Joint Estimation-Detection of Cyber Attacks in Smart Grids: Bayesian and Non-Bayesian Formulations

Ali Gaber†, Karim G. Seddik‡, and Ayman Y. Elezabi‡

†Department of Electrical Engineering, Alexandria University, Alexandria 21544, Egypt.
‡Electronics and Communications Engineering Department, American University in Cairo, AUC Avenue, New Cairo 11835, Egypt.
email: aligaber@alexu.edu.eg, kseddik@aucegypt.edu, aelezabi@aucegypt.edu

***Abstract*— Smart grid operations face a significant threat from the presence of cyber attacks or bad data that may contaminate the system observations. Therefore, in this paper, we are interested in introducing a new strategy for detecting the presence of bad data in smart grids and we also try to simultaneously estimate it in order to be able to separate the bad data from the system observations. We aim to obtain the attack free observations which reflect the true state of the smart grid. This can be done by defining a joint detection-estimation strategy based on Bayesian and non-Bayesian settings where the costs in general will be functions of the observation. We start with Bayes approach and derive the detector (which, in general, may not be a LRT) and then we set the problem by defining some maximum constraint under the null hypothesis based on the derived detector and minimize certain cost under the alternative hypothesis. Our results reveal that the proposed model is applicable on some cases that other models reported in previous works failed to deal with.**

I. Introduction

Smart grid state estimation has become an important issue in smart grid communications as it plays a key role in controlling the performance of power grids. It can be considered as a signal processing technique that makes use of the sensors and smart meters readings and convert them into an estimate of the system state [1]. System state is actually a term that describes the magnitudes and phase angles of the grid buses currents and voltages at different points. The state of these buses should be monitored accurately in order to have full control over the system operations. State estimation is also used to construct a real-time model of the grid network [2], [3]. A necessary function of state estimators is therefore to detect the existence of attacks (injected bad data) and/or measurement errors, and estimate this bad data in order to get rid of it.

Attack detection and estimation can be considered as a combined detection and estimation problem, and this problem has previously been treated in two different ways. The first one is to treat each subproblem separately using appropriate optimum test. For example, the detection subproblem can be established using the Neyman-Pearson optimum test, and the unknown parameter estimation subproblem can be established using the optimum Bayesian estimator. However, this "disjoint" approach does not guarantee optimum results [4]. The second approach is using the generalized likelihood ratio test (GLRT) which performs the detection and estimation simultaneously using the maximum likelihood (ML) estimator for the estimation part. However, this method too does not result in optimum performance as well since GLRT is generally a heuristic approach [4].

For the joint estimation and detection approach, there is little work in the literature, e.g. [4], [5], where, according to specific formulations and cost criteria, optimum solutions are obtained. In [4], [5], a purely Bayesian technique was reported which enhances the quality of detection and estimation by defining generalized cost functions. In [6], a technique is presented that combines between Bayesian and Neyman-Pearson methodologies by replacing the error probabilities under the two alternative hypothesis by estimation costs, restricting the cost under the nominal hypothesis, and optimizing the cost under the alternative hypothesis. In [7] the estimation cost under the hypothesis of bad data is minimized subject to constraints on false alarm and misdetection probabilities.

In this paper, we consider the joint detection-estimation problem based on Bayesian- and Neyman-Pearson-like formulations with application to smart grids. First, we consider the Bayesian approach by defining the detector cost functions as some measure of the estimation error. The cost functions are, in general, functions of the observations. We derive the optimum detector under this Bayesian formulation, which in general is not a LRT. Then, we consider a Neyman-Pearson-like formulation in which we minimize a generalized cost under some hypothesis given a cost constraint under the other hypothesis. Our formulation allows us to consider cases that the formulation in [7], [8] is not able to deal with. As will be explained later, there are singular cases in which the formulation in [7], [8] cannot be applied, e.g. the important case where the observations and the quantities to be estimated are jointly Gaussian when we apply the mean-square error (MSE) or minimum absolute-error (MAE) cost functions.

This work was supported by a grant from Intel Corporation - University Collaborative Research.

Note that the classical Neyman-Pearson approach considers the minimization of the probability of misdetection under a probability of false alarm constraint, which is different from our Neyman-Pearson-like approach in which we consider minimizing a generalized cost function under one hypothesis under a generalized cost constraint under the other hypothesis. Hence, the Neyman-Pearson approach may be considered a special case of our formulation with a specific choice of costs.

II. System Model

Assume that we have a system with linear dynamics. Therefore, the bad data injection can be modeled as follows:

$$\mathbf{y}_t = \mathbf{H}\mathbf{u} + \mathbf{x}_t + \mathbf{z}, \tag{1}$$

where $\mathbf{y}_t$ is the observation vector, $\mathbf{u}$ is the state vector of the system, $\mathbf{H}$ is the Jacobian matrix which indicates the dynamics of the system, $\mathbf{z}$ represents the noise term and $\mathbf{x}_t$ is the attack vector injected in the grid.

The attacker target is to focus all its energy, the "bad" data vector, in the range space of $\mathbf{H}$ and this occurs when the system dynamics $\mathbf{H}$ are fully known to the attacker, but what actually happens is that the attacker has partial information about the elements of $\mathbf{H}$ so there will be a non-zero projection of the bad data in the null space of $\mathbf{H}$ [7]. Then the bad data $\mathbf{x}_t$ can be decomposed into two projections one in the range space of $\mathbf{H}$ and the other one in the null space of $\mathbf{H}$. We will focus on the non-zero projection of the bad data vector in the null space of $\mathbf{H}$ and develop a joint detection and estimation technique to mitigate its effect. All the information about the bad data vector $\mathbf{x}_t$ will be embedded in its projection on the null space of the Jacobian matrix $\mathbf{H}$, so the projection of the observation vector on the null space of $\mathbf{H}$, denoted by $\mathbf{y}$, is given by:

$$\mathbf{y} = \mathbf{x} + \mathbf{n} \tag{2}$$

where $\mathbf{x}$ and $\mathbf{n}$ are the projections of $\mathbf{x}_t$ and $\mathbf{z}_t$ on the null space of $\mathbf{H}$, respectively.

Here our objective is to detect the existence of an attack and estimate the bad data injected in the system due to this attack. Therefore, our problem is a binary hypothesis testing problem with H_0 indicating the no bad data hypothesis and H_1 reflecting the bad data existence hypothesis. Thus, our observation model will be:

$$H_0 : \mathbf{y} = \mathbf{n} \tag{3}$$

$$H_1 : \mathbf{y} = \mathbf{x} + \mathbf{n}. \tag{4}$$

We make the standard assumptions that the conditional distributions of $\mathbf{y}$, under each hypothesis, denoted by $f_0(\mathbf{y})$ under H_0 and $f_1(\mathbf{y}|\mathbf{x})$ under H_1, are perfectly known and the marginal distribution of $\mathbf{x}$, denoted by $g(\mathbf{x})$, is also perfectly known.

III. Problem Formulation

We start our analysis by defining a joint detection-estimation strategy in a Bayesian setting where the cost of deciding hypothesis i given that the true hypothesis is j can be denoted as follows:

$$c_{ij}(\mathbf{y}) = E[C(\alpha, \mathbf{x})\ |\mathbf{y}, H_j] \tag{5}$$

where α is the estimate of $\mathbf{x}$ that in general depends on the decision i, and $C(\alpha, \mathbf{x})$ is some cost criterion, e.g. mean-square error. Therefore, the costs will be in general functions of the observation vector $\mathbf{y}$. We seek to define our joint detection-estimation strategy in a Bayesian setting where we derive a general expression for the detector decision rule and then obtain an estimate of the bad data injected in the system. We also formulate a second approach which is similar in some sense to the Neyman-Pearson approach by minimizing the expected cost under H_1 subject to a constraint on the average cost under H_0.

A. Bayesian Approach

In this subsection we pose our joint detection-estimation problem in a Bayesian setting, therefore this approach is applicable when the prior probabilities of each hypothesis can be obtained and when the costs have a meaningful assignment. Here our objective is to derive a general formula of the detector decision rule that minimizes the average cost which is given by:

$$E[C_{D(\mathbf{y})}] = \sum_{i=0}^{1}\sum_{j=0}^{1} c_{ij}(\mathbf{y}) \Pr(D(\mathbf{y}) = H_i, T = H_j) \tag{6}$$

where T is the true hypothesis and $D(\mathbf{y})$ is the detector decision. Since the costs are non-negative, the expected cost can be minimized by making the decision as follows:

$$c_{01}(\mathbf{y}) \Pr(H_1|\mathbf{y}) + c_{00}(\mathbf{y}) \Pr(H_0|\mathbf{y}) \underset{H_0}{\overset{H_1}{\gtrless}} c_{11}(\mathbf{y}) \Pr(H_1|\mathbf{y}) + c_{10}(\mathbf{y}) \Pr(H_0|\mathbf{y}), \tag{7}$$

where each term represents the cost of the corresponding decision.

The last equation can be simplified using Bayes rule to give the well-known general formula of our detector but where the costs are functions of the observation

$$\frac{P(\mathbf{y}|H_1)}{p(\mathbf{y}|H_0)} \frac{c_{01}(\mathbf{y}) - c_{11}(\mathbf{y})}{c_{10}(\mathbf{y}) - c_{00}(\mathbf{y})} \underset{H_0}{\overset{H_1}{\gtrless}} \eta, \tag{8}$$

where η is the decision threshold with optimum value equal to $\frac{P_0}{P_1}$ in the Bayesian formulation, and p_0 and p_1 are the prior probabilities of H_0 and H_1, respectively. It is obvious that this test is, in general, not a LRT.

B. Neyman-Pearson-Like Approach

In this subsection we formulate our joint detection-estimation problem as an optimization problem, where the objective is to minimize the average cost under H_1 under a constraint on the average cost under H_0. The problem can be formulated as follows:

$$\begin{aligned} &\min E[C(\alpha, \mathbf{x})|H_1] \\ &\text{s.t. } E[C(\alpha, \mathbf{x})|H_0] \leqslant \zeta. \end{aligned} \tag{9}$$

The average cost constraint under H_0 can be written as follows:

$$E[C(\alpha, \mathbf{x})|H_0] = \int_{\{\mathbf{y}:\, D(\mathbf{y})=H_0\}} c_{00}(\mathbf{y}) P(\mathbf{y}|H_0) d\mathbf{y} + \int_{\{\mathbf{y}:\, D(\mathbf{y})=H_1\}} c_{10}(\mathbf{y}) P(\mathbf{y}|H_0) d\mathbf{y} \leqslant \zeta. \tag{10}$$

where $\{\mathbf{y}: \; D(\mathbf{y}) = H_0\}$ and $\{\mathbf{y}: \; D(\mathbf{y}) = H_1\}$ are the regions where the detector decides in favor of H_0 and H_1, respectively. The optimum test in this case can be easily shown to be the Bayesian test in (8) for some η, where η in that case will be determined by the constraint.

Note that the formulation in (9) is different from the formulation in [7], [8] in which the joint estimation-detection problem is formulated as a minimization of the estimation cost under H_1 and making a decision of H_1 with false alarm and miss-detection probabilities constraints. As will be explained later, our formulation allows us to consider important cases that can not be considered with the formulation in [7], [8], as they form singular cases for the formulation in [7], [8].

IV. Case Study

In this section we apply our Bayesian approach on two different cost functions, namely, the mean square error (MSE) and the absolute error (AE). We show that our model is able to deal with some singular cases that the model reported in [7] fails to deal with. For the Neyman-Pearson-like approach, we consider the case where the costs under H_0 are defined in terms of the detection error probability and the costs under H_1 are defined in terms of the estimation error. This maps the constraint on the average cost under H_0 to be a false alarm probability (P_{fa}) constraint.

The observation model for our case study can be formulated as follows. Under H_0, the observation vector $\mathbf{y}$ is an $N \times 1$ noise vector and the variance of the i-th noise element is σ_{i0}^2 and the noise vector elements are assumed to be independent. Therefore, the distribution of the vector $\mathbf{y}$ under H_0, $f_0(\mathbf{y})$, is given by:

$$f_0(\mathbf{y}) = \frac{1}{(2\pi)^{\frac{N}{2}} \prod_{i=1}^N \sigma_{i0}} e^{-\sum_{i=1}^N \frac{y_i^2}{2\sigma_{i0}^2}}, \tag{11}$$

where the noise is assumed to have zero mean.

Under H_1, $\mathbf{y} = \mathbf{x} + \mathbf{n}$. The bad data vector is assumed to be Gaussian and is independent of the noise, and thus $\mathbf{x}$ and $\mathbf{y}$ are jointly Gaussian vectors. The conditional distribution of $\mathbf{x}$ conditioned on a certain observation $\mathbf{y}$ and assuming that the true hypothesis is H_1 is given by:

$$f(\mathbf{x}|\mathbf{y}, H_1) = \frac{1}{(2\pi)^{\frac{N}{2}} \left|\mathbf{C}_{\mathbf{x}|\mathbf{y},H_1}\right|^{\frac{1}{2}}} e^{(\mathbf{x}-\mu_{\mathbf{x}|\mathbf{y},H_1})^T \mathbf{C}_{\mathbf{x}|\mathbf{y},H_1}^{-1} (\mathbf{x}-\mu_{\mathbf{x}|\mathbf{y},H_1})}, \tag{12}$$

where $\mathbf{C}_{\mathbf{x}|\mathbf{y},H_1}$ is the conditional covariance matrix and $\mu_{\mathbf{x}|\mathbf{y},H_1}$ is the conditional mean vector and are given as follows:

$$\mathbf{C}_{\mathbf{x}|\mathbf{y},H_1} = \mathbf{C}_{\mathbf{xx}} - \mathbf{C}_{\mathbf{xy},H_1} \mathbf{C}_{\mathbf{yy},H_1}^{-1} \mathbf{C}_{\mathbf{yx},H_1} \tag{13}$$

$$\mu_{\mathbf{x}|\mathbf{y},H_1} = \mu_{\mathbf{x}} + \mathbf{C}_{\mathbf{xy},H_1} \mathbf{C}_{\mathbf{yy},H_1}^{-1} (\mathbf{y} - \mu_{\mathbf{y},H_1}) \tag{14}$$

where $\mathbf{C}_{\mathbf{xy},H_1}$ is the cross-covariance matrix between $\mathbf{x}$ and $\mathbf{y}$ under H_1.

A. Bayesian Approach for Gaussian Data with MSE Costs

For MSE, the cost is given by $c(\alpha, \mathbf{x}) = ||\alpha - \mathbf{x}||^2$ and the general formulas of the costs will be given as

$$c_{00} = 0 \tag{15}$$

$$c_{01}(\mathbf{y}) = E(||\mathbf{x}||^2|\mathbf{y}, H_1) \tag{16}$$

$$c_{10}(\mathbf{y}) = \|E(\mathbf{x}|\mathbf{y}, H_1)\|^2 \tag{17}$$

$$c_{11} = \text{Trace}\left(\mathbf{C}_{\mathbf{x}|\mathbf{y},H_1}\right) \tag{18}$$

where $c_{01}(\mathbf{y})$ is the cost of making a decision in favor of H_0 (i.e., $\alpha = \mathbf{0}$) while the true hypothesis is H_1. Note that $c_{01}(\mathbf{y}) = E(||\mathbf{0} - \mathbf{x}||^2|\mathbf{y}, H_1)$. Similarly c_{10} is the cost of making a decision in favor of H_1 (i.e., $\alpha = E(\mathbf{x}|\mathbf{y}, H_1)$) while the true hypothesis is H_0. Note that for the jointly Gaussian case and MSE, c_{11} is not a function of $\mathbf{y}$ and therefore this case cannot be addressed by the formulation presented in [7].

Note that for the case of MMSE, the term $\frac{c_{01}(\mathbf{y})-c_{11}(\mathbf{y})}{c_{10}(\mathbf{y})-c_{00}(\mathbf{y})}$ is equal to one so the decision rule is a likelihood ratio test (LRT). For Gaussian noise and bad data with prior Gaussian distribution, the decision rule will be given by:

$$-\frac{1}{2}(\mathbf{y} - \mu_{\mathbf{y},H_1})^T \mathbf{C}_{\mathbf{yy},H_1}^{-1} (\mathbf{y} - \mu_{\mathbf{y},H_1}) + \sum_{i=1}^N \frac{y_i^2}{2\sigma_{i0}^2} \underset{H_0}{\overset{H_1}{\gtrless}} \ln(\eta) - \ln \prod_{i=1}^N \sigma_{i0} + \ln |\mathbf{C}_{\mathbf{yy},H_1}|^{1/2} \tag{19}$$

B. Bayesian Approach for Gaussian Data with AE Costs

For AE, the cost function is defined as $c(\alpha, \mathbf{x}) = \sum_{i=1}^N |\alpha_i - x_i|$. In this case, the decision rule will not be a likelihood ratio test as we can readily show that if we substitute in the term $\frac{c_{01}(\mathbf{y})-c_{11(\mathbf{y})}}{c_{10}(\mathbf{y})-c_{00}(\mathbf{y})}$ it will not be equal to one. Therefore our detector will be:

$$-\frac{1}{2}(\mathbf{y} - \mu_{\mathbf{y},H_1})^T \mathbf{C}_{\mathbf{yy},H_1}^{-1} (\mathbf{y} - \mu_{\mathbf{y},H_1}) + \sum_{i=1}^N \frac{y_i^2}{2\sigma_{i0}^2} + \ln\left(\frac{c_{01}(\mathbf{y}) - c_{11}}{c_{10}(\mathbf{y})}\right) \underset{H_0}{\overset{H_1}{\gtrless}} \ln(\eta) - \ln \prod_{i=1}^N \sigma_{i0} + \ln |\mathbf{C}_{\mathbf{yy},H_1}|^{1/2}. \tag{20}$$

where it is difficult to obtain closed-form expressions for the costs in the vector case. In this case, the minimum AE estimator is the median of the conditional distribution of $\mathbf{x}$ on $\mathbf{y}$, which is equal to the conditional mean for the Gaussian case. Note that in this case c_{11} is again not a function of the observation $\mathbf{y}$, and the formulation in [7] cannot address this case either.

For the scalar case, and after some mathematical manipulations, we have

$$c_{00} = 0 \tag{21}$$

$$c_{01}(y) = \frac{2}{\sqrt{2\pi}} \sigma_{x|y,H_1} e^{\frac{\mu_{x|y,H_1}}{2\sigma_{x|y,H_1}}} + \mu_{x|y,H_1} \left(1 - 2Q\left(\frac{\mu_{x|y,H_1}}{\sigma_{x|y,H_1}}\right)\right) \tag{22}$$

Note that the MMSE estimator is the conditional mean estimator under H_1, which is applied throughout this subsection.

$$c_{10}(y) = \mu_{x|y,H1}\left(1 - 2Q\left(\frac{\mu_{x|y,H_1} - \mu_{x|y,H_0}}{\sigma_{x|y,H_0}}\right)\right) \quad (23)$$

$$c_{11} = \frac{2}{\sqrt{2\pi}}\sigma_{x|y,H_1} \quad (24)$$

where $Q(.)$ is the Gaussian Q-function, $\mu_{x|y,H_i}$ is the conditional mean of x given y under the hypothesis H_i, and $\sigma^2_{x|y,H_i}$ is the conditional variance of x given y under H_i. Clearly, c_{11} is not a function of y.

C. Neyman-Pearson-Like Approach for scalar Gaussian Data with MSE Costs

In this subsection we apply our model to the scalar Gaussian case with MSE estimation cost and assume the costs under H_0 are defined in terms of the estimation error. According to our cost assignment using MSE, $c_{00}(y)$ is equal to zero. The constraint on the cost under H_0 is given by

$$E[C(\alpha, x)|H_0] = \int_{\{y:\, D(y)=H_1\}} c_{10}(y)P(y|H_0)dy \leqslant \zeta. \quad (25)$$

In this case the optimum detector can be readily proved to be given by

$$y^2 \underset{H_0}{\overset{H_1}{\gtrless}} \beta \quad (26)$$

where β is related to the detector threshold η as follows:

$$\beta = \left(\ln(\eta) - \ln\left(\frac{\sigma_0}{\sigma_1}\right)\right)\frac{2\sigma_0^2\sigma_1^2}{\sigma_1^2 - \sigma_0^2}. \quad (27)$$

The relation between the constraint ζ on the average cost under H_0 and the threshold of detection β can be obtained as follows:

$$\sqrt{\pi}\left(\frac{1 - erf(\sqrt{\beta/(2\sigma_0^2)})}{4(1/2\sigma_0^2)^{\frac{3}{2}}}\right) + \frac{\sqrt{\beta}e^{(-\beta/2\sigma_0^2)}}{2(1/2\sigma_0^2)} = \zeta\frac{\sqrt{2\pi}\sigma_0}{2}\left(1 + \frac{\sigma_0^2}{\sigma_x^2}\right)^2 \quad (28)$$

For a given ζ we can obtain β which is used to give the minimum average cost under H_1. It is worth mentioning that the average cost under H_0 is monotonically decreasing in $\ln(\eta)$, and the average cost under H_1 is monotonically increasing in η. This means that the optimal detector will select a threshold β that always satisfies the constraint on average cost under H_0 in (25) with equality.

1) Cost assignment with false alarm probability constraint: Our formulation allows us to set the costs in any general form. One way to set the costs is to define the costs under H_1 to be estimation costs and the costs under H_0 to be detection costs. Clearly under H_0, x is not present so it is more meaningful to care more about detection errors under H_0 which is not the case under H_1 where we care more about estimation errors. We can reformulate our problem by defining the costs under H_0 as functions of the detection error as

$$\begin{aligned} c_{00} &= 0 \\ c_{10} &= 1 \end{aligned} \quad (29)$$

and the costs under H_1 as functions of the estimation error

$$\begin{aligned} c_{01}(y) &= E[C(0, x)|y, H_1] \\ c_{11}(y) &= E[C(\alpha, x)|y, H_1]. \end{aligned} \quad (30)$$

This modification allows us to be able to minimize the average cost under H_1 subject to a certain constraint on the probability of false alarm, as in this case the average cost under H_0 will map to the false alarm probability P_{fa}. The constraint can be written as follows:

$$E[C(\alpha, x)|H_0] = \int_{y \in \hat{H}=H_1} P(y|H_0)dy = P_{fa} \leqslant \zeta \quad (31)$$

and then our problem will be as follows:

$$\begin{aligned} &\min E[C(\alpha, x)|H_1] \\ &\quad \text{s.t. } P_{fa} \leqslant \zeta. \end{aligned} \quad (32)$$

For the case of scalar Gaussian MSE and by solving (31), the relation between detector threshold β and ζ is found to be

$$\beta = \sigma_0 Q^{-1}\left(\frac{\zeta}{2}\right). \quad (33)$$

In this case β is given by the solution of the following equation:

$$y^2\left(\frac{\sigma_1^2 - \sigma_0^2}{2\sigma_0^2\sigma_1^2}\right) + 2\ln|y| = \ln(\eta) + \ln\left(\frac{\sigma_1}{\sigma_0}\left(\frac{\sigma_x^2 + \sigma_0^2}{\sigma_x^2}\right)^2\right). \quad (34)$$

Note that the formulation in [7] aims at minimizing the cost c_{11} under a false alarm and miss-detection probabilities constraints. The model in [7] considers the estimation error in the case of c_{11} only, but our model is more general as we propose a general expression for the costs in (5), and according to the definition of the costs we can consider different types of errors such as estimation error or detection error in many cases.

Therefore the model in [7] can not deal with the Gaussian model reported under different estimation error costs. If we try to apply the model in [7] to our Gaussian model, we will find that the probability of detection will have two discrete values, either zero or one, and so is the probability of false alarm. So P_{det} and P_{fa} will be equal to each other and we will not be able to trade off the detection quality by increasing P_{mis} or (decreasing P_{det}) to enhance the estimation quality as the values P_{mis} can take are zero or one.

V. Numerical Results

In this section, we present our simulations. We let N be the number of measuring units that transmit their measurements to the system controller; this number of measuring units corresponds to the length of the received observation vector $\mathbf{y}_t$.

For the case of Bayesian approach with Gaussian measurements and MSE cost, we run our simulations for $N = 10$ measuring units with $p_0 = 0.8$ and $p_1 = 0.2$. The bad data variance is set to be 20 and the noise variance is set

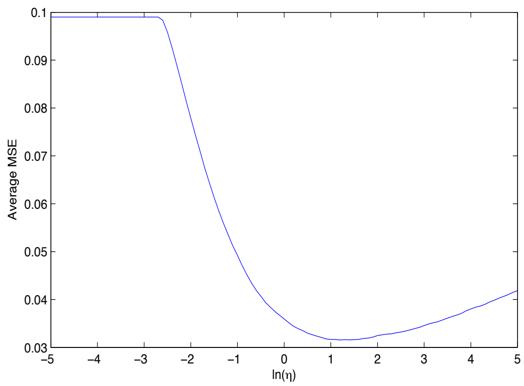

Fig. 1: Average MSE vs $\ln(\eta)$ for $N = 10$ and $p_0 = 0.8$ and $p_1 = 0.2$.

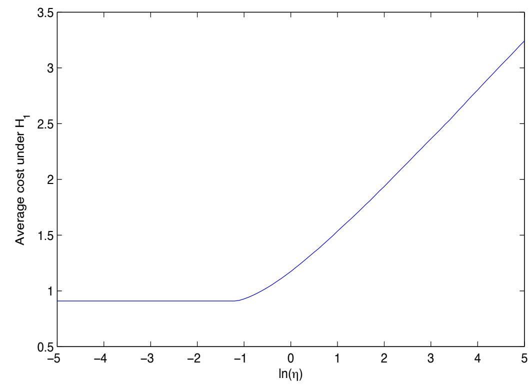

Fig. 3: Average cost under H_1 is monotonically increasing with $\ln(\eta)$.

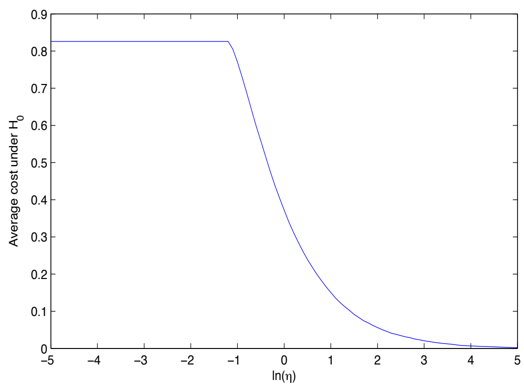

Fig. 2: Average cost under H_0 is monotonically decreasing with $\ln(\eta)$.

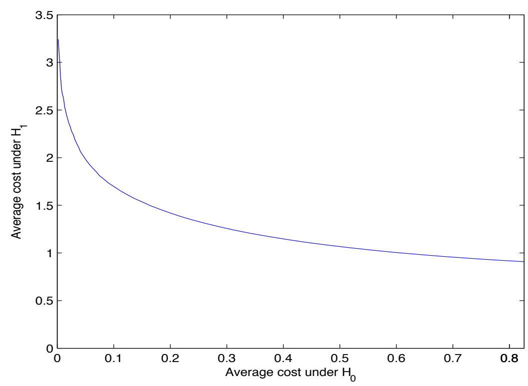

Fig. 4: Average cost under H_1 vs Average cost under H_0.

to be 1. The minimum average mean square error occurs at $\ln(\eta_{opt}) = \ln(\frac{0.8}{0.2}) = 1.38$ as shown in Fig. 1, which is the optimum threshold as derived above.

Next, we consider the case of Neyman-Pearson-like setting. In the sequel, the bad data variance is set to be 10 and the noise variance is set to be 1. First, we consider the scalar Gaussian MSE model and run the simulations for $\zeta = 0.27$ (which is the MSE cost constraint under the null hypothesis). Again, in this case, and as expected, the optimum threshold as derived above matches the optimum threshold obtained from the simulation results as will be shown below.

As shown in Fig. 2, the average cost under H_0 is monotonically decreasing with $\ln(\eta)$, but the average cost under H_1 in Fig. 3 is monotonically increasing with $\ln(\eta)$ as mentioned above. Fig. 4 shows that if the average cost under H_0 is lower than ζ then the average cost under H_1 will increase, therefore, the minimum value of the average cost under H_1 occurs when the average cost under H_0 equals ζ, i.e. the constraint is satisfied with equality.

Next, we consider the Neyman-Pearson-like case and reformulate our problem as in eqn. (32) to map the average cost under H_0 to the false alarm probability (P_{fa}). We run the simulations for $\zeta = 0.45$ and as expected the optimum threshold is $\ln(\eta_{opt}) = -1.7$ which is consistent with the optimum threshold as derived above. Using the same reasoning as before, we can deduce that the average cost under H_1 is monotonically increasing with $\ln(\eta)$ as shown in Fig. 5 and from Fig. 6 we can see that the probability of false alarm (P_{fa}) is monotonically decreasing with $\ln(\eta)$. and the relation between the average cost under H_1 and P_{fa} is shown in Fig. 7.

For the same problem stated in eqn. (32), we run the simulation but for the vector case where $N = 10$ measuring units and $P_{fa} = 0.45$. As expected from the analysis above, the value of the optimum threshold is $\ln(\eta_{opt}) = -4.9$ and this value gives the minimum average cost under H_1 according to our constraint on P_{fa} as shown in Fig. 8.

Note that we do not compare the performance of our formulation with the formulation in [7], [8] for the following two reasons. Firstly, the two formulations consider different cost functions so comparing the two formulations, each will be better than the other if compared based on the cost function used in the former formulation. Secondly, and more impor-

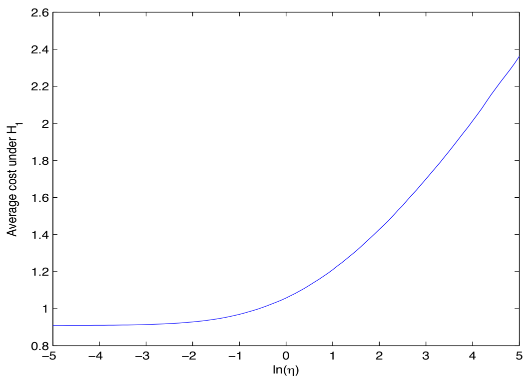

Fig. 5: Average cost under H_1 is monotonically incresing with $\ln(\eta)$ for the new formulation.

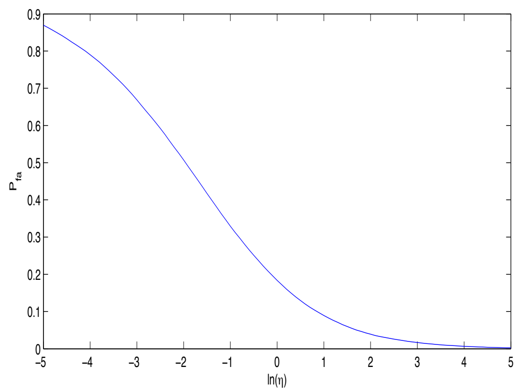

Fig. 6: Average cost under H_0 "P_{fa}" is monotonically decreasing with $\ln(\eta)$ for the new formulation.

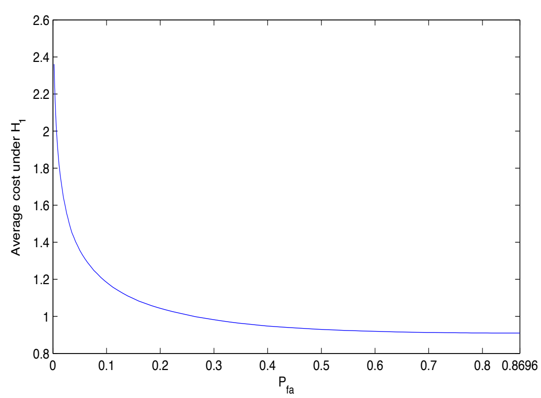

Fig. 7: Average cost under H_1 vs P_{fa} for the new formulation.

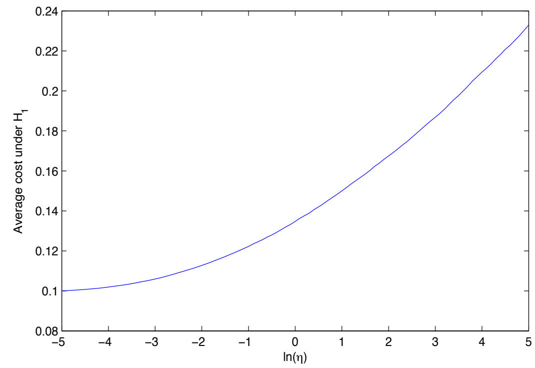

Fig. 8: Average cost under H_1 vs $ln(\eta)$ for the vector case $N = 10$ measuring units.

tantly all the cases considered in our simulations represent singular cases for the formulation in [7], [8] as explained above.

VI. Conclusion

In this paper we have introduced a joint detection-estimation strategy for the bad data injected in the smart grid system. We have defined two approaches based on the Bayesian setting and Neyman-Pearson-like formulation using general assignment of the costs by setting a maximum constraint on the cost under nominal hypothesis and optimize the cost under the alternative hypothesis to trade off between the quality of estimation and the accuracy of detection. Our formulation allows us to consider some singular cases that the previous works in the same track failed to deal with. These cases include the important case where the observations and the quantities to be estimated are jointly Gaussian when we apply the mean-square error (MSE) or minimum absolute-error (MAE) cost functions.

References

[1] F.C. Schweppe, "Power system static-state estimation, part iii: Implementation," *Power Apparatus and Systems, IEEE Transactions on*, vol. PAS-89, no. 1, pp. 130–135, Jan 1970.

[2] F.F. Wu, K. Moslehi, and A. Bose, "Power system control centers: Past, present, and future," *Proceedings of the IEEE*, vol. 93, no. 11, pp. 1890–1908, Nov 2005.

[3] A. Monticelli, "Electric power system state estimation," *Proceedings of the IEEE*, vol. 88, no. 2, pp. 262–282, Feb 2000.

[4] D. Middleton and R. Esposito, "Simultaneous optimum detection and estimation of signals in noise," *Information Theory, IEEE Transactions on*, vol. 14, no. 3, pp. 434–444, May 1968.

[5] M. Rezaeian and Ba-Ngu Vo, "Error bounds for joint detection and estimation of a single object with random finite set observation," *Signal Processing, IEEE Transactions on*, vol. 58, no. 3, pp. 1493–1506, March 2010.

[6] George V Moustakides, "Finite sample size optimality of glr tests," *arXiv preprint arXiv:0903.3795*, 2009.

[7] A. Tajer, S. Kar, H.V. Poor, and Shuguang Cui, "Distributed joint cyber attack detection and state recovery in smart grids," in *Smart Grid Communications (SmartGridComm), 2011 IEEE International Conference on*, Oct 2011, pp. 202–207.

[8] G.V. Moustakides, G.H. Jajamovich, A. Tajer, and Xiaodong Wang, "Joint detection and estimation: Optimum tests and applications," *Information Theory, IEEE Transactions on*, vol. 58, no. 7, pp. 4215–4229, July 2012.