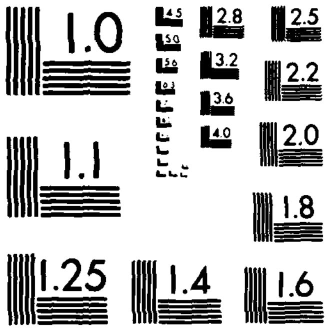
AD-A170 222 ON AVERAGE MUTUAL INFORMATION FOR SPHERICALLY INVARIANT 1/1
MEASURES AND SOME (U) NORTH CAROLINA UNIV AT CHAPEL
HILL DEPT OF STATISTICS A F GUALTIEROTTI MAR 85

UNCLASSIFIED N00014-84-C-0212 F/G 9/4 NL


1.0
1.1
1.25
1.4
1.6
1.8
2.0
2.2
2.5
2.8
3.2
3.6
4.0

# On Average Mutual Information for Spherically Invariant Measures and Some Other Mixtures of Gaussian Measures(*)

A.F. Gualtierotti

IDHEAP, BFSH 1
University of Lausanne
CH 1015 Lausanne
Switzerland

and

Department of Statistics
University of North Carolina
Chapel Hill, N.C. 27514

LISS 16
March, 1986



(*) Research supported by ONR Contract N00014-84-C-0212

**6a. NAME OF PERFORMING ORGANIZATION**
Department of Statistics

**6b. OFFICE SYMBOL** *(If applicable)*

**7a. NAME OF MONITORING ORGANIZATION**

**6c. ADDRESS** *(City, State and ZIP Code)*
University of North Carolina
Chapel Hill, North Carolina 27514

**7b. ADDRESS** *(City, State and ZIP Code)*

**8a. NAME OF FUNDING/SPONSORING ORGANIZATION**
Office of Naval Research

**8b. OFFICE SYMBOL** *(If applicable)*

**9. PROCUREMENT INSTRUMENT IDENTIFICATION NUMBER**
N00014-84-C-0212

**8c. ADDRESS** *(City, State and ZIP Code)*
Statistics & Probability Program
Arlington, VA 22217

**10. SOURCE OF FUNDING NOS.**

| PROGRAM ELEMENT NO. | PROJECT NO. | TASK NO. | WORK UNIT NO. |
|---|---|---|---|
| NR | 042 | 269 | SRO 105 |

**11. TITLE** *(Include Security Classification)*
On Average Mutual Information for Spherically . . .

**12. PERSONAL AUTHOR(S)**
A.F. Gualtierotti

**13a. TYPE OF REPORT**
TECHNICAL

**13b. TIME COVERED**
FROM ________ TO ________

**14. DATE OF REPORT** *(Yr., Mo., Day)*
March, 1986

**15. PAGE COUNT**
29

**16. SUPPLEMENTARY NOTATION**

**17. COSATI CODES**

| FIELD | GROUP | SUB. GR. |
|---|---|---|
| | | |
| | | |

**18. SUBJECT TERMS** *(Continue on reverse if necessary and identify by block number)*
information theory; channel capacity; nonGaussian stochastic processes; Shannon theory

**19. ABSTRACT** *(Continue on reverse if necessary and identify by block number)*

Average mutual information is studied for nonGaussian measures of two types: spherically invariant and Gaussian mixtures (a generalization of spherically-invariant). A complete answer is obtained for finite mixtures. These results are applicable to communication channels perturbed by nonGaussian noise processes described by these measures.

TITLE CONT.: Invariant Measures and Some Other Mixtures of Gaussian Measures.

**20. DISTRIBUTION/AVAILABILITY OF ABSTRACT**
UNCLASSIFIED/UNLIMITED ☒ SAME AS RPT. ☒ DTIC USERS ☐

**21. ABSTRACT SECURITY CLASSIFICATION**

**22a. NAME OF RESPONSIBLE INDIVIDUAL**
C.R. Baker

**22b. TELEPHONE NUMBER** *(Include Area Code)*
(919) 962-2189

**22c. OFFICE SYMBOL**

# ABSTRACT

Average mutual information is studied for nonGaussian measures of two types: spherically invariant and Gaussian mixtures (a generalization of spherically-invariant). A complete answer is obtained for finite mixtures. These results are applicable to communication channels perturbed by nonGaussian noise processes described by these measures.

## On average mutual information for spherically invariant measures and some other mixtures of Gaussian measures

### 1. Introduction

The actual transmission capacity of a given channel is a parameter of basic importance in any communication system since it limits the rate at which information can be transmitted reliably. There has thus been an effort which started with Shannon (1948) to compute the capacity of transmission for different channel and transmission models. In the case of a continuous channel, most results have been obtained for a Gaussian noise (Baker (1978); Hitsuda and Ihara (1975); Kadota, Zakai and Ziv (1971)). Some attempts to steer away from the Gaussian case have also been made (Gualtierotti (1980)) and these indicate that new methods may be required. Indeed, in the Gaussian case, most quantities of interest can be explicitely obtained, whereas these computations are almost always impossible in other instances. Furthermore the computation of mutual information requires that the joint law of the message and the received signal be absolutely continuous with respect to the product of the marginals and that the Radon-Nikodym derivative be computed : though the Gaussian case is well known (Baker (1973)), this knowledge is again unavailable for most other models.

Spherically invariant distributions, when the mean is zero, are mixtures of Gaussian ones and through mixing a number of well known distributions can be obtained, such as the double exponential and the Student distributions (Keilson and Steutel, (1974)). There is also some evidence that some real life noises can be described through spherically invariant probabilities, particularly in underwater acoustics. It is thus natural to investigate the problems of absolute continuity, of calculation of mutual information and channel capacity for spherically invariant noises and mixtures of Gaussian probabilities.

This is the subject of the present paper. Spherically invariant distributions are abbreviated as "SIM's" and Gaussian mixtures as "GMM's". We start by studying admissible translates of SIM's and GMM's for two reasons : the first is to assess the difference betwen the two cases and the second is that the computation of channel capacity can in certain cases be achieved by a reduction to a family of translates (Baker (1979)). The conclusion is that admissible translates for GMM's are harder to come by than for SIM's : the mean of the Gaussian measure being mixed must belong to the range of the square root of its covariance operator. We then consider the equivalence problem for certain Gaussian measures when one of these is a product of marginals : they arise when one studies the equivalence problem for joint SIM's and GMM's with respect to the product of their marginals. It turns out that joint measures and products of marginals are most often orthogonal, so that one is lead to believe that joint SIM's and GMM's must be orthogonal to the product of their marginals. This is indeed the case when the mixing function is smooth. So one is left with the problem of finding an "explanation" for this orthogonality. By restricting attention to a finite mixture, one can find a formula for the average mutual information : it is given by the average mutual information of the Gaussian measure being mixed to which the entropy of the mixing function is added. The latter can be made as large as one wishes by having the discrete mixing distribution approach a continuous one. The framework within which the problem is addressed may be found in Baker (1978).

## 2. Notation, definitions and useful results

*2.1.*

$H_1$ and $H_2$ are real and separable Hilbert spaces with respective inner products $< h_1^1, h_2^1 >_1$ and $< h_1^2, h_2^2 >_2$. $H$ is $H_1 \times H_2$ and has elements $\vec{h} = (h^1, h^2)$. The inner product

$$< \vec{h}_1, \vec{h}_2 > \quad = \quad < h_1^1, h_2^1 >_1 \quad + \quad < h_1^2, h_2^2 >_2$$

makes $H$ into a real and separable Hilbert space. The Borel sets of $H, H_1, H_2$ are connected by the equality

$$\mathcal{B}[H] = \mathcal{B}[H_1] \otimes \mathcal{B}[H_2].$$

(Partharasarathy (1967) : p.6, Theorem 1.10.).

Let $p_1$ and $p_2$ be the projections of $H$ with respective ranges $H_1 \times \{\theta_2\}$ and $\{\theta_1\} \times H_2$, $J_1$ and $J_2$ be the maps $J_1(h^1, h^2) = h^1$ and $J_2(h^1, h^2) = h^2$. From "first principles" one has :

$$J_1^* h^1 = (h^1, \theta_2), \quad J_2^* h^2 = (\theta_1, h^2),$$
$$J_1 J_1^* = id_{H_1}, \quad J_2 J_2^* = id_{H_2}, \quad J_1^* J_1 = p_1, \quad J_2^* J_2 = p_2,$$
$$J_1 J_2^* = 0, \quad J_2 J_1^* = 0, \quad \text{and } p_1 + p_2 = id_H.$$

*2.2.*

Let $P$ be a probability measure on $\mathcal{B}[H]$. $P^1 = P \circ J_1^{-1}$ and $P^2 = P \circ J_2^{-1}$ are the marginals of $P$, defined on $\mathcal{B}[H_1]$ and $\mathcal{B}[H_2]$ respectively, and $P^{\otimes}$ stands for $P^1 \otimes P^2$.

*2.3.*

Let $K$ be a real and separable Hilbert space. A Gaussian probability measure $Q$, defined on $\mathcal{B}[K]$, is denoted $Q \sim N_K(m_Q, R_Q)$.

For $P \sim N_H(\vec{m}_P, R_P)$ write :

$$R_P^1 = J_1\, R_P\, J_1^*,\;\; R_P^2 = J_2\, R_P\, J_2^*,\;\; R_P^{\otimes} = p_1 R_P p_1 + p_2 R_P p_2.$$

Then (Baker (1973) p.280 : Proposition 2) :

$$P^i \sim N_{H_i}(m_P^i, R_P^i),\;\; P^{\otimes} \sim N_H(\vec{m}_P, R_P^{\otimes}),$$

and $\mathcal{R}\{(R_P)^{1/2}\} \subseteq \mathcal{R}(\sqrt{R_P^{\otimes}})$ where $\mathcal{R}(T)$ denotes the range of the operator $T$. Furthermore, $\vec{m}_P \in \mathcal{R}\{(R_P)^{1/2}\}$ implies

$$(\alpha m_P^1, \beta m_P^2) \in \mathcal{R}(\sqrt{R_P^{\otimes}}), \quad \alpha \geq 0, \quad \beta \geq 0, \text{ and } \vec{m}_P \in \mathcal{R}(\sqrt{R_P^{\otimes}})$$

if and only if $m_P^1 \in \mathcal{R}(\{(R_P^1)^{1/2}\})$ and $m_P^2 \in \mathcal{R}(\{(R_P^2)^{1/2}\})$. Indeed,

$$< (\beta m_P^1, \gamma m_P^2), \vec{h} >^2 \leq (\beta+\gamma)^2 \{< m_P^1, h^1 >_1^2 + < m_P^2, h^2 >_2^2\}$$

and $< \vec{m}_P, \vec{h} >^2 \leq C\cdot < R_P \vec{h}, \vec{h} >$ (Baker (1970) : p.5, Corollary 2,e)). So, choosing for $\vec{h}$ the element $J_1^* h^1$ for example, one has

$$< m_P^1, h^1 >_1^2 \leq C\cdot < R_P^1 h^1, h^1 >_1 .$$

Finally, $< R_P^{\otimes} \vec{h}, \vec{h} > = < R_P^1 h^1, h^1 >_1 + < R_P^2 h^2, h^2 >_2$ (Baker (1973), p.280 ). Thus :

$$< (\beta m_P^1, \gamma m_P^2), \vec{h} >^2 \leq C(\beta+\gamma)^2 < R_P^{\otimes} \vec{h}, \vec{h} >,$$

which is enough to prove the first claim (Baker (1970) : p.5, Corollary 2,e). The second is proved similarly. For example,

$$< \vec{m}_P, \vec{h} >^2 \leq 2\{< m_P^1, h^1 >_1^2 + < m_P^2, h^2 >_2^2\}$$
$$\leq 2max\{C_1, C_2\}(< R_P^1 h^1, h^1 >_1 + < R_P^2 h^2, h^2 >_2).$$

*2.4. Gaussian mixture measures (GMM's)*

Let $Q \sim N_K(m_Q, R_Q)$ and $T_{K,\alpha}k = \alpha k, k \in K, \alpha \geq 0$. Set $Q_\alpha = Q \circ T_{K,\alpha}^{-1}$ and $q_Q(\alpha, B) = Q_\alpha(B)$, $B \in \mathcal{B}[K]$, $\alpha \geq 0$. Then, through the explicit form of a Gaussian measure (Parthasarathy (1967) : p.179, Theorem 4.9) and Fubini's theorem, one has that

$Q_\alpha \sim N_K(\alpha m_Q, \alpha^2 R_Q)$ and $q_Q$ is a transition function on $\mathbb{R}_+ \times \mathcal{B}[K]$. Let $F$ be a probability on $\mathcal{B}[\mathbb{R}_+]$. A GMM is a probability of the form :

$$Q_F(B) = \int_0^\infty q_Q(\alpha, B)F(d\alpha), B \in \mathcal{B}[K]$$

*2.5. Spherically invariant measures (SIM's)*

$Q$ and $T_{K,\alpha}$ are as in 2.4. Furthermore $\mathring{Q}_\alpha \sim N_K(m_Q, \alpha^2 R_Q)$ and $\mathring{q}_Q(\alpha, B) = \mathring{Q}_\alpha(B)$, $B \in \mathcal{B}[K]$ , $\alpha \geq 0$. $\mathring{q}_Q$ is again a transition function and a SIM is a probability of the form :

$$\mathring{Q}_F(B) = \int_0^\infty \mathring{q}_Q(\alpha, B)F(d\alpha), B \in \mathcal{B}[K].$$

*2.6. Second order properties of GMM's and SIM's*

Let $A$ be a random variable with law $F$. One may check that, provided $EA^2$ exists,

i) $Q_F$ has mean $EA \cdot m_Q$ and covariance $EA^2 \cdot R_Q + VA \cdot m_Q \otimes m_Q$,

ii) $\mathring{Q}_F$ has mean $m_Q$ and covariance $EA^2 \cdot R_Q$.

Furthermore, $\mathcal{R}\{(R_Q)^{1/2}\} \subseteq \mathcal{R}\{(R_{Q_F})^{1/2}\}$ , with equality if and only if $m_Q \in \mathcal{R}\{(R_Q)^{1/2}\}$. This may be seen as follows. The square root of $m_Q \otimes m_Q$ is $\{m_Q/\sqrt{\|m_Q\|_K}\} \otimes \{m_Q/\sqrt{\|m_Q\|_K}\}$ , so that $\mathcal{R}\{(R_{Q_F})^{1/2}\}$ is the linear manifold generated by $\mathcal{R}\{(R_Q)^{1/2}\}$ and $m_Q$ (Baker (1970) : p.6, Corollary 2; Sytaya (1969) : p.507, Lemma 2) .

*2.7. GMM's and SIM's on product spaces*

Let $P \sim N_H(\vec{m}_P, R_P)$ and $F$ be a probability on $B[\mathbb{R}_+]$ . $P_F$ and $\dot{P}_F$ are defined as in 2.4 and 2.5 respectively. Using again the explicit form of the characteristic function of a Gaussian measure (Parthasarathy (1967) : p.179, Theorem 4.9), one has that :

i) $(P \circ T_{H,\alpha}^{-1})^i = P^i \circ T_{H_i,\alpha}^{-1}$, the result begin written $P_\alpha^i$,

ii) $P_\alpha^1 \otimes P_\alpha^2 = P^\otimes \circ T_{H,\alpha}^{-1}$, the result begin written $P_\alpha^\otimes$ .

The following notation shall be used :

$$\dot{P}_\alpha^i = \dot{P}_\alpha \circ J_i^{-1},\ \dot{P}_\alpha^\otimes = \dot{P}_\alpha^1 \otimes \dot{P}_\alpha^2,$$

$$P_{\alpha,\beta}^\otimes = P_\alpha^1 \otimes P_\beta^2,\ \dot{P}_{\alpha,\beta}^\otimes = \dot{P}_\alpha^1 \otimes \dot{P}_\beta^2,$$
$$R_{\alpha,\beta}^\otimes = \alpha^2 p_1 R_P p_1 + \beta^2 p_2 R_P p_2.$$

Then one has also :

*j)* $$P_{\alpha,\beta}^\otimes \sim N_H((\alpha m_P^1, \beta m_P^2), R_{\alpha,\beta}^\otimes),$$

*jj)* $$\dot{P}_{\alpha,\beta}^\otimes \sim N_H(\vec{m}_P, R_{\alpha,\beta}^\otimes),$$

*k)* $$P_F^i(B_i) = \int_0^\infty P_\alpha^i(B_i)F(d\alpha),\ B_i \in B[H_i],$$

kk) $$\hat{P}^i_F(B_i) = \int_0^\infty \hat{P}^i_\alpha(B_i)F(d\alpha),\ B_i \in B[H_i],$$

l) $$P^{\otimes}_F(B) = \int_0^\infty \int_0^\infty P^{\otimes}_{\alpha,\beta}(B)F \otimes F(d\alpha,\ d\beta), B \in B[H],$$

ll) $$\hat{P}^{\otimes}_F(B) = \int_0^\infty \int_0^\infty \hat{P}^{\otimes}_{\alpha,\beta}(B)F \otimes F(d\alpha,\ d\beta), B \in B[\hat{H}].$$

*2.8. Translates of GMM's and SIM's*

Let $Q \sim N_K(m_Q,\ R_Q)$ and $S_{K,a}k = k + a,\ k \in K,\ a \in K$. Then one writes $Q^a_\alpha$ for $(Q \circ T^{-1}_{K,\alpha}) \circ S^{-1}_{K,a} = (Q \circ S^{-1}_{K,a}) \circ T^{-1}_{K,\alpha}$. One has :

$$Q^a_F(B) = (Q_F \circ S^{-1}_{K,a})(B) = \int_0^\infty Q^a_\alpha(B)F(d\alpha),\ B \in \mathcal{B}[K].$$

## 3. Equivalence and singularity results

### *3.1. The case of translates of GMM's and SIM's*

a) Let $Q \sim N_K(m_Q, R_Q)$ and $R_Q$ have the representation $R_Q = \sum_i \rho_i e_i \otimes e_i$, where $R_Q e_i = \rho_i e_i, \rho_i > 0$, the $e_i$'s are orthonormal and $\sum_i \rho_i < \infty$. Set :

$$\psi_n(k) = (1/n)\sum_{i=1}^{n}\{< k, e_i >_K^2 / \rho_i\},$$

$$\psi(k) = \limsup_n \psi_n(k).$$

Then,

i) if $m_Q \in \mathcal{R}(\sqrt{R_Q})$,

$$Q\alpha\{k \in K : \psi(k) = \alpha^2\} = \dot{Q}\alpha\{k \in K : \psi(k) = \alpha^2\} = 1,$$

ii) if $m_Q \notin \mathcal{R}(\sqrt{R_Q})$, nothing can be generally asserted about the behaviour of $\psi_n$.

Indeed :

i) Let $\lambda_i =< m_Q, e_i >_K / \sqrt{\rho_i}$ and $X_i(k) =< k - \alpha m_Q, e_i >_K / \sqrt{\rho_i}$. Then:

$$\psi_n(k) = \psi_n^1(k) + \psi_n^2(k) + \psi_n^3(k),$$

$$\psi_n^1(k) = (1/n)\sum_{i=1}^{n} X_i^2(k),$$

$$\psi_n^2(k) = 2\alpha(1/n)\sum_{i=1}^{n}\lambda_i X_i,$$

$$\psi_n^3(k) = \alpha^2(1/n)\sum_{i=1}^{n}\lambda_i^2.$$

Let $p_n = \sum_{i=1}^{n} e_i \otimes e_i$. Then

$$\sum_{i=1}^{n} \lambda_i e_i = (\sqrt{R_Q})^{-1} p_n m_Q,$$

so that

$$\sum_{i=1}^{n} \lambda_i^2 = \|(\sqrt{R_Q})^{-1} p_n m_Q\|_K^2.$$

Since $m_Q \in \mathcal{R}(\sqrt{R_Q})$, $\|(\sqrt{R_Q})^{-1} m_Q\|_K^2 < \infty$, so that $\sum_{i=1}^{\infty} \lambda_i^2 < \infty$ and $\lim_n \psi_n^3(k) = 0$.

With respect to $Q_N$, $\{Y_i(k) = \lambda_i X_i(k)\}$ is a sequence of independent random variables, each of which has mean equal to zero. Furthermore,

$$VY_i = \lambda_i^2 VX_i = \alpha^2 \lambda_i^2, \text{ so that } \sum_{i=1}^{\infty} (VY_i/i^2) < \infty.$$

Thus $\lim_n \psi_n^2(k) = 0, Q_\alpha - a.s.$ (Neveu (1965): p.146, Proposition IV.6.1).

Finally, $\lim_n \psi_n^1(k) = \alpha^2, Q_\alpha - a.s.$ by the law of large numbers.

ii) Let now $K$ be $l_2$, $\vec{e}_i$ be that element of $l_2$ which has components equal to zero, except for the i-th, which is equal to one. $\vec{m}_Q$ has components $m_i$ and $\vec{m}_Q \notin \mathcal{R}(\sqrt{R_Q})$ means $\sum_{i=1}^{\infty} (m_i^2/\rho_i) = \infty$. Choose for $m_i$ the value $\sqrt{\rho_i/i}$. Then $\vec{m}_Q \notin \mathcal{R}(\sqrt{R_Q})$. However,

$$\sum_{i=1}^{n} \lambda_i^2 = \sum_{i=1}^{n} (1/i) = \log n + C + 0(1)$$

(Pisot and Zamansky (1959) : p.523), so that

$\lim_n \psi_n^3(k) = 0$. Now $(VY_i/i^2) = \alpha^2/i^3$, which has a convergent sum, and thus

$\lim_n \psi_n^2(k) = 0, Q_\alpha - a.s.$. So, $\psi(k) = \alpha^2, Q_\alpha - a.s.$. Choose now for $m_i$ the value $\beta\sqrt{\rho_i}$. Since

$$\sum_{i=1}^{n}(m_i^2/\rho_i) = \beta^2 n, \ \tilde{m}_Q \notin \mathcal{R}(\sqrt{R_Q}).$$

Furthermore, $\lim_n \psi_n^3(k) = \alpha^2\beta^2$. Now $(VY_i/i^2) = \alpha^2\beta^2/i^2$, which has still a convergent sum, so that $\lim_n \psi_n^2(k) = 0$. Consequently, $\psi(k) = \alpha^2(1+\beta^2)$. Finally, if $m_i$ has the value $\sqrt{\rho_i \cdot i^\beta}$ , $0 < \beta < 1$,

$$\sum_{i=1}^{n}(m_i^2/\rho_i) = \sum_{i=1}^{n} i^\beta \geq n^{\beta+1}/(\beta+1),$$

so that $\lim_n \psi_n^3(k) = \infty$. But $(VY_i/i^2) = \alpha^2/i^{2-\beta}$, which is the term of a convergent series, so that, still, $\lim_n \psi_n^2(k) = 0$.

Consequently, $\lim_n \psi_n(k) = \infty, k \in K$.

b) $Q_F^a \perp Q_F$ or $\dot{Q}_F^a \perp \dot{Q}_F$ implies $a \notin \mathcal{R}\{(R_Q)^{1/2}\}$.

Indeed :

Let $B \in \mathcal{B}[K]$ be such that $Q_F^a(B) = Q_F(B^c) = 0$. Define :

$$C = \{\alpha \geq 0 : Q_\alpha(B^c) = 0\}, C_a = \{\alpha \geq 0 : Q_\alpha^a(B) = 0\}.$$

$C$ and $C_a$ may be assumed to be measurable. Thus, since $F(C) = F(C_a) = 1$, $F(C \cap C_a) = 1$ and consequently $C \cap C_a \neq \emptyset$. So there is some $\alpha$ for which $Q_\alpha^a \perp Q_\alpha$. Now, since $Q_\alpha \sim N_K(\alpha m_Q, \alpha^2 R_Q)$ and

$Q_\alpha^a \sim N_K(a + \alpha m_Q, \alpha^2 R_Q)$, one must have $a \notin \mathcal{R}(\sqrt{R_Q})$ (Rao and Varadarajan (1968) : p.305, Remark; p.308, Theorem 4.1; p.312, Theorem 5.1).

*c) i)* $a \notin \mathcal{R}(\sqrt{R_Q})$ implies $\dot{Q}_F^a \perp \dot{Q}_F$

*ii)* $a \notin \mathcal{R}(\sqrt{R_Q})$ and $m_Q \in \mathcal{R}(\sqrt{R_Q})$ imply $Q_F^a \perp Q_F$

Indeed :

Let $X_n(k) = \langle k, k_n \rangle_K$ and $B_n = \{k \in K : X_n(k) > c_n\}$. Then : $Q_F(B_n) = \int_0^\infty Q_\alpha(B_n)F(d\alpha)$ and, with respect to $Q_\alpha$,

$$X_n \sim N_K(\alpha \langle m_Q, k_n \rangle_K, \alpha^2 \langle R_Q\, k_n, k_n \rangle_K).$$

Thus, for $\alpha > 0$ and $\langle R_Q\, k_n, k_n \rangle_K \neq 0$, $Q_\alpha(B_n) = 1 - \phi(s_n)$, where $\phi$ is the distribution fonction of a random variable which is distributed $N(0,1)$, and $s_n = (c_n - \alpha \langle m_Q, k_n \rangle_K)/\alpha\|\sqrt{R_Q}k_n\|_K$. Similarly, $Q_\alpha^a(B_n) = 1 - \phi(t_n)$, where

$$t_n = (c_n - \alpha \langle m_Q, k_n \rangle_K - \langle a, k_n \rangle_K)/\alpha\|\sqrt{R_Q}k_n\|_K.$$

So, if $c_n$ can be chosen so that $\lim_n s_n = \infty$ and $\lim_n t_n = -\infty$, one has $Q_F^a \perp Q_F$. For SIM's, one need only drop the $\alpha$ in front of $\langle m_Q, k_n \rangle_K$. The corresponding constants shall be denoted $\dot{s}_n$ and $\dot{t}_n$.

The assumption $m_Q \in \mathcal{R}(\sqrt{R_Q})$ implies that

$$\langle m_Q, k_n \rangle_K / \|\sqrt{R_Q}k_n\|_K$$

is bounded, and $a \notin \mathcal{R}(\sqrt{R_Q})$ implies that $k_n$ can be so chosen that $\langle a, k_n \rangle_K > d_n \cdot \langle R_Q\, k_n, k_n \rangle_K$, with $d_n$ arbitrarily large and $\|k_n\|_K = 1$ (Baker (1970) : p.5, Corollary 2,e)). Thus, choosing for GMM's, $c_n = \langle a, k_n \rangle_K /2$, and, for SIM's,

$$c_n = (\langle a, k_n \rangle_K /2) + \langle m_Q, k_n \rangle_K,$$

one has $\lim_n s_n = \lim_n \hat{s}_n = \infty$ and $\lim_n t_n = \lim_n \hat{t}_n = -\infty$.

So, one is left to consider the cases $\alpha = 0$ and $< R_Q k_n, k_n >_K = 0$, some n. Suppose thus that $k_{n_0} = k_0$ is such that $< R_Q k_0, k_0 >_K = 0$. Let $X(h) =< k, k_0 >_K$. Since $m_Q \in \mathcal{R}(\sqrt{R_Q})$ implies $< m_Q, k_0 >_K = 0$ (Baker (1970) : p.5, Corollary 2,e)), $X$ is, with respect to $Q_\alpha$, concentrated at 0, and, for the same reason, at $< a, k_0 >_K > 0$, with respect to $Q_\alpha^a$. In the case of SIM's the respective points are $< m_Q, k_0 >_K$ and

$< m_Q, k_0 >_K + < a, k_0 >_K$. Thus

$$Q_F(X = 0) = \hat{Q}_F(X =< m_Q, k_0 >_K) = 1$$

and

$$Q_F^a(X = 0) = \hat{Q}_F^a(X =< m_Q, k_0 >_K) = 0.$$

Now, if $\alpha = 0$, $X_n$, with respect to $Q_0$ is concentrated at 0, and, with respect to $Q_0^a$, at $< a, k_n >_K$. Thus, since

$< a, k_n >_K > 0$, $Q_0(X_n > c_n) = 0$ and $Q_0^a(X_n > c_n) = 1$. In the SIM's case, the values are, respectively,

$< m_Q, k_0 >_K$ and $< m_Q, k_n > + < a, k_n >_K$, with $< m_Q, k_n >$ bounded, and the same argument applies.

**Remark 1:** The argument used fails in the case of GMM's for which $m_Q \notin \mathcal{R}(\sqrt{R_Q})$ . In fact, if $k$ is such that $< R_Q k, k >_K = 0$, $< a, k >_K > 0$, and $m_Q = a$, let $T_a$ be the map $T_a \alpha = \alpha a$, $\alpha \in \mathbb{R}_+$ and $S$ be the map $S\alpha = \alpha + 1$, $\alpha \in \mathbb{R}_+$. Then, $Q_F(B) = F \circ T_a^{-1}(B)$ and $Q_F^a(B) = F \circ S^{-1} \circ T_a^{-1}(B)$. Thus $Q_F^a \perp Q_F$ implies $F \perp F \circ S^{-1}$. But the distribution function of $F \circ S^{-1}$ is that of $F$ shifted to the right by 1, and in general, the supports of $F$ and $F \circ S^{-1}$ overlap. This also shows that $Q_F^a \equiv Q_F$ does not obtain either.

The argument used in proving singularity can be found in Pan ((1973): p.12) where it is applied to the Gaussian case. Its application to SIM's induced on $I\!R^I$ by real processes is to be found in Huang and Cambanis ((1979) : Part 3.). The zero mean SIM case has been addressed in Sytaya ((1969) : p.508, Theorem 1).

d) If $a \in \mathcal{R}\{(R_Q)^{1/2}\}$ and $F$ has no mass at the origin, $Q_F^a \equiv Q_F$ and $\dot{Q}_F^a \equiv Q_F$

For $\alpha > 0$, one has $Q_\alpha^a \equiv Q_\alpha$ and $\dot{Q}_\alpha^a \equiv Q_\alpha$ (Rao and Varadarajan (1968): p.312; Theorem 5.1). For $\alpha = 0$, $Q_\alpha^a$ is concentrated at $a$ and $Q_\alpha$ at 0, $\dot{Q}_F^a$ at $a + m_Q$ and $\dot{Q}_F$ at $m_Q$, so that the assumption on $F$ is required.

e) $\dot{Q}_F^a \equiv \dot{Q}_F$ implies $a \in \mathcal{R}\{(R_Q)^{1/2}\}$ and both $Q_F^a \equiv Q_F$ and $m \in \mathcal{R}\{(R_Q)^{1/2}\}$ imply $a \in \mathcal{R}\{(R_Q)^{1/2}\}$

f) Using a) and (Skorohod (1974) : p.99, Theorem 2), one has :

i) if $a$ and $m_Q$ belong to $\mathcal{R}\{(R_Q)^{1/2}\}$, then

$$[dQ_F^a/dQ_F](x) =$$
$$exp\{[\psi(x)]^{-1} \lim_n \sum_{i=1}^{n} \rho_i^{-1} < x - \sqrt{\psi(x)}m_Q, e_i >_K < a_i, e_i >_K$$
$$- \|(\sqrt{R_Q})^{-1} a\|_K^2/2\},$$

ii) if furthermore a belongs to $\mathcal{R}(R_Q)$, then

$$[dQ_F^a/dQ_F](x) = exp\{[\psi(x)]^{-1} < x - \sqrt{\psi(x)}m_Q,\ R_Q^{-1}a >_K$$
$$- \|\{(R_Q)^{1/2}\}^{-1} a\|_K^2/2\}.$$

The formulae for SIM's are obtained by dropping the $\sqrt{\psi(x)}$ in front of $m_Q$.

**Remark 2:** The latter can be found, for the zero mean case, in Sytaya ((1969): p.509, Theorem 2).

*3.2. The case of joint measures and measures which are products of marginals : the Gaussian case*

The notation is that of 2.3, 2.4 and 2.5. The basic assumption is that $P \sim N_H(\hat{m}_P, R_P)$.

a) If $|\alpha - \gamma| > 0$ or $|\beta - \delta| > 0$, $P^{\otimes}_{\alpha,\beta} \perp P^{\otimes}_{\gamma,\delta}$. Consequently, $P^{\otimes}_{\alpha,\beta} \equiv P^{\otimes}_{\gamma,\delta}$ implies $\alpha = \gamma$ and $\beta = \delta$

Indeed :

If $Q \sim N_K(m_Q, R_Q)$ and $|\alpha - \beta| > 0$, $Q_\alpha \perp Q_\beta$, for then

$$R_{Q_\alpha} = \sqrt{R_{Q_\beta}}(id_K + T)\sqrt{R_{Q_\beta}} \text{ and } T = [(\alpha^2/\beta^2) - 1]id_K,$$

which is not Hilbert-Schmidt (Rao and Varadarajan (1968) : p.312, Theorem 5.1). So, in case $|\alpha - \gamma| > 0$ but $\beta = \delta$, choose $B = A \times H_2$, with $A$ such that $P^1_\alpha(A) = P^1_\gamma(A^c) = 0$, and in case $|\beta - \delta| > 0$ also, choose

$B = (A \times H_2) \cup (H_1 \times C)$, with $P^2_\beta(C) = P^2_\delta(C^c) = 0$, to obtain $P^{\otimes}_{\alpha,\beta}(B) = P^{\otimes}_{\gamma,\delta}(B^c) = 0$.

b) Let $\alpha$, $\beta$ and $\gamma$ be positive and assume $P \perp P^{\otimes}$. Then $P_\alpha \perp P^{\otimes}_{\beta,\gamma}$ and $\hat{P}_\alpha \perp \hat{P}^{\otimes}_{\beta,\gamma}$

Indeed:

Since all the measures considered are Gaussian, if the result is false, one must have, for example, $P_\alpha \equiv P^{\otimes}_{\beta,\gamma}$ (Rao and Varadarajan (1968) : p.308, Theorem 4.1). The proof consists in showing that the latter implies $P \equiv P^{\otimes}$. Since $P$ and $P^{\otimes}$ have the same mean (2.3), one must only check that

$$R^{\otimes}_P = \sqrt{R_P}(id_H + S)\sqrt{R_P},$$

S Hilbert-Schmidt, self-adjoint, and such that $\sigma(S) > -1$ (Rao and Varadarajan (1968) : p.312, Theorem 5.1) : this condition will be referred to as the "H-S condition" henceforth.

Since $P_\alpha \sim N_H(\alpha \hat{m}_P, \alpha^2 R_P)$ and $P^{\otimes}_{\beta,\gamma} \sim N_H((\beta m^1_P, \gamma m^2_P), R^{\otimes}_{\beta,\gamma})$ (2.4 and 2.7) and since it is assumed that $P_\alpha \equiv P^{\otimes}_{\beta,\gamma}$, the H-S condition obtains :

$$\text{i)} \qquad R^{\otimes}_{\beta,\gamma} = \alpha^2 \{(R_P)^{1/2}\}(id_H + T)\{(R_P)^{1/2}\},$$

T Hilbert-Schmidt, self-adjoint, $\sigma(T) > -1$. By 2.7 one has that

$p_1 R^{\otimes}_{\beta,\gamma} p_2 = 0$, or, equivalently, using $i$),

$$\text{ii)} \qquad p_1 R_P p_2 + p_1\{(R_P)^{1/2}\}T\{(R_P)^{1/2}\}p_2 = 0.$$

Similarly, using 2.3., one has that $R^{\otimes}_P - R_P = -p_1 R_P p_2 - p_2 R_P p_1$, so that, from $ii$), one gets :

$$\text{iii)} \quad R^{\otimes}_P - R_P = p_1\{(R_P)^{1/2}\}T\{(R_P)^{1/2}\}p_2 + p_2\{(R_P)^{1/2}\}T\{(R_P)^{1/2}\}p_1.$$

2.7 also yields

$$< p_1 R_P p_1 \hbar, \hbar > \leq < R^{\otimes}_{\beta,\gamma} \hbar, \hbar > / \beta^2,$$

so that, by i),

$$< p_1 R_P p_1 \hbar, \hbar > \leq \Gamma \cdot < R_P \hbar, \hbar >,$$

for some appropriate $\Gamma$. This, in turn, implies (Douglas (1966) : p.413, Theorem 1) :

iv) $p_1 R_P p_1 = \{(R_P)^{1/2}\}U\{(R_P)^{1/2}\}$, $U : H \to H$ bounded, self-adjoint and $\geq 0$. One has similarly :

v) $p_2 R_P p_2 = \{(R_P)^{1/2}\}V\{(R_P)^{1/2}\}$, $V : H \to H$ bounded, self-adjoint and $\geq 0$. The polar decomposition (Weidmann (1980) : p.197, Theorem 7.20) now yields :

vi) $\{(R_P)^{1/2}\}p_1 = A(p_1 R_P p_1)^{1/2}$, $A$ a partial isometry with initial set $L_1$ which is the closure of $\mathcal{R}([p_1 R_P p_1]^{1/2})$ and final set $L_2$ which is the closure of $\mathcal{R}(\{(R_P)^{1/2}\}p_1)$. Similarly, one has :

vii) $\sqrt{U}\{(R_P)^{1/2}\} = B(\{(R_P)^{1/2}\}U\{(R_P)^{1/2}\})^{1/2}$.

Hovever $A^*A$ is the projection $p_{L_1}$ onto $L_1$ and $AA^*$ the projection $p_{L_2}$ onto $L_2$. vi) thus yields $A^*\{(R_P)^{1/2}\}p_1 = (p_1 R_P p_1)^{1/2}$, and, similarly, vii) yields $B^*\sqrt{U}\{(R_P)^{1/2}\} = (\{(R_P)^{1/2}\}U\{(R_P)^{1/2}\})^{1/2}$. It follows then from iv) that $A^*\{(R_P)^{1/2}\}p_1 = B^*\sqrt{U}\{(R_P)^{1/2}\}$, which can be rewritten as

viii) $p_1\{(R_P)^{1/2}\} = \{(R_P)^{1/2}\}\sqrt{U}BA^*$.

Similarly, one has :

ix) $p_2\{(R_P)^{1/2}\} = \{(R_P)^{1/2}\}\sqrt{V}DC^*$.

So, using viii) and ix) in iii), one has, setting $\tilde{T} = \sqrt{U}BA^*TCD^*\sqrt{V}$,

x) $R_P^{\otimes} = \{(R_P)^{1/2}\}(id_H + \tilde{T} + \tilde{T}^*)\{(R_P)^{1/2}\}$

One must finally show that $\sigma(\tilde{T} + \tilde{T}^*) > -1$ , or, equivalently, that $\sqrt{R_P^{\otimes}}$ and $\{(R_P)^{1/2}\}$ have the same range. Now $\sqrt{R_P^{\otimes}}$ and $\sqrt{R_{\beta,\gamma}^{\otimes}}$ have the same range, since

$$0 < min(\beta^2, \gamma^2)R_P^{\otimes} \le R_{\beta,\gamma}^{\otimes} \le max(\beta^2, \gamma^2)R_P^{\otimes}$$

(Douglas (1966) : p.413, Theorem 1). The H-S condition due to $P_\alpha \equiv P_{\beta,\gamma}^{\otimes}$ insures similarly that $\alpha\sqrt{R_P}$ and $\sqrt{R_{\beta,\gamma}^{\otimes}}$ have the same range. Since $\alpha\sqrt{R_P}$ and $\sqrt{R_P}$ have the same range, the H-S condition on $P$ and $P^{\otimes}$ obtains, leading to a contradiction.

c) Let $P \equiv P^{\otimes}$ and assume that the closures of $\mathcal{R}(\sqrt{R_P}p_1)$ and $\mathcal{R}(\sqrt{R_P}p_2)$ have infinite dimension. Then $P_\alpha \equiv P_{\beta,\gamma}^{\otimes}$ if and only if $\alpha = \beta = \gamma$ (thus otherwise $P_\alpha \perp P_{\beta,\gamma}^{\otimes}$).

**Remark 3:** That the closure of $\mathcal{R}(\sqrt{R_P} p_1)$ has positive dimension means that the range of $R_P$ is not contained in $\{\theta_1\} \times H_2$ : indeed, if the dimension is zero, the kernel of $\sqrt{R_P}$ contains $H_1 \times \{\theta_1\}$, and, since $R_P$ and $\sqrt{R_P}$ have the same kernel, $H_1 \times \{\theta_1\}$ is contained in the orthogonal to the closure of $\mathcal{R}(R_P)$, so that $\{\theta_1\} \times H_2$ contains the closure of $\mathcal{R}(R_P)$. Infinite dimension is obtained by choosing, for example, $R_P$ to be an injection.

Proof of c)

Assume thus that $P_\alpha \equiv P^{\otimes}_{\beta,\gamma}$. The H-S condition for $P_\alpha$ and $P^{\otimes}_{\beta,\gamma}$ implies in particular that $\sqrt{R_{P_\alpha}}$ and $\sqrt{R^{\otimes}_{\beta,\gamma}}$ have identical ranges, so that there exist positive constants, $\Gamma_1$ and $\Gamma_2$, for which

$$\Gamma_1 \cdot < R_{P_\alpha} \vec{h}, \vec{h} > \quad \leq \quad < R^{\otimes}_{\beta,\gamma} \vec{h}, \vec{h} > \quad \leq \quad \Gamma_2 \cdot < R_{P_\alpha} \vec{h}, \vec{h} >$$

*Choosing for $\vec{h}$, $p_1 \vec{h}'$ one gets :*

$$\alpha^2 \Gamma_1 \|R_P p_1 \vec{h}'\|^2 \quad \leq \quad \beta^2 \|R_P p_1 \vec{h}'\|^2 \quad \leq \quad \alpha^2 \Gamma_2 \|R_P p_1 \vec{h}'\|^2.$$

Since there exists at least one $\vec{h}'$ such $\|R_P p_1 \vec{h}'\| \quad > 0$, one has

$\alpha^2 \Gamma_1 \quad \leq \quad \beta^2 \quad \leq \quad \alpha^2 \Gamma_2$, so that $\beta = 0$ if and only if $\alpha = 0$. The same is true for $\gamma$. One may thus assume that $\alpha > 0$.

Let $K_1^{\otimes}$ be the closure of the range of the square root of $J_1 R^{\otimes}_{\beta,\gamma} J_1^*$, $L_1^{\otimes}$ be that of the range of $\sqrt{R^{\otimes}_{\beta,\gamma}} J_1^*$, $U_1$ be the partial isometry with $K_1^{\otimes}$ as initial set and $L_1^{\otimes}$ as final set. Similarly, $K_1$ corresponds to $J_1 R_{P_\alpha} J_1^*$, $L_1$ to $\sqrt{R_{P_\alpha}} J_1^*$ and the related isometry is $V_1$. The polar decomposition (Weidmann (1980) : p.197, Theorem 7.20) yields :

$$\sqrt{R^{\otimes}_{\beta,\gamma}} J_1^* = U_1 (J_1 R^{\otimes}_{\beta,\gamma} J_1^*)^{1/2}, \quad \sqrt{R_{P_\alpha}} J_1^* = V_1 (J_1 R_{P_\alpha} J_1^*)^{1/2}.$$

But $J_1 R^{\otimes}_{\beta,\gamma} J_1^* = \beta^2 J_1 R_P J_1^*$, so that $\sqrt{R^{\otimes}_{\beta,\gamma}} = \beta U_1 (J_1 R_P J_1^*)^{1/2}$. Similarly, $\sqrt{R_{P_\alpha}} J^* = \alpha V_1 (J_1 R_P J_1^*)^{1/2}$. Since $(J_1 R_{P_\alpha} J_1^*)^{1/2}$ and $(J_1 R_P J_1^*)^{1/2}$ have the same range, $(J_1 R_P J_1^*)^{1/2} = V_1^* \sqrt{R_{P_\alpha}} J_1^* / \alpha$, and , consequently,

$$\sqrt{R^{\otimes}_{\beta,\gamma}} J_1^* = (\beta/\alpha) U_1 V_1^* \sqrt{R_{P_\alpha}} J_1^*.$$

Let $S_1 = (\beta/\alpha) U_1 V_1^*$. Then

$$\sqrt{R^{\otimes}_{\beta,\gamma}} p_1 = S_1 \sqrt{R_{P_\alpha}} p_1. \quad \text{Similarly,} \quad \sqrt{R^{\otimes}_{\beta,\gamma}} p_2 = S_2 \sqrt{R_{P_\alpha}} p_2.$$

Thus $\sqrt{R^{\otimes}_{\beta,\gamma}} = S_1 \sqrt{R_{P_\alpha}} p_1 + S_2 \sqrt{R_{P_\alpha}} p_2$.

The H-S condition on $P_\alpha$ and $P^{\otimes}_{\beta,\gamma}$ can thus be rewritten, letting $b^2 = \beta^2/\alpha^2$ and $c^2 = \gamma^2/\alpha^2$,

$$\begin{aligned} R_P - b^2 p_1 R_P p_1 - c^2 p_2 R_P p_2 = \\ (p_1 \sqrt{R_P} S_1^* + p_2 \sqrt{R_P} S_2^*) T (S_1 \sqrt{R_P} p_1 + S_2 \sqrt{R_P} p_2), \end{aligned}$$

which yields, after pre-and post-multiplication by $p_1$,

$$(1 - b^2) \|\sqrt{R_P} p_1 \hat{h}\|^2 = < S_1^* T S_1 \sqrt{R_P} p_1 \hat{h}, \sqrt{R_P} p_1 \hat{h} >,$$

which becomes in turn, by continuity,

$$(1 - b^2) \|\hat{h}\|^2 = < S_1^* T S_1 \hat{h}, \hat{h} >, \; \hat{h} \in L_1,$$

for $\sqrt{R_P} J_1^*$ and $\sqrt{R_{P_\alpha}} J_1^*$ have the same range and $J_1^* H_1 = p_1 H$ .

Assume now $1 - b^2 \neq 0$. Then either $(S_1^* T S_1)^{1/2}$ or $(-S_1^* T S_1)^{1/2}$ has the same range as the identity on $L_1$ (Douglas (1966) : p.413, Theorem 1), that is, either $(S_1^* T S_1)^{1/2}$ or $(-S_1^* T S_1)^{1/2}$ has closed range. But these operators are compact and can only have closed range if $L_1$ has finite dimension. Thus $1 - b^2 = 0$, that is $\alpha = \beta$. $\alpha = \gamma$ is obtained similarly.

*3.3. The case of joint measures and measures which are products of marginals : the case of GMM's and SIM's with smooth mixing.*

Let $F$ be continuous and without mass at the origin. Assume further that $\vec{m}_P \in \mathcal{R}(\sqrt{R_P^{\otimes}})$. Then $P_F \perp P_F^{\otimes}$ and $\hat{P}_F \perp \hat{P}_F^{\otimes}$ .

Proof :

By 2.3, $m_P^1 \in \mathcal{R}(\sqrt{R_P^1})$ and $m_P^2 \in \mathcal{R}(\sqrt{R_P^2})$ . Define successively,

$$R_P^1 e_i = \lambda_i e_i, R_P^2 f_i = \mu_i f_i,$$

$$\vec{g}_i = J_1^*(\sqrt{R_P^1})^{-1} e_i, \vec{h}_i = J_2^*(\sqrt{R_P^2})^{-1} f_i,$$

$$X_i(\vec{x}) =< \vec{x}, \vec{g}_i >, Y_i(\vec{x}) =< \vec{x}, \vec{h}_i >,$$

$$\phi_n(\vec{x}) = (1/n)\sum_{i=1}^{n} X_i^2(\vec{x}), \psi_n(\vec{x}) = (1/n)\sum_{i=1}^{n} Y_i^2(\vec{x}),$$

$$\phi(\vec{x}) = \limsup_n \phi_n(\vec{x}), \psi(\vec{x}) = \limsup_n \psi_n(\vec{x}).$$

Now, with respect to $P_\alpha$, $X_i$ has mean equal to

$$\alpha < \vec{m}_P, \vec{g}_i >=< (\sqrt{R_P^1})^{-1} m_P^1, e_i >_1 .$$

Furthermore, cov $(X_i, X_j) = \alpha^2 < R_P \vec{g}_i, \vec{g}_j >= \alpha^2 < e_i, e_j >_1 = \alpha^2 \delta_{i,j}$.

Consequently, as in 3.1,a), $\phi(\vec{x}) = \psi(\vec{x}) = \alpha^2$, a.s. . Similarly, with respect to $P_{\beta,\gamma}^{\otimes}$, $\phi(\vec{x}) = \beta^2$ and $\psi(\vec{x}) = \gamma^2$, a.s. . Let now $\Delta$ be the diagonal in $\mathbb{R}_+^2$ and set $B = (\phi, \psi)^{-1}(\Delta)$. Then

$$P_F(B^c) = \int_0^\infty P_\alpha(B^c) F(d\alpha) = 0$$

and

$$P_F^{\otimes}(B) = \int_0^{\infty}\int_0^{\infty} P_{\beta,\gamma}^{\otimes}(B) F\otimes F(d\beta, d\alpha) = F\otimes F(\Delta).$$

We are thus left to show that $F\otimes F(\Delta) = 0$. Fix an integer $p$ and let $\Delta_p = \{(x,x) : 0 < x \le p\}$. Let $B_i$ be the box centered at $([2i+1]/2n, [2i+1]/2n)$ and equal to $]i/n, (i+1)/n]\times]i/n, (i+1)/n]$. Then $\Delta_p \subseteq \cup_{i=0}^{q} B_i$, where $q = np - 1$. Then

$$\begin{aligned} F\otimes F(\Delta_p) &\le \sum_{i=0}^{q}\{F([i+1]/n) - F(i/n)\}^2 \\ &\le F(p) \max_{1\le i\le np}\{F([i+1]/n) - F(i/n)\} \end{aligned}$$

Since $F$ is continuous, it is uniformly continuous on $[0, p]$ and, consequently,

$$\limsup_{n} \max_{1\le i\le np}\{F([i+1]/n) - F(i/n)\} = 0.$$

Thus $F\otimes F(\Delta) = F\otimes F(\cup_p \Delta_p) = 0$.

*3.4. The case of joint measures and measures which are products of marginals : the case of GMM's and SIM's with finite support.*

Let $0 \le \alpha_1 < \alpha_2 < \cdots < \alpha_n$ be the support of $F$ with mass $p_i \in ]0,1[$ at $\alpha_i, 1 \le i \le n$. We write $P_i$ and $P_{i,j}^{\otimes}$ for, respectively, $P_{\alpha_i}$ and $P_{\alpha_i,\alpha_j}^{\otimes}$.

a) Let $P_1, \cdots, P_m$ and $Q_1, \cdots, Q_n$ be probability measures on $(\Omega, \mathcal{A})$ such that $P_i \perp Q_j$, $1 \le i \le m$, $1 \le j \le n$. Then there exists a set $A \in \mathcal{A}$ such that $P_i(A) = Q_j(A^c) = 0$, $1 \le i \le \mathrm{m}, 1 \le j \le n$.

Indeed :

For measures $\lambda, \mu$ and $\nu$ such that $\lambda \perp \nu$ and $\lambda \perp \mu$, $\lambda \perp \mu + \nu$ (Ash (1972) : p.67, 2.2.5 Lemma). By induction, one obtains that

$$\sum_{i=1}^{m} P_i \perp \sum_{j=1}^{n} Q_j$$

and $A$ is a set such that

$$(\sum_{i=1}^{m} P_i)(A) = (\sum_{j=1}^{n} Q_j)(A^c) = 0.$$

b) $P \equiv P^{\otimes}$ implies $P_F << P_F^{\otimes}$, but never $P_F \equiv P_F^{\otimes}$ whenever $\mathcal{R}(\sqrt{R_P} p_1)$ and $\mathcal{R}(\sqrt{R_P} p_2)$ have infinite dimension.

Indeed :

By 3.2, c), $P_i \perp P_{j,k}^{\otimes}, 1 \leq i,j,k \leq n, j \neq k$. Choose $B$ such that, for $P_i \in \{P_1, \cdots, P_n\}$ and

$$P_{j,k}^{\otimes} \in \{P_{j,k}^{\otimes}, j \neq k, 1 \leq j,k \leq n,\}, P_i(B) = P_{j,k}^{\otimes}(B^c) = 0$$

which can be done because of 3.4,a). Then $P_F(B) = 0$, but

$P_F^{\otimes}(B) = \sum_{j \neq k} p_j p_k > 0$. Absolute continuity follows from the assumption $P \equiv P^{\otimes}$ and 2.7.

c) if $P \perp P^{\otimes}$ and $\alpha_1 > 0$, $P_F \perp P_F^{\otimes}$. If $\alpha_1 = 0$, neither $P_F \perp P_F^{\otimes}$, nor $P_F \equiv P_F^{\otimes}$ obtain.

Indeed :

If $\alpha_1 = 0$, $P_1$ is concentrated at the mean, which is $\vec{\theta}$, and similarly for $P_{1,1}^{\otimes}$. So, if $P_F(B) = 0$, $B^c \supseteq \{\vec{\theta}\}$ and $P_F^{\otimes}(B^c) \geq p_1^2 > 0$.

[illegible] $> 0$, from 3.2 , b), we have that $P_i \in \{P_1, \cdots, P_n\}$ and

$$P_{j,k}^{\otimes} \in \{P_{j,k}^{\otimes}, 1 \leq j,k \leq n,\}$$

are orthogonal, so that the assertion follows from 3.4,a).

d) If $P_F \perp P_F^{\otimes}$, $P \perp P^{\otimes}$.

Indeed :

$P_F(B) = P_F^{\otimes}(B^c)$ implies $P_i(B) = P_{i,i}^{\otimes}(B^c) = 0$ and then one applies 2.7.

e) If $P_F << P_F^{\otimes}$ and $\alpha_1 > 0$, $P \equiv P^{\otimes}$.

Indeed :

If $P \not\equiv P^{\otimes}$, $P \perp P^{\otimes}$ (Rao and Varadarajan (1968) : p.308, Theorem 4.1). But, by 3.4,c), we then have $P_F \perp P_F^{\otimes}$. So, if

$$P_F(B) = P_F^{\otimes}(B^c) = 0, \; P_F(B^c) = 0$$

and $P_F(H) = 0$ which is impossible.

f) If $\alpha_1 > 0$, $\mathcal{R}(\sqrt{R_P}p_1)$ and $\mathcal{R}(\sqrt{R_P}p_2)$ have infinite dimensional range and $P_F << P_F^{\otimes}$, there exist Borel sets $B_i$, $1 \le i \le n$, such that

$$[dP_F/dP_F^{\otimes}](\vec{h}) = \sum_{i=1}^{n}(1/p_i)\{1 - I_{B_i}(\vec{h})\}[dP_i/dP_{i,i}^{\otimes}](\vec{h}).$$

Indeed :

For each fixed $i$, $1 \le i \le n$, choose a Borel set $B_i$ such that

$$P_i(B_i) = P_{j,k}^{\otimes}(B_i^c) = 0, \; 1 \le j,k \le n, (j,k) \ne (i,i).$$

This is possible because of 3.4,d), 3.2,c) and 3.4,a) successively. From 2.7, we furthermore have

$P_i \equiv P_{i,i}^{\otimes}$, $1 \le i \le n$. Let thus

$$\Delta = \sum_{i=1}^{n}(1/p_i)\{1 - I_{B_i}(\vec{h})\}[dP_i/dP_{i,i}^{\otimes}](\vec{h}).$$

Then, for $B \in \mathcal{B}[H]$,

$$\int_B \Delta dP_F^{\otimes} = \sum_{i=1}^{n}\sum_{j=1}^{n} p_i p_j \int_B \Delta dP_{i,j}^{\otimes}.$$

But

$$\int_B \Delta dP^{\otimes}_{i,j} = \sum_{i=1}^{n}(1/p_k)\int_{B\cap B_k^c}[dP_k/dP^{\otimes}_{k,k}](\vec{h})P^{\otimes}_{i,j}(d\vec{h}).$$

Now, $P^{\otimes}_{i,j}(B_k^c) = 0$, except possibly when $i = j = k$. Thus, for $i \neq j$, $\int_B \Delta dP^{\otimes}_{i,j} = 0$.

Furthermore,

$$\int_B \Delta dP^{\otimes}_{i,i} = (1/p_i)\int_{B\cap B_i^c}[dP_i/dP^{\otimes}_{i,i}](\vec{h})P^{\otimes}_{i,i}(d\vec{h})$$
$$= (1/p_i)P_i(B\cap B_i^c) = (1/p_i)P_i(B).$$

Consequently,

$$\int_B \Delta dP^{\otimes}_F = \sum_{i=1}^{n} p_i^2(1/p_i)P_i(B) = P_F(B).$$

**Remark 4:** The case of $\dot{P}_F$ and $\dot{P}^{\otimes}_F$ is treated in the exact same way so that in a)-f), one may replace $P_F$ and $P^{\otimes}_F$ by, respectively, $\dot{P}_F$ and $\dot{P}^{\otimes}_F$.

## 4. Average mutual information

*4.1. Calculation of mutual information in the case of 3.4*

If $Q$ is any probability measure on $\mathcal{B}[H]$, $Q^{\otimes}$ the product of the marginals of $Q$, one has that the average mutual information of $Q$ is given, provided $Q << Q^{\otimes}$, by the formula $I(Q) = \int_H \log[dQ/dQ^{\otimes}]dQ$. The entropy of $F$ is the quantity $H(F) = -\sum_{i=1}^{n} p_i \log p_i$. One has :

under the assumptions of 3.4,f), and the assumption $\vec{m}_P \in \mathcal{R}(\sqrt{R_P})$,

$$i) \qquad I(P_F) = I(P) + H(F)$$

$$ii) \qquad I(\hat{P}_F) = I(P) + H(F)$$

Indeed :

Let $\vec{e}_i$ be the eigenvector of $R_P$ associated with the eigenvalue $\rho_i$. Let also

$$\psi_n(\vec{h}) = (1/n)\sum_{i=1}^{n}\{< \vec{h}, \vec{e}_i > /\sqrt{\rho_i}\}^2,\ \psi(\vec{h}) = \limsup_n \psi_n(\vec{h})$$

and $A_i = \psi^{-1}(\alpha_i^2)$, $1 \le i \le n$. As in 3.1,a)

$$P_i(A_i) = 1,\ P_i(A_j) = 0,\ i \neq j,\ 1 \le i, j \le n.$$

One may thus assume that $dP_i/dP_{i,i}^{\otimes}$ is zero outside of $A_i$. Consequently,

$$\begin{aligned}\int_H \log[dP_F/dP_F^{\otimes}]dP_F &= \sum_{i=1}^{n} p_i \int_H \log[dP_F/dP_F^{\otimes}]dP_i,\\ &= \int_H \log[(1/p_i)[dP_i/dP_{i,i}^{\otimes}]]dP_i,\\ &= -\log p_i + I(P_i).\end{aligned}$$

But, because of Lemma 4 in (Baker (1978) : p.76), $I(P_i) = I(P)$. Consequently

$$I(P_F) = \sum_{i=1}^{n} p_i\{-\log p_i + I(P)\} = I(P) + H(F).$$

*4.2. The case of a more general mixing function*

Every measure on the Borel sets of a separable metric space can be weakly approximated by a measure whose support is a finite number of points (Partha-rasarathy (1967) : p.44, Theorem 6.3). One may thus try to use 4.1 to obtain bounds for the case of a mixing $F$ which does not have finite support.

i) Let $B = \{\vec{h} \in H : ||\vec{h} - \vec{a}|| < \alpha\}$ and $I = \{t : t\vec{h} \in B\}$, $\vec{h} \in B$. Then $I$ is an open interval.

Indeed :

Suppose $t_1 \in I$, $t_2 \in I$ and $t_1 < t < t_2$. Let $\lambda = (t_2 - t)/(t_2 - t_1)$ and $1 - \lambda = (t - t_1)(t_2 - t_1)$. Then $t = \lambda t_1 + (1 - \lambda)t_2$ and

$$||t\vec{h} - \vec{a}|| = ||[\lambda t_1 + (1-\lambda)t_2]\vec{h} - a|| \leq \lambda||t_1\vec{h} - \vec{a}|| + (1-\lambda)||t_2\vec{h} - \vec{a}|| < \alpha,$$

so that $I$ is an interval. Let now $t \in I$ and $\beta = \alpha - ||t\vec{h} - \vec{a}||$. Then, for

$$s < \beta/||\vec{h}||, ||(t \pm s)\vec{h} - \vec{a}|| \leq ||t\vec{h} - \vec{a}|| + s||\vec{h}|| < \alpha,$$

so that $I$ is open and contains points other that 1.

b) Suppose $F_n$ converges weakly to $F$, that $Q_{F_n}(B) = \int_0^\infty Q_F(B)F_n(d\alpha)$ and that $Q_F(B) = \int_0^\infty Q_\alpha(B)F(d\alpha)$, $B \in \mathcal{B}[K]$. Then $Q_{F_n}$ converges weakly to $Q_F$ .

Indeed :

Let

$$g(\vec{h}) = \int_0^\infty I_B(\alpha\vec{h})F(d\alpha)$$

and

$$g_n(\vec{h}) = \int_0^\infty I_B(\alpha\vec{h}) F_n(d\alpha).$$

Then

$$Q_F(B) = \int_K g(\vec{h}) P(d\vec{h})$$

and

$$Q_{F_n}(B) = \int_K g_n(\vec{h}) P(d\vec{h}).$$

Let $B$ be open : it is a union of open balls, so that, $I_B(\alpha\vec{h})$, as a function of $\alpha$, is the indicator of an open set on the real line. Consequently, by weak convergence

$$g(\vec{h}) \leq \liminf_n g_n(\vec{h}),$$

and thus, by Fatou's lemma,

$$Q_F(B) \leq \liminf_n Q_{F_n}(B).$$

c) Let the assumptions of 4.1 hold for $P$ and $F_n$, where $F_n$ is a sequence of probabilities with finite support converging to $F$ weakly.

Then

$$I(P_F) \leq \liminf_n H(F_n) + I(P),$$

since $I$ is lower semi-continuous for weak convergence (Bretagnolle (1979) : p.36, 2.3).

## References


R.B. Ash (1972)

Real Analysis and Probability

Academic Press, New York

C.R. Baker (1970)

On covariance operators

Institute Of Statistics Mimeo Series No. 715

Chapel Hill, North Carolina

C.R. Baker (1973)

Joint measures and cross-covariance operators

Trans. Amer. Math. Soc. 186, 273-289

C.R. Baker (1978)

Capacity of the Gaussian channel without feedback

Information and Control 37, 70-89

C.R. Baker (1979)

Calculation of the Shannon information

J. Math. Anal. Applic. 69, 115-123

J. Bretagnolle (1979)

Formule de Chernoff pour les lois empiriques de variables aléatoires à valeurs dans des espaces généraux

Astérisque 68, 33-52

R.C. Douglas (1966)

On majorization, factorization and range inclusion of operators in Hilbert space

Proc. Amer. Math. soc. 17, 413-415

A.F. Gualtierotti (1980)

On average mutual information and capacity for a channel without feedback and contaminated Gaussian noise

Information and Control 46, 46-70

M. Hitsuda and S. Ihara (1975)

Gaussian channels and the optimal coding

J. Multivariate Anal. 5, 106-118

S.T. Huang and S. Cambanis (1979)

Spherically invariant processes: their non-linear structure, discrimination and estimation

J. Multivariate Anal. 9, 59-83

T.T. Kadota, M. Zakai and J. Ziv (1971)

Mutual information of the white Gaussian channel with and without feedback

IEEE Trans. Information Theory 17, 368-371

J. Keilson and F.W. Steutel (1974)

Mixtures of distributions, moment inequalities and measures of exponentiality and normality

Ann. Prob. 2, 112-130

J. Neveu (1965)

Mathematical Foundations of the Calculus of Probability

Holden-Day, San Francisco

Y.M. Pan (1973)

Simple proofs of equivalence conditions for measures induced by

Gaussian processes

Selected Transl. in Math. Statist. Probability 12, 109-118

Amer. Math. Soc. , Providence RI

K.R. Partharasarathy (1967)

Probability Measures on Metric Spaces

Academic Press, New York

C. Pisot and M. Zamansky (1959)

Mathématiques Générales

Dunod, Paris

C.R. Rao and V.S. Varadarajan (1968)

Discrimination of Gaussian processes

Sankhyà A 25, 303-330

C.E. Shannon (1948)

A mathematical theory of communication

Bell Syst. Tech. J. 27, 379-423 and 623-656

A.V. Skorohod (1974)

Integration in Hilbert Space

Springer Verlag, New York

G.N. Sytaya (1969)

On admissible displacements of suspended Gaussian measures

Theory Probability Applic. 14, 506-509

J. Weidmann (1980)

Linear Operators in Hilbert Spaces

Springer Verlag, New York

END

DTIC

9-86