

# Performance Analysis of (TDD) Massive MIMO with Kalman Channel Prediction

Salil Kashyap, Christopher Mollén, Björnson Emil and Erik G. Larsson

**Conference Publication**



N.B.: When citing this work, cite the original article.

Original Publication:

Salil Kashyap, Christopher Mollén, Björnson Emil and Erik G. Larsson, Performance Analysis of (TDD) Massive MIMO with Kalman Channel Prediction, 2017, International Conference on Acoustics, Speech, and Signal Processing.

Copyright: [www.ieee.org](http://www.ieee.org)

Postprint available at: Linköping University Electronic Press

<http://urn.kb.se/resolve?urn=urn:nbn:se:liu:diva-135788>



# PERFORMANCE ANALYSIS OF (TDD) MASSIVE MIMO WITH KALMAN CHANNEL PREDICTION

Salil Kashyap, Christopher Mollén, Emil Björnson, and Erik G. Larsson

Department of Electrical Engineering (ISY), Linköping University, 581 83 Linköping, Sweden  
Email: salilkashyap@gmail.com, {christopher.mollen, emil.bjornson, erik.g.larsson}@liu.se

## ABSTRACT

In massive MIMO systems, which rely on uplink pilots to estimate the channel, the time interval between pilot transmissions constrains the length of the downlink. Since switching between up- and downlink takes time, longer downlink blocks increase the effective spectral efficiency. We investigate the use of low-complexity channel models and Kalman filters for channel prediction, to allow for longer intervals between the pilots. Specifically, we quantify how often uplink pilots have to be sent when the downlink rate is allowed to degrade by a certain percentage. To this end, we consider a time-correlated channel aging model, whose spectrum is rectangular, and use autoregressive moving average (ARMA) processes to approximate the time-variations of such channels. We show that ARMA-based predictors can increase the interval between pilots and the spectral efficiency in channels with high Doppler spreads. We also show that Kalman prediction is robust to mismatches in the channel statistics.

**Index Terms**—channel aging, channel estimation, channel prediction, Kalman estimation, massive MIMO.

## 1. INTRODUCTION

Massive MIMO base stations are equipped with hundreds of antennas, which enable them to communicate with tens of users over the same time-frequency resource. Such systems can handle larger volumes of data and numbers of users than existing systems and are, therefore, a leading technology for future communication systems [1]. We want to answer the following questions:

- Can channel prediction increase the interval between pilot transmissions in massive MIMO? What is the rate loss if the precoding matrix is based on a predicted channel, instead of a recently estimated channel?
- How robust is the predictor to imperfect knowledge of the channel statistics, such as a mismatch in Doppler spread? and what performance loss does a mismatch incur?
- Can we use low-complexity predictor models to approximate a given channel spectrum?

**Contributions:** We consider low-complexity ARMA models to approximate the time-variations of a channel whose true spectrum is rectangular. The downlink achievable rates are computed for predicted channel state information (CSI) acquired using Kalman channel prediction. We also investigate cases, where either the channel spectrum is not fully known or the Doppler spread is not fully known. We

This research is funded by the European Union Seventh Framework Programme under grant agreement number ICT-619086 (MAMMOET).

S. Kashyap was with Linköping University during the course of this work. He is now with Marvell, India.

present numerical results that quantify the loss in rate incurred due to prediction errors and due to imperfect knowledge of the channel statistics.

**Related Literature:** The effect of channel aging on massive MIMO systems assuming matched filter and an infinite number of antennas at the base station was investigated in [2]. However, there an AR(1) model was used to approximate the time-variations of a channel whose true spectrum is the Jakes spectrum. Another paper that investigated the effects of channel aging assuming Jakes spectrum is [3], where the sum-rate for massive MIMO systems with matched filter and zero-forcing receivers in the presence of channel aging was derived. In [4], the authors showed that the Doppler shift due to relative movement of users as well as the phase noise due to noisy local oscillators contribute to channel aging. They incorporated both these effects in their channel aging analysis based on random matrix theory for massive MIMO systems. The problem of optimizing the throughput in terms of the number of transmit antennas, data and pilot energies in large point-to-point MIMO systems with channel aging was analyzed in [5], where it was shown that the effective channel coherence time increases with increasing number of antennas.

In contrast, in this paper, we consider Kalman filters based low-order ARMA models to estimate the time-varying channel coefficients whose true spectrum is rectangular.

## 2. SYSTEM MODEL

We consider a single-cell massive MIMO OFDM system, where the bandwidth is divided into  $S$  orthogonal subcarriers. The base station is equipped with an array of  $M$  antennas and there are  $K$  single-antenna users in the cell. The  $L$ -tap channel from the  $k^{\text{th}}$  user to the  $m^{\text{th}}$  base station antenna over the  $n^{\text{th}}$  OFDM symbol is denoted by  $\tilde{\mathbf{g}}_k^m[n] = [\tilde{g}_k^m[n, 0] \tilde{g}_k^m[n, 1] \dots \tilde{g}_k^m[n, L-1]]^T$ . For any user-antenna pair, the taps are assumed to be independent, but need not be identically distributed. We assume that the path-loss from a user is the same to all the base station antennas. Furthermore, we consider uncorrelated Rayleigh fading with  $\tilde{\mathbf{g}}_k^m[n] \sim \mathcal{CN}(\mathbf{0}, \mathbf{\Lambda}_k)$ , where  $\mathbf{\Lambda}_k = \text{diag}(\Lambda_k[0], \dots, \Lambda_k[L-1])$  is a diagonal matrix representing the channel power delay profile and large-scale fading of the  $k^{\text{th}}$  user.

### 2.1. Uplink Pilot Signaling and Channel Estimation

The frequency-domain signal  $\mathbf{y}_m[n] \in \mathbb{C}^{N_p \times 1}$  received over the  $n^{\text{th}}$  OFDM symbol at the  $m^{\text{th}}$  base station antenna during uplink pilot transmission is

$$\mathbf{y}_m[n] = \sum_{k=1}^K \sqrt{p_k^u} \mathbf{\Upsilon}_k^t \mathbf{\Omega}_r \tilde{\mathbf{g}}_k^m[n] + \mathbf{w}_m[n], \quad (1)$$

where  $p_k^u$  is the uplink pilot SNR per subcarrier and per OFDM symbol of the  $k^{\text{th}}$  user and  $\mathbf{\Upsilon}_k^t \in \mathbb{C}^{N_p \times N_p}$  is a diagonal matrix with the

$N_p$ -length pilot sequence  $\mathbf{x}'_k$  of that user on its diagonal. The matrix  $\mathbf{\Omega}_r \in \mathbb{C}^{N_p \times L}$  consists of the first  $L$  columns and  $N_p$  rows of the  $S$ -point discrete Fourier transform (DFT) matrix  $\mathbf{\Omega} \in \mathbb{C}^{S \times S}$ , where  $[\mathbf{\Omega}]_{m,n} = e^{-j2\pi(m-1)(n-1)/S}$ . The rows in  $\mathbf{\Omega}_r$  correspond to the set of subcarriers on which the  $N_p$  pilots are sent. The pilots are equally spaced in frequency and  $N_p \leq S$ . The noise vector is denoted by  $\mathbf{w}_m[n] \sim \mathcal{CN}(\mathbf{0}, \mathbf{I}_{N_p})$  and is independent and identically distributed (i.i.d.) across antennas  $m$  and time  $n$ . The pilot sequences are orthogonal between the users, in the sense that  $\mathbf{\Omega}_r^H \mathbf{\Upsilon}_k^H \mathbf{\Upsilon}_i \mathbf{\Omega}_r = N_p \mathbf{I}_L \delta_{ki}$ , where  $\delta_{ki} = 1$  if  $k = i$  and  $\delta_{ki} = 0$  otherwise. For this condition to hold  $N_p \geq KL$ . A sufficient statistics for estimating  $\tilde{\mathbf{g}}_k^m[n]$  is

$$\tilde{\mathbf{y}}_k^m[n] = \frac{1}{\sqrt{N_p}} \mathbf{\Omega}_r^H \mathbf{\Upsilon}_k^H \mathbf{y}_m[n] = \sqrt{p_k^u N_p} \tilde{\mathbf{g}}_k^m[n] + \tilde{\mathbf{w}}_k^m[n], \quad (2)$$

where  $\tilde{\mathbf{w}}_k^m[n] \sim \mathcal{CN}(\mathbf{0}, \mathbf{I}_L)$  is i.i.d. across  $k$  and  $m$ .

## 2.2. Time-Correlated Channel Fading Model

The autocorrelation function (ACF) determines how the wireless channel varies over time. The ACF depends on the propagation geometry, the velocity with which the user moves and the antenna characteristics. In a scenario with isotropic scattering in all three dimensions, it is shown in [6] that the power spectral density (PSD) has flat band-limited characteristics with a normalized ACF  $r_{gg}[\tau] = \text{sinc}(2f_d T_s \tau)$ , where  $f_d$  is the maximum Doppler frequency and  $T_s$  is the OFDM symbol duration. This is the PSD that we consider in this paper, motivated by the fact that massive MIMO enables 3D beamforming, but we stress that the predictor can be used on channels of any ACF.

The objective is to estimate the channel taps at different time instants. To this end, the true ACF of the time variations is approximately modeled by a finite-order ARMA model. An ARMA( $p, p$ ) model for  $\tilde{\mathbf{g}}_k^m[n, l]$  can be written as [7]

$$\tilde{\mathbf{g}}_k^m[n, l] = \sum_{i=1}^p \phi_i \tilde{\mathbf{g}}_k^m[n-i, l] + \sum_{i=0}^p \psi_i \tilde{\eta}_k^m[n-i, l], \quad (3)$$

where  $p$  is the model order. One way to closely approximate the rectangular spectrum of the ACF  $r_{gg}[\tau] = \text{sinc}(2f_d T_s \tau)$  is to select the coefficients  $\phi_i$  and  $\psi_i$  in (3) from the transfer function of a Butterworth low pass filter of order  $p$  with cutoff frequency  $f_d T_s$ .

The ARMA model can be equivalently given as a state-space model with the state transitions

$$\tilde{\mathbf{X}}_k^m[n+1, l] = \mathbf{A} \tilde{\mathbf{X}}_k^m[n, l] + \mathbf{B} \tilde{\mathbf{u}}_k^m[n+1, l], \quad (4)$$

where  $\tilde{\mathbf{X}}_k^m[n, l] \triangleq [\tilde{g}_k^m[n, l], \dots, \tilde{g}_k^m[n-p+1, l]]^T$  is the state of the system of the  $l^{\text{th}}$  channel tap at time  $n$ , and  $\tilde{\mathbf{u}}_k^m[n+1, l]$  is the white Gaussian process noise. The matrices  $\mathbf{A}$  and  $\mathbf{B}$  in (4) are

$$\mathbf{A} = \begin{pmatrix} \phi_1 & \phi_2 & \dots & \phi_p \\ 1 & 0 & \dots & 0 \\ 0 & 1 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & 1 \end{pmatrix} \in \mathbb{C}^{p \times p}, \quad (5)$$

$$\mathbf{B} = \begin{pmatrix} \psi_0 & \psi_1 & \dots & \psi_p \\ 0 & 0 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & 0 \end{pmatrix} \in \mathbb{C}^{p \times (p+1)}. \quad (6)$$

From (2), the observations of the state of channel tap  $l$  at time  $n$  can be represented by a linear equation

$$\tilde{\mathbf{y}}_k^m[n, l] = \mathbf{S} \tilde{\mathbf{X}}_k^m[n, l] + \tilde{\mathbf{w}}_k^m[n, l], \quad (7)$$

where

$$\mathbf{S} = [\sqrt{p_k^u N_p}, 0, \dots, 0] \quad (8)$$

and  $\tilde{\mathbf{w}}_k^m[n, l] \sim \mathcal{CN}(0, 1)$  is the additive measurement noise.

Given a set of observations  $\tilde{\mathbf{y}}_k^m[1, l], \tilde{\mathbf{y}}_k^m[2, l], \dots, \tilde{\mathbf{y}}_k^m[n+1, l]$ , the task is to determine the filter that at the  $(n+1)^{\text{th}}$  time instant generates an estimate  $\hat{\tilde{\mathbf{X}}}_k^m[n+1, l]$  of the state  $\tilde{\mathbf{X}}_k^m[n+1, l]$ . This motivates the use of a Kalman filter. The following steps obtain the Kalman estimate of the  $l^{\text{th}}$  channel tap:

1. *Initialization:* We begin by initializing  $\hat{\tilde{\mathbf{X}}}_k^m[0, l] | [0] = \mathbf{0}$  and the prediction error covariance matrix  $\mathbf{P}_{0|0} = \Lambda_k[l] \mathbf{R}$ , where

$$\mathbf{R} = \begin{bmatrix} r_{gg}[0] & r_{gg}[1] & \dots & r_{gg}[p-1] \\ r_{gg}[1] & r_{gg}[0] & \ddots & r_{gg}[p-2] \\ \vdots & \ddots & \ddots & \vdots \\ r_{gg}[p-1] & \dots & \dots & r_{gg}[0] \end{bmatrix} \quad (9)$$

and  $r_{gg}[\tau] = \text{sinc}(2f_d T_s \tau)$  for the exemplified ACF.

2. *One-step-ahead prediction:* This involves estimating the state at  $n+1$  based on observations up to time instant  $n$ :

$$\hat{\tilde{\mathbf{X}}}_k^m[n+1, l] | [n] \triangleq \mathbb{E}[\tilde{\mathbf{X}}_k^m[n+1, l] | \tilde{\mathbf{y}}_k^m[n]] = \mathbf{A} \hat{\tilde{\mathbf{X}}}_k^m[n, l] | [n], \quad (10)$$

where  $(\tilde{\mathbf{y}}_k^m)^n = \tilde{\mathbf{y}}_k^m[1, l], \dots, \tilde{\mathbf{y}}_k^m[n, l]$ .

3. *Computing the prediction error covariance matrix:* The prediction error covariance matrix is given by

$$\begin{aligned} \mathbf{P}_{n+1|n} &\triangleq \mathbb{E}[(\tilde{\mathbf{X}}_k^m[n+1, l] - \hat{\tilde{\mathbf{X}}}_k^m[n+1, l] | [n]) \\ &\quad \times (\tilde{\mathbf{X}}_k^m[n+1, l] - \hat{\tilde{\mathbf{X}}}_k^m[n+1, l] | [n])^H | (\tilde{\mathbf{y}}_k^m)^n] \\ &= \mathbf{A} \mathbf{P}_{n|n} \mathbf{A}^H + \mathbf{B} \mathbf{B}^H. \end{aligned} \quad (11)$$

4. *Kalman update:* Given the prediction  $\hat{\tilde{\mathbf{X}}}_k^m[n+1, l] | [n]$ , suppose we take another observation  $\tilde{\mathbf{y}}_k^m[n+1, l]$ , then this can be used to update the predictive estimate as

$$\begin{aligned} \hat{\tilde{\mathbf{X}}}_k^m[n+1, l] | [n+1] &= \hat{\tilde{\mathbf{X}}}_k^m[n+1, l] | [n] \\ &\quad + \mathbf{K}_{n+1} (\tilde{\mathbf{y}}_k^m[n+1, l] - \mathbf{S} \hat{\tilde{\mathbf{X}}}_k^m[n+1, l] | [n]), \end{aligned} \quad (12)$$

where  $\tilde{\mathbf{y}}_k^m[n+1, l]$  denotes the observation at the current time instant  $n+1$  and  $\mathbf{S} \hat{\tilde{\mathbf{X}}}_k^m[n+1, l] | [n]$  denotes the predicted observation. The Kalman gain matrix  $\mathbf{K}_{n+1}$  that minimizes the mean square error is given by

$$\mathbf{K}_{n+1} = \mathbf{P}_{n+1|n} \mathbf{S}^H (\mathbf{S} \mathbf{P}_{n+1|n} \mathbf{S}^H + 1)^{-1}. \quad (13)$$

5. *Updated error covariance matrix:* The updated error covariance matrix is given by:

$$\begin{aligned} \mathbf{P}_{n+1|n+1} &\triangleq \mathbb{E}[(\tilde{\mathbf{X}}_k^m[n+1, l] - \hat{\tilde{\mathbf{X}}}_k^m[n+1, l] | [n+1]) \\ &\quad \times (\tilde{\mathbf{X}}_k^m[n+1, l] - \hat{\tilde{\mathbf{X}}}_k^m[n+1, l] | [n+1])^H | (\tilde{\mathbf{y}}_k^m)^{n+1}] \\ &= (\mathbf{I}_p - \mathbf{K}_{n+1} \mathbf{S}) \mathbf{P}_{n+1|n} (\mathbf{I}_p - \mathbf{K}_{n+1} \mathbf{S})^H + \mathbf{K}_{n+1} \mathbf{K}_{n+1}^H. \end{aligned} \quad (14)$$

This Kalman filter will be used for channel prediction in the next sections. We focus on the downlink where the channel ages but no new uplink pilots are available.

### 3. ACHIEVABLE DOWNLINK RATE ANALYSIS

In this section, we derive the achievable downlink rates, when using predicted channels. The signal  $\mathbf{x}_d[n, s] \in \mathbb{C}^{M \times 1}$  transmitted by the base station in the downlink over the  $s^{\text{th}}$  subcarrier and the  $n^{\text{th}}$  OFDM symbol is

$$\mathbf{x}_d[n, s] = \sqrt{p_d} \hat{\mathbf{A}}[n, s] \mathbf{q}[n, s], \quad (15)$$

where  $p_d$  is the downlink SNR,  $\hat{\mathbf{A}}[n, s] \in \mathbb{C}^{M \times K}$  is the precoding matrix that depends on the predicted CSI at the  $n^{\text{th}}$  OFDM symbol index and the  $s^{\text{th}}$  subcarrier, and  $\mathbf{q}[n, s] \sim \mathcal{CN}(\mathbf{0}, \mathbf{I}_K)$  contains the information symbols that is transmitted to the  $K$  users.

The precoding matrix  $\hat{\mathbf{A}}[n, s]$  should be selected based on the predicted estimates of the channels  $\hat{\mathbf{g}}_1^m[n], \dots, \hat{\mathbf{g}}_K^m[n]$  that are available at the base station at time  $n$ . Recall that the first entry of the state  $\tilde{\mathbf{X}}_k^m[n, l]$  contains the prediction of the  $l^{\text{th}}$  channel tap at time  $n$ . We can gather these predictions in a vector  $\hat{\mathbf{g}}_k^m[n]$  for  $l = 0, \dots, L - 1$ , which is the prediction of  $\mathbf{g}_k^m[n]$ . The predicted channel matrix  $\hat{\mathbf{G}}[n, s] \in \mathbb{C}^{M \times K}$  at subcarrier  $s$  is then formed by setting  $[\hat{\mathbf{G}}[n, s]]_{m,k} = \omega_s^H \hat{\mathbf{g}}_k^m[n]$ , where  $\omega_s^H$  consists of the  $L$  first elements at the  $s^{\text{th}}$  row in the DFT matrix  $\mathbf{\Omega}$ .

We consider zero-forcing, where the precoding matrix is

$$\hat{\mathbf{A}}[n, s] = \alpha_{ZF} \hat{\mathbf{G}}[n, s] (\hat{\mathbf{G}}[n, s]^H \hat{\mathbf{G}}[n, s])^{-1}. \quad (16)$$

The factor  $\alpha_{ZF}$  is chosen such that  $\text{tr}(\hat{\mathbf{A}}[n, s] \hat{\mathbf{A}}[n, s]^H) = 1$ , which makes  $\mathbb{E}[\|\mathbf{x}[n, s]\|^2] = p_d$ .

The signal vector  $\mathbf{y}_d[n, s] \in \mathbb{C}^{K \times 1}$  received collectively at the  $K$  users is given by

$$\mathbf{y}_d[n, s] = \mathbf{G}^H[n, s] \mathbf{x}_d[n, s] + \mathbf{w}_d[n, s], \quad (17)$$

where  $\mathbf{w}_d[n, s] \sim \mathcal{CN}(\mathbf{0}, \mathbf{I}_K)$  denotes the additive white Gaussian noise. Then, the signal  $y_{dk}[n, s]$  received on the downlink at the  $k^{\text{th}}$  user over the  $n^{\text{th}}$  OFDM symbol and the  $s^{\text{th}}$  subcarrier is

$$y_{dk}[n, s] = \sqrt{p_d} \mathbf{g}_k^H(n, s) \hat{\mathbf{a}}_k[n, s] q_k[n, s] + \sqrt{p_d} \sum_{i \neq k} \mathbf{g}_k^H[n, s] \hat{\mathbf{a}}_i[n, s] q_i[n, s] + w_{dk}[n, s], \quad (18)$$

where  $\hat{\mathbf{a}}_k[n, s]$  is column  $k$  of  $\hat{\mathbf{A}}[n, s]$ . Let us define  $\hat{\mathbf{W}}_{ki} \triangleq \mathbf{g}_k^H[n, s] \hat{\mathbf{a}}_i[n, s]$ . The users are assumed to only have statistical CSI, since there are typically no downlink pilots in massive MIMO. Therefore, we use the technique in [8] to obtain the downlink rate over subcarrier  $s$  and OFDM symbol  $n$  as

$$R_k(n, s) = \log_2 \left( 1 + \frac{p_d \mathbb{E} [|\hat{\mathbf{W}}_{kk}|]^2}{1 + p_d \text{var} [|\hat{\mathbf{W}}_{kk}|] + p_d \sum_{i \neq k} \mathbb{E} [|\hat{\mathbf{W}}_{ki}|^2]} \right). \quad (19)$$

Note that we compute a separate rate for each subcarrier  $s$  and OFDM symbol  $n$ . On average, the rate over any time-frequency grid will be  $\frac{1}{SN_d} \sum_n \sum_s R_k(n, s)$ , where  $N_d$  is the number of OFDM symbols used for downlink data transmission.

### 4. NUMERICAL RESULTS

In this section, we present numerical results to understand how channel prediction can be used to improve the zero-forcing precoder as the channel ages. Unless mentioned otherwise, we take  $M = 100$ ,  $K = 8$ ,  $L = 8$ ,  $S = 256$  and  $p_d = p_d^u = -5$  dB. We consider a uniform power

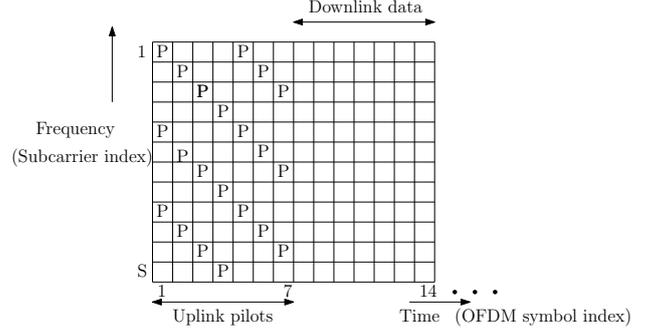


Fig. 1: Uplink pilots and downlink data transmission.

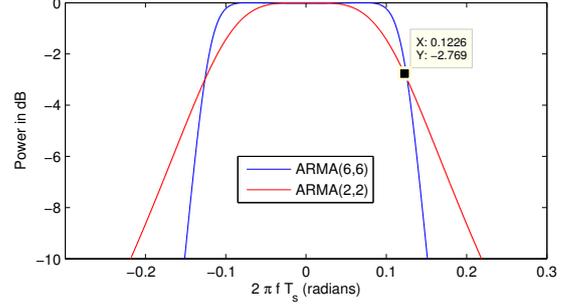


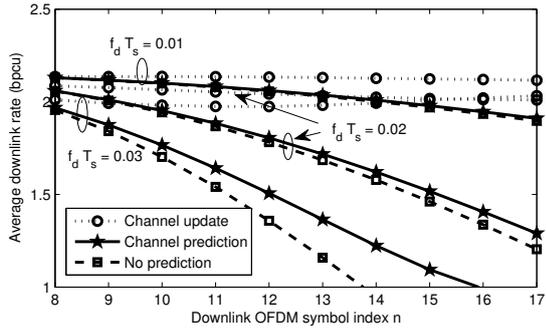
Fig. 2: Power spectral density of ARMA predictors of different order,  $f_d T_s = 0.02$

delay profile<sup>1</sup> and we take the number of pilot subcarriers  $N_p = KL$ . We further assume that the pilots are distributed over OFDM symbols 1 to 7 as shown in Fig. 1. We consider the normalized Doppler spread values 0.01, 0.02 and 0.03, which correspond to mobile scenarios at speeds of 81, 160 and 240 km/h with 2 GHz carrier frequency and  $T_s = 66.67 \mu\text{s}$  OFDM symbol duration (15 kHz inter-subcarrier spacing).

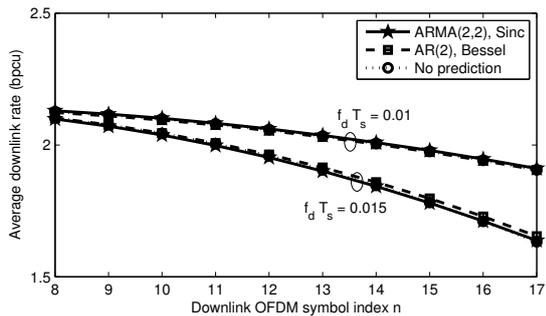
Fig. 2 shows the PSDs of ARMA predictors of different orders, whose ARMA coefficients are obtained from the transfer function of a Butterworth low pass filter of order  $p$  with cutoff frequency  $f_d T_s$ . These PSDs are used to approximate the rectangular spectrum of the true PSD of the channel. We observe that, as the model order increases, the spectrum falls off more sharply at the transition from the passband to the stopband.

Fig. 3 plots the average downlink rate  $(\frac{1}{S} \sum_s R_k(n, s))$  as a function of the OFDM symbol index with an ARMA(2,2) predictor for different Doppler spreads  $f_d T_s$ . Kalman estimates of the channel matrices are obtained from the uplink pilots located over symbols 1 to 7. For symbol indices 8 to 17, channel prediction is performed as only downlink data is transmitted over these symbols. The zero-forcing precoder computations over symbols 8 to 17 are based on the predicted channel matrices. It can be observed that the downlink rate decreases as  $f_d T_s$  increases or as time elapses with the increase in OFDM symbol index, since the channel estimates become more and more outdated. We also plot the case of no prediction, where instead of predicting the channel from 8<sup>th</sup>–17<sup>th</sup> OFDM symbol, we just continue to use the zero-forcing matrix computed at the 7<sup>th</sup> OFDM symbol index for precoding at subsequent OFDM symbols. While no prediction performs as well as ARMA(2,2) prediction at lower

<sup>1</sup>Note that channels with uniform power delay profile represent the worst case scenario [9]. Therefore, the study of such channels gives us an insight into the performance under the worst case conditions.



**Fig. 3:** Downlink rate with an ARMA(2, 2) predictor ( $M = 100$ ,  $K = 8$ ,  $L = 8$ ,  $S = 256$ ,  $p_d = p_k^u = -5$  dB)



**Fig. 4:** Robustness of predictor to mismatched spectrum ( $M = 100$ ,  $K = 8$ ,  $L = 8$ ,  $S = 256$ ,  $p_d = p_k^u = -5$  dB)

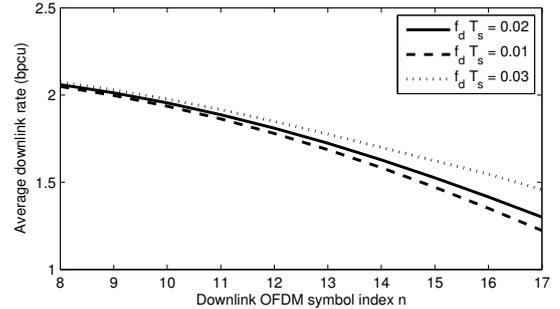
values of the normalized Doppler frequency  $f_d T_s$ , the gain in rate due to prediction increases as  $f_d T_s$  increases. Also plotted is the channel update approach, where zero-forcing matrix computations are obtained assuming the existence of uplink pilot transmissions over all the OFDM symbols 1 to 17.

This plot gives us an idea about how often we need to send uplink pilots before the rate has dropped below a certain percentage. For example, four downlink OFDM symbols can be sent in case of channel prediction and  $f_d T_s = 0.02$  if the system can tolerate a rate reduction by about 8%.

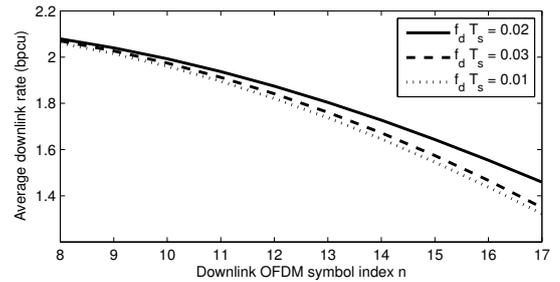
Fig. 4 plots the average rate as a function of the OFDM symbol index for the case when the channel spectrum is not fully known. We consider the ARMA(2,2) predictor from the previous figure, as well as an AR(2) predictor that is designed as if the ACF were a Bessel function (i.e., a Jakes spectrum). It is observed that the mismatched AR(2) predictor is almost as good as the ARMA(2,2) predictor and that no prediction results in the worst downlink rate. This indicates that the prediction can work despite mismatches in the statistics.

Figs. 5a and 5b plot the average downlink rate as a function of the OFDM symbol index for the case when there is a mismatch between the Doppler spread of the channel and the Doppler spread with which the predictor is designed, for ARMA(2,2) and ARMA(6,6) predictors respectively. There is a slight rate reduction in case of a mismatch, both when  $f_d T_s$  is smaller or larger than the true value, particularly at higher ARMA model orders.

Fig. 6 plots the average downlink rate over the 14<sup>th</sup> OFDM symbol as a function of the ARMA model order. It can be observed that the downlink rate increases marginally with the increase in the model order. It is, therefore, justified to use predictor models of order 2 which are computationally less expensive without compromising the performance.

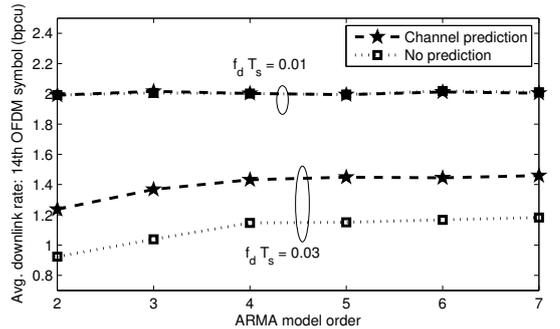


(a) ARMA(2, 2) predictor



(b) ARMA(6, 6) predictor

**Fig. 5:** Effect of mismatch of  $f_d T_s$ , Rect channel with Sinc ACF and  $f_d T_s = 0.02$ , ( $M = 100$ ,  $K = 8$ ,  $L = 8$ ,  $S = 256$ ,  $p_d = p_k^u = -5$  dB)



**Fig. 6:** Rate at 14<sup>th</sup> OFDM symbol vs. ARMA order ( $M = 100$ ,  $K = 8$ ,  $L = 4$ ,  $S = 256$ ,  $p_d = p_k^u = -5$  dB)

## 5. CONCLUSIONS

We investigated how channel prediction can be used in the downlink to improve the performance of the zero-forcing precoder as the channel ages. To this end, we designed a Kalman filter and exemplified it for channel aging modeled by a rectangular spectrum. We observed that an ARMA-based predictor can improve the spectral efficiency over no prediction, particularly at higher Doppler spreads. At low Doppler spreads, no prediction works reasonably well and performs poorer than channel prediction only when the channel becomes highly outdated. We also looked at how robust the predictors are to mismatches in the Doppler spreads. We found that, for a channel with rectangular PSD, the performance loss is marginal, which is explained by the fact that the channel decorrelates relatively slowly over time.

## 6. REFERENCES

- [1] E. G. Larsson, O. Edfors, F. Tufvesson, and T. Marzetta, "Massive MIMO for next generation wireless systems," *IEEE Commun. Mag.*, vol. 52, no. 2, pp. 186–195, Feb. 2014.
- [2] K. T. Truong and R. W. Heath Jr., "Effects of channel aging in massive MIMO systems," *J. Commun. and Networks*, vol. 15, no. 4, pp. 338–351, Aug. 2013.
- [3] C. Kong, C. Zhong, A. K. Papazafeiropoulos, M. Matthaiou, and Z. Zhang, "Sum-rate and power scaling of massive MIMO systems with channel aging," *IEEE Trans. Commun.*, vol. 63, no. 12, pp. 4879–4893, Dec. 2015.
- [4] A. K. Papazafeiropoulos, "Impact of general channel aging conditions on the downlink performance of massive MIMO," *IEEE Trans. Veh. Technol.*, May 2016.
- [5] R. Chopra, C. R. Murthy, and H. A. Suraweera, "On the throughput of large MIMO beamforming systems with channel aging," *IEEE Signal Process. Lett.*, Nov. 2016.
- [6] R. H. Clarke and W. L. Khoo, "3-D mobile radio channel statistics," *IEEE Trans. Veh. Technol.*, vol. 46, no. 3, pp. 798–799, May 1997.
- [7] Steven M. Kay, *Fundamentals of Statistical Signal Processing: Estimation Theory*, vol. 1, Prentice Hall, 1993.
- [8] J. Jose, A. Ashikhmin, T. L. Marzetta, and S. Vishwanath, "Pilot contamination and precoding in multi-cell TDD systems," *IEEE Trans. Wireless Commun.*, vol. 10, no. 8, pp. 2640–2651, Aug. 2011.
- [9] S. Stanczak, G. Wunder, and H. Boche, "On pilot-based multipath channel estimation for uplink CDMA systems: an overloaded case," *IEEE Trans. Signal Process.*, vol. 54, no. 2, pp. 512–519, Feb. 2006.