Kalman Smoothing for Irregular Pilot Patterns; A Case Study for Predictor Antennas in TDD Systems

Rikke Apelfröjd, Joachim Björsell, Mikael Sternad
Signals and Systems, Uppsala University
Box 534, 751 21 Uppsala, Sweden
Email: {rikke.apelfrojd, joachim.bjorsell, mikael.sternad}@angstrom.uu.se

Dinh-Thuy Phan-Huy
Orange, Paris, France
Email: dinhthuy.phanhuy@orange.com

***Abstract*—For future large-scale multi-antenna systems, channel orthogonal downlink pilots are not feasible due to extensive overhead requirements. Instead, channel reciprocity can be utilized in time division duplex (TDD) systems so that the downlink channel estimates can be based on pilots transmitted during the uplink. User mobility affects the reciprocity and makes the channel state information outdated for high velocities and/or long downlink subframe durations. Channel extrapolation, e.g. through Kalman prediction, can reduce the problem but is also limited by high velocities and long downlink subframes.**

An alternative solution has been proposed where channel predictions are made with the help of an extra antenna, e.g. on the roof of a car, so called predictor antenna, with the primary objective to measure the channel at a position that is later encountered by the rearward antenna(s). The predictor antenna is not directly limited by high velocities and allows the channel in the downlinks to be interpolated rather than extrapolated.

One remaining challenge here is to obtain a good interpolation of the uplink channel estimate, since a sequence of uplink reference signals (pilots) will be interrupted by downlink subframes. We here evaluate a Kalman smoothing estimate of the downlink channels and compare it to a cubic spline interpolation. These results are also compared to results where uplink channels are estimated through Kalman filters and predictors. Results are based on measured channels and show that with Kalman smoothing, predictor antennas can enable accurate channel estimates for a longer downlink period at vehicular velocities. The gaps in the uplink pilot stream, due to downlink subframes, can have durations that correspond to a vehicle movement of up to 0.75 carrier wavelengths in space, for Rayleigh-like non-line-of-sight fading.

I. Introduction

For many wireless transmission schemes, accurate channel state information at the transmitter (CSIT) of a downlink is crucial to achieve desirable gains. Such schemes include adaptive modulation and coding, channel aware scheduling and multi user multiple-input multiple-output (MIMO) transmission, e.g. zero forcing, [1], [2]. If the CSIT becomes outdated the gains associated with advanced transmission schemes may be lost, but this could be compensated by extrapolating channel estimates from the uplink, e.g. by using Kalman predictions [3]. However, prediction of the small scale fading through extrapolation of past estimates is very difficult for a prediction horizon that corresponds to a travelled distance beyond $0.1\lambda - 0.3\lambda$, where λ is the carrier wavelength [4]. For pedestrian users, the required prediction horizon is usually within this limitation, whereas for high mobility users a prediction horizon beyond 0.3λ is often required.

This limit of prediction horizon has been circumvented by the predictor antenna concept. It uses an extra antenna, *predictor antenna*, placed in front of the *main antenna*, e.g. on the roof of the vehicle, to scout the channel that will later be encountered by the main antenna, as originally proposed in [5]. Further experimental studies have shown a prediction normalized mean squared error of about -10 dB for velocities up to 50 km/h for all measured predictions horizons up to 3λ, ten times longer than the limit for channel extrapolation [6], [7]. To produce accurate predictions, the concept requires channel estimates for the predictor antenna from positions close to where the main antenna will transmit or receive signals. For a frequency division duplex (FDD) system, this requires dense enough downlink channel estimates, which was the case in [6], [7], [8], where orthogonal frequency division multiplexing (OFDM) pilots, evenly distributed in time and frequency, were used.

For a time division duplex (TDD) system, the predictor antenna would transmit known pilots in the uplink subframes, that are used for channel estimation on the network side. Assuming channel reciprocity, these channel estimates are then used to calculate predictions of the channel to the main antenna during a subsequent downlink subframe. The uplink/downlink ratio of the TDD frame might be adjusted so that the downlink transmission of the main antenna occurs close to a position where the predictor antenna already has measured the channel, as proposed in [9] and evaluated in [10]. How-

978-1-5386-6009-6/18/$31.00

ever, such a scheme would require individually adaptable uplink/downlink ratios based on the velocity of each user which is problematic from a system perspective. Instead, as suggested in [10], interpolation can be used for any given uplink/downlink ratio to generate channel estimates for the gaps in the uplink pilot sequence.

Interpolations can be used not only in TDD systems but also for FDD systems when the pilots are sparsely distributed in time relative to the vehicle velocity. However, the TDD case will be more challenging as the pilots are irregularly spaced in time and the longest durations without any pilots (due to downlink frames) are longer than for the FDD case. This motivates us to study the performance of interpolation schemes in this setting.

The Kalman smoother is the minimum mean squared error (MMSE)-optimal linear interpolator of noisy data, for known second order statistics of signal and noise [11]. It has been studied in applications such as the compensation for packet loss in wireless sensor network system [12] and channel equalization on the receiver side [13], both where partial observation losses occur.

In this paper we apply the Kalman smoother for the interpolation of the predictor antenna for TDD systems with a two-filter approach, using two state space models. The measurement based simulation results provided show that in the presence of predictor antennas when CSI is interpolated based on Kalman smoothing, the CSI quality of the downlink slot can be improved so that the downlink slot duration can be extended significantly compared to when only extrapolation of CSI into the downlink slot is utilized. The results can be used to determine how the flexibility that is left open within developing 5G standards in terms of reference signal rates and TDD subframe duration, can be used for vehicular users, some of which may use predictor antennas.

Notations

We here use diag$\{\cdot\}$ to represent a block diagonal matrix, a_j represents the j^{th} element of the vector $\mathbf{a}$ and $\mathbf{A}_k$ is the part of the matrix (or vector) $\mathbf{A}$ associated with the k^{th} channel component. The average is denoted $E[\cdot]$ and the covariance matrix of a zero mean vector is denoted $\text{cov}(\mathbf{a}) = E[\mathbf{a}\mathbf{a}^*]$. The estimate of a vector $\mathbf{a}(\tau_1)$ at time τ_1 based on measurements $\mathbf{y}(1), \ldots, \mathbf{y}(\tau_2)$ will be denoted $\hat{\mathbf{a}}_f(\tau_1|\tau_2)$ while $\hat{\mathbf{a}}_{bp}(\tau_1|\tau_1{+}N)$ is the estimate based on future measurements $\mathbf{y}(\tau_1{+}1), \ldots, \mathbf{y}(\tau_1{+}N)$.

II. System model

For simplicity, we outline the scheme for a single antenna base station and a vehicle with a single predictor antenna and a single main antenna. We assume an OFDM system where known pilots may be transmitted on OFDM symbols separated by an integer time index τ. The predictor antenna is located on the roof of the vehicle, a distance d forward from the main antenna. This means that prediction horizons in time of up to $T_{\max} = d/v_{\max}$, where $v_{\max}$ is the maximal vehicle velocity, can be accommodated. At times indexed by τ, the predictor antenna may transmit uplink pilots. If it does so, the main antenna is assumed not to transmit at the time-frequency resources used by these pilots. The baseband measurements at the base station antenna at a set of K potentially pilot-bearing subcarriers at time τ can then be written as

$$\mathbf{y}(\tau) = \mathbf{\Phi}(\tau)\mathbf{h}(\tau) + \mathbf{v}(\tau), \tag{1}$$

where the channel $\mathbf{h}(\tau)$ is a channel vector consisting of the zero mean scalar complex-valued channels at K subcarriers and $\mathbf{v}(\tau)$ is the measurement noise vector which may include many different components such as thermal noise, background noise and pilot contamination from other cells. The matrix $\mathbf{\Phi}(\tau)$ is a known pilot matrix which will be an all zero matrix for all τ when no uplink pilots are transmitted, e.g. during the TDD downlink subframes. We assume that the noise term $\mathbf{v}(\tau)$ is uncorrelated over time and uncorrelated with the state vector, and that it is zero mean Gaussian with known covariance matrix $\mathbf{R}_v = \text{cov}(\mathbf{v}(\tau))$.

A. Channel modelling

The small scale fading of a channel vector $\mathbf{h}(\tau)$ can be modelled as an autoregressive (AR) model on state space form

$$\begin{aligned} \mathbf{x}(\tau{+}1) &= \mathbf{A}\mathbf{x}(\tau) + \mathbf{B}\mathbf{w}(\tau), \\ \mathbf{h}(\tau) &= \mathbf{C}\mathbf{x}(\tau). \end{aligned} \tag{2}$$

Here, $\mathbf{x}(\tau)$ is the zero mean complex-valued state vector with covariance matrix $\mathbf{\Pi} = \text{cov}(\mathbf{x}(\tau))$, $\mathbf{w}(\tau)$ is the zero mean complex-valued Gaussian process noise with covariance matrix $\mathbf{Q} = \text{cov}(\mathbf{w}(\tau))$ and $\mathbf{A}$, $\mathbf{B}$, $\mathbf{C}$ are state space matrices.

The model (2) can be estimated from previous channel estimates. For this purpose, each component $h_k(\tau)$ in $\mathbf{h}(\tau)$ is individually modelled by an $n_{AR}{}^{\text{th}}$ order AR model

$$h_k(\tau) = -\Sigma_{i=1}^{n_{AR}} \alpha_i h_k(\tau{-}i) + w_k(\tau), \tag{3}$$

where $\{\alpha_i\}_{i=1,\ldots,n_{AR}}$ are the complex-valued model coefficients. By multiplying both sides of (3) by $h_k^*(\tau{-}t)$ for $t = 1, \ldots, n_{AR}$ and taking the expected value we obtain a set of equations

$$r_k(t) = -\Sigma_{i=1}^{n_{AR}} \alpha_i r_k(t{-}i), \qquad t = 1, \ldots, n_{AR}, \tag{4}$$

where $r_k(t)$ is the autocorrelation function of $h_k(\tau)$. Solving (4) provides the coefficients $\{\alpha_i\}_{i=1,\ldots,n_{AR}}$ through which the poles $\{p_i\}_{i=1,\ldots,n_{AR}}$ of the AR model are found by solving

$$z^{n_{AR}} + \Sigma_{i=1}^{n_{AR}} \alpha_i z^{n_{AR}-i} = 0. \tag{5}$$

Next, we set up the AR model (3) on state space form

$$\begin{aligned} \mathbf{x}_k(\tau{+}1) &= \mathbf{A}_k\mathbf{x}_k(\tau) + \mathbf{b}_k w_k(\tau), \\ h_k(\tau) &= \mathbf{c}_k\mathbf{x}_k(\tau), \end{aligned} \tag{6}$$

where $\mathbf{A}_k \in \mathbb{C}^{n_{AR}\times n_{AR}}$, $\mathbf{b}_k \in \mathbb{C}^{n_{AR}\times 1}$ and $\mathbf{c}_k \in \mathbb{C}^{1\times n_{AR}}$. We use the *diagonal* state space form, which reduces the complexity of the Kalman filters [3]. Then

$$\begin{aligned} \mathbf{A}_k &= \text{diag}\{p_i\}_{i=1,...,n_{AR}}, \\ b_{k_j} &= \Pi_{i=1,...,n_{AR},i\neq j}(p_j - p_i)^{-1}, \\ c_{k_j} &= p_j^{n_{AR}-1} \end{aligned} \tag{7}$$

for $j = 1, ..., n_{AR}$. The state vector in (2) is of dimension $n = Kn_{AR}$ and given by $\mathbf{x}(\tau) = [\mathbf{x}_1(t), ..., \mathbf{x}_K(\tau)]$ and the state transition matrices are given by

$$\begin{aligned} \mathbf{A} &= \text{diag}\{\mathbf{A}_k\}_{k=1,..,K}, \\ \mathbf{B} &= \text{diag}\{\mathbf{b}_k\}_{k=1,..,K}, \\ \mathbf{C} &= \text{diag}\{\mathbf{c}_k\}_{k=1,..,K}. \end{aligned} \tag{8}$$

The correlation of the fading of $h_k(\tau)$ at different subcarriers k is modelled by the process noise covariance matrix $\mathbf{Q}$. If $\mathbf{R}_h = \text{cov}(\mathbf{h}(\tau))$ has been estimated from previous measurement and if $\mathbf{A}$, $\mathbf{B}$ and $\mathbf{C}$ in (2) are know, then a covariance matrix of $\mathbf{w}(\tau)$, $\mathbf{Q} \in \mathbb{C}^{n\times n}$ which generates the current channel covariance matrix $\mathbf{R}_h \in \mathbb{C}^{K\times K}$ is given by Theorem 4.2 in [3] as

$$\mathbf{Q} = \mathbf{R}_h \oslash \mathbf{C}(\mathbf{B1B}^* \oslash (\mathbf{1} - \mathbf{aa}^*))\mathbf{C}^*. \tag{9}$$

Here $\mathbf{1}$ is an all one matrix of appropriate dimension, $\mathbf{a}$ is a vector given by the diagonal of the state space matrix $\mathbf{A}$ and $\oslash$ represents element wise division. We also obtain the covariance matrix of the state space vector as

$$\mathbf{\Pi} = \mathbf{BQB}^* \oslash (\mathbf{1} - \mathbf{aa}^*). \tag{10}$$

The expressions (9) and (10) assume a state space model on diagonal form (7), based on a stable AR model with distinct and nonzero poles $\{p_i\}_{i=1,...,n_{AR}}$. The element wise divisions appearing in (9) are then guaranteed to be nonsingular, since $\mathbf{1} - \mathbf{aa}^* \neq 0$ and all elements of $\mathbf{C}(\mathbf{B1B}^* \oslash (\mathbf{1} - \mathbf{aa}^*))\mathbf{C}^*$ will be nonzero. Additional aspects on AR modeling of fading multipoint MIMO channel models, where the use of (9) can be problematic in some FDD downlink situations, can be found in [14].

B. Backwards recursions

Similarly to (3)-(6) we can set up a backwards recursive model of the channel. The time inversion for the backwards recursive filter requires the introduction of an inverted complex plane $\zeta = z^{-1}$ where the the z-transform of the backwards time shift gives $x(\tau) \rightarrow x(\tau-1) \leftrightarrow \zeta X(\zeta)$. The time reversed AR model of order n_{AR} with zero mean i.i.d. process noise $w_k^b(\tau)$ is defined by

$$h_k(\tau) = -\Sigma_{i=1}^{n_{AR}} \alpha_i^b h_k(\tau+i) + w_k^b(\tau). \tag{11}$$

The model coefficients $\{\alpha_i^b\}_{i=1,...,n_{AR}}$ are found by solving

$$r_k(t) = -\Sigma_{i=1}^{n_{AR}} \alpha_i^b r_k(i+t), \qquad t = 1, \ldots, n_{AR}. \tag{12}$$

The similarity between (4) and (12) and the autocorrelation symmetry $r_k(t) = r_k^*(-t)$ gives us the backwards recursive model coefficients $\alpha_i^b = \alpha_i^*$ for $i = 1, \ldots, n_{AR}$. The model poles in the ζ-domain are given by solving

$$\zeta^{n_{AR}} + \Sigma_{i=1}^{n_{AR}} \alpha_i^b \zeta^{n_{AR}-i} = \zeta^{n_{AR}} + \Sigma_{i=1}^{n_{AR}} \alpha_i^* \zeta^{n_{AR}-i} = 0, \tag{13}$$

which with the similarity to (5) gives us the poles $\{p_i^b\}_{i=1,...,n_{AR}}$ which will be related to the poles of the forward AR model by $p_i^b = p_i^*$.

We can then set up a backwards recursive model for the channel vector

$$\begin{aligned} \mathbf{x}^b(\tau) &= \mathbf{A}^b\mathbf{x}^b(\tau+1) + \mathbf{B}^b\mathbf{w}^b(\tau), \\ \mathbf{h}(\tau) &= \mathbf{C}^b\mathbf{x}(\tau). \end{aligned} \tag{14}$$

Assuming diagonal form, then through (7)-(10) we have

$$\begin{aligned} &\mathbf{A}^b = \mathbf{A}^*, \qquad \mathbf{B}^b = (\mathbf{B}^T)^*, \qquad \mathbf{C}^b = (\mathbf{C}^T)^*, \\ &\mathbf{Q}^b = \mathbf{R}_h \oslash \mathbf{C}^b(\mathbf{B}^b\mathbf{1}(\mathbf{B}^b)^* \oslash (\mathbf{1} - \mathbf{a}^b(\mathbf{a}^b)^*))(\mathbf{C}^b)^*, \\ &\mathbf{\Pi}^b = \mathbf{B}^b\mathbf{Q}^b(\mathbf{B}^b)^* \oslash (\mathbf{1} - \mathbf{a}^b(\mathbf{a}^b)^*), \end{aligned} \tag{15}$$

where $\mathbf{Q}^b = \text{cov}(\mathbf{w}^b(\tau))$, $\mathbf{\Pi}^b = \text{cov}(\mathbf{x}^b(\tau))$ and $\mathbf{a}^b$ is a vector given by the diagonal elements of $\mathbf{A}^b$.

C. Upsampling

It can be convenient to use the Kalman filter, not only to interpolate over those time steps τ when there are no pilot transmissions, but also in between the pilot transmission, e.g. in cases where pilots are transmitted sparsely and accurate CSIT is needed in between these. If the channel needs to be upsampled by a factor n then we can use a Kalman filter based on AR models (2) and (14) with a time index spaced by τ/n. The poles of the forward model are then given by $\{p_i^{1/n}\}_{i=1,...,n_{AR}}$, [3], and likewise $\{(p_i^*)^{1/n}\}_{i=1,...,n_{AR}}$ for the time reversed model.

III. Kalman smoothing

Based on the measurements $\mathbf{y}(1), \ldots, \mathbf{y}(\tau)$ through (1) we can estimate the channel vector $\mathbf{h}_f(\tau)$ through the model (2) as

$$\hat{\mathbf{h}}_f(\tau|\tau) = \mathbf{C}\hat{\mathbf{x}}_f(\tau|\tau), \tag{16}$$

where the vector of state filter estimates $\hat{\mathbf{x}}_f(\tau|\tau)$ can be estimated recursively

$$\hat{\mathbf{x}}_f(t|t) = \mathbf{K}(t)\mathbf{y}(t) + (\mathbf{I} - \mathbf{K}(t)\mathbf{\Phi}(t)\mathbf{C})\mathbf{A}\hat{\mathbf{x}}_f(t-1|t-1)), \tag{17}$$

for $t = 1, \ldots, \tau$. Here, the $n \times K$ matrix $\mathbf{K}(t)$ is the Kalman filter gain based on the model (3).

Similarly, based on $\mathbf{y}(\tau+1), \ldots, \mathbf{y}(\tau+N)$ through (1) we can estimate the channel $\mathbf{h}(\tau)$ through the backwards recursive model (14) as

$$\hat{\mathbf{h}}_{bp}(\tau|\tau+N) = \mathbf{C}\hat{\mathbf{x}}_{bp}^b(\tau|\tau+N), \tag{18}$$

where the state vector $\hat{\mathbf{x}}^b_{bp}(\tau|\tau{+}N)$ can be estimated recursively backwards through

$$\hat{\mathbf{x}}^b_{bp}(t{-}1|\tau{+}N) = \mathbf{A}^b(\mathbf{K}^b(t)\mathbf{y}(t) + (\mathbf{I} - \mathbf{K}^b(t)\mathbf{\Phi}(t)\mathbf{C}^b)\hat{\mathbf{x}}^b_{bp}(t|\tau{+}N)), \tag{19}$$

for $t = \tau{+}N{+}1, \ldots, \tau{+}1$ and where the $n \times k$ matrix $\mathbf{K}^b(\tau)$ is the Kalman filter gain based on the model (11).

Details on how to initialize the filters and how to calculate the Kalman gains and the error covariance matrices $\mathbf{\Gamma}(\tau|\tau) = \text{cov}(\mathbf{h}(\tau) - \hat{\mathbf{h}}_f(\tau|\tau))$ and $\mathbf{\Gamma}_{bp}(\tau|\tau{+}N) = \text{cov}(\mathbf{h}(\tau) - \hat{\mathbf{h}}_{bp}(\tau|\tau{+}N))$ are provided in the Appendix.

Provided that the models (2) and (14) are accurate, the forward recursive Kalman filter provides the linear optimal estimate of $\mathbf{h}(\tau)$ through (16) based on all measurements $\mathbf{y}(1), \ldots, \mathbf{y}(\tau)$. The backwards Kalman filter provides the linear optimal estimate of $\mathbf{h}(\tau)$ through (18), based on all measurements $\mathbf{y}(\tau{+}1), \ldots, \mathbf{y}(\tau{+}N)$. As the two estimates are based on different measurements and, as the measurement noise $\mathbf{v}(\tau)$ in (1) is assumed independent over time and uncorrelated with $\mathbf{h}(\tau)$, the estimation errors of $\hat{\mathbf{h}}_f(\tau|\tau)$ and $\hat{\mathbf{h}}_{bp}(\tau|\tau{+}N)$ are uncorrelated. The mean squared error (MSE) optimal smoothed estimate of the channel vector is therefore given by the weighed average of these two estimates [11]

$$\hat{\mathbf{h}}_f(\tau|\tau{+}N) = \mathbf{W}_3(\mathbf{W}_1\hat{\mathbf{h}}_f(\tau|\tau) + \mathbf{W}_2\hat{\mathbf{h}}_{bp}(\tau|\tau{+}N)), \tag{20}$$

with weighting $\mathbf{W}_1 = \mathbf{\Gamma}^{-1}(\tau|\tau)$, $\mathbf{W}_2 = \mathbf{\Gamma}^{-1}_{bp}(\tau|\tau{+}N)$ and

$$\begin{aligned}\mathbf{W}_3 = \mathbf{\Gamma}(\tau|\tau{+}N) &= \text{cov}(\mathbf{h}(\tau) - \hat{\mathbf{h}}_f(\tau|\tau{+}N)) \\ &= \left(\mathbf{W}_1 + \mathbf{W}_2 - \mathbf{R}_h^{-1}\right)^{-1}.\end{aligned} \tag{21}$$

(The term $-\mathbf{R}_h^{-1}$ in (21) come in as both $\mathbf{W}_1$ and $\mathbf{W}_2$ include the prior information of $\mathbf{R}_h^{-1}$, and hence one of these must be removed.) For details, please see Sections 3.4.3 and 10.4 of [11].

A comment on the computational complexity of the Kalman smoother can be found in Appendix D.

IV. Case study: Predictor antennas for TDD

We consider a channel for one base station antenna in a TDD system where pilots are transmitted in the uplink for three consecutive pilot slots, indexed by τ, followed by a downlink with three consecutive slots without pilots. Two antennas, spaced by 2λ are placed in a straight line in the direction of travel on a vehicle. The required prediction horizon is 5 time slots and we assume that the vehicle is traveling with a velocity such that it takes 10 time slots for the main antenna to reach the position of the predictor antenna. The prediction horizon of 5τ then corresponds to a prediction horizon of λ in space. This is illustrated in Fig. 1, where we at time index $\tau = 9$ would need to predict the channel for the main antenna at time index $\tau = 14$. We then have CSIT based on the predictor antenna pilots at positions 1-3 and 7-9 and but want to know the CSIT for the main antenna when it reaches position 4. There is no measurement for that position, so interpolation will be needed.

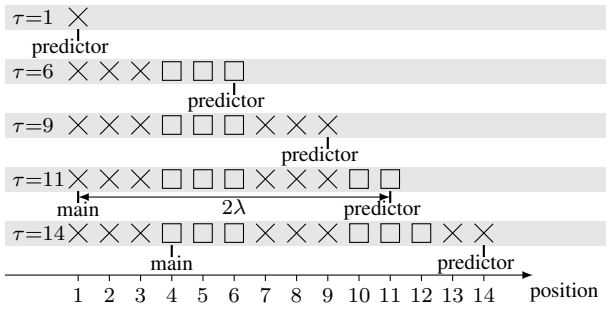

Fig. 1. A TDD system where pilots are transmitted in the uplink for three consecutive pilot slots followed by a downlink subframe with three consecutive slots without pilots. The figure shows the position of the predictor antenna and the main antenna at different times τ. Crosses mark positions for which the predictor antenna has transmitted pilots and thus made channel estimates available.

In situations when the positions of the main antenna and predictor antenna do not overlap (or are sufficiently close), the CSIT from the predictor antenna would also have to be interpolated in a denser pattern. This can be achieved by the upsampling described in Section II-C. Such upsampling would be required since the predictions are based on the property that the antennas are located at exactly the same position in space but at different times. Upsampling would typically be required at higher velocities as this reduces the spatial resolution of the channel estimate time series.

A. Measurements and simulation assumptions

To evaluate our scheme we use measurements collected while driving at a velocity of 25-39 km/h in downtown Dresden. The measurements were collected by TU Dresden. Channel sounding pilots were transmitted every 0.5 ms from four antennas mounted on the roof of the car at a carrier frequency of 2.53 GHz (λ = 119 mm) and were received at a base station. Three measurement sets are here used. These represent three different scenarios; A line of sight (LOS) scenario, a non-LOS (NLOS) scenario with a Doppler spectrum similar to that of Rayleigh fading and a NLOS scenario with a relatively flat Doppler spectrum. The measurements are selected from a larger set, which is described in greater detail in [6].

The original channel measurements were here subsampled (by a factor 4 or 6 depending on the original velocity) such that they would correspond to the channels collected when pilots are transmitted every 1 ms at a velocity of 75-83 km/h. Further, the channels have

been filtered before subsampling to reduce noise, using a 20th order finite impulse response (FIR) low pass filter. The filtered channels are estimated to have a channel-to-estimation error power ratio (here denoted SNR) of 28-32 dB. We assume these to be an accurate representation of the channel in our simulations.

We form 22 channel vectors $\mathbf{h}(\tau)$ that each consists of four subcarriers spaced by 60 kHz (in total 88 subcarriers that span 5.28 MHz). For each channel vector, we simulate measurement signals through (1) by adding i.i.d. circular symmetric Gaussian noise with a power set to simulate a pilot SNR of -5, 5 and 15 dB. To model a TDD system with frame structure as in Figure 1, the pilot matrix $\mathbf{\Phi}(\tau)$ of (1) is a unit matrix for $\tau = \{1, 2, 3\}+3m$ and an all zero matrix for $\tau = \{4, 5, 6\} + 3m$ with m being an integer. If pilots are transmitted at the very beginning and at the end of the downlink subframe, then this corresponds to an uplink subframe of length 2 ms and a downlink subframe of length 4 ms. This downlink subframe duration would at 80 km/h correspond to a distance of 89 mm, or 0.75λ.

For the example in Fig. 1 the channel prediction for the main antenna will be given by $\hat{\mathbf{h}}_f(\tau|\tau{+}5)$ by (20). The remaining signal to estimation noise error of the smoothed estimate of the predictor antenna SNR_p translates to the inverse of the normalized MSE (NMSE). Through the NMSE of the smoothing interpolation error at the predictor antenna NMSE_p, we can calculate the NMSE of the channel prediction for the main antenna NMSE_m as given by equation (14) in [6]

$$\text{NMSE}_m = 1 - \frac{|b|^2 \text{SNR}_p}{\text{SNR}_p + 1} = 1 - \frac{|b^2|}{1 + \text{NMSE}_p}, \quad (22)$$

where b is the maximum normalized cross-correlation between the channel of the main antenna and the delayed channel of the predictor antenna. This parameter is $|b| = 1$ in the ideal case, and it determines the ultimate performance for error-free measurements of the channel of the predictor antenna [6], [7]. If $|b|$ is close to $|b| = 1$, then by (22) $\text{NMSE}_m \approx \text{NMSE}_p$ for $\text{NMSE}_p \leq -8$ dB.

We model the small scale fading by using fourth order AR models (2) and (14). The autocorrelation of a single carrier channel, $r_k(t)$ and the correlation matrix of the channel vector $\mathbf{R}_h$ are found by averaging over all subcarriers and over 210 OFDM symbols for each antenna separately.

B. Results

In the scenario described above, the channel estimate of interest (used as predictor for the main antenna channels) is at the position where the predictor antenna was 5 time steps before it reached its current position. We therefore evaluate the Kalman smoothing estimate $\hat{\mathbf{h}}_f(\tau|\tau{+}5)$ for the predictor antenna. The resulting NMSE_p will depend on the SNR and on the location of τ inside or outside of the downlink subframes. For $\tau = \{1, 2, 3\}$ it is within an uplink subframe, where pilots are available. For $\tau = \{4, 5, 6\}$ we target a point that was passed during a downlink subframe. We compare this to the Kalman filter estimate $\hat{\mathbf{h}}_f(\tau|\tau)$. Note that the filter estimate will constitute a predictor from past estimates for $\tau = \{4, 5, 6\}$. We also compare to an estimate gained by using the measurement at $\tau = \{1, 2, 3\} + 3m$ and performing smoothing cubic interpolation through the matlab function `csaps`[1].

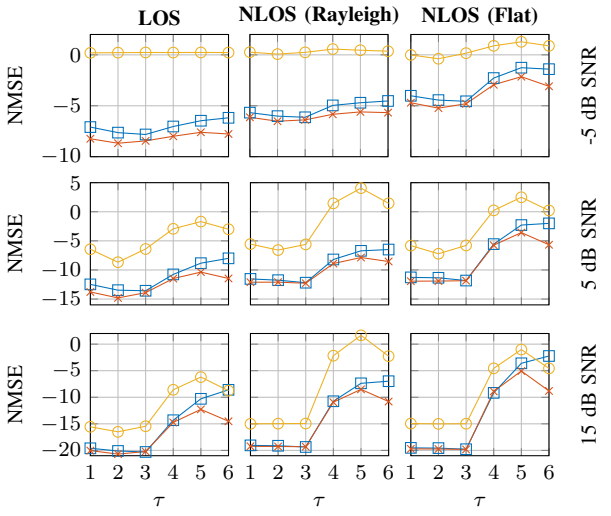

Fig. 2. The average NMSE for $\tau = 1, \ldots, 6 + 6m$ where m is an integer. Results are shown for Kalman filtering (squares), Kalman smoothing (crosses) and interpolation through smoothing cubic interpolation (circles).

Fig. 2 shows the average NMSE over all subcarriers and transmit antennas for the different fading scenarios and SNR levels. As a benchmark for estimation performance, we can use an NMSE of ≤ -8 dB, which will result in good prediction performance for a prediction horizon of one wavelength assuming that $|b|$ is close to $|b| = 1$ in (22).

For this particular scenario, our findings can be summarised as follows:

- Smoothing cubic interpolation is insufficient for estimating all downlink channels in any of the scenarios, whereas Kalman smoothing ensures good estimation performance for both the LOS scenario and the NLOS with Rayleigh fading at an SNR of 5 dB and above.
- An added bonus from Kalman smoothing is the noise reduction at low SNR. For the LOS scenario this ensures an NMSE below -8 dB for time slots $\tau = \{1, 2, 3\}$ at an SNR of -5 dB, which the filter/predictor does not achieve.

[1]This requires a smoothing factor $s \in [0, 1]$ where 1 means no smoothing and 0 gives the mean of the measurements as an estimate for all τ. For each SNR value and fading scenario we here choose the $s \in \{0, 0.1, 0.5, 0.9, 1\}$ that gives the best average NMSE.

- When the temporal correlation is lower, as in the case of the flat Doppler spectrum, it is harder to achieve the here targeted NMSE, even with smoothing.
- The performance difference between smoothing and filtering is largest for τ within downlink subframes $\tau \in \{4, 5, 6\}$. Here, the filter/predictor $\hat{\mathbf{h}}_f(\tau|\tau)$ is forced to perform a prediction based on previous pilot bearing symbols, while the smoother $\hat{\mathbf{h}}(\tau|\tau{+}5)$ also utilizes measurements from the next uplink subframe.

V. Conclusions

The present paper illustrates and evaluates an advantage of having a "predictor antenna" available. In a TDD system, where uplink and downlink channels are estimated based on uplink transmissions of known pilot symbols, such transmissions would be unavailable during the downlink transmission subframes. The only way to predict the downlink channels would then be to extrapolate forward in time from pilot-based measurements in previous uplink slots. The presence of a predictor antenna makes estimates of the "future" channel available. We can use this to improve the estimate of the channels that will be encountered within the next downlink subframe. The results here indicate that for many realistic fading scenarios, user mobility of 80 km/h, uplink pilot spacing of 1 ms and a downlink subframe duration of 4 ms, during which the vehicle in our example has travelled 0.75 wavelengths, Kalman smoothing can provide good interpolation estimates of the downlink of TDD channels at relatively low SNR (of 5 dB). Thus, with Kalman smoothing, predictor antennas can push the bound for advanced transmission scheme at high mobility in TDD systems. However, for some fading scenarios, here represented by the NLOS scenario with a flat Doppler spectrum, the TDD downlink channels will be very difficult to estimate, for so long downlink subframes.

The simulation in this paper is restricted to a few selected channels and a larger simulation with a larger variety of fading scenarios, pilot SNRs and downlink subframe lengths is required to draw conclusions as to what system requirements are needed to achieve sufficient estimation performances through Kalman smoothing. The effect of how the correlation between the channels at different base station antennas affects the estimates should also be investigated. This is the objective of a larger ongoing study where the results here are combined with the method for estimation predictor antenna system performance in [6].

Appendix A
Forward recursion Kalman filter

Assume that $\mathbf{P}(\tau_1|\tau_2) = \text{cov}(\mathbf{x}(\tau_1) - \hat{\mathbf{x}}(\tau_1|\tau_2))$ is the covariance matrix of the estimation error of $\hat{\mathbf{x}}(\tau_1|\tau_2)$. The Kalman filter can be divided into two parts, the one step prediction and the state update. Based on a recursive channel model as in (2) the one step prediction is given by

$$\hat{\mathbf{x}}(\tau|\tau{-}1) = \mathbf{A}\hat{\mathbf{x}}(\tau{-}1|\tau{-}1), \tag{23}$$

$$\mathbf{P}(\tau|\tau{-}1) = \mathbf{A}\mathbf{P}(\tau{-}1|\tau{-}1)\mathbf{A}^* + \mathbf{B}\mathbf{Q}\mathbf{B}^*, \tag{24}$$

and the update, given the current measurement $\mathbf{y}(\tau)$, is

$$\hat{\mathbf{x}}(\tau|\tau) = \mathbf{K}(\tau)\mathbf{y}(\tau) + (\mathbf{I} - \mathbf{K}(\tau)\mathbf{\Phi}(\tau)C)\hat{\mathbf{x}}(\tau|\tau{-}1)), \tag{25}$$

$$\mathbf{P}(\tau|\tau) = \mathbf{P}(\tau|\tau{-}1) - \mathbf{K}\mathbf{\Phi}(\tau)\mathbf{C}\mathbf{P}(\tau|\tau{-}1). \tag{26}$$

Here $\mathbf{K}(\tau)$ is the Kalman gain given by

$$\mathbf{K}(\tau) = \mathbf{P}(\tau|\tau{-}1)\mathbf{C}^*\mathbf{\Phi}(\tau)^* \cdot (\mathbf{\Phi}(\tau)\mathbf{C}\mathbf{P}(\tau|\tau{-}1)\mathbf{C}^*\mathbf{\Phi}^*(\tau) + \mathbf{R}_v)^{-1}.$$

Substituting (23) into (25) and assuming available measurements from times $t = 1, \ldots, \tau$ yields the expression (17).

Through the state vector estimation error matrix (26) we can find the covariance matrix of the estimation error for the filtered channel estimate (16) through

$$\mathbf{\Gamma}(\tau|\tau) = \text{cov}(\mathbf{h}(\tau) - \hat{\mathbf{h}}_f(\tau|\tau)) = \mathbf{C}\mathbf{P}(\tau|\tau)\mathbf{C}^*. \tag{27}$$

Appendix B
Backwards recursion Kalman filter

By reversing the time we can calculate the one step backwards Kalman prediction based on the backwards AR model (14) and given an estimate $\hat{\mathbf{x}}^b(\tau{+}1|\tau{+}1)$ as

$$\hat{\mathbf{x}}^b(\tau|\tau{+}1) = \mathbf{A}^b\hat{\mathbf{x}}^b(\tau{+}1|\tau{+}1), \tag{28}$$

$$\mathbf{P}^b(\tau|\tau{+}1) = \mathbf{A}^b\mathbf{P}^b(\tau{+}1|\tau{+}1)(\mathbf{A}^b)^* + \mathbf{B}^b\mathbf{Q}^b(\mathbf{B}^b)^*. \tag{29}$$

The backwards filter update, based on an estimate $\hat{\mathbf{x}^b}(\tau{+}1|\tau{+}1)$ and the current measurement $\mathbf{y}(\tau)$ is

$$\hat{\mathbf{x}}^b(\tau|\tau) = \hat{\mathbf{x}}^b(\tau|\tau{+}1){+}\mathbf{K}^b(\mathbf{y}(\tau){-}\mathbf{\Phi}(\tau)\mathbf{C}^b\hat{\mathbf{x}}^b(\tau|\tau{+}1)), \tag{30}$$

$$\mathbf{P}^b(\tau|\tau) = \mathbf{P}^b(\tau|\tau{+}1){-}\mathbf{K}^b\mathbf{\Phi}(\tau)\mathbf{C}^b\mathbf{P}^b(\tau|\tau{+}1), \tag{31}$$

where $\mathbf{P}^b(\tau_1|\tau_2) = \text{cov}(\mathbf{x}^b(\tau_1) - \hat{\mathbf{x}}^b(\tau_1|\tau_2))$ and $\mathbf{K}^b(\tau)$ is the Kalman gain given by

$$\mathbf{K}^b(\tau) = \mathbf{P}^b(\tau|\tau{+}1)(\mathbf{C}^b)^*\mathbf{\Phi}(\tau)^* \cdot (\mathbf{\Phi}(\tau)\mathbf{C}^b\mathbf{P}^b(\tau|\tau{+}1)(\mathbf{C}^b)^*\mathbf{\Phi}^*(\tau) + \mathbf{R}_v)^{-1}.$$

Substituting (30) into (28) and assuming available measurements from times $t = \tau{+}1, \ldots, \tau{+}N$ gives (19).

Through the state vector estimation error matrix (29) we can calculate the covariance matrix of the estimation error for the backwards one step channel prediction (18) through

$$\begin{aligned}\mathbf{\Gamma}_p(\tau|\tau{+}N) &= \text{cov}(\mathbf{h}(\tau) - \hat{\mathbf{h}}_p(\tau|\tau{+}N)) \\ &= \mathbf{C}^b\mathbf{P}^b(\tau|\tau{+}1)(\mathbf{C}^b)^*.\end{aligned} \tag{32}$$

APPENDIX C
INITIATING THE FILTERS

As the forward filters (17) and backwards filter (19) are both recursive, they must be initiated, which we choose to do by setting $\hat{\mathbf{x}}(0|0)$ and $\hat{\mathbf{x}}^b(\tau+N+1|\tau+N)$ to all zero vectors. Then the covariance matrices of the estimation error of the state vectors are given by

$$\mathbf{P}(0|0) = \mathbf{\Pi}, \qquad \mathbf{P}^b(\tau+N+1|\tau+N) = \mathbf{\Pi}^b.$$

APPENDIX D
COMPUTATIONAL COMPLEXITY

In Chapter 4.3 in [3] the required number of (arithmetic complex) operations, here referring to one multiplication and one addition, for the Kalman filter with a model on diagonal form is presented. From there we can derive the required number of operations for one Kalman recursion (23)-(26), $\mathbb{C}_{\text{Kal\_rec}}$. The Kalman smoother, as presented here, will require one forward recursion and N backward recursions for every time slot. The total number of operations related to the Kalman filter will then be

$$\mathbb{C}_{\text{Kal\_tot}} = (N+1)\mathbb{C}_{\text{Kal\_rec}}. \tag{33}$$

The smoothed estimate requires further calculation by (20) which in turn requires calculations of (16), (18), (27) and it's inverse, (32) and it's inverse and (21), in total $\mathbb{C}_{\text{smo}}$. The total number of operation by the Kalman smoother per time slot will then be

$$\mathbb{C}_{\text{smo\_tot}} = \mathbb{C}_{\text{Kal\_tot}} + \mathbb{C}_{\text{smo}}. \tag{34}$$

The exact number of operations is not presented here due to lack of space. However, $\mathbb{C}_{\text{smo\_tot}}$ will have a complexity of $\mathcal{O}((Kn_{AR})^2)$.

Both $\mathbb{C}_{\text{Kal\_tot}}$ and $\mathbb{C}_{\text{smo}}$ can be reduced by pre-calculating the converged Kalman matrices by Riccati equations. These matrices will be cyclo-stationary functions when the pilot patterns are cyclic over time and the environment is stationary [3]. The complexity is in this case $\mathcal{O}(n_{AR}, K^2)$.

A. Numerical example

For a setup as in the case study in Section IV, $K_{\text{tot}} = 88$, $K = 4$ and $n_{AR} = 4$, the Kalman smoother would require $\mathbb{C}_{\text{smo\_tot}} = 172480 + 11352 = 183832$ operations per step τ. The cyclo-stationary version would require $\mathbb{C}_{\text{smo\_tot}} = 50688 + 1848 = 52536$ operations, without the required pre-calculations taken into account.

The case study also makes a comparison to a smoothing spline method which has a complexity of $\mathcal{O}(n_x^3)$ where $n_x + 1$ is the number of measurement of which the spline coefficients are calculated from. When using five measurement every sixth time slot to smooth and interpolate a total of eight time slots, the smoothing spline method requires approximately 15000 operations every sixth time slot.

In this example the stationary Kalman filter requires 3.5 times more operations (and six times as often) compared to smoothing spline and the regular Kalman smoother requires 3.5 more operations than the stationary Kalman. It is worth mentioning that many of these calculations are parallelized by nature which is beneficial implementation wise.

ACKNOWLEDGEMENTS

We thank the 5G Lab at TU Dresden (Prof. Gerhard Fettweis) for providing the test-bed hardware and software used for the channel measurements. The work was supported by Orange Labs under contract F03131.

REFERENCES

[1] M. Sternad, S. Falahati, T. Svensson, and D. Aronsson, "Adaptive TDMA/OFDMA for wide area coverage and vehicular velocities," in *Proc. IST Mobile and Vehicular*, Dresden, Germany, Jun. 2005.

[2] N. Ravindran and N. Jindal, "Multi-user diversity vs. accurate channel state information in MIMO downlink channels," *IEEE Trans. Wireless Commun.*, vol. 11, pp. 3037–3046, 2012.

[3] D. Aronsson, "Channel estimation and prediction for MIMO OFDM systems - key design and performance aspects of kalman-based algorithms," Ph.D. dissertation, Uppsala University, Uppsala, Sweden, Mar. 2011. [Online]. Available: http://www.signal.uu.se/Publications/ptheses.html

[4] T. Ekman, "Predictions of mobile radio channels," Ph.D. dissertation, Uppsala University, Uppsala, Sweden, Oct. 2002. [Online]. Available: http://www.signal.uu.se/Publications/ptheses.html

[5] M. Sternad, M. Grieger, R. Apelfröjd, T. Svensson, D. Aronsson, and A. B. Martinez, "Using "predictor antennas" for long-range prediction of fast fading for moving relays," in *Proc. of IEEE WCNC 2012*, Paris, France, Apr. 2012.

[6] J. Björsell, M. Sternad, and M. Grieger, "Using predictor antennas for the prediction of small-scale fading provides an order-of-magnitude improvement of prediction horizons," in *Proc. of IEEE ICC 2017 Workshop WDN-S6*, Paris France, Jul. 2017.

[7] J. Björsell, M. Sternad, and M. Grieger, "Predictor antennas in action," in *Proc. of IEEE PIMRC 17*, Montreal, Canada, Oct. 2017.

[8] D.-T. Phan-Huy, S. Wesemann, J. Björsell, and M. Sternad, "Adaptive massive MIMO for fast moving connected vehicles: It will work with predictor antennas!" in *22nd International Workshop on Smart Antennas (WSA2018)*, Bochum, Germany, March 2018.

[9] D. T. Phan-Huy and M. Helard, "Large MISO beamforming for high speed vehicles using separate receive and training antennas," in *Proc. of IEEE Int. Symposium on Wireless Vehicular Com.*, Dresden, Germany, Jun. 2013.

[10] D. T. Phan-Huy, M. Sternad, and T. Svensson, "Making 5G adaptive antennas work for very fast moving vehicles," *IEEE Intelligent Transportation System Magazine*, pp. 71–84, Summer 2015.

[11] T. Kailath, A. H. Sayed, and B. Hassibi, *Linear Estimation*. Upper Saddle River: Prentice Hall, 2000.

[12] X. Lu, H. Wang, and M. Li, "Kalman fixed-interval and fixed-lag smoothing for wireless sensor systems with multiplicative noises," in *Proc. of IEEE CCDC 2012*, Taiyuan, China, Jul. 2012.

[13] S. Park and S. Choi, "Iterative equalizer based on kalman filtering and smoothing for MIMO-ISI channels," *IEEE Trans. Signal Process.*, vol. 63, pp. 5111–5120, 2015.

[14] R. Apelfröjd, "Kalman predictions for multipoint OFDM downlink channels," Signals and Systems, Uppsala University, Sweden, Tech. Rep., March 2018. [Online]. Available: http://www.signal.uu.se/Publications/pdf/c1402.pdf