

Spectral Efficiency of Time-Variant Massive MIMO Using Wiener Prediction

David Löschenbrand¹, Member, IEEE, Markus Hofer², Member, IEEE,
and Thomas Zemen¹, Senior Member, IEEE

Abstract—Massive multiple-input multiple-output (MIMO) serves as technological pillar in current 5G deployments. The spectral efficiency (SE) reduction due to time-variance of the wireless channel, and respective mitigation strategies, are still not sufficiently understood. In this paper, we propose a multi-step Wiener predictor with a general temporal covariance function to quantify the impact of time-variance on the SE for different numbers of consecutive pilot symbols and prediction horizons. We provide asymptotic expressions of the SE for maximum ratio combining (MRC) beam-forming that show excellent agreement to numerical simulations. We investigate the channel hardening capability of massive MIMO systems in time-variant propagation channels and show that it is largely independent of the prediction horizon. The numerical evaluations show that channel prediction allows to double SE utilizing four instead of one uplink pilot symbol for a prediction horizon of 0.3λ for both MRC and regularized zero-forcing (RZF) beam-forming.

Index Terms—massive MIMO, time variance, channel aging, prediction, channel hardening,

I. INTRODUCTION

MASSIVE multiple-input multiple-output (MIMO) systems are an important enabling technology in current 5G deployments, providing improved spectral efficiency (SE) and reliability. However, they still struggle to unlock their potential in high mobility scenarios, e.g., in automotive and industrial use cases.

The fundamental cause for this is the systems dependency on high quality and timely channel state information (CSI) for the communication channel from all users to all base station (BS) antenna elements for accurate beam-forming [1]. Currently, the only viable and scalable solution is the application of a time division duplex (TDD) method and relying on channel reciprocity. In high mobility scenarios, this approach inherently causes service degradation due to outdated CSI, also termed *channel aging* [2]. Solely increasing pilot overhead to overcome channel aging quickly becomes prohibitively expensive. Therefore, channel prediction is widely considered as the method of choice to combat outdated CSI [3]–[7].

In [3], the effects of channel aging are thoroughly derived, but the utilized autoregressive temporal correlation model of order one (AR(1)) and the consideration of only one-step prediction limits the applicability of the results to band-limited real-world fading processes. In [4], the authors model

achievable sum-rates of MIMO systems in the presence of channel aging and derive prediction strategies. However, again a simplified AR(1) model is utilized for the temporal correlation. A more suitable correlation is found in [8], where a Clarke's model is employed for modeling channel aging. However, no prediction algorithms are introduced or analyzed in [8]. Recent studies on high mobility massive MIMO utilize advanced prediction methods, such as Kalman filtering [5], [6] and machine learning (ML) [6], [7]. While providing state-of-the-art channel prediction quality, the work in [5]–[7] is not easily accessible to analytic and asymptotic considerations, especially concerning the variance of the achieved received signal power around its mean (which becomes negligible in case of *channel hardening*).

In this paper, we analyze the uplink SE of a massive MIMO system employing channel prediction to quantify its degradation due to user mobility. Further, we investigate the variance of the received signal power around its mean for different prediction horizon lengths, i.e., how channel prediction influences channel hardening. This facilitates the reliability analysis of a massive MIMO system in terms of received signal power fluctuation.

Scientific Contribution:

- We derive the instantaneous and asymptotical signal to interference and noise ratio (SINR) and the SE of a TDD massive MIMO system for a time-correlated fading process utilizing multi-step minimum mean square error (MMSE) (Wiener) prediction. This result generalizes the work of [3] for single-step AR(1) process prediction.
- We prove and show by numerical simulation that the capability of massive MIMO to reduce the variance of the received signal power is independent of the CSI age, i.e., channel hardening is independent of channel aging. The CSI age mainly affects the SINR and thereby the SE.

Notation: Throughout this work, letters in italic font (a) denote scalars. Bold lower case letters (\mathbf{a}) denote vectors, while bold upper case letters (\mathbf{A}) denote matrices. $(\cdot)^T$ marks a transpose and $(\cdot)^H$ marks the Hermitian transpose. The Kronecker product of two matrices \mathbf{A} , \mathbf{B} is denoted by $\mathbf{A} \otimes \mathbf{B}$. The element-wise expectation of a random vector or matrix \mathbf{a} is denoted as $\mathbb{E}\{\mathbf{a}\}$ and its variance is written as $\text{Var}\{\mathbf{a}\}$.

II. SIGNAL MODEL

We consider an uplink massive MIMO system, where K users with single-antenna terminals $k \in \{1, \dots, K\}$ send pilots and data to a BS deploying A antenna elements, without

Manuscript received ...
David Löschenbrand, Markus Hofer and Thomas Zemen are with the AIT Austrian Institute of Technology GmbH, Vienna, Austria (e-mail: david.loeschenbrand@ait.ac.at; markus.hofer@ait.ac.at, thomas.zemen@ait.ac.at)
Digital Object Identifier

neighboring cells. The notation we introduce for this system is largely based on [3], with the exception that only one cell is considered as our focus mainly lies on the aspect of time variance of the bandlimited fading process and its impact on the SINR and SE.

Due to space constraints we discuss the uplink (UL) only. However all results on SE apply for the downlink (DL) too, through the UL-DL duality [9, Theorem 4.8].

The vector collecting the received symbols from all K users at time index m is

$$\hat{\mathbf{x}}_m = \mathbf{W}_m^H \mathbf{H}_m \mathbf{x}_m + \frac{1}{\sqrt{P}} \mathbf{z}_m \in \mathbb{C}^{K \times 1}, \quad (1)$$

with P the average transmit power of each user, $\mathbf{x}_m \in \mathbb{C}^{K \times 1}$ the vector collecting the transmit symbols of all users, and $\frac{1}{\sqrt{P}} \mathbf{z}_m = \frac{1}{\sqrt{P}} \mathbf{W}_m^H \tilde{\mathbf{z}}_m \sim \mathcal{CN}(0, \frac{\sigma_z^2}{P} \mathbf{I}_K)$ denoting filtered complex Gaussian noise. The channel matrix

$$\mathbf{H}_m = [\mathbf{h}_{0,m}, \mathbf{h}_{1,m}, \dots, \mathbf{h}_{K-1,m}] \in \mathbb{C}^{A \times K} \quad (2)$$

contains the individual channel vectors of all users. Similarly, the beam-forming matrix

$$\mathbf{W}_m = [\mathbf{w}_{0,m}, \mathbf{w}_{1,m}, \dots, \mathbf{w}_{K-1,m}] \in \mathbb{C}^{A \times K} \quad (3)$$

contains the individual beam-forming vectors $\mathbf{w}_{k,m}$ for user k and all A BS antenna elements.

The spatial correlation matrix of the channel vector for user k is

$$\mathbf{R}_k = \mathbb{E} \{ \mathbf{h}_{k,m} \mathbf{h}_{k,m}^H \}. \quad (4)$$

Further, we assume local wide-sense stationarity of the fading process and a joint spatial and temporal covariance matrix

$$\mathbb{E} \{ \mathbf{h}_{k,m-\ell} \mathbf{h}_{k,m-\ell}^H \} = r_{\mathbf{h},|\ell-\ell'|} \mathbf{R}_k, \quad (5)$$

with $r_{\mathbf{h},\ell}$ being the temporal correlation coefficient of the channel vector with time delay ℓ .

Note that the scalar approach in (5) intrinsically assumes, that the temporal correlation is the same for all antenna pairs a, a' , which is valid if BS antenna elements are collocated.

The received symbol estimate from one single user k at the BS is

$$\hat{x}_{k,m} = \mathbf{w}_{k,m}^H \mathbf{h}_{k,m} x_{k,m} + \frac{1}{\sqrt{P}} z_{k,m} + \sum_{k' \neq k} \mathbf{w}_{k,m}^H \mathbf{h}_{k',m} x_{k',m}, \quad (6)$$

where the first term is the scaled signal, the second term is filtered and scaled Gaussian noise $\frac{1}{\sqrt{P}} z_{k,m} \sim \mathcal{CN}(0, \frac{\sigma_z^2}{P})$, and the third term is interference from other users $k' \neq k$.

A. Channel Estimation

For channel estimation, a single-cell approach similar to [3, (9)] is used to obtain noisy observations of the channel vector

$$\mathbf{y}_{k,m} = \mathbf{h}_{k,m} + \frac{1}{\sqrt{P}} \mathbf{z}_m. \quad (7)$$

They are filtered by the MMSE estimator [3] to obtain the channel vector estimate

$$\hat{\mathbf{h}}_{k,m} = \mathbf{R}_k \left(\frac{\sigma_z^2}{P} \mathbf{I}_A + \mathbf{R}_k \right)^{-1} \mathbf{y}_{k,m}, \quad (8)$$

distributed as $\hat{\mathbf{h}}_{k,m} \sim \mathcal{CN}(\mathbf{0}, \boldsymbol{\Theta}_{k,1,0})$ with

$$\boldsymbol{\Theta}_{k,1,0} = \mathbf{R}_k \left(\frac{\sigma_z^2}{P} \mathbf{I}_A + \mathbf{R}_k \right)^{-1} \mathbf{R}_k \quad (9)$$

denoting the covariance matrix.

B. Multi-Step Channel Prediction

Given the following matrices to ease notation

$$\boldsymbol{\delta}_{N,\ell} = [\mathbf{r}_{\mathbf{h},\ell}, \mathbf{r}_{\mathbf{h},\ell+1}, \dots, \mathbf{r}_{\mathbf{h},\ell+N-1}], \quad (10)$$

$$\boldsymbol{\Delta}_{N,\ell} = \begin{bmatrix} \mathbf{r}_{\mathbf{h},\ell} & \mathbf{r}_{\mathbf{h},\ell+1} & \dots & \mathbf{r}_{\mathbf{h},\ell+N-1} \\ \mathbf{r}_{\mathbf{h},\ell+1} & \mathbf{r}_{\mathbf{h},\ell} & \dots & \mathbf{r}_{\mathbf{h},\ell+N-2} \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{r}_{\mathbf{h},\ell+N-1} & \mathbf{r}_{\mathbf{h},\ell+N-2} & \dots & \mathbf{r}_{\mathbf{h},\ell} \end{bmatrix}, \quad (11)$$

$$\mathbf{T}_{k,N,\ell} = \left[\boldsymbol{\Delta}_{N,\ell} \otimes \mathbf{R}_k + \frac{\sigma_z^2}{P} \mathbf{I}_{AN} \right]^{-1}, \quad (12)$$

$$\boldsymbol{\Theta}_{k,N,\ell} = (\boldsymbol{\delta}_{N,\ell} \otimes \mathbf{R}_k) \mathbf{T}_{k,N,0} (\boldsymbol{\delta}_{N,\ell} \otimes \mathbf{R}_k)^H, \quad (13)$$

we find the linear Wiener multi-step predictor by considering the framework provided in [3] and the joint spatial and temporal correlation defined in (5).

Theorem II.1 (Wiener multi-step predictor). *Given N consecutive pilot symbols, the optimal linear Wiener predictor $\mathbf{V}_{k,\ell}$ with prediction horizon ℓ is*

$$\mathbf{V}_{k,\ell} = (\boldsymbol{\delta}_{N,\ell} \otimes \mathbf{R}_k) \mathbf{T}_{k,N,0}, \quad (14)$$

and the predicted channel vector $\tilde{\mathbf{h}}_{k,m+1}$ is

$$\tilde{\mathbf{h}}_{k,m+1} = \mathbf{V}_{k,\ell+1} \tilde{\mathbf{y}}_{k,m-\ell}, \quad (15)$$

with

$$\tilde{\mathbf{y}}_{k,m-\ell} = \begin{bmatrix} \mathbf{y}_{k,m-\ell} \\ \mathbf{y}_{k,m-\ell-1} \\ \vdots \\ \mathbf{y}_{k,m-\ell-N+1} \end{bmatrix} \in \mathbb{C}^{AN \times 1} \quad (16)$$

stacking N received pilot symbols delayed by ℓ .

Proof. We follow the approach in [3], but extend it to a general temporal fading model and multi-step prediction. We assume that N consecutive pilot symbols are used for prediction.

Suppose that the pilot symbols are outdated, e.g., only pilot symbols $\tilde{\mathbf{y}}_{k,m-\ell}$ with CSI acquisition delay $\ell > 0$ are available to predict $\mathbf{h}_{k,m+1}$. The optimal predictor [3] extended to multi-step prediction is then found by solving

$$\mathbb{E} \{ (\mathbf{h}_{k,m+1} - \mathbf{V}_{k,\ell+1} \tilde{\mathbf{y}}_{k,m-\ell}) \tilde{\mathbf{y}}_{k,m-\ell}^H \} = \mathbf{0} \quad (17)$$

or, after reformulating,

$$\mathbf{V}_{k,\ell+1} = \mathbb{E} \{ \mathbf{h}_{k,m+1} \tilde{\mathbf{y}}_{k,m-\ell}^H \} \mathbb{E} \{ \tilde{\mathbf{y}}_{k,m-\ell} \tilde{\mathbf{y}}_{k,m-\ell}^H \}^{-1}. \quad (18)$$

By using the channel estimation signal model in (7), the joint spatial and temporal correlation (5), and assuming that uncorrelated noise terms cancel for $\ell > 0$, we find the covariance matrix of two received pilot symbols with CSI acquisition delay ℓ to be

$$\mathbb{E} \{ \mathbf{y}_{k,m} \mathbf{y}_{k,m-\ell}^H \} = \begin{cases} \mathbf{r}_{\mathbf{h},0} \mathbf{R}_k + \frac{\sigma_z^2}{P}, & \ell = 0; \\ \mathbf{r}_{\mathbf{h},\ell} \mathbf{R}_k, & \text{otherwise.} \end{cases} \quad (19)$$

The covariance matrix of the stacked pilot symbols is then calculated with (16) and (19) as

$$\begin{aligned} \mathbb{E} \{ \tilde{\mathbf{y}}_{k,m-\ell} \tilde{\mathbf{y}}_{k,m-\ell}^H \} &= \mathbb{E} \{ \tilde{\mathbf{y}}_{k,m} \tilde{\mathbf{y}}_{k,m}^H \} \\ &= \mathbf{\Delta}_{N,0} \otimes \mathbf{R}_k + \frac{\sigma_z^2}{P} \mathbf{I}_{AN} = \mathbf{T}_{k,N,0}^{-1}. \end{aligned} \quad (20)$$

The first line of (20) holds since wide-sense stationarity is assumed, i.e., the statistics are independent of an arbitrary time shift ℓ for the considered frame duration and prediction horizon [10].

With (7), (5) and (16), the cross-correlation between the true channel and the stacked pilot symbols is

$$\begin{aligned} \mathbb{E} \{ \mathbf{h}_{k,m+1} \tilde{\mathbf{y}}_{k,m-\ell}^H \} &= [\mathbf{r}_{h,\ell+1} \mathbf{R}_k, \dots, \mathbf{r}_{h,\ell+N} \mathbf{R}_k] \\ &= \boldsymbol{\delta}_{N,\ell+1} \otimes \mathbf{R}_k. \end{aligned} \quad (21)$$

Substituting (20) and (21) into (18) gives (14).

The predicted channel vector is thus

$$\tilde{\mathbf{h}}_{k,m+1} = \mathbf{V}_{k,\ell+1} \tilde{\mathbf{y}}_{k,m-\ell}, \quad (22)$$

and its distribution $\tilde{\mathbf{h}}_{k,m+1} \sim \mathcal{CN}(\mathbf{0}, \boldsymbol{\Theta}_{k,N,\ell+1})$ depends on the prediction horizon ℓ [3]. \square

Note that in the case of $N = 1$, i.e., only one pilot symbol is used, (10) and (11) each reduce to the scalar temporal correlation coefficient and the predicted channel becomes $\tilde{\mathbf{h}}_{k,m+1} = \mathbf{V}_{k,\ell+1} \tilde{\mathbf{y}}_{k,m-\ell} = r_{h,\ell} \hat{\mathbf{h}}_{k,m-\ell+1}$. Simply put, with only one pilot at hand, the evolution of the channel cannot be tracked and the best prediction is the weighted aged channel estimate.

III. SPECTRAL EFFICIENCY AND CHANNEL HARDENING

By introducing the predicted channel $\tilde{\mathbf{h}}_{k,m+1}$ in (6), we obtain the estimated received symbol from one single user k at the BS after receive combining as

$$\begin{aligned} \hat{x}_{k,m+1} &= \mathbf{w}_{k,m+1}^H \tilde{\mathbf{h}}_{k,m+1} x_{k,m+1} + \frac{1}{\sqrt{P}} \tilde{z}_{k,m+1} \\ &+ \mathbf{w}_{k,m+1}^H (\mathbf{h}_{k,m+1} - \tilde{\mathbf{h}}_{k,m+1}) x_{k,m+1} \\ &+ \sum_{k' \neq k} \mathbf{w}_{k,m+1}^H \mathbf{h}_{k',m+1} x_{k',m+1}. \end{aligned} \quad (23)$$

The first term is the intended signal multiplied with the effective channel gain

$$\gamma_{k,m+1} = \mathbf{w}_{k,m+1}^H \tilde{\mathbf{h}}_{k,m+1}, \quad (24)$$

the second term is additive noise, the third term is the channel prediction error considered as noise, and the fourth term is interference from other users $k' \neq k$.

The choice of the beam-forming vectors in \mathbf{W}_m greatly affects the capability of a massive MIMO system to suppress the interference from other users. In this work, we revert to maximum ratio combining (MRC) and regularized zero-forcing (RZF) as common choices of beam-forming vectors [1], [4]:

$$\mathbf{W}_m = \begin{cases} \tilde{\mathbf{H}}_m, & \text{MRC;} \\ \tilde{\mathbf{H}}_m (\tilde{\mathbf{H}}_m^H \tilde{\mathbf{H}}_m + \frac{\sigma_z^2}{P} \mathbf{I}_K)^{-1}, & \text{RZF;} \end{cases} \quad (25)$$

where the columns of the predicted channel matrix $\tilde{\mathbf{H}}_m$ are composed of the predicted channel vectors similar to (2).

A. Instantaneous and Asymptotic SINR

Considering the signal model (23) and assuming channel prediction with prediction horizon ℓ at the BS, the instantaneous signal power is given as the squared absolute value of the effective channel gain $\gamma_{k,m+1}$ and yields

$$S_{k,\ell} = |\gamma_{k,m+1}|^2 = |\mathbf{w}_{k,m+1}^H \tilde{\mathbf{h}}_{k,m+1}|^2. \quad (26)$$

The instantaneous interference and noise power is defined as the sum of all interference/noise contributions in (23)

$$\begin{aligned} I_{k,\ell} &= \frac{\sigma_z^2}{\sqrt{P}} \mathbf{w}_{k,m+1}^H \tilde{\mathbf{h}}_{k,m+1} \\ &+ |\mathbf{w}_{k,m+1}^H (\mathbf{h}_{k,m+1} - \tilde{\mathbf{h}}_{k,m+1})|^2 \\ &+ \sum_{k' \neq k} |\mathbf{w}_{k,m+1}^H \mathbf{h}_{k',m+1}|^2. \end{aligned} \quad (27)$$

The instantaneous SINR is defined as the ratio of instantaneous signal to interference and noise power

$$\eta_{k,\ell} = \frac{S_{k,\ell}}{I_{k,\ell}}. \quad (28)$$

Theorem III.1. *The asymptotic deterministic SINR $\bar{\eta}_{k,\ell} = \lim_{A \rightarrow \infty} \eta_{k,\ell}$ with MRC is calculated as seen in (29).*

Proof. Since MRC is assumed, the beamforming vector is computed as $\mathbf{w}_{k,m+1} = \tilde{\mathbf{h}}_{k,m+1}$. For $A \rightarrow \infty$, the instantaneous signal power (26) converges towards its asymptotic deterministic equivalent [3]

$$\begin{aligned} \bar{S}_{k,\ell} &= \lim_{A \rightarrow \infty} \frac{1}{A^2} S_{k,\ell} \\ &= |\text{tr}\{\boldsymbol{\Theta}_{k,N,\ell+1}\}|^2, \end{aligned} \quad (30)$$

with the covariance matrix $\boldsymbol{\Theta}_{k,N,\ell+1}$ defined in (13). Similarly, the instantaneous interference and noise power (27) converges toward its asymptotic deterministic equivalent [3]

$$\begin{aligned} \bar{I}_{k,\ell} &= \lim_{A \rightarrow \infty} \frac{1}{A^2} I_{k,\ell} \\ &= \frac{\sigma_z^2}{\sqrt{P}} \text{tr}\{\boldsymbol{\Theta}_{k,N,\ell+1}\} \\ &+ \text{tr}\{\boldsymbol{\Theta}_{k,N,\ell+1} (\mathbf{R}_k - \boldsymbol{\Theta}_{k,N,\ell+1})\} \\ &+ \sum_{k' \neq k} \text{tr}\{\boldsymbol{\Theta}_{k,N,\ell+1} \mathbf{R}_{k'}\} \end{aligned} \quad (31)$$

as the number of BS antenna elements grows.

The asymptotic deterministic SINR is therefore found as the ratio $\bar{S}_{k,\ell}/\bar{I}_{k,\ell}$ as specified in (29). \square

B. Ergodic and Asymptotic Spectral Efficiency

Similar to [3], we define the uplink ergodic achievable SE of user k with prediction horizon ℓ as

$$R_{k,\ell} = \mathbb{E} \{ \log_2 (1 + \eta_{k,\ell}) \}, \quad (32)$$

where the expectation is over channel realizations. In the asymptotic case with the number of BS antenna elements $A \rightarrow \infty$, the SINR is deterministic and the uplink asymptotic achievable SE of user k with prediction horizon ℓ is [3]

$$\bar{R}_{k,\ell} = \log_2 (1 + \bar{\eta}_{k,\ell}). \quad (33)$$

$$\bar{\eta}_{k,\ell} = \frac{\bar{S}_{k,\ell}}{\bar{I}_{k,\ell}} = \frac{|\text{tr}\{\mathbf{\Theta}_{k,N,\ell+1}\}|^2}{\frac{\sigma_z^2}{\sqrt{P}} \text{tr}\{\mathbf{\Theta}_{k,N,\ell+1}\} + \text{tr}\{\mathbf{\Theta}_{k,N,\ell+1}(\mathbf{R}_k - \mathbf{\Theta}_{k,N,\ell+1})\} + \sum_{k' \neq k} \text{tr}\{\mathbf{\Theta}_{k,N,\ell+1} \mathbf{R}_{k'}\}} \quad (29)$$

C. Channel Hardening

Channel hardening refers to the phenomenon of the effective channel gain $\gamma_{k,m+1} = \mathbf{w}_{k,m+1}^H \hat{\mathbf{h}}_{k,m+1}$ in (24) fluctuating only marginally around its mean [11],

$$\frac{\gamma_{k,m+1}}{\mathbb{E}\{\gamma_{k,m+1}\}} \rightarrow 1, \text{ as } A \rightarrow \infty. \quad (34)$$

To assess if channel hardening is occurring, we investigate the ratio of the effective gains variance to its squared mean

$$\beta_{k,\ell} = \frac{\text{Var}\{\gamma_{k,m+1}\}}{\mathbb{E}\{\gamma_{k,m+1}\}^2}, \quad (35)$$

similar to [11]–[13]. Low values of $\beta_{k,\ell}$ indicate a low probability of the channel gain $\gamma_{k,m+1}$ deviating significantly from its mean.

Theorem III.2 (Channel hardening). *Let $N = 1$ and $\mathbf{w}_{k,m+1} = \mathbf{r}_{\mathbf{h},\ell+1} \hat{\mathbf{h}}_{k,m-\ell}$ (i.e., MRC is applied), then the channel hardening metric $\beta_{k,\ell}$ in (35) is independent of the prediction horizon ℓ .*

Proof. We acknowledge that, for $N = 1$, the predicted channel vector $\hat{\mathbf{h}}_{k,m+1}$ is the outdated channel estimate decreased by its temporal correlation coefficient, i.e., $\hat{\mathbf{h}}_{k,m+1} = r_{\mathbf{h},\ell} \hat{\mathbf{h}}_{k,m-\ell+1}$. Further, by applying MRC, the beam-forming vector becomes $\mathbf{w}_{k,m+1} = r_{\mathbf{h},\ell} \hat{\mathbf{h}}_{k,m-\ell+1}$.

We then calculate

$$\begin{aligned} \mathbb{E}\{\gamma_{k,m+1}\} &= \mathbb{E}\left\{\mathbf{w}_{k,m+1}^H \hat{\mathbf{h}}_{k,m+1}\right\} \\ &= r_{\mathbf{h},\ell}^2 \mathbb{E}\left\{\hat{\mathbf{h}}_{k,m-\ell+1}^H \hat{\mathbf{h}}_{k,m-\ell+1}\right\} \\ &= r_{\mathbf{h},\ell}^2 \text{tr}\{\mathbf{\Theta}_{k,1,0}\}, \end{aligned} \quad (36)$$

where the third line follows from (8), and

$$\begin{aligned} \text{Var}\{\gamma_{k,m+1}\} &= \mathbb{E}\left\{\left(\mathbf{w}_{k,m+1}^H \hat{\mathbf{h}}_{k,m+1}\right)^2 - \mathbb{E}\left\{\mathbf{w}_{k,m+1}^H \hat{\mathbf{h}}_{k,m+1}\right\}^2\right\} \\ &= r_{\mathbf{h},\ell}^4 \mathbb{E}\left\{\left(\hat{\mathbf{h}}_{k,m-\ell+1}^H \hat{\mathbf{h}}_{k,m-\ell+1}\right)^2 - \text{tr}\{\mathbf{\Theta}_{k,1,0}\}^2\right\} \\ &= r_{\mathbf{h},\ell}^4 \mathbb{E}\left\{\left(\hat{\mathbf{h}}_{k,m}^H \hat{\mathbf{h}}_{k,m}\right)^2 - \text{tr}\{\mathbf{\Theta}_{k,1,0}\}^2\right\}. \end{aligned} \quad (37)$$

We plug (36) and (37) into (35) to get

$$\beta_{k,\ell} = \frac{\mathbb{E}\left\{\left(\hat{\mathbf{h}}_{k,m}^H \hat{\mathbf{h}}_{k,m}\right)^2 - \text{tr}\{\mathbf{\Theta}_{k,1,0}\}^2\right\}}{\text{tr}\{\mathbf{\Theta}_{k,1,0}\}^2}, \quad (38)$$

which is independent of the prediction horizon ℓ . \square

While $\beta_{k,\ell}$ is time-dependent in general when employing RZF, we show empirically in the next section that its variation is negligible compared to the impact of the number of BS antenna elements on channel hardening.

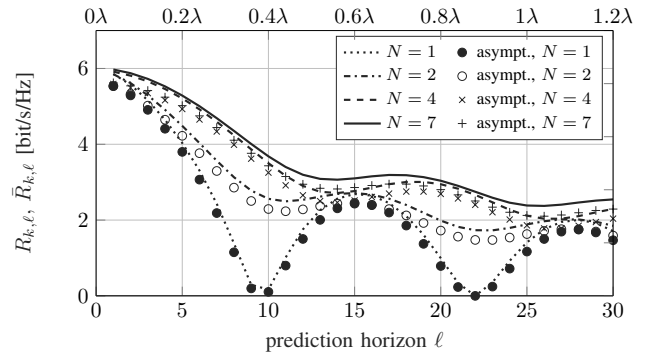


Fig. 1. Ergodic and asymptotic achievable SE $R_{k,\ell}$ over prediction horizon ℓ for a normalized Doppler $f_D T_s = 0.04$, an SNR $\frac{P}{\sigma_z^2} = 12$ dB, MRC beam-forming, and $N \in [1, 2, 4, 7]$ pilots. The distance in terms of the wavelength λ corresponding to a given prediction horizon is given as second x-axis.

IV. RESULTS

We use a Monte-Carlo simulation to verify the uplink ergodic and asymptotic SE and channel hardening results under the influence of aged CSI and multi-step Wiener prediction. The expectation and variance in (32) and (35) are numerically estimated as the mean over random channel realizations. The random variable realizations are drawn according to the distributions specified in Sec. II. The number of runs for the simulation is 20000.

Channel vector realizations with proper statistics in (6) are calculated using the approach outlined in [14], [15] with 400 paths and one tap. We consider Clarke's model $r_{\mathbf{h},\ell} = J_0(2\pi f_D T_s \ell)$ for the temporal correlation coefficient, where J_0 denotes the Bessel function of the first kind and $f_D T_s$ denotes the maximum Doppler shift f_D normalized by the symbol duration T_s . This is more suitable for modeling a band limited fading processes, compared to the unbounded Doppler spectrum of the AR(1) model used in [3]. It also provides a direct relation between correlation in time and maximum Doppler shift. In the simulations, we set the normalized Doppler shift to $f_D T_s = 0.04$, which corresponds to a velocity of 172 km/h at a carrier frequency of 3.5 GHz (wavelength $\lambda = 8.6$ cm) and a symbol duration $T_s = 71.4 \mu\text{s}$. The number of BS antenna elements is $A = 64$ unless stated otherwise.

We note an excellent match between the SE in (32), simulated with $A = 64$ BS antenna elements, and the asymptotic SE in (33) for MRC beam-forming, shown in Fig. 1. Channel prediction with a higher number of pilots $N > 1$ increases the SE considerably and prevents sudden drops (caused by the temporal correlation function $r_{\mathbf{h},\ell}$ approaching zero). For a prediction horizon $\ell = 7$ (equivalent to a movement of the user by 0.28λ since CSI acquisition), utilizing four instead of one pilot symbol doubles the achievable SE from 2.2 bit/(s Hz) to 4.3 bit/(s Hz).

Figure 2 shows significant achievable SE improvements in the most relevant region $\ell \leq 10$ when choosing RZF over MRC. For a prediction horizon $\ell = 7$, utilizing four

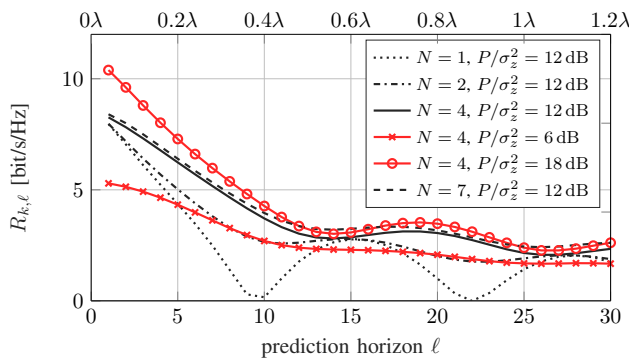


Fig. 2. SE $R_{k,\ell}$ over prediction horizon ℓ for a normalized Doppler $f_D T_s = 0.04$, RZF beam-forming, and $N \in [1, 2, 4, 7]$ pilots. The distance in terms of the wavelength λ corresponding to a given prediction horizon is given as second x -axis.

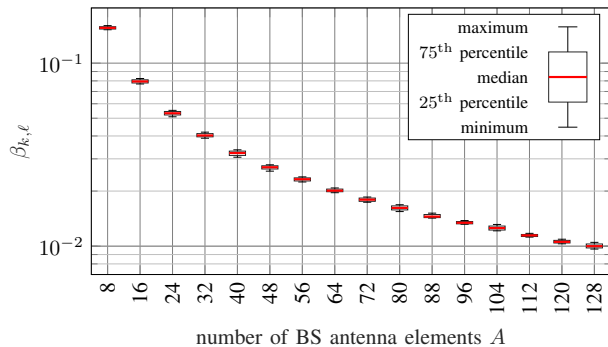


Fig. 3. Boxplot of $\beta_{k,\ell}$ distribution over prediction horizon $\ell \in \{0, \dots, 30\}$ for different number of BS antennas, $A \in [8, 128]$. The normalized Doppler $f_D T_s = 0.04$, the SNR $P/\sigma_z^2 = 12$ dB, the number of pilots $N = 4$, and the beam-forming method is RZF. The small variance of $\beta_{k,\ell}$ demonstrates that channel hardening is basically independent of the channel age.

instead of one pilot symbol increases the achievable SE from 2.5 bit/(s Hz) to 5.2 bit/(s Hz). We also note that for $\ell \geq 5$, the SE with an SNR of 6 dB and $N = 4$ is superior to the SE with an SNR of 12 dB and $N = 1$, highlighting the importance of channel prediction.

Figure 3 shows a boxplot of the distribution of $\beta_{k,\ell}$ over the prediction horizon ℓ for different numbers of BS antenna elements $A \in [8, 128]$. We verify that channel hardening is indeed largely independent of the prediction horizon ℓ and is mostly determined by the number of BS antenna elements.

V. CONCLUSION

In this paper, we addressed the SE of time-variant massive MIMO using Wiener prediction. We derived asymptotic expressions of the achievable SE in an uplink massive MIMO system for an arbitrary prediction horizon $\ell \geq 1$, assuming a general temporal covariance matrix. Our results are also directly applicable to the downlink, due to UL-DL duality.

Given a prediction horizon of 0.3λ , utilizing four instead of one pilot symbol is shown to double the SE, both for MRC and RZF. Utilizing more than four pilots shows no clear advantages.

Further, channel hardening, i.e., the capability of a massive MIMO system to eliminate small-scale fading and create quasi-deterministic effective channel gains, is shown to be (largely) independent of channel aging.

Our work allows to assess effects of channel aging due to mobility, and channel prediction as mitigation strategy, for arbitrary time delays.

ACKNOWLEDGEMENT

This work was supported by the project DEDICATE (Principal Scientist grant) at the AIT Austrian Institute of Technology. The authors acknowledge the fruitful discussions in COST Action CA20120 INTERACT.

REFERENCES

- [1] T. L. Marzetta, "Noncooperative cellular wireless with unlimited numbers of base station antennas," *IEEE Trans. Wireless Commun.*, vol. 9, no. 11, pp. 3590–3600, 2010.
- [2] D. Löschenbrand, M. Hofer, L. Bernadó, S. Zelenbaba, and T. Zemen, "Towards cell-free massive MIMO: A measurement-based analysis," *IEEE Access*, vol. 10, pp. 89 232–89 247, 2022.
- [3] K. T. Truong and R. W. Heath, "Effects of channel aging in massive MIMO systems," *J. Commun. Netw.*, vol. 15, no. 4, pp. 338–351, 2013.
- [4] 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, 2015.
- [5] S. Kashyap, C. Mollén, E. Björnson, and E. G. Larsson, "Performance analysis of (TDD) massive MIMO with Kalman channel prediction," in *2017 IEEE Int. Conf. Acoustics, Speech and Signal Process. (ICASSP)*, 2017, pp. 3554–3558.
- [6] H. Kim, S. Kim, H. Lee, C. Jang, Y. Choi, and J. Choi, "Massive MIMO channel prediction: Kalman filtering vs. machine learning," *IEEE Trans. Commun.*, vol. 69, no. 1, pp. 518–528, 2021.
- [7] C. Wu, X. Yi, Y. Zhu, W. Wang, L. You, and X. Gao, "Channel prediction in high-mobility massive MIMO: From spatio-temporal autoregression to deep learning," *IEEE J. Sel. Areas Commun.*, vol. 39, no. 7, pp. 1915–1930, 2021.
- [8] A. K. Papazafeiropoulos, "Impact of general channel aging conditions on the downlink performance of massive MIMO," *IEEE Trans. Veh. Technol.*, vol. 66, no. 2, pp. 1428–1442, 2016.
- [9] E. Björnson, J. Hoydis, and L. Sanguinetti, "Massive MIMO Networks: Spectral, Energy, and Hardware Efficiency," *Foundations and Trends® in Signal Processing*, vol. 11, no. 3–4, pp. 154–655, 2017. [Online]. Available: <http://www.nowpublishers.com/article/Details/SIG-093>
- [10] L. Bernadó, T. Zemen, F. Tufvesson, A. F. Molisch, and C. F. Mecklenbrauker, "Delay and Doppler spreads of nonstationary vehicular channels for safety-relevant scenarios," *IEEE Trans. Veh. Technol.*, vol. 63, no. 1, pp. 82–93, Jan. 2014.
- [11] H. Q. Ngo and E. G. Larsson, "No downlink pilots are needed in TDD massive MIMO," *IEEE Trans. Wireless Commun.*, vol. 16, no. 5, pp. 2921–2935, 2017.
- [12] S. Gunnarsson, J. Flordelis, L. Van der Perre, and F. Tufvesson, "Channel hardening in massive MIMO – a measurement based analysis," in *2018 IEEE 19th Int. Workshop Signal Process. Advances Wireless Commun. (SPAWC)*. IEEE, 2018, pp. 1–5.
- [13] T. Zemen, D. Löschenbrand, M. Hofer, C. Pacher, and B. Rainer, "Orthogonally precoded massive MIMO for high mobility scenarios," *IEEE Access*, vol. 7, pp. 132 979–132 990, 2019.
- [14] Y. R. Zheng and C. Xiao, "Simulation models with correct statistical properties for Rayleigh fading channels," *IEEE Trans. Commun.*, vol. 51, no. 6, pp. 920–928, 2003.
- [15] T. Zemen and C. F. Mecklenbrauker, "Time-variant channel estimation using discrete prolate spheroidal sequences," *IEEE Trans. Signal Process.*, vol. 53, no. 9, pp. 3597–3607, 2005.