INSTITUT NATIONAL DE LA STATISTIQUE ET DES ETUDES ECONOMIQUES
Série des Documents de Travail du CREST
(Centre de Recherche en Economie et Statistique)

n° 2002-08

Testing for the Mean of Random Curves : from Penalization to Dimension Selection

A. MAS[1]

[1] CREST et Université de Toulouse 3. Email : mas@cict.fr

Testing for the mean of random curves : from penalization to dimension selection.

André Mas

Université Toulouse III - CREST-INSEE, Paris.

Résumé : Soit $X_1, ..., X_n$ un échantillon i.i.d. de courbes aléatoires de moyenne m. On propose une procédure asymptotique de test de

$$H_0 : m = m_0 \text{ contre } H_a : m \neq m_0.$$

Cele-ci est basée sur des principes similaires au contexte fini-dimensionnel. L'estimation de l'inverse de l'opérateur de covariance de l'échantillon nous amène à résoudre un probl ème linéaire inverse mal posé original. On montre que le test est convergent et on donne un minorant de la puissance sous une alternative g énérale.

Mots-clés : Statistique fonctionnelle, opérateurs de covariance, problème linéaire inverse, distance de Prokhorov.

Abstract : Let $X_1, ..., X_n$ be an i.i.d. sample of random curves (of Hilbert space valued random elements) with mean m. An asymptotic test of

$$H_0 : m = m_0 \text{ vs } H_a : m \neq m_0$$

is proposed. This procedure is based on the same principles as in the finite dimensional case. Inverting the covariance operator leads to a non standard linear inverse problem. Convergence of the test is obtained as well as a bound for the power.

Key-words : Functional statistics, covariance operators, linear inverse problem, Prohorov's metric.

1. Introduction

1.1. The functional statistics setting. Inference on random curves is undoubtedly a soaring area in nonparametric statistics. Modern computational techniques makes it possible to deal with "high dimensional random vectors". Data that are obtained from an underlying continuous-time process, for instance, are extremely common : in finance, climatology, medicine amongst many others. Random curves collected from (independent or not) experiments are also likely to be studied by non parametric techniques. Some references are Franck and Friedman (1993), Cavallini et alii (1994), Besse et alii (2000) for applications in respectively chemometrics, industry and meteorology.
This trend revealed that, conversely to probabilists, statisticians sometimes lack theoretical results for studying random functions. Many authors anyway paid attention to these topics, developping methods to link this rather formal framework with the statistician's "everyday's life". Amongst these are Ramsay and Silverman (1997) and their famous monograph. Antoniadis and Beder (1989) focused on the gaussian case. Dauxois, Pousse and Romain (1982) in an earlier paper investigated the principal component analysis for Hilbert-valued random variables. Recently several authors generalized to the functional framework standard models in finite dimension : the linear regression model in Cardot, Ferraty, Sarda (1999), the autoregressive model for time series in Bosq (2000), the infinite moving average model in Mas (2002b).
We propose an asymptotic procedure to test for the mean of a random function. Let $X_1, ..., X_n$ be an i.i.d. sample of random curves with mean m (note that m is also a curve). We should write $X_i(\omega, t)$ where for fixed ω $X_i(\omega, .)$ is a the path of the curve and where for fixed t $X_i(., t)$ is a real random variable. But for the sake of simplicity both indices will be dropped. Estimating the curve m is a typical non parametric problem (involving techiques like kernel, splines, wavelets, etc). We refer for instance to Rice and Silverman (1991), Antoniadis, Gregoir and Mc Keagie (1994) amongst many others.
The test may be written :

$$\begin{cases} H_0 : m = m_0 \\ H_a : m \neq m_0. \end{cases}$$

Although the interest for such a question is crucial in order to validate the estimate of the preceding step, it seems that it was not adressed in the literature. In a recent work Cardot et alii (2002) proposed a test for the regression operator in a linear model with functional inputs. Here, the procedure under concern is quite different. It remains asymptotic and *truly* infinite dimensional : it does not rely on a finite dimensional approximation of m. The limiting distribution is however extremely simple. Some may view it as goodness of fit test. We will see further that the main difficulties arise from the fact that the background of the test is connected with an inverse linear problem. The studied random curves will always be seen as random variables with values in an infinite dimensional, real and separable Hilbert space H endowed with inner product

$\langle .,. \rangle$ and norm $\|\cdot\|$. The
Hilbert space setting enables to consider different sorts of basis, makes computations easier (especially as far as the central limit theorem is concerned). In the special case when $H = L^2([0,1]$
and for all random functions u and v in H (u and v are consequently defined on $\Omega \times [0,1]$ where Ω
is an abstract probability space) $\langle u, v \rangle$ is a
random number :

$$\langle u, v \rangle (\omega) = \int_0^1 u(\omega, t) v(\omega, t) dt. \tag{1.1}$$

Spaces of smooth functions, such as Sobolev spaces, may (and are often) prefered to L^2 spaces for stability reasons. We refer to Silverman (1996) for developments of this approach. Formally, this does not change much the inner product defined in (1.1) : Lebesgue's measure dt
is replaced with another non-finite measure $d\mu(t)$.

1.2. **Preliminary facts.** Let $X_1, ..., X_n$ be an i.i.d. sample of Hilbert-valued random variables
with mean m.
We denote Γ (resp. Γ_n) the covariance operator of X_1
(resp. the empirical covariance operator of the sample). These operators are bounded linear mappings from H to H. They are defined this way : for all
x in H,

$$\begin{aligned} \Gamma(x) &= E\left[\langle X_1 - m, x \rangle (X_1 - m)\right], \\ \Gamma_n(x) &= \frac{1}{n} \sum_{k=1}^{n} \left[\langle X_k - m, x \rangle (X_k - m)\right]. \end{aligned}$$

In the following, for all u, v in H, $u \otimes v$ just stands for the rank-one operator defined for all x in H by $u \otimes v(x) = \langle$
u,xv. With these notations, we get :

$$\begin{aligned} \Gamma &= E\left[(X_1 - m) \otimes (X_1 - m)\right], \\ \Gamma_n &= \frac{1}{n} \sum_{k=1}^{n} \left[(X_k - m) \otimes (X_k - m)\right]. \end{aligned}$$

Remind that the notation $Y_n = O_P(t_n)$ means that the
sequence of random variables $Y_n t_n^{-1}$ is bounded in probability.
For any linear operator T defined on and with values in H, $\|$
T$_\infty$ stands for the usual norm of continuous linears
operators i.e. $\|T\|_\infty = \sup \|Tx\|$ for
x $intheunitballof$H.$Atthispointweenouncethefourfollowing$
facts (references to check these points are given below). The first one is a consequence of the central limit theorem for Hilbert space-valued random variables :

Fact 1 : If $E\|X_1\|^2 < +\infty$,

$$S_n = \frac{1}{\sqrt{n}} \sum_{k=1}^{n} (X_k - m) \underset{n \to +\infty}{\overset{w}{\to}} G, \tag{1.2}$$

where $\overset{w}{\to}$ means "converges weakly" or "in distribution" and G is a centered gaussian random function (a H-valued random element) with covariance operator Γ.

Fact 2 : If $E\|X_1\|^4 < +\infty$,

$$\|\Gamma_n - \Gamma\|_\infty = O_P\left(\frac{1}{\sqrt{n}}\right). \tag{1.3}$$

Under the moment assumptions on X mentioned just above, it is well-known that Γ is a positive selfadjoint compact operator. Once and for all by $\lambda_1 \geq \lambda_2 \geq ... \geq 0$ we denote the ordered sequence of its eigenvalues (associated to the eigenvectors $e_1, e_2, ...$), the series $\sum_p \lambda_p$ is finite. We will always suppose that Γ is one to one, which means that the sequence of its strictly positive eigenvalues is infinite and that the set ker $\Gamma = \{h \in H : \Gamma h = 0\}$ is empty.

Fact 3 : The random element G may be written :

$$G = \sum_{k=1}^{+\infty} \eta_k \sqrt{\lambda_k} e_k \quad a.s. \tag{1.4}$$

where the η_k's are i.i.d. centered normal real random variables with unit variance.

Fact 4 : The operator $\Gamma^{1/2}$ is defined by $\Gamma^{1/2} = \sum_{k=1}^{+\infty} \sqrt{\lambda_k} (e_k \otimes e_k$. It is a compact operator. When Γ is one to one, its inverse $\Gamma^{-1/2}$ is defined on a domain $\mathcal{D}$ in H (D$isadensevectorsubspaceof$H)$.The linear mapping \Gamma^{-1/2}$ is unbounded (i.e. $\Gamma^{-1/2}$ is continuous point *at no point of* $\mathcal{D}$*) which also means* $\left\|\Gamma^{-1/2}\right\|_\infty = +\infty$ *. As a consequence* $\Gamma^{-1/2} X_1$ *cannot be considered as a (bounded) random variable.*

The proof of the first point in Fact 2 may be found for instance in Dauxois, Pousse and Romain (1982). The well-known Fact 3 is mentioned in Grenander (1963). More information about Fact 4 may be found in Dunford-Schwartz (1988).

1.3. **Formulating the problem.** *Now, in order to understand the particularities of the situation, let us* *consider first what happens in the finite dimensional setting. Let $Y_1, ..., Y_n$ be a sample of i.i.d. vectors of R^p with mean y. The test is the following*

$$\begin{cases} H_0 : y = y_0 \\ H_a : y \neq y_0. \end{cases}$$

The operator Γ_n is replaced with a square matrix of size p, say $M_n(p)$, the emprirical covariance matrix of the sample, which is often supposed to be invertible for n large enough. Then the test statistics (usually chi-square) is derived from :

$$M_n(p)^{-1/2}\left(\frac{1}{\sqrt{n}}\sum_{k=1}^{n}(Y_k - y_0)\right)$$

which converges under the null hypothesis to a gaussian random variable whose covariance matrix is the identity.
*In our framework we still have **(Fact 1)** : S_n converges weakly to G but the distribution of G depends on the unknown λ_i's. These eigenvalues may be viewed as nuisance parameters. Besides Γ_n, conversely to M_n is never invertible in H even for large n since its range (the linear span of $X_1, ..., X_n$) is of finite dimension. Considering "$\Gamma_n^{-1/2} S_n$" makes no sense, as well as "$\Gamma^{-1/2} G$" **(Fact 4)**. Consequently our goal is double.*
First we should estimate (approximate would be more accurate here) the operator $\Gamma^{-1/2}$ by, say, L_n which should be a random linear and continuous mapping depending on Γ_n. Then we will have to study weak convergence for $L_n S_n$. This approach will be made more precise in the next section. Anyway we should expect the norm of L_n to be a non decreasing sequence tending to $+\infty$. We could also say that the sequence of operators L_n is unstable or ill-conditioned.
Finally, copying the finite dimensional approach to our test procedure leads us to a non standard ill-posed inverse problem : G will be approximated by S_n and $\Gamma^{-1/2}$ by L_n but S_n and L_n are connected via the sample $X_1, ..., X_n$. Fortunately we will see that it is possible to propose a convergent test procedure and even to obtain a rate of convergence for this statistic.

1.4. **The Prokhorov metric for probability measures.** *In the sequel we will need a metric defined on spaces of measures. This (Prokhorov) metric, in many cases, metrizes the topology of weak convergence for probability measures on a metric space E. It will be denoted π_E. For the definition and main properties, we refer to Billingsley (1968), Dudley (1968) and Araujo and Giné (1980). We will make use of the two following results keeping the notations and assumptions made above.*

Theorem 1. *(Yurinski, 1977, p.244) If $\lambda_k = O\left(k^{-(1+\varepsilon)}\right)$ for some $\varepsilon > 0$, then*

$$\pi_H(S_n, G) = O\left(n^{-\varepsilon/(6+8\varepsilon)} \log n\right).$$

For further purpose we may propose an analogous result in the case of an exponential decay of the eigenvalues :

Proposition 1. *If* $\lambda_k = O(\exp(-\mu k))$ *for some* $\mu > 0$ *then*

$$\pi_H(S_n, G) = O\left(\frac{(\log n)^{3/4}}{n^{1/8}}\right).$$

This proposition will be proved, following Yurinskii's method, at the end of the article.

2. Main result

2.1. **Pseudo-inverse estimators.** *The literature coping with ill-posed linear problems or inverse linear problems is especially rich. The problem of approximating inverses of selfadjoint compact operators is absolutely not new. It is adressed in Nashed and Wahba (1974), Tikhonov and Arsenin (1977), Groetsch (1993) amongst many others. The main point is always to regularize a matrix* M *(resp. an operator* S*) which is invertible but "not by much" (resp. unbounded). This property implies that for any vector* x, Mx *(resp.* Sx*) may have large variations even when* x *does not vary much. Numerous procedures were proposed. We will keep two of them since they are suited to our problem and may be easily implemented.*
In the basis of eigenvectors of Γ *we may write :*

$$\Gamma^{-1/2} = diag\left(\lambda_1^{-1/2}, \lambda_2^{-1/2}, ..., \lambda_k^{-1/2}, ...\right).$$

As mentioned above, we denote $\mathcal{D}$ *the domain of this operator i.e. the set of points* x *in* H *for which* $\Gamma^{-1/2}x$ *has a finite norm.*
Clearly if x *is expressed in the basis* e_i, $x = \sum x_k e_k$. *We have :*

$$\mathcal{D} = \left\{ x \in H : \sum_k \frac{x_k^2}{\lambda_k} < +\infty. \right\}.$$

We recall that we are aiming at approximating $\Gamma^{-1/2}$. *A first idea consists in deleting all the terms for large* k. *Indeed let us fix an integer* p. *In the basis* e_i *we set*

$$\Gamma_p^{-1/2} = diag\left(\lambda_1^{-1/2}, \lambda_2^{-1/2}, ..., \lambda_p^{-1/2}, 0, 0, ...\right).$$

It is clear that $\Gamma_p^{-1/2}$ *is a bounded operator with norm* $\lambda_p^{-1/2}$. *We will say that* $\Gamma_p^{-1/2}$ *is the S.C. estimator of order* p *of* $\Gamma^{-1/2}$ *(S.C. stands for Spectral Cut).*
Another well-known approximation is given by the penalized operator

$$(\Gamma + \alpha I)^{-1/2} = diag\left((\lambda_1 + \alpha)^{-1/2}, (\lambda_2 + \alpha)^{-1/2}, ..., (\lambda_k + \alpha)^{-1/2}, ...\right)$$

in the basis of eigenvectors of Γ. *The penalization term is* α I *where*α *is a strictly postive real number and* I *is the identity operator on* H. *This operator is continuous for all strictly postive* α *with norm* $1/\alpha$.
Note that both operators depend on a parameter : p *for the first one and* α *for the second. In the sequel* p *will increase whereas* α *will tend to zero and the norm of these operators will go to infinity. We could easily prove that for all* x *in* $\mathcal{D}$, $\Gamma_p^{-1/2}x$ *as well as* $(\Gamma+\alpha I)^{-1/2}x$ *both tend to* $\Gamma^{-1/2}x$ *in* H. *We should expect the parameters to be indexed by* n, *the size of the sample, if* Γ *is replaced with* Γ_n.

2.2. **Scheme of the proof.** *Let us set* $\Gamma_p^{-1/2}$ *be the S.C. approximation of* $\Gamma^{-1/2}$ *defined in the basis of eigenvectors of* Γ *by :*

$$\Gamma_p^{-1/2} = diag\left(\lambda_1^{-1/2}, \lambda_2^{-1/2}, ..., \lambda_p^{-1/2}, 0, 0, ...\right).$$

It is easily seen that $\Gamma_p^{-1/2}G$ *is a gaussian random function with a degenerated covariance operator namely* Π^p *(i.e. the projection operator on the* p *first eigenvectors of* Γ*). Now in view of (1.4) :*

$$\left\|\Gamma_p^{-1/2}G\right\|^2 = \sum_{k=1}^{p}\eta_k^2.$$

Applying the central limit theorem on the real line for i.i.d variables, the η_k^2*s, we get*

$$\frac{\left\|\Gamma_p^{-1/2}G\right\|^2 - p}{\sqrt{p}} \underset{p\to+\infty}{\overset{w}{\to}} N(0,2). \tag{2.1}$$

This fact (which is nothing but the classical chi-square approximation when the degree of freedom tends to infinity) is the starting point for our test procedure. The first idea consists in replacing G *by* S_n. *The next step would consist in replacing* $\Gamma_p^{-1/2}$ *by its S.C. approximation based on* Γ_n. *But it turns out that the penalized estimator of* Γ_n *is more appropriate for at least three reasons :*

- *The S.C. estimator is based on the functional Principal Component Analysis of* Γ_n. *It is necessary to estimate the eigenvectors and eigenvalues of this random operator before projecting the sample* $X_1 - m, ..., X_n - m$ *on these eigenvectors (to obtain principal components as a by-product). The estimation procedure is consequently not that simple and usually entails serious stability problems. Estimating the penalized estimate just requires a* ***nonrandom basis****, (e.g. spline or sinusoidal).*
- *As it may be seen in Cardot et alii (1999), and Bosq (2000), convergence rates of estimates steming from the S.C. procedure always depend on the speed of decay of the (unknown) eigenvalues of* Γ. *Assumption are usually made, restricting the generality of the results.*

- *Speed of convergence of estimates involving the S.C. estimator are usually not good, due to the very slow rate of uniform convergence of the empirical eigenvectors of* Γ_n *to the eigenvectors of* Γ.
  Also note that, conversely to the S.C. estimator, the norm of the penalized one ($1/\alpha_n$*) is* ***nonrandom*** *and does not depend on the rate of decay of the eigenvalues of* Γ.

Suppose that k_n *is an increasing sequence of real numbers. The test statistic to be considered should then be :*

$$\frac{1}{\sqrt{k_n}}\left(\left\|(\Gamma_n+\alpha_n I)^{-1/2}S_n\right\|^2-k_n\right)$$

where we set once and for all :

$$S_n=\frac{1}{\sqrt{n}}\sum_{k=1}^{n}(X_k-m_0)$$

and we aim at proving that such a statistics converges weakly to a N (*0,2 distribution under the null hypothesis. The proof takes the following steps :*

First step : *We prove that*

$$\frac{1}{\sqrt{k_n}}\left(\left\|(\Gamma_n+\alpha_n I)^{-1/2}S_n\right\|^2-\left\|(\Gamma+\alpha_n I)^{-1/2}S_n\right\|^2\right)\xrightarrow{P}0. \tag{2.2}$$

The deterministic operator Γ *replaces the random operator* Γ_n.

Second step : *We prove that the limit in distribution of*

$$\frac{1}{\sqrt{k_n}}\left(\left\|(\Gamma+\alpha_n I)^{-1/2}S_n\right\|^2-k_n\right)$$

is the same as the limit in distribution of $\frac{1}{\sqrt{k_n}}$ (
$\left\|(\Gamma+\alpha_n I)^{-1/2}G\right\|$
${}^2-k_n$.

Third step : *We prove that*

$$\frac{1}{\sqrt{k_n}}\left(\left\|(\Gamma+\alpha_n I)^{-1/2}G\right\|^2-\left\|\Gamma_{k_n}^{-1/2}G\right\|^2\right)\xrightarrow{P}0.$$

The weak convergence of the test statistic to a $N(0,2)$ *distribution will then follow from (2.1).*

Remark 1. : *This last step links the penalization approximation of* $\Gamma^{-1/2}$ *with its S.C. estimate, which explains the title of the article.*

Before giving the main results, we recapitulate the assumptions made above :

H1 *:* $E\left\|X_1\right\|^4 < +\infty.$
H2 *:* Γ *is one to one.*

The following subsection both contains a main theorem, followed by the consistence of this test. These results are made more precise when the rate of decay of the eigenvalues of Γ *is known or estimated.*

2.3. **Weak convergence and consistence.** *The first result announces the existence of the test procedure. It remains formal since no information is available on the choice of the penalization and dimension parameters. But it should be stressed that it is given -conversely to many areas dealing with inverse problems (deconvolution, image anlaysis, etc.)- under very mild assumptions on the* X_i*'s (their distribution is unknown) as well as on the spectrum of the compact operator* Γ.

Theorem 2. *: Under **H1** and **H2** there exists sequences* k_n *(of integers) and* α_n *(of nonnegative numbers) respectively increasing to infinity and decreasing to zero such that under the null hypothesis* $m = m_0$ *:*

$$T_{pen}(n) = \frac{1}{\sqrt{k_n}}\left(\left\|\left(\Gamma_n + \alpha_n\right)^{-1/2} S_n\right\|^2 - k_n\right) \underset{n\to+\infty}{\overset{w}{\to}} N(0,2).$$

Remark 2. *:*
In fact the proofs even yield a more precise result : an explicit bound for the Prokhorov distance between the distributions of T_{pen} *and* $N(0,2)$. *This bound is made of three terms (corresponding to the three steps mentioned above) all depending on* k_n *and* α_n. *Hopefully, it is possible to find compatible conditions, connecting both parameters, so that this bound tends to zero.*

Proposition 2. *: The test, described above and based on the computation of* T_{pen}*, is consistent.*

It turns out that, when the rate of decay of the eigenvalues λ_k *is known,* α_n *and* k_n *may be explicitely computed.*

Theorem 3. *: When **H1** and **H2** hold and when the eigenvalues decay at a geometric rate, say* $\lambda_k = mk^{-($
1+ε for some positive constant m *and* ε
¿$\mu > 0$, *then one can take*

$$\alpha_n = n^{-1/4} \text{ and } k_n = n^{\mu(1+\varepsilon)^{-2}/4}.$$

Theorem 4. *: When **H1** and **H2** hold and when the eigenvalues decay at an exponential rate, say* $\lambda_k = c\exp($
-δk *for some positive constant* c *and* $\delta > \mu > 0$, *then*

$$\alpha_n = n^{-1/4} \text{ and } k_n = \frac{\mu}{4\delta^2} \log n.$$

Remark 3. *:*
*In the special case when the curves are gaussian random functions, it is easily seen that the second step mentioned in the previous section always holds (as well as assumption **H1**) since the distribution of S_n is, for all n similar to the distribution of G. But this does not change the final results. Indeed, it turns out that the condition obtained at this stage is weaker than the one obtained at the first step.*

2.4. **Test procedure and bounds for the power.** *For the sake of completeness, the test procedure is given below :*

- *Fix m_0 and compute T_{pen} where $S_n = n^{-1/2} \sum_{k=1}^{n}$ $(X_k - m_0)$ and*

$$\Gamma_n = \frac{1}{n} \sum_{k=1}^{n} (X_k - m_0) \otimes (X_k - m_0) .$$

- *Fix a level of significance κ anc compute u_κ such that $P(|N(0,1)| \leq u_\kappa)$ =1-κ*
- *If $|T_{pen}| \leq \sqrt{2} u_\kappa$ H_0 is accepted otherwise, it is refused.*

Although a thorough study of the power of this test or of local alternatives goes beyond the scope of this article (since it involves, for instance, large or moderate deviations techniques on infinite dimensional spaces), it seems of a real intest to give an easily obtained lower bound. This bound will precise the rate of convergence to one of the power.
Before stating the following Proposition, note that if $m \neq m_0$ it is always possible to find an integer j such that $\langle$ m,e_j $\neq \langle m_0, e_j \rangle$.

Proposition 3. *: When **H1** and **H2** hold and H_a is true :*

$$P\left(|T_{pen}| > \sqrt{2} u_\kappa\right) \geq \left(1 - \frac{c}{\alpha_n^2 n^{3/2}}\right)$$

where c is some constant.

Remark 4. *:*
*This bound is obtained, as **H1** holds, by elementary (Chebyshev and Markov) inequalities. When the assumptions of Theorems 3 and expo hold $\alpha_n^2 n^{3/2} = n$.*

3. Proofs

The proof of Theorem 2 relies, as announced on the forthcoming propositions.

Proposition 4. *: **(First step)** We have :*

$$\frac{1}{\sqrt{k_n}}\left|\left\|(\Gamma_n+\alpha_n I)^{-1/2}S_n\right\|^2-\left\|(\Gamma+\alpha_n I)^{-1/2}S_n\right\|^2\right| = O_P\left(\frac{1}{\alpha_n^2\sqrt{nk_n}}\right). \tag{3.1}$$

Proof of the Proposition :

$$\begin{aligned}
&\frac{1}{\sqrt{k_n}}\left|\left\|(\Gamma_n+\alpha_n I)^{-1/2}S_n\right\|^2-\left\|(\Gamma+\alpha_n I)^{-1/2}S_n\right\|^2\right|\\
= {} & \frac{1}{\sqrt{k_n}}\left|\left\langle(\Gamma_n+\alpha_n I)^{-1}S_n,S_n\right\rangle-\left\langle(\Gamma+\alpha_n I)^{-1}S_n,S_n\right\rangle\right|\\
= {} & \frac{1}{\sqrt{k_n}}\left|\left\langle\left((\Gamma_n+\alpha_n I)^{-1}-(\Gamma+\alpha_n I)^{-1}\right)S_n,S_n\right\rangle\right|\\
\leq {} & \frac{1}{\sqrt{k_n}}\left\|\left((\Gamma_n+\alpha_n I)^{-1}-(\Gamma+\alpha_n I)^{-1}\right)S_n\right\|\|S_n\|\\
= {} & \frac{1}{\sqrt{k_n}}\left\|\left((\Gamma+\alpha_n I)^{-1}(\Gamma-\Gamma_n)(\Gamma_n+\alpha_n I)^{-1}\right)S_n\right\|\|S_n\|\\
\leq {} & \frac{1}{\alpha_n^2\sqrt{k_n}}\|\Gamma-\Gamma_n\|_\infty\|S_n\|^2 = O_P\left(\frac{1}{\alpha_n^2\sqrt{nk_n}}\right).
\end{aligned}$$

which is the intended result. We used Cauchy-Schwarz inequality from line three to four and Fact 2 for the last inequality.

Proposition 5. ***(Second step)*** *: It is possible to find a sequence α_n such that the limiting distribution of*

$$T_n' = \frac{1}{\sqrt{k_n}}\left(\left\|(\Gamma+\alpha_n I)^{-1/2}S_n\right\|^2-k_n\right)$$

is the same (if it exists) as the limiting distribution of

$$T_{G,n} = \frac{1}{\sqrt{k_n}}\left(\left\|(\Gamma+\alpha_n I)^{-1/2}G\right\|^2-k_n\right).$$

In terms of the Prokhorov metric the distance between the distributions of these random variables is bounded by $\alpha_n^{-1}k_n^{-1/2}\rho_n$ where ρ_n is a sequence of positive numbers tending to zero. This sequence depends only on the Prokhorov distance between the distribution of S_n and G.

Proof of the Proposition :

The proof of this lemma is based on a recent result by Mas (2002a). It should be first stressed that calculations show that T_n' cannot be written as a sum of n random variables. Consequently, it appears *that it is not possible to prove the weak convergence of* T_n' *by standard methods. The famous theorems of preservation of weak convergence by mappings (e.g. Billingsley, 1968 p.30 and 31) cannnot be applied either. Reformulating the Proposition and identifiying a random variable on* H *with the measure it induces we will get the desired result if*

$$\pi_R\left(T_n', T_{G,n}\right) \rightarrow 0.$$

Or equivalently

$$\frac{1}{\sqrt{k_n}}\pi_R\left(\left\|\left(\Gamma+\alpha_n I\right)^{-1/2} S_n\right\|^2, \left\|\left(\Gamma+\alpha_n I\right)^{-1/2} G\right\|^2\right) \rightarrow 0.$$

First we rewrite $\|(\Gamma+\alpha_n I)$
$^{-1/2}S_n^2$ *as* $tr\left((\Gamma+\alpha_n I)\right.$
$^{-1}S_n \otimes S_n$*. Now let us define the sequence of functions* $\tau_n : H \rightarrow R$ *by*

$$\tau_n(h) = tr\left(\left(\Gamma+\alpha_n I\right)^{-1} h \otimes h\right).$$

We have

$$\pi_R\left(T_n', T_{G,n}\right) = \frac{1}{\sqrt{k_n}}\pi_R\left(\tau_n(S_n \otimes S_n), \tau_n(G \otimes G)\right).$$

Remind that we know that $\pi\left(S_n, G\right) \rightarrow 0$*. Mas (2002a) proved that it is possible to bound* $\pi_R\left(T_n'\right.$
$,T_{G,n}$ *by a function of* $\pi_H\left(S_n, G\right)$*. This function depends on the modulus of continuity of the function* τ_n *on the balls (centered at zero) of* H*. More precisely we obtain successively :*

$$\begin{aligned}
\pi_R\left(\tau_n(S_n \otimes S_n), \tau_n(G \otimes G)\right) &\leq \alpha_n^{-1}\pi_R\left(\|S_n\|^2, \|G\|^2\right) \\
&\leq c\alpha_n^{-1}\pi_H\left(S_n, G\right)\sqrt{-\ln\left[\pi_H\left(S_n, G\right)\right]}.
\end{aligned}$$

The first inequality stems from the fact that $tr(S_n \otimes$
$S_n) = \|S_n\|^2$ *and that* τ_n *is a linear mapping (see also Whitt (1974)). The last inequality is true for n large enough and is precisely the one given in Case 1 of paragraph 4.1 in Mas (2002a). We will denote*

$$\rho_n = \pi_H\left(S_n, G\right)\sqrt{-\ln\left[\pi_H\left(S_n, G\right)\right]}$$

and finally

$$\pi_R\left(T_n', T_{G,n}\right) = O\left(\frac{\rho_n}{\alpha_n\sqrt{k_n}}\right).$$

which is the expected result.

Proposition 6. ***(Third step)*** *: There exists a function* φ *defined and with values on the set of positive real numbers such that* $\varphi(0) = 0$ *and* φ *is continuous on a neighborhood of* 0 *with*

$$\frac{1}{\sqrt{k_n}}E\left|\left\|(\Gamma+\alpha_n I)^{-1/2}G\right\|^2-\left\|\Gamma_{k_n}^{-1/2}G\right\|^2\right|=o_P(\frac{\varphi(\alpha_n)}{\lambda_{k_n}\sqrt{k_n}}).$$

Remark 5. *The rate of decay to* 0 *of the function* φ *depends on the speed of decay to zero of the (unknown)* λ_p*'s, as will be shown in the proofs. Note anyway that, since* α_n *does not depend on* k_n*, the rate of convergence to zero of the baove expression just depends on the choice of an accurate* α_n.

Proof of the Proposition :
We have successively :

$$\begin{aligned}
R_n &= \frac{1}{\sqrt{k_n}}E\left|\left\|(\Gamma+\alpha_n I)^{-1/2}G\right\|^2-\left\|\Gamma_{k_n}^{-1/2}G\right\|^2\right| \\
&= \frac{1}{\sqrt{k_n}}E\left|\left\langle(\Gamma+\alpha_n I)^{-1}G,G\right\rangle-\langle\Gamma_{k_n}^{-1}G,G\rangle\right| \\
&= \frac{1}{\sqrt{k_n}}E\left|\left\langle\left((\Gamma+\alpha_n I)^{-1}-\Gamma_{k_n}^{-1}\right)G,G\right\rangle\right| \\
&= \frac{1}{\sqrt{k_n}}E\left|\left\langle\left(I-\Gamma_{k_n}^{-1}(\Gamma+\alpha_n I)\right)(\Gamma+\alpha_n I)^{-1/2}G,(\Gamma+\alpha_n I)^{-1/2}G\right\rangle\right|.
\end{aligned}$$

The last inequality is due to the fact that $\Gamma_{k_n}^{-1}$ *and* $(\Gamma+\alpha_n I)^{-1}$ *commute. By Cauchy-Schwarz inequality we obtain :*

$$\begin{aligned}
R_n &\le \frac{1}{\sqrt{k_n}}\left\|I-\Gamma_{k_n}^{-1}(\Gamma+\alpha_n I)\right\|_\infty E\left\|(\Gamma+\alpha_n I)^{-1/2}G\right\|^2 \\
&= \frac{\alpha_n}{\lambda_{k_n}\sqrt{k_n}}\sum_{p=1}^{+\infty}\frac{\lambda_p}{\lambda_p+\alpha_n}E\eta_p^2 \\
&= \frac{\alpha_n}{\lambda_{k_n}\sqrt{k_n}}\sum_{p=1}^{+\infty}\frac{\lambda_p}{\lambda_p+\alpha_n}. \qquad (3.2)
\end{aligned}$$

We know that there exists a nondecreasing sequence a_p *such that* a_p *goes to infinity and* $\sum_{p=1}^{+\infty}\lambda_p a_p<+\infty$. *We write :*

$$\sum_{p=1}^{+\infty}\frac{\lambda_p}{\lambda_p+\alpha_n}=\sum_{p=1}^{+\infty}\frac{a_p\lambda_p}{a_p(\lambda_p+\alpha_n)}. \qquad (3.3)$$

We are aiming at applying Lebesgue's dominated convergence theorem to the previous series. First we need the following lemma :

Lemma 1. *:*
Let $M(\alpha_n)=a_p^{-1}(\lambda_p+\alpha$
$_n)^{-1}$, $M(\alpha_n)$ *is strictly positive and* α

${}_nM\left(\alpha_n\right)$ *tends to zero as* n *tends to infinity for any sequence* α_n *tending to zero. In the previous Proposition we took* $\varphi\left(\alpha_n\right)=\alpha_nM\left(\alpha\right.$

n

Proof of the Lemma :
Suppose that the positive sequence $\alpha_nM(\alpha_n)$ *does not tend to zero. For some* $\varepsilon>0$*, one can find a subsequence, say* n'*, such that* $\alpha_{n'}M(\alpha_{n'}$
)¿ε *which may be rewritten*

$$\sup_p \frac{\alpha_{n'}}{a_p(\lambda_p+\alpha_{n'})}>\varepsilon$$

which implies the existence of some $p(n')$ *such that*

$$\frac{\alpha_{n'}}{a_{p(n')}(\lambda_{p(n')}+\alpha_{n'})}>\varepsilon/2.$$

Next we must consider two cases. First suppose that the sequence $p(n')$ *is bounded by* M *then*

$$\frac{\alpha_{n'}}{a_{p(n')}(\lambda_{p(n')}+\alpha_{n'})}\leq\frac{\alpha_{n'}}{a_1(\lambda_M+\alpha_{n'})}\to 0$$

as n' *tends to infninty which contradicts the previous inequality. Let us turn now to the case when the sequence* $p(n')$ *is not bounded. This implies the existence of another subsequence* n'
$'$ *such that* $p(n'')$ *tends to infinity. Then*

$$\frac{\alpha_{n''}}{a_{p(n'')}(\lambda_{p(n'')}+\alpha_{n''})}\leq\frac{1}{a_{p(n'')}}\to 0.$$

This proves that the sequence $\alpha_nM(\alpha_n)$ *tends to zero and enables us to define* $\varphi\left(t\right)=tM\left(t\right)$ *as enounced above which finishes the proof of the lemma.*

Now we rewrite (3.3)

$$M(\alpha_n)\sum_{p=1}^{+\infty}\frac{1}{M(\alpha_n)}\frac{a_p\lambda_p}{a_p(\lambda_p+\alpha_n)}.$$

Applying Lebesgue's dominated convergence we get

$$\frac{1}{M(\alpha_n)}\sum_{p=1}^{+\infty}\frac{a_p\lambda_p}{a_p(\lambda_p+\alpha_n)}\to 0$$

as n *tends to infinity since for fixed* p $(M(\alpha$
${}_n)a_p\left(\lambda_p+\alpha_n\right)^{-1}$ *tends to zero. At last :*

$$R_n=o_P\left(\frac{\varphi(\alpha_n)}{\lambda_{k_n}\sqrt{k_n}}\right).$$

Proof of Theorem 2 :
By (3.1) the first step is achieved if $\alpha_n^2\sqrt{nk_n}$
$\to +\infty$ *and so does the second if* $\alpha_n\sqrt{k_n}\rho$
$_n^{-1} \to +\infty$ *(see Proposition 5). Both conditions are fulfilled if we set*

$$\alpha_n^* = \min\left(n^{-1/4}, \rho_n\right). \tag{3.4}$$

At last, choosing k_n *such that* $\varphi(\alpha_n^*$
$= \lambda_{k_n^*}$ *in Proposition 6 entails the convergence to zero of the last term. Note that clearly* α_n^* *tends to zero whereas* k_n^* *goes to infinity. This finishes the proof of Theorem 2.*

Remark 6. *:*
Note for further reference that one may also easily prove that k_n/n *tends to zero.*

Proof of Proposition 2 :
It suffices to prove that, whenever $EX_1 = m \neq m_0$*, for all* $\varepsilon > 0$

$$\lim_{n\to+\infty} P\left(|T_{pen}| > \varepsilon\right) = 1.$$

But

$$\begin{aligned} P\left(|T_{pen}| > \varepsilon\right) &\geq P\left(\left\|(\Gamma_n + \alpha_n I)^{-1/2} S_n\right\|^2 > \varepsilon\sqrt{k_n} + k_n\right) \\ &\geq P\left(\left\|(\Gamma_n + \alpha_n I)^{-1/2} S_n\right\|^2 > 2k_n\right) \end{aligned}$$

for a sufficiently large n *since* k_n *goes to infinity. Denoting*
$U_n = \left\|(\Gamma_n + \alpha_n I)^{-1/2} S_n\right\|$
^{2} *and* $V_n = \|(\Gamma + \alpha_n I)$
$^{-1/2}S_n^2$*, we may write (*m_n *is a non decreasing sequence of postive real numbers) :*

$$\begin{aligned} & P\left(\left\|(\Gamma_n + \alpha_n I)^{-1/2} S_n\right\|^2 \leq 2k_n\right) \\ = \;& P\left(U_n - V_n + V_n \leq 2k_n\right) \\ \leq \;& P\left(U_n - V_n + V_n \leq 2k_n, |U_n - V_n| \leq m_n\right) \\ & + P\left(U_n - V_n + V_n \leq 2k_n, |U_n - V_n| > m_n\right) \\ \leq \;& P\left(V_n \leq 2k_n + m_n\right) + P\left(|U_n - V_n| > m_n\right) \end{aligned} \tag{3.5}$$

Taking $m_n = k_n$*, the second term on the right clearly tends to 1 and the first is bounded below by* $P\left(V_n \leq 3k_n\right)$*.*
Consequently we just need to prove that

$$P\left(\left\|(\Gamma+\alpha_n I)^{-1/2}S_n\right\|^2 \leq 3k_n\right) \to 0$$

under the alternative hypothesis. But note that since $EX_1 = m \neq m_0$, *one may pick an eigenvector* e_j *such that* $\langle$
$m, e_j \neq \langle m_0, e_j \rangle$. *Then*

$$\begin{aligned}
& P\left(\left\|(\Gamma+\alpha_n I)^{-1/2}S_n\right\|^2 \leq 3k_n\right) \\
\leq\ & P\left(\left\langle(\Gamma+\alpha_n I)^{-1/2}S_n, e_j\right\rangle^2 \leq 3k_n\right) \\
=\ & P\left(\langle S_n, e_j\rangle^2 \leq 3(\lambda_j+\alpha_n)k_n\right) \\
=\ & P\left(\left[\frac{1}{n}\sum_{k=1}^{n}\langle X_k, e_j\rangle - \langle m_0, e_j\rangle\right]^2 \leq 3(\lambda_j+\alpha_n)\frac{k_n}{n}\right). \qquad (3.6)
\end{aligned}$$

Now k_n/n *tends to zero (see Remark at the end of the proof of Theorem 2). The weak law of large numbers on the real line ensures that* $\frac{1}{n}\sum_{k=1}^{n}\langle X_k, e_j\rangle$ *tends in probability to* $\langle m, e_j\rangle$ *which is enough to conclude that the above probability tends to zero and that the proposed test is consistent.*

Corollary 1. *: If* $\lambda_p = \dfrac{m}{p^{1+\varepsilon}}$, *for all*
0¡$\mu < \varepsilon$,

$$\frac{1}{\sqrt{k_n}}\left|\left\|(\Gamma+\alpha_n I)^{-1/2}G\right\|^2 - \left\|\Gamma_{k_n}^{-1/2}G\right\|^2\right| = o_P\left(\alpha_n^{\mu/(1+\varepsilon)}k_n^{1/2+\varepsilon}\right). \qquad (3.7)$$

Proof of the Corollary :
Calculations below aim, in fact, at computing the functions M *and* φ *mentionedabove.Obviously* :

$$\sum_{p=1}^{+\infty}\frac{\lambda_p}{\lambda_p+\alpha_n} = \sum_{p=1}^{+\infty}\frac{m}{p^{1+\varepsilon}}\frac{1}{\frac{m}{p^{1+\varepsilon}}+\alpha_n}.$$

We take $a_p = p^{\mu}$ *where* $0 < \mu < \varepsilon$. *Note that* $\sum_{p=1}^{+\infty} a_p\lambda_p < +\infty$. *We have to find :*

$$\begin{aligned}
\sup_p \frac{1}{p^{\mu}\left(\frac{m}{p^{1+\varepsilon}}+\alpha_n\right)} &= \sup_p \frac{p^{1+\varepsilon-\mu}}{m+p^{1+\varepsilon}\alpha_n} \\
&\leq \sup_{x\geq 1}\frac{x^{1+\varepsilon-\mu}}{m+x^{1+\varepsilon}\alpha_n}.
\end{aligned}$$

The argmax of the above quantity is $x_n^* = K_0\alpha_n^{-1/(1+\varepsilon)}$ *with*

$$K_0 = \left(m \frac{1+\varepsilon-\mu}{\mu} \right)^{1/(1+\varepsilon)}.$$

Elementary algebra yields :

$$\sup_p \frac{1}{p^\mu (\frac{m}{p^{1+\varepsilon}} + \alpha_n)} \leq K_1 \alpha_n^{-1+\mu/(1+\varepsilon)} = M(\alpha_n)$$

with K_1 *depending on* m *and* K_0*. By (3.2) we find*
$R_n = o_P \left(\alpha_n^{\mu/(1+\varepsilon)} \right.$
$k_n^{1/2+\varepsilon}$.

Corollary 2. *: If* $\lambda_p = c \exp(-\delta p)$ *and*
0¡$\mu < \delta$

$$\text{(3.8)} \qquad \frac{1}{\sqrt{k_n}} \left| \left\| (\Gamma + \alpha_n I)^{-1/2} G \right\|^2 - \left\| \Gamma_{k_n}^{-1/2} G \right\|^2 \right| = o_P \left(k_n^{-1/2} \alpha_n^{\mu/\delta} \exp(\delta k_n) \right).$$

Proof of the Corollary :
This time

$$\sum_{p=1}^{+\infty} \frac{\lambda_p}{\lambda_p + \alpha_n} = \sum_{p=1}^{+\infty} \frac{c \exp(-\delta p)}{c \exp(-\delta p) + \alpha_n}.$$

We take $a_p = \exp(\mu p)$ *and we have to find*

$$\sup_p \frac{\exp(-\mu p)}{c \exp(-\delta p) + \alpha_n} \leq \sup_{x \geq 1} \frac{\exp(-\mu x)}{c \exp(-\delta x) + \alpha_n}.$$

This supremum is obtained for $x_n^* = K - \delta^{-1} \log \alpha_n$
(with $K = \delta^{-1} \log(c(\delta/\mu) - 1)$*).*
We get

$$\sup_{x \geq 1} \frac{\exp(-\mu x)}{c \exp(-\delta x) + \alpha_n} = K^* \alpha_n^{\mu/\delta - 1} = M(\alpha_n),$$

where K^* *depends on* K*,* δ*,* μ*. We get the desired reult once more by (3.2).*

Proof of Theorem 3 :
By Theorem 1 we know that

$$\rho_n = n^{-\varepsilon/(6+8\varepsilon)} \log n \cdot \sqrt{-\log\left(n^{-\varepsilon/(6+8\varepsilon)} \log n\right)} = O\left(n^{-\varepsilon/(6+8\varepsilon)} (\log n)^{3/2}\right)$$

and since $n^{-\varepsilon/(6+8\varepsilon)} (\log n)$
$^{3/2} \geq n^{-1/4}$*, we are sure that, at least for a sufficiently large* n*,* $\alpha_n^* = \min\left(n^{-1/4}, \rho_n\right) = n^{-1/4}$
by (3.4). Consequently, the final condition (on R_n*), namely equation (3.7), is* $n^{\mu/(4+4\varepsilon)}$
$=k_n^{1+\varepsilon}$ *hence*

$$\begin{aligned}\alpha_n &= n^{-1/4},\\ k_n &= n^{\mu(1+\varepsilon)^{-2}/4}.\end{aligned}$$

Proof of Theorem 4 :
The scheme of the proof is very imilar to the previous Theorem. We first invoke Proposition 1 to get

$$\rho_n = O\left(\frac{(\log n)^{5/4}}{n^{1/8}}\right)$$

and again $\alpha_n^* = \min\left(n^{-1/4}, \rho_n\right)$
$=n^{-1/4}$. *At last, by(3.8) we get :*

$$\begin{aligned}\alpha_n &= n^{-1/4},\\ k_n &= \frac{\mu}{4\delta^2}\log n.\end{aligned}$$

Proof of Proposition 3 :
We refer to the proof of Proposition 2, especially to (3.5).
This time we keep a general m_n *since we hope to find an an accurate sequence balancing both terms in this equation.*

$$\begin{aligned}&P\left(|T_{pen}| > \sqrt{2}u_\kappa\right)\\ \geq\ & 1 - P\left(V_n \leq 2k_n + m_n\right) + P\left(|U_n - V_n| > m_n\right).\end{aligned}$$

It was shown (see Proposition 4 above) that the second term may be bounded.

$$P\left(|U_n - V_n| > m_n\right) = O\left(\frac{1}{\alpha_n^2 m_n \sqrt{n}}\right). \tag{3.9}$$

The first term may also be bounded copying the arguments of the above proofs :

$$\begin{aligned}P\left(V_n \leq 2k_n + m_n\right) &\leq P\left(\langle S_n, e_j\rangle^2 \leq (\lambda_j + \alpha_n)(2k_n + m_n)\right)\\ &\leq P\left(\frac{1}{n\lambda_j}\langle S_n, e_j\rangle^2 \leq (4k_n + 2m_n)/n\right)\end{aligned}$$

and the last inequality holds for a sufficiently large n : $(\lambda$
$_j + \alpha_n)/\lambda_j$ *was bounded by* 2.
Now we set $x_k = x_{k,j} =$
$\sqrt{\lambda_j}$ *and* $h_j = \langle m - m_0, e_j\rangle$.
Note that x_k *is a centered real random variable with unit variance.*

$$\begin{aligned}
& P\left(\frac{1}{n\lambda_j}\left\langle S_n, e_j\right\rangle^2 \leq \left(4k_n+2m_n\right)/n\right) \\
= \;& P\left(\left|\frac{1}{n}\sum_{k=1}^{n} x_k + \frac{h_j}{\sqrt{\lambda_j}}\right| \leq \left(4k_n+2m_n\right)/n\right) \\
\leq \;& P\left(\left|\left|\frac{1}{n}\sum_{k=1}^{n} x_k\right| - \left|\frac{h_j}{\sqrt{\lambda_j}}\right|\right| \leq \left(4k_n+2m_n\right)/n\right) \\
\leq \;& P\left(\left|\frac{1}{n}\sum_{k=1}^{n} x_k\right| \geq \frac{|h_j|}{\sqrt{\lambda_j}} - \left(4k_n+2m_n\right)/n\right)
\end{aligned}$$

Then applying Markov inequality at order 4 (remind that the x_k*'s are i.i.d. with finite fourth moment by* ***H1****) with* $m_n = n \mid h_j / \left(4\sqrt{\lambda_j}\right)$ *and for some constant c :*

$$P\left(V_n \leq 2k_n + m_n\right) \leq \frac{c}{n^2}$$

for a sufficiently large n*.*
Finally, comparing the last inequality with (3.9) the bound for the power is given by,

$$P\left(\left|T_{pen}\right| > \sqrt{2}u_\kappa\right) \geq 1 - \frac{c'}{\alpha_n n^{3/2}}$$

where c' *is another constant.*

Proof of Proposition 1 :
We first refer to Theorem 1 p.236 of Yurinskii (1977) which provides the Prokhorov distance between the sum of n independent random vectors in R^k *and an associated gaussian vector. Straightforward calculations lead to* $ck^{1/4}n^{-1/8}\left(\log n\right)^{1/2}$ *. The link with Hilbertian identically distributed random elements is made more explicit at the beginning of the second section p.244. These random variables are first projected on the k first eigenvectors of their common covariance operator. The Prokhorov distance between the initial (infinite dimensional) vectors and their projections is then computed. Chebyshev's inequality yields*

$$P\left(\left\|S_n - \Pi_k S_n\right\| > \gamma\right) \leq \frac{1}{\gamma^2}\sum_{i=k+1}^{+\infty} \lambda_i \leq \frac{c'}{\gamma^2}\exp\left(-\mu k\right).$$

By Theorem 1 in Dudley (1968), which implies that

$$\pi_H\left(X, Y\right) \leq \inf\left\{\varepsilon > 0 : P\left(\left\|X - Y\right\| > \varepsilon\right) \leq \varepsilon\right\},$$

we finally obtain

$$\pi_H\left(S_n, \Pi_k S_n\right) \leq c''\exp\left(-\mu k/3\right).$$

Combining both previous estimates provides an optimal k satisfying

$$\exp\left(-\mu k/3\right) = ck^{1/4}n^{-1/8}\left(\log n\right)^{1/2}.$$

Choosing $k = \dfrac{3}{8\mu} \log n$ *leads to* $\pi_H\left(S_n, G\right)$ $= O\left(n^{-1/8} \left(\log n\right)^{3/4}\right)$ *which is the desired result*

References

[1] Antoniadis, A. and Beder, J.M. (1989). Joint estimation of
the mean and the covariance structure of a Banach valued Gaussian vector. *Statistics* **20,** 77–93.

[2] Antoniadis, A., Gregoir, G. and Mc Keagie, I.W. (1994).
Wavelet methods for curve estimation. *J. Amer. Statist. Assoc.* **89,** 1340–1353.

[3] Araujo, A. and Gine, E. (1980)
for Real and Banach Valued Random Variables.*The Central Limit Theorem for Real and Banach Valued Random Variables. Wiley and Sons.*

[4] Arsenin, V. and Tikhonov, A. (1977).
*ill-posed problems*Solutions of
ill-posed problems. Winston and Sons, Washington D.C.

[5] Besse, P., Cardot, H. and Stephenson, D.B. (2000).
Autoregressive forecasting of some functional climatic variations.
Scand. Journal Stat.
Scand. Journal Stat. ***27,*** *673–687.*

[6] Bosq, D. (1991). Modelization, non-parametric estimation and prediction for continuous time processes. In Roussas, G., editor, Nonparametric Functional Estimation and Related Topics
Nonparametric Functional Estimation and Related Topics, NATO, ASI Series, 509–529.

[7] Bosq, D. (2000). *Linear processes in function spaces*
. Lecture Notes in Statistics. Springer-Verlag.

[8] Cardot, H., Ferraty, F. and Sarda, P. (1999). Functional
linear model. *Statist. & Prob. Letters* **45,** 11–22.

[9] Cardot, H., Ferraty, F., Mas ,A. and Sarda, P. (2002).
Testing hypotheses in the functional linear model. To appear in
Journal Stat.*Scand.*
Journal Stat.

[10] Cavallini, A., Montanari, G.C., Loggini, M., Lessi, O. and Cacciari, M. (1994). Nonparametric prediction of harmonic levels in electrical networks. Proc. IEEE ICHPS, *6, Bologna.*

[11] Dauxois, J., Pousse, A. and Romain, Y. (1982). Asymptotic theory for the principal component analysis of a random vector function: some applications to statistical inference. Journal Mult. Anal.
12, *136–154.*

[12] Dudley, R.M. (1968). Distances of Probability Measures and Random Variables. Ann. Math. Statist. ***39,*** *1563–1572.*

[13] Dunford, N. and Schwartz, J.T. (1962). Linear operators
, Vol. 2. Wiley, New-York.

[14] Franck, I.E., Friedman, J.H. (1993). A statistical view of some chemometrics regression tools. Technometrics ***35,*** *109–148.*

[15] Grenander, U. (1963).
*Structures*Probabilities on Algebraic
Structures. Almquist and Wiksell.

[16] Groetsch, C. (1993).
Mathematical Sciences*Inverse Problems in the Mathematical Sciences. Vieweg, Wiesbaden.*

[17] Lehmann, E.L. (1986). Testing Statistical Hypotheses
. John Wiley and Sons.

[18] Mas, A. (2002a). Rates of weak convergence for images of measures by families of mappings. Stat. Proba. Letters ***56,*** *7–12.*

[19] Mas, A. (2002b). Weak convergence for the covariance operators of a Hilbertian linear process. To appear in

*and Appl.*Stoch. Process
and Appl.

[20] Nashed, M.Z. and Wahba, G. (1974). Generalized inverses in reproducing kernel spaces: An approach to regularization of linear operator equations. *SIAM Journal on Math. Analysis* **6,** 974–987.

[21] Ramsay, J. O. and Silverman, B.W. (1997).
Data Analysis*Functional*
Data Analysis. Springer-Verlag, Berlin.

[22] Rice, J.A. and Silverman, B.W. (1991). Estimating the mean and covariance structure nonparametrically when the data are curves.
J. Roy. Stat. Soc
J. Roy. Stat. Soc. Ser. B, **53,** 233–243.

[23] Silverman, B.W. (1996). Smoothed principal compoenent Analysis by choice of norm. *Ann. Statist.* **24,** 1–24.

[24] Whitt, W. (1974). Preservation of rates of convergence under mappings. *Z. Wahr. verw. Gebiete* **29,** 39–44.

[25] Yurinskii, V.V. (1977). On the error of the Gaussian approximation for convolutions. *Theory Prob. Applications* **2,** 236–247.