Application of a Machine Learning Model to the realization of a wireless communication system

Wieslaw Citko, Adam Trzebiatowski, Wieslaw Sienko

Abstract—This paper presents the author's proposal for a neural detector realization of a Massive-MIMO-OFDM system using extended Hopfield neural circuits. An important feature of such an implementation is that the system can be learned without the need to solve multi-parameter optimization tasks requiring high computational power.

Keywords—wireless machine communication systems: learning; Massive-MIMO-OFDM systems

I. INTRODUCTION

ACHINE learning algorithms, implemented in neural circuit structures, have revolutionized many areas of significant theoretical and practical interest. First and foremost, data processing systems such as image processing, speech recognition, automatic translation or autonomous systems should be mentioned here. A product of recent years are artificial intelligence (A. I.) systems. However, it can be noted that in such an important field as wireless communication, machine learning algorithms have not yet found satisfactory practical implementations. Traditional communication systems are based on statistical models describing transmission processes, wave propagation, noise levels and channel interference. Such models are complex and do not always sufficiently describe and adapt to dynamically developing communication systems (e.g. Internet of Things). As an alternative to statistical models, machine learning systems can be used, i.e. based on the availability of I/O data to form training sets [1].



Fig. 1. Basic structures, conventional and neural, of communication systems T- training set.

 $P_{v|s}$ - statistical model estimating the relation of the sets $\{y_i\}$ and $\{S_i\}$

Figure 1 illustrates the basic communication structures realized by conventional and machine learning technology. The main

Wieslaw Sienko is with the Department of Marine Telecommunications, Faculty of Electrical Engineering, Gdynia Maritime University, Poland (email: w.sienko@we.umg.edu.pl).

difference is in the means of detecting symbols transmitted over time-varying channels. Receivers in conventional systems require the implementation of symbol detection based on the knowledge of statistical relationships [2]. It should be emphasized that the effective realizability of a neural communication system requires solving two problems: - channel estimation,

- realization of the neural detector.

The realization of these tasks is based on the appropriate generation of training sets. The realization of neural detector learning is more difficult compared to, for example, neural image processing due to channel variability and time constraints for the use of stationarity of channel parameters (coherence interval).

II. CHANNEL ESTIMATION AND DETECTION MIMO-OFDM

OFDM systems are currently the dominant structure of broadband communication systems. Such systems are analyzed in this paper. The availability of channel state information (CSI) is critical to the realization of MIMO-OFDM communication systems. It is worth noting that the signal transmission structure of an OFDM system is given in the form of a table as in Fig. 2.



Fig. 2. OFDM system time-frequency tables • pilot vector element symbol

o symbol of the information signal vector element.

The data blocks are transmitted in parallel on M orthogonal subcarriers. The OFDM channel estimation using machine learning algorithms has been the subject of research for a decade. One of the most interesting proposals is to treat the array in Fig. 2 as an image defined by the distribution of symbols (pixels). A reasonable solution is therefore to use the structures of convolutional neural circuits (CNNs), which, according to the established opinions in the literature, are best suited for image processing. Training sets for learning CNNs

Adam Trzebiatowski is with the Department of Marine Telecommunications, Faculty of Electrical Engineering, Gdynia Maritime University, Poland (e-mail: a.trzebiatowski@we.umg.edu.pl).



Wieslaw Citko is with the Department of Marine Telecommunications, Faculty of Electrical Engineering, Gdynia Maritime University, Poland (email: w.citko@we.umg.edu.pl).

are obtained using 'images' containing pilot signal symbols [3]-[6]. The detection of a MIMO system under the assumptions of stationarity and flat frequency characteristics of the channel, is described in the baseband by the equation:

$$y = Hx + n \tag{1}$$

where: $\mathbf{x} \in C^{N_t}$ is a vector of transmitted symbols belonging to a finite constellation of dimension M,

 $y \in C^{N_r}$ is the received vector,

 $H \in C^{N_r \times N_t}$ is the composite channel matrix,

 $\mathbf{n} \in C^{N_r}$ is additive white noise with Gaussian distribution.

Assuming knowledge of the channel matrix H, the solution of (1) is given by the solution of the optimization problem:

$$\widehat{\boldsymbol{x}} = \arg\min_{\boldsymbol{x}} \|\boldsymbol{y} - \boldsymbol{H}\boldsymbol{x}\|^2 \tag{2}$$

where: \hat{x} – estimator of x.

Obtaining such a solution would require an analysis of all the x vectors transmitted in the channel. Such an analysis is practically unfeasible with large constellations and in large MIMO systems. Hence the attempt to find the solution of (2) by detectors based on neural circuit structures [7].

III. APPLICATION OF AN EXTENDED HOPFIELD NEURAL CIRCUIT MODEL IN A WIRELESS COMMUNICATION SYSTEM

One important neural structure are Hopfield-type systems, which are both physical models and algorithms used in neural computing. The earlier paper proposed an extended model of the systems defined by the following differential equation [8]:

$$\dot{\boldsymbol{x}} = (\eta \boldsymbol{W} - \boldsymbol{w}_0 \boldsymbol{1} + \varepsilon \boldsymbol{W}_s) \boldsymbol{\theta}(\boldsymbol{x}) + \boldsymbol{I}_d \tag{3}$$

where: W— skew symmetric orthogonal matrix of weights connections,

 W_s — real symmetric matrix,

1—unitary matrix,

 $\theta(x)$ —vector of activation functions,

 I_d — input vector,

 ε , w_0 , η — parameters.

Equation (3) at equilibrium of the system takes the form:

$$(\eta W - w_0 \mathbf{1} + \varepsilon W_s) \boldsymbol{\theta}(\mathbf{x}) + I_d = \mathbf{0}$$
(4)

Equation (4) provides the basis for universal machine learning models based on biorthogonal transformations that enable typical functions of learning systems. One of these functions is the implementation of associative memories. The application of the system to reconstruct and recognize distorted/noisy images using associative memory is described in the paper [8]. On the other hand, the implementation of a machine learning system for solving inverse problems (Inverse Problem) was investigated in the paper [9]. In the aforementioned works, the original image was processed by a linear matrix operator whose size was not quadratic. Thus, there was no inverse operator in the sense of matrix algebra. A suitably designed machine learning system performed the reconstruction of the original image based on its projection. This paper proposes the application of the above machine learning model to the implementation of a Massive-MIMO-OFDM communication system [10].

IV. STRUCTURE OF THE MACHINE LEARNING MODEL

The equilibrium (4) is the basis for creating the structures of universal machine learning models. It should be noted that assuming: W_s is a Hermitian matrix, a complex data (vector) processing model is obtained. Thus, the solution (4) transforms to the form:

$$(\boldsymbol{W} - 2\boldsymbol{1} + \boldsymbol{W}_{H})\boldsymbol{\theta} + \boldsymbol{I}_{d} = \boldsymbol{0}$$
(5)

where: W — skew symmetric orthogonal matrix,

 \boldsymbol{W}_{H} — Hermitian matrix $\boldsymbol{W}_{H} = \boldsymbol{W}_{H}^{+}$,

1 — unitary matrix,

 $\boldsymbol{\theta}$ — activation function vector, \boldsymbol{I}_d — input vector,

 $\varepsilon = 1, w_0 = 2, \eta = 1.$

The machine learning model realizes the mapping: $F: X \to Y$, of the I/O type, where the sets X and Y are realized by pairs of complex training vectors $\{x_i, y_i\}_{i=1}^N, x_i \in X \subset C^m, y_i \in Y \subset C^n$. The realizations of the mapping F can be obtained by transforming (5) to the form:

$$(\boldsymbol{W} - 2\boldsymbol{1} + \boldsymbol{W}_H)\boldsymbol{m}_i + \boldsymbol{u}_i = \boldsymbol{0}$$
(6)

where: $\boldsymbol{u}_i = \begin{bmatrix} \boldsymbol{x}_i \\ \boldsymbol{y}_i \end{bmatrix}, i = 1, ..., N; dim \boldsymbol{u}_i = m + n,$ $\boldsymbol{m}_i = \frac{1}{2}(W + 1)\boldsymbol{u}_i,$ $\boldsymbol{m}_i \in \boldsymbol{M} = [\boldsymbol{m}_1, \boldsymbol{m}_2, ..., \boldsymbol{m}_N],$ $\boldsymbol{W}_H = \boldsymbol{M}(\boldsymbol{M}^T \boldsymbol{M})^{-1} \boldsymbol{M}^T,$ \boldsymbol{M} — the spectral matrix of the vectors \boldsymbol{u}_i .

Hence, Eq.(6) has N - stable solutions, which are the centers of attraction of the mapping F. The block structure of the machine learning model is shown in Fig. 3.



Fig. 3. Block structure of the machine learning model

V. RECOGNITION AND RECONSTRUCTION OF A VECTOR (IMAGE) AS AN INVERSE PROBLEM

The paper [9] gives a method for reconstructing a vector (image) from the knowledge of its projection, i.e.:

$$Ax = \widetilde{y}; \ A(x) = \widetilde{y} \tag{7}$$

where: $A(\cdot)$ — linear operator,

- A complex rectangular matrix: dim $A = (m \times n), m > n$,
- x original image, $x \in C^n$,
- \tilde{y} available projection of the original $\tilde{y} \in C^m$.

The solution of (7) belongs to the solution of the inverse problem:

$$\boldsymbol{x} = A^{-1}(\boldsymbol{\tilde{y}}) \tag{8}$$

Most of the solutions to (8) known from the literature use optimization methods, for example:

$$\min_{\mathbf{x}} \|\widetilde{\mathbf{y}} - \mathbf{A}\mathbf{x}\|_{2}^{2}, \text{ s.t. } \mathbf{x} \in K$$

$$\min_{\mathbf{x}} \|\widetilde{\mathbf{y}} - \mathbf{A}\mathbf{x}\|_{2}^{2} + \beta R(\mathbf{x})$$
(9)

where: K — set of admissible solutions, $R(\mathbf{x})$ — regularize, β — regularization parameter.

Equation (7) was solved using the model in Fig.3, with the system vectors u_i in (6) being of the form:

$$\boldsymbol{u}_{i} = \begin{bmatrix} \boldsymbol{x}_{i} \\ \boldsymbol{y}_{i} \end{bmatrix}, i = 1, \dots, N; N = n$$
(10)

where:: $x_i \in C^n$, $y_i \in C^m$ are pairs of training random vectors:

$$Ax_i = y_i \tag{11}$$

The solution to the inverse problem is realized by the model shown in Fig. 4.



Fig. 4. Structure of the model implementing the inverse mapping

It should further be noted that the model performs the solution of the inverse problem by superposition of random training vectors. The detection of the MIMO system described by (1) can be interpreted as the solution of the inverse problem after the estimation of the *H*-channel matrix.

VI. MACHINE LEARNING MODEL FOR THE MASSIVE-MIMO-OFDM System

A. Uplink Channel

The Massive-MIMO-OFDM wireless communication system is considered to be the most scalable communication structure [10]. The architecture of such a system is shown in Fig. 5.



Fig. 5. Uplink of a Massive-MIMO system ($N_r > N_t$) (number of receiving BS antennas N_r greater than the number of transmitting terminal antennas N_t)

The system uses a TDD (time-division duplex) operation. The base station learns the channel (uplink) using pilots. The estimation of the downlink channel, on the other hand, is obtained by transposing the uplink channel. Using the structure realizing the $A^{-1}(\cdot)$ operator in Fig. 4, for the *H*-channel estimation and vector detection $\mathbf{x} \in C^{N_t}$, requires the transmission of N_t pilot signals. Thus, according to (6), the system vectors \mathbf{u}_i take the form:

$$\boldsymbol{u}_i = \begin{bmatrix} \boldsymbol{y}_i \\ \boldsymbol{x}_i \end{bmatrix}, i = 1, \dots, N_t; \dim \boldsymbol{y}_i = N_r, \dim \boldsymbol{x}_i = N_t \ (12)$$

where:

$$\boldsymbol{y}_i = \boldsymbol{H}\boldsymbol{x}_i; i = 1, \dots, N_t \tag{13}$$

 x_i — pilot vectors,

 y_i — vectors received at base station,

 $H(\cdot)$ — physical model of noisy channel (H is static in the coherence interval).

Statement:

The detection of any vector $\mathbf{x}_t \in C^{N_t}$ transmitted through the channel i.e.

$$\mathbf{y} = \mathbf{H}\mathbf{x}_t, \mathbf{x}_t \neq \mathbf{0} \tag{14}$$

hence:

$$\widehat{\boldsymbol{x}}_t = \boldsymbol{H}^{-1}\boldsymbol{y} \tag{15}$$

leads to an estimate of \hat{x}_t at the output of the neural detector with the MSE error:

$$\|\widehat{\boldsymbol{x}}_t - \boldsymbol{x}_t\|^2 \approx 0 \tag{16}$$

The channel estimation and detector model is shown in Fig. 6.



Fig. 6. Neural channel and detector model, \hat{x} vector estimation of x

The learning of the above model is described by Eq. (6). The limitation of such a model is the number: $N_t + N_r = 2^k$, $k = 3, 4, ..., N_r > N_t$ (massive model)

B. Downlink Channel

Assuming the operation of the communication system using TDD (Time Division Duplex), the uplink channel estimation can be used for downlink estimation due to the reversibility of the channel. From the considerations carried out in the previous section, the base station has an uplink channel model as shown in Fig. 6. In particular, the X matrix of the pilot vectors and the Y matrix of their transmission through the channel are fixed. Thus, the numerical value of the channel matrix \hat{H} can be determined from the following relationship:

$$\widehat{\boldsymbol{H}} = \boldsymbol{X} (\boldsymbol{X}^T \boldsymbol{X} + \gamma \boldsymbol{1})^{-1} \boldsymbol{X}^T$$
(17)

where: $\gamma \mathbf{1}$ — regularization component ($\gamma > 0$)

It can be shown that the evaluation occurs:

$$\left\|\widehat{\boldsymbol{H}} - \boldsymbol{H}\right\|^2 \approx 0 \tag{18}$$

where: *H* — test matrix (random matrix)

The Massive-MIMO model for downlink transmission is shown in Fig. 7.



Fig. 7. Downlink of a Massive-MIMO system ($N_r > N_t$) (number of receiving BS antennas N_r greater than the number of transmitting terminal antennas N_t)

In conventional down-link transmission of the symbol vectors s (Fig. 7), the base station uses H-channel estimators. It uses H^T to linear precoding the symbols and transmit them to all terminals. Each terminal should have a CSI for coherent detection of the transmitted symbols. The literature proposes beamforming of pilots generated at the base stations for path estimation to each terminal. In the authors' further research, it is assumed that an attempt will be made to solve the down-link transmission problem using the neural model of Fig. 6. using beamforming or spatial filtering for each terminal.

VII. COMPUTATIONAL EXAMPLE OF NEURAL UPLINK DETECTOR

The computational example assumes a Massive-MIMO communication system consisting of 12 single-antenna terminals and a base station containing 20 receive antennas. A model of such a system is illustrated in Fig. 8.



Fig. 8. Massive-MIMO system model analyzed in the example x_i – pilots vectors y_i - vectors received at the base station

According to (13), the matrix operator $H(\cdot)$ modeling the physical channel was assumed as a random matrix $H(20 \times$

physical channel was assumed as a random matrix
$$H(20 \times 12)$$
 with coefficients taking complex values with a normal distribution. If x_1 is a fixed vector from the set $\{x_i\}_{i=1}^{12}$, according to the relation:

$$\boldsymbol{y}_1 = \boldsymbol{H}\boldsymbol{x}_1 \tag{19}$$

it corresponds to the received after the transmission vector y_1 . The selected exemplary pair of vectors x_1 and y_1 take the form: W. CITKO, A. TRZEBIATOWSKI, W. SIENKO

$$\boldsymbol{x}_{1}^{T} = \begin{bmatrix} 1-1i, & -1-3i, & -1+3i, & 1+3i, & 3-1i, & 1+1i, \\ 1-3i, & -3+1i, & -1+1i, & -1-1i, & -3-1i, & 3-1i \end{bmatrix}$$

$$\boldsymbol{y}_{1}^{T} = \begin{bmatrix} -16.5-0.5i, & -9.5+10.2i, & -1.4-9.1i, & -10.2-21.9i, \\ 5.5-30.9i, & -12.4+0.9i, & -19.5+0.7i, & 2.5+9.2i, \\ -6.6-16.3i, & -5.9-8.7i, & -2.3-0.5i, & 21.8+1.2i, \\ 9.2-13.6i, & 1.0-1.1i, & -4.4+10.1i, & 10.1-1.1i, \\ 3.3-10.9i, & -3.5+7.8i, & -9.3-4.9i, & -2.0+1.5i \end{bmatrix}$$

where: x_1 – transmitted vector, y_1 – vector received after transmission in the channel.



Fig. 9. Constellation of vectors

a) x_1 – transmitted vector

b) y_1 – received vector after transmission in the channel

All 12 pilot vectors were generated as permutations of the x_1 vector. According to (10), the system vectors u_i take the form:

$$\boldsymbol{u}_i = \begin{bmatrix} \boldsymbol{x}_i \\ \boldsymbol{y}_i \end{bmatrix}$$
, $i = 1, \dots, 12$; dim $\boldsymbol{y}_i = 20$, dim $\boldsymbol{x}_i = 12$

The neural detector implementation for the above data is presented in Fig. 6. The detection of any vector $x_t \in C^{12}$ transmitted through the channel, i.e.

$$\mathbf{y} = H\mathbf{x}_t, \mathbf{x}_t \neq \mathbf{0} \tag{20}$$

hence:

$$\hat{\boldsymbol{x}}_t = \boldsymbol{H}^{-1} \boldsymbol{y} \tag{21}$$

leads to an estimate of \hat{x}_t at the output of the neural detector system shown in Fig. 6. with the MSE error:

$$\|\widehat{\boldsymbol{x}}_t - \boldsymbol{x}_t\|^2 \approx 0 \tag{22}$$

It is interesting to note that for non-static channel (due to the noisy component) the solution of Eq. (19) can be illustrated by the constellations as shown in Fig. 10 where $\hat{x}_t = H^{-1}(y + n)$; n - noise vector.



Fig. 10. The constellations of \hat{x}_t obtained at different SNR levels

The analyses, the results of which are shown in Fig. 10, were carried out to investigate the properties of the neural decoder in the presence of Gaussian noise in the communication channel.

Simulations were performed for different signal-to-noise ratios ranging from 60dB to 20dB. In each case, 10,000 trials were carried out. It can be concluded that the decoder shows some

resistance to channel interference. By analyzing the individual results, it can be concluded that correct symbol detection is possible when the signal-to-noise ratio in the communication channel is not less than 20dB.

A full assessment of the suitability of the communication system using the neural detector described in this paper requires further research.

VIII. CONCLUSION

According to the available literature, the application of machine learning to the realization of wireless communication systems is in the early stages of technological development. Research in this area, indicates the potential to realize the technology that is competitive to the conventional technology. This paper focuses on describing the implementation rationale for the Massive-MIMO-OFDM system. The literature indicates that the research has been conducted at a number of research centers using a variety of neural circuit architectures (e.g. CNNs). The final choice of learning methods and architecture does not appear to be fixed. This paper presents the author's contributions to the neural detector realization of the Massive-MIMO-OFDM system using extended Hopfield neural circuits. An important feature of such implementations is learning without the need to solve multi-parameter optimization tasks that require high computational power. The neural detector model presented in this paper requires further simulation studies.

Appendix

Algorithm of Machine Learning Model Design [8]

1. Declaration:

Input the set of training points:

 $S = \{x_i, y_i\}, i = 1, 2, ..., N,$

$$x_i \in C^n, y_i \in C^m, n + m = 2^k, k = 3, 4, ...$$

2. System design:

Create system vectors \boldsymbol{u}_i :

$$\boldsymbol{u}_i = \begin{bmatrix} \boldsymbol{x}_i \\ \boldsymbol{y}_i \end{bmatrix}$$
, dim $\boldsymbol{u}_i = n + m$.

Calculate the spectrum \boldsymbol{m}_i of system vectors \boldsymbol{u}_i :

$$\boldsymbol{m}_i = \frac{1}{2} \big(\boldsymbol{W}_{2^k} + \mathbf{1} \big) \boldsymbol{u}_i$$

Create spectrum matrix **M**:

 $\boldsymbol{M} = [\boldsymbol{m}_1, \boldsymbol{m}_2, \dots, \boldsymbol{m}_N]$ Calculate Hermitian matrix \boldsymbol{W}_H :

$$\boldsymbol{W}_{H} = \boldsymbol{M}(\boldsymbol{M}^{T}\boldsymbol{M})^{-1}\boldsymbol{M}^{T}$$

Calculate orthogonal transformation $T(\cdot)$: $T(\cdot) \equiv T = \frac{1}{2} (W_{2^k} + 1)$ Calculate biorthogonal transformation $T_s(\cdot)$: $T_s(\cdot) \equiv T_s = (2 \cdot 1 - W_H - W_{2^k})^{-1}$.

for
$$l = 1:N$$

 $\widetilde{\mathbf{x}}_{i}^{(0)} = \mathbf{0}$
while $\|\widetilde{\mathbf{x}}_{i}^{(l)} - \widetilde{\mathbf{x}}_{i}^{(l-1)}\| \ge eps$
 $\begin{bmatrix} \widetilde{\mathbf{x}}_{i} \\ \mathbf{y}_{i} \end{bmatrix}^{(l)} = \mathbf{T}^{-1}\mathbf{T}_{s}\left(\begin{bmatrix} \mathbf{0} \\ \mathbf{y}_{i} \end{bmatrix} + \begin{bmatrix} \widetilde{\mathbf{x}}_{i} \end{bmatrix}^{(l-1)} \right)$
end

end

 $(l = 1, 2, \dots$ steps of recurrence)

Final results of recurrence: $\tilde{x}_i = x_i$.

REFERENCES

- Y. C. Eldar, A. Goldsmith, D. Gunduz, H. V. Poor H (Eds.), *Machine Learning and Wireless Communication*, Cambridge University Press, 2022.
- [2] N. Shlezinger, N. Farsad, Y. C. Eldar, A. Goldsmith, "Model-Based Machine Learning for Communications in *Machine Learning and Wireless Communication*," Cambridge University Press, 2022.
- [3] H. He, H. Ye, S. Jin, G. Y. Li, "Channel Estimation, Feedback, and Signal Detection," in *Machine Learning and Wireless Communication*, Cambridge University Press, 2022.
- [4] S. Coleri, M. Ergen, A. Puri, A. Bahai, "Channel Estimation Techniques Based on Pilot Arrangement in OFDM Systems," *IEEE Transactions on Broadcasting*, vol. 48, no. 3, pp. 223-229, Sept. 2002, https://doi.org/10.1109/TBC.2002.804034
- [5] M. Soltani, V. Pourahmadi, A. Mirzaei, H. Sheikhzadeh, "Deep Learning-Based Channel Estimation," *IEEE Communications Letters*, vol. 23, no. 4, pp. 652-655, April 2019, https://doi.org/10.1109/LCOMM.2019.2898944
- [6] C. Dong et al., "Image Super-Resolution Using Deep Convolutional Networks," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 38 (2014): 295-307.
- [7] O. Sholev, H. H. Permuter, E. Ben-Dror, W. Liang, "Neural Network MIMO Detection for Coded Wireless Communication with Impairments," 2020 IEEE Wireless Communications and Networking Conference (WCNC), Seoul, Korea (South), 2020, pp. 1-8, https://doi.org/10.1109/WCNC45663.2020.9120517
- [8] W. Citko, W. Sienko, "Inpainted Image Reconstruction Using an Extended Hopfield Neural Network Based Machine Learning System," *Sensors*, 22 (2022), no. 3, https://doi.org/10.3390/s22030813
- [9] W. Citko, W. Sienko, "Image Recognition and Reconstruction With Machine Learning: An Inverse Problem Approach," in *IEEE Access*, vol. 11, pp. 107463-107471, 2023, https://doi.org/10.1109/ACCESS.2023.3315831
- [10]T. Marzetta, E. Larsson, H. Yang, H. Hgo, *Fundamentals of Massive MIMO*, Cambridge University Press, 2016.