A NEW APPROACH FOR SEMIPARAMETRIC DETECTION

Asmaa Amehraye[1] *and Lionel Fillatre*[2]

[1]ESIGETEL, Laboratoire de Recherche et d'Innovation Technologique,
1 rue de port de Valvins, 77210, Avon, France
[2]STMR, UMR CNRS 6279, Université de technologie de Troyes
12 rue Marie Curie - BP 2060 - 10010, Troyes, France,

ABSTRACT

Semiparametric detection consists of combining the statistical optimality of a parametric test to the robustness regarding the data of a nonparametric test. This approach is specially interesting in presence of statistical hypotheses depending on unknown probability distributions. The proposed semiparametric approach consists of splitting the measurement vector into two parts such that the first part has a known statistical distribution. Then, it is proposed to calculate a likelihood ratio test based both on the first part and the detection result of a nonparametric test applied to the second part. The statistical performance of the proposed test is analytically established.

Index Terms— Semiparametric detection, Likelihood ratio test, Nonparametric test, Support vector machine.

1. INTRODUCTION

The nature of statistical hypothesis testing depends upon what is known about the data used in the test. If the probability of the data sample conditioned upon the hypothesis is known to within a finite set of parameters then parametric hypothesis tests can be utilized [1]. However, if the aforementioned probability is unknown, or depends on a too large number of unknown parameters, then nonparametric tests must be used [2]. In general, with all other things being equal, the parametric tests perform better than the nonparametric tests due to the additional information imparted by the probability distribution but nonparametric tests have robust performance due to the rather mild assumptions made regarding the data [3].

In practice, there is a considerable interest to propose some tests, namely the semiparametric tests, which combine the advantages of each approach. Previous works [4, 5] have already studied such an approach. There are mainly two trends in the literature. The first one [4, 6] consists of estimating the nonparametric part of the model in a first step and to design a test in a second step based on this estimate. The second one [5] consists of combining two statistical tests, a parametric one and a nonparametric one, obtained from two different sources of data. This is also known as a fusion test.

This work was supported in part by the Grant ANR-08-SECU-013-02.

This paper proposes a new approach. Often, a measurement vector comes from a measurement model which contains both a parametric part and a nonparametric part. Hence, it is proposed to split the measurement vector into two subvectors such that each subvector is either associated to a pure parametric model or a pure nonparametric one. The nonparametric subvector is then processed by a nonparamatric test in order to produce a binary decision value whose statistical distribution is well estimated. This nonparametric decision value is then combined with the parametric subvector to produce the final decision. This combination is parametric in essence since it is based on the well-known likelihood ratio. From this way, the nonparametric component endows the semiparametric test with the flexibility necessary to capture complex regularities in the data and the parametric part of the semiparametric test provides a robust description of some of the patterns present in the data to ensure a constraint of the false alarm probability.

The paper is organized as follows. Section 2 presents the detection problem and the classical approaches that can be used to solve it. It is then proposed to solve this detection problem by the semiparametric test described in Section 3. Section 4 studies the performances of the proposed test. Section 5 concludes this paper.

2. DETECTION PROBLEM

Let $\boldsymbol{z} = \boldsymbol{\theta} + \boldsymbol{\xi}$ be the random measurement vector of $\mathbb{R}^n$ where $\boldsymbol{\xi}$ is a random vector with the Gaussian distribution $\mathcal{N}(0, \sigma^2 \boldsymbol{I}_n)$ where σ is known and $\boldsymbol{I}_n$ is the identity matrix. The goal is to decide whether $\boldsymbol{\theta}$ is zero or not. If not zero, the mean value $\boldsymbol{\theta}$ is assumed to approximatively belong to a subspace of $\mathbb{R}^n$ with a small dimension, i.e.,

$$\boldsymbol{z} = \boldsymbol{P}\boldsymbol{x} + \boldsymbol{Q}\boldsymbol{y} + \boldsymbol{\xi}, \tag{1}$$

where the matrices $\boldsymbol{P}$, of size $n \times p$, and $\boldsymbol{Q}$, of size $n \times q$, are known such that $p < q$. The space spanned by $\boldsymbol{P}$ is supposed to contain the major part of $\boldsymbol{\theta}$ and $\boldsymbol{Q}$ spans the complementary subspace of $\boldsymbol{P}$ in $\mathbb{R}^n$. This complementary subspace can contain some traces of $\boldsymbol{\theta}$ which are not well modeled. Such

a model can occur for example in [7] where the signal of interest belongs to a subspace of $\mathbb{R}^n$. For simplicity, the values of the vectors $\boldsymbol{x} \in \mathbb{R}^p$ and $\boldsymbol{y} \in \mathbb{R}^q$ depend on the two hypotheses :

$$\begin{aligned} \mathcal{H}_0 &: \{\boldsymbol{x} = 0\,,\ \ \boldsymbol{y} = 0\}, \\ \mathcal{H}_1 &: \{\boldsymbol{x} = \boldsymbol{x}_0\,,\ \ \boldsymbol{y} \sim P_1\} \end{aligned} \tag{2}$$

where $\boldsymbol{x}_0$ is known and P_1 is an unknown probability distribution. Hence, under $\mathcal{H}_1$, the distribution of the random vector $\boldsymbol{y} \in \mathbb{R}^q$ is partially unknown. Without any loss of generality, it is assumed that $\mathrm{rank}\,(\boldsymbol{P}) = p$, $\mathrm{rank}\,(\boldsymbol{Q}) = q$ and $\mathrm{rank}\,([\boldsymbol{P}\ \boldsymbol{Q}]) = p+q = n$. The column vectors of $\boldsymbol{P}$ and those of $\boldsymbol{Q}$ form a family of orthonormal vectors. It is supposed that a learning data set

$$\mathcal{S} = \Big\{(\boldsymbol{z}^1, \ell_1), (\boldsymbol{z}^2, \ell_2), \ldots, (\boldsymbol{z}^N, \ell_N)\Big\}$$

of N independent and identically distributed measurements vectors $\boldsymbol{z}^i$, where $\ell_i \in \{0, 1\}$ is the label of $\boldsymbol{z}^i$, is available. The detection problem (2) is difficult because hypothesis $\mathcal{H}_1$ is composite : the vector $\boldsymbol{y}$ is unknown. A statistical test is a function of $\boldsymbol{z}$ into $\{0, 1\}$ which takes the value i if hypothesis $\mathcal{H}_i$ is accepted. Two classical tests are possible to solve this problem : a parametric approach, which ignores $\mathcal{S}$, and a nonparametric approach, which exploits $\mathcal{S}$.

2.1. Parametric detection approach

Using a parametric detection approach consists of assuming that $\boldsymbol{y}$ is an unknown deterministic vector under $\mathcal{H}_1$. This yield to the parametric detection problem

$$\begin{aligned} \mathcal{H}_0 &: \{\boldsymbol{z} \sim \mathcal{N}(0, \sigma^2 \boldsymbol{I}_n)\}, \\ \mathcal{H}_1 &: \{\boldsymbol{z} \sim \mathcal{N}(\boldsymbol{\theta}, \sigma^2 \boldsymbol{I}_n),\ \boldsymbol{\theta} \neq 0\} \end{aligned} \tag{3}$$

where $\boldsymbol{\theta} = \boldsymbol{P}\boldsymbol{x}_0 + \boldsymbol{Q}\boldsymbol{y}$ is an unknown non-zero vector. Hence the goal is to detect any non-zero deviation $\boldsymbol{\theta}$ in the expectation of $\boldsymbol{z}$. The optimal test $\delta_p(\boldsymbol{z})$ is given in [8] :

$$\delta_p(\boldsymbol{z}) = \begin{cases} 0 \ \text{ if } \ \Lambda_p(\boldsymbol{z}) \overset{\text{def.}}{=} \frac{\|\boldsymbol{z}\|^2}{\sigma^2} \leq \lambda_p, \\ 1 \ \text{ else,} \end{cases} \tag{4}$$

where $\|.\|$ is the Euclidean norm and λ_p is a threshold to satisfy a false alarm probability $\alpha \in [0, 1]$, i.e., $\mathrm{Pr}_0(\Lambda_p(\boldsymbol{z}) \geq \lambda_p) = \alpha$ where $\mathrm{Pr}_k(A)$ stands for the probability of A when $\boldsymbol{z}$ is distributed according to $\mathcal{H}_k$. It is well known that the performance of $\delta_p(\boldsymbol{z})$, especially its probability of correct detection, depends only on the norm of $\boldsymbol{\theta}$, which may be very restrictive.

2.2. Nonparametric detection approach

Using a nonparametric detection approach consists of assuming that $\boldsymbol{y}$ is an unknown random vector whose distribution can be learned from $\mathcal{S}$. To simplify the presentation, it is assumed that the problem is solved by using a Support Vector Machines (SVM) detector. The SVM detector has the form [9] :

$$\delta_n(\boldsymbol{z}) = \begin{cases} 0 \ \text{ if } \ \Lambda_n(\boldsymbol{z}) \leq \lambda_n, \\ 1 \ \text{ else,} \end{cases} \tag{5}$$

where the decision function has the form

$$\Lambda_n(\boldsymbol{z}) = \sum_{i=1}^{N} \gamma_i \, K(\boldsymbol{z}^i, \boldsymbol{z}). \tag{6}$$

Here, $K(\boldsymbol{z}^i, \boldsymbol{z})$ is the kernel function using a similarity measure between the observations $\boldsymbol{z}^{(i)}$ and $\boldsymbol{z}$. The weighted coefficients γ_i are obtained by minimizing a prefixed cost based on the learning set $\mathcal{S}$ (see details in [9]). The threshold λ_n is automatically adjusted by the minimization step. In practice, it is very difficult to fix it in order to respect a prescribed false alarm probability.

These two approaches have their own advantages but, overall, their main disadavantages. The parametric detector has well controlled error probabilities but there is a loss of optimality since the random nature of $\boldsymbol{y}$ is ignored. The nonparametric exploits the random nature of $\boldsymbol{y}$ but its statistical performances are not well controlled. Hence, it is proposed to combine these two approaches in a unified framework to obtain a better detector, called the semiparametric detector.

3. SEMIPARAMETRIC DETECTION

The principle of the semiparametric test can be described as follows. The measurement vector is assumed to be split into two statistically independent subvectors. The first subvector is processed with a nonparametric detector of the form (5) which produces a decision value 0 or 1. The first subvector together with the decision value of the nonparametric test form a couple whose distribution is known under each hypothesis. The likelihood ratio test is calculated from this couple of random variables. This yields to the semiparametric test.

3.1. Statement of the semiparametric detection problem

The measurement vector can be decomposed into two subvectors $\boldsymbol{z}_p$ and $\boldsymbol{z}_q$ defined by

$$\boldsymbol{z}_p \overset{\text{def.}}{=} \boldsymbol{P}^\top \boldsymbol{z} = \boldsymbol{x} + \boldsymbol{\xi}_p, \tag{7}$$

$$\boldsymbol{z}_q \overset{\text{def.}}{=} \boldsymbol{Q}^\top \boldsymbol{z} = \boldsymbol{y} + \boldsymbol{\xi}_q \tag{8}$$

where $\boldsymbol{P}^\top$ (resp. $\boldsymbol{Q}^\top$) is the transpose of $\boldsymbol{P}$ (resp. $\boldsymbol{Q}$), $\boldsymbol{\xi}_p \sim \mathcal{N}(0, \sigma^2 \boldsymbol{I}_p)$ and $\boldsymbol{\xi}_q \sim \mathcal{N}(0, \sigma^2 \boldsymbol{I}_q)$. It is obvious that knowing the vector $\boldsymbol{z}$ is statistically equivalent to know the couple $(\boldsymbol{z}_p, \boldsymbol{z}_q)$. From this decomposition, $\boldsymbol{z}_p$ has a well-defined distribution under both $\mathcal{H}_0$ and $\mathcal{H}_1$. Hence it can be exploited to design a parametric test. On the contrary, the distribution of

$\boldsymbol{z}_q$ is badly characterized under both hypothesis, hence it must be exploited by a nonparametric test. we note that $\boldsymbol{z}_p$ and $\boldsymbol{z}_q$ are statistically independent since the noise $\boldsymbol{\xi}$ is Gaussian distributed with a diagonal covariance matrix.

Let $\delta_n(\boldsymbol{z}_q)$ the SVM detector defined by (5)-(6) be based on the subvector $\boldsymbol{z}_q$ instead of the complete vector $\boldsymbol{z}$. This SVM detector is learned from the restricted data set

$$\mathcal{S}_q = \Big\{(\boldsymbol{z}_q^1, \ell_1), (\boldsymbol{z}_q^2, \ell_2), \ldots, (\boldsymbol{z}_q^N, \ell_N)\Big\}$$

where $\boldsymbol{z}_q^i = \boldsymbol{Q}^\top \boldsymbol{z}^i$. The false alarm probability of $\delta_n(\boldsymbol{z}_q)$ is $\widetilde{\alpha}$=$\mathrm{Pr}_0(\delta_n(\boldsymbol{z}_q)$=1) and its probability of correct detection is $\widetilde{\beta}$=$\mathrm{Pr}_1(\delta_n(\boldsymbol{z}_q)$=1). These two probabilities are assumed to be known (or well estimated). Consequently, under $\mathcal{H}_0$, the random variable $\delta_n(\boldsymbol{z}_q)$ follows the Bernoulli distribution $\mathcal{B}(\widetilde{\alpha})$: it takes the value 1 with probability $\widetilde{\alpha}$ and the value 0 with probability $1-\widetilde{\alpha}$. Under $\mathcal{H}_1$, the random variable $\delta_n(\boldsymbol{z}_q)$ follows the Bernoulli distribution $\mathcal{B}(\widetilde{\beta})$. The proposed semiparametric approach consist of solving the decision problem

$$\begin{aligned}\underline{\mathcal{H}}_0 &: \Big\{\boldsymbol{z}_p \sim \mathcal{N}(0, \sigma^2\boldsymbol{I}_p),\ \delta_n(\boldsymbol{z}_q) \sim \mathcal{B}(\widetilde{\alpha})\Big\},\\ \underline{\mathcal{H}}_1 &: \Big\{\boldsymbol{z}_p \sim \mathcal{N}(\boldsymbol{x}_0, \sigma^2\boldsymbol{I}_p),\ \delta_n(\boldsymbol{z}_q) \sim \mathcal{B}(\widetilde{\beta})\Big\} \quad (9)\end{aligned}$$

by using the famous likelihood ratio test. This test is optimal for testing $\underline{\mathcal{H}}_0$ and $\underline{\mathcal{H}}_1$ since these hypotheses whose statistical distributions are well known are simple [1]. Obviously, the problem (9) is not equivalent to the initial decision problem (3) : the statistical information in vector $\boldsymbol{z}_q$ is condensed, and certainly reduced, by the nonparametric test $\delta_n(\boldsymbol{z}_q)$.

3.2. Semiparametric test

The decision problem (9) can be solved by the well-known log-likelihood ratio test given by

$$\delta_s(\boldsymbol{z}) = \begin{cases} 0 \text{ if } \Lambda_s(\boldsymbol{z}) \leq \lambda_s, \\ 1 \text{ else,} \end{cases} \quad (10)$$

with the decision function $\Lambda_s(\boldsymbol{z})$:

$$\Lambda_s(\boldsymbol{z})=\Lambda_s(\boldsymbol{z}_p, \delta_n(\boldsymbol{z}_q))=\log \frac{f_{\boldsymbol{x}_0}(\boldsymbol{z}_p)\, \widetilde{\alpha}^{\delta_n(\boldsymbol{z}_q)}\, (1-\widetilde{\alpha})^{1-\delta_n(\boldsymbol{z}_q)}}{f_0(\boldsymbol{z}_p)\, \widetilde{\beta}^{\delta_n(\boldsymbol{z}_q)}\, (1-\widetilde{\beta})^{1-\delta_n(\boldsymbol{z}_q)}}$$

where $f_{\boldsymbol{x}}(\boldsymbol{z}_p)$ is the probability density function of the Gaussian distribution $\mathcal{N}(\boldsymbol{x}, \sigma^2\boldsymbol{I}_p)$. Calculation yields to

$$\Lambda_s=\Lambda_s(\boldsymbol{z}_p, \delta_n(\boldsymbol{z}_q))=\varrho\left(\frac{\boldsymbol{x}_0^\top \boldsymbol{z}_p}{\sigma\|\boldsymbol{x}_0\|} + \frac{\gamma}{\varrho}\,\delta_n(\boldsymbol{z}_q) + \frac{\omega}{\varrho}\right) \quad (11)$$

where $\gamma = \log\left(\dfrac{\widetilde{\beta}(1-\widetilde{\alpha})}{\widetilde{\alpha}(1-\widetilde{\beta})}\right)$ and $\omega = \log\dfrac{1-\widetilde{\alpha}}{1-\widetilde{\beta}}$

and ϱ^2=$\|\boldsymbol{x}_0\|^2/\sigma^2$ is the Parametric Signal-to-Noise Ratio (PSNR) associated to the parametric part of the measurement model. It must be noted that the test (10) can be simplified as follows : it is sufficient to compare the decision function $\Lambda_s^*(\boldsymbol{z})$ to a threshold λ_s^* instead of comparing $\Lambda_s(\boldsymbol{z})$ to λ_s, where $\Lambda_s^*(\boldsymbol{z})$ is given by :

$$\Lambda_s^*(\boldsymbol{z})=\Lambda_s^*(\boldsymbol{z}_p, \delta_n(\boldsymbol{z}_q))=\frac{\boldsymbol{x}_0^\top \boldsymbol{z}_p}{\|\boldsymbol{x}_0\|\sigma} + \frac{\gamma}{\varrho}\,\delta_n(\boldsymbol{z}_q). \quad (12)$$

From (12), it appears that the parameter γ measures the global performances of the nonparametric test $\delta_n(\boldsymbol{z}_q)$. When γ is large, it can seriously modify the performance of the semiparametric test. Hence, the ratio γ/ϱ automatically quantifies the tradeoff between the parametric part $\boldsymbol{x}_0^\top \boldsymbol{z}_p/(\|\boldsymbol{x}_0\|\sigma)$ of the test and the nonparametric decision value $\delta_n(\boldsymbol{z}_q)$. Furthermore, from (12), it is straightforward to obtain

$$\alpha_s \overset{\text{def.}}{=} \mathrm{Pr}_0(\Lambda_s^* \geq \lambda_s^*)=Q(\lambda_s^*)(1-\widetilde{\alpha})+Q\left(\lambda_s^*-\frac{\gamma}{\varrho}\right)\widetilde{\alpha}, \quad (13)$$

$$\beta_s \overset{\text{def.}}{=} \mathrm{Pr}_1(\Lambda_s^* \geq \lambda_s^*)=Q(\lambda_s^*-\varrho)(1-\widetilde{\beta})+Q\left(\lambda_s^*-\frac{\gamma}{\varrho}-\varrho\right)\widetilde{\beta} \quad (14)$$

$$\text{where } Q(x) = \frac{1}{\sqrt{2\pi}}\int_x^{+\infty} \exp(-\frac{u^2}{2})du. \quad (15)$$

The theoretical calculation of the threshold λ_s^* to satisfy a prescribed level α is not an easy task but its numerical computation is very easy. It permits to obtain a good control of the false alarm probability α_s.

3.3. Robustness to the learning data set

Obviously, in practice, the performance indices α_s and β_s of the semiparametric test are some random variables $\widehat{\alpha}_s$ and $\widehat{\beta}_s$ which depend on the learning data set $\mathcal{S}_q$ through $\widetilde{\alpha}$ and $\widetilde{\beta}$. In fact, the "true" probabilities $\widetilde{\alpha}$ and $\widetilde{\beta}$ are unknown and the only available values are the estimate $\widehat{\alpha}$=$\widehat{\alpha}(\delta_n)$ and $\widehat{\beta}$=$\widehat{\beta}(\delta_n)$ given in [3] by :

$$\widehat{\alpha}=\frac{1}{N_0}\sum_{\substack{i=1,\\ \ell_i=0}}^{N} \mathbf{1}_{\{\delta_n(\boldsymbol{z}_q^i)=1\}} \text{ and } \widehat{\beta}=\frac{1}{N_1}\sum_{\substack{i=1,\\ \ell_i=1}}^{N} \mathbf{1}_{\{\delta_n(\boldsymbol{z}_q^i)=1\}}$$

where N_k is the number of samples $\boldsymbol{z}_q^i$ in $\mathcal{S}_q$ with label $\ell_i = k$ and $\mathbf{1}_{\{\cdot\}}$ is the indicator function. It is well known that $\widetilde{\alpha}$ and $\widetilde{\beta}$ can be well approximated by $\widehat{\alpha}(\delta_n)$ and $\widehat{\beta}(\delta_n)$ provided that N_0 and N_1 are sufficiently large [3]. In fact, given two constants $\eta_0, \eta_1 > 0$, there exist $\epsilon_0 = \epsilon_0(N_0, \eta_0)$ and $\epsilon_1 = \epsilon_1(N_1, \eta_1)$ such that :

$$\mathrm{Pr}\Big(|\widehat{\alpha} - \widetilde{\alpha}| > \epsilon_0(N_0, \eta_0)\Big) \leq \eta_0 \quad (16)$$

and

$$\mathrm{Pr}\Big(|\widehat{\beta} - \widetilde{\beta}| > \epsilon_1(N_1, \eta_1)\Big) \leq \eta_1. \quad (17)$$

The values η_0 and η_1 decay exponentially fast as function of increasing ϵ_0 and ϵ_1 (see details in [3]).

Using the "Probably Approximately Correct" (PAC) bounds (16)-(17) together with (13)-(14) immediately yields to the following result : given a constant $\eta > 0$, there exists $\epsilon = \epsilon(N_0, N_1, \eta)$ such that

$$\Pr\Big(|\alpha_s - \widehat{\alpha}_s| > \epsilon\Big) \leq \eta \text{ and } \Pr\Big(|\beta_s - \widehat{\beta}_s| > \epsilon\Big) \leq \eta. \quad (18)$$

Hence, the theoretical, but unknown, Receiver Operating Curve (ROC) curve of the semiparametric test, which corresponds to the plot of β_s against α_s (see details in [2]), can be bounded, with a high probability $1 - \eta$, by a lower and upper bound depending essentially on the estimates $\widehat{\alpha}_s$ and $\widehat{\beta}_s$ plus or minus a constant ϵ depending on the sizes N_0 and N_1 of the learning data set. If the learning data set is sufficiently large, the statistical performances of the semiparametric test can be well characterized.

4. EXPERIMENT RESULTS

To illustrate the relevance of the proposed approach, numerical simulations were conducted for the measurement model (1) where $n = 3$, $p = 1$, $q = 2$,

$$\boldsymbol{P} = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}, \quad \boldsymbol{Q} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \\ 0 & 0 \end{pmatrix},$$

$\sigma = 3$, $x_0 = 5$ and $\boldsymbol{y} = (y_1, y_2)^\top$ such that y_1, resp. y_2, follows the exponential distribution with expectation 2, resp. 1. For evaluation purpose, the performance of the parametric test $\delta_p(\boldsymbol{z})$ in (4), the nonparametric test $\delta_n(\boldsymbol{z})$ in (5) with Gaussian kernel and the proposed semiparametric test $\delta_s(\boldsymbol{z})$ in (10) are compared. The false alarm rate and the correct detection rate of each test are estimated from 10^5 random samples $\boldsymbol{z}$. The learning data base is composed of 100 random samples $\boldsymbol{z}^i$ with label ℓ_i=0 and 100 random samples with label ℓ_i=1. For the semiparametric test, γ is obtained from the values $\widetilde{\alpha}$ and $\widetilde{\beta}$ achieved by the nonparametric test $\delta_n(\boldsymbol{z}_q)$, applied to the restricted subvector $\boldsymbol{z}_q$, when its threshold is zero.

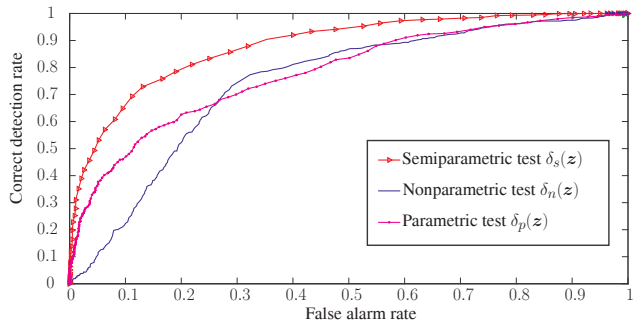

Fig. 1. The correct detection rate as a function of the false alarm rate for the three tests $\delta_s(\boldsymbol{z})$, $\delta_n(\boldsymbol{z})$ and $\delta_p(\boldsymbol{z})$.

The results of the simulation are plotted on Fig. 1. This figure shows that the nonparametric test $\delta_n(\boldsymbol{z})$ and the parametric test $\delta_p(\boldsymbol{z})$ have a comparable level of performances. They are clearly outperformed by the proposed semiparametric test $\delta_s(\boldsymbol{z})$. The theoretical correct detection probability β_s of the semiparametric test is not plotted on the figure because it coincides with the proposed empirical curve. Contrary to the nonparametric test, the false alarm rate of the semiparametric test can be easily controlled from (13).

5. CONCLUSION

This paper proposes a semiparametric test based on a likelihood ratio test dependent of a nonparametric decision value. Numerical simulations show that this test outperforms both a pure parametric test and a pure nonparametric test.

6. REFERENCES

[1] E. L. Lehmann and J. P. Romano, *Testing statistical hypotheses*, 3rd ed., ser. Springer Texts in Statistics. New York : Springer, 2005.

[2] H. V. Poor, *An introduction to Signal Detection and Estimation*. New York, NY : Springer-Verlag Inc., 1988.

[3] L. Devroye, L. Györfi, and G. Lugosi, *A Probabilistic Theory of Pattern Recognition*. Springer, New York, 1996.

[4] W. K. Härdle, M. Müller, S. Sperlich, and A. Werwatz, *Nonparametric and Semiparametric Models*. Springer, 2004.

[5] P. Singer, "The fusion of parametric and non-parametric hypothesis tests," in *Proc. of the Sixth International Conference of Information Fusion*, vol. 2, 2003, pp. 780–784.

[6] L. Bruzzone and D. Prieto, "An adaptive semiparametric and context-based approach to unsupervised change detection in multitemporal remote-sensing images," *IEEE Trans. Image Process.*, vol. 11, no. 4, pp. 452 –466, 2002.

[7] P. Stoica, P. Babu, and J. Li, "New method of sparse parameter estimation in separable models and its use for spectral analysis of irregularly sampled data," *IEEE Trans. Signal Process.*, vol. 59, no. 1, pp. 35–47, jan. 2011.

[8] L. Fillatre and I. Nikiforov, "Non-bayesian detection and detectability of anomalies from a few noisy tomographic projections," *IEEE Trans. Signal Process.*, vol. 55, no. 2, pp. 401–413, 2007.

[9] J. Shawe-Taylor and N. Cristianini, *Kernel Methods for Pattern Analysis*. New York, NY, USA : Cambridge University Press, 2004.