Article

Complex-Valued Phase Transmittance RBF Neural Networks for Massive MIMO-OFDM Receivers

Jonathan Aguiar Soares ^{1,*} , Kayol Soares Mayer ¹ , Fernando César Comparsi de Castro ²
and Dalton Soares Arantes ¹

¹ Department of Communications, School of Electrical and Computer Engineering, University of Campinas, Campinas 13083-852, Brazil; kayol@decom.fee.unicamp.br (K.S.M.); dalton@unicamp.br (D.S.A.)

² Department of Electronics and Computing, Technology Center, Federal University of Santa Maria, Santa Maria 97105-900, Brazil; fernando.castro@ufsm.br

* Correspondence: jonathan@decom.fee.unicamp.br

Abstract: Multi-input multi-output (MIMO) transmission schemes have become the techniques of choice for increasing spectral efficiency in bandwidth-congested areas. However, the design of cost-effective receivers for MIMO channels remains a challenging task. The maximum likelihood detector can achieve excellent performance—usually, the best performance—but its computational complexity is a limiting factor in practical implementation. In the present work, a novel MIMO scheme using a practically feasible decoding algorithm based on the phase transmittance radial basis function (PTRBF) neural network is proposed. For some practical scenarios, the proposed scheme achieves improved receiver performance with lower computational complexity relative to the maximum likelihood decoding, thus substantially increasing the applicability of the algorithm. Simulation results are presented for MIMO-OFDM under 5G wireless Rayleigh channels so that a fair performance comparison with other reference techniques can be established.

Citation: Soares, J.A.; Mayer, K.S.; de Castro, F.C.C.; Arantes, D.S. Complex-Valued Phase Transmittance RBF Neural Networks for Massive MIMO-OFDM Receivers. *Sensors* **2021**, *21*, 8200. <https://doi.org/10.3390/s21248200>

Academic Editor: Gianmarco Romano

Received: 29 October 2021

Accepted: 6 December 2021

Published: 8 December 2021

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Copyright: © 2021 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

Keywords: artificial neural networks; phase transmittance radial basis function; massive MIMO; MIMO decoding; 5G

1. Introduction

In recent years, with the increasing demand for the real-time processing of big data, the Internet of Things (IoT), and 4K video streaming, technologies to increase area throughput [1] in base station (BS) coverage and hotspot tiers [2] have become increasingly important. In general, the system throughput can be improved by three independent factors: the number of BSs, bandwidth, and spectral efficiency. While the number of base stations is a complicated variable to handle, there are substantial bandwidths in the millimeter wavelength (mmWave) bands that could be employed for BS hotspot tiers. On the other hand, as objects and human bodies easily block mmWaves, increasing the spectral efficiency (SE) of BS coverage tiers arises as a potential solution for wide-area coverage. In order to increase SE, advanced techniques are necessary to use the available BSs and bandwidth more efficiently. In view of this, both BSs and user equipment (UE) currently operate with multiple antennas and orthogonal frequency-division multiplexing (OFDM) [3–5] to increase spectral efficiency.

Multicarrier modulation schemes, such as OFDM, have been widely employed in digital communications systems due to their low susceptibility to intersymbol interference (ISI) [6–9]. OFDM divides the channel bandwidth into K orthogonal subcarriers [10]. The serial stream at a high data rate applied to the OFDM input is first converted to multiple parallel low transmission rate sub-streams. Each of the K parallel sub-streams modulates one of the K subcarriers. In this way, the OFDM symbol duration is K times longer than the symbol duration of the equivalent single carrier system, thus avoiding ISI [11]. Another

important characteristic of OFDM systems is that multiple users can be multiplexed in frequency, using K subcarriers.

Multiple-input and multiple-output (MIMO) technologies use multiple antennas on the transmitter and receiver sides, increasing the wireless channel capacity without extra bandwidth, extra power transmission, or both [11]. Usually, as BSs have more computational power than UEs, a larger number of antennas is applied at the transmitter. In this context, when the number of antennas exceeds the number of users, the term massive MIMO (mMIMO) is frequently used. Generally, mMIMO systems operate with 16 or more antennas in BSs. In addition, the uplink can be composed of one or more UEs, where the former characterizes a single-user mMIMO (SU-mMIMO) [12] and the latter a multi-user mMIMO (MU-mMIMO) [13]. For 6G technologies, ultra-massive MIMO (UM-MIMO) schemes have been explored to support data throughputs of Terabits [14]. MIMO and mMIMO communications are broadly implemented either by beamforming, space–time block codes (STBC), or both [15,16]. While beamforming techniques (either analog, digital, or hybrid [17]) may be more attractive to BSs, because of the necessity of many transmitting antennas, STBC, on the other hand, is feasible for use in both downlink and uplink connections.

The combination of OFDM and MIMO is a major trend in mobile communication systems [18–22], such as 5G and next-generations, which are based on the so-called MIMO-OFDM and mMIMO-OFDM approaches [23]. Nonetheless, digital communication systems over wireless channels may suffer severe signal degradations due to multipath propagation, additive white Gaussian noise (AWGN) [24,25], and Doppler effects [26]. Moreover, another frequent impairment in OFDM systems is signal distortion, characterized by the PAPR (peak-to-average power ratio), due to nonlinearities at the high-power transmitter amplifier [27,28]. Since nonlinear impairments usually degrade the performance of linear filters, the search for robust nonlinear filters at the receiver is essential to circumvent this issue [11,27,29].

Along with the application of nonlinear filters designed for specific problems in telecommunications, artificial neural networks (ANNs) have been extensively studied in various challenging areas of digital communications, including soft and hard fault detection, channel estimation, equalization, and beamforming [30–40]. Neural networks can operate like nonlinear filters, in a structure that can be modeled by nonlinear activation functions, as in multilayer perceptrons (MLPs), or by Gaussian neurons in radial basis function neural networks (RBFNN) [35]. RBFNN Gaussian neurons have two free parameters, namely the Gaussian centers and the variances. Moreover, there is a linear free parameter vector of weights, which linearly weighs the neuron outputs to yield the network output [35,41]. With these three independent sets of parameters, RBFNNs are able to represent high-order nonlinear surfaces without increasing the number of layers, thereby reducing complexity compared with deep neural networks [11,35].

In this context, this work proposes a novel complex-valued RBF neural network architecture: a multiple-input multiple-output phase transmittance RBF (MIMO-PTRBF) neural network for channel estimation and symbol detection in massive MIMO-OFDM communication systems. The PTRBF neural network model was chosen due to its lower computational complexity when compared with deep neural networks and due to its crucial role in avoiding the phase invariance which occurs in a standard complex-valued RBFNN [11,35,41]. The proposed MIMO-PTRBF is an extension to multiple outputs of the single-output PTRBF neural network presented in [41]. This work is based on [35], in which the authors redesigned the PTRBF neural network of [41] to obtain a low-complexity MIMO beamforming transmitter with the complex-MIMO RBF (CMM-RBF). In the present paper, using the Gaussian neuron output bounds presented in [35], we prove convergence in the mean of the PTRBF neural network, relaxing the condition of the trans-dimensional transformation of the complex-valued Gaussian neuron layer. Preliminary results indicate that the proposed architecture competes quite favorably with the conventional MIMO quasi-orthogonal space–time block coding (QOSTBC) with maximum likelihood (ML)

decoding based on linear processing [11]. In addition, an STBC coding algorithm with full-rate and half-diversity was developed for massive MIMO coding with square matrices. However, because of its high computational complexity, the ML decoding may be unfeasible for mMIMO with a large number of antennas, as opposed to the MIMO-PTRBF neural network proposed here. Although other decoding techniques with lower computational complexities could be taken into account, for performance comparison, we chose ML decoding because of its optimal performance in linear channels. For example, the sphere decoding has lower computational complexity compared to the ML decoding, but its performance is only lower-bounded by the ML decoding [42]. Furthermore, the MIMO-PTRBF is able to both estimate the channel and decode the received signal, relying on a training sequence. Simulation results show that the proposed architecture achieves significantly improved BER figures when compared with MIMO QOSTBC in an equal scenario, either with linear or with nonlinear impairments under 5G channels.

This paper is an extension of J. A. Soares' MSc. dissertation developed at the School of Electrical and Computer Engineering, University of Campinas, in the area of Telecommunications and Telematics [11]. In addition to the complex-valued RBF-based MIMO system proposed in the dissertation [11], in this paper, the MIMO-PTRBF receiver is further elaborated with additional results and with the use of 5G channel models.

The remainder of this work is organized as follows. In Section 2, a brief review of multi-antenna systems is presented. The proposed STBC coding scheme and the MIMO-PTRBF for channel estimation and symbol detection in massive MIMO-OFDM systems are presented in Section 3. In Section 4, simulation results of the MIMO-PTRBF are compared with the results obtained with the OSTBC under maximum likelihood decoding in 5G channel models. Computational complexities are presented in Section 5 and conclusions are discussed in Section 6.

2. Background

The main goal in multi-antenna systems is to increase the channel capacity with M_T transmit and M_R receive antennas by a factor of $\min(M_T, M_R)$ without using additional transmit power or spectral bandwidth [43]. Considering the MIMO digital communication system $\mathbf{r}(k) = \mathbf{H}^T \mathbf{x}(k) + \boldsymbol{\eta}(k)$, with transmitted signal $\mathbf{x}(k) \in \mathbb{C}^{M_T}$, received signal $\mathbf{r}(k) \in \mathbb{C}^{M_R}$, and additive white Gaussian noise (AWGN) vector $\boldsymbol{\eta}(k) \in \mathbb{C}^{M_R}$, the channel capacity of $\mathbf{H} \in \mathbb{C}^{M_T \times M_R}$ is expressed as [44]

$$C = \log_2 \left[\det \left(\mathbf{I}_{M_R} + \frac{E_s}{M_T E_0} \mathbf{H}^H \mathbf{R}_{xx} \mathbf{H} \right) \right], \quad (1)$$

in which $\mathbf{I}_{M_R}$ is an $M_R \times M_R$ identity matrix, $[\cdot]^T$ is the transpose operator, $[\cdot]^H$ is the conjugate transpose operator, E_s is the total transmitted signal power, E_0 is the AWGN power, $\mathbf{R}_{xx} = E\{\mathbf{x}(k)\mathbf{x}^H(k)\}$ is the correlation matrix of $\mathbf{x}(k)$, and $E\{\cdot\}$ is the expectation operator.

However, if no channel state information (CSI) is available at the transmitter, we can assume that the channel components are equally probable. In this case, we consider that power is equally divided among the transmitting antennas, which implies $\mathbf{R}_{xx} = \mathbf{I}_{M_T}$. The capacity in such a case is then given by [11,45]

$$C = \log_2 \left[\det \left(\mathbf{I}_{M_R} + \frac{E_s}{M_T N_0} \mathbf{H}^H \mathbf{H} \right) \right]. \quad (2)$$

Note that Equation (2) can be outperformed if the channel information is available at the transmitter (leading to a coding gain). However, Equation (2) is the maximum diversity capacity without channel knowledge at the transmitter. Furthermore, if $M_T = M_R = 1$, Equation (2) represents the Shannon capacity for SISO systems [11].

In order to increase capacity, the concepts of diversity [46], coding [47], and array [48] gains play key roles in MIMO systems. The array gain is the average increase in the signal-to-noise ratio (SNR) at the receiver that arises from the coherent combining effect of multiple antennas at the transmitter, receiver, or both. Multiple antenna systems require

perfect channel knowledge at the transmitter, receiver, or both to achieve this array gain [48]. On the other hand, diversity gain is obtained by the provision of replicas of the transmitted signal at the receiver [46]. Diversity gain techniques are used to mitigate degradations in the error performance due to wireless fading channels (e.g., due to multipath). Since the probability that statistically independent fading channels simultaneously experience deep fading is insignificant, there are various ways of performing diversity gain and space diversity. To accomplish this, it is necessary to use sufficiently separated antennas in the array (by more than 10λ on base stations and 2λ to 5λ on mobile devices [44]) to guarantee independent wireless channels. In contrast, coding gain is usually provided by temporal channel coding, e.g., convolutional and block codes [11].

Space–time code is a digital communication technique used to transmit multiple copies of a data stream via multiple antennas to compensate for fading and AWGN. At the receiver side, these multiple copies of the signal are received by one or more antennas, improving the communication reliability. Depending on the encoder algorithm at the transmitter, we can have different space–time codes. Space–time trellis codes (STTCs) combine modulation and trellis coding to transmit signals over a MIMO channel. Although STTCs provide both coding gain and diversity gain, the computational complexity is higher than other space–time codes, mainly in the receiver, where a Viterbi decoder is necessary [49]. Space–time block codes (STBCs) combine multiple symbols from a digital modulation, creating a block of symbols. The components of this block (i.e., matrix of symbols) are indexed by the transmitting antenna and the transmitting time. At the transmitter, STBC decoding is performed in linear processing. Another technique is the space–time labeling diversity (STLD), a variation of STBC that takes two bit-streams and outputs two pairs of symbols. Two symbols in each pair are transmitted by two transmit antennas in two time slots, which results in full-diversity and half-rate [50]. In addition, STLD only works with a limited number of transmit antennas.

2.1. Space–Time Block Coding and OFDM

MIMO systems are mainly designed for narrowband or flat channels. Applying MIMO systems in the wideband frequency selective channel implies a constant penalty factor in the coding gain compared with that in flat-frequency channels. Furthermore, at high SNRs, an irreducible error rate floor is inevitable [51]. This irreducible error rate floor is due to the existence of multipath delay spread, and it persists even if we increase the number of antennas. Since the ISI is the root cause of the error floor, in principle, it can be mitigated by resorting to adaptive equalization, but this can be too complex to implement in such an environment. Another option that is widely used is to resort to OFDM, which naturally converts a frequency-selective fading channel into a frequency-nonselective fading channel. The subcarriers (i.e., tones) in an OFDM symbol are essentially narrowband signals. Since these tones fit perfectly as vehicles for space–time codes, OFDM is an enabler for this efficient coding technique [11].

Figure 1 shows the coding scheme for a generic coding matrix $\mathbf{X}[k] \in \mathbb{C}^{M_T \times P}$, where P is the number of time samples for the transmission of one block of coded symbols and $k = 1, 2, \dots, K$ is the carrier index of the k th MIMO-STBC encoded symbol matrix $\mathbf{X}[k]$ along the OFDM symbol. $\mathbf{R}[k] \in \mathbb{C}^{M_R \times P}$ is the matrix of received symbols, $\hat{\mathbf{s}} \in \mathbb{C}^{M_S}$ is the decoded vector, and M_S is the number of modulated symbols in a MIMO-STBC matrix.

The transmitting space–time block coder (STBC) encodes the data symbol vector $\mathbf{s}[k] \in \mathbb{C}^{M_S}$ using the code matrix to construct the transmitting matrix $\mathbf{X}[k]$ of length K . The streams $\mathbf{x}_{m_T,p}[k]$ are fed to the IFFT modulator of each m_T transmitting antenna, at each p period of time relative to the OFDM symbol sequence. In this manner, the information is transmitted in $\mathbf{X}[k]$ blocks of M_T antennas and P OFDM symbols in each k -th carrier. To illustrate this scheme, Figure 2 shows an example for a two-transmitting antenna system using Alamouti coding. Consequently, the channel is given by $\mathbf{H} \in \mathbb{C}^{M_T \times M_R \times P \times K}$. It should be emphasized that in the simulation in Section 4 of this work, the channel is not assumed to be static over the entire MIMO-STBC block since it spreads over time in

P -consecutive OFDM symbols. This is particularly necessary for the proposed work, as the receiver will fit and adapt to the characteristics and variations of the channel over time. This is also necessary for a massive number of broadcast antennas due to the length of the long block coding P that transmits over time in consecutive OFDM symbols [11].

Figure 1. Coding scheme for a MIMO-OFDM system in which k is the index of the k th MIMO-STBC-encoded symbol matrix $\mathbf{X}[k]$ along the OFDM symbol [11].

Figure 2. Detailed MIMO-OFDM coding system. $\{\cdot\}^*$ denotes the conjugate operator.

As in OFDM systems, MIMO-OFDM also requires the channel state information to decode the received symbols. One of the most popular and widely used approaches to MIMO channel estimation is to employ pilot signals (also referred to as training sequences) and then estimate the channel based on the received data and the knowledge of the training sequence, as detailed in Figure 3. Based on pilot signals, in [52,53], the least-squares (LS) channel estimation technique is applied for orthogonal frequency-division multiplexing systems with multiple transmit antennas [11].

A generalized coding scheme referred to as space–time block codes (STBCs) [16,44,54], based on the theory of orthogonal matrix designs, can achieve the full-transmit diversity of $M_T M_R$ employing the maximum likelihood decoding algorithm at the receiver [44]. The idea is to transmit M_T orthogonal streams, which implies that the receiver antennas receive M_T orthogonal streams. This special class of space–time block codes is the so-called orthogonal STBC (OSTBC) [11,54,55].

An OSTBC example of coding matrix for $M_T = 4$ [44] is given by

$$\text{OSTBC}_{4,8} = \begin{bmatrix} s[1] & -s[2] & -s[3] & -s[4] & s[1]^* & -s[2]^* & -s[3]^* & -s[4]^* \\ s[2] & s[1] & s[4] & -s[3] & s[2]^* & s[1]^* & s[4]^* & -s[3]^* \\ s[3] & -s[4] & s[1] & s[2] & s[3]^* & -s[4]^* & s[1]^* & s[2]^* \\ s[4] & s[3] & -s[2] & s[1] & s[4]^* & s[3]^* & -s[2]^* & s[1]^* \end{bmatrix}, \quad (3)$$

in which $s[m_s]$ is the transmitted signal in the discrete symbol index m_s . Notice that, as proved by Tarokh et al. [54], the inner product of any two distinct rows of this matrix is equal to zero (i.e., the matrix is orthogonal) and of full-rank, yielding full-diversity [11].

Figure 3. MIMO-OFDM model system with channel estimation.

One of the disadvantages of OSTBC is the code rate. Let P represent the number of time samples to convey one block of coded symbols and M_s represent the number of symbols transmitted per block. The space–time block code rate is defined as the ratio between the number of symbols that the encoder receives at its input and the number of space–time coded symbols transmitted from each antenna, given by $R = M_s/P$. This implies that Equation (3) has a code rate $R = 1/2$, which consequently reduces the spectral efficiency. Supplementary to the diversity gain, the OSTBC leads to a secondary linear coding gain $G_c = 10 \log(R)$ at the receiver due to the coherent detection of multiple copies of the signal over time. Furthermore, the multi-antenna system, as presented in Figure 1, will lead to an array gain $G_a = 10 \log(M_R)$ due to the coherent combination of multiple received signals over the receiving antennas [11].

2.2. Quasi-Orthogonal Special Case

In order to increase the spectral efficiency in orthogonal codes, Jafarkhani [56] proposed quasi-orthogonal STBC (QOSTBC) of rate one, relaxing the requirement of orthogonality. However, when compared with orthogonal codes, the diversity gain is reduced by a factor of two. Besides, in contrast to orthogonally designed codes that process one symbol at a time at the decoder, quasi-orthogonal codes process pairs of transmitted symbols, which exponentially increases the computational complexity of decoding [11].

Jafarkhani [56] proposed a coding matrix of rate one for $M_T = 4$, given by

$$\mathbf{QOSTBC}_{4,4} = \begin{bmatrix} s[1] & s[2] & s[3] & s[4] \\ -s[2]^* & s[1]^* & -s[4]^* & s[3]^* \\ -s[3]^* & -s[4]^* & s[1]^* & s[2]^* \\ s[4] & -s[3] & -s[2] & s[1] \end{bmatrix}. \quad (4)$$

In the literature, related approaches with a maximum of $M_T = 6$ antennas were proposed for quasi-orthogonal codes [57–59]. In [58], the authors developed an architecture similar to [56]; however, this presents full-diversity at the cost of more processing and is limited to $M_T = 4$ antennas. In the same way, by increasing the decoding processing, Sindhu and Hameed [59] proposed two quasi-orthogonal schemes with $M_T = 5$ and 6 antennas [11].

2.3. Decoding for Space–Time Block Codes

Maximum likelihood (ML) detection calculates the Euclidean distance among the received signal matrix $\mathbf{R}$ and the product of all possible transmitted signal vectors by the channel matrix $\mathbf{H}$. Considering $\mathbb{A}$, the set of constellation symbols of the transmitted signal, and M_s , the number of transmitted symbols per MIMO block, ML detection determines the estimation of the conveyed signal vector $\mathbf{s}$ as [11]

$$\hat{\mathbf{s}} = \underset{\mathbf{s} \in \mathbb{A}^{M_s}}{\operatorname{argmin}} \|\mathbf{R} - \mathbf{H}^T \mathbf{X}\|^2. \quad (5)$$

As in maximum a posteriori (MAP) detection, ML detection achieves the optimal performance when all transmitted vectors are equally probable. However, the number of ML computation metrics is A^{M_s} , where A is the modulation order. Thus, the ML complexity increases exponentially with the modulation order or the number of transmit symbols, or both [44,54,60]. Although this method has a high computational complexity, the ML decoding is used as a benchmark due to its optimal performance [11].

For orthogonal coding schemes, the ML metric can be simplified, decoding symbol by symbol [54]. Via this simplification, it is possible to circumvent the issue of exponential computational complexity. However, even with this simplification, the computational complexity can be considerably high. In QOSTBC, the ML metric can be also simplified, but the computational complexity remains higher than the orthogonal case, because QOSTBC is decoded in pairs of symbols [11].

3. Proposed Approach

3.1. Coding Scheme

Similarly to the work of [56], the present work is derived from the full-rate full-diversity complex-valued space–time block code scheme proposed by Alamouti [61]. The transmission matrix proposed in [61] is given by [11]

$$\mathbf{A}_{i,j} = \begin{bmatrix} s[i] & s[j] \\ -s[j]^* & s[i]^* \end{bmatrix}, \quad (6)$$

in which $s[i]$ is the i th input symbol to be encoded.

Based on [61], Jafarkhani [56] proposed a quasi-orthogonal coding scheme using four antennas and consequently four encoded symbols as [11]

$$\mathbf{S}_4^4 = \begin{bmatrix} \mathbf{A}_{1,2} & \mathbf{A}_{3,4} \\ -[\mathbf{A}_{3,4}]^* & [\mathbf{A}_{1,2}]^* \end{bmatrix} = \begin{bmatrix} s[1] & s[2] & s[3] & s[4] \\ -s[2]^* & s[1]^* & -s[4]^* & s[3]^* \\ -s[3]^* & -s[4]^* & s[1]^* & s[2]^* \\ s[4] & -s[3] & -s[2] & s[1] \end{bmatrix}, \quad (7)$$

where $\mathbf{S}$ is the quasi-orthogonal coding matrix. The main idea behind the work of [56] is to build a 4×4 matrix from two 2×2 matrices, keeping a fixed transmission rate [11].

In the present paper, we generalize the idea presented in [56] to a new recursive method of generating coding schemes, as given by [11]

$$\mathbf{S}_{M_s}^{M_T} = \begin{bmatrix} \mathbf{S}_{M_s-M_T/2}^{M_T/2} & \mathbf{S}_{M_s}^{M_T/2} \\ -[\mathbf{S}_{M_s}^{M_T/2}]^* & [\mathbf{S}_{M_s-M_T/2}^{M_T/2}]^* \end{bmatrix}, \quad (8)$$

in which $M_T = 2^n$, $\forall n \geq 1$ is the number of transmitting antennas and M_s is the number of encoded symbols. In the proposed scheme, $M_s \triangleq M_T$ and the code rate is $R = M_T/M_s = 1$ [11]. The recurrence is performed until we find $\mathbf{S}_n^1 = s[n]$, $\forall n \in [1, 2, \dots, M_s]$ in Equation (8).

For example, with four transmitting antennas, Equation (8) results in [11]

$$\mathbf{S}_4^4 = \begin{bmatrix} \mathbf{S}_{4-4/2}^{4/2} & \mathbf{S}_4^{4/2} \\ -[\mathbf{S}_{4-4/2}^{4/2}]^* & [\mathbf{S}_{4-4/2}^{4/2}]^* \end{bmatrix} = \begin{bmatrix} \mathbf{S}_2^2 & \mathbf{S}_4^2 \\ -[\mathbf{S}_4^2]^* & [\mathbf{S}_2^2]^* \end{bmatrix}. \quad (9)$$

From the recurrent structure of Equation (8), $\mathbf{S}_2^2$ and $\mathbf{S}_4^2$ are

$$\mathbf{S}_2^2 = \begin{bmatrix} \mathbf{S}_1^1 & \mathbf{S}_2^1 \\ -[\mathbf{S}_2^1]^* & [\mathbf{S}_1^1]^* \end{bmatrix}, \quad (10)$$

and

$$\mathbf{S}_4^2 = \begin{bmatrix} \mathbf{S}_3^1 & \mathbf{S}_4^1 \\ -[\mathbf{S}_4^1]^* & [\mathbf{S}_3^1]^* \end{bmatrix}. \quad (11)$$

Thus, substituting Equations (10) and (11) into Equation (9),

$$\mathbf{S}_4^4 = \begin{bmatrix} \begin{bmatrix} \mathbf{S}_1^1 & \mathbf{S}_2^1 \\ -[\mathbf{S}_2^1]^* & [\mathbf{S}_1^1]^* \end{bmatrix} & \begin{bmatrix} \mathbf{S}_3^1 & \mathbf{S}_4^1 \\ -[\mathbf{S}_4^1]^* & [\mathbf{S}_3^1]^* \end{bmatrix} \\ -\begin{bmatrix} \mathbf{S}_3^1 & \mathbf{S}_4^1 \\ -[\mathbf{S}_4^1]^* & [\mathbf{S}_3^1]^* \end{bmatrix}^* & \begin{bmatrix} \mathbf{S}_1^1 & \mathbf{S}_2^1 \\ -[\mathbf{S}_2^1]^* & [\mathbf{S}_1^1]^* \end{bmatrix}^* \end{bmatrix} = \begin{bmatrix} \mathbf{S}_1^1 & \mathbf{S}_2^1 & \mathbf{S}_3^1 & \mathbf{S}_4^1 \\ -\mathbf{S}_2^{1*} & \mathbf{S}_1^{1*} & -\mathbf{S}_4^{1*} & \mathbf{S}_3^{1*} \\ -\mathbf{S}_3^{1*} & -\mathbf{S}_4^{1*} & \mathbf{S}_1^{1*} & \mathbf{S}_2^{1*} \\ \mathbf{S}_4^1 & -\mathbf{S}_3^1 & -\mathbf{S}_2^1 & \mathbf{S}_1^1 \end{bmatrix}. \quad (12)$$

Replacing $\mathbf{S}_n^1 = s[n]$, $\forall n \in [1, 2, 3, 4]$, into Equation (12),

$$\mathbf{S}_4^4 = \begin{bmatrix} s[1] & s[2] & s[3] & s[4] \\ -s[2]^* & s[1]^* & -s[4]^* & s[3]^* \\ -s[3]^* & -s[4]^* & s[1]^* & s[2]^* \\ s[4] & -s[3] & -s[2] & s[1] \end{bmatrix}. \quad (13)$$

Note that Equation (13) is equal to the coding scheme proposed by [56] with four antennas, as in Equation (4). However, in contrast to the work of [56], our coding scheme, presented in Equation (8), can generate coding matrices for any $M_T = 2^n$, $\forall n \geq 1$, and $M_s \triangleq M_T$ [11]. For the case of $n = 1$, Equation (8) is equal to Equation (6), the full-rate full-diversity complex-valued space-time block code scheme proposed in [61].

The main issue of the proposed coding scheme is that we cannot define a simplified ML decoding method as in the former cases. Then, it is here that the system proposed in this paper takes shape, with the MM-PTRBF decoding, making the joint solution feasible. We have observed, by extensive simulations, that Equation (8) achieves half of the diversity presented by the orthogonal coding schemes but keeps full-rate (i.e., $R = 1$), which is essentially the characteristics of the quasi-orthogonal scheme proposed by [11,56].

3.2. Complex MIMO-PTRBF Neural Network for Massive MIMO Decoding

In the proposed system, the maximum likelihood decoder is replaced by a neural network, the MIMO-PTRBF, to decode the received symbols, as shown in Figure 4. The MIMO-PTRBF has a supervised learning stage, in which a training sequence is used to fit the hyper-parameters of the neural network. A pseudo-random generator creates this training sequence, which is known both at the transmitter and receiver sides. When the neural network output achieves the desired MSE, it switches from the learning stage to the decoding stage. At this time, the information data are then effectively transmitted over the system, and the BER is computed. These two stages are implemented as in Figure 4, with the input switch of the MIMO STBC Encoder block and the output switch of the Neural Network Decoder block. The switches have two states represented by (a) and (b), which shift between the training and decoding stages [11].

Figure 4. Complete vision of the proposed MIMO-OFDM model [11].

As in the maximum likelihood detector, the input signal to the MIMO-PTRBF algorithm is the set of received vectors $\mathbf{r}$, as shown in Figure 4. The MIMO-PTRBF architecture, with N neurons, has three free parameters: the matrix of synaptic weights $\mathbf{W} \in \mathbb{C}^{M_s \times N}$, the matrix of center vectors $\mathbf{\Gamma} \in \mathbb{C}^{M_R P \times N}$, and the vector of variances $\sigma^2 \in \mathbb{C}^{N \times 1}$. The MIMO-PTRBF is an extension of the PTRBF for multiple outputs. The key difference between both architectures is the multiple-output layer, which fits each output individually. Figure 5 shows a closer view of the receiver side using the MIMO-PTRBF neural network for decoding [11].

Figure 5. Closer view of the system with neural network decoding [11].

The output vector is thus given by

$$\hat{\mathbf{s}}[u] = \mathbf{W}[u]\boldsymbol{\phi}[u]. \quad (14)$$

Following the complex-valued radial basis function presented in [41], the n th neuron output of the MIMO-PTRBF (ϕ_n), for the p th output vector of $\mathbf{r}$, is [11]

$$\begin{aligned}
w_{m_s,n}[u+1] &= w_{m_s,n}[u] + \eta_w e_{m_s}[u] \phi_n^*[u], \\
\gamma_n[u+1] &= \gamma_n[u] + \eta_\gamma [\text{Re}(\xi_n[u])\text{Re}(\alpha_n[u]) - j\text{Im}(\xi_n[u])\text{Im}(\alpha_n[u])], \\
\sigma_n[u+1] &= \sigma_n[u] + \eta_\sigma [\text{Re}(\xi_n[u])\text{Re}(\beta_n[u]) - j\text{Im}(\xi_n[u])\text{Im}(\beta_n[u])],
\end{aligned} \quad (19)$$

in which $[\cdot]^*$ denotes the complex conjugate operator and $\xi_n[u]$ is the n th synaptic transmittance, given by

$$\xi_n[u] = \sum_{m_s=1}^{M_s} e_{m_s}^*[u] w_{m_s,n}[u]. \quad (20)$$

Furthermore, $\alpha_n[u] \in \mathbb{C}^R$ is the m th vector of the matrix of weighted centers ($\mathbf{A}[u] \in \mathbb{C}^{N \times R}$):

$$\alpha_n[u] = \text{Re}(\phi_n[u]) \frac{[\text{Re}(\mathbf{x}[u]) - \text{Re}(\gamma_n[u])]}{\text{Re}(\sigma_n[u])} + j\text{Im}(\phi_n[u]) \frac{[\text{Im}(\mathbf{x}[u]) - \text{Im}(\gamma_n[u])]}{\text{Im}(\sigma_n[u])}. \quad (21)$$

In a similar way, $\beta_n[u] \in \mathbb{C}$ is the n th element of the vector of weighted kernel ($\beta[u] \in \mathbb{C}^N$):

$$\beta_n[u] = \text{Re}(\phi_n[u]) \frac{\|\text{Re}(\mathbf{x}[u]) - \text{Re}(\gamma_n[u])\|_2^2}{[\text{Re}(\sigma_n[u])]^2} + j\text{Im}(\phi_n[u]) \frac{\|\text{Im}(\mathbf{x}[u]) - \text{Im}(\gamma_n[u])\|_2^2}{[\text{Im}(\sigma_n[u])]^2}. \quad (22)$$

Generalizing Equation (19) to matrix structures results in

$$\begin{aligned}
\mathbf{W}[u+1] &= \mathbf{W}[u] + \eta_w \mathbf{e}[u] \Phi^H[u], \\
\mathbf{\Gamma}[u+1] &= \mathbf{\Gamma}[u] + \eta_\gamma \mathbf{A}^*[u], \\
\boldsymbol{\sigma}[u+1] &= \boldsymbol{\sigma}[u] + \eta_\sigma \boldsymbol{\alpha}^*[u],
\end{aligned} \quad (23)$$

where $\mathbf{A}[u]$ and $\boldsymbol{\alpha}[u]$ are auxiliary variables used to reduce the computational complexity. $\mathbf{A}[u]$ and $\boldsymbol{\alpha}[u]$ are given by

$$\begin{aligned}
\mathbf{A}[u] &= \text{Re}(\Xi[u])\text{Re}(\mathbf{A}[u]) + j\text{Im}(\Xi[u])\text{Im}(\mathbf{A}[u]) \in \mathbb{C}^{N \times R}, \\
\boldsymbol{\alpha}[u] &= \text{Re}(\Xi[u])\text{Re}(\beta[u]) + j\text{Im}(\Xi[u])\text{Im}(\beta[u]) \in \mathbb{C}^N,
\end{aligned} \quad (24)$$

in which $\Xi[u]$ is the diagonal matrix of synaptic transmittance:

$$\Xi[u] = \begin{bmatrix} \xi_1[u] & 0 & \cdots & 0 \\ 0 & \xi_2[u] & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \xi_N[u] \end{bmatrix} \in \mathbb{C}^{N \times N}. \quad (25)$$

Each training update is given by Equation (23); however, for $u = 0$, the MIMO-PTRBF free parameters are initialized following some criterion defined by the user (e.g., based on the probability distribution of the input data). Although (23) minimizes the error between the output vector $\hat{\mathbf{s}}$ and the reference vector $\mathbf{s}$, as the neurons are dependent on exponential functions, a risk of instability is assumed if the exponential argument is positive. In order to circumvent this issue, based on Theorem A1 of [35], the real and imaginary parts of each scalar component of the vector of variances are lower-bounded by the limit $\mu > 0$, which, consequently, bounds the real and imaginary parts of the neurons output from 0 to 1 [11]. In addition, taking into account Theorem A2 (see Appendix A), the adaptive step of the matrix of synaptic weights is limited by $\eta_w < 1/N$ for all simulations, to guarantee convergence in the mean. In addition, in the Appendix, Corollaries A1 and A2 are of utmost importance to prove Theorem A1. In addition, Definition A1 is used to prove Corollary A2.

4. Simulation Results

Using the formerly mentioned OSTBC [54] and QOSTBC [56] coding schemes, several setups are compared with the proposed approach to validate and assess their performance in massive MIMO-OFDM. OSTBC and QOSTBC are simulated with the maximum likelihood (ML) decoding with perfect channel knowledge. This configuration achieves the maximum diversity gain $G_d = M_T M_R$ at the cost of half of the theoretical bandwidth efficiency, since $R = 1/2$ in this case. Considering a practical QOSTBC application, we also implement the EQOSTBC with the least-squares (ML-LS) channel estimation [11]. Figure 7 illustrates the simulated system.

Figure 7. Block model of the simulated systems.

In Figure 7, the binary input data are created by a pseudo-random generator with uniform distribution. The bit-stream is then modulated according to the M -QAM or M -PSK modulation scheme used in the simulation. Subsequently, using the coding scheme proposed in Section 3, the modulated symbols are encoded in the STBC block. In the IDFT block, the STBC symbols are frequency-multiplexed for OFDM transmission. At the receiver side, after the transmitted signal passes through the channel, the DFT is applied to demultiplex the STBC symbols. In the decoder, the proposed ANN-based technique presented in Section 3 and the ML algorithm (either with perfect channel knowledge at the receiver or with channel estimation by LS) are employed to assess the system performance. In the sequel, the decoder output symbols are demodulated, and BER is computed.

For the sake of comparison, the BER as a function of E_b/N_0 (energy per bit to noise power spectral density ratio) is used in the simulations. By adjusting the transmitting power for each antenna, the received signals are normalized by M_T transmitting antenna, by the receiver array gain M_R , and by the code rate gain R , implying

$$SNR_{(dB)} = E_b/N_{0(dB)} + 10 \log_{10} \left(\frac{b}{RM_T M_R} \right), \quad (26)$$

in which b is the number of bits per QAM symbol.

In Figure 8, aiming to validate the simulator shown in Figure 7, we compare the obtained results with the theoretical performance of OSTBC for 4th, 8th, 16th, and 64th diversity orders using 4-QAM modulation for a Rayleigh channel with AWGN. For all OSTBC diversity orders, theoretical and simulated results were approximately the same, validating the framework. In addition, Figure 9 presents the reference results of [56] with $M_T = 4$ antennas and $M_R = 1$ antenna for 16-QAM OSTBC and 4-QAM QOSTBC and the obtained results for the same scenarios. The simulated results are in line with theoretical results, which also corroborates the framework's reliability.

Figure 8. Simulated and theoretical results for 4th, 8th, 16th, and 64th diversity orders.

Figure 9. Simulated, reference, and theoretical results for equal diversity order and bitrate [11].

With the simulation framework validated, we can compare the proposed coding algorithm with results from the literature. Firstly, Figure 10 shows the results of the proposed coding algorithm and theoretical results for 2nd, 3rd, 4th, 5th, 8th, and 10th order diversity. As addressed by [56], quasi-orthogonal transmitting schemes with four antennas achieve at least half of the theoretical diversity ($D_o = \frac{M_T}{2} = 2$) of the orthogonal four antenna scheme ($D_o = M_T = 4$). This can be seen in the solid blue curve with squares of Figure 10, which is located between the theoretical second and third-order curves. Using simulations, we extend the concept introduced by [56] for QOSTBC with 8×1 and 16×1 antennas. In order to simulate these scenarios, we employ the proposed coding algorithm presented in Equation (8). As expected, the solid green curve with diamonds and the solid orange curve with circles are between the theoretical 4th and 5th order and the 8th and 10th order, respectively. Then, utilizing this analysis, we validate the proposed code algorithm and show that it is a suitable approach for generating QOSTBC matrices for at least 16 antennas. Higher-order QOSTBC architectures using Equation (8) are not simulated because of the extensive time required to perform maximum likelihood detection.

Figure 10. Simulation for the proposed coding scheme against theoretical results for equal bitrate and $M_T = 4$ and 8 with $M_R = 1$ [11].

In order to represent more practical scenarios, we set the simulation system with a 3GPP TS 38.211 specification [62] for 5G Physical channels and modulation. The Subcarrier Spacing (Δf) scales from 15 kHz to 240 kHz. The number of active carriers is 256, and the pilot sample rate (when applicable) is $M_T \times 8 \times f_{Doppler}$ with the conventional block-based pilot scheme [63]. We perform simulations in the extremes to demonstrate the robustness of the proposed approach.

The radio channel realizations are created using the 3GPP TR 38.901 report on 5G: Study on channel model for frequencies from 0.5 GHz to 100 GHz [64]. The 3GPP channel models [64] are applicable for frequency bands in the range of 0.5 GHz to 100 GHz. From Tapped Delay Line (TDL) models in [64], TDL-B is selected from Table 7.7.2-2 (depicted in Table 1) for the channel model simulated in this work.

Table 1. Table 7.7.2-2. TDL-B.

Tap #	Normalized Delay	Power (dB)	Tap #	Normalized Delay	Power (dB)
1	0.0000	0.00	13	1.1021	−4.80
2	0.1072	−2.20	14	1.2756	−5.70
3	0.2155	−4.00	15	1.5474	−7.50
4	0.2095	−3.20	16	1.7842	−1.90
5	0.2870	−9.80	17	2.0169	−7.60
6	0.2986	−1.20	18	2.8294	−12.2
7	0.3752	−3.40	19	3.0219	−9.80
8	0.5055	−5.20	20	3.6187	−11.4
9	0.3681	−7.60	21	4.1067	−14.9
10	0.3697	−3.00	22	4.2790	−9.20
11	0.5700	−8.90	23	4.7834	−11.3
12	0.5283	−9.00			

In Table 1, as the channel model delays are normalized, they need to be scaled according to a desired delay spread in nanoseconds (ns):

$$\tau_{scaled} = \tau_{model} DS_{ns} \quad (27)$$

in which τ_{model} is the normalized delay value of the TDL model, τ_{scaled} is the new delay value (in [ns]), and DS_{ns} is the desired delay spread (in [ns]). From Table 7.7.3-1 [64], examples of scaling delay spreads are very short ($DS_{ns} = 10$ ns), short ($DS_{ns} = 30$ ns), nominal ($DS_{ns} = 100$ ns), long ($DS_{ns} = 300$ ns), and very long ($DS_{ns} = 1000$ ns). In this work, we use a delay spread of $DS_{ns} = 50$ ns.

From the channel model of Table 1, a Rayleigh distribution is used to compute each sub-channel of $\mathbf{H} \in \mathbb{C}^{M_T \times M_R}$ (MIMO channel matrix). The M-QAM BER figure for the AWGN channel is also used to define a lower bound on BER vs. E_b/N_0 performance. Additionally, it is assumed that all received signals are uncorrelated [11].

A realistic scenario to assess the performance of a MIMO-OFDM system must also include the nonlinear effects of the transmitter power amplifiers [11]. This is necessary because the OFDM signal can have relatively high peak values (i.e., high PAPR) in the time domain since many subcarrier components are added via an IFFT operation. A high PAPR is one of the most detrimental aspects of the OFDM system, as it decreases the SQNR (signal-to-quantization noise ratio) of ADCs (analog-to-digital converters) and DACs (digital-to-analog converters), while also imposing a back-off that degrades the efficiency of the power amplifier in the transmitter. The PAPR issue is usually more critical in the uplink since the efficiency of the power amplifier is critical due to the limited battery power in a mobile terminal. For this purpose, from now on, the results assume mild amplifier nonlinearities, represented by a first-grade power amplifier or an appropriate back-off operating point [11]. Based on [41], the nonlinearity vector $\rho = [\rho_1 \ \rho_2 \ \rho_3]^T = [0.9 \ 0.1 \ 0.05]^T$ implies 90%, 10%, and 5% first, second, and third-order coefficients, respectively.

With the model properly validated and the specified 5G channel model, we are now able to analyze the proposed complex-valued ANN-based decoder for MIMO-OFDM systems. First, we present the MSE convergence curves during the learning process. The MSE curves are averaged over 10 subsequent simulation traces, and a 4-QAM modulation with $E_b/N_0 = 12$ dB is employed. Figure 11a–d show the MSE evolution for $M_T = M_R = 4$, $M_T = M_R = 8$, $M_T = M_R = 16$, and $M_T = M_R = 32$ antennas, respectively. The red bottom and top curves in Figure 11 refer to the MSE standard deviations over the 10 subsequent simulation traces, and the green curves refer to the mean values. Although the steady-state MSEs decrease slightly as the number of antennas increases, one may notice that the decays of the standard deviations are more conspicuous as $M_T = M_R$ increases. This is due to the MIMO characteristics that mitigate the channel effects by sending several samples of the same signal to the receiver. Thus, sudden channel variations are smoothed, suggesting that the PTRBF learning process presents a robust and cohesive behavior.

The 4-QAM scatter plots presented in Figure 12 show the convergence of the proposed neural network decoder for the first 35 training epochs. For this sequence of scatter plots, each training epoch corresponds to one OFDM symbol; i.e., 256 4-QAM symbols. As shown in Figure 12, the proposed algorithm has a fast convergence rate since only 10 training epochs are sufficient to separate the 4-QAM constellation symbols efficiently [11].

The 16-QAM scatter plots presented in Figure 13 show the convergence of the proposed neural network decoder for the first 140 training epochs, spaced in intervals of 20 OFDM symbols. In this case, the number of training epochs necessary for algorithm convergence is greater than for the 4-QAM case, given the intrinsic complexity of the higher-order constellation. Nevertheless, with only 20 training epochs, it is already possible to visually identify the 16-QAM constellation symbols and, with 80 training epochs, to see the correctly grouped symbols in the scatter plot [11].

The 64-QAM scatter plots presented in Figure 14 show the convergence of the proposed neural network decoder for the first 7000 training epochs, spaced in intervals of 1000 OFDM symbols. In this case, the number of training epochs necessary for algorithm convergence is

greater than for the former 16/4-QAM cases, in view of the intrinsic complexity of the much higher-order 64-QAM constellation. Nevertheless, with 3000 training epochs, it is already possible to visually identify the 64-QAM constellation symbols and, with 5000 training epochs, to see the correctly grouped symbols in the scatter plot [11]. Although an abrupt increase in training epochs occurs, when compared with 4-QAM in Figure 12, it represents a time interval of only 80 ms (5000 OFDM symbols with 240 kHz sub-carrier spacing).

After analyzing the MSE curves and constellations of the MMPTRBF, we can further investigate the BER vs. E_b/N_0 of the proposed approach. Figure 15 shows the BER vs. E_b/N_0 results of the QOSTBC, EQOSTBC, and MIMO-PTRBF systems operating with $M_T = M_R = 4$ antennas and 4-QAM modulation. The 4-QAM AWGN curve defines a lower bound for all MIMO systems that use 4-QAM modulation. The simulation results in Figure 15 indicate that the QOSTBC system outperforms the proposed work when perfect channel knowledge is available at the receiver, which is impractical. Although the EQOSTBC system is a feasible and practical version of the QOSTBC, due to the channel estimation block at the receiver, simulations using the least-squares channel estimation show that the EQOSTBC performance is degraded by more than 2.5 dB when compared with the QOSTBC [11]. Furthermore, even with a perfect channel estimator, it is computationally expensive to decode QOSTBC codes with maximum likelihood for more than four antennas, as addressed by [56].

Figure 11. Evolution of MSE values averaged over 10 realization sequences of the proposed MMPTRBF network decoder, using 4-QAM and $E_b/N_0 = 12$ dB for (a) $M_T = M_R = 4$ antennas, (b) $M_T = M_R = 8$ antennas, (c) $M_T = M_R = 16$ antennas, (d) $M_T = M_R = 32$ antennas.

Figure 12. Scatter plots for the 4-QAM symbols at the output of the MMPTRBF during the training period, for $M_T = M_R = 4$ antennas and $E_b/N_0 = 12$ dB. (a) 1 epoch, (b) 5 epochs, (c) 10 epochs, (d) 15 epochs, (e) 20 epochs, (f) 25 epochs, (g) 30 epochs, and (h) 35 epochs.

Figure 13. Scatter plots for the 16-QAM symbols at the output of the MMPTRBF during the training period, for $M_T = M_R = 4$ antennas and $E_b/N_0 = 20$ dB. (a) 1 epoch, (b) 20 epochs, (c) 40 epochs, (d) 60 epochs, (e) 80 epochs, (f) 100 epochs, (g) 120 epochs, and (h) 140 epochs.

Figure 14. Scatter plots for the 64-QAM symbols at the output of the MMPTRBF in the training phase, for $M_T = M_R = 4$ antennas and $E_b/N_0 = 26$ dB. (a) 1 epoch, (b) 1000 epochs, (c) 2000 epochs, (d) 3000 epochs, (e) 4000 epochs, (f) 5000 epochs, (g) 6000 epochs, and (h) 7000 epochs.

Figure 15. Systems with $M_T = M_R = 4$ antennas for 4-QAM modulation.

Figures 16 and 17 show the $\text{BER} \times E_b/N_0$ results of the QOSTBC, EQOSTBC, and MIMO-PTRBF systems operating with $M_T = 4$ and $M_T = 8$ with $M_R = 1$ antennas and 4-QAM modulation. Figures 16 and 17 highlight the diversity gain of the proposed system when compared with the EQOSTBC. Although the mathematical derivation of the proposed system diversity gain has not been obtained yet, simulations indicate a significant diversity gain [11]. Contrasting Figure 15 with Figure 16, one can see the increase in the diversity gain as the number of transmitting antennas increases.

Figure 16. Systems with $M_T = 4$ and $M_R = 1$ antennas for 4-QAM modulation.

Figure 17. Systems with $M_T = 8$ and $M_R = 1$ antennas for 4-QAM modulation.

Figure 18 shows the BER vs. E_b/N_0 results of the MMPTRBF systems operating with $M_T = 4, 6$, and 8 antennas for a 4-QAM modulation. Figure 18 also highlights the increase in the diversity gain as the number of transmitting antennas increases. For a $\text{BER} = 10^{-2}$, there is a gain of approximately 2 dB when doubling M_T .

To further investigate the effects of a larger number of transmitting and receiving antennas on the performances of BER, in Figure 19, we present the simulation results for a higher-order system with $M_T = M_R = 8$. It can be seen that the performance of the quasi-orthogonal code with channel estimation is worse for $M_T = 8$ than for $M_T = 4$ antennas, as shown in Figure 19. It is shown in [53] that the performance of the linear estimator decreases proportionally with the number of transmitting antennas, which adds a constraint to the number of transmitting antennas for linearly decoded systems.

Figure 18. MMPTRBF with $M_T = 4$, $M_T = 8$, and $M_R = 1$ antennas for 4-QAM modulation.

Figure 19. Systems with $M_T = M_R = 4$, $M_T = M_R = 8$ antennas for 4-QAM modulation.

Figure 20 presents the simulation results for $M_T = M_R = 4$ ($N = 100$), $M_T = M_R = 8$ ($N = 150$), $M_T = M_R = 16$ ($N = 200$), and $M_T = M_R = 32$ ($N = 600$) to examine the extent of the proposed work for massive MIMO operations. These results show the potential of the proposed work to efficiently operate with a massive number of transmitting and receiving antennas. In addition, taking as reference the $\text{BER} = 10^{-3}$, one can notice that the gain increment is reduced when doubling the number of antennas as M_T increases. For instance, increasing from $M_T = M_R = 4$ to $M_T = M_R = 8$ yields a gain of 1.22 dB, while increasing from $M_T = M_R = 16$ to $M_T = M_R = 32$ yields a gain of only 0.2 dB.

It is important to highlight that, in contrast to the maximum likelihood detection, the proposed MMPTRBF decoding is able to operate with more than 16 antennas due to its reduced computational complexity, as discussed below in Section 5.

Figure 20. MMPTRBF with $M_T = M_R = 4$, $M_T = M_R = 8$, $M_T = M_R = 16$, $M_T = M_R = 32$, antennas for 4-QAM modulation.

Figure 21 shows the BER vs. E_b/N_0 results of the ML-QOSTBC and ML-LS-QOSTBC with 16-PSK and MIMO-PTRBF with 16-QAM to further examine the extent of the proposed work for higher-order modulations with $M_T = M_R = 4$ antennas. This result shows that the proposed work operates efficiently with a higher modulation order of 16-QAM. It is important to note, however, that different modulation formats are used in this scenario because the maximum likelihood QOSTBC decoding is not capable of dealing with quadrature amplitude modulation, as addressed in [56]. For this reason, in order to keep 16-order modulation, a 16-PSK modulation format for QOSTBC is used. Figure 21 shows that the proposed work outperforms the QOSTBC using maximum likelihood (which is the optimal decoder for 16-PSK) by about 2 dB and outperforms the channel estimated scenario by more than 4 dB. Although this robust result seems to show a great advantage of using the proposed approach, we should be careful as it is not quite fair to compare 16-QAM and 16-PSK formats under the proposed nonlinear scenario.

Figure 21. Systems with $M_T = M_R = 4$ antennas operating at the same bitrate with 16-QAM (MMPTRBF) and 16-PSK (ML-QOSTBC and ML-LS-QOSTBC).

Figure 22 shows the BER vs. E_b/N_0 results of the MIMO-PTRBF systems operating with $M_T = M_R = 4$ and $M_T = M_R = 8$ for 16-QAM modulation. Figure 22 highlights the increase in the diversity gain as a result of the increase in the number of transmitting antennas. Although the MMPTRBF with $M_T = M_R = 8$ has the worst results for low values of E_b/N_0 , for values above $E_b/N_0 = 10$ dB, the performance with $M_T = M_R = 8$ surpasses the $M_T = M_R = 4$ results by about 1.6 dB for BER = 10^{-3} .

Figure 22. MMPTRBF with $M_T = M_R = 4$ and $M_T = M_R = 8$ antennas operating at the same bitrate with 16-QAM modulation.

Figure 23 shows the BER vs. E_b/N_0 results of the MIMO-PTRBF with 64-QAM modulation to further investigate the proposed work for higher-order modulations with $M_T = M_R = 4$ antennas. Figure 23 shows the potential of the proposed approach of working with high-order modulation and highlights the gains over the QOSTBC with 64-PSK and perfect channel estimation, by about 5 dB, and for the channel estimated scenario by about 7 dB. It is important to emphasize, once again, that it is not quite fair to compare 64-QAM and 64-PSK formats under the proposed nonlinear scenario.

Figure 23. Systems with $M_T = M_R = 4$ antennas operating at the same bitrate with 64-QAM (MMPTRBF) and 64-PSK (ML-QOSTBC and ML-LS-QOSTBC).

Figure 24 shows the BER vs. E_b/N_0 results of the MIMO-PTRBF systems operating with $M_T = M_R = 4$ and $M_T = M_R = 8$ for 64-QAM modulation. Figure 24 highlights the increase in the diversity gain as a result of the increase in transmitting antennas. When doubling the number of antennas from $M_T = M_R = 4$ to $M_T = M_R = 8$, the gain is about 0.87 dB for $\text{BER} = 10^{-3}$.

Figure 24. MMPTRBF with $M_T = M_R = 4$ and 8 antennas operating at the same bitrate with 64-QAM modulation.

5. Computational Complexities

Table 2 presents the computational complexities of the OSTBC and EOSTBC with ML decoding, both with the additional complexity of channel estimation, and the proposed scheme with MIMO-PTRBF for training and decoding operation modes. M_T and M_R are the number of transmitting and receiving antennas, M_s is the number of transmitted symbols per MIMO block, P is the number of time samples per block of coded symbols, A is the constellation order (e.g., $A = 4$ in the case of 4-QAM), and N is the number of neurons used in the PTRBF neural network. Since the $\exp(\cdot)$ function can be easily implemented in hardware by lookup tables, multiplication is the most costly operation. One may note that, for $M_T \leq 8$, the complexity of the proposed algorithm is similar to the complexity of the OSTBC, for which no code exists for $M_T > 8$ [54]. The case of QOSTBC is similar, for which no simplified ML metric exists for $M_T \neq 4$ [11,56].

Table 2. Computational complexities.

Decoder	Multiplications	Additions	$\exp(\cdot)$
OSTBC [†]	$(2PM_R + 4)M_sA + ch$	$(3PM_R + 3)M_sA + ch$	ch
QOSTBC [†]	$24M_R(M_s/2)A^2 + ch$	$16(M_R - 1)(M_s/2)A^2 + ch$	ch
ML* with $R = 1$	$M_R M_T P(A^{M_s}) + ch$	$M_R M_T (P - 1)(A^{M_s}) + ch$	ch
MIMO-PTRBF (train)	$12NM_s + 12N + 8NM_R P + 2M_s$	$6NM_R P + 12NM_s - 2N$	$2N$
MIMO-PTRBF (decoding)	$2NM_R P + 2N + 4NM_s$	$4NM_R P + 4NM_s - 2N - 2M_s$	$2N$

[†] ch refers to the additional complexity of channel estimation.

Table 3 presents the computational complexities for the OSTBC, QOSTBC, and the proposed system, for $M_T = M_R = 4$. In Table 3, $N = 100$ neurons are used in the neural network, and maximum likelihood decoding is simulated with $R = 1$. Note that generic maximum likelihood decoding refers to the minimization of Equation (5) for a rate one ($R = 1$) coding scheme (e.g., it could decode the QOSTBC for the case $M_T = 4$ at a higher computational cost), and it will be assumed as an upper bound for the computational

complexities of the other quasi-orthogonal systems. Furthermore, appropriate modulation schemes are used to provide the desired transmission rate for the evaluated systems; i.e., 4-QAM for rate one code ($R = 1$) and 16-QAM for half-rate code ($R = 1/2$) [11].

Table 3. Computational complexities for $M_T = M_R = 4$ and 2 bits/s/Hz.

Decoder	Multiplications	Additions	exp($\cdot$)
OSTBC [†]	$4.35 \times 10^3 + ch$	$6.34 \times 10^3 + ch$	ch
QOSTBC [†]	$3.07 \times 10^3 + ch$	$1.54 \times 10^3 + ch$	ch
ML for $R = 1$ [†]	$1.64 \times 10^4 + ch$	$1.23 \times 10^4 + ch$	ch
MIMO-PTRBF (train)	1.88×10^4	1.46×10^4	200
MIMO-PTRBF (test)	5.00×10^3	7.79×10^3	200

[†] ch refers to the additional complexity of channel estimation.

Table 4 displays the computational complexities for $M_T = M_R = 8$, when the PTRBF is equipped with $N = 150$ neurons. QOSTBC is defined as not applicable since no simplified ML decoding metric has been presented in the literature. Thus, we need to rely on the usual ML metric to perform decoding, which implies the limitation of using QOSTBC combined with ML for a higher number of antennas in practical approaches.

Table 4. Computational complexities for $M_T = M_R = 8$ and 2 bits/s/Hz.

Decoder	Multiplications	Additions	exp($\cdot$)
OSTBC [†]	$3.33 \times 10^4 + ch$	$4.95 \times 10^4 + ch$	ch
QOSTBC [†]	Not applicable	Not applicable	Not applicable
ML for $R = 1$ [†]	$3.36 \times 10^7 + ch$	$2.94 \times 10^7 + ch$	ch
MIMO-PTRBF (train)	9.30×10^4	7.23×10^4	300
MIMO-PTRBF (test)	2.43×10^4	4.29×10^4	300

[†] ch refers to the additional complexity of channel estimation.

Table 5 displays the computational complexities for $M_T = M_R = 32$, when the PTRBF is equipped with $N = 600$ neurons. As the OSTBC coding matrix is limited to eight antennas (see [54]), it is not applicable for $M_T = M_R = 32$. As already mentioned, since there is no simplified ML metric to perform ML decoding with QOSTBC, it results in an explosion of computational complexity for $M_T = M_R = 32$. On the other hand, the proposed approach can expand the number of antennas, maintaining a reasonable compromise between computational complexity and BER, as discussed in Section 4.

Table 5. Computational complexities for $M_T = M_R = 32$ and 2 bits/s/Hz.

Decoder	Multiplications	Additions	exp($\cdot$)
OSTBC [†]	Not applicable	Not applicable	Not applicable
QOSTBC [†]	Not applicable	Not applicable	Not applicable
ML for $R = 1$ [†]	$6.04 \times 10^{23} + ch$	$5.86 \times 10^{23} + ch$	ch
MIMO-PTRBF (train)	5.15×10^6	3.92×10^6	1200
MIMO-PTRBF (test)	1.31×10^6	2.53×10^6	1200

[†] ch refers to the additional complexity of channel estimation.

The decoding computational complexities, shown in Figure 25, are addressed in terms of real-valued multiplications per MIMO symbol, as a function of $M_T = M_R$ antennas. The orthogonal and quasi-orthogonal systems are not illustrated for the entire simulation range in Figure 25 due to the absence of coding matrices and the simplification of the ML metric for configurations with $M_T = M_R > 8$ [11].

Figure 25. Computational complexities as a function of $M_T = M_R$.

6. Conclusions

This work proposes a novel MIMO scheme for M-QAM systems that aims to achieve diversity gain for any number of antennas and at a lower computational cost when compared with traditional methods. The presented architecture is based on existing systems but with substantial improvements in the coding and decoding methods, based on conventional MIMO-OFDM systems with quasi-orthogonal coding but implemented with complex-valued Radial Basis Functions neural networks. The state-of-the-art algorithms and the proposed approach have been simulated in MATLAB to measure their relative performance under fading scenarios.

Based on the synergistic combination of the coding and decoding algorithms presented in Section 2, the proposed MIMO-PTRBF system is discussed and analyzed in Section 3. The main functional features of the proposed architecture can be summarized as follows: (1) the proposed coding algorithm generalizes the generation of quasi-orthogonal coding matrices, (2) the MIMO-PTRBF algorithm decodes the signal with satisfactory performance and feasible computational cost, presenting low steady-state MSE with fast convergence, and (3) the proposed approach seems practically feasible, at least for 32×32 MIMO systems, which are simulated in this work. We conjecture the practical feasibility of higher-order systems if faster hardware, such as FPGAs, is used.

The MIMO-PTRBF algorithm has been proposed in this work to implement massive MIMO schemes as an alternative to the classic MIMO-OSTBC systems under maximum likelihood detection. Simulations have shown that the proposed technique has a great potential to improve the signal-to-noise ratio at the receiver, with competitive computational complexity. Although there are recent works in the literature proposing techniques for MIMO decoding, they are focused on reducing computational complexity at the cost of performance and are limited by the simulated ML decoding. In this work, results show that the proposed approach achieves better results than ML decoding for higher-order modulation schemes with nonlinearities from power amplifiers, keeping a competitive computational complexity. Moreover, the proposed system is easily scalable in terms of the number of antennas, meaning that a wide range of transmitting and receiving antennas can be used. This is especially important for the next generations of mobile communications, such as 5G, 6G, and probably beyond.

The proposed architectures and algorithms find potential applications in some configurations of the next generations of wireless systems. For example, some specialized hardware improvements currently aim exclusively at real-time neural network algorithms. These are intended to be implemented in low-power graphical processing units (LPGPUs), favoring the speed and energy consumption of these algorithms. Therefore, the proposed

architecture will be able to work with low-power consumption devices, with the ability to handle the distortions of nonlinear power amplifiers while maintaining a fast convergence rate. It should be emphasized that a fast convergence characteristic is essential for wireless channels with dynamic fluctuations.

This paper addresses some crucial aspects of MIMO-OFDM coding and decoding schemes for quasi-static channels. A complementary analysis of dynamic scenarios is also presented. We conjecture that the proposed work may be further improved using additional techniques, such as a mathematical approach for designing an optimum adaptive configuration.

Furthermore, it would be interesting to study and validate the proposed architecture for dynamic scenarios. In addition, as challenging and promising future work, the proposed algorithm can be adapted and implemented in advanced optical communication systems with Spatial Division Multiplexing (SDM), which is similar to a MIMO wireless system.

7. Patents

A patent application with the results presented in this paper is being prepared by Inova—UNICAMP Innovation Agency.

Author Contributions: All the authors have contributed in various degrees to ensuring the quality of this work: system conceptualization, J.A.S.; MIMO-PTRBF conceptualization, K.S.M.; methodology, J.A.S. and K.S.M.; formal analysis, J.A.S., K.S.M., F.C.C.d.C., and D.S.A.; writing—original draft preparation, J.A.S.; theorems, corollaries, and proofs, K.S.M.; writing—review and editing, J.A.S., K.S.M., F.C.C.d.C., and D.S.A.; supervision, D.S.A. All authors have read and agreed to the published version of the manuscript.

Funding: Jonathan A. Soares is supported in part by the National Council for Scientific and Technological Development, Brasil (CNPq), with Grant Number 132545/2019-5.

Kayol S. Mayer is supported in part by the Coordenação de Aperfeiçoamento de Pessoal de Nível Superior, Brasil (CAPES), with Finance Code 001.

Institutional Review Board Statement: Not Applicable.

Informed Consent Statement: Not Applicable.

Data Availability Statement: Not applicable.

Acknowledgments: We would like to thank Andrew Miller for the brilliant publication on his website on 9 August 2015, entitled “Trans-dimensional change of variables and the integral Jacobian”. <https://andymiller.github.io/2015/08/09/integral-jacobian.html> accessed on 8 May 2021.

Conflicts of Interest: The authors declare no conflict of interest.

Appendix A

Definition A1. A non-monotone and differentiable transformation $\mathbf{g} : \mathbb{R}^N \mapsto \mathbb{R}^N$, composed of M -real solutions, yields a multivariate change of probability density function:

$$f_Y(\mathbf{y}) = \sum_{m=1}^M f_X(\mathbf{x}) \left| J(\mathbf{g}_m^{-1}(\mathbf{y})) \right|,$$

in which $\mathbf{x}, \mathbf{y} \in \mathbb{R}^N$ are random vectors, $J(\cdot)$ is the determinant of the Jacobi operator, and $\mathbf{g}_m^{-1}$ is the inverse transformation of $\mathbf{g}$, regarding the m th real solution (see [65], p. 322 and [66], pp. 69–70).

Theorem A1. If $\mathbf{x} \in \mathbb{R}^N$ is an input random vector of a kernel function $g(\mathbf{x}) = y \triangleq \exp \left\{ -\frac{\|\mathbf{x} - \gamma\|_2^2}{z} \right\}$, with constant $\gamma \in \mathbb{R}^N$ and $z \in \mathbb{R}$; then, g performs a trans-dimensional transformation of probability density function from $f_X(\mathbf{x})$ to $f_Y(y)$.

Proof. Let the Gaussian kernel function $g(\mathbf{x}, \gamma, z) \triangleq \exp \left\{ -\frac{\|\mathbf{x} - \gamma\|_2^2}{z} \right\}$, with constant center vector γ and variance z , be denoted by $y = g(\mathbf{x})$, since it is only dependent on the random vector $\mathbf{x} \in \mathbb{R}^N$. Furthermore, let $\mathbf{y} \triangleq [g(\mathbf{x}) \ y_2 \ \cdots \ y_N]^T$ be the auxiliary expanded vector of mapping. With Definition A1, we have the following transformation of the probability density function (PDF):

$$f_Y(\mathbf{y}) = \sum_{m=1}^M f_X(\mathbf{x}) \left| J(\mathbf{g}_m^{-1}(\mathbf{y})) \right|. \quad (\text{A1})$$

However, we need to marginalize the extra dimensions of $\mathbf{y}$ in (A1) to obtain the PDF of the desired random variable $f_Y(y)$:

$$f_Y(y) = \sum_{m=1}^M \int f_X(\mathbf{x}) \left| J(\mathbf{g}_m^{-1}(\mathbf{y})) \right| dy_2 \cdots dy_N. \quad (\text{A2})$$

Then, finding some non-monotone function $h(y) = h(g(\mathbf{x})) = f_X(\mathbf{x})$, in which $h(y)$ is not a function of the expanded auxiliary terms $y_2, y_3, \dots, y_N$, we can rewrite (A2) as

$$f_Y(y) = \sum_{m=1}^M h_m(y) \int \left| J(\mathbf{g}_m^{-1}(\mathbf{y})) \right| dy_2 \cdots dy_N, \quad (\text{A3})$$

where the right-hand term under the integral is called integral Jacobian, and it is responsible for the volume correction [67] of the trans-dimensional transformation.

In addition, as the Gaussian kernel is a non-monotone function with two real solutions (due to the Euclidean norm), (A3) results in

$$f_Y(y) = h_1(y) \int \left| J(\mathbf{g}_1^{-1}(\mathbf{y})) \right| dy_2 \cdots dy_N + h_2(y) \int \left| J(\mathbf{g}_2^{-1}(\mathbf{y})) \right| dy_2 \cdots dy_N, \quad (\text{A4})$$

which is the trans-dimensional transformation of PDF performed by a Gaussian kernel function. $\square$

Corollary A1. For any complex-valued vectors of same size $\mathbf{x}, \mathbf{y} \in \mathbb{C}^N$ and a scalar $z \in \mathbb{C}$, a kernel function $f(\mathbf{x}, \mathbf{y}, z) \triangleq \exp \left\{ -\frac{\|\Re(\mathbf{x}) - \Re(\mathbf{y})\|_2^2}{\Re(z)} \right\} + j \exp \left\{ -\frac{\|\Im(\mathbf{x}) - \Im(\mathbf{y})\|_2^2}{\Im(z)} \right\}$ have real and imaginary boundaries between 0 and 1, if $\Re(z) > 0$ and $\Im(z) > 0$.

Proof. The proof is straightforward using Theorem A1 of [35], where the real-valued analyses must be independently performed for the real and imaginary components of $\mathbf{x}$, $\mathbf{y}$, and z . $\square$

Corollary A2. If $\mathbf{x} \in \mathbb{C}^N$ is a complex-valued input random vector of a complex-valued kernel function $g(\mathbf{x}) = y \triangleq \exp \left\{ -\frac{\|\Re(\mathbf{x}) - \Re(\gamma)\|_2^2}{\Re(z)} \right\} + j \exp \left\{ -\frac{\|\Im(\mathbf{x}) - \Im(\gamma)\|_2^2}{\Im(z)} \right\}$, with constant $\gamma \in \mathbb{C}^N$ and $z \in \mathbb{C}$, then g performs a complex-valued trans-dimensional transformation of probability density function from $f_{\Re(\mathbf{x})}(\Re(\mathbf{x})) + j f_{\Im(\mathbf{x})}(\Im(\mathbf{x}))$ to $f_{\Re(\mathbf{y})}(\Re(\mathbf{y})) + j f_{\Im(\mathbf{y})}(\Im(\mathbf{y}))$, with independent real and imaginary components.

Proof. The proof is straightforward using Theorem A1, where the real-valued analyses must be independently performed for the real and imaginary components of $\mathbf{x}$ and y . $\square$

Theorem A2. If $\mathbf{x} \in \mathbb{C}^K$ is a stationary complex-valued input random vector of a PT-RBF, then the matrix of synaptic weights $\mathbf{W}[k] \in \mathbb{C}^{M \times N}$ converges in the mean to the optimum matrix of synaptic weights $\mathbf{W}_o$ when $k \rightarrow \infty$, if the matrix of center vectors and the vector of variances are constants.

Proof. Let the PT-RBF output $\mathbf{y}[k] = \mathbf{W}[k]\boldsymbol{\phi}[k] \in \mathbb{C}^M$, where $\boldsymbol{\phi}[k] \in \mathbb{C}^N$ is the vector of the neuron output. Furthermore, let the update function of the matrix of synaptic weights $\mathbf{W}[k+1] = \mathbf{W}[k] + \eta_w \mathbf{e}[k]\boldsymbol{\phi}^H[k] \in \mathbb{C}^{M \times N}$, where η_w is the adaptive step and $\mathbf{e}[k] \in \mathbb{C}^M$ is the vector of errors. We can assume without loss in generality that the desired response is

$$\mathbf{d}[k] = \mathbf{W}_o \boldsymbol{\phi}[k] + \mathbf{q}[k] \in \mathbb{C}^M, \quad (\text{A5})$$

where $\mathbf{q}[k] \in \mathbb{C}^M$ is a complex-valued white Gaussian noise vector, with zero mean and variance σ_q^2 , which is uncorrelated with $\boldsymbol{\phi}[k]$. Then, substituting (A5) into the error equation $\mathbf{e}[k] = \mathbf{d}[k] - \mathbf{y}[k]$, we have

$$\mathbf{e}[k] = \mathbf{W}_o \boldsymbol{\phi}[k] + \mathbf{q}[k] - \mathbf{W}[k]\boldsymbol{\phi}[k]. \quad (\text{A6})$$

Thus, replacing (A6) into the update function of the matrix of synaptic weights:

$$\mathbf{W}[k+1] = \mathbf{W}[k] + \eta_w \mathbf{W}_o \boldsymbol{\phi}[k]\boldsymbol{\phi}^H[k] + \eta_w \mathbf{q}[k]\boldsymbol{\phi}^H[k] - \eta_w \mathbf{W}[k]\boldsymbol{\phi}[k]\boldsymbol{\phi}^H[k]. \quad (\text{A7})$$

Subtracting $\mathbf{W}_o$ from both sides of (A7), the synaptic weights error matrix $\mathbf{V}[k] = \mathbf{W}[k] - \mathbf{W}_o$ can be given by

$$\mathbf{V}[k+1] = \mathbf{V}[k] - \eta_w \mathbf{V}[k]\boldsymbol{\phi}[k]\boldsymbol{\phi}^H[k] + \eta_w \mathbf{q}[k]\boldsymbol{\phi}^H[k]. \quad (\text{A8})$$

Applying the expectation operator to both sides of (A8):

$$\mathbf{E}(\mathbf{V}[k+1]) = \mathbf{E}\left(\mathbf{V}[k] - \eta_w \mathbf{V}[k]\boldsymbol{\phi}[k]\boldsymbol{\phi}^H[k] + \eta_w \mathbf{q}[k]\boldsymbol{\phi}^H[k]\right), \quad (\text{A9})$$

and employing the independence assumptions of $\mathbf{W}[k] \perp \boldsymbol{\phi}[k]$ and $\mathbf{q}[k] \perp \boldsymbol{\phi}[k]$:

$$\mathbf{E}(\mathbf{V}[k+1]) = \mathbf{E}(\mathbf{V}[k]) - \eta_w \mathbf{E}(\mathbf{V}[k])\mathbf{E}(\boldsymbol{\phi}[k]\boldsymbol{\phi}^H[k]) + \eta_w \mathbf{E}(\mathbf{q}[k])\mathbf{E}(\boldsymbol{\phi}^H[k]), \quad (\text{A10})$$

which results in:

$$\mathbf{E}(\mathbf{V}[k+1]) = \mathbf{E}(\mathbf{V}[k])\left[\mathbf{I} - \eta_w \mathbf{E}(\boldsymbol{\phi}[k]\boldsymbol{\phi}^H[k])\right] = \mathbf{E}(\mathbf{V}[k])\left[\mathbf{I} - \eta_w \mathbf{R}_{\phi\phi}\right], \quad (\text{A11})$$

where $\mathbf{I}$ is the identity matrix and $\mathbf{R}_{\phi\phi}$ is the correlation matrix of the neuron outputs. As $\mathbf{R}_{\phi\phi}$ is Hermitian and positive semidefinite (see [68], pp. 387, 469), it can be rotated into a diagonal matrix by the unitary transformation $\mathbf{R}_{\phi\phi} = \mathbf{Q}\boldsymbol{\Lambda}\mathbf{Q}^H$, in which $\boldsymbol{\Lambda} = \text{diag}(\lambda_1 \ \lambda_2 \ \dots \ \lambda_N)$ is the diagonal matrix of real and positive eigenvalues of $\mathbf{R}_{\phi\phi}$ (see [68], p. 471), in the form of $\lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_N$, and $\mathbf{Q} \in \mathbb{C}^{N \times N}$ is the orthonormal matrix of eigenvectors that diagonalizes $\mathbf{R}_{\phi\phi}$ through a similarity transformation [69].

Then, rotating $\mathbf{E}(\mathbf{V}[k])$ by the matrix of eigenvectors $\mathbf{Q}$ —i.e., $\tilde{\mathbf{V}}[k] = \mathbf{E}(\mathbf{V}[k])\mathbf{Q}$ —decouples the evolution of its coefficients. By means of this rotation, we can express the modes of convergence (see [70], p. 77) of (A11) as

$$\tilde{\mathbf{V}}[k+1] = \tilde{\mathbf{V}}[k](\mathbf{I} - \eta_w \boldsymbol{\Lambda}). \quad (\text{A12})$$

As $(\mathbf{I} - \eta_w \boldsymbol{\Lambda})$ is diagonal and the n th row of $\tilde{\mathbf{V}}[k]$ represents the projection of $\mathbf{E}(\mathbf{V}[k])$ onto the n th eigenvector of $\mathbf{R}_{\phi\phi}$, all elements of $\tilde{\mathbf{V}}[k]$ evolve independently. Hence, (A12) converges to zero if $|1 - \eta_w \lambda_n| < 1$ (see [69], p. 84). As the fastest mode of convergence corresponds to the maximum eigenvalue λ_{max} , using the identity $\lambda_1 = \lambda_{max} \leq \text{tr}(\mathbf{R}_{\phi\phi})$ (see [70], p. 77), the condition for the convergence in the mean becomes

$$0 < \eta_w < \frac{2}{\lambda_{max}} \approx \frac{2}{\text{tr}(\mathbf{R}_{\phi\phi})}. \quad (\text{A13})$$

As the trace of $\mathbf{R}_{\phi\phi}$ is equal to the product of the number of neurons outputs and the respective signal power, the adaptive step bound is given by

$$0 < \eta_w < \frac{2}{NE(|\phi[k]|^2)}. \quad (\text{A14})$$

Note that the convergence in the mean of both PT-RBF and complex LMS is similar (see [70], p. 77), and it is natural in some way, considering that after the trans-dimensional transformation step of the PT-RBF, both algorithms have comparable architectures.

However, in view of Corollary A2, $E(|\phi[k]|^2)$ can be difficult to obtain. To circumvent this issue, using Corollary A1, we can replace $\phi[k]$ by its maximum value $(1 + j)$ into (A14), which yields

$$0 < \eta_w < \frac{2}{N|1 + j|^2} = \frac{1}{N}, \quad (\text{A15})$$

which is the adaptive step bound for the convergence in the mean. $\square$

References

- Asif, R.M.; Arshad, J.; Shakir, M.; Noman, S.M.; Rehman, A.U. Energy efficiency augmentation in massive MIMO systems through linear precoding schemes and power consumption modeling. *Wirel. Commun. Mob. Comput.* **2020**, *2020*, 1–13. [\[CrossRef\]](#)
- Shang, B.; Liu, L.; Chen, H.; Zhang, C.J.; Pudlewski, S.; Bentley, E.S.; Ashdown, J.D. Spatial spectrum sensing in uplink two-tier user-centric deployed HetNets. *IEEE Trans. Wirel. Commun.* **2020**, *19*, 7957–7972. [\[CrossRef\]](#)
- Temiz, M.; Alsusa, E.; Baidas, M.W. A dual-functional massive MIMO OFDM communication and radar transmitter architecture. *IEEE Trans. Veh. Technol.* **2020**, *69*, 14974–14988. [\[CrossRef\]](#)
- Mirfarshbafan, S.H.; Shabany, M.; Nezamalhoseini, S.A.; Studer, C. Algorithm and VLSI design for 1-Bit data detection in massive MIMO-OFDM. *IEEE Open J. Circuits Syst.* **2020**, *1*, 170–184. [\[CrossRef\]](#)
- Ghzaoui, M.E.; Hmamou, A.; Foshi, J.; Mestoui, J. Compensation of non-linear distortion effects in MIMO-OFDM systems using constant envelope OFDM for 5G applications. *J. Circuits Syst. Comput.* **2020**, *29*, 1–21. [\[CrossRef\]](#)
- Yoon, E.; Kwon, S.; Yun, U.; Kim, S.Y. LDPC Decoding With Low Complexity for OFDM Index Modulation. *IEEE Access* **2021**, *9*, 68435–68444. [\[CrossRef\]](#)
- Marey, M.; Mostafa, H. Turbo Modulation Identification Algorithm for OFDM Software-Defined Radios. *IEEE Commun. Lett.* **2021**, *25*, 1707–1711. [\[CrossRef\]](#)
- Hwang, S.; Seo, J.; Park, J.; Kim, H.; Jeong, B.J. Compressive Sensing-Based Radar Imaging and Subcarrier Allocation for Joint MIMO OFDM Radar and Communication System. *Sensors* **2021**, *21*, 2382. [\[CrossRef\]](#)
- Condoluci, M.; De Valgas, J.F.; Monserrat, J.F.; Arslan, H. Flexible Numerology in 5G NR: Interference Quantification and Proper Selection Depending on the Scenario. *Mob. Inf. Syst.* **2021**, *2021*, 1–9. [\[CrossRef\]](#)
- Hassan, A.Y. A Frequency-Diversity System with Diversity Encoder and OFDM Modulation. *IEEE Access* **2021**, *9*, 2805–2818. [\[CrossRef\]](#)
- Soares, J.A. Complex Phase-Transmittance RBF Neural Network for Massive MIMO-OFDM Decoding. Master's Thesis, Department of Communications, School of Electrical and Computer Engineering, University of Campinas, Campinas, Brasil, 2021.
- Ko, K.; Byun, I.; Lee, J.; Shin, W. Joint power allocation and scheduling techniques for BER minimization in multiuser MIMO systems. *IEEE Access* **2021**, *9*, 66675–66686. [\[CrossRef\]](#)
- Dilli, R. Performance analysis of multi user massive MIMO hybrid beamforming systems at millimeter wave frequency bands. *Wirel. Netw.* **2021**, *27*, 1925–1939. [\[CrossRef\]](#)
- Jamali, V.; Tulino, A.M.; Fischer, G.; Müller, R.R.; Schober, R. Intelligent Surface-Aided Transmitter Architectures for Millimeter-Wave Ultra Massive MIMO Systems. *IEEE Open J. Commun. Soc.* **2021**, *2*, 144–167. [\[CrossRef\]](#)
- He, B.; Su, H.; Huang, J. Joint beamforming and power allocation between a multistatic MIMO radar network and multiple targets using game theoretic analysis. *Digit. Signal Process.* **2021**, *115*, 1–13. [\[CrossRef\]](#)
- Li, F.; Jiang, Y.; Du, C.; Wang, X. Construction of Golay complementary matrices and its applications to MIMO omnidirectional transmission. *IEEE Trans. Signal Process.* **2021**, *69*, 2100–2113. [\[CrossRef\]](#)
- Han, M.; Du, J.; Zhang, Y.; Li, X.; Rabie, K.M.; Nauryzbayev, G. Efficient Hybrid Beamforming Design in mmWave Massive MU-MIMO DF Relay Systems With the Mixed-Structure. *IEEE Access* **2021**, *9*, 66141–66153. [\[CrossRef\]](#)
- Sokal, B.; Gomes, P.R.B.; Almeida, A.L.F.d.; Haardt, M. Tensor-Based Receiver for Joint Channel, Data, and Phase-Noise Estimation in MIMO-OFDM Systems. *IEEE J. Sel. Top. Signal Process.* **2021**, *15*, 803–815. [\[CrossRef\]](#)
- Elnakeeb, A.; Mitra, U. Bilinear Channel Estimation for MIMO OFDM: Lower Bounds and Training Sequence Optimization. *IEEE Trans. Signal Process.* **2021**, *69*, 1317–1331. [\[CrossRef\]](#)

20. Majumder, M.; Saxena, H.; Srivastava, S.; Jagannatham, A.K. Optimal Bit Allocation-Based Hybrid Precoder-Combiner Design Techniques for mmWave MIMO-OFDM Systems. *IEEE Access* **2021**, *9*, 54109–54125. [\[CrossRef\]](#)
21. Yerrapragada, A.K.; Kelley, B. On the Application of K-User MIMO for 6G Enhanced Mobile Broadband. *Sensors* **2020**, *20*, 6252. [\[CrossRef\]](#) [\[PubMed\]](#)
22. Guerreiro, J.A.; Dinis, R.; Campos, L. On the Achievable Capacity of MIMO-OFDM Systems in the CathLab Environment. *Sensors* **2020**, *20*, 938. [\[CrossRef\]](#)
23. Saad, W.; Bennis, M.; Chen, M. A Vision of 6G Wireless Systems: Applications, Trends, Technologies, and Open Research Problems. *IEEE Netw.* **2020**, *34*, 134–142. [\[CrossRef\]](#)
24. Mayer, K.S.; Moreira, V.R.; Soares, J.A.; Arantes, D.S. High data-rates and high-order DP-QAM optical links can be efficiently implemented with concurrent equalization. In Proceedings of the 22nd Photonics North (PN), Niagara Falls, ON, Canada, 26–28 May 2020; p. 1. [\[CrossRef\]](#)
25. Mayer, K.S.; Müller, C.; De Castro, F.C.C.; De Castro, M.C.F. A new CPFSK demodulation approach for software defined radio. *J. Circuits, Syst. Comput.* **2019**, *28*, 1–14. [\[CrossRef\]](#)
26. Osman, A.M.; Rahman, S.M.A.; Borny, Z.N.; Baki, A.K.M. A Modified Method of Filtering for FBMC Based 5G Communications on Minimizing Doppler Shift. In Proceedings of the 2021 6th International Conference for Convergence in Technology (I2CT), Mumbai, India, 2–4 April 2021; pp. 1–4. [\[CrossRef\]](#)
27. Mayer, K.S.; De Oliveira, M.S.; Müller, C.; De Castro, F.C.; De Castro, M.C. Blind Fuzzy Adaptation Step Control for a Concurrent Neural Network Equalizer. *Wirel. Commun. Mob. Comput.* **2019**, *2019*, 1–11. [\[CrossRef\]](#)
28. Zou, F.; Liu, Z.; Hu, X.; Wang, G. A novel PAPR reduction scheme for OFDM systems based on neural networks. *Wirel. Commun. Mob. Comput.* **2021**, *2021*, 1–8. [\[CrossRef\]](#)
29. Mayer, K.S.; Müller, C.; de Castro, M.C.F.; de Castro, F.C.C. Nonlinear Modified Concurrent Equalizer. *J. Commun. Inf. Syst.* **2019**, *34*, 201–205. [\[CrossRef\]](#)
30. De Sousa, T.F.B.; Fernandes, M.A.C. Butterfly Neural Equalizer Applied to Optical Communication Systems with Two-Dimensional Digital Modulation. *Opt. Express* **2018**, *26*, 30837–30850. [\[CrossRef\]](#) [\[PubMed\]](#)
31. De Sousa, T.F.B.; Arantes, D.S.; Fernandes, M.A.C. Adaptive Beamforming Applied to OFDM Systems. *Sensors* **2018**, *18*, 3558. [\[CrossRef\]](#) [\[PubMed\]](#)
32. De Sousa, T.F.B.; Fernandes, M.A.C. Butterfly Neural Filter Applied to Beamforming. *IEEE Access* **2019**, *7*, 96455–96469. [\[CrossRef\]](#)
33. Dong, Z.; Huang, H. A training algorithm with selectable search direction for complex-valued feedforward neural networks. *Neural Netw.* **2021**, *137*, 75–84. [\[CrossRef\]](#) [\[PubMed\]](#)
34. Enriconi, M.P.; de Castro, F.C.C.; Müller, C.; de Castro, M.C.F. Phase Transmittance RBF Neural Network Beamforming for Static and Dynamic Channels. *IEEE Antennas Wirel. Propag. Lett.* **2020**, *19*, 243–247. [\[CrossRef\]](#)
35. Mayer, K.S.; Soares, J.A.; Arantes, D.S. Complex MIMO RBF Neural Networks for Transmitter Beamforming over Nonlinear Channels. *Sensors* **2020**, *20*, 378. [\[CrossRef\]](#)
36. Mayer, K.S.; Soares, J.A.; Pinto, R.P.; Rothenberg, C.E.; Arantes, D.S.; Mello, D.A.A. Soft failure localization using machine learning with SDN-based network-wide telemetry. In Proceedings of the 46th European Conference on Optical Communication (ECOC 2020), Brussels, Belgium, 6–10 December 2020; pp. 1–4. [\[CrossRef\]](#)
37. Mayer, K.S.; Soares, J.A.; Pinto, R.P.; Rothenberg, C.E.; Arantes, D.S.; Mello, D.A.A. Machine-learning-based soft-failure localization with partial software-defined networking telemetry. *J. Opt. Commun. Netw.* **2021**, *13*, E122–E131. [\[CrossRef\]](#)
38. Shimizu, D.Y.; Mayer, K.S.; Soares, J.A.; Arantes, D.S. A deep neural network model for link failure identification in multi-path ROADM based networks. In Proceedings of the 22nd Photonics North (PN), Niagara Falls, ON, Canada, 26–28 May 2020; p. 1. [\[CrossRef\]](#)
39. Swain, R.R.; Khilar, P.M.; Dash, T. Multifault diagnosis in WSN using a hybrid metaheuristic trained neural network. *Digit. Commun. Netw.* **2020**, *6*, 86–100. [\[CrossRef\]](#)
40. Pinto, R.P.; Mayer, K.S.; Soares, J.A.; Arantes, D.S.; Mello, D.A.A.; Cavalcante, V.; Santos, L.L.; Moraes, F.D.; Rothenberg, C.E. Demonstration of machine-intelligent soft-failure localization using SDN telemetry. In Proceedings of the Optical Fiber Communications Conference and Exhibition (OFC 2021), San Francisco, CA, USA, 6–10 June 2021; pp. 1–3.
41. Loss, D.; De Castro, M.C.F.; Franco, P.R.G.; DeCastro, F.C.C. Phase Transmittance RBF Neural Networks. *Electron. Lett.* **2007**, *43*, 882–884. [\[CrossRef\]](#)
42. Durga, R.V.; McLauchlin, A. The proposed novel sphere decoder for MIMO detection. In Proceedings of the 2021 International Conference on Artificial Intelligence and Smart Systems (ICAIS), Tamilnadu, India, 25–27 March 2021; pp. 1240–1245. [\[CrossRef\]](#)
43. Yao, G.; Chen, H.; Hu, J. An improved expectation propagation based detection scheme for MIMO systems. *IEEE Trans. Commun.* **2021**, *69*, 2163–2175. [\[CrossRef\]](#)
44. Jankiraman, M. *Space-Time Codes and MIMO Systems*, 1st ed.; Artech House: Norwood, MA, USA, 2004.
45. Zhao, X.; Jia, P.; Zhang, Q.; Li, Y.; Xing, R.; Liu, Y. Analysis of a distributed MIMO channel capacity under a special scenario. *Eurasip J. Wirel. Commun. Netw.* **2019**, *2019*, 1–7. [\[CrossRef\]](#)
46. Zhang, W.; Jiang, S. Effect of node mobility on MU-MIMO transmissions in mobile Ad Hoc networks. *Wirel. Commun. Mob. Comput.* **2021**, *2021*, 1–9. [\[CrossRef\]](#)
47. Chopra, S.R.; Gupta, A. Error analysis of grouped multilevel space-time trellis coding with the combined application of massive MIMO and cognitive radio. *Wirel. Pers. Commun.* **2021**, *117*, 461–482. [\[CrossRef\]](#)

48. Toka, M.; Kucur, O. Performance of MRT/RAS MIMO-NOMA with residual hardware impairments. *IEEE Wirel. Commun. Lett.* **2021**, *10*, 1071–1074. [\[CrossRef\]](#)
49. Shr, K.T.; Chen, H.D.; Huang, Y.H. A low-complexity Viterbi decoder for space-time trellis codes. *IEEE Trans. Circuits Syst. I Regul. Pap.* **2010**, *57*, 873–885. [\[CrossRef\]](#)
50. Xu, H.; Pillay, N. Multiple Complex Symbol Golden Space-Time Labeling Diversity. *IEEE Access* **2021**, *9*, 70233–70241. [\[CrossRef\]](#)
51. Gong, Y.; Letaief, K.B. Performance evaluation and analysis of space-time coding in unequalized multipath fading links. *IEEE Trans. Commun.* **2000**, *48*, 1778–1782. [\[CrossRef\]](#)
52. Li, Y. Optimum training sequences for OFDM systems with multiple transmit antennas. In Proceedings of the Global Telecommunications Conference. Conference Record (Cat. No.00CH37137), San Francisco, CA, USA, 27 November–1 December 2000; Volume 3, pp. 1478–1482. [\[CrossRef\]](#)
53. Biguesh, M.; Gershman, A.B. Training-based MIMO channel estimation: A study of estimator tradeoffs and optimal training signals. *IEEE Trans. Signal Process.* **2006**, *54*, 884–893. [\[CrossRef\]](#)
54. Tarokh, V.; Jafarkhani, H.; Calderbank, A.R. Space-time block codes from orthogonal designs. *IEEE Trans. Inf. Theory* **1999**, *45*, 1456–1467. [\[CrossRef\]](#)
55. Hu, Z.; Zhao, H.; Xue, J. Error exponent for Nakagami-m fading massive MIMO channels. In Proceedings of the 2020 IEEE 6th International Conference on Computer and Communications (ICCC), Chengdu, China, 11–14 December 2020; pp. 59–63. [\[CrossRef\]](#)
56. Jafarkhani, H. A quasi-orthogonal space-time block code. *IEEE Trans. Commun.* **2001**, *49*, 1–4. [\[CrossRef\]](#)
57. Tirkkonen, O.; Boariu, A.; Hottinen, A. Minimal non-orthogonality rate 1 space-time block code for 3+ Tx antennas. In Proceedings of the 2000 IEEE Sixth International Symposium on Spread Spectrum Techniques and Applications. ISSTA 2000. Proceedings (Cat. No.00TH8536), Parsippany, NJ, USA, 6–8 September 2000; Volume 2, pp. 429–432. [\[CrossRef\]](#)
58. Weifeng, S.; Xia, X.G. Quasi-orthogonal space-time block codes with full diversity. In Proceedings of the Global Telecommunications Conference, Taipei, Taiwan, 17–21 November 2002; Volume 2, pp. 1098–1102. [\[CrossRef\]](#)
59. Sindhu, P.; Hameed, A. Efficient quasi-orthogonal space-time block codes for five and six transmit antennas. In Proceedings of the 2015 IEEE International Conference on Electronics, Computing and Communication Technologies (CONECCT), Bangalore, India, 10–11 July 2015; pp. 1–5. [\[CrossRef\]](#)
60. Cho, Y.S.; Kim, J.; Yang, W.Y.; Kang, C.G. *MIMO-OFDM Wireless Communications with MATLAB*, 1st ed.; Wiley: New York, NY, USA, 2010.
61. Alamouti, S.M. A simple transmit diversity technique for wireless communications. *IEEE J. Sel. Areas Commun.* **1998**, *16*, 1451–1458. [\[CrossRef\]](#)
62. ETSI. TS 138 211—5G; NR; Physical Channels and Modulation (3GPP TS 38.211 Version 16.2.0 Release 16); ETSI: Sophia Antipolis, France, 2020.
63. Mei, K.; Liu, J.; Zhang, X.; Cao, K.; Rajatheva, N.; Wei, J. A low complexity learning-based channel estimation for OFDM systems with online training. *IEEE Trans. Commun.* **2021**, *69*, 6722–6733. [\[CrossRef\]](#)
64. ETSI. TR 138 901—5G; Study on Channel Model for Frequencies from 0.5 to 100 GHz (3GPP TR 38.901 Version 16.1.0 Release 16); ETSI: Sophia Antipolis, France, 2020.
65. Held, L.; Bové, D.S. *Applied Statistical Inference: Likelihood and Bayes*, 1st ed.; Springer: Berlin, Germany, 2014; p. 322.
66. Dolecek, G.V. *Random Signals and Processes Primer with MATLAB*, 1st ed.; Springer: New York, NY, USA, 2013; pp. 69–70.
67. Sambridge, M.; Gallagher, K.; Jackson, A.; Rickwood, P. Trans-dimensional inverse problems, model comparison and the evidence. *Geophys. J. Int.* **2006**, *167*, 528–542. [\[CrossRef\]](#)
68. Tenoudji, F.C. *Analog and Digital Signal Analysis: From Basics to Applications*, 1st ed.; Springer: Paris, France, 2016; pp. 387, 469, 471.
69. Diniz, P.S.R. *Adaptive Filtering: Algorithms and Practical Implementation*, 4th ed.; Springer: New York, NY, USA, 2013; pp. 83–84.
70. Mandic, D.P.; Goh, V.S.L. *Complex Valued Nonlinear Adaptive Filters: Noncircularity, Widely Linear and Neural Models*, 1st ed.; John Wiley & Sons, Inc.: Chippingham, UK, 2009; pp. 73, 77.