2014 IEEE International Conference on Acoustic, Speech and Signal Processing (ICASSP)

DESIGNING DISCRETE SEQUENTIAL TESTS VIA MIXED INTEGER PROGRAMMING

Michael Fauß and Abdelhak M. Zoubir

Signal Processing Group, Technische Universität Darmstadt
Merckstraße 25, 64283 Darmstadt, Germany
{mfauss,zoubir}@spg.tu-darmstadt.de

ABSTRACT

We show that the optimal design of non-randomized discrete sequential tests, i.e., tests whose test statistics take on only a countable number of states, can be modeled as a mixed integer linear problem. This is done by reformulating the difference equations describing the random walk on the integer lattice in terms of linear mixed integer constraints. We outline the general procedure and give a simple example to show how the proposed method can be used in practice.

Index Terms— detection, mixed integer programming, random walk, sequential analysis

1. INTRODUCTION

Due to its superior performance in time-critical environments, sequential testing [1] has attracted increased attention in recent years [2]. In many areas, sequential techniques have successfully been applied to reduce delays and/or increase detection performance. Among the driving applications in signal processing are spectrum sensing in cognitive radio [3, 4] and distributed detection in sensor networks [5, 6]. On the theoretical side, progress has been made in establishing performance bounds and asymptotic optimality [7]. The design of strictly optimal sequential tests, however, is still an open problem and is rarely addressed in the literature.

In this paper, we present an approach to design optimal discrete sequential tests using the framework of mixed integer programming. The recent progress in this branch of optimization theory resulted in a multitude of free and commercial solvers that are able to deal with small to medium sized problems in a sufficiently efficient way. By embedding discrete sequential testing in this rapidly evolving framework, we provide a generic design procedure that, for better or worse, does not rely on specific characteristics of a particular application. To the best of our knowledge, this is the first attempt to establish a connection between mixed integer programming and sequential testing.

This work was supported by the LOEWE Priority Program Cocoon (http://www.cocoon.tu-darmstadt.de) supported by the LOEWE research initiative of the state of Hesse/Germany.

We use the term discrete sequential tests to refer to tests whose test statistic can only take on a countable number of states such that every state can be mapped to an integer value. Theoretically, this holds true for every digital implementation of a sequential test. Realistically, however, this work targets scenarios where only a few bits are used for quantization. In distributed sensor networks, for example, each sensor might transmit a binary valued local decision to a fusion center, which in turn performs a sequential test on these messages – see the example in Section 4. In fact, there seems to be a tendency in distributed detection to use rough quantizations in order to reduce the communication load in large networks. The widely predicted rise in the prevalence of such networks emphasizes the need for efficient and powerful design algorithms for discrete tests.

The paper is organized as follows. In Section 2, we present the problem formulation and state the difference equations describing the general discrete sequential test. Based on these equations, we derive a mixed integer programming formulation of the test design problem in Section 3. The procedure is illustrated with an example in Section 4.

A word on notation: Boldface lower case letters $\boldsymbol{x}$ denote vectors and boldface upper case letters $\boldsymbol{X}$ matrices. The ith entry of a vector is written as $[\boldsymbol{x}]_i$. The identity matrix is denoted by $\boldsymbol{I}$ and the all-ones vector by $\mathbf{1}$. For the sake of a more compact notation, the dimensions are not always stated explicitly, but should be obvious from the context.

2. SYSTEM MODEL AND PROBLEM FORMULATION

We consider a sequence of random variables $(X_n)_{n\geq 1}$, defined on some product probability space $(\Omega^\infty, \mathcal{A}^\infty, P)$. Our aim is to perform a sequential statistical test between the two simple hypotheses

$$\mathcal{H}_0 : P = P_0 \quad \text{and} \quad \mathcal{H}_1 : P = P_1.$$

The test is performed by observing a sequence of discrete test statistic $(T_n)_{n\geq 1}$, where each $T_n : \Omega^n \to \mathbb{Z}$ maps the observations $X_1, \dots, X_n$ to an integer z_n. Without loss of generality, we assume that positive integers are associated with a decision for $\mathcal{H}_1$ while negative integers are associated with a decision for $\mathcal{H}_0$, meaning that if the test stops at a state T_n, it decides for $\mathcal{H}_1$, if $T_n > 0$, and for $\mathcal{H}_0$, if $T_n < 0$. The state $T_n = 0$ corresponds to events that do not allow any inference about the true hypothesis.

In accordance with Wald's sequential test [1], one can think of T_n as a quantized log likelihood ratio. This, however, is by no means the only meaningful test statistic. To interpret binary messages as quantized log likelihood ratios, for example, is usually more misleading than helpful. If Ω is a finite set, one can even avoid the use of a statistic and simply enumerate all possible outcomes directly. In order not to introduce unnecessary conceptual limitations, we do not specify the test statistic explicitly.

The more relevant question in sequential detection usually is not what test statistic to use, but when to stop the test. In this work, we consider only non-randomized tests. This means that at every state T_n the test either stops or continues with probability one. The possibility to include randomized stopping rules is briefly discussed in Section 3. The stopping rule is in the following denoted by $\delta : \mathbb{Z} \to \{0, 1\}$, where $\delta(z) = 1$ corresponds to stopping the test if the test statistic equals z and $\delta(z) = 0$ to continuing it.

Finally, we assume that the probability measures P_0 and P_1 can be specified via a set of transition probabilities $\{q_k(z, y) : z, y \in \mathbb{Z}\}$, where $k = 0, 1$ and

$$q_k(z, y) = P_k(T_{n+1} = y \,|\, T_n = z) \quad \forall n \geq 0.$$

Note that the transition probabilities are assumed to be independent of the time index n. This implies stationarity, but not necessarily independence of the sequence $(X_n)_{n\geq 1}$.

To describe the dynamic behavior of the test, we introduce the quantities $p_k(z)$, which denote the probability under P_k that the test ends with a decision for $\mathcal{H}_0$, given that the current test statistic equals z. Due to the stationarity assumption, these probabilities are time invariant and related via

$$p_k(z) = \begin{cases} (1 - \delta(z)) \sum_{y \in \mathbb{Z}} q_k(z, y) p_k(y), & z > 0 \\ \sum_{y \in \mathbb{Z}} q_k(z, y) p_k(y), & z = 0 \\ \delta(z) + (1 - \delta(z)) \sum_{y \in \mathbb{Z}} q_k(z, y) p_k(y), & z < 0. \end{cases} \tag{1}$$

Equation (1) is the well-known difference equation describing the dynamics of a generalized random walk on the integer lattice [8]. The role of the stopping rule $\delta(z)$ in (1) is to declare terminal states. If $\delta(z) = 1$, the state z terminates the random walk, otherwise it continues, according to the transition probabilities. Note that we do not allow the test to stop if there is no preference for either hypothesis, i.e., $\delta(0) = 0$. The error probabilities of the first and second kind are given by $1 - p_0(z_0)$ and $p_1(z_0)$, respectively, where $z_0 = T_0$ denotes the initial state of the test.

In a similar way, the expected run length $n(z)$ of a test starting at state z is governed by the difference equation

$$n(z) = (1 - \delta(z)) \left(1 + \sum_{y \in \mathbb{Z}} q(z, y) n(y) \right). \tag{2}$$

Again, $\delta(z)$ determines whether the test stops immediately, resulting in an average run length of zero, or makes a transition to some state y with corresponding average run length $n(y)$. Note that in (2), we leave the probability measure underlying the test unspecified. Common choices are P_0 or P_1 to consider the run length under each hypothesis. In a Bayesian framework, one would choose the sum measure $\pi_0 P_0 + (1 - \pi_0) P_1, \pi_0 \in (0, 1)$. In general, any measure that can be represented by time-invariant transition probabilities can be considered.

The difference equations (1) and (2) sufficiently describe the sequential test in terms of its run length and error probabilities. They provide the basis for the mixed integer formulation given in the next section. In particular, we present solutions to the classic sequential detection problem

$$\min_{\delta} \; n(z_0) \quad \text{s.t.} \quad p_0(z_0) \geq 1 - \alpha, \quad p_1(z_0) \leq \beta \tag{3}$$

and the sequential equivalent to the Neyman-Pearson test

$$\min_{\delta} \; p_1(z_0) \quad \text{s.t.} \quad p_0(z_0) \geq 1 - \alpha, \quad n(z_0) \leq \gamma \tag{4}$$

where $\alpha, \beta \in (0, 1)$ and $\gamma > 0$. The solution of (3) and (4) is an optimal stopping rule δ^* or, equivalently, an optimal stopping region $\mathcal{S}^*$ of the form

$$\mathcal{S}^* = \{z \in \mathbb{Z} : \delta^*(z) = 1\}.$$

At every time instant n, the optimal test continues, if $T_n \in \overline{\mathcal{S}^*} = \mathbb{Z} \setminus \mathcal{S}^*$, and stops if $T_n \in \mathcal{S}^*$.

So far we have implicitly assumed that the optimal test stops with probability one, i.e., eventually hits the stopping region. Tests for which this assumption does not hold exist, but are rather pathological and require an entirely different problem formulation, since their average run length is infinite. We therefore exclude them from our treatment.

3. SEQUENTIAL DETECTION AS A MIXED INTEGER PROBLEM

The mixed integer structure of the sequential test is obvious from (1) and (2), where δ and p_k are the unknown quantities. However, the bilinear terms on the right-hand side cause problems in the sense that they are hard to tackle numerically and are not supported by standard solvers. In this section, we restate the equations in a form that can be solved readily via mixed integer linear programming (MILP).

3.1. Reduction to a finite-dimensional state space

Before applying any numerical optimization technique to the sequential detection problem, we have to restrict its state space to a set of finite cardinality. This means that we have to determine some $\mathcal{N} \subset \mathbb{Z}$, with $|\mathcal{N}| < \infty$, for which we leave the stopping rule undetermined. For all $z \in \mathbb{Z} \setminus \mathcal{N}$, however, the stopping rule has to be chosen in advance. In other words, a subset $\tilde{\mathcal{S}} \subset \mathcal{S}^*$ of the optimal stopping region has to be known a priori.

Without loss of generality, we assume the initial estimate of the stopping region to be of the form

$$\delta(z) = 1, \quad \text{for} \quad |z| \geq A$$

where A is a positive integer. Since this renders all states $|z| \geq A$ equivalent to the states A or $-A$, it allows us to consider only the finite number of states $z \in \mathcal{N}_A = \{z \in \mathbb{Z} : |z| \leq A\}$. The transition probabilities have to be adjusted accordingly to

$$\tilde{q}_k(z, y) = q_k(z, y), \quad |y| < A$$
$$\tilde{q}_k(z, \pm A) = \sum_{l=A}^{\infty} q_k(z, \pm l).$$

Choosing A may appear critical at first glance, but usually does not pose a major problem. In the case of a quantized probability ratio test, for example, any upper bound on the absolute value of the optimal thresholds can be used. Such bounds can be obtained from Wald's approximations [1] and similar results [9]. In cases where bounds are not available, one might even work with just an initial guess, which can then be altered, depending on the outcome of the subsequent optimization. Given sufficient computational resources, one merely needs to choose A "large enough".

3.2. Reformulation of the system equations

Given a finite state space $\mathcal{N}_A$, we can collect $p_k(z)$, $n(z)$ and $\delta(z)$, with $z \in \mathcal{N}_A$, in vectors $\boldsymbol{p}_0, \boldsymbol{p}_1, \boldsymbol{n} \in \mathbb{R}^{2A+1}$ and $\boldsymbol{\delta} \in \{0,1\}^{2A+1}$ with entries

$$[\boldsymbol{p}_k]_i = p_k(i - A - 1)$$
$$[\boldsymbol{n}]_i = n(i - A - 1)$$
$$[\boldsymbol{\delta}]_i = \delta(i - A - 1)$$

where $i = 1, \dots, 2A+1$. The transition probabilities are accordingly collected in a matrix $\boldsymbol{Q} \in \mathbb{R}^{(2A+1)\times(2A+1)}$ with entries

$$[\boldsymbol{Q}]_{ij} = \tilde{q}(i - A - 1, j - A - 1).$$

In matrix vector notation, (1) and (2) can now be written as

$$[\boldsymbol{p}_k]_i = \begin{cases} (1 - [\boldsymbol{\delta}]_i) \cdot [\boldsymbol{Q}_k \boldsymbol{p}_k]_i, & i > A+1 \\ [\boldsymbol{Q}_k \boldsymbol{p}_k]_i, & i = A+1 \\ [\boldsymbol{\delta}]_i + (1 - [\boldsymbol{\delta}]_i) \cdot [\boldsymbol{Q}_k \boldsymbol{p}_k]_i, & i < A+1 \end{cases} \tag{5}$$

and

$$[\boldsymbol{n}]_i = (1 - [\boldsymbol{\delta}]_i) \cdot (1 + [\boldsymbol{Q}\boldsymbol{n}]_i). \tag{6}$$

In a more compact form, (5) and (6) become

$$\boldsymbol{p}_k = \boldsymbol{I}_A \boldsymbol{\delta} + (\boldsymbol{1} - \boldsymbol{\delta}) \odot \boldsymbol{Q}_k \boldsymbol{p}_k \tag{7a}$$
$$\boldsymbol{n} = (\boldsymbol{1} - \boldsymbol{\delta}) \odot (\boldsymbol{1} + \boldsymbol{Q}\boldsymbol{n}) \tag{7b}$$

where $\odot$ denotes element-wise multiplication,

$$\boldsymbol{I}_A = [\boldsymbol{e}_1, \dots, \boldsymbol{e}_A, \boldsymbol{0}, \dots, \boldsymbol{0}] \in \{0,1\}^{(2A+1)\times(2A+1)}$$

and $\boldsymbol{e}_i$ denotes the ith canonical basis vector of the $\mathbb{R}^{2A+1}$.

The general procedure to reformulate bilinear integer problems as MILP problems has recently been outlined in [10]. Naturally, our approach does not essentially differ, but exploits some characteristics of the problem at hand to simplify the expressions. The reformulation is stated in the following Lemma.

Lemma. Equations (5) and (6) can equivalently be formulated as

$$|(\boldsymbol{I} - \boldsymbol{Q}_k)\boldsymbol{p}_k| \leq \boldsymbol{\delta} \tag{8a}$$
$$|(\boldsymbol{I} - \boldsymbol{Q})\boldsymbol{n} - \boldsymbol{1}| \leq M\boldsymbol{\delta} \tag{8b}$$
$$\boldsymbol{I}_A \boldsymbol{\delta} \leq \boldsymbol{p}_k \leq \boldsymbol{1} - (\boldsymbol{I} - \boldsymbol{I}_{A+1})\boldsymbol{\delta} \tag{8c}$$
$$\boldsymbol{0} \leq \boldsymbol{n} \leq M(\boldsymbol{1} - \boldsymbol{\delta}) \tag{8d}$$

where $M \gg 0$ is a large, but otherwise arbitrary constant and all inequalities have to be read element-wise.

Proof. A proof of the lemma can be given by inspection. Setting $[\boldsymbol{\delta}]_i = 0$ in (8a) to (8d), yields

$$[\boldsymbol{p}_k]_i = [\boldsymbol{Q}_k \boldsymbol{p}_k]_i \tag{9a}$$
$$[\boldsymbol{n}]_i = 1 + [\boldsymbol{Q}\boldsymbol{n}]_i \tag{9b}$$
$$0 \leq [\boldsymbol{p}_k]_i \leq 1 \tag{9c}$$
$$0 \leq [\boldsymbol{n}]_i \leq M. \tag{9d}$$

Since $[\boldsymbol{p}_k]_i$ is a probability and M can be chosen arbitrarily large, the inequalities in (9c) and (9d) are non-binding. The equalities (9a) and (9b), on the other hand, correspond to (5) and (6) evaluated at $[\boldsymbol{\delta}]_i = 0$.

In the case $[\boldsymbol{\delta}]_i = 1$, equations (8a) to (8d) become

$$|[\boldsymbol{p}_k - \boldsymbol{Q}_k \boldsymbol{p}_k]_i| \leq 1 \tag{10a}$$
$$|[\boldsymbol{n} - \boldsymbol{Q}\boldsymbol{n}]_i - 1| \leq M \tag{10b}$$
$$[\boldsymbol{p}_k]_i = 0, \qquad \text{for} \quad i > A+1 \tag{10c}$$
$$[\boldsymbol{p}_k]_i = 1, \qquad \text{for} \quad i < A+1 \tag{10d}$$
$$[n]_i = 0. \tag{10e}$$

Again, inequalities (10a) and (10b) are non-binding because $|[\boldsymbol{p}_k - \boldsymbol{Q}_k \boldsymbol{p}_k]_i|$ is the absolute difference of two probabilities and M can be chosen arbitrarily large. Since the test stops at state i, if $[\boldsymbol{\delta}]_i = 1$, the equality constraints (10c) to (10e) enforce $[\boldsymbol{n}]_i = 0$ and set $[\boldsymbol{p}_k]_i$ to zero or one, depending on the decision associated with the state i. Note that $[\boldsymbol{\delta}]_{A+1} = \delta(0) = 0$ by definition. □

3.3. Restating the sequential testing problem

We are now in a position to state the sequential detection problems in Section 2 as mixed integer programs. Problem (3) is equivalent to

$$\min_{\boldsymbol{\delta}_k,\boldsymbol{p}_k} [\boldsymbol{n}]_{i_0} \text{ s.t. } \boldsymbol{\delta}_k \in \{0,1\}^{A-1},\ \boldsymbol{\delta}^{\mathrm{T}} = [\,1\ \ \boldsymbol{\delta}_0^{\mathrm{T}}\ \ 0\ \ \boldsymbol{\delta}_1^{\mathrm{T}}\ \ 1\,],$$
$$\boldsymbol{p}_k \in \mathbb{R}^{2A+1},\ [\boldsymbol{p}_0]_{i_0} \geq 1-\alpha,\ [\boldsymbol{p}_1]_{i_0} \leq \beta,$$
$$\text{and (8a) to (8d)}$$

where $k = 0, 1$ and i_0 denotes the initial state. Analogously, Problem (4) can be stated as

$$\min_{\boldsymbol{\delta}_k,\boldsymbol{p}_k} [\boldsymbol{p}_1]_{i_0} \text{ s.t. } \boldsymbol{\delta}_k \in \{0,1\}^{A-1},\ \boldsymbol{\delta}^{\mathrm{T}} = [\,1\ \ \boldsymbol{\delta}_0^{\mathrm{T}}\ \ 0\ \ \boldsymbol{\delta}_1^{\mathrm{T}}\ \ 1\,],$$
$$\boldsymbol{p}_k \in \mathbb{R}^{2A+1},\ [\boldsymbol{p}_0]_{i_0} \geq 1-\alpha,\ [\boldsymbol{n}]_{i_0} \leq \gamma,$$
$$\text{and (8a) to (8d).}$$

Some remarks are in order at this point:

- The mixed integer problem in fact optimizes over the *boundary conditions* of a system of difference equations. Only the on-off nature of the stopping rules allows us to express them as a vector of binary variables.
- Such non-randomized discrete sequential tests can only realize certain combinations of error probabilities and run lengths. To meet the constraints exactly, randomized stopping rules $\delta \in (0, 1)$ have to be used. Including the latter in the optimization is not straightforward, since the bilinear terms in the difference equations can no longer be handled case by case. Unfortunately, bilinear programming is even more computationally demanding than MILP and currently available algorithms are relatively inefficient. Further progress in this area might change the picture, though.
- The free parameter M can be interpreted as an upper bound on the worst-case average run length. Therefore, any large integer is a valid choice. In contrast to A, which determines the size of the state space, M does not influence the performance of the algorithm.

4. EXAMPLE

In order to illustrate the mixed integer approach to sequential detection we present a simple example from the context of distributed detection. Think of a remote sensor that makes i.i.d. observations $X_1, X_2, \ldots$ and forwards them to a fusion center, where a sequential test $\mathcal{H}_0 : X_n \sim \mathcal{N}(0, 1)$ vs. $\mathcal{H}_1 : X_n \sim \mathcal{N}(\mu, 1)$ is performed. The optimal test in this case is a classical likelihood ratio test. In order to save bandwidth and energy, however, the sensor might send binary decisions to the fusion center instead of accurate observations or likelihood ratios. In addition, the sensor could apply some censoring rule [11] such that only a fraction $c < 1$ of the decisions is actually sent to the fusion center.

We assume the following situation: The supposedly rare event $\mathcal{H}_1$ should be detected with high accuracy, say $\beta = 10^{-3}$, whereas the more frequent event $\mathcal{H}_0$ is allowed to be misclassified more often, say $\alpha = 0.03$. We want to design a sequential test that works on the censored, binary decision messages from the sensor and achieves at least the same α as the optimal test. We further choose to minimize β under the constraint that the expected number of *transmissions* under $\mathcal{H}_0$ is the same as for the optimal test. The optimization problem is accordingly given by

$$\min_{\delta} p_1(z_0) \quad \text{s.t.} \quad p_0(z_0) \geq 1-\alpha, \quad c \cdot n_0(z_0) \leq n_0^{\text{opt}}$$

where n_0^{opt} denotes the average run length of the optimal test under $\mathcal{H}_0$. As a discrete test-statistic, we use the sum of the binary messages, which we assume to be either 1 or -1, or 0 in the case of a censored transmission. The initial state is $z_0 = 0$. We further choose $\mu = 0.4$ and a censoring region symmetric around 0.5μ such that $c = 0.5$ under both $\mathcal{H}_0$ and $\mathcal{H}_1$. Using Wald's approximations for the design of the optimal test, we get the thresholds $T_{\text{up}}^{\text{opt}} \approx 3.51$ and $T_{\text{low}}^{\text{opt}} \approx -6.88$ and a lower bound on the average run length of $n_0^{\text{opt}} > 84.7$.

To solve the mixed integer problem, we used Gurobi 5.5 [12] via the CVX [13] frontend and set $M = 500$, $A = 30$. The result of the optimization are thresholds $T_{\text{up}} = 7$ and $T_{\text{low}} = -22$, yielding a 0.0267 probability of first kind errors and an average of 84.45 transmissions under $\mathcal{H}_0$. The probability of errors of the second kind, however, is reduced by almost two orders of magnitude to $p_1(0) = 1.23 \cdot 10^{-5}$. Moreover, even though it has not been considered in the optimization, the expected number of transmissions under $\mathcal{H}_1$ decreases from $n_1^{\text{opt}} > 43.86$ to $c\,n_1(0) = 27.91$.

With this example, we want to emphasize two aspects of discrete sequential tests and our proposed design algorithm. First, it shows that under non-standard design objectives, discrete tests can outperform their continuous counter parts, not only in terms of energy and spectral efficiency, but also in terms of test performance. Under the communication constraints considered here, even the very rough binary quantization leads to a surprisingly good result. Second, the example shows the flexibility that the mixed integer approach offers the test designer by allowing a wide variety of objectives in sequential testing to be handled within one consistent framework.

5. CONCLUSIONS

We have presented a mixed integer programming approach to the design of sequential tests and demonstrated its applicability by means of an example. Based on this work, we strongly advocate a closer connection between statistical test design and optimization theory in future research.

6. REFERENCES

[1] A. Wald, *Sequential Analysis*, John Wiley and Sons, 1st edition, 1947.

[2] T. L. Lai, "Sequential Analysis: Some Classical Problems and new Challenges," *Statistica Sinica*, vol. 11, pp. 303–408, 2001.

[3] L. Lai, Y. Fan, and H. V. Poor, "Quickest Detection in Cognitive Radio: A Sequential Change Detection Framework," in *Proc. of the IEEE Global Telecommunications Conference (GLOBECOM)*, 2008.

[4] F. Y. Suratman and A. M. Zoubir, "Collaborative Spectrum Sensing in Cognitive Radio Using Hard Decision Combining with Quality Information," in *Proc. of the IEEE Workshop on Statistical Signal Processing (SSP)*, 2011.

[5] R.S. Blum and B.M. Sadler, "Energy Efficient Signal Detection in Sensor Networks Using Ordered Transmissions," *IEEE Transactions on Signal Processing*, vol. 56, pp. 3229–3235, 2008.

[6] M. Fauß and A. M. Zoubir, "Performance Analysis of Sequential Detection for Collision Avoidance in Sensor Networks," in *Proc. of the IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP)*, 2013.

[7] Y. Mei, "Asymptotic Optimality Theory for Decentralized Sequential Hypothesis Testing in Sensor Networks," *IEEE Transactions on Information Theory*, vol. 54, pp. 2072–2089, 2008.

[8] W. Feller, *An Introduction to Probability Theory and its Applications. Vol. II*, Second edition. John Wiley & Sons Inc., New York, 1971.

[9] B. K. Ghosh and P. K. Sen, Eds., *Handbook of Sequential Analysis*, CRC, 1 edition, 1991.

[10] A. Gupte, S. Ahmed, M. Cheon, and S. Dey, "Solving Mixed Integer Bilinear Problems Using MILP Formulations," *SIAM Journal on Optimization*, vol. 23, pp. 721–744, 2013.

[11] C. Rago, P. Willett, and Y. Bar-Shalom, "Censoring Sensors: A Low-communication-rate Scheme for Distributed Detection," *Aerospace and Electronic Systems, IEEE Transactions on*, vol. 32, pp. 554–568, 1996.

[12] Gurobi Optimization, Inc., "Gurobi Optimizer Reference Manual," 2013.

[13] M. Grant and S. Boyd, "CVX: Matlab Software for Disciplined Convex Programming, version 2.0 beta," 2013.