Modular Cross-Layer Optimization Based on Layer Descriptions

Johannes Brehmer and Wolfgang Utschick

Institute for Circuit Theory and Signal Processing
Munich University of Technology, 80333 Munich, Germany
{brehmer, utschick}@tum.de

***Abstract*— Future wireless communication systems require the adaptation of the parameters of all layers in order to efficiently provide a multitude of different applications. The optimality of a parameter setup is determined at the application layer. A parameter adaptation can be achieved by discarding the layered architecture and merging all layer parameters into a single optimization. In order to preserve the layered structure, a concept for parameter optimization based on the passing of layer descriptions from bottom to top layer is proposed. Theoretically optimum and practically feasible descriptions are discussed. It is shown that multiobjective optimization plays a key role in the description-based optimization of layered communication systems. The proposed concept is applied to an exemplary system.**

1. Introduction

A key challenge in the design of future wireless communication systems is the efficient provision of a multitude of different applications. In this work, we assume that the ultimate goal when setting up the various parameters of such a system is to maximize the application layer quality of service experienced by the end users, under the constraint of limited resources. Thus, the optimality of a parameter setup at each layer is determined based on an application layer criterion. Optimum performance can only be achieved if the parameter setup in each layer takes into account the properties of the application. This observation leads directly to the field of cross-layer optimization.

In a first approach to cross-layer optimization, the parameters of several adjoining layers are optimized jointly. *Joint optimization* requires full transparency between layers. Thus, the modular structure provided by a layered architecture is abandoned in favor of increased performance – but sacrificing the advantages of a layered architecture may lead to unintended side effects. Modular cross-layer optimization limits the amount of transparency required. *Modularity* signifies that the parameter setup at a certain layer is independent of the "inner" details of the other layers. In the parameter setup process, such abstraction can be achieved by an interface-based information exchange between layers. In order to achieve modularity, interfaces are designed such that the details of a parameter setup remain hidden to the other layers. An example for an interface are the utility functions used in [1]. The utility function used in [1] represents an abstraction of the properties of the upper layers that is provided to the lower layers ("top-down"information exchange). Importantly, the formulation of the utility function is independent of the properties of the lower layers.

We propose a concept in which each layer describes its capabilities by a finite set of feasible operation points. This set, denoted as *layer description*, serves as an interface between adjacent layers. Each layer provides its description to the neighboring upper layer. In this way, information about feasible operation points is propagated from bottom layer to top layer, giving rise to the term "bottom-up" optimization.

A layer description represents an abstract description of feasible parameter setups, i.e., while a layer description represents the common language between adjacent layers, it does not matter to the upper layer how the operation points in the description are realized by the lower layer. In the same manner, for the lower layer it does not matter how the upper layers map an operation point from its description into application layer performance. Thus, except for the interface specifications (i.e. how to describe operation points), the designs of all layers remain independent of each other.

The description-based approach can be interpreted as the reversal of utility-based top-down approaches: While a utility function is formulated independent of the lower layers, a layer description is generated independent of the properties of the upper layers (except for a monotonicity constraint in both cases, see [1] and Section 5).

A generic mathematical model for the parameter setup problem in a layered system is introduced in Section 3. Parameter optimality is defined in Section 4. Modular optimization based on layer descriptions is described in detail in Section 5. Concerning the generation of layer descriptions, two design objectives can be identified: On the one hand, we desire descriptions for which the best operation point obtained by optimization based on layer descriptions coincides with the optimum operation point achievable by joint optimization. Such optimum descriptions are discussed in Section 6. On the other hand, description size is limited due to complexity reasons. In Section 7, we discuss how to minimize the deviation from the optimum solution for a finite number of elements in a description.

Throughout the paper, we use an exemplary wireless communication system to illustrate the application of the proposed generic concept. The exemplary system is introduced in Section 2. For the exemplary system, simulation results provided in Section 8 demonstrate that modular optimization based on layer descriptions can achieve close to optimum performance.

2. Exemplary system

We consider the optimization of the layer parameters at the transmitter in a cellular system with one central transmitter and two non-cooperating receivers. Based on the assumption of a cellular structure, we only consider a single hop in the wireless domain and neglect routing issues in the optimization. Moreover, we assume a system with an ARQ-free protocol at the transport layer and without queuing of frames at the MAC layer, modeling traffic with tight delay constraints. Based on these assumptions, we employ a simplified architecture consisting of only two layers: layer 1 comprises physical layer signal processing and FEC, layer 2 corresponds to the application layer of the OSI reference model.

At layer 1, assuming an orthogonal multiple access scheme such as TDMA, the wireless link is modeled as parallel channels $h_k \in \mathbb{C}$, $k = 1, 2$, to the two users. The channels h_k are assumed to be i.i.d. block fading with coherence time T_c and Rayleigh distributed amplitude. The average gain of channel h_k is given by $\rho_k = \text{E}\left[|h_k|^2\right]$. The received signals are distorted by complex-valued AWGN with variance $\sigma^2_{\text{n},k}$.

The physical layer signal processing at the transmitter corresponds to allocating a transmit power δ_k to each of the 2 users, subject to a transmit power constraint $\delta_1 + \delta_2 \leq P_\text{tr}$. We consider the case that the transmitter knows the average gains ρ_k, but not the channel realizations h_k. Under the given assumptions, the probability that the instantaneous capacity of the channel to the kth user is smaller than a prescribed rate R_k is given by the outage probability [2]

$$\varepsilon_k(R_k, \delta_k) = 1 - \exp\left(-\frac{\sigma^2_{\text{n},k}}{\rho_k \delta_k}\left(2^{R_k} - 1\right)\right).$$

We make the (idealizing) assumption of capacity-achieving coding and signaling. In this case, the packet error probability for a packet of length T_c transmitted to user k is given by ε_k.

At layer 2, the transmission of quantized data is considered. As the channel is lossy, the application layer distortion of the kth user results from distortion due to quantization $(D_{\text{Q},k})$ and from distortion due to packet loss $(D_{\text{L},k})$. While $D_{\text{Q},k}$ depends on the source quantization rate $R_{\text{Q},k}$, $D_{\text{L},k}$ depends on the packet error probability ε_k. Adopting the model from [3], we assume that both distortions can be combined as follows:

$$\tilde{D}_k(R_{\text{Q},k}, \varepsilon_k) = D_{\text{Q},k}(R_{\text{Q},k})\,(1 - \varepsilon_k) + D_{\text{Q},k}(0)\varepsilon_k.$$

Taking into account that the source quantization rate $R_{\text{Q},k}$ cannot be larger than the transmission rate, the overall application layer distortion for the kth user is modeled as

$$D_k(R_{\text{Q},k}, \varepsilon_k, R_k) = \begin{cases} \tilde{D}_k(R_{\text{Q},k}, \varepsilon_k), & R_{\text{Q},k} \leq R_k \\ \tilde{D}_k(0, 1), & R_{\text{Q},k} > R_k. \end{cases}$$

The choices for a rate-distortion function $(D_{\text{Q},k})$ are manifold. Rate-distortion models for video data are presented, e.g., in [4]. For simplicity, we employ the following rate-distortion function:

$$D_{\text{Q},k}(R_{\text{Q},k}) = 2^{-2R_{\text{Q},k}},$$

which, under the assumption of infinite block length, can be interpreted as the rate-distortion function of a real-valued Gaussian source with unit variance.

3. Layer model

The problem of finding the optimum setup for the parameters of each layer is considered. We start by defining a mathematical model for the parameter setup problem in a layered system. Let N denote the number of layers. The nth layer is modeled by a set $\mathcal{X}_n$ of feasible parameter values and a so-called layer function

$$\boldsymbol{f}_n : \mathcal{X}_n \times \mathcal{D}_{n-1} \rightarrow \mathbb{R}^{N_{\text{o},n}}. \tag{1}$$

The set $\mathcal{D}_{n-1}$ is provided by the $(n-1)$th layer and is used for information exchange between layers. Given an element $\boldsymbol{d} \in \mathcal{D}_{n-1}$ and a feasible parameter choice $\boldsymbol{x} \in \mathcal{X}_n$, the layer function provides an abstract representation of $(\boldsymbol{x}, \boldsymbol{d})$ in terms of $N_{\text{o},n}$ quantities. The $N_{\text{o},n}$ quantities are to be chosen such that all relevant characteristics of a layer are captured. Without loss of generality, it is assumed that each quantity corresponds to a cost, i.e., the layer function assigns a cost vector $\boldsymbol{f}_n(\boldsymbol{x}, \boldsymbol{d})$ to each parameter choice. Moreover, it is assumed that the layer function of the top layer is scalar valued, i.e., $N_{\text{o},N} = 1$.

The layer output set is given by

$$\mathcal{Y}_n = \left\{\boldsymbol{f}_n(\boldsymbol{x}_n, \boldsymbol{d}_{n-1}) \in \mathbb{R}^{N_{\text{o},n}} : \boldsymbol{x}_n \in \mathcal{X}_n, \boldsymbol{d}_{n-1} \in \mathcal{D}_{n-1}\right\}.$$

A description $\mathcal{D}_n$ of the nth layer is generated by selecting elements from $\mathcal{Y}_n$, i.e., $\mathcal{D}_n \subseteq \mathcal{Y}_n$, see Sections 6 and 7. The bottom layer receives no input, i.e., $\mathcal{D}_0 = \emptyset$.

Applied to our exemplary system, layer parameters and layer functions are given as follows: At layer 1, we collect the parameters into a vector

$$\boldsymbol{x}_1 = (\delta_1, \delta_2, R_1, R_2).$$

For simplicity, we exclude the special cases $\delta_k = 0$. The feasible layer parameters are then given by

$$\mathcal{X}_1 = \left\{\boldsymbol{x}_1 \in \mathbb{R}^4_+ : \delta_1 + \delta_2 \leq P_\text{tr}\right\}.$$

The layer function of layer 1 is chosen as

$$\boldsymbol{f}_1(\boldsymbol{x}_1) = \big(\varepsilon_1(\delta_1, R_1), -R_1, \varepsilon_2(\delta_2, R_2), -R_2\big).$$

At layer 2, we have

$$\boldsymbol{x}_2 = (R_{\text{Q},1}, R_{\text{Q},2}) \quad \text{and} \quad \mathcal{X}_2 = \mathbb{R}^2_{0,+}.$$

As the layer function for layer 2 we choose the weighted sum of the distortions D_k:

$$f_2(\boldsymbol{x}_2, \boldsymbol{d}_1) = \sum_{k=1}^{2} \alpha_k D_k(R_{\text{Q},k}, \varepsilon_k, R_k),$$

with $\alpha_k \geq 0$ and $\alpha_1 + \alpha_2 = 1$. The weights α_k are included to model different application profiles, e.g., if some users have higher priority than others.

4. Parameter optimality

Optimality is determined at the application layer, i.e., an application layer criterion is used to measure the cost of a parameter setup. We introduce the following notation:

$$\boldsymbol{x}_{k:\ell} = (\boldsymbol{x}_k, \ldots, \boldsymbol{x}_\ell), \quad \mathcal{X}_{k:\ell} = \mathcal{X}_k \times \ldots \times \mathcal{X}_\ell.$$

Let $\boldsymbol{f}_{k:\ell}, k < \ell$ denote the concatenation of the layer functions $\boldsymbol{f}_k, \ldots, \boldsymbol{f}_\ell$ resulting from substitution of $\boldsymbol{d}_n$ by $\boldsymbol{f}_n(\boldsymbol{x}_n, \boldsymbol{d}_{n-1})$ for $n = \ell - 1, \ell - 2, \ldots, k$. The cost of a parameter setup $\boldsymbol{x}_{1:N}$ is given by

$$c\left(\boldsymbol{x}_{1:N}\right) = f_{1:N}\left(\boldsymbol{x}_{1:N}\right).$$

Assuming that all feasible parameter sets $\mathcal{X}_n$ and layer functions $\boldsymbol{f}_n$ are globally known, we can directly write down the optimization problem that defines the optimum parameter setup:

$$\tilde{\boldsymbol{x}}_{1:N} = \underset{\boldsymbol{x}_{1:N} \in \mathcal{X}_{1:N}}{\text{argmin}}\ c\left(\boldsymbol{x}_{1:N}\right). \tag{2}$$

The optimization problem formulated in Equation (2) can be interpreted as the joint optimization of the parameters of all layers. In the formulation of Equation (2), properties such as abstraction and modularization provided by a layered architecture are abandoned.

5. Optimization based on layer descriptions

In order to achieve a close to optimum parameter setup while preserving modularity, we propose a cross-layer optimization based on information exchange between layers. In the proposed concept, neighboring layers exchange information in terms of layer descriptions $\mathcal{D}_n$ and description elements $\boldsymbol{d}_n \in \mathcal{D}_n$.

The parameter setup takes place in two stages: In the first stage, the layer descriptions $\mathcal{D}_1, \ldots, \mathcal{D}_{N-1}$ of the first $N-1$ layers are generated in an iterative fashion. Starting with $n=1$, a description $\mathcal{D}_n$ of each layer is generated by *selecting* elements from $\mathcal{Y}_n$. That is, layer $(n-1)$ passes its description $\mathcal{D}_{n-1}$ to layer n. Layer n computes its layer output set $\mathcal{Y}_n$, generates a description $\mathcal{D}_n$ by selecting elements from $\mathcal{Y}_n$, and finally passes $\mathcal{D}_n$ to layer $(n+1)$, and so on. The selection of elements from a layer output set is discussed in detail in the following sections. After the description $\mathcal{D}_{N-1}$ has been computed, the top layer N can determine the optimum setup of its parameters, given description $\mathcal{D}_{N-1}$, as well as the optimum element from $\mathcal{D}_{N-1}$:

$$\left(\hat{\boldsymbol{x}}_N, \hat{\boldsymbol{d}}_{N-1}\right) = \underset{\boldsymbol{x}_N \in \mathcal{X}_N, \boldsymbol{d}_{N-1} \in \mathcal{D}_{N-1}}{\operatorname{argmin}} f_N(\boldsymbol{x}_N, \boldsymbol{d}_{N-1}). \quad (3)$$

At the beginning of the second stage, the optimum element $\hat{\boldsymbol{d}}_{N-1}$ from the description $\mathcal{D}_{N-1}$ is passed down to layer $(N-1)$. Note that $\hat{\boldsymbol{d}}_{N-1}$ corresponds to all parameter setups $\hat{\boldsymbol{x}}_{N-1} \in \mathcal{X}_{N-1}$ and description elements $\hat{\boldsymbol{d}}_{N-2} \in \mathcal{D}_{N-2}$ such that

$$f_{N-1}(\hat{\boldsymbol{x}}_{N-1}, \hat{\boldsymbol{d}}_{N-2}) = \hat{\boldsymbol{d}}_{N-1}. \quad (4)$$

All pairs $(\hat{\boldsymbol{x}}_{N-1}, \hat{\boldsymbol{d}}_{N-2})$ that fulfill Equation (4) are equivalent, as they lead to the same minimum cost at layer N. One pair is chosen, yielding the parameter setup $\hat{\boldsymbol{x}}_{N-1}$ of layer $(N-1)$. The corresponding description element $\hat{\boldsymbol{d}}_{N-2}$ is passed down to layer $(N-2)$.

In the same fashion, from layer $(N-2)$ down to layer 1, after receiving $\hat{\boldsymbol{d}}_n$ from layer $(n+1)$, layer n chooses parameters $\hat{\boldsymbol{x}}_n \in \mathcal{X}_n$ and a description element $\hat{\boldsymbol{d}}_{n-1} \in \mathcal{D}_{n-1}$ such that

$$f_n(\hat{\boldsymbol{x}}_n, \hat{\boldsymbol{d}}_{n-1}) = \hat{\boldsymbol{d}}_n$$

holds and then passes $\hat{\boldsymbol{d}}_{n-1}$ down to layer $(n-1)$. When layer 1 is reached, the parameters of all layers are set up.

6. Optimum descriptions

Let $\tilde{c}$ denote the minimum cost achievable by *joint optimization* of all layer parameters according to Section 4. Let $\hat{c}(\mathcal{D}_n)$ denote the minimum cost achievable by joint optimization of the parameters of layers $n+1, \ldots, N$, given a description $\mathcal{D}_n$ of the nth layer:

$$\hat{c}(\mathcal{D}_n) = \min_{\substack{\boldsymbol{x}_{n+1:N} \in \mathcal{X}_{n+1:N}, \\ \boldsymbol{d}_n \in \mathcal{D}_n}} \boldsymbol{f}_{n+1:N}(\boldsymbol{x}_{n+1:N}, \boldsymbol{d}_n).$$

For a description $\mathcal{D}_n$ to be optimum, we require

$$\tilde{c} = \hat{c}(\mathcal{D}_n). \quad (5)$$

Recall that in order to achieve modularity, $\mathcal{D}_n$ is generated independent of the properties of the upper layers. In other words, we require Eq. (5) to hold for all possible $\mathcal{X}_{n+1:N}$ and $\boldsymbol{f}_{n+1:N}$.

Due to practical reasons, we also desire descriptions that contain the smallest possible number of elements. This property is expressed by requiring that for a descriptions to be optimum, the removal of a single element from $\mathcal{D}_n$ implies that Eq. 5 does no longer hold for all possible $\mathcal{X}_{n+1:N}$ and $\boldsymbol{f}_{n+1:N}$.

In [5] it is argued that it is necessary to constrain the set of all possible $\boldsymbol{f}_{n+1:N}$. In particular, it is shown that optimization based on layer descriptions requires the layer functions $\boldsymbol{f}_n(\boldsymbol{x}_n, \boldsymbol{d}_n)$ to be *monotone* (or order-preserving) in the variable $\boldsymbol{d}_n$. Recall that a layer function expresses a cost of an operation point. Monotonicity basically states that a parameter setup $\boldsymbol{x}_n$ that yields a higher cost at layer n than a parameter setup $\boldsymbol{x}'_n$ cannot lead to a lower cost at a higher layer. If performance measures are properly chosen, the restriction to monotone layer functions therefore does not represent a limitation in real systems.

As shown in [5], under the condition of monotone layer functions, the optimum description of the nth layer is given by the efficient set of the *multiobjective optimization* (MOO) problem

$$\min_{\boldsymbol{x}_n \in \mathcal{X}_n, \boldsymbol{d}_{n-1} \in \mathcal{D}_{n-1}} \boldsymbol{f}_n(\boldsymbol{x}_n, \boldsymbol{d}_{n-1}). \quad (6)$$

The cost function in Eq. (6) is, in general, vector-valued (hence multiobjective optimization). The partial order

$$\boldsymbol{y} \preceq \boldsymbol{y}' \Leftrightarrow y_k \leq y'_k, \forall k$$

is used to compare elements of $\mathcal{Y}_n$. Due to the fact that the employed order is not total, the MOO problem generally has a set of equivalent solutions (see, e.g., [6]). These solutions constitute the efficient set. The set $\mathcal{Y}_n$ can be interpreted as an achievable cost region. The efficient elements are those elements on the boundary of the achievable region for which one component can only be further decreased by increasing at least on other component. The parameter choices $\boldsymbol{x}_n$ that lead to efficient elements $\boldsymbol{d}_n$ are denoted as *Pareto optimal*.

In our example system, the set of Pareto optimal parameters $\boldsymbol{x}_1$ at layer 1 is given by [5]:

$$\mathcal{P}_1 = \left\{\boldsymbol{x}_1 \in \mathbb{R}_+^4 : \delta_1 + \delta_2 = P_{\text{tr}}\right\}.$$

Thus, all parameter configurations that are optimal from the perspective of layer 1 fully exploit the transmit power budget. Notably, based on the knowledge available at layer 1, all combinations of non-negative rates are Pareto optimal. Note that the corresponding optimum description

$$\mathcal{D}_1 = \boldsymbol{f}_1(\mathcal{P}_1)$$

contains an infinite number of elements and is not compact. This issue is addressed in the following section.

7. Approximate descriptions

For monotone layer functions, optimum descriptions (as defined in Section 6) correspond to efficient sets. However, a description in terms of efficient sets may not be feasible, for example if it is not possible to find a parametrized description of the efficient set or, in case finite sets are to be used for information exchange between layers, if the number of elements in the efficient set is simply too large.

In this section, we consider *approximate* descriptions $\mathcal{D}_n$ generated by selecting a finite number of elements from an efficient set $\mathcal{E}_n$. Approximate descriptions do not guarantee an optimum parameter setup. However, it is possible to design approximate descriptions such that the probability of a sub-optimum parameter setup is minimized for a given number $N_{\mathrm{d},n}$ of elements in $\mathcal{D}_n$.

From the optimality of efficient sets we can conclude that approximate descriptions should be obtained by sampling the efficient set. If monotonicity is the only condition imposed on the properties of the layer functions, each element in $\mathcal{E}_n$ is equally likely to yield the optimum parameter setup. As a result, a good sampling should yield an approximately equidistant distribution of samples.

It may happen that a uniform distribution of the $N_{\mathrm{d},n}$ samples on the efficient set is not possible due to the properties of the efficient set. In the example system, the efficient set $\mathcal{E}_1$ cannot be equidistantly sampled with a finite number of samples, due to the fact that $\mathcal{E}_1$ is not compact (rates R_k can take any positive value). This problem can be solved by providing additional knowledge about the upper layers that allows for choosing a compact subset $\hat{\mathcal{E}}_n \subset \mathcal{E}_n$. The compact set $\hat{\mathcal{E}}_n$ can be properly sampled with a finite number of elements. The set $\hat{\mathcal{E}}_n$ can be interpreted as a coarse estimate ("ball park") of the region in which the optimum operating point lies.

Depending on the properties of the layer function, in some cases it is much simpler to determine the set of Pareto optimal parameters $\mathcal{P}_n$. If the Pareto set is known, elements of the efficient set can be computed by collecting $N_{\mathrm{d},n}$ elements from the Pareto set in a sampling set $\mathcal{S}_n \subset \mathcal{P}_n$ and then evaluating the layer function for each element in the sampling set:

$$\mathcal{D}_n = f_n(\mathcal{S}_n). \tag{7}$$

Ideally, elements from $\mathcal{P}_n$ are chosen such that the samples in $\mathcal{D}_n$ are distributed in the desired manner. In [7], a first-order Taylor approximation of the layer function is employed to achieve a nearly equidistant distribution of samples.

In our example, we choose a particularly simple sampling scheme: Let $N_{\mathrm{d},1} = S^3$. A sampling set is generated by

$$\mathcal{S}_1 = \mathcal{S}_\delta \times \mathcal{S}_R, \tag{8}$$

with

$$\mathcal{S}_\delta = \left\{ (\delta_1, P_{\mathrm{tr}} - \delta_1) : \delta_1 = \frac{P_{\mathrm{tr}}}{S+1} s, s = 1, \dots, S \right\}, \tag{9}$$

$$\mathcal{S}_R = \{ (R_1, R_2) : R_k = As + B, s = 1, \dots, S \}, \tag{10}$$

where

$$A = \frac{2}{S-1}, \quad \text{and} \quad B = \frac{S-3}{S-1}.$$

It is easily verified that this corresponds to choosing

$$\hat{\mathcal{E}}_1 = \{ \boldsymbol{d}_1 \in \mathcal{E}_1 : 1 \le R_k \le 3 \}.$$

8. Simulation results

In this section, the performance of the proposed description-based cross-layer design for different description sizes is compared with the performance of two strategies that are based on layer separate optimization.

The average gains of the users' channels are chosen as $\rho_1 = 1$ and $\rho_2 = 2$. The noise powers are equal, i.e., $\sigma^2_{\mathrm{n},1} = \sigma^2_{\mathrm{n},2}$. We investigate three different sizes of the description $\mathcal{D}_1$:

$$N_{\mathrm{d},1} \in \{S^3 : S = 3, 5, 11\}.$$

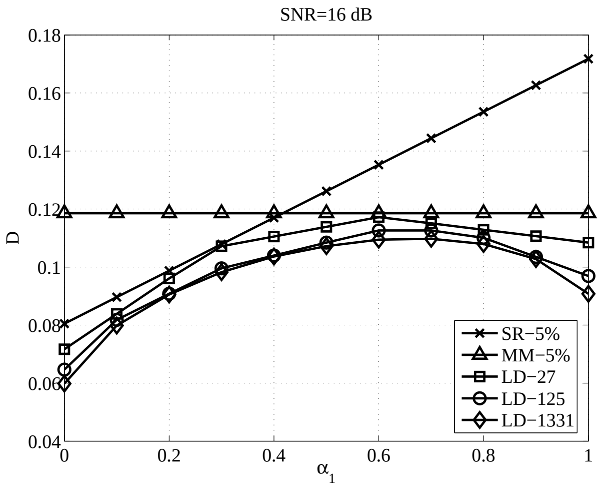

Figure 1: Overall distortion for different application profiles

For a given description size, the approximate description $\mathcal{D}_1$ is generated according to Eqs. (7)-(10).

Two approaches for choosing the parameters of layer 1 without information exchange between layers are considered. Both strategies fix the outage probability of both users to a target outage probability $\bar{\varepsilon}$. In the *sum-rate* (SR) strategy, the parameters of layer 1 are chosen such that the sum of the users' rates is maximized, under the constraint that the outage probability of each user equals $\bar{\varepsilon}$:

$$\boldsymbol{x}_{1,\mathrm{SR}} = \operatorname*{argmax}_{\boldsymbol{x}_1 \in \mathcal{X}_1} R_1 + R_2 \quad \text{s.t.} \quad \varepsilon_k = \bar{\varepsilon}.$$

In the second approach, denoted as *max-min* (MM) strategy, the parameters of layer 1 are chosen such that the smallest among the users' rates is maximized, under the constraint that the outage probability of each user equals $\bar{\varepsilon}$. The max-min strategy follows from the sum-rate strategy by adding the constraint that all rates R_k are equal.

Fig. 1 shows the overall distortion $D = f_2(\hat{\boldsymbol{x}}_2, \hat{\boldsymbol{d}}_1)$ for different weighting factors α_1 and a transmit SNR of 16 dB. The transmit SNR is defined as

$$\mathrm{SNR} = 10 \log_{10} \frac{P_{\mathrm{tr}}}{\sigma^2_{\mathrm{n},1} + \sigma^2_{\mathrm{n},2}}.$$

Depicted are the performance of a cross-layer optimization based on layer descriptions with $N_{\mathrm{d},1} = 27$ (LD-27), $N_{\mathrm{d},1} = 125$ (LD-125), and $N_{\mathrm{d},1} = 11^3$ (LD-1331) elements in $\mathcal{D}_1$, as well as the performance of the two layer-separate optimizations according to the sum-rate strategy with target outage probability $\bar{\varepsilon} = 5\%$ (SR-5%) and the max-min strategy, also with target outage probability $\bar{\varepsilon} = 5\%$ (MM-5%). The description-based optimization of layer parameters can adapt to the different application profiles and provide close to optimum performance, even for small description sizes. In the scenario under consideration, a description size of $N_{\mathrm{d},1} = 125$ is sufficient to well-approximate the efficient set. In contrast to description-based optimization, the parameter setups found based on optimizing layer 1 criteria only do not take into account the properties of the application. This can lead to severe performance degradations. For example, the sum-rate strategy allocates more power to the user with the stronger channel, which, in our scenario, is user 2. The higher the priority of the user with the weaker channel, the

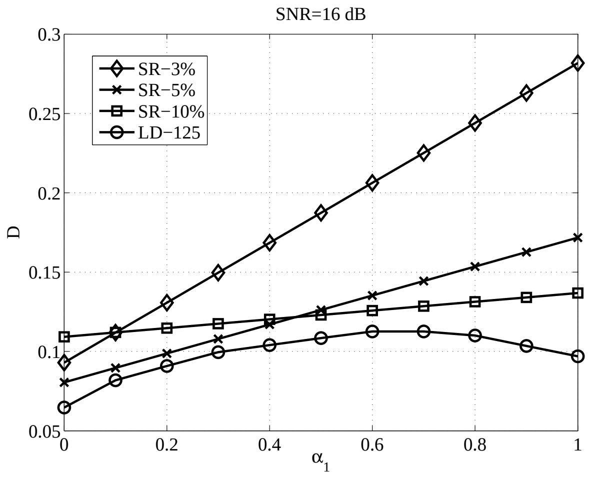

Figure 2: LD-based optimization vs. sum-rate strategy

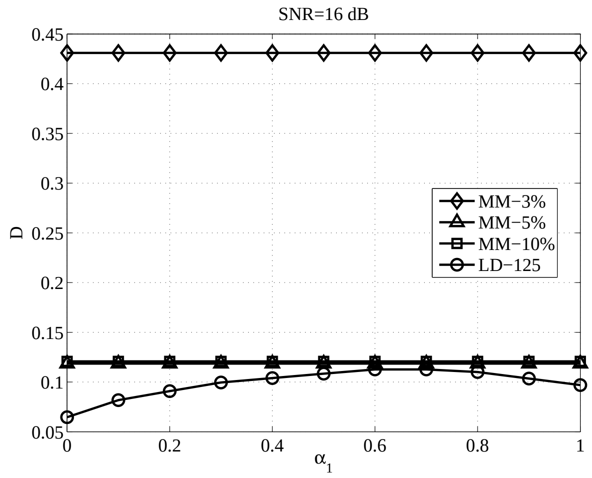

Figure 3: LD-based optimization vs. max-min strategy

larger the performance loss induced by the sum-rate strategy. The max-min strategy provides equal QoS to both users. Again, if the optimization strategy at layer 1 contradicts the user priorities at the application layer, a significant performance degradation is inevitable. In case of the max-min-strategy, this situation occurs if the priority of one user is much larger than that of the other one.

In Fig. 2 optimization based on layer descriptions with $N_{\text{d},1} = 125$ (LD-125) is compared with the sum-rate strategy. The performance of a parameter setup at layer 1 based on the sum-rate strategy is evaluated for target outage probabilities $\bar{\varepsilon} = 3\%$ (SR-3%), $\bar{\varepsilon} = 5\%$ (SR-5%), and $\bar{\varepsilon} = 10\%$ (SR-10%). The curves in Fig. 2 emphasize the impact of the choice of a target outage probability on the performance of the sum-rate strategy. It can be observed that the performance of the sum-rate strategy is highly sensitive to the choice of $\bar{\varepsilon}$. In the case of transmission of Gaussian sources, $\bar{\varepsilon} = 3\%$ represents a poor choice. However, it may be a good choice for a different application. As a result, choosing physical layer parameters without knowledge about the upper layers can be considered a game of chance.

In Fig. 3 optimization based on layer descriptions with $N_{\text{d},1} = 125$ (LD-125) is compared with the max-min strategy. The performance of the max-min strategy is shown for target outage probabilities $\bar{\varepsilon} = 3\%$ (SR-3%), $\bar{\varepsilon} = 5\%$ (SR-5%), and $\bar{\varepsilon} = 10\%$ (SR-10%). Similarly to the sum-rate strategy, a high sensitivity to the choice of a target outage probability can be observed. For a target outage probability $\bar{\varepsilon} = 3\%$, the max-min strategy fails completely, while the choice $\bar{\varepsilon} = 5\%$ provides the best performance among the three choices. The same arguments apply concerning the impossibility of a reliable choice of physical layer parameters without knowledge of the application.

9. Conclusions

A concept for modular cross-layer optimization of layer parameters was presented. The layer parameters are optimized according to an optimality criterion defined at the application layer. The proposed concept is based on a "bottom-up" exchange of layer descriptions between layers. By providing a description of the relevant characteristics of a layer to the neighboring upper layer, a modular architecture can be preserved. Under the assumption of monotonicity, multiobjective optimization was shown to play a key role in the description-based optimization of layered communication systems. At each layer, the optimum description is provided by an efficient set. Good approximate descriptions are obtained by sampling the efficient set.

Optimization based on layer descriptions can provide close to optimum performance by adapting the layer parameters to varying application profiles. In contrast, a layer-separate optimization of parameters can lead to a severe performance degradation.

10. References

[1] 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.

[2] L. H. Ozarow, S. Shamai, and A. D. Wyner, "Information Theoretic Considerations for Cellular Mobile Radio," *IEEE Transactions on Vehicular Technology*, vol. 43, no. 2, pp. 359–378, May 1994.

[3] L. U. Choi, M. T. Ivrlac, E. Steinbach, and J. A. Nossek, "Bottom-Up Approach to Cross-Layer Design for Video Transmission Over Wireless Channels," in *Proc. VTC Spring 2005*, May 2005.

[4] K. Stuhlmuller, N. Farber, M. Link, and B. Girod, "Analysis of Video Transmission over Lossy Channels," *IEEE Journal on Selected Areas in Communications*, vol. 18, pp. 1012–1032, June 2000.

[5] J. Brehmer and W. Utschick, "Bottom-Up Optimization of Layered Communication Systems Based on Description-Passing," Tech. Rep. TUM-LNS-TR-05-02, Munich University of Technology, June 2005.

[6] S. Boyd and L. Vandenberghe, *Convex Optimization*, Cambridge University Press, 2004.

[7] J. Brehmer and W. Utschick, "Physical Layer Characterization of MIMO Systems by Means of Multiobjective Optimization," in *Proc. ICASSP 2005*, March 2005, vol. III, pp. 1141–1144.