

Machine Learning-based Channel Prediction for Widely Distributed Massive MIMO with Real-World Data

David Löschenbrand*, Markus Hofer*, Lukas Eller^{†‡}, Markus Rupp[†], Thomas Zemen*

*Center for Digital Safety & Security, AIT Austrian Institute of Technology GmbH, Austria

[†]Institute of Telecommunications, TU Wien, Austria

[‡]Christian Doppler Laboratory for Digital Twin Assisted AI for Sustainable Radio Access Networks, Institute of Telecommunications, TU Wien, Austria
Email: david.loeschenbrand@ait.ac.at

Abstract—Widely distributed massive multiple input multiple output (WD-MIMO) systems are promising candidates for future mobile networks, given their improved energy efficiency, coverage and throughput. To spatially separate the users, WD-MIMO relies heavily on accurate and timely channel state information (CSI), which is hard to obtain in high mobility scenarios. To reduce the amount of pilot overhead necessary for obtaining CSI, we investigate linear and machine learning (ML)-based CSI prediction techniques and compare them in terms of achievable spectral efficiency (SE). The considered methods are constant continuation, Wiener prediction, dense, and long short term memory (LSTM) neural networks (NNs). Real-world data from a widely distributed massive MIMO channel measurement campaign with various base station (BS) antenna array aperture sizes is utilized for NN training and validation purposes. The capability of the considered CSI prediction methods to mitigate the effects of channel aging in realistic high-mobility scenarios is analyzed for different geometries of the massive MIMO BS antenna arrays. We can demonstrate a SE improvement of 2 bit/s/Hz for the LSTM NN compared to a Wiener predictor.

Index Terms—AI/ML, CSI prediction, massive MIMO, widely distributed, spectral efficiency (SE)

I. INTRODUCTION

While fifth generation (5G) wireless communication systems are currently being deployed and in use world-wide, the search for and evaluation of 6G candidate technology solutions already started in the scientific community. One of the most promising concepts is widely distributed (or cell-free) massive multiple input multiple output (MIMO) systems [1], which dissolves the cellular network structure and follows a user-centric approach, with different subsets of distributed base station (BS) antennas (also called access points) serving different users in an optimized and cooperative way. This approach offers several advantages over conventional massive MIMO, such as (i) a strong transmit energy reduction due to reduced distance to the user equipment (UE), (ii) a consistent throughput over the coverage area avoiding the strong throughput drop at cell edges, and (iii) mitigation of large-scale fading [2].

For conventional and widely distributed massive MIMO systems alike, accurate channel state information (CSI) at the BS side is a key requirement to avoid degraded performance for

mobile users [3] when operating in time division duplex (TDD) mode. Hence, robust channel prediction algorithms, exploiting the sounding reference signals (SRSs), are of strong interest to avoid the expiration of the CSI. Moreover, these prediction algorithms must be tested with real-world data to verify their robustness.

The effects of outdated CSI on the spectral efficiency (SE) of a communication system and mitigation strategies using prediction are thoroughly investigated in, e.g., [4]–[7]. However, no real-world data is used in the analysis. There have been many recent studies on high mobility users in massive MIMO systems that utilize advanced prediction methods, such as Kalman filtering [8], [9] and machine learning (ML) [9], [10]. However, the methods are validated using only simulated data.

To the best of the authors knowledge, [11] shows the first validation of a ML based CSI predictor using real-world data. However, only one conventional collocated BS antenna configuration is considered.

In this paper, we provide the following scientific contributions:

- We present two ML algorithms based on a dense neural networks (NNs) and long short term memory (LSTM) architecture for predicting the CSI from uplink data for the downlink transmission.
- Training and testing is performed with CSI obtained from the AIT distributed massive MIMO testbed with 32 antenna elements spread in different patterns over an aperture of 46.5 m.
- We present a comparison in terms of SE for linear prediction algorithms and ML based methods, demonstrating a SE improvement of 2 bit/s/Hz for the LSTM NN compared to a Wiener predictor.

II. SIGNAL MODEL

The notation we use is largely based on [7]. We consider an uplink massive MIMO system, where K users $k \in \{1, \dots, K\}$ with single-antenna terminals send pilots and data to a BS deploying A antenna elements $a \in \{1, \dots, A\}$, without neighboring cells. We assume an orthogonal frequency-division

multiplexing (OFDM) system where the individual subcarriers q are orthogonal and are therefore treated individually. The dependency on frequency is thus omitted in the considerations below to facilitate notation.

The effects of the wireless channel on the scalar transmitted information symbol $s_{k,m}$ at time index m are modelled by a complex scalar channel coefficient $h_{k,a,m}$. The channel coefficients of one user k to all A BS antennas are collected in a channel vector

$$\mathbf{h}_{k,m} = [h_{k,1,m} \ h_{k,2,m} \ \dots \ h_{k,A,m}]^T \in \mathbb{C}^{A \times 1}. \quad (1)$$

The channel matrix

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

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

$$\mathbf{W}_m = [\mathbf{w}_{1,m} \ \mathbf{w}_{2,m} \ \dots \ \mathbf{w}_{K,m}] \in \mathbb{C}^{A \times K} \quad (3)$$

contains the individual beam-forming vectors $\mathbf{w}_m \in \mathbb{C}^{A \times 1}$ for user k and all A BS antenna elements.

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

$$\hat{\mathbf{s}}_m = \mathbf{W}_m^H \mathbf{y}_m = \mathbf{W}_m^H \mathbf{H}_m \mathbf{s}_m + \mathbf{W}_m^H \mathbf{n}_m, \quad (4)$$

with $\mathbf{s}_m \in \mathbb{C}^{K \times 1}$ the vector collecting the transmitted information symbols of all users, and $\mathbf{W}_m^H \mathbf{n}_m \sim \mathcal{CN}(0, \frac{\sigma^2}{P} \mathbf{I}_K)$ denoting filtered complex Gaussian noise. P is the average transmit power of each user.

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

$$\begin{aligned} \hat{s}_{k,m} &= \mathbf{w}_{k,m}^H \mathbf{y}_m = \mathbf{w}_{k,m}^H \mathbf{H}_m \mathbf{s}_m + \mathbf{w}_{k,m}^H \mathbf{n}_m \\ &= \mathbf{w}_{k,m}^H \mathbf{h}_{k,m} s_{k,m} + \mathbf{w}_{k,m}^H \mathbf{n}_m + \sum_{k' \neq k} \mathbf{w}_{k,m}^H \mathbf{h}_{k',m} s_{k',m}, \end{aligned} \quad (5)$$

where the first term in the second line is the scaled signal, the second term is filtered and scaled Gaussian noise $\mathbf{w}_{k,m}^H \mathbf{n}_m \sim \mathcal{CN}(0, \frac{\sigma^2}{P})$, and the third term is interference from other users $k' \neq k$.

A. Multi-Step CSI Prediction

It is shown theoretically, by simulation, and with measured data in numerous papers (e.g., [4], [7], [11], [12]) that high mobility of the users causes severe degradation of SE in massive MIMO systems, often referred to as *channel aging*. It is caused by the acquired CSI at the BS being outdated when it is used for down-link beam-forming. To mitigate this effect, CSI prediction is a proven approach [13].

Given N consecutive pilot symbols and assuming perfect channel estimation, we obtain the predicted channel vector $\tilde{\mathbf{h}}_{k,m+1}$ with prediction horizon ℓ as

$$\tilde{\mathbf{h}}_{k,m+1} = f(\tilde{\mathbf{h}}_{k,m-\ell}), \quad (6)$$

with f the prediction function and

$$\tilde{\mathbf{h}}_{k,m-\ell} = \begin{bmatrix} \mathbf{h}_{k,m-\ell} \\ \mathbf{h}_{k,m-\ell-1} \\ \vdots \\ \mathbf{h}_{k,m-\ell-N+1} \end{bmatrix} \in \mathbb{C}^{AN \times 1}. \quad (7)$$

stacking N consecutively received pilot.

We consider the following methods to predict the CSI in this work:

- constant continuation, i.e., $\tilde{\mathbf{h}}_{k,m+1} = \mathbf{h}_{k,m-\ell}$,
- Wiener prediction as introduced in [7], and
- ML approaches that will be detailed in Section III, specifically:
 - a small dense network with one hidden layer and four nodes in total that serves as baseline algorithm, and
 - a more complex LSTM-based network.

For each of this methods, we evaluate the prediction quality by evaluating the achievable SE.

B. Achievable Spectral Efficiency for Multiple Users

By introducing the predicted channel $\tilde{\mathbf{h}}_{k,m+1}$ in (5), we obtain the estimated received symbol from one single user k at the BS after receive combining as

$$\begin{aligned} \hat{s}_{k,m+1} &= \mathbf{w}_{k,m+1}^H \tilde{\mathbf{h}}_{k,m+1} s_{k,m+1} + \mathbf{w}_{k,m}^H \mathbf{n}_{m+1} \\ &\quad + \mathbf{w}_{k,m+1}^H (\mathbf{h}_{k,m+1} - \tilde{\mathbf{h}}_{k,m+1}) s_{k,m+1} \\ &\quad + \sum_{k' \neq k} \mathbf{w}_{k,m+1}^H \mathbf{h}_{k',m+1} s_{k',m+1}. \end{aligned} \quad (8)$$

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

$$\gamma_{k,m+1} = \mathbf{w}_{k,m+1}^H \tilde{\mathbf{h}}_{k,m+1}, \quad (9)$$

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 utilize regularized zero-forcing (RZF) as a common choice of beam-forming vectors [5], [14]:

$$\begin{aligned} \mathbf{W}_m &= [\mathbf{w}_{1,m} \ \mathbf{w}_{2,m} \ \dots \ \mathbf{w}_{K,m}] \\ &= \tilde{\mathbf{H}}_m \left(\tilde{\mathbf{H}}_m^H \tilde{\mathbf{H}}_m + \frac{\sigma^2}{P} \mathbf{I}_K \right)^{-1}, \end{aligned} \quad (10)$$

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

Considering the signal model (8) 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

$$S_{k,\ell} = |\gamma_{k,m+1}|^2 = \left| \mathbf{w}_{k,m+1}^H \tilde{\mathbf{h}}_{k,m+1} \right|^2. \quad (11)$$

The instantaneous interference and noise power is defined as the sum of all interference/noise contributions in (8)

$$I_{k,\ell} = \frac{\sigma^2}{\sqrt{P}} \mathbf{w}_{k,m+1}^H \tilde{\mathbf{h}}_{k,m+1} + \left| \mathbf{w}_{k,m+1}^H (\mathbf{h}_{k,m+1} - \tilde{\mathbf{h}}_{k,m+1}) \right|^2 + \sum_{k' \neq k} \left| \mathbf{w}_{k,m+1}^H \mathbf{h}_{k',m+1} \right|^2. \quad (12)$$

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

$$\eta_{k,\ell} = \frac{S_{k,\ell}}{I_{k,\ell}} \quad (13)$$

Similar to [4], we define the uplink ergodic achievable SE of user k with prediction horizon ℓ as

$$SE_{k,\ell} = \mathbb{E} \{ \log_2 (1 + \eta_{k,\ell}) \}, \quad (14)$$

where the expectation is over channel realizations.

III. CSI PREDICTION WITH NNs

In this section, we introduce two NN-based models for CSI prediction of measured data to mitigate channel aging. For both models, the input tensors consist of batches of dimension $N \times A \times 2Q$, where N is the number of consecutive pilot symbols, A is the number of BS antennas, and Q is the total number of subcarriers considered. We split the pilot symbols in real and imaginary part and append them in the last dimension of the input tensor, hence its last dimension is $2Q$.

A. Baseline Dense Algorithm

This model is chosen in the most simplistic form to serve as a baseline for other CSI prediction methods. It consists of a reshaping layer to manipulate the dimensions of the input tensor to fit the subsequent dense layer. The dense layer consists of only N fully connected nodes and comprises the only trainable parameters in the model. The following reshaping layer serves the purpose of restoring the output tensor dimensions to match the input dimensions. Note that each of the A antenna samples is processed individually, therefore correlation among BS antennas can inherently not be exploited. Figure 1 shows the architecture of the baseline algorithm.

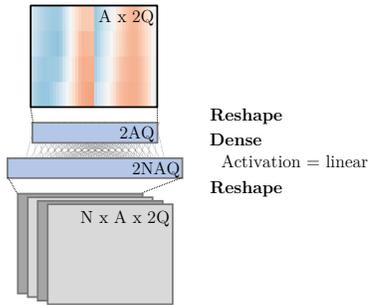


Fig. 1: Baseline dense model architecture. The dense layer is comprised of only N nodes, resulting in a total of $N(N+1)$ trainable parameters.

B. LSTM Algorithm

The proposed model architecture is common in video frame prediction and is repurposed here to serve as CSI prediction that uses "frames" of dimension $A \times 2Q$. With this approach, spatial correlation among BS antennas can be exploited. A sequence of three convolutional LSTM layers is followed by a max-pooling layer with pool size 1×3 . After flattening, a dense layer with $2NAQ$ nodes and subsequent reshaping comprises the last processing steps. Figure 2 shows the architecture of the LSTM algorithm.

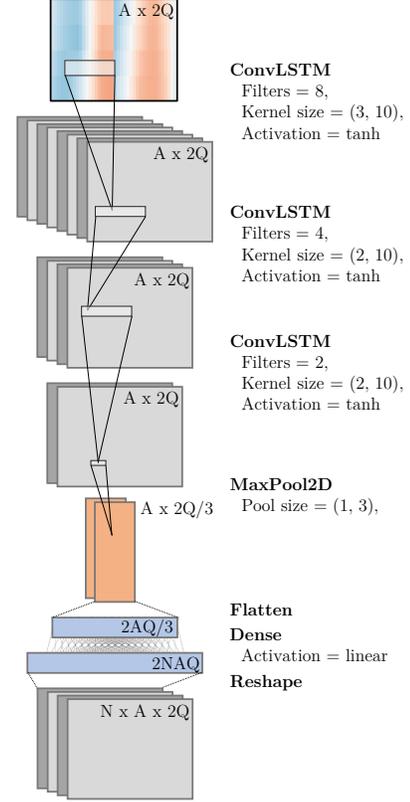


Fig. 2: LSTM model architecture with the total number of trainable parameters $\mathcal{O}(2NAQ(4AQ/3 + 1))$ largely determined by the dense layer.

C. Training and Testing

For each BS configuration, 75% of data points in time are utilized for training purposes, while 25% are used for testing. Table I shows the average duration of training and inference of the proposed models on the CPU of a standard laptop, as well as their number of trainable parameters. The inference times of the constant continuation and Wiener prediction methods are included for comparison.

The training and inference times in Table I are obtained for the values $N = 4$, $A = 32$, and $Q = 20$, which are used in the remainder of this paper. For those parameters, the exact number of trainable parameters is 20 (baseline dense NN) and 4 278 456 (LSTM).

TABLE I: Average training and inference duration on the CPU of a standard laptop, as well as the number of trainable parameters where applicable.

	Epochs	Train. [s]	Inf. [s]	param.
baseline	200	430	0.1	$N(N+1)$
LSTM	30	23100	7	$\mathcal{O}(2NAQ(4AQ/3+1))$
constant	-	-	0.04	-
Wiener	-	-	0.3	-

IV. SCENARIO DESCRIPTION AND MEASUREMENT FRAMEWORK

We conducted a widely distributed massive MIMO vehicle to infrastructure (V2I) channel sounding campaign in March 2022 at the premises of AIT in Vienna, Austria (48.269 080°N, 16.427 637°E). In this campaign, three different array geometries with apertures ranging from 2 m to 46.5 m were installed and tested. Further details of the measurement campaign can be found in [12].

A. Measurement Scenario

Widely distributed and cell-free massive MIMO systems are envisioned to be deployed in urban environments, as they potentially mitigate the burden of large-scale fading, i.e., blocking by buildings, vegetation, or cars. The widely distributed massive MIMO channel sounding campaign is designed to capture these urban channel characteristics, including mobility, multipath propagation, blocking and transition from line of sight (LOS) to non line of sight (NLOS).

Two UE antennas are mounted on the roof of a van. They are referenced in the following as UE 1 and UE 2, respectively. Both UEs are following a fixed trajectory for all measurements, with velocities ranging from 15 km/h to 60 km/h. Figure 3 shows a top view of the scenario under consideration, with the UE trajectory indicated in blue. The starting position of the UE is marked with a white van icon.

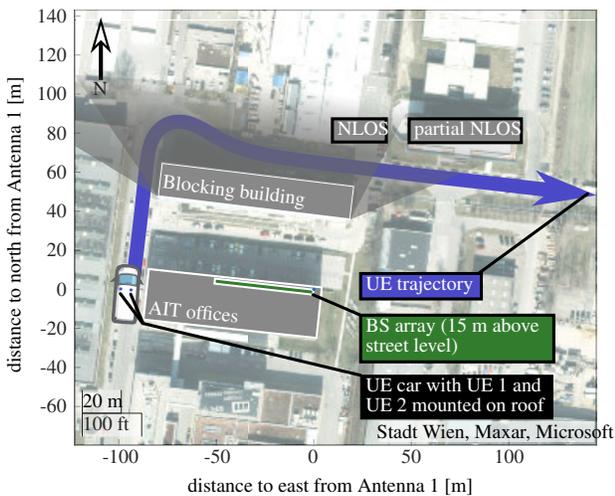


Fig. 3: Top view of the urban scenario of the measurement campaign. The UE trajectory is indicated in blue. The BS antenna array is located on the roof of an office building and indicated in green. The individual BS antennas are facing a large office building to the north, dividing the UE trajectory into LOS and NLOS regions, indicated by gray shadows.

On the BS side, 32 array elements consisting of single patch antennas are positioned on the roof top of an office building at a height of 15 m as a horizontal linear array, with vertical polarization and their individual main lobe facing north. The BS antennas are receiving the signal transmitted by the UE. The green area in Fig. 3 shows the position of the BS array. Since there is an office building of similar height to the north of the linear BS antenna array (in the directions the individual patch antenna array elements are pointing to), parts of the UE trajectory are exhibiting NLOS conditions. The measurements are conducted at a carrier frequency $f_c = 3.2$ GHz, i.e., with a wavelength of $\lambda = c_0/f_c = 9.38$ cm.

Three different linear horizontal BS array configurations are implemented, with significant variation in the array aperture and antenna element spacing. Figure 3 shows in green the position where all three array configurations are located. The exact positioning and spacing of the patch antenna elements for each BS array configuration is detailed in Fig. 4 (from [12]). Photos of the three different BS antenna array setups on top of the AIT office building are shown in Fig. 5.

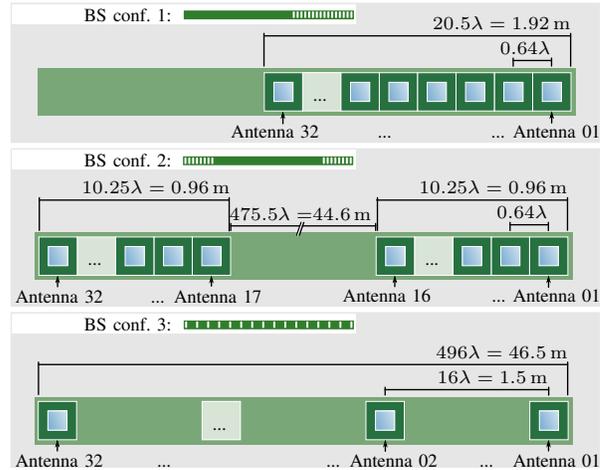


Fig. 4: 32 BS antennas assembled in one of three BS array configurations with total aperture sizes ranging from 2 m to 46.5 m.

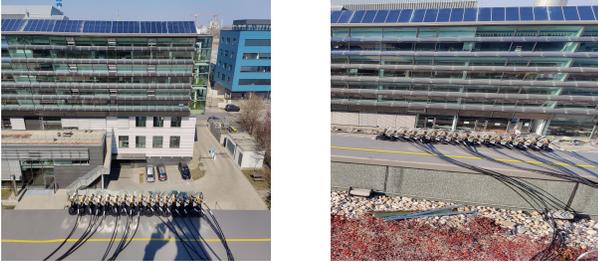
V. RESULTS

We utilize the CSI data obtained in the measurement campaign outlined in Section IV to construct the channel matrix according to (2). Subsequently, we perform CSI prediction according to the methods outlined in Section II-A (constant continuation and Wiener prediction) and Section III (dense NN and LSTM NN). The predicted CSI (6) is then used to calculate the ergodic achievable SE using (14).

The CSI is acquired and predicted with a sampling rate of $T_s = 1$ ms, i.e., a prediction horizon of $\ell = 3$ corresponds to predicting $(\ell + 1)T_s = 4$ ms into the future. The temporal correlation of the fading process [7] for the Wiener filter are assumed stationary for the whole test data-set. For the temporal correlation, a rectangular Doppler spectrum is assumed to obtain a robust predictor, see also [15]. The spectral support in this paper is determined by maximizing the SE with an exhaustive search over the maximum Doppler frequency. We



(a) BS conf. 1 with 32 patch antennas in one collocated linear antenna array.



(b) BS conf. 2 with two distributed linear antenna arrays (16 elements each).



(c) BS conf. 3 with 32 equispaced patch antennas over an aperture of 46.5 m.
Fig. 5: BS configurations with apertures ranging from 2 m to 46.5 m.

chose a maximum Doppler frequency of $\nu_D = 30$ Hz, which corresponds to a relative velocity of the users of 10 km/h. Utilizing a bathtub-like Doppler spectrum according to a Clarke’s model does not yield any improvements compared to the rectangular spectrum. No spatial correlation is assumed between antennas.

A. Spectral Efficiency

In Fig. 6, we show $SE_{k,\ell}$ as defined in (14) versus the prediction horizon ℓ . We use $N = 4$ pilot symbols, BS conf. 1 and compare all considered channel prediction methods. The constant continuation, i.e., utilizing the last known CSI value for beam-forming, yields low SE and is not suitable for high mobility scenarios. The baseline dense NN algorithm works astonishingly well for low prediction horizons and yields comparable SE values to the Wiener prediction method. The LSTM approach, however, clearly outperforms the other considered methods and improves the SE by about 2 bit/s/Hz even for high prediction horizons. For $\ell = 3$ and using the LSTM approach, the SE is 4.47 bit/s/Hz, which is almost four times more than the constant continuation approach (1.32 bit/s/Hz) and a 160% increase over the Wiener predictor (2.79 bit/s/Hz).

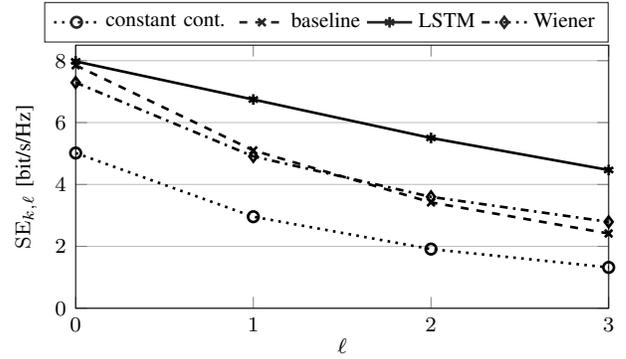


Fig. 6: $SE_{k,\ell}$ over prediction horizon ℓ , when using the BS conf. 1 data set for training and testing.

In Fig. 7 we show the results for BS conf. 2, where we increased the spatial diversity separating two antenna groups by about 45 m. The basic behaviour is similar to the results in Fig. 6. However, the LSTM algorithm shows a small improvement by about 10% in Fig. 7.

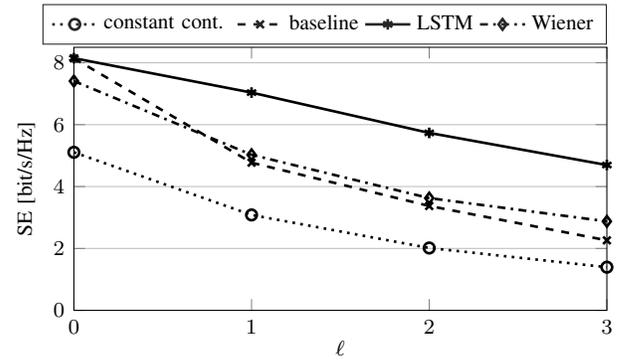


Fig. 7: $SE_{k,\ell}$ over prediction horizon ℓ , when using the BS conf. 2 data set for training and testing.

Finally, in Fig. 8 we show the results for the 32 widely distributed antenna elements over an aperture of 46.5 m and a spacing of 1.5 m between the antenna elements. Again, the LSTM-based NN is able to accurately predict the CSI. However, the baseline dense NN is considerably worse than for BS conf. 1 and 2. This can be attributed to the fact that there is a great variety in channel realizations for BS conf. 3 that cannot be reproduced by the small number of nodes in the baseline dense NN.

B. Testing with Different BS Conf. Data Sets

We tested the SE performance of the considered NNs with the respective BS configuration data sets that have not been used for training to evaluate their capability to handle new channel realization data and evaluate their reliance on correlation among BS antennas. Table II list the corresponding results for a prediction horizon $\ell = 1$.

For the baseline dense NN, we notice similar SE for all test data sets, regardless of the training BS configuration. For the LSTM-based NN, we notice a significant drop in SE of approximately 25% when training with either BS conf. 1 or 2 and testing with BS conf. 3. However, when training with

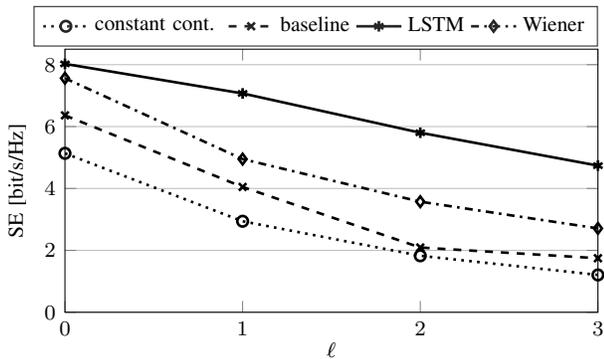


Fig. 8: $SE_{k,\ell}$ over prediction horizon ℓ , when using the BS conf. 3 data set for training and testing.

TABLE II: SE for prediction horizon $\ell = 1$ and all combinations of BS antenna configurations as training and test data.

SE [bit/s/Hz]	Test data set		
	BS conf. 1	BS conf. 2	BS conf. 3
constant continuation	2.96	3.08	2.94
Wiener predictor	4.91	5.03	4.95
baseline dense NN			
Training data set:			
BS conf. 1	5.10	4.83	4.90
BS conf. 2	4.89	4.77	4.70
BS conf. 3	4.19	3.95	4.05
LSTM NN			
Training data set:			
BS conf. 1	6.75	5.45	5.08
BS conf. 2	5.86	7.04	5.20
BS conf. 3	6.55	6.56	7.07

BS conf. 3, the SE for all test data sets stays almost constant. We conclude that that the great variety of observed channel realizations in the widely distributed BS conf. 3 data set is beneficial for LSTM training purposes. Moreover, correlation among BS antennas is either not leveraged by the NNs or does not play a significant role in this specific case, as it is shown in [12] that BS conf. 3 hardly shows spatial correlation.

VI. CONCLUSION

We introduced a simple dense NN and a more complex LSTM-based method to perform CSI prediction on real-world cell-free (widely distributed) massive MIMO CSI data. We evaluated the NN predictors in terms of achievable SE and compared them to two standard approaches, namely constant continuation and Wiener prediction. The LSTM-based NN shows superior prediction quality, with an increase in SE of 160% over the Wiener filter for a prediction horizon of 4 ms. Further, we show that a widely distributed BS antenna setup yields favorable training data due to the large channel realization variety. Spatial BS antenna correlation does not significantly contribute to the prediction quality.

ACKNOWLEDGEMENTS

The work of David Löschenbrand, Markus Hofer and Thomas Zemen is funded within the Principal Scientist grant Dependable Wireless 6G Communication Systems (DEDICATE 6G) at the AIT Austrian Institute of Technology. The work

of Lukas Eller has been funded by the Christian Doppler Laboratory for Digital Twin assisted AI for sustainable Radio Access Networks, Institute of Telecommunications, TU Wien. The financial support by the Austrian Federal Ministry for Labour and Economy and the National Foundation for Research, Technology and Development and the Christian Doppler Research Association is gratefully acknowledged.

REFERENCES

- [1] Ö. T. Demir, E. Björnson, L. Sanguinetti *et al.*, “Foundations of user-centric cell-free massive MIMO,” *Foundations and Trends in Signal Processing*, vol. 14, no. 3-4, pp. 162–472, 2021.
- [2] J. Zhang, S. Chen, Y. Lin, J. Zheng, B. Ai, and L. Hanzo, “Cell-free massive MIMO: A new next-generation paradigm,” *IEEE Access*, vol. 7, pp. 99 878–99 888, 2019.
- [3] H. Jin, K. Liu, M. Zhang, L. Zhang, G. Lee, E. N. Farag, D. Zhu, E. Onggosanusi, M. Shafi, and H. Tataria, “Massive MIMO evolution toward 3GPP release 18,” *IEEE Journal on Selected Areas in Communications*, vol. 41, no. 6, pp. 1635–1654, 2023.
- [4] K. T. Truong and R. W. Heath, “Effects of channel aging in massive MIMO systems,” *Journal of Communications and Networks*, vol. 15, no. 4, pp. 338–351, 2013.
- [5] 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 Transactions on Communications*, vol. 63, no. 12, pp. 4879–4893, 2015.
- [6] A. K. Papazafeiropoulos, “Impact of general channel aging conditions on the downlink performance of massive MIMO,” *IEEE Transactions on Vehicular Technology*, vol. 66, no. 2, pp. 1428–1442, 2016.
- [7] D. Löschenbrand, M. Hofer, and T. Zemen, “Spectral efficiency of time-variant massive MIMO using Wiener prediction,” *IEEE Communications Letters*, vol. 27, no. 4, pp. 1225 – 1229, April 2023.
- [8] 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.
- [9] H. Kim, S. Kim, H. Lee, C. Jang, Y. Choi, and J. Choi, “Massive MIMO channel prediction: Kalman filtering vs. machine learning,” *IEEE Transactions on Communications*, vol. 69, no. 1, pp. 518–528, 2021.
- [10] 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 Journal on Selected Areas in Communications*, vol. 39, no. 7, pp. 1915–1930, 2021.
- [11] M. K. Shehzad, L. Rose, S. Wesemann, and M. Assaad, “ML-based massive MIMO channel prediction: Does it work on real-world data?” *IEEE Wireless Communications Letters*, vol. 11, no. 4, pp. 811–815, 2022.
- [12] 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.
- [13] S. Schwarz and M. Rupp, “Adaptive channel direction quantization—enabling multi user MIMO gains in practice,” in *IEEE International Conference on Communications (ICC)*, Ottawa, Canada, June 2012, pp. 6947–6952.
- [14] T. L. Marzetta, “Noncooperative cellular wireless with unlimited numbers of base station antennas,” *IEEE Transactions on Wireless Communications*, vol. 9, no. 11, pp. 3590–3600, 2010.
- [15] T. Zemen, C. F. Mecklenbräuker, B. H. Fleury, and F. Kaltenberger, “Minimum-energy band-limited predictor with dynamic subspace selection for time-variant flat-fading channels,” *IEEE Transactions on Signal Processing*, vol. 55, no. 9, pp. 4534–4548, September 2007.