REBEL: CONVEX RELAXATIONS FOR POISSON GLRT

Zachary Harmany[†] *Matthew Malloy*[†] *Aniruddha Bhargava*[†] *Nikhil Rao*[†] *Stephen Wright* [⋆]

[†]Department of Electrical and Computer Engineering
[⋆] Department of Computer Sciences
University of Wisconsin - Madison

ABSTRACT

This paper proposes an adaptive method for detection of sparse signals generated from a Poisson distribution, motivated by the problem of preventing nuclear fuel from crossing secure borders. The procedure, termed REBEL, is based on a convex relaxation of a generalized likelihood ratio test (GLRT). The relaxed problem is solved using iterative methods. In the case of 1-sparse signals, the proposed algorithm is shown to be optimal in terms of minimizing the false positive and false negative rates as the sample size grows.

***Index Terms*—** sparsity, anomaly detection, iterative algorithms, adaptivity

1. INTRODUCTION

Smuggling of nuclear material across secure borders is a concern for many nations. As techniques to hide illicit materials get more sophisticated, detection methods must also evolve and adapt in order to remain effective. To this end, recent developments in construction and design of large-scale detectors have made great progress in terms of hardware advancements [1]. From an estimation and detection perspective, breakthroughs in sparse signal recovery have shown great promise as a means to improve sensitivity of many detection problems [2, 3].

In this paper, we consider the problem of detecting illegal movements of nuclear fuel. Specifically, we consider a scenario where it is necessary to determine whether nuclear contraband is being taken across a security checkpoint. In such cases, vehicles are typically required to pass through an area that is sensed using Geiger counters. The personnel at the checkpoint need to determine whether nuclear contraband is present in the vehicle being scanned. It may be the case that a decision cannot be made in one pass of measurements. In such cases, one may sense the vehicle multiple times until a confident decision can be made about its contents.

In this paper we proposes an adaptive and iterative procedure, termed REBEL (Robust Estimation By Enhanced Likelihood), which relies on a convex relaxation of a generalized likelihood ratio test. The procedure is developed in both a non-adaptive and adaptive framework.

We model the observations collected by the sensors as draws from independent Poisson distributions whose intensity varies over space. This intensity is composed of a constant background count irrespective of the presence of a radioactive source, and a foreground component when there is a radioactive source. As we expect the presence of the contraband to be spatially concentrated in the vehicle, we may model this component as sparse. This leads to the formulation of a penalized log-likelihood ratio where we use the ℓ_0-norm to encourage sparsity.

Our testing approach is based on a sequential Generalized Likelihood Ratio Test (GLRT). At every time step, the proposed procedure decides to either make a decision or request more samples. This decision is made based on the result of an iterative soft thresholding procedure which enforces the sparsity constraint and reliably models the Poisson likelihood. We show that this problem can be relaxed and can be solved in real time. The proposed relaxed generalized likelihood ratio test (ReGLRT) is one of the novel contributions in this paper.

The remainder of the paper is organized as follows. In Sec. 2 we set up the problem we wish to solve in a mathematical framework, and summarize prior art in the field. In Sec 3 we introduce our method, REBEL: Robust Estimation By Enhanced Likelihood. We then analyze REBEL in Sec 4. We present experiments and report results in Sec 5. Finally, we conclude our paper in Sec 6 and provide avenues for further research.

2. PROBLEM STATEMENT

This paper considers the following problem. Assume a number of sensors are positioned in such a way that they measure the level of radioactivity emitted from a particular location. The response of the n sensors is a stochastic vector $\boldsymbol{y}_t \in \mathbb{Z}_+^n$ modeled as an inhomogeneous Poisson process. Specifically,

$$\boldsymbol{y}_t \sim \text{Poisson}(\lambda) \tag{1}$$

where the subscript t denotes multiple samples gathered sequentially in time. The parameter $\lambda \in \mathbb{R}_+^n$ represents the mean level of radioactivity observed at each sensors. The goal of the problem is to scan a vehicle in a region equipped with sensors and identify existence of contraband nuclear material.

The mean level of the Poisson process depends on whether or not the vehicle under test contains nuclear material. We postulate the following composite hypothesis testing problem, to detect the presence of contraband nuclear material:

$$\begin{aligned} H_0 &: \lambda = \mu \mathbf{1} \\ H_1 &: \lambda = \mu \mathbf{1} + \boldsymbol{s} \end{aligned} \tag{2}$$

where $\mu \in \mathbb{R}_+$ is a known scalar representing the ambient level of background radiation, $\mathbf{1}$ is the all ones vector, and $\boldsymbol{s} \in \mathbb{R}_+^n$ is an *unknown* k-sparse vector representing a heightened level of radioactivity in a small subset of the n sensors that are measuring the radioactive decay. The assumption that $\boldsymbol{s}$ is sparse reflects the fact that the radioactive material is not diverse spatially, but rather localized to a small portion of the area covered by the sensors. H_0 is termed the null hypothesis and indicates the absence of nuclear material, while H_1 is the alternative hypothesis, and indicates the presence of nuclear material.

The goal of the problem is to reliably identify vehicles that are attempting to smuggle radioactive material past the sensors. The performance of any test used to declare the presences or absence of nuclear material is characterized by the number of samples of the Poisson process, and the false negative and false positive rates:

$$\beta = \mathbb{P}(\text{declare } H_0 | H_1)$$

and

$$\alpha = \mathbb{P}(\text{declare } H_1 | H_0)$$

respectively.

2.1. Prior Work

Fixed sample size procedures for tests of composite hypothesis have been given extensive treatment in traditional estimation and detection literature [4]. Of particular relevance to the developments in this paper is the generalized likelihood ratio test (GLRT), which is optimal as the sample size grows [5,6].

In the sequential or adaptive setting the decision to take an additional sample of the random process at time t depends on the realization of the prior samples, $\{\boldsymbol{y}_1, \ldots, \boldsymbol{y}_{t-1}\}$. In the case of a simple binary hypothesis test, the sequential probability ratio test (SPRT) is optimal in terms of simultaneously minimizing both the expected number of samples and α and β. In the composite setting studied here, there are also a number of procedures based on variations of the GLRT [7,8]. To the best of our knowledge, there have been no adaptive procedures based on convex relaxations, that can be efficiently solved iteratively, of the GLRT for the problem of composite Poisson tests.

3. PROPOSED METHOD

We propose two methods, the first method highlights our key idea of using convex relaxations for the generalized likelihood ratio test. Assuming that $\boldsymbol{s}$ is k-sparse, the GLRT for this problem would form a statistic

$$\Lambda = \frac{\max_{\|\boldsymbol{s}\|_0 = k} \mathbb{P}(\boldsymbol{y}|\boldsymbol{s})}{\mathbb{P}(\boldsymbol{y}|\boldsymbol{s} = 0)}. \tag{3}$$

As it stands, (3) is intractable to solve efficiently, due to the presence of the ℓ_0 pseudo-norm. To alleviate this, we propose the relaxed GLRT:

$$\Lambda = \frac{\max_{\|\boldsymbol{s}\|_1 \leq \tau} \mathbb{P}(\boldsymbol{y}|\boldsymbol{s})}{\mathbb{P}(\boldsymbol{y}|\boldsymbol{s} = 0)} \tag{4}$$

where λ is a threshold parameter. The $\ell 1$ norm acts as a convex heuristic for the sparsity of the signal, and can be solved using iterative methods. If Λ is above a particular threshold τ, we declare H_1, and declare H_0 if below. We demonstrate the effectiveness of this procedure by constructing ROC curves in the experiments in Sec 5.

In the case where $k = 1$ in the unrelaxed version (3) we merely need to pick the maximal element from the observed variable $\boldsymbol{y}$, precluding the need to choose an appropriate λ and solve (4). For $k > 1$, we maximize the numerator of the GLRT in (4) by solving a sparsity-regularized Poisson inverse problem. In this minimization, estimating $\boldsymbol{s}$ from $\boldsymbol{y}$ can be a significant computational challenge. Previous work [9] developed methods for solving this inverse problem, and generalizations where there may be a linear observation matrix that aggregates photons for compressive measurements. A future direction for work would be adaptive design of the measurement process. Our optimization formulation uses a penalized negative Poisson log-likelihood objective function with nonnegativity constraints (since Poisson intensities are naturally nonnegative) in an iterative minimization framework. In particular, our approach incorporates key ideas of using separable quadratic approximations to the objective function at each iteration and penalization terms related to ℓ_1 norms of coefficient vectors. This approach is easily extended to other sparsity models than those considered here, such as total variation seminorms, and partition-based multi-scale estimation methods. Our approach is very stable relative to alternative iterative methods based on Augmented Lagrangians (AL) because there is no need to choose a good value of the AL parameter, which is a significant practical challenge [10].

To extend this to an measurement adaptive scenario, we consider the presence of two thresholds on the GLRT, τ_l and τ_u, $\tau_l < \tau_u$. We estimate $\boldsymbol{s}$ from the data and decide on hypothesis H_0 (low nuclear activity) if $\Lambda \leq \tau_l$. We decide on hypothesis H_1 (high nuclear activity) if $\Lambda \geq \tau_u$. Otherwise, we decide to obtain more samples and recompute Λ from additional data as to make a decision.

To summarize, we propose an adaptive algorithm that determines the following:

- The vehicle being scanned displays low levels of radioactivity, meaning that its safe to let it pass
- The vehicle being scanned displays high levels of radioactivity, meaning that there is contraband being smuggled, and certain action needs to be taken (in accordance with the laws of the land)
- It is not clear if there exists contraband in the vehicle, and we need to take more measurements, which aggregates photons across multiple measurements.

4. ANALYSIS

We define $S \subset \{1, 2, \ldots, n\}$ to be the set of active indices of $\boldsymbol{s}$. To simplify the analysis, we consider the 1-sparse case here. We denote the non-zero value in $\boldsymbol{s}$ by η. In the 1-sparse case, the log likelihood ratio is given by:

$$\begin{aligned}
\Lambda &= \frac{\arg\max_{\boldsymbol{s}:||\boldsymbol{s}||_0=1} \mathbb{P}(\boldsymbol{y}|\eta)}{\mathbb{P}(\boldsymbol{y}|\boldsymbol{s}=0)} \\
&= \frac{\prod_{i\neq i^*} \frac{\mu^{\boldsymbol{y}_i} e^{-\mu}}{\boldsymbol{y}_i!} \frac{(\mu+\eta)^{\boldsymbol{y}_S} e^{-(\mu+\eta)}}{\boldsymbol{y}_S!}}{\prod \frac{\mu^{\boldsymbol{y}_i} e^{-\mu}}{\boldsymbol{y}_i!}} \\
&= \frac{\frac{(\mu+\eta)^{\boldsymbol{y}_S} e^{-(\mu+\eta)}}{\boldsymbol{y}_S!}}{\frac{\mu^{\boldsymbol{y}_S} e^{-\mu}}{\boldsymbol{y}_S!}} \\
&= \left(1 + \frac{\eta}{\mu}\right)^{\boldsymbol{y}_S} e^{\eta}
\end{aligned}$$

This gives us an explicit formula for the log-likelihood:

$$\log \Lambda = \boldsymbol{y}_S \log\left(1 + \frac{\eta}{\mu}\right) \tag{5}$$

where in (5), we ignore the constant terms since they will not play a part in the optimization, when we look to maximize the likelihood ratio.

Notice that in the 1-sparse case, our MLE for the sparsity index is:

$$\hat{S} = \max_i \boldsymbol{y}_i$$

$$\boldsymbol{y}_{\hat{S}} = \eta + \mu \tag{6}$$

Substituting (6) in (5), we get the GLRT:

$$\log \Lambda = (\eta + \mu) \log\left(1 + \frac{\eta}{\mu}\right) \tag{7}$$

From [11] and the fact that we have a one-sided composite hypothesis test, we know that plugging in the MLE is optimal.

5. EXPERIMENTS AND NUMERICAL RESULTS

In this section we detail an experiment showing the promise of the REBEL algorithm for detecting sparse signals from Poisson observations. In this experiment, we consider a length $n = 100$ observation vector. We consider two levels of sparsity, one in which the sparse vector is 1-sparse (i.e., there is only a single location that contains appreciable foreground count), and one in which we have a $k = 10$ sparse vector. In both cases we have a background count of $\mu = 0.1$, and consider a sparse vector with $\sum_i s_i = 0.5$. This normalization ensures that we are not boosting the effective signal-to-noise ratio with decreasing sparsity, as well as ensuring that the resulting observations y are sufficiently photon limited.

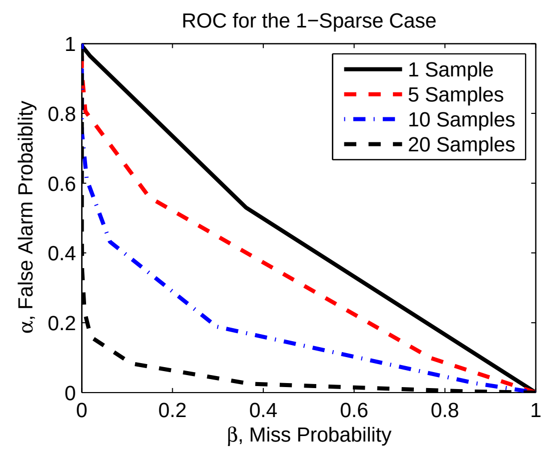

Fig. 1. ROC for the 1-sparse non-sequential case

In the first set of experiments, we consider the nonsequential GLRT algorithm where we are able to acquire multiple random measurements. Figures 1 and 2 show the detection performance achieved with this method as we sweep the threshold of the GLRT forming a ROC. Notice, as expected,

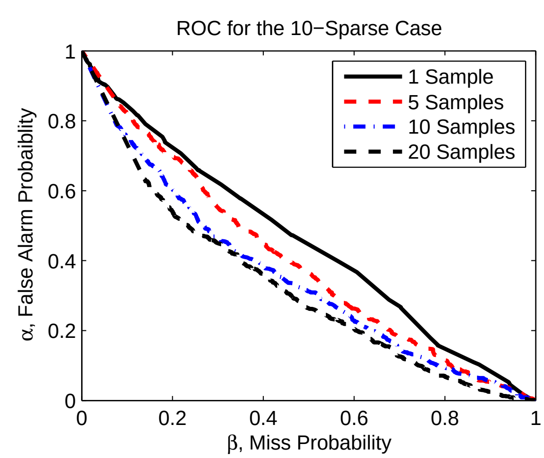

Fig. 2. ROC for the 10-sparse non-sequential case

an increasing number of measurements increase the overall performance of the detection method. In particular, we show results for 1, 5, 10, and 20 samples of the Poisson process. Note that for the $k = 10$ sparse case we solve an ℓ_1 regularized formulation of the problem, yielding a convex objective. This necessitates the selection of a regularization parameter, which was selected as $\tau = 1\text{e-}9$. While additional tuning of this parameter may improve performance, we simply fixed this parameter in all the numerical experiments. Additionally, this iterative method converges very quickly, in less than ten iterations.

In the second set of experiments, we demonstrate the performance of the adaptive REBEL - the sequential version of our algorithm. We consider the same experimental setup as before and select upper and lower thresholds of $\tau_u = 100$ and $\tau_l = 2$.

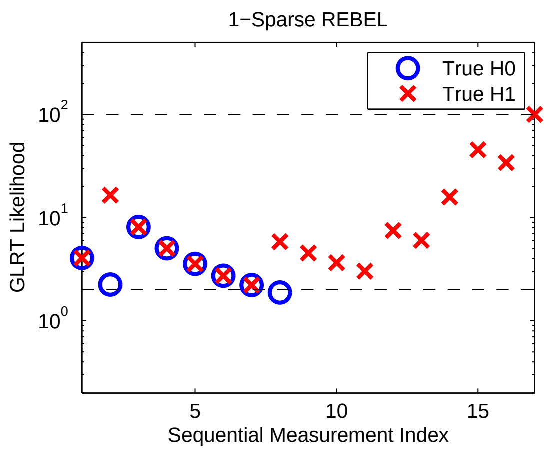

Fig. 3. Example for the 1-sparse adaptive case

Fig. 3 we consider the $k = 1$ sparse case. We consider two scenarios, one with the target absent (true H_0) and target present (true H_1). We see that we successfully accept H_0 after 8 sequential measurements, and accept H_1 after 17 measurements. In the $k = 10$ case, we successfully accept H_0 after 9 sequential measurements and sucessfully accept H_1 after only 4 sequential measurements.

As we can clearly see from our figures, REBEL is able to detect the contraband in the low SNR regime. This is true not only for the case of 1-sparsity but also when we have a higher sparsity level. This shows that methods like REBEL that include both and iterative (SPRT) and the GLRT components provide good recovery in practice. The combination of both of these has not been looked at closely to the best of our knowledge.

6. CONCLUSIONS AND FUTURE WORK

In this paper we proposed an adaptive and iterative method based on sequential composite tests, used to identify rebel vehicles at a security checkpoint. The method leverages a Poisson model to determine if vehicles contain nuclear material or not. In the case of one-sparse signals, we showed the proposed method is optimal.

One of the first extensions that we would like to consider is the case k-sparse Poisson signals. Extend the optimality results to this case would be of great interest. We would also like to consider situations where a k- sparse signal is subject to blurring kernel before being measured. The iterative algorithm used to constrained the sparsity can be extended to this scenario. Another extension is to aggregate responses from (possibly overlapping) groups of spatially co-located sensors. We conjecture that models used for group sparsity will aid in identifying much weaker activity in the alternate hypothesis.

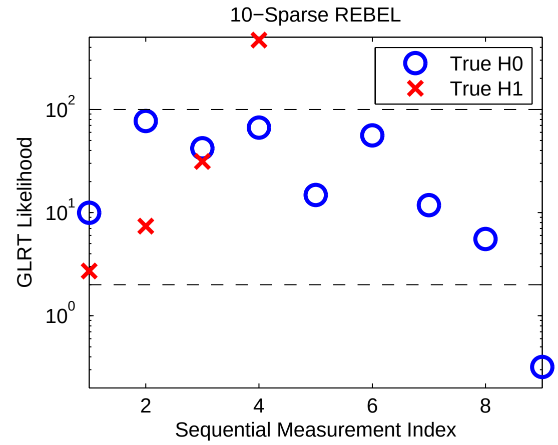

Fig. 4. Example for the 10-sparse adaptive case

7. REFERENCES

[1] M. Hohlmann, K. Gnanvo, L. Grasso, JB Locke, A. Quintero, and D. Mitra, "Design and construction of a first prototype muon tomography system with GEM detectors for the detection of nuclear contraband," in *Nuclear Science Symposium Conference Record (NSS/MIC), 2009 IEEE*. IEEE, 2009, pp. 971–975.

[2] N. Rao, B. Recht, and R. Nowak, "Universal measurement bounds for structured sparse signal recovery," in *Proceedings of AISTATS*, 2012, vol. 2102.

[3] M. Malloy and R. Nowak, "Sequential analysis in high-dimensional multiple testing and sparse recovery," in *Information Theory Proceedings (ISIT), 2011 IEEE International Symposium on*, 31 2011-aug. 5 2011, pp. 2661 –2665.

[4] S.M. Kay, "Fundamentals of statistical signal processing, volume ii: Detection theory," *Upper Saddle River (New Jersey)*, vol. 7, 1998.

[5] T.L. Lai, "Likelihood ratio identities and their applications to sequential analysis," *Sequential Analysis*, vol. 23, no. 4, pp. 467–497, 2004.

[6] G. Lorden, "Open-ended tests for koopman-darmois families," *The Annals of Statistics*, pp. 633–643, 1973.

[7] G. Fellouris and A. G. Tartakovsky, "Almost minimax sequential tests of composite hypotheses," *ArXiv e-prints*, Apr. 2012.

[8] D.R. Cox, "Sequential tests for composite hypotheses," in *Proc. Camb. Phil. Soc*. Cambridge Univ Press, 1952, vol. 48, pp. 290–299.

[9] Z. T. Harmany, R. Marcia, and R. Willett, "This is SPIRAL-TAP: Sparse Poisson Intensity Reconstruction ALgorithms Theory and Practice," *IEEE Transactions on Image Processing*, vol. PP, no. 99, 2011, DOI:10.1109/TIP.2011.2168410.

[10] Z. T. Harmany, J. Mueller, Q. Brown, N. Ramanujam, and R. Willett, "Tissue quantification in photon-limited microendoscopy," in *Proc. SPIE Optics and Photonics*, 2011.

[11] O. Zeitouni, J. Ziv, and N. Merhav, "When is the generalized likelihood ratio test optimal?," *Information Theory, IEEE Transactions on*, vol. 38, no. 5, pp. 1597 – 1602, sep 1992.