Weighted ℓ_1 Minimization for Event Detection in Sensor Networks

Zainul Charbiwala, Younghun Kim, Sadaf Zahedi, Rahul Balani, Mani B. Srivastava
Dept. of Electrical Engineering, University of California, Los Angeles

***Abstract*—Event detection is an important application of wireless sensor networks. When the event signature is sparse in a known domain, mechanisms from the emerging area of Compressed Sensing (CS) can be applied for estimation with average measurement rates far lower than the Nyquist requirement. A recently proposed algorithm called IDEA uses knowledge of *where* the signal is sparse combined with a greedy search procedure called Orthogonal Matching Pursuit (OMP) to demonstrate that detection can be performed in the sparse domain with even fewer measurements. A different approach called Basis Pursuit (BP), which uses ℓ_1 norm minimization, provides better performance in reconstruction but suffers from a larger sampling cost since it tries to estimate the signal completely.**

In this paper, we introduce a mechanism that uses a modified BP approach for detection of sparse signals with known signature. The modification is inspired from a novel development that uses an adaptively weighted version of BP. We show, through simulation and experiments on MicaZ motes, that by appropriately weighting the coefficients during ℓ_1 norm minimization, detection performance exceeds that of an unweighted approach at comparable sampling rates.

I. Introduction

Detecting interesting events in a noisy and interference prone environment is a critical application of sensor networks [1], [2]. Classic methods such as *Hypothesis Testing* focus on minimizing error by testing the statistical significance of the event signature while taking noise characteristics into account [3]. Besides being robust to noise, simple hypothesis testing has the advantage of being computationally economical [4]. Although it performs well in many cases, it suffers in interference prone environments. For example, if the detection scheme monitors the power level of a signal, strong interference would trigger a false alarm.

Transform domain analysis of the signal helps differentiate the event against interference [5], [6], but requires expensive transformation steps before detection can be performed. Furthermore, in order to ensure that the transformed signal has sufficient information, most approaches *implicitly* assume a Nyquist sampling rate for sensing. While this is required for some applications, sampling at the Nyquist rate is often wasteful, especially when the event signature is known to be *sparse*, i.e. has few non-zero values.

Advances in the emerging area of Compressed Sensing (CS) [7], [8] suggest that if the signal is sparse in a known domain, very few measurements need to be taken in an incoherent domain to *reconstruct* the signal perfectly with overwhelming probability. If it were possible to transform the signal to the sparse domain at the source, this would be no different than transform domain analysis. The key difference underpinning CS mechanisms is that the structure of these sparse signals can be described quite compactly even in the incoherent domain if the so-called restricted isometry property (RIP [9], [10]) is obeyed. Practically, this implies, first, that the signal has to be transformed into this incoherent domain at the source and second, that the measurements must obey the RIP. Researchers have shown [11] that by taking suitable random projections of the signal, both these requirements are met.

CS mechanisms are especially beneficial to sensor networks because the nodes are usually constrained in sensing, computation, memory and communications bandwidth. For example, in applications based on acoustic signals [12], [13], low-end sensor network platforms, including MicaZ motes, lack the ability to sample data at Nyquist rates. Resultingly, Allen et. al. [12] use an ARM processor based sensor node for detecting marmot calls and Wang. et. al. [13] use higher end sensor nodes partly due to sampling rate requirements. In addition to sensing cost, high sampling rates are detrimental to the energy efficiency of both computation and communication.

Proposed CS reconstruction algorithms are computationally intensive today, but this poses little practical hindrance since a notable characteristic of many sensor network deployments is an asymmetry in architecture with a data collection and fusion center endowed with a considerable amount of computing and storage ability. Thus, if the sensor nodes are able to take random projections of the sampled signal, the fusion center will be able to reconstruct the signal with high probability using only a fraction of what the Nyquist rate would have required.

While many CS mechanisms have focused on signal reconstruction, interesting results for event detection have also been reported [14]. For example, Davenport et. al. [15] introduce the smashed filter method that exploits the fact that a particular object has a unique set of manifolds, which efficiently describes object characteristics. They have shown that object recognition in images is possible with far fewer samples than traditional detection schemes. Dang et. al. [16] describe an event detection scenario to detect and classify cane-toads in northern Australia using a combination of randomized sampling and matched filters.

A key observation that researchers have made for CS based event detection schemes is that the number of measurements required to reliably detect the signal can be considerably lower than for CS reconstruction. Recently, a new algorithm proposed by Duarte, et. al. (IDEA [17]) demonstrates this by utilizing knowledge of *where* the event may be present in the sparse domain. The mechanism employed in the algorithm is an iterative greedy search procedure called orthogonal matching pursuit (OMP [18], [19], [20]), which conceptually looks for the single component that best describes the signal. It then removes this component from the signal and repeats the procedure for a fixed number of iterations. IDEA builds upon OMP in that by knowing which components represent the event signature, OMP can terminate as soon as these components are found. This results in not only fewer OMP iterations, but also requires fewer sample measurements.

An alternative to OMP for reconstruction from incoherent projections is called basis pursuit (BP). BP uses a relaxation of the ideal ℓ_0 norm sparse recovery procedure to an ℓ_1 norm minimization [9], and has been demonstrated in [8] to perform better than OMP in practice. Intuitively, this is because BP attempts to find the global minimum while OMP might get caught in a local dip. There are two drawbacks of using BP over IDEA for detection, however. The first is that though ℓ_1 norm minimization can complete in polynomial time, the computation requirement is far higher than OMP [20]. The second is that since BP attempts to reconstruct the signal completely, the number of measurements required for comparable detection performance may actually be higher than IDEA. In a practical sensor network deployment with a capable back-end fusion center, the first drawback can be overlooked. This paper focuses on overcoming the second drawback.

Our solution to the problem was inspired by the recent work of Candes, et. al. [21], which applies an iterative procedure around ℓ_1 norm recovery. In each iteration of the loop, an adaptive weighting matrix is applied while performing the minimization. The weights used in each iteration is an inverse function of the values computed in the previous iteration. The effect of this form of weighting is an equalization of the penalty faced by each component, making the re-weighted ℓ_1 norm minimization a much closer approximation to ℓ_0 norm minimization.

The solution we propose applies the same insight, but in a novel way. *If we know where the event's components lie in the sparse domain, we can bias those components so that they are artificially enhanced against background noise*. This is done

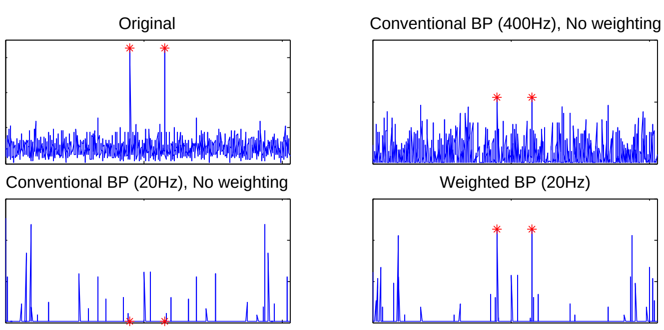

Fig. 1. The effects of weighted ℓ_1 reconstruction of a 450Hz tone of 20dB SNR using 20 Hz random sampling.

by applying a weighting matrix similar to [21], but choosing weights such that detection probability improves. The solution to both the unweighted and the weighted problems will adhere to the measurement constraints, but the effect of weighting is that of improving the 'contrast' of the signal against noise. This is visually depicted in Figure 1 for the detection of a single frequency tone at 450 Hz in the presence of white noise. The reconstruction is performed in the Fourier domain using an average sampling rate of 20 Hz. Notice that as the weighting factor is increased, the frequency tones progressively stand out.

It is important to mention here that enhancement through biased weighting may hurt false alarm performance. If the noise or interference reside at the same locations as the event, they will be erroneously enhanced too. For most realistic event signatures and scenarios, however, the structure of the signals and noise will be distinct enough that false detections will be minimal. We evaluate this issue in detail in Section IV. To complete the solution, we also propose to use the Winner Takes All (WTA) strategy for detection. This strategy tests whether the top components resulting from the weighted ℓ_1 norm minimization correspond to the event signature locations. If not, the event is declared absent, however large the component values are. We compare this strategy to that proposed in IDEA [17] that tests against some precomputed threshold, γ. We show in Section IV why thresholding performs better than WTA in simulation, but results in reduced performance with a real implementation.

An additional contribution of this paper is an implementation of the proposed solution on off-the-shelf MicaZ sensor nodes. In Section III, we show how RIP compliant random projections can be taken on these motes. In Section IV we evaluate both simulated and experimental performance on a sensor network testbed for a number of sampling rates in various noisy environments. In the following section, we formulate our detection problem.

II. WEIGHTED ℓ_1 NORM MINIMIZATION

Before we determine the context of the detection procedure, we first outline the BP estimation problem as follows: Assume that the signal of interest $x \in \mathbb{R}^n$ and that a set of measurements $z \in \mathbb{R}^k$, $k \ll n$ are available to us, such that $z = Ax$, where $A \in \mathbb{R}^{k \times n}$ is the measurement projection matrix. Then, under the condition that x is sufficiently sparse, the solution to the following combinatorial optimization problem recovers the signal exactly:

$$\hat{x} = \underset{\tilde{x}}{\text{argmin}} \, \|\tilde{x}\|_{\ell 0} \quad \text{s.t.} \quad z = A\tilde{x} \tag{1}$$

where $\|x\|_{\ell 0} \triangleq |\{i : x_i \neq 0\}|$. Equation 1 is NP-complete in general. Instead, a relaxed version of the problem that is convex is proposed:

$$\hat{x} = \underset{\tilde{x}}{\text{argmin}} \, \|\tilde{x}\|_{\ell 1} \quad \text{s.t.} \quad z = A\tilde{x} \tag{2}$$

where $\|x\|_{\ell 1} \triangleq \sum_{i=1}^{n} |x_i|$. It is shown in [7] that under the sparsity condition and that A satisfies the restricted isometry property, the reconstruction $\hat{x}$ is exact with overwhelming probability [9]. Practically, this means that if the signal is sparse in the sensing domain, then taking k measurements through a suitable linear transformation A will be sufficient to reconstruct the signal. If the signal is not sparse in the sensing domain, but in another known domain, the reconstruction must be performed in two steps. Assume a separate invertible linear transformation F which renders the signal sparse, that is, $y = Fx$ where $F \in \mathbb{C}^{n\times n}$. For example, if the signal x was a single frequency tone, then its time domain representation is not sparse, but with F as the Fourier transform, y is sparse. The equivalent reconstruction procedure is then:

$$\hat{y} = \underset{\tilde{y}}{\operatorname{argmin}} \|\tilde{y}\|_{\ell 1} \quad \text{s.t. } z = AF^{-1}\tilde{y} \tag{3}$$

$$\hat{x} = F^{-1}\hat{y} \tag{4}$$

An additional requirement for the reconstruction to succeed is that F be *incoherent* [9] with the sensing domain. For many real-world signals, the sensing domain is temporal or spatial and the incoherent transform that (approximately) 'sparsifies' the signal is well known and there exist fast algorithms for computing both Fx and $F^{-1}x$.

It is understood that ℓ_1 regularization performs quite well when the sparsity condition is satisfied, but the question we wish to investigate here is whether an event signature can be identified from fewer measurements. To aid our understanding of the problem, we refer to a recent development by Candes, et. al. [21] that attempts to bridge the gap between Equation 1 and 2, without losing the convexity property of the latter. A key property of the ℓ_0 minimization problem is that it treats all non-zero values the same, whereas the ℓ_1 version of the problem penalizes component values based on their magnitude. The authors propose an iterative solution to the problem by modifying Equation 3 as follows:

$$\hat{y} = \underset{\tilde{y}}{\operatorname{argmin}} \|W\tilde{y}\|_{\ell 1} \quad \text{s.t. } z = AF^{-1}\tilde{y} \tag{5}$$

The weighting matrix W in [21] is an iteratively updated diagonal matrix with entries that are inversely proportional to the magnitude of the solution $\hat{y}$ from the previous iteration. The algorithm is initiated with $W = I$. The authors demonstrate both analytically and numerically that this inverse relationship between weights and signal magnitudes renders the procedure a better approximation to ℓ_0 minimization.

Returning to our detection problem, we propose to exploit weighted ℓ_1 norm minimization as follows: if we know the set Ω of component indices of y that represent the event to be detected, we construct the diagonal matrix W in Equation 5 as :

$$W = diag(\frac{1}{w_1}, ..., \frac{1}{w_n}), \quad \begin{array}{ll} w_j \geq 1 & \forall j \in \Omega \\ w_j = 1 & \forall j \notin \Omega \end{array} \tag{6}$$

Such a weighting reduces the penalty of only those components that form part of the event signature. This improves their chances of being selected compared to the unweighted components. One way to view the effect of this biased weighting is that it artificially enhances the contrast of the event against background noise. Candes et. al. [21] particularly support this interpretation when the event signature is present. One interesting observation we see in Fig. 1 is that as the weighting value $\frac{1}{w_i}$ decreases the results tend to stabilize, which means that a weighting less than a certain value gives an almost identical result. This is partly because the norm ball $\|W\tilde{y}\|_{\ell 1}$ hits the same point on the polyhedra $z = AF^{-1}\tilde{y}$ beyond a point.

In an event detection scenario, the more interesting question arises when the event is not present as opposed to the case in [21], because then the weighting violates the implicit assumption that the event signature is present at those entries. If the event is absent, noise or interference at indices Ω, will be erroneously enhanced, therefore choosing the right weights w_Ω is a critical decision in the procedure. This particular interest suggests that we cannot directly use the approach by Candes et. al. Further, we conjecture that the event signature has a structure distinct from noise and if the detection function accounts for this distinction, the number of false alarms will be minimal. In general, however, a trade-off exists between the number of missed detections and false alarms and in Section IV, we show that that choice of weights depends heavily on the detection function and the signal-to-noise ratio (SNR) and to an extent, on the average sampling rate.

A. *Detection Functions*

We declare the hypothesis of an event being present ($\mathcal{H}_1$) or absent ($\mathcal{H}_0$) by computing $\mathcal{D}(\hat{y}, \Omega)$, where $\hat{y}$ is the solution to Equation 5 with weighting using (6), $|\Omega| \geq 1$ and $|\cdot|$ is the cardinality operator. In general, the function $\mathcal{D}$ is non-linear and in this paper we consider two alternatives: precomputed threshold testing (PTT) and winner takes all (WTA), which are defined as follows:

$$\mathcal{D}_{PTT}(y, \Omega) = \begin{cases} 1 & \text{if } y_j > \theta_j \;\; \forall j \in \Omega \\ 0 & \text{otherwise} \end{cases} \tag{7}$$

where θ_Ω represents a threshold for each component in y that the event signature is composed of. PTT is an adaptation of the γ thresholding procedure described in IDEA [17]. WTA is defined as:

$$\mathcal{D}_{WTA}(y, \Omega) = \begin{cases} 1 & \text{if } \mathcal{M}(y) \in \Omega \\ 0 & \text{otherwise} \end{cases} \tag{8}$$

where $\mathcal{M}(y)$ returns the *index* of the maximum component in y. The WTA procedure, as defined, declares the event present if the index of $\max(y)$ belongs to Ω. A more conservative version, kWTA may compare the largest k values of y. When $|\Omega| = 1$, these are equivalent.

Detection performance is measured in terms of the probability of missed detections, P_{MD} and probability of false alarms, P_{FA}, which are defined as:

$$P_{MD} = \Pr[\mathcal{D}(\hat{y}, \Omega) = 0 \mid \mathcal{H}_1] \tag{9}$$

$$P_{FA} = \Pr[\mathcal{D}(\hat{y}, \Omega) = 1 \mid \mathcal{H}_0] \tag{10}$$

We extensively evaluate the performance of WTA and PTT in both simulation and through experiments in Section IV, but first describe some nuances that are specific to implementation on low-end sensor network platforms.

III. Implementation

In order to test our proposition in practice, we implemented the solution using MicaZ sensor motes running the TinyOS operating system. To motivate the application of a CS based approach, we selected the use of the microphone for acoustic signature detection similar to [22]. We chose to detect a single frequency tone at 450 Hz using the Fourier basis for reconstruction. It is arguable whether single tone detection has any significant practical uses. However, we chose this seemingly simple detection scenario as a case study for three reasons. First, it provides us with a baseline for comparison using a well known basis. Second, the solution is easily extended to event signatures sparse in other bases, with no change to the TinyOS code. And finally, single frequency tones have little structure to be exploited by the detection function and thus the false alarm rates reported in Section IV may be considered worst case.

MicaZ sensor motes contain an 8 MHz 8-bit ATMEGA128 processor with built-in ADC and an IEEE 802.15.4 compliant radio, but sustain sampling rates of only a few hundred Hz, partly due to the absence of a DMA unit. Clearly, detecting the tone via Fourier domain analyis on the mote itself or through sampling and collecting data wirelessly at the Nyquist rate would have been infeasible. Our goal was to detect the tone reliably at a fusion center at the lowest sampling rate.

A. Taking Random Projections

Perhaps the most important aspect of implementing CS is the construction of the projection matrix A (in Equation 5) through which the sensor node collects incoherent sample measurements. From [11], we learn that when the elements of A are independent realizations of a Gaussian random variable, such that $A_{ij} = \mathcal{N}(0, \frac{1}{n})$ or when they are independent realizations of an equiprobable $\pm\frac{1}{\sqrt{n}}$ Bernoulli random variable, the restricted isometry property is obeyed. Practically, however, this requires the sensor node to not just sample at the Nyquist rate but to also perform floating point multiply and add operations at that rate to compute the random projections. The device described in [23] is a novel hardware based approach to computing random projections, but we desired a lightweight software approach for our implementation.

Following the approach of uniform random sampling from [24], [15] and studying the procedure in [16], we sought to model our implementation to theirs. Uniform random sampling, however, is non-causal if the random numbers are generated on-the-fly. To ensure causality, one would have generate, sort and store the numbers somewhere in memory. Further, this technique has the disadvantage that two sample times may be closer together than the hardware can handle. Dang, et. al. [16] circumvented this problem by applying a scaling factor before and after generating the random sample indices. To avoid quantization effects, they add a normally distributed jitter to the resulting values.

A simpler technique that solves both issues and is a good approximation to the uniform distribution is mentioned in Bilinskis and Mikelsons [25] and was first suggested by

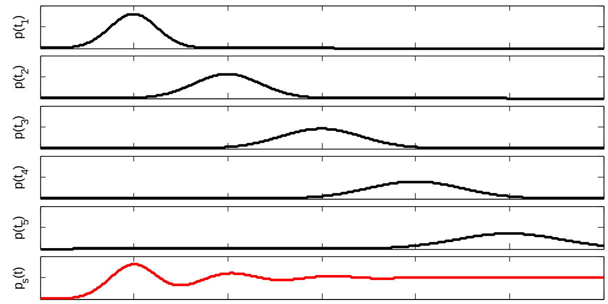

Fig. 2. The effect of an additive Gaussian random sampling process

Shapiro and Silverman [26]. If we define the k sampling instants as $t_i,\ i \in \{1, ..., k\}$, then the sampling instants are generated using:

$$t_{i+1} = t_i + \tau_i \qquad (11)$$

where τ_i are independent realizations of a Gaussian random variable $\sim \mathcal{N}(\mu, r^2\mu^2)$. Here, μ represents the required averaged sampling intervals and r determines the width of the bell and the resulting speed of convergence to a uniform distribution. The effect of the additive random sampling procedure is visually depicted in Figure 2 (adapted from [25]) with $r = 0.25$. The top 5 plots represent the PDFs of each $t_i,\ i \in \{1, .., 5\}$ and the bottom plot represents the PDF of realizations of all t_i, which approximates the uniform distribution as required.

We use this procedure in our implementation to generate random sampling times on-the-fly, with μ decided by the sampling rate and r fixed at 0.25. Ironically, we use the uniform random number generator provided in TinyOS to generate the Gaussian random variable by approximating it to an Irwin-Hall distribution of order 12 as described in [27].

IV. Results

Figure 3 depicts a schematic representation of the detection process used for evaluation. A host machine generates the signal and white noise at a specific SNR at a high sampling rate. This audio stream is played out over a speaker and recorded through the microphone of a *sensing* MicaZ mote using random projections as described in Section III-A. The recorded samples are then wirelessly transmitted to a *basestation* mote connected to the fusion center, which performs weighted ℓ_1 minimization to recover the signal in the frequency domain. The FFT coefficients are fed into the detection function alongwith the indices Ω to produce the hypothesis decision. We also run a simulation version of the process, which emulates the recording and collection process by applying the same random projection matrix as would have been computed on the sensing mote.

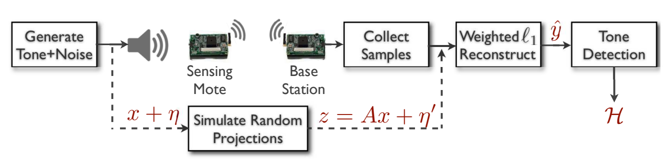

Fig. 3. Schematic representation of detection process with MicaZ motes and in simulation

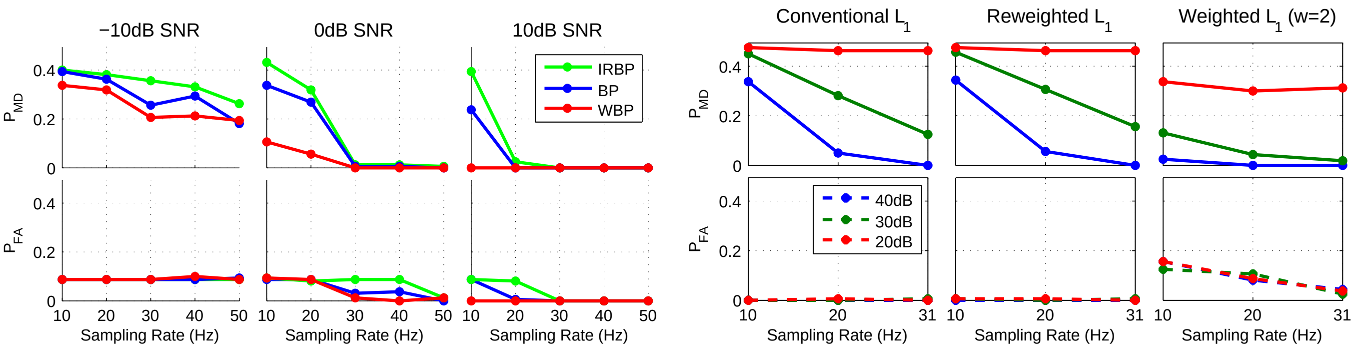

Fig. 4. P_{MD} and P_{FA} for PTT (left) and WTA (right) for various SNR and sampling rates for Reweighted ℓ_1[21].

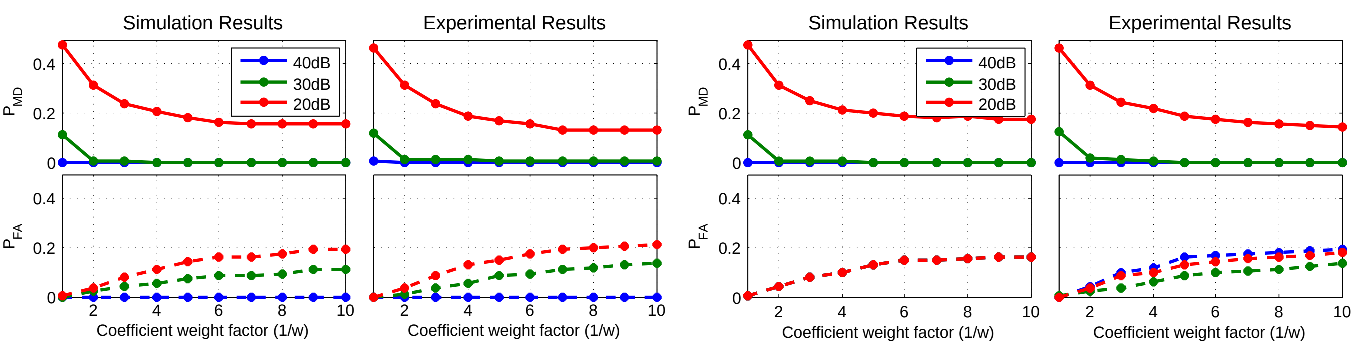

Fig. 5. P_{MD} and P_{FA} for PTT (left) and WTA (right) for various SNR at 31 Hz sampling.

We show results from Monte Carlo simulations and experiments at three different sampling rates (10, 20 and 31 Hz) and at three different SNRs (20, 30 and 40 dB) each. The set of plots on the left illustrate performance using PTT and on the right, using WTA, as defined in Section II-A. The threshold θ_Ω for PTT were computed similar to a likelihood ratio test in hypothesis testing, i.e $\theta_\Omega = c \cdot 10^{\frac{SNR}{10}}$, where the constant parameter c, $(0 < c < 1)$ was tuned to reduce overall error probability. We first evaluate whether using a biased weighting approach enhances performance (in terms of P_{MD} and P_{FA}) compared to a conventional ℓ_1 technique and while using the iterative reweighting technique described in [21]. This comparison is shown in Figure 4.

We observe some general trends right away – increasing SNR or sampling rate reduces both P_{MD} and P_{FA} for all three techniques. This is expected, since a higher quality signal (or the lack of it) as well as additional samples improve both the detection and rejection performance of the system. Note, however, a slight discrepancy at 31 Hz and 20 dB for weighted ℓ_1 minimization. We believe this is due to inevitable differences in experimental setup across the runs. Simulation results in Figure 5 tend to follow the trend correctly.

Comparing first conventional and reweighted ℓ_1, we observe that reweighted ℓ_1 is always better (or no worse) than conventional ℓ_1 in terms of P_{MD} performance at all SNRs and sampling rates. False alarm performance is near perfect for both, but at the cost of high missed detection rates. Weighted ℓ_1, with a weight factor $w = 2$, results in lower P_{MD}, even at low SNRs. In particular, we see that weighted ℓ_1 with PTT at 40dB SNR performs better than conventional and reweighted ℓ_1 at 20 Hz and above. Again, this is not surprising because weighted ℓ_1, even with a slight bias, gravitates the solution of the minimization to a favorable state. This comes at the cost of higher P_{FA}, though, since there many cases where noise components coincide with the indices Ω. We focus more on this lost P_{FA} performance below.

Figure 5 illustrates results at 31 Hz across varying coefficient weights. Included in this set of plots are results from simulated runs of the same experiments. A few trends are visible here too – P_{MD} is a monotonically *non-increasing* function of w for fixed SNR and sampling rate and P_{FA} is a monotonically *non-decreasing* function of w. This conlusion is intuitive, owing to the fact that as w increases, the penalty on the indices Ω is subsequently reduced, promoting those indices in the solution (even if the signal was not present). Note that at high SNR, false alarm rates are negligible even at $w = 10$ for PTT, but are quite high for WTA. However, when SNR is low, WTA outperforms PTT slightly. We believe this is because WTA picks the maximum component in the FFT coefficients and even in low SNR regimes, the signal component is inclined to stand out.

It must be mentioned here that the improved performance of PTT has a cost associated with it. Selecting the right thresholds θ_Ω is non-trivial in cases where the signal is not completely captured within the samples being processed. Since the reconstruction is performed on a block of received data (in our case, we used 1 sec worth of samples) each time, the event signature may not be aligned with the block. In order

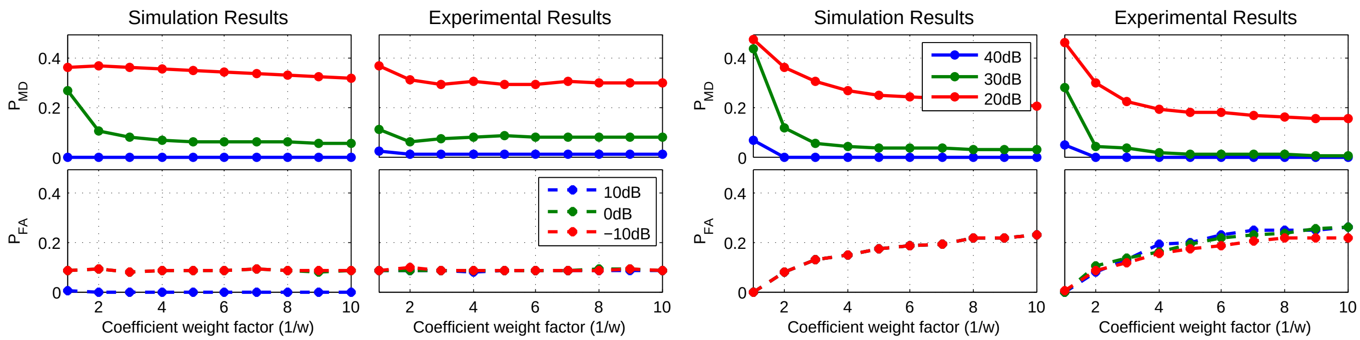

Fig. 6. P_{MD} and P_{FA} for PTT (left) and WTA (right) for various SNR at 20 Hz sampling.

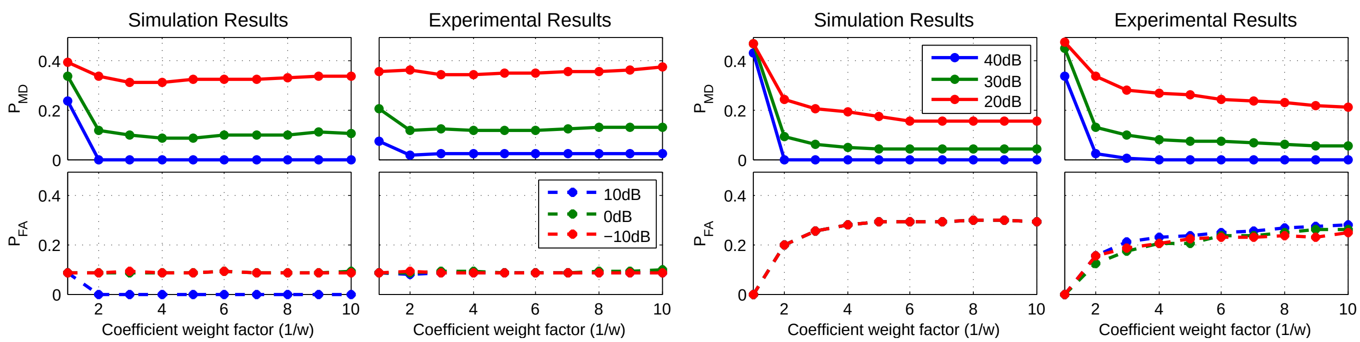

Fig. 7. P_{MD} and P_{FA} for PTT (left) and WTA (right) for various SNR at 10 Hz sampling.

to avoid this complexity and to demonstrate the asymptotic performance of the system, we applied a preprocessing step for experimental data to recover the correct alignment by padding the event signature with silence.

Another aspect worth noting is that with WTA, all SNRs result in the same P_{FA} performance, in simulation results. This is not a discrepancy. As the SNR increases, the chances of detecting the signal are higher if it is present (resulting in lower P_{MD}). But, when there is only noise (which happens for false alarms), SNR is effectively 0, so the same noise component gets picked everytime regardless of SNR. Note that this result is also a side effect of maintaining the same random noise seeds across the Monte Carlo simulation runs. This behavior is not present in the experimental results though the random noise generated is be the same because the noise in the recorded samples is affected by multiple factors, including ambient and circuit noise.

Figures 6 and 7 show the performance of the system for average sampling rates of 20 Hz and 10 Hz respectively. We observe that both P_{MD} and P_{FA} performance worsens as the sampling rate is reduced. This is because the feasible solution space that conforms to the polyhedra $z = AF^{-1}\tilde{y}$ expands to include many points that may be classified wrongly.

A. Event Signatures with Structure

In Figure 8, we illustrate results that test our conjecture that signals with structure may be detected more easily, with possibly fewer false alarms. A trivial signal structure results from two frequency components, so we add an equal amplitude tone at 150 Hz to the original event signature and repeat the experiments. To ensure that we recognize and exploit the structure correctly, we used a support vector machine (SVM [28]) classifier for detection rather than PTT or WTA.

Using this classifier required model selection, which was performed through randomized 10-fold cross validation and training, which was conducted with data from 50% of the simulation runs. The features used for classification were the magnitudes and angles of the complex FFT coefficients. To establish a fair comparison, the same detection procedure was also performed on the single frequency tone signature. Both events were randomly sampled at 31 Hz.

Surprisingly, we observe little or no improvement in the dual tone detection scenario. In fact, we observe some deterioration in P_{MD} performance. Upon inspection, we understand two reasons for this behavior. Firstly, since SNR is computed using the total signal power, which is now shared between two frequency components, individual indices constitute a lower power contribution against the same noise power. And secondly, by introducing an extra frequency tone, we have changed the sparsity of the event signature. From [7], we know that the number of measurements required for reconstruction is proportional to the sparsity of the signal. We surmise that it is the simple structure in the signature that offsets the deterioration resulting from both these issues.

V. Conclusion

We have presented a novel modification to the basis pursuit reconstruction procedure for known-signature event detection from sparse incoherent measurements. This modification was inspired by a recent development that uses an iterative

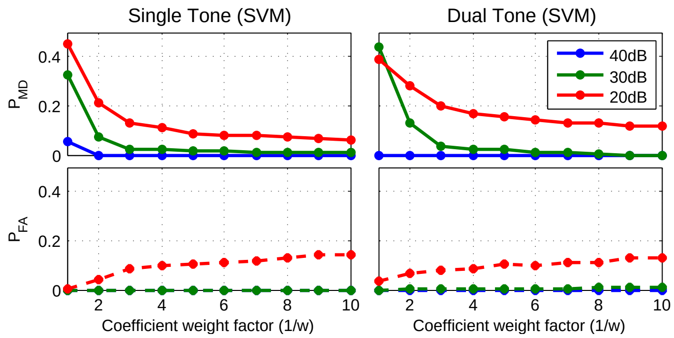

Fig. 8. P_{MD} and P_{FA} detecting a dual tone signal using an SVM classifier.

reweighting technique to equalize the penalty of each component during ℓ_1 minimization. We show through simulations and an implementation on MicaZ motes that this strategy is not only feasible at rates $30\times$ below the Nyquist requirement but that it results in higher detection rates at the cost of minimal false alarm performance compared to both conventional ℓ_1 and reweighted ℓ_1 regularization. We show that threshold testing performs better for detection when the event signature is aligned with the sample block, but that simply testing the maximum component's index works well in practice.

A key advantage of parameterized weighted ℓ_1 minimization is that it allows designers to tune the performance of the reconstruction procedure based on prior event probabilities, if known, and for acceptable detection and false alarm rates. This was not possible with either conventional ℓ_1 or reweighted ℓ_1 minimization, both of which rely on tuning parameters within the detection function. We also show results using a more sophisticated SVM event classifier that signals with some structure may be detected and rejected more reliably.

VI. Acknowledgments

This material is supported in part by the U.S. ARL and the U.K. MOD under Agreement Number W911NF-06-3-0001 and the U.S. Office of Naval Research under MURI-VA Tech Award CR-19097-430345. Any opinions, findings and conclusions or recommendations expressed in this material are those of the authors and do not necessarily reflect the views of the listed funding agencies. The U.S. and U.K. Governments are authorized to reproduce and distribute reprints for Government purposes not withstanding any copyright notation herein.

References

[1] J. Sauvageon, A. Agogino, A. Mehr, and I. Tumer, "Comparison of event detection methods for centralized sensor networks," *Sensors Applications Symposium, 2006. Proceedings of the 2006 IEEE*, pp. 95–100, 2006.

[2] P. Dutta, M. Grimmer, A. Arora, S. Bibyk, and D. Culler, "Design of a wireless sensor network platform for detecting rare, random, and ephemeral events," in *Proceedings of the 4th international symposium on Information processing in sensor networks*, IEEE Press Piscataway, NJ, USA, 2005.

[3] R. N. McDonough and A. Whalen, *Detection of Signals in Noise, 2nd edition, Academic Press*. 1995.

[4] V. Guralnik and J. Srivastava, "Event detection from time series data," in *KDD '99: Proceedings of the fifth ACM SIGKDD international conference on Knowledge discovery and data mining*, (New York, NY, USA), pp. 33–42, ACM, 1999.

[5] S. N. Patel, T. Robertson, J. A. Kientz, M. S. Reynolds, and G. D. Abowd, "At the flick of a switch: Detecting and classifying unique electrical events on the residential power line," *Ubicomp,*, 2007.

[6] S. Leeb, S. Shaw, and J. Kirtley, J.L., "Transient event detection in spectral envelope estimates for nonintrusive load monitoring," *Power Delivery, IEEE Transactions on*, vol. 10, pp. 1200–1210, Jul 1995.

[7] E. Candes, J. Romberg, and T. Tao, "Robust uncertainty principles: Exact signal reconstruction from highly incomplete frequency information," *IEEE Transactions on Information Theory*, vol. 52, no. 2, pp. 489–509, 2006.

[8] S. Chen, D. Donoho, and M. Saunders, "Atomic Decomposition by Basis Pursuit," *SIAM Journal on Scientific Computing*, vol. 20, no. 1, pp. 33–61, 1998.

[9] E. Candes and M. Wakin, "People hearing without listening: An introduction to compressive sampling," *IEEE Signal Processing Magazine*.

[10] E. Candes and T. Tao, "Decoding by linear programming," *IEEE Transactions on Information Theory*, vol. 51, no. 12, pp. 4203–4215, 2005.

[11] R. Baraniuk, M. Davenport, R. DeVore, and M. Wakin, "A simple proof of the restricted isometry property for random matrices," *Constructive Approximation*, 2008.

[12] M. Allen, L. Girod, R. Newton, S. Madden, D. T. Blumstein, and D. Estrin, "Voxnet: An interactive, rapidly-deployable acoustic monitoring platform," *IPSN08*.

[13] H. B. Wang, C. E. Chen, A. M. Ali, S. Asgari, R. E. Hudson, K. Yao, D. Estrin, and C. E. Taylor, "Acoustic sensor networks for woodpecker localization," 2005.

[14] J. Haupt and R. Nowak, "Compressive Sampling for Signal Detection," in *ICASSP*, 2007.

[15] M. Davenport, M. Duarte, M. Wakin, J. Laska, D. Takhar, K. Kelly, and R. Baraniuk, "The smashed filter for compressive classification and target recognition," in *SPIE*, 2007.

[16] T. Dang, N. Bulusu, and W. Hu, "Lightweight Acoustic Classification for Cane-Toad Monitoring," *Asilomar Conference on Signals, Systems and Computers*, 2008.

[17] M. Duarte, M. Davenport, M. Wakin, and R. Baraniuk, "Sparse Signal Detection from Incoherent Projections," in *ICASSP*, 2006.

[18] S. Mallat and Z. Zhang, "Matching pursuits with time-frequency dictionaries," *IEEE Transactions on Signal Processing*, vol. 41, no. 12, pp. 3397–3415, 1993.

[19] Y. Pati, R. Rezaiifar, and P. Krishnaprasad, "Orthogonal matching pursuit: recursive function approximation withapplications to wavelet decomposition," in *Signals, Systems and Computers, 1993. 1993 Conference Record of The Twenty-Seventh Asilomar Conference on*, pp. 40–44, 1993.

[20] J. Tropp and A. Gilbert, "Signal Recovery From Random Measurements Via Orthogonal Matching Pursuit," *Information Theory, IEEE Transactions on*, 2007.

[21] E. Candes, M. Wakin, and S. Boyd, "Enhancing sparsity by reweighted l1 minimization," *Journal of Fourier Analysis and Applications*, vol. 14, no. 5-6, pp. 877–905, 2008.

[22] A. Griffin and P. Tsakalides, "Compressed Sensing of Audio Signals Using Multiple Sensors," *EUSIPCO08*.

[23] S. Kirolos, J. Laska, M. Wakin, M. Duarte, D. Baron, T. Ragheb, Y. Massoud, and R. Baraniuk, "Analog-to-information conversion via random demodulation," in *DCAS*, 2006.

[24] F. Boyle, J. Haupt, G. Fudge, and C. Yeh, "Detecting Signal Structure from Randomly-Sampled Data," in *Statistical Signal Processing*, 2007.

[25] I. Bilinskis and A. Mikelson, *Randomized Signal Processing*. Prentice-Hall, Inc. Upper Saddle River, NJ, USA, 1992.

[26] H. Shapiro and R. Silverman, "Alias-free sampling of random noise," *J. Soc. Indust. Appl. Math*, vol. 8, no. 2, pp. 225–248, 1960.

[27] N. Johnson, S. Kotz, and N. Balakrishnan, *Continuous univariate distributions. Volume 2*. New York: Wiley, 1995.

[28] C.-C. Chang and C.-J. Lin, *LIBSVM: a library for support vector machines*, 2001. Software available at http://www.csie.ntu.edu.tw/~cjlin/libsvm.