Channel Estimation with Reduced Phase Allocations in RIS-Aided Systems

Benedikt Fesl, Andreas Faika, Nurettin Turan, Michael Joham, and Wolfgang Utschick
School of Computation, Information and Technology, Technical University of Munich, Germany
Email: {benedikt.fesl, andreas.faika, nurettin.turan, joham, utschick}@tum.de

***Abstract*—We consider channel estimation in systems equipped with a reconfigurable intelligent surface (RIS). In order to illuminate the additional cascaded channel as compared to systems without a RIS, commonly an unaffordable amount of pilot sequences has to be transmitted over different phase allocations at the RIS. However, for a given base station (BS) cell, there exist immanent structural characteristics of the environment which can be leveraged to reduce the necessary number of phase allocations. We verify this observation by a study on discrete Fourier transform (DFT)-based phase allocations where we exhaustively search for the best combination of DFT columns. Since this brute-force search is unaffordable in practice, we propose to learn a neural network (NN) for joint phase optimization and channel estimation because of the dependency of the optimal phase allocations on the channel estimator, and vice versa. We verify the effectiveness of the approach by numerical simulations where common choices for the phase allocations and the channel estimator are outperformed. By an ablation study, the learned phase allocations are shown to be beneficial in combination with a different state-of-the-art channel estimator as well.**

***Index Terms*—Reconfigurable intelligent surface, channel estimation, phase optimization, convolutional neural network.**

I. Introduction

RIS-aided systems are enabling enhanced communication performance and are thus considered to be a key technology in 6G systems [1]. Having accurate estimates of both the direct and the cascaded channel including the RIS is crucial. Since the RIS only consists of passive elements, processing the impinging waves is not possible. Consequently, no separate channel estimation can be conducted at the RIS. To fully illuminate the cascaded channel, a large number of training sequences has to be transmitted over different phase allocations at the RIS. However, the transmission of these long training sequences drastically diminishes the available time for data transmission and hence decreases the achievable rate. Since the number of training phase allocations scales both in the number of RIS elements and in the number of pilot sequences for a MIMO system, it is generally considered to be unaffordable to fully illuminate the cascaded channel [2].

A variety of approaches for phase optimization and channel estimation is considered in the literature. An on/off strategy was proposed in, e.g. [3], where the direct and the cascaded channels are estimated subsequently, which is known to be suboptimal [4]. As shown in [4], [5], the DFT matrix is the optimal phase allocation matrix when employing the least squares (LS) estimator for full illumination. Unfortunately, the optimal phase allocations are unknown in general for the minimum mean square error (MMSE) channel estimator or when having reduced phase allocations. Therefore, some prior works have considered phase optimization or reduction for specific instances. The work in [6] discusses optimization of discrete phase shifts, and [7] investigates joint pilot and phase optimization for the MMSE estimator in the full illumination case. In [8], a projected gradient descent algorithm for optimizing the phase allocation matrix is proposed which only holds for a sparse geometry-modeled channel. An element grouping strategy was proposed in [9] whose disadvantage is the loss of degrees of freedom at the RIS. To summarize, none of the existing prior works considers phase reduction and optimization jointly for a generally unknown and arbitrary complex channel distribution which we consider in this paper.

Contributions: We investigate the optimization of phase shifts at the RIS for channel estimation inside a specific BS cell, especially for a reduced number of phase allocations. To analyze the potential of phase optimization with respect to a given radio propagation environment, we perform a study on reduced DFT-based phases where we exhaustively search for the best combination of DFT columns as phase matrix, which is shown to be heavily dependent on the considered scenario.

Motivated by this observation, we propose a NN which jointly learns the phase matrix and the channel estimator. The first part of the NN emulates the observed signal by interpreting the angles of the reduced phase matrix as parameterizable weights. The phase matrix module by design fulfills the unit magnitude constraint enforced by the passive nature of the RIS elements that is problematic in classical RIS optimization algorithms. This allows to adjust the reduced phase matrix to the propagation scenario by training. The second part of the NN consists of a convolutional neural network (CNN) for channel estimation. We show in numerical experiments that the proposed approach outperforms DFT-based and random phase allocations together with state-of-the-art channel estimators. We further perform an ablation study to evaluate the properties of the optimized phase allocations, i.e., the performance with respect to a different channel estimator.

II. System and Channel Model

We consider a RIS-aided single-input multiple-output (SIMO) system where we denote the direct channel between

The authors acknowledge the financial support by the Federal Ministry of Education and Research of Germany in the programme of “Souverän. Digital. Vernetzt.”. Joint project 6G-life, project identification number: 16KISK002.

a single-antenna mobile terminal (MT) and an M-antenna BS by $\boldsymbol{h}_0 \in \mathbb{C}^M$. The channel between the RIS with L passive elements and the MT is denoted by $\boldsymbol{h}_1 \in \mathbb{C}^L$, whereas the channel between the RIS and the BS is denoted by $\boldsymbol{H}_2 \in \mathbb{C}^{M \times L}$. The received uplink signal is then given by

$$\boldsymbol{y}' = \boldsymbol{h}_0 + \boldsymbol{H}_2 \boldsymbol{\Phi} \boldsymbol{h}_1 + \boldsymbol{n}' \tag{1}$$

where $\boldsymbol{\Phi} = \operatorname{diag}(\boldsymbol{v}) \in \mathbb{C}^{L \times L}$ comprises the unimodular phase shift coefficients at the RIS elements and $\boldsymbol{n}' \sim \mathcal{N}_{\mathbb{C}}(\mathbf{0}, \sigma^2 \mathbf{I})$ is additive white Gaussian noise (AWGN). Due to the passive elements at the RIS, the amplitudes of the reflected signals are not changed. Hence, $v_\ell = \mathrm{e}^{\mathrm{j}\theta_\ell}$ with the angle $\theta_\ell \in [0, 2\pi)$ and unit-magnitude entries $|v_\ell| = 1$ for $\ell = 1, \ldots, L$. With $\boldsymbol{H} = [\boldsymbol{h}_0, \boldsymbol{h}_1^{\mathrm{T}} \circledast \boldsymbol{H}_2] \in \mathbb{C}^{M \times L+1}$, where $\circledast$ denotes the Khatri-Rao product, the system in (1) can be written as $\boldsymbol{y}' = \boldsymbol{H}\boldsymbol{v}' + \boldsymbol{n}'$ where $\boldsymbol{v}' = [1, \boldsymbol{v}^{\mathrm{T}}]^{\mathrm{T}}$, see e.g., [5].

Note that $\boldsymbol{h}_1$ and $\boldsymbol{H}_2$ of the cascaded channel $\boldsymbol{H}_2 \boldsymbol{\Phi} \boldsymbol{h}_1$ cannot be estimated explicitly [5]. Therefore, N_v different phase allocations are considered, that are collected in $\boldsymbol{V} = [\boldsymbol{v}_1', \ldots, \boldsymbol{v}_{N_v}']$, to illuminate the channel. This yields

$$\boldsymbol{Y} = \boldsymbol{H}\boldsymbol{V} + \boldsymbol{N} \in \mathbb{C}^{M \times N_v} \tag{2}$$

as the training sequence where the N_v different observations are collected as the columns of $\boldsymbol{Y}$. After vectorization, we get

$$\boldsymbol{y} = (\boldsymbol{V}^{\mathrm{T}} \otimes \mathbf{I})\boldsymbol{h} + \boldsymbol{n} = \boldsymbol{A}\boldsymbol{h} + \boldsymbol{n} \in \mathbb{C}^{M N_v}, \tag{3}$$

with the vectorized expressions $\boldsymbol{h} = \operatorname{vec}(\boldsymbol{H})$, $\boldsymbol{y} = \operatorname{vec}(\boldsymbol{Y})$, $\boldsymbol{n} = \operatorname{vec}(\boldsymbol{N})$, and the observation matrix $\boldsymbol{A} = \boldsymbol{V}^{\mathrm{T}} \otimes \mathbf{I}$, where $\otimes$ denotes the Kronecker product. We define the signal-to-noise ratio (SNR) as $\mathrm{SNR} = 1/\sigma^2$ where we normalize the channels to $\mathrm{E}[\|\boldsymbol{h}\|_2^2] = M(L+1)$.

For the construction of a scenario-specific channel dataset, we use the QuaDRiGachannel simulator [10]. We consider an urban macrocell (UMa) scenario following the 3GPP 38.901 specification, where the BS is placed at a height of 25m and covers a sector of 120°. The RIS is placed opposite to the BS with a distance of 500m at the same height. Note that, opposite to the MT with possibly non-line-of-sight (NLOS) channels, the channel between RIS and BS has line-of-sight (LOS) condition. We want to highlight that, although the position of the BS and RIS is fixed, the corresponding channel is not constant within the dataset but slightly changes according to the UMa conditions. The generated channels are post-processed to remove the path gain.

III. Reference Methods

A. Channel Estimation

We briefly introduce the Gaussian mixture model (GMM) and the conditional mean estimator (CME) based thereon from [11], [12]. A GMM with K components is a probability density function (PDF) of the form $f_{\boldsymbol{h}}^{(K)}(\boldsymbol{h}) = \sum_{k=1}^{K} p(k) \mathcal{N}_{\mathbb{C}}(\boldsymbol{h}; \boldsymbol{\mu}_k, \boldsymbol{C}_k)$ consisting of a weighted sum of K Gaussian PDFs. Given data samples, an expectation-maximization (EM) algorithm can be used to fit a K-components GMM [13, Sec. 9.2]. In [11], [12], a CME is formulated based on a GMM, which is proven to asymptotically converge to the true CME when K grows large. The estimator is formulated as a convex combination of linear minimum mean square error (LMMSE) terms, given as

$$\hat{\boldsymbol{h}}^{(K)} = \sum_{k=1}^{K} p(k \mid \boldsymbol{y})(\boldsymbol{\mu}_k + \boldsymbol{C}_k \boldsymbol{A}^{\mathrm{H}} \boldsymbol{C}_{\boldsymbol{y},k}^{-1}(\boldsymbol{y} - \boldsymbol{A}\boldsymbol{\mu}_k)) \tag{4}$$

where the responsibilities $p(k \mid \boldsymbol{y})$ are computed by

$$p(k \mid \boldsymbol{y}) = \frac{p(k) \mathcal{N}_{\mathbb{C}}(\boldsymbol{y}; \boldsymbol{A}\boldsymbol{\mu}_k, \boldsymbol{C}_{\boldsymbol{y},k})}{\sum_{i=1}^{K} p(i) \mathcal{N}_{\mathbb{C}}(\boldsymbol{y}; \boldsymbol{A}\boldsymbol{\mu}_i, \boldsymbol{C}_{\boldsymbol{y},i})} \tag{5}$$

with $\boldsymbol{C}_{\boldsymbol{y},k} = \boldsymbol{A}\boldsymbol{C}_k \boldsymbol{A}^{\mathrm{H}} + \sigma^2 \mathbf{I}$, cf. (3). In order to reduce the online computational complexity of the estimator, structural features of the covariances can be utilized, cf. [14], which is out of the scope of this paper.

A LMMSE estimator is evaluated as baseline using a cell-wide sample covariance matrix $\boldsymbol{C} = \frac{1}{N} \sum_{n=1}^{N} \boldsymbol{h}_n \boldsymbol{h}_n^{\mathrm{H}}$ with $N = 19 \cdot 10^4$ training samples to compute

$$\boldsymbol{h}_{\text{sample cov.}} = \boldsymbol{C}\boldsymbol{A}^{\mathrm{H}} (\boldsymbol{A}\boldsymbol{C}\boldsymbol{A}^{\mathrm{H}} + \sigma^2 \mathbf{I})^{-1} \boldsymbol{y}. \tag{6}$$

Finally, the LS estimator is $\hat{\boldsymbol{h}}_{\mathrm{LS}} = \boldsymbol{A}^{\dagger} \boldsymbol{y} = (\boldsymbol{V}^{\dagger} \otimes \mathbf{I}) \boldsymbol{y}$, where $\boldsymbol{V}^{\dagger}$ is the pseudoinverse of $\boldsymbol{V}$.

B. Phase Allocations

A simple choice for the phase allocations is to use random phase shifts for every MT. We therefore construct a phase matrix by sampling i.i.d. Gaussian realizations from $\mathcal{N}_{\mathbb{C}}(0, 1)$ per entry and dividing each entry by its absolute value to fulfill the unit magnitude constraint. Note that these phase allocations might be difficult to implement in a practical system because of the very limited processing ability at the RIS.

Since DFT-based phases are optimal in the full-illumination case for the LS estimator [4], [5], we evaluate the use of a DFT submatrix for reduced phase allocations, i.e., the m,nth entry is given as $V_{m,n}^{\text{sub-DFT}} = \exp((m-1)(n-1)\mathrm{j}2\pi/N_v)$ with $m = 1, \ldots, L+1$ and $n = 1, \ldots, N_v$. Note that the columns of the DFT submatrix are not orthogonal for $N_v < L+1$.

IV. DFT-Based Phase Allocation Study

In this section, we investigate the potential of the phase allocation optimization based on a DFT grid. We consider a BS with a uniform linear array (ULA) consisting of $M = 8$ antennas serving single-antenna MTs supported by a RIS with $L = 4 \times 4 = 16$ elements. Instead of full illumination with $L+1$ phase allocations at the RIS we set $N_v = 8$ in order to simulate a reduced phase allocation situation. We then exhaustively search for the best combination of eight columns drawn from the full $(L+1)$-dimensional DFT matrix for $10{,}000$ uniformly sampled MTs in the BS cell. Note that this procedure is infeasible in practical systems since in general $\binom{L+1}{N_v}$ combinations of DFT columns have to be tested for every MT which drastically increases for higher numbers of RIS patches. In the considered case this already yields $\binom{17}{8} = 24{,}310$ combinations. For each MT, we choose

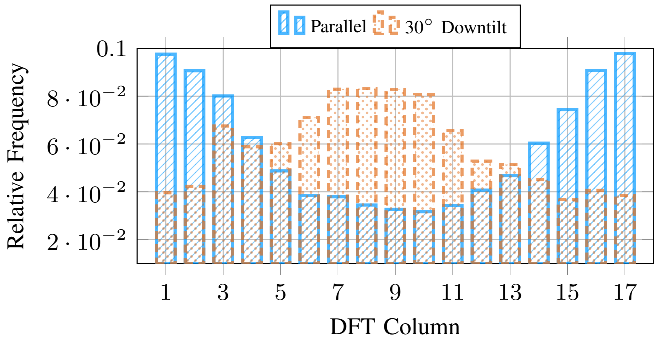

Fig. 1: Histogram of the occurrence of the DFT columns in the exhaustive search approach for different RIS configurations with $M = 8$, $L = 16$, and $N_v = 8$ at a SNR of 40dBm.

the combination which yields the best channel estimation performance based on the GMM estimator introduced in Section III-A at an SNR of 40dBm. The histogram in Fig. 1 shows how often each DFT column occurs relatively in the exhaustive search of DFT column combinations over all MTs for two different scenarios.

In the first scenario, the RIS array is placed in parallel to the BS array where it can be observed that especially the first and last DFT columns occur more frequently in the best-performing combinations. On average, the best combination of DFT columns for this scenario is $\{1, 2, 3, 4, 14, 15, 16, 17\}$. In contrast to that, for a scenario where the RIS has a downtilt of 30°, the middle DFT columns occur primarily in the best combinations and $\{3, 5, 6, 7, 8, 9, 10, 11\}$ is the best combination on average.

The conclusions of this study are twofold. First, we have seen that there is great potential for optimizing the phases since for a given setting, some DFT columns are much more important than others. Second, we showed that the optimization of the phase allocations heavily depends on the considered scenario. We further evaluate the optimization based on the DFT grid in the numerical experiments section where we compare this exhaustive brute-force search approach to our proposed optimization procedure.

V. Learning-Based Joint Phase Optimization and Channel Estimation

In Section IV, we have seen that the choice of the phase allocation matrix is depending heavily on the underlying system setup, i.e., the configuration of the RIS, as well as on the propagation environment that induces structural properties which can be exploited for reduced phase allocations. However, on the one hand, the optimization procedure from Section IV is infeasible in practice because of the combinatorial search, on the other hand, it is limited to a search on the DFT grid which may be sub-optimal in general.

Thus, we propose to utilize machine learning for joint phase optimization and channel estimation via a specific NN architecture in RIS-aided systems. In essence, we parametrize the phase allocations of the matrix $\boldsymbol{V}$. Due to the passiveness

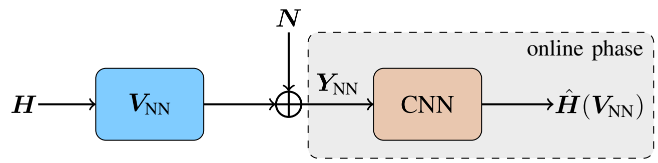

Fig. 2: Flowchart of the proposed NN architecture for joint phase optimization and channel estimation. In the online phase only the CNN is evaluated for a fixed $\boldsymbol{V}_{\text{NN}}$.

of the RIS which enforces the unit magnitude constraint, we only train with respect to the angles of the phase matrix. In particular, the phase matrix is constructed as

$$\boldsymbol{V}_{\text{NN}} = \cos(\boldsymbol{\Phi}) + \text{j} \sin(\boldsymbol{\Phi}) \tag{7}$$

where $\boldsymbol{\Phi} \in \mathbb{R}^{L+1 \times N_v}$. A similar approach for the optimization of a sensing matrix with a magnitude constraint was employed in [15] which serves as a motivation for our considerations.

The training procedure is summarized as follows. The parametrized phase matrix $\boldsymbol{V}_{\text{NN}}$ given by (7) is multiplied with a channel realization from the training dataset. Afterwards, we artificially add AWGN, yielding an emulated observation $\boldsymbol{Y}_{\text{NN}}$ following the model in (2). The emulated observation $\boldsymbol{Y}_{\text{NN}}$ then serves as the input of a CNN which yields a channel estimate $\hat{\boldsymbol{H}}(\boldsymbol{V}_{\text{NN}})$ at the output. Therefore, the complex-valued input of the CNN is split into its real and imaginary part as different convolution channels and each layer employs 2D convolutions. Since the phase optimization and the training of the CNN for channel estimation depend on each other, it is not possible to separately update their parameters. Thus, we jointly optimize the phase matrix $\boldsymbol{V}_{\text{NN}}$ and the CNN for which we exploit the efficient framework of NNs with powerful gradient-based optimization techniques. As such, we can interpret the phase matrix $\boldsymbol{V}_{\text{NN}}$ as a layer with a specific structure, cf. (7), of a larger NN that contains the CNN as further layers. The described architecture is summarized as a flowchart in Fig. 2. We utilize labeled data from the constructed dataset, cf. Section II, to compute gradients with the mean square error (MSE)

$$\text{MSE} = \text{E}[\|\boldsymbol{H} - \hat{\boldsymbol{H}}(\boldsymbol{V}_{\text{NN}})\|_F^2] \tag{8}$$

as cost function. Note that a single forward pass propagates through both NN parts and, therefore, all network parameters are updated simultaneously. After training, the optimized phase allocations are given by (7) and the trained CNN is extracted as the channel estimator.

We initialize the weights of the phase matrix randomly at the beginning of the training and we perform a random hyperparameter search for the NN parameters, i.e., the batch size ($\in [2^5, 2^{11}]$), activation functions (ReLU, Tanh, Sigmoid, SiLU, ELU), batch normalization, learning rate ($\in [10^{-5}, 10^{-1}]$), number of kernels ($\in [16, 512]$) and layers ($\in [3, 9]$) for 3×3 convolution kernels, where we choose the best setting over 100 random initializations.

The optimized phase matrix $\boldsymbol{V}_{\text{NN}}$ after training is further evaluated by an ablation study where it is used in combination

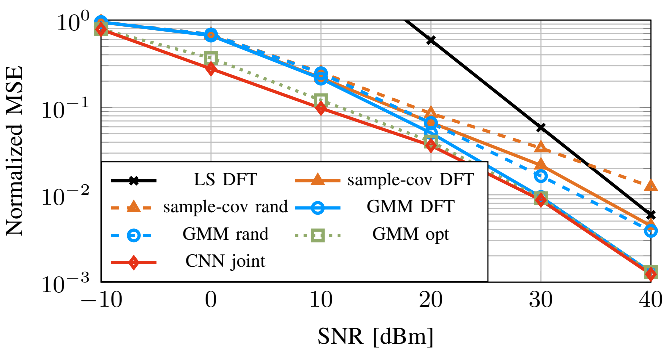

Fig. 3: $M = 8$ ULA BS antennas, $L = 4 \times 4 = 16$ URA RIS patches and single-antenna MTs with $N_v = L + 1$.

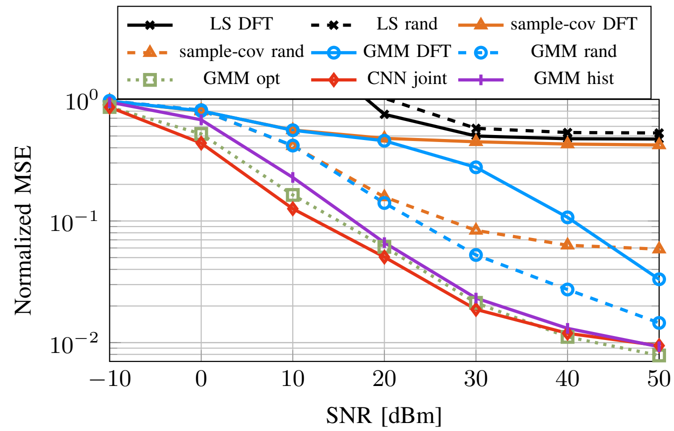

Fig. 4: $M = 8$ ULA BS antennas, $L = 4 \times 4 = 16$ URA RIS patches and single-antenna MTs with $N_v = 8$.

with a different channel estimator, i.e., the GMM estimator from Section III-A, instead of the trained CNN. Since the resulting performance is better in comparison to the baseline phase allocations (cf. Section VI), although the GMM estimator is not jointly trained with the phase matrix, we conclude that the optimized phase matrix $\boldsymbol{V}_{\text{NN}}$ is beneficial for the whole BS cell and exploits its structural properties.

The online computational complexity of the proposed CNN estimator is determined by a single forward pass, which depends on the chosen hyper-parameters, since the optimized phase matrix $\boldsymbol{V}_{\text{NN}}$ is fixed after training, cf. Fig. 2. Note that we also employ the trained phase matrix $\boldsymbol{V}_{\text{NN}}$ when using the GMM estimator whose online complexity is given in [11], [14].

VI. Numerical Results

We present numerical results for the described setting in Section II. We utilize a dataset consisting of $19 \cdot 10^4$ data samples for fitting the GMM with $K = 128$ components and training the NN. Each method is evaluated using 10^4 samples which are not part of the training data. For all plots, we evaluate the scenario with a parallel RIS with a uniform rectangular array (URA) opposite to the BS with a ULA since the results are qualitatively the same for both depicted scenarios in Section IV. The curves labeled "LS", "sample-cov", or "GMM" refer to the baseline estimators from Section III-A, whereas "CNN joint" refers to the proposed approach from Section V. The additional labeling "DFT", "rand", "opt", or "hist" refers to the choice of the phase allocation matrix based on the DFT (sub)matrix or on random allocations, cf. Section III-B, the optimized phase allocations from the NN, cf. Section V, or the histogram based search from Section IV, respectively.

A. Full Illumination

In Fig. 3, we depict results for the case of full illumination, i.e., $N_v = L + 1$ with $M = 8$ ULA BS antennas and $L = 4 \times 4$ URA RIS patches. In case of the LS estimator, the DFT phase matrix is shown to be optimal, cf. [4], [5]. In contrast, a randomly chosen phase matrix might be ill-conditioned in general which results, in combination with the LS estimator, in a normalized MSE that is larger than one. Therefore, it is not shown in Fig. 3. When using the GMM or the sample-covariance based estimator, the DFT phase allocations yield a better performance as compared to random allocations.

Interestingly, it can be observed that the channel estimators with the optimized phase matrix ("CNN joint" and "GMM opt") are able to outperform the DFT matrix in the low SNR regime with a vanishing gap in the high SNR, where the LS estimator is reasonable. This means that phase optimization is in fact useful also for the full illumination case for low SNR values. Furthermore, since the performance is very similar for both the GMM and CNN estimator it can be concluded that the optimized phase matrix is not only useful for the jointly trained CNN, but is generally adapted to the scenario.

B. Reduced Phase Allocations

In Fig. 4, we show the same setting as in Section VI-A but with a reduced number of $N_v = 8$ phase allocations, i.e., less than 50% of the fully illuminated case. First, it can be observed that the LS estimator performs poorly due to the underdetermined system. Since the random phases in combination with the LS estimator show no performance gains, we leave out this combination in the following. Second, the random phase allocations outperform the sub-DFT allocations when using the GMM or sample covariance estimator where the GMM estimator performs significantly better than the sample covariance estimator. Note that the performance of the sub-DFT phases is not consistent for an increasing number of phases which can be explained through the insights from Section IV. Finally, the CNN and GMM estimator with optimized phase allocations show a similar performance which is better than all baseline methods, including the GMM based on the histogram search method from Section IV. This demonstrates the great potential of optimization for reduced phase allocations.

Fig. 5 depicts the same setting as before for a fixed SNR of 40dBm with a varying number of phase allocations. Similarly as before, the methods with optimized phase allocations outperform all baseline algorithms. For the GMM estimator it

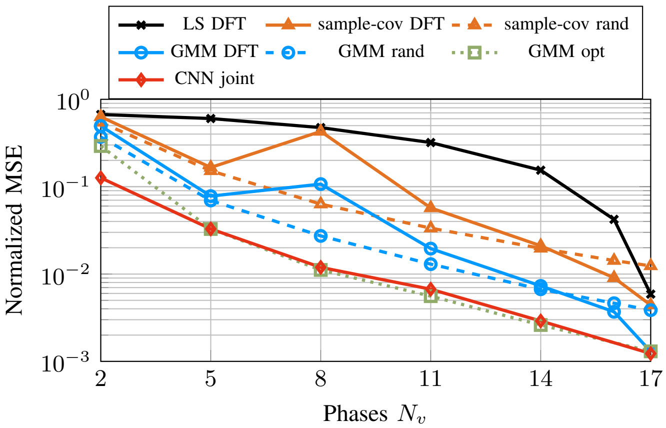

Fig. 5: $M = 8$ ULA BS antennas, $L = 4 \times 4 = 16$ URA RIS patches and single-antenna MTs with SNR $= 40$dBm.

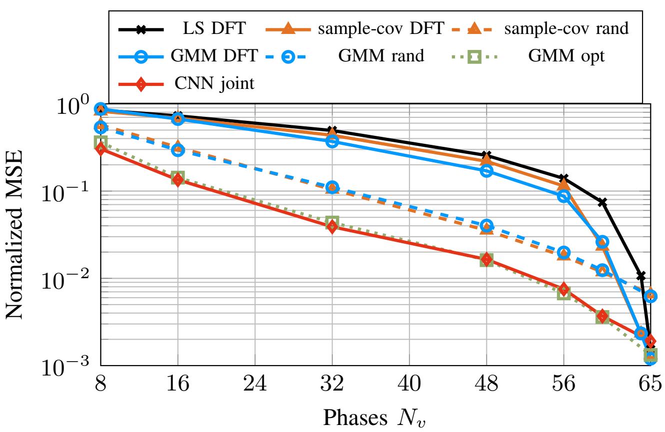

Fig. 6: $M = 16$ ULA BS antennas, $L = 8 \times 8 = 64$ URA RIS patches and single-antenna MTs with SNR $= 40$dBm.

is possible to achieve a normalized MSE of 10^{-2} with only $N_v = 9$ phase allocations, whereas more than $N_v = 12$ phase allocations are needed to achieve the same MSE when having random or DFT-based phase allocations. In the case of full illumination, i.e., $N_v = 17$, the CNN and GMM estimators with optimized or DFT-based phase allocations show the same performance which is in accordance with Fig. 3.

Finally, in Fig. 6, we show results for a larger system setup with $M = 16$ ULA BS antennas and $L = 8 \times 8$ URA RIS patches for a fixed SNR of 40dBm. It can be observed that especially the gap to the DFT-based phase allocations increases drastically which perform poorly for this larger system setup. Once again, the optimized phase allocations allow for drastic performance gains, which is equivalent to requiring less phase allocations to achieve the same estimation quality. In conclusion, the optimization of the phase allocations has increasing potential for larger systems which is in compliance with the trend to massive multiple-input multiple-output (MIMO) systems.

VII. Conclusion

In this work, we investigated the potential of optimizing the reduced phase allocation matrix for channel estimation in RIS-aided systems which tackles the problem of unaffordable large pilot overhead for full illumination. With a study based on a selection of DFT columns, we found that the system setup drastically influences the choice of the optimal phase allocations. We proposed a NN which jointly learns a phase allocation matrix together with a channel estimator. The proposed approach outperforms the baseline approaches over the whole SNR range by a large margin. In addition, when using the optimized phase allocation matrix for a different state-of-the-art channel estimator, its performance is significantly increased. This leads to the conclusion that the optimized phase allocation matrix is able to leverage the inherent structure of the BS' environment to performance gains.

References

[1] Hexa-X, "Deliverable D4.1 AI-driven Communication & Computation Co-Design: Gap Analysis and Blueprint," 5G PPP, Tech. Rep. 101015956, 2021.

[2] B. Zheng, C. You, W. Mei, and R. Zhang, "A Survey on Channel Estimation and Practical Passive Beamforming Design for Intelligent Reflecting Surface Aided Wireless Communications," *IEEE Commun. Surveys Tuts.*, vol. 24, no. 2, pp. 1035–1071, 2022.

[3] Z. Wang, L. Liu, and S. Cui, "Channel Estimation for Intelligent Reflecting Surface Assisted Multiuser Communications: Framework, Algorithms, and Analysis," *IEEE Trans. Wireless Commun.*, vol. 19, no. 10, pp. 6607–6620, 2020.

[4] T. L. Jensen and E. De Carvalho, "An Optimal Channel Estimation Scheme for Intelligent Reflecting Surfaces Based on a Minimum Variance Unbiased Estimator," in *IEEE Int. Conf. Acoust., Speech, Signal Process.*, 2020, pp. 5000–5004.

[5] M. Joham, H. Gao, and W. Utschick, "Estimation Of Channels In Systems With Intelligent Reflecting Surfaces," in *IEEE Int. Conf. Acoust., Speech, Signal Process.*, 2022, pp. 5368–5372.

[6] C. You, B. Zheng, and R. Zhang, "Channel Estimation and Passive Beamforming for Intelligent Reflecting Surface: Discrete Phase Shift and Progressive Refinement," *IEEE J. Sel. Areas Commun.*, vol. 38, no. 11, pp. 2604–2620, 2020.

[7] J.-M. Kang, "Intelligent Reflecting Surface: Joint Optimal Training Sequence and Refection Pattern," *IEEE Commun. Lett.*, vol. 24, no. 8, pp. 1784–1788, 2020.

[8] S. Noh, H. Yu, and Y. Sung, "Training Signal Design for Sparse Channel Estimation in Intelligent Reflecting Surface-Assisted Millimeter-Wave Communication," *IEEE Trans. Wireless Commun.*, vol. 21, no. 4, pp. 2399–2413, 2022.

[9] Y. Yang, B. Zheng, S. Zhang, and R. Zhang, "Intelligent Reflecting Surface Meets OFDM: Protocol Design and Rate Maximization," *IEEE Trans. Commun.*, vol. 68, no. 7, pp. 4522–4535, 2020.

[10] S. Jaeckel, L. Raschkowski, K. Börner, and L. Thiele, "QuaDRiGa: A 3-D Multi-Cell Channel Model With Time Evolution for Enabling Virtual Field Trials," *IEEE Trans. Antennas Propag.*, vol. 62, no. 6, pp. 3242–3256, 2014.

[11] M. Koller, B. Fesl, N. Turan, and W. Utschick, "An Asymptotically MSE-Optimal Estimator Based on Gaussian Mixture Models," *IEEE Trans. Signal Process.*, vol. 70, pp. 4109–4123, 2022.

[12] ——, "An Asymptotically Optimal Approximation of the Conditional Mean Channel Estimator Based on Gaussian Mixture Models," in *IEEE Int. Conf. Acoust., Speech, Signal Process.*, 2022, pp. 5268–5272.

[13] C. M. Bishop, *Pattern Recognition and Machine Learning (Information Science and Statistics)*. Berlin, Heidelberg: Springer-Verlag, 2006.

[14] B. Fesl, M. Joham, S. Hu, M. Koller, N. Turan, and W. Utschick, "Channel Estimation based on Gaussian Mixture Models with Structured Covariances," in *56th Asilomar Conf. Signals, Syst., Comput.*, 2022.

[15] M. Koller and W. Utschick, "Learning a Compressive Sensing Matrix with Structural Constraints via Maximum Mean Discrepancy Optimization," *Signal Process.*, vol. 197, p. 108553, 2022.