Reconfigurable Intelligent Surface-Aided OFDM Wireless Communications: Hardware Aspects of Reflection Optimization Methods

Dimitris Vordonis, Dimitris Kompostiotis, and Vassilis Paliouras
Electrical and Computer Engineering Department, University of Patras, Greece
Email: d.vordonis@upnet.gr, d.kompostiotis@upnet.gr, paliuras@ece.upatras.gr

***Abstract*—A Reconfigurable Intelligent Surface (RIS) is a passive system node, envisioned as a new physical layer technology in sixth-generation (6G) infrastructure. A RIS enables smart propagation environments by tuning the signal reflection in real time. Reflection optimization is an active field of research, regarding the RIS integration into wireless networks. A real-time response is required, while the unit-modulus constraint on the phase shift introduced by RIS elements makes the optimization problem even more challenging. In this paper, both low-complexity and efficient, in terms of achievable rate, passive beamforming methods are evaluated over a frequency-selective fading channel. Practical aspects of a RIS, including binary phase shifts with unbalanced amplitudes and mutual coupling, are considered. A hardware-oriented reflection optimization method is proposed and is implemented on a Zynq UltraScale+ multiprocessor system-on-a-chip (MPSoC) device. The proposed architecture leads to an extremely low execution time, useful especially for high mobility scenarios, in which coherence time is a tough constraint.**

***Index Terms*—6G, Reconfigurable Intelligent Surfaces, OFDM, reflection optimization, discrete phase shifts, mutual coupling, unimodular quadratic problems, hardware implementation.**

I. Introduction

Reconfigurable Intelligent Surface (RIS) technology has emerged as a prominent candidate for enhancing the spectral and the energy efficiency of future wireless communication systems [1], [2]. RIS is strategically placed in the infrastructure and its main goal is to send the incident signal, at each element of the surface, to the transmitter introducing an optimal phase shift so that the signal power is maximized at the receiving point. Most research works, solving the RIS reflection optimization problem, assume frequency-flat fading channels, but practical channels are frequency selective. The passive beamforming solution must be common for all sub-carriers in an Orthogonal Frequency Division Multiplexing (OFDM) block, despite that the RIS configuration is directly correlated to the channel coefficients at each frequency [1]–[4]. Therefore, the lack of the frequency-selective RIS reflection generates additional constraints on the reflection optimization problem for the case of OFDM transmission.

This research has been co-financed by the European Union and Greek national funds through the Operational Program Competitiveness, Entrepreneurship and Innovation, under the call RESEARCH-CREATE-INNOVATE (project code: T1EDK-02551).

A common objective for each passive beamforming method is either to maximize the achievable rate or to maximize the received signal power at the end-user. Under certain conditions, the latter leads to a uni-modular quadratic program (UQP) optimization that is mainly a non-deterministic polynomial time hard (NP-hard) problem [5]. Uni-modular constraints, which apply on the phase shift solution vector, are non-convex [1], [5]. Power iteration-based techniques are proposed for local optimization of UQP [3], [6]–[8], providing a sub-optimal solution but also meeting the real-time requirements. Some state-of-the-art reflection optimization techniques include semidefinite relaxation (SDR) or successive convex approximation (SCA) [4], [9]. These methods derive an approximate upper bound in the maximization problem, but the complexity remains high. On the other hand, Strongest Tap Maximization (STM) in the time domain [2] and Greedy Fast Beamforming Algorithm (GFBA) [10] are low-complexity heuristic methods, utilized in RIS-aided OFDM systems.

Fast time-varying channels pose a challenge for RIS technology, due to the required real-time adaptive reflection. Channel estimation and adjustment of the reflection coefficients should not cumulatively exceed the coherence time. Typical channel coherence time is on the order of millisecond (ms) [1], but in future 6G use cases, including high-mobility scenarios, the coherence time may be smaller. Therefore, dedicated hardware is an enabler to achieve such stringent requirements. A field-programmable gate array (FPGA) implementation is feasible placed either at RIS or at the base station side.

In this paper, we study the efficiency of power iteration-based methods in the received signal power maximization problem, assuming a RIS-aided OFDM system. Apart from state-of-the-art passive beamforming methods, general techniques that solve the UQP problem, are also examined. Binary phase shifts with unbalanced amplitudes and mutual coupling (MC) are considered, while a method that is easily implemented on hardware is proposed. Furthermore, proposed architectural optimizations, related to memory and loop dependencies management, lead to a low latency-oriented solution. Thus, the implementation of the proposed method, on a Zynq UltraScale+ multiprocessor system on a chip (MPSoC) device, results in an extremely low execution time.

The remainder of this paper is structured as follows: Sec-

tion II reviews the system model and formulates the examined optimization problem. Section III evaluates power-iteration-based reflection optimization methods, in terms of achievable rate, and proposes a hardware-oriented method. Section IV details the hardware architecture of the proposed method. Architectural optimizations lead to a low-latency solution, while implementation results are also revealed. Finally, Section V discusses conclusions.

II. System Model

In this work, to facilitate reproducibility of the results, the system model of IEEE Signal Processing Cup 2021 [3], [11] is considered. Specifically, one RIS of $N = 4096$ elements is deployed between a single-antenna access point (AP) and multiple single-antenna users. Provided that only one user is served at a time, each user can be considered individually. Let N_V denote the number of rows and N_H the number of columns of the RIS, where $N_V \times N_H = N$. The employed RIS has the form of a uniform planar array with $N_V = N_H = 64$. Björnson states that the AP has a line-of-sight (LoS) channel to the RIS, while the users have either a LoS or a non-line-of-sight (NLoS) channel to the RIS in [3], [11]. Between the AP and the users, we consider a very weak channel. The transmission of an OFDM block is described by

$$\overline{\boldsymbol{z}} = \overline{\boldsymbol{h}}_{\boldsymbol{\theta}} \odot \overline{\boldsymbol{x}} + \overline{\boldsymbol{w}}, \tag{1}$$

where $\overline{\boldsymbol{z}} = [\overline{z}[0], \ldots, \overline{z}[K-1]]^{\mathsf{T}}$, and $\overline{z}[i]$ denote the received signal for each OFDM sub-carrier, $\overline{\boldsymbol{x}} = [\overline{x}[0], \ldots, \overline{x}[K-1]]^{\mathsf{T}}$ is the transmitted signal, $\overline{\boldsymbol{w}} = [\overline{w}[0], \ldots, \overline{w}[K-1]]^{\mathsf{T}}$ is the Additive White Gaussian Noise (AWGN), K is the number of sub-carriers and $\odot$ denotes the Hadamard product. The frequency domain channel model $\overline{\boldsymbol{h}}_{\boldsymbol{\theta}}$ is given by

$$\overline{\boldsymbol{h}}_{\boldsymbol{\theta}} = \boldsymbol{F}(\boldsymbol{h}_d + \boldsymbol{V}^{\mathsf{T}}\boldsymbol{\omega}_{\boldsymbol{\theta}}) \tag{2}$$

where $\boldsymbol{F} \in \mathbb{C}^{K \times M}$, $\boldsymbol{h}_d \in \mathbb{C}^{M}$, $\boldsymbol{V}^{\mathsf{T}} \in \mathbb{C}^{M \times N}$ and $\boldsymbol{\omega}_{\boldsymbol{\theta}} = [e^{\mathrm{j}\theta_1}, \ldots, e^{\mathrm{j}\theta_N}]^{\mathsf{T}} \in \mathbb{C}^{N}$ denotes the Discrete Fourier Transform (DFT) matrix, the direct channel from AP to the user equipment (UE), the cascaded from the AP to RIS and from RIS to UE channel, and the RIS configuration, respectively. M denotes the number of channel filter taps in the time domain.

As shown from (2), when the signal is transmitted from the AP to the UE, a part of the channel impulse response is determined by the RIS configuration. The maximum received power formulation of the problem, where power iteration methods are based, is given by

$$\underset{\boldsymbol{\omega}_{\boldsymbol{\theta}}}{\operatorname{argmax}} \|\overline{\boldsymbol{h}}_{\boldsymbol{\theta}}\|^2 = \underset{\boldsymbol{\omega}_{\boldsymbol{\theta}}}{\operatorname{argmax}} \, \boldsymbol{c}^{\mathsf{H}}\boldsymbol{B}\boldsymbol{c} =$$

$$\underset{\boldsymbol{\omega}_{\boldsymbol{\theta}}}{\operatorname{argmax}} \begin{bmatrix} 1 \\ \frac{\boldsymbol{A}^{\mathsf{T}}\boldsymbol{\omega}_{\boldsymbol{\theta}}}{N_V} \end{bmatrix}^{\mathsf{H}} \underbrace{[\boldsymbol{h}_d, N_V\boldsymbol{V}_{\text{row}}^{\mathsf{T}}]^{\mathsf{H}}[\boldsymbol{h}_d, N_V\boldsymbol{V}_{\text{row}}^{\mathsf{T}}]}_{\boldsymbol{B}} \begin{bmatrix} 1 \\ \frac{\boldsymbol{A}^{\mathsf{T}}\boldsymbol{\omega}_{\boldsymbol{\theta}}}{N_V} \end{bmatrix}, \tag{3}$$

where $\boldsymbol{V} = \underbrace{(\boldsymbol{1}_{N_V} \otimes \boldsymbol{I}_{N_H})}_{\boldsymbol{A}} \boldsymbol{V}_{\text{row}}$. This is a reduced dimensionality model, suggested from Björnson [3]. It is assumed that each column of the RIS has the same configuration. Matrix $\boldsymbol{B}$ in (3), is alternatively defined as in [9]. The statistical correlation model, described in [3], is employed in this work. Thus, practical MC effect is introduced between the adjacent RIS elements. Finally, finite resolution phase shifters are assumed, where the phase shifting value θ_n for each n-th RIS element is $\theta_n \in \{0, \pi\}$.

Algorithm 1: Proposed Beamforming Method

Input: $\boldsymbol{V}_{\text{row}}^{\mathsf{T}}, \boldsymbol{h}_d, \boldsymbol{a}$
Output: The passive beamforming solution $\boldsymbol{a}_{\text{new}}$
$\boldsymbol{B} = \boldsymbol{D}^H\boldsymbol{D}$, where $\boldsymbol{D} = [\boldsymbol{h}_d, N_V\boldsymbol{V}_{\text{row}}^{\mathsf{T}}]$;
$\boldsymbol{a}_{\text{new}} = \text{quantize}(\boldsymbol{B}\boldsymbol{a})$;

III. Reflection Optimization Methods

The experimental results derived in this paper are based on the dataset of IEEE Signal Processing Cup 2021 [11]. This work considers a RIS-aided OFDM system and focuses in low-complexity yet efficient in terms of achievable rate, reflection optimization methods. Perfect channel state information is assumed for all the examined reflection optimization techniques.

Apart from state-of-the-art passive beamforming techniques, namely SDR, SCA, STM and GFBA, power iteration-based methods are utilized in reflection optimization. These methods maximize suboptimally the UQP problem, defined in (3), with polynomial time complexity. Ragi *et al.* [7] propose Dominant Eigenvector Matching (DEM) method. This method computes the dominant eigenvector of $\boldsymbol{B}$, before the matching process [7]. The method maximizes the objective value $\boldsymbol{c}^{\mathsf{H}}\boldsymbol{B}\boldsymbol{c}$ in a non-iterative way, provided that the dominant eigenvector is computed. Yu *et al.* [8] and Soltanalian *et al.* [6] define Fixed Point Iteration (FPI) method and Power method respectively, in order to maximize a UQP problem. The objective value increases through iterations until convergence in both power iteration-based techniques. FPI in each iteration computes $\mathbf{c}^{(\mathbf{t+1})}$, dividing each element of $\mathbf{B}\mathbf{c}^{(\mathbf{t})}$ vector with its norm metric, where t denotes the iteration step. On the other hand, Power method computes $e^{j \arg(\mathbf{B}\mathbf{c}^{(\mathbf{t})})}$ in each iteration. Furthermore, Björnson *et al.* [3] proposes a power iteration-based technique, called in this paper as Björnson's method. This iterative method rotates the result of $\mathbf{B}\mathbf{c}^{(\mathbf{t})}$ in each repetition, so that $\mathbf{c}^{(\mathbf{t})}$ agrees with (3). In addition, the quantization is applied after each iteration of the algorithm, in contrast with the other power iteration-based methods, investigated in this paper, where the quantization takes place after convergence.

A. Proposed Method

Going from software to a hardware implementation, there are some features of algorithms in [3], [6]–[8] that do not make them a hardware friendly choice. DEM, apart from a module that estimates the dominant eigenvector of $\mathbf{B}$, requires the $\arctan$ computation [7]. This results in both high complexity and the necessity of making a linear approximation of the function. Power method, also requires the $\arctan$ and $\exp$ computation [6]. On the other hand, FPI requires N_H divisions at each iteration [8]. These divisions add an important overhead to the architecture and as a result latency is increased. Finally, the iterative part of Björnson's method [3], does not ensure that the rate value is increased per iteration. Furthermore, the calculation of the estimated rate performance is needed at each iteration.

The proposed passive beamforming algorithm is described in Alg. 1. The best pilot configuration $\boldsymbol{a}$ is utilized as an

initial solution $\mathbf{c}^{(\mathbf{0})}$. The low-complexity algorithm in Alg. 1 does not include an iterative part and it does not require a hardware module for the computation of a complex mathematical function. There is also no configuration rotation stage as described in [3], since the result does not change if the process is skipped. In fact the rotation results in the reflected signals arriving with different phases than before, but the phase differences between the signals remain the same. Therefore, the overall received signal does not change.

B. Achievable rate

Table I depicts the achievable rate of the proposed method, using the dataset in [11]. Both state-of-the-art reflection optimization methods and techniques for UQP problems are also examined. The rate is computed according to [3], [11]. As expected LoS users achieve a higher rate, compared to NLoS ones. The weight of NLoS users is double for the computation of average metric, in agreement with [3], [11]. Apart from upper bound, all depicted methods assume binary phase elements and same reflection response in each RIS column. The upper bound theoretical value, defined in [2], considers reflection optimization of each sub-carrier individually, proving the performance degradation due to the lack of frequency-selective reflection of RIS technology. Furthermore, Table I shows the effect of MC. Lower rates are achieved, when neighboring RIS elements are correlated. The proposed method, together with FPI [8] and Power method [6] are superior compared to the examined methods, in the case of LoS users. Contrariwise, it slightly lags behind in the NLoS scenario. Based on simulation results presented here and regarding the optimization problem of this paper, FPI requires two iterations to converge, while for Power method one iteration is enough. However, the examined techniques in [8] and [6] are more complex. As best pilot method, we denote the RIS configuration that led to the highest measured received power, during channel estimation process. This solution vector used to initialize all the examined power iteration-based methods, including the proposed one. Finally, the significant contribution of RIS node is proved, observing the score metric of the uniform surface scenario, in which all reflecting coefficients equal to -1.

IV. Hardware Implementation

A. Proposed Architecture

Fig. 1 shows the overall diagram of the proposed reflection optimization architecture, containing the RAMs (Vt.re, Vt.im, a), where the input data are stored and the main Power Method Hardware (PMHw) component. PMHw consists of three modules, namely the matrix-matrix multiplication unit (MMMU), the matrix-vector multiplication unit (MVMU) and the quantizer. Communication between the modules is achieved through First In First Out (FIFO) logic. First, inputs $\boldsymbol{V}_{\text{row}}^{\mathsf{T}}$ and $\boldsymbol{h}_d$ are processed to obtain matrix $\boldsymbol{D}$. Then the $\boldsymbol{B} = \boldsymbol{D}^H \boldsymbol{D}$ computation follows. After that, a matrix-vector multiplication of $\boldsymbol{B}$ with $\boldsymbol{a} = \left[1 \ (\frac{\boldsymbol{A}^{\mathsf{T}} \boldsymbol{\omega_\theta}}{N_V})^T\right]^{\mathsf{T}}$ takes place and finally the resulting configuration is quantized so that each value $\in \{-1, +1\}$. Fixed-point arithmetic is used to reduce complexity and power consumption and further increase speed.

Because $\boldsymbol{D} \in \mathbb{C}^{M \times (N_H+1)}$, calculating $\boldsymbol{B}$ requires complex operations. Since only the real part of the result is required to produce the output, the multiplications were limited to two real for each complex multiplication. In addition $\boldsymbol{D}$ is partitioned into M smaller memories, so that the whole column of the matrix $\boldsymbol{D}$ can be read in only one clock cycle. Also each partial memory is a dual port RAM, so that two columns of matrix $\boldsymbol{D}$ can be read simultaneously in one cycle and therefore, the computation of their inner product is accelerated. Finally, the MVMU is pipelined, so that reading the memory locations to be multiplied, performing the operations and storing the result in memory are done at overlapping times. Therefore the throughput of the architecture is further increased.

B. Implementation Results

The implementation results, depicted in Table II, are obtained by synthesis using Mentor Graphics' Catapult [12] as a front-end high-level synthesis tool. For the measurements, a Zynq UltraScale+ ZU11EG MPSoC device was used. Both a generic and a latency-optimized solution are implemented. Table II shows the maximum clock frequency ($f_{\max}$), the required clock cycles in order to execute the passive beamforming algorithm, as well as the execution time in ms. The required area of the design, the number of multipliers and adders, and the employed word format are also reported. Furthermore, the power consumption is depicted.

Due to the proposed architectural optimizations detailed in Section IV-A, a significant decrease in latency is observed in the case of the optimized solution, compared to the generic one. Aiming to a more parallel architecture, more hardware resources are required. Therefore an increase in the number of computation units and consequently in total area utilization is observed. Loop unrolling applied in the MMMU, provokes an increase of the critical path delay of the design. As a result, $f_{\max}$ drops, but the overall achieved execution time improves. Due to the increased area of the parallel architecture, power consumption is increased, but this is covered, because $f_{\max}$ drops. So power consumption is preserved (Table II).

As mentioned, execution time of passive beamforming algorithm should not exceed the channel coherence time. Processing regarding channel estimation takes also place within the coherence time, hence a small execution time is crucial. Execution time results for both an SDR (via CVX [13]) and a Deep Learning (DL) based approach are reported in [14], utilizing an Intel i7-8700 CPU. Signal-to-Noise Ratio (SNR) maximization in a multi-input-single-output (MISO) system with a single-antenna user leads to a UQP problem, similar to (3). The noticeable difference is in the dimensions of the problem, since in our case we depend on M while in MISO system on the number of antennas Q at AP. Assuming $Q = 8$ and $N = 64$, the DL-based method achieved ≈ 0.12 ms, considering offline training though [14]. Contrariwise, SDR achieved 715 ms execution time, without real-time constraints being met. In this work, due to the assumption of same reflection response in each RIS column, N can be considered equal to 64 but $M = 20$ that is $> Q$, compared to [14]. Assuming a system model similar to this work, He *et al.* reveal

TABLE I
ACHIEVABLE RATE FOR EACH REFLECTION OPTIMIZATION TECHNIQUE USING THE DATASET IN [11]

Method	Average (Mbit/s)		LoS users (Mbit/s)		NLoS users (Mbit/s)	
	no MC	under MC	no MC	under MC	no MC	under MC
Upper Bound [2]	150.17	–	139.84	–	88.36	–
Fixed Point Iteration [8]	126.42	115.81	124.27	114.51	65.97	59.58
Power method [6]	126.40	115.77	124.27	114.51	65.93	59.52
Proposed method	125.93	115.33	124.26	114.49	65.11	58.74
Björnson's method [3]	125.50	114.79	124.09	114.52	64.56	57.73
STM [2]	125.20	114.22	123.81	114.15	64.39	57.20
GFBA [10]	125.13	114.29	124.07	114.57	63.93	56.78
Dominant Eigenvector Matching [7]	124.31	113.31	122.48	112.67	64.51	57.48
Best Pilot [3]	113.18	103.56	117.61	107.98	50.90	46.08
Uniform Surface	37.24	–	32.03	–	25.32	–

TABLE II
HARDWARE IMPLEMENTATION RESULTS

Architecture	$f_{\max}$ (MHz)	Latency (cycles)	Exec.time (ms)	Total Area*	Multipliers	Adders	Word format†	Power consumption (W)
Generic	403.23	363218	≈ 0.90	2856.59	3	12	1.17	0.850
Optimized**	227.27	32146	≈ 0.14	22459.46	22	174	1.17	0.936

*Measured in area units [12]. †Word format $i.f$ denotes i integral and f fractional bits. **Latency-oriented solution

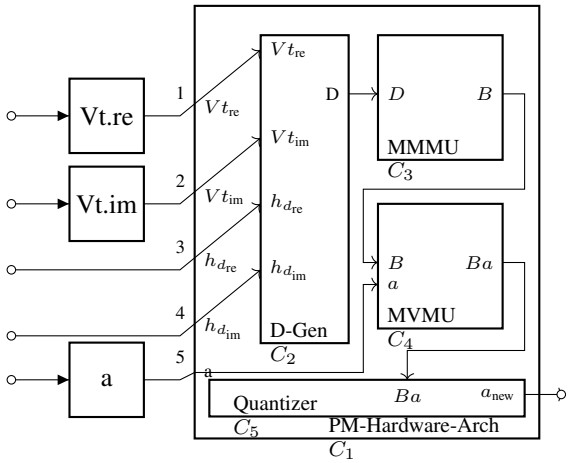

Fig. 1. High-level block diagram

the average running time of both SCA and a majorization-minimization (MM)-based iterative approach, based on an i7-4790 CPU [15]. SCA does not meet the real-time response, while the MM-based approach achieves 130 ms, assuming $M = 30$ and $N = 32$. Yang *et al.* report execution time results of beamforming algorithms, concerning a multi-user multiple-input-multiple-output (MIMO) RIS-aided system [16]. Finally, Pei *et al.* [10] implement GFBA on FPGA RIS controller, reporting 1.5 W power consumption. This measurement concerns a RIS prototype with 1100 elements, of which every five in each column share the same bias voltage.

V. CONCLUSION

This paper proves the efficiency of power iteration-based methods as a RIS reflection optimization solution in UQP problems. A low-complexity hardware-oriented passive beamforming method and its architectural optimizations are proposed, achieving an execution of 0.14 ms. The disclosed results demonstrated the great potential of a hardware implementation, in the effort of meeting real-time constraints.

VI. ACKNOWLEDGEMENTS

D. Kompostiotis' work was financially supported by the "Andreas Mentzelopoulos Foundation."

REFERENCES

[1] Q. Wu, S. Zhang, B. Zheng, C. You, and R. Zhang, "Intelligent reflecting surface-aided wireless communications: A tutorial," *IEEE Transactions on Communications*, vol. 69, no. 5, pp. 3313–3351, 2021.

[2] E. Björnson, H. Wymeersch, B. Matthiesen, P. Popovski, L. Sanguinetti, and E. de Carvalho, "Reconfigurable intelligent surfaces: A signal processing perspective with wireless applications," *IEEE Signal Processing Magazine*, vol. 39, no. 2, pp. 135–158, 2022.

[3] E. Björnson, "Optimizing a binary intelligent reflecting surface for OFDM communications under mutual coupling," *arXiv preprint arXiv:2106.04280*, 2021.

[4] Y. Yang, B. Zheng, S. Zhang, and R. Zhang, "Intelligent reflecting surface meets OFDM: Protocol design and rate maximization," *IEEE Transactions on Communications*, vol. 68, no. 7, pp. 4522–4535, 2020.

[5] N. K. Kundu and M. R. McKay, "RIS-assisted MISO communication: Optimal beamformers and performance analysis," in *2020 IEEE Globecom Workshops (GC Wkshps*. IEEE, 2020, pp. 1–6.

[6] M. Soltanalian and P. Stoica, "Designing unimodular codes via quadratic optimization," *IEEE Transactions on Signal Processing*, vol. 62, no. 5, pp. 1221–1234, 2014.

[7] S. Ragi, E. K. Chong, and H. D. Mittelmann, "Polynomial-time methods to solve unimodular quadratic programs with performance guarantees," *IEEE Transactions on Aerospace and Electronic Systems*, vol. 55, no. 5, pp. 2118–2127, 2018.

[8] X. Yu, D. Xu, and R. Schober, "MISO wireless communication systems via intelligent reflecting surfaces," in *2019 IEEE/CIC International Conference on Communications in China (ICCC)*. IEEE, 2019, pp. 735–740.

[9] Q. Wu and R. Zhang, "Intelligent reflecting surface enhanced wireless network via joint active and passive beamforming," *IEEE Transactions on Wireless Communications*, vol. 18, no. 11, pp. 5394–5409, 2019.

[10] X. Pei, H. Yin, L. Tan, L. Cao, Z. Li, K. Wang, K. Zhang, and E. Björnson, "RIS-aided wireless communications: Prototyping, adaptive beamforming, and indoor/outdoor field trials," *IEEE Transactions on Communications*, vol. 69, no. 12, pp. 8627–8640, 2021.

[11] E. Björnson, "Configuring an intelligent reflecting surface for wireless communications," *arXiv preprint arXiv:2106.03497*, 2021.

[12] M. Fingeroff, *High-level synthesis: blue book*. Xlibris Corporation, 2010.

[13] M. Grant and S. Boyd, "CVX: Matlab software for disciplined convex programming, version 2.1," 2014.

[14] J. Gao, C. Zhong, X. Chen, H. Lin, and Z. Zhang, "Unsupervised learning for passive beamforming," *IEEE Communications Letters*, vol. 24, no. 5, pp. 1052–1056, 2020.

[15] Z. He, H. Shen, W. Xu, and C. Zhao, "Low-cost passive beamforming for RIS-aided wideband OFDM systems," *IEEE Wireless Communications Letters*, 2021.

[16] H. Yang, X. Yuan, J. Fang, and Y.-C. Liang, "Reconfigurable intelligent surface aided constant-envelope wireless power transfer," *IEEE Transactions on Signal Processing*, vol. 69, pp. 1347–1361, 2021.