JOINT MEASURES AND CROSS-COVARIANCE OPERATORS

BY

CHARLES R. BAKER

ABSTRACT. Let $H_1$ (resp., $H_2$) be a real and separable Hilbert space with Borel $\sigma$-field $\Gamma_1$ (resp., $\Gamma_2$), and let $(H_1 \times H_2, \Gamma_1 \times \Gamma_2)$ be the product measurable space generated by the measurable rectangles. This paper develops relations between probability measures on $(H_1 \times H_2, \Gamma_1 \times \Gamma_2)$, i.e., joint measures, and the projections of such measures on $(H_1, \Gamma_1)$ and $(H_2, \Gamma_2)$. In particular, the class of all joint Gaussian measures having two specified Gaussian measures as projections is characterized, and conditions are obtained for two joint Gaussian measures to be mutually absolutely continuous. The cross-covariance operator of a joint measure plays a major role in these results and these operators are characterized.

Introduction. Let $H_1$ (resp., $H_2$) be a real and separable Hilbert space with inner product $\langle \cdot, \cdot \rangle_1$ (resp., $\langle \cdot, \cdot \rangle_2$) and Borel $\sigma$-field $\Gamma_1$ (resp., $\Gamma_2$). Let $\Gamma_1 \times \Gamma_2$ denote the $\sigma$-field generated by the measurable rectangles $A \times B, A \in \Gamma_1, B \in \Gamma_2$. Define $H_1 \times H_2 = \{(u, v): u \in H_1, v \in H_2\}$. $H_1 \times H_2$ is a real linear space, with addition and scalar multiplication defined by $(u, v) + (z, y) = (u + z, v + y)$ and $k(u, v) = (ku, kv)$. $H_1 \times H_2$ is a separable Hilbert space under the inner product $\langle \cdot, \cdot \rangle$ defined by $\langle (u, v), (t, z) \rangle = \langle u, t \rangle_1 + \langle v, z \rangle_2$; moreover, the open sets under the norm obtained from this inner product generate $\Gamma_1 \times \Gamma_2$ [10]. Let $\| \cdot \|_1$ (resp., $\| \cdot \|_2$) denote the norm in $H_1$ (resp., $H_2$) obtained from the inner product, and let $\| \| \cdot \| \|$ denote the norm in $H_1 \times H_2$ obtained from the inner product. A probability measure on $(H_1 \times H_2, \Gamma_1 \times \Gamma_2)$ will be called a joint measure.

A probability measure $\mu_i$ on $(H_i, \Gamma_i)$ ($i = 1$ or $2$) that satisfies

$$\int_{H_i} \| x \|^2 d\mu_i(x) < \infty$$

defines an operator $R_i$ in $H_i$ and a mean element $m_i$ of $H_i$ by

$$\langle m_i, u \rangle_i = \int_{H_i} \langle x, u \rangle_i d\mu_i(x)$$

Received by the editors June 3, 1971.


Key words and phrases. Joint measures, Gaussian measures, absolute continuity of measures, covariance operators, mutual information.

(1) This research was supported by the National Science Foundation under grant GU-2059 and by the U. S. Air Force Office of Scientific Research under contract AFOSR-68-1415.

Copyright © 1974, American Mathematical Society
and
\[ (R_i u, v) = \int_{H_i} (x - m_i, u, v) d\mu_i(x); \]

\( R_i \) is a covariance operator; i.e., it is linear, bounded, nonnegative, selfadjoint, and trace-class. \( \mu_i \) is Gaussian if the probability distribution on the Borel sets of the real line induced from \( \mu_i \) by every bounded linear functional on \( H_i \) is Gaussian. If \( \mu_i \) is Gaussian, (**) is satisfied; moreover, to every covariance operator \( R \) and element \( m \) in \( H \) there corresponds a unique Gaussian measure [9]. All measures on \( (H, \Gamma) \) considered in this paper are probability measures that satisfy (**).

We are interested in determining relations between joint measures and their projections on \( (H, \Gamma) \). In particular, the following questions are answered:

1. What is the relation between the covariance operator of a joint measure and the covariance operators of its projections?
2. Given two Gaussian measures \( \mu_i \) on \( (H_i, \Gamma_i) \), \( i = 1 \) and \( 2 \), how can one characterize the set of all joint Gaussian measures having \( \mu_1 \) and \( \mu_2 \) as projections?
3. What are conditions for equivalence of two joint Gaussian measures, given in terms of operators on \( H_1 \) and \( H_2 \)?

The answers to all three questions involve cross-covariance operators, and a characterization of such operators is given.

Joint measures. A probability measure on \( (H \times H, \Gamma_1 \times \Gamma_2) \) will be called a joint measure. Suppose that \( \mu_{XY} \) is a joint measure; the projection \( \mu_X \) is the probability measure on \( (H_1, \Gamma_1) \) induced from \( \mu_{XY} \) by the \( \Gamma_1 \times \Gamma_2 / \Gamma_1 \) measurable map \( \mathcal{P}_1 \), \( \mathcal{P}_1(x, y) = x \). Similarly, the projection \( \mu_Y \) is the measure on \( (H_2, \Gamma_2) \) induced from \( \mu_{XY} \) by the map \( \mathcal{P}_2 \), \( \mathcal{P}_2(x, y) = y \). Note that there will, in general, not be a unique joint measure having \( \mu_X \) and \( \mu_Y \) as projections; the notation \( \mu_{XY} \) is used to relate the joint measure to its (unique) projections.

A joint measure \( \mu_{XY} \) is Gaussian if the probability distribution on the Borel sets of the real line defined by

\[ \mathcal{P}^{(u, v)}(A) = \mu_{XY}\{(x, y): (x, y, (u, v)) \in A\} \]

is Gaussian for all \( (u, v) \) in \( H_1 \times H_2 \). \( \mathcal{P}^{(u, v)} \) is clearly Gaussian for all \( (u, v) \) in \( H_1 \times H_2 \) if and only if the distribution \( \mathcal{P}_0^{(u, v)} \) on \( B[R^2] \) defined by

\[ \mathcal{P}_0^{(u, v)}(A \times B) = \mu_{XY}\{(x, y): (x, u) \in A, (y, v) \in B\} \]

is Gaussian for all \( (u, v) \) in \( H_1 \times H_2 \).

\( \mu_{XY} \) will have a covariance operator \( \mathbb{R}_{XY} \) and a mean \( m_{XY} \) in \( H_1 \times H_2 \) if
this is always the case if \( \mu_{XY} \) is Gaussian [9]. We will assume that (**) is satisfied for all joint measures considered in this paper. This is consistent with the assumptions for the measures on \( (H_i, \Gamma_i) \), since if \( \mu_{XY} \) is a joint measure with projections \( \mu_X \) and \( \mu_Y \), then

\[
\int_{H_1 \times H_2} \| (u, v) \|^2 \, d\mu_{XY}(u, v) = \int_{H_1 \times H_2} \| u \|^2 + \| v \|^2 \, d\mu_{XY}(u, v)
= \int_{H_1} \| u \|^2 \, d\mu_X(u) + \int_{H_2} \| v \|^2 \, d\mu_Y(v)
\]

so that \( \mu_{XY} \) satisfies (**) if and only if both \( \mu_X \) and \( \mu_Y \) satisfy (*). Given a joint measure \( \mu_{XY} \), we will use \( R_X \) and \( m_X \) (resp., \( R_Y \) and \( m_Y \)) to denote the covariance operator and mean element of the projection \( \mu_X \) (resp., \( \mu_Y \)).

**Cross-covariance operators.** Suppose \( \mu_{XY} \) is a joint measure satisfying (**) Define a functional \( G \) on \( H_1 \times H_2 \) by

\[
G(u, v) = \int_{H_1 \times H_2} \langle x - m_X, u \rangle \langle y - m_Y, v \rangle \, d\mu_{XY}(x, y).
\]

For fixed \( u \) (resp., \( v \)), \( G \) is a linear functional on \( H_2 \) (resp., \( H_1 \)). Moreover, \( |G(u, v)|^2 \leq \| R_X^2 u \|_1^2 \| R_Y^2 v \|_2^2 \), where \( R_X \) and \( R_Y \) are the covariance operators of \( \mu_X \) and \( \mu_Y \). Hence, for fixed \( u \), there exists by Riesz’ theorem a unique element \( q_u \in H_2 \) such that \( G(u, v) = \langle q_u, v \rangle \) for every \( v \in H_2 \). Similarly, for fixed \( v \in H_2 \), there exists a unique element \( g_v \in H_1 \) such that \( G(u, v) = \langle g_v, u \rangle \) for all \( u \in H_1 \). Define a map \( R_{XY}: H_2 \to H_1 \) by \( R_{XY} v = g_v \). \( R_{XY} \) is single-valued, by the fact that \( g_v \) is unique. \( R_{XY} \) is defined everywhere in \( H_2 \), is clearly linear, and is bounded since

\[
\| R_{XY} v \|_1^2 = \| g_v \|_1^2 \leq \sup_{u \in H_1} \left( \frac{\langle g_v, u \rangle_1^2}{\| u \|_1^2} \right) = \| R_X \|_1 \| R_Y \|_2 \| v \|_2^2 \leq \sup_{u \in H_1} \left( \frac{\| R_X^2 u \|_1^2}{\| u \|_1^2} \right) \| R_Y^2 v \|_2^2 \leq \| R_X \|_1 \| R_Y \|_2 \| v \|_2^2.
\]

Clearly \( R_{XY}^*: H_1 \to H_2 \) is defined by \( R_{XY}^* u = q_u^* \). Thus \( G(u, v) = \langle R_{XY}^* u, v \rangle_1 = \langle v, R_{XY} u \rangle_2 \) for all \( u \in H_1 \) and \( v \in H_2 \). We define \( R_{XY}^* = R_{YX}^* \). The operator \( R_{XY} \) will be called the **cross-covariance** operator of \( \mu_{XY} \). A partial characterization of the cross-covariance operator was given in [1] for the case where \( R_X \) and

License or copyright restrictions may apply to redistribution; see http://www.ams.org/journal-terms-of-use
Let \( P_X \) (resp., \( P_Y \)) be the projection operator mapping \( H_1 \) onto range(\( R_X \)) (resp., \( H_2 \) onto range(\( R_Y \))), where range(\( R \)) denotes the closure of range(\( R \)). We then have the following result.

**Theorem 1.** (A) If \( \mu_{XY} \) is a joint measure with a covariance operator and mean element, then the cross-covariance operator \( R_{XY} \) has a representation as \( R_{XY} = R_X^2 V R_Y^2 \), where \( V \) is a unique bounded linear operator such that \( V: H_2 \to H_1, \|V\| \leq 1, \) and \( V = P_X V P_Y \).

(B) If \( R: H_2 \to H_1 \) is a bounded linear operator of trace class, then there exists a joint Gaussian measure \( \mu_{XY} \) such that \( R \) is the cross-covariance operator of \( \mu_{XY} \).

**Proof.** (A) Let \( s \) be any fixed element in range(\( R_Y^2 \)), with \( z \) any element of \( H_2 \) satisfying \( R_Y z = s \). Define a linear functional \( f_s \) on range(\( R_X^2 \)) by

\[
f_s(R_X^2 u) = \int_{H_1 \times H_2} (x - m_X, u) (y - m_Y, z)^2 \, d\mu_{XY}(x, y)
\]

Since \( |f_s(R_X^2 u)| \leq \|s\| \cdot \|R_X^2 u\|_1 \), \( f_s \) is bounded on range(\( R_X^2 \)) and thus can be extended by continuity to a bounded linear functional on range(\( R_X^2 \)) (= \( P_X[H_1] \)). Note that the extension has norm \( \leq \|s\|_2 \). By Riesz' theorem, there exists a unique element \( h \) in \( P_X[H_1] \) such that \( f_s(w) = \langle h, w \rangle_1 \) for all \( w \) in \( P_X[H_1] \) and \( \|h\|_1 \leq \|s\|_2 \). Define a map \( V': H_2 \to H_1 \) by \( V' s = h \). \( V' \) is defined for all \( s \) in range(\( R_Y^2 \)), is clearly linear and single-valued, and is bounded because \( \|V' s\|_1 \leq \|s\|_2 \). \( V' \) can thus be extended by continuity to a bounded linear operator \( V \) defined on \( P_X[H_2] \); note that \( Vs = P_X V P_Y s \) for \( s \) in \( P_X[H_2], \|V\| \leq 1, \) and \( f_s(w) = \langle Vs, w \rangle_1 \). We extend the domain of \( V \) to all of \( H_2 \) by defining \( Vu = 0 \) for \( u \) in \( P_X[H_2]^\perp \). Thus, for any \( z \) in \( H_2 \), for \( s = R_Y^2 z \), and for any \( u \) in \( H_1 \), one has \( f_s(R_X^2 u) = \langle R_{XY} z, u \rangle_1 = \langle VR_Y^2 z, R_X^2 u \rangle_1 \), so that \( R_{XY} = R_X^2 V R_Y^2, \|V\| \leq 1, \) and \( V = P_X V P_Y \).

To see that \( V \) is unique, suppose that \( R_{XY} = R_X^2 G R_Y^2 \), with \( \|G\| \leq 1 \) and \( G = P_X G P_Y \). Then \( (V - G)R_Y u = R_X^2 (V - G)R_Y^2 u = 0, \) all \( u \in H_2, \) so that \( Vu = Gu, \) all \( u \) in \( P_Y[H_2]. \) Since \( Vu = Gu = 0 \) for \( u \perp P_Y[H_2], \) \( V = G \) on \( H_2. \)

(B) By the polar decomposition theorem, \( R = UT^2 \), where \( T: H_2 \to H_2, \) \( T^2 = R^* R \), and \( U: H_2 \to H_1 \) is partially isometric, isometric on \( P_T[H_2] \) and zero on \( (P_T[H_2])^\perp, \) with range(\( U \)) = range(\( R \)) (\( P_T = \) the projection operator in \( H_2 \)).
with range equal to \( \text{range}(T) \). Since \( R \) is trace-class, \( T \) and \( UT \) are Hilbert-Schmidt. Further, \( TU^* = W(UT^2U^*)^{1/2} \) for a partially isometric \( W: H_1 \rightarrow H_2 \), \( W \) isometric on \( \text{range}(UT^2U^*) \), and with \( \text{range}(W) \subset \text{range}(T) = \text{range}(R^2R) \). Thus \( R = UT^2 = (UT^2U^*)^{1/2}W^*T \).

Since \( T \) is selfadjoint and Hilbert-Schmidt, there exists a Gaussian measure \( \mu_Y \) on \( (H_2, \Gamma_2) \) with covariance operator \( T^2 \) and null mean element. Define \( Y: H_2 \rightarrow H_2 \) as the identity map, and \( X: H_2 \rightarrow H_1 \) by \( Xv = Uv \). \( X \) is \( \Gamma_2 / \Gamma_1 \) measurable as a continuous map, and thus induces from \( \mu_Y \) a probability measure \( \mu_X \) on \( (H_1, \Gamma_1) \), \( \mu_X(A) = \mu_Y(Uv \in A), A \in \Gamma_1 \). \( \mu_X \) is Gaussian, with null mean element and covariance operator \( R_X = UT^2U^* \).

The map \( (X, Y): H_2 \rightarrow H_1 \times H_2, (X, Y)(v) = (Uv, v) \), is \( \Gamma_2 / \Gamma_1 \times \Gamma_2 \) measurable, thus induces from \( \mu_Y \) a measure \( \mu_{XY} \) on \( (H_1 \times H_2, \Gamma_1 \times \Gamma_2) \), defined by \( \mu_{XY}(C) = \mu_Y((Uu, u) \in C), C \in \Gamma_1 \times \Gamma_2 \). Moreover,

\[
\langle R_{XY}u, v \rangle_1 = \int_{H_1 \times H_2} \langle x, v \rangle_1(y, u)_2 \, d\mu_{XY}(x, y) = \int_{H_2} \langle y, U^*v \rangle_2(y, u)_2 \, d\mu_Y(y) = \langle R_YU^*v, u \rangle_2 \quad \text{for all } u \in H_2, v \in H_1.
\]

Hence \( R_{XY} = UR_Y = UT^2 = R \). Finally, it is clear from the definitions that \( \mu_{XY} \) is Gaussian. This completes the proof.

Covariance operators for joint measures. Suppose that \( \mu_{XY} \) is a joint measure satisfying (**). We proceed to determine the relations between the covariance operator and mean element of \( \mu_{XY} \), and the covariance operators and mean elements of the projections \( \mu_X \) and \( \mu_Y \).

**Proposition 1.** Let \( \mu_{XY} \) be a joint measure such that

\[
\int_{H_1 \times H_2} \| (u, v) \|^2 \, d\mu_{XY}(u, v) < \infty.
\]

Let \( R_{XY} \) and \( m_{XY} \) be the covariance operator and mean element of \( \mu_{XY} \), and denote by \( R_X \) and \( m_X \) (resp., \( R_Y \) and \( m_Y \)) the covariance operator and mean element of the projection \( \mu_X \) (resp., \( \mu_Y \)). Then, \( m_{XY} = (m_X, m_Y) \), and \( R_{XY}(u, v) = (R_Xu + R_Yv, R_Yv + R_YXu) \) for all \( (u, v) \) in \( H_1 \times H_2 \).

**Proof.** It is clear that \( m_{XY} = (m_X, m_Y) \); for example,

\[
\langle m_X, u \rangle_1 = \int_{H_1} \langle x, u \rangle_1 \, d\mu_X(x) = \int_{H_1 \times H_2} [(x, y), (u, 0)] \, d\mu_{XY}(x, y) = [m_{XY}, (u, 0)].
\]
To describe the covariance operator $R_{XY}$, we can assume that $\mu_{XY}$ has mean element. Then

$$\begin{align*}
[R_{XY}(u, v), (t, z)] &= \int_{H_1 \times H_2} [(x, u) + (y, v), (t, t)] \, d\mu_{XY}(x, y) \\
&= \int_{H_1 \times H_2} \{ (x, u) + (y, v), (t, t) \} \, d\mu_{XY}(x, y) \\
&= \langle Rxu, t \rangle + \langle R_yv, t \rangle + \langle Rxv, u \rangle + \langle Ryu, v \rangle \\
&= \langle Rxu + R_yv, R_yv + Rxv \rangle.
\end{align*}$$

One thus sees that $R_x$, $R_y$, and $R_{xy}$ completely characterize $\mu_{XY}$.

Characterization of joint Gaussian measures. One can now characterize the set of all Gaussian joint measures having given Gaussian measures $\mu_x$ and $\mu_y$ as projections.

**Theorem 2.** Suppose that $\mu_x$ and $\mu_y$ are Gaussian measures on $(H_1, \Gamma_1)$ and $(H_2, \Gamma_2)$, respectively. Let $R_x$ and $m_x$ (resp., $R_y$ and $m_y$) denote the covariance operator and mean element of $\mu_x$ (resp., $\mu_y$).

(A) A joint Gaussian measure $\mu$, having $R$ as covariance operator and $m$ as mean element, has $\mu_x$ and $\mu_y$ as projections if and only if $m = (m_x, m_y)$ and $R(u, v) = (Rxu + R_{xy}v, R_yv + Rxv)$ for all $(u, v)$ in $H_1 \times H_2$, where $R_{xy} = R_{xy}^{1/2}VR_{xy}^{1/2}$ for a bounded linear operator $V: H_2 \to H_1$ with $\|V\| \leq 1$.

(B) Let $V$ be any bounded linear operator mapping $H_2$ into $H_1$, with $\|V\| \leq 1$. Define an operator $R_{XY}: H_2 \to H_1$ by $R_{XY} = R_x^{1/2}VR_{xy}^{1/2}$, and define $\tilde{R}: H_1 \times H_2 \to H_1 \times H_2$ by $\tilde{R}(u, v) = (Rxu + R_{xy}v, R_yv + Rxv)$ for all $(u, v)$ in $H_1 \times H_2$. $\tilde{R}$ is then a covariance operator, and the Gaussian joint measure having $\tilde{R}$ as covariance operator and $(m_x, m_y)$ as mean element has $\mu_x$ and $\mu_y$ as projections.

**Proof.** (A) If $\mu$ is a joint Gaussian measure with projections $\mu_x$ and $\mu_y$, then the assertion on the form of $R$ and $m$ follows from Proposition 1.

Conversely, suppose that $\mu$ is a joint Gaussian measure with covariance operator and mean element having the form given in the statement of (A). Suppose also that $\mu$ has projections $\mu_Z$ and $\mu_W$, and that these measures have covariance operators $R_Z$ and $R_W$ and mean elements $m_Z$ and $m_W$. The projections must be Gaussian and, by Proposition 1, $m = (m_Z, m_W)$ and also $R(u, v) = (R_Zu + RZWv, R_Wv + RZWu)$ for all $(u, v)$ in $H_1 \times H_2$, where $R_{ZW}$ is the cross-covariance operator of $\mu$. It is clear that $m_Z = m_Z$ and $m_W = m_W$. Now let $\{u_n\}$ be a c.o.n. set in $H_1$; for any $u$ in $H_1$, one has that $\left[\tilde{R}(u, 0), (u_k, 0)\right] = \langle Rxu, u_k \rangle$ for each $u_k$ in $\{u_n\}$, so that $R_x = R_Z$. Similarly, $R_y = R_W$. Hence, $\mu_x = \mu_Z$ and $\mu_y = \mu_W$, by the unique correspondence between a Gaussian measure and its covariance operator and mean element.
(B) It is sufficient to show that the operator $\mathcal{R}$ is a covariance operator, with $\mathcal{R}$ as defined in the theorem, and with $V$ any bounded linear operator mapping $H_2$ into $H_1$ with $\|V\| \leq 1$.

It is straightforward to verify that $\mathcal{R}$ is linear and selfadjoint; since $\mathcal{R}$ is defined everywhere in $H_1 \times H_2$, $\mathcal{R}$ is also bounded, by the closed-graph theorem. To see that $\mathcal{R}$ is nonnegative,

$$
\langle \mathcal{R}(u, v), (u, v) \rangle = (R_X^T u, u) + (R_Y v, v) + 2(R_{XY} v, u)
$$

$$
= \|R_X^T u\|^2 + \|R_Y v\|^2 + 2(V R_{XY}^T v, R_X^T u)
$$

$$
\geq \|R_X^T u\|^2 + \|R_Y v\|^2 - 2\|R_X^T u\|_1 \|R_Y v\|_2
$$

$$
= (\|R_X^T u\|_1 - \|R_Y v\|)^2 \geq 0.
$$

It remains to show that $\mathcal{R}$ is trace-class. Let $\{u_n\}, n = 1, 2, \cdots$, be c.o.n. in $H_1$, and let $\{v_n\}, n = 1, 2, \cdots$, be c.o.n. in $H_2$. Then the set $\{(u_n, 0), n = 1, 2, \cdots\} \cup \{(0, v_n), n = 1, 2, \cdots\}$ is c.o.n. in $H_1 \times H_2$, and

$$
\sum_n \langle \mathcal{R}(u_n, 0), (u_n, 0) \rangle + \sum_n \langle \mathcal{R}(0, v_n), (0, v_n) \rangle
$$

$$
= \sum_n \langle R_X^T u_n, u_n \rangle + \sum_n \langle R_Y v_n, v_n \rangle = \text{Trace } R_X + \text{Trace } R_Y.
$$

Thus, $\mathcal{R}$ is a covariance operator, and by (A) the joint Gaussian measure having $\mathcal{R}$ as covariance operator and $(m_X, m_Y)$ as mean element has $\mu_X$ and $\mu_Y$ as projections.

Corollary. Let $\mathcal{R}$ be a bounded linear operator in $H_1 \times H_2$. If there exist bounded linear operators $R_1$, $R_2$ and $R_3$, with $R_1 : H_1 \to H_1$, $R_2 : H_2 \to H_2$, and $R_3 : H_2 \to H_1$, such that $\mathcal{R}(u, v) = (R_1 u + R_2 v, R_1 v + R_3 u)$ for all $(u, v)$ in $H_1 \times H_2$, then these operators are unique.

Proof. Follows directly from the second part of the proof of (A).

Further properties of the covariance operator. This section contains more details on the covariance operator $\mathcal{R}_{XY}$ for a joint measure $\mu_{XY}$. Included is an explicit expression for the square root of $\mathcal{R}_{XY}$, and a description of spectral properties of operators related to the cross-covariance operator.

The following result will be used in this and succeeding sections.

Lemma 1 [2]. Suppose that $H$ is a real Hilbert space, and that $R_1$ and $R_2$ are bounded linear operators, $R_1 : H_1 \to H$, $R_2 : H_2 \to H$. Let $P_1$ (resp., $P_2$) be the projection operator mapping $H_1$ (resp., $H_2$) onto $\text{range}(R_1)$ (resp., $\text{range}(R_2)$).
Then range(\text{R}_1) \subseteq \text{range}(\text{R}_2) if and only if there exists a bounded linear operator \text{G}: H_1 \to H_2 such that \text{R}_1 = \text{R}_2 \text{G}, \text{G} = \text{P}_2 \text{GP}^{-1}_1. Moreover, range(\text{R}_1) \subseteq \text{range}(\text{R}_2) if and only if there exists \(k < \infty\) such that \(\|\text{R}_1^* u\|^2 \leq k \|\text{R}_2^* u\|^2\) for all \(u \in H_1\).

Corollary. (a) \text{Range}(\text{R}_1) \subseteq \text{range}(\text{R}_2) \iff there exists a bounded linear operator \text{Q}: H_2 \to H_2 such that \text{R}_1 \text{R}_1^* = \text{R}_2 \text{Q}\text{R}_2^*.

(b) \text{Range}(\text{R}_1) = \text{range}(\text{R}_2) \iff there exists a bounded linear operator \text{Q} having bounded inverse with \((\text{R}_1 \text{R}_1^*)^{1/2} = (\text{R}_2 \text{R}_2^*)^{1/2}\) \text{Q} \iff there exists a bounded linear operator \text{T} having bounded inverse with \((\text{R}_1 \text{R}_1^*)^{1/2} = (\text{R}_2 \text{R}_2^*)^{1/2} \text{T}(\text{R}_2 \text{R}_2^*)^{1/2}\).

(c) \((\text{R}_1 \text{R}_1^*)^{1/2} = \text{R}_1 \text{A}^*\), where \text{A} is partially isometric, isometric on \text{range}(\text{R}_1^*), and \(\text{Au} = 0\) for \(u \perp \text{range}(\text{R}_1^*)\).

In many applications, such as determining equivalence of two joint Gaussian measures, one must verify conditions involving the square root of a covariance operator. We now obtain an explicit representation for the square root of the covariance operator of a joint measure \(\mu_{XY}\).

Let \(R_{X \otimes Y}\) denote the covariance operator of \(\mu_X \otimes \mu_Y\), the product measure for \(\mu_X\) and \(\mu_Y\) on \((H_1 \times H_2, \Gamma_1 \times \Gamma_2)\). For \(\mu_X \otimes \mu_Y\), the cross-covariance operator is the null operator, so that

\[R_{X \otimes Y}(u, v) = (R_X^* u, R_Y^* v),\]

Hence \(R_{X \otimes Y}^{1/2}(u, v) = (R_X^{1/2} u, R_Y^{1/2} v)\), and one can write \(R_{XY}\) as \(R_{XY}(u, v) = (R_X^* u, R_Y^* v) + (R_X^* V R_Y^* v, R_Y^* V R_X^* u)\) (where \(R_{XY} = R_X^* V R_Y^*\), \(V: H_2 \to H_1, \|V\| \leq 1\) = \(R_{X \otimes Y}(u, v) + R_{X \otimes Y}^{1/2} \Omega R_{X \otimes Y}(u, v)\), with \(\Omega(u, v) = (Vv, V^* u)\). \(\Omega\) is a selfadjoint bounded linear operator with \(\|\Omega\| = \|V\|\), as can be easily verified.

We have established the following result.

Proposition 2. \(R_{XY} = \Omega \Omega^{1/2} + R_{X \otimes Y}^{1/2},\) and \(R_{XY} = \Omega \Omega^{1/2} + R_{X \otimes Y}^{1/2}\), where \(\Omega\) is a partially isometric operator, isometric on \text{range}(\text{R}_{XY}), and zero on the null space of \text{R}_{XY}, and \(\Omega\) is the identity operator in \(H_1 \times H_2\).

The second part of this result follows from the corollary to Lemma 1, and the fact that \text{range}(\text{R}_{XY}) = \text{range}(\text{R}_{X \otimes Y}).

Note that Proposition 2 and Lemma 1 imply that \(\text{range}(\text{R}_{XY}^{1/2}) \subseteq \text{range}(\text{R}_{X \otimes Y}^{1/2})\), with equality if and only if \(\Omega + \Omega\) has bounded inverse. This implies \(\|\Omega\|\) that the support of a joint Gaussian measure \(\mu_{XY}\) is always contained in the support of \(\mu_X \otimes \mu_Y\) with equality if and only if \(\Omega + \Omega\) is nonsingular on \text{range}(\text{R}_{X \otimes Y}).

The operators \(\Omega + \Omega\) and \(\Omega\), defined above, play an important role in the study
of equivalence of joint Gaussian measures. We proceed to examine the spectral properties of these operators. Recall that the set of limit points of the spectrum of a selfadjoint bounded linear operator consists of all points of the continuous spectrum, all limit points of the point spectrum, and all eigenvalues of infinite multiplicity. The identity operator in $H_1$ and the identity operator in $H_2$ will both be denoted by the symbol $I$; the appropriate space will be clear from the context.

**Theorem 3.** (a) $\mathcal{V} \mathcal{V}^*$ has a c.o.n. set of eigenvectors if and only if $\mathcal{V}^* \mathcal{V}$ has a c.o.n. set of eigenvectors.

(b) $\mathcal{O}(a, b) = \lambda(a, b) \Leftrightarrow \mathcal{V} \mathcal{V}^* a = \lambda^2 a$, $\mathcal{V}^* \mathcal{V} b = \lambda^2 b$, and $\mathcal{V}^* b = \lambda b \Leftrightarrow \mathcal{O}(a, -b) = -\lambda(a, -b)$. $\mathcal{O}$ has a c.o.n. set of eigenvectors if and only if $\mathcal{V} \mathcal{V}^*$ has a c.o.n. set of eigenvectors.

(c) $\mathcal{O}$ is compact if and only if $\mathcal{V}$ is compact.

(d) $\mathcal{O} + \mathcal{O}$ is compact if and only if both $H_1$ and $H_2$ are finite-dimensional spaces.

(e) If there exists a scalar $\alpha < 1$ such that $\|\mathcal{V}^* a\|^2 \leq \alpha \|a\|^2$ for all $a$ in $\text{range}(\mathcal{I} - \mathcal{V} \mathcal{V}^*)$, then $\mathcal{O} + \mathcal{O}$, $\mathcal{I} - \mathcal{V}^* \mathcal{V}$, and $\mathcal{I} - \mathcal{V} \mathcal{V}^*$ each has closed range, and $\mathcal{O} + \mathcal{O}$ is a limit point of the spectrum of $\mathcal{I} - \mathcal{V} \mathcal{V}^*$.

Proof. (a) Suppose $\mathcal{V} \mathcal{V}^* a = \lambda^2 a$, with $\{a\}$ c.o.n. in $H_1$. Define $\mathcal{V}^* a = \lambda^2 a$ for all $a$ such that $\lambda^2 \neq 0$. Then $\mathcal{V}^* \mathcal{V} a = \lambda^2 a$. Now suppose that there exists $v$ in $H_2$ such that $(v, v) = 0$ for all $a$. Then $\mathcal{V} v$ is in the null space of $\mathcal{V}^* \mathcal{V}$, so that $\mathcal{V}^* \mathcal{V} v = 0$. Hence $\{v\}$ contains a set that is c.o.n. in $H_2$. The converse follows by symmetry.

(b) $\mathcal{O}(a, b) = \lambda(a, b) \Leftrightarrow \mathcal{V} \mathcal{V}^* a = \lambda a$ and $\mathcal{V} \mathcal{V}^* b = \lambda b$ and $\mathcal{V}^* \mathcal{V} a = \lambda^2 a$ and $-\lambda(-b) = \mathcal{V} \mathcal{V}^* b = -\lambda(b, -a)$. If $\{(m, z), (u, v)\}$ is a complete set in $H_1 \times H_2$, then $\{u\}$ must be complete in $H_1$, so that $\mathcal{V} \mathcal{V}^*$ has a complete set of eigenvectors if $\mathcal{O}$ has a complete set. Conversely, if $\mathcal{V} \mathcal{V}^* a = \lambda^2 a$, define $v$ by $\lambda v = \mathcal{V}^* v = \lambda^2 v$ and $\mathcal{V} v = \mathcal{V} \mathcal{V}^* v$. If $\{u, v\}$ is a complete set in $H_1 \times H_2$, then $\mathcal{O}(u, v)$ is a limit point of the spectrum of $\mathcal{O} + \mathcal{O}$, and $\mathcal{O}(u, v)$ is c.o.n. in $H_1$. In this case, the eigenvectors of $\mathcal{V}^* \mathcal{V}$ are complete in $H_2$, by (a), and the proof of (a) shows that the nonzero point spectrum of $\mathcal{V}^* \mathcal{V}$ is identical to the nonzero point spectrum of $\mathcal{V} \mathcal{V}^*$. Thus, the element $v$ above must belong to the null space of $\mathcal{V}^* \mathcal{V}$. If $\{x_n\}$ is c.o.n. in the null space of $\mathcal{V}^* \mathcal{V}$, the set $\{(0, x_n)\}$ are eigenvectors of $\mathcal{O}$ corresponding to the eigenvalue zero. Hence, the union of this set with the eigenvectors $\{(u, v)\}$
C. R. Baker

derived as above from the eigenvectors of $VV^*$ constitutes a complete set in $H_1 \times H_2$.

(c) Follows directly from (b).

(d) From (b), $\mathcal{G} + \mathcal{O}$ cannot have zero as the only limit point of its nonzero eigenvalues, since $\lambda$ is an eigenvalue of $\mathcal{G} + \mathcal{O}$ if and only if $2 - \lambda$ is an eigenvalue. $\mathcal{G} + \mathcal{O}$ is thus compact if and only if range($\mathcal{G} + \mathcal{O}$) is finite-dimensional.

The above eigenvalue relation also shows that the dimension of the null space of $\mathcal{G} + \mathcal{O}$ is equal to the multiplicity of the eigenvalue $\lambda = 2$. Hence, the dimension of the range of $\mathcal{G} + \mathcal{O}$ will be finite-dimensional if and only if $H_1 \times H_2$ is finite-dimensional. This occurs if and only if both $H_1$ and $H_2$ are finite-dimensional.

(e) Suppose there exists $\alpha < 1$ such that $\|V^*u\|_2^2 \leq \alpha \|u\|_2^2$ on range($I - VV^*$). Then, $VV^* \leq \alpha I$ on range($I - \mathcal{G}$), $V^*V \leq \alpha I$ on range($I - \mathcal{G}^*$), and hence $\mathcal{O} \leq \alpha \mathcal{G}$ on range($\mathcal{G} - \mathcal{O}^2$). If $u$ is in range($\mathcal{G} - \mathcal{O}^2$), the above inequalities yield $((\mathcal{G} + \mathcal{O})u, u) \geq (1 - \alpha \frac{1}{2}) \|u\|_2^2$. Suppose that $u \perp \text{range}(\mathcal{G} - \mathcal{O}^2)$; then either $\mathcal{O} \perp \text{range}(\mathcal{G} - \mathcal{O})$ or else $(\mathcal{G} - \mathcal{O})u = 2u$. Noting that the null space of $\mathcal{G} + \mathcal{O}$ is contained in the null space of $\mathcal{G} - \mathcal{O}^2$, one sees that $\mathcal{G} + \mathcal{O} \geq (1 - \alpha \frac{1}{2}) \mathcal{G}$ on range($\mathcal{G} - \mathcal{O}$). This implies that $\mathcal{G} + \mathcal{O}$ has closed range. The fact that $I - VV^*$ and $I - V^*V$ each has closed range follows from the preceding inequalities.

(f) $\alpha$ is a limit point of the spectrum of $\mathcal{G} + \mathcal{O}$ if and only if there exists a normalized sequence $(u_n, v_n)$ in $H_1 \times H_2$ which is weakly convergent to zero, and such that $\|(\mathcal{G} + \mathcal{O} - \alpha \mathcal{G})(u_n, v_n)\| \to 0$ [12]. $(\mathcal{G} + \mathcal{O} - \alpha \mathcal{G})(u_n, v_n) \to (0, 0)$ if and only if $(1 - \alpha)u_n + Vv_n \to 0$, $(1 - \alpha)v_n + V^*u_n \to 0$, and $(1 - \alpha)^2 u_n - VV^*u_n \to 0$. If $\|u_n\|_1$ is bounded away from zero, it is clear that $1 - (1 - \alpha)^2$ is a limit point for $I - VV^*$, since $\|u_n\|_1$ must be weakly convergent to zero. Thus, suppose that $u_n \to 0$ for some subsequence $\{u_{n_k}\}$. Then $V^*u_{n_k} \to 0$ and hence $v_{n_k} \to 0$; this contradicts the assumption that $\{u_n, v_n\}$ is normalized in $H_1 \times H_2$. Hence, $\|u_n\|_1$ must be bounded away from zero, and thus $1 - (1 - \alpha)^2$ is a limit point for $I - VV^*$.

Conversely, suppose that $\{u_n\}$ is a normalized sequence in $H_1$, weakly convergent to zero, with $(1 - \alpha)^2 u_n - VV^*u_n \to 0$. Define positive scalars $\{k_n\}$ by $k_n^2 = 1 + (1 - \alpha)^{-2} \|VV^*u_n\|^2$. $\{k_n^{-1}(u_n, (\alpha - 1)^{-1}V^*u_n)\}$ is a normalized sequence in $H_1 \times H_2$, weakly convergent to zero, with $((\mathcal{G} + \mathcal{O} - \alpha \mathcal{G})(k_n^{-1}u_n, k_n^{-1}(\alpha - 1)^{-1}V^*u_n)) \to (0, 0)$.

$\alpha$ is thus a limit point of $\mathcal{G} + \mathcal{O}$ when $1 - (1 - \alpha)^2$ is a limit point of $I - VV^*$. This completes the proof of the theorem.

Equivalence of joint Gaussian measures. For two Gaussian measures on a real and separable Hilbert space, necessary and sufficient conditions for equiva-
lence (mutual absolute continuity) have been the subject of extensive investigations. It is known that the two measures are either equivalent (denoted by ~) or orthogonal \[3, \] \[6\]; the necessary and sufficient conditions for equivalence can be expressed in terms of the covariance operators and mean elements of the two measures (see, e.g. \[11, \text{Theorem 5.1}\], or \[13\]). Since these results apply to any two Gaussian measures on any real and separable Hilbert space, they can be used to determine equivalence or orthogonality of two joint Gaussian measures on \((H_1 \times H_2, \Gamma_1 \times \Gamma_2)\). Our objective here is to formulate conditions for equivalence in terms of operators and elements in the individual spaces \(H_1\) and \(H_2\).

Throughout this section, we will consider two Gaussian joint measures, \(\mu_{XY}\) and \(\mu_{ZW}\). The cross-covariance operators will have the form \(R_{XY} = R_{X}^{1/2} V R_{Y}^{1/2}\) and \(R_{ZW} = R_{Z}^{1/2} T R_{W}^{1/2}\), where \(V\) and \(T\) are linear, \(\|V\| < 1\) and \(\|T\| < 1\). \(\mu_{XY}\) has mean \(m_{XY} = (m_x, m_y)\), and \(\mu_{ZW}\) has mean \(m_{ZW} = (m_z, m_w)\). We will also assume that range\((R_{X}^{1/2} \otimes \Theta_Y)\) = range\((R_{Z}^{1/2} \otimes \Theta_W)\) = \(H_1 \times H_2\), where \(R_{X}^{1/2} \otimes \Theta_Y\) is the covariance operator of \(\mu_X \otimes \mu_Y\), similarly \(R_{Z}^{1/2} \otimes \Theta_W\) for \(\mu_Z \otimes \mu_W\). This is no restriction, for the following reasons: (1) range\((R_{X}^{1/2} \otimes \Theta_Y)\) \(\subset\) range\((R_{Z}^{1/2} \otimes \Theta_W)\), as shown previously; (2) conditions for equivalence of \(\mu_{ZW}\) and \(\mu_{XY}\) depend only on elements in, and operators defined on, range\((R_{X}^{1/2} \otimes \Theta_Y)\) and range\((R_{Z}^{1/2} \otimes \Theta_W)\), by Lemma 2 below; (3) \(\mu_X \otimes \mu_Y \perp \mu_Z \otimes \mu_W\) if range\((R_{X}^{1/2} \otimes \Theta_Y)\) \(\neq\) range\((R_{Z}^{1/2} \otimes \Theta_W)\), as one can easily show, and we will see that \(\mu_{XY} \perp \mu_{ZW}\) if \(\mu_X \otimes \mu_Y \perp \mu_Z \otimes \mu_W\), independently of the assumption that range\((R_{X}^{1/2} \otimes \Theta_Y)\) = range\((R_{Z}^{1/2} \otimes \Theta_W)\) = \(H_1 \times H_2\).

The two lemmas below are fundamental to our results.

**Lemma 2** \[11\]. Suppose \(\mu_1\) and \(\mu_2\) are two Gaussian measures on the Borel \(\sigma\)-field of a real and separable Hilbert space \(H\). Let \(R_1\) and \(m_1\) be the covariance operator and mean element of \(\mu_1\). Then \(\mu_1 \sim \mu_2\) if and only if

(a) \(m_1 - m_2\) is in the range of \(R_1^{1/2}\);

(b) \(R_1 = R_2 + R_2^{1/2} W R_2^{1/2}\), where \(W\) is a Hilbert-Schmidt operator that does not have \(-1\) as an eigenvalue, and \(W\) is identically zero on the null space of \(R_2^{1/2}\).

This is simply a restatement of Rao-Varadarajan Theorem 5.1 \[11\], slightly modified to allow for range\((R_2^{1/2}) \neq H\). The extension follows easily from Theorem 4.1 of \[11\], and the fact that the support of a zero-mean Gaussian measure is the closure of the range of its covariance operator \[8\] (note that range\((R_2^{1/2}) = range(R_2))\). Since the values of \(W\) on the null space of \(R_2\) do not affect the values of \(R_1 = R_2 + R_2^{1/2} W R_2^{1/2}\) on \(H\), it is clear that if one defines \(W\) by \(W = W'\) on range\((R_2)\), and \(W = 0\) on range\((R_2)^{1/2}\), then \(R_1 = R_2 + R_2^{1/2} W R_2^{1/2}\).

**Lemma 3.** \(\mu_X \otimes \mu_Y \sim \mu_Z \otimes \mu_W\) if and only if \(\mu_X \sim \mu_Z\) and \(\mu_Y \sim \mu_W\).
Proof. Suppose $\mu_X \sim \mu_Z$ and $\mu_Y \sim \mu_W$. Then

$$
\mu_X \otimes \mu_Y [A] = \int_A \mu_X(x) \mu_Y(y),
$$

so that $\mu_X \otimes \mu_Y [A] = 0$ if and only if $\mu_X|A; (x, y) \in A = 0 \text{ a.e. } d\mu_Y(y)$ and $\mu_Y|A; (x, y) \in A = 0 \text{ a.e. } d\mu_X(x)$. It follows easily that $\mu_X \otimes \mu_Y \sim \mu_Z \otimes \mu_W$.

For the converse, one notes that $\mu_X(B) = \mu_X \otimes \mu_Y(B \times H_2)$. We now examine the equivalence or singularity of the joint Gaussian measures $\mu_{XY}$ and $\mu_{ZW}$. We have seen that $\mathcal{R}_{XY} = \mathcal{R}_{XY}^\perp + \mathcal{U}$ and $\mathcal{R}_{ZW} = \mathcal{R}_{ZW}^\perp + \mathcal{U}$, where $\mathcal{U}(u, v) = (\mathcal{V}_u, \mathcal{V}_v)$, $\mathcal{F}(u, v) = (\mathcal{V}_u, \mathcal{V}_v)$, and $\mathcal{Q}(u, v) = (\mathcal{V}_u, \mathcal{V}_v)$. If $\text{range}(\mathcal{R}_{XY}^\perp) = \text{range}(\mathcal{R}_{ZW}^\perp)$ and $\text{range}(\mathcal{R}_{ZW}) = \text{range}(\mathcal{R}_{XY})$, then by Lemma 1 there exist bounded linear operators $B_1: H_1 \to H_1$ and $B_2: H_2 \to H_2$ such that both $B_1$ and $B_2$ have bounded inverse, and $\mathcal{R}_{ZW}^\perp = \mathcal{R}_{XY}^\perp B_1$, $\mathcal{R}_{ZW} = \mathcal{R}_{XY}^\perp B_2$. We can then define an operator $\mathcal{B}$ in $H_1 \times H_2$ by $\mathcal{B}(u, v) = (B_1 u, B_2 v)$. $\mathcal{B}$ is linear, bounded, and has bounded inverse.

Theorem 4. $\mu_{XY} \sim \mu_{ZW}$ if and only if

(a) $\mu_X \sim \mu_Z$ and $\mu_Y \sim \mu_W$;

(b) there exists a Hilbert-Schmidt operator $\mathcal{B}$ in $H_1 \times H_2$ such that $\mathcal{B}$ does not have $-1$ as an eigenvalue, and

$$
\mathcal{J} + \mathcal{U} = \mathcal{B} \mathcal{B}^* + \mathcal{B} \mathcal{B}^* \mathcal{J} \mathcal{B} \mathcal{B}^* + \mathcal{J} \mathcal{B} \mathcal{B}^* \mathcal{J} \mathcal{B} \mathcal{B}^*;
$$

(c) $m_{XY} - m_{ZW}$ belongs to the range of $\mathcal{R}_{XY}^\perp \mathcal{B} \mathcal{B}^* + \mathcal{J} \mathcal{B} \mathcal{B}^* \mathcal{J}$.

Remark. Note that condition (a) implies (Lemma 2) that the operator $\mathcal{B}$ (defined above) exists and is linear, bounded, and has bounded inverse.

Proof. (A) Suppose $\mu_{XY} \sim \mu_{ZW}$. Condition (a) is obviously satisfied since, for example, $\mu_X(A) = \mu_{XY}(A \times H)$. This implies (Lemma 3) that $\mu_X \otimes \mu_Y \sim \mu_Z \otimes \mu_W$; since $\mathcal{R}_{ZW} = \mathcal{R}_{XY}^\perp + \mathcal{U}$, the operator $\mathcal{J} - \mathcal{R}_{ZW}$ is Hilbert-Schmidt and does not have $+1$ as an eigenvalue, by Lemma 2.

$\mu_{XY} \sim \mu_{ZW}$ also implies that $\mathcal{R}_{XY} = \mathcal{R}_{ZW}^\perp + \mathcal{K}$, where $\mathcal{K}$ is Hilbert-Schmidt, $\mathcal{K}$ can be taken as zero on the null space of $\mathcal{R}_{ZW}$, and $-1$ is not an eigenvalue of $\mathcal{K}$. Hence $\mathcal{R}_{XY} = \mathcal{R}_{XY}^\perp + \mathcal{R}_{XY} \mathcal{B} \mathcal{B}^* + \mathcal{R}_{XY} \mathcal{B} \mathcal{B}^* \mathcal{J}$, where $\mathcal{R}_{XY} = \mathcal{R}_{XY}^\perp + \mathcal{J}$. $\mathcal{R}_{XY}$ also can be written as $\mathcal{R}_{XY} = \mathcal{R}_{XY}^\perp + \mathcal{J} + \mathcal{J} \mathcal{R}_{XY}$. Hence

$$
\mathcal{J} + \mathcal{U} = \mathcal{B} \mathcal{B}^* \mathcal{B}^* + \mathcal{B} \mathcal{B}^* \mathcal{J} \mathcal{B} \mathcal{B}^* + \mathcal{J} \mathcal{B} \mathcal{B}^* \mathcal{J} \mathcal{B} \mathcal{B}^*;
$$

where $\mathcal{B} = \mathcal{B} \mathcal{B}^* \mathcal{B}^* + \mathcal{B} \mathcal{B}^* \mathcal{J} \mathcal{B} \mathcal{B}^*$ the partially isometric operator satisfying $\mathcal{R}_{ZW} = \mathcal{R}_{XY}^\perp + \mathcal{J} \mathcal{R}_{XY} \mathcal{B} \mathcal{B}^* + \mathcal{J} \mathcal{R}_{XY} \mathcal{B} \mathcal{B}^* \mathcal{J}$, $\mathcal{B} = \mathcal{J} \mathcal{B} \mathcal{B}^*$. Condition (b) is satisfied.

Finally, $\mu_{XY} \sim \mu_{ZW}$ implies that $m_{XY} - m_{ZW}$ belongs to range($\mathcal{R}_{XY}^\perp$). Since $\mathcal{R}_{XY} = \mathcal{R}_{XY}^\perp + \mathcal{J}$, $\mathcal{B}$ partially isometric and isometric on range($\mathcal{R}_{XY}^\perp$), one
sees that $m_{XY} - m_{ZW} = R^{1/2}_{XY} [f + \Omega]^{1/2} m$ for some $m$ in $H_1 \times H_2$.

(B) Suppose conditions (a), (b) and (c) are satisfied. Condition (a) implies the existence of the operator $\mathcal{B}, R_{Z\omega W} = R^{1/2}_{X\omega Y} \mathcal{B}^* R^{1/2}_{X\omega Y}, \mathcal{B}^* - \mathcal{B}^*$ Hilbert-Schmidt, and \(\mathcal{B}^*\) having bounded inverse. Assumption (b) then implies that $R_{XY} - R_{ZW} = K_{ZW} \mathcal{B}^*$ where $K = D \otimes D^*$, $D$ partially isometric, isometric on range($R_{ZW}$) and zero on range($R_{ZW}^{1/2}$). $K$ is thus Hilbert-Schmidt, vanishes on the null space of $R_{ZW}$, and does not have $-1$ as an eigenvalue. Assumption (c) implies $m_{XY} - m_{ZW}$ in range($R^{1/2}_{XY}$), by Proposition 2. Conditions (a), (b) and (c) thus imply $\mu_{XY} \sim \mu_{ZW}$ by Lemma 2, and the theorem is proved.

The necessary and sufficient conditions given in Theorem 4 are completely general. However, they are stated in terms of operators in $H_1 \times H_2$, whereas one might prefer conditions given in terms of operators in $H_1$ and in $H_2$. The next result gives such conditions for equivalence for a wide class of joint Gaussian measures.

**Theorem 5.** Suppose that $\|V^* x\|_2 < \alpha \|x\|_1$ for all $x$ in range($I - V V^*$), some $\alpha < 1$. Then $\mu_{XY} \sim \mu_{ZW}$ if and only if

(a) $\mu_X \sim \mu_Z$ and $\mu_Y \sim \mu_W$;
(b) $V - B_1 T B_2^*$ is Hilbert-Schmidt;
(c) there exists a finite and strictly positive scalar $\lambda$ such that

$$\lambda B_1 (I - T T^*) B_1^* \leq I - V V^* \leq \lambda^{-1} B_1 (I - T T^*) B_1^*$$

and

$$\lambda B_2 (I - T^* T) B_2^* \leq I - V^* V \leq \lambda^{-1} B_2 (I - T^* T) B_2^*.$$ 

(d) there exist elements $u$ in $H_1$ and $v$ in $H_2$ such that $m_X - m_Z = R^{1/2}_X (u + V v)$ and $m_Y - m_W = R^{1/2}_Y (v + V^* u)$.

**Proof.** (A) Suppose $\mu_{XY} \sim \mu_{ZW}$. Condition (a) is necessary, from Theorem 4, and implies that $\mathcal{B} - \mathcal{B}^*$ is Hilbert-Schmidt. Condition (b) is then implied by condition (b) of Theorem 4. To see that (c) holds, one notes that range($R^{1/2}_{XY}$) = range($R^{1/2}_{ZW}$) implies the existence of strictly positive and finite scalars $\beta_1$ and $\beta_2$ such that $\beta_1 R_{ZW} \leq R_{XY} \leq \beta_2 R_{ZW}$, from Lemma 1. This is equivalent to $\beta_1 B (\mathcal{B} + \mathcal{B}) B^* \leq \mathcal{B} + \mathcal{B}^* \leq \beta_2 B (\mathcal{B} + \mathcal{B}) B^*$, which is equivalent to range($\mathcal{B} (\mathcal{B} + \mathcal{B}) B^*$) = range($\mathcal{B} (\mathcal{B} + \mathcal{B}) B^*$). This implies that the null space of $I - V V^*$ is identical to the null space of $B_1 (I - T T^*) B_1^*$ and that the null space of $I - V^* V$ is identical to the null space of $B_2 (I - T^* T) B_2^*$. The assumption that $\|V^* x\|_2$ is bounded away from $\|x\|_1$ for $x$ in range($I - V V^*$) implies, by Theorem 3, that $\mathcal{B} + \mathcal{B} = \mathcal{B} (\mathcal{B} + \mathcal{B}) B^*$, and $I - V V^*$ each has closed range. Hence, $(\mathcal{B} + \mathcal{B}) B^*$ has closed range, and range($\mathcal{B} + \mathcal{B}$) = range($\mathcal{B}^*$), so that range($\mathcal{B} (\mathcal{B} + \mathcal{B}) B^*$) is closed. This
implies that the range spaces of $B_1[I - TT^*]B_1^*$ and $B_2[I - TT^*]B_2^*$ are closed, so that $\text{range}(I - VV^*) = \text{range} B_1[I - TT^*]B_1^*$ and $\text{range}(I - V^*V) = \text{range} B_2[I - TT^*]B_2^*$. Since these range spaces are all closed, condition (c) follows.

Finally, $\mu_{XY} \sim \mu_{ZW}$ implies that $m_{XY} - m_{ZW}$ belongs to the range of $R_{XY}^{1/2}$, so that $m_{XY} - m_{ZW} = R_{X \otimes Y}^{1/2}(I + \mathcal{O})^{1/2}m$ for some $m$ in $H_1 \times H_2$. This implies condition (d), since $\text{range}(I + \mathcal{O}^{1/2}) = \text{range}(I + \mathcal{O})$.

(B) Suppose conditions (a)-(d) are satisfied. (a) implies that $R_{X \otimes Y} = R_{Z \otimes W}^* R_{Z \otimes W}^*$, where $I - \mathcal{B}^*$ has bounded inverse. The assumption that $\|V^*x\|_2 \leq \alpha \|x\|_1$ for all $x$ in range$(I - VV^*)$, some $\alpha < 1$, implies that $I - VV^*$, $I - V^*V$, and $I + \mathcal{O}$ have closed range. Condition (c) then implies $\text{range}[B_1(I - TT^*)B_1^*] = \text{range}(I - VV^*)$ and $\text{range}(I - V^*V) = \text{range}(B_2[I - TT^*]B_2^*)$, so that $\text{range}(B[I + \mathcal{O}]^*) = \text{range}(I + \mathcal{O})$. Since range$(I + \mathcal{O}^{1/2})$ is closed, this implies that $\text{range}(I + \mathcal{O}^{1/2}) = \text{range}(B[I + \mathcal{O}]^*)^{1/2}$, so that by Lemma 1 there exist finite and strictly positive scalars $\beta_1, \beta_2$ such that

$$\beta_1[I + \mathcal{O}]^* \leq I + \mathcal{O} \leq \beta_2[I + \mathcal{O}]^*.$$ 

This is equivalent to $\beta_1 R_{ZW} \leq R_{XY} \leq \beta_2 R_{ZW}$, which implies $\text{range}(R_{XY}^{1/2}) = \text{range}(R_{ZW}^{1/2})$. Thus, there exists a bounded linear operator $K$ such that $K = 0$ on the null space of $R_{ZW}^{1/2}$, with $R_{X \otimes Y} = R_{Z \otimes W}^{1/2} I + K R_{ZW}^{1/2}$, and $I + K$ is bounded away from zero on $\text{range}(R_{ZW}^{1/2})$.

One now has

$$I + \mathcal{O} = K[I + \mathcal{O}]^* + B \mathcal{K} \mathcal{A} \mathcal{B}^*$$

where $\mathcal{G} \mathcal{A}^* = I + \mathcal{F}$, or $I - \mathcal{B}^* \mathcal{C} + \mathcal{O} - \mathcal{B} \mathcal{B}^* = B \mathcal{K} \mathcal{A} \mathcal{B}^*$. Since $B$ has bounded inverse, and $I - \mathcal{B}^* \mathcal{C}$ is Hilbert-Schmidt, condition (b) implies that $\mathcal{K} \mathcal{A} \mathcal{B}^*$ is Hilbert-Schmidt. The null space of $K$ contains the null space of $R_{ZW}^{1/2}$, so that $\text{range}(K) \subseteq \text{range}(R_{ZW}^{1/2})$. But $R_{ZW}^{1/2} = \mathcal{G} \mathcal{A}^* R_{ZW}^{1/2}$, so that $\text{range}(K) \subseteq \text{range}(\mathcal{G} \mathcal{A}^*)$. The assumption that $\|Vu\|_2$ is bounded away from $\|u\|_1$ for $u$ in range$(I - V^*V)$, and condition (c), imply that $\|\mathcal{G}(u, v)\|_2^2 \geq k \|(u, v)\|_2$ for all $(u, v)$ in range$(\mathcal{G} \mathcal{A}^*)$ and some $k > 0$. Moreover, range$(\mathcal{G} \mathcal{A}^*) = \text{range}(\mathcal{G} \mathcal{A} \mathcal{B}^*)^2) = \text{range}(\mathcal{G} \mathcal{A}^*)$. Hence if $\{(u_n, v_n)\}$ is a c.o.n. set in range$(\mathcal{G} \mathcal{A}^*)$, then

$$\sum_n \|\mathcal{G} \mathcal{A}^*(u_n, v_n)\|_2^2 \geq k \sum_n \|K \mathcal{A}^*(u_n, v_n)\|_2^2 \geq k^2 \sum_n \|K(u_n, v_n)\|_2^2,$$

so that $K$ is Hilbert-Schmidt and does not have $-1$ as an eigenvalue (the latter because $I + K$ is invertible).

Finally, it is easy to show that condition (d) implies $m_{XY} - m_{ZW}$ belongs
to the range of $\mathcal{R}_{XY}^{\frac{1}{2}}$. This concludes the proof.

In the case where $\mu_{ZW} = \mu_X \otimes \mu_Y$, a simple and general condition for equivalence can be given.

**Theorem 6.** $\mu_{XY} \sim \mu_X \otimes \mu_Y$ if and only if $\|V\| < 1$ and $V$ is Hilbert-Schmidt.

**Proof.** $\mathcal{R}_{XY} = \mathcal{R}_{X}^{\frac{1}{2}} \otimes (I + \mathcal{V}) \mathcal{R}_{Y}^{\frac{1}{2}}$. From Theorem 3, $\mathcal{V}$ is Hilbert-Schmidt if and only if $V$ is Hilbert-Schmidt, and $\mathcal{O}$ has $-1$ as an eigenvalue if and only if $\mathcal{V}V^*$ has $+1$ as an eigenvalue. If $V$ is Hilbert-Schmidt, $\|V\| = 1$ if and only if $\mathcal{V}V^*$ has $1$ as an eigenvalue, since the set of limit points of the spectrum of a selfadjoint compact operator contains only zero. The result follows from Lemma 2.

**Corollary.** $\mu_{XY} \sim \mu_{ZW}$ if

(a) $\mu_X \sim \mu_Z$ and $\mu_Y \sim \mu_W$;

(b) $\|V\| < 1$ and $V$ is Hilbert-Schmidt;

(c) $\|T\| < 1$ and $T$ is Hilbert-Schmidt.

**Proof.** (b) implies $\mu_{XY} \sim \mu_X \otimes \mu_Y$, (c) implies $\mu_{ZW} \sim \mu_Z \otimes \mu_W$, and (a) implies $\mu_X \otimes \mu_Y \sim \mu_Z \otimes \mu_W$.

We give an example of two equivalent joint Gaussian measures which satisfy neither condition (b) nor condition (c) of the corollary. Suppose that $\mu_X \sim \mu_Z$, and let $W$ be any bounded linear operator mapping $H_1 \rightarrow H_2$. Define $\mu_Y$ by

$$\mu_Y[B] = \mu_X[x; Wx \in B] \text{ for } B \in \mathcal{F}_2,$$

and define $\mu_W$ by $\mu_W[B] = \mu_Z[x; Wx \in B]$. Also define $\mu_{XY}$ and $\mu_{ZW}$ by

$$\mu_{XY}[C] = \mu_X[x; (x, Wx) \in C], \quad C \in \mathcal{F}_1 \times \mathcal{F}_2,$$

$$\mu_{ZW}[C] = \mu_Z[x; (x, Wx) \in C].$$

Then $\mu_{XY} \sim \mu_{ZW}$, and both $\mu_X$ and $\mu_Z$ are Gaussian if $\mu_X$ and $\mu_Z$ are Gaussian. However, $\mathcal{R}_{XY} = \mathcal{R}_X^{\frac{1}{2}} U \mathcal{R}_Y^{\frac{1}{2}}$ ($\mathcal{R}_Y = \mathcal{W} \mathcal{R}_X \mathcal{W}^*$, and $\mathcal{R}_{XY} = \mathcal{R}_Z \mathcal{W}^*$), where $U$ is a partially isometric map of $H_2$ into $H_1$, similarly, $\mathcal{R}_{ZW} = \mathcal{R}_Z^2 \mathcal{S} \mathcal{R}_W^2$, where $S$ is a partial isometry of $H_2$ into $H_1$. Thus $\|U\| = \|S\| = 1$, and neither $U$ nor $S$ is Hilbert-Schmidt if range($\mathcal{R}_Y$) is infinite dimensional.

From the proof of the corollary, it is clear that $\mu_{XY} \perp \mu_{ZW}$ if (b) or (c) but not both (b) and (c) of the corollary is satisfied.

**Applications to information theory.** The average mutual information (AMI) of a joint measure $\mu_{XY}$ is defined as

$$\text{AMI}(\mu_{XY}) = \int_{H_1 \times H_2} \frac{d\mu_{XY}}{d\mu_X \otimes \mu_Y}(u, v) \log \left[ \frac{d\mu_{XY}}{d\mu_X \otimes \mu_Y}(u, v) \right] d\mu_X \otimes \mu_Y(u, v)$$
if $\mu_{XY} \ll \mu_X \otimes \mu_Y$, and equal to $+\infty$ otherwise. For Gaussian $\mu_{XY}$, AMI($\mu_{XY}$) is clearly finite if and only if $\mu_{XY} \sim \mu_X \otimes \mu_Y$. The most comprehensive results on AMI are due to Gel’fand and Yaglom [5]. In [1], necessary and sufficient conditions for AMI($\mu_{XY}$) to be finite, stated in terms of covariance operators, were obtained for $\mu_{XY}$ Gaussian, and $H_1 = H_2$. That result was an extension of the work of Gel’fand and Yaglom. Theorem 6 gives necessary and sufficient conditions for AMI($\mu_{XY}$) to be finite without requiring $H_1 = H_2$, provided $\mu_{XY}$ is Gaussian, and is independent of the results of Gel’fand and Yaglom.

Our final result is of interest in information theory applications. Let $H_1 = H_2 = H, \Gamma_1 = \Gamma_2 = \Gamma$, and consider the measure $\mu_{Y-X}$ defined by $\mu_{Y-X}(A) = \mu_{XY}^{-1}(\{x, y): y - x \in A\}$ (this is a well-defined Gaussian measure, since $f(u, v) = v - u$ is a continuous linear map from $(H \times H, \Gamma \times \Gamma)$ to $(H, \Gamma)$). In information theory applications, $\mu_{Y-X}$ represents "noise", $\mu_X$ "signal" and $\mu_Y$ "signal-plus-noise". Typically, these measures are induced by measurable mean-square continuous stochastic processes, with $H_1 = H_2 = H = L_2[0, T]$ for some finite $T$. Of interest is the AMI($\mu_{XY}$) (the "average information about the signal obtained by observing signal-plus-noise") and the equivalence or orthogonality of $\mu_X$ and $\mu_{Y-X}$. For, if AMI($\mu_{XY}$) $< \infty$, one might intuitively expect that $\mu_{Y-X} \sim \mu_Y$, and conversely, since $\mu_Y \perp \mu_{Y-X}$ allows one to discriminate perfectly between noise and signal-plus-noise. It is known that $\mu_{Y-X} \sim \mu_Y$ does not imply AMI($\mu_{XY}$) $< \infty$ [7], [1]. In the case where $\mu_{X-Y,X} = \mu_{Y-X} \otimes \mu_X$, Hájek [7] has shown that AMI($\mu_{XY}$) $< \infty$ if and only if $\mu_Y$ and $\mu_{Y-X}$ are strongly equivalent; i.e., $R_{Y-X} = R_Y^Q(I + W)R_Y^Q$, where $W$ is trace-class and does not have $-1$ as an eigenvalue.

The significance of strong equivalence is partly because one can then explicitly express the Radon-Nikodym derivative in series form [7], [11]. In the following, we show that not only is strong equivalence of $\mu_{Y-X}$ and $\mu_Y$ not equivalent to $\mu_{XY} \sim \mu_X \otimes \mu_Y$ when $\mu_{Y-X} \neq \mu_{Y-X} \otimes \mu_X$, but also that $\mu_{XY} \sim \mu_X \otimes \mu_Y$ does not even imply that $\mu_{Y-X} \sim \mu_Y$.

Theorem 7. If $\mu_{XY} \sim \mu_X \otimes \mu_Y$, and $\|m_X\| = \|m_Y\| = 0$, then $\mu_{Y-X} \sim \mu_Y$ if and only if $R_Y = R_Y^Q Q R_Y^Q$ for $Q$ Hilbert-Schmidt.

Proof. Define $\mu_{Y \otimes X}$ by $\mu_{Y \otimes X}(A) = \mu_X \otimes \mu_Y^{-1}(x, y): y - x \in A, A \in \Gamma$. Note that $\mu_{XY} \sim \mu_X \otimes \mu_Y \Rightarrow \mu_{Y-X} \sim \mu_{Y \otimes X}$.

Suppose first that $\mu_{Y-X} \sim \mu_Y$, then $\mu_Y \sim \mu_{Y \otimes X}$, and $\mu_{Y} \sim \mu_{Y \otimes X} = R_Y \otimes \mu_Y \sim R_Y \otimes Q R_Y^Q$, $Q$ Hilbert-Schmidt, $-1$ not an eigenvalue of $Q$. But $R_Y \otimes Q \sim R_Y + R_Y Q R_Y^Q$, hence $R_Y = R_Y Q R_Y^Q$, where $Q$ is Hilbert-Schmidt.

Conversely, suppose $R_Y = R_Y^Q Q R_Y^Q$, $Q$ Hilbert-Schmidt. Then $\mu_{Y \otimes X} \sim \mu_Y$, since $Q$ is nonnegative. Hence $\mu_Y \sim \mu_{Y-X}$.

As an example where $\mu_{XY} \sim \mu_X \otimes \mu_Y$ but $\mu_{Y-X} \perp \mu_Y$, one can take any co-
variance operator $R_Y$ such that $\| R_Y \| < 1$, and define $R_X = R_Y$, $R_{XY} = R_Y^{3/2}$. Then the Gaussian joint measure $\mu_{XY}$ having $\mu_X$ and $\mu_Y$ as projections and $R_Y^{3/2}$ as cross-covariance operator is equivalent to $\mu_X \otimes \mu_Y$. However, $\mu_{Y-X} \perp \mu_Y$ by Theorem 7, since $R_X = R_Y^{3/2}R_Y^{3/2}$.

Although the result of Hájek quoted above does not hold when $\mu_{Y-X,X} \neq \mu_{Y-X} \otimes \mu_X$, it is true that if $R_Y \geq R_{Y-X}$, then $\mu_{XY} \sim \mu_X \otimes \mu_Y \Rightarrow \mu_{Y-X}$ is strongly equivalent to $\mu_Y$ [1].

These results indicate that the concept of mutual information is not completely satisfactory, since one may be able to discriminate perfectly between $\mu_X$ and $\mu_{Y-X}$ while having only a finite value of $\text{AMI}(\mu_{XY})$, or $\text{AMI}(\mu_{XY})$ can be infinite while it is impossible to discriminate perfectly between $\mu_X$ and $\mu_{Y-X}$.

REFERENCES


DEPARTMENT OF STATISTICS, UNIVERSITY OF NORTH CAROLINA AT CHAPEL HILL, CHAPEL HILL, NORTH CAROLINA 27514