Modulation Classification of MIMO-OFDM Signals by Independent Component Analysis and Support Vector Machines

H. Agirman-Tosun*, Yu Liu*, A.M. Haimovich*, Osvaldo Simeone*, Wei Su**
Jason Dabin***, Emmanuel Kanterakis****,

* New Jersey Institute of Technology, Newark, NJ 07102
** US Army CERDEC Aberdeen Proving Ground, MD 21005
***US Navy SPAWAR SCP, San Diego, CA
**** CACI International, Eatontown, NJ 07724 USA

***Abstract*— A modulation classification scheme based on Independent Component Analysis (ICA) in conjunction with either maximum likelihood or Support Vector Machines (SVM) is proposed for MIMO-OFDM signals over frequency selective, time varying channels. The method is blind in the sense that it is assumed that the receiver has no information about the channel and transmitted signals other than that the spatial streams of signals are statistically independent. The processing consists of separation of the MIMO streams followed by modulation classification of the separated signals. While in general, blind separation of signals over frequency selective channels is a difficult problem, the non-frequency selective nature of the channel experienced by individual symbols in a MIMO-OFDM system enables the application of well-known ICA algorithms. Modulation classification is implemented by maximum likelihood and by an SVM-based modulation classification method relying on pre-selected modulation-dependent features. To improve performance in time varying channels, the invariance of the is exploited across the coherence bandwidth and the time coherence. The proposed method is shown to perform with high probability of correct classification over realistic ITU pedestrian and vehicular channels.**

I. Introduction

To meet the growing demand for high-data rates in communications systems, new wireless applications rely on multiple-input multiple-output (MIMO) technologies. MIMO can support increased data capacity through, spatial multiplexing, i.e., the transmission of data in parallel streams. Orthogonal Frequency Division Multiplexing (OFDM) is a multicarrier transmission technique where the frequency band is divided into several orthogonal sub-bands, such that the symbols transmitted on each sub-band experience frequency non-selective fading. Channel equalization is then reduced to a one-tap filter per data symbol. The combination of MIMO transmission and OFDM data modulation is central to fourth generation (4G) wireless technologies, such as WiMax, LTE, and IEEE 802.22.

Recognition of the modulation of unknown received signals has obvious military applications. As for civilian applications, attempts to reduce overhead of reference signals required for channel estimation has motivated research in blind and semi-blind MIMO techniques. Blind techniques are also expected to play a role in software defined radio and cognitive radio. Configuration information required by a software defined radio system is transmitted as overhead to the data. However, intelligent receivers capable of extracting this information blindly may improve transmission efficiency through reductions in overhead. For example, automatic modulation classification eliminates the need for supplementary information on the modulation type.

In general, classification requires preprocessing of the received signals for acquiring signal parameters, such as carrier frequency and symbol rate. This paper focuses on the modulation classification of MIMO-OFDM signals assuming that frequency and time synchronization have already been attained.

Modulation classification methods for single-input single-output SISO systems are generally classified as likelihood-based or feature-based. Likelihood-based classification is optimal in the sense of attaining minimum probability of misclassification, but is computationally complex often requiring exhaustive searches through parameter values. With feature-based methods, specific features are extracted from the signal and compared with pre-calculated values. Feature-based methods are usually ad-hoc, but computationally efficient. A detailed survey of automatic modulation classification methods for SISO systems is given in [1].

MIMO modulation classification relies on the blind channel estimation of the MIMO channel. Blind MIMO channel estimation has been an active area of research (e.g. [2] and [3]). Blind channel estimation for MIMO-OFDM has been studied in [4]. A likelihood-based approach to MIMO modulation classification is proposed in [5], where the channel matrix required for the calculation of the likelihood is first estimated blindly by independent component analysis (ICA).

The main contributions of the current paper are: (1) exploit the frequency non-selective channel experienced by the MIMO-OFDM data symbols and the finite frequency and time selectivity to perform modulation classification on groups of data symbols with a common channel; (2) develop a low complexity SVM-based modulation classifier.

The rest of the paper is organized as follows: the next section introduces the signal model, the proposed MIMO-OFDM modulation classification methods are presented in Section III, numerical examples are provided in Section IV, and Section V wraps up with conclusions.

II. SIGNAL MODEL

Consider a MIMO-OFDM system with M_t transmit antennas and M_r receive antennas. Identifiability conditions of the MIMO channel require, $M_t \le M_r$. The system transmits frames of OFDM symbols with $\mathbf{s}^{(i)}(k,n)$ denoting the length M_t symbol vector of modulation $\Omega^{(i)}$, transmitted on subcarrier index n and frame index k. A frame is an OFDM block of data symbols. The transmitted symbols are of unknown PSK/QAM modulation, but are assumed statistically independent between antennas, subcarriers and frames. In addition, ideal time synchronization as well as ideal carrier frequency synchronization is assumed at the receiver side. A block diagram of the MIMO-OFDM system is shown in Fig. 1.

Assuming a cyclic prefix that ensures inter-carrier interference-free observations, the received length M_r vector in the frequency domain, $\mathbf{y}(k,n)$, is expressed

$$\mathbf{y}(k,n) = \mathbf{H}(k,n)\mathbf{s}(k,n) + \mathbf{z}(k,n) \tag{1}$$

where $\mathbf{H}(k,n)$ is the MIMO channel matrix associated with the subcarrier and frame, and $\mathbf{z}(k,n)$ is additive white Gaussian noise. The noise is complex-valued, zero mean, has known variance $\sigma^2/2$ for both real and imaginary parts, and is independent between receive antennas, subcarriers, and frames.

Blind estimation of the MIMO channel relies on channel values that remain static over multiple observations. We exploit the coherence bandwidth and time coherence of the channel, assumed known at the receiver, to form a set of K frames and N subcarriers over which the channel is fixed, and denote the channel matrix $\mathbf{H}$. For notational convenience, the KN observation vectors $\mathbf{y}(k,n)$, signal vectors $\mathbf{s}(k,n)$, and noise vectors $\mathbf{z}(k,n)$ associated with channel $\mathbf{H}$, are re-indexed $\mathbf{y}_k$, $\mathbf{s}_k$, and $\mathbf{z}_k$, respectively, for $k = 1,\ldots,KN$. With that, the received signal is written

$$\mathbf{y}_k = \mathbf{H}\mathbf{s}_k + \mathbf{z}_k \tag{2}$$

For future use, denote $\mathbf{Y} = \{\mathbf{y}_k\}_{k=1}^{KN}$ the set of all observations and $\mathbf{S}^{(i)} = \{\mathbf{s}_k\}_{k=1}^{KN}$ the set of all transmitted vectors belonging to modulation $\Omega^{(i)}$.

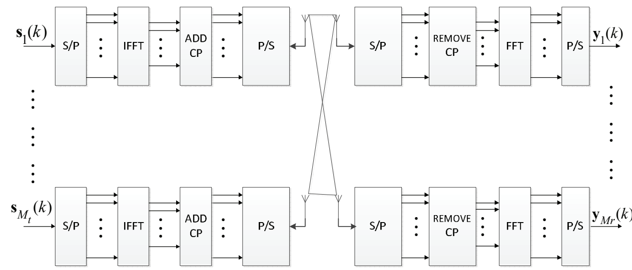

Fig. 1. MIMO-OFDM system model

A MIMO-OFDM system can be considered a set of instantaneous mixtures of transmitted signals. The problem of separating MIMO-OFDM signals becomes a blind source separation problem (BSS) at each subcarrier.

III. PROPOSED METHODS

The proposed method for the blind classification of the MIMO-OFDM signals has three stages as summarized in Fig. 2. The first stage groups subcarriers and frames to maximize the number of observations for a fixed channel matrix. The second stage applies an ICA algorithm to separate the MIMO signals. Finally, modulation classification methods are applied to the separated data streams.

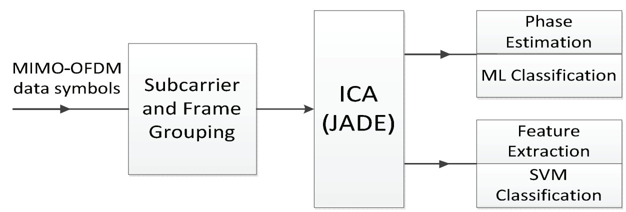

Fig. 2. Block diagram of MIMO-OFDM modulation classification

A. MIMO Separation by ICA

The received MIMO signal at each subcarrier is a linear mixture of the transmitted sources. Reconstructing unobserved signals from their linear mixtures is possible if the signals are statistically independent. ICA is a class of techniques for achieving that. Since the mutual interference between signals in the mixture dominates additive noise effects, noise is usually ignored in ICA formulations. Given the signal model (2), ICA methods [6-8] seek to compute a demixing matrix $\mathbf{W} = \mathbf{H}^{-1}$. Application of $\mathbf{W}$ to the data vector $\mathbf{y}_k$ enables to recover the vector of transmitted symbols $\mathbf{s}_k$, according to $\hat{\mathbf{s}}_k = \mathbf{W}\mathbf{y}_k$. Since $\mathbf{H} = \mathbf{W}^{-1}$, computation of the demixing matrix is equivalent to the estimation of the channel matrix.

ICA relies on statistical independence, a condition stricter than requiring the components of $\mathbf{s}_k$ to be only orthogonal. It is well known that if the components of a linear mixture are statistically orthogonal, methods relying on second order moments can separate the signals up to an invertible matrix.

For signals that are non-Gaussian and statistically independent, it is possible to resolve this ambiguity up to a complex scaling and permutation. In other words, ICA estimates the channel matrix

$$\hat{\mathbf{H}} = \mathbf{DPH} \tag{3}$$

where $\mathbf{D}$ is a diagonal, complex-valued matrix, and $\mathbf{P}$ is a permutation matrix. The scaling ambiguity can be reduced to a phase only ambiguity by setting $E[\mathbf{s}_k\mathbf{s}_k^H] = \mathbf{I}$, where the superscript denotes Hermitian operation, and $\mathbf{I}$ is the identity matrix.

A number of algorithms are available for implementing ICA. They have in common the aim to minimize a contrast function, which measures the statistical dependency of the sources in the mixture. Minimization of the contrast function yields the desired independent signals. In this paper, we use the JADE algorithm [7, 8] due to its relative fast speed of convergence. As discussed above, the signal separation is up to a permutation and phase ambiguity.

All ICA algorithms, JADE included, require multiple observations to estimate the demixing matrix $\mathbf{W}$. Thus, it is implicit in the estimation process that the target channel matrix $\mathbf{H}$ is constant over the processed observations. In practice, due to data support and noise limitations, the estimated demixing matrix $\mathbf{W}$ is only an approximation of the true inverse channel matrix. Thus each stream of separated signals contains also multiuser interference, which affects the modulation classification.

B. Maximum Likelihood Modulation Classification

For the MIMO-OFDM system (2), under the assumption of statistically independent received symbols, the likelihood function of the observations is given by

$$L\left(\boldsymbol{Y} \mid \mathbf{S}^{(i)}, \mathbf{H}\right) = \frac{1}{(\pi\sigma^2)^{NK\cdot M_r}} \prod_{k=1}^{NK} \exp\left[-\frac{1}{\sigma^2}\left\|\mathbf{y}_k - \mathbf{H}\mathbf{s}_k^{(i)}\right\|^2\right] \tag{4}$$

where the norm is Euclidean. The likelihood is conditioned on the transmitted symbols $\mathbf{S}^{(i)}$ and channel $\mathbf{H}$, which are unknown. These unknown quantities are addressed in different ways. The ICA processing yields an estimate $\hat{\mathbf{H}}$ of the channel matrix, which is then used in (4). The difficulty posed by the unknown symbols $\mathbf{S}^{(i)}$ is resolved by assuming a uniform a priori distribution, and averaging over the symbols from each constellation $\Omega^{(i)}$. With these modifications, the likelihood function for constellation $\Omega^{(i)}$ is given by

$$L^{(i)}\left(\boldsymbol{Y} \mid \hat{\mathbf{H}}\right) = E_{\Omega^{(i)}}\left[L^{(i)}\left(\boldsymbol{Y} \mid \mathbf{S}^{(i)}, \hat{\mathbf{H}}\right)\right] \tag{5}$$

The ICA stage of the processing produces a demixing matrix $\mathbf{W}$ or equivalently, an estimate of the channel matrix $\hat{\mathbf{H}} = \mathbf{W}^{-1}$. As discussed previously and indicated by (3), the channel estimate produced by ICA has inherent permutation and phase indeterminacies. It is shown in [5] that the likelihood function (5) is invariant to a permutation matrix applied to $\hat{\mathbf{H}}$, i.e. $L^{(i)}(\boldsymbol{Y} \mid \hat{\mathbf{H}}) = L^{(i)}(\boldsymbol{Y} \mid \mathbf{P}\hat{\mathbf{H}})$.

Unlike its invariance to unknown permutations of the signals, the likelihood function for modulation classification is dependent on the unknown phase offsets contained in the diagonal of the matrix $\mathbf{D}$. These phase offsets need to be estimated for correct modulation classification. If the unknown phase offset for one of the separated MIMO streams is ϕ, the log likelihood function for estimating it is

$$\ell(\phi) = \sum_{k=1}^{KN} \log\left\{\sum_{\Omega^{(i)}} \exp\left[-\frac{1}{2\sigma_x^2}\left|x(k) - s(k)e^{j\phi}\right|^2\right]\right\} \tag{6}$$

where $x(k)$ is a data symbol at the output of the ICA algorithm, $s(k)$ is one of the symbols in the constellation, and σ_x^2 is the interference term (multiuser interference due to imperfect MIMO channel equalization by the ICA). In [9] it is shown that as the SNR tends to zero, the maximum likelihood phase estimator takes on the blind form

$$\hat{\phi} = \frac{1}{P}\arg\left[E\left[s^{*P}\right]\sum_{k=1}^{KN} x^P(k)\right] \tag{7}$$

where P is the number of symbols of the constellation that are rotationally symmetric. For example, for QAM constellations, $P = 4$. According to [9], even after the phase offset estimation according to (7), there remains a leftover phase ambiguity corresponding to a multiple of the phase difference between the constellation symbols. It is an easy argument to make that this remaining ambiguity does not interfere with the modulation classification.

Substituting the estimated channel matrix in the likelihood expression, averaging over the symbols of each hypothesized constellation, and estimating the phase offsets leads to the following likelihood-based modulation classification

$$\hat{\Omega} = \arg\max_{\Omega^{(i)}} \ln\{L^{(i)}(\boldsymbol{Y} \mid \hat{\mathbf{H}})\}, \tag{8}$$

C. SVM Modulation Classification

Feature-based modulation classification methods are of interest since they have lower complexity than likelihood-based methods. Here, we propose an SVM modulation classification method that combines multiple features.

The fourth order cumulants, $|C_{40}|$ and C_{42} were previously proposed for modulation classification [10]. It is known that fourth order cumulants can be applied to distinguish between modulations, and are robust to noise effects. Given M samples of a signal $s(k)$, cumulant C_{40} is defined [10, 11]

$$C_{40} = \frac{\frac{1}{M}\sum_{k=0}^{M-1} s(k)^4 - 3\left(\frac{1}{M}\sum_{k=0}^{M-1} s(k)^2\right)^2}{\left(\frac{1}{M}\sum_{k=0}^{M-1} |s(k)|^2\right)^2} \tag{9}$$

whereas C_{42} is defined

$$C_{42} = \frac{\frac{1}{M}\sum_{k=0}^{M-1} s(k)^4 - \left|\frac{1}{M}\sum_{k=0}^{M-1} s(k)^2\right|^2 - 2\left(\frac{1}{M}\sum_{k=0}^{M-1} |s(k)|^2\right)^2}{\left(\frac{1}{M}\sum_{k=0}^{M-1} |s(k)|^2\right)^2} \quad (10)$$

It is shown in [10] that $|C_{40}|$ and C_{42} are invariant to carrier phase offset, which in our case corresponds to the phase ambiguity inherited from the ICA algorithm. Therefore, classification based on these features is not affected by the unknown phase offsets introduced by the ICA algorithm.

SVM is an important pattern recognition method, in which each pattern is represented by D features [12-13]. The aim in SVM is to find the best separating hyperplane in D-dimensional space to discriminate between the patterns. SVM processing has two main steps: training and testing. In the training step, training data obtained from known sources is processed to find the optimum hyperplane separating data of different modulations. In the second stage of SVM, the test data is compared with the separating hyperplane and then classified accordingly. The training data in the proposed method corresponds to a number of $|C_{40}|$ and C_{42} values for each candidate modulation and for each SNR value. Since thresholding between the modulations is SNR dependent, SVM modulation classification requires knowledge of the SNR at the receiver.

IV. Numerical Examples

Numerical simulations were carried out to demonstrate the ability of the proposed MIMO-OFDM modulation classification methods to discriminate between QPSK and 16QAM modulations over standard channel models. The simulations addressed a MIMO-OFDM system with two transmit antennas, $M_t = 2$, and four receive antennas, $M_r = 4$. The total number of subcarriers was set to $N = 512$, and the signal bandwidth was chosen 6.4 MHz.

Performance was evaluated over the ITU pedestrian B and vehicular A channel models with the parameters given in Table I. The maximum speeds were assumed 3 km/hr for pedestrians, and 60 km/hr for vehicles. Time variations of the channels were modeled according to the Clarke and Guns model [14].

In the training stage of the SVM, 150 symbols from each modulation QPSK and 16QAM were generated synthetically. The average power of each modulation was set to unity. White Gaussian noise with a set variance corresponding to the SNR was added to the modulated signal. Finally, the features $|C_{40}|$ and C_{42} of the noisy training data were calculated according to (9) and (10). SVM was implemented with the Matlab Bioinformatics Toolbox. The probability of correct classification was computed based on 1000 Monte Carlo trials for each SNR value.

Fig. 3 shows the probability of correct classification for pedestrian B channels, and both stationary and moving terminals. In this figure, performance was evaluated using observations collected from 50 OFDM frames, without exploiting the similarity in channel response across the coherence bandwidth. Classification scores are shown for both likelihood-based and SVM-based methods. It is observed that both methods perform well for stationary and moving pedestrians. This is not surprising since at 3 km/hr, a pedestrian experiences a negligible Doppler spread. To achieve 85% correct classification rate, the likelihood-based approach outperformed the SVM approach by approximately 10 dB.

It is interesting to compare the computational complexity of the two methods. The number of basic operations required for the SVM approach is $O(KN \times M_t)$, where KN is the number of data samples processed in the ICA algorithm. In contrast, the likelihood-based approach requires $O(KN \times M_u^{M_t})$ basic operations, where M_u is the maximum number of hypothetical states of the assumed constellations. In our case, $M_u = 16$, since the 16QAM modulation has 16 states. Thus the complexity of the likelihood-based algorithm can be considerably higher than that of the SVM algorithm.

The second simulation was performed for the vehicular A channel. In Fig. 4 are shown the successful classification rates for stationary and moving terminals. At high Doppler, the performance is significantly degraded. This degradation is due to the time varying nature of the channel, which renders the channel estimation by ICA useless.

In order to improve the performance over time varying channel, the number of OFDM frames should be limited according to the coherence time of the channel. Regarding the OFDM system simulated, for a moving vehicle at the speed of 60 km/hr, the coherence time spans approximately 35 OFDM frames. But from the simulation results, we observed that JADE needs at least 50 samples to separate the MIMO data streams. One way to overcome this problem is through grouping subcarriers according to the coherence bandwidth of the channel. In the simulations, we grouped the data of five subcarriers and 10 frames. The performance is shown in Fig. 5. This method significantly improved the classification rate. For example, even at high speed, likelihood-based correct classification is close to guaranteed for SNR greater than 10 dB, while classification with SVM has a success rate higher than 80% for SNR greater than 10 dB.

V. Conclusions

This paper studies two classification modulation methods for MIMO-OFDM signals. First, the ICA JADE algorithm is applied to separate the data streams of the MIMO-OFDM signal. The modulation of the separated data streams is subsequently detected by maximum likelihood and by SVM. Modulation classification of MIMO signals is more challenging than for SISO signals due to the residual multiuser interference resulting from imperfect signal separation. Our approach to modulation classification relies on the frequency non-selective nature of the channel experienced by individual OFDM data symbols, and exploits the invariance of the MIMO channel matrix across the coherence bandwidth and time coherence. The proposed modulation classification methods were demonstrated over ITU channels. Over the slowly time-varying ITU pedestrian channel, the proposed SVM method achieved 85% classification rates for SNR higher than 15 dB. Over the fast fading channel, high probabilities of correct classification were maintained by grouping signals according to the coherent bandwidth and time coherence of the channel. In this case, performance was close to the static channel.

References

[1] O. A. Dobre, A. Abdi, Y. Bar-Ness, and W. Su, "A survey of automatic modulation classification techniques: classical approaches and new developments," IET Communications vol.1, pp. 137-156, April 2007.

[2] K. Abed-Meraim, P. Loubaton, and E. Moulines, "A subspace algorithm for certain blind identification problems," IEEE Trans. Inform. Theory, vol. 43, pp. 499–511, Mar. 1997.

[3] Jing Liang; Zhi Ding; , "Blind MIMO system identification based on cumulant subspace decomposition," Signal Processing, IEEE Transactions on , vol.51, no.6, pp. 1457- 1468, June 2003

[4] C. Shin, R. W. Heath, and E. J. Powers, "Blind channel estimation for MIMO-OFDM systems," IEEE Trans. Veh. Technol., vol. 56, no. 2, pp. 670–685, Mar. 2007.

[5] V. Choqueuse, S. Azou, K. Yao, L. Collin and G. Burel, "Blind Modulation recognition for MIMO Systems," MTA Review, pp. 183-195, June 2009.

[6] A. Hyvarinen, J. Karhunen, E. Oja, *Independent Component Analysis*, John Wiley & Sons, 2001.

[7] JF. Cardoso, "Blind Beamforming of Non-Gaussian Signals," IEEE Proceeding of Radar and Signal Processing, vol.140, pp.362-370, December 1993.

[8] JF. Cardoso, "High-Order Contrast for Independent Component Analysis," Neural Computation, vol.11, pp.157-192, 1999.

[9] M. MOENECLAEY, G. de JONGHE – ML-Oriented NDA Carrier Synchronization for General Rotationally Symmetric Signal Constellations, IEEE Transactions on Communications, Vol. 42, No. 8, pp. 2531-2533, Aug. 1994

[10] A. Swami and B.M. Sadler, "Hierarchical Digital Modulation Classification Using Cumulants," IEEE Transactions on Communications, vol.48, pp.416-429, March 2000.

[11] H. Mustafa, "Algorithms for Digital Modulation Classification," PhD Thesis, The George Washington University, April 2007.

[12] C.J.C. Burges, "A Tutorial on Support Vector Machines for Pattern Recognition," Data Mining and Knowledge Discovery, vol.2, pp-121-167, 1998.

[13] T. Fletcher, Support Vector Machines Explained (online), http://www.tristanfletcher.co.uk/SVM%20Explained.pdf

[14] T.S, Rappaport, *Wireless Communications: Principles and Practice*, Prentice Hall, 2002.

TABLE I
ITU TAPPED DELAY LINE CHANNEL PARAMETERS FOR PEDESTRIAN B AND VEHICULAR A TEST ENVIRONMENT

Tap	Pedestrian B		Vehicular A		Doppler spectrum
	Relative delay (ns)	Average power (dB)	Relative delay (ns)	Average power (dB)	
1	0	0	0	0.0	Classic
2	200	–0.9	310	-1.0	Classic
3	800	–4.9	710	-9.0	Classic
4	1 200	–8.0	1090	-10.0	Classic
5	2 300	–7.8	1730	-15.0	Classic
6	3 700	–23.9	2510	-20.0	Classic

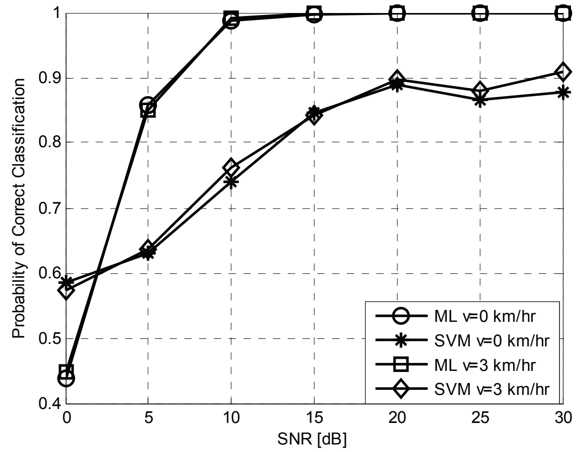

Fig. 3. Probability of Correct Classification over ITU Pedestrian B channel.

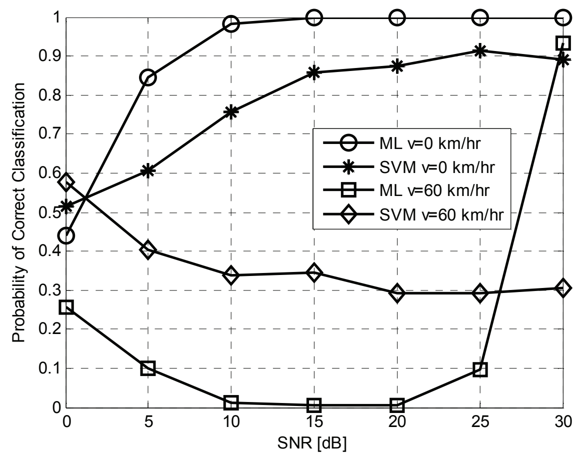

Fig. 4. Probability of Correct Classification over ITU Vehicular A channel.