OCTOBER 1988 LIDS-P-1827

MAXIMUM A-POSTERIORI ESTIMATION OF RANDOM FIELDS - PART II: NON-GAUSSIAN FIELDS[1]

by

Amir Dembo
Information Systems Laboratory
Stanford University
Stanford, CA 94305

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

ABSTRACT

The "prior density for path" (Onsager-Machlup functional) is defined for solutions of nonlinear elliptic type PDE's driven by white noise. Applying a general theorem of Ramer on equivalence of measures, we prove the existence of this functional. We then consider the Maximum A-Posteriori (MAP) estimation problem, where the solution of the elliptic PDE is observed via nonlinear noisy sensors. We prove the existence of the optimal estimator which is then represented by means of another elliptic stochastic PDE.

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

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

[1]This work we supported by the Air Force Office of Scientific Research grant AFOSR-85-0227B, the National Science Foundation under grant MIP86-19169, the U.S. Army Research Office under contracts DAAL03-86-K-0171 and DAAL03-86-K-0045, and NASA Headquarters, Center for Aeronautics and Space Information Sciences under grant NAGW-419-S1.3. Parts of this research were carried out while the second author visited the ISL, Stanford University.

1. INTRODUCTION

In this paper we extend our previous work [2] on Maximum a-Posteriori (MAP) estimation of random fields, beyond the Gaussian case.

For one dimensional diffusions, the change from Gaussian to non-Gaussian processes involved only adding a finite correction term in the "prior density for path" [4]. However, in the non-Gaussian multidimensional case, the technique which yielded the prior density in the Gaussian case fails to converge. To remedy this situation, we have to impose smoothness conditions on the solution of the PDE which are beyond the minimal ones required for the existence of continuous solutions, c.f. below. In what follows we use the same notations as in [2].

We will deal here with random fields generated by the solution of semi-linear elliptic PDE's, over (nice) bounded domains in $\mathbf{R}^d$, $d \geq 2$. Let P be a strongly elliptic linear operator of order 2k with smooth coefficients, P_∂ an associated linear boundary operator (of order k-1), and F a non-linear operator of order m such that $2k > \frac{d}{2} + m$. For example, $F(\phi) = f(\phi, D_x\phi, D_x^2\phi, \ldots, D_x^m\phi)$ for some smooth $f: R^{m+1} \to R$. The field model we consider is:

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

where n is white noise (a somewhat more general form of (F) is used in Section II). For simplicity, <u>we will concentrate on the Dirichlet problem</u>, and denote by B the Dirichlet form associated with (P, P_∂) (c.f. [1]). We use the observation model as in [2] (where we assumed $F \equiv 0$), i.e., that of white noise corrupted nonlinear observations:

$$y(\underline{x}) = \int_0^{x_1} \cdots \int_0^{x_d} h(u(\underline{\theta}))d\underline{\theta} + \tilde{w}(\underline{x}) \tag{1.2}$$

where $\tilde{w}(\underline{x})$ is a Brownian sheet independent of n, and $h(\cdot): \mathbf{R} \to \mathbf{R}$ is a C^{2k+1} function with all derivatives up to order 2k+1 bounded.

Our prototype example is:

$$
(P)\begin{cases}
\Delta^{1+\delta}u + F(u) = n, & x\in D = [0,1]x[0,1] \quad \mathbf{R}^2, \ \delta>0 \\
u\big|_{\partial D} = 0 & \\
y(x_1, x_2) = \int_0^{x_1}\int_0^{x_2} u(\theta_1, \theta_2)\, d\theta_1 d\theta_2 + \tilde{w}(x_1,x_2) &
\end{cases}
$$

here $F(\cdot)$ is a smooth nonlinear function.

Under certain smoothness conditions on P, F and D, equation (1.1) admits a unique solution in $W^{m,2}(D)$, (c.f. Theorem 2.1). For $\phi(x)$ smooth, with $P_\partial\phi(x) = 0$ on ∂D, we apply a general theorem of Ramer [7] to evaluate the Radon-Nykodim derivative between the measure P_1 defined by solutions of:

$$
(P_1)\begin{cases}
Pu(x) + P\phi(x) + F(u(x) + \phi(x)) = n(x) & x\in D \\
P_\partial u(x) = 0 & x\in \partial D \\
y(\underline{x}) = \int_{0\le\underline{\theta}\le\underline{x}} h(u(\underline{\theta}))d\underline{\theta} + \tilde{w}(\underline{x}) & x\in D
\end{cases} \tag{1.3}
$$

and the reference measure P_0 for which $\phi\equiv0$, $h\equiv0$ (consider theorem 2.2). Let this functional be denoted by $\Lambda_\phi(u(x))$, then a plausible definition of the "posterior probability" of the path ϕ given the observation σ-field $\sigma(y)$ is:

$$
J_y(\phi) = \lim_{\varepsilon\to0} \frac{P_F(\|u-\phi\|_{m,2}<\varepsilon|\sigma(y))}{P_F(\|u\|_{m,2}<\varepsilon|\sigma(y)} = \lim_{\varepsilon\to0} E_0(\Lambda_\phi(u(x))\big|\|u\|_{m,2} < \varepsilon,\sigma(y)) \tag{1.4}
$$

provided that $Jy(\phi)$ is well defined. Here $\|\cdot\|_{m,2}$ is the Sobolev norm of $W^{m,2}(D)$ and P_F denotes the probability measure generated by (1.1, 1.2). We note that in [2,9] we have used the $L^\infty(D)$ norm to define the ε-neighborhoods. We can use this norm here for m=0, but when m>0 we need some control on the derivatives of u as well (i.e., $\|u\|_{m,2} \le K(D)\varepsilon$). While for $F\equiv0$, $J_y(\phi)$ is well defined for almost all y in the support of (F) for an appropriate class of $\phi(x)$ whenever $2k>\frac{d}{2}$ (i.e., (1.1)

(1.2) make sense), for non-zero, non-linear $F(\cdot)$ this is not the case. While for $2k>d+m$ we obtain the same result as in the linear case (compare Theorem 3.1 in both papers), when $d\geq 2k-m>\frac{d}{2}$, $Jy(\phi)$ in general diverges. The "likelihood ratio" (involving the change of $h(\cdot)$ to zero, i.e., making u and y independent) converges exactly as in the linear case (consider lemma 2.2), but the problematic term is:

$$\eta^{\varepsilon}(\phi) \triangleq E_0[\exp\{"\int_D (F(\phi) - F(u+\phi))\circ n + Tr(D_{\phi}F(P+D_0F)^{-1})"\} \quad \Big| \, \|u\|_{m,2}<\varepsilon] \qquad (1.5)$$

which arises from Ramer's theorem, where E_0 denotes expectations w.r.t. the reference measure (P_0), and $D_{\phi}F$ denotes a Frechet derivative at ϕ. In the linear case this term is part of the term $E_0[\exp-\int_D (P\phi)n | \|u\|<\varepsilon]$ that converges to 1 as $\varepsilon\to 0$ (see equation (3.4) in [2]). This is exactly the term that contributed a (finite) correction to the Onsager-Machlup functional for diffusions in [4].

To understand intuitively the behavior of $\eta^{\varepsilon}(\phi)$ for $\varepsilon\to 0$, let us replace $F(u+\phi)-F(\phi)$ by the linearized operator $D_{\phi}F\,(P+D_0F)^{-1}n$ (since $\|u\|_{m,2}\leq\varepsilon$, this should not change much the results), expand n and u w.r.t. ψ_i, the eigenfunctions of the operator $(P+D_0F)^{-1}$ and ignore the cross terms. So our "approximation" of $\eta^{\varepsilon}(\phi)$ will be:

$$\tilde{\eta}^{\varepsilon}(\phi) \triangleq E[\exp \sum_{i=1}^{\infty} f_i/\lambda_i(a_i^2-1) | \sum_{i=1}^{\infty} \frac{a_i^2}{\lambda_i^2} \leq \varepsilon^2] \qquad (1.6)$$

where $f_i \triangleq -\int_D (D_{\phi}F\psi_i)\psi_i$, and a_i are i.i.d. N(0,1) random variables. For $\varepsilon\to 0$, for those i with $\lambda_i\varepsilon << 1$, a_i is very small due to the conditioning, while when $\lambda_i\varepsilon >> 1$ the conditioning has little effect on a_i. So, again "approximately", the value of $\tilde{\eta}_{\varepsilon}\phi)$ behaves like $\exp \sum_{\lambda_i\leq 1/\varepsilon}[f_i/\lambda_i]$.

From [1] we have the estimate $\lambda_i \sim i^{2k/d}$ and moreover the f_i are bounded by $i^{m/\partial}$ under the smoothness assumptions we will impose on DF. Therefore $\sum_{i=1}^{\infty} f_i/\lambda_i$ converges in general only when $2k-m>d$. When $d\geq 2k-m>d/2$ the most "likely" paths ϕ are those with minimal $\sum_{\lambda_i \leq 1/\varepsilon} f_i/\lambda_i$. Thus, had we had a convenient characterization of those ϕ, we could have constrained our estimation problem to this subspace, and normalize $\Lambda_\phi(u(\underline{x}))$ accordingly. Since we were unable to find a convenient substitute for $\sum_{\lambda_i \leq 1/\varepsilon} f_i/\lambda_i$, we pursue here only the case of smooth P, F, i.e., 2k-m>d. In particular, our prototype example (P) with $\delta = 0$ is excluded!

In the next section, we collect all the results that hold true for 2k-m>d/2, namely, existence of solutions of (1.1) (Theorem 2.1), as well as the expression of the Radon-Nykodim derivative $\Lambda_\phi(u(\underline{x}))$ (Theorem 2.2), and the convergence of the "likelihood ratio" for any $\phi \in W^{2k,2}(D)$, (Lemma 2.3).

The third section concentrate on the results in which we need 2k-m>d, namely, the existence of $J_y(\phi)$ for $\phi \in W_0^{\ell,2}(D)$, $\ell > 2k+d/2+m$, (Theorem 3.1), and that of a solution $\hat{\phi} = \underset{\phi \in W_0^{2k,2}(\phi)}{\operatorname{argmax}} J_y(\phi)$ (Theorem 3.2). Finally, in Theorem 3.3 we represent $\hat{\phi}$ by means of a weak solution of an appropriate stochastic PDE, and check that indeed $\hat{\phi} \in W_0^{\ell,2}(D)$, $\ell < 4k - d/2$.

We note that the existence of $J_y(\phi)$ for $\phi \in W_0^{\ell,2}(D)$ requires conditioning on $\|u\|_{m+4k-\ell,2} < \in$, instead of $\|u\|_{m,2} < \in$ as in the Gaussian case. We elaborate more in this issue in Section III.

II. PRIOR MODEL, AND EQUIVALENCE OF MEASURES

In the sequel $\| \|_{\ell,2}$ will denote the Sobolev norm in $W^{\ell,2}(D)$, and $\| \|_2$ the usual $L^2(D)$ norm. $W_0^{\ell,2}(D)$ will denote the closure of $C_0^\infty(D)$ w.r.t. the norm $W^{\ell,2}(D)$, and $W_0^{-\ell,2}(D)$ will denote the space of distributions which is the dual of $W_0^{\ell,2}(D)$. We denote by n the random distribution valued white noise in D, i.e., for each $\phi \in C_0^\infty(D)$, $n(\phi)$ is a Normal random variable of zero mean and variance $\|\phi\|_2^2$. Note that n is $W_0^{-\ell,2}(D)$ valued $\forall \ell > d/2$ [8], and that for $\phi \in C_0^\infty(D)$ and any basis e_i of $L^2(D)$, $n(\phi) = \sum_{i=1}^\infty a_i(\phi, e_i)$ (in q.m.), where $a_i = n(e_i)$ are i.i.d. N(0,1) random variables.

By a solution to (1.1) we mean a $W^{m,2}(D)$ valued random variable u such that u has a continuous version and $B[\phi,u] + (F(u),\phi) = n(\phi)$ for all $\phi \in C_0^\infty(D)$. Note that the boundary conditions need not be classically defined. We assume throughout that F is a continuous operator from $W^{m,2}(D)$ to $L^2(D)$, whose $W^{m,2}(D)$-
derivative $D_uF: W_0^{m,2}(D) \to L^2(D)$ exists, the mapping $u \to D_uF$ is continuous , and uniformly bounded, i.e., $\sup_{u \in W^{m,2}(D)} \|D_uF\| \triangleq \|DF\| < \infty$. We further assume $F(0) = 0$ (if this is not the case one can always accomodate F(0) into P). For example, $f \in C^1(\mathbf{R})$ with bounded derivative, is such an operator for m=0. Now we can prove using Picard iterations the following basic existence theorem:

Theorem 2.1: Let $2k > \frac{d}{2} + m$, $|B[\phi,\phi]| \geq c_P\|\phi\|_{k,2}^2$, and $c_P > \|DF\|$, then (1.1) has a unique solution in $W_0^{m,2}(D)$, which is in $C^{m+\alpha}(D)$ for some $\alpha > 0$.

Proof: Consider the linear PDE:

$$\begin{cases} P\tilde{u} = n & \text{on } D \\ P_\partial \tilde{u} = 0 & \text{on } \partial D \end{cases} \tag{2.1}$$

This PDE admits a unique solution in $W_0^{m,2}(D)$ since $2k-\frac{d}{2} > m$ (the proof is an easy extension of the proof in [8,2]), which is in $C^{m+\alpha}(D)$ for some $\alpha>0$. Thus, w.l.o.g., we can consider the solutions of:

$$\begin{cases} P\bar{u} = -F(\bar{u} + \tilde{u}) & \text{on } D \\ P_\partial \bar{u} = 0 & \text{on } \partial D \end{cases} \tag{2.2}$$

where P_∂ denotes Dirichlet conditions, and let the solution of (1.1) be $u = \bar{u} + \tilde{u}$. We shall prove by Picard iterations that (2.2) admits a unique solution $\bar{u}$ in $W_0^{2k,2}(D)$.

For $v, u \in W_0^{m,2}(D)$ and $0 \le x \le 1$,

let $g(x) \overset{\Delta}{=} \|F(u+xv) - F(u)\|_2 - x\|v\|_{m,2}\|DF\|$. Note that

$$\begin{aligned} g(x+\varepsilon) - g(x) &= \|F(u+(x+\varepsilon)v) - F(u)\|_2 - \|F(u+xv) - F(u)\|_2 - \varepsilon\|v\|_{m,2}\|DF\| \\ &\le \|F(u+xv + \varepsilon v) - F(u+xv)\|_2 - \varepsilon\|v\|_{m,2}\|DF\| \\ &\le \|F(u+xv + \varepsilon v) - F(u+xv)\|_2 - \varepsilon\|D_{u+xv} F(v)\|_2 \\ &\le \|F(u+xv+\varepsilon v) - F(u+xv) - \varepsilon D_{u+xv}F(v)\|_2 \end{aligned}$$

and therefore, $\limsup_{\varepsilon\to 0} \frac{1}{\varepsilon}(g(x+\varepsilon) - g(x)) \le 0$. Since $g(0) = 0$, one concludes that $g(1) \le 0$, i.e. $\|F(u+v) - F(u)\|_2 \le \|DF\| \, \|v\|_{m,2}$. In particular, substituting u=0 one has $\|F(v)\|_2 \le \|DF\| \, \|v\|_{m,2}$, $\forall v \in W_0^{m,2}(D)$.

<u>**Existence:**</u> We construct a sequence of functions $\bar{u}^{(\ell)} \in W_0^{2k,2}(D)$ by: $\bar{u}^{(0)} = 0$, and $\bar{u}^{(\ell)}$ is the solution of the linear PDE:

$$\begin{cases} P\bar{u}^{(\ell)} = -F(\bar{u}^{(\ell-1)} + \tilde{u}) & \text{on } D \\ P_\partial \bar{u}^{(\ell)} = 0 & \text{on } \partial D \end{cases} \tag{2.3}$$

Since $-F(\bar{u}^{(\ell-1)} + \tilde{u}) \in L^2(D)$, there exists a unique solution $\bar{u}^{(\ell)}$ of (2.3) in $W_0^{2k,2}(D)$ by the classical theory of elliptic PDE's (c.f. [1]).

Now, let $\delta^{(\ell)} \triangleq \bar{u}^{(\ell)} - \bar{u}^{(\ell-1)}$, then $\delta^{(\ell)}$ are the unique solutions of:

$$\begin{cases} P\delta^{(\ell)} = F(\bar{u}^{(\ell-2)} + \tilde{u}) - F(\bar{u}^{(\ell-2)} + \delta^{(\ell-1)} + \tilde{u}) & \text{on } D \\ P_\partial \delta^{(\ell)} = 0 & \text{on } \partial D \end{cases} \tag{2.4}$$

in $W_0^{2k,2}(D)$. Now, $\exists \varepsilon>0$, s.t.

$$\|P\delta^{(\ell)}\|_2 = \|F(\bar{u}^{(\ell-2)} + \tilde{u}) - F(\bar{u}^{(\ell-2)} + \tilde{u} + \delta^{(\ell-1)})\|_2 \le c_p(1-\varepsilon)\|\delta^{(\ell-1)}\|_{2k,2}$$

However, $|B[\phi,\phi]| \ge c_P\|\phi\|_{k,2}^2 \ \forall \phi \in C_0^\infty(D)$ implies that $\|P\delta\|_2 \ge c_P\|\delta\|_{2k,2}$ $\forall \delta \in W_0^{2k,2}(D)$. In particular, we obtain:

$$\|\delta^{(\ell)}\|_{2k,2} \le (1-\varepsilon)\|\delta^{(\ell-1)}\|_{2k,2} \tag{2.5}$$

So that, $\forall \ell > \ell'$:

$$\|\bar{u}^{(\ell)} - \bar{u}^{(\ell')}\|_{2k,2} \le \sum_{r=\ell'+1}^{\ell} \|\delta^{(r)}\|_{2k,2} \le \frac{(1-\varepsilon)^{\ell'}}{\varepsilon} \|\bar{u}^{(1)}\|_{2k,2} \tag{2.6}$$

Thus, $\bar{u}^{(\ell)} \underset{\ell \to \infty}{\longrightarrow} \bar{u}^*$ in $W_0^{2k,2}(D)$, and since F is continuous we have:

$$0 = P\bar{u}^{(\ell)} + F(\bar{u}^{(\ell-1)} + \tilde{u}) \underset{\ell \to \infty}{\longrightarrow} P\bar{u}^* + F(\bar{u}^* + \tilde{u}) \tag{2.7}$$

Note that:

$$\|u\|_{m,2} = \|\bar{u}^* + \tilde{u}\|_{m,2} \le \|\tilde{u}\|_{m,2} + \|\bar{u}^* - \bar{u}^{(1)}\|_{2k,2} + \|\bar{u}^{(1)}\|_{2k,2} \le \|\tilde{u}\|_{m,2} + \frac{1}{\varepsilon}\|\bar{u}^{(1)}\|_{2k,2}$$

$$\leq \|\tilde{u}\|_{m,2} + \frac{1}{\varepsilon c_P}\| F(\tilde{u})\|_2 \leq K\|\tilde{u}\|_{m,2} \tag{2.8}$$

and therefore $\|u\|_{m,2}$ is bounded by $\|\tilde{u}\|_{m,2}$, where $\tilde{u}$ is the Gaussian field corresponding to $F \equiv 0$.

__Uniqueness :__ Assume $\bar{u}_1, \bar{u}_2$ are two solutions of (2.2) in $W_0^{m,2}(D)$ and let $v = \bar{u}_1 - \bar{u}_2$, then v is the solution of:

$$\begin{cases} Pv = F(\tilde{u} + \bar{u}_2) - F(\tilde{u} + \bar{u}_2 + v) \\ P_\partial v = 0 \end{cases} \tag{2.9}$$

Now:

$$c_P\|v\|_{m,2} \leq \|Pv\|_2 = \|F(\tilde{u} + \bar{u}_2) - F(\tilde{u} + \bar{u}_2 + v)\|_2 < c_P\|v\|_{m,2} \tag{2.10}$$

a contradiction unless $\|v\|_{m,2} = 0$.

We next turn to the computation of $\Lambda_\phi(u(x))$. We first show the following preliminary lemma:

__Lemma 2.1:__

Let $\phi \in W_0^{2k}(D)$. Define $Ku \triangleq [P\phi + F(\phi+u) - F(u)]$ and define

$$\tilde{K}v \triangleq K((P+F)^{-1}v)$$

where $(P+F)^{-1}: L^2(D) \to W_0^{2k,2}(D)$ exists due to our conditions on P,F.

Throughout, let $Pu + F(u) = v$, $v \in L^2(D)$. Then:

a) $$(D_v\tilde{K}) = (D_{u+\phi}F - D_uF)(P+D_uF)^{-1} \tag{2.11}$$

in the sense that both sides of (2.11) exist.

b) $D_v\tilde{K}$ is a Hilbert-Shmidt operator.

c) If $2k > d+m$, $D_v\tilde{K}$ is a trace class operator.

Proof: a) Let $u = \tilde{u}+\bar{u}$ be the decomposition as in theorem (2.1), i.e. $\tilde{u}$ is the weak solution of:

$$\begin{cases} P\tilde{u} = \nu & \text{on } D \\ P_{\partial}\tilde{u} = 0 & \text{on } \partial D \end{cases}$$

and $\bar{u}$ is the unique solution in $W_0^{2k,2}(D)$ of:

$$P\bar{u} + F(\bar{u}+\tilde{u}) = 0 \tag{2.12}$$

Let now $h \in L^2(D)$ and $\bar{\bar{u}}, \bar{\bar{\bar{u}}}$ be the unique solutions in $W_0^{2k,2}(D)$ of:

$$P\bar{\bar{u}} + F(\bar{\bar{u}} + \tilde{u}) = h \tag{2.13}$$

$$P\bar{\bar{\bar{u}}} + D_{\bar{u}+\tilde{u}}F(\bar{\bar{\bar{u}}}) = h \tag{2.14}$$

Now, let

$$A(h) \overset{\Delta}{=} \|\tilde{K}(\nu+h) - \tilde{K}\nu - (D_{u+\phi}F - D_uF)(P+D_uF)^{-1}h\|_2 \leq \|F(\phi+\tilde{u} + \bar{\bar{u}}) - F(\phi+\tilde{u}+\bar{u}) - D_{\phi+\tilde{u}+\bar{u}}F\bar{\bar{\bar{u}}}\|_2$$

$$+ \|F(\tilde{u} + \bar{\bar{u}}) - F(\tilde{u} +\bar{u}) - D_{\tilde{u}+\bar{u}}F\bar{\bar{\bar{u}}}\|_2 \tag{2.15}$$

But:

$$\|F(\tilde{u} + \bar{\bar{u}}) - F(\tilde{u} + \bar{u}) - D_{\tilde{u}+\bar{u}}F\bar{\bar{\bar{u}}}\|_2 \leq \|F(\tilde{u} + \bar{\bar{u}}) - F(\tilde{u} + \bar{u}) - D_{\tilde{u}+\bar{u}}F(\bar{\bar{u}} - \bar{u})\|_2 +$$
$$+ \|D_{\tilde{u}+\bar{u}}F(\bar{\bar{u}} - \bar{u} - \bar{\bar{\bar{u}}})\|_2 \leq o(\|\bar{\bar{u}} - \bar{u}\|_{2k,2}) + \|DF\| \, \|\bar{\bar{u}} - \bar{u} - \bar{\bar{\bar{u}}}\|_{2k,2}$$

and similarly for the other term in the RHS of (2.15).

Therefore, to complete the proof we only have to show that:

$$\|\bar{\bar{u}} - \bar{u}\|_{2k,2} \leq c\, \|h\|_2 \tag{2.16a}$$

$$\|\bar{\bar{u}} - \bar{u} - \bar{\bar{\bar{u}}}\,\|_{2k,2} = o(\|h\|_2) \tag{2.16b}$$

However, let $\delta = \bar{\bar{u}} - \bar{u}$, then

$$(P+D_uF)\delta = h - [F(u+\delta) - F(u) - (D_uF)\delta]$$

Since $\inf[\lambda_i(P)] \geq c_P > \|DF\| \geq \|D_uF\|$, $(P+D_uF)^{-1}$ is a bounded linear operator from $L_2(D)$ to $W_0^{2k,2}(D)$, and therefore: $\|\delta\|_{2k,2} \leq \|(P+D_uF)^{-1}\|\{\|h\|_2 + \|F(u+\delta)-F(u) - D_uF\delta\|_2\}$. For $\|h\|_2$ small, $\|F(u+\delta) - F(u) - D_uF\delta\|_2 = o(\|\delta\|_{2k,2}) = o\,(\|h\|_2)$, and (2.16a) is established. Now let $\tilde{\delta} \triangleq \delta - \bar{\bar{\bar{u}}}$, then,

$$(P+D_uF)\tilde{\delta} = -(F(u+\delta) - F(u) - D_uF\delta) \tag{2.17}$$

And thus, $\|\tilde{\delta}\|_{2k,2} \leq \|(P+D_uF)^{-1}\|\{\|F(u+\delta) - F(u) - D_uF\delta\|_2\} = o\,(\|h\|_2)$ due to (2.16a).

b) We need to show that $Tr(D_v\tilde{K}^*D_v\tilde{K}) < \infty$. Let e_i be the complete orthonormal basis of $L^2(D)$ composed of the generalized eigenvectors of $P+D_uF$. One has

$$Tr(D_v\tilde{K}^*D_vK) = \sum_{i=1}^{\infty} \|(D_{u+\phi}F - D_uF)(P+D_uF)^{-1}e_i\|_2^2 \leq 4\|DF\|^2 \sum_{i=1}^{\infty} \frac{1}{\lambda_i^2} \tag{2.18}$$

Note that the behavior at infinity of λ_i is the same as that of $\lambda_i(P)$, whereas $\lambda_i \geq \varepsilon > 0$ for some ε since $\|D_uF\| < \inf \lambda_i(P)$. Therefore, the RHS of (2.18) is bounded by $c \sum_{i=1}^{\infty} i^{-4k/d} < \infty$.

c) We need to show that $Tr(D_v\tilde{K}) < \infty$. Repeating the argument as above, one obtains that

$$Tr(D_v\tilde{K}) \leq c_1 \sum_{i=1}^{\infty} \frac{1}{\lambda_i} \leq c_2 \sum_{i=1}^{\infty} i^{-2k/d} < \infty \tag{2.19}$$

whenever 2k>d.

Before we may state the theorem concerning the structure of $\Lambda_\phi(u(x))$, we introduce some notations.

Definitions:

1. Let A be a Hilbert-Schmidt operator. The Carlemen-Fredholm determinant of A is defined as

$$\Delta(A) \triangleq \prod_{i=1}^{\infty} (1+\lambda_i) \exp(-\lambda_i) \tag{2.20}$$

and the product in (2.20) converges absolutely (see [3, XI.9.22]).

2. Let A be a trace-class operator. The determinant of A is defined by

$$|(I+A)| = \det(I+A) \triangleq \prod_{i=1}^{\infty} (1+\lambda_i) \tag{2.21}$$

and the product (2.21) converges absolutely (see [3, XI, 9.22]). We quote from [3, XI.9.18,19,22,23] the following lemma:

Lemma 2.2:

a) Let A,B be trace class operators. Then

$$\frac{d}{dz} \det(I+A+zB)\Big|_{z=0} = \det(I+A) \cdot \operatorname{tr}(I+A)^{-1}B \tag{2.22}$$

and det(I+A+zB) is analytic in z.

b) $\ln \det(AB) = \ln \det(A) + \ln \det(B)$

c) $\Delta(A) = \exp(-\operatorname{tr}(A))\det(I+A)$

Finally, we introduce the following stochastic integral which was first defined by Ramer (c.f. [7, lemma 4.2]):

Definition: Let A be an L^2 differentiable nonlinear operator whose derivative, A_x, is Hilbert-Schmidt. Let n, a white noise, be represented by $n = \sum_i a_i e_i$, where $a_i \sim N(0,1)$ are independent and e_i is a complete, orthonormal basis in $L^2(D)$. Then

$$I_n \triangleq \left(A \sum_{i=1}^{n} a_i e_i ,\ \sum_{i=1}^{n} a_i e_i\right) - \sum_{i=1}^{n} (A_x e_i, e_i) \tag{2.23}$$

converges in $L^2(\mu)$ (where μ = Gaussian measure which makes $\{a_i\}$ i.i.d. N(0,1)) to a random variable I which is independent of the basis e_i. We use the notation

$$I = \int An \circ n - \operatorname{tr} D_n A \tag{2.24}$$

Note that when A_x is trace class, both terms in the RHS of (2.24) exist, the first as an Ogawa integral (c.f. [6,5 and 2]), and their difference indeed equals I. We have completed the preliminaries required to compute $\Lambda_\phi(u(x))$:

Theorem 2.2. Let $h(\cdot) \in C^1(\mathbf{R})$ with a bounded derivative. Assume the conditions of theorem 2.1, let $u \in W^{m,2}(D)$ denote the solution of (1.1), and let $\phi \in W_0^{2k,2}(D)$ be independent of n. Then

$$\Lambda_\phi(u(x)) = \exp\Big(\int_D h(u(\underline{\theta}))d\tilde{w}(\underline{\theta}) - \frac{1}{2}\int_D h^2(u(\underline{\theta}))d\underline{\theta}\Big)$$

$$\exp\Big[-\frac{1}{2}\int_D (P\phi + F(u+\phi) - F(u))^2$$

$$-\int_D (P\phi + F(u+\phi) - F(u)) \circ n + \operatorname{tr}((D_{u+\phi}F - D_uF)(P+D_uF)^{-1})\Big]$$

$$\cdot \Delta((D_{u+\phi}F - D_uF)(P+D_uF)^{-1}) \tag{2.25}$$

Proof: Define $P_{\tilde{w}}$ as the measure P_1 with h≡0, i.e. y independent of u. Exactly as in the linear case, one has

$$\frac{dP_1}{dP_{\tilde{w}}} = \exp\left[\int_D h(u(\underline{\theta}))d\tilde{w}(\underline{\theta}) - \frac{1}{2}\int_0 h^2(u(\underline{\theta}))d\theta\right] \tag{2.26}$$

where the stochastic integral, taken under $P_{\tilde{w}}$, is well defined since $u(\underline{x}) \in L^2(D)$ is independent of $\sigma(\tilde{w}(\theta), \theta \in D)$. This part of $\Lambda_\phi(u(x))$ is denoted as the likelhood ratio. In the sequel, therefore, we can assume h≡0. Let now $u \in W_0^{m,2}(D)$ be the unique solution of (1.1), and associate it with $v = Pu + F(u)$ (v is the "white noise"). Under P_0, v=n whereas under $P_{\tilde{w}}$, $v + \tilde{K}v = n$. By lemma (2.1) and Ramer [7], we have

$$\frac{dP_{\tilde{w}}}{dP_0} = \Delta(D_v\tilde{K}) \exp[-\frac{1}{2}\int (\tilde{K}v)^2 - \int (\tilde{K}v)\circ n + tr(D_v\tilde{K})] \tag{2.27}$$

where again, when $D_v\tilde{K}$ is Hibert-Schmidt then the sum of the two RHS terms in the argument of the exponent is well defined and further when $D_v\tilde{K}$ is trace class each is defined individually. Note that $\tilde{K}:W^{-d/2-\delta,2}(D) \to L^2(D)$ where $\delta \leq 2k - \frac{d}{2} - m$ and therefore $\tilde{K}v \in L^2(D)$. Combining now (2.26) and (2.27), the proof is completed.

We conclude this section with the first step towards computing the Onsager - Machlup functional:

<u>**Lemma 2.3**</u>: Under the conditions of Theorem 2.1, for $h(\cdot) \in C^{2k+1}(\mathbf{R})$ with all derivatives up to order (2k+1) bounded, and $\phi(\underline{x}) \in W^{2k,2}(D)$:

$$E_{\tilde{w}}\left[\frac{dP_1}{dP_{\tilde{w}}} | \sigma(y), \|u-\phi\|_{m,2} \leq \varepsilon\right] \underset{\varepsilon\to 0}{\to} \exp\left[\int_D h(\phi(\underline{\theta}))dy(\underline{\theta}) - \frac{1}{2}\int_D h^2(\phi(\underline{\theta}))d\underline{\theta}\right]$$

where the stochastic integral in the RHS is well defined for all $\phi \in W^{2k,2}(D)$ and $\tilde{w}$ in the support set of the measure defining y by the pairing between $W^{-d/2-\delta,2}(D)$ and $W^{2k,2}(D)$ $(2k>d/2+\delta)$. So, ϕ may be stochastic and depend on y.

<u>**Proof**</u>: This is lemma 4.1 in [2] and it admits exactly the same proof, as $\|u-\phi\|_2 \leq \|u-\phi\|_{m,2}$. We note here that for this proof we assumed that $E_{\tilde{w}}[\exp\ \varepsilon \|u\|_2^2] < \infty$ for $\varepsilon>0$ small enough. This, however, follows from (2.8) since for $\varepsilon>0$ small enough, the Gaussian field $\tilde{u}$ obtained in (1.3) for F=0 admits $E_{\tilde{w}}[\exp \varepsilon \|\tilde{u}\|_{m,2}^2] < \infty$.

III. THE ONSAGER-MACHLUP FUNCTIONAL, AND THE MAP ESTIMATOR

Let

$$J_y^\varepsilon(\phi) \triangleq \frac{P(\|u-\phi\|_{m+d/2+\delta} < \varepsilon \mid \sigma\{y\})}{P(\|u\|_{m+d/2+\delta} < \varepsilon)} \tag{3.1}$$

$$= E_0(\Lambda_\phi(u(x)) \mid \|u\|_{m+d/2+\delta} < \varepsilon, \sigma\{y\})$$

where $0<\delta<2k-m-d \triangleq \delta_0$ and $2\delta>\delta_0$. As mentioned earlier, we assume throughout $\delta_0>0$, and consider $J_y(\phi) = \lim_{\varepsilon\to 0} J_y^\varepsilon(\phi)$. We postpone the discussion on the specific neighborhoods appearing in (3.1) and on the reason for $2\delta > \delta_0 > \delta$ to the remarks immediately following the proof of theorem 3.1 below. Note that $\delta_0 > 0$ implies that $Tr(D_uFP^{-1}) \leq \|DF\|\, Tr(P^{-1}) < \infty$ and therefore $D_uF\,P^{-1}$ is a trace class operator. Moreover,

$$|\lambda_i(D_uFP^{-1})| \leq \frac{\|DF\|}{c_P} < 1$$

Define:

$$\hat{J}_y(\phi) \triangleq \int_D h(\phi(\underline{\theta}))dy(\underline{\theta}) - \frac{1}{2}\int_D h(\phi(\underline{\theta}))^2 d\underline{\theta} - \frac{1}{2}\int_D (P\phi + F(\phi))^2 \tag{3.2}$$

$$+ \ln\left|I + (D_\phi FP^{-1})\right| - \ln\left|I + (D_0FP^{-1})\right|$$

where the stochastic integral on the RHS (and hence the whole RHS) is defined for any $\phi \in W_0^{2k,2}(D)$, and $\tilde{w}$ in the support set of the measure defining y by the pairing between $W^{2k,2}(D)$ and $W^{-d/2-\delta_0,2}(D)$.

Theorem 3.1: Assume that $\|D_vF-D_uF\| \leq K\|v-u\|_{m,2}$ for some $K<\infty$. Then, for any $\phi \in W_0^{4k-d/2-\delta,2}(D)$,

$$\lim_{\varepsilon\to 0} [J_y^\varepsilon(\phi)] = \exp \hat{J}_y(\phi) \qquad \text{a.s. in } \tilde{n} \tag{3.3}$$

Proof: In view of Theorem (2.2) and Lemma (2.3), we can without loss of generality assume $h \equiv 0$, and then assume ϕ deterministic.

Note that

$$\mathrm{Tr}((D_{u+\phi}F - D_uF)(P+D_uF)^{-1}) \leq 2\|DF\| \, \mathrm{Tr}(P^{-1}(I+D_uFP^{-1})^{-1}) \tag{3.4}$$

$$\leq \left(\frac{2\|DF\|c_P}{c_P - \|DF\|}\right) \mathrm{Tr}(P^{-1}) < \infty.$$

On the other hand, note that for every $\delta^` > 0$ u, $u+\phi \in W_0^{m+d/2+\delta^`}(D)$ and

therefore the stochastic integral w.r.t. n in (2.25) is defined without the need of the trace term correction. Combining the above facts, and using lemma (2.2c), we rewrite (2.25) (again, with $h \equiv 0$!) as

$$\Lambda_\phi(u(x)) = \det(I+(D_{u+\phi}F - D_uF)(P+D_uF)^{-1}) \tag{3.5}$$

$$\exp\left(-\frac{1}{2}\int_D (P\phi + F(u+\phi) - F(u))^2 - \int_D \left[P\phi + (F(u+\phi) - F(u))\right] \mathrm{o}n\right)$$

Note now that

$$(I + (D_{u+\phi}F - D_uF)(P+D_uF)^{-1}) = (I+D_{u+\phi}FP^{-1})(I+D_uFP^{-1})^{-1} \tag{3.6}$$

and since both operators on the RHS of (3.6) are trace class, we use lemma (2.2b) and (3.6) to rewrite (3.5) as:

$$\Lambda_\phi(u(x)) = \exp(\hat{J}_y(\phi)) \; A_1 \, A_2 \, A_2 \, A_4 \, A_5 \, A_6 \tag{3.7}$$

where

$$\ln A_1 \triangleq \ln \det(I+D_0FP^{-1}) - \ln \det (I+D_uFP^{-1}) \tag{3.8a}$$

$$\ln A_2 \triangleq \ln \det(I+D_{u+\phi}FP^{-1}) - \ln \det(I+D_{\phi}FP^{-1}) \tag{3.8b}$$

$$\ln A_3 \triangleq -\int_D P\phi \text{ on} + \int_D P\phi\, F(u) \tag{3.8c}$$

$$\ln A_4 \triangleq -\int_D (F(u+\phi) - F(u))\text{on} + \int_D (F(u+\phi) - F(u))F(u) \tag{3.8d}$$

$$\ln A_5 \triangleq -\int_D \Big(F(u+\phi) - F(\phi)\Big)\left(P\phi + \frac{F(\phi) + F(u+\phi)}{2}\right) \tag{3.8e}$$

$$\ln A_6 \triangleq +\int_D \frac{1}{2} F^2(u) \tag{3.8f}$$

Consider first $\ln A_1$. Using lemma (2.2b,c), one has

$$\ln \det (I+D_0FP^{-1}) - \ln \det(I+D_uFP^{-1}) = \ln \det (I+D_0FP^{-1})(I+D_uFP^{-1})^{-1}$$

$$= \ln \det(I+(D_0F-D_uF)(P+D_uF)^{-1})$$

$$= \sum_{n=1}^{\infty} \frac{1}{n} \operatorname{tr}\{(D_uF - D_0F)\,(P+D_uF)^{-1}\}^n$$

Therefore,

$$|\ln A_1| \leq \sum_{n=1}^{\infty} \frac{1}{n} \|(I+D_uFP^{-1})^{-1}\|^n \|D_0F-D_uF\|^n \|P^{-1}\|^{n-1} \operatorname{Tr}(P^{-1})$$

$$\leq c_P^1 \operatorname{Tr}(P^{-1}) \sum_{n=1}^{\infty} \frac{1}{n} \left(\frac{c_P^2}{c_P - \|DF\|}\right)^n \|D_0F-D_uF\|^n \leq c_1\, \|D_0F-D_uF\|$$

provided $\|D_0F-D_uF\| < \dfrac{c_P - \|DF\|}{c_P^2}$.

Therefore, for $\|u\|_{m,2}$ small enough (under P_0!),

$$|\ln A_1| < c_1\|u\|_{m+d/2+\delta,2} \tag{3.9a}$$

and repeating the argument,

$$|\ln A_2| < c_2\|u\|_{m+d/2+\delta,2} \tag{3.9b}$$

where c_1, c_2 are <u>deterministic.</u>

Since $\phi \in W_0^{4k-\frac{d}{2}-\delta,2}(D)$, one has

$$|\ln A_3| = |(P^*P\phi,u)| \le \|u\|_{m+\frac{d}{2}+\delta,2}\|P^*P\phi\|_{-(m+d/2+\delta,2)} \le c_3(\phi)\|u\|_{m+\frac{d}{2}+\delta,2} \tag{3.9c}$$

with $c_3(\phi) < \infty$. Turning to (3.8d), note that

$$|\ln A_4| = |(P^*(F(\phi+u) - F(u)),u)| \le \|u\|_{m+d/2+\delta,2}\left(\|P^*F(\phi+u)\|_{-(\frac{d}{2}+m+\delta),2} + \|P^*F(u)\|_{-(\frac{d}{2}+m+\delta),2}\right) \tag{3.9d}$$

$$\le 2\|u\|_{m+\frac{d}{2}+\delta,2}\|P^*\|\,\|DF\|\left(\|u\|_{m+\frac{d}{2}+\delta,2} + \|u+\phi\|_{\frac{d}{2}+m+\delta,2}\right) \le c_4(\phi)\|u\|_{m+d/2+\delta,2}$$

where $\|P^*\|$ is the operator norm of $P^*: W^{d/2+\delta,2} \to W^{-(d/2+m+\delta),2}$ which is clearly bounded.

Next, note that

$$|\ln A_5| \le \tilde{c}_5(\phi)\|F(u+\phi) - F(\phi)\|_2 \le \tilde{c}_5(\phi)\|DF\|\,\|u\|_{m,2}$$

$$\le c_5(\phi)\|u\|_{m+d/2+\delta,2} \tag{3.9e}$$

Similarly,

$$|\ln A_6| \leq \tilde{c}_6\, \|DF\|^2 \|u\|^2_{m,2} \leq c_6\, \|u\|^2_{m+d/2+\delta,2} \tag{3.9f}$$

Combining (3.9) and (3.7) yields the theorem.

Remarks. We explain below why we used the $W^{m+d/2+\delta,2}(D)$ neighborhoods in the conditioning in (3.1) and why $2\delta>\delta_0$.

Recall (c.f. the introduction immediately below (1.5)) that the requirement $\delta_0>0$ (i.e., $2k>m+d$) was necessary in order to be able to have a nice (i.e., without additional constraints) characterization of the limit $\hat{J}_y(\phi)$. Obviously, had we allowed in the conditioning distributional neighborhoods, this problem could have been avoided. However, we choose not to do so because then interpretation of the results in terms of MAP estimation is not clear. Note that in the 1-D case the problem is avoided since no singularities occur (i.e., all correction terms are finite).

Next, note that in order to have (3.9c) with ϕ-s having only the degree of smoothness guaranteed for the solution of the estimation problem, we had to take $W^{m+d/2+\delta,2}$ neighborhoods, any $0<\delta<\delta_0$, or alternatively, impose structural conditions (which are satisfied by the optimizer) as in the linear case [2, theorem 3.2] or 1-D case [9]. However, here structural conditions would not help due to the term A_4: the integrand being dependent on u, one cannot apply an integration by parts as in [2, theorem 3.2]. In the 1-D case, the problem is avoided basically using the Ito calculus. Here, since we don't have such a powerful tool, we have to require the right degree of smoothness to be able to write (3.9d). This leads to $2\delta>\delta_0$.

We turn now to the existence issue and claim:

Theorem 3.2: There exists a solution to the problem:

$$\hat{\phi} = \underset{\phi\in W^{2k,2}(D)}{\operatorname{argmax}}\ \hat{J}_y(\phi) \tag{3.10}$$

Proof: We follow the same steps as in [2, 10]; we therefore give the details below only of those parts of the argument which are new:

(a) For $y\in C(D)$, $\hat{J}_y(0) = \int_D h(0)dy(\underline{\theta}) - \frac{1}{2}h(0)^2 \mathrm{Vol}(D) > -\infty$.

(b) We have bounded $|\ln(|I+D_\phi FP^{-1}|)|$ by a uniform bound (independent of ϕ).

Further, $\|F(\phi)\|_2 \leq \|DF\| \, \|\phi\|_{2k,2}$, $\|P\phi+F(\phi)\|_2^2 \geq \|P\phi-\|DF\|\phi\|_2^2$ while $(P-\|DF\|I)$

is bounded away from zero, so we can follow the proof of [2] to show that

$$\lim_{\|\phi\|_{2k,2} \to \infty} \hat{J}(\phi) = -\infty.$$

(c) To show that $\hat{J}(\phi)$ is lower semi continuous w.r.t. the weak topology in $W^{2k,2}(D)$, we note that the additional terms (w.r.t. [2]) in the non-linear case do not cause any problem, since:

(1) $F: W^{2k,2}(D) \to W^{2k-m,2}(D)$ is a continuous mapping, and since $(2k-m) > d$, $\phi_n \overset{\omega}{\to} \phi$ implies $F(\phi_n) \to F(\phi)$ pointwise, and also in $W^{d/2,2}(D)$, so the cross term $P_\phi \cdot F(\phi)$ will be continuous as well as the $F(\phi)^2$-term.

(2) Note that

$$|\ln \det(I+D_{\phi_n}FP^{-1}) - \ln \det(I+D_{\phi}FP^{-1})| \leq$$
$$\leq c_1 \|\phi_n - \phi\|_{m,2}$$

due to the argument preceding (3.9a). However, $\phi_n \underset{W^{2k,2}(D)}{\overset{\omega}{\to}} \phi$, implies, because $2k > m + d/2$, that $\|\phi_n - \phi\|_{m,2} \to 0$. Combining all the above, we conclude that $\hat{J}(\phi)$ is lower semicontinous w.r.t. the weak topology in $W^{2k,2}(D)$ which, combined with (a), (b) and [2], yields the theorem.

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

Theorem 3.3: Let k,m be integers. Any maximizer of (3.10) is a weak solution of:

$$\begin{cases} (P+D_{\hat\phi} F)^*(P\hat\phi + F(\hat\phi)) = -\, h'(\hat\phi)h(\hat\phi) + h'(\hat\phi)o\dot y + G(\hat\phi) & \text{on } D \\ P_\partial \hat\phi = 0 & \text{on } \partial D \\ (P+D_{\hat\phi} F)^*_\partial (P\hat\phi + F(\hat\phi)) = 0 & \end{cases} \qquad (3.11)$$

where $(P+DF\hat\phi)^*_\partial$ denotes the boundary operator defined by $(P+D\hat\phi\, F)^*$ i.e., for $\phi_1, \phi_2 \in C^\infty(D)$,

$$\left((P+D_{\hat\phi} F)^*\phi_1, \phi_2\right) - \left(\phi_1, (P+D_{\hat\phi} F)\phi_2\right)$$

$$= \sum_{j=1}^{2k-1} \left[\overline{(P+D_{\hat\phi} F)^*_\partial \phi_1}\right]_j \frac{\partial^j \phi_2}{\partial \underline{n}^j}$$

with $\underline{n}$ being the exterior normal to D, $\overline{(P+D_{\hat\phi} F)^*_\partial}$ above defined by Green's formula, and

$$(P+D_{\hat\phi} F)^*_\partial \phi_1 = \left[\left(\overline{(P+D_{\hat\phi} F)^*_\partial}\, \phi_1\right)_k, \left(\overline{P+D_{\hat\phi} F)^*_\partial}\, \phi_1\right)_{k+1}, \ldots, \left(\overline{(P+D_{\hat\phi} F)^*_\partial}\phi_1\right)_{2k-1}\right]$$

Here $G(\hat\phi) \triangleq \sum_{i=1}^{\infty} Tr(I+D_{\hat\phi} FP^{-1})^{-1}[D^2_{\hat\phi} F <e_i, P^{-1}\cdot >]e_i$, with e_i any orthogonal basis of $W_0^{2k,2}(D)$ Note that we assumed the existence of the $W_0^{2k,2}(D)$-derivative of $D_\phi F$, w.r.t. ϕ, i.e., $\lim_{\|u\|_{2k,2}\to 0} \{\|D_{(\phi+u)}F(v) - D_\phi F(v) - D^2_\phi F <u,v>\|_2 / \|v\|_{m,2} \cdot \|u\|_{2k,2}\} = 0$, for every $\phi \in W_0^{2k,2}(D)$, $v \in W_0^{m,2}(D)$, and further assumed that $\|D^2_\phi F<u,v>\|_2 \leq K\|v\|_{m,2}\, \|u\|_{2k,2}$.

Proof: The proof is an easy application of the necessary conditions of the calculus of variations and is therefore omitted. The only interesting part is

$$\ln |(I+D_{\phi+\delta\phi}FP^{-1}| - \ln|I+D_{\phi}FP^{-1}| = \mathrm{Tr}\{(I+D_{\phi}FP^{-1})^{-1}D_{\phi}^{2}F < \delta\phi, P^{-1}\cdot >\} = \quad (3.12)$$

$$= \int_{0}(\sum_{i=1}^{\infty} \mathrm{Tr}(I+D_{\phi}FP^{-1})^{-1}D_{\phi}^{2}F \ <e_i, P^{-1}\cdot >e_i)\delta\phi = \int_{D} G(\phi)\delta\phi$$

where our assumptions guarantee that:

$$|\mathrm{Tr}(I+D_{\phi}FP^{-1})^{-1}[D_{\phi}^{2}F \ <\delta\phi, P^{-1}\cdot >]| = |\sum_{i=1}^{\infty}\frac{1}{\lambda_i}\int_{D}(I+D_{\phi}FP^{-1})^{-1}D_{\phi}^{2}F \ <\delta\phi, \psi_i>\psi_i| \leq$$

$$1/[1-\|DF\|/c_p]\sum_{i=1}^{\infty}\frac{1}{\lambda_i}\|\delta\phi\|_{2k,2}\|\psi_i\|_{m,2}\|\psi_i\|_2 < \infty$$

Here ψ_i are the generalized eigenfunctions of P in $W_0^{m,2}(D)$ and λ_i the corresponding eigenvalues. Note that using lemma 3.1 in [2] we can understand (3.11) as a pathwise equation for each $y \in C(D)$, and also the solution $\hat{\phi} \in W_0^{4k-d/2-\delta}(D)$, $\forall\delta>0$, i.e., $\hat{\phi} \in W_0^{2k+d/2+\delta_1}(D)$ $\forall\delta_1 < \delta_0$ as well, and so $\hat{\phi}$ satisfies the conditions of theorem 3.1.

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 I: Elliptic Gaussian Fields Observed Via a Noisy Nonlinear Channel, submitted.

[3] Dunford, N. and Schwartz, J.T., Linear Operators, Part 2, John Willey, New York, 1963.

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

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

[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] J.B. Walsh, An introduction to stochastic partial differential questions, Lecture Notes in Math. #1180, pp. 265-439, Springer 1986.

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

[10] 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.