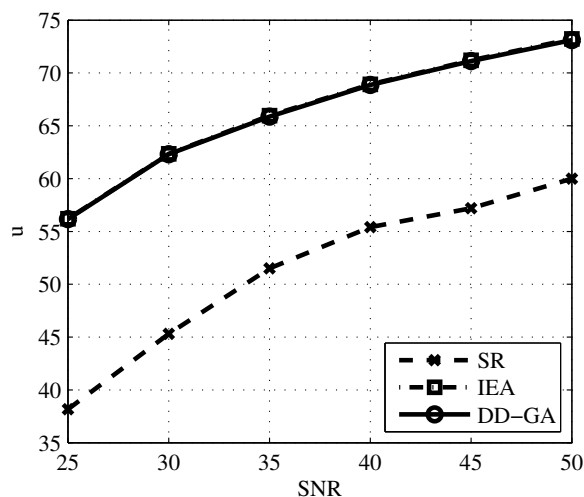


Fig. 3. Convex Interpolation

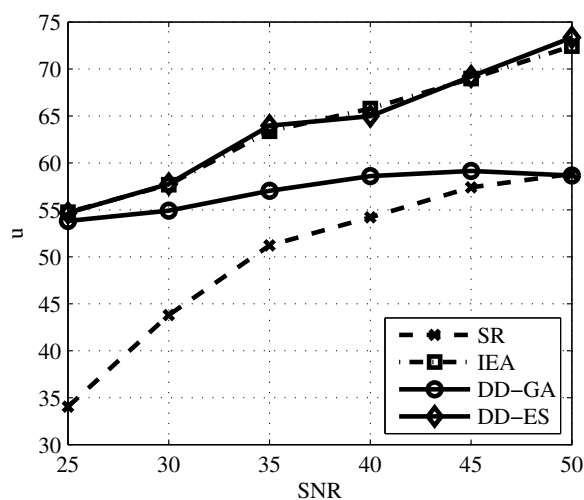


Fig. 4. Nonconvex Interpolation

performance gain compared to the maximum sum rate strategy.

VII. CONCLUSIONS

Utility maximization in an OFDMA downlink is investigated. Two methods for modular parameter optimization are applied to solve the utility maximization problem. Moreover, two cases are considered: A convex and a nonconvex utility model. For a convex model, both methods provide the optimum solution, and significantly outperform the sum rate maximization strategy. In the nonconvex case, the local convergence properties of the IEA method lead to good performance of this approach. The performance of the dual decomposition strongly depends on whether the application layer subproblem

can be solved to optimality. If a global maximizer is found, the dual decomposition, combined with a heuristic rule for primal recovery, yields good performance in the OFDMA downlink, even for nonconcave utility. On the other hand, if the optimum dual variable cannot be found, the dual decomposition shows significantly lower performance than the IEA method.

REFERENCES

- [1] S. Stanczak, M. Wiczanowski, and H. Boche, *Resource Allocation in Wireless Networks - Theory and Algorithms*, ser. Lecture Notes in Computer Science (LNCS 4000). Springer-Verlag, 2006.
- [2] M. Chiang, S. H. Low, A. R. Calderbank, and J. C. Doyle, "Layering as optimization decomposition: a mathematical theory of network architectures," *Proceedings of the IEEE*, vol. 95, 2007.
- [3] X. Lin, N. B. Shroff, and R. Srikant, "A tutorial on cross-layer optimization in wireless networks," *IEEE Journal on Selected Areas in Communications*, vol. 24, pp. 1452–1463, 2006.
- [4] J. W. Lee, R. R. Mazumdar, and N. B. Shroff, "Nonconvexity issues for internet rate control with multiclass services: stability and optimality," in *IEEE INFOCOM*, vol. 1, 2004, p. 34.
- [5] K. Seong, M. Mohseni, and J. M. Cioffi, "Optimal resource allocation for OFDMA downlink systems," in *IEEE International Symposium on Information Theory (ISIT)*, 2006, pp. 1394–1398.
- [6] G. Song and Y. Li, "Cross-layer optimization for OFDM wireless networks – part I: theoretical framework," *IEEE Transactions on Wireless Communications*, vol. 4, pp. 614–624, 2005.
- [7] J. Huang, V. Subramanian, R. Agrawal, and R. Berry, "Downlink scheduling and resource allocation for OFDM systems," in *Proceedings of the Conference on Information Sciences and Systems (CISS)*, March 2006.
- [8] T. C. Ng, W. Yu, J. Zhang, and A. Reid, "Joint optimization of relay strategies and resource allocations in cooperative cellular networks," in *Proceedings of the Conference on Information Sciences and Systems (CISS)*, March 2006.
- [9] J. Brehmer and W. Utschick, "A decomposition of the downlink utility maximization problem," in *Conference on Information Sciences and Systems (CISS)*, Baltimore, USA, March 2007.
- [10] M. Fazel and M. Chiang, "Network utility maximization with nonconcave utilities using sum-of-squares method," in *Decision and Control, 2005 and 2005 European Control Conference. CDC-ECC '05. 44th IEEE Conference on*, 2005, pp. 1867–1874.
- [11] H. Schwarz, D. Marpe, and T. Wiegand, "Overview of the scalable video coding extension of the H.264/AVC standard," *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 17, 2007.
- [12] JVT, "JSVM Software Manual," 2007, version JSVM 8.9.
- [13] D. Bertsekas, A. Nedic, and A. Ozdaglar, *Convex analysis and optimization*. Athena Scientific, 2003.
- [14] J. Brehmer and W. Utschick, "Utility maximization strategies in the multi-user MIMO downlink," in *International Workshop on Cross Layer Design (IWCLD)*, Jinan, China, September 2007.
- [15] D. G. Luenberger, *Linear and Nonlinear Programming*, 2nd ed. Addison-Wesley, Inc., 1984.