Rate Prediction and Receding Horizon Power Minimization in Block-Fading Broadcast Channels

Christoph Hellings, Paskal Kiefer, and Wolfgang Utschick

2012 IEEE 13th International Workshop on Signal Processing Advances in Wireless Communications (SPAWC)

Çeşme, Turkey, 17th - 20th June, 2012

Technische Universität München
Associate Institute for Signal Processing
http://www.msv.ei.tum.de

RATE PREDICTION AND RECEDING HORIZON POWER MINIMIZATION IN BLOCK-FADING BROADCAST CHANNELS

Christoph Hellings, Paskal Kiefer, and Wolfgang Utschick

Associate Institute for Signal Processing, Technische Universität München, 80290 Munich, Germany
Telephone: +49 89 289-28516, Fax: +49 89 289-28504, {hellings,paskal.kiefer,utschick}@tum.de

ABSTRACT

We consider the minimization of the average transmit power in a block-fading broadcast channel with time division multiple access under constraints on the average rates of all users. The globally optimal solution of this problem would require noncausal channel knowledge such that all blocks can be optimized jointly in advance, which is of course impossible in a practical system. However, it is possible to predict future channel realizations based on their statistical properties and on the observations of the current and past realizations. Therefore, we study a receding horizon optimization, where future fading blocks are incorporated into the optimization by means of an MMSE channel prediction or by means of a rate prediction method proposed in this paper. While the optimization based on channel prediction does not lead to the desired reduction of the average transmit power, the rate-prediction-based method is able to achieve a notable reduction.

1. INTRODUCTION

In many practical scenarios, the future development of the channel state of time-varying channels can be predicted accurately in the near future based on past and current channel measurements (e.g., [1, 2], and the references therein). A possible benefit of such a prediction is that a delay caused by the necessity of feeding back the channel estimation to the transmitter can be compensated such that the use of outdated channel state information is avoided [2]. In [3], scheduling based on such a prediction of the channels in the upcoming time block was studied for a multicarrier MIMO broadcast channel. The authors distinguished between predictable and non-predictable users, where the channels of the former were predicted with a parametric approach (such as ESPRIT, e.g., [4]) or a MMSE predictor [1] while the rate achievable for the latter by means of omnidirectional transmission was predicted using the a priori knowledge of the statistical parameters of the channels.

In our work, we also apply an MMSE channel predictor as well as a rate prediction, which is either performed based on a priori information or conditioned on former observations. However, instead of predicting the channel for the current block, for which a scheduling decision is to be made, we focus on the exploitation of the prediction of future blocks for scheduling purposes.

Including future blocks in the optimization of the current transmit strategy has already been studied in [5–10]. However, the authors of these works assumed that there exists a mechanism that allows to predict the rate that will be supported by the physical layer for a user in a certain time block in the future, and they concentrated on the pure scheduling problem. By contrast, our work focuses on the physical layer and studies how such a rate prediction can be performed in a way that is beneficial for the considered optimization.

As in [5–9], we incorporate the predicted future rates into the optimization by means of a so-called receding horizon approach, a concept which is adopted from control theory (e.g., [11]). In the context of predictive scheduling, the term receding horizon optimization refers to the following procedure: the current step is optimized jointly with a certain number of future steps, but the strategies obtained for the future steps are discarded as they will be re-optimized later when the respective step becomes the current step. A more detailed description of this concept can be found in Section 3 after the introduction of the system model in Section 2.

In this paper, predictive optimization is applied with the aim of reducing the average transmit power needed to fulfill constraints on the average rates of the users. The motivation behind such an optimization are systems with elastic traffic, where the additional freedom (compared to inelastic traffic) can be exploited to reduce the energy consumption of the transmitter. Note that this kind of optimization implies that we allow a variable transmit power, which is different from the proportional fair scheduling studied in [3,6–8] and the throughput maximization in [5, 9, 10], where the transmit power was fixed. The additional challenge of a variable transmit power is that the prediction of the rate is no longer a certain value, but a function of the power, which itself is an optimization variable. Methods to predict such rate functions will be discussed in Section 5. These expected rate functions can then be plugged into the optimization method discussed in Section 4. The potential of the various prediction methods is evaluated in numerical simulations in Section 6.

Notation: We use $\bullet^{\mathrm{H}}$ for the conjugate transpose of a vector or matrix, $\mathbf{I}_M$ for the identity matrix of size M, $\mathrm{E}[\bullet]$ for the expected value, and $\otimes$ for the Kronecker product.

2. SYSTEM MODEL AND PROBLEM FORMULATION

We consider a time-varying broadcast channel with an M-antenna base station and K single-antenna receivers, where the channel of each user is characterized by a noise power $\sigma^2_{\mathrm{n},k}(t)$ and a complex channel vector $\boldsymbol{h}_k^{\mathrm{H}}(t) \in \mathbb{C}^{1\times M}$. Without loss of generality, we assume that $\sigma^2_{\mathrm{n},k}(t) = 1$ for all k and t since any other value could be treated by introducing an effective channel $\tilde{\boldsymbol{h}}_k^{\mathrm{H}}(t) = \sigma^{-1}_{\mathrm{n},k}(t)\boldsymbol{h}_k^{\mathrm{H}}(t)$.

Throughout the paper, we assume correlated block-fading channels, where the channel coefficients remain constant during intervals of length T, i.e.,

$$\boldsymbol{h}_k^{\mathrm{H}}(t) = \boldsymbol{h}_k^{\mathrm{H}}[n] \text{ for } t \in [(n-1)T, nT[\qquad (1)$$

while the vectors $\boldsymbol{h}_k^{\mathrm{H}}[n]$ in neighboring blocks are correlated. However, the channel coefficients belonging to different antennas or different users are assumed to be independent. Furthermore, we assume the channels to form a stationary random sequence, and we assume that the temporal correlations are the same for all antennas

978-1-4673-0971-4/12/$31.00 ©2012 IEEE

of a user. As a result, the channels have a separable crosscovariance according to the definition in [1], and, assuming the coefficients to be circularly symmetric Gaussian with zero mean and unit variance, the joint distribution of the channel vectors in a group of N blocks $\boldsymbol{x}_k[N] = [\boldsymbol{h}_k^{\mathrm{H}}[1], \ldots, \boldsymbol{h}_k^{\mathrm{H}}[N]]^{\mathrm{H}}$ can be described by

$$\boldsymbol{x}_k[N] \sim \mathcal{CN}(\mathbf{0}, \boldsymbol{C}_k[N] \otimes \mathbf{I}_M). \tag{2}$$

The element of $\boldsymbol{C}_k[N]$ in the mth column and nth row equals the correlation of the channel coefficients in the mth and nth block, i.e., $\mathrm{E}[h_{k,\ell}[m]h_{k,\ell}[n]^*] = [\boldsymbol{C}_k[N]]_{m,n} = c_k[m,n]$ for all ℓ, where $h_{k,\ell}[n]$ is the ℓth component of $\boldsymbol{h}_k[n]$. Note that due to stationarity, $\boldsymbol{C}_k[N]$ is a Toeplitz matrix, and $\boldsymbol{C}_k[N] \otimes \mathbf{I}_M$ is a block Toeplitz matrix.

At some points in this paper, we will make use of the conditional distribution of the channel vector $\boldsymbol{h}_k[i]$ of user k in block i conditioned on the realizations of the vectors in the blocks $1, \ldots, n$, which is a Gaussian distribution described by the conditional mean

$$\begin{aligned}\boldsymbol{\mu}_k[i|n] &= (\boldsymbol{c}_k^{\mathrm{H}}[n,i] \otimes \mathbf{I}_M)\,(\boldsymbol{C}_k[n] \otimes \mathbf{I}_M)^{-1}\,\boldsymbol{x}_k[n] \\ &= \Big((\boldsymbol{c}_k^{\mathrm{H}}[n,i]\boldsymbol{C}_k^{-1}[n]) \otimes \mathbf{I}_M\Big)\,\boldsymbol{x}_k[n] \end{aligned}\tag{3}$$

and the conditional covariance

$$\begin{aligned}\boldsymbol{C}_k[i|n] &= \mathbf{I}_M - (\boldsymbol{c}_k^{\mathrm{H}}[n,i] \otimes \mathbf{I}_M)\,(\boldsymbol{C}_k[n] \otimes \mathbf{I}_M)^{-1}\,(\boldsymbol{c}_k[n,i] \otimes \mathbf{I}_M) \\ &= \underbrace{(1 - \boldsymbol{c}_k^{\mathrm{H}}[n,i]\,\boldsymbol{C}_k^{-1}[n]\,\boldsymbol{c}_k[n,i])}_{\sigma_k^2[i|n]}\,\mathbf{I}_M \end{aligned}\tag{4}$$

with $\boldsymbol{c}_k[n,i] = [c_k[1,i], \ldots, c_k[n,i]]^{\mathrm{T}}$.

The aim of our studies is to minimize the average transmit power subject to the constraint that a certain average per-user rate ρ_k is achieved for each user. To avoid long delays, we enforce the average rate constraints of all users to be fulfilled after a group of N blocks. Afterwards, the same optimization can be performed for the next N blocks. To keep the stochastic expressions simple, such that they can be evaluated analytically, we restrict ourselves to the case of time division multiple access (TDMA), i.e., at each time instant, only one user is served. The method could, however, be extended to the case of spatial multiplexing (e.g., using MMSE or ZF beamforming) by deriving a method to evaluate the expectations in Section 5 for the case of spatial multiplexing and by replacing the inner optimization in Section 4 with a monotonic optimization as in [12] or by using a heuristic optimization method.

Note that we allow each block n to be arbitrarily subdivided into K subintervals where a certain user k is served in each subinterval. Thus, the overall optimization reads as

$$\begin{aligned}&\min_{(L_k[n]\geq 0,\ p_k[n]\geq 0)_{\forall n, \forall k}} \quad \frac{1}{N}\sum_{n=1}^{N}\sum_{k=1}^{K} L_k[n]\,p_k[n] \\ &\text{s.t.}\ \ \frac{1}{N}\sum_{n=1}^{N} L_k[n]\,r_k[n](p_k[n]) \geq \rho_k\ \forall k \\ &\qquad \text{and}\ \ \sum_{k=1}^{K} L_k[n] = 1\ \ \forall n\end{aligned}\tag{5}$$

where $TL_k[n]$ is the length of the subinterval reserved for user k in block n, and $r_k[n](p_k[n])$ is the achievable rate for user k during the corresponding subinterval of block n given by

$$r_k[n](p_k[n]) = \log_2(1 + p_k[n]\,\|\boldsymbol{h}_k[n]\|_2^2) \tag{6}$$

which is achievable with transmit beamforming (e.g., [13]).

3. RECEDING HORIZON OPTIMIZATION

The main issue of the optimization in (5) is that in a practical implementation, the channels $\boldsymbol{h}_k^{\mathrm{H}}[i]$ of blocks $i > n$ are not known when a decision has to be made for the strategy to be applied in block n. On the other hand, in order to not deal with channel estimation and to concentrate on the issues of prediction, we assume that a perfect estimation of the channels in block $i = n$ is available when the final decision for the strategy for block n is made. The same assumption was made, e.g., in [5].

To deal with the lack of knowledge about future channels, we replace the rate $r_k[i](p_k[i])$ for $i > n$ by a prediction $\hat{r}_k[i|n](p_k[i])$ based on the knowledge available in block n. In Section 5, we will discuss various realistic as well as idealized prediction methods, and we will compare their performance in Section 6.

Assuming that a prediction $\hat{r}_k[i|n](p_k[i])$ is available for $i \in \{n+1, \ldots, n+N_h\}$, we can optimize the system in a receding horizon fashion as in [5–9]. To this end, we have to solve the following optimization in order to decide for a strategy for block n:

$$\begin{aligned}&\min_{(L_k[i]\geq 0,\ p_k[i]\geq 0)_{i\in\{n,\ldots,\tilde{n}\}, \forall k}} \quad \sum_{i=n}^{\tilde{n}}\sum_{k=1}^{K} L_k[i]\,p_k[i] \\ &\text{s.t.}\ \ \sum_{i=n}^{\tilde{n}} L_k[i]\,\hat{r}_k[i|n](p_k[i]) \geq \tilde{\rho}_k[n]\ \forall k \\ &\qquad \text{and}\ \ \sum_{k=1}^{K} L_k[i] = 1 \text{ for } i \in \{n, \ldots, \tilde{n}\}\end{aligned}\tag{7}$$

with $\tilde{n} = \min\{n + N_h, N\}$, $\hat{r}_k[n|n](p_k[n]) = r_k[n](p_k[n])$, and $\tilde{\rho}_k[n] = \tilde{n}\rho_k - \sum_{j=1}^{n-1} L_k[j]\,r_k[j]$. Note that the subinterval lengths $L_k[j]$ and the rates $r_k[j]$ for blocks $j < n$ can no longer be changed when the strategy for block n is optimized, i.e., they are constants. On the other hand, the strategies for blocks $i > n$, which are part of the optimization in (7), have to be considered as virtual strategies, which can still be changed later on. They are only a vehicle to optimize the transmission in the current block and not an actual scheduling decision. Due to the assumption of perfect estimation of the channels in the current block, the overall rate constraints are surely fulfilled after the Nth block.

Following the nomenclature of [7,9], we will call N_h the *prediction horizon*. The same quantity was called *look-ahead* in [5], *maximum scheduling range* in [6], and *prediction window* in [8]. The case $N_h = 0$ corresponds to the straightforward approach where only the current block is optimized and no prediction is needed at all.

4. OPTIMIZATION METHOD

As in [12], the optimization in (7) can be solved by the dual decomposition approach

$$\begin{aligned}&\max_{(\lambda_k)_{\forall k}} \quad \min_{(p_k[i]\geq 0)_{i\in\{n,\ldots,\tilde{n}\}, \forall k}} \\ &\qquad \sum_{i=n}^{\tilde{n}}\sum_{k=1}^{K} (p_k[i] - \lambda_k \hat{r}_k[i|n](p_k[i])) + \sum_{k=1}^{K} \lambda_k \tilde{\rho}_k[n] \\ &\text{s.t.}\ \ p_k[i]p_\ell[i] = 0 \text{ for } \ell \neq k,\ \forall i\end{aligned}\tag{8}$$

where λ_k is the dual variable for the rate constraint of user k. The optimal subinterval lengths $L_k[i]$ can then be obtained from the primal recovery (cf., e.g., [12]), and the constraint in the last line of (8) ensures that only one user is active in each subinterval.

The outer problem can be solved, e.g., using the cutting plane algorithm (e.g., [14]) as was done in [12]. In the following, we will discuss the solution of the inner problem, which can be solved independently for each i.

Due to the constraint $p_k[i]p_\ell[i] = 0$ for $\ell \neq k$, only one user can be active, and we can rewrite the inner problem for block i as

$$\min_{k\in\{1,\dots,K\}} \min_{P\geq 0} P - \lambda_k \hat{r}_k[i|n](P) \tag{9}$$

which is a scalar program in P combined with a linear search for the optimal k. As $\hat{r}_k[i|n](P)$ is concave for the prediction methods presented in Section 5, the optimization over P is a convex program and can be, e.g., solved by means of the interval of uncertainty approach described in [14, Section 8.1].

To extend the method studied in this paper to the case of spatial multiplexing, i.e., to allow simultaneous transmission to multiple users, the inner optimization could be replaced by a monotonic optimization as in [12].

5. PREDICTION METHODS

We will now study various choices for the predicted rate $\hat{r}_k[i|n]$ as a function of the transmit power $p_k[i]$ for $n < i \leq \tilde{n}$. As already stated in Section 4, all prediction functions presented in this section are concave functions of $p_k[i]$. This can be easily verified for the approaches considered in Subsections 5.1 and 5.2 and will be shown for the approaches proposed in Subsections 5.3 and 5.4.

5.1. Genie-Aided Prediction

As a benchmark for the other methods, we will, first of all, introduce a so-called genie-aided prediction, where perfect prediction is assumed for all channels in blocks i within the prediction horizon $i \leq \tilde{n}$ with $\tilde{n} = \min\{n + N_h, N\}$. In this case, we have $\hat{r}_k[i|n](p_k[i]) = r_k[i](p_k[i])$ not only for $i = n$, but for all $i \in \{n, \dots, \tilde{n}\}$.

Note that for $N_h = N - 1$ the receding horizon optimization (7) with the genie-aided predictor becomes equivalent to the optimization in (5). This can be used to compute the ultimate minimum of the average transmit power for TDMA with average rate constraints, which is a benchmark for any realistic scheduler.

5.2. MMSE Channel Prediction

According to [1], the optimal prediction of the channel $\boldsymbol{h}_k[i]$ of user k in block i in the MMSE sense based on information available in block n is given by $\boldsymbol{\mu}_k[i|n]$ from (3). Using this result, a possible prediction of the rate of user k in block i is obtained by plugging this predicted channel into the rate equation (6), which yields

$$\hat{r}_k[i|n](p_k[i]) = \log_2(1 + p_k[i]\,\|\boldsymbol{\mu}_k[i|n]\|_2^2). \tag{10}$$

However, due to the nonlinearity of the rate equation, this does not lead to an accurate prediction of the achievable rate since

$$\mathrm{E}[\log_2(1 + p_k[i]\,\|\boldsymbol{h}_k[i]\|_2^2)] \neq \log_2(1 + p_k[i]\,\|\mathrm{E}[\boldsymbol{h}_k[i]]\|_2^2). \tag{11}$$

As a result, the rate is in most cases underestimated, as can be exemplarily observed for the extreme case of vanishing correlation, where the conditional expectation tends towards the a priori expectation, i.e., $\boldsymbol{\mu}_k[i|n] \to \mathrm{E}[\boldsymbol{h}_k[i]]$. Since $\mathrm{E}[\boldsymbol{h}_k[i]] = \mathbf{0}$, this yields $\hat{r}_k[i|n](p_k[i]) \to 0$ for any $p_k[i]$, which is obviously not a sensible prediction of the achievable rate.

As will turn out in the numerical simulations in Section 6, this kind of prediction is indeed not helpful to reduce the average transmit power in the considered system.

5.3. A Priori Rate Prediction

To overcome the problem observed above, we propose to use the expected value of the rate as rate prediction. We will first discuss the a priori expectation, i.e.,

$$\hat{r}_k[i|n](p_k[i]) = \mathrm{E}\left[\log_2(1 + p_k[i]\,\|\boldsymbol{h}_k[i]\|_2^2)\right] \tag{12}$$

which does not rely on the current and past observations. In the next subsection, we will then extend the method to the conditional expectation based on the knowledge available in block n.

Due to the linearity of the expectation and of the differential operator, we have

$$\begin{aligned}\frac{\partial^2}{\partial P^2}\hat{r}_k[i|n](P) &= \frac{\partial^2}{\partial P^2}\,\mathrm{E}\left[\log_2(1 + P\,\|\boldsymbol{h}_k[i]\|_2^2)\right] \qquad (13)\\ &= \mathrm{E}\Bigg[\underbrace{\frac{\partial^2}{\partial P^2}\log_2(1 + P\,\|\boldsymbol{h}_k[i]\|_2^2)}_{<0}\Bigg] < 0\end{aligned}$$

which shows that the function $\hat{r}_k[i|n](p_k[i])$ is concave.

To evaluate the expected value, we note that the random variable $X = 2\|\boldsymbol{h}_k[i]\|_2^2$ is central χ^2_{2M} distributed since the real and imaginary parts of all components of $\sqrt{2}\boldsymbol{h}_k[i]$ are independent Gaussians with zero mean and unit variance. Consequently, we have

$$\begin{aligned}\hat{r}_k[i|n](P) &= \mathrm{E}\left[\log_2\left(1 + \frac{PX}{2}\right)\right] \qquad (14)\\ &= \int_0^\infty f_X(x)\log_2\left(1 + \frac{Px}{2}\right)\mathrm{d}x\end{aligned}$$

where

$$f_X(x) = \begin{cases}\frac{1}{2^M\Gamma(M)}x^{M-1}\,\mathrm{e}^{-\frac{x}{2}} & \text{if } x \geq 0,\\ 0 & \text{otherwise}\end{cases} \tag{15}$$

is the probability density function of the χ^2_{2M} distribution, and $\Gamma(\bullet)$ is the gamma function [15]. The integral has to be evaluated numerically. For the special case $M = 1$, we get $\hat{r}_k[i|n](P) = -\frac{1}{\ln 2}\,\mathrm{e}^{\frac{1}{P}}\,\mathrm{Ei}(-\frac{1}{P})$ where $\mathrm{Ei}(\bullet)$ is the exponential integral [15].

Note that the a priori rate prediction does not depend on the block index i. Thus, the optimization in (9) does not have to be performed for each block i, but it suffices to perform the optimization once for each given dual variable λ_k and to use the result for all $i \in \{n + 1, \dots, \tilde{n}\}$, which reduces the computational complexity.

5.4. Conditional Rate Prediction

The derivation of the conditional expected rate

$$\hat{r}_k[i|n](p_k[i]) = \mathrm{E}\left[\log_2(1 + p_k[i]\,\|\boldsymbol{h}_k[i]\|_2^2)\,\middle|\,n\right] \tag{16}$$

where $\mathrm{E}[\bullet\,|\,n]$ is used to denote conditioning on all knowledge available in block n, is more involved since the conditional distribution of the channel vector $\boldsymbol{h}_k[i]$ has nonzero mean [cf. (3)] so that the conditional distribution of $2\|\boldsymbol{h}_k[i]\|_2^2$ is not a centered χ^2_{2M} distribution.

Instead of deriving the expression for the expected value, we show that the desired expectation is equivalent to a special case of the expression for the mean of the mutual information of MIMO

Rician channels derived in [16]. Just like the conditional distribution under consideration, the channels of the Rician model used in [16] have nonzero mean and a scaled identity as covariance matrix. The difference is, however, that due to the assumption of lacking channel state information at the transmitter, the results from [16] are based on the assumption that the transmit covariance matrix is a scaled identity while we have assumed that the base station obtains perfect channel state information before the actual transmission is started and can, thus, perform transmit beamforming.

To overcome this model mismatch, we switch to a dual uplink formulation, where the user transmits data to the base station with the same power and rate as in the downlink. Due to the assumption of single-antenna user terminals, the transmit covariance is then a scalar, which is nothing but a special case of the scaled identity matrix from [16].

Thus, (16) can be evaluated using the expression for the mean of the mutual information in [16], by setting the number of transmit antennas to one, the number of receive antennas to M, the mean of the channel vector to $\boldsymbol{\mu}_k[i|n]$ from (3), and the covariance matrix to $\boldsymbol{C}_k[i|n] = \sigma_k^2[i|n]\,\mathbf{I}_M$ from (4):

$$\hat{r}_k[i|n](P) = \frac{\mathrm{e}^{-\gamma_k[i|n]}}{\Gamma(M)} \cdot \int_0^\infty x^{M-1} \log_2(1 + P\sigma_k^2[i|n]\,x) e^{-x}{}_0F_1(M, x\gamma_k[i|n])\mathrm{d}x \tag{17}$$

with $\gamma_k[i|n] = \sigma_k^{-2}[i|n]\,\|\boldsymbol{\mu}_k[i|n]\|_2^2$. The integral has to be evaluated numerically again, and ${}_0F_1(\bullet,\bullet)$ is a generalized hypergeometric function [15]. Note that the method from Subsection 5.3 is equivalent to plugging zero mean and unit variance into (17), i.e., setting $\gamma_k[i|n] = 0$ and $\sigma_k^2[i|n] = 1$.

To show concavity of $\hat{r}_k[i|n](p_k[i])$ from (16), the same reasoning as in (13) can be applied. Thus, (9) is again a convex optimization problem. However, unlike for the a priori rate prediction, the optimization now has to be explicitly performed for each $i \in \{n+1, \ldots, \tilde{n}\}$ since the prediction now depends on the index i.

5.5. Extension to Spatial Multiplexing

In the case of spatial multiplexing, the methods from this paper could be applied to optimize the system in the dual uplink if a rate prediction $\hat{r}_k[i|n](q_1[i], \ldots, q_K[i])$ as a function of the uplink powers $q_1[i], \ldots, q_K[i]$ is available.

The predictions discussed in Subsections 5.1 and 5.2 can be easily extended to this case by replacing (6) by an appropriate rate equation. However, it is not obvious how to efficiently evaluate the expectations from Subsections 5.3 and 5.4 in the case of spatial multiplexing. One way to apply the methods from this paper would be to numerically evaluate the expectations by sampling channel realizations from the a priori distribution or from the conditional distribution, respectively.

However, especially in combination with the optimization from Section 4, which would also become more involved for the spatial multiplexing case, such a sampling would lead to a very high computational complexity. Therefore, finding easy ways to compute these expectations in the case of spatial multiplexing or finding appropriate approximations would be an interesting question for future research.

6. NUMERICAL RESULTS AND DISCUSSION

To evaluate the potential of the prediction methods discussed in the last section, we have performed numerical simulations with a so-called AR1 correlation model (e.g., [17]) with $c_k[i,j] = \mathrm{e}^{-\alpha|i-j|}$ for some value of α and a Jakes correlation model (e.g., [17]) with $c_k[i,j] = J_0(2\pi\beta|i-j|)$ for some β, where $J_0(\bullet)$ is the Bessel function of first kind and order zero [15]. All simulation results are averaged over 200 realizations in the dB domain.

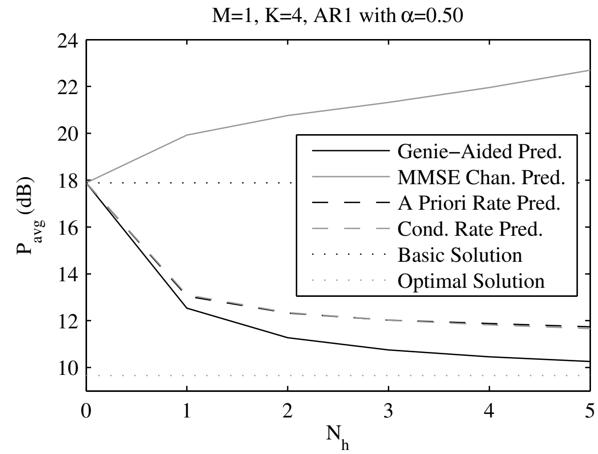

Fig. 1. Average power achieved by receding horizon optimization.

For the simulations in Fig. 1, we have considered a system with $K = 4$ users with average rate requirement $\rho_k = 1$ bit for all k and a total number of $N = 20$ time blocks. For simplicity, we have considered a single-antenna system with $M = 1$. For the channels, we have used the AR1 correlation model with $\alpha = 0.5$. The two dotted lines in the plot correspond to the straightforward approach without prediction ($N_h = 0$) and to the global optimum with perfect noncausal channel knowledge (genie-aided prediction with $N_h = 19$), respectively. It can be seen that optimization with the genie-aided predictor leads to a significant decrease of the average transmit power even for short prediction horizons N_h, which suggests that receding horizon optimization can be beneficial in block-fading channels. However, when replacing the unrealistic perfect prediction by the MMSE channel prediction discussed in Subsection 5.2, the average transmit power is increased instead of decreased. The reason for this is that plugging a channel prediction into the nonlinear rate equation does not deliver an accurate prediction of the expected rate as discussed in Section 5.2. As the future rate is mostly underestimated, too much power is invested in early blocks. This shows that the predictive optimization is very sensitive to the quality of the prediction. On the other hand, it can be seen that the a priori rate prediction and the conditional rate prediction are precise enough to achieve a significant fraction of the possible gain even for short prediction horizons. Due to the weak correlation of the channels, conditional rate prediction does not have an advantage compared to a priori rate prediction.

In Fig. 2, we have considered the same system, but we have used stronger correlations with $\alpha = 0.05$. For higher correlations, knowledge of the current realizations delivers more information about the future ones. Therefore, we can now observe a difference between a priori rate prediction and conditional rate prediction, and the disadvantage of the optimization based on channel prediction is less pronounced (though still observable). However, for some channel realizations, it now happens that the a priori rate prediction (and in very few cases also the conditional rate prediction) leads to a transmit power much higher than the basic solution (note that the plot only shows the average performance). The reason is that in case of a bad channel in a certain block n, the algorithm might decide to schedule low rates in that block and high rates in the future since the

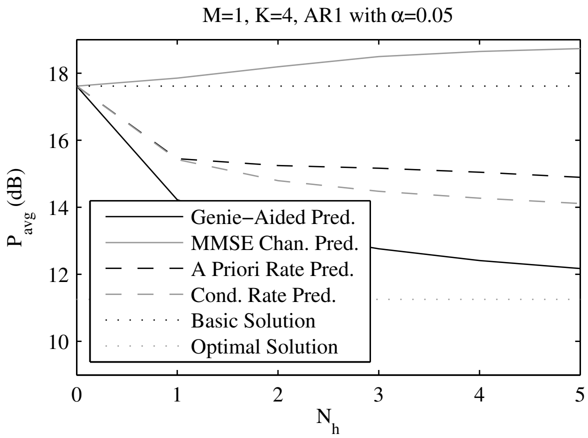

Fig. 2. Average power achieved by receding horizon optimization.

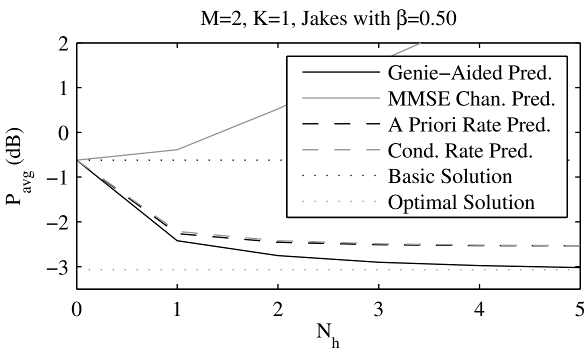

Fig. 3. Average power achieved by receding horizon optimization.

prediction (especially the a priori rate prediction) suggests that the channel will become better in the future. However, as channels with high temporal correlations change slowly, it might happen that this improvement does not occur before the limit of N blocks is reached. In this case, a very high power is needed in the Nth block. Therefore, for a practical implementation, it might be sensible to loosen the constraint that the average rates have to match the given minimum exactly after N steps whenever such a situation occurs.

Simulation results for a single-user multiantenna system with $M = 2$ transmit antennas, $N = 10$ blocks, average rate requirement $\rho_1 = 1$ bit, and a Jakes correlation model with weak correlations specified by $\beta = 0.5$ are shown in Fig. 3. The results are qualitatively the same as in the multiuser system with AR1 channel model.

7. CONCLUSION

We have demonstrated that a notable reduction of the transmit power can be achieved in a block-fading broadcast channel with elastic traffic if a prediction of the future rates is exploited for the optimization of the current block in a receding horizon style. However, we have also observed that imperfect predictions might lead to a worsening instead of an improvement. In particular, predicting the rate based on MMSE channel predictions turned out to not be a sensible strategy. The high potential of predictive optimization revealed in this paper by also considering an idealized perfect estimation motivates research on finding prediction methods that come close the idealized solution in the considered system as well as under different system assumption such as, e.g., spatial multiplexing.

8. REFERENCES

[1] F. A. Dietrich, *Robust signal processing for wireless communications*, vol. 2 of *Foundations in Signal Processing, Communications and Networking*, Springer, Berlin, Germany, 2008.

[2] A. Duel-Hallen, "Fading Channel Prediction for Mobile Radio Adaptive Transmission Systems," *Proc. IEEE*, vol. 95, no. 12, pp. 2299–2313, Dec. 2007.

[3] H. Shirani-Mehr, D. N Liu, and G. Caire, "Channel state prediction, feedback and scheduling for a multiuser MIMO-OFDM downlink," in *Conf. Record 42nd Asilomar Conf. Signals, Syst., Comput. (ACSSC)*, Oct. 2008, pp. 136–140.

[4] M. C. Vanderveen, A.-J. v. d. Veen, and A. Paulraj, "Estimation of multipath parameters in wireless communications," *IEEE Trans. Signal Process.*, vol. 46, no. 3, pp. 682–690, Mar. 1998.

[5] G. Sahin, F. Jin, A. Arora, and H.-A. Choi, "Predictive scheduling in multi-carrier wireless networks with link adaptation," in *Proc. Veh. Technol. Conf. (VTC) 2004-Fall*, Sept. 2004, vol. 7, pp. 5015–5020.

[6] T. Min and C. G. Kang, "Multi-Phase predictive proportional fairness scheduling for a Multi-Channel wireless packet system," in *Proc. IEEE 18th Int. Symp. Personal, Indoor, Mobile Radio Commun. (PIMRC)*, Athens, Greece, Sept. 3–7, 2007.

[7] H. J Bang, T. Ekman, and D. Gesbert, "Channel predictive proportional fair scheduling," *IEEE Trans. Wireless Commun.*, vol. 7, no. 2, pp. 482–487, Feb. 2008.

[8] J. Hajipour and V. C. M. Leung, "Proportional Fair Scheduling in Multi-Carrier Networks Using Channel Predictions," in *Proc. Int. Conf. Commun. (ICC)*, Cape Town, South Africa, May 23–27, 2010.

[9] J. F. Schmidt, J. E. Cousseau, R. Wichman, and S. Werner, "Prediction based resource allocation in OFDMA," in *Proc. 45th Annual Conf. Inf. Sci. Syst. (CISS)*, Baltimore, MD, USA, Mar. 23–25, 2011.

[10] R. Pitic and A. Capone, "An Opportunistic Scheduling Scheme with Minimum Data-Rate Guarantees for OFDMA," in *Proc. Wireless Commun. Netw. Conf. (WCNC)*, Mar./Apr. 2008, pp. 1716–1721.

[11] W. H. Kwon and S. Han, *Receding horizon control: model predictive control for state models*, Springer, London, UK, 2005.

[12] C. Hellings, M. Joham, M. Riemensberger, and W. Utschick, "Minimal Transmit Power in Parallel Vector Broadcast Channels with Linear Precoding," *IEEE Trans. Signal Process.*, vol. 60, no. 4, pp. 1890–1898, Apr. 2012.

[13] D. Tse and P. Viswanath, *Fundamentals of Wireless Communication*, Cambridge University Press, Cambridge, UK, 2005.

[14] M. S. Bazaraa, H. D. Sherali, and C. M. Shetty, *Nonlinear Programming: Theory and Algorithms*, Wiley-Interscience, Hoboken, NJ, USA, 3rd edition, 2006.

[15] I. S. Gradshteyn and I. M. Ryzhik, *Table of Integrals, Series, and Products*, Elsevier, 7th edition, 2007.

[16] M. Kang and M.-S. Alouini, "Capacity of MIMO Rician Channels," *IEEE Trans. Wireless Commun.*, vol. 5, no. 1, pp. 112–122, Jan. 2006.

[17] B. C. Banister and J. R. Zeidler, "Feedback Assisted Stochastic Gradient Adaptation of Multiantenna Transmission," *IEEE Trans. Wireless Commun.*, vol. 4, no. 3, pp. 1121–1135, May 2005.