PROBABILITY
AND
MATHEMATICAL STATISTICS

Vol. 9.1 (1988), p. 95-114

TIGHTNESS CRITERIA FOR RANDOM MEASURES WITH APPLICATION TO THE PRINCIPLE OF CONDITIONING IN HILBERT SPACES

BY
ADAM JAKUBOWSKI (TORUŃ)

Abstract. Suppose that $\{\mu_n\}$ is a sequence of random probability measures on a real and separable Hilbert space such that, for each $n \in N$, μ_n is a pointwisely convergent convolution of some sequence $\{\mu_{nk} | k \in N\}$ of random measures. The sequence $\{\mu_n\}$ is said to be *shift-tight* if one can find random vectors $\{A_n\}$ such that the "centered" sequence $\{\mu_n * \delta_{-A_n}\}$ is tight.

It is proved that for a shift-tight sequence $\{\mu_n\}$ there exists a "progressively measurable" centering which changes $\{\mu_n\}$ into a tight sequence.

As an application, Principle of Conditioning and Martingale Central Limit Theorem in a Hilbert space are proved.

1. INTRODUCTION

Principle of Conditioning, proved in [9], gives a method of derivation of a certain type of limit theorems for sums of dependent random variables. The main idea in the proof of the Principle of Conditioning is to find, for every sum of random variables, a random probability measure in such a way that convergence in probability of those random measures gives an information about the asymptotic behaviour of laws of respective sums (details are given in Section 4).

In fact this method is deeper: it is proved in [8] that tightness of the accompanying random measures implies tightness of sums of Hilbert space valued random vectors. This observation is developed in Section 4, and the considerations are based on some tightness criteria for random measures proved in Sections 2 and 3.

In Section 2 a general Theorem 2.1 on tightness of random finite measures defined on the Polish space is given. It is proved that a family of random measures $\{M_i\}$ is tight iff it is possible to approximate $\{M_i\}$ with an arbitrary exactness by a family $\{M_i'\}$ of random measures with finite expectations and such that the respective family $\{EM_i'\}$ of expectations is relatively compact.

Kallenberg [10] gives some methods of identification of weak limit for a sequence of random measures on a *locally compact space*, and his methods can be easily adapted to the Polish spaces. Hence Theorem 2.1 can be treated as a convenient complementary tool for Kallenberg's theory on the Polish spaces.

Section 3 is devoted to the shift-tightness of random measures on a Hilbert space. A sequence $\{\mu_n\}$ of random measures is said to be *shift-tight* if there exist random vectors A_n such that the "centered" sequence $\{\mu_n * \delta_{-A_n}\}$ is tight. The main result of this section, Theorem 3.1, deals with the sequences of the special form: for each n, $\mu_n = \prod^{*} \mu_{nk}$ $(k = 1, 2, \ldots)$, where the sequence $\{\mu_{nk}: k \in N\}$ of random measures is such that the infinite convolution $\prod^{*} \mu_{nk}(\omega)$, $k = 1, 2, \ldots$, converges pointwisely. Theorem 3.1 asserts that $\{\mu_n\}$ is shift-tight iff there exists a progressively measurable centering, i.e. one can find an array $\{A_{nk}: k \in N,\ n \in N\}$ of random vectors such that, for each n and k, A_{nk} is $\sigma(\mu_{n1}, \mu_{n2}, \ldots, \mu_{nk})$-measurable and the centered sequence $\mu_n * \delta_{-\sum_k A_{nk}}$ is tight.

The results of Sections 2 and 3 lead to a new tightness criterion for sums of random elements in a Hilbert space (Theorem 4.1) and allow us to extend the Principle of Conditioning to the infinite-dimensional case (Theorem 4.2).

In Section 5, Martingale Central Limit Theorem in a Hilbert space (Theorem 5.1) is obtained as a corollary from the Principle of Conditioning.

In the sequel H will denote a real and separable Hilbert space with the inner product $\langle \cdot, \cdot \rangle$ and the σ-algebra of Borel subsets $\mathscr{B}_H$.

2. GENERAL CRITERION

Let $\mathscr{H}$ be a complete separable metric space. Denote the set of finite Borel measures on $\mathscr{B}_{\mathscr{H}}$ by $\mathscr{M}(\mathscr{H})$. Weak convergence in $\mathscr{M}(\mathscr{H})$ can be metrised (e.g. by Prokhorov metric), so we will treat $\mathscr{M}(\mathscr{H})$ as a complete separable metric space.

We say that a measurable mapping M: $(\Omega, \mathscr{F}, P) \to \mathscr{M}(\mathscr{H})$ (i.e. a random finite measure M) is *integrable* iff $E(M(\mathscr{H})) < +\infty$. In such a case an element EM of $\mathscr{M}(\mathscr{H})$ is defined by $(EM)(A) = E(M(A))$.

A family $\{M_i: (\Omega, \mathscr{F}, P) \to \mathscr{M}(\mathscr{H}): i \in I\}$ of random measures is *tight* iff for each $\delta > 0$ one can find measurable subsets $\{A_{i,\delta}: i \in I\}$ such that

(2.1) $$P(A_{i,\delta}) > 1 - \delta, \quad i \in I;$$

(2.2) the set $\bigcup_{i \in I} \{M_i(\cdot, \omega): \omega \in A_{i,\delta}\} \subset \mathscr{M}(\mathscr{H})$ is relatively compact in $\mathscr{M}(\mathscr{H})$.

By Prokhorov theorem, condition (2.2) is equivalent to the following conditions:

(2.3) there exists a constant $\gamma > 0$ such that

$$\sup_{i \in I} \sup_{\omega \in A_{i,\delta}} M_i(\mathscr{H}, \omega) \leqslant \gamma;$$

(2.4) there exists an increasing sequence $\{K_{n,\delta}: n \in N\}$ of compact subsets of $\mathscr{H}$ such that, for each $n \in N$,

$$\sup_{i \in I} \sup_{\omega \in A_{i,\delta}} M_i(K_{n,\delta}^c, \omega) \leqslant n^{-1}.$$

By means of Prokhorov theorem we will prove another, more useful criterion for tightness of random measures.

2.1. THEOREM. *Let $\{M_i: i \in I\}$ be a family of random measures on $\mathscr{H}$. Then $\{M_i: i \in I\}$ is tight iff for each $\delta > 0$ there exists a family $\{M_i^\delta: i \in I\}$ of integrable random measures such that*

(2.5) $$\sup_{i \in I} P(M_i^\delta \neq M_i) < \delta,$$

(2.6) *the set $\{EM_i^\delta\} \subset \mathscr{M}(\mathscr{H})$ is relatively compact.*

In particular, if the family $\{M_i: i \in I\}$ is tight and the random variables $\{M_i(\mathscr{H}, \cdot): i \in I\}$ are uniformly integrable, then $\{EM_i: i \in I\} \subset \mathscr{M}(\mathscr{H})$ is relatively compact.

Proof. If $\{M_i: i \in I\}$ is tight and the sets $\{A_{i,\delta}: i \in I\}$ are taken from (2.1) and (2.2), then defining

$$M_i^\delta(\cdot, \omega) = I_{A_{i,\delta}}(\cdot, \omega) M_i(\cdot, \omega),$$

we get the desired family.

Conversely, suppose that the family $\{M_i^\delta: i \in I\}$ has properties (2.5) and (2.6). Let $\gamma = \delta^{-1} \sup_{i \in I} EM_i^\delta(\mathscr{H})$. Choose an increasing family of compacts K_n such that

$$\sup_{i \in I} EM_i^\delta(K_n^c) \leqslant \delta n^{-1} \cdot 2^{-n}, \qquad n \in N,$$

and define

$$A_{i,3\delta} = \{M_i^\delta = M_i\} \cap \{M_i^\delta(\mathscr{H}) \leqslant \gamma\} \cap \bigcap_{n=1}^{\infty} \{M_i^\delta(K_n^c) \leqslant n^{-1}\}.$$

Thus, $P(A_{i,3\delta}^c) < 3\delta$ and for $\{A_{i,3\delta}\}$ conditions (2.3) and (2.4) are satisfied.

Suppose that $\{M_i: i \in I\}$ is tight and $\{M_i(\mathscr{H}): i \in I\}$ is uniformly integrable. Let $\varepsilon > 0$. Choose $\delta > 0$ in such a way that $P(A) \leqslant \delta$ implies $\sup_{i \in I} EM_i(\mathscr{H}) I_A \leqslant \varepsilon/2$.

For $\{M_i: i \in I\}$ take measures $\{M_i^\delta: i \in I\}$ satisfying (2.5) and (2.6). Let a compact K be such that $\sup_{i\in I} EM_i^\delta(K^c) \leqslant \varepsilon/2$. Then

$$EM_i(K^c) = EM_i(K^c)I(M_i = M_i^\delta)+EM_i(K^c)I(M_i \neq M_i^\delta)$$
$$\leqslant EM_i^\delta(K^c)+EM_i(\mathscr{H})I(M_i \neq M_i^\delta) \leqslant \varepsilon/2+\varepsilon/2 = \varepsilon.$$

Since also $\sup_{i\in I} EM_i(\mathscr{H}) < +\infty$, the relative compactness of $\{EM_i\colon i\in I\}$ is proved.

In the sequel $\mathscr{P}(\mathscr{H}) \subset \mathscr{M}(\mathscr{H})$ will denote the space of probability measures on $\mathscr{H}$.

2.2. Corollary. *A family $\{\mu_i\colon i\in I\}$ of random probability measures on $\mathscr{H}$ is tight iff $\{E\mu_i\colon i\in I\} \subset \mathscr{P}(\mathscr{H})$ is relatively compact.*

Recall that $(H, \langle\cdot,\cdot\rangle)$ denotes the Hilbert space.

If $\mu(\cdot,\cdot)\colon B_H\times\Omega \to [0, 1]$ is a random probability measure, then

$$\hat{\mu}(y, \omega) = \int \exp i\langle y, x\rangle \mu(dx, \omega)$$

is its characteristic function. Applying Corollary 2.2, we shall describe the convergence in probability of random probability measures on finite-dimensional Hilbert space in terms of the convergence of their characteristic functions.

2.3. Proposition. *Suppose that* $\dim H < +\infty$. *Then a sequence $\{\mu_n\}$ of random probability measures on H converges to some random probability measure $\mu(\cdot,\omega)$: $\mu_n \underset{\mathrm{P}}{\Rightarrow} \mu$ iff*

$$\hat{\mu}_n(y,\cdot) \underset{\mathrm{P}}{\rightarrow} \hat{\mu}(y,\cdot), \qquad y\in H', \tag{2.7}$$

where $H' \subset H$ is a dense subset such that $\{x\colon \|x\| < \varepsilon\} \subset H'$ for some $\varepsilon > 0$.

2.4. Lemma. *Let $(\mathscr{H}, \varrho)$ and $(\mathscr{H}', \varrho')$ be complete separable metric spaces and let $\{f_i\colon \mathscr{H} \to \mathscr{H}', i\in I\}$ be a countable family of continuous mappings which separates the points of $\mathscr{H}$. Let $\{X_n\colon (\Omega, \mathscr{F}, P) \to (\mathscr{H}, \mathscr{B}_{\mathscr{H}})\colon n\in N\}$ be a sequence of measurable mappings such that.*

(a) *$\{X_n\colon n\in N\}$ is tight,*

(b) *for each $i\in I$ there exists a measurable mapping $Y_i\colon (\Omega, \mathscr{F}, P) \to (\mathscr{H}', \mathscr{B}_{\mathscr{H}'})$ such that $f_i(X_n) \underset{\mathrm{P}}{\rightarrow} Y_i$.*

Then there exists $X\colon (\Omega, \mathscr{F}, P) \to (\mathscr{H}, \mathscr{B}_{\mathscr{H}})$ such that $X_n \underset{\mathrm{P}}{\rightarrow} X$ and, in particular, $Y_i = f_i(X)$ a.s.

Proof. It suffices to prove that X_n is a fundamental in probability sequence. Suppose the contrary case. Thus one can find an $\varepsilon > 0$ and a subsequence $\{n_k\} \subset N$ such that, for each $k\in N$,

$$P\big(\varrho(X_{n_{2k-1}}, X_{n_{2k}}) > \varepsilon\big) > \varepsilon. \tag{2.8}$$

By passing to a further subsequence, we may assume that $f_i\big(X_{n_k}(\omega)\big)$

$\to Y_i(\omega)$ for each $i \in I$ and each ω from some subset $\Omega_1 \in F$, $P(\Omega_1) = 1$, while (2.8) is still valid. Since $\{X_n\}$ is tight, there exists a compact subset $K \subset \mathscr{H}$ such that

$$\inf_n P(X_n \in K) > 1 - \varepsilon/3.$$

Let $A_k = \Omega_1 \cap \{X_{n_{2k-1}} \in K\} \cap \{X_{n_{2k}} \in K\} \cap \{\varrho(X_{n_{2k-1}}, X_{n_{2k}}) > \varepsilon\}$. Then $P(A_k) > \varepsilon/3$ and, in particular,

$$P(\operatorname{Lim\,sup}_k A_k) = P(A_k \text{ i.o.}) \geqslant \varepsilon/3.$$

Hence $\operatorname{Lim\,sup}_k A_k \neq \emptyset$. If $\omega \in \operatorname{Lim\,sup}_k A_k$, then $\omega \in A_{k'}$ for some increasing sequence $\{k'\} = \{k'(\omega)\}$. But K is compact, hence for $\{k''\} \subset \{k'\}$ we have

$$X_{n_{2k''-1}}(\omega) \to x, \ X_{n_{2k''}}(\omega) \to y, \qquad \text{where } x, y \in \mathscr{H}.$$

By the continuity of f_i, $f_i(x) = f_i(y) = Y_i(\omega)$ for all $i \in I$. Hence $x = y$ and $\varrho(X_{n_{2k''-1}}(\omega), X_{n_{2k''}}(\omega)) \to 0$. This leads to a contradiction since $\omega \in A_{k''}$.

Proof of Proposition 2.3. Clearly, $\mu_n(\cdot, \omega) \Rightarrow \mu(\cdot, \omega)$ implies (2.7) for each $y \in H$. Conversely, let (2.7) be satisfied. Then, for y from some neighbourhood of zero,

$$(E\hat{\mu}_n)(y) = E\hat{\mu}_n(y, \omega) \to E\hat{\mu}(y, \omega) = (E\hat{\mu})(y). \tag{2.9}$$

But (2.9) implies relative compactness of $\{E\mu_n\colon n \in N\} \subset \mathscr{P}(H)$ and, by Corollary 2.2, also tightness of $\{\mu_n\}$. Now it is sufficient to apply Lemma 2.4 with the countable family of separating continuous functions of the form $\mathscr{P}(H) \ni \mu \mapsto f_y(\mu) = \hat{\mu}(y) \in \mathbf{C}$, where y varies over some countable dense subset $H'' \subset H'$.

2.5. Remark. A linear operator $S\colon H \to H$ is called *S-operator* ($S \in \mathscr{S}$) iff S is nonnegative, hermitian and with finite trace,

$$\operatorname{tr} S = \sum_{j=1}^{\infty} \langle Se_j, e_j \rangle < +\infty,$$

where $\{e_j\}$ is an orthonormal basis for H. By the mapping $I^{1/2}\colon \mathscr{S} \to \mathscr{S}^{1/2}$, $I^{1/2}(S) = S^{1/2}$, we may embed $\mathscr{S}$ into the space of Hilbert–Schmidt operators on H. With induced metric, $\mathscr{S}$ becomes a complete and separable metric space and, for a sequence $\{S_n\} \subset \mathscr{S}$, $S_n \to S$ in $\mathscr{S}$ iff

$$\operatorname{tr} S_n \to \operatorname{tr} S \ \text{ and } \ \langle S_n y, y \rangle \to \langle Sy, y \rangle, \qquad y \in H. \tag{2.10}$$

Moreover, a subset $\{S_i\colon i \in I\} \subset \mathscr{S}$ is relatively compact iff

$$\sup_{i \in I} \operatorname{tr} S_i < +\infty, \tag{2.11}$$

$$\limsup_{N\to\infty} \sum_{i\in I} \sum_{k=N}^{\infty} \langle S_i e_k, e_k\rangle = 0. \tag{2.12}$$

Define the mapping $\mathscr{S} \ni S \mapsto M^S \in \mathscr{M}(N)$ by the formula $M^S(\{k\}) = \langle Se_k, e_k\rangle$. It is clear that $\{S_i\colon i\in I\} \subset \mathscr{S}$ is relatively compact iff $\{M^{S_i}\colon i\in I\} \subset \mathscr{M}(N)$ is relatively compact. So to examine the tightness of random S-operators one can use Theorem 2.1.

3. SHIFT-TIGHTNESS OF CONVOLUTIONS OF RANDOM MEASURES

The shift-tightness of random probability measures is the problem we will consider next. We shall outline it briefly.

A sequence $\{\mu_n\colon n\in N\} \subset \mathscr{P}(\mathscr{H})$ is *shift compact* iff there exists a centering sequence $\{x_n\colon n\in N\} \subset H$ such that $\{\mu_n * \delta_{-x_n}\colon n\in N\}$ is relatively compact. It is well-known (see e.g. [12]) that the shift-compactness of the sequence μ_n is equivalent to the relative compactness of the sequence of symmetrisations $\{|\mu_n|^2\colon n\in N\}$.

In the particular case $H = \boldsymbol{R}^1$, Doob ([5], p. 408 - 409) gave the explicite formula for the centering constants as numbers x_n satisfying the equation

$$\int \arctan(x-x_n)\,\mu_n(dx) = 0. \tag{3.1}$$

It follows that in the one-dimensional case, if the centering exists, then it can be obtained in the form $x_n = f(\mu_n)$, where $f\colon \mathscr{P}(\boldsymbol{R}^1) \to \boldsymbol{R}^1$ is a measurable function.

If the sequence $\{\mu_n\}$ is formed by convolutions in the rows of a certain array $\{\mu_{nk}\}$ of measures, $\mu_n = \mu_{n1} * \mu_{n2} * \ldots * \mu_{nk_n}$, and if, for example, the array $\{\mu_{nk}\}$ is uniformly infinitesimal,

$$\max_{1\leqslant k\leqslant k_n} \mu_{nk}(\|x\| > \varepsilon) \to 0, \qquad \varepsilon > 0,$$

then the shift compactness of $\{\mu_n\}$ implies the existence of centering constants in the particular form

$$x_n = \sum_{1\leqslant k\leqslant k_n} x_{nk}, \tag{3.2}$$

where $x_{nk} = f(\mu_{nk})$ and f is a measurable function on $\mathscr{P}(H)$.

The author does not know whether in the case of infinite-dimensional H there exists a universal function f such that, for any array $\{\mu_{nk}\}$ with the shift compact convolutions in rows, the centering constants may be chosen in the form (3.2). Hence, to handle the respective problem for arrays of random measures, the technic of measurable sections seems to be adequate.

Recall that the sequence $\{\mu_n\}$ of random probability measures on H is *shift-tight* if one can find the random vectors A_n such that the centered sequence $\{\mu_n * \delta_{-A_n}\}$ is tight.

In the sequel the notation of [1] is used.

3.1. THEOREM. *Let $\{\mu_{nk}(\cdot, \omega): k \in N, n \in N\}$ be a double array of random probability measures on H. For $N \in N$ define*

$$\mu_n(N)(\omega) = \prod_{1 \leqslant k \leqslant N}^{*} \mu_{nk}(\omega).$$

The following statements are equivalent:

(i) *The set of symmetrisations $\{|\mu_n(N)|^2: N \in N, n \in N\}$ is a tight family of random measures in $\mathscr{P}(H)$.*

(ii) *There exist $\sigma(\mu_{n1}, \ldots, \mu_{nk})$-measurable H-valued random variables B_{nk} such that the random S-operators $\{T^n: n \in N\}$, defined by the quadratic forms*

$$\langle T^n y, y \rangle = \sum_{k=1}^{\infty} \int \langle y, h(x - B_{nk}(\omega)) \rangle^2 \mu_{nk}(dx, \omega),$$

are tight in $\mathscr{S}$, and the random finite measures $\{M^n: n \in N\}$ such that

$$M^n(A) = \sum_{k=1}^{\infty} [\mu_{nk}(\cdot, \omega) * \delta_{-B_{nk}(\omega)}](A \cap \{\|x\| > 1\})$$

are tight in $\mathscr{M}(H)$.

(iii) *There are $\sigma(\mu_{n1}, \ldots, \mu_{nk})$-measurable random variables C_{nk} such that the set of random probability measures*

$$\{\mu_n(N) * \delta_{-\sum_{1 \leqslant k \leqslant N} C_{nk}}: N \in N, n \in N\}$$

is tight in $\mathscr{P}(H)$.

Moreover, if B_{nk} are taken from (ii), *then the centering random variables C_{nk} can be defined by*

$$C_{nk} = B_{nk}(\omega) + \int_{\|x\| \leqslant 1} x [\mu_{nk}(\cdot, \omega) * \delta_{-B_{nk}(\omega)}](dx). \tag{3.3}$$

The function $h: H \to H$ appearing in (ii) *is any (but fixed for ever) bounded continuous function such that*

$$h(x) = \begin{cases} x & \text{if } \|x\| \leqslant 1, \\ 0 & \text{if } \|x\| \geqslant 2. \end{cases}$$

Proof. Implications (ii) $\Rightarrow$ (iii) $\Rightarrow$ (i) follow from the respective properties of convolution of non-random measures. Indeed, (iii) $\Rightarrow$ (i) stems from the continuity of the operation $\mu \mapsto |\mu|^2$. Now, suppose that (ii) holds. By (2.1) and (2.2) we may choose subsets $A_{n,\delta}$ such that $P(A_{n,\delta}) > 1 - \delta$, $n \in N$, and the sets

$$\bigcup_{n \in N} \{T^n(\omega): \omega \in A_{n,\delta}\} \subset \mathscr{S}, \qquad \bigcup_{n \in N} \{M^n(\omega): \omega \in A_{n,\delta}\} \subset \mathscr{M}(H)$$

are relatively compact. By Theorem 6.1, Chapter VI, [12], the set of probability measures

$$\bigcup_{n\in N} \{\mathrm{Poiss}(\sum_{1\leqslant k\leqslant N} \mu_{nk}(\cdot, \omega) * \delta_{-B_{nk}(\omega)}): N\in N, \omega\in A_{n,\delta}\}$$

is relatively shift compact. Hence, by Theorem 4.8, Chapter 3, [1], the set

$$\bigcup_{n\in N} \{ \prod^{*}_{1\leqslant k\leqslant N} \mu_{nk}(\cdot, \omega) * \delta_{-C_{nk}(\omega)}: N\in N, \omega\in A_{n,\delta}\}$$

is relatively compact, where C_{nk} are given by (3.3).

In a similar way, condition (i) implies that the sequence $\{(T')^n\}$ of random S-operators, given by the quadratic forms

$$\langle (T')^n y, y\rangle = \sum_{k=1}^{\infty} \int \langle y, h(x)\rangle^2 |\mu_{nk}|^2 (dx),$$

is tight, and the sequence $\{(M')^n\}$ of finite random measures on H,

$$(M')^n(A) = \sum_{k=1}^{\infty} |\mu_{nk}|^2 (A\cap\{\|x\|>1\}), \quad A\in\mathscr{B}_H,$$

is also tight.

Fix an orthogonal basis $\{e_i\}$ in H and, for $\mu\in\mathscr{P}(H)$, define the finite measure $\tilde{M}(\mu)$ on $\mathscr{H} = \{1\}\times N\cup\{2\}\times H$ by the formula

(3.4)

$$\tilde{M}(\mu)(\{1\}\times A_1\cup\{2\}\times A_2) = \sum_{i\in A_1} \int \langle e_i, h(x)\rangle^2 \mu(dx) + \mu(A_2\cap\{\|x\|>1\}),$$

where $A_1\subset N$ and $A_2\in\mathscr{B}_H$.

Observe that, for every $\tilde{A}\in\mathscr{B}_{\tilde{\mathscr{H}}}$,

$$\int \tilde{M}(\mu*\delta_{-x})(\tilde{A})\,\mu(dx) = \tilde{M}(|\mu|^2)(\tilde{A}). \tag{3.5}$$

Due to Remark 2.5 it is clear that the tightness of $\{(T')^n\}$ and $\{(M')^n\}$, i.e. condition (i), is equivalent to the tightness of random measures

$$\{\sum_{k=1}^{\infty} \tilde{M}(|\mu_{nk}|^2): n\in N\}$$

and that for (ii) to hold we have to guarantee the existence of random variables B_{nk}, $\sigma(B_{nk})\subset\sigma(\mu_{n1}, \ldots, \mu_{nk})$ such that the random measures

$$\{\sum_{k=1}^{\infty} \tilde{M}(\mu_{nk}*\delta_{-B_{nk}}): n\in N\}$$

are tight.

3.2. Lemma. *Let $b>0$, $0<c<2^{-1}(\sum_{r=1}^{\infty} r^{-2})^{-1}$ and let $\{K_r: r\in N\}$ be an* *increasing sequence of compact subsets of* $\mathscr{H}$. *There exists a Borel-measurable mapping* $B\colon \mathscr{K} \to H$ *such that* Graph $B = \{(\mu, B\mu)\colon \mu \in \mathscr{K}\} \subset \mathscr{K}_H$ *and* $B(\delta_x) = x$, *where*

$$\mathscr{K} = \mathscr{K}(\{K_r\}, b) = \{\mu \in \mathscr{P}(H)\colon \tilde{M}(|\mu|^2(\mathscr{H}) < b,\ \tilde{M}(|\mu|^2)(K_r^c) \leqslant r^{-3},\ r \geqslant 1\},$$

$$\begin{aligned}\mathscr{K}_H &= \mathscr{K}_H(\{K_r\}, b)\\ &= \{(\mu, x) \in \mathscr{K} \times H\colon \tilde{M}(\mu * \delta_{-x})(\mathscr{H}) \leqslant 2\tilde{M}(|\mu|^2)(\mathscr{H}),\\ &\qquad \tilde{M}(\mu * \delta_{-x})(K_r^c) \leqslant (r^2/c)\,\tilde{M}(|\mu|^2)(K_r^c),\ r \geqslant 1\}.\end{aligned}$$

Proof. Clearly, $\mathscr{K}$ and $\mathscr{K}_H$ are Borel subests of $\mathscr{P}(H)$ and $\mathscr{P}(H) \times H$, respectively. Suppose we know that for each $\mu \in \mathscr{K}$ the section $(\mathscr{K}_H)_\mu = \{x \in H\colon (\mu, x) \in \mathscr{K}_H\} \subset H$ is non-empty and compact. Then, applying Theorem 21b, Chap. III, [4], we can find a Borel subset $U \subset \mathscr{K}_H$ such that the projection $\pi_{\mathscr{P}(H)}$ of $\mathscr{P}(H) \times H$ onto $\mathscr{P}(H)$ is one-to-one when restricted to U and $\pi_{\mathscr{P}(H)} U = \mathscr{K}$. By the Kuratowski theorem, $(\pi_{\mathscr{P}(H)}|_U)^{-1}\colon \mathscr{K} \to U \subset \mathscr{K}_H$ is measurable, hence $B = \pi_H \circ (\pi_{\mathscr{P}(H)}|_U)^{-1}\colon \mathscr{K} \to H$ has the property Graph $B = U$.

To see that the section $(\mathscr{K}_H)_\mu$ is non-empty for $\mu \in \mathscr{K}$, suppose that, for some $r \geqslant 1$, $M(|\mu|^2)(K_r^c) > 0$. Then, by (3.5),

$$\mu(\{x\colon \tilde{M}(\mu * \delta_{-x})(K_r^c) > (r^2/c)\,\tilde{M}(|\mu|^2)(K_r^c)\}) \leqslant c/r^2.$$

If $0 = \tilde{M}(|\mu|^2)(K_r^c) = \int \tilde{M}(\mu * \delta_{-x})(K_r^c)\,\mu(dx)$, the integral being taken over the set $\{x\colon \tilde{M}(\mu * \delta_{-x})(K_r^c) > 0\}$, then also

$$\begin{aligned}&(\{x\colon \tilde{M}(\mu * \delta_{-x})(K_r^c) > (r^2/c)\,\tilde{M}(|\mu|^2)(K_r^c)\})\\ &\qquad = \mu(\{x\colon \tilde{M}(\mu * \delta_{-x})(K_r^c) > 0\}) = 0 < c/r^2.\end{aligned}$$

Similarly, we prove that

$$\mu\{x\colon \tilde{M}(\mu * \delta_{-x})(\mathscr{H}) > 2\tilde{M}(|\mu|^2)(\mathscr{H})\} < 1/2.$$

So

$$\mu((\mathscr{K}_H)_\mu^c) < \frac{1}{2} + \sum_{r=1}^{\infty} \frac{c}{r^2} < 1 \quad \text{and} \quad (\mathscr{K}_H)_\mu \neq \emptyset.$$

Now suppose that $\mu \in \mathscr{K}$, $(\mu, x_n) \in \mathscr{K}_H$ for $n \in N$ and $x_n \to x$. Then $\mu * \delta_{-x_n} \Rightarrow \mu * \delta_{-x}$ and, for each $r \geqslant 1$,

$$\tilde{M}(\mu * \delta_{-x})(K_r^c) \leqslant \liminf_n \tilde{M}(\mu * \delta_{-x_n})(K_r^c) \leqslant \frac{r^2}{c}\,\tilde{M}(|\mu|^2)(K_r^c)$$

and

$$\tilde{M}(\mu * \delta_{-x})(\mathscr{H}) \leqslant \liminf_n \tilde{M}(\mu * \delta_{-x_n})(\mathscr{H}) \leqslant 2\tilde{M}(|\mu|^2)(\mathscr{H}).$$

Hence $x \in (\mathscr{K}_H)_\mu$, i.e. $(\mathscr{K}_H)_\mu$ is closed. It is also relatively compact. Indeed, it suffices to prove that $\{\tilde{M}(\mu * \delta_{-x})\colon x \in (\mathscr{K}_H)\,\mu\} \subset \mathscr{M}(\mathscr{H})$ is relatively compact. If the supremum below is taken over $x \in (\mathscr{K}_H)_\mu$, then

$$\sup \tilde{M}(\mu * \delta_{-x})(\mathscr{H}) \leqslant 2\tilde{M}(|\mu|^2)(\mathscr{H}) \leqslant 2b,$$

$$\sup \tilde{M}(\mu * \delta_{-x})(\mathscr{H}) \leqslant 2\tilde{M}(|\mu|^2)(K_r^c) \leqslant 2b,$$

3.3 Lemma. *Let $\{\mu_1(\cdot, \omega), \mu_2(\cdot, \omega), \ldots\}$ be a sequence of random probability measures on H, defined on $(\Omega, \mathscr{F}, P)$. Let $A_1 \subset A_2 \subset \ldots$ be an increasing sequence of elements of $\mathscr{F}$. Suppose that for each $q \in N$ there are given a number $b_q > 0$ and a sequence $\{K_{q,r}: r \in N\}$ of compact subsets of $\mathscr{H} = \{1\} \times N \cup \{2\} \times H$ such that*

$$b_{q+1} \geqslant b_q, \quad q \in N, \tag{3.6}$$

$$K_{q,r} \subset K_{q,r+1} \subset K_{q+1,r} (*) \tag{3.7}$$

and, for each $\omega \in A_q$, the inequalities

$$\sum_{k=1}^{\infty} \tilde{M}(|\mu_k(\cdot, \omega)|^2)(\mathscr{H}) \leqslant b_q, \tag{3.8}$$

$$\sum_{k=1}^{\infty} \tilde{M}(|\mu_k(\cdot, \omega)|^2)(K_{q,r}^c) \leqslant r^{-3} \tag{3.9}$$

hold, where $\tilde{M}$ is defined by (3.4).

Then there exist random variables B_k such that

$$\sigma(B_k) \subset \sigma(\mu_1, \mu_2, \ldots, \mu_k), \quad k \in N, \tag{3.10}$$

and, for $\omega \in A_q$, random measures $\{\mu_k(\cdot, \omega) * \delta_{-B_k(\omega)}\}$ satisfy the conditions

$$\sum_{k=1}^{\infty} \tilde{M}(\mu_k(\cdot, \omega) * \delta_{-B_k(\omega)})(\mathscr{H}) \leqslant 2b_q, \tag{3.11}$$

$$\sum_{k=1}^{\infty} \tilde{M}(\mu_k(\cdot, \omega) * \delta_{-B_k(\omega)})(K_{q,r}^c) \leqslant q/cr. \tag{3.12}$$

Proof. Set $A_{k,0} = \emptyset$, $k \in N$, and, for $q \in N$,

$$A_{k,q} = \{\omega: \sum_{j=1}^{k} \tilde{M}(|\mu_j(\omega)|^2)(\tilde{\mathscr{H}}) \leqslant b_q, \quad \sum_{j=1}^{k} \tilde{M}(|\mu_j(\omega)|^2 (K_{q,r}^c) \leqslant r^{-3}, r \geqslant 1\}$$

Then, by (3.6) and (3.7), $A_{k,q+1} \supset A_{k,q} \supset A_{k+1,q}$ and, by (3.8) and (3.9), $A_q \subset \bigcap_{k=1}^{\infty} A_{k,q}$.

Consider the random measures

$$\mu_{k,q}(\cdot, \omega) = \mu_k(\cdot, \omega) I_{A_{k,q}}(\omega) + \delta_0 I_{A_{k,q}^c}(\omega).$$

(*) Hence, for $q \geqslant p$, $K_{p,r+q-p} \subset K_{q,r}$.

$\{\mu_{k,q}(\cdot,\omega): k\in N\}$ take values in the set $\mathscr{K}(\{K_{q,r}: r\in N\}, b_q)$ defined in Lemma 3.2. Hence there exists a measurable mapping B^q: $\mathscr{K}\to H$ such that Graph $B^q \subset \mathscr{K}_H(\{K_{q,r}\}, b_q)$ (recall that $B^q(\delta_0)=0$), i.e. random measures

$$\nu_{k,q}(\cdot,\omega) = [\mu_k(\cdot,\omega) * \delta_{-B^q(\mu_k(\cdot,\omega))}]\, I_{A_{k,q}}(\omega) + \delta_0 I_{A^c_{k,q}}(\omega)$$

satisfy the conditions

$$\tilde{M}\big(\nu_{k,q}(\cdot,\omega)\big)(\mathscr{H}) \leqslant 2\tilde{M}\big(|\mu_{k,q}(\cdot,\omega)|^2\big)(\mathscr{H}), \tag{3.13}$$

$$M\big(\nu_{k,q}(\cdot,\omega)\big)(K^c_{q,r}) \leqslant \frac{r^2}{c}\tilde{M}\big(|\mu_{k,q}(\cdot,\omega)|^2\big)(K^c_{q,r}), \qquad r\geqslant 1. \tag{3.14}$$

Let

$$B_k = \sum_{p=1}^{\infty} B^p\big(\mu_k(\cdot,\omega)\big) \mathrm{I}_{(A_{k,p}-A_{k,p-1})}(\omega).$$

Observe that $\sigma(B_k)\subset\sigma(\mu_j: 1\leqslant j\leqslant k)$ since $A_{k,p}\in\sigma(\mu_j: 1\leqslant j\leqslant k)$ for $p\in N_0$. Measures $\{\mu_k(\cdot,\omega)*\delta_{-B_k(\omega)}\}$ satisfy conditions (3.11) and (3.12). Indeed,

$$I_{A_q}\sum_{k=1}^{\infty}\tilde{M}\big(\mu_k(\cdot,\omega)*\delta_{-B_k(\omega)}\big)(\mathscr{H})$$

$$= I_{A_q}\sum_{k=1}^{\infty}\sum_{p=1}^{\infty} I_{(A_{k,p}-A_{k,p-1})}\tilde{M}(\nu_{k,p})(\mathscr{H})$$

$$= \sum_{k=1}^{\infty}\sum_{p=1}^{\infty} I_{A_q} I_{(A_{k,p}-A_{k,p-1})}\tilde{M}(\nu_{k,p})(\mathscr{H})$$

$$= I_{A_q}\sum_{k=1}^{\infty}\sum_{p=1}^{q} I_{(A_{k,p}-A_{k,p-1})}\tilde{M}(\nu_{k,p})(\mathscr{H})$$

(since $A_q\subset A_{k,p-1}$ for $p>q$)

$$\leqslant I_{A_q}\sum_{k=1}^{\infty}\sum_{p=1}^{q} I_{(A_{k,p}-A_{k,p-1})}\cdot 2\tilde{M}(|\mu_{k,p}|^2)(\mathscr{H})$$

(by (3.13))

$$= I_{A_q}\sum_{k=1}^{\infty}\sum_{p=1}^{q} I_{(A_{k,p}-A_{k,p-1})}\cdot 2\tilde{M}(|\mu_k|^2)(\mathscr{H})$$

(since $\mu_{k,p}=\mu_k$ on $A_{k,p}$)

$$= 2I_{A_q}\sum_{k=1}^{\infty}\tilde{M}(|\mu_k|^2)(\mathscr{H})\, I_{A_{k,q}} \leqslant 2I_{A_q} b_q$$

(by the definition of $A_{k,q}$).

Similarly,

$$I_{A_q} \sum_{k=1}^{\infty} \tilde{M}(\mu_k(\omega) * \delta_{-B_k(\omega)})(K_{q,r}^c)$$

$$= I_{A_q} \sum_{k=1}^{\infty} \sum_{p=1}^{q} I_{(A_{k,p}-A_{k,p-1})} \tilde{M}(\nu_{k,p})(K_{q,r}^c)$$

$$\leqslant I_{A_q} \sum_{k=1}^{\infty} \sum_{p=1}^{q} I_{A_{k,p}} \tilde{M}(\nu_{k,p})(K_{q,r}^c)$$

$$\leqslant I_{A_q} \sum_{k=1}^{\infty} \sum_{p=1}^{q} I_{A_{k,p}} \tilde{M}(\nu_{k,p})(K_{p,r+q-p}^c)$$

(since $K_{p,r+q-p} \subset K_{q,r}$ by (3.7))

$$\leqslant I_{A_q} \sum_{p=1}^{q} \sum_{k=1}^{\infty} I_{A_{k,p}} ((r+q-p)^2/c) \tilde{M}(|\mu_k|^2)(K_{p,r+q-p}^c)$$

(by (3.14))

$$= I_{A_q} \sum_{p=1}^{q} ((r+q-p)^2/c) \sum_{k=1}^{\infty} \tilde{M}(|\mu_k|^2)(K_{p,r+q-p}^c) I_{A_{k,p}}$$

$$\leqslant I_{A_q} \sum_{p=1}^{q} (r+q-p)^2 (r+q-p)^{-3}/c$$

(by the definition of $A_{k,p}$)

$$\leqslant I_{A_q} q/rc.$$

Now the proof of Theorem 3.1 can be easily completed. By the tightness of

$$\{\sum_{k=1}^{\infty} \tilde{M}(|\mu_{nk}|^2): n \in N\}$$

for each $q > 0$ we can find sets A_q^n such that $A_q^n \subset A_{q+1}^n$, $P(A_q^n) > 1 - q^{-1}$ and

$$\mathcal{M}_q = \bigcup_{n=1}^{\infty} \{\sum_{k=1}^{\infty} \tilde{M}(|\mu_{nk}(\omega)|^2): \omega \in A_q\}$$

is relatively compact. Then for each $\mathcal{M}_q$ we can find $b_q > 0$ and a sequence of compact sets $\{K_{q,r}: r \in N\}$ satisfying (3.6) and (3.7) and such that, for each n, (3.8) and (3.9) are fulfilled on A_q^n. Finally, we can apply Lemma 3.3 to each sequence $\{\mu_{nk}: k \in N\}$ separately.

3.4. COROLLARY. *If we consider the arrays $\{\mu_{nk}\}$ of non-random probability measures, then it follows from Lemma 3.2 that for every array $\{\mu_{nk}: 1 \leqslant k \leqslant k_n, n \in N\}$, for which convolutions in rows are relatively shift compact, there exists* *a measurable centering function B, defined on some closed subset $\mathscr{K}_0 \subset \mathscr{P}(H)$, $\mathscr{K}_0 \supset \{\mu_{nk}\}$, such that the set of all partial convolutions*

$$\{\prod^{*} \mu_{nj} * \delta_{-B(\mu_{nj})}\colon 1 \leqslant k \leqslant k_n,\ n \in N\}$$

is relatively compact. The set $\mathscr{K}_0$ may be chosen to be closed with respect to the shifts.

4. PRINCIPLE OF CONDITIONING IN A HILBERT SPACE

Consider an array $X = \{X_{nk}\colon k \in N,\ n \in N\}$ of random variables defined on a probability space $(\Omega, \mathscr{F}, P)$ with values in H, and an array $F = \{\mathscr{F}_{nk}\colon k \in N_0 = N \cup \{0\},\ n \in N\}$ of σ-subalgebras of $\mathscr{F}$ such that each row of F, i.e. $\{\mathscr{F}_{nk}\colon k \in N_0\}$, forms a filtration ($\mathscr{F}_{n,k-1} \subset \mathscr{F}_{n,k}\colon k \in N$). Say that X is adapted to F (and denote this by $\mathrm{Ad}(X, F)$) if

$$\sigma(X_{nk}) \subset \mathscr{F}_{nk}, \quad k \in N,\ n \in N. \tag{4.1}$$

For each n, k choose and fix a version $\mu_{nk}(\cdot, \cdot)\colon \mathscr{B}_H \times \Omega \to [0, 1]$ of the regular conditional distribution (r.c.d.) of X_{nk} given $\mathscr{F}_{n,k-1}$:

$$\mu_{nk}(B, \omega) = P_{X_{nk}}(B \mid \mathscr{F}_{n,k-1})(\omega) \text{ a.s.}$$

Then the kernel $\Omega \ni \omega \mapsto \mu_{n1}(\cdot, \omega) \times \mu_{n2}(\cdot, \omega) \times \ldots = \tilde{\mu}_n(\cdot, \omega) \in \mathscr{P}(H^\infty)$ defines a probability measure $\tilde{Q}$ on $(\Omega \times H^\infty, \mathscr{F} \otimes \mathscr{B}_{H^\infty})$ by the formula

$$\tilde{Q}_n(A \times B) = \int_A \tilde{\mu}_n(B, \omega) P(d\omega).$$

Note that every random variable on $(\Omega, \mathscr{F}, P)$ can be redefined on this new probability space without changing its law.

It is easy to see that the coordinate projections $X^*_{nk}\colon \Omega \times H^\infty \to H$, $X^*_{nk}(\omega, (x_1, x_2, \ldots)) = x_k$ have the following properties:

(4.2) for each $n \in N$, $\{X^*_{nk}\colon k \in N\}$ are $\tilde{Q}_n$-conditionally independent over $\tilde{\mathscr{F}} = \mathscr{F} \otimes \{\emptyset, H^\infty\}$;

(4.3) r.c.d. of X^*_{nk} given $\tilde{\mathscr{F}}$ and r.c.d. of X_{nk} given $\mathscr{F}_{n,k-1}$ coincide: $P_{X^*_{nk}}(B|\tilde{\mathscr{F}})(\omega) = P_{X_{nk}}(B \mid \mathscr{F}_{n,k-1})(\omega)$, $B \in \mathscr{B}_H$, $\omega \in \Omega$.

Now, for each $n \in N$, let $\sigma_n\colon (\Omega, \mathscr{F}, P) \to N_0$ be a stopping time with respect to $\{\mathscr{F}_{nk}\colon k \in N_0\}$. Define

$$S_n = S_n(\sigma_n) = \sum_{1 \leqslant k \leqslant \sigma_n} X_{nk}, \quad S^*_n = S^*_n(\sigma_n) = \sum_{1 \leqslant k \leqslant \sigma_n} X^*_{nk}. \tag{4.4}$$

By the definition of X^*, r.c.d. of $S^*_n(\sigma_n)$, given $\tilde{\mathscr{F}}$, satisfies

$$P_{S^*_n}(\cdot|\tilde{\mathscr{F}})(\omega) = \mu_{n1}(\cdot, \omega) * \mu_{n2}(\cdot, \omega) * \ldots * \mu_{n\sigma_n(\omega)}(\cdot, \omega) =: \mu_n(\sigma_n)(\cdot, \omega). \tag{4.5}$$

Theorems 4.1, 4.2 and 4.3 establish several connections between the laws of random measures $\{\mu_n(\sigma_n)\colon n \in N\}$ and the laws of sums $\{S_n(\sigma_n)\colon n \in N\}$. Theorem 4.2 has grown from the observation made by the author in his paper [7]. By the reasons detaily described in [9], Theorem 4.2 may be considered as a formal counterpart of the Principle of Conditioning for sums of depedent random vectors with values in a real and separable Hilbert space.

4.1. THEOREM. (i) *Suppose that there exist $\mathscr{F}$-measurable random variables $\{A_n\colon n \in N\}$ such that the sequence $\{S_n^*(\sigma_n) - A_n\colon n \in N\}$ is tight.*

Then there exist random variables C_{nk} such that

(a) *for every $n, k \in N$, C_{nk} is $\mathscr{F}_{n,k-1}$-measurable, i.e.*

$$C_n(k) = \sum_{1 \leqslant j \leqslant k} C_{nk} \tag{4.6}$$

is a predictable sequence with respect to $\{\mathscr{F}_{nk}\colon k \in N_0\}$;

(b) *the sequence $\{S_n^*(\sigma_n) - C_n(\sigma_n)\colon n \in N\}$ is tight;*

(c) *the sequence $\{S_n(\sigma_n) - C_n(\sigma_n)\colon n \in N\}$ is tight.*

(ii) *In particular, if the sequence $\{S_n^*(\sigma_n)\colon n \in N\}$ is tight, then $\{S_n(\sigma_n)\colon n \in N\}$ is tight.*

Proof. Suppose that there are measurable random variables $\{A_n\colon n \in N\}$ such that the sequence $\{S_n^*(\sigma_n) - A_n\colon n \in N\}$ is tight. Since

$$\mathscr{L}(S_n^*(\sigma_n) - A_n) = EP_{S_n^*(\sigma_n)}(\cdot \mid \mathscr{F})(\omega) * \delta_{-A_n(\omega)},$$

by Corollary 2.2 the sequence of random measures under expectation is tight. Hence

$$\{|P_{S_n^*(\sigma_n)}(\cdot \mid \mathscr{F})(\omega)|^2\colon n \in N\}$$

is tight. Define the array $\{\mu'_{nk}(\cdot, \omega)\colon k \in N,\ n \in N\}$ of random measures by stopping the array $\{\mu_{nk}(\cdot, \omega) = P_{X_{nk}}(\cdot \mid \mathscr{F}_{n,k-1})(\omega)\}$:

$$\mu'_{nk}(\cdot, \omega) = \begin{cases} \mu_{nk}(\cdot, \omega) & \text{if } k \leqslant \sigma_n(\omega), \\ \delta_0 & \text{if } k > \sigma_n(\omega). \end{cases}$$

Since

$$\mathop{\boxed{*}}_{k=1}^{\infty} \mu'_{nk}(\cdot, \omega) = P_{S_n^*(\sigma_n)}(\cdot \mid \mathscr{F})(\omega)$$

for the array $\{\mu'_{nk}\}$, condition (i) of Theorem 3.1 is satisfied. Hence, by Theorem 3.1 (ii), one can find the random variables B_{nk}, $\sigma(B_{nk}) \subset \sigma(\mu'_{n1}, \mu'_{n2}, \ldots, \mu'_{nk}) \subset \sigma(\mu_{n1}, \mu_{n2}, \ldots, \mu_{nk}, \sigma_n \wedge k-1) \subset \mathscr{F}_{n,k-1}$, such that the sequences of random S-operators $\{T^n(\sigma_n)\colon n \in N\}$ and random finite measures

$\{M^n(\sigma_n)\colon n \in N\}$ are tight, where

$$\langle T^n(\sigma_n)y, y\rangle = \sum_{k=1}^{\infty} \int \langle y, hx - B_{nk}(\omega)\rangle^2 \mu'_{nk}(dx, \omega)$$
$$= \sum_{1 \leqslant k \leqslant \sigma_n} E_{n,k-1}\big(\langle y, h(X_{nk} - B_{nk})\rangle^2\big),$$
$$M^n(\sigma_n)(A) = \sum_{k=1}^{\infty} \mu'_{nk} * \delta_{-B_{nk}}(A \cap \{\|x\| > 1\})$$
$$= \sum_{1 \leqslant k \leqslant \sigma_n} P_{n,k-1}(\|X_{nk} - B_{nk}\| > 1,\ X_{nk} - B_{nk} \in A).$$

In particular, if

$$C_{nk} = B_{nk} + \int_{\{\|x\| \leqslant 1\}} x(\mu'_{nk} * \delta_{-B_{nk}})(dx),$$

then, by Theorem 3.1 (iii), the sequence $\{\mu_n(\sigma_n)(\cdot, \omega) * \delta_{-\sum_{1 \leqslant k \leqslant \sigma_n} C_{nk}}\}$ is tight and, by Corollary 2.2, $\{S_n^*(\sigma_n) - \sum_{1 \leqslant k \leqslant \sigma_n} C_{nk}\}$ is tight. Note that

$$C_n(\sigma_n) = \sum_{1 \leqslant k \leqslant \sigma_n} C_{nk} = \sum_{1 \leqslant k \leqslant \sigma_n} B_{nk} + E_{n,k-1}\big((X_{nk} - B_{nk}) I(\|X_{nk} - B_{nk}\| \leqslant 1)\big).$$

Now consider the random variables

$$Y'_{nk} = (X_{nk} - B_{nk}) I(\|X_{nk} - B_{nk}\| > 1),$$
$$Y''_{nk} = (X_{nk} - B_{nk}) I(\|X_{nk} - B_{nk}\| \leqslant 1) - E_{n,k-1}\big((X_{nk} - B_{nk}) I(\|X_{nk} - B_{nk}\| \leqslant 1)\big).$$

Clearly,

$$S_n(\sigma_n) - C_n(\sigma_n) = \sum_{1 \leqslant k \leqslant \sigma_n} Y'_{nk} + \sum_{1 \leqslant k \leqslant \sigma_n} Y''_{nk} = S'_n(\sigma_n) + S''_n(\sigma_n). \tag{4.7}$$

Hence, in order to check the tightness of $\{S_n(\sigma_n) - C_n(\sigma_n)\}$, it suffices to check the tightness of $\{S'_n(\sigma_n)\}$ and $\{S''_n(\sigma_n)\}$ separately.

Consider an array $\{\delta_{Y'_{nk}}(\cdot)\}$ of random measures. The random finite measures

$$N_n(\sigma_n)(A) = \sum_{1 \leqslant k \leqslant \sigma_n} \delta_{Y'_{nk}}(\omega)(A \setminus \{0\})$$
$$= \sum_{1 \leqslant k \leqslant \sigma_n} I(X_{nk} - B_{nk} \in A) I(\|X_{nk} - B_{nk}\| > 1)$$

are tight. Indeed, let

$$\tau_n^C = \sigma_n \wedge \sup\{k \sum_{j=1}^{k} P_{n,j-1}(\|X_{nk} - B_{nk}\| > 1) \leqslant C\}.$$

By the tightness of $\{M^n(\sigma_n)\}$ we always can find such a $C > 0$ that $\sup_n P(\tau_n^C < \sigma_n)$ is arbitrarily small. Then the measures $M^n(\tau_n^C)$ $(\leqslant M^n(\sigma_n))$ are tight and $M^n(\tau_n^C)(H) \leqslant C$, so, by Theorem 2.1, $\{EM(\tau_n^C)\colon n \in N\}$ is relatively compact. But τ_n^C is a stopping time with respect to $\{\mathscr{F}_{nk}\colon k \in N_0\}$, hence $EM^n(\tau_n^C) = EN_n(\tau_n^C)$. Thus we may approximate the measures $N_n(\sigma_n)$ uniformly in probability by tight random measures, hence $\{N_n(\sigma_n)\}$ is tight. Since $N_n(\sigma^n)(\|x\| \leqslant 1) = 0$, it follows from Theorem 3.1(iii) that

$$\mathop{\boxed{*}}_{1 \leqslant k \leqslant \sigma_n} \delta_{Y_n'^k(\omega)}(\cdot) = \delta_{S_n'(\sigma_n)(\omega)}(\cdot)$$

is a tight sequence of random measures and, by Corollary 2.2, the sequence $E\delta_{S_n'(\sigma_n)} = \mathscr{L}(S_n'(\sigma_n))$ is relatively compact.

Now consider the sequence $\{S_n''(\sigma_n)\}$. For each $n \in N$ the sequence $\{\sum_{j=1}^{k} Y_{nj}''\colon k \in N\}$ is a martingale. Let again

$$\tau_n^C = \sigma_n \wedge \sup\{k\colon \sum_{j=1}^{k} E_{n,j-1}(\|h(X_{nk}-B_{nk})\|^2) \leqslant C\}.$$

By the arguments similar to those used for $N_n(\tau_n^C)$, the sequence of covariance operators

$$\begin{aligned}\langle U_n y, y\rangle = E\langle y, S_n''(\tau_n^C)\rangle^2 &= E \sum_{1 \leqslant k \leqslant \tau_n^C} \langle y, Y_{nk}''\rangle^2 \\ &\leqslant E \sum_{1 \leqslant k \leqslant \tau_n^C} E_{n,k-1}\langle y, h(X_{nk}-B_{nk})\rangle^2\end{aligned}$$

is relatively compact. Hence $\{S_n''(\tau_n^C)\}$ is tight. And so $\{S_n''(\sigma_n)\}$ is also tight.

4.2. THEOREM. *Suppose that*

$$\mu_n(\sigma_n)(\cdot, \omega) \underset{P}{\Rightarrow} \mu, \tag{4.8}$$

where $\mu \in \mathscr{P}(H)$ *is non-random. Then:*

(i) *The sequence* $\{S_n(\sigma_n)\colon n \in N\}$ *is tight.*

(ii) *If* ν *is a limit for some subsequence of* $\{\mathscr{L}(S_n(\sigma_n))\colon n \in N\}$, *then* ν *satisfies the equation*

$$\nu * \mu = \mu * \mu. \tag{4.9}$$

(iii) *If equation* (4.9) *has the unique solution* $\nu = \mu$, *then*

$$S_n(\sigma_n) \underset{\mathscr{D}}{\rightarrow} \mu. \tag{4.10}$$

Proof. Assume (4.8). Clearly, $\{\mathscr{L}(S_n(\sigma_n)) = EP_{S_n^*(\sigma_n)}(\cdot \mid \tilde{\mathscr{F}})\colon n \in N\}$ is relatively compact, so it follows from Theorem 4.1 that $\{\mathscr{L}(S_n(\sigma_n)) = EP_{S_n(\sigma_n)}(\cdot \mid \mathscr{F}_{n0})\colon n \in N\}$ is relatively compact and $\{P_{S_n(\sigma_n)}(\cdot \mid \mathscr{F}_{n0})\colon n \in N\}$ is a tight sequence of random probability measures.

Following the proof of Theorem 1.1, [9] (with application of Lemma 3.2 instead of Lemma 1.2), we get the convergence E_{n0} exp $i\langle y, S_n(\sigma_n)\rangle \underset{P}{\to} \hat{\mu}(y)$ provided $\hat{\mu}(y) \neq 0$. In particular, for each $y \in H$,

$$\hat{\mu}(y) E_{n0} \exp i \langle y, S_n(\sigma_n)\rangle \to [\hat{\mu}(y)]^2$$

and (4.9) follows.

Suppose that $\nu = \mu$ is the unique solution of (4.9). We shall prove a little bit more than (4.10), namely that

$$P_{S_n(\sigma_n)}(\cdot \mid \mathscr{F}_{n0}) \underset{P}{\Rightarrow} \mu. \tag{4.11}$$

Observe that we have the tightness of random measures $\{\mu * P_{S_n(\sigma_n)}(\cdot \mid \mathscr{F}_{n0})(\omega)\colon n \in N\}$ and the convergence in probability of their characteristic functions on H, and so on some countable dense subset of H. By Lemma 2.4

$$\mu * P_{S_n(\sigma_n)}(\cdot \mid \mathscr{F}_{n0})(\omega) \underset{P}{\Rightarrow} \mu * \mu.$$

Now the convergence $P_{S_n(\sigma_n)}(\cdot \mid \mathscr{F}_{n0})(\omega) \Rightarrow \mu$ follows by passing to a.s. convergent subsequences.

Say that $\mathrm{Ad}(X, F)$ is obtained by *scaling* a single sequence if $X_{nk} = B_n(X_k - A_n)$, $\mathscr{F}_{nk} = \mathscr{F}_k$, where $\{X_k\colon k \in N\}$ is adapted to $\{\mathscr{F}_k\colon k \in N_0\}$, $\{B_n\colon H \to H\colon n \in N\}$ is a sequence of bounded linear operators on H, strongly converging to 0 (i.e. $B_n(x) \to 0$ for each $x \in H$), and $\{A_n\colon n \in N\}$ is a sequence of vectors such that $B_n(A_n) \to 0$.

4.3. THEOREM. *Suppose that* $\mathrm{Ad}(X, F)$ *is obtained by scaling a single sequence and that*

$$\mu_n(\sigma_n)(\cdot, \omega) \underset{P}{\Rightarrow} \mu(\cdot, \omega), \tag{4.12}$$

where $\mu(\cdot, \omega)$ *is a random measure with the property*

$$P(\hat{\mu}(y, \cdot) = 0) = 0, \quad y \in H', \tag{4.13}$$

for some dense subset H' *of* H.

Then there exists a non-decreasing sequence $\{k_n\} \subset N$, $k_n \to \infty$, *such that*

$$P_{S_n(\sigma_n)}(\cdot \mid \mathscr{F}_{k_n}) \underset{P}{\Rightarrow} \mu(\cdot, \omega). \tag{4.14}$$

Proof. For each $k \in N$, $B_n(X_k - A_n) \to 0$ a.s. and it is easy to see that $\mu_{nk}(\cdot, \omega) = P_{B_n(X_k - A_n)}(\cdot \mid \mathscr{F}_{k-1}) \Rightarrow \delta_0$ a.s. Hence it is possible to find a non-decreasing sequence $k_n \to \infty$ such that $\mu_n(k_n) \Rightarrow \delta_0$ a.s. Just the same way as in Theorem 3.1, [9], we obtain, for each $y \in H'' \subset H$,

$$E_{k_n} \exp i \langle y, S_n(\sigma_n)\rangle \underset{P}{\to} \hat{\mu}(y, \omega),$$

where H'' is a countable dense subset of H'.

The application of Lemma 2.4 completes the proof of (4.14).

Note that at the present moment we make use of Lemma 2.4 in nontrivial way (since $\mu(\cdot, \omega)$ is not constant), while in the proof of Theorem 4.2 the convergence (4.11) may be immediately obtained from the fact that the convergence in distribution to a constant random element is equivalent to the convergence in probability.

5. MARTINGALE CENTRAL LIMIT THEOREM IN HILBERT SPACE

The Central Limit Theorem for martingale difference arrays is the most important application of the Principle of Conditioning.

Usually, CLT for martingales involves the convergence of conditional variances (Brown's Theorem — see [3] and [9]) or the convergence of quadratic variation (McLeish [11]). While in the first case there are many results for the Hilbert space-valued martingales ([2], [7], [13]), it seems that there is no infinite-dimensional generalization of the latter approach.

We shall apply Theorem 4.2 in the proof of the Hilbert space version of CLT due to Ganssler and Hausler [6] (which is an improved McLeish CLT).

Martingale difference array (MDA) is a system $\mathrm{Ad}(X, F)$ satisfying $E\|X_{nk}\| < +\infty$, $E_{n,k-1}(X_{nk}) = 0$, $n, k \in \boldsymbol{N}$.

As in Section 4, let for each $n \in \boldsymbol{N}$, σ_n: $(\Omega, \mathscr{F}, P) \to \boldsymbol{N}_0$ be a stopping time with respect to $\{\mathscr{F}_{nk}: k \in \boldsymbol{N}_0\}$.

5.1. Theorem. *Let* $\mathrm{Ad}(X, F)$ *be an MDA and let* $\{e_i: i \in \boldsymbol{N}\}$ *be an orthonormal basis in* H.

Suppose the following conditions to hold:

$$E\Big(\sup_{1 \leqslant k \leqslant \sigma_n} \|X_{nk}\|\Big) \to 0, \tag{5.1}$$

$$\sum_{1 \leqslant k \leqslant \sigma_n} \langle X_{nk}, e_i\rangle \langle X_{nk}, e_j\rangle \to \psi_{ij} \in \boldsymbol{R}^1, \qquad i, j \in \boldsymbol{N}, \tag{5.2}$$

$$\lim_{N\to\infty} \limsup_{n\to\infty} P\Big(\sum_{1 \leqslant k \leqslant \sigma_n} r_N^2(X_{nk}) > \varepsilon\Big) = 0, \qquad \varepsilon > 0, \tag{5.3}$$

where $r_N^2(x) = \sum_{i=N}^{\infty} \langle x, e_i\rangle^2$.

Then there exists a Gaussian symmetric distribution $G(0, S)$ *with the covariance operator* S *such that* $\langle Se_i, e_j\rangle = \psi_{ij}$, i, $j \in \boldsymbol{N}$, *and* $\mathscr{L}(S_n(\sigma_n)) \Rightarrow G(0, S)$.

Proof. Let $\mathscr{S}$ be the metric space of S-operators (see Section 2). Define the random S-operators U_n by the following formula:

$$\langle U_n y, y\rangle = \sum_{1 \leqslant k \leqslant \sigma_n} \langle X_{nk}, y\rangle^2.$$

Then, by (5.2) and (5.3), the sequence $\{U_n\}$ is tight in $\mathscr{S}$. By (5.2) and

Lemma 2.4 there exists an $S \in \mathscr{S}$ such that $\langle Se_i, e_j \rangle = \psi_{ij}$ and $U_n \underset{P}{\to} S$. In particular,

$$\operatorname{tr} U_n = \sum_k \|X_{nk}\|^2 \underset{P}{\to} \operatorname{tr} S. \tag{5.4}$$

Define $Y_{nk} = X_{nk} I(\|X_{nk}\| \leqslant 1)$.

Following exactly the lines of the proof of Theorem 5.2, [9], one can show that (5.1)–(5.4) imply the following conditions:

$$\sum_k E_{n,k-1}(Y_n) \underset{P}{\to} 0, \tag{5.5}$$

$$\sum_k P_{n,k-1}(\|X_{nk}\| > \varepsilon) \underset{P}{\to} 0 \quad \text{for every } \varepsilon > 0, \tag{5.6}$$

$$\sum_k E_{n,k-1}\|Y_{nk}\|^2 - \|E_{n,k-1}(Y_{nk})\|^2 \underset{P}{\to} \operatorname{tr} S, \tag{5.7}$$

$$\sum_k E_{n,k-1}\langle Y_{nk}, e_i\rangle \langle Y_{nk}, e_j\rangle - E_{n,k-1}\langle Y_{nk}, e_i\rangle E_{n,k-1}\langle Y_{nk}, e_j\rangle \underset{P}{\to} \langle Se_i, e_j\rangle, \quad i, j \in N. \tag{5.8}$$

Applying CLT for independent random vectors in the way discribed in [7], p. 185, or in [9], p. 905, we get the convergence $\mu_n(\sigma_n) \underset{P}{\Rightarrow} G(0, S)$. By our Theorem 4.2 we get $\mathscr{L}(S_n(\sigma_n)) \Rightarrow G(0, S)$.

REFERENCES

[1] A. Araujo and E. Giné, *The Central Limit Theorem for real and Banach valued random variables*, Wiley, New York 1980.

[2] E. Berger, *Asymptotic behaviour of a class of stochastic approximation procedures*, Prob. Th. Rel. Fields 71 (1986), p. 517-552.

[3] B. M. Brown and G. K. Eagleson, *Martingale convergence to infinitely divisible laws with finite variances*, Trans. Amer. Math. Soc. 162 (1971), p. 449-453.

[4] A. Dellacherie and P. Meyer, *Probabilities and potential*, North-Holland, Amsterdam 1978.

[5] J. L. Doob, *Stochastic processes*, Wiley, New York 1953.

[6] P. Ganssler and E. Hausler, *Remarks on the functional central limit theorem for martingales*, Z. Wahrsch. verw. Gebiete 50 (1979), p. 237-243.

[7] A. Jakubowski, *On limit theorems for sums of dependent Hilbert space valued random variables*, Lecture Notes in Stat. 2 (1980), p. 178-187.

[8] — *Limit theorems for sums of dependent random variables with values in a Hilbert space*, Thesis (in Polish), Toruń 1982.

[9] — *Principle of conditioning in limit theorems for sums of random variables*, Ann. Probab. 14 (1986), p. 902-915.

[10] O. Kallenberg, *Random measures*, Akademie-Verlag, Berlin 1975.

[11] D. L. McLeish, *Dependent central limit theorems and invariance principles*, Ann. Probab. 2 (1974), p. 620-628.

[12] K. R. Parthasarathy, *Probability measures on metric spaces*, Academic Press, New York 1967.

[13] H. Walk, *An invariance principle for the Robbins-Monroe process in a Hilbert space*, Z. Wahrsch. verw. Gebiete 39 (1977), p. 135 - 150.

Institute of Mathematics
Nicholas Copernicus University
ul. Chopina 12/18
87-100 Toruń, Poland

Received on 3. 2. 1986;
revised version on 1. 12. 1986