Towards the handling of uncertainties in statistical FDI

Michèle Basseville∗ Albert Benveniste ∗∗ Qinghua Zhang ∗∗

∗ IRISA/CNRS, Rennes, F. (e-mail: michele.basseville@irisa.fr).
∗∗ INRIA Rennes (e-mail: albert.benveniste,qinghua.zhang@inria.fr)

Abstract: The paper provides fundamental steps towards addressing the theoretically and practically important issue of handling uncertainties in statistical model-based approaches to fault detection and isolation (FDI). Two main types of uncertainties are distinguished: uncertainties in parameters on the one hand; uncertainties due to under-modeling as well as those resulting from changes in functioning modes, on the other hand. It is suggested that two different theories should be used.

Keywords: Fault detection and isolation, statistical inference, uncertainty, reduced order models.

1. INTRODUCTION

The emergence of stronger safety and environmental norms, the need for early decision mechanisms, together with the widespread diffusion of sensors of all kinds, result in an increased need for research investigations in the field of fault detection and isolation (FDI). In particular, efficient and robust methods for fault and damage detection, diagnostics and localization, are necessary for condition-based maintenance and for fatigue and aging prevention. Many approaches to FDI problems assume that a model of the monitored system is available. This is a reasonable assumption since many machines and processes rely on physical principles and models can be built on that basis (Ljung (1999)). However, it is often argued that a major drawback of model-based FDI methods lies in the requirement of a precise model of the system to be monitored. In particular, changes in the functioning modes and under-modeling of the system to be monitored, are often reported as being crucial issues for model-based FDI algorithms; see Frank (1990), Patton et al. (2000). Consequently, increasing the robustness of FDI methods with respect to unmodeled disturbances and model parameter uncertainties has been a very active topic during the last twenty years.

Three main types of investigations can be found in the literature. The first type involves system theoretic approaches to the design of unknown input observers and disturbance decoupling rejection methods (Chen and Patton (1999), Chen and Speyer (2002), Douglas and Speyer (1996), Frank (1994b), Nikoukhah (1994), Yoshimura et al. (1997)). In (Besancon (2003)), a high gain observer smooths down the uncertainty on the state equation; this results in robust residuals provided that there is no uncertainty on the output equation. In (Henry and Zolghadri (2005)) and (Zhong et al. (2003)), an $H_{\infty}$ approach is used for optimizing the filters generating residuals maximally sensitive to faults and minimally sensitive to perturbations; see also Mangoubi (1998) for robust game theoretic $H_{\infty}$ filtering theory. Other investigations of and discussions on robustness issues for FDI can be found in (Bokor (2009), Casavola et al. (2008), De Persis and Isidori (2002), Frisk and Nielsen (2006), Johansson et al. (2006), Li and Zhou (2009), Ma et al. (2007), Petersen and McFarlane (2004), Stoustrup and Zhou (2008), Tan et al. (2008)).

The second type is based on modal interval analysis and consistency techniques (Puig et al. (2008)). The fault detection problem is stated as a constraint satisfaction problem with a high number of variables and constraints. Consistency techniques based on interval arithmetic achieve consistency checks of the analytical redundancy relations dealing with uncertain measurements and parameters (Gelso et al. (2007)).

The third type of investigations is based on a statistical approach both to detect the faults and to handle the disturbances and uncertainties. Using stochastic modeling and statistical decision tools is indeed a natural approach for that; see Basseville and Nikiforov (1993), Gustafsson (2000). In particular, the statistical local approach has shown useful for designing on-board FDI algorithms in various application domains. Representing the monitored system by a probability distribution $p_0$ parametrized by the vector $\theta$ characterizing the system, this approach consists in detecting small deviations in $\theta$ w.r.t. a reference parameter $\theta_0$ and can be applied to a wide class of processes (Basseville (1998), Zhang et al. (1994)). In practice, when the uncertainty on $\theta_0$, if not on the distribution $p_0$ itself, is large w.r.t. the magnitude of the changes in $\theta$ to be detected, the efficiency of the local approach may be affected. It is then necessary to design a mechanism to translate the FDI algorithm more robust to those uncertainties. In (Basseville et al. (1994), Basseville (1997)), such a mechanism has been proposed. The main idea consists in designing residuals and decision functions together by using confidence ellipsoids for inserting the uncertainty in the shape of the (no-fault and fault) hypotheses between which one is willing to test. The resulting algorithm mainly handles uncertainties on the reference parameter $\theta_0$.

In this paper, we propose to use the recent theory of uncertain distributions (Peng (2009), Peng (2010)) to design
another mechanism, still in the framework of the statistical local approach, that we believe more appropriate for handling uncertainties resulting from under-modeling, as it is the case when the model used for on-board FDI is of (much) smaller dimension than the full model.

The paper is organized as follows. In section 2, we recall the key components of the statistical local approach to FDI and of the first solution for handling the uncertainties in the reference parameter. Section 3 summarizes the concepts and results of the theory of uncertain distributions that are useful for our purpose. Section 4 discusses how to combine both approaches to design FDI algorithms robust both to parametric and structural uncertainties. Section 5 concludes with future research investigations.

2. THE STATISTICAL LOCAL APPROACH TO FDI

The statistical local approach to FDI can be applied to a large class of parameterized nonlinear dynamical processes. This approach aims at detecting small deviations with respect to a reference parameter $\theta_0$, and basically transforms such a quite general detection problem into the well studied problem of detecting a change in the mean of a Gaussian vector; see Basseville (1998), Zhang et al. (1994). The approach relies on the design of a residual tightly connected to a parameter estimating function, and on a central limit theorem under both no-fault and fault hypotheses. In this section, we recall the key components of that theory, design and Taylor expansion of the residual and central limit theorem. Then we and describe the first solution for handling uncertainties in $\theta_0$.

2.1 Residual

For designing a residual in the general case of semi-Markov processes, an auxiliary process $(Z_k)_k$ is needed, that is a finite-dimensional vector-valued output of some causal dynamical system parameterized by vector $\theta$ and having as inputs the known or measured inputs and the outputs of the monitored system. It is assumed that a reference parameter $\theta_0$ is available, either known or identified on data recorded on the system under safe condition.

Let $K(\theta,Z_i)$ be an asymptotically stationary estimating function, differentiable w.r.t. $\theta$, and satisfying local identifiability hypotheses, namely there exists a neighborhood $\omega(\theta_0)$ such that

$$\mathbb{E}_{\theta} K(\theta_0,Z_i) = 0 \text{ if } \theta = \theta_0$$

$$\mathbb{E}_{\theta} K(\theta_0,Z_i) \neq 0 \text{ if } \theta \in \omega(\theta_0) \setminus \theta_0,$$

where $\mathbb{E}_{\theta}$ is the expectation when the parameter is $\theta$.

Assume that a $N$-size sample of data is available. The residual $\zeta_N(\theta_0)$ is defined as:

$$\zeta_N(\theta_0) \triangleq \sum_{i=1}^{N} X_i / \sqrt{N}$$

where:

$$X_i \triangleq K(\theta_0, Z_i)$$

Define also:

$$Y_i \triangleq \partial/\partial \theta K(\theta, Z_i)|_{\theta=\theta_0}$$

Introduce the mean deviation (Jacobian matrix):

$$M(\theta_0) \triangleq - \mathbb{E}_{\theta_0} (Y_i) = - \partial / \partial \theta \mathbb{E}_{\theta_0} K(\theta, Z_i)|_{\theta=\theta_0}$$

and the covariance matrix: $\Sigma(\theta_0) \triangleq \lim_{N \to \infty} \Sigma_N(\theta_0)$.

$$\Sigma_N(\theta_0) \triangleq \mathbb{E}_{\theta_0} (\zeta_N(\theta_0) \zeta_N^T(\theta_0))$$

2.2 Taylor expansion

For detecting small deviations w.r.t. $\theta_0$, the considered null and alternative hypotheses are closed to each other in the following sense (local hypotheses):

$$H_0 : \theta = \theta_0,$$

$$H_1 : \theta = \theta_0 + \Upsilon / \sqrt{N}$$

where $\Upsilon$ is unknown but fixed. Since small deviations are of interest, it is natural to consider the Taylor expansion of the residual (3) around $\theta_0$. As displayed in Table 1, the first term of that expansion converges to a zero-mean Gaussian vector under the null hypothesis $H_0$, thanks to the central limit theorem (CLT). Under the same hypothesis, and thanks to the law of large numbers (LLN), the second term converges to a constant vector $-M(\theta_0) \Upsilon$.

2.3 CLT under local hypotheses

The elements displayed in Table 1 are the key roots for the following asymptotic Gaussianity result under both the null (no-fault) and local (small fault) alternative hypotheses (Hall and Mathiason (1990), Basseville (1998)):

$$\zeta_N(\theta_0) \overset{N \to \infty}{\sim} \begin{cases} N(0, \Sigma(\theta_0)) \text{ under } H_0 \\ N(0, \Sigma(M(\theta_0) \Upsilon, \Sigma(\theta_0))) \text{ under } H_1 \end{cases}$$

Then the hypothesis testing problem in (7), where $\Upsilon$ is unknown, is solved with the aid of the generalized log-likelihood ratio (GLR) for an unknown change in the mean of a Gaussian vector. The resulting $\chi^2$-test writes:

$$\chi^2 \overset{N \to \infty}{\sim} \Sigma(\theta_0)^{-1} M(\theta_0) \Sigma(\theta_0)^{-1} M^T \Sigma(\theta_0)^{-1}$$

where the dependence in $\theta_0$ is omitted for simplification. The estimation of $M$ in (6) from a data sample is obtained by replacing the expectation by a sample average. The estimation of $\Sigma(\theta)$ is more tricky. The reader is referred to (Zhang and Basseville (2003)) for an efficient implementation of that test.

This theory can also be used for addressing the fault isolation problem; see Basseville (1998).

2.4 Different types of uncertainties

Referring to Table 1, uncertainties on the monitored system may affect the three key components in the residual $\zeta_N$, namely $\theta_0$ and $X_i$. $Y_i$. It should be clear that $X_i$ and $Y_i$ may be affected by uncertainties fully different from the uncertainties on $\theta_0$. This is explained next.

Uncertainties in the reference parameter. When using the local approach for change detection, the reference parameter $\theta_0$ is generally estimated from data recorded on the system assumed to be under safe condition. Then the uncertainty on $\theta_0$ is the estimation error. Since $\theta_0$ is inferred from data, the uncertainty about $\theta_0$ relates to the length $N$ of the data sample. This is adequately captured...
by formulating the uncertainty as being local, i.e. of the order $1/\sqrt{N}$, which fits well with the present framework of the local approach.

The main idea consists in designing residuals and decision functions together by using confidence ellipsoids for inserting the uncertainty in the shape of the (no-fault and fault) hypotheses between which one is willing to test, instead of using a confidence interval around the decision function itself often reported (Frank (1994a), Isermann (1994)), for example with adaptive threshold selection (Emami-Naeini et al. (1988), Frank (1994b)). Actually, this idea can be viewed as a multi-parameter extension of the idea of a scalar minimum magnitude of fault to be detected, which in turn has been proven to be an efficient alternative to the more intuitive idea of using multiple thresholds.

It is assumed that the uncertainty on the reference parameter $\theta_0$ is local, namely with order of magnitude $1/\sqrt{N}$. This uncertainty is handled in the formulation of the two no-fault and fault hypotheses to be tested against each other, and represented by two ellipsoids:

$$
\Theta_0 : \Upsilon^T M^T \Sigma^{-1} M \Upsilon \leq \rho_0^2 \\
\Theta_1 : \Upsilon^T M^T \Sigma^{-1} M \Upsilon \geq \rho_1^2
$$

(10)

where $\Upsilon = \sqrt{N}(\theta - \theta_0)$ and metric $M^T \Sigma^{-1} M$ specifies the variance of the error on $\theta_0$. The GLR test for those hypotheses writes (Basseville (1998), Basseville et al. (1994)):

$$
\frac{\sum^2}{N} \triangleq \begin{cases} 
-(\chi - \rho_1)^2 & \text{if } \chi \leq \rho_0 \\
-(\chi - \rho_0)^2 + (\chi - \rho_0)^2 & \text{if } \rho_0 \leq \chi \leq \rho_1 \\
(\chi - \rho_0)^2 & \text{if } \chi \geq \rho_1 
\end{cases}
$$

(11)

with $\chi \triangleq \chi_N/\sqrt{N}$, $\chi_N$ defined in (9). In practice, the “magnitudes” $\rho_0$ and $\rho_1$ are tuned from the available specifications about the changes to be monitored. Since in practice the sample size $N$ is finite and thus $(\theta - \theta_0)$ is finite too, those tuned values depend on $N$.

This simple idea has turned out to be a useful tool for obtaining robustness against small changes in the functioning modes and against under-modeling of the system to be monitored; see Basseville et al. (1994), Cussenot et al. (1996). It also helps for robustness against the nominal parameter uncertainty.

Uncertainties due to under-modeling. In contrast, uncertainties not related to the data set cannot depend on its length and thus cannot be modeled as local. Uncertainties affecting the $X_i$’s or the $Y_i$’s without affecting $\theta_0$ may occur in case of under-modeling. This is the case of the vibration monitoring applications discussed in (Benveniste et al. (2006)) for example. In those cases, the model used for FDI is of much smaller dimension than the finite element model (FEM) used for the design of the structure. When it comes to localize and diagnose damages in terms of the latter model, the idea is to use specific Jacobian matrices to be plugged as product terms in the mean value of the Gaussian vector in (8). The uncertainty then comes from the use of the underlying FEM model (and not of data as above), and furthermore from the model reduction operated within the computation of the Jacobian matrices.

The aim of the paper is to suggest another route for handling those uncertainties, based on the recent theory of uncertain distributions.

Finally, the uncertainties resulting from unmodeled nonlinearities are of the same kind, and may be handled with the approach proposed in the present paper.

3. THE THEORY OF UNCERTAIN DISTRIBUTIONS

The main components of Peng’s theory of uncertain distributions are now presented (Peng (2009), Peng (2010)). To begin with, the motivations for and key ideas of that theory are briefly sketched.

3.1 Introduction

The main motivation for the theory of uncertain distributions is to provide a LLN and CLT, and thus an explanation for the wide use of the Gaussian distribution, in the numerous practical situations where the i.i.d. assumption is not satisfied. The first idea for handling model uncertainty in a statistical setting is to consider that the true system parameter $\theta^*$ lies within a set of possible values $\Theta^*$. The beautiful idea of S. Peng, already suggested by P. Huber, consists in adopting drastically different angles. The two essential novelties are the following. The first one consists in considering primarily the linear map $X \mapsto \mathbb{E}(X)$ where $X$ ranges over the set of random variables, instead of focusing on $(\Omega, \mathcal{F}, \mathbb{P})$. The second novelty concerns the normal distribution $N(\mu, \Sigma)$, which is not defined through its density as usual, but is characterized by two classical properties (Kagan et al. (1973)) instead:

(1) It is a stable distribution with finite covariance: if $X, X \sim N(\mu, \Sigma)$ and $X$ is independent of $X$, then:

$$aX + bX \overset{d}{=} \sqrt{a^2 + b^2} (X - \mu) + (a + b)\mu$$

where $\overset{d}{=}$ means equality in distribution.
(2) For every suitable test function $\varphi$, the expectation $E(\varphi(X))$ can be computed using the fact that $u(t,x,y) = \frac{\|u\|_{\infty}}{\partial_t u} = \varphi$ and $D_y = (\partial_y, \partial_x)$, where $D_y = (\partial_y, \partial_x)_{i,j}$ and $G(p, A) \triangleq p^T \mu + 1/2 Tr(A)\Sigma$.

The above formulations generalize to the case of uncertainty of the form $(\Omega, F, P) \in \Theta$ in the following manner. First, $X \to \hat{E}(X) \triangleq \sup_{\theta \in \Theta} E_{\theta}(X)$ is convex – in fact it is monotonic and sublinear, which is somehow stronger in that constants are preserved. Thanks to Hahn-Banach theorem, the converse holds true: if $X \to \hat{E}(X)$ is monotonic and sublinear, then it is of the form $\hat{E}(X) = \sup_{\theta \in \Theta} E_{\theta}(X)$ for some $\Theta$. The operator $X \to \hat{E}(X)$ is thus called a sublinear expectation. In this framework, the equality in distribution of two vectors $X$ and $Y$, namely $X \overset{d}{=} X$, means $\hat{E}(f(X)) = \hat{E}(f(Y))$ for any function $f$.

Second, the so-called $G$-normal distribution is characterized by the fact that it is stable with finite covariance. Also, for every suitable test function $\varphi$, the expectation $\hat{E}(\varphi(X))$ can be computed using the fact of a $G$-heat equation, where $G(p, A) \triangleq \sup_{\theta \in \Theta} E_{\theta}(X)$, for some uncertainty set $\Gamma$ for the mean and covariance.

The main concepts are now introduced.

### 3.2 Sublinear expectation and normal distributions

**Definition.** A sublinear expectation is a functional $\hat{E} : \mathcal{H} \to \mathbb{R}$ satisfying the properties:

- (a) Monotonicity: if $X \geq Y$, then $\hat{E}(X) \geq \hat{E}(Y)$.
- (b) Constant preservation: $\hat{E}(c) = c$.
- (c) Sub-additivity: $\hat{E}(X + Y) \leq \hat{E}(X) + \hat{E}(Y)$.
- (d) Positive homogeneity: $\hat{E}(\lambda X) = \lambda \hat{E}(X), \forall \lambda \geq 0$.

It was introduced under the name of upper expectation by Huber (Huber (1981)[Chap.10]); see also Walley (1991).

**Robust representation.** If $\hat{E}$ is a sublinear expectation over $(\Omega, \mathcal{H})$, thanks to Hahn-Banach theorem there exists a family of linear expectations $\{E_{\theta} : \theta \in \Theta\}$ on $\Omega, \mathcal{H}$ s.t.

$$\hat{E}(X) = \sup_{\theta \in \Theta} E_{\theta}(X), \forall X \in \mathcal{H}. \quad (12)$$

In other words, instead of a unique distribution, one considers a family indexed by a convex set $\Theta$. The size of $\Theta$ represents the degree of uncertainty on the model.

**Identically distributed vectors.** Let $X_1$ and $X_2$ be two vectors in $\mathbb{R}^d$ defined on the sublinear expectation spaces $(\Omega_1, \mathcal{H}_1, \hat{E}_1)$ and $(\Omega_2, \mathcal{H}_2, \hat{E}_2)$, respectively. They are said to be identically distributed, written as $X_1 \overset{d}{=} X_2$, if $\hat{E}_1(\varphi(X_1)) = \hat{E}_2(\varphi(X_2)), \forall \varphi \in \mathcal{C}_{L1}(\mathbb{R}^d)$.

In that case, the sets of distributions associated with their respective uncertainties are identical: with $\hat{E}_X(\varphi) \triangleq \hat{E}(\varphi(X))$, $\{\hat{E}_{X_1}(\theta_1, \cdot) : \theta_1 \in \Theta_1\} = \{\hat{E}_{X_2}(\theta_2, \cdot) : \theta_2 \in \Theta_2\}$.

**Four typical parameters.** Define:

$$\mu \triangleq \hat{E}(X), \quad \mu \triangleq -\hat{E}(-X) \quad (14)$$

$$\sigma^2 \triangleq \hat{E}(X^2), \quad \sigma^2 \triangleq -\hat{E}(-X^2) \quad (15)$$

The subsets $[\mu, \mu]$ and $[\sigma^2, \sigma^2]$ characterize the mean-uncertainty and the variance-uncertainty of the distribution of $X \in \mathcal{H}$.

**Independent vectors.** On a sublinear expectation space $(\Omega, \mathcal{H}, \hat{E})$, a vector $Y = (Y_1, \ldots, Y_d)$, $Y_i \in \mathcal{H}$, is independent of another vector $X = (X_1, \ldots, X_m)$, $X_i \in \mathcal{H}$ under $\hat{E}$ if, for all $\varphi \in \mathcal{C}_{L1}(\mathbb{R}^m \times \mathbb{R}^d)$:

$$\hat{E}(\varphi(X, Y)) = \hat{E}(\hat{E}(\varphi(X, Y))|_X = X) \quad (16)$$

The condition (16) means that the uncertainty on $Y$ is not modified by the realization of $X$, in other words that the conditional sublinear expectation of $Y$ knowing $X$ is $\hat{E}(\varphi(X, Y))|_{X = X}$. In the case of a linear expectation, the classical independence is recovered. Note that $Y$ being independent of $X$ does not automatically implies that $X$ is independent of $Y$.

**Normal distributions.** Recall that, in the classical case:

$$X \overset{d}{=} N(0, \Sigma) \iff aX + bX \overset{d}{=} \sqrt{a^2 + b^2}X, \forall a, b \geq 0 \quad (17)$$

where $X$ is an independent copy of $X$.

**G-normal distribution with zero mean.** A vector $X = (X_1, \ldots, X_d)$ in a sublinear expectation space $(\Omega, \mathcal{H}, \hat{E})$ is said G-normal distributed if:

$$aX + bX \overset{d}{=} \sqrt{a^2 + b^2}X, \forall a, b \geq 0 \quad (17)$$

It is easy to check that, if $X$ satisfies (17), $\hat{E}(X_i) = 0$ and similarly $\hat{E}(-X_i) = 0, \forall i = 1, \ldots, d$. The distribution of $X$ is characterized by the sublinear function $\hat{G} = \hat{G}_X$:

$$\hat{G}(A) \triangleq 1/2 \hat{E}(\langle AX, X \rangle), \forall A \in S(d) \quad (18)$$

Thanks to (12) a bounded subset $\hat{\Theta} \subset S(d)$ exists s.t.

$$\hat{G}(A) = 1/2 \sup_{q \in \hat{\Theta}} Tr(AQ), \forall A \in S(d) \quad (19)$$

$\hat{G}(A)$ is monotonic; hence: $\hat{\Theta} \subset S_{+}(d) \triangleq \{\theta \in S(d) : \theta \geq 0\} = \{BB^T : B \in \mathbb{R}^d\}$. If $\hat{\Theta} = \{Q\}$, the classical case $X \sim N(0, Q)$ is recovered.

$\hat{\Theta}$ characterizes the covariance uncertainty of $X$. The notation is $X \overset{d}{=} N(\{0\} \times \hat{\Theta})$.

**g-normal distribution.** A vector $Y = (Y_1, \ldots, Y_d)$ in a sublinear expectation space $(\Omega, \mathcal{H}, \hat{E})$ is g-normal distributed if:

$$a^2Y + b^2Y \overset{d}{=} (a^2 + b^2)Y, \forall a, b \in \mathbb{R} \quad (20)$$

The distribution of $Y$ is characterized by a sublinear function $\tilde{g} = g_Y : \mathbb{R}^d \to \mathbb{R}$ defined by:

$$\tilde{g}(p) \triangleq \hat{E}(\langle p, Y \rangle), \quad p \in \mathbb{R}^d \quad (21)$$

Thanks to (12), there exists a bounded subset $\mathcal{F} \subset \mathbb{R}^d$ s.t.

$$\tilde{g}(p) = \sup_{q \in \mathcal{F}} \langle p, q \rangle, \quad p \in \mathbb{R}^d \quad (22)$$
It can be shown that the distribution of \( Y \) is given by:
\[
\hat{F}_Y(\varphi) = \hat{E}(\varphi(Y)) = \sup_{v \in \Theta} \varphi(v) = \sup_{v \in \Theta} \int \varphi(x) \mu_v(dx),
\]
\( \forall \varphi \in C_{l,Lip}(\mathbb{R}^d) \), where \( \mu_v \) is the Dirac measure in \( \{v\} \).
In other words, \( Y \) has the worst distribution for the uncertainty set defined by Dirac measures on \( \Theta \).
The notation is \( Y \triangleq \mathcal{N}(\Theta \times \{0\}) \).

The two distributions above can be generalized in a nontrivial manner in the following situation.

**G-normal distribution with uncertain mean.** A pair of \( d \)-dimensional vectors \( X, Y \) in a sublinear expectation space \((\Omega, \mathcal{F}, \mathbb{E})\), where \( X \) is \( G \)-normal and \( Y \) is \( g \)-normal distributed, is said \( G \)-normal distributed if, \( \forall a, b \geq 0 \):
\[
(aX + bX, a^2Y + b^2Y) \triangleq (\sqrt{a^2 + b^2}X, (a^2 + b^2)Y)
\]
(23)
The distribution of \( (X, Y) \) can be characterized by the function \( G : \mathbb{R}^d \times \mathcal{S}(d) \rightarrow \mathbb{R} \) defined by:
\[
G(p, A) \triangleq \mathbb{E}\left((1/2 \left< AXX^T + (p, p) \right>)\right)
\]
(24)
For \( (t, x, y) \in [0, \infty) \times \mathbb{R}^d \times \mathbb{R}^d \), let:
\[
u(t, x, y) \triangleq \mathbb{E}\left(\varphi(x + \sqrt{t}X, y + tY)\right)
\]
(26)
The distribution of \( (X, Y) \) is generated by the parabolic PDE in \( \nu \) on \( [0, \infty) \times \mathbb{R}^d \times \mathbb{R}^d \):
\[
\partial_t \nu - G(D_x \nu, D_y^2 \nu) = 0 \quad \nu|_{t=0} = \varphi
\]
(27)
Eq.(27) is the \( G \)-heat equation. Because of the presence of the supremum in the definition of \( G \), the \( G \)-heat equation is of Hamilton-Jacobi-Belman (HJB) type following P.-L. Lions (Crandall et al. (1992)).

This allows to show that, if two pairs \( (X, Y) \) and \( (\bar{X}, \bar{Y}) \) are \( G \)-normal with the same function \( G \), namely both satisfy (24), then they are identically distributed in the sense of the present context, i.e. \( (X, Y) \triangleq (\bar{X}, \bar{Y}) \).

### 3.3 Combined CLT and LLN

Let \( \{X_i, Y_i\}_{i=1}^\infty \) be a sequence in the space \( (\Omega, \mathcal{F}, \mathbb{E}) \), with values in \( \mathbb{R}^d \times \mathbb{R}^d \) and i.i.d. in the following sense:
\[
(X_{i+1}, Y_{i+1}) \overset{d}{=} (X_i, Y_i)
\]
(28)

\( (X_{i+1}, Y_{i+1}) \) is independent of \( \{(X_1, Y_1), \ldots, (X_i, Y_i)\} \) for all \( i = 1, 2, \ldots \). Note that (28) allows for nonstationarity in \( (X, Y) \), thanks to the property of the sets distributions associated with two identically distributed uncertain variables displayed just below (13). Assume also that
\[
\mathbb{E}(X_1) = \mathbb{E}(\bar{X}_1) = 0
\]
(29)
Note that (29) implies that \( \mathbb{E}(X) \) is time invariant. Define:
\[
S_N \triangleq \sum_{i=1}^N \left( X_i / \sqrt{N} + Y_i / N \right)
\]
(30)
Then for all \( \varphi \in C(\mathbb{R}^d) \) satisfying a linear growth condition:
\[
\frac{\mathbb{E}\left(\varphi(S_N)\right)}{N \rightarrow \infty} \rightarrow \mathbb{E}\left(\varphi(\xi + \zeta)\right)
\]
(31)
where \((\xi, \zeta)\) is a pair of \( G \)-normal distributed vectors and where the sublinear function \( G : \mathbb{R}^d \times \mathcal{S}(d) \rightarrow \mathbb{R} \) is defined, for all \( (p, A) \in \mathbb{R}^d \times \mathcal{S}(d) \), by:
\[
G(p, A) \triangleq \mathbb{E}\left((1/2 \left< AX_1, X_1 + (p, Y_1) \right>)\right)
\]
(32)

4. **COMBINING BOTH THEORIES IN HANDLING UNCERTAINTIES FOR FDI**

It is useful to recall that CLT and LLN are jointly used in the derivation of the local approach in Table 1. Then, facing Table 1 and Eq. (8) with Eqs. (30), (33) suggests:
- To define, for \( \Upsilon \) fixed:
\[
X_i \overset{d}{=} K(\theta_0, Z_i)
\]
(34)
\[
Y_i \overset{d}{=} \partial/\partial \theta K(\theta, Z_i)|_{\theta=\theta_0} \Upsilon
\]
- To see \( (X_i, Y_i) \) as uncertain variables, namely admitting a representation of the type (12);
- Assuming that the sequence \( \{(X_i, Y_i)\}_{i=1}^{\infty} \) satisfies (28), to apply (33) to (34) under \( \mathbb{H}_0 \) and \( \mathbb{H}_1 \) in (7).

In other words, the whole picture of Table 1 is kept except that the CLT and LLN are now those of Peng’s theory of uncertain distributions. The aim is to obtain a result similar to (8) for the residual (3), where \( \Sigma \) and \( M \) are now defined in (25). Writing the counterpart of the GLR and thus of test (9) for uncertain distributions is not straightforward and is currently under study.

5. **CONCLUSION**

In this paper, the key issue of handling uncertainties in statistical FDI is discussed. Two different types of uncertainties are distinguished. It is argued that two theories should be used accordingly: the statistical local approach previously developed by the authors and the recent theory of uncertain distributions due to S. Peng. The fundamental steps of each theory are described. While the detailed combination of both theories is still under investigation, it is believed that, among other elements of the paper, bringing the key steps of the theory of uncertain distributions used for financial pricing and risk management to the attention of the modeling and identification community is useful.
REFERENCES


