Distributed Asynchronous Modulation Classification Based on Hybrid Maximum Likelihood Approach

Thakshila Wimalajeewa†, Jithin Jagannath*, Pramod K. Varshney†, Andrew Drozd* and Wei Su‡
†Dept. of EECS, Syracuse University, Syracuse, NY, {twwewelw, varshney}@syr.edu
*ANDRO Advanced Applied Technology, ANDRO Comput. Solutions, LLC, Rome, NY, {jjagannath, adrozd}@androcs.com
‡Army CERDEC, Aberdeen Proving Ground, MD, wei.su@ieee.org

***Abstract*—In this paper, we consider the problem of automatic modulation classification (AMC) with multiple sensors. A distributed hybrid maximum likelihood (HML) based algorithm in the presence of unknown time offset, phase offset and channel gain is presented. The proposed distributed algorithm that employs the generalized expectation maximization (GEM) algorithm is robust to initialization of unknown parameters, computationally efficient and require much less communication overhead compared to performing GEM in a centralized setting. Simulation and experimental results depict the efficacy of the proposed algorithm.**

***Keywords*—Modulation classification, hybrid maximum likelihood, generalized expectation maximization algorithm, distributed decision fusion.**

I. Introduction

The problem of automatic modulation classification (AMC) is becoming an integral part of various intelligent communication systems [1]–[3]. While there has been a substantial amount of work in the literature (please see [2], [3] for an extensive overview), there are still some important aspects to be further considered in the AMC problem to fully utilize it in practice. More specifically, the development of computationally efficient algorithms to perform AMC in the presence of unknown parameters is a challenging problem.

Various likelihood-based (LB) AMC techniques have been proposed in the literature depending on how the unknown parameters are treated [4]–[8]. However, most of this work ignores the time offset. In the presence of time offset, feature based techniques, which are suboptimal, have been developed in [9]. In a recent work [10], [11], the authors developed a hybrid maximum likelihood (HML) based approach to AMC in the presence of unknown time offset, phase offset and signal amplitude. To find unknown parameters via ML estimation, a computationally efficient numerical algorithm based on generalized expectation maximization (GEM) was proposed. The GEM algorithm developed in [10], [11] assumes that all the observations are available at a central fusion center and provides promising results when there is a good initialization technique for unknown parameters available. This is because, when jointly estimating the unknown parameters with the joint probability density function, the GEM algorithm gets trapped in local optima and is highly susceptible to the choice of the initial values. This leads to poor performance especially in the mid and high signal-to-noise ratio (SNR) regions. To alleviate this problem, we propose a distributed method where we use the GEM algorithm independently at each node to estimate unknowns. Once the unknowns are estimated, individual classification decisions are obtained and fused to perform the final AMC. This method, which does not require a sophisticated initialization technique for unknown parameters, is shown to perform better at medium to high SNR both by simulations and actual experiments. By estimating unknowns independently via the GEM algorithm at each node, we gain two main advantages over joint processing as considered in [11]; (i). While the joint estimation of unknowns via the GEM algorithm requires a sophisticated initialization technique especially in the mid-high SNR regions, the individual estimation of unknowns provides acceptable results with widely used initialization methods. (ii). Since not all the raw observations are transmitted to the fusion center and only a summary is sent, the communication overhead between sensors and the fusion center is much less than that in [11].

II. Problem Formulation

Let there be L radio receivers (or alternatively we call them sensors) observing a linearly modulated communication signal that undergoes block fading. The received baseband signal at the l-th radio (sensor) can be expressed as

$$y_l(t) = a_l e^{j\theta_l} \sum_n I_n g(t - nT - \varepsilon_l T) + w_l(t), \quad 0 \le t \le T_0 \quad (1)$$

for $l = 1, \cdots, L$ where T_0 is the observation interval, T is the symbol duration, $g(t)$ is the transmitted pulse, I_n is the n^{th} complex constellation of the transmitted symbol, $w_l(t)$ is the additive complex zero-mean white Gaussian noise process at the l-th radio with two-sided power spectral density (PSD) $N_0/2$, $a_l > 0$ is the channel gain between the transmitter and the l-th node, $\theta_l \in [-\pi, \pi)$ is the channel phase between the transmitter and the l-th node, and $\varepsilon_l T$ is the residual time offset at the l-th radio. We assume that the estimation of $g(\cdot)$, T and the carrier frequency has been accomplished at each receiver. Without loss of generality, we also assume $\varepsilon_l \in [0, 1)$. In this model, $\{a_l, \theta_l, \varepsilon_l\}_{l=1}^{L}$ for $l = 1, \cdots, L$ and $\{I_n\}_{n=0}^{N-1}$ are the unknown signal parameters. Let $\mathbf{u}_l \triangleq [a_l, \theta_l, \varepsilon_l]^T$ represent the deterministic unknown parameter vector at the l-th node for $l = 1, \cdots, L$ and $\mathbf{u} \triangleq [\mathbf{u}_1(1), \cdots, \mathbf{u}_L(1), \mathbf{u}_1(2), \cdots, \mathbf{u}_L(2), \mathbf{u}_1(3), \cdots, \mathbf{u}_L(3)]^T$ where $\mathbf{u}_l(j)$ denotes the j-th element of $\mathbf{u}_l$ for $j = 1, 2, 3$. Suppose there are S candidate modulation formats under consideration and let $I_n^{(i)}$ denote the n^{th} constellation symbol corresponding to the i-th modulation format

[1]Distribution Statement A: Approved for Public Release; Distribution is Unlimited

where $i \in \{1,\ldots,S\}$. Let $\mathcal{H}_i$ denote the hypothesis associated with the i-th modulation format. Further, let $\mathbf{I} \triangleq [I_0,\ldots,I_{N-1}]^T$ and $\mathbf{y} \triangleq [\mathbf{y}_1^T,\ldots,\mathbf{y}_L^T]^T$ where $\mathbf{y}_l$ denotes a vector representation of $y_l(t)$ and $(\cdot)^T$ denotes vector/matrix transpose. The goal is to identify the correct modulation format from S candidate formats (or equivalently the corresponding hypothesis) based on $\mathbf{y}$.

In [11], the authors have considered an HML approach [2] assuming that all the observation vectors $\mathbf{y}_l$'s are available at a central processing unit. The joint likelihood function of $\mathbf{y}$ is marginalized over the unknown random constellation symbols I_n and then maximized over the remaining unknown parameter vector $\mathbf{u}$. To alleviate the computational complexity of HML method, the GEM algorithm was used to estimate unknown parameters. When all the unknowns are estimated jointly via GEM, it is observed that the performance depends highly on parameter initialization. In particular, when simulated annealing (SA) with a coarse grid is used to initialize unknown parameters, it has been observed that the performance does not improve and even degrades as the SNR and the number of sensors increase. However, when the initial values are not far away from the actual values of unknown parameters, the GEM algorithm in a centralized setting provides promising results. Rather than finding sophisticated initialization schemes, we explore alternative techniques to exploit the presence of multiple sensor measurements to perform AMC with a given initialization technique for the GEM algorithm. To that end, the goal of this paper is to illustrate the performance gain of a distributed algorithm in which the individual nodes estimate unknown parameters corresponding to the particular node independently compared to centralized AMC via GEM. In the distributed approach, the number of unknown parameters estimated at each sensor is small, and the impact of initialization of parameters on the overall performance is not significant. Decision statistics are computed based on individual estimates of unknowns and individual decisions are transmitted to a fusion center to make the final decision.

III. Distributed AMC via GEM

In the proposed algorithm, the deterministic unknown vector $\mathbf{u}_l$ for each l is estimated separately at the l-th node and a test statistic computed based on the estimates is transmitted to the fusion center to perform final classification. Let $\hat{\mathbf{u}}_l{}^i$ be the estimated unknown vector under $\mathcal{H}_i$ and $\boldsymbol{\gamma}_l$ be a test statistic computed at the l-th node for $l = 1,\cdots,L$ and $i = 1,\cdots,S$. The l-th node transmits $\boldsymbol{\gamma}_l$ to the fusion center so that the final decision is made as $\hat{i}^* = f(\{\boldsymbol{\gamma}_l\}_{l=1}^L)$.

Conditioned on $\mathbf{I}$ and $\mathbf{u}_l$, the likelihood function based on $\mathbf{y}_l$ under $\mathcal{H}_i$ can be written in the following form [11] (note that we omit the use of superscript i on $\mathbf{u}_l$ that corresponds to $\mathcal{H}_i$ when there is no ambiguity):

$$p_i(\mathbf{y}_l|\mathbf{u}_l,\mathbf{I}) \propto \exp\left\{-\frac{E_g}{N_0}\sum_{n=0}^{N-1}|I_n|^2 a_l^2\right\}\cdot \exp\left\{\frac{2}{N_0}\sum_{n=0}^{N-1} a_l \Re\left\{I_n^* e^{-j\theta_l}\int_0^{T_0} y_l(t)g^*(t-nT-\varepsilon_l T)dt\right\}\right\}. \quad (2)$$

Then the marginalized likelihood function over $\mathbf{I}$ based on $\mathbf{y}_l$ is given by $p_i(\mathbf{y}_l|\mathbf{u}_l) = \sum_{\mathbf{I}^{(i)}} p_i(\mathbf{y}_l|\mathbf{u}_l,\mathbf{I}^{(i)})P(\mathbf{I}^{(i)})$. The resulting log-likelihood function (LLF) $\Lambda_i(\mathbf{u}_l) \triangleq \ln p_i(\mathbf{y}_l|\mathbf{u})$ under $\mathcal{H}_i$ at the l-th node based on $\mathbf{y}_l$ is given in (3) assuming $P(I_n^{(i)}) = 1/M_i$, where M_i is the cardinality of the constellation symbol set for hypothesis i. The maximum likelihood estimate (MLE) of $\mathbf{u}_l$ at the l-th node under $\mathcal{H}_i$ is given as

$$\hat{\mathbf{u}}_l^i = \arg\max_{\mathbf{u}_l} \Lambda_i(\mathbf{u}_l). \quad (4)$$

To reduce the computational complexity associated with (4), the GEM algorithm is used. Since the details are given in [11], we only provide relevant equations here. At each iteration r of the GEM algorithm, an estimate for $\mathbf{u}_l$ under $\mathcal{H}_i$ is found as,

$$\hat{\mathbf{u}}_l{}^i(r+1) = [\hat{a}_l(r+1)\ \hat{\theta}_l(r+1)\ \hat{\epsilon}_l(r+1)]^T \text{ with}$$

$$\hat{a}_l(r+1) = \frac{1}{E_g\hat{E}_I^l(r)}\sum_{n=0}^{N-1}\Re\left\{\hat{I}_n^{l*}(r)e^{-j\hat{\theta}_l(r+1)} \int_0^{T_0} y_l(t)g^*(t-nT-\hat{\varepsilon}_l(r+1)T)dt\right\} \quad (5)$$

$$\hat{\varepsilon}_l(r+1) = \arg\max_{\varepsilon_l}\sum_{n=0}^{N-1}\Re\left\{\hat{I}_n^{l*}(r)e^{-j\hat{\theta}_l(r)} \int_0^{T_0} y_l(t)g^*(t-nT-\varepsilon_l T)dt\right\} \text{ and} \quad (6)$$

$$\hat{\theta}_l(r+1) = \tan^{-1}\left(\frac{\Im(\hat{\mathbf{I}}^{l^H}(r)\mathbf{y}_l(r+1))}{\Re(\hat{\mathbf{I}}^{l^H}(r)\mathbf{y}_l(r+1))}\right), \quad (7)$$

where $\Re\{.\}$ and $\Im\{.\}$ denote real and imaginary parts respectively and we define $y_{n,l}(r)$ as

$$\begin{aligned} y_{n,l}(r) &\triangleq y_l(nT+\hat{\varepsilon}_l(r)T) \\ &= \int_0^{T_0} y_l(t)g^*(t-nT-\hat{\varepsilon}_l(r)T)dt, \end{aligned} \quad (8)$$

with $\hat{\mathbf{I}}^l(r) \triangleq \left[\hat{I}_0^l(r),\ldots,\hat{I}_{N-1}^l(r)\right]^T$ and $\mathbf{y}_l(r+1) \triangleq [y_{0,l}(r+1),\ldots,y_{N-1,l}(r+1)]^T$, in which $y_{n,l}(r+1)$ is obtained from (8), i.e., $y_{n,l}^{(r)} \triangleq y_l(nT+\hat{\varepsilon_l}^{(r)}T)$. The other relevant quantities required to compute (5), (6) and (7) are given by,

$$\hat{I}_n^l(r) \triangleq \sum_{m=1}^{M_i}\alpha_n^{m,l}(r)I_n^m \quad (9)$$

$$\hat{E}_I^l(r) \triangleq \sum_{n=0}^{N-1}\sum_{m=1}^{M_i}\alpha_n^{m,l}(r)|I_n^m|^2 \text{ with} \quad (10)$$

$$\alpha_n^{m,l}(r) = \frac{\exp\left(-|y_{n,l}(r)-\hat{a}_l(r)e^{j\hat{\theta}_l(r)}I^m|^2/N_0\right)}{\sum_{k=1}^{M_i}\exp\left(-|y_{n,l}(r)-\hat{a}_l(r)e^{j\hat{\theta}_l(r)}I^k|^2/N_0\right)}. \quad (11)$$

The iterations are continued by each node until a stopping criterion is satisfied.

$$\Lambda_i(\mathbf{u}_l) \;=\; \sum_{n=0}^{N-1} \ln\left(\sum_{k=1}^{M_i} \exp\left\{\frac{2}{N_0} a_l \mathrm{Re}\left\{I_n^{k*} e^{-j\theta_l} \int_0^{T_0} y_l(t) g^*(t - nT - \varepsilon_l T) dt\right\} - \frac{E_g}{N_0} a_l^2 |I_n^k|^2\right\}\right) - N \ln M_i \tag{3}$$

Once the estimates $\hat{\mathbf{u}}_l^i$'s are found for each $\mathcal{H}_i$, the corresponding marginalized log likelihood function $\Lambda_i(\hat{\mathbf{u}}_l^i)$ is computed at each node. After computing individual log likelihood functions, each node makes a decision on which modulation format is present and that decision is sent to the fusion center. Let $\hat{i}_l$ be the decision made at the l-th node based on

$$\hat{i}_l = \arg\max_i \left\{\Lambda_i(\hat{\mathbf{u}}_l^i)\right\}. \tag{12}$$

The transmitted information by each node is $\gamma_l = \hat{i}_l$ and the final decision at the fusion center on the modulation format is then given by,

$$\hat{i}^* = f_m(\{\hat{i}_l\}_{l=1}^L) \tag{13}$$

where f_m is a function that yields the modulation format with largest number of votes over $\hat{i}_l$ for $l = 1, \cdots, L$. It is noted that, to have an acceptable performance based on this scheme, the number of sensors should be larger than the number of modulation formats in the dictionary. The proposed algorithm is summarized in Algorithm 1

Algorithm 1 Asynchronous AMC with individual estimates

At the each node (radio) l for $l = 1, \cdots, L$

1) Set stopping criterion δ.
2) FOR $i = 1, \ldots, S$
3) Set $r = 0$, initialize $\hat{\mathbf{u}}_l^i(0)$
4) Compute $\alpha_n^{m,l}(r)$ from (11), $\hat{I}_n^l(r)$ from (9) for $n = 0, \ldots, N-1$; $m = 1, \ldots, M_i$. Compute $\hat{E}_I^r(l)$ from (10).
5) Set $r = r + 1$
6) Compute $\hat{a}^{(r+1)}$, $\hat{\theta}_l(r+1)$ and $\hat{\varepsilon}_l(r+1)$ using (5), (7), and (6), respectively.
7) If stopping criteria is not met, go to Step 4, else set $\hat{\mathbf{u}}_l^i = \hat{\mathbf{u}}_l^i(r+1)$ and continue
8) Compute $\hat{i}_l$ as in (12)
9) ENDFOR

At the fusion center

Receive $\hat{i}_l$ for $l = 1, \cdots, L$. Obtain $\hat{i}^* = f_m(\{\hat{i}_l\}_{l=1}^L)$ as in (13).

Final decision is $\hat{i}^*$

IV. Numerical Results

In this section, we illustrate the performance of the proposed distributed AMC algorithm and compare it with the GEM based AMC in a centralized setting as considered in [11]. For numerical results, we assume that $g(t)$ is a symmetrically truncated RRC pulse, with a roll-off factor of 0.3 and duration $8T$. We consider that $\mathbb{E}\{|I_n|^2\} = 1$, $N_0 = 1$, and a_l to be a Rayleigh distributed random variable with scale parameter σ for $l = 1, \cdots, L$. With these assumptions, the channel signal-to-noise ratio (SNR) is $\mathbb{E}\{a_l^2|I_n|^2\}/N_0 = 2\sigma^2$. We further take $T = 1$, $\theta_l \sim \mathcal{U}[-\pi, \pi)$ and $\varepsilon_l \sim \mathcal{U}[0, 1)$, for $l = 1, \cdots, L$ where $\mathcal{U}[a, b)$ denotes uniform distribution with support $[a, b)$. The observation interval is set as $T_p = NT$.

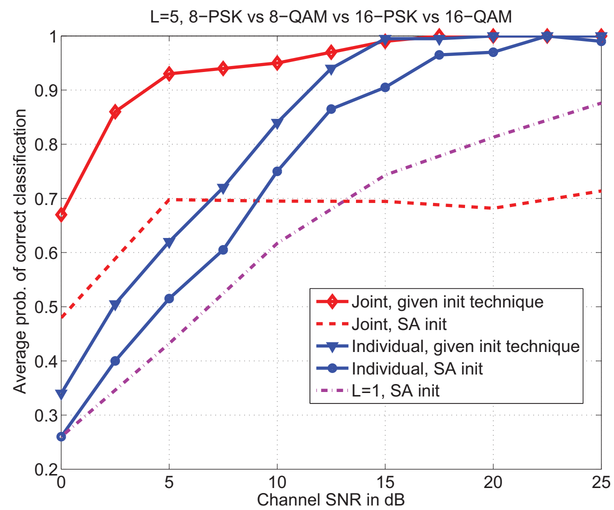

Fig. 1. 8-PSK vs 8-QAM vs 16-PSK vs 16-QAM. Initialization for GEM: (i) with a given initialization technique with $\delta_a = 5$, $\delta_\theta = \pi/10$, $\delta_\varepsilon = 0.1$ (ii) SA, $L = 5$, $N = 100$

In Fig. 1, we consider a quaternary classification scenario where the modulation formats to be classified are 8-PSK, 8-QAM, 16-PSK, and 16-QAM. We plot the AMC performance in terms of the average probability of correct classification (assuming all possible formats in the dictionary appear with equal probability) vs channel SNR. For the GEM algorithm, we consider two initialization techniques: (i) Initialization points of unknown parameters are taken as the true values plus some error. More specifically, we consider that the initial values for the unknown parameters a_l, θ_l and ε_l can take any random values uniformly distributed in the regions $[0, a_l + \delta_a]$, $[\theta_l - \delta_\theta, \ \theta_l + \delta_\theta]$, and $[\varepsilon_l - \delta_\varepsilon, \ \varepsilon_l + \delta_\varepsilon]$, respectively, for $l = 1, \cdots, L$ where $\delta_a, \delta_\theta, \delta_\varepsilon > 0$ are the maximum errors for each unknown. These error bounds determine how close the initial points are to the true values. (ii) Simulated annealing (SA) as the initialization technique for GEM. For SA initialization, we consider a coarse uniform grid as considered in [11].

In Fig. 1, we let $L = 5$ and $N = 100$. We also plot the performance with joint estimation of unknowns based on GEM in a centralized setting as considered in [11]. As a reference, we show the performance with having only one sensor. It is noted that when $L = 1$, the proposed distributed algorithm and the one developed in [11] coincide with each other. As discussed in detail in [11], it can be seen in Fig. 1 that the GEM algorithm in a centralized setting performs well only when a good initialization technique is available. With SA based initialization scheme with a coarse grid, which is a computationally quite efficient approach for initialization, the performance of the centralized scheme does not improve and even somewhat degrades as SNR increases. However, in the

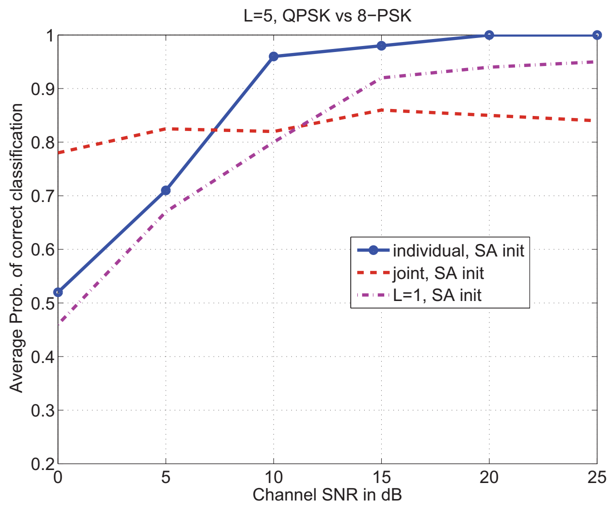

Fig. 2. QPSK vs 8-PSK when SA is taken as the initialization scheme for GEM, $L = 5$, $N = 100$

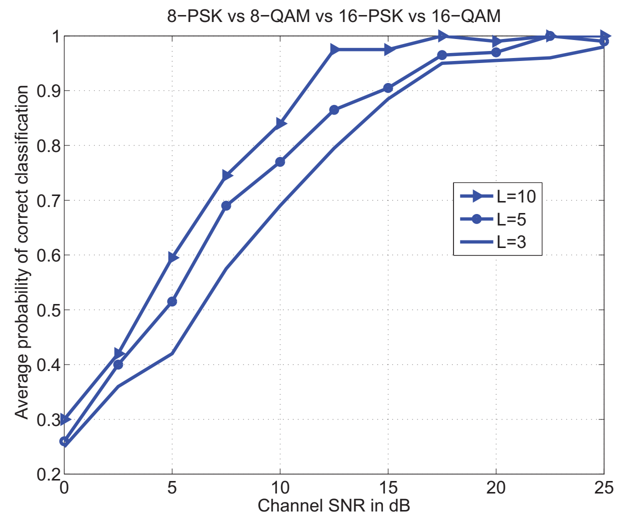

Fig. 3. 8-PSK vs 8-QAM vs 16-PSK vs 16-QAM. GEM with SA based initialization, $L = 5$, $N = 100$

distributed algorithm considered in this paper, where unknowns are estimated independently at each node, the performance of AMC in the high SNR region is promising. In particular, in the distributed algorithm, the performance with SA based initialization is comparable with that when the initial values are given with a small error (compared to true value). As discussed to a certain extent in [11], when all the unknowns are estimated jointly in a centralized setting, the GEM algorithm gets trapped at local maxima in the high SNR region with multiple sensors leading to poor performance. Thus, in the mid and high SNR regions, to better exploit the GEM algorithm with multiple sensors, it is better to perform unknown parameter estimation independently at each node, as considered in this paper. Simulation results with the distributed AMC algorithm based on GEM in classifying QPSK vs 8-PSK is presented in Fig. 2 where parameter initialization is done via SA. It again shows the improved performance of the GEM algorithm with individual estimate of unknowns compared to that with the joint estimates in the mid-high SNR regions.

It is also worth commenting on the performance in the low SNR region. As can be seen in both Fig. 1 and Fig. 2, the performance of the GEM algorithm with joint estimates is superior to performance with individual estimates in the low SNR region with SA based initialization scheme. However, since the performance is not increasing as the SNR increases when adding more sensors, the use of the centralized scheme is limited to the cases when the SNR is low or when a 'good' initialization scheme for unknown parameters is available. The proposed distributed algorithm is capable of improving the overall classification performance compared to a single sensor irrespective of the value of the SNR and its use is significant in the mid-high SNR regions compared to the centralized scheme.

Fig. 3 shows the performance of the distributed AMC algorithm as the number of sensors varies when SA is taken as the initialization technique for GEM. It is noted that the performance improvement with the addition of more sensors is not that significant in low SNR regions. In the proposed algorithm, after the individual decisions are made based on individual estimates of unknowns, fusion is performed via

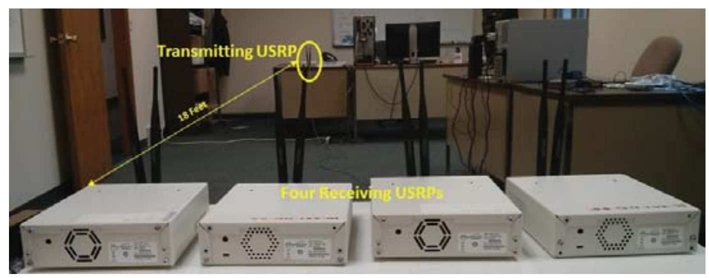

Fig. 4. Experimental set-up with Universal Software Radio Peripheral N210 (USRP N210) and GNU radio platform

the majority rule. Thus, the addition of sensors is exploited only at the fusion stage of the individual decisions. On the other hand, if we were to perform AMC jointly via GEM with all the raw observations (where fusion is performed with raw observations), the GEM algorithm gets trapped in local maxima leading to poor performance. Thus, there is a trade-off between the expected performance improvement and the use of multiple sensors when AMC is performed via GEM. However, as the SNR increases, a significant performance improvement is observed in Fig. 3 as the number of sensors increases.

Experimental validation

In the following, we provide experimental results to illustrate the AMC performance with individual estimates for unknowns using a simple initialization scheme for GEM.

The algorithm was evaluated on a software defined radio (SDR) testbed. The testbed comprised of five Universal Software Radio Peripheral N210 (USRP N210) controlled using open-source signal processing software called GNU radio and two Linux based host PCs. The SDR framework used to implement the AMC algorithm is described in [12]. The experimental set-up is shown in Fig. 4 which is the same as that used for experimental results in [13]. However, fusion and parameter initialization techniques used in [13] are different from the algorithm presented in this paper. The power output of USRP is 15 dBm and the noise figure is 5 dB. The daughterboard performs mixing, amplification and low pass filtering of the signals. The on-board FPGA performs digital up/down conversion, interpolation and decimation before/after the dual DAC/ADC. Each USRP was paired with an omni-directional VERT2450 vertical antennas. One USRP platform was designated as the transmitter. The transmitter is first set to transmit 8-PSK symbols and QPSK symbols respectively for evaluating the performance of the classifier for the respective modulation scheme. The transmitter (Tx) and receiver (Rx) gain is set to 1 dB. As shown in Table I we conducted the experiments with $L = 1, 2, 4$. We also repeated the experiments for $L = 1, 2$ with different combination of USRP to capture the variability between data collection of hardware itself (frequency offset, location, noise ect). The probability of correct classification (P_c) was determined from 100 runs for each experimental setup. The initial values of unknowns are selected randomly and unknown parameters corresponding to each node are estimated independently using GEM. During the implementation we did not estimate timing offset because we could not find any degradation in performance that was caused by timing error. The final classification decision is made by fusing the individual likelihood functions. As we can see in Table I, the performance of classifier improves monotonically as the number of sensors increases. Thus, the experimental results corroborate theoretical results that show performance improvement by adding multiple sensors for AMC via GEM with individual estimates of unknown parameters using a simple initialization scheme.

TABLE I. PROBABILITY OF CORRECT CLASSIFICATION: QPSK VS 8PSK

Tx mod	$L = 1$	USRP ID	$L = 2$	USRP ID	$L = 4$	USRP ID
QPSK	0.54	1	0.67	1 & 2	0.97	1,2,3,4
	0.89	2	0.93	1 & 3		
	0.83	3	0.80	3 & 2		
	0.90	4	0.98	3 & 4		
8-PSK	0.31	1	0.79	1 & 2	0.87	1,2,3,4
	0.22	2	0.70	1 & 4		
	0.23	3	0.82	3 & 4		
	0.56	4	0.67	1 & 3		

V. CONCLUSION

We have proposed a distributed algorithm to perform AMC with multiple sensors in the presence of unknown channel gain, channel phase and time offset. In the proposed scheme, each radio (senor node) estimates the unknowns associated with it independently by maximizing the marginalized LF based on the GEM algorithm. Then, a summary is transmitted to a fusion center to make the global decision. The proposed distributed algorithm is shown to less susceptible for initial values of unknowns (required by GEM) compared to performing AMC by joint estimation of unknowns as in [11]. Further, in the proposed approach, each radio is required to transmit only a small amount of information to a central processing unit compared to the amount raw observations. The performance of the distributed AMC algorithm is evaluated by via both simulation and experimental results.

REFERENCES

[1] J. Hamkins, M. K. Simon, and J. H. Yuhen, *Autonomous Software-Defined Radio Receivers for Deep Space Applications (JPL Deep-Space Communications and Navigation Series)*. Wiley-Interscience, 2006.

[2] O. A. Dobre, A. Abdi, Y. Bar-Ness, and W. Su, "Survey of automatic modulation classification techniques: classical approaches and new trends," *IET Communications*, vol. 1, no. 2, pp. 137–159, Apr. 2007.

[3] Z. Zhu and A. K. Nandi, *Automatic Modulation Classification: Principles, Algorithms and Applications*. ISBN: 978-1-118-90649-1, 2014, published online.

[4] U. Mengali and A. N. D'Andrea, *Synchronization Techniques for Digital Receivers*. New York: Plenum, 1997.

[5] A. E. El-Mahdy and N. M. Namazi, "Classification of multiple M-ary frequency-shift keying signals over a Rayleigh fading channel," *IEEE Trans. Commun.*, vol. 50, no. 6, pp. 967–974, Jun. 2002.

[6] C. Huang and A. Polydoros, "Likelihood methods for MPSK modulation classification," *IEEE Trans. Commun.*, vol. 43, no. 234, pp. 1493–1504, 1995.

[7] Q. Shi, Y. Gong, and L. Guan, "Asynchronous classification of high-order QAMs," in *Proc. IEEE Wireless Comm. and Networking Conf. (WCNC)*, Apr. 2008.

[8] W. C. Headley and C. R. C. M. da Silva, "Asynchronous classification of digital amplitude-phase modulations in flat-fading channels," *IEEE Trans. Commun.*, vol. 59, no. 1, pp. 7–12, Jan. 2011.

[9] B. F. Beidas and C. L. Weber, "Asynchronous classification of MFSK signals using the higher order correlation domain," *IEEE Trans. Commun.*, vol. 46, no. 4, pp. 480–493, Apr. 1998.

[10] O. Ozdemir, P. K. Varshney, and W. Su, "Asynchronous hybrid maximum likelihood classification of linear modulations," in *Globecom*, Dec. 2013.

[11] O. Ozdemir, T. Wimalajeewa, B. Dulek, P. K. Varshney, and W. Su, "Hybrid maximum likelihood based linear modulation classification with multiple sensors via generalized EM algorithm," *IEEE Trans. Wirless Commun.*, submitted 2014, now in second round review with minor revision, arXiv preprint arXiv:1409.8370.

[12] J. Jagannath, H. M. Saarinen, and A. L. Drozd, "Framework for automatic signal classification techniques (FACT) for software defined radios," in *IEEE Symposium on Computational Intelligence in Security and Defense Applications (CISDA)*, Verona, New York, May 2015.

[13] S. Foulke, J. Jagannath, A. Drozd, T. Wimalajeewa, P. K. Varshney, and W. Su, "Multisensor Modulation Classification (MMC) Implementation considerations USRP case study," in *Proc. 2014 IEEE MILCOM*, Baltimore, MD, Oct 2014, pp. 1663 – 1668.