

# Krein Factorization of Covariance Operators of 2-Parameter Random Fields and Application to the Likelihood Ratio

Rob Luesink and Arunabha Bagchi, *Member, IEEE*

**Abstract**—The covariance operator of a signal-plus-noise observation for a 2-parameter random field is factorized as products of Volterra type operators. As a result of this factorization we can express the determinant of this operator in terms of the system parameters. This determinant appears in the likelihood ratio that plays a significant role in identification and detection problems.

**Index Terms**—Random fields, likelihood ratio, Krein factorization.

## I. INTRODUCTION

CONSIDER a 2-parameter random field  $\{X_t\}$  on some compact set  $\mathcal{D}$  in  $\mathbb{R}^2$ . Let  $\{Y_t\}$  be the observation process of this field. For purposes of identification and signal detection, we are interested in determining the likelihood ratio of the measure induced by the observation process with respect to the standard Gaussian measure. In [7], Shepp gives an expression for the likelihood ratio in the 1-parameter case. This formula of Shepp involves the determinant  $\det(I+R)$  of the covariance operator of the observation process. In practice, however, this term is not readily computable.

For 1-parameter two-point boundary value process models, Bagchi and Westdijk [3] succeeded in expressing this determinant in terms of the system parameters. The key was to use Krein factorization, which has already been applied earlier for this purpose in the Markov case [4].

In this paper we give an explicit Krein factorization formula for the operator  $(I+R)$  in the 2-parameter case when  $\mathcal{D}$  is a rectangle. It turns out that we can factorize the operator in this case into a product of four "Volterra" type operators. An extension of our result to the multiparameter case is straightforward. For an abstract of this type of factorization, see Bromley and Kallianpur [6].

The likelihood ratio for the Gaussian signal-plus-noise model is derived next. The expression is very similar to that obtained by Shepp [7] in the 1-parameter case. The determinant term in the likelihood ratio may be expressed in the system parameters using the Krein factorization of  $(I+R)$ . In the 2-parameter situation, Wong and Zakai [8] gave an expression for the likelihood ratio using the 2-parameter Brownian motion model for the observation. This formula has to be modified in practice by a correction term when using real data. This is due to the fact that the observation

noise can never be a true 2-parameter Brownian motion (in the integrated form). Using a direct white noise model for the observation noise, Bagchi [1] derived an expression for the likelihood ratio which differed from the Wong-Zakai formula. In analogy with the 1-parameter case, it may be inferred that this difference is precisely the correction term previously mentioned. The determinant term calculated in the present paper may give new possibilities for expressing the correction term. We do not pursue this point any further in this paper.

## II. KREIN FACTORIZATION FOR 2-PARAMETER COVARIANCE OPERATORS

We first consider a 1-parameter stochastic process  $\{X_t\}$  in  $L_2(\Omega)$ ,  $t \in [0, T]$ , taking values in  $\mathbb{R}^n$ , with Hilbert-Schmidt covariance operator  $R: L_2^n([0, T]) \rightarrow L_2^n([0, T])$ , i.e.,  $Rf(t) = \int_0^T r(t, s)f(s)ds$  with  $n \times n$  matrix valued kernel  $r(t, s) = \text{cov}(X_t, X_s)$ . Let  $\{Y_t\}$ ,  $t \in [0, T]$  be the measurement process

$$Y_t = X_t + n_t \quad (2.1)$$

with  $\{n_t\}$ ,  $t \in [0, T]$ , a Gaussian standard white noise independent of  $X$ . Then the observation process has covariance operator  $R_y = I + R$ . Following a result of Krein [5, p. 130] we can factorize  $(I+R)^{-1}$  as

$$(I+R)^{-1} = (I-L^*)(I-L) \quad (2.2)$$

where  $L$  is a Volterra operator on  $L_2^n([0, T])$  and  $L^*$  is its adjoint, with  $Lf(t) = \int_0^t l(t, s)f(s)ds$ . Suppose further that  $L+L^*$  is trace class. Then one can prove [4, p. 232] that

$$\log(\det(I+R)) = \text{tr}(L+L^*) = \int_0^T \text{tr} l(t, t) dt, \quad (2.3)$$

where  $\text{tr}$  denotes the trace of an operator and  $\text{tr}$  denotes the trace of a matrix.

Our purpose is to extend this result to the 2-parameter case when the domain  $\mathcal{D}$  is a rectangle. Let  $\mathcal{D} = [0, T_1] \times [0, T_2]$ ,  $\mathcal{H} = L_2^n(\mathcal{D})$  and consider a 2-parameter  $n$ -vector value random field  $\{X_t\}$ ,  $t = (t_1, t_2) \in \mathcal{D}$ . Let  $R: \mathcal{H} \rightarrow \mathcal{H}$  be the H-S covariance operator of the random field  $\{X_t\}$ ; that is,  $Rf(t) = \int_{\mathcal{D}} r(t, s)f(s)ds$  ( $ds = ds_2 ds_1$ ), with  $r(t, s) = \text{cov}(X_t, X_s)$  continuous in  $\mathcal{D}$  and  $\int_{\mathcal{D}} \int_{\mathcal{D}} \|r(t, s)\|^2 ds dt < \infty$ . Then we obtain the following factorization.

*Theorem 2.1:*

$$(I+R) = (I-P_2)(I-P_1)(I-P_1^*)(I-P_2^*) \quad (2.4)$$

Manuscript received November 8, 1989; revised April 20, 1990.  
The authors are with the Department of Applied Mathematics, University of Twente, P.O. Box 217, 7500 AE Enschede, The Netherlands.  
IEEE Log Number 9038859.

with  $P_1$  and  $P_2$  given by

$$\begin{aligned} P_1 f(t) &= \int_0^{t_1} \int_0^{t_2} p(t, s) f(s) ds, \\ P_2 f(t) &= \int_0^{t_1} \int_{t_2}^{T_2} p(t, s) f(s) ds \end{aligned} \quad (2.5)$$

and  $P_1^*$  and  $P_2^*$  their adjoints,

$$\begin{aligned} P_2^* f(t) &= \int_{t_1}^{T_1} \int_0^{t_2} p^T(s, t) f(s) ds, \\ P_1^* f(t) &= \int_{t_1}^{T_1} \int_{t_2}^{T_2} p^T(s, t) f(s) ds \end{aligned}$$

$p(t, s)$  is a continuous kernel, which will be derived in the proof.

*Proof:* The proof depends on the following three lemmas.

*Lemma 2.1:* The integral equation

$$g(q; t, s) + \int_0^q \int_0^{T_2} r(t, z) g(q; z, s) dz = r(t, s) \quad (2.6)$$

in the area  $(q, t, s) \in [0, T_1] \times \mathcal{L} \times \mathcal{L}$ , has a unique  $n \times n$  matrix valued solution  $g(q; t, s)$ . Furthermore, this solution is continuous in  $(t, s)$  on  $\mathcal{L} \times \mathcal{L}$  and is continuously differentiable in  $q$  in  $[0, T_1]$ . In fact,  $g(q; t, s)$  satisfies the implicit equation

$$-\frac{\partial}{\partial q} g(q; t, s) = \int_0^{T_2} g(q; t, (q, \sigma)) g(q; (q, \sigma), s) d\sigma. \quad (2.7)$$

*Proof:* Let  $R_q: \mathcal{H} \rightarrow \mathcal{H}$  be defined by  $R_q f(t) = \int_0^q \int_0^{T_2} r(t, z) f(z) dz$ . Then (2.6) can be rewritten as  $(I + R_q)g(q; \cdot, s)(t) = r(t, s)$ . Since  $R_q$  is in fact the covariance operator of the system restricted to  $[0, q] \times [0, T_2]$ , we have  $-1 \notin \sigma(R_q)$ , where  $\sigma(\cdot)$  denotes the spectrum of an operator. This implies that (2.6) has a unique solution for any fixed  $s$  and  $q$ .

The kernel  $r(t, s)$  is continuous in  $t$ , so that  $\int_0^q \int_0^{T_2} r(t, z) g(q; z, s) dz$  is continuous in  $t$ , and hence (2.6) implies that  $g(q; t, s)$  is continuous in  $t$ . To prove continuity in  $s$  we note that

$$\begin{aligned} g(q; t, s_1) - g(q; t, s_2) &+ \int_0^q \int_0^{T_2} r(t, z) (g(q; z, s_1) - g(q; z, s_2)) dz \\ &= r(t, s_1) - r(t, s_2). \end{aligned}$$

Hence,

$$(I + R_q)(g(q; \cdot, s_1) - g(q; \cdot, s_2))(t) = r(t, s_1) - r(t, s_2).$$

Since  $\|(I + R_q)\|_1 \geq 1$  we obtain

$$\|g(q; \cdot, s_1) - g(q; \cdot, s_2)\|_2 \leq \|r(\cdot, s_1) - r(\cdot, s_2)\|_2$$

where  $\|P\|_1 = \sup \sigma(P)$  and  $\|\cdot\|_2$  denotes the  $L_2$  norm over  $t \in \mathcal{L}$ . The facts that  $r(t, s)$  is continuous in  $s$ , and  $g(q; t, s)$  is continuous in  $t$  imply that  $g(q; t, s)$  is continuous in  $s$ .

Further,

$$\begin{aligned} g(q + \Delta; t, s) - g(q; t, s) &+ \int_0^q \int_0^{T_2} r(t, z) (g(q + \Delta; z, s) - g(q; z, s)) dz \\ &= - \int_q^{q+\Delta} \int_0^{T_2} r(t, z) g(q + \Delta; z, s) dz. \end{aligned}$$

Hence

$$\begin{aligned} (I + R_q)(g(q + \Delta; \cdot, s) - g(q; \cdot, s))(t) &= - \int_q^{q+\Delta} \int_0^{T_2} r(t, z) g(q + \Delta; z, s) dz. \end{aligned}$$

for any fixed  $s$  and  $q$ . Since  $\|(I + R_q)\|_1 \geq 1$  we obtain

$$\|g(q + \Delta; \cdot, s) - g(q; \cdot, s)\|_2 \leq \left\| \int_q^{q+\Delta} \int_0^{T_2} r(\cdot, z) g(q + \Delta; z, s) dz \right\|_2$$

and

$$\lim_{\Delta \rightarrow 0} \|g(q + \Delta; \cdot, s) - g(q; \cdot, s)\|_2 = 0.$$

by the continuity of  $g(q; t, s)$  in  $t$ . Hence  $g(q; t, s)$  is continuous in  $q$ . Moreover

$$\begin{aligned} \|(g(q + \Delta; \cdot, s) - g(q; \cdot, s))/\Delta\|_2 &\leq \left\| - \int_q^{q+\Delta} \int_0^{T_2} r(\cdot, z) g(q + \Delta; z, s) dz / \Delta \right\|_2. \end{aligned}$$

Since  $r(t, s)$  and  $g(q; t, s)$  are continuous in all the arguments and  $[0, T_2]$  is bounded, we obtain

$$\begin{aligned} \lim_{\Delta \rightarrow 0} \|(g(q + \Delta; \cdot, s) - g(q; \cdot, s))/\Delta\|_2 &\leq \lim_{\Delta \rightarrow 0} \left\| - \int_q^{q+\Delta} \int_0^{T_2} r(\cdot, z) g(q + \Delta; z, s) dz / \Delta \right\|_2 \\ &= \left\| \int_0^{T_2} r(\cdot, (q, \sigma)) g(q; (q, \sigma), s) d\sigma \right\|_2 < \infty, \end{aligned}$$

for all fixed  $q$  and  $s$ . Hence  $\partial/\partial q g(q; t, s)$  exists. Finally we deduce from (2.6) that

$$\begin{aligned} -\frac{\partial}{\partial q} g(q, t, s) &= \int_0^{T_2} r(t, (q, \sigma)) g(q, (q, \sigma), s) d\sigma \\ &+ \int_0^q \int_0^{T_2} r(t, z) \frac{\partial}{\partial q} g(q, z, s) dz, \end{aligned} \quad (2.8)$$

$$\begin{aligned} \int_0^{T_2} r(t, (q, \sigma)) g(q, (q, \sigma), s) d\sigma &= \int_0^{T_2} g(q, t, (q, \sigma)) g(q, (q, \sigma), s) d\sigma \\ &+ \int_0^{T_2} \left( \int_0^q \int_0^{T_2} r(t, z) g(q, z, (q, \sigma)) dz \right) d\sigma. \end{aligned} \quad (2.9)$$

Comparing (2.8) and (2.9) we find (2.7) that immediately implies that  $\partial/\partial q g(q, t, s)$  is continuous.  $\square$

*Lemma 2.2:* Let  $L: \mathcal{H} \rightarrow \mathcal{H}$  be defined by  $Lf(t) = \int_0^{t_1} \int_0^{T_2} l(t, s) f(s) ds$  with  $l(t, s) = g(t_1, t, s)$  and  $g$  solution of (2.6). Then

$$(I + R)^{-1} = (I - L^*)(I - L). \quad (2.10)$$

*Proof:* Let  $L^*$  be the adjoint of  $L$ ,  $L^* f(t) = \int_{t_1}^{T_1} \int_0^{T_2} l^T(s, t) f(s) ds$ . Then, since  $l^T(s, t) = g^T(s_1, s, t) =$

$g(s_1, t, s)$ ,

$$\begin{aligned} (L + L^* - L^*L)f(t) &= \int_0^{t_1} \int_0^{T_2} g(t_1; t, s) f(s) ds + \int_{t_1}^{T_1} \int_0^{T_2} g(s_1; t, s) f(s) ds \\ &\quad - \int_{t_1}^{T_1} \int_0^{T_2} g(s_1; t, s) \left( \int_0^{s_1} \int_0^{T_2} g(s_1; s, z) f(z) dz \right) ds \\ &= \int_{\mathcal{D}} k(t, s) f(s) ds \end{aligned}$$

with

$$k(t, s) = \begin{cases} g(t_1; t, s) - \int_{t_1}^{T_1} \int_0^{T_2} g(z_1; t, z) g(z_1; z, s) dz, & \text{if } s_1 \leq t_1 \\ g(s_1; t, s) - \int_{s_1}^{T_1} \int_0^{T_2} g(z_1; t, z) g(z_1; z, s) dz & \text{if } s_1 > t_1. \end{cases}$$

Using (2.7) we derive

$$k(t, s) = \begin{cases} g(t_1; t, s) + \int_{t_1}^{T_1} \frac{\partial}{\partial z_1} g(z_1; t, s) dz_1 & \text{if } s_1 \leq t_1 \\ g(s_1; t, s) + \int_{s_1}^{T_1} \frac{\partial}{\partial z_1} g(z_1; t, s) dz_1, & \text{if } s_1 > t_1 \end{cases} = g(T_1; t, s).$$

This implies, using (2.6) that

$$\begin{aligned} (I + R)(I - L - L^* + L^*L)f(t) &= (I + R) \left( f(\cdot) - \int_{\mathcal{D}} g(T_1, \cdot, s) f(s) ds \right) (t) \\ &= f(t) + \int_{\mathcal{D}} \left( -g(T_1, t, s) + r(t, s) - \int_{\mathcal{D}} r(t, z) g(T_1, z, s) dz \right) f(s) ds \\ &= f(t), \end{aligned}$$

or

$$(I + R)^{-1} = (I - L^*)(I - L). \quad \square$$

*Remark:* Lemma 2.2 appears to be the ordinary Krein factorization in the 2-parameter case. But the form of  $L$  clearly shows that it does not have the ‘‘Volterra’’ structure, as in the 1-parameter case. We show that  $(I - L)^{-1}$  and  $(I - L^*)^{-1}$  can, in turn, be factorized and this will lead to the desired 2-parameter Krein factorization.

*Lemma 2.3:* The integral equation

$$h(q; t, s) - \int_{s_1}^{t_1} \int_0^q l(t, z) h(q; z, s) dz = -l(t, s) \quad (2.11)$$

in the area  $(q, t, s) \in [0, T_2] \times \mathcal{D} \times \mathcal{D}$ , has a unique solution. Furthermore, this solution  $h(q; t, s)$  is continuous in  $(t, s)$  on  $\mathcal{D} \times \mathcal{D}$  and is continuously differentiable in  $q$  in  $[0, T_2]$ . In fact,  $h(q; t, s)$  satisfies the implicit equation

$$-\frac{\partial}{\partial q} h(q; t, s) = \int_{s_1}^{t_1} h(q; t, (q, \sigma)) h(q; (q, \sigma), s) d\sigma. \quad (2.12)$$

*Proof:* The proof is analogous to that of Lemma 2.1. We give here a brief outline of the proof. Define

$$F_{q, s_1}: L_2^2(\mathcal{D}) \rightarrow L_2^2(\mathcal{D})$$

$$\text{as } F_{q, s_1} f(t, s) = \int_{s_1}^{t_1} \int_0^q l(t, z) f(z, s) dz.$$

Then

$$l \notin \sigma(F_{q, s_1}) \quad \text{and} \quad (I - F_{q, s_1})h(q; t, s) = -l(t, s)$$

has a unique solution. Continuity of  $h(q, t, s)$  w.r.t.  $t$  follows immediately from the continuity of  $l(t, s)$  w.r.t.  $t$ . Furthermore,

$$\begin{aligned} (I - F_{q, s_1})(h(q; t, (s_1 + \Delta_1, s_2 + \Delta_2)) - h(q; t, (s_1, s_2))) \\ = -(l(t, (s_1 + \Delta_1, s_2 + \Delta_2)) - l(t, (s_1, s_2))) \\ - \int_{s_1}^{s_1 + \Delta_1} \int_0^q l(t, z) h(q; z, (s_1 + \Delta_1, s_2 + \Delta_2)) dz. \end{aligned}$$

Hence,  $h(q; t, s)$  is continuous in  $s$ . Finally,

$$\begin{aligned} (I - F_{q, s_1})(h(q + \Delta; t, s)) - h(q; t, s) \\ = \int_{s_1}^{t_1} \int_q^{q + \Delta} l(t, z) h(q; z, s) dz \end{aligned}$$

implies continuity and differentiability in  $q$ . From (2.11) we derive

$$\begin{aligned} \frac{\partial}{\partial q} h(q; t, s) &= \int_{s_1}^{t_1} l(t, (\sigma, q)) h(q; (\sigma, q), s) d\sigma \\ &\quad + \int_{s_1}^{t_1} \int_0^q l(t, z) \frac{\partial}{\partial q} h(q; z, s) dz, \quad (2.13) \end{aligned}$$

and

$$\begin{aligned} & \int_{s_1}^{t_1} l(t, (\sigma, q)) h(q; (\sigma, q), s) d\sigma \\ &= - \int_{s_1}^{t_1} h(q; t, (\sigma, q)) h(q; (\sigma, q), s) d\sigma \\ &+ \int_{s_1}^{t_1} \int_0^q l(t, z) \left( \int_{s_1}^{z_1} h(q; z, (\sigma, q)) \right. \\ &\quad \left. \cdot h(q; (\sigma, q), s) d\sigma \right) dz, \end{aligned} \quad (2.14)$$

which together imply (2.12).

*Proof of Theorem 2.1:* Let

$$p(t, s) = \begin{cases} h(t_2; t, s), & \text{if } s_2 \leq t_2 \\ h(s_2; t, s), & \text{if } s_2 > t_2 \end{cases}, \quad s, t \in \mathcal{L}.$$

With  $P_1$  and  $P_2$  defined in (2.5) we derive

$$\begin{aligned} \log(\det(I + R)) &= -\text{tr}(P_1 + P_2 + P_1^* + P_2^*) \\ &= - \int_{\mathcal{L}} \text{tr} h(t_2, t, t) dt \\ &= \int_{\mathcal{L}} \text{tr} l(t, t) dt \\ &= \int_0^{t_1} \int_0^{T_2} k(t, s) f(s) ds, \end{aligned}$$

with

$$k(t, s) = \begin{cases} h(t_2; t, s) - \int_{s_1}^{t_1} \int_{t_2}^{T_2} h(z_2; t, z) h(z_2; z, s) dz, & \text{if } s_2 \leq t_2 \\ h(s_2; t, s) - \int_{s_1}^{t_1} \int_{s_2}^{T_2} h(z_2; t, z) h(z_2; z, s) dz, & \text{if } s_2 > t_2. \end{cases}$$

Using (2.12) we obtain

$$k(t, s) = \begin{cases} h(t_2; t, s) + \int_{t_2}^{T_2} \frac{\partial}{\partial \sigma} h(\sigma; t, s) d\sigma, & \text{if } s_2 \leq t_2 \\ h(s_2; t, s) + \int_{s_2}^{T_2} \frac{\partial}{\partial \sigma} h(\sigma; t, s) d\sigma, & \text{if } s_2 > t_2 \end{cases} = h(T_2; t, s).$$

Hence, (2.11) gives us

$$\begin{aligned} (I - L)(I - P_1 - P_2 + P_2 P_1) f(t) &= (I - L) \left( f(\cdot) - \int_0^{t_1} \int_0^{T_2} k(\cdot, s) f(s) ds \right) (t) \\ &= f(t) + \int_0^{t_1} \int_0^{T_2} \left( -1(t, s) - h(T_2, t, s) \right. \\ &\quad \left. + \int_{s_1}^{t_1} \int_0^{T_2} 1(t, z) h(T_2, z, s) dz \right) f(s) ds \\ &= f(t), \end{aligned}$$

and thus  $(I - L)^{-1} = (I - P_2)(I - P_1)$ . Furthermore,

$$(I - L_1^*)^{-1} = ((I - L_1)^{-1})^* = (I - P_1^*)(I - P_2^*).$$

This leads to the result.  $\square$

Using Theorem 2.1, we obtain the following expression for the determinant of the operator  $(I + R)$ .

*Theorem 2.2:* Suppose that  $P_1$  and  $P_2$ , appearing in Theorem 2.1 are such that  $P_1 + P_2 + P_1^* + P_2^*$  is trace class. Then

$$\begin{aligned} \log(\det(I + R)) &= -\text{tr}(P_1 + P_2 + P_1^* + P_2^*) \\ &= - \int_{\mathcal{L}} \text{tr} h(t_2, t, t) dt \\ &= \int_{\mathcal{L}} \text{tr} l(t, t) dt \end{aligned}$$

where  $h$  is solution of (2.11) and  $l$  is defined as in Lemma 2.2.

*Proof:* The proof depends on two lemmas stated next.  $\square$

*Lemma 2.4:* Let  $K: \mathcal{H} \rightarrow \mathcal{H}$  be defined by  $Kf(t) = \int_{\mathcal{L}} k(t, s) f(s) ds$  with  $k(t, s)$  continuous on  $\mathcal{L} \times \mathcal{L}$ . If  $K$  is nuclear (trace class), then  $\text{tr} K = \int_{\mathcal{L}} \text{tr} k(t, t) dt$ .

*Proof:* The proof is analogous to the proof in the 1-parameter case given in [5, p. 116].  $\square$

*Lemma 2.5:* Let  $K$  be defined by

$$Kf(t) = \int_0^{t_1} \int_0^{T_2} k(t, s) f(s) ds$$

with continuous kernel  $k(t, s)$ . Then  $K^n$ ,  $n > 1$  is nuclear and  $\text{tr} K^n = 0$   $n > 1$ .  $\square$

*Proof:* It is easily seen that

$$K^n f(t) = \int_0^{t_1} \int_0^{T_2} k_n(t, s) f(s) ds,$$

where

$$k_1(t, s) = k(t, s)$$

and

$$k_n(t, s) = \int_{s_1}^{t_1} \int_0^{T_2} k(t, z) k_{n-1}(z, s) dz, \quad n > 1.$$

$K^n$ ,  $n > 1$  is a product of H-S operators and hence nuclear. Then by Lemma 2.3,

$$\text{tr} K^n = \int_{\mathcal{L}} \text{tr} k_n(t, t) dt = 0,$$

since

$$k_n(t, t) = \int_{t_1}^{t_1} \int_0^{T_2} k(t, z) k_{n-1}(z, t) dz = 0. \quad \square$$

*Proof of Theorem 2.2:* Define  $K = P_1 + P_2 - P_2 P_1$  and  $B = K$ ,  $A = -K^*(I - K^*)^{-1}$ . Then

$$\begin{aligned} (I - B)(I - A)^{-1} &= (I - P_2)(I - P_1)(I - P_1^*)(I - P_2^*) \\ &= (I + R). \end{aligned}$$

Further, let  $A(\xi)$  (resp.  $B(\xi)$ ) be the solution of  $I + \xi A(\xi) = (I - \xi A)^{-1}$  (resp.  $I + \xi B(\xi) = (I - \xi B)^{-1}$ ). Then

$$\begin{aligned} \text{tr}(A(\xi) - B(\xi)) &= \text{tr}((I - \xi B)(I - \xi A)^{-1})^{-1} (A - B)(I - \xi A)^{-2} \\ &= \frac{d}{d\xi} \log(\det((I - \xi B)(I - \xi A)^{-1})), \end{aligned}$$

and hence

$$\begin{aligned} \int_0^1 \text{tr}(A(\xi) - B(\xi)) d\xi &= \log(\det((I - B)(I - A)^{-1})) \\ &= \log(\det(I + R)). \end{aligned}$$

Furthermore,

$$\begin{aligned} \text{tr}(A(\xi) - B(\xi)) &= \text{tr}((I - \xi A)^{-1} - (I - \xi B)^{-1}) / \xi \\ &= \text{tr} \left( \sum_0^{\infty} (\xi A)^i - (\xi B)^i \right) / \xi \\ &= \sum_1^{\infty} (\text{tr} A^i - \text{tr} B^i) \xi^{i-1}. \end{aligned}$$

Here,

$$\begin{aligned} \text{tr} A^i &= \text{tr}(-K^*(I - K^*)^{-1})^i = -\text{tr} \left( \sum_1^{\infty} K^{*j} \right)^i \\ &= \begin{cases} -\text{tr} K^*, & \text{if } i = 1 \\ 0, & \text{if } i > 1 \end{cases} \end{aligned}$$

and

$$\text{tr} B^i = \text{tr} K^i = \begin{cases} \text{tr} K, & \text{if } i = 1 \\ 0, & \text{if } i > 1. \end{cases}$$

Hence

$$\begin{aligned} \int_0^1 \text{tr}(A(\xi) - B(\xi)) d\xi &= -\text{tr}(K + K^*) \\ &= -\text{tr}(P_1 + P_2 - P_2 P_1 + P_1^* + P_2^* - P_1^* P_2^*). \end{aligned}$$

Here,  $P_2 P_1 + P_1^* P_2^*$  is nuclear with kernel

$$\rho(t, s) = \begin{cases} \int_{s_1}^{t_1} \int_{t_2}^{T_2} h(z_2; t, z) h(z_2; z, s) dz, & s_1 \leq t_1 \quad s_2 \leq t_2 \\ \int_{s_1}^{t_1} \int_{s_2}^{T_2} h(z_2; t, z) h(z_2; z, s) dz, & s_1 \leq t_1 \quad s_2 > t_2 \\ \int_{t_1}^{s_1} \int_{s_2}^{T_2} h^T(z_2; z, t) h^T(z_2; s, z) dz, & s_1 > t_1 \quad s_2 \leq t_2 \\ \int_{t_1}^{s_1} \int_{t_2}^{T_2} h^T(z_2; z, t) h^T(z_2; s, z) dz, & s_1 > t_1 \quad s_2 > t_2. \end{cases}$$

Lemma 2.4 implies that  $\text{tr} P_2 P_1 + P_1^* P_2^* = \int_{\mathcal{D}} \text{tr} \rho(t, t) dt = 0$ . Therefore,

$$\begin{aligned} \text{tr}(K + K^*) &= \text{tr}(P_1 + P_2 + P_1^* + P_2^*) \\ &= \int_{\mathcal{D}} \text{tr} h(t_2; t, t) dt. \end{aligned}$$

Finally, using (2.11) we obtain

$$\begin{aligned} \log(\det(I + R)) &= \int_0^1 \text{tr}(A(\xi) - B(\xi)) d\xi \\ &= \int_{\mathcal{D}} \text{tr} l(t, t) dt. \quad \square \end{aligned}$$

### III. THE LIKELIHOOD FUNCTIONAL

In detection and identification problems, one is interested in obtaining an expression for the likelihood ratio of the observation process with respect to the noise process. The usual way to proceed is to integrate the observation equation and work with the 2-parameter Brownian motion model. Since we are working with Gaussian models, one can as well directly use the observation equation

$$y_t = x_t + n_t, \quad (3.1)$$

with  $\{n_t\}$  standard Gaussian 2-parameter white noise. The only problem is that white noise does not induce a countably additive measure on any appropriate function space. The way out is to work with cylindrical measures on  $\mathcal{H} = L_2^0(\mathcal{D})$ . For details on the Radon-Nikodym derivative and the associated likelihood ratio, see Bagchi [1]. We mention only some facts of direct relevance next.

Let  $\mu_G$  be the canonical Gaussian measure on cylinder sets of  $\mathcal{H}$  defined by means of the characteristic function  $C(h) = \exp(-1/2[h, h])$ ,  $h \in \mathcal{H}$ , where  $[\cdot, \cdot]$  denotes the inner product in  $\mathcal{H}$ . This measure cannot be extended to a countably additive measure on the Borel sets of  $\mathcal{H}$ . By a finitely additive white noise on  $(\mathcal{H}, \mu_G)$  we mean the process with sample paths  $\eta$  in  $\mathcal{H}$ . Let  $f: \mathcal{H} \rightarrow \mathbb{R}$  be any Borel measurable function. Inverse images of Borel sets in  $\mathbb{R}$  are not necessarily cylinder sets in  $\mathcal{H}$ , so that  $f$  needs not induce a measure on Borel sets of  $\mathbb{R}$ . Let  $\{P_k\}$  be a sequence of finite dimensional projections on  $\mathcal{H}$  such that  $P_k \rightarrow I$  strongly. Then for each  $k$ ,  $f(P_k \eta)$  is a random variable on  $\mathbb{R}$ . If  $\{f(P_k(\eta))\}$  is Cauchy in probability, then  $C(h) = \lim C_k(h)$ ,  $h \in \mathbb{R}$ , where  $C_k(h) = E \exp i[f(P_k \eta), h]$  defines a countably additive probability measure on  $\mathbb{R}$ . If  $C(\cdot)$  is independent of the choice of sequence  $\{P_k\}$ , then we say that  $f$  is a physical random variable (p.r.v.). If  $f$  is a p.r.v., then it induces a countably additive probability measure on  $\mathbb{R}$ .

A cylindrical measure  $\mu$  on  $\mathcal{H}$  is absolutely continuous w.r.t.  $\mu_G$  if, for any  $\epsilon > 0$ , there exists a  $\delta > 0$  such that for any cylinder set  $A$  in  $\mathcal{H}$ ,  $\mu_G(A) < \delta$  implies  $\mu(A) < \epsilon$ . Suppose  $f: \mathcal{H} \rightarrow \mathbb{R}$  is a p.r.v. such that for any sequence of projections  $P_k \rightarrow I$  strongly,

$$\mu(A) = \lim \int_A f(P_k h) d\mu_G(h)$$

uniformly. We then call  $f$  the Radon-Nikodym derivative of  $\mu$  w.r.t.  $\mu_G$  and we denote  $f = d\mu/d\mu_G$ .

In the 1-parameter Gaussian case the likelihood ratio for the integrated Brownian motion model is given by Shepp [7]. The white noise version of this formula is given by

$$\begin{aligned} LF(Y + EY) &= d\mu_Y / d\mu_G(Y + EY) \\ &= d^{-1/2} \exp([\![HY, Y]\!]/2 + [Y, EY] \\ &\quad + [EY, EY]/2), \quad (3.2) \end{aligned}$$

with  $\mu_Y$  the density induced by  $Y$ . Moreover,  $EY$  denotes the expectation of  $Y$  w.r.t.  $\mu_G$ ,  $H = R(I + R)^{-1}$  and  $d$  the Carleman-Fredholm determinant of  $(I + R)$ .

Here we show that the same formula holds for 2-parameter random fields. Further, using Theorem 2.2, we can express the determinant term appearing in formula (3.2), in terms of the system parameters. The work in an extension of the result of Bagchi and Westdijk [3] for 1-parameter boundary value processes. For similar formula in the case of elliptic stochastic system models, see Bagchi and Aihara [2].

Consider the 2-parameter random field as introduced in Section II, with covariance operator  $R$  having continuous kernel  $r(t, s)$ . Let  $\{\phi_i\}$  the complete set of normalized eigenfunctions of  $R$ , that is  $R\phi_i = \lambda_i\phi_i$ ,  $(\phi_i, \phi_j) = \delta_{ij}$ . Note that such a set exists since  $R$  is self adjoint. Define

$$\begin{aligned} Y_i &= [\phi_i, Y] = \int_{\mathcal{J}} \phi_i^T(t) Y_t dt; \\ EY_i &= [\phi_i, EY] = \int_{\mathcal{J}} \phi_i^T(t) EY_t dt; \\ d &= \prod_1^\infty (1 + \lambda_i) = \det(I + R); \\ LF_i(Y) &= (1 + \lambda_i)^{-1/2} \\ &\quad \cdot \exp\left(-\frac{(Y_i - EY_i)^2}{2(1 + \lambda_i)}\right) \exp(Y_i^2/2); \\ LF(Y) &= \prod_1^\infty LF_i(Y). \end{aligned}$$

Then the following hold true.

*Lemma 3.1:*

- I:  $\{\gamma_i = \lambda_i/(1 + \lambda_i)\}$  is the set of eigenvalues of  $H = R(I + R)^{-1}$ ,  $\gamma_i$  with corresponding eigenfunction  $\phi_i$ .
- II:  $r_Y(t, s) = \sum_1^\infty (1 + \lambda_i) \phi_i(t) \phi_i^T(s)$  is the kernel of  $R_Y = (I + R)$ .
- III:  $h(t, s) = \sum_1^\infty \gamma_i \phi_i(t) \phi_i^T(s)$  is the kernel of  $H$ .
- IV:  $E_G(Y_i - EY_i)(Y_j - EY_j)^T = (1 + \lambda_i) \delta_{ij}$ .
- IV:  $E_G(\sum_1^\infty (Y_i - EY_i) \phi_i(t) - (Y_t - EY_t)) (\sum_1^\infty (Y_j - EY_j) \phi_j(s) - (Y_s - EY_s))^T = 0$ .
- VI:  $Y_t - EY_t = \sum_1^\infty (Y_i - EY_i) \phi_i(t)$ , almost surely  $\mu_G$ .

*Proof:* I is obvious. To prove II, note that  $\{\phi_i\}$  is complete and, for any  $j$ ,

$$\begin{aligned} &\int_{\mathcal{J}} \left( \sum_1^\infty (1 + \lambda_i) \phi_i(t) \phi_i^T(s) \right) \phi_j(s) ds \\ &= \sum_1^\infty (1 + \lambda_i) \phi_i(t) \left( \int_{\mathcal{J}} \phi_i^T(s) \phi_j(s) ds \right) \\ &= \sum_1^\infty (1 + \lambda_i) \phi_i(t) [\phi_i, \phi_j] \\ &= (1 + \lambda_j) \phi_j(t) = R_Y \phi_j(t). \end{aligned}$$

III follows in the same way as II. IV can be established as follows.

$$\begin{aligned} &E_G(Y_i - EY_i)(Y_j - EY_j)^T \\ &= E_G \left( \int_{\mathcal{J}} \phi_i^T(t) (Y_t - EY_t) dt \right) \left( \int_{\mathcal{J}} \phi_j^T(s) (Y_s - EY_s) ds \right)^T \\ &= \int_{\mathcal{J}} \int_{\mathcal{J}} \phi_i^T(t) E_G(Y_t - EY_t)(Y_s - EY_s)^T \phi_j(s) ds dt \\ &= \int_{\mathcal{J}} \phi_i^T(t) \left( \int_{\mathcal{J}} r_Y(t, s) \phi_j(s) ds \right) dt \\ &= [\phi_i, R_Y \phi_j] = (1 + \lambda_j) [\phi_i, \phi_j] = (1 + \lambda_j) \delta_{ij}. \end{aligned}$$

Moreover,

$$\begin{aligned} &E_G \left( \sum_1^\infty (Y_i - EY_i) \phi_i(t) - (Y_t - EY_t) \right) \\ &\quad \cdot \left( \sum_1^\infty (Y_j - EY_j) \phi_j(s) - (Y_s - EY_s) \right)^T \\ &= E_G \left( \sum_1^\infty (Y_i - EY_i) \phi_i(t) \right) \left( \sum_1^\infty (Y_j - EY_j) \phi_j(s) \right)^T \\ &\quad + E_G((Y_t - EY_t)(Y_s - EY_s)^T) \\ &\quad - E_G \left( \sum_1^\infty (Y_i - EY_i) \phi_i(t) \right) (Y_s - EY_s)^T \\ &\quad - E_G(Y_t - EY_t) \left( \sum_1^\infty (Y_j - EY_j) \phi_j(s) \right)^T, \end{aligned}$$

where

$$\begin{aligned} E_G \left( \sum_1^\infty (Y_i - EY_i) \phi_i(t) \right) \left( \sum_1^\infty (Y_j - EY_j) \phi_j(s) \right)^T \\ = \sum_1^\infty (1 + \lambda_i) \phi_i(t) \phi_i^T(s) = r_Y(t, s), \end{aligned}$$

$$\begin{aligned} &E_G \left( \sum_1^\infty (Y_i - EY_i) \phi_i(t) \right) (Y_s - EY_s)^T \\ &= E_G \sum_1^\infty \left( \int_{\mathcal{J}} \phi_i(t) \phi_i^T(z) (Y_z - EY_z) dz \right) (Y_s - EY_s)^T \\ &= \sum_1^\infty \phi_i(t) \left( \int_{\mathcal{J}} r_Y(s, z) \phi_i(z) dz \right)^T \\ &= \sum_1^\infty (1 + \lambda_i) \phi_i(t) \phi_i^T(s) = r_Y(t, s), \\ &E_G((Y_t - EY_t)(Y_s - EY_s)^T) = r_Y(t, s). \end{aligned}$$

This proves V. VI follows immediately from V.  $\square$

*Theorem 3.1:*

- I:  $LF(Y) = \partial \mu_Y / \partial \mu_G(Y)$ .
- II:  $LF(Y + EY) = d^{-1/2} \exp([HY, Y]/2 + [Y, EY] + [EY, EY]/2)$ .

*Proof:*

*Part I)* First note that  $Y_i, Y_j$  are independent w.r.t.  $\mu_Y$  if  $i \neq j$ .  $LF_i$  and  $LF_j$  are independent, since  $Y_i$  and  $Y_j$  are. Let  $A = \{Y: Y_i < a_i, i = 1 \cdots N\}$  be a cylinder set in  $\mathcal{Y}$ . Then

$$\begin{aligned} &\int_A LF(Y) d\mu_G(Y) \\ &= \prod_1^N \int_{-\infty}^{a_i} (1 + \lambda_i)^{-1/2} \\ &\quad \cdot \exp\left(-\frac{(y_i - EY_i)^2}{2(1 + \lambda_i)}\right) \exp(y_i^2/2) \\ &\quad \cdot (2\pi)^{-1/2} \exp(-y_i^2/2) dy_i \\ &= \prod_1^N \Phi((a_i - EY_i)/(1 + \lambda_i)^{1/2}) = \mu_Y(A), \end{aligned}$$

where  $\Phi(\cdot)$  denotes the standard Gaussian distribution on  $\mathbb{R}$ . Hence  $\mu_Y(A) = \int_A LF(Y) d\mu_G(Y)$  for the given cylinder set  $A$ , and hence for all  $A$  in the algebra of cylinder sets in  $\mathcal{H}$ . Furthermore,  $LF(Y)$  is easily shown to be a physical random variable. This establishes Part I).

Part II)

$$\begin{aligned} LF(Y + EY) &= \prod_1^\infty LF_i(Y + EY) \\ &= \prod_1^\infty (1 + \lambda_i)^{-1/2} \\ &\quad \cdot \exp(-Y_i^2/2(1 + \lambda_i)) \exp((Y_i + EY_i)^2/2) \\ &= d^{-1/2} \exp \sum_1^\infty (\gamma_i Y_i^2/2 + Y_i EY_i + EY_i^2/2). \end{aligned}$$

On the other hand,

$$\begin{aligned} \sum_1^\infty \gamma_i Y_i^2 &= \sum_1^\infty \gamma_i \left( \int_{\mathcal{J}} Y_s^T \phi_i(s) ds \right) \\ &\quad \cdot \left( \int_{\mathcal{J}} \phi_i^T(t) Y_t dt \right) = [HY, Y]. \\ \sum_1^\infty Y_i EY_i &= \sum_1^\infty \left( \int_{\mathcal{J}} Y_s^T \phi_i(s) ds \right) \\ &\quad \cdot \left( \int_{\mathcal{J}} \phi_i^T(t) EY_t dt \right) = [Y, EY]. \\ \sum_1^\infty EY_i^2 &= \sum_1^\infty \left( \int_{\mathcal{J}} EY_s^T \phi_i(s) ds \right) \\ &\quad \cdot \left( \int_{\mathcal{J}} \phi_i^T(t) EY_t dt \right) = [EY, EY]. \end{aligned}$$

Hence Part II follows.  $\square$

Theorem 3.2:

$$\begin{aligned} LF(Y + EY) &= \exp \left( \int_{\mathcal{J}} \text{tr} l(t, t) dt / 2 \right. \\ &\quad + [LY, Y] - [LY, LY]/2 + [Y, EY] \\ &\quad \left. + [EY, EY]/2 \right). \end{aligned} \quad (3.2)$$

Proof:

$$\begin{aligned} LF(Y + EY) &= d^{-1/2} \exp([HY, Y]/2 \\ &\quad + [Y, EY] + [EY, EY]/2) \\ &= \exp(-\log(\det(I + R))/2 \\ &\quad + [HY, Y]/2 + [Y, EY] + [EY, EY]/2). \end{aligned}$$

Note that

$$\begin{aligned} H &= R(I + R)^{-1} = I - (I + R)^{-1} = I - (I - L^*)(I - L) \\ &= L + L^* - L^*L, \end{aligned}$$

so

$$\begin{aligned} [HY, Y] &= [(L + L^* - L^*L)Y, Y] \\ &= 2[LY, Y] - [LY, LY]. \end{aligned}$$

With Theorem 2.2 the result follows.  $\square$

#### IV. CONCLUSION

In this paper we derived a factorization of the covariance operator  $(I + R)$  for the observation process of a 2-parameter random field. This result can be applied to express the determinant term appearing in Shepp's expression for the likelihood ratio in terms of the system parameters. This means that in practice, one of the problems in computing the likelihood ratio for random fields is solved. Extensions to the multiparameter case are straightforward.

The expression of the determinant of  $(I + R)$  in terms of the system parameters may also be used to reexpress the Wong-Zakai correction term. The application of our results to specific 2-parameter system models is being investigated at present.

#### REFERENCES

- [1] A. Bagchi, "Cylindrical measures in Hilbert space and likelihood ratio for 2-parameter signals in additive white noise," *Contr. Theory Adv. Technol.*, vol. 1, pp. 139-153, 1985.
- [2] A. Bagchi and S. I. Aihara, "Smoothing and identification for random fields," in *Advances in Communication and Control Systems*, N. DeClaric, Ed. New York: Optimization Software Inc., 1988.
- [3] A. Bagchi and H. Westdijk, "Smoothing and likelihood ratio for Gaussian boundary value processes," *IEEE trans. Automat. Contr.*, vol. AC-34, pp. 954-962, 1989.
- [4] A. V. Balakrishnan, *Stochastic Differential Systems I, Lecture Notes in Economics and Mathematical Sciences*, vol. 84, M. Beckmann et al., Eds. Berlin: Springer Verlag, 1973.
- [5] A. V. Balakrishnan, *Applied Functional Analysis*. Berlin: Springer Verlag, 1976.
- [6] C. Bromley and G. Kallianpur, "Gaussian random fields," *Appl. Math. Optimization*, vol. 6, pp. 361-376, 1980.
- [7] L. A. Shepp, "Radon-Nikodym derivatives of Gaussian measures," *Ann. Math. Statist.*, vol. 37, 321-354, 1966.
- [8] E. Wong and M. Zakai, "Likelihood ratios and transformations of probability associated with two parameter Wiener processes," *Z. Wahrscheinlichkeitstheorie*, vol. 40, pp. 283-308, 1977.