TAMKANG JOURNAL OF MATHEMATICS
Volume 30, Number 4, Winter 1999

WEAK CONVERGENCE OF COMPOUND PROBABILITY MEASURES ON UNIFORM SPACES

JUN KAWABE

Abstract. We obtain a convergence theorem of compound probability measures on a uniform space for a net of uniformly equicontinuous transition probabilities. This theorem contains convergence theorems of product or convolution measures. We also show that for Gaussian transition probabilities on a Hilbert spaces, our assumptions in the convergence theorem can be expressed in terms of mean and covariance functions.

1. Introduction

In Billingsley [1] it was proved that the product of probability measures on separable metric spaces is jointly continuous with respect to the weak topology of measures, i.e. $\mu_\alpha \xrightarrow{w} \mu$ and $\nu_\alpha \xrightarrow{w} \nu$ imply $\mu_\alpha \times \nu_\alpha \xrightarrow{w} \mu \times \nu$, where the symbol $\xrightarrow{w}$ denotes the weak convergence. This result has been extended by Vakhania *et al.* [17] to τ-smooth probability measures on arbitrary completely regular spaces. On the other hand, it was shown in Csiszár [2] that the joint continuity of convolution measures on an arbitrary topological group, i.e., $\mu_\alpha \xrightarrow{w} \mu$ and $\nu_\alpha \xrightarrow{w} \nu$ imply $\mu_\alpha * \nu_\alpha \xrightarrow{w} \mu * \nu$. These results are important in the study of weak convergence of measures. The main aim of the present paper is to obtain a convergence theorem which contains the results above. For this end, we introduce the notion of compound probability measures.

Let X and Y be topological spaces. For a probability measure μ on X and a transition probability λ on $X \times Y$, we define a compound probability measure $\mu \circ \lambda$ by

$$\mu \circ \lambda(D) = \int_X \lambda(x, D_x)\mu(dx).$$

Compound probability measures can be viewed as a generalization of product or convolution measures. In fact, if a transition probability λ is given by $\lambda(x, B) = \nu(B)$ for all $x \in X$ and all Borel subsets B of Y, where ν is a probability measure on X, then $\mu \circ \lambda$ is the product measure $\mu \times \nu$. On the other hand, if $X = Y$ is a topological group and λ

Received February 20, 1998; revised March 30, 1999.
1991 *Mathematics Subject Classification.* Primary 60B05, 60B10, Secondary 28A33, 28C15.
Key words and phrases. Uniform equicontinuity, transition probability, compound probability measure, Gaussian transition probability.
Research supported by Grant-in-Aid for General Scientific Research No. 11640160, the Ministry of Education, Science, Sports and Culture, Japan.

is given by $\lambda(x, B) = \nu(x^{-1}B)$, then the projection $\mu\lambda$ of $\mu \circ \lambda$ onto Y is the convolution measure $\mu * \nu$.

After recalling notation and necessary definitions and results, in Section 3 we obtain a main theorem concerning weak convergence of compound probability measures, and give the joint continuity of product or convolution measures as its immediate consequences. In this section, uniform equicontinuity of transition probabilities plays an important role. As another application, we also show the lower semicontinuity of mutual information for a uniformly continuous channel. In Section 4, we show that for Gaussian transition probabilities on a Hilbert space, our assumptions in the theorem can be expressed in terms of the corresponding mean and covariance functions.

Throughout this paper, we suppose that all the topological spaces, all the topological groups and all the uniform spaces are Hausdorff. We denote by $\mathbb{N}$, $\mathbb{R}$ and $\mathbb{C}$ the set of all natural numbers, real numbers and complex numbers, respectively.

2. Preliminaries and Notation

Let X be a topological space and $\mathcal{B}(X)$ be the σ-algebra of all Borel subsets of X. By a *Borel measure* on X we mean a finite measure defined on $\mathcal{B}(X)$ and we denote by $\mathcal{P}(X)$ the set of all Borel probability measures on X.

A family $\mathcal{A}$ of subsets of X is said to be *filtering downwards* (resp. *upwards*) to a subset A_0 of X, and we write $\mathcal{A} \downarrow A_0$ (resp. $\mathcal{A} \uparrow A_0$) if for any $A_1, A_2 \in \mathcal{A}$ we can find $A_3 \in \mathcal{A}$ such that $A_3 \subset A_1 \cap A_2$ (resp. $A_3 \supset A_1 \cup A_2$) and $A_0 = \bigcap_{A \in \mathcal{A}} A$ (resp. $A_0 = \bigcup_{A \in \mathcal{A}} A$). In this paper, the following concept of regularity for Borel measures is useful. We say that a Borel measure μ on X is *τ-smooth* if for any family $\mathcal{F}$ of closed subsets of X with $\mathcal{F} \downarrow F_0$ we have $\mu(F_0) = \inf_{F \in \mathcal{F}} \mu(F)$, and this is equivalent to the condition that for any family $\mathcal{G}$ of open subsets of X with $\mathcal{G} \uparrow G_0$, we have $\mu(G_0) = \sup_{G \in \mathcal{G}} \mu(G)$. We denote by $\mathcal{P}_\tau(X)$ the set of all τ-smooth probability measures on X. Every Radon measure is τ-smooth, and if X is regular every τ-smooth measure is regular (see e.g., Proposition I.3.1 of [17]). We also know that every Borel measure on a Suslin space is Radon (see Schwartz [13], Theorem II.10 in Part I, page 122), and hence τ-smooth. Here we recall that a topological space is called a *Suslin space* if it is the continuous image of some Polish space (see Definition II.3 in Part I of [13]).

If X is completely regular, we equip $\mathcal{P}(X)$ with the weakest topology for which the functionals

$$\mu \in \mathcal{P}(X) \mapsto \int_X f(x)\mu(dx), \quad f \in C_b(X),$$

are continuous. Here $C_b(X)$ denotes the set of all bounded, continuous real-valued functions on X. This topology on $\mathcal{P}(X)$ is called the *weak topology*, and we say that a net $\{\mu_\alpha\}$ in $\mathcal{P}(X)$ *converges weakly* to $\mu \in \mathcal{P}(X)$ and we write $\mu_\alpha \xrightarrow{w} \mu$, if

$$\lim_\alpha \int_X f(x)\mu_\alpha(dx) = \int_X f(x)\mu(dx)$$

for every $f \in C_b(X)$, and this is equivalent to the condition that for each open subset G (resp. closed subset F)of X,

$$\liminf_{\alpha} \mu_\alpha(G) \geq \mu(G) \ \Big(\text{resp. } \limsup_{\alpha} \mu_\alpha(F) \leq \mu(F)\Big),$$

provided that $\mu \in \mathcal{P}_\tau(X)$ (see e.g., Theorem I.3.5 of [17]).

Let X be a topological space and Y be a completely regular topological space. A (Borel) *transition probability* λ on $X \times Y$ is defined to be a mapping from X into $\mathcal{P}(Y)$ such that for every $B \in \mathcal{B}(Y)$, the function $x \in X \mapsto \lambda(x, B)$ is Borel measurable.

We say that a transition probability λ is *τ-smooth* if the probability measure $\lambda_x \equiv \lambda(x, \cdot)$ is τ-smooth for each $x \in X$, that is, it is a mapping from X into $\mathcal{P}_\tau(Y)$.

3. Uniform Equicontinuity of Transition Probabilities

Let Y be a completely regular space. Let Γ be a non-empty subset of $C_b(Y)$. The *weak topology generated by* Γ is the weakest topology for which all the functions in Γ are continuous. The *uniform structure on* $\mathcal{P}(Y)$ *generated by* Γ is the uniform structure in which a uniformity base is formed by the family of the sets

$$W = \left\{(\mu, \nu) \in \mathcal{P}(Y) \times \mathcal{P}(Y) : \left| \int_Y g_i d(\mu - \nu) \right| < \varepsilon, \ i = 1, 2, \ldots, n \right\},$$

where $\varepsilon > 0$, $n \in \mathbb{N}$, and $g_i \in \Gamma$ $(i = 1, 2, \ldots, n)$. We say that a net $\{\nu_\alpha\}$ in $\mathcal{P}(Y)$ *converges* Γ*-weakly* to $\nu \in \mathcal{P}(Y)$ and we write $\nu_\alpha \xrightarrow{\Gamma} \nu$ if

$$\lim_{\alpha} \int_Y g d\nu_\alpha = \int_Y g d\nu \quad \text{for all } g \in \Gamma.$$

It is readily proved that Γ-weak convergence is equivalent to the convergence with respect to the uniform topology determined by the uniform structure on $\mathcal{P}(Y)$ generated by Γ.

Let $F(Y)$ be the vector lattice of all real-valued functions on Y with the usual pointwise ordering and lattice operations: $f \leq g \Leftrightarrow f(y) \leq g(y)$ for all $y \in Y$, $f \vee g = \max(f, g)$ and $f \wedge g = \min(f, g)$. In the rest of this paper we often assume that $\Gamma \subset F(Y)$ satisfies the following conditions:

(Γ1) For any closed subset F of Y and any $a \notin F$, there exists $g \in \Gamma$ such that $0 \leq g \leq 1$, $g(a) = 0$ and $g(y) = 1$ for all $y \in F$.

(Γ2) g_1, $g_2 \in \Gamma$ implies $g_1 \wedge g_2 \in \Gamma$.

If Γ satisfies (Γ1) and $\Gamma \subset C_b(Y)$, then it generates the topology on Y, i.e., the original topology on Y coincides with the weak topology generated by Γ. We also note that if Γ is a linear subspace of $F(Y)$, then (Γ2) implies that Γ is a vector sublattice of $F(Y)$, i.e., it is a linear subspace with the property that g_1, $g_2 \in \Gamma$ implies that $g_1 \vee g_2$, $g_1 \wedge g_2 \in \Gamma$.

In case when we have to treat characteristic functions of probability measures, in addition to (Γ1) and (Γ2), we shall assume the following somewhat technical condition:

(Γ3) $g \in \Gamma$ implies $\mathrm{Re}\{e^{ig}\}$, $\mathrm{Im}\{e^{ig}\} \in \Gamma$.

For any uniform space Y denote by $U_b(Y)$ the vector lattice of all bounded, uniformly continuous real-valued functions on Y. For any metric space (Y, d) denote by $BL(Y, d)$ the vector lattice of all bounded, real-valued Lipschitz functions on Y with the norm $\|g\|_{BL} = \|g\|_L + \|g\|_\infty$, where $\|g\|_L = \sup_{x\neq y} |g(x) - g(y)|/d(x, y)$ and $\|g\|_\infty = \sup_{y\in Y} |g(y)|$. The following example indicates that we can utilize the spaces $U_b(Y)$ and $BL(Y, d)$ for concrete examples of Γ.

Example 1. (1) Let Y be a uniform space. Then $U_b(Y)$ satisfies (Γ1), (Γ2) and (Γ3).

(2) Let (Y, d) be a metric space. Then $BL(Y, d)$ satisfies (Γ1), (Γ2) and (Γ3).

Proof. The verification of (Γ2) and (Γ3) is easy. So we only prove (Γ1). Fix a closed subset F of Y and $a \notin F$. (1) By a variant of a construction for non-constant functions due to Urysohn (see. e.g., Proposition 11.5 of James [8]), we can find a function $g \in U_b(Y)$ such that $0 \le g \le 1$, $g(a) = 0$ and $g(y) = 1$ for all $y \in F$. Hence $U_b(Y)$ satisfies (Γ1).

(2) Put $g(y) = 0 \vee (1 - d(y, F)/\alpha)$, $y \in Y$, where $\alpha = d(a, F) > 0$. Then $g \in BL(Y, d)$ such that $0 \le g \le 1$, $g(a) = 0$ and $g(y) = 1$ for all $y \in F$. Hence $BL(Y, d)$ satisfies (Γ1).

We collect some probably known facts about Γ-weak topology defined above and give their proofs for the sake of completeness.

Proposition 1. *Let Y be a completely regular space. Assume that a non-empty subset Γ of $C_b(Y)$ satisfies (Γ1) and (Γ2). Then the following statements are valid.*

(1) *For any net $\{\nu_\alpha\} \subset \mathcal{P}(Y)$ and $\nu \in \mathcal{P}_\tau(Y)$, $\nu_\alpha \xrightarrow{w} \nu$ if and only if $\nu_\alpha \xrightarrow{\Gamma} \nu$.*
(2) *The usual weak topology on $\mathcal{P}_\tau(Y)$ coincides with the uniform topology determined by the uniform structure on $\mathcal{P}_\tau(Y)$ generated by Γ.*

Proof. (2) follows from (1), and *only if* part of (1) is obvious. So we prove *if* part of (1). Assume that $\nu_\alpha \xrightarrow{\Gamma} \nu$ and fix a closed subset F of Y. Then by Portmanteau theorem (see e.g., Theorem 8.1 of [15]), in order to prove that $\nu_\alpha \xrightarrow{w} \nu$ it is sufficient to show that

$$\limsup_\alpha \nu_\alpha(F) \le \nu(F). \tag{3.1}$$

Since (3.1) is trivial for $G = \emptyset$ and Y, we assume that $\emptyset \subsetneqq F \subsetneqq Y$. Put

$$\mathcal{F} = \{g \in \Gamma : 0 \le g \le 1 \text{ and } g(y) = 1 \text{ for all } y \in F\}.$$

By (Γ1), $\mathcal{F}$ is a uniformly bounded, non-empty subset of $C_b(Y)$, and by (Γ2) it is filtering to the left, i.e., for any g_1, $g_2 \in \mathcal{F}$ we can find $g_3 \in \mathcal{F}$ such that $g_3 \le g_1 \wedge g_2$. Further, it is easy to prove that $\inf\{g : g \in \mathcal{F}\} = 1_F$, where 1_F is the indicator function of F. Since $\nu_\alpha \xrightarrow{\Gamma} \nu$ and ν is τ-smooth, by P15 of [15, page XIII] we have

$$\nu(F) = \inf_{g\in\mathcal{F}} \int_Y g d\nu = \inf_{g\in\mathcal{F}} \lim_\alpha \int_Y g d\nu_\alpha \ge \limsup_\alpha \int_Y 1_F d\nu_\alpha = \limsup_\alpha \nu_\alpha(F),$$

and the proof is complete.

Let X be a uniform space with the uniformity $\mathcal{U}_X$. Let Y be completely regular space and $\Gamma \subset C_b(Y)$. We say that a τ-smooth transition probability λ on $X \times Y$ is Γ-*uniformly continuous* if for each $g \in \Gamma$, the mapping $x \in X \mapsto \int_Y g(y)\lambda(x, dy)$ is uniformly continuous. Denote by $U_\Gamma(X, Y)$ the set of all τ-smooth, Γ-uniformly continuous transition probabilities on $X \times Y$. We need the following proposition in order to define the notion of compound probability measures:

Proposition 2 ([9]). *Let X be a topological space and Y be a completely regular topological space. Let λ be a mapping from X into $\mathcal{P}_\tau(Y)$. Assume that $\mathcal{P}_\tau(Y)$ is equipped with the usual weak topology of measures.*

(1) *λ is continuous if and only if for each open subset U of $X \times Y$, the function $x \in X \mapsto \lambda(x, U_x)$ is lower semi-continuous on X.*
(2) *If λ is continuous then for each Borel subset D of $X \times Y$, the function $x \in X \mapsto \lambda(x, D_x)$ is Borel measurable.*

Here for a subset D of $X \times Y$ and $x \in X$, D_x denotes the section determined by x, that is, $D_x = \{y \in Y : (x, y) \in D\}$.

If we assume that $\Gamma \subset C_b(Y)$ satisfies (Γ1) and (Γ2), then by (1) of Proposition 1, every $\lambda \in U_\Gamma(X, Y)$ is a continuous mapping from X with the uniform topology into $\mathcal{P}_\tau(Y)$ with the usual weak topology of measures. Consequently, by (2) of Proposition 2, for any $\mu \in \mathcal{P}(X)$ and any $\lambda \in U_\Gamma(X, Y)$, we can define a Borel probability measure $\mu \circ \lambda$ on $X \times Y$, which is called the *compound probability measure* of μ and λ, by

$$\mu \circ \lambda(D) = \int_X \lambda(x, D_x)\mu(dx) \quad \text{for all } D \in \mathcal{B}(X \times Y).$$

Denote by $\mu\lambda$ the projection of $\mu \circ \lambda$ onto Y, that is, $\mu\lambda(B) = \mu \circ \lambda(X \times B)$ for all $B \in \mathcal{B}(Y)$. If μ is τ-smooth, then $\mu \circ \lambda$ and $\mu\lambda$ are also τ-smooth (see Proposition 2 of [9]). By a standard argument, we can verify that the Fubini's theorem remains valid for all Borel measurable and $\mu \circ \lambda$-integrable functions h on $X \times Y$;

$$\int_{X \times Y} h(x, y)\mu \circ \lambda(dx, dy) = \int_X \int_Y h(x, y)\lambda(x, dy)\mu(dx).$$

We say that a subset Q of $U_\Gamma(X, Y)$ is Γ-*uniformly equicontinuous* if for each $g \in \Gamma$, the set of mappings $x \in X \mapsto \int_Y g(y)\lambda(x, dy)$, $\lambda \in Q$, is uniformly equicontinuous. It is obvious that the Γ-uniform equicontinuity of Q is equivalent to the condition that Q is a uniformly equicontinuous set of mappings from the uniform space $(X, \mathcal{U}_X)$ into the uniform space $\mathcal{P}_\tau(Y)$ with the uniform structure generated by Γ.

We have typical examples of uniformly equicontinuous transition probabilities below. As another important example of them, we shall consider Gaussian transition probabilities on a Hilbert space in Section 4.

Example 2. (1) Let X and Y be uniform spaces and $\{\nu_\alpha\} \subset \mathcal{P}_\tau(Y)$. Put for all α,

$$\lambda_\alpha(x, B) = \nu_\alpha(B) \quad \text{for all } x \in X \text{ and all } B \in \mathcal{B}(Y).$$

Then $\{\lambda_\alpha\}$ is $U_b(Y)$-uniformly equicontinuous.

(2) Let G be a topological group and $\{\nu_\alpha\} \subset \mathcal{P}_\tau(G)$. Put for all α,

$$\lambda_\alpha(x, B) = \nu_\alpha(x^{-1}B) \quad \text{for all } x \in X \text{ and all } B \in \mathcal{B}(G).$$

Then $\{\lambda_\alpha\}$ is $U_b(G)$-uniformly equicontinuous with respect to the right uniform structure on G.

(3) Let T and Y be uniform spaces with their uniformities $\mathcal{U}_T$ and $\mathcal{U}_Y$, respectively. Let $(\Omega, \mathcal{A}, P)$ be a probability measure space and $Z_\alpha(t, \omega)$, $t \in T$ and $\omega \in \Omega$, be $\mathcal{B}(T) \times \mathcal{A}$-measurable, Y-valued, τ-smooth stochastic processes which are uniformly equicontinuous in probability, that is, their distributions on Y are τ-smooth, and for every $\varepsilon > 0$ and every $V \in \mathcal{U}_Y$ there exists $U \in \mathcal{U}_T$ such that if $(t_1, t_2) \in U$ then $P(\{\omega \in \Omega : (Z_\alpha(t_1, \omega), Z_\alpha(t_2, \omega)) \in V\}) > 1 - \varepsilon$ for all α. Put for all α,

$$\lambda_\alpha(t, B) = P(\{\omega \in \Omega : Z_\alpha(t, \omega) \in B\}) \quad \text{for all } t \in T \text{ and all } B \in \mathcal{B}(Y).$$

Then $\{\lambda_\alpha\}$ is $U_b(Y)$-uniformly equicontinuous.

Proof. (1) is obvious. (2) Let G be a topological group with the right uniform structure. Since the ν_α's are τ-smooth, it follows that the λ_α's are τ-smooth. So we only show that $\{\lambda_\alpha\}$ is $U_b(G)$-uniformly equicontinuous.

Fix $\varepsilon > 0$ and $g \in U_b(G)$. We put

$$\varphi_\alpha(x) = \int_G g(y)\lambda_\alpha(x, dy) = \int_G g(xz)\nu_\alpha(dz), \quad x \in G.$$

Then we can find an open neighborhood U of the neutral element of G such that $x_2x_1^{-1} \in U$ implies $|g(x_1) - g(x_2)| < \varepsilon$. Since $(x_2z)(x_1z)^{-1} = x_2x_1^{-1} \in U$ for all $z \in G$, $x_2x_1^{-1} \in U$ implies $|g(x_1z) - g(x_2z)| < \varepsilon$ for all $z \in G$, and from this it follows that

$$|\varphi_\alpha(x_1) - \varphi_\alpha(x_2)| \leq \int_G |g(x_1z) - g(x_2z)|\nu_\alpha(dz) \leq \varepsilon \text{ for all } \alpha.$$

Consequently $\{\varphi_\alpha\}$ is uniformly equicontinuous with respect to the right uniform structure on G, and this implies $U_b(G)$-uniform equicontinuity of $\{\lambda_\alpha\}$.

(3) It can be proved using the definition of uniform equicontinuity in probability for stochastic processes, and hence we omit its proof.

Now we can state our main theorem concerning the weak convergence of compound probability measures.

Theorem 1. *Let X be a uniform space. Let Y be a completely regular space and let a linear subspace Γ of $C_b(Y)$ satisfy (Γ1), (Γ2) and (Γ3). Assume that a net $\{\lambda_\alpha\} \subset U_\Gamma(X, Y)$ satisfies*

(a) *$\{\lambda_\alpha\}$ is Γ-uniformly equicontinuous, and*

(b) *there exists $\lambda \in U_\Gamma(X,Y)$ such that $\lambda_\alpha(x,\cdot)\xrightarrow{\Gamma}\lambda(x,\cdot)$ for each $x \in X$.*

Then for any net $\{\mu_\alpha\}$ in $\mathcal{P}(X)$ converging weakly to $\mu \in \mathcal{P}_\tau(X)$, we have $\mu_\alpha \circ \lambda_\alpha \xrightarrow{w} \mu \circ \lambda$.

Remark 1. In Theorem 1 of [9] we assume that $\{\lambda_\alpha\}$ is equicontinuous on every compact subset of X, i.e, for each $g \in C_b(Y)$ the set of mappings $x \in X \to \int_Y g(y)\lambda(x,dy)$, $\lambda \in Q$, is equicontinuous on every compact subset of X. Then it is easy to see that the equicontinuity above is equivalent to $C_b(Y)$-*uniform* equicontinuity on every compact subset of X, i.e., for each $g \in C_b(Y)$ the set of mappings $x \in X \to \varphi_\lambda(x) \equiv \int_Y g(y)\lambda(x,dy)$, $\lambda \in Q$, is *uniformly* equicontinuous on every compact subset of X. (cf. Theorem 2.4.5 of [4]). Consequently, our assumption imposed on $\{\lambda_\alpha\}$ of Theorem 1 is stronger than that of Theorem 1 in [9] as to the extent of the domain of definition, in which $\{\varphi_\lambda\}$ is uniformly continuous. However, it has an advantage that the same result of Theorem 1 of [9], as well as the convergence of convolution measures and product measures, holds without *any* assumptions concerning the restriction of spaces and uniform tightness of probability measures and transition probabilities.

Further, our assumptions in Theorem 1 do not seem to be very strong conditions for Gaussian transition probabilities on a Hilbert space, which are one of the most important examples, because they are derived from uniform equicontinuity of the corresponding mean and covariance functions (see Theorems 2 and 3 in Section 4).

Before starting to prove Theorem 1, we give some applications.

(I) *Applications to Weak Convergence of Product or Convolution Measures:* Let G be a topological group and let $\mu \in \mathcal{P}(G)$ and $\nu \in \mathcal{P}_\tau(G)$. By [2] (or Propositions 1 and 2 of [9]), we can define the *convolution* $\mu * \nu$ of μ and ν by

$$\mu * \nu(B) = \mu \circ \lambda(G \times B) \text{ for all } B \in \mathcal{B}(G),$$

where $\lambda(x,B) = \nu(x^{-1}B)$ for all $x \in G$ and all $B \in \mathcal{B}(G)$. Then we have

Corollary 1. *Let G be a topological group, and let $\{\mu_\alpha\}$ be a net in $\mathcal{P}(G)$ and $\{\nu_\alpha\}$ be a net in $\mathcal{P}_\tau(G)$. If $\mu_\alpha \xrightarrow{w} \mu$ and $\nu_\alpha \xrightarrow{w} \nu$, where $\mu, \nu \in \mathcal{P}_\tau(G)$, then $\mu_\alpha * \nu_\alpha \xrightarrow{w} \mu * \nu$.*

Proof. We put $\lambda_\alpha(x,B) = \nu_\alpha(x^{-1}B)$ and $\lambda(x,B) = \nu(x^{-1}B)$ for all $x \in G$ and $B \in \mathcal{B}(G)$. Then $\{\lambda_\alpha\}$ is $U_b(G)$-uniformly equicontinuous with respect to the right uniform structure by (2) of Example 2, and it is clear that $\lambda_\alpha(x,\cdot)\xrightarrow{w}\lambda(x,\cdot)$ for each $x \in G$. Therefore, if we set $X = Y = G$ and $\Gamma = U_b(G)$ in Theorem 1, then we have $\mu_\alpha \circ \lambda_\alpha \xrightarrow{w} \mu \circ \lambda$ and hence $\mu_\alpha\lambda_\alpha \xrightarrow{w} \mu\lambda$. This implies $\mu_\alpha * \nu_\alpha \xrightarrow{w} \mu * \nu$.

Remark 2. The corollary above is Corollary to Theorem 1 of [2].

If we put $\lambda(x,B) = \nu(B)$ for all $x \in X$ and $B \in \mathcal{B}(Y)$, where $\nu \in \mathcal{P}_\tau(Y)$, then we have $\mu \circ \lambda = \mu \times \nu$ for each $\mu \in \mathcal{P}(X)$, and hence by (1) of Example 2 and Theorem 1, we can easily prove the following

Corollary 2. *Let X and Y be uniform spaces. Let $\{\mu_\alpha\}$ be a net $\mathcal{P}(X)$ and $\{\nu_\alpha\}$ be a net in $\mathcal{P}_\tau(Y)$. If $\mu_\alpha \xrightarrow{w} \mu \in \mathcal{P}_\tau(X)$ and $\nu_\alpha \xrightarrow{w} \nu \in \mathcal{P}_\tau(Y)$ then $\mu_\alpha \times \nu_\alpha \xrightarrow{w} \mu \times \nu$.*

Remark 3. Corollary 2 was known in case when X and Y are separable metric spaces (see e.g. [1, Theorem 3.2]), and has been extended by Vakhania *et al.* [17, Proposition I.4.1] to completely regular spaces. But their technique is that the weak convergence $\mu_\alpha \xrightarrow{w} \mu$ can be proved by showing that $\mu_\alpha(A) \to \mu(A)$ for some special class of sets A, and is different from ours.

(II) *An Application to Information Theory:* In information theory, a transition probability $\lambda : X \to \mathcal{P}(Y)$ is called a channel from an input space X to an output space Y, and $\mu \in \mathcal{P}(X)$, $\mu\lambda \in \mathcal{P}(Y)$, and $\mu \circ \lambda \in \mathcal{P}(X \times Y)$ are called an input source, an output source, and a compound source, respectively (see e.g., Umegaki [16]). Then our result shows that for a net of uniformly equicontinuous channels, the weak convergence of input sources and channels assures the weak convergence of output sources and compound sources.

To evaluate channel capacity, we use the mutual information $I(\mu, \lambda)$ between an input source and an output source, which is defined using the notion of relative entropy as follows:

$$I(\mu, \lambda) = H(\mu \circ \lambda, \mu \times \mu\lambda),$$

where for any probability measures μ and ν on a measurable space

$$H(\mu, \nu) = \begin{cases} \int \log \frac{d\mu}{d\nu} d\mu & \text{if } \mu \ll \nu \\ \infty & \text{otherwise,} \end{cases}$$

where $\mu \ll \nu$ denotes that μ is absolutely continuous with respect to ν.

We know that the relative entropy $H(\mu, \nu)$ is a lower semicontinuous function of μ and ν in the weak topology of measures on a metric space (this fact can be proved as in the proof of corollary to Lemma 2.1 of Donsker and Varadhan [3]). Consequently, by our result we obtain the lower semicontinuity of mutual information for a *uniformly continuous* channel:

$$\mu_\alpha \xrightarrow{w} \mu \text{ implies } \liminf_\alpha I(\mu_\alpha, \lambda) \geq I(\mu, \lambda),$$

which will be of use in information theory.

We start to prove Theorem 1. The following lemma is a slight generalization of Lemma 1 of [2].

Lemma 1. *Let X be a uniform space with its uniformity $\mathcal{U}_X$ and $\mu \in \mathcal{P}_\tau(X)$. Then for every $U \in \mathcal{U}_X$ and $\varepsilon > 0$ there exist $x_1, x_2, \ldots, x_n \in X$ and a function $g \in U_b(X)$ such that $0 \leq g \leq 1$, $g(x_i) = 0$ for $i = 1, 2, \ldots, n$, $g(x) = 1$ for $x \notin \bigcup_{i=1}^n U(x_i)$ and $\int_X g(x)\mu(dx) < \varepsilon$, where $U(x_i) = \{x \in X : (x, x_i) \in U\}$.*

Proof. Let $U \in \mathcal{U}_X$ and fix $\varepsilon > 0$. Since X is a uniform space, for each $a \in X$ we can find a function $g_a \in U_b(X)$ which satisfies $0 \leq g_a \leq 1$, $g_a(a) = 0$, and $g_a(x) = 1$ if

$x \notin U(a)$ (the existence of such a function g_a follows from a variant of a construction for non-constant functions due to Urysohn; see e.g., Proposition 11.5 of [8]). Let α range over the finite subsets of X and for $\alpha = \{x_1, x_2, \dots, x_n\}$ we put

$$g_\alpha(x) = \min_{1\le i\le n} g_{x_i}(x), \quad x \in X.$$

Then $\{g_\alpha\}$ is a decreasing net of uniformly continuous functions on X converging pointwise to 0 (the ordering for the α's being the set-theoretical inclusion). Since μ is τ-smooth, $\int_X g_\alpha(x)\mu(dx) < \varepsilon$ for some $\alpha = \{x_1, x_2, \dots, x_n\}$ by Proposition I.3.2 of [17], and this function g_α is a required one.

Denote by $U_b(X;\mathbb{C})$ the set of all bounded, uniformly continuous complex-valued functions on X. We can prove the following lemma as in the proof of Theorem 1 of [2], if we shall use Lemma 1 of this paper instead of Lemma 1 of [2]. So we omit its proof.

Lemma 2. *Let X be a uniform space and let $\{\mu_\alpha\}$ be a net in $\mathcal{P}(X)$. Suppose that a net $\{\varphi_\alpha\} \subset U_b(X,\mathbb{C})$ satisfies*

(a) *$\{\varphi_\alpha\}$ is uniformly bounded, and*
(b) *$\{\varphi_\alpha\}$ is uniformly equicontinuous.*

If $\mu \in \mathcal{P}_\tau(X)$ and $\mu_\alpha \xrightarrow{w} \mu$, and if $\varphi \in U_b(X;\mathbb{C})$ and $\varphi_\alpha(x) \to \varphi(x)$ for each $x \in X$, then we have

$$\lim_\alpha \int_X \varphi_\alpha(x)\mu_\alpha(dx) = \int_X \varphi(x)\mu(dx).$$

Lemma 3. *Let X be a uniform space. Let Y be a completely regular space and $\Gamma \subset C_b(Y)$. Let $\{\mu_\alpha\}$ be a net in $\mathcal{P}(X)$. Assume that $\{\lambda_\alpha\} \subset U_\Gamma(X,Y)$ is Γ-uniformly equicontinuous. If $\mu \in \mathcal{P}_\tau(X)$ and $\mu_\alpha \xrightarrow{w} \mu$, and if $\lambda \in U_\Gamma(X,Y)$ and $\lambda_\alpha(x,\cdot) \xrightarrow{\Gamma} \lambda(x,\cdot)$ for each $x \in X$, then $\mu_\alpha\lambda_\alpha \xrightarrow{\Gamma} \mu\lambda$. Consequently, if Γ satisfies (Γ1) and (Γ2), then we also have $\mu_\alpha\lambda_\alpha \xrightarrow{w} \mu\lambda$.*

Proof. Fix $g \in \Gamma$ and put for each $x \in X$,

$$\varphi_\alpha(x) = \int_Y g(y)\lambda_\alpha(x,dy) \text{ and } \varphi(x) = \int_Y g(y)\lambda(x,dy).$$

Then the net $\{\varphi_\alpha\}$ satisfies assumptions of Lemma 2. Since $\mu_\alpha \xrightarrow{w} \mu$, by Lemma 2 we have

$$\lim_\alpha \int_Y g(y)\mu_\alpha\lambda_\alpha(dy) = \lim_\alpha \int_X \varphi_\alpha(x)\mu_\alpha(dx) = \int_X \varphi(x)\mu(dx) = \int_Y g(y)\mu\lambda(dy),$$

and this implies $\mu_\alpha\lambda_\alpha \xrightarrow{\Gamma} \mu\lambda$. If Γ satisfies (Γ1) and (Γ2), it follows from (1) of Proposition 1 that $\mu_\alpha\lambda_\alpha \xrightarrow{w} \mu\lambda$, since $\mu\lambda$ is τ-smooth as stated in Section 3.

Lemma 4. *Let X be a regular space. Assume that a family $\mathcal{G}$ of subsets of X satisfies the following two conditions:*

(1) *$\mathcal{G}$ is closed under finite unions.*
(2) *$\mathcal{G}$ contains an open basis for the topology of X.*

Then, for any non-empty open subset G of X we can find a subfamily $\mathcal{H}$ of $\mathcal{G}$ such that $\mathcal{H} \uparrow G$.

Proof. We put $\mathcal{H} = \{H \in \mathcal{G} : H \subset \overline{H} \subset G\}$. Then, using assumption (1), it is easy to see that $\mathcal{H}$ is filtering upwards. Hence we have only to show that the equality $\bigcup_{H \in \mathcal{H}} H = G$ holds. Fix $x \in G$. Since X is regular, there exists an open subset U of X such that $x \in U \subset \overline{U} \subset G$. On the other hand, we can find $H \in \mathcal{G}$ such that $x \in H \subset U$ by assumption (2). Consequently we have $x \in H \in \mathcal{H}$, and this implies $\bigcup_{H \in \mathcal{H}} H \supset G$. The reverse inclusion is obvious.

For each $\mu \in \mathcal{P}(X)$, define its characteristic function $\hat{\mu}$ by

$$\hat{\mu}(f) = \int_X e^{if(x)} \mu(dx), \quad f \in C(X),$$

where $C(X)$ denotes the set of all continuous real-valued functions on X. Then the following lemma asserts that a τ-smooth probability measure, as well as a Radon probability measure, is uniquely determined by its characteristic function (c.f. Theorem IV.2.2 of [17]).

Lemma 5. *Let X be a completely regular space and μ, $\nu \in \mathcal{P}_\tau(X)$. Assume that a linear subspace Σ of $C(X)$ generates the topology on X, that is, the original topology on X coincides with the weak topology generated by Σ. If $\hat{\mu}(f) = \hat{\nu}(f)$ for all $f \in \Sigma$ then we have $\mu = \nu$ on $\mathcal{B}(X)$.*

Proof. Let $\mathcal{C}$ be the σ-field generated by the cylinder sets of the form

$$\{x \in X : (f_1(x), f_2(x), \ldots, f_n(x)) \in B\},$$

where $n \in \mathbb{N}, f_1, f_2, \ldots, f_n \in \Sigma$ and $B \in \mathcal{B}(\mathbb{R}^n)$. Here $\mathbb{R}^n$ denotes the n-dimensional Euclidean space. Denote by μ_0 and ν_0 the restrictions of μ and ν to $\mathcal{C}$, respectively. Then by Lemma I.3.1 of [17], μ_0 and ν_0 are regular on $\mathcal{C}$, and it is obvious that μ_0 and ν_0 are τ_0-smooth on $\mathcal{C}$ (see also [17] for necessary definitions). Consequently, noting that $\mathcal{C}$ contains an open basis for the topology of X, by Theorem I.3.2-(a) of [17], μ_0 and ν_0 admit unique τ-smooth Borel extensions $\tilde{\mu}_0$ and $\tilde{\nu}_0$, respectively. Since $\hat{\mu}(f) = \hat{\nu}(f)$ for all $f \in \Sigma$, we have $\mu_0 = \nu_0$ on $\mathcal{C}$ by Theorem IV. 2.2-(a) of [17], and by the uniqueness of the extension, we have $\tilde{\mu}_0 = \tilde{\nu}_0$ on $\mathcal{B}(X)$. Thus we complete the proof if we show that $\mu = \tilde{\mu}_0$ and $\nu = \tilde{\nu}_0$ on $\mathcal{B}(X)$.

It is obvious that $\mu(C) = \tilde{\mu}_0(C)$ for all $C \in \mathcal{C}$. Let G be an arbitrary non-empty open subset of X. Since $\mathcal{C}$ satisfies conditions (1) and (2) of Lemma 4, we can find a subfamily $\mathcal{H}$ of $\mathcal{C}$ with $\mathcal{H} \uparrow G$. Then we have $\mu(G) = \sup_{H \in \mathcal{H}} \mu(H) = \sup_{H \in \mathcal{H}} \tilde{\mu}_0(H) = \tilde{\mu}_0(G)$ since μ and $\tilde{\mu}_0$ are τ-smooth. Therefore μ and $\tilde{\mu}_0$ coincide on the open subsets of X, and hence they coincide on $\mathcal{B}(X)$ since μ and $\tilde{\mu}_0$ are also regular. Similarly we can show that $\nu = \tilde{\nu}_0$ on $\mathcal{B}(X)$.

We shall prove the theorem.

Proof of Theorem 1. Assume that $\{\mu_\alpha\}$ and $\{\lambda_\alpha\}$ satisfy conditions of Theorem 1. We first show that every subnet of $\{\mu_\alpha \circ \lambda_\alpha\}$ contains a further subnet converging to a τ-smooth probability measure on $X \times Y$. To do this, by Theorem 6 of [14], we have only to show that the condition of "τ-smoothness"

$$\inf_{F \in \mathcal{F}} \limsup_{\alpha} \mu_\alpha \circ \lambda_\alpha(F) = 0 \tag{3.2}$$

holds for every family $\mathcal{F}$ of closed subsets of $X \times Y$ with $\mathcal{F} \downarrow \emptyset$. (We remark that in Theorem 6 of [14], we need not assume that a net (μ_α) is in $\mathcal{M}_+(X;\tau)$— it is enough to consider a net in $\mathcal{M}_+(X)$; then the conditions of the theorem are the necessary and sufficient conditions that every subnet of (μ_α) contains a further subnet converging to a measure in $\mathcal{M}_+(X;\tau)$).

Fix $\varepsilon > 0$ and let $\mathcal{F}$ be an arbitrary family of closed subsets of $X \times Y$ with $\mathcal{F} \downarrow \emptyset$. If we put $\mathcal{E} = \{G^c : G \text{ is open and } G^c \supset F \text{ for some } F \in \mathcal{F}\}$, then we can show that $\mathcal{E} \downarrow \emptyset$ as in the proof of Lemma 4. Since μ and λ are τ-smooth, by Proposition 2 of [9], $\mu \circ \lambda$ is also τ-smooth, and hence we have $\inf_{E \in \mathcal{E}} \mu \circ \lambda(E) = 0$, which implies that there exists an open subset G_ε of $X \times Y$ such that $G_\varepsilon^c \supset F_\varepsilon$ for some $F_\varepsilon \in \mathcal{F}$ and

$$\mu \circ \lambda(G_\varepsilon^c) \le \frac{\varepsilon}{3}. \tag{3.3}$$

We put

$$\mathcal{G} = \left\{ \bigcup_{i=1}^{n} (U_i \times V_i) : \text{The } U_i\text{'s are open subsets of } X \text{ and the } V_i\text{'s are open subsets of } Y \right\},$$

then $\mathcal{G}$ satisfies conditions (1) and (2) of Lemma 4. Consequently we can find a subfamily $\mathcal{H}$ of $\mathcal{G}$ such that $\mathcal{H} \uparrow G_\varepsilon$, and this and (3.3) implies

$$\mu \circ \lambda(G_\varepsilon) = \sup_{H \in \mathcal{H}} \mu(H) > 1 - \frac{\varepsilon}{3}.$$

Thus there exists $H_\varepsilon \in \mathcal{H}$ such that $H_\varepsilon \subset G_\varepsilon$ and

$$\mu \circ \lambda(H_\varepsilon) > 1 - \frac{\varepsilon}{2}. \tag{3.4}$$

Since $H_\varepsilon \in \mathcal{G}$, it can be expressed by the form $H_\varepsilon = \bigcup_{i=1}^{n} (U_i \times V_i)$, where the U_i's are open subsets of X and the V_i's are open subsets of Y.

Since Γ satisfies (Γ1) and (Γ2), by Lemma 3 and assumptions of Theorem 1, we have

$$\mu_\alpha \xrightarrow{w} \mu \quad \text{and} \quad \mu_\alpha \lambda_\alpha \xrightarrow{w} \mu\lambda. \tag{3.5}$$

Consequently, if we notice that $F_\varepsilon \subset G_\varepsilon^c \subset H_\varepsilon^c$ and the equality

$$H_\varepsilon^c = \left\{\left(\bigcap_{i=1}^n U_i^c\right) \times Y\right\} \cup \left\{X \times \left(\bigcap_{i=1}^n V_i^c\right)\right\}$$

holds, putting $K = \bigcap_{i=1}^n U_i^c$ and $L = \bigcap_{i=1}^n V_i^c$, then by (3.4) and (3.5) we have

$$\begin{aligned}
\limsup_\alpha \mu_\alpha \circ \lambda_\alpha(F_\varepsilon) &\leq \limsup_\alpha \mu_\alpha \circ \lambda_\alpha(H_\varepsilon^c) \\
&= \limsup_\alpha \mu_\alpha \circ \lambda_\alpha((K \times Y) \cup (X \times L)) \\
&\leq \limsup_\alpha \mu_\alpha \circ \lambda_\alpha(K \times Y) + \limsup_\alpha \mu_\alpha \circ \lambda_\alpha(X \times L) \\
&= \limsup_\alpha \mu_\alpha(K) + \limsup_\alpha \mu_\alpha\lambda_\alpha(L) \\
&\leq \mu(K) + \mu\lambda(L) \\
&= \mu \circ \lambda(K \times Y) + \mu \circ \lambda(X \times L) \\
&\leq 2\mu \circ \lambda(H_\varepsilon^c) < \varepsilon,
\end{aligned}$$

and this implies that (3.2) holds.

For each $\gamma \in \mathcal{P}(X \times Y)$, define its characteristic function $\hat{\gamma}$ by

$$\hat{\gamma}(h) = \int_{X\times Y} e^{ih(x,y)}\gamma(dx, dy), \quad h \in C(X \times Y).$$

We put $\Sigma = U_b(X) \oplus \Gamma \equiv \{f \oplus g : f \in U_b(X), g \in \Gamma\}$, where $(f \oplus g)(x, y) = f(x) + g(y)$, $(x, y) \in X \times Y$. Since Γ is a linear subspace of $C_b(Y)$ and satisfies (Γ1), it is easy to see that Σ is a linear subspace of $C(X \times Y)$ and the product topology on $X \times Y$ is generated by Σ. Then by Lemma 5 and a standard argument (see Theorem IV.3.1 of [17]), in order to complete the proof it is sufficient to show that for each $f \in U_b(X)$ and $g \in \Gamma$, we have

$$(\mu_\alpha \circ \lambda_\alpha)\hat{}(f \oplus g) \to (\mu \circ \lambda)\hat{}(f \oplus g). \tag{3.6}$$

Fix $f \in U_b(X)$ and $g \in \Gamma$, and put

$$\varphi_\alpha(x) = e^{if(x)} \int_Y e^{ig(y)}\lambda_\alpha(x, dy) \text{ and } \varphi(x) = e^{if(x)} \int_Y e^{ig(y)}\lambda(x, dy).$$

Since Γ satisfies (Γ3), by assumptions (a) and (b) of Theorem 1, it is easily verified that $\{\varphi_\alpha\}$ and φ satisfy assumptions of Lemma 2. Therefore, by Lemma 2 we have

$$\lim_\alpha \int_X \varphi_\alpha(x)\mu_\alpha(dx) = \int_X \varphi(x)\mu(dx),$$

and this implies (3.6). Hence the proof of Theorem 1 is complete.

4. Uniform Equicontinuity of Gaussian Transition Probabilities

In this section we consider uniform equicontinuity of Gaussian transition probabilities on a Hilbert space. Let H be a real Hilbert space with norm $\|\cdot\|$ and inner product $\langle\cdot,\cdot\rangle$. A Borel probability measure μ on H is said to be *Gaussian* if for each $u \in H$, the function $w \in H \to \langle w, u\rangle$ is a (possibly degenerate) Gaussian random variable on the probability measure space $(H, \mathcal{B}(H), \mu)$. Since every Gaussian measure μ satisfies

$$\int_H \|w\|^2 \mu(dw) < \infty,$$

we can define a *mean vector* $m_\mu \in H$ and a *covariance operator* $S_\mu \in \mathcal{S}(H)$ by

$$\langle m_\mu, u\rangle = \int_H \langle w, u\rangle \mu(dw), \quad u \in H,$$

and

$$\langle S_\mu, v\rangle = \int_H \langle w - m_\mu, u\rangle\langle w - m_\mu, v\rangle \mu(dw), \quad u, v \in H.$$

Here $\mathcal{S}(H)$ denotes the set of all positive and symmetric trace class operators on H; it is endowed with the metric topology derived from the trace norm $\|\cdot\|_{tr}$. See e.g. [17], [11] and [12] for definitions, properties and related facts on trace class operators. Since a Gaussian measure μ is uniquely determined by its mean vector m_μ and covariance operator S_μ, we write $\mu = \mathcal{N}(m_\mu, S_\mu)$ (see e.g., Theorem IV.2.4 of [17]).

Let $(X, \mathcal{A})$ be a measurable space. A transition probability λ on $X \times H$ is said to be *Gaussian* if for each $x \in X$, $\lambda_x(\cdot) \equiv \lambda(x, \cdot)$ is a Gaussian measure on H. For a Gaussian transition probability λ on $X \times H$, we can define weakly measurable mappings $m_\lambda : X \to H$ and $S_\lambda : X \to \mathcal{S}(H)$ such that

$$\langle m_\lambda(x), u\rangle = \int_H \langle w, u\rangle \lambda(x, dw), \quad x \in X, \ u \in H, \tag{4.1}$$

and

$$\langle S_\lambda(x)u, v\rangle = \int_H \langle w - m_\lambda(x), u\rangle\langle w - m_\lambda(x), v\rangle \lambda(x, dw), \quad x \in X, \ u, v \in H. \tag{4.2}$$

See Hille and Phillips [7, page 74] for the definitions and results concerning the measurability of vector and operator valued functions. We note here that weak measurability and strong measurability of m_λ and S_λ are equivalent in our case since H is separable. The functions $m_\lambda : X \to H$ and $S_\lambda : X \to \mathcal{S}(H)$ defined by (4.1) and (4.2) are called the *mean function* and the *covariance function* of λ, respectively.

Conversely, if mappings $m_\lambda : X \to H$ and $S_\lambda : X \to \mathcal{S}(H)$ are weakly measurable and $\lambda_x = \mathcal{N}(m_\lambda(x), S_\lambda(x))$ for each $x \in X$, then $\lambda = \lambda_x(\cdot)$ is a transition probability on $X \times H$ by Proposition 1 of [10]. From this, together with the fact that a Gaussian measure is uniquely determined by its mean vector and covariance operator, it is readily verified that a Gaussian transition probability λ is also uniquely determined by its mean function m_λ and covariance function S_λ, and hence we write $\lambda = \mathcal{TN}[m_\lambda, S_\lambda]$.

Denote by $BL(H)$ the Banach space of all bounded, real-valued Lipschitz functions on H. We also denote by $U(X;H)$ (resp. $U(X;\mathcal{S}(H))$) the set of all uniformly continuous mappings from a uniform space X into H (resp. $\mathcal{S}(H)$) with the metric topology derived from the norm on H (resp. the trace norm on $\mathcal{S}(H)$).

In this section we show that the following theorems, which state that for Gaussian transition probabilities we can express assumptions of Theorem 1 in terms of the corresponding mean and covariance functions. We note here that every transition probability on $X \times H$ are τ-smooth since any probability measure on H is Radon, and hence τ-smooth.

Theorem 2. *Let X be a uniform space.*

(1) *Let $\lambda = \mathcal{TN}[m_\lambda, S_\lambda]$ be a Gaussian transition probability on $X \times H$. Then λ is $BL(H)$-uniformly continuous if $m_\lambda \in U(X;H)$ and $S_\lambda \in U(X;\mathcal{S}(H))$.*
(2) *Let Q be a set of Gaussian transition probabilities on $X \times H$ with $\lambda = \mathcal{TN}[m_\lambda, S_\lambda]$, $\lambda \in Q$. Then Q is $BL(H)$-uniformly equicontinuous if $\{m_\lambda\} \subset U(X;H)$ and $\{S_\lambda\} \subset U(X,\mathcal{S}(H))$ are uniformly equicontinuous.*

Theorem 3. *Let X be a uniform space. Let $\lambda_\alpha = \mathcal{TN}[m_\alpha, S_\alpha]$ be a net of Gaussian transition probabilities on $X \times H$ and $\lambda = \mathcal{TN}[m, S]$ a Gaussian transition probability on $X \times H$. Assume that the following conditions are satisfied:*

(a) *$\{m_\alpha\} \subset U(X;H)$ and $\{S_\alpha\} \subset U(X;\mathcal{S}(H))$ are uniformly equicontinuous.*
(b) *$m \in U(X;H)$ and $S \in U(X;\mathcal{S}(H))$.*
(c) *$\lim_\alpha \|m_\alpha(x) - m(x)\| = 0$ and $\lim_\alpha \|S_\alpha(x) - S(x)\|_{tr} = 0$ for each $x \in X$.*

Then for any net $\{\mu_\alpha\} \subset \mathcal{P}(X)$ converging weakly to $\mu \in \mathcal{P}_\tau(X)$, we have $\mu_\alpha \circ \lambda_\alpha \xrightarrow{w} \mu \circ \lambda$.

To prove theorems above we need some information on β-distance and L^2 Wasserstein distance between probability measures, and an inequality for covariance operators.

Let $\mu, \nu \in \mathcal{P}(H)$. The *$\beta$-distance* between μ and ν is defined by

$$\beta(\mu, \nu) = \sup\left\{ \left| \int_H g d(\mu - \nu) \right| : \|g\|_{BL} \leq 1 \right\}.$$

The *L^2 Wasserstein distance* between μ and ν is defined by

$$W_2(\mu, \nu) = \left\{ \inf \int_{H \times H} \|u - v\|^2 \gamma(du, dv) \right\}^{1/2},$$

where the infimum is taken over the family of all probability measures on $H \times H$ with marginals μ and ν respectively. See Dudley [4], and Givens and Shortt [6] for a discussion on these and other distances on probability measures. We collect some results needed in the sequel, whose proofs can be found in [6] and [5].

Theorem 4. (1) *Let $\mu, \nu \in \mathcal{P}(H)$. Then we have*

$$\beta(\mu,\nu) \leq W_2(\mu,\nu). \tag{4.3}$$

(2) *For any two Gaussian measures $\mu = \mathcal{N}[m_\mu, S_\mu]$ and $\nu = \mathcal{N}[m_\nu, S_\nu]$ on H, we have an explicit expression (Gelbrich [5])*

$$W_2(\mu,\nu)^2 = \|m_\mu - m_\nu\|^2 + tr(S_\mu + S_\nu - 2(S_\mu^{1/2} S_\nu S_\mu^{1/2})^{1/2}). \tag{4.4}$$

The following trace inequality for covariance operators will be used in proving uniform equicontinuity of Gaussian transition probabilities.

Lemma 6. *For any S, $T \in \mathcal{S}(H)$, we have*

$$tr(S+T-2(S^{1/2}TS^{1/2})^{1/2}) \leq tr((S^{1/2}-T^{1/2})^2) \leq \|S-T\|_{tr}. \tag{4.5}$$

Proof. We first prove the left-hand inequality in (4.5). Note that $S^{1/2}$ and $T^{1/2}$ are positive and symmetric operators of Hilbert-Schmidt class, and hence $S^{1/2}T^{1/2}$ and $T^{1/2}S^{1/2}$ are trace class operators. Then we have

$$S^{1/2}TS^{1/2} = S^{1/2}T^{1/2}T^{1/2}S^{1/2} = (T^{1/2}S^{1/2})^*(T^{1/2}S^{1/2}) = |T^{1/2}S^{1/2}|^2,$$

and this implies that

$$tr((S^{1/2}TS^{1/2})^{1/2}) = tr(|T^{1/2}S^{1/2}|) \geq |tr(T^{1/2}S^{1/2})| \geq tr(T^{1/2}S^{1/2}). \tag{4.6}$$

Since $tr(T^{1/2}S^{1/2}) = tr(S^{1/2}T^{1/2})$, by (4.6) we have

$$\begin{aligned} tr(S+T-2(S^{1/2}TS^{1/2})^{1/2}) &\leq tr(S) + tr(T) - 2tr(T^{1/2}S^{1/2}) \\ &= tr((S^{1/2}-T^{1/2})^2). \end{aligned}$$

Next we prove the right-hand inequality in (4.5). Put $A = S^{1/2} - T^{1/2}$ and $B = S^{1/2} + T^{1/2}$. Then A and B are symmetric operators and $AB + BA = 2(S-T)$. By positivity of $S^{1/2}$ and $T^{1/2}$, we have $B \geq \pm A$.

Since A is a symmetric operator of Hilbert-Schmidt class, it has an representation

$$Au = \sum_{i=1}^{\infty} \lambda_i \langle u, e_i \rangle e_i, \quad u \in H, \tag{4.7}$$

where $\{e_i\}$ is an orthonormal basis in H, $\{\lambda_i\}$ is a set of real numbers with $\sum_{i=1}^{\infty} \lambda_i^2 < \infty$, and the sum in (4.7) converges in norm. Then we have

$$\begin{aligned} \|S-T\|_{tr} &= tr(|S-T|) \\ &= \frac{1}{2}\sum_{i=1}^{\infty}\langle |AB+BA|e_i, e_i\rangle \end{aligned}$$

$$\geq \frac{1}{2}\sum_{i=1}^{\infty}|\langle (AB+BA)e_i, e_i\rangle| \quad (\text{since } |AB+BA| \geq \pm(AB+BA))$$
$$= \sum_{i=1}^{\infty}|\lambda_i\langle Be_i, e_i\rangle| \quad (\text{since } Ae_i = \lambda_i e_i \text{ for all } i)$$
$$\geq \sum_{i=1}^{\infty}|\lambda_i||\langle Ae_i, e_i\rangle| \quad (\text{since } B \geq \pm A)$$
$$\geq \sum_{i=1}^{\infty}\lambda_i^2 \quad (\text{since } Ae_i = \lambda_i e_i \text{ for all } i)$$
$$= \sum_{i=1}^{\infty}\langle A^2 e_i, e_i\rangle = tr(A^2) = tr((S^{1/2} - T^{1/2})^2),$$

and the proof is complete.

By Theorem 4 and Lemma 6, we have the following

Proposition 3. *For any two Gaussian measures $\mu = \mathcal{N}[m_\mu, S_\mu]$ and $\nu = \mathcal{N}[m_\nu, S_\nu]$, we have*

$$W_2(\mu,\nu)^2 \leq ||m_\mu - m_\nu||^2 + ||S_\mu - S_\nu||_{tr}.$$

Proof of Theorem 2. (1) follows from (2), and hence we only prove (2). Fix $\lambda \in Q$, $g \in BL(H)$ and $x_1, x_2 \in X$. Then by (1) of Theorem 4, Proposition 3 and definition of β-distance, we have

$$\begin{aligned}
&\left|\int_H g(y)\lambda(x_1,dy) - \int_H g(y)\lambda(x_2,dy)\right| \\
&\leq ||g||_{BL}\cdot\beta(\lambda(x_1,\cdot),\lambda(x_2,\cdot)) \\
&\leq ||g||_{BL}\cdot W_2(\lambda(x_1,\cdot),\lambda(x_2,\cdot)) \\
&\leq ||g||_{BL}\cdot\{||m_\lambda(x_1) - m_\lambda(x_2)||^2 + ||S_\lambda(x_1) - S_\lambda(x_2)||_{tr}\}^{1/2}.
\end{aligned}$$

From this it follows that Q is $BL(H)$-uniformly equicontinuous since $\{m_\lambda\}$ and $\{S_\lambda\}$ are uniformly equicontinuous by assumption of Theorem 2.

Proof of Theorem 3. Since $BL(H)$ satisfies (Γ1), (Γ2) and (Γ3) by (2) of Example 1, we put $\Gamma = BL(H)$ in Theorem 1. Then we have only to verify that $\{\lambda_\alpha\}$ and λ satisfy assumptions (a) and (b) of Theorem 1.

It follows from Theorem 2 and assumptions (a) and (b) of Theorem 3 that $\{\lambda_\alpha\}$ is Γ-uniformly equicontinuous and $\lambda \in U_\Gamma(X,H)$.

Let $x \in X$ and $g \in \Gamma = BL(H)$. Then by (1) of Theorem 4, Proposition 3 and the definition of β-distance, we have

$$\begin{aligned}
&\left|\int_H g(y)\lambda_\alpha(x,dy) - \int_H g(y)\lambda(x,dy)\right| \\
&\leq ||g||_{BL}\cdot\beta(\lambda_\alpha(x,\cdot),\lambda(x,\cdot))
\end{aligned}$$

$$\leq \|g\|_{BL} \cdot W_2(\lambda_\alpha(x,\cdot), \lambda(x,\cdot))$$
$$\leq \|g\|_{BL} \cdot \{\|m_\alpha(x) - m(x)\|^2 + \|S_\alpha(x) - S(x)\|_{tr}\}^{1/2}.$$

From this and assumption (c) of Theorem 3 it follows that $\lambda_\alpha(x,\cdot) \xrightarrow{\Gamma} \lambda(x,\cdot)$ for each $x \in X$. Consequently, $\{\lambda_\alpha\}$ and λ satisfy assumptions (a) and (b) of Theorem 1, and the proof of Theorem 3 is complete.

Acknowledgements

The author would like to thank the referee for helpful comments, which led me to consider uniform equicontinuity of Gaussian transition probabilities on Hilbert spaces.

References

[1] P. Billingsley, Convergence of Probability Measures, Wiley, New York, 1968.
[2] I. Csiszár, *On the weak* continuity of convolution in a convolution algebra over an arbitrary topological group*, Studia Sci. Math. Hungar., **6**(1971), 27-40.
[3] M. D. Donsker and S. R. S. Varadhan, *Asymptotic evaluation of certain Markov process expectations for large time I*, Comm. Pure Appl. Math., **28**(1975), 1-47.
[4] R. M. Dudley, Real Analysis and Probability, Wadsworth & Brooks/Cole, California, 1989.
[5] M. Gelbrich, *On a formula for the L^2 Wasserstein metric between measures on Euclidean and Hilbert spaces*, Math. Nachr., **147**(1990), 185-203.
[6] C. R. Givens and R. M. Shortt, *A class of Wasserstein metrics for probability distances*, Michigan Math. J., **31**(1984), 231-240.
[7] E. Hille and R. S. Phillips, Functional Analysis and Semi-Groups, Amer. Math. Soc., Providence, Rhode Island, 1957.
[8] I. M. James, Topological and Uniform Spaces, Springer-Verlag, New York, 1987.
[9] J. Kawabe, *Convergence of compound probability measures on topological spaces*, Colloq. Math., **67**(1994), 161-176.
[10] ——, *Uniform tightness for transition probabilities on topological spaces*, Tamkang J. Math., **26**(1995), 283-298.
[11] M. Reed and B. Simon, Methods of Modern Mathematical Physics Vol.1: Functional Analysis, Academic Press, New York, 1980.
[12] R. Schatten, Norm Ideals of Completely Continuous Operators, Springer-Verlag, New York, 1970.
[13] L. Schwartz, Radon Measures on Arbitrary Topological Spaces and Cylindrical Measures, Oxford University Press, 1973.
[14] F. Topsøe, *Compactness in spaces of measures*, Studia Math., **36**(1970), 195-212.
[15] ——, Topology and Measure, Lecture Notes in Math. 133, Springer-Verlag, New York, 1970.
[16] H. Umegaki, *Representations and extremal properties of averaging operators and their applications to information channels*, J. Math. Anal. Appl., **25**(1969), 41-73.

[17] N. N. Vakhania, V. I. Tarieladze and S. A. Chobanyan, Probability Distributions on Banach Spaces, D. Reidel Publishing Company, 1987.

Department of Mathematics, Faculty of Engineering, Shinshu University, 4-17-1 Wakasato, Nagano 380-8553, Japan.

E-mail: jkawabe@gipwc.shinshu-u.ac.jp