Canonical Training is Bad for Reconfigurable Intelligent Surfaces

Bharath Shamasundar and Aria Nosratinia
The University of Texas at Dallas, Richardson, TX, USA
Email: {bharath.shamasundar, aria}@utdallas.edu

2022 IEEE International Symposium on Information Theory (ISIT) | 978-1-6654-2159-1/22/$31.00 ©2022 IEEE | DOI: 10.1109/ISIT50566.2022.9834861

***Abstract*—Channel training in reconfigurable intelligent surfaces (RIS) is different from MIMO; in addition to pilots, it also requires setting *RIS training states*. Several RIS training schemes are studied in the literature, but the effect of training overhead and accuracy on capacity has not received the deserved attention. This is necessary for guiding the choice of RIS dimensionality and the parameters of modulation and coding. The present work fills this gap and shows that the spectral efficiency of RIS with *DFT training* (that uses the columns of DFT matrix for the RIS training states) is higher than that with the *canonical training* (that uses the canonical basis for RIS training states). Specifically, with 32 RIS elements, DFT training achieves a 2 bits/s/Hz gain compared to canonical training when the coherence interval is 150 time slots. Our results also reveal that beyond a certain critical RIS array size, further increase in size is not beneficial and that the optimal array size goes down with increase in SNR.**

I. Introduction

Reconfigurable intelligent surfaces (RIS) are passive, controllable arrays of small reflectors that allow them to direct radio waves toward or away from a target node, thus enabling a better management of signals and interference in wireless networks. Their interpretation as a means of controlling the propagation environment, or as a passive beamforming mechanism that significantly magnifies the scope of capabilities in wireless networking, has fascinated and attracted the research community [1], [2], [3], [4], [5], [6].

Most RIS applications aim to improve SNR at a desired user or reduce interference on other users, requiring passive beamforming by the RIS. Toward this end, several beamforming techniques for RIS are proposed in the literature (see [7], [8], [9], [10] and the references therein). RIS-beamforming needs channel estimates that are difficult to measure because of the large number of links created by the RIS elements requiring estimation. Further, the passive RIS is unable to transmit nor receive pilots for channel estimation. The exploration of options for addressing this problem includes [11], [12], [13], [14], [15], [16], [17].

The effect of training overhead and training error on the spectral efficiency of RIS-aided systems, however, has not received the attention it deserves. This is essential for guiding the choice of RIS array size, as it reveals the trade-offs between the RIS dimensionality and the performance. Specifically, a large RIS can consume too much time for training and leave less time for data, whereas, a small RIS may not effectively capitalize on the passive beamforming, both of which resulting in sub-optimal performance. Motivated by this, the present work analyzes the spectral efficiency of RIS with channel training.

The training in RIS is different from that in MIMO since, in addition to transmitting pilot signals, it also requires setting training amplitudes and phases at the RIS, referred as the *RIS training states*. The present work shows that, for the same pilot sequence, different choices of RIS training states result in different achieved spectral efficiencies. Specifically, we analyze the spectral efficiency with two training schemes, namely, the *canonical training* and the *DFT training*. Canonical training is inspired by MIMO training in which channels from the transmitter to the receiver via each of the RIS elements are estimated individually by activating one RIS element in each training slot with the rest of the elements deactivated[1] [11]. Although this is intuitive, the pilot signal in canonical training is reflected by only one RIS element (smaller area) in each training slot, leading to less received pilot power. The DFT training uses the columns of the DFT matrix as the RIS training states [17]. Since DFT training does not deactivate RIS elements, the received pilot power is higher than that in canonical training, resulting in better training quality. We also derive optimal power allocation between training and data for both canonical and DFT training.

Our results show that, for RIS with 32 elements and a channel with coherence interval of 150 time slots, DFT training achieves a gain of 3.5 bits/s/Hz compared to canonical training with equal power allocation between training and data, and a gain of 2 bits/s/Hz with optimal power allocation. Our results also reveal the trade-offs between RIS array size and spectral efficiency. Specifically, beyond a certain critical RIS size that depends on SNR, further increase of array size is not beneficial. At low SNR, power is at a premium while degrees of freedom are not, therefore a larger RIS is beneficial. At high-SNR, training overhead becomes more critical and hence larger array is unhelpful.

II. System Model

Consider a single antenna transmitter and receiver assisted by an RIS with N passive elements. Let $\mathbf{h}_1 =$

This work was supported in part by the NSF grant 1956213.

[1]A deactivated RIS element reflects no energy. This assumption is widely used, but its practical possibility is still unclear in the literature.

$[h_1(1)\ h_1(2)\ \ldots\ h_1(N)]^T$ denote the $N \times 1$ channel between the transmitter and the RIS, $\mathbf{h}_2^H = [h_2(1)\ h_2(2)\ \ldots\ h_2(N)]$ denote the $1 \times N$ channel between RIS and the receiver. The channel gains follow i.i.d $\mathcal{CN}(0,1)$[2]. We assume that there is no direct path between transmitter and receiver. Let $\mathbf{\Phi} = \text{diag}[\alpha_1 e^{j\phi_1}, \alpha_2 e^{j\phi_2}, \ldots, \alpha_N e^{j\phi_N}]$ denote the RIS matrix with $\alpha_i \in [0,1]$ being the amplitude and $\phi_i \in [0, 2\pi)$ being the phase induced by the ith RIS element on the impinging signal. Let x denote the transmit baseband signal such that $\mathbb{E}(|x|^2) = 1$. Then, the system model is given by

$$y = \sqrt{\rho}\mathbf{h}_2^H \mathbf{\Phi} \mathbf{h}_1 x + w, \tag{1}$$

where w is additive noise distributed $\mathcal{CN}(0,1)$. The system model can be written as

$$y = \sqrt{\rho} \sum_{i=1}^{N} h_1(i) h_2(i) \alpha_i e^{j\phi_i} x + w. \tag{2}$$

Let $g(i) \triangleq h_1(i)h_2(i)$ denote the product channel[3] corresponding to ith RIS element and let $\mathbf{g}^H \triangleq [g(1)\ g(2)\ \ldots\ g(N)]$ denote the $1 \times N$ product channel vector. Also, let $\boldsymbol{\theta} \triangleq [\alpha_1 e^{j\phi_1}\ \alpha_2 e^{j\phi_2}\ \ldots\ \alpha_N e^{j\phi_N}]^T$ denote the $N \times 1$ RIS vector. With this, the system model can be alternatively written as

$$y = \sqrt{\rho}\mathbf{g}^H \boldsymbol{\theta} x + w. \tag{3}$$

From (3), it is sufficient to estimate the product channel $\mathbf{g}$ for passive beamforming at RIS and for signal detection at receiver.

III. Channel Estimation

In this section, we present two training schemes that differ in their choice of RIS training states.

A. Canonical Training

In canonical training, the pilot signal $x = 1$ is transmitted over N training slots and the RIS elements are sequentially activated one at a time in each training slot with the remaining $N-1$ elements deactivated. In the ith training slot, $\alpha_i = 1$, $\phi_i = 0$, and $\alpha_j = 0, j \neq i$, and therefore, the RIS training state is given by

$$\boldsymbol{\theta}_i = [0\ \ldots \underbrace{1}_{i\text{th element}} \ldots\ 0]^T.$$

Let ρ_τ denote the training power. Then, from (3), the received signal in the ith training slot is given by

$$y(i) = \sqrt{\rho_\tau} g(i) + w(i).$$

From [18, Eq. (26)], the MMSE estimate for $g(i)$ is given by

$$\hat{g}(i) = \frac{\sqrt{\rho_\tau} y(i)}{\rho_\tau \mathbb{E}(|g(i)|^2) + 1}.$$

Since $\mathbb{E}(|g(i)|^2) = \mathbb{E}(|h_1(i)|^2 |h_2(i)|^2) = 1$,

$$\hat{g}(i) = \frac{\sqrt{\rho_\tau} y(i)}{\rho_\tau + 1}. \tag{4}$$

[2]As there are no well accepted, experimentally verified baseband channel models for RIS yet, we assume i.i.d Gaussian gains in this work.

[3]Also referred as cascaded channel in the literature.

B. DFT Training

In DFT training, pilot signal $x = 1$ is transmitted over the N training slots and the RIS training states are set equal to the columns of $N \times N$ DFT matrix. That is,

$$[\boldsymbol{\theta}_1\ \boldsymbol{\theta}_2\ \ldots\ \boldsymbol{\theta}_N] = \mathbf{F}_N,$$

where $\mathbf{F}_N$ is the $N \times N$ DFT matrix. Let $\mathbf{y}_\tau^H = [y(1)\ y(2)\ \ldots\ y(N)]$ be the received pilot sequence. Then, we have

$$\mathbf{y}_\tau^H = \sqrt{\rho_\tau}\mathbf{g}^H \mathbf{F}_N + \mathbf{w}^H,$$

where $\mathbf{w}^H = [w(1)\ w(2)\ \ldots\ w(N)]$. The MMSE estimate for $\mathbf{g}$ is obtained as [18]

$$\begin{aligned} \hat{\mathbf{g}}^H &= \mathbf{y}_\tau^H \left(\rho_\tau \mathbf{F}_N^H \mathbb{E}(\mathbf{g}\mathbf{g}^H)\mathbf{F}_N + \mathbf{I}_N\right)^{-1} \sqrt{\rho_\tau}\mathbf{F}_N^H \mathbb{E}(\mathbf{g}\mathbf{g}^H) \\ &= \frac{\sqrt{\rho_\tau}\mathbf{y}_\tau^H \mathbf{F}_N^H}{N\rho_\tau + 1}, \end{aligned} \tag{5}$$

where we have used $\mathbb{E}(\mathbf{g}\mathbf{g}^H) = \mathbf{I}_N$ which follows since

$$\mathbb{E}(g(i)g(j)^*) = \begin{cases} \mathbb{E}(|g(i)|^2) = 1 & \text{for } i = j \\ \mathbb{E}(h_1(i)h_1^*(j))\mathbb{E}(h_2(i)h_2^*(j)) = 0 & \text{for } i \neq j \end{cases}$$

IV. Spectral Efficiency with Training

Let $\tilde{\mathbf{g}} = \mathbf{g} - \hat{\mathbf{g}}$ denote the channel estimation error and ρ_d denote the data transmission power. Then, the received signal can be written as

$$\begin{aligned} y &= \sqrt{\rho_d}\hat{\mathbf{g}}^H \boldsymbol{\theta} x + \sqrt{\rho_d}\tilde{\mathbf{g}}^H \boldsymbol{\theta} x + w \\ &= \sqrt{\rho_d}\hat{\mathbf{g}}^H \boldsymbol{\theta} x + w', \end{aligned} \tag{6}$$

where $w' = \sqrt{\rho_d}\tilde{\mathbf{g}}^H \boldsymbol{\theta} x + w$. From the property of MMSE estimator, $\hat{\mathbf{g}}$ and $\tilde{\mathbf{g}}$ are uncorrelated. We assume that the estimated CSI at the receiver is communicated to the RIS through an error free channel for passive beamforming. With this we have the following results on the spectral efficiency of RIS-based system with canonical and DFT training.

Proposition 1. *The capacity of RIS-assisted system using canonical training satisfies the following lower bound:*

$$C \geq \mathbb{E}_{\hat{\mathbf{g}}} \log_2 \left(1 + \frac{\rho_d \left(\sum_{i=1}^{N} |\hat{g}(i)|\right)^2}{1 + \frac{N\rho_d}{1+\rho_\tau}}\right) \tag{7}$$

Proof. The capacity of RIS-assisted system with estimated CSIR is

$$C = \max_{p(x,\boldsymbol{\theta})} I(x; y, \hat{\mathbf{g}}),$$

where the maximization is over the joint distribution of input x and the RIS phase vector $\boldsymbol{\theta}$. We have,

$$\begin{aligned} I(x; y, \hat{\mathbf{g}}) &= I(x; y|\hat{\mathbf{g}}) + I(x; \hat{\mathbf{g}}) \\ &= I(x; y|\hat{\mathbf{g}}), \end{aligned}$$

where the last equality follows since x is independent of $\hat{\mathbf{g}}$. With this, we have

$$\begin{aligned} C &= \max_{p(x,\boldsymbol{\theta})} I(x; y|\hat{\mathbf{g}}) \\ &\geq \max_{\boldsymbol{\theta}(\hat{\mathbf{g}}), p(x)} I(x; y|\hat{\mathbf{g}}), \end{aligned} \tag{8}$$

where $\boldsymbol{\theta}(\hat{\mathbf{g}})$ is a deterministic function of the estimated channel $\hat{\mathbf{g}}$ and the inequality follows since we no longer jointly optimize over the distributions of x and $\boldsymbol{\theta}$. That is, we consider a scheme where RIS phases are set based on estimated channel, which is practical but not proven optimal and hence the inequality. From the worst case uncorrelated noise theorem [19, Theorem 1], if $\mathbb{E}(w'x^*|\hat{\mathbf{g}}) = 0$, then

$$\max_{p(x)} I(x; y|\hat{\mathbf{g}}) \geq \mathbb{E}_{\hat{\mathbf{g}}} \log_2 \left(1 + \frac{\rho_d |\hat{\mathbf{g}}^H \boldsymbol{\theta}(\hat{\mathbf{g}})|^2}{\sigma_{w'}^2}\right), \quad (9)$$

where $\sigma_{w'}^2$ is the variance of w'. To see that $\mathbb{E}(w'x^*|\hat{\mathbf{g}}) = 0$, we have

$$\begin{aligned}
\mathbb{E}(w'x^*|\hat{\mathbf{g}}) &= \mathbb{E}((\tilde{\mathbf{g}}^H \boldsymbol{\theta}(\hat{\mathbf{g}})x + w)x^*|\hat{\mathbf{g}}) \\
&= \mathbb{E}(\tilde{\mathbf{g}}^H \boldsymbol{\theta}(\hat{\mathbf{g}})|\hat{\mathbf{g}})\mathbb{E}(|x|^2|\hat{\mathbf{g}}) + \mathbb{E}(wx^*|\hat{\mathbf{g}}) \\
&= \mathbb{E}(\tilde{\mathbf{g}}^H|\hat{\mathbf{g}})\boldsymbol{\theta}(\hat{\mathbf{g}}) \\
&= \mathbb{E}((\mathbf{g} - \hat{\mathbf{g}})^H|\hat{\mathbf{g}})\boldsymbol{\theta}(\hat{\mathbf{g}}) \\
&= 0 \cdot \boldsymbol{\theta}(\hat{\mathbf{g}}) = 0,
\end{aligned}$$

where the final equality follows from the property of MMSE estimator and hence (9) holds. Now,

$$\begin{aligned}
\sigma_{w'}^2 &= \mathbb{E}(w'^H w'|\hat{\mathbf{g}}) \\
&= \mathbb{E}\big((\sqrt{\rho_d}\tilde{\mathbf{g}}^H \boldsymbol{\theta}(\hat{\mathbf{g}})x + w)^H (\sqrt{\rho_d}\tilde{\mathbf{g}}^H \boldsymbol{\theta}(\hat{\mathbf{g}})x + w)|\hat{\mathbf{g}}\big) \\
&= \rho_d \boldsymbol{\theta}(\hat{\mathbf{g}})^H \mathbb{E}(\tilde{\mathbf{g}}\tilde{\mathbf{g}}^H|\hat{\mathbf{g}})\boldsymbol{\theta}(\hat{\mathbf{g}}) + 1.
\end{aligned}$$

From [18, Eq. 27], we have

$$\mathbb{E}(\tilde{\mathbf{g}}\tilde{\mathbf{g}}^H|\hat{\mathbf{g}}) = \left(\frac{1}{1+\rho_\tau}\right)\mathbf{I}_N, \quad (10)$$

and hence

$$\begin{aligned}
\sigma_{w'}^2 &= \frac{\rho_d \|\boldsymbol{\theta}(\hat{\mathbf{g}})\|^2}{1+\rho_\tau} + 1 \\
&= \frac{N\rho_d}{1+\rho_\tau} + 1. \quad (11)
\end{aligned}$$

Substituting (11) in (9) and combining with (8), we have

$$C \geq \max_{\boldsymbol{\theta}(\hat{\mathbf{g}})} \mathbb{E}_{\hat{\mathbf{g}}} \log_2 \left(1 + \frac{\rho_d |\hat{\mathbf{g}}^H \boldsymbol{\theta}(\hat{\mathbf{g}})|^2}{1 + \frac{N\rho_d}{1+\rho_\tau}}\right). \quad (12)$$

It can be seen that the choice of RIS states (and hence $\boldsymbol{\theta}(\hat{\mathbf{g}})$) that maximize the rate term above is $\phi_i = -\angle \hat{g}(i),\ \alpha_i = 1,\ i = 1, \ldots, N$. Substituting these values in (12) proves the proposition. □

Proposition 2. *The capacity of RIS-assisted system using DFT training satisfies the following lower bound:*

$$C \geq \mathbb{E}_{\hat{\mathbf{g}}} \log_2 \left(1 + \frac{\rho_d \big(\sum_{i=1}^N |\hat{g}(i)|\big)^2}{1 + \frac{N\rho_d}{1+N\rho_\tau}}\right) \quad (13)$$

Proof. The proof follows similar steps as that of Proposition 1 with the only difference being that the covariance of estimation error for scheme 2 is

$$\mathbb{E}(\tilde{\mathbf{g}}\tilde{\mathbf{g}}^H|\hat{\mathbf{g}}) = \left(\frac{1}{1+N\rho_\tau}\right)\mathbf{I}_N, \quad (14)$$

and hence the effective noise variance becomes

$$\sigma_{w'}^2 = \frac{N\rho_d}{1+N\rho_\tau} + 1. \quad (15)$$

Using (15) in (9) and following similar arguments as in proposition 1 completes the proof. □

Remark 1. In canonical training, the transmitted pilot is reflected by only one RIS element in each training slot resulting in reduced received pilot energy. Whereas, in DFT training, the transmitted pilot is reflected by all the RIS elements in all the training slots, which shows up in the spectral efficiency expression for DFT training as a factor N multiplied to the pilot power ρ_τ.

Let T be the coherence interval (block length) of the product channel in number of channel uses. In both the training schemes discussed above, training consumes N channel uses. Let T_d denote the number of channel uses available for data transmission. Then, by conservation of time, we have:

$$T = N + T_d.$$

Also, by conservation of energy we have:

$$\rho T = \rho_\tau N + \rho_d T_d.$$

Accounting for the training time, the training-based capacity using canonical training is lower bounded as:

$$C_\tau \geq \left(1 - \frac{N}{T}\right)\mathbb{E}_{\hat{\mathbf{g}}} \log_2 \left(1 + \frac{\rho_d \big(\sum_{i=1}^N |\hat{g}(i)|\big)^2}{1 + \frac{N\rho_d}{1+\rho_\tau}}\right). \quad (16)$$

Similarly, the training-based capacity using DFT training is lower bounded as:

$$C_\tau \geq \left(1 - \frac{N}{T}\right)\mathbb{E}_{\hat{\mathbf{g}}} \log_2 \left(1 + \frac{\rho_d \big(\sum_{i=1}^N |\hat{g}(i)|\big)^2}{1 + \frac{N\rho_d}{1+N\rho_\tau}}\right). \quad (17)$$

Lemma 1. *The optimal data and training powers with canonical training are given by $\rho_d T_d = \beta_1^* \rho T$ and $\rho_\tau N = (1 - \beta_1^*)\rho T$, where*

$$\beta_1^* = \frac{\sqrt{\left(1 + \frac{\rho T}{N}\right)\left(1 + \frac{N\rho T}{T-N}\right)} - \left(1 + \frac{\rho T}{N}\right)}{\left(\frac{N\rho T}{T-N} - \frac{\rho T}{N}\right)} \quad (18)$$

Proof. The power allocation enters the spectral efficiency in (16) through the factor

$$\gamma \triangleq \frac{\rho_d \sigma_{\hat{g}}^2}{1 + \frac{N\rho_d}{1+\rho_\tau}},$$

where $\sigma_{\hat{g}}^2$ denotes the variance of estimate. If $\sigma_{\tilde{g}}^2$ denotes the variance of estimation error, then by orthogonality of MMSE estimator, $\sigma_{\hat{g}}^2 = 1 - \sigma_{\tilde{g}}^2 = \frac{\rho_\tau}{1+\rho_\tau}$. Therefore,

$$\gamma = \frac{\rho_d \rho_\tau}{1 + \rho_\tau + N\rho_d}.$$

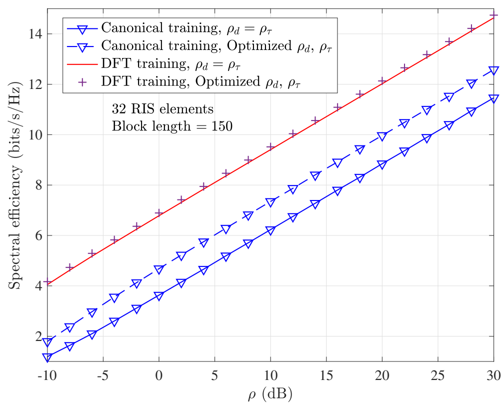


Fig. 1: Training-based bounds on capacity with canonical and DFT training.

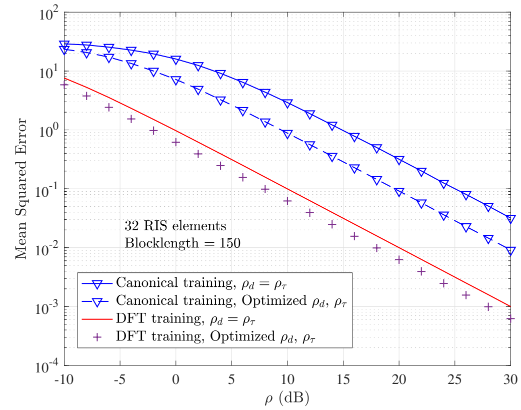


Fig. 2: MSE for canonical and DFT training.

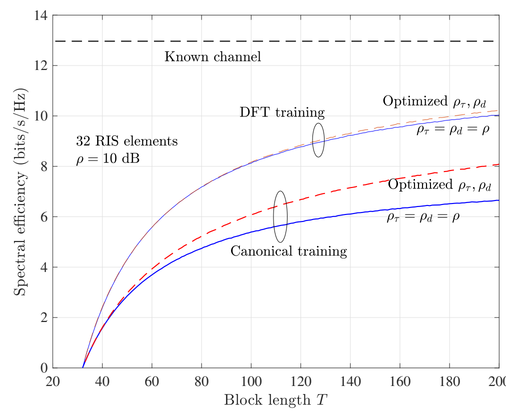


Fig. 3: Training-based bounds on capacity as a function of block length.

Let β_1 denote the fraction of total power allocated for data. That is, $\rho_d T_d = \beta_1 \rho T$. Similarly, we have $\rho_\tau N = (1-\beta_1)\rho T$ as the training power. With this, we have

$$\gamma = \frac{\frac{\beta_1 \rho T}{T_d} \frac{(1-\beta_1)\rho T}{N}}{1 + \frac{N\beta_1 \rho T}{T_d} + \frac{(1-\beta_1)\rho T}{N}}$$

Maximizing γ with respect to β_1 through differentiation proves the proposition. We skip these details for brevity. □

Lemma 2. *The optimal data and training powers with DFT training are given by* $\rho_d T_d = \beta_2^* \rho T$ *and* $\rho_\tau N = (1-\beta_2^*)\rho T$*, where*

$$\beta_2^* = \frac{\sqrt{(1+\rho T)\left(1 + \frac{N\rho T}{T-N}\right)} - (1+\rho T)}{\left(\frac{N\rho T}{T-N} - \rho T\right)}. \quad (19)$$

Proof. The proof follows similar ideas as that of Lemma 1 and hence skipped for brevity. □

V. Simulations and Discussions

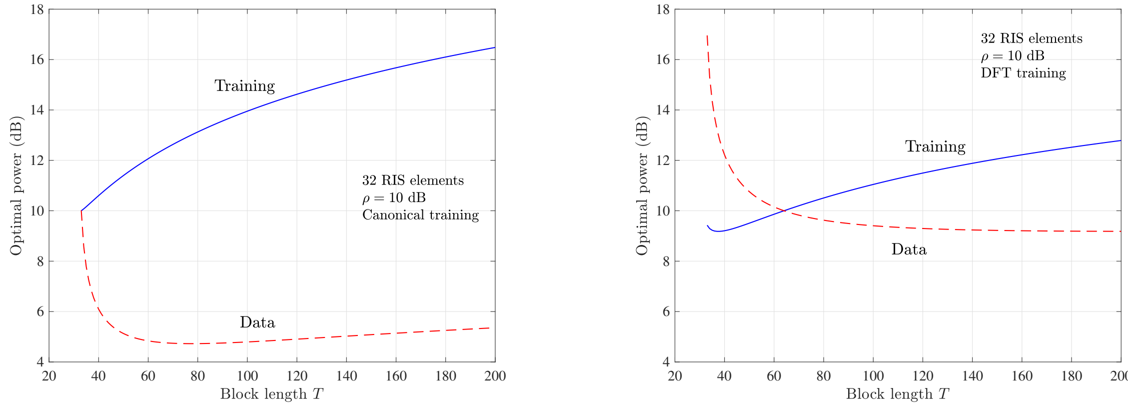
Figure 1 shows the training-based bounds on the capacity of RIS-assisted system with 32 RIS elements when the channel coherence interval is $T = 150$ slots. It can be seen that the spectral efficiency with DFT training is higher than that with canonical training. Specifically, with equal power allocation between training and data, DFT training achieves a gain of 3.5 bits/s/Hz compared to canonical training and with optimal power allocation it achieves a gain of 2 bits/s/Hz. This is due to the fact that the training power is not efficiently utilized in canonical training since the pilot signal is reflected by only one RIS element in each training slot. On the other hand, DFT training activates all the RIS elements in all the training slots, thereby efficiently utilizing the available training power.

Figure 2 shows the mean squared error (MSE) for canonical and DFT training schemes with equal and optimal power allocations. It can be seen that, the optimal power allocation, in addition to improving the spectral efficiency (as seen in Fig. 1), also leads to reduced MSE compared to equal power allocation for both the schemes.

Figure 3 shows the training-based bounds on capacity as a function block length T for RIS with 32 elements at $\rho = 10$ dB. It can be seen that DFT training achieves better performance across all T. Also, for both the schemes, the optimal power allocation results in higher gains in spectral efficiency at larger block lengths.

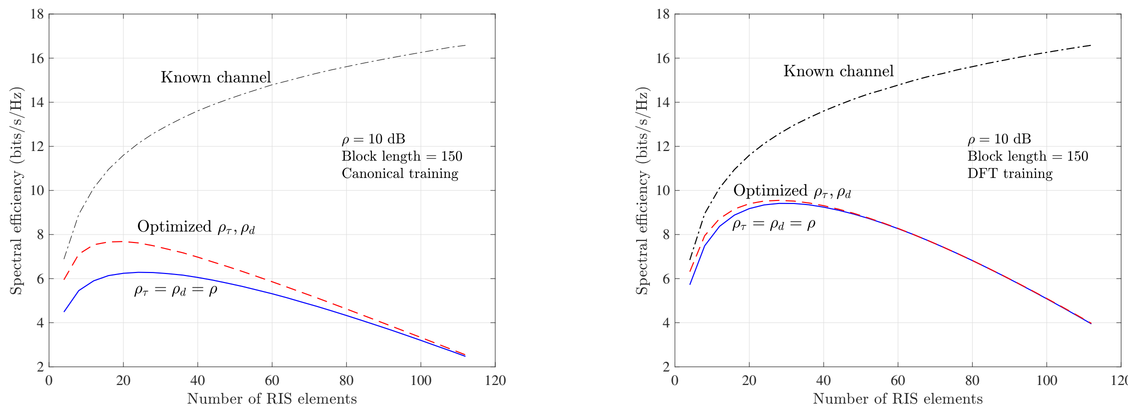
Figure 4 shows optimal training and data power allocation as a function of block length T for RIS with 32 elements at $\rho = 10$ dB. It can be seen that for canonical training, the optimal training power is greater than the optimal data power for all T. This is because, the power allocation compensates for the lesser received pilot power by a higher transmit pilot power for achieving good training quality. Whereas, with DFT training, it can be seen that $\rho_d > \rho_\tau$ for $T < 64$ and $\rho_\tau > \rho_d$ for $T > 64$. This can be explained as follows. For $T < 64$, the number of data slots $T_d < 32$ since $N = 32$ and $T = N + T_d$. That is, more than half the block length is used for training, which is compensated by allocating higher power for data. Whereas for $T > 64$, more slots are available for data and hence higher power is allocated for training to obtain reliable channel estimates, which can then be used during data transmission.

Figure 5 shows the training-based bounds on capacity as a



(a) Canonical training (b) DFT training

Fig. 4: Optimal power allocation between training and data as a function of block length.



(a) Canonical training (b) DFT training

Fig. 5: Training-based bounds on capacity as a function of number of RIS elements.

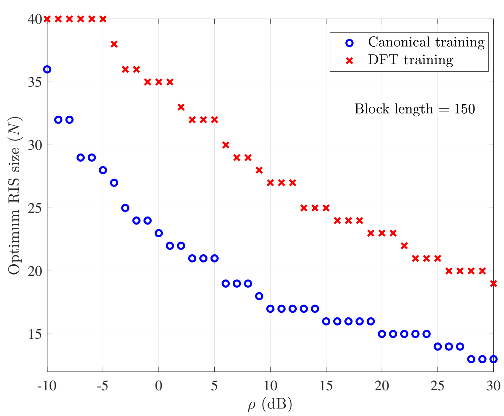


Fig. 6: Optimum RIS size that maximizes spectral efficiency.

function of number of RIS elements for $\rho = 10$ dB and block length $T = 150$. From the figure, it can be seen that for both canonical and DFT training schemes, the spectral efficiency increases up to a certain critical value of N beyond which increasing the RIS size is unhelpful. This is because, beyond the critical RIS array size, training overhead consumes more time, leaving less time for RIS beamforming to be exploited, as a result the spectral efficiency suffers. Figure 6 shows the optimum RIS array size as a function of SNR. It can be seen that the optimum array size goes down with increase in SNR. This is because, at low SNR, power is at a premium while degrees of freedom are not, therefore a larger RIS is beneficial. At high-SNR, training overhead becomes more important, and therefore larger array is unhelpful.

VI. Conclusions

We analyzed the spectral efficiency of RIS with canonical and DFT training and showed that canonical training, although intuitive, is inefficient compared to DFT training. Our simulations revealed that, beyond a certain critical RIS size, further increase in number of elements in unhelpful and that the optimum RIS array size goes down with increase in SNR. The present work did not consider optimization over training time, which could be a potential future work. Extending the analysis to multiple antenna systems and reduced overhead training schemes are also topics of future research.

REFERENCES

[1] M. Di Renzo, A. Zappone, M. Debbah, M.-S. Alouini, C. Yuen, J. de Rosny, and S. Tretyakov, "Smart radio environments empowered by reconfigurable intelligent surfaces: How it works, state of research, and the road ahead," *IEEE Journal on Selected Areas in Communications*, vol. 38, no. 11, pp. 2450–2525, 2020.

[2] Y. Liu, X. Liu, X. Mu, T. Hou, J. Xu, M. Di Renzo, and N. Al-Dhahir, "Reconfigurable intelligent surfaces: Principles and opportunities," *IEEE Communications Surveys Tutorials*, vol. 23, no. 3, pp. 1546–1577, 2021.

[3] E. Basar, M. Di Renzo, J. De Rosny, M. Debbah, M.-S. Alouini, and R. Zhang, "Wireless communications through reconfigurable intelligent surfaces," *IEEE Access*, vol. 7, pp. 116 753–116 773, 2019.

[4] Q. Wu and R. Zhang, "Towards smart and reconfigurable environment: Intelligent reflecting surface aided wireless network," *IEEE Communications Magazine*, vol. 58, no. 1, pp. 106–112, 2020.

[5] C. Pan, H. Ren, K. Wang, J. F. Kolb, M. Elkashlan, M. Chen, M. Di Renzo, Y. Hao, J. Wang, A. L. Swindlehurst, X. You, and L. Hanzo, "Reconfigurable intelligent surfaces for 6G systems: Principles, applications, and research directions," *IEEE Communications Magazine*, vol. 59, no. 6, pp. 14–20, 2021.

[6] E. Björnson, O. Özdogan, and E. G. Larsson, "Reconfigurable intelligent surfaces: Three myths and two critical questions," *IEEE Communications Magazine*, vol. 58, no. 12, pp. 90–96, 2020.

[7] X. Ma, S. Guo, H. Zhang, Y. Fang, and D. Yuan, "Joint beamforming and reflecting design in reconfigurable intelligent surface-aided multi-user communication systems," *IEEE Transactions on Wireless Communications*, vol. 20, no. 5, pp. 3269–3283, 2021.

[8] B. Di, H. Zhang, L. Song, Y. Li, Z. Han, and H. V. Poor, "Hybrid beamforming for reconfigurable intelligent surface based multi-user communications: Achievable rates with limited discrete phase shifts," *IEEE Journal on Selected Areas in Communications*, vol. 38, no. 8, pp. 1809–1822, 2020.

[9] J. Wang, H. Wang, Y. Han, S. Jin, and X. Li, "Joint transmit beamforming and phase shift design for reconfigurable intelligent surface assisted MIMO systems," *IEEE Transactions on Cognitive Communications and Networking*, vol. 7, no. 2, pp. 354–368, 2021.

[10] J. Ye, S. Guo, and M.-S. Alouini, "Joint reflecting and precoding designs for ser minimization in reconfigurable intelligent surfaces assisted MIMO systems," *IEEE Transactions on Wireless Communications*, vol. 19, no. 8, pp. 5561–5574, 2020.

[11] X. Wei, D. Shen, and L. Dai, "Channel estimation for RIS assisted wireless communications—part i: Fundamentals, solutions, and future opportunities," *IEEE Communications Letters*, vol. 25, no. 5, pp. 1398–1402, 2021.

[12] Z. Zhou, N. Ge, Z. Wang, and L. Hanzo, "Joint transmit precoding and reconfigurable intelligent surface phase adjustment: A decomposition-aided channel estimation approach," *IEEE Transactions on Communications*, vol. 69, no. 2, pp. 1228–1243, 2021.

[13] H. Liu, X. Yuan, and Y.-J. A. Zhang, "Matrix-calibration-based cascaded channel estimation for reconfigurable intelligent surface assisted multiuser MIMO," *IEEE Journal on Selected Areas in Communications*, vol. 38, no. 11, pp. 2621–2636, 2020.

[14] L. Wei, C. Huang, G. C. Alexandropoulos, C. Yuen, Z. Zhang, and M. Debbah, "Channel estimation for RIS-empowered multi-user MISO wireless communications," *IEEE Transactions on Communications*, vol. 69, no. 6, pp. 4144–4157, 2021.

[15] J. He, H. Wymeersch, and M. Juntti, "Channel estimation for RIS-aided mmwave MIMO systems via atomic norm minimization," *IEEE Transactions on Wireless Communications*, pp. 1–1, 2021.

[16] Y. Liu, S. Zhang, F. Gao, J. Tang, and O. A. Dobre, "Cascaded channel estimation for RIS assisted mmwave MIMO transmissions," *IEEE Wireless Communications Letters*, pp. 1–1, 2021.

[17] T. L. Jensen and E. De Carvalho, "An optimal channel estimation scheme for intelligent reflecting surfaces based on a minimum variance unbiased estimator," in *ICASSP 2020 - 2020 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, 2020, pp. 5000–5004.

[18] M. Biguesh and A. Gershman, "Training-based MIMO channel estimation: a study of estimator tradeoffs and optimal training signals," *IEEE Transactions on Signal Processing*, vol. 54, no. 3, pp. 884–893, 2006.

[19] B. Hassibi and B. Hochwald, "How much training is needed in multiple-antenna wireless links?" *IEEE Transactions on Information Theory*, vol. 49, no. 4, pp. 951–963, 2003.