

# Supervised Machine Learning Classifiers for Diversity Combined Signals in 6G Massive MIMO Receivers

J. S. Daba<sup>1,\*</sup>, O. M. Abdul-Latif<sup>2</sup>

<sup>1</sup>Department of Electrical Engineering, University of Balamand, Lebanon

<sup>2</sup>Department of Electrical Engineering, Rochester Institute of Technology, Dubai, United Arab Emirates

Received July 24, 2020; Revised September 2, 2020; Accepted September 29, 2020

## Cite This Paper in the following Citation Styles

(a): [1] J. S. Daba, O. M. Abdul-Latif, "Supervised Machine Learning Classifiers for Diversity Combined Signals in MIMO Receivers," *Universal Journal of Electrical and Electronic Engineering*, Vol. 7, No. 6, pp. 320 - 327, 2020. DOI: 10.13189/ujeee.2020.070604.

(b): J. S. Daba, O. M. Abdul-Latif (2020). *Supervised Machine Learning Classifiers for Diversity Combined Signals in MIMO Receivers*. *Universal Journal of Electrical and Electronic Engineering*, 7(6), 320 - 327. DOI: 10.13189/ujeee.2020.070604.

Copyright©2020 by authors, all rights reserved. Authors agree that this article remains permanently open access under the terms of the Creative Commons Attribution License 4.0 International License

**Abstract** Support Vector Machine (SVM) is a statistical learning tool that was initially developed by Vapnik in 1979 and later developed to a more complex concept of structural risk minimization (SRM). SVM, as a supervised machine learning tool, is playing an increasing role in applications to detection problems in various engineering problems, notably in statistical signal processing, pattern recognition, image analysis, and 6G wireless communication networks. In this paper, SVM is applied to signal detection in 6G communication systems in the presence of channel noise in the form of fully developed Rayleigh multipath fading and receiver noise generalized as additive color Gaussian noise (ACGN). The structure and performance of SVM in terms of the bit error rate (BER) metric is derived and simulated for these advanced stochastic noise models and the computational complexity of the implementation, in terms of average computational time per bit, is also presented. The performance of SVM is then compared to conventional  $M$ -ary signaling optimal model-based detector driven by  $M$ -ary phase shift keying (MPSK) modulation. We show that the SVM performance is superior to that of conventional detectors which require as much as 7 bits-coding ( $M \geq 128$ ) to produce comparable results to those of SVM. Finally, the SVM-based detector is implemented in an uplink SIMO system using both Equal Gain Combiner (EGC) technique and Root Mean Square Gain Combiner (RMSGC) technique in which the later technique will be proven to be superior to the earlier.

**Keywords** Least Square-Support Vector Machine,

Massive MIMO,  $M$ -ary Phase Shift Keying, Orthogonal Frequency Division Multiplexing, Root Mean Square Gain Combining, Single Input Multiple Output, 6G Networks

## 1. Introduction

Support Vector Machine (SVM) is a recent class of statistical classification and regression techniques getting an increased attention on its application to classification problems in various engineering areas. SVM is based on the statistical learning theory initially developed by Vapnik [1] in 1979 and later developed to a more complex concept of structural risk minimization (SRM). SVM is formulated on the structural risk minimization (SRM) principle which minimizes an upper bound on the generalization error, as opposed to the classical empirical risk minimization (ERM) approach which minimizes the error on the training data and is embodied in statistical learning.

In a broad sense, two classes of classifiers are widely used in the literature: (1) *model-based* classifiers such as the maximum likelihood (ML) and maximum a posteriori (MAP) detectors and (2) *boundary-based* classifiers such as support vector machine, neural networks and fuzzy logic.

SVM claims to guarantee generalization, i.e., the decision rules reflect the regularities of the training data rather than the incapability of the learning machine.

SVM has been widely used in solving classification and function estimation problems due to its many attractive

features and promising empirical performance with many successful applications in synthetic aperture radar image classification [2] and pattern recognition. Recently, SVM has been introduced to digital communication systems as a new method for *channel equalization* [3] – [5] and has proved to be very effective in overcoming intersymbol interference (ISI) and co-channel interference (CCI). SVM was also applied for the equalization of burst time division multiple access (TDMA) transmission [6]. To the best of our knowledge, SVM has not been implemented yet for *receiver detection* in digital communication systems in the presence of advanced additive colour receiver noise and multiplicative channel fading noise. Notable exceptions are the initial work of Dubois and Abdel-Latif [7] who applied SVM to OOK-infrared channels in a local fading environment with partially developed multipath fading and additive white Gaussian receiver noise (AWGN), and the work of Mokbel and Hashem [8] who applied SVM to a BNRZ detector (sampler and comparator) using multiple samples per binary period in the presence of AWGN in wire-line communication systems.

## 2. Wireless Communication System

### A. Modulation Scheme

When making a decision on the choice of model-based or boundary-based classifiers, we must take into consideration the driving modulation scheme.  $M$ -ary phase shift keying (MPSK) is increasingly being adopted as a physical layer modulation technique for cellular systems, most notably, 3G UMTS (3<sup>rd</sup> Generation Universal Mobile Telecommunications System) which uses QPSK (quadrature phase shift keying), and 2.5G EDGE (Enhanced Data Rates for Global Evolution) which uses 8-PSK for high data rates (replacing the classical GMSK (Gaussian minimum shift keying)) [9].

MPSK has a higher bit rate than BPSK, yet it needs a more complex receiver, its carrier recovery process is much more complicated, and most importantly, it has a worse bit error rate (BER) performance (but better symbol error rate). Since SVM is essentially a binary classifier, it is only logical to apply SVM to MPSK to improve its BER performance.

### B. Stochastic Noise Model

The Rayleigh fading channel is widely assumed in the literature for wireless systems, especially for mega- and macro-cells in the absence of dominant line of sight [9]. The fading envelope obeys the scattering stochastic model

$$\gamma_L = \left| \sum_{k=1}^L A_k e^{j\phi_k} \right|, \quad (1)$$

with fading power being  $\nu_L = \gamma_L^2$ .  $A_k$  is the random amplitude of the  $k^{\text{th}}$  scatterer,  $\phi_k$  is the random phase of the  $k^{\text{th}}$  scatterer assumed to be uniformly distributed

between  $[0, 2\pi)$ , and  $L$  is the random number of scatterers in the channel. When  $L$  is sufficiently large so that the central limit theorem (CLT) holds and the scattered field is approximately a circular complex Gaussian random variable, we say that the fading is *fully developed*. Even for statistically dependent scatterers, the scattered field would still be asymptotically circular Gaussian hinging on the fact that the sequence of scattering random variables satisfies the  $\alpha$ -mixing property or the conditions of the Lindeberg-Feller CLT. A simple random variable transformation results in the fading envelope  $\gamma_L$  asymptotically Rayleigh distributed (with parameter  $P_{\text{dif}} = E(\gamma_L^2)$ ), and the fading power  $\nu_L$  exponentially distributed (central chi-square with 2 degrees of freedom).

In general, the received wireless QPSK signal is corrupted by two types of noises: (1) channel fading *multiplicative* noise  $\gamma_L$  and (2) receiver *additive* colour Gaussian noise (ACGN) which characterizes more complicated channel and receiver interference. This stochastic model is widely used in wireless optical communication systems [10] in the presence of interfering ambient or incandescent/fluorescent light and non-ideal photo-detectors. Technically, these noise models respectively correspond to shot noise and “random telegraph signal” or laser-phase noise. Most commonly, ambient light induces shot noise in the photo-detector of optical receivers.

By studying colour noise, we are assuming a severe noise environment. Without loss of generality, we will consider the autocorrelation function of a random telegraph signal

$$R_N(\tau) = \varepsilon^2 \exp(-2\lambda|\tau|), \quad (2)$$

where  $2\varepsilon$  is the peak-to-peak noise amplitude and  $\lambda$  is the rate of the influencing underlying Poisson point process [10]. The power spectral density (PSD) of the noise is given by

$$P_N(f) = \frac{\varepsilon^2 \lambda}{1 + \pi^2 f^2}, \quad (3)$$

which has a de-emphasis filter-like characteristic.

## 3. Support Vector Machine

In this section, we provide a succinct introduction to the SVM approach. The reader is referred to the initial work of Vapnik [1] and the book of Christianini [11] for more in-depth treatment of the SVM theory.

The relation between the capacity of a learning machine and its performance is ruled by a set of boundaries, which is referred to as the bound on the generalization performance. Statistical pattern recognition techniques face two problems: the identification problem and the parameters estimation problem. The identification problem is the problem of determination of the degree of freedom or complexity of the model and is generally the more complex

problem [12]. The estimation problem is how to get an optimal estimate of the model parameters regarding the training data set.

Let us consider a mapping  $\Phi: \mathcal{X}^d \mapsto H$ , which maps the training data from  $\mathcal{X}^d$  to a higher Euclidean space  $H$ , that may have an infinite dimension. In this high dimension space, the data are linearly separable, hence linear SVM formulation above can be applied for any type of data [2]. In the SVM formulations, the training data only appear in the form of dot products  $\mathbf{x} \cdot \mathbf{x}$ . These can be replaced by dot products in the Euclidean space  $H$ , i.e.,  $\varphi(\cdot) \cdot \varphi(\cdot)$ .

The dot product in the high dimension space can also be replaced by a kernel function. By computing the dot product directly using a kernel function, one avoids the mapping  $\Phi(x)$ . This is desirable because  $H$  has possibly infinite dimensions and  $\Phi(x)$  can be tricky or impossible to compute. Using a kernel function, a SVM that operates in infinite dimensional space can be constructed [11].

Given a training set of  $N$  data points  $\{y_k, x_k\}^N$ , where  $x_k$  denotes the  $k^{\text{th}}$  input pattern and  $y_k$  the  $k^{\text{th}}$  output pattern, the SVM aims at constructing a decision function or classifier

$$\begin{aligned} f(x) &= \text{sign} \left[ \mathbf{w}^T \boldsymbol{\varphi}(x) + b \right] \\ &= \text{sign} \left[ \sum_{k=1}^N \alpha_k y_k K(x, x_k) + b \right], \end{aligned} \quad (4)$$

where  $\mathbf{w}$  is the weight vector in the reproducing kernel Hilbert space (RKHS),  $\alpha_k$  are support values (Lagrangian multipliers),  $b$  is the bias term, and the kernel function

$$K(x, x_k) = \varphi(x) \varphi(x_k). \quad (5)$$

For every new test data, the kernel functions for each SV (support vector) need to be recomputed.

For any kernel function suitable for SVM, there must exist at least one pair of  $\{H, \Phi\}$ , such that (5) is satisfied. The kernel that has these properties is said to obey the Mercer's condition, i.e., for any  $g(x)$  with finite  $L_2$  norm,

$$\int g^2(x) dx < \infty, \quad (6)$$

$$\iint K(x, y) g(x) g(y) dx dy \geq 0 \quad (7)$$

By choosing different kernel functions, the SVM can emulate some well known classifiers [13], as shown in Table 1.

**Table 1.** kernel functions' classifiers

Kernel Function	Type of Classifier
$K(x, y) = xy$	Linear
$K(x, y) = \exp\left(-\frac{\ x - y\ _2^2}{\sigma^2}\right)$	Gaussian radial bias function (RBF)
$K(x, y) = (xy + \tau)^d$	Polynomial of degree $d$
$K(x, y) = \tanh(\kappa xy + \theta)$	Multi layer perceptron

While standard SVM solutions involve solving quadratic or linear programming problems, the least square version of SVM (LS-SVM), which has been adopted for this

research, corresponds to solving a set of linear equations. In LS-SVM, the Mercer's condition is still applicable. Hence several types of kernels can be used, yet the RBF is the adopted one since it gives a Gaussian distribution for the errors in the feature space yielding an optimal estimate of the support values [14]. Many reasons could be stated for preferring LS-SVM over other models of SVM, yet the most important one is that LS-SVM is an iterative method that could be used to solve large scale problems with robustness in the sense of the choice of the regularization and smoothing parameters. Moreover, it offers a fast method for obtaining classifiers with good generalization performance in many real-life applications [15].

So far, the formulation of SVM was based on a two-class problem (SVM is essentially a binary classifier). Various schemes can be applied to the basic SVM algorithm to handle the  $M$ -class pattern classification problem. Some of these schemes [2], [16], for solving the multi-class problem are:

- Using  $M$  one-to-rest classifiers.
- Using  $M(M-1)/2$  pair-wise classifiers with one of the voting schemes: Majority voting; Pairwise coupling.
- Extending the formulation of SVM to support the  $M$  class problem: By considering all classes at once; By considering each class with only the training data points belonging to that particular class.

The derivation of the multi-class LS-SVM is based on the Lagrangian optimization formulation

$$\min_{\mathbf{w}, b, \xi} \mathfrak{J}_{LS}(\mathbf{w}, \xi) = \frac{1}{2} \mathbf{w}^T \mathbf{w} + \gamma \sum_{k=1}^N \xi_k, \quad (8)$$

subject to the equality constraint

$$y_k \left[ \mathbf{w}^T \boldsymbol{\varphi}(x_k) + b \right] = 1 - \xi_k, \quad \xi_k \geq 0, \quad k = 1, \dots, N \quad (9)$$

where  $\gamma$  is the regularization factor and  $\xi_k$  is the difference between the output  $y_k$  and discriminant function  $f(x_k)$ . Using standard techniques, the Lagrangian for (8) and (9) is

$$\mathcal{L}(\mathbf{w}, b, \xi; \alpha) = \mathfrak{J}_{LS}(\mathbf{w}, \xi) + \sum_{k=1}^N \alpha_k \left[ y_k \left( \mathbf{w}^T \boldsymbol{\varphi}(x_k) + b \right) - 1 + \xi_k \right], \quad (10)$$

where  $\alpha_k$  are the Lagrangian multipliers corresponding to (9). The saddle point is obtained from

$$\max_{\alpha} \min_{\mathbf{w}, b, \xi} \mathcal{L}(\mathbf{w}, b, \xi; \alpha). \quad (11)$$

This yields the Karush-Kuhn-Tucker optimality conditions:

$$\begin{cases} \frac{\partial \mathcal{L}}{\partial \mathbf{w}} = 0 \rightarrow \mathbf{w} = \sum_{k=1}^N \alpha_k y_k \boldsymbol{\varphi}(x_k) \\ \frac{\partial \mathcal{L}}{\partial b} = 0 \rightarrow \sum_{k=1}^N \alpha_k y_k = 0 \\ \frac{\partial \mathcal{L}}{\partial \xi_k} = 0 \rightarrow \alpha_k = 2\gamma \xi_k \\ \frac{\partial \mathcal{L}}{\partial \alpha_k} = 0 \rightarrow y_k \left[ \mathbf{w}^T \boldsymbol{\varphi}(x_k) + b \right] = 1 - \xi_k \end{cases} \quad (12)$$

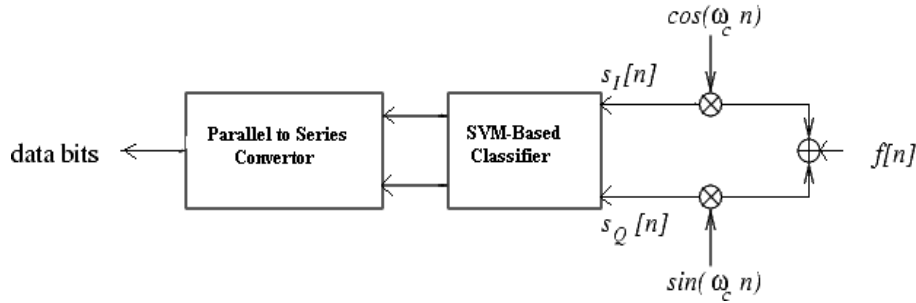


Figure 1. SVM-Based Detector for QPSK system

### 4. SVM Based QPSK Detector

Since QPSK modulation is widely used in existing wireless communication systems [9], we will describe the SVM-based QPSK detector. As shown in Fig. 1, the QPSK receiver locally generates sine and cosine carriers  $S_I[n]$  and  $S_Q[n]$ , referred to as in-phase and quadrature signals, respectively. Provided that a coherent receiver system is employed, both the in-phase and quadrature signals can be recovered exactly, allowing the system to transmit twice the amount of signal information at the same carrier frequency that could be achieved by a single oscillator.  $S_I[n]$  and  $S_Q[n]$  are then fed into the SVM classifier in order to classify the received signal into a parallel bit stream that will be next converted into a series bit stream representing the received desired data.

A similar diagram can be adopted for higher classes of MPSK where the only change will be in the number of local oscillators ( $\log_2 M$ ).

### 5. Massive Multiple Input Multiple Output (m-MIMO)

Massive MIMO systems can be defined simply as follow: Given an arbitrary wireless communication system, we consider a link for which the transmitting end, as well as the receiving end, are equipped with large scale antenna array. The idea behind MIMO is that the signals on the transmitting (TX) antennas at one end and the receive (RX) antennas at the other end are combined in such a way that the quality (bit-error rate or BER of the communication for each m-MIMO user will be improved [17]. Such an advantage can be used to increase both the network's quality of service and the operator's revenues significantly.

The channel model adopted for the m-MIMO system will be the multipath Rayleigh flat fading with the presence of Additive White Gaussian Noise (AWGN) and Co-Channel Interferer (CCI) in which the code word encompasses M-fading blocks. The complex fading gains are constant over one fading block but are independent from block to block. We model the CCI as a white Gaussian process whose variance is constant across one block and changes independently from block to block. We further adopt a simple two state model for the CCI noise

variance. The additive noise samples are, therefore, independent samples of a zero-mean complex Gaussian random variable. The variance of the noise is assumed to be constant across the fading block but changes independently from block to block. While this model may not be very realistic. In fact, this model is widely used in the literature to model the CCI as mentioned in [18] and [19].

There are many different schemes to implement the MIMO receiver, yet in this paper only two will be adopted to simulate and compare: Equal Gain Combiner (EGC) and Root-Mean-Square Gain Combiner (RMSGC). The first scheme (EGC) receives the signal on  $N$  Antennas and magnifies it by a gain  $G$  (that is equal for all  $N$  Antennas), then it combines all received signals from the  $N$  Antennas and passes them to the detector as shown in figure 2(a). The second scheme (RMSGC) also receives the signal on  $N$  Antennas and squares each one of the received signals, then all the signals are combined and averaged, then passed through a square-root device and submitted to the detector as shown in figure 2(b).

The new proposed detector which is the SVM-based detector has proved to be a better detector than the classical ones. Once again the SVM-based detector has proven to be a better one from a QoS point of view. The outcome presented in figure 3 is the result of a simulating each of the EGC and RMSGC schemes with Maximum Likelihood (ML) detector and SVM-based detector.

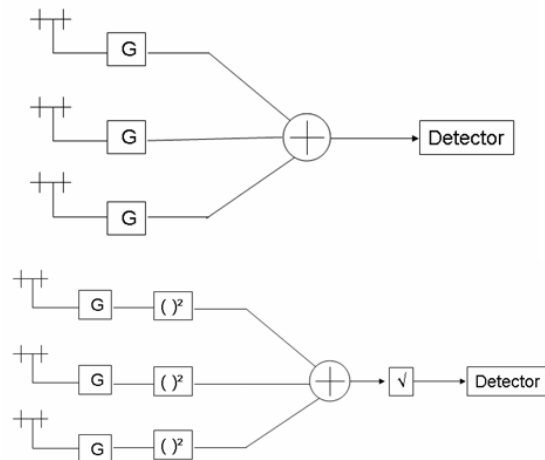
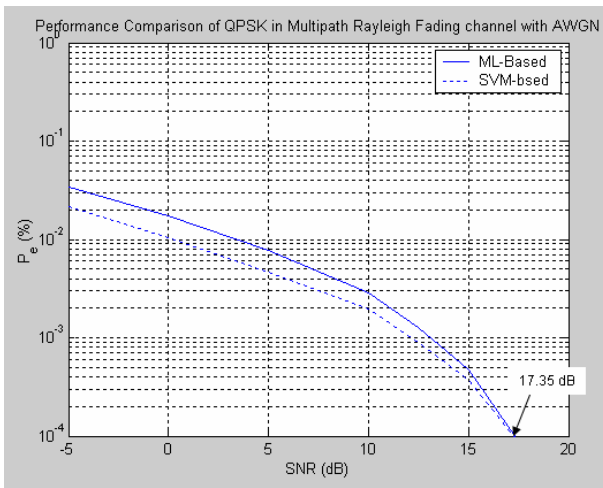


Figure 2. (a) MIMO EGC schematic (b)MIMO RMSGC schematic

## 6. Simulation Results and Discussions

For simulation purposes, Matlab is used due to its enhanced mathematical capabilities and engineering-based structure. The LS-SVM model was simulated using Matlab code on a 1.7 GHz Pentium IV computer with 256 MB RAM, to ensure that the comparison with classical detectors is fair since it is the main scope of this simulation. Without loss of generality (wlg) and for the purpose of simulation, we assumed  $P_{diff} = 1$  (Rayleigh driving parameter) and  $\lambda = \varepsilon = 1$  (colour noise parameters). In order to take full advantage of the SVM scheme, we consider several samples of the QPSK signal in the bit period. This offers a generalization since SVM is applied in a wider space.



**Figure 3.** Comparative performance of a QPSK wireless system in multipath Rayleigh fading with additive colour Gaussian noise (ACGN)

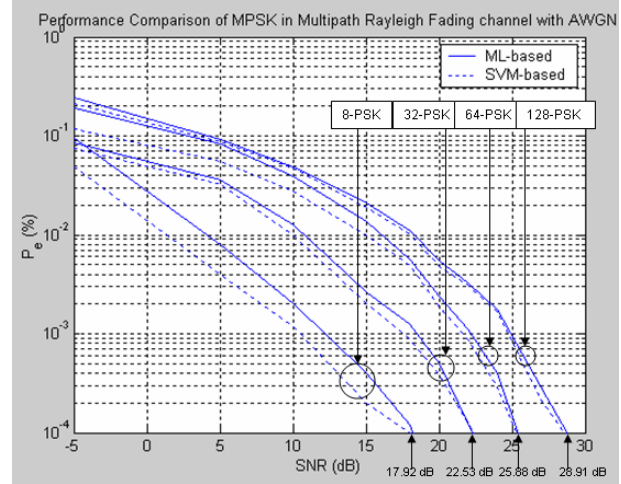
The results of the simulated LS-SVM-based QPSK system are shown in Fig. 3. We observe that the SVM-based detector outperforms the classical ML-based detector for low SNR, while for high SNR, both systems seem to produce similar results and converge at 17.35 dB.

Yet this superior performance occurs at cost of processing time as shown in the Table 2. This drawback is expected because the SVM method is a block-data based method.

**Table 2.** Processing Time for the Different Schemes of QPSK

Adopted Scheme	Processing Time (micro Sec/bit)
Classical	0.0945
SVM	0.5054

After QPSK, we simulated higher levels of  $M$ -ary PSK in order to assess the performance of the SVM-based detector and compare it with the classical detectors for  $M = 8, 32, 64$  and 128. The results of the simulation are shown in Fig. 4.



**Figure 4.** Performance of MPSK in Rayleigh fading with ACGN for different  $M$

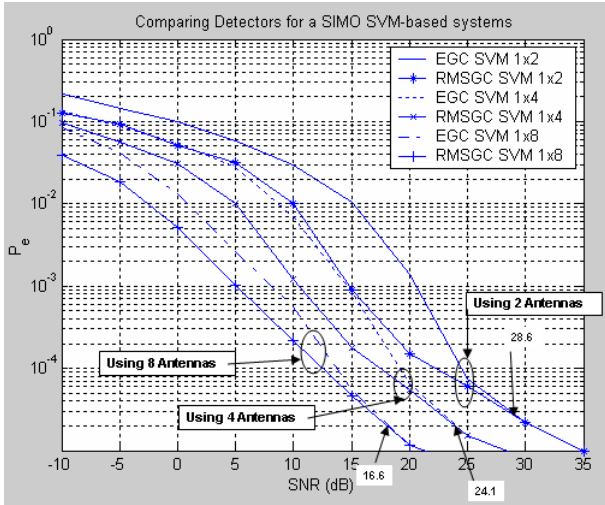
Again, the SVM outperformed the ML detector for low SNR and it achieved the ML performance asymptotically. For large SNR, there are diminishing differences in the BER curves. For high SNR, the BER is very weak and cannot be measured with sufficient precision for both methods, so a much larger training data block must be used. The converging dB levels of the SNR for both SVM and ML detectors are displayed in Table 3.

**Table 3.** Working Region of SVM-Based Detector for MPSK System

M-ary Value	SNR Value (dB)
8-PSK	17.92
32-PSK	22.53
64-PSK	25.88
128-PSK	28.91

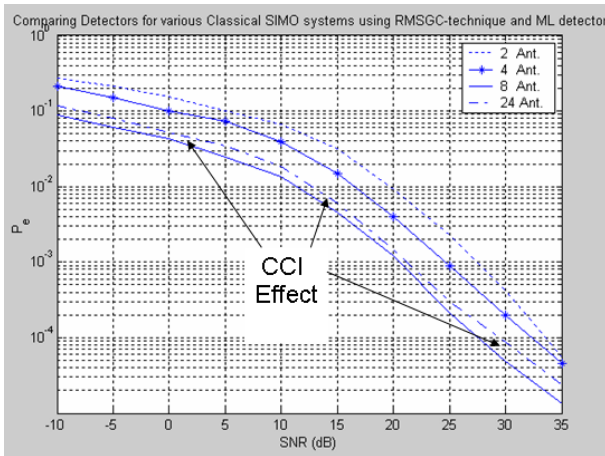
We also observe from Fig. 4 that for  $M > 64$  ( $M = 128$  or 7-bits coding) the SVM-based system does not give much better results than that of the classical systems.

For the MIMO system, first we need to establish the better technique for the simulation. The two techniques, which are EGC and RMSGC have been simulated in a SIMO system using different number of receiving antennas.



**Figure 5.** Comparing Performance of EGC vs. RMSGC for a SIMO system

The outcome presented in figure 5 is self-explanatory; it shows definite supremacy of the RMSGC technique over the EGC technique for low SNR; however, for relatively high SNR the performance of the two techniques seems to converge.

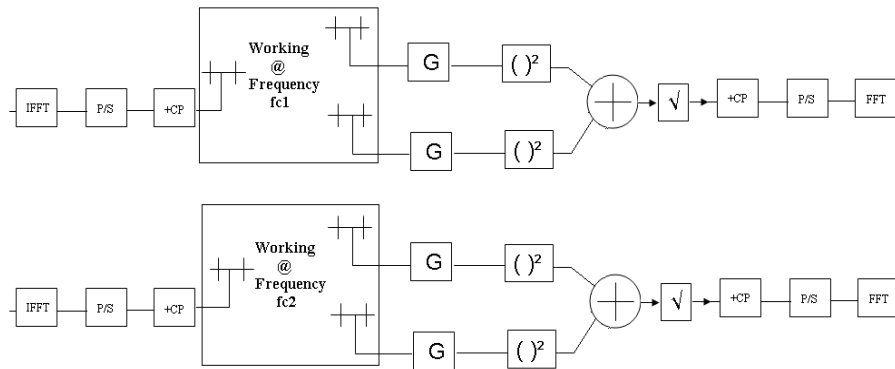


**Figure 6.** The effect of the additive noise on the multiple antennas' phenomena

At first glance, it looks like the more number of implemented antennas, the better Quality-of-Server it could be achieved. However, after conducting a simulation, the outcome shown in figure as shows that this is not accurately the case, and that behavior is due to the assumption that has been taken to suppress the additive noise  $n(t)$ , yet it had a limited effect on the transmitted signal till the number of simultaneously transmitted signals exceeds a certain limit after which the effect of the “added additive noise” on the transmitted signals will be completely destructive and it will deteriorate the effect of the multiple antenna phenomena.

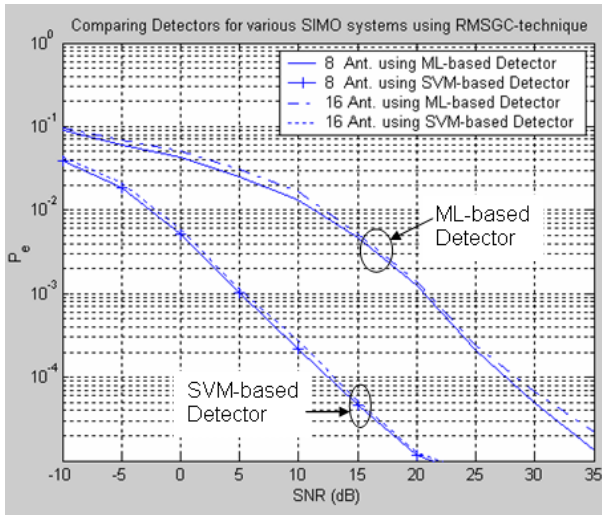
Figure 6 shows clearly that the optimum number of receiving antennas is eight since the performance of the system for a QoS viewpoint shows that the system with eight receivers is relatively the best choice and even though the system with ten receiving antennas gives a relatively better performance, it is not cost-efficient to pay for two more antennas for that relatively small enhancement in the behavior of the system.

To prevail over the influence of the “Added Additive Noise”, we can adopt the Orthogonal Frequency Division Multiplexing (OFDM). OFDM inherently provides frequency diversity over sub-channels (or tones), which offers an opportunity for both interference averaging and interference avoidance in the frequency domain [20]. The main motivation for using OFDM in a MIMO channel is the fact that OFDM modulation turns a frequency-selective MIMO channel into a set of parallel frequency-flat MIMO channels. This renders multi-channel equalization particularly simple, since for each OFDM-tone only a constant matrix has to be inverted. The user signals pass through IFFT, parallel-to-serial (P/S) conversion, and cyclic prefix insertion (+CP) at the transmitter side, and the corresponding inverse processing at the receiver. The operations of CP insertion and removal make the effective channel responses into circulant matrices as indicated in Figure 7, which can be diagonalized by normalized (unitary) IFFT and FFT matrices.



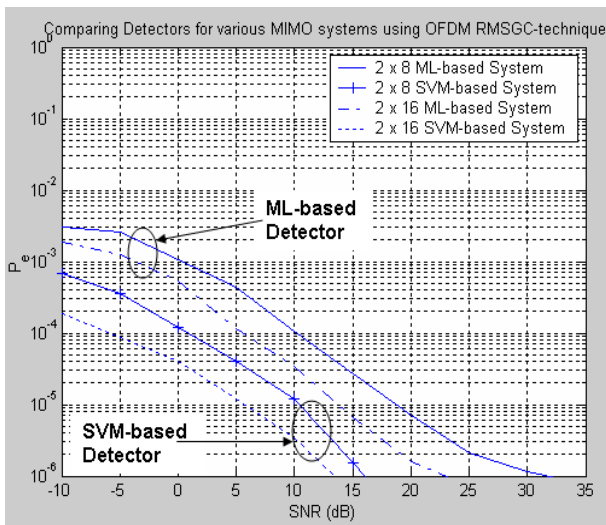
**Figure 7.** MIMO-OFDM system block diagram

Again, to study the behavior of the MIMO-OFDM, we need to show its performance compared to the MIMO system previously mentioned in the paper. The outcome presented in figure 8 shows the performance of a SIMO system without the use of OFDM technology for 8 and 16 receiving antennas, and figures 9 shows the performance of the MIMO-OFDM system for the 2x8 and 2x16 systems.



**Figure 8.** Comparing detectors for various SIMO systems using RMSGC

The results shown in figure 8 prove once again the superiority of the SVM-based detector over the classical ML detector, and show the performance of the SIMO system that doesn't adopt the OFDM technology and emphasizes on the effect of the additive noise on the multiple antennae.



**Figure 9.** Comparing MIMO-OFDM Performance for different systems

On the other hand, the results shown in figure 9 show that MIMO-OFDM overcome this problem and combat the deterioration effect of the accumulating additive noise especially the Co-Channel Interference (CCI).

## 7. Conclusions

In this paper, we first applied SVM to single-input-single-output MPSK detection in wireless systems with severe noise conditions modeled by Rayleigh fading channel noise and colour receiver noise. SVM is found to be a learning machine suitable for wireless communication with the ability to handle data coming out from a relatively hostile wireless channel at a low SNR with a considered superiority to the classical ML detector at the cost of relatively longer processing time, especially for higher classes. For large SNR, the performance of the SVM detector was similar to that of the ML detector.

Then we simulated the SVM for a MIMO system. The main part was to introduce the SVM-based detector and prove its superiority over the classical detectors adopted in the general design and implementation of the MIMO system. However, in order to improve its supremacy, we first wanted to decide the best schematic to implement the MIMO system; we compared the reputed EGC with the RMSGC and proved that the later outperform. Then we had to decide on the optimum number of antennas to be used at the receiver, after simulating the system for different numbers of antennas, the proper choice was taken to be eight. Finally, the system faced the problem of additive noise, which will restrict the enhancement of the system, and the solution was found by employing the technology of OFDM which results a splendid outcome.

One major weakness of SVM that could be realized from Fig. 4 is that it does not give much better results than the ML detector for classes above 64. This is expected since SVM is essentially a binary classifier. However, the approaches mentioned in section III to extend the SVM to multi-class detection did give reasonable results for classes up to 64.

We also showed that conventional detectors require as much as 7 bits-coding ( $M \geq 128$ ) to produce comparable results to those of SVM. Therefore, SVM can be successfully applied to  $M$ -ary signaling with 1 to 6 bits-coding. Since QPSK and 8-PSK are widely used in existing wireless communication systems, notably in cellular UMTS and land mobile satellite, we expect this research to trigger a newer generation of SVM-based wireless systems.

Although the simulations were conducted for a medium-scale antenna array system due to computational limitations, the same results and conclusions can be extended to large-scale antenna systems.

As future work, we propose to adopt one of the many pre-designed SVM chips [21] and implement a real-time system to compare results with the simulation outputs. As processors technology becomes faster SVM will be able to meet real-time computational requirements for nowadays high-speed data communications.

We also propose to apply relevance vector machine (RVM) [22] to MPSK detection for Giga-bits Ultra Wide Band (UWB) wireless systems [23, 24]. RVM has an

identical functional form to SVM and has been demonstrated to have a comparable generalization performance to SVM while requiring dramatically fewer kernel functions, thus enjoying faster computation.

---

## REFERENCES

- [1] V. Vapnik, "Estimation of Dependences Based on Empirical Data," *Nauka English Translation*, Springer Verlag, 1982.
- [2] J. S. Daba and M. R. Bell, "Segmentation of Speckled Images Using a Likelihood Random Field Model," *Optical Engineering*, vol. 47, no. 1, pp. 017005-1 to 017005-20, 2008.
- [3] S. Chen, S. Gunn, and C. Harris, "Decision Feedback Equalizer Design Using Support Vector Machines," *Inst. Elect. Eng. Proc. Vision, Image and Signal Processing*, vol. 147, no. 3, 2000, pp. 213-219.
- [4] D. Sebald and J. Bucklew, "Support Vector Machine Techniques for Non Linear Equalization," *IEEE Trans. Sig. Proc.*, vol. 48, 2000, pp. 3217-66.
- [5] F. Albu and D. Martinez, "The Application of SVM with Gaussian Kernels for Overcoming CCI," *IEEE Int. Workshop Neural Net.*, 1999, pp. 49-57.
- [6] F. Perez-Cruz, A. Vazquez, P. Dianna, and A. Rodriguez, "SVC-Based Equalizer for Burst TDMA Transmission," *Signal Processing*, 2000.
- [7] J. P. Dubois and O. M. Abdul-Latif, "SVM-Based Detection of SAR Images in Partially Developed Speckle Noise," *IEC (Prague)*, Czech Republic, 2005.
- [8] C. Mokbel and F. Hashem, "Support Vector Machines in Digital Communication," Master Thesis, Univ. of Balamand, Lebanon, 2003.
- [9] M. Simon and M. Alouini, *Communication over Fading Channels*. New Jersey: Wiley-IEEE Press, 2005.
- [10] L. Gagliardi and S. Karp, *Optical Communications*. New York: Wiley, 1995.
- [11] N. Christianini and J. Taylor, "Support Vector Machine and Other Kernel Learning Methods". London: Cambridge University Press, 2003.
- [12] J. Christopher and C. Burges, "A Tutorial on Support Vector Machines for Pattern Recognition", *Kulwer Publishers*, vol. 12, no. 2, 1998, pp. 121-167.
- [13] T. Joachims, "Support Vectors and Kernel Methods", Cornell University.
- [14] J. Suykers and J. Vardewalle, "Multi-Class Least Square-Support Vector Machine," Universite Catholique de Louvain, Belgium, unpublished.
- [15] K. Lin and C. Lin, "A Study on Reduced Support Vector Machines," *IEEE Transaction on Neural Networks*, vol. 14, no. 6, 2003.
- [16] C. Hsu and C. Lin, "A Comparison of Methods for Multi-Class Support Vector Machines," *IEEE Trans. Neural Net.*, vol. 13, 2002, pp. 415-425.
- [17] Ayman Naguib, Da-shan Shiu, David Gesbert, Mansoor Shafi; "From Theory to Practice: An Overview of MIMO Space-Time Coded Wireless Systems", *IEEE Journal on selected areas in communications*, vol 21, No 3, pp. 281-302, April 2003.
- [18] Paulraj, Gesbert, Papadias, "Smart Antennas for Mobile Communication". Encyclopedia for electrical engineering, John Wiley Publishing Co., 2000
- [19] Anand Arunachalam, Hesham El Gamal, "Space-Time Coding for MIMO Systems with Co-Channel Interference". Conference on Information Sciences and Systems, September 2002.
- [20] L. H'aring, A. Czylik, "Synchronization in MIMO OFDM systems", *Advances in Radio Science* (2004) 2: 147-153 © Copernicus GmbH 2004
- [21] D. Anguita, A. Boni, and S. Ridella, "A Digital Architecture for SVM," *IEEE Trans. Neural Networks*, vol. 14, no. 5, 2003, pp. 993-1000.
- [22] M. Tipping, "The Relevance Vector Machine," in *Advances in Neural Information Processing Systems*. Cambridge: MIT Press, 2000.
- [23] J. Daba, "Performance of Ultra Wide Band Systems in High Speed Wireless Personal Area Networks," *Advances in Electrical and Computer Technologies, Lecture Notes in Electrical Engineering Series*, Springer Nature, vol. 672, pp. 911-935, September 2020.
- [24] J. S. Daba, "Mean Bit Error Analysis of IEEE 802.15.3.a UWB Channels," *Universal Journal of Electrical and Electronic Engineering*, Horizon Research Publishing, 2020.