$\mathcal{E}$-Optimal Anomaly Detection in Parametric Tomography

Lionel FILLATRE, Igor NIKIFOROV and Florent RETRAINT
Université de Technologie de Troyes,
Institut Charles Delaunay, Troyes, France

Abstract. The paper concerns the radiographic non-destructive testing of well-manufactured objects. The detection of anomalies is addressed from the statistical point of view as a binary hypothesis testing problem with nonlinear nuisance parameters. A new detection scheme is proposed as an alternative to the classical GLR test. It is shown that this original decision rule detects anomalies with a loss of a negligible (ε) part of optimality with respect to an optimal invariant test designed for the "closest" hypothesis testing problem with linear nuisance parameters.

1. Introduction

For radiographic inspection of industrial objects (nuclear fuel rods, for example), it is desirable to detect defects, inclusions or any unexpected cavities in order to assure the safety and reliability of installations.

1.1 State-of-the-art

Existing methods for detecting internal defects in radiographies fall into three groups : 1) methods without *a priori* knowledge, 2) reference methods and 3) Computerized Tomography (CT) methods. The actual paper belongs to this last group.

The first group includes the methods of defect detection without prior knowledge of the imaged object – in other words, the methods without (statistical) model. Such approaches typically use image processing tools, pattern recognition, expert systems and artificial neural networks. The prerequisite for these methods is the existence of common properties which consistently define all kinds of anomalies and distinguish them from the character features of the non-anomalous imaged object. Often, these methods are noise-sensitive since they do not include explicitly a random noise in the measurement model.

In the second group, it is assumed that a reference radiography (model) is available. If a significant difference is identified by comparing the tested image with a reference one, then the inspected piece is classified as defective. This approach is efficient to deal with a well-known inspected piece but it is heavily based on a reference model that is not always possible in practice.

The third group (CT methods) consists of detecting defects from radiographies by reconstructing the imaged object, which are composed of an (partially) unknown non-anomalous background with possibly hidden defects. Since the number of projections and/or angles of view available for inspection is very limited and the pixel-by-pixel reconstruction is impossible, the introduction of prior information on the unknown background (i.e. on the nuisance parameter) is inevitable to fill up the gap in the missing data. Two methods of prior information introduction are available in the literature according to their nature – deterministic and statistical. A purely deterministic regularization of this ill-posed problem has some drawbacks including artifacts in the resulting image and noise sensitivity among others. The statistical CT approaches to anomaly detection can be divided into two groups : Bayesian and non-Bayesian. The dominant trend in the literature is the Bayesian statistical approach. In the case of anomaly detection problem, it is assumed that : i) the considered hypotheses, $H_0 =$ {the inspected object is defect-free} and $H_1 =$ {the inspected object contains defects}, are random events with known prior probabilities; ii) the non-anomalous background (or its structure) and the anomalies are random and the parameters of their (usually Gaussian) *a priori* distribution are known. In this paper which continues our previous publications [1,2], another non-Bayesian statistical philosophy is adopted : it is assumed that the non-anomalous background is a deterministic nuisance parameter.

1.2 Motivation of this study

Let us briefly discuss the relative advantages of the Bayesian and non-Bayesian approaches to anomaly detection with nuisance parameters in order to show the relevance of the proposed non-Bayesian approach to the problem of nuclear fuel rod inspection.

The advantages of Bayesian approach to the problem of statistical anomaly detection are the following : it allows the natural incorporation of prior knowledge; the Bayesian decision rules are simple and efficient (lower probabilities of false alarm and non-detection). A typical application of the Bayesian approach is the biomedical X-ray examinations. Here, the statistical models of non-anomalous object and anomalies are well-known. The usage of Bayesian approach seems to be justified especially understanding that the biomedical X-ray examinations deal with low-dose imaging: the statistical decision procedure has to preserve its efficiency with low SNR signals.

If the geometrical and/or physical properties of the inspected objects cannot be described by an *a priori* known probabilistic model then a more convenient working hypothesis about the non-anomalous object is the assumption that its geometrical and/or physical properties are deterministic unknown nuisance parameter composed of known basis functions. Often, this is the case in the non-destructuve testing of industrial equipment components, in welding defects detection for instance.

1.3 Contribution

Since the number of projections and/or view angles available for inspection is very limited, the defect detection problem is based on the assumption that the imaged medium is composed of an (partially) unknown background with a possibly hidden anomaly. It is considered as a parametric hypotheses testing problem between two composite alternatives with nonlinear nuisance parameters. A key assumption is the existence of a nonlinear parametric parsimonious model of the non-anomalous background to counterbalance the lack of observations.

The Generalized Likelihood Ratio (GLR) test [3,4], which is usually used to solve this kind of problem, has three major drawbacks: 1) this tool is relevant when the number of observations is very large but it is often suboptimal for a limited number of observations; 2) the GLR test requires to estimate the unknown parameters before taking a decision, which is difficult in a nonlinear case and 3) the GLR test makes no distinction between the nuisance parameters with respect to their impact on the nonlinearity of the model, which is not relevant from the practical point of view. A new detection scheme is proposed as an alternative to the GLR test: it detects anomalies with a loss of a negligible (ε) part of optimality with respect to an optimal invariant test designed for the "closest" hypotheses testing problem with linear nuisance parameters.

This paper is organized as follows. First, a parametric-based approach which includes the nonlinear parsimonious parametric model of the inspected object and radiographic process is presented in section 2. Secondly, in section 3, an ε-optimal test is designed to detect anomalies in the presence of nonlinear nuisance parameters. Finally, some experimental results with real radiographies show the relevance of the theoretical developments in section 4.

2. Problem statement: anomaly detection in parametric tomography

2.1 Nuclear fuel rod inspection

A nuclear fuel rod is composed of a body and a plug as shown in Fig. 1. The body is manufactured separately from the plug and, before its use, the plug is welded with the body. The goal of the nuclear fuel rod inspection is to detect defects (anomalies) in the welding zone which corresponds to a tangential part of the fuel rod (see Fig. 1). During the monitoring process, the nuclear fuel rod is imaged with a tomographic system composed of a X-source and a planar detector. The fuel rod is put into a compensator which is made of the same material to avoid the high contrast of radiography near the edges of the fuel rod. The goal is to decide between the two possible situations: $H_0 =$ {there is no anomaly} and $H_1 =$ {there is at least one anomaly}.

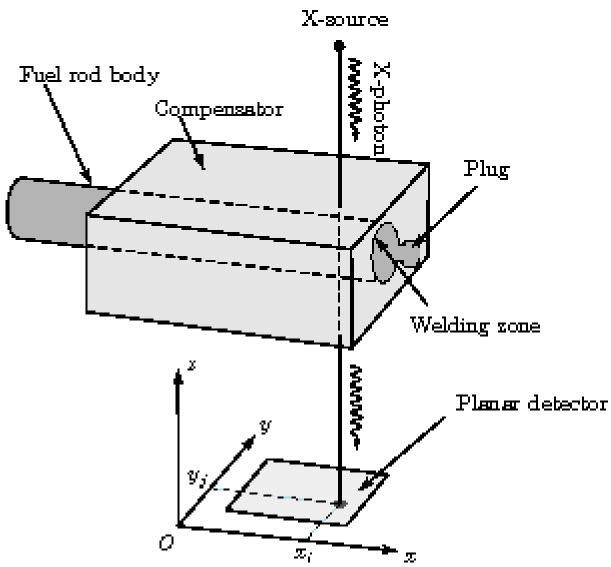

Fig. 1. Geometry of the nuclear fuel rod inspection system.

2.2 Physical background

To simplify the problem, the parallel-beam geometry is used in the paper and the X-rays are all oriented along the z-axis (see Fig. 1). The planar detector coincides with the xOy-plane. The measurements $\zeta(x, y)$ at different points (x, y) of the detector are modeled as independently distributed random variables [5] such that:

$$\zeta(x,y) \sim \Pi(m(x,y)) = \Pi(\mu(x,y)+\omega(x,y)), \quad (1)$$

where $\Pi(m)$ denotes the Poisson law with parameter $m>0$. The *unknown* quantity $\mu(x,y)$ (resp. $\omega(x,y)$) represents the mean number of photons passing through the object (resp. the mean number of extra photons, caused primarily by scattered radiations) at the position (x,y).

Let r be the radius of the fuel rod and $l(x,y;r)$ be the material (the fuel rod together with the compensator) thickness corresponding to the location (x,y) on the detector (see Fig. 1). It is assumed that an unknown value of r belongs to the interval $I=[r_0-\varrho;r_0+\varrho]$, where ϱ is a small positive constant and r_0 is exactly known. It is assumed that the quantity $\mu(x,y)$ can be well approximated [6] by the polynomial function:

$$\mu(x,y) \approx \widehat{\mu}(x,y;r,a_0,\mathbf{a}) = a_0 + \sum_{k=1}^{n_{\mathbf{a}}} a_k\, l^k(x,y;r), \quad (2)$$

where $\mathbf{a}=(a_1\, a_2 \ldots a_{n_{\mathbf{a}}})^T$ is the vector of coefficients, and the impact of scattered radiations can be approximated by a bivariate polynomial function:

$$\omega(x,y) \approx \widehat{\omega}(x,y;\mathbf{b}) = \sum_{u=0}^{n_u}\sum_{v=0}^{n_v} b_{u,v} x^u y^u, \quad (3)$$

where $\mathbf{b}=(b_{0,0}\, b_{1,0} \ldots b_{n_u,n_v})^T$. To avoid the redundancy with the term $b_{0,0}$ in (3), the term a_0 from equation (2) is omitted in the rest of the paper. It is assumed that the vector $\mathbf{a}$ belongs to a compact set $K_{\mathbf{a}} \subset R^{n_{\mathbf{a}}}$ and the vector $\mathbf{b}$ belongs to a compact set $K_{\mathbf{b}} \subset R^{n_{\mathbf{b}}}$ with $n_{\mathbf{b}}=(n_u+1)(n_v+1)$ to warrant the validity of the approximation given by equations (2) and (3).

2.3 Measurement model

By considering equations (2) and (3), equation (1) can be rewritten as:

$$\zeta(x,y) \sim \begin{cases} \Pi(m(x,y)) & \text{under } H_0 \\ \Pi(\theta(x,y)+m(x,y)) & \text{under } H_1 \end{cases} \quad (4)$$

where $\theta(x,y)$ represents the local (at (x,y)) variation of the mean number of X-photons arrived on the planar detector due to the anomaly at the position (x,y). For the considered problem, the exposure time and the X-flux intensity are high enough to warrant a good signal-to-noise ratio. Consequently, the Gaussian approximation of the Poisson distribution is relevant, which leads to a more tractable detection problem when anomalies are unspecified. Hence, the measurement model (4) is approximated by the following one:

$$\zeta(x,y) = \begin{cases} \widehat{m}(x,y;\mathbf{c})+\xi(x,y) & \text{under } H_0 \\ \theta(x,y)+\widehat{m}(x,y;\mathbf{c})+\xi(x,y) & \text{under } H_1 \end{cases},$$

with $\widehat{m}(x,y;\mathbf{c})=\widehat{\mu}(x,y;r,\mathbf{a})+\widehat{\omega}(x,y;\mathbf{b})$, $\mathbf{c}=(r,\mathbf{a},\mathbf{b})\in K$, $K=I\times K_{\mathbf{a}}\times K_{\mathbf{b}} \subset R^{n_{\mathbf{c}}+1}$, $n_{\mathbf{c}}=n_{\mathbf{a}}+n_{\mathbf{b}}$ and $\xi(x,y) \sim N(0,\sigma^2(x,y))$. The standard deviation $\sigma(x,y)$ is defined by $\sigma(x,y)=(\eta\overline{m}(x,y))^{\frac{1}{2}}$ where η is a known experimental coefficient independent of (x,y) and $\overline{m}(x,y)$ is an experimental mean value for $m(x,y)$.

The planar detector, which is composed of $n=n_x\, n_y$ discrete sensors, can be viewed as a $n_x \times n_y$ matrix. Let us note $\zeta_{i,j}$ the sensor measurement at the row i and the column j. By denoting $vec\left(\{\zeta_{i,j}\}\right)$ the lexicographical ordering of measurements $\zeta_{i,j}$, the above approximated measurement model can be rewritten:

$$\Xi = vec\left(\{\zeta_{i,j}\}\right) = \begin{cases} M(\mathbf{c}) + \xi & \text{under } H_0 \\ \theta + M(\mathbf{c}) + \xi & \text{under } H_1 \end{cases}, \tag{5}$$

where $\theta = vec\left(\{\theta_{i,j}\}\right)$, $M(\mathbf{c}) = vec\left(\{\widehat{m}_{i,j}(\mathbf{c})\}\right)$ and $\xi = vec\left(\{\xi_{i,j}\}\right)$. The random vector $\xi \sim N(0,\Sigma)$ follows the n-dimensional Gaussian law with a zero mean and a known diagonal positive definite covariance matrix Σ. A bit of algebra shows that:

$$M(\mathbf{c}) = F(r)\mathbf{a} + G\mathbf{b}, \tag{6}$$

where $F(r) = \left(F_1(r) \ldots F_{n_\mathbf{a}}(r)\right)$ is an $n \times n_\mathbf{a}$ matrix, $G = (G_1 \ldots G_{n_\mathbf{b}})$ is an $n \times n_\mathbf{b}$ matrix, with $F_s(r) = vec\left(\{l_{i,j}^s(r)\}\right)$ and $G_k = vec\left(\{p_i^u q_j^v\}\right)$ for k verifying $k = u\,(n_v + 1) + v + 1$.

3. Anomaly detection: hypotheses testing with nuisance parameters

The anomaly detection problem consists in solving the hypotheses testing problem with nonlinear nuisance parameters. Because a number of obstacles stand in the way, it is proposed to linearize and simplify the model of nuisance parameters by using additional nuisance parameters. To examine the impact of the proposed approximation, the almost-optimality of the decision algorithm is studied by comparing the proposed test with an optimal one obtained for the “closest” linear model of nuisance parameters. The main contribution is given by Proposition 1.

3.1 Motivation

Since the matrix Σ is known, the testing problem (5) consists of choosing between the two alternatives:

$$H_0 = \{\mathbf{y} \sim N(\theta + H(\mathbf{c}), I_n); \theta = 0, \mathbf{c} \in K\} \tag{7}$$

$$H_1 = \{\mathbf{y} \sim N(\theta + H(\mathbf{c}), I_n); \theta \neq 0, \mathbf{c} \in K\}, \tag{8}$$

with $\mathbf{y} = \Sigma^{-\frac{1}{2}}\Xi$, $H(\mathbf{c}) = \Sigma^{-\frac{1}{2}}M(\mathbf{c})$ and $\Sigma^{-\frac{1}{2}}$ is the square-root matrix of Σ^{-1} such that $\Sigma^{-\frac{1}{2}}\Sigma^{-\frac{1}{2}} = \Sigma^{-1}$.

The quality of a binary statistical test $\delta : R^n \mapsto \{H_0, H_1\}$ is defined with the probability of false alarm α and the power of the test β [7]. The subtlety of the above mentioned hypotheses testing problem consists in the existence of the unknown nonlinear nuisance parameter $\mathbf{c}$. Distinguishing two subsets of components of the parameter vector, the parameters of interest (anomaly) θ and the nuisance parameter (non anomalous object) $\mathbf{c}$, is necessary for the following reason : nuisances parameters are of no interest for inspection. The performance indexes (α, β) of statistical tests are functions of both the parameter of interest θ and the nuisance parameters $\mathbf{c}$. The desirable relation between the performance indexes (α, β) of a test and the parameter of interest θ usually results from the application and the statistical nature of the problem, in order to achieve optimal properties of the test. The main difference between θ and $\mathbf{c}$ is the following : in contrast to the parameter of interest, the nuisance parameter $\mathbf{c}$ has no desirable impact on the performance indexes. When designing a test for deciding between hypotheses in the presence of a nuisance parameter, the goal is to achieve the performance indexes independent of the actual value of $\mathbf{c}$.

3.2 Problem statement

To summarize the decision problem statement, let us define the class $K_\alpha = \{\delta : \sup_{\mathbf{c}\in K} \Pr_{\theta=0,\mathbf{c}}(\delta = H_1) \leq \alpha\}$ of tests with upper-bounded maximum false alarm probability, where the probability $\Pr_{\theta,\mathbf{c}}$ stands for the vector of observations $\mathbf{y}$ being generated by the distribution $N(\theta + H(\mathbf{c}), I_n)$ and α is the prescribed probability of false alarm. The power function β is defined with the probability of detection : $\beta(\theta;\mathbf{c}) = \Pr_{\theta\neq 0,\mathbf{c}}(\delta = H_1)$, while considering $\mathbf{c}$ as an unknown vector. Roughly speaking, $\beta(\theta;\mathbf{c})$ should be as large as possible for every $\theta \neq 0$ and $\mathbf{c} \in K$ for the prescribed probability of false alarm α. In the case of a vector parameter θ, the crucial issue is to find an optimal solution over a set of alternatives which is rich enough. Unfortunately, Uniformly Most Powerful (UMP) tests scarcely exist, except when the parameter θ is scalar, the family of distributions has a monotone likelihood ratio, and the test is one-sided [7]. Because of the presence of the nuisance parameter $\mathbf{c}$ and composite hypotheses H_0 and H_1, the design of an UMP test becomes impossible.

In the case of a linear model with nuisance parameters $\mathbf{y} = H\mathbf{c} + L\theta + \xi$, where H and L are known matrices of suitable dimensions, the optimal tests based on the theory of Wald [3] and on the theory of invariance have been discussed in [1,2,8,9]. Unfortunately, due to the fact that the vector-function $\mathbf{c} \mapsto H(\mathbf{c})$ is nonlinear and the nuisance parameter vector $\mathbf{c}$ belongs to a compact K, the direct application of the theory of invariant tests to the problem given by equations (7)-(8) is also compromised. At first glance, the only solution to overcome the above-mentioned difficulties is to apply the GLR test, which is applicable without the prerequisites :

$$\hat{\delta}(\mathbf{y}) = \begin{cases} H_0 & \text{if} \quad 2\log\hat{\Lambda}(\mathbf{y}) = 2\log\dfrac{\sup_{\theta\neq 0,\mathbf{c}\in K} f_{\theta+H(\mathbf{c})}(\mathbf{y})}{\sup_{\mathbf{c}\in K} f_{H(\mathbf{c})}(\mathbf{y})} < \hat{\gamma} \\ H_1 & \text{else} \end{cases}, \tag{9}$$

where $f_{\theta+H(\mathbf{c})}(\mathbf{y}) = \frac{1}{(2\pi)^{\frac{n}{2}}}\exp\left\{-\frac{1}{2}\|\mathbf{y} - \theta - H(\mathbf{c})\|_2^2\right\}$ is the pdf of the distribution $N(\theta + H(\mathbf{c}), I_n)$ and the threshold $\hat{\gamma}$ is chosen to satisfy the false alarm level α : $\sup_{\mathbf{c}\in K} \Pr_{\theta=0,\mathbf{c}}(2\log\hat{\Lambda}(\mathbf{y}) \geq \hat{\gamma}) = \alpha$. But this solution suffers from several disadvantages. First of all, the problem of non-asymptotic optimality of $\hat{\delta}(\mathbf{y})$ remains unsolved. Second, the numerical computation of the Maximum Likelihood (ML) estimate $(\hat{\mathbf{c}}, \hat{\theta}) = \arg\inf_{\theta,\mathbf{c}\in K} \|Y - \theta - H(\mathbf{c})\|_2^2$ meets serious difficulties : a poor and unstable convergence of the minimization process; a high sensitivity to initial conditions. Our experience with the problem of nuclear fuel rod inspection shows that the numerical application of the GLR test given by equation (9) is impossible. To overcome the above-mentioned disadvantages, the following approach is developed in the paper : a linearized measurement model is obtained and an invariant test (optimal in the case of a linear model) is designed by using this model. The proposed linearized test does not require to study the nuisance nonlinearity and it is also free from any numerical instability, which is very attractive from the practical point of view. The loss of optimality of the proposed linearized test is precisely established and negligible.

3.3 Optimal and suboptimal tests

The vector of unknown nuisance parameters $\mathbf{c}=(r,\mathbf{a}^T,\mathbf{b}^T)^T$ is partitioned into three blocks; the external radius $r\in I$ defines the geometrical properties of the nominal fuel rod body and the vectors $\mathbf{a}$ and $\mathbf{b}$ define its physical properties. The model is nonlinear due the function $r\mapsto F(r)$. It is straightforward to verify that the function $r\mapsto F(r)$ is infinitely differentiable on the interval $I=[r_0-\varrho;r_0+\varrho]$, where ϱ is a small constant. Therefore, by using the second-order Taylor expansion and equation (6), the vector function $\mathbf{c}\mapsto H(\mathbf{c})$ can be re-written in the following manner :

$$H(\mathbf{c})=\Sigma^{-\frac{1}{2}}\left[F(r_0)+(r-r_0)\dot{F}(r_0)+\frac{1}{2}(r-r_0)^2\ddot{F}(r_0)+o\left((r-r_0)^2 1\right)\right]\mathbf{a}+\Sigma^{-\frac{1}{2}}G\mathbf{b} \quad (10)$$

where $\dot{F}(r)=(\dot{F}_1(r)\ \dot{F}_2(r)\ \dots\ \dot{F}_{n_\mathbf{a}}(r))$ is the $n\times n_\mathbf{a}$ matrix of first-order derivatives of F, with $\dot{F}_s(r)=vec\left(\left\{\frac{dl_{i,j}^s(r)}{dr}\right\}\right)$, $\ddot{F}(r)=(\ddot{F}_1(r)\ \ddot{F}_2(r)\ \dots\ \ddot{F}_{n_\mathbf{a}}(r))$ is the $n\times n_\mathbf{a}$ matrix of second-order derivatives of F with $\ddot{F}_s(r)=vec\left(\left\{\frac{d^2l_{i,j}^s(r)}{dr^2}\right\}\right)$ and 1 is a $n\times n_\mathbf{a}$ matrix composed of ones. Let us now define the following linearized model of nuisance parameters :

$$H(\mathbf{c})\simeq\Sigma^{-\frac{1}{2}}\left[\dot{F}(r_0)\,|\,F(r_0)\,|\,G\right]\begin{pmatrix}(r-r_0)\mathbf{a}\\ \mathbf{a}\\ \mathbf{b}\end{pmatrix}=H_r\mathbf{x}. \quad (11)$$

The peculiarity of this linearized model is the fact that the radius r is replaced by the vector $(r-r_0)\mathbf{a}$ to avoid the simultaneous nonlinear estimation of r and $\mathbf{a}$ in the case of GLR (see equation (9)). Therefore, the components of the vector $\mathbf{x}\in R^{2n_\mathbf{a}+n_\mathbf{b}}$ are considered as independent variables, ignoring their intrinsic relations in the model of nuisance.

In the rest of the paper, the following idea is developed. The linearized model of nuisance given by equation (11) allows us to overcome the above-mentioned obstacles linked to numerical instability of GLR and to design an optimal invariant test based on a linear rejection of nuisance parameters. Nevertheless, two problems remain unsolved. It is necessary to estimate : *i)* the impact of the second-order term $\left[\frac{1}{2}(r-r_0)^2\ddot{F}(r_0)+o\left((r-r_0)^2 1\right)\right]\mathbf{a}$ on the power of the test; *ii)* the loss of optimality of this "linearized" test with respect to the linear nuisance case when the true value of radius is exactly known $r=r_0$.

Let us denote the column space of the matrix H_r by $R(H_r)$ and its orthogonal complement by $R(H_r)^\perp$. In the rest of the paper, it is assumed that $\theta\in\Theta=R(H_r)^\perp$ to warrant the detectability of the anomaly θ. A more detailed discussion of the detectability problem can be found in [2]. Hence, in the case of linearized model, the hypotheses testing problem consists in deciding between

$$\overline{H}_0=\{Y\sim N(\theta+H_r\mathbf{x},\sigma^2 I_n);\theta=0,\mathbf{x}\in R^{2n_\mathbf{a}+n_\mathbf{b}}\} \quad (12)$$

and

$$\overline{H}_1=\{Y\sim N(\theta+H_r\mathbf{x},\sigma^2 I_n);\theta\in\Theta,\mathbf{x}\in R^{2n_\mathbf{a}+n_\mathbf{b}}\}, \quad (13)$$

Let us note $P_{H_r}^\perp=I_n-H_r(H_r^T H_r)^{-1}H_r^T$ the orthogonal projection on the null space of the matrix H_r. The "linearized" test δ_r taking in account the unknown radius r is defined by :

$$\delta_r(\mathbf{y}) = \begin{cases} \overline{H}_0 \;\; if \;\; \Lambda_r(\mathbf{y}) = \left\| P_{H_r}^{\perp}\mathbf{y} \right\|_2^2 < \gamma_r \\ \overline{H}_1 \;\; else \end{cases}, \tag{14}$$

where the threshold $\gamma_r = \gamma_r(\alpha; n - 2n_{\mathbf{a}} - n_{\mathbf{b}})$ is chosen to satisfy the false alarm level α : $\Pr_{\theta=0}(\Lambda_r(\mathbf{y}) \geq \overline{\gamma}_r) = \alpha$. The hypotheses testing problem given by equations (12) - (13) is used to design the optimal invariant test $\delta_r(\mathbf{y})$ and to choose the threshold γ_r. But to estimate the impact of the second-order term of Taylor expansion (see equation (10)) and the loss of optimality, the original "nonlinear" hypotheses testing problem given by equations (7) - (8) is considered in the rest of the paper. For this reason, the power function $\beta_{\delta_r}(\theta;\mathbf{c})$ of the "linearized" test $\delta_r(\mathbf{y})$ depends on the nuisance parameters $\mathbf{c}$.

3.4 Bound on the loss of optimality

Let us now suppose that the external radius $r = r_0$ is known and only the parameters of the beam hardening and X-scattering models $\mathbf{a}, \mathbf{b}$ are unknown. In such a case the nuisance is described by a linear model :

$$H(r_0, \mathbf{a}, \mathbf{b}) = \Sigma^{-\frac{1}{2}}\left[F(r_0) \mid G\right]\begin{pmatrix}\mathbf{a}\\ \mathbf{b}\end{pmatrix} = H_0\begin{pmatrix}\mathbf{a}\\ \mathbf{b}\end{pmatrix}$$

The associated hypotheses testing problem consists in deciding between

$$\widetilde{H}_0 = \{Y \sim N(\theta + H_0(\mathbf{a}^T, \mathbf{b}^T)^T, \sigma^2 I_n); \theta = 0, (\mathbf{a}^T, \mathbf{b}^T)^T \in R^{n_{\mathbf{a}}+n_{\mathbf{b}}}\} \tag{15}$$

and

$$\widetilde{H}_1 = \{Y \sim N(\theta + H_0(\mathbf{a}^T, \mathbf{b}^T)^T, \sigma^2 I_n); \theta \in \Theta, (\mathbf{a}^T, \mathbf{b}^T)^T \in R^{n_{\mathbf{a}}+n_{\mathbf{b}}}\}. \tag{16}$$

The optimal invariant test δ_0 associated to this linear problem (see details in [2]) is obtained directly from (14) by replacing H_r with H_0:

$$\delta_0(\mathbf{y}) = \begin{cases} \widetilde{H}_0 \;\; if \;\; \Lambda_0(\mathbf{y}) = \left\| P_{H_0}^{\perp}\mathbf{y} \right\|_2^2 < \gamma_0 \\ \widetilde{H}_1 \;\; else \end{cases}, \tag{17}$$

where the threshold $\gamma_0 = \gamma_0(\alpha; n - n_{\mathbf{a}} - n_{\mathbf{b}})$ is chosen to satisfy the false alarm level α : $\Pr_{\theta=0}(\Lambda_0(\mathbf{y}) \geq \gamma_0) = \alpha$. Let us denote the power of this optimal invariant test by $\beta_{\delta_0}(\theta)$. This power function will be used as a reference in the rest of the paper.

Definition 1. Let us define a class of all tests K_α with the false alarm level α. A test $\delta \in K_\alpha$ is called ε-optimal over the set Θ with respect to an optimal invariant test $\delta_0 \in K_\alpha$ if there exists a (small) positive constant $\varepsilon > 0$ such that

$$\sup_{\theta \in \Theta, \mathbf{c} \in K} \left| \beta_\delta(\theta;\mathbf{c}) - \beta_{\delta_0}(\theta) \right| \leq \varepsilon.$$

Proposition 1. The test δ_r is ε-optimal over the set Θ with respect to the optimal invariant test $\delta_0 \in K_\alpha$.

Let us discuss briefly the relation between the rejection of nuisance parameters and their estimation. The principal obstacle to implementing the GLR test is the nonlinear estimation of unknown informative and nuisance parameters θ and $\mathbf{c}$ in equation (9). In contrast to such a nonlinear optimization, the rejection of the linearized nuisance is much simpler : the intrinsic dependence between estimated parameters, which leads to a poor and unstable convergence in the case of nonlinear estimation, is not an obstacle for the nuisance rejection problem which performs even better in the case where the matrix H_r is not full rank. Hence, due to a conveniently chosen linearized nuisance parameter model, the disadvantages of GLR test are transformed in the advantages of invariant test based on the linear rejection. The price of this transformation is an augmentation of the dimension of nuisance parameter vector that leads to a certain loss, ε, of optimality of the "linearized" test δ_r.

4. Experimental results with real radiographies

Because of the limited volume of the paper, experimental results are not described in details. The planar detector is composed of $n_x = 50 \times n_y = 100$ sensors, i.e. $n = 5000$, with a resolution of 0.003 cm, $\eta = 0.1749$, $\alpha = 10^{-2}$, $\varrho = 0.005$ cm, $n_{\mathbf{a}} = 2$ and $n_x = n_y = 3$. It is assumed that $\mathbf{a} = (a_1\ a_2)^T$ verifies $-2000 \leq a_1 \leq 2000$ and $-20000 \leq a_2 \leq 20000$. Radiographies have the estimated signal-to-noise ratio $\mathrm{SNR}_{\mathrm{dB}} = 10 \log(\mathrm{SNR}) \approx 70.1$ dB with $\mathrm{SNR} = H(\mathbf{c})^T \Sigma^{-1} H(\mathbf{c})$.

When the inspected object is anomaly-free (see radiography Ξ_1 in Fig. 2(a)), the unknown background is properly rejected and the residuals are close to a stationary "white noise" (see Fig. 2(b)). Fig. 2(c) presents a radiography Ξ_2 with an anomaly. This leads to the residuals with an "anomaly signature" (white and black spots) as shown in Fig. 2(d). Under H_0, the decision function is $\overline{\Lambda}(\mathbf{y}_1) = 4986.02 < \gamma_{0.01}$ with $\gamma_{0.01} = 5222.27$. Under H_1, its value is $\overline{\Lambda}(\mathbf{y}_2) = 5883.24 > \gamma_{0.01}$. Since anomalies are assumed to belong to the detectable space Θ_m and the nuisance parameter space K is bounded, the upper bound $\varepsilon \approx 10^{-3}$ is estimated by sampling K to find the largest difference between the power functions β^* and $\beta_{\overline{\delta}}$. The loss of optimality is almost negligible and the false alarm probability holds an acceptable level.

5. Conclusion

A parsimonious nonlinear parametric model is proposed to describe the radiographic non-destructive inspections. This approach can be called "parametric tomography". It allows us to avoid a complete physical study of the radiographic process. Moreover, to overcome numerical instabilities of the classical GLR test, a new ε-optimal statistical test is proposed and studied to solve the problem of anomaly detection. The experimental results on real radiographic data confirm the relevance and efficiency of the proposed solution.

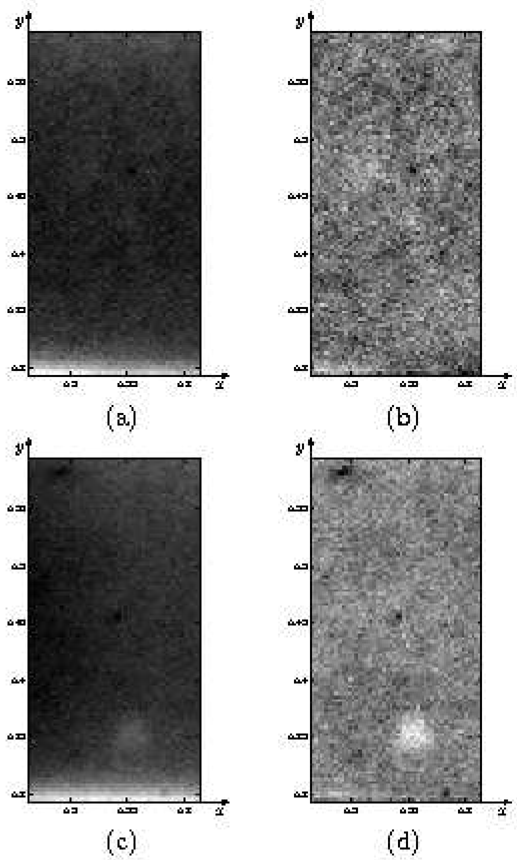

Fig. 2. (a) Radiography Ξ_1 of a safe fuel rod, (b) residuals $\bar{\mathbf{y}}_1 = P_{H_0}^{\perp}\mathbf{y}_1$ of the radiography (a), (c) radiography Ξ_2 of a fuel rod with an anomaly, (d) residuals $\bar{\mathbf{y}}_2 = P_{H_0}^{\perp}\mathbf{y}_2$ of the radiography (c).

References

[1] Fillatre and Nikiforov, "A statistical detection of an anomaly from a few noisy tomographic projections," *Journal of Applied Signal Processing*, no. 14, pp. 2215–2228, 2005.
[2] L.Fillatre, F.Retraint, and I.Nikiforov, ``Detectability of anomalies from a few noisy tomographic projections," in *International Federation Symposium on Automatic Control World Congress (IFAC2005),* Prague, July 2005.
[3] A. Wald, "Tests of statistical hypotheses concerning several parameters when the number of observations is large," *Trans. Amer. Math. Soc.*, vol. 54, pp. 426–482, 1943.
[4] Seber and Wild, *Nonlinear regression*.Wiley, 1989.
[5] Elbakri and Fessler, "Statistical image for reconstruction for polyenergetic x-ray computed tomography," *IEEE Trans. Med. Imag.*, vol. 21, no. 2, pp. 89–99, 2002.
[6] Herman, "Correction for beam hardening in computed tomography," *Phys. med. biol.*, vol. 24, no. 1, pp. 81–106, 1979.
[7] Lehman, *Testing Statistical Hypotheses, Second Edition*.Chapman & Hall, 1986.
[8] Scharf and Friedlander, "Matched subspace detectors," *IEEE Trans. Signal Processing*, 1994.
[9] M. Fouladirad and I. Nikiforov, "Optimal statistical fault detection with nuisance parameters," *Automatica*, vol. 41, no. 7, pp. 1157 – 1171, jul 2005.