AUGUST 1988 LIDS-P-1803

MAXIMUM A-POSTERIORI ESTIMATION OF RANDOM FIELDS - ELLIPTIC GAUSSIAN FIELDS OBSERVED VIA A NOISY CHANNEL

by

Amir Dembo[1]
Information Systems Laboratory
Stanford University
Stanford, CA 94305

Ofer Zeitouni[2]
Laboratory for Information and Decision Systems
Massachusetts Institute of Technology
Cambridge, MA 02139

ABSTRACT

An extension of the "prior density for path" (Onsager-Machlup functional) is defined and shown to exist for Gaussian fields generated by solutions of elliptic PDE's driven by white noise. This functional is then used to define and solve the MAP estimation of such fields observed via nonlinear noisy sensors. Existence results and a representation of the estimator are derived for this model.

AMS (1980) Subject Classification: Primary 60H15, 60G60 Secondary 93E10.

KEYWORDS: Onsager-Machlup, Stochastic PDE, Random Fields, MAP Estimation.

---

[1]This work was partially supported by the U.S. Army Research Office, under contract DAAL03-86-K-0045, NASA Headquarters grant NAGW-419-S1.3, and the Weizman post doctoral fellowship.

[2] This work was partially supported by Weizmann post doctoral fellowship, Air Force Office of Scientific Research grant AFOSR-85-0227B, U.S. Army Research Office contract DAAL03-86-K-0171 and National Science Foundation grant MIP86-19169. Parts of this research were carried out while the second author visited the ISL, Stanford University.

1. INTRODUCTION

In this paper, we consider the Maximum a-Posteriori estimation of random fields observed via nonlinear sensors in multidimensional white noise. Our goal is to extend the one dimensional results obtained for diffusion processes in [11]. Due to singularities which appear in the case of non-Gaussian random fields, we defer the treatment of those to a companion paper ([2]).

The basic model we will deal with is that of random fields generated by the solution of noise driven elliptic PDE's: loosely speaking, let D be a nice bounded domain in R^d and let P be a strongly elliptic operator of order 2k, with P_∂ an associated boundary operator, the field model we consider is

$$\begin{cases} Pu(x) = n(x) & x \in D \\ P_\partial u(x) = 0 & x \in \partial D \end{cases} \tag{1.1}$$

where n is white noise. For simplicity, we will concentrate on the Dirichlet problem.

The observation model will be that of white noise corrupted nonlinear observations, i.e.

$$\dot{y} = h(u) + \tilde{n} \tag{1.2}$$

(an exact definition of the model involved is given in section 2 below).

A typical example of the model is the following problem which we consider as our prototype example. It motivates our study since it seems suitable for image analysis applications.

$$\text{(P)} \quad \Delta u - \alpha^2 u = n \qquad x \in D = [0,1]^2$$

$$u\big|_{\partial D} = 0$$

$$y(x_1,x_2) = \int_0^{x_1}\int_0^{x_2} u(\theta_1, \theta_2)d\theta_1 d\theta_2 + \tilde{w}(x_1,x_2)$$

where $\tilde{w}(x_1,x_2)$ is a Brownian sheet on D independent of n. Note that in (P), k=1 and d=2.

Following the approach of [11], we define the "posterior probability" of the process u given the observation σ-field $\sigma\{y\}$ as

$$J_y(\phi) = \lim_{\varepsilon\to 0} \frac{P(\|u-\phi\| < \varepsilon | \sigma\{y\})}{P(\|u\| < \varepsilon)} \tag{1.3}$$

where $\| \ \|$ denotes the sup norm. We refer to $J_y(\phi)$ as the "conditional Onsager-Machlup" functional. We will show in Section 3 that for suitable ϕ, $J(\phi)$ is well defined, at least for almost all y in the support of the measure generated by (1.2). Note however that unlike in [11], the normalizing constant $P(\|u\| < \varepsilon)$ depends on the model (1.1) used, although not on ϕ. Actually, in the problem (P) described above, had we tried to define $J_y(\phi)$ using for example $P(\|\Delta^{-1}n\| < \varepsilon)$ as the normalizing constant, the limit in (1.3) would have in general failed to exist, even for the case $h(\cdot) \equiv 0$. This is the main new difficulty in the d>1 case, and this fact forces us to treat the non-Gaussian case separately, for in that case no ϕ-independent normalizing constant can be found.

Once (1.3) is well defined, an obvious candidate for a "sample" MAP estimator is

$$\hat{\phi} = \operatorname{argmax} J_y(\phi) \tag{1.4}$$

From this point on, the treatment is similar to the 1-D case: the existence of estimators is proved and a representation result derived. In the case of linear $h(\cdot)$, a convexity argument yields also the uniqueness of the estimator. When specified to our prototype problem (P), the results read:

$$\begin{cases} \Delta^2\hat{\phi} = -h'(\hat{\phi})h(\hat{\phi}) + h'(\hat{\phi})\,\dot{y} \text{ in } D \\ \hat{\phi} = 0 \text{ on } \partial D \\ \dfrac{\partial\hat{\phi}}{\partial v} = 0 \text{ on } \partial D \end{cases} \tag{1.5}$$

and the term $h'(\hat{\phi})\dot{y}$ is to be understood in the Ogawa sense, c.f. definition 3.1 below.

Finally, we remark that we have not tried to treat the most general case possible: thus we consider only bounded domains, we do not consider pseudo-differential operators (which for d such that $\frac{d+1}{4}$ is not an integer are natural candidates, since $\Delta^{(d+1)/4}$ creates in $\mathbf{R}^d$ the Levy motion) etc. The results concerning existence of the limit $J(\phi)$ and of solutions $\hat{\phi}$ do however seem to carry over.

The organization of the paper is as follows: in section 2 below we define rigorously our model as well as the notations used. In section 3 we define the conditional Onsager-Machlup functional and prove it's existence. Finally, in section 4 the estimation problem is finally attacked: existence and representation results are derived.

2. Model Definition

Let D be a closed bounded domain in $\mathbf{R}^d$, with a smooth boundary ∂D. Let P be a strongly elliptic differential operator of order 2k with smooth coefficients, and let P_∂ be the boundary operator (of order k-1). We denote by B the Dirichlet form associated with (P,P_∂) (c.f. [1]).

In the sequel, $W^{m,2}(D)$ will denote the usual Sobolev space of order (m,2) based on D. The Sobolev norm in $W^{m,2}(D)$ is denoted by $\| \ \|_{m,2}$ and $\| \ \|_2$ denotes the usual $L^2(D)$ norm. $W_0^{m,2}(D)$ will denote the closure of $C_0^\infty(D)$ w.r.t. the norm $\| \ \|_{m,2}$. $W_0^{-m,2}(D)$ is the space of distributions which is the dual of $W_0^{m,2}(D)$. We denote by n the random distribution valued white noise in D, i.e. the random distribution n such that for each smooth $\phi \in C_0^\infty(D)$, $n(\phi)$ is a Normal random variable of mean zero and variance $\|\phi\|_2^2$. Note that n is $W^{-m,2}(D)$ valued for $m>d/2$ [9] and that, for $\phi \in C_0^\infty(D)$ and any basis e_i of $L^2(D)$,

$$n(\phi) = \sum_{i=1}^{\infty} a_i(\phi,e_i) \qquad \text{(in q.m.)} \tag{2.1}$$

where the a_i in the R.H.S. in (2.1) are i.i.d. N(0,1) random variables, with $a_i = n(e_i)$.

By a solution to the equation

$$\begin{cases} Pu = n & \text{in } D \\ P_\partial u = 0 & \text{on } \partial D \end{cases} \tag{2.2}$$

we mean a distribution valued random variable u such that u has a continuous version and $B[\phi,u] = n(\phi)$ for all $\phi \in C_0^\infty(D)$, where B is the associated Dirichlet form [9]. Note that u is not a classical solution and moreover even the boundary conditions need not be classically satisfied. Only if $u \in W_0^{k-1,2}(D)$ will one have a classical generalized Dirichlet problem in the sense of [1, ch.8].

We can show the following basic theorem:

Theorem 2.1: Assume that $2k > \frac{d}{2}$. Further assume that $|B[\phi,\phi]| \geq C\|\phi\|^2_{k,2}$. Then

(2.2) has a unique $L^2(D)$ solution which is Holder continuous with some exponent > 0. Moreover, any two such solutions are equal in the sense of distributions in $W_0^{-k,2}(D)$.

Proof. The theorem follows by an easy application of the machinery developed in [1], [9]. We therefore give below only it's sketch.

Uniqueness. Let $\phi \in W_0^{k,2}$ and let $P^*\psi = \phi$, $P^*_\partial\ \psi = 0$ on ∂D. By the usual theory PDE's, $\psi \in W_0^{2k,2}(D)$. Let U_1, U_2 be solutions of (2.2). Then

$$0 = n(\psi) - n(\psi) = U_1(P^*\psi) - U_2(P^*\psi) = U_1(\phi) - U_2(\phi)$$

Since ϕ is arbitrary in $W_0^{k,2}(D)$, one deduces that $U_1 = U_2$ in $W_0^{-k,2}(D)$.

Existence. We show the existence only for the case of formally self adjoint operators, the general case requiring a different but similar construction. We first quote the following lemma, which is a combination of theorems (16.5) and (15.1) in [1].

Lemma 2.1. Let $\{e_i^k, \lambda_i\}$ denote the generalized eigenfunctions and the eigenvectors associated with B, respectively, i.e.

$$(P - \lambda_i I)^k e_i^k = 0\,, \quad i \in J, \quad k = 1,2,\ldots,k(i) \tag{2.3}$$

Then

(a) J is countable and k(i) is finite for all i.

(b) Let $N(\lambda)$ be the number of eigenvalues (counting multiplicty), such that $\mathrm{Re}(\lambda_i) \leq \lambda$. Then

$$N(\lambda) = c\lambda^{d/2k} + 0(\lambda^{d/2k}) \tag{2.4}$$

(Note that (b) implies that when arranged by increasing size, $\lambda_i \sim i^{2k/d}$).

(c) e_i^k, $i \in Z^+$, $k=1,2,\ldots,k(i)$ span $L^2(D)$.

Note that the assumptions of the theorem also guarantee that $|\lambda_i| \geq c$. Let now $a_{i,k}$ be the i.i.d N(0,1) r.v. given by $a_{i,k} = n(e_i^k)$. Let A_i be the k(i) dimensional matrix.

$$A_i \triangleq \begin{pmatrix} \lambda_i & 1 & & 0 \\ & \cdot & \cdot & \\ & & \cdot & 1 \\ 0 & & & \lambda_i \end{pmatrix} ,$$

and $b_{i,k}$ the k-th element of the vector

$$(A_i)^{-1} \begin{pmatrix} a_{i,1} \\ \vdots \\ a_{i,k(i)} \end{pmatrix} .$$

We claim that

$$u = \sum_{i \in Z^+} \sum_{k=1}^{k(i)} b_{i,k} e_i^k$$

Note that $E(\|u\|_2^2) = \sum_{i,k} E(b_{i,k}^2) \leq C \sum_i (\frac{1}{i^{4k/d}}) < \infty$, and therefore $u \in L^2(D)$. It is easy to check that u does indeed satisfy (2.2), and the existence of a continuous version follows from Kolmogorov's criteria. We ommit the details.

We conclude this section by defining the observation model: let $h(\cdot):\mathbf{R} \to \mathbf{R}$ be a C^{2k+1} function with all derivatives up to order 2k+1 bounded. Let $\tilde{n}$ be a d-dimensional white noise, independent of n. Note that

$$\tilde{w}(x_1,\ldots,x_d) = \int_0^{x_1} \cdots \int_0^{x_d} \tilde{n}(\theta_1,\ldots,\theta_d) d\theta_1 \ldots d\theta_d \tag{2.5}$$

is a Brownian sheet. We define the observation model y(x), $x \in D$ as

$$y(\underline{x}) = \int_0^{x_1} \cdots \int_0^{x_d} h(u(\theta_1,\ldots,\theta_d))d\underline{\theta} + \tilde{w}(\underline{x}) \tag{2.6}$$

Note that as in [10], the measures generated by y(x) and by w(x), x∈D are equivalent with Radon-Nikodym derivative:

$$\frac{dP_y}{dP_{\tilde{w}}} = \exp\left(\int_D h(u(\underline{\theta}))d\tilde{w}(\underline{\theta}) - \frac{1}{2}\int_D h^2(u(\underline{\theta}))d\underline{\theta}\right) \tag{2.7}$$

where the first integral in (2.4) is a stochastic integral which is well defined since $u(\underline{x})$ is independent of $\sigma\{\tilde{w}(\underline{\theta}),\ \underline{\theta} \in D\}$. We refer below to (2.7) as the "likelihood ratio".

3. Prior Density - The Onsager-Machlup Functional For Gaussian Random Fields

We recall that in [3], [8], the following definition was introduced for the prior density of a diffusion process satisfying $dx_t = f(x_t)dt + dw_t$ with f a given C_b^2 function:

$$I(\phi) \triangleq \lim_{\varepsilon \to 0} \frac{P(\|x - \phi\| < \varepsilon)}{P(\|w\| < \varepsilon)} \tag{3.1}$$

where $\| \;\|$ denotes the sup norm (here, in [0,1], say). Note that the denominator is a "standard reference" and does not depend on f.

For $\phi \in C^{1+\alpha}$, $\alpha>0$, one can show ([13]) that in the 1-D case,

$$\log I(\phi) = -\frac{1}{2}\int_0^1 \left(\dot{\phi} - f(\phi)\right)^2 dt - \frac{1}{2}\int_0^1 f'(\phi)dt \tag{3.2}$$

For linear f, the second term in the RHS of (3.2) is a constant. It's origin can be traced to the fact that in the Radon-Nikodym derivatives between the two measures in the RHS of (3.1), a term of the form $\int f'(\phi)\, w dw = \int f'(\phi) \frac{d(w^2-t)}{2}$ appears; under the conditioning of w small, the dw^2 term drops and one recovers (3.2).

A natural approach to the multidimensional case would try to compute (3.1). However, here instead of $\int f'(\phi) w dw$, one would have a term of the form $\int f'(\phi)P^{-1}(n)n$. Proceeding formally, it is easy to check that this term is infinite, even under the conditioning of $\|w\|$ small. We postpone the treatment of the case of nonlinear drifts to [2]. For Gaussian fields u generated by P, however, one has:

Theorem 3.1:

Let $J^\varepsilon(\phi) \triangleq \frac{P(\|u-\phi\|<\varepsilon)}{P(\|u\|<\varepsilon)}$. Then

a) $J(\phi) \triangleq \lim_{\varepsilon \to 0} J^\varepsilon(\phi)$ exists for $\phi \in W^{4k,2}(D)$ which satisfy the boundary conditions

$P_\partial \phi = 0.$

b) $J(\phi) = \exp -\frac{1}{2}\int_D (P\phi)^2$

Proof. Let Λ denote the Radon-Nikodym derivative between the measure induced by u and that induced by u-ϕ. One then has [7]:

$$\Lambda = \exp\left(\int -(P\phi)n - \frac{1}{2}\int_D (P\phi)^2\right) \tag{3.3}$$

Note now that

$$J^{\varepsilon}(\phi) = E_0(\Lambda \,|\, \|u\| < \varepsilon)$$

where E_0 denotes expectations w.r.t. the measure defined by $\phi = 0$. Therefore, to prove the theorem we have to show that for $\phi \in w^{4k,2}(D)$,

$$E_0(\exp \int_D -(P\phi)n \,|\, \|u\|<\varepsilon) \overset{\varepsilon\to 0}{\longrightarrow} 1 \tag{3.4}$$

Note that since $P\phi \in W^{2k,2}(D)$ $L^2(D)$, and is deterministic (and in general, independent of σ(n)) the stochastic integral in (3.4) is well defined. Moreover,

$$\int_D P\phi\, n = \int_D P\phi P u$$

$$= \int_D P^*(P\phi)u \tag{3.5}$$

However, $\psi \overset{\Delta}{=} P^*P\phi \in L^2(D)$ is independent of u. One has therefore $|\int_D (P^*P\phi)u| \le k\|u\|\|\psi\|_2$ from which (3.4) follows.

As was the case in [11], theorem (3.1) will not be quite enough, for the optimal estimate will turn out not to be in $W^{4k,2}(D)$. In order to define the class of functions in which the estimator exists, we need to introduce some new machinery, which is reminiscent the "Ogawa integration" in 1-D (c.f. [4], [6]):

Definition 3.1. Let $\phi \in L^2(D)$ be random and generated by the white noise $\tilde{n}$. Let e_i be an arbitrary orthonormal base in $L^2(D)$ which satisfies the boundary conditions of (2.2). Recall that $\tilde{n}$ can be represented as $\tilde{n} = \sum_i a_i e_i$ in the sense of distributions in $W^{-d,2}(D)$ where a_i are i.i.d. N(0,1) r.v., and $a_i = \int_D e_i\tilde{n}$ is well defined since e_i is deterministic.

If the sum

$$I = \sum_{i=1}^{\infty} \left(\int_D \phi e_i \right) a_i \tag{3.6}$$

converges in $L^2(\Omega)$ and its value does not depend on the choice of the basis e_i, we say that ϕ is Ogawa integrable and denote its integral $I = \int_D \phi \circ n$.

It is easy to see that if $\phi \in W_0^{d,2}(D)$, ϕ is Ogawa integrable; actually, we can have more:

Lemma 3.1. Let $u \in L^2(D)$ be a deterministic bounded function. Let $K \in W_0^{2k,2}(D)$ be a random function generated by $\tilde{n}$, with $2k > d/2$. Then

$$\int_D (uK\circ\tilde{n}) = \int_D P^*(K)P^{-1}(u\,\tilde{n}) \tag{3.7}$$

where $P^{-1}(u\tilde{n})$ denotes the unique $L^2(D\times\Omega)$ solution to the S.P.D.E.

$$\begin{cases} Pv = u\,\tilde{n} \\ P_\partial v = 0 \end{cases} \tag{3.8}$$

which exists, and the equality (3.7) is in the sense that each side exists.

Proof. Assume first that a unique solution to (3.8) exists in $L^2(D\times\Omega)$. By our assumptions, the R.H.S. of (3.7) exists as a r.v. in $L^2(\Omega)$.

Let now e_i denote an $L^2(D)$ basis which satisfies the boundary conditions associated with P. One has that:

$$\sum_{i=1}^{M} \left(\int_D uKe_i \right) \left(\int_D e_i \circ \tilde{n} \right) = \sum_{i=1}^{M} \left(\int_D P^*(K)\, P^{-1}\,(ue_i) \right) \left(\int_D e_i \circ \tilde{n} \right)$$

$$= \int_D P^*(K) \sum_{i=1}^{M} \left[P^{-1}(ue_i) \int_D (e_i o \tilde{n}) \right] \tag{3.9}$$

By our assumption on the existence of a unique $L^2(D\text{x}\Omega)$ solution to (3.8), one checks easily that the sum in the R.H.S. of (3.9) converges in $L^2(D\text{x}\Omega)$ to such a solution, and (3.7) is proved.

It remains therefore to show that a unique $L^2(\Omega\text{x}D)$ solution to (3.8) exists. Note that if a solution which satisfies the boundary conditions exists then it is unique, since if v_1 and v_2 are two such solutions $P(v_1-v_2) = 0$ and $v_1-v_2 = 0$ by the classical PDE theory. We therefore proceed to construct a solution:

Let (λ_i, ψ_i) be the eigenvalues and the generalized eigenfunctions associated with P. For simplicity, we assume that the eigenvalues have all single geometric multiplicity - the general case follows easily since by [1] the eigenvalues have always finite multiplicity. Note that since u is bounded, $u\psi_k \in L^2(D)$ and one has

$$u\psi_k(x) = \sum_{\ell} b_k^{\ell}\, \psi_{\ell}(x)$$

where $b_k^{\ell} = \int_D u\psi_k\psi_{\ell}$. Let $a_k \stackrel{\Delta}{=} \int_D \psi_k o\tilde{n}$. Define

$$C_{\ell} \stackrel{\Delta}{=} \frac{1}{\lambda_{\ell}} \left(\sum_k b_k^{\ell}\, a_k \right)$$

Note that

$$E \sum_{\ell} C_{\ell}^2 = \sum_{\ell} \frac{1}{\lambda_{\ell}^2} \sum_k \left(b_k^{\ell} \right)^2$$

However, $b_k^{\ell} = \int_D u\psi_k\psi_{\ell}$ and therefore

$$\sum_k \left(b_k^{\ell} \right)^2 = \int_D u^2\, \psi_{\ell}^2(x) \le \|u\|^2 < c$$

Therefore, $E\sum_{\ell} C_{\ell}^2 < c\sum_{\ell} \frac{1}{\lambda_{\ell}^2} < \infty$ for $2k > d/2$. Define now

$$v = \sum_{\ell} C_{\ell}\psi_{\ell}(x)$$

Clearly, $v \in L^2(D)$. It is easy to check that v such defined is a solution to (3.8). The proof is completed.

We are ready now to state the extension of Theorem (3.1) which we will need in the sequel.

Theorem 3.2.

Let $\phi \in W^{2k,2}(D)$ be random and generated by some white noise $\tilde{n}$ independent of n. Assume that ϕ satisfies the boundary conditions associated with P. Let $\psi = P\phi \in L^2(D)$ and assume that

$$P^*\psi = A(\phi,\psi) + K(\phi)o\tilde{n} \qquad \text{a.s.} \qquad (3.10)$$

where $A(\phi,\psi) \in L^2(D)$ and $K(\cdot) \in C_b^{2k}(R \to R)$, where by a solution we mean that for any $g \in C^\infty(D)$, $\int_D \psi Pg = \int_D gK(\phi)o\tilde{n} + \int_D A(\phi,\psi)g$. Then the conclusion of theorem (3.1) still holds a.s. in $\tilde{n}$.

Proof. We have only to show that

$$E_0(\exp\int_D \psi\, on \,|\, \|u\| < \varepsilon,\tilde{n}) \overset{\varepsilon\to 0}{\to} 1 \qquad \text{a.s. } \tilde{n}$$

Note that by lemma (3.1) (with u=1, $K = (P^*)^{-1}A(\phi,\psi)$), the $A(\phi,\psi)$ term in (3.10) does not cause any difficulty. Therefore, we actually need to show that

$$\mu_\varepsilon = E_0(\exp\int_D \tilde{\psi}\, on \,|\, \|u\|<\varepsilon,\tilde{n}) \overset{\varepsilon\to 0}{\to} 1 \qquad \text{a.s. } \tilde{n}$$

where

$$P^*\tilde{\psi} = K(\phi)o\tilde{n}$$

Indeed, $\tilde{\psi} = \psi - (P^*)^{-1}A(\phi,\psi)$. Note that

$$\mu_\varepsilon \geq \exp E_0\left(\int_D \tilde{\psi}\, on \,\middle|\, \|u\| < \varepsilon, \tilde{n}\right) = 1 \tag{3.11}$$

due to the symmetry in n on the conditioning set $\|u\| < \varepsilon$. On the other hand, let e_i be the eigenfunction expansion of P, which spans $L^2(D)$, and assume for simplicity that all eigenvalues are simple (the general case does not pose any difficulty). Then

$$\int_D \tilde{\psi}\, on \triangleq \sum_{i=1}^{\infty} (\tilde{\psi}, e_i)(e_i, n) = \sum_{i=1}^{\infty} (P^*\tilde{\psi}, P^{-1}e_i)(P^*e_i, P^{-1}n)$$

$$= \sum_{i=1}^{\infty} (P^*\tilde{\psi}, e_i)(e_i, u) = \sum_{i=1}^{\infty} (P^*\tilde{\psi}, (e_i, u)e_i) = \int_D uK(\phi)o\tilde{n} \tag{3.12}$$

where we have used the fact that $e_i \in C_b^{\infty}(D)$ and u is bounded. Applying lemma 3.1, we have therefore that

$$\int_D \tilde{\psi}\, on = \int_D P^*(K(\phi))\; P^{-1}(u\tilde{n})$$

Therefore,

$$\left|\int_D \tilde{\psi}\, on\right| \leq C\, \|P^{-1}(u\tilde{n})\|_2 \tag{3.13}$$

Note that $\theta \triangleq P^{-1}(u\tilde{n})$ is conditioned on u a Gaussian process with $E(\|\theta\|_2^2 | u) < C\|u\|^2$ converging to zero as $\|u\| \to 0$ (see lemma 3.1). Therefore, for any c>0

$$
\begin{aligned}
&E_0(\mu_\varepsilon^c \mid \|u\| < \varepsilon) \\
&\quad \leq E_0(\exp c\int_D \tilde{\psi}\, on \mid \|u\| < \varepsilon) \\
&\quad \leq E_0(\exp c_1 \|\theta\|_2 \mid \|u\| < \varepsilon) \\
&\quad \leq E_0(E_0(\exp c_1 \|\theta\|_2 \mid u) \mid \|u\| < \varepsilon) \\
&\quad \leq E_0(\exp c_2 \|u\|^2 \mid \|u\| < \varepsilon) \\
&\quad \leq \exp c_2 \varepsilon^2 \underset{\varepsilon \to 0}{\longrightarrow} 1
\end{aligned} \tag{3.14}
$$

Combining (3.14) with (3.11) one obtains the theorem.

In the sequel, we say that $\phi \in \mathcal{L}$ if ϕ satisfies the conditions of theorem (3.2).

4. MAP Estimator

We start by finally defining our MAP estimator: Let u denote the solution of (2.2) and let y denote the observation defined by (2.6). Define

$$J_y^{\varepsilon}(\phi) \triangleq \frac{P(\|u-\phi\|<\varepsilon|\sigma\{y\})}{P(\|u\|<\varepsilon)}$$

We make the following definition:

Definition (4.1):

Let $J_y(\phi) = \lim_{\varepsilon \to 0} J_y^{\varepsilon}(\phi)$, and assume a (not necessarily unique) solution exists to the following stochastic optimization problem

$$\hat{\phi} \triangleq \arg\max_{\phi \in \mathcal{L}} J_y(\phi) \tag{4.1}$$

Then $\hat{\phi}$ is called the MAP estimator of u given $\sigma\{y\}$.

Remark. As we will see below, a version of $J_y(\phi)$ can be defined for all y in the support of the law of y defined by (2.6). Therefore, the optimization problem of definition (4.1) is well posed. It will also hold true that the optimization of the expression we have for $J_y(\phi)$ over a space larger than $\mathcal{L}$ (specifically, over $W_0^{2k}(D)$) still yields a solution in $\mathcal{L}$. It therefore justifies the fact that we look for a solution in $\mathcal{L}$ to start with.

In order to compute $J_y(\phi)$, we start with the following lemma, whose proof we defer to the appendix:

Lemma (4.1)

Let $\frac{dP_y}{dP_{\tilde{w}}}$ be as in (2.7). Then

$$E_{\tilde{w}}\left(\frac{dP_y}{dP_{\tilde{w}}}\Big| y, \|u-\phi\| < \varepsilon\right) \underset{a.s}{\overset{\varepsilon\to 0}{\longrightarrow}} \exp\left(\int_D h(\phi(\theta))dy(\theta) - \frac{1}{2}\int_D h^2(\phi(\theta))d\theta\right) \quad (4.2)$$

where the stochastic integral in the RHS of (4.2) is well defined for all $\phi \in W^{2k,2}(D)$ and $\tilde{w}$ in the support set of the measure defining $\tilde{y}$ by the pairing between $W^{-d/2-\delta,2}(D)$, $(\delta<2k-d/2)$, and $W^{2k,2}(D)$ (i.e., it is well defined even for ϕ stochastic which depends on y).

Combining now lemma (4.1) and theorem (3.2), one has the following easy corollary.

Corollary 4.1.

$\forall \phi \in \mathcal{L}$,

$$\hat{J}_y(\phi) \triangleq \log J_y(\phi) = \left[\int_D h(\phi(\theta))dy(\theta) - \frac{1}{2}\int_D h^2(\phi(\theta))d\theta - \frac{1}{2}\int_D (P\phi(\theta))^2 d\theta\right] \quad (4.3)$$

Having defined the cost functional $\hat{J}_y(\phi)$, we can turn to the existence issue. We claim:

Theorem 4.2.

Assume that the conditions of theorem 2.1 hold together with $h \in C^{2k+1}(R)$ and $h' \in C_0^{2k}(R)$.

Then a solution to the problem

$$\hat{\phi} = \arg \max_{\phi\in W^{2k,2}(D)} \hat{J}_y(\phi)$$

exists. Moreover, in the case h(x) is linear, this solution turns out to be uniqe.

Proof. The proof follows closely the lines of [12]. We proceed in two steps:

a) We show that for each $y \in C(D)$,

$$\lim_{\|\phi\|_{2\kappa,2}\to\infty} \hat{J}_y(\phi) = -\infty, \quad \hat{J}_y(0) \neq -\infty$$

b) We show that $\hat{J}_y(\phi)$ is lower semi continuous w.r.t. the weak topology in $W^{2k,2}(D)$.

Note that a) and b) imply the first part of the theorem, for by a) there exists for each $y \in C(D)$ a number $R(y)<\infty$ such that the supremum of $\hat{J}_y(\phi)$ is achieved inside a ball of radius R(y) in $W^{2k,2}(D)$. This ball being weakly compact, b) then implies that the supremum is actually achieved. Note that due to the strict convexity in ϕ of $\hat{J}_y(\phi)$ when $h(\cdot)$ is linear, the second part of the theorem follows once the first part is proved.

We turn now to the proof of a): Note first that $\hat{J}_y(0) = \int h(0)dy - \frac{1}{2}$ $h^2(0)\mathrm{Vol}(D) > -\infty$. Due to the ellipticity of P and our assumptions on P, there exists a $c_1 > 0$ such that

$$\int_D (P(\phi))^2 d\theta \geq (c_1\|\Delta^k\phi\|_2^2 - c_2\|\phi\|_2^2) \vee c_3\|\phi\|_2^2$$

On the other hand,

$$\left|\int_D h(\phi)dy\right| \leq \|y\|_{d/2+\delta,2} \|h(\phi)\|_{d/2+\delta,2} \leq c_4(y)(\|\phi\|_2 + \|\Delta^k\phi\|_2)$$

with $c_4(y) < \infty$, where $\delta < (2k-d/2)$. Let $\|\Delta^k\phi\|_2 = x$, $\|\phi\|_2 = z$. Note that $\|\phi\|_{2k,2} \leq C(x+z)$; one has then

$$\hat{J}_y(\phi) \leq -((c_1x^2 - c_2z^2)\vee c_3z^2) + c_4(x+z) \overset{x+z\to\infty}{\longrightarrow} -\infty.$$

and the proof of a) is completed.

We finally show b). Note first that since 2k>d/2, weak convergence in $W^{2k,2}(D)$ implies strong convergence in $L^2(D)$ and therefore the second integral in the R.H.S. of (4.3) is weakly continuous. Considering the first integral, note that it is defined by the pairing $<W^{-2k,2}, W^{2k,2}>$ and since $dy \in W^{-d/2-\delta/2}(D)$ $W^{-2k,2}(D)$, the weak continuity follows immediately. Finally, considering the third integral, let $\phi_n \overset{\omega}{\to} \phi$ in $W^{2k,2}(D)$, then clearly $B[\phi_n, e_i] \to B[\phi, e_i]$ where e_i is any member of the orthonormal basis associated with P and B in the associated Dirichlet form. But

$$\liminf_{n\to\infty} \|P\phi_n\|_2 = \liminf_{n\to\infty} \sum_{i=1}^{\infty} (P\phi_n, e_i)^2$$

$$= \liminf_{n\to\infty} \sum_{i=1}^{\infty} (B[\phi_n, e_i])^2$$

$$\geq \sum_{i=1}^{\infty} \liminf_{n\to\infty} (B[\phi_n, e_i])^2$$

$$= \sum_{i=1}^{\infty} (B[\phi, e_i])^2 = \|P\phi\|_2 \tag{4.3}$$

where we used Fatous' inequality. This completes the proof of the theorem.

We conclude this section by the following representation result for the estimator $\hat{\phi}$:

Theorem 4.3:

Let k be an integer, k>d/2. Any maximizer of (4.1) is a weak solution of

$$\begin{cases} P^*P\hat{\phi} = -h'(\hat{\phi})h(\hat{\phi}) + h'(\hat{\phi})o\dot{y} & \text{on } D \\ P_\partial\hat{\phi} = 0 & \text{on } \partial D \\ P_\partial^* P\hat{\phi} = 0 & \text{on } \partial D \end{cases} \tag{4.4}$$

where P_∂^* denotes the boundary operator defined by P* i.e., for $\phi_1, \phi_2 \in C^\infty(D)$

$$(P^*\phi_1, \phi_2) - (\phi_1, P\phi_2) = \sum_{j=1}^{2k-1} (\tilde{P}_\partial^* \phi_1)_j \frac{\partial^j \phi_2}{\partial \underline{n}^j}$$

with $\underline{n}$ being the exterior normal to D, $\tilde{P}_\partial^*$ above defined by Green's formula and

$(P_\partial^* \phi_1) = [(\tilde{P}_\partial^* \phi_1)_k, (\tilde{P}_\partial^* \phi_1)_{k+1}, \ldots, (\tilde{P}_\partial^* \phi_1)_{2k-1}]$.

The proof of theorem 4.3 follows by an easy application of the necessary conditions of the calculus of variations and is therefore ommitted. Note that by lemma 3.1 we can understand (4.4) as a pathwise equation defined for each $y \in C(D)$! Note also that by (4.4), $\hat{\phi} \in \mathcal{L}$.

Remark: Note that by [5], the Ogawa integral in (4.4) may be replaced by a Stratonovich integral, by using a Haar basis. This basis also yields an approximation to (4.4) by means of difference equations.

Appendix

Proof of lemma 4.1:

Clearly, all we have to show is that

$$E_{\tilde{w}}(\exp \int_D \big(h(\phi(\underline{\theta})) - h(u(\underline{\theta}))\big) \circ \dot{y}(\theta)) \big| \|\phi - u\| < \varepsilon, y) \overset{\varepsilon \to 0}{\to} 1 \tag{A.1}$$

Let $u^\delta(\theta) \triangleq u(\theta) * j^\delta(\theta)$, $\phi^\delta(\theta) \triangleq \phi(\theta) * j^\delta(\theta)$ where $j^\delta(\theta)$ is a $\delta(\varepsilon)$ molifier (say, a Poisson kernel). We will choose $\delta(\varepsilon) \underset{\varepsilon \to 0}{\to} 0$ below. Note that (A-1) will be proved if we can show that

$$\mu^1 \triangleq \exp(c \int_D (h(\phi) - h(\phi^\delta)) \circ \dot{y}) \overset{\delta(\varepsilon) \to 0}{\to} 1 \tag{A.2a}$$

$$\mu^2 \triangleq E_{\tilde{w}}(\exp(\hat{c} \int_D (h(\phi^\delta) - h(u^\delta)) \circ \dot{y}) \big| \|\phi - u\| < \varepsilon, y) \overset{\delta(\varepsilon) \to 0}{\to} 1, \ \forall \hat{c} \in \mathbf{R} \tag{A.2b}$$

$$\mu^3 \triangleq E_{\tilde{w}}(\exp(\hat{c} \int_D (h(u^\delta) - h(u)) \circ \dot{y}) \big| \|\phi - u\| < \varepsilon, y) \overset{\delta(\varepsilon) \to 0}{\to} 1, \ \forall \hat{c} \in \mathbf{R} \tag{A.2c}$$

Consider first (A-2a): Note that since $\phi \in W^{2k,2}(D)$, $\|h(\phi) - h(\phi^\delta)\|_{2k,2} \le c\delta \overset{\delta\to 0}{\longrightarrow} 0$ which implies (A-2a). Consider next (A-2b): Note that $\|h(\phi^\delta) - h(u^\delta)\|_{2k,2} \le c \ \|\phi^\delta - u^\delta\|_{2k,2} \le c \frac{\varepsilon}{\delta^{2k}}$. By choosing $\delta^{-2k} = o(\varepsilon^{-1})$, one obtains (A-2b).

We finally consider (A.2c): Since under $P_{\tilde{w}}$, $\dot{y}$ is white noise and u is independent of $\dot{y}$, one has:

$$\forall \tilde{c}, |\tilde{c}| \ge 1 \ E_{\tilde{w}}\left((\mu^3)^{\tilde{c}}\right) \le E_{\tilde{w}}\left[\exp c'\|h(u) - h(u^\delta)\|_2^2\right] \le \exp(c'\delta^2) \overset{\delta(\varepsilon)\to 0}{\longrightarrow} 1$$

which concludes the proof of the lemma.

REFERENCES

[1] S. Agmon, Lectures on elliptic boundary value problems, Van Nostrand, 1965.

[2] A. Dembo and O. Zeitouni, Maximum a-Posteriori estimation of random fields, part II: Non-Gaussian fields, submitted.

[3] N. Ikeda and S. Watanabe, Stochastic differential equations and diffusion processes, North Holland, 1981.

[4] D. Nualart and M. Zakai, Generalized stochastic integrals and the Malliavin calculus, Probability Theory and Related Fields, 73 (1986) pp. 255-280.

[5] D. Nualart and M. Zakai, On the relation between the Stratonovich and Ogawa integrals, Preprint 1988.

[6] S. Ogawa, The stochastic integral of noncausal type as an extension of the symmetric integrals, Japan J. Applied Math 2, (1984) pp. 229-240.

[7] R. Ramer, On nonlinear transformations of Gaussian measures, J. Funct. Analysis, 15(1974), pp. 166-187.

[8] R.L. Stratonovich, On the probability functional of diffusion processes, 1960. In: Selected Trans. in Math. Statistics and Prob. 10 (1971).

[9] J.B. Walsh, An introduction to stochastic partial differential equestions, Lecture Notes in Math. #1180, pp. 265-939, Springer 1986.

[10] E. Wong, A likelihood ratio formula for two dimensional random fields, IEEE Trans. Inf. Theory, 20(1979), pp. 418-422.

[11] O. Zeitouni and A. Dembo, A. Maximum A Posteriori estimator for trajectories of diffusion processes, Stochastics 20 (1987) pp. 221-246.

[12] O. Zeitouni and A. Dembo, An Existence theorem and some properties of maximum a posteriori estimators of trajectories of diffusions, Stochastics 23 (1988) pp. 197-218.

[13] O. Zeitouni, On the Onsager-Machlup functional around non C^2 curves, to appear, The annals of Probability.