# Measurement-Based Evaluation of Robust Linear Precoding for Downlink CoMP

Rikke Apelfröjd, Mikael Sternad and Daniel Aronsson
Signals and Systems, Uppsala University
Box 534, 751 21 Uppsala, Sweden, {riab,ms}@signal.uu.se, daniel.at.aronsson@gmail.com

***Abstract*—We study the design and evaluation of joint processing coordinated multipoint (CoMP) downlink transmission. Precoders will then be designed based on outdated channel state information (CSI), so interference cannot be eliminated completely as by an ideal zero-forcing (ZF) solution. We here strive to design and evaluate realistic linear transmit schemes. Kalman predictors are used for orthogonal frequency-division multiplexing (OFDM) channels. They provide optimal linear predictions and also estimates of their uncertainty. Robust linear precoders are designed based on these uncertainty estimates. We introduce and use robust linear quadratic optimal feedforward control, with the criterion averaged (marginalized) over the CSI uncertainty. This flexible solution performs minimum mean square error (MSE) minimization. It can also iteratively optimize other criteria, such as sum-rate. The prediction- and transmission performance is evaluated using measured data on 20 MHz OFDM downlinks from three base stations, for users at fast pedestrian velocities. Downlink CoMP is here also compared to cellular transmission, that uses orthogonal resources within cells but allows uncontrolled interference between cells.**

## I. Introduction

Future wireless broadband systems need to improve the transmission capacity at low capacity locations. Shadowed areas and interference at cell borders then pose particular challenges. A potentially powerful but challenging tool is to use joint coherent transmission, using remote radio heads or coordination between cellular base station sites [1], [2]. Joint coherent transmission enables spatial interference avoidance. Multiple users can then simultaneously be offered a larger bandwidth per user than with e.g. joint time-frequency scheduling designed for interference avoidance. This makes it an interesting option for highly loaded systems. We here study the design and evaluation of joint processing coordinated multipoint (CoMP) transmissions, focusing on the (more challenging) downlink problem.

Global optimization of a joint coherent transmission strategy with respect to a very large set of widely spaced transmitters is unrealistic, due to a high computational load and very high demands on the backhaul capacity. In addition, if distances between cooperating base stations are very large, receivers would have trouble with symbol-level synchronization due to transmission delays longer than the cyclic prefix. However, sets of $N$ transmitters could form coordinated clusters of limited size, aiming to suppress the intracluster interference when jointly transmitting to $M$ users. Such clusters may be designed to also limit intercluster interference [3].

Ideally, the intracluster interference can then be eliminated when $N \geq M$. However, due to signaling delays, the precoders (beamformers) will be designed based on outdated channel state information (CSI), which limits the performance. We here use imperfect channel estimates and strive to design and evaluate the best realistic linear precoding schemes.

Kalman predictors are used for OFDM channels. They provide optimal linear predictions and also, importantly, estimates of their uncertainty. All $M$ users of an OFDM frequency-division duplexing (FDD) downlink are here assumed to feed back the predicted CSI and the prediction uncertainty via the base stations to a central unit (CU), which calculates the joint precoding matrix. Robust precoders are designed using linear quadratic optimal feedforward control, by optimizing a criterion averaged (marginalized) over the CSI uncertainty. Additional corruption due to e.g. quantization is not explicitly considered here, but can easily be included in the design. This flexible solution can iteratively optimize a desired criterion, such as the sum-rate. After the linear design, a scaling is adjusted to satisfy per-antenna transmit power constraints.

The robust precoder for flat fading OFDM subcarriers is a special case of results for dynamic (frequency-selective) systems, previously developed in [4]. Robust linear precoder design by averaging over CSI uncertainty has more recently been highlighted for multiple-input single-output (MISO) transmit schemes by [5], [6] and for multi-user and multiple-input multiple-output (MIMO) downlinks in [7], [8].

A main contribution of the present work is the evaluation of predictor and robust precoder performance based on broadband channel measurements. We use measured data on 20 MHz OFDM downlinks from three base stations for users at pedestrian velocities. We test schemes in fully loaded systems for groups of 3 transmitters and 3 users, with channels having widely varying powers and realistically attainable prediction errors for pedestrian users. Downlink CoMP is here also compared to cellular transmission, that uses orthogonal resources within cells but allows uncontrolled interference between cells.

## II. Channel model

We assume an OFDM downlink for $K$ subcarriers, $M$ single-antenna users and $N$ transmitters that may represent antennas or fixed beams at different base stations. The received signals, $y^k \in \mathbb{C}^{M\times 1}$, at the $k'th$ subcarrier are given by

$$y^k = H^k u^k + n^k, \tag{1}$$

where $n^k \in \mathbb{C}^{M\times 1}$ is the sum of noise and out-of-cluster interference (we will henceforth call it noise), modeled as

i.i.d. random variables with zero mean and known variance. The vector $u^k \in \mathbb{C}^{N\times 1}$ represents the transmitted signals and $H^k \in \mathbb{C}^{M\times N}$ is the channel matrix with complex channel gains $h_{ij}^k$. The true channel is a sum of the predicted channel matrix $\hat{H}^k$ and the prediction error $\Delta H^k$

$$H^k = \hat{H}^k + \Delta H^k. \tag{2}$$

In the following, $\bar{E}[\cdot]$ averages over the distribution of prediction errors, $E[\cdot]$ averages over the statistics of noise and message symbols, $(\cdot)^*$ denotes the Hermitian transpose and $(\cdot)_{ij}$ denotes element $(i,j)$ of a matrix.

## III. Channel predictions

### A. Short term fading models

The downlink transmissions contain known common reference symbols (CRS), or pilots, with regular time and frequency spacing, $\Delta t$ and $\Delta f$. For each user $i$, the frequency domain channels $h_{ij}^k$, from all transmitters $j = 1,..N$ in the cooperation cluster and all CRS-bearing subcarriers $f_0 + k\Delta f$, are predicted over a prediction horizon $\vartheta\Delta t$ where $\vartheta \in \mathbb{N}$. Here, $f_0$ is the carrier frequency for the first subcarrier in the system.

We consider FDD system downlinks, so predictions must be based on downlink measurements. We will assume predictors to be localized in the user terminals. Each terminal predicts a row of $H^k$, and feeds back results to the CU via uplink control channels. The required prediction horizon $\vartheta\Delta t$ corresponds to the delay that would be allowed for the entire transmission control loop, including channel predictions, feedback, scheduling, joint precoding and any additional delays.

In the following, Kalman predictors will be implemented in the frequency domain. Their design requires statistical models of the correlation properties of the channels over time and frequency. To represent the correlation over time, we use autoregressive (AR) processes of order $n_a$, adjusted to approximate the short-term fading statistics of the channel coefficients. The prediction accuracy is improved by predicting $w > 1$ channels at several correlated subcarriers jointly, since the influence of the noise can then be reduced.

The AR-models of the channels to the $M$ users at $w$ subcarriers can then be realized in state-space form as

$$\begin{aligned} x_i(t+1) &= A_i x_i(t) + B_i e_i(t), \\ h_i(t) &= C_i x_i(t), \qquad\qquad i = 1,...,M. \end{aligned} \tag{3}$$

Here, the integer $t$ represents the time steps spaced by $\Delta t$, $x_i(t) \in \mathbb{C}^{(w\cdot n_a\cdot N)\times 1}$ is the vector of state variables and $e_i(t) \in \mathbb{C}^{(w\cdot N)\times 1}$ is the zero mean process noise. The vector $h_i(t) = \left[h_{i1}^{qw}(t),\cdots,h_{i1}^{(q+1)w-1}(t), h_{i2}^{qw}(t),\cdots,h_{iN}^{(q+1)w-1}(t)\right]^T$ represents the channels to user $i$, from all $N$ transmitters, over a set of $w$ CRS-bearing subcarriers indexed by $q$. The correlation between the fading channels at different subcarriers is represented by the covariance matrix $Q$ of the process noise $e_i(t)$ and by the matrices $A_i$, $B_i$ and $C_i$. Please see the Appendix for additional details on the modeling.

### B. Kalman filter and predictor

For each user, the channels over the whole bandwidth are predicted by a set of Kalman predictors, each of which is using measurements at groups of $w$ CRS bearing subcarriers

$$y_i(t) = \Phi h_i(t) + n_i(t), \quad i = 1,..,M, \tag{4}$$

where the vector $y_i(t) = \left[y_i^{qw}(t),\cdots,y_i^{(q+1)w-1}(t)\right]^T$ contains the received signals at the $w$ subcarriers for the different predictors $q = 0, 1, ....$ The measurement noise $n_i(t) \in \mathbb{C}^{w\times 1}$ is assumed zero mean with known covariance matrix $R_n$. The matrix $\Phi \in \mathbb{C}^{w\times w\cdot N}$ contains only known reference symbols and zeros. Reference symbols may be transmitted at orthogonal time-frequency resources by different transmitters. Alternatively, to reduce the CRS overhead, we may use quasi-orthogonal CRS, "overlapping pilots", as e.g. proposed in [9]. All $N$ transmitters then send CRSs simultaneously. Their CRS at the $w$ subcarriers are designed orthogonal. If $w \geq N$, the users can then fully separate $N$ flat fading channels.

The equations for updating one of the Kalman filters are given by e.g. (4.3.1)-(4.3.8) in [10]. The estimates provided by the filters are then used in the predictors to update the channel prediction vector $\hat{h}_i(t+\vartheta|t)$ at the $i'th$ user through

$$\begin{aligned} \hat{h}_i(t+\vartheta|t) &= C_i\hat{x}_i(t+\vartheta|t) = C_i A_i^{\vartheta}\hat{x}_i(t|t), \\ P_i(t+\vartheta|t) &= A_i P_i(t+\vartheta-1|t)A_i^* + B_i Q B_i^*. \end{aligned} \tag{5}$$

Here, $\hat{x}_i(t_1|t_2)$ and $P_i(t_1|t_2)$ are the estimate of the state vector and of the covariance matrix of the state vector at time $t_1$, given measurements up to time $t_2$.

We will in the following assume that the errors in the channel predictions $\hat{h}_i(t+\vartheta|t)$ have zero means and prediction error covariances described entirely by the matrices $(C_i P_i(t+\vartheta|t) C_i^*)$. This is an approximation since it assumes that the estimated model (3) describes the fading perfectly.

The prediction performance will be evaluated using the normalized mean square prediction error (NMSE) for the channel from the $j'th$ transmitter to the $i'th$ user

$$NMSE_{ij} = \frac{\sum_{t=1}^{T}|h_{ij}^k(t) - \hat{h}_{ij}^k(t|t-\vartheta)|^2}{\sum_{t=1}^{T}|h_{ij}^k(t)|^2}, \tag{6}$$

where $T$ is an appropriate averaging interval. An NMSE of $-x$ dB implies for example that, if the corresponding channel represents an interferer, it can be suppressed by at most $x$ dB.

For simplicity, we will in the following exclude time and subcarrier indices so that $h_{ij} \triangleq h_{ij}^k(t)$, $\hat{h}_{ij} \triangleq \hat{h}_{ij}^k(t|t-\vartheta)$, $n \triangleq n^k$, $u \triangleq u^k$ and $y \triangleq y^k$.

## IV. Precoding schemes

A CU, with full information of the predicted channels and of the covariances of the prediction errors, designs precoding matrices $R \in \mathbb{C}^{N\times M}$ for time-frequency resource blocks that contain at least one CRS-bearing subcarrier. The transmitted vector, $u \in \mathbb{C}^{N\times 1}$, is then generated by linear precoding

$$u = \frac{1}{c}Rs, \tag{7}$$

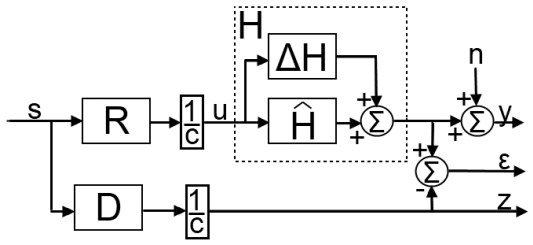


Figure 1. System model used for precoder design.

where $s \in \mathbb{C}^{M\times 1}$ is the message vector, assumed to have zero mean, unit covariance and to be uncorrelated with the noise $n$. We assume per-antenna transmit power constraints, $P_{j,max}$, that apply to all subcarriers individually. The scaling constant $c$ is selected to assure that $E\left[|u_j|^2\right] \leq P_{j,max}$ for $j = 1, ..., N$.

An overview of the system model used for precoder design is shown in Figure 1, where $\varepsilon, z \in \mathbb{C}^{M\times 1}$. The desired received noise-free vector is $z = \frac{1}{c}Ds$. The target matrix $D$ is diagonal, representing the ideal of a complete interference suppression. The targeted received signal magnitudes (the diagonal elements of $D$) should be set to realistic levels. We here do this by adjusting them to the channel amplitude between each user and its strongest base station

$$d_{ii} = \max_j |\hat{h}_{ij}|,\ i = 1, ..., M. \tag{8}$$

### A. Zero forcing

For full rank channels when $M = N$, the ZF precoding matrix, $R_{ZF}$, is designed as

$$R_{ZF} = \hat{H}^{-1}D. \tag{9}$$

The received signal, by (1), (2) and (9),

$$y = \frac{1}{c}(I + \Delta H\hat{H}^{-1})Ds + n, \tag{10}$$

is then corrupted by residual interference due to prediction errors, and by noise. Furthermore, we encounter the phenomenon of "power normalization loss": Ill-conditioned matrices $\hat{H}$ lead to precoders with large elements, resulting in a large scaling factor $c$. This does not affect the SIR but will reduce the power of the signal as compared to that of the noise. The SNR is hence reduced, which might lead to poor performance.

### B. Robust linear precoding

The robust linear precoding (RLP) scheme is based on a robust linear quadratic optimal feedforward control solution. The precoding matrix is designed to minimize the sum of the weighted mean square control error $\varepsilon = Hu - z$ and a transmit power penalty term, averaged over all prediction errors, noise and measurement statistics

$$J = \bar{E}\left[E\left\|V\varepsilon\right\|^2\right] + E\left\|Su\right\|^2. \tag{11}$$

Here, $V$ is a positive definite matrix and $S$ is a positive semidefinite matrix. As a special case of the result in section V of [4], it follows that when $\bar{E}\left[\Delta H^* V^* V \Delta H\right]$ is known and $\bar{E}\left[\Delta H\right] = 0$, the precoding matrix minimizing $J$ is given by

$$R_{RLP} = \left(\hat{H}^* V^* V \hat{H} + S^* S + \overline{E}[\Delta H^* V^* V \Delta H]\right)^{-1} \hat{H}^* V^* V D. \tag{12}$$

The gain constant $c$ is then adjusted, as in the ZF case, to satisfy the per-antenna power constraint. This does not affect the the optimal $R$ w.r.t. (11).

The basic RLP scheme (11), (12) is explicitly designed to take the second order statistics of the channel prediction errors into account. The influence of the prediction errors is quantified as follows. The relation

$$\left(\overline{E}[\Delta H^* V^* V \Delta H]\right)_{jn} = \mathrm{tr}\left(V^* V \bar{E}\left[\Delta H_n \Delta H_j^*\right]\right),$$

follows from the proof of Lemma 1 of [4]. Here, $(\cdot)_j$ denotes column $j$ of a matrix. Since the users predict the channels independently we have $\overline{E}[\Delta h_{ij}^* \Delta h_{mn}] = 0$ when $i \neq m$. Using this, and assuming Kalman predictions (5), we find that $\overline{E}[\Delta H_n \Delta H_j^*]$ is a diagonal matrix where element $(i,i)$ is given by

$$\left(\left(C_i P_i\left(t+\vartheta|t\right) C_i^*\right)_k\right)_{nj}.$$

Here $(\cdot)_k$ denotes the submatrix of $\left(C_i P_i\left(t+\vartheta|t\right) C_i^*\right)$ from (3) and (5) corresponding to the relevant subcarrier $k$.

Other covariance terms may be added to the inverse in (12), describing e.g. the effects of feedback quantization errors.

When $M = N$, $\hat{H}$ is invertible, $V = I$, $S = 0$ and $\Delta H = 0$, then (12) reduces to the ZF precoder (9).

Two out of many possible adjustments of the RLP will be evaluated and compared here.

*1) Minimizing intracluster interference:* With $V = I$ and $S = 0$ in (11), the transmit power is not penalized and the errors at all receivers are considered equally important. This setup minimizes the sum of interference powers.

*2) Iteratively optimizing a different criterion:* The minimization criterion in (11) represents an auxillary criterion. Often what we really wish to optimize is a function

$$f(\bar{E}\left[E\left[P_{S,i}\right]\right], \bar{E}\left[E\left[P_{I,i}\right]\right], \bar{E}\left[E\left[P_{N,i}\right]\right]), \tag{13}$$

where $P_{s,i}$ is the signal power, $P_{I,i}$ the interference power and $P_{N,i}$ the noise power at user $i$. The weighting matrices $V$ and/or $S$ can then be used as design variables, to be adjusted by optimizing (13). For example we may set $V = I$ and set $S$ diagonal, with real-valued adjustable diagonal elements. This yields a $N$ dimensional search as opposed to global optimization of $R$ which corresponds to a $2{\cdot}N{\cdot}M$ dimensional search for all the real and imaginary parts of the elements of $R$. To optimize (13) we can then iteratively adjust the diagonal real-valued elements $\rho_{jj}$ of the penalty matrix

$$S_\rho = diag(\rho_{jj}). \tag{14}$$

In each iteration we calculate the expected values of the powers as

$$\begin{aligned}
\bar{E}\left[E\left[P_{N,i}\right]\right] &= \sigma_{n,i}^2, \\
\bar{E}\left[E\left[P_{I,i}\right]\right] &= \frac{1}{c^2}\sum_{m\neq i}^{M} \bar{E}\left[|HR|_{im}^2\right], \\
\bar{E}\left[E\left[P_{S,i}\right]\right] &= \frac{1}{c^2}\bar{E}\left[|HR|_{ii}^2\right],
\end{aligned} \tag{15}$$

where $\sigma^2_{n,i}$ is the variance of the noise at user $i$ and, assuming $\bar{E}\left[\Delta H\right] = 0$,

$$\bar{E}\left[|HR|^2_{im}\right] = \bar{E}\left[\left|\sum_{j=1}^{N}\hat{H}_{ij}R_{jm} + \sum_{j=1}^{N}\Delta H_{ij}R_{jm}\right|^2\right]$$

$$= \left|\left(\hat{H}R\right)_{im}\right|^2 + \sum_{j=1}^{N}\sum_{n=1}^{N}\left(\left(C_i P_i\left(t+\vartheta|t\right)C_i^*\right)_k\right)_{jn}R_{jm}R^*_{nm}.$$

To limit computational complexity and ensure a more practical scheme we, in this paper, focus on a one dimensional search of $S_\rho$. This is carried out in two steps. First we calculate $R_{RLP}$ as in section IV-B1. Second we find the base station that is given the highest transmit power with this solution, $n_{max}$ and penalize the transmit power of that base station by setting the elements in (14) to

$$\rho_{jj} = \begin{cases} \rho & j = n_{max} \\ 0 & \text{otherwise} \end{cases}. \qquad (16)$$

This reduces the difference in required transmit powers at different transmitters, which reduces the rescaling $c$ needed to satisfy the power constraints. Then the total signal power, and thus the SNR, increases, at the price of a decreasing SIR. The optimal balance with respect to (13) is found in the range of $0 \leq \rho < \rho_{max}$ where $\rho_{max}$ is the smallest $\rho$ that causes a new base station to have highest transmit power.

In this paper we use an estimate of the sum-rate

$$\sum_{i=1}^{M}\log\left(1 + \frac{\bar{E}\left[|HR|^2_{ii}\right]}{\sum_{m\neq i}^{M}\bar{E}\left[|HR|^2_{im}\right] + c^2\sigma^2_{n,i}}\right), \qquad (17)$$

as our criterion (13) to be maximized.

## V. Evaluations Based on Measured Channels

### A. Channel measurements

All simulations are here based on channel sounding measurements carried out by Ericsson Research as described in [11]. Three single antenna base stations, located at different sites with 350-600 m distance, were used to transmit to a measurement van in an outdoor suburban environment. The measurements, which were obtained with high SNR over a 400 s interval along the measurement route, are used as true channels in the following simulations. Figure 2 shows the received signal powers from the base stations. To simulate a lower maximum velocity and to make the model more 3GPP-LTE like, the data has been upsampled $25 \times 3$ times in time×frequency using the fast Fourier transform. The measurement parameters and the parameters for the upsampled system are presented in Table I.

### B. Simulation method and assumptions

Kalman predictors were used, and adjusted based on the channel statistics. The fading statistics in time and frequency, represented by fourth order AR processes, were estimated separately and periodically every $0.5$ s for the downlink

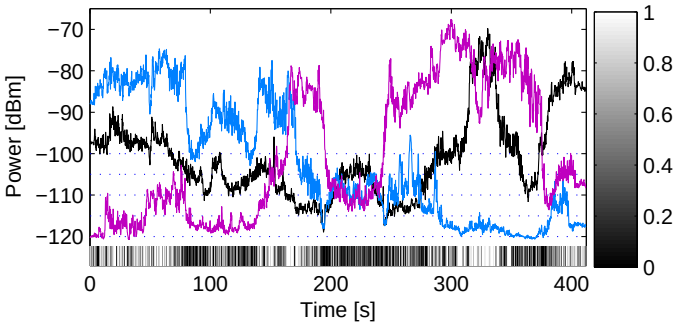


Figure 2. Power of the received signals transmitted from the base stations (full lines), noise floors of $-120$ to $-100$ dB used for simulations (dotted lines) and the result of the simulations in Section V-E (gray-scaled bars).

Table I
Measurement and Simulation Parameters.

| Parameter | Value for measurements | Value for simulations |
|---|---|---|
| Carrier frequency | 2.66GHz | 2.66GHz |
| Subcarrier bandwidth | 45kHz | 15kHz |
| Number of base stations | 3 | 3 |
| Sampling time (time between CRS) | 5.3ms | 0.64ms |
| Maximum speed | 30km/h | 10km/h |
| Prediction horizon | | 5.1ms |

channels from $N = 3$ single antenna base stations. We here based the AR-models on noise-free channel data from the past $0.5$ s. The channel correlation in frequency used to create the matrix $Q$ in (5) was estimated as the sample mean of $h_i^k\left(h_i^{k+f}\right)^*$ for $k = 1, ..K - w + 1$ and $f = 0, ..., w - 1$. Computational complexity increases with $w$, so we have used a low value of $w = 4$ at the cost of larger prediction errors. The prediction errors are reduced in a second step, by Wiener smoothing over estimates from multiple Kalman predictors, $\{q\}$ working on adjacent frequency groups. Sets of signal measurements spanning an appropriate range of SNR values were created through (4) by the use of quasi-orthogonal CRSs and by adding white Gaussian noise of five different power levels, $\sigma_n^2$, see Figure 2. The noise power was set equal for all users, yielding $R_n = \sigma_n^2 I$. The channels were then predicted using a prediction horizon of 72 OFDM-symbols, corresponding to $5.1$ ms in time and $0.08 - 0.12$ wavelengths in space at the given carrier frequency and user speed (Table I). The predictions were carried out for $432$ adjacent subcarriers, using CRS on every ninth OFDM-symbol and every subcarrier. The NMSE (6) was averaged over each $0.5$ s interval for every subcarrier separately. We have observed that the Kalman filter converges fast in this setting. Therefore we assume the use of a stationary Kalman filter.

For evaluation of the precoding methods, $M = 3$ single antenna users were scattered over the measurement route in Figure 2. The precoding schemes were evaluated for the $432$ subcarriers, using the predictions obtained for all five $\sigma_n^2$ levels, as the users moved forward along the route for $0.5$ s. The performance was averaged over the $0.5$ s for each subcarrier at $545$ different sets of user starting positions along the measurement route.

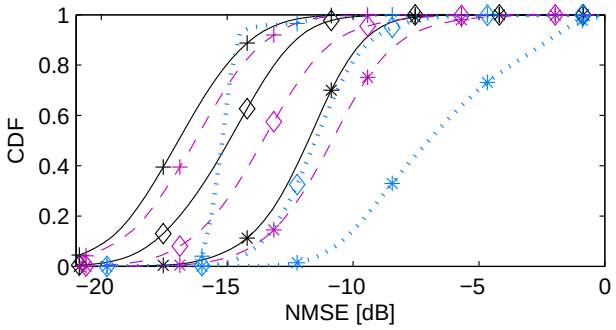


Figure 3. CDF for the prediction NMSE of the channel of the best (black solid lines), the second best (purple dashed lines) and the worst (blue dotted lines) base station, grouped w.r.t. SNR as 0 <SNR< 10 dB (stars), 10 <SNR< 20 dB (diamonds), and 20 <SNR< 30 dB (pluses) .

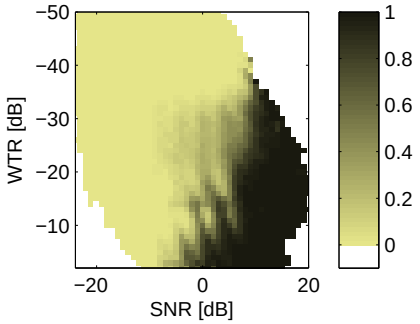


Figure 4. The prediction performance for the channel from the weakest base station as a function of the SNR for that base station and the WTR. Black indicates that all realizations achieved an $NMSE \leq -8dB$ and pale yellow that none did so. White areas correspond to no data.

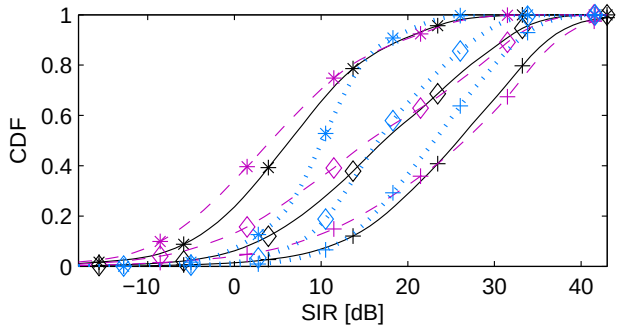


Figure 5. CDF of the SIR of the individual users, using ZF (blue dotted lines), RLP with $S = 0$ (black solid lines), RLP with $S = S_\rho$ (purple dashed lines). The users are sorted into groups of the SNR each user would experience if it received non-coherent JT from all base stations with no intracluster interference. The intervals are 0 <SNR< 10dB (stars), 10 <SNR< 20dB (diamonds), and 20 <SNR< 30dB (pluses).

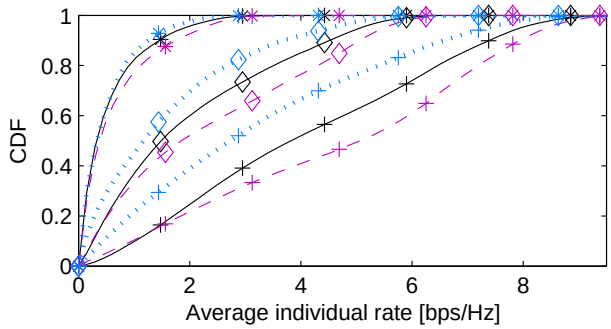


Figure 6. CDF of the rate of the individual users. Lines and markers represent the same as in Figure 5.

### C. Prediction performance

Channel prediction for downlink CoMP has to be performed for channels with very different powers. Figure 3 shows cumulative distribution functions (CDFs) for the NMSE of the predictions. The results are sorted into groups, best (with highest SNR), second best and worst base station every 0.5 s, since different base stations provide the strongest signal. The SNR, which here represents the ratio of the CRS power to the noise power, is averaged over 0.5 s. The channels from best and second best base stations usually have a good NMSE, much lower than for the channels from the worst base station, even when they all belong in the same SNR range.

In Figure 4 we see the performance of the predictions of the channel from the worst base station as a function of its SNR and of the weakest-to-total signal ratio (WTR). The WTR is the SNR when only the worst base station transmits divided by the SNR when all three base stations transmit non-coherently. It is clear that the prediction performance depends both on SNR and WTR. It is the WTR dependence that causes the large gap between the CDFs of the best and worst base stations in Figure 3. This dependence arises due to our use of quasi-orthogonal downlink CRSs. For frequency selective channels, the received CRS vectors will not be orthogonal over the $w$ subcarriers ($w = 4$ corresponds to 60 kHz). Therefore, since the measured channels are not perfectly flat fading, there will be some interference between the predictions of the three channels. This effect is most severe for the channel from the weakest base station and it increases with a decreasing WTR. A similar behavior can be seen for the second best base station.

### D. Precoding schemes

The Shannon spectral efficiency (rate) for the $i'th$ user, $C_i = \log_2(1 + SINR_i)$, is calculated for all five noise power levels, every subcarrier, and each set of starting positions.

The resulting average performance of the precoding schemes, averaged over 545 sets of user positions, is shown in Table II. We see that the RLP scheme with $S = 0$ attains a higher average SIR than both ZF and RLP with an iteratively adjusted $S = S_\rho$. This is expected since a minimization of the interference, averaged over the prediction errors, should also increase the SIR. However, with respect to the sum-rate, which was the optimization criterion used for the iterations, RLP with $S = S_\rho$ outperforms RLP with $S = 0$ and ZF by 13% and 41% respectively. Studies have also been made on optimizing more than one diagonal element of $S_\rho$ but these have only resulted in minor improvements.

The differences of the schemes can be studied more closely in Figure 5 and Figure 6, which shows the CDFs for the SIR and rate respectivly, for the individual users. Here users are sorted by the SNR they would experience if all three base stations would transmit non coherently to that user only. From the rate results we observe large gains with RLP at

Table II
AVERAGE PERFORMANCE FOR THE PRECODING METHODS.

| Method | $\overline{\text{SIR}}$ [dB] | $\overline{\sum_i C_i}$ [bps/Hz] |
|---|---|---|
| ZF | 21.0 | 9.4 |
| RLP, $V = I$,$S = 0$ | 22.1 | 11.8 |
| RLP, $V = I$ $S = S_\rho$ | 21.4 | 13.3 |

high SNR and small gains at low SNR. This indicates that RLP would work best in combination with clustering schemes that reduce the interference floor [3]. From the SIR CDFs we see that the RLP schemes let users with low channel quality experience a very low SIR in order to improve the SIR for users with high channel quality. In contrast, ZF does the opposite. However, the effort by ZF to improve the SIR does not results in noticeable improvement in the *rates* for the weak users. Their channel quality remains noise limited. In contrast, the users with high channel quality are often interference limited, so from a sum-rate perspective the system gains more by decreasing their interference.

There are three reasons as to why the RLP schemes prioritize users with high channel quality.

- High channel gain users will have smaller prediction errors, causing the RLP to prioritize them.
- Our choice of $V = I$ indicates that all interference components are equally important to minimize. The resulting residual interference power then tends to be equal for different users and will affect the users with low channel gain more than users with high channel gain.
- When we use $S_\rho$, the received signal power is part of the criterion and the scheme does not waste transmit power overcoming a poor channel. This effect is not present when using $S = 0$.

One could argue that from a fairness point of view, ZF is superior to RLP. However, the SIR improvement that ZF achieves for the worst users is still not sufficient to improve the rate for those users. We found that the situations for which the RLP suppress the weakest user were often characterized by one base station having a very poor channels to all three users. In these cases we basically have two transmit antennas trying to send to three receiver antennas. It would then make sense to instead remove one user and let it use other transmission resources.

### *E. Comparison of joint transmission and cellular transmission*

Both RLP and ZF are joint transmission (JT) CoMP schemes where we assume that all transmitters have access to the same information. We will now compare JT, using the RLP scheme with $S = S_\rho$, to localized transmission where every user is then allocated to its strongest base station, at all subcarriers for the full $0.5$ s. Three users are randomly positioned, as in the previous section. When $m$ users belong to the same base station they are served on orthogonal resources, using every $m'th$ subcarrier. Our test environment in Figure 2 then comprises three single antenna base stations, each transmitting at maximum power independently of the other base stations, in a cellular system assumed to use frequency reuse 1.

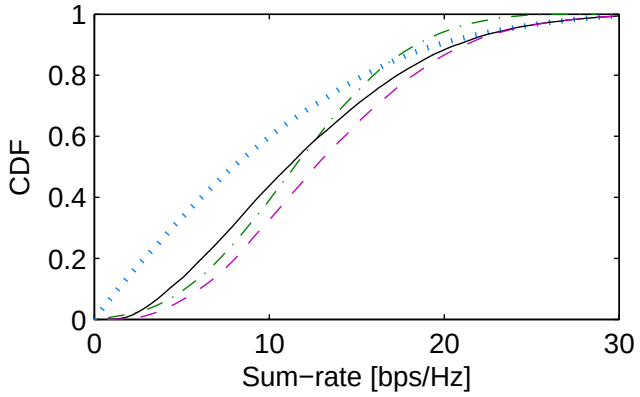


Figure 7. CDF of sum-rate over sets of user positions, noise levels and subcarriers for ZF (blue dotted lines), RLP with $S = 0$ (black solid lines), RLP with $S = S_\rho$ (purple dashed lines) and cellular transmission (green dashed-dotted lines).

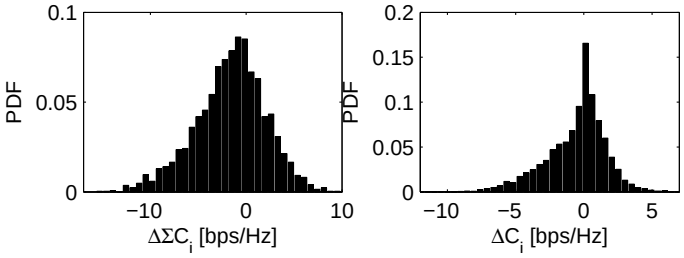


Figure 8. The distribution of difference between the rate performance of a cellular approach and the CoMP scheme (RLP with $S = S_\rho$) from a system, i.e. sum-rate, perspective (left) and a user perspective (right). Columns to the left of zero indicate that CoMP is the better of the two and columns to the right of zero indicate that the cellular approach is the better of the two.

In Figure 7, we see that the sum-rate is improved when using RLP JT instead of cellular transmission from an average of $11.7$ bps/Hz to $13.3$ bps/Hz, corresponding to a $13\%$ improvement. The improvement for the lowest sum-rates ($5\%$ percentile) is $24\%$.

However, all user groups are not helped by switching from cellular to RLP JT. The average rate over all subcarriers (including those subcarriers where the rate is zero due to no transmission) of the cellular approach $\bar{C}_i^{cell}$ is now compared to the average rate over all subcarriers for the RLP scheme, $\bar{C}_i^{RLP}$. Figure 8 shows the distribution of the difference between the two schemes from a system perspective (left) and from a user perspective (right). The system perspective is represented as the difference in sum-rate, $\Delta\left(\Sigma_i^M C_i\right) = \Sigma_i^M \bar{C}_i^{cell} - \Sigma_i^M \bar{C}_i^{RLP}$, and the user perspective is represented as the difference in Shannon rate for each user, $\Delta C_i = \bar{C}_i^{cell} - \bar{C}_i^{RLP}$. We find that that although RLP JT outperforms cellular transmission on average, $48\%$ of the users would benefit from cellular transmission and in $33\%$ of the situations simulated, the system would benefit from cellular transmission.

A system that could switch between the two schemes at a $0.5$ s basis could hence achieve a higher sum-rate than a system using only one of the schemes. In our simulation, switching would result in an average sum-rate of $14.0$ bits/s/Hz which correspond to a $20\%$ improvement from cellular trans-

mission and a 5% improvement from RLP JT. Calculating the 5% percentile of the sum-rate over all subcarriers (i.e. without averaging over all subcarriers), we find a 45% gain compared to always using cellular transmission and a 17% gain compared to always using RLP JT.

The gray-scaled bar at the bottom of Figure 2 shows the fraction of users at a given position that benefit from RLP JT. CoMP tends to be most beneficial to users who have at least two transmitters with good SNR. Otherwise, it is not obvious from these results that CoMP would always be better or worse at a specific geographical position. Therefore it is important to have a tool to evaluate if a user should belong to a joint or a localized transmission scheme.

## VI. Conclusions

We have studied how well we can simultaneously predict channels from three transmitters with different fading statistics, using Kalman predictors and quasi-orthogonal CRSs. We also introduce robust feedforward control precoder designs and evaluate them on the measured channels. Prediction accuracy is in general high for the considered 5.1 ms horizons at high pedestrian velocity. However, non-orthogonality between reference signals due to frequency selective channels reduces the performance for the weakest channels. Fully orthogonal CRS patterns would remove this effect. One major benefit of the Kalman predictor is that it provides channel error statistics. We have seen, by comparing RLP to ZF, that taking such statistics into account in the precoder design improves the user rates. The scheme can be further improved by iterative adjustment of the transmit power penalty matrix, $S$.

Though JT RLP outperforms cellular transmission on average, results indicate that further gains are possible with a system that can switch between the two approaches every 0.5 s.

*Open issues :* The possibilities for influencing the performance by using the penalty matrices $S$ and $V$ in (11) were only touched upon in this study. Studies are ongoing on how to adjust these matrice's parameters for different purposes. Moreover, our results will be compared to global optima with respect to all elements of $R$.

Of the possible optimization criteria, the sum-rate is widely used and is therefore convenient for producing results that are comparable. However, the use of the sum-rate produces a tendency to suppress the users with weak channels that would need the rate improvement most. In ongoing work, other criteria are investigated.

With more than $N$ potential users, an appropriate user grouping and scheduling scheme, that allocates the frequency domain resource blocks to sets of spatially compatible users, should be used in combination with the precoder. Such joint designs are studied in ongoing work.

## Acknowledgments

The research leading up to these results has received funding from the European Commission's seventh framework program FP7-ICT-2009 under grant of arrangement no 247223 also referred to as ARTIST4G. It is also partially funded by the Swedish Research Council via the framework program Dynamic Multipoint Wireless Transmission. We thank Ericsson Research for providing the channel measurements.

## Appendix

AR model parameter estimates are obtained at the receiver using autocorrelation over time and the Yule-Walker equations. To improve the accuracy, AR models are estimated for data series subsampled with the step $\vartheta$ equal to the prediction horizon. Poles $\{p_l\}$ of the $\Delta t$-spaced model for $h_{ij}^k$ are calculated from the subsampled model. As described in Chapter 4 of [10], the model (3) is constructed as follows. First, the AR process is realized in diagonal state space form, with $\mathrm{diag}(p_l)$ as state update matrix. The vector of channels from one transmitter at all $w$ subcarriers is then formed in block diagonal form. The same AR parameters are used for all subcarriers. These submodels are combined in block diagonal form to obtain the model (3) for all channels from the $N$ transmitters. Submodels for transmitters at different sites require different AR parameters.

The computational complexity is reduced to a level that is realistic for use in terminals by the matrix $A_i$ being diagonal and by using the stationary Kalman filter, see Chapter 8.6 of [10]. The process noise covariance matrix is found by combining equations (4.2.28) and (4.2.42) in [10].

## References

[1] K.M. Karakayali, G.J. Foschini and R.A. Valenzuela, "Network coordination for spectrally efficient communications in cellular systems," *IEEE Wireless Communications*, vol. 13, no. 4, Aug. 2006, pp. 56-61.

[2] R. Irmer, H. Droste, P. Marsch, M. Grieger, G. Fettweis, S. Brueck, H-P. Mayer, L. Thiele and V. Jungnickel, "Coordinated Multipoint: Concepts, performance and field trial results," *IEEE Communications Magazine*, Feb 2011, pp. 102-111.

[3] W. Mennerich and W. Zirwas, "Reporting effort for cooperative systems applying interference floor shaping," *IEEE PIMRC 2011*, Toronto, Canada, Sept. 2011.

[4] K. Öhrn, A. Ahlén and M. Sternad, "A probabilistic approach to multivariable robust filtering and open-loop control," *IEEE Trans. on Automatic Control*, vol 40, March 1995, pp 405-418.

[5] R. Hunger, F.A. Dietrich, M. Joham and W.Utschick, "Robust transmit zero-forcing filter," *2004 ITG Workshop on Smart Antennas*, Munich, Mar 2004, pp 130-137.

[6] P.M Castro, M. Joham, L.Castedo and W. Utschick, "Robust MMSE linear precoding for Multiuser MISO systems with limited feedback and channel prediction," *IEEE PIMRC 2008*, Sept. 2008.

[7] M.B. Shenouda, and T.N. Davidson, "On the design of linear transceivers for multiuser systems with channel uncertainty," *IEEE J. on Select. Areas Comm*, vol. 26, no. 6, Aug. 2008, pp. 1015-1024.

[8] X Zhang, D.P. Palomar, and B. Ottersten, "Statistically robust design of linear MIMO transceivers," *IEEE Trans. Signal Process*, vol. 56, no 8, Aug. 2008, pp 3678-3689.

[9] D. Aronsson and M. Sternad, "Kalman predictor design for frequency-adaptive scheduling of FDD OFDM uplinks," *IEEE PIMRC 2007*, Athens, Greece, Sept. 2007.

[10] D. Aronsson, *Channel Estimation and Prediction for MIMO OFDM Systems - Key Design and Performance Aspects of Kalman-based Algorithms*, Department of Engineering Sciences, Uppsala University, March 2011. Available: www.signal.uu.se/Publications/pdf/a112.pdf.

[11] J. Medbo, I. Siomina, A. Kangas and J. Furuskog, "Propagation channel impact on LTE positioning accuracy - A study based on real measurements of observed time difference of arrival," *IEEE PIMRC 2009*, Tokyo, Japan Sept. 2009.