

AUTOMATIC MODULATION CLASSIFICATION USING A WAVEFORM SIGNATURE

William H. Clark IV (saikou@vt.edu), Joseph M. Ernst (jmernst@vt.edu), and
Robert W. McGwier (rwmcgwi@vt.edu)

Ted and Karyn Hume Center for National Security and Technology,
Virginia Tech ECE Department,
Blacksburg, VA, USA

ABSTRACT

Cognitive Radios (CRs) build upon Software Defined Radios (SDRs) to allow for autonomous reconfiguration of communication architectures. In recent years, CRs have been identified as an enabler for Dynamic Spectrum Access (DSA) applications in which secondary users opportunistically share licensed spectrum. A major challenge for DSA is accurately characterizing the spectral environment, which requires blind signal classification. Existing work in this area has focused on simplistic channel models; however, more challenging fading channels (e.g., frequency selective fading channels) cause existing methods to be computationally complex or insufficient. This paper develops a novel blind modulation classification algorithm, which uses a set of higher order statistics to overcome these challenges. The set of statistics forms a signature, which can either be used directly for classification or can be processed using big data analytical techniques, such as principle component analysis (PCA), to learn the environment. The algorithm is tested in simulation on both flat fading and selective fading channel models. Results of this blind classification algorithm are shown to improve upon those which use single value higher order statistical methods.

keywords - Modulation Classification, Cumulants, Selective Fading, Time Correlation, Modulation Signature, Big RF

1. INTRODUCTION

Automatic Modulation Classification (AMC) is the meter by which CRs are able to dynamically change their receiver architecture based upon observed signals. The AMC algorithms are therefore between the detection of a signal and the demodulation of symbols back to bits at baseband. This functionality has a diverse array of applications in military and civil communications systems. One of the most recent applications is within DSA systems to allow for adaptive secondary users, which can take advantage of white spaces in under utilized spectrum bands.

There are two major trends in developing AMC algorithms: Likelihood-Based (LB) and Feature-Based (FB) [1]. While the LB algorithms can be shown to be the optimal solution in the Bayesian sense [2], the computational complexity required for

achieving that optimum is often too intense for many mobile application to handle in real time while maintaining a reasonable power budget. For this reason, this paper focuses on the FB algorithms [3–10] instead of the LB approach [2, 11]. This paper makes use of multiple cumulants from second to tenth order for a total of twenty features, which form a Waveform Signature (WS). While the higher order cumulants suffer from an increased variance in the estimator, there is also an increased separation between ideal cumulant values that can be useful in the WS. In contrast to approaches present in the literature that make use of a Decision Tree (DT) and operate on one feature at a time [3–7], or Support Vector Machines (SVM) [9, 10], which use a set of features in a one-vs-all DT, the proposed algorithm makes use of L1 norm distance to classify the observed signal for its simplicity. The classification for the WS algorithm presented here is performed using the observed signal’s WS and minimizing the L1 norm distance between a known WS learned in a supervised database.

The paper is organized as follows: Section 2 describes the channel models considered. Section 3 explains the features used in this paper and an algorithm that exists in the literature for AMC that use similar features. Section 4 then covers the proposed WS generation and AMC algorithm. Section 5 then compares the proposed algorithm with the existing algorithm on the different channel models, and Section 6 concludes.

2. SYSTEM MODEL

The sampled received signal can be represented as

$$r(\nu) = \sum_{\tau=0}^{L-1} h(\nu, \tau)x(\nu - \tau) + n(\nu) \quad (1)$$

where $r(\nu)$ is the received signal, $x(\nu)$ is the unit variance transmitted symbols with rectangular pulse shape equally distributing the energy over three samples per symbol. $n(\nu)$ is AWGN with circularly-symmetric $\mathcal{CN}(0, \sigma_n^2)$ distribution. The channel effects are contained in $h(\nu, \tau)$ where ν is the sample instance and τ is delay relative to ν .

The first channel model considered in this paper is given in Section 2.1 where both flat fading and block fading are assumed.

The next channel examined is given in Section 2.2 where the flat fading assumption is relaxed and the channel undergoes frequency selective fading. The final channel model examined uses the flat fading model, but relaxes the block fading assumption. Using Clarke's model [12], time correlation is added to the flat fading channel model based on Doppler Spread. This channel model is described in Section 2.3.

2.1. Flat, Block Fading

When the assumptions of flat fading and block fading are made, the channel is approximated as a single constant value over an observation period. These two assumptions remove the channel dependence on ν and τ and h becomes a complex scalar. Equation (1) collapses to (2) where the channel is multiplied with the transmitted signal. To generate each channel realization, Clarke's sum of sinusoids narrowband model is used combining 200 components with uniform angle of arrival $[0, 2\pi)$.

$$r(\nu) = h \cdot x(\nu) + n(\nu) \quad (2)$$

The generation of this channel model is the average of the 200 components, each a circularly-symmetric $\mathcal{CN}(0, 1)$, representative of a Rayleigh flat fading environment. With every new trial, a new value for h is generated.

2.2. Selective, Block Fading

By removing the assumption of flat fading, the channel becomes a set of multiple paths from the transmitter to the receiver, with each path having independent amplitudes and phases. Using Turin's model with an arrival rate $\lambda = 3.55$ within an observation period of $56\mu s$, a channel is modeled with 4 total paths [13]. Each path has a Rayleigh magnitude with parameter $\sigma^2 = 0.05$ as was used in [4, 5, 7, 8] with a uniform random phase $\theta \sim U[0, 2\pi)$. Under a more idealized model, these four paths translate directly, $L = L_{paths}$ as each delay is exactly a sample delay. In this paper a more practical generation occurs where the L_{paths} have delays found using a Poisson process with Turin's parameters. These Poisson delays are then combined with the Rayleigh weighted taps. The taps are then sinc interpolated to the discrete sampling instances ν where the sinc pulse is truncated for significance such that the ripples whose magnitude is less than 0.01 are ignored. This is simulated for a narrowband signal to observe the difference between the selective model and the flat fading assumption of the first channel model approach. The symbol period $T_S = 156.25\mu s$ is greater than the expected maximum path delay, $T_S > 63\mu s$, for the Poisson process with Turin's parameters. The given model results in $L = 10$ on average and is represented as

$$r(\nu) = \sum_{\tau=0}^{L-1} h(\tau)x(\nu - \tau) + n(\nu) \quad (3)$$

where h is conditioned such that $\sum_{\tau=-L}^{L-1} |h(\tau)|^2 = 1$.

2.3. Flat, Doppler Spread Fading

The final channel model examined again assumes flat fading; however, the assumption of block fading is removed. Each sampling instant, ν , is the mean of 200 independent components with uniform angle of arrival $[0, 2\pi)$ utilizing Clarke's sum of sinusoids narrowband model. The received signal from the time varying channel model is given as

$$r(\nu) = h(\nu) \cdot x(\nu) + n(\nu) \quad (4)$$

The flat, Doppler spread channel model is examined with three different Doppler values: 5, 70, and 200 Hz. These values were chosen based on their use in the LTE specification.

3. CUMULANT BASED ALGORITHM BACKGROUND

In order to examine the benefit of the proposed algorithm, first background on the higher order cumulants is provided and one application that uses a single feature cumulant is given. The cumulant is a higher order statistic that was proposed by Swami and Sadler in [3] as a feature to be used in a hierarchical classification method by using the set of fourth order cumulants. Xi and Wu [5] then extended the residual channel effect, $\beta_{4,2}$, shown in [3] to the multipath channel model using $\beta_{4,2}$ directly to normalize the loss seen by the residual channel; however, the channel estimation technique used in [5] is shown to be unstable in [7]. Orlic and Dukic [7] propose a new channel estimation method and a modified sixth order cumulant, discussed in Section 3.2, for modulation classification.

3.1. Cumulant

The cumulant is defined as:

$$\kappa(\mathbf{r}_1, \dots, \mathbf{r}_n) \triangleq \sum_P (-1)^{N_p-1} (N_p - 1)! \sum_{b=1}^{N_B} \prod_{p=1}^{N_p} E \left[\prod_{i \in p} \mathbf{r}_i^{(*)} \right] \quad (5)$$

where P is a list of all integer partitions of n on the indices $i = 1, \dots, n$. Each partition is made up of a list of non-repeating nonempty blocks, B , of length N_B , where N_p is the number of parts, p , in the partition. A block, b , is one unique arrangement of indices separated into N_p parts. The expectation operation is given as $E[\mathbf{r}] \approx \frac{1}{N} \sum_{\nu=0}^{N-1} r(\nu)$ and assumes that \mathbf{r} is ergodic. By splitting the cumulant order, n , into $n - q$ copies of the samples and q copies of sample conjugates the notation $\kappa_{n,q}$ is used to represent the cumulant of order n with q conjugates. That is

$$\mathbf{r}_i^{(*)} = \begin{cases} \mathbf{r}, & 1 \leq i \leq n - q \\ \mathbf{r}^*, & n - q < i \leq n \end{cases} \quad (6)$$

The cumulant $\kappa_{3,0}$ is given below as an example.

$$\begin{aligned}\kappa_{3,0} = & (-1)^{1-1}(1-1)! \cdot \{\langle(m_{3,0})\rangle\} \\ & + (-1)^{2-1}(2-1)! \cdot \{\langle(m_{2,0})(m_{1,0})\rangle + \\ & \langle(m_{2,0})(m_{1,0})\rangle + \langle(m_{2,0})(m_{1,0})\rangle\} \\ & + (-1)^{3-1}(3-1)! \cdot \{\langle(m_{1,0})(m_{1,0})(m_{1,0})\rangle\}\end{aligned}$$

where $\{\cdot\}$ is a partition, $\langle\cdot\rangle$ is a block within the block list for the partition, and (\cdot) is a part in the block. The use of $m_{n,q}$ is shorthand for the moment $E[r^{n-q}r^q]$. Without the descriptive bracket notation used, the equation simplifies to.

$$\kappa_{3,0} = m_{3,0} - 3m_{2,0}m_{1,0} + 2 \cdot m_{1,0}^3$$

When fourth order cumulants were proposed by Swami and Sadler [3] they provided means to account for unknown amplitude and phase of the received signal caused by the channel effects. The first adjustment normalizes all higher order cumulants in such a way that they are energy neutral when used. This energy normalization is given by

$$\tilde{\kappa}_{n,q} = \frac{\kappa_{n,q}}{(\kappa_{2,1} - \sigma_n^2)^{n/2}} \quad (7)$$

where $\tilde{\kappa}_{n,q}$ represents the energy normalized cumulant value. For ideal performance at lower Signal to Noise Ratio (SNR) the noise power, σ_n^2 , is needed to properly normalize the higher order cumulants. In this paper the noise power is assumed known as it has been shown to be blindly estimated given a long enough observation in [14, 15]. To account for unknown phase in the received signal, Swami and Sadler proposed the use of the absolute value of the cumulant, this is employed here even at the potential cost of loss of separation seen for certain modulation comparisons [3]. The normalization of the cumulant is denoted

$$\hat{\kappa}_{n,q} = |\tilde{\kappa}_{n,q}| \quad (8)$$

where $\hat{\kappa}_{n,q}$ is the magnitude of the cumulant value from (7).

For reduced computation of the cumulant values, and by using the property that the received signals are assumed to be zero-mean random variables, any moment that is multiplied by $E[r]$ is assumed to be zero and removed from the computation. Further making use of the symmetric placement of symbols on the IQ plane, any moment of an odd power, $E[r^{x_{odd}}]$, is assumed to be zero. These assumptions allow for higher order cumulants to be calculated under ideal conditions perfectly with reduced computation.

3.2. Single Feature Approach

The single feature approach used as a comparison for this paper's proposed algorithm is the algorithm proposed by Orlic and Dukic [7]. The cumulant feature they proposed to use is a modified version of $\tilde{\kappa}_{6,3}$ before applying the channel normalization.

The difference between $\tilde{\kappa}_{6,3}$ and $\check{\kappa}_{6,3}$ used by Orlic and Dukic is given by

$$\tilde{\kappa}_{6,3} = \frac{m_{6,3} - 6\mathcal{R}\{m_{2,0}m_{4,1}^*\} - 9m_{4,2}m_{2,1} + 18|m_{2,0}|^2m_{2,1} + 12m_{2,1}^3}{(\kappa_{2,1} - \sigma_n^2)^3} \quad (9)$$

$$\check{\kappa}_{6,3} = \frac{m_{6,3} - 9m_{4,2}m_{2,1} + 12|m_{2,0}|^2m_{2,1} + 12m_{2,1}^3}{(\kappa_{2,1} - \sigma_n^2)^3} \quad (10)$$

where $\mathcal{R}\{\cdot\}$ is the real of the complex value. The modified cumulant removes the phase dependence, $\mathcal{R}\{m_{2,0}m_{4,1}^*\}$, from the sixth order cumulant.

Before the modified cumulant $\check{\kappa}_{6,3}$ is used for classification it first must be normalized with a $\beta_{6,3}$ factor defined as

$$\beta_{6,3} = \frac{\sum_{l=0}^{L-1} |h(l)|^6}{\left(\sum_{l=0}^{L-1} |h(l)|^2\right)^3} \quad (11)$$

from [7].

The cumulant estimator is then defined as

$$\check{\kappa}_{6,3} = \frac{1}{\beta_{6,3}} \check{\kappa}_{6,3}. \quad (12)$$

This estimator, $\check{\kappa}_{6,3}$, is then compared to the ideal cumulant value for each waveform being considered. The modified cumulant in (10) has a different value from the true $\kappa_{6,3}$ found from (9) for PAM modulations of order four and greater, but otherwise has same value as the true $\kappa_{6,3}$ for the remaining modulations considered in this paper.

As the generation of $\check{\kappa}_{6,3}$ relies on $\beta_{6,3}$, which in turn relies on the channel taps, the channel taps must be estimated from the received signal. The relative channel estimation procedure is defined as

$$\hat{h}(k) = \frac{m_{4,0}(f, f, f, k)}{m_{4,0}(f, f, f, f)}, \quad f, k = 0, \dots, L_R - 1 \quad (13)$$

where L_R is the estimated number of taps, and $m_{4,0}(f, f, f, k)$ is defined by

$$\begin{aligned}m_{4,0}(f, f, f, k) &= E[r(\nu - k) \cdot r(\nu - f)^3] \\ &= \left(\sum_{l=0}^{L_R-1} h^3(l)h(k - f + l) \right) E[x^4] \\ &\quad + \text{Res}(f, f, f, k)\end{aligned} \quad (14)$$

with the $\text{Res}(f, f, f, k) \approx 0$ the estimated values of the taps are given in [7] as

$$\hat{h}(k) \approx \frac{h(k - f + m)}{h(m)} \quad (15)$$

where $h(m)$ corresponds to the strongest tap and the desired value of $f = m$. As the strongest tap is not known *a priori* the

channel must be estimated L_R times in order to have the best normalized channel estimate. Then in order to determine which value of f should be used, this procedure is repeated N_E times. Whichever value from the $L_R \times N_E$ matrix of estimates minimizes the standard deviation of the cumulant estimator, $\hat{\kappa}_{6,3}$, from the ideal cumulant values over N_E is chosen as f . This means a second order minimization must occur across all L_R taps, N_E estimates, and $|\mathcal{M}|$ modulations to be considered. For simplicity in this paper $N_E = 1$ and the classification results in the minimization of the residual of the estimate from the ideal, and the modulation is chosen that achieves this minimum from an $|\mathcal{M}| \times L_R$ matrix.

4. ALGORITHM DEVELOPMENT

This paper's contribution is the Waveform Signature (WS), which makes use of the base application proposed by Swami and Sadler and seeks to circumvent the run time channel estimation needed by both Xi's and Wu's, and Orlic's and Dukic's algorithms. Unlike using a single feature for classification, or using a set of features individually or jointly in a hierarchical decision tree, the proposed algorithm makes use of the entire set of features for classification. The set of features used in this algorithm are the cumulants given in (16), a 20-dimensional vector. The equations for generating each cumulant value prior to energy and phase normalization in (7) and (8) respectively are given in Appendix A. This set of features is denoted as a Waveform Signature (WS) and is the estimator used for modulation classification.

$$\begin{aligned} \mathbf{ws} = & [\hat{\kappa}_{2,0}, \hat{\kappa}_{2,1}, \hat{\kappa}_{4,0}, \hat{\kappa}_{4,1}, \hat{\kappa}_{4,2}, \hat{\kappa}_{6,0}, \hat{\kappa}_{6,1}, \hat{\kappa}_{6,2}, \hat{\kappa}_{6,3}, \\ & \hat{\kappa}_{8,0}, \hat{\kappa}_{8,1}, \hat{\kappa}_{8,2}, \hat{\kappa}_{8,3}, \hat{\kappa}_{8,4}, \\ & \hat{\kappa}_{10,0}, \hat{\kappa}_{10,1}, \hat{\kappa}_{10,2}, \hat{\kappa}_{10,3}, \hat{\kappa}_{10,4}, \hat{\kappa}_{10,5}] \end{aligned} \quad (16)$$

To emphasize the use of the WS for modulation classification, databases are made for each modulation of interest on every considered channel. These databases are denoted $D_{M,C}$ where M is the modulation in the database and C is the channel that the database was collected on. This represents a supervised learning database for the signals of interest within the operation environment. Each database is a collection of 2,000 waveforms on an AWGN channel with 20 dB E_s/N_o . For use in classification, a simple approach takes the average of all WS within the database and creates a single signature, $\mathbf{ws}_{M,C} = \frac{1}{2000} \sum_{\mathbf{ws} \in D_{M,C}} \mathbf{ws}$, to compare against for each unknown waveform, $\hat{\mathbf{ws}}$. Any unknown WS is classified by (17).

$$\hat{M} = \min_{M \in \mathcal{M}} \|\mathbf{ws}_{M,C} - \hat{\mathbf{ws}}\|_1 \quad (17)$$

where \mathcal{M} contains all modulations being considered and $\|\cdot\|_1$ is the vector 1-norm.

By taking ideal captures of the modulations under consideration, $h = 1, \sigma_n^2 = 0$, such that the only variance is caused

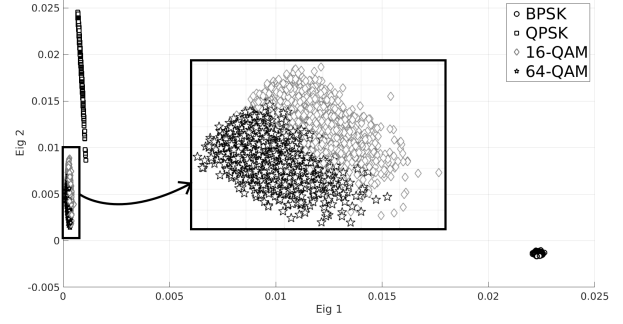


Figure 1: Principle Component Analysis in 2-Dimensions of the Waveform Signatures as seen on 20 dB E_s/N_o AWGN Channels for BPSK, QPSK, 16QAM, and 64QAM. Subplot shows 16QAM and 64QAM when zoomed in.

by using the sample mean instead of the exact moments of each waveform. The ideal database is then generated by

$$\bar{D}_{\mathcal{M}, Ideal} = [D_{M_1, Ideal}^T, D_{M_2, Ideal}^T, \dots, D_{|\mathcal{M}|, Ideal}^T]^T \quad (18)$$

where $|\mathcal{M}|$ is the number of modulations being considered and $(\cdot)^T$ is the transpose. Performing Principle Component Analysis (PCA) on the ideal database and using the greatest ρ component vectors a reduction in dimensions can be performed on the WS. The reduction matrix is defined in (19).

$$W^{\{\rho\}} = [\mathbf{w}_{(1)}, \dots, \mathbf{w}_{(\rho)}] \quad (19)$$

Each $\mathbf{w}_{(i)}$ is the i^{th} loading column vector of the PCA on the ideal database in (18). The reduced WS is then found in (20).

$$\mathbf{ws}^{\{\rho\}} = \mathbf{ws} \cdot W^{\{\rho\}} \quad (20)$$

To illustrate this reduction, the WS of four modulations {BPSK, QPSK, 16-QAM, 64-QAM} on a AWGN channel are shown in Figure 1. The subplot zooms into the region around the 16-QAM and 64-QAM data points to show the separation between the two QAM modulations. The PCA was performed on the ideal database that consists of 14 modulations, {BPSK, QPSK, 8PSK, 16PSK, 4PAM, 8PAM, 16PAM, 4QAM, 8QAM, 16QAM, 32QAM, 64QAM, 128QAM, 256QAM}.

Modulation classification can be performed on $\mathbf{ws}^{\{\rho\}}$ by projecting the database into the reduced space $D_{M,C}^{\{\rho\}} = D_{M,C} \cdot W^{\{\rho\}}$ and then following the same procedure of finding the average reduced WS, $\mathbf{ws}_{M,C}^{\{\rho\}} = \frac{1}{2000} \sum_{\mathbf{ws} \in D_{M,C}} \mathbf{ws} \cdot W^{\{\rho\}}$, and minimizing the unknown reduced WS, $\hat{\mathbf{ws}}^{\{\rho\}}$, as shown in (21).

$$\hat{M} = \min_{M \in \mathcal{M}} \|\mathbf{ws}_{M,C}^{\{\rho\}} - \hat{\mathbf{ws}}^{\{\rho\}}\|_1 \quad (21)$$

It should be noted that this reduction of dimensions does come at a small performance loss at low SNR as not all of the variance is contained in the first few components; however, by filtering out less useful components, a greater performance can be

seen at higher SNR. This difference will be examined in the next section. Using the reduction matrix of the idealized 14 modulations, $D_{(14),C}^{\{\rho\}}$, when only a subset of the modulations are being considered, $D_{(\mathcal{M}),C}^{\{\rho\}}$, will show a greater performance loss and this will be examined in the next section for the most notable case.

5. PERFORMANCE ANALYSIS

Simulations are conducted to analyze the performance of the algorithms discussed in the paper by comparing the average probability of correct classification, P_{cc} , as defined in (22).

$$P_{cc} = \sum_{i=1}^{|\mathcal{M}|} P(\widehat{M} = M_i | M_i) P(M_i) \quad (22)$$

where $P(M_i) = \frac{1}{|\mathcal{M}|} \forall i$, and $P(\widehat{M} = M_i | M_i)$ is the probability that M_i is the classification choice given that M_i was transmitted.

Three sets of modulations are considered, which have commonly been used in AMC literature for observing the AMC performance. The first set, Ω_1 , consists of {BPSK, QPSK} [5–7]. The second set, Ω_2 , consists of {QPSK, 16QAM, 64QAM} [5–8], while the third set, Ω_3 consists of {BPSK, QPSK, 8PSK, 4PAM, 16QAM}, which is a combination of sets from [3, 8, 11]. These sets are examined on the three channel models discussed in Section 2 for narrowband signals: Flat, Block Fading; Selective, Block Fading; and Flat, Doppler Spread Fading.

For all the simulations the block length, N_b , is 1920 samples which means that the number of symbols, N_s , is 640 since there are three samples per symbol. The estimated tap length, L_R , is set to 10 in [7], which is the average tap length seen in the Turin model. This is chosen because the considered channel models are all being used to examine the P_{cc} on similar narrowband channel models. The P_{cc} performance is the average of 2,000 Monte Carlo trials. The reduced waveform uses the first three loading vectors, $\rho = 3$ from (20), of the PCA for classification in this section. A summary of the performance is given in Section 5.4.

5.1. Flat, Block Fading

The first channel to test the WS upon is the standard Flat, Block Fading channel model. This uses a scalar Rayleigh random variable with parameter, σ_h , equal to $\sqrt{0.5}$ and is generated uniquely for each simulation run. Figure 2 shows the performance of the proposed algorithm in comparison to the algorithm in [7] for modulation space Ω_1 . The three performance curves are for the full WS, ws (square), the reduced WS utilizing the 14-modulation database from section 4, $ws_{(14),C}^{\{3\}}$ (o), and the algorithm in [7], $\kappa_{6,3}$ (dashed). At low SNR values there is a performance loss from using the reduced WS when compared to the full WS, but beyond 1dB SNR the loss is minimal. In contrast

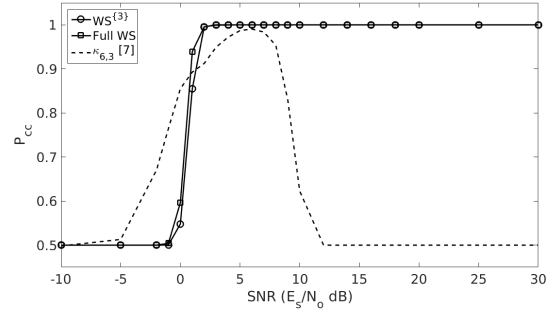


Figure 2: The P_{cc} using modulation set Ω_1 on a Flat, Block Fading channel using the Clarke model.

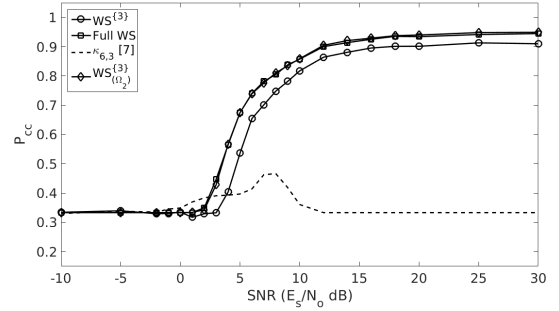


Figure 3: The P_{cc} using modulation set Ω_2 on a Flat, Block Fading channel using the Clarke model.

the algorithm from [7] has better performance for SNR values ≤ 0 dB.

For the modulation space Ω_2 , the effect of using the 14-modulation database, $D_{(14),C}^{\{3\}}$ (o), versus the 3-modulation database, $D_{(\Omega_2),C}^{\{3\}}$ (diamond), has a significant difference in classification performance, shown in Figure 3. This difference is due to the separation between 16QAM and 64QAM being a difficult task for cumulants. The difficulty is addressed in [3] when a conservative estimate of the number of symbols needed to reach 90% accuracy was $>10,000$ symbols. To reach the same accuracy for most other 2-case modulation classifications, <300 symbols are required to achieve the same threshold. By only considering the modulations to be examined in the PCA reduction, more degrees of freedom can be applied to their separation than when all modulation are considered. Using the 14-modulation database to do the WS reduction results in approximately 1-2 dB loss in performance along with a 3.5% decrease in maximum classification accuracy.

Examining Ω_3 shows that the reduced WS results in a 0.5 dB loss in performance for a more diverse modulation set, but achieves equal performance to the full WS for $\text{SNR} \geq 9$ dB.

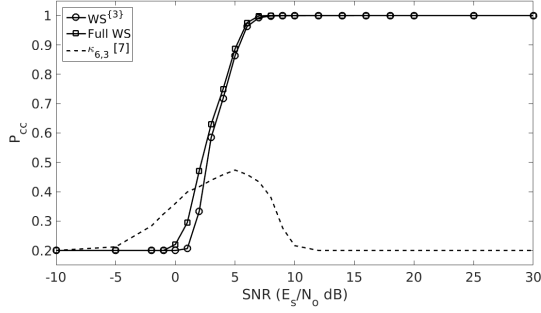


Figure 4: The P_{cc} using modulation set Ω_3 on a Flat, Block Fading channel using the Clarke model.

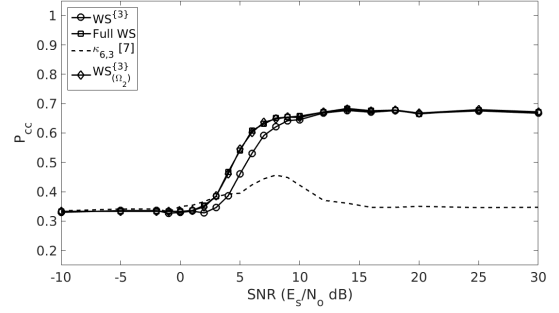


Figure 6: The P_{cc} using modulation set Ω_2 on a Selective, Block Fading channel using the Turin model.

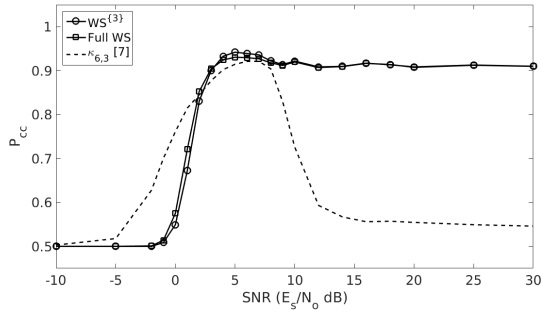


Figure 5: The P_{cc} using modulation set Ω_1 on a Selective, Block Fading channel using the Turin model.

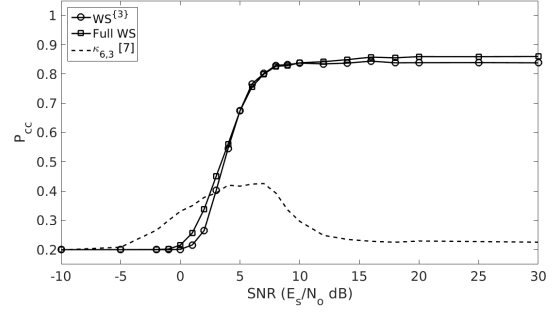


Figure 7: The P_{cc} using modulation set Ω_3 on a Selective, Block Fading channel using the Turin model.

5.2. Selective, Block Fading

The second channel model examines the effect of frequency correlation in the channel. The performance of the different modulations sets, Ω_1 , Ω_2 , and Ω_3 , are shown in Figures 5, 6, and 7 respectively. For Ω_1 there is a loss in performance over the Flat fading channel, but 90% accuracy is reached and maintained at 3dB SNR and maintained using both the full and reduced WS. In Figure 6 the difference between using $D_{(14),C}^{\{3\}}$ and $D_{(\Omega_2),C}^{\{3\}}$ in the reduced WS shows as a 1 dB loss in performance; however, there is no longer a reduction of the maximum classification achieved by using $D_{(14),C}^{\{3\}}$. For Ω_3 there is a loss in performance for the reduced WS when compared to the full WS.

5.3. Flat, Doppler Spread Fading

The first two channel models assumed that the channel is constant for the duration of the observation, which is reasonable for relatively stationary links; however, the channel changes over time when there is relative motion in the link. The third channel examines the effect the time correlation of the channel has on the performance of the system. Two maximum Doppler values examined here are 70 Hz, and 200 Hz, which are also used in the LTE standard's channel models. The third maximum Doppler value, 5Hz, is omitted here for brevity due to performance being

relatively close to that seen in the Flat, Block Fading figures.

5.3.3. Maximum Doppler 70Hz

In Figure 8 the effect on the modulation set Ω_1 is seen to reduce the performance of the WS classification by 2% but otherwise has similar performance to Figure 2. Ω_2 in Figure 9 shows that the reduced WS using $D_{(\Omega_2),C}^{\{3\}}$ results in a single percentage greater maximum performance. For Ω_3 in Figure 10, much like in Ω_1 , there is reduction in the maximum performance observed by 5%.

5.3.3. Maximum Doppler 200Hz

As the maximum Doppler increases, the channel changes faster in time and causes a greater reduction in the performance of the WS. Figure 11 is the only condition that was tested where the algorithm in [7] exceeds the performance of the WS at higher SNR once the WS has outperformed [7]. The performance of the WS for Ω_1 drops to the lowest of 90% accuracy in this channel. The performance Ω_2 is shown in Figure 12. Using the reduced WS from $D_{(14),C}^{\{3\}}$ reaches a greater maximum accuracy at higher SNR than either the full WS or the reduced WS with the more selective database; however, a right shift is still seen at low SNR. Figure 13 shows a reduction in performance for modulation set Ω_3 to 82% with the reduced WS slightly outperforming the full WS.

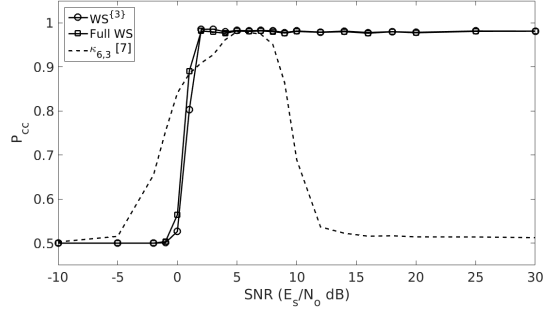


Figure 8: The P_{cc} using modulation set Ω_1 on a Flat, Doppler Spread Fading channel using the Clarke model with maximum Doppler of 70Hz (Clarke70).

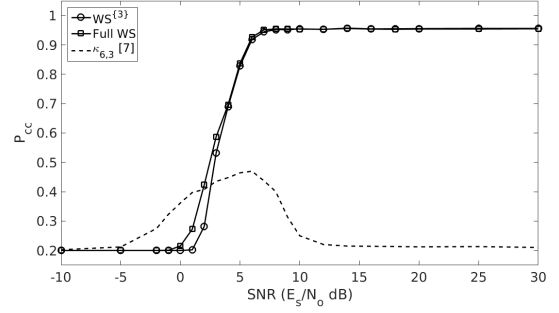


Figure 10: The P_{cc} using modulation set Ω_3 on a Flat, Doppler Spread Fading channel using the Clarke model with maximum Doppler of 70Hz (Clarke70).

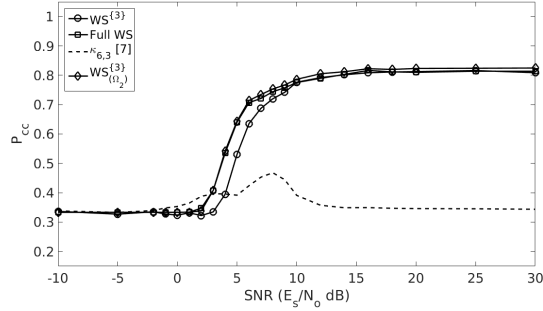


Figure 9: The P_{cc} using modulation set Ω_2 on a Flat, Doppler Spread Fading channel using the Clarke model with maximum Doppler of 70Hz (Clarke70).

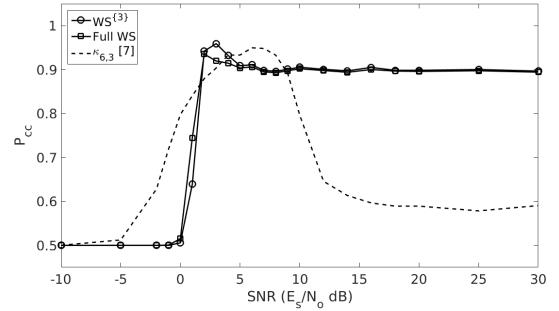


Figure 11: The P_{cc} using modulation set Ω_1 on a Flat, Doppler Spread Fading channel using the Clarke model with maximum Doppler of 200Hz (Clarke200).

Table 1: Performance Summary at 5 dB SNR

Channel Model	Ω	$P_{cc} @ 5\text{dB} (\%)$			
		ws	$ws_{(14)}^{(3)}$	$ws_{(\Omega)}^{(3)}$	$\ddot{\kappa}_{6,3}$
Clarke	Ω_1	100	100	-	98.7
Clarke	Ω_2	67.5	53.6	67.5	39.7
Clarke	Ω_3	88.6	86.3	-	47.4
Turin	Ω_1	93.1	94.2	-	91.5
Turin	Ω_2	54.0	46.1	54.5	39.5
Turin	Ω_3	67.4	67.5	-	41.7
Clarke5	Ω_1	100	100	-	98.5
Clarke5	Ω_2	67.4	54.0	67.3	39.5
Clarke5	Ω_3	89.4	86.8	-	46.8
Clarke70	Ω_1	98.2	98.3	-	98.0
Clarke70	Ω_2	64.0	53.1	64.4	39.2
Clarke70	Ω_3	83.8	82.9	-	46.4
Clarke200	Ω_1	90.4	91.0	-	93.3
Clarke200	Ω_2	57.5	49.0	57.8	39.7
Clarke200	Ω_3	82.1	84.0	-	26.8

5.4. Performance Summary

The P_{cc} performance, which is shown in Figures 2 - 13, is displayed in Tables 1, 2, and 3 for SNR values 5, 10, 16 dB respectively. The performance on the Clarke model with a 5Hz maximum Doppler is shown as well.

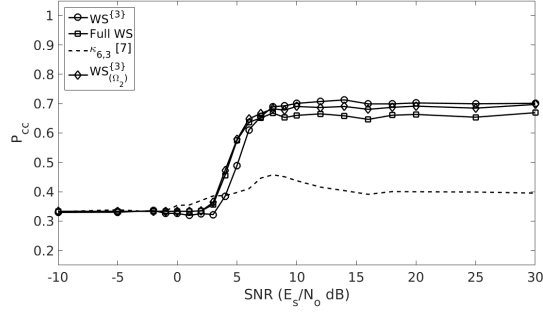


Figure 12: The P_{cc} using modulation set Ω_2 on a Flat, Doppler Spread Fading channel using the Clarke model with maximum Doppler of 200Hz (Clarke200).

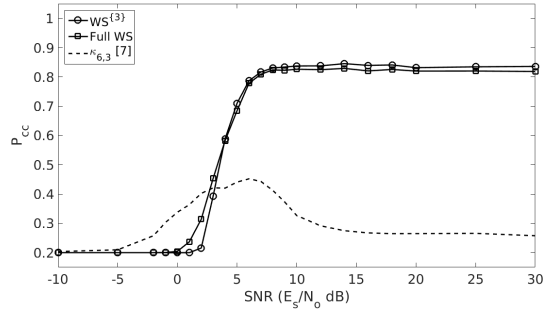


Figure 13: The P_{cc} using modulation set Ω_3 on a Flat, Doppler Spread Fading channel using the Clarke model with maximum Doppler of 200Hz (Clarke200).

Table 2: Performance Summary at 10dB SNR

Channel Model	Ω	$P_{cc} @ 10\text{dB} (\%)$			
		ws	$ws_{(14)}^{(3)}$	$ws_{(\Omega)}^{(3)}$	$\ddot{\kappa}_{6,3}$
Clarke	Ω_1	100	100	-	62.3
Clarke	Ω_2	85.8	81.7	85.9	36.1
Clarke	Ω_3	100	100	-	21.7
Turin	Ω_1	92.0	92.2	-	72.7
Turin	Ω_2	65.8	64.5	65.3	42.2
Turin	Ω_3	83.8	83.7	-	29.7
Clarke5	Ω_1	100	100	-	62.9
Clarke5	Ω_2	85.5	81.9	86.3	35.8
Clarke5	Ω_3	100	100	-	21.4
Clarke70	Ω_1	98.1	98.2	-	69.0
Clarke70	Ω_2	77.5	77.6	78.6	39.1
Clarke70	Ω_3	95.4	95.4	-	25.0
Clarke200	Ω_1	90.2	90.6	-	79.7
Clarke200	Ω_2	66.0	70.1	69.0	43.8
Clarke200	Ω_3	82.6	83.8	-	32.7

Table 3: Performance Summary at 16dB SNR

Channel Model	Ω	$P_{cc} @ 16\text{dB} (\%)$			
		ws	$ws_{(14)}^{(3)}$	$ws_{(\Omega)}^{(3)}$	$\ddot{\kappa}_{6,3}$
Clarke	Ω_1	100	100	-	50.0
Clarke	Ω_2	92.5	89.6	92.9	33.3
Clarke	Ω_3	100	100	-	20.0
Turin	Ω_1	91.7	91.7	-	55.6
Turin	Ω_2	67.6	67.0	67.5	34.7
Turin	Ω_3	85.7	84.4	-	22.9
Clarke5	Ω_1	100	100	-	50.0
Clarke5	Ω_2	92.4	89.8	93.3	33.3
Clarke5	Ω_3	100	100	-	20.0
Clarke70	Ω_1	97.7	97.8	-	51.6
Clarke70	Ω_2	81.7	80.9	82.2	34.9
Clarke70	Ω_3	95.4	95.5	-	21.4
Clarke200	Ω_1	90.0	90.1	-	59.7
Clarke200	Ω_2	64.6	69.9	68.1	39.1
Clarke200	Ω_3	68.5	71.0	-	44.1

6. CONCLUSION

In this paper an algorithm was developed for AMC using multiple higher order cumulant values in a Waveform Signature (WS). This algorithm outperforms the single cumulant estimator from [7] that relies on channel estimation techniques to recover ideal cumulant values at higher SNR levels. The WS can be classified with high accuracy by minimizing L1 norm distance and without relying on more complex classification techniques like SVM or decision trees. The WS can be used to track how the spectrum is being used without having to record raw IQ samples and is instead represented as only 20 floats. While this WS was used in a supervised learning scenario for classification purposes, it is possible to use the WS in an unsupervised sense where clustering techniques could be used to classify unknown captures. Future work will consider the effect of channel mismatch on the classification accuracy, and blind channel selection based on the observed WS.

REFERENCES

- [1] O. Dobre, A. Abdi, Y. Bar-Ness, and W. Su, "Survey of automatic modulation classification techniques: classical approaches and new trends," *Communications, IET*, vol. 1, no. 2, pp. 137–156, April 2007.
- [2] W. Wei and J. Mendel, "Maximum-likelihood classification for digital amplitude-phase modulations," *Communications, IEEE Transactions on*, vol. 48, no. 2, pp. 189–193, Feb 2000.
- [3] A. Swami and B. Sadler, "Hierarchical digital modulation classification using cumulants," *Communications, IEEE Transactions on*, vol. 48, no. 3, pp. 416–429, Mar 2000.

- [4] L. Liu and J. Xu, "A novel modulation classification method based on high order cumulants," in *Wireless Communications, Networking and Mobile Computing, 2006. WiCOM 2006. International Conference on*, Sept 2006, pp. 1–5.
- [5] S. Xi and H.-C. Wu, "Robust automatic modulation classification using cumulant features in the presence of fading channels," in *Wireless Communications and Networking Conference, 2006. WCNC 2006. IEEE*, vol. 4, April 2006, pp. 2094–2099.
- [6] H.-C. Wu, M. Saquib, and Z. Yun, "Novel automatic modulation classification using cumulant features for communications via multipath channels," *Wireless Communications, IEEE Transactions on*, vol. 7, no. 8, pp. 3098–3105, August 2008.
- [7] V. Orlic and M. Dukic, "Automatic modulation classification: Sixth-order cumulant features as a solution for real-world challenges," in *Telecommunications Forum (TELFOR), 2012 20th*, Nov 2012, pp. 392–399.
- [8] P. Liu and P.-L. Shui, "A new cumulant estimator in multipath fading channels for digital modulation classification," *Communications, IET*, vol. 8, no. 16, pp. 2814–2824, 2014.
- [9] A. E. Sherme, "A novel method for automatic modulation recognition," *Applied Soft Computing*, vol. 12, no. 1, pp. 453 – 461, 2012. [Online]. Available: <http://www.sciencedirect.com/science/article/pii/S1568494611003085>
- [10] Y. e Wang, T. qi Zhang, J. Bai, and R. Bao, "Modulation recognition algorithms for communication signals based on particle swarm optimization and support vector machines," in *Intelligent Information Hiding and Multimedia Signal Processing (IIH-MSP), 2011 Seventh International Conference on*, Oct 2011, pp. 266–269.
- [11] W. Headley, V. Chavali, and C. da Silva, "Maximum-likelihood modulation classification with incomplete channel information," in *Information Theory and Applications Workshop (ITA), 2013*, Feb 2013, pp. 1–4.
- [12] R. Clarke, "A statistical theory of mobile-radio reception," *Bell System Technical Journal, The*, vol. 47, no. 6, pp. 957–1000, July 1968.
- [13] G. Turin, F. Clapp, T. Johnston, S. Fine, and D. Lavry, "A statistical model of urban multipath propagation," *Vehicular Technology, IEEE Transactions on*, vol. 21, no. 1, pp. 1–9, Feb 1972.
- [14] J. Vartiainen, H. Saarnisaari, J. Lehtomaki, and M. Juntti, "A blind signal localization and snr estimation method," in *Military Communications Conference, 2006. MILCOM 2006. IEEE*, Oct 2006, pp. 1–7.
- [15] H. Xiao, Y. Shi, W. Su, and J. Kosinski, "An investigation of non-data-aided snr estimation techniques for analog modulation signals," in *Sarnoff Symposium, 2010 IEEE*, April 2010, pp. 1–5.

A CUMULANT GENERATION

Table 4: The WS Cumulants, $\kappa_{n,q}$ part 1

1	$\kappa_{2,0} = m_{2,0}$
2	$\kappa_{2,1} = m_{2,1}$
3	$\kappa_{4,0} = m_{4,0} - (3m_{2,0}^2)$
4	$\kappa_{4,1} = m_{4,1} - (3m_{2,0}m_{2,1})$
5	$\kappa_{4,2} = m_{4,2} - (m_{2,0} ^2 + 2m_{2,1}^2)$
6	$\kappa_{6,0} = m_{6,0} - (15m_{2,0}m_{4,0}) + 2 \cdot (15m_{2,0}^3)$
7	$\kappa_{6,1} = m_{6,1} - (10m_{2,0}m_{4,1} + 5m_{2,1}m_{4,0}) + 2 \cdot (15m_{2,0}^2m_{2,1})$
8	$\kappa_{6,2} = m_{6,2} - (6m_{2,0}m_{4,2} + 8m_{2,1}m_{4,1} + m_{2,0}^*m_{4,0}) + 2 \cdot (3 m_{2,0} ^2m_{2,0} + 12m_{2,1}^2m_{2,0})$
9	$\kappa_{6,3} = m_{6,3} - (3m_{2,0}m_{4,1}^* + 3m_{2,0}^*m_{4,1} + 9m_{2,1}m_{4,2}) + 2 \cdot (9 m_{2,0} ^2m_{2,1} + 6m_{2,1}^3)$
10	$\kappa_{8,0} = m_{8,0} - (28m_{2,0}m_{6,0} + 35m_{2,0}^2m_{4,0}) + 2 \cdot (210m_{2,0}^2m_{2,0}) - 6 \cdot (105m_{2,0}^4)$
11	$\kappa_{8,1} = m_{8,1} - (21m_{2,0}m_{6,1} + 7m_{2,0}m_{6,2} + 35m_{2,0}m_{6,0}) + 2 \cdot (105m_{2,0}m_{6,0} + 105m_{2,0}m_{6,2}) - 6 \cdot (105m_{2,0}^2m_{2,1})$
12	$\kappa_{8,2} = m_{8,2} - (15m_{2,0}m_{6,2} + 12m_{2,1}m_{6,1} + m_{2,0}^*m_{6,0} + 15m_{4,0}m_{4,2} + 20m_{4,1}^2) + 2 \cdot (45m_{2,0}^2m_{4,2} + 120m_{2,0}m_{2,1}m_{4,1} + 15 m_{2,0} ^2m_{4,0} + 30m_{2,1}^2m_{4,0}) - 6 \cdot (15m_{2,0}^2 m_{2,0} ^2 + 90m_{2,0}^2m_{2,1}^2)$
13	$\kappa_{8,3} = m_{8,3} - (10m_{2,0}m_{6,3} + 15m_{2,1}m_{6,2} + 3m_{2,0}^*m_{6,1} + 5m_{4,0}m_{4,1}^* + 30m_{4,1}m_{4,2}) + 2 \cdot (15m_{2,0}^2m_{4,1}^* + 90m_{2,0}m_{2,1}m_{4,2} + 30 m_{2,0} ^2m_{4,1} + 15m_{2,0}^*m_{2,1}m_{4,0} + 60m_{2,1}^2m_{4,1}) - 6 \cdot (45 m_{2,0} ^2m_{2,0}m_{2,1} + 60m_{2,0}m_{2,1}^3)$
14	$\kappa_{8,4} = m_{8,4} - (6m_{2,0}m_{6,2}^* + 16m_{2,1}m_{6,3} + 6m_{2,0}^*m_{6,2} + m_{4,0} ^2 + 16 m_{4,1} ^2 + 18m_{4,2}^2) + 2 \cdot (3m_{2,0}^3m_{4,0}^* + 48m_{2,0}m_{2,1}m_{4,1}^* + 36 m_{2,0} ^2m_{4,2} + 72m_{2,1}^2m_{4,2} + 48m_{2,0}^*m_{2,1}m_{4,1} + 3(m_{2,0}^*)^2m_{4,0}) - 6 \cdot (9 m_{2,0} ^4 + 72 m_{2,0} ^2m_{2,1}^2 + 24m_{2,1}^4)$

Table 5: The WS Cumulants, $\kappa_{n,q}$ part 2

15	$\kappa_{10,0} = m_{10,0} - (45m_{2,0}m_{8,0} + 210m_{4,0}m_{6,0})$ $+ 2 \cdot (630m_{2,0}^2m_{6,0} + 1575m_{2,0}m_{4,0}^2)$ $- 6 \cdot (3150m_{2,0}^3m_{4,0}) + 24 \cdot (945m_{2,0}^5)$
16	$\kappa_{10,1} = m_{10,1} - (36m_{2,0}m_{8,1} + 9m_{2,1}m_{8,0})$ $+ 126m_{4,0}m_{6,1} + 84m_{4,1}m_{6,0})$ $+ 2 \cdot (378m_{2,0}^2m_{6,1} + 252m_{2,0}m_{2,1}m_{6,0})$ $+ 1260m_{2,0}m_{4,0}m_{4,1} + 315m_{2,1}m_{4,0}^2)$ $- 6 \cdot (1260m_{2,0}^3m_{4,1}$ $+ 1890m_{2,0}^2m_{2,1}m_{4,0})$ $+ 24 \cdot (945m_{2,0}^4m_{2,1})$
17	$\kappa_{10,2} = m_{10,2} - (28m_{2,0}m_{8,2} + 70m_{4,0}m_{6,2})$ $+ 16m_{2,1}m_{8,1} + 112m_{4,1}m_{6,1}$ $+ m_{2,0}^*m_{8,0} + 28m_{4,2}m_{6,0})$ $+ 2 \cdot (210m_{2,0}^2m_{6,2}$ $+ 420m_{2,0}m_{4,0}m_{4,2} + 28 m_{2,0} ^2m_{6,0}$ $+ 35m_{2,0}^*m_{4,0}^2 + 560m_{2,1}m_{4,0}m_{4,1}$ $+ 56m_{2,1}^2m_{6,0} + 336m_{2,0}^2m_{2,1}m_{4,1}$ $+ 560m_{2,0}m_{4,1}^2)$ $- 6 \cdot (1680m_{2,0}^2m_{2,1}m_{6,1}$ $+ 210m_{2,0} m_{2,0} ^2m_{4,0}$ $+ 420m_{2,0}^3m_{4,2} + 840m_{2,0}m_{2,1}^2m_{4,0})$ $+ 24 \cdot (105m_{2,0}^3 m_{2,0} ^2 + 840m_{2,0}^3m_{2,1}^2)$
18	$\kappa_{10,3} = m_{10,3} - (21m_{2,0}m_{8,3} + 35m_{4,0}m_{6,3})$ $+ 7m_{4,1}^*m_{6,0} + 3m_{2,0}^*m_{8,1} + 63m_{4,2}m_{6,1})$ $+ 21m_{2,1}m_{8,2} + 105m_{4,1}m_{6,2})$ $+ 2 \cdot (105m_{2,0}^2m_{6,3} + 315m_{2,0}m_{2,1}m_{6,2}$ $+ 63 m_{2,0} ^2m_{6,1} + 126m_{2,1}^2m_{6,1}$ $+ 21m_{2,0}^*m_{2,0}m_{6,0} + 105m_{2,0}m_{4,0}m_{4,1}^*$ $+ 630m_{2,0}m_{4,1}m_{4,2} + 315m_{2,1}m_{4,0}m_{4,2}$ $+ 420m_{2,1}m_{4,1}^2 + 105m_{2,0}^*m_{4,0}m_{4,1})$ $- 6 \cdot (206m_{2,0}^3m_{4,1}^* + 945m_{2,0}^2m_{2,1}m_{4,2}$ $+ 315m_{2,0} m_{2,0} ^2m_{4,1}$ $+ 1260m_{2,0}m_{2,1}^2m_{4,1} + 315 m_{2,0} ^2m_{2,1}m_{4,0}$ $+ 210m_{2,1}^3m_{4,0})$ $+ 24 \cdot (315m_{2,0}^2 m_{2,0} ^2m_{2,1} + 630m_{2,0}^2m_{2,1}^3)$

 Table 6: The WS Cumulants, $\kappa_{n,q}$ part 3

19	$\kappa_{10,4} = m_{10,4} - (15m_{2,0}m_{8,4} + 15m_{4,0}m_{6,2}^*)$ $+ 24m_{2,1}m_{8,3} + 80m_{4,1}m_{6,3} + 6m_{2,0}^*m_{8,2}$ $+ 90m_{4,2}m_{6,2} + 24m_{4,1}^*m_{6,1} + m_{4,0}^*m_{6,0})$ $+ 2 \cdot (3(m_{2,0}^*)^2m_{6,1} + 72m_{2,0}^*m_{2,1}m_{6,1})$ $+ 180m_{2,1}^2m_{6,2} + 90 m_{2,0} ^2m_{6,2}$ $+ 240m_{2,0}m_{2,1}m_{6,3} + 45m_{2,0}^2m_{6,2}^*$ $+ 90m_{2,0}^*m_{4,0}m_{4,2} + 120m_{2,0}^*m_{4,1}^2$ $+ 120m_{2,1}m_{4,0}m_{4,1}^* + 720m_{2,1}m_{4,1}m_{4,2}$ $+ 240m_{2,0} m_{4,1} ^2 + 15m_{2,0} m_{4,0} ^2$ $+ 270m_{2,0}m_{4,2}^2) - 6 \cdot (15m_{2,0}^3m_{4,0}^*$ $+ 360m_{2,0}^2m_{2,1}m_{4,1}^* + 270m_{2,0} m_{2,0} ^2m_{4,2}$ $+ 720 m_{2,0} ^2m_{2,1}m_{4,1} + 480m_{2,1}^3m_{4,1}$ $+ 180m_{2,0}^*m_{2,1}^2m_{4,0} + 45 m_{2,0} ^2m_{2,0}^*m_{4,0}$ $+ 1080m_{2,0}m_{2,1}^2m_{4,2})$ $+ 24 \cdot (45m_{2,0} m_{2,0} ^4 + 528m_{2,0} m_{2,0} ^2m_{2,1}^2$ $+ 372m_{2,0}m_{2,1}^4)$
20	$\kappa_{10,5} = m_{10,5} - (10m_{2,0}m_{8,3}^* + 10m_{2,0}^*m_{8,3})$ $+ 25m_{2,1}m_{8,4} + 5m_{4,0}m_{6,1}^* + 5m_{4,0}^*m_{6,1}$ $+ 50m_{4,1}m_{6,2}^* + 50m_{4,1}^*m_{6,2} + 100m_{4,2}m_{6,3})$ $+ 2 \cdot (15m_{2,0}^2m_{6,2}^* + 15(m_{2,0}^*)^2m_{6,2}$ $+ 100 m_{2,0} ^2m_{6,3} + 150m_{2,0}m_{2,1}m_{6,2}^*$ $+ 150m_{2,0}^*m_{2,1}m_{6,2} + 200m_{2,1}^2m_{6,3}$ $+ 50m_{2,0}m_{4,0}^*m_{4,1} + 50m_{2,0}^*m_{4,0}m_{4,1}^*$ $+ 300m_{2,0}m_{4,1}^*m_{4,2} + 300m_{2,0}^*m_{4,1}m_{4,2}$ $+ 25m_{2,1} m_{4,0} ^2 + 400m_{2,1} m_{4,1} ^2$ $+ 450m_{2,1}m_{4,2}^2) - 6 \cdot (75(m_{2,0}^*)^2m_{2,1}m_{4,0}$ $+ 75m_{2,0}^2m_{2,1}m_{4,0}^* + 150m_{2,0} m_{2,0} ^2m_{4,1}^*$ $+ 150m_{2,0}^* m_{2,0} ^2m_{4,1}$ $+ 900 m_{2,0} ^2m_{2,1}m_{4,2} + 600m_{2,0}m_{2,1}^2m_{4,1}^*$ $+ 600m_{2,0}^*m_{2,1}^2m_{4,1} + 600m_{2,1}^3m_{4,2})$ $+ 24 \cdot (120m_{2,1}^5 + 600 m_{2,0} ^2m_{2,1}^3$ $+ 225 m_{2,0} ^4m_{2,1})$