Exponential estimates for convex norms and some applications

L.A. Shepp[1] and O. Zeitouni[2]

Abstract The role of correlation inequalities and martingale arguments in establishing conditional exponential bounds is reviewed. Applications to the computation of the Onsager Machlup functional for diffusions under non supremum norms follow.

1 Introduction

Many questions of interest related to diffusion processes boil down to the asymptotic evaluation of conditional exponential expectations. A problem in which such computations make their appearance is the evaluation of the Onsager-Machlup functional. Roughly, let $x(t)$ be the diffusion process which is the solution of the Stochastic Differential Equation (SDE)

$$dx(t) = f(x(t))dt + dw(t), \quad x(0) = 0, \quad x_t \in \mathbb{R}^d \quad . \tag{1}$$

The Onsager-Machlup functional is the limit

$$L(\phi, \dot{\phi}) = \log \lim_{\varepsilon \to 0} \frac{P(\|x - \phi\|_\infty < \varepsilon)}{P(\|w\|_\infty < \varepsilon)} = -\frac{1}{2}\int_0^T |\dot{\phi}(t) - f(\phi(t))|^2 dt - \frac{1}{2}\int_0^T \nabla f(\phi(t))dt \tag{2}$$

where $\|\cdot\|_\infty$ is the supremum norm on $[0, T]$. This functional was evaluated, in various degrees of generality, in [6],[7],[11],[13],[15].

The standard method of proof makes use of the fact that one deals with the supremum norm, and uses a radial decomposition to achieve the crucial exponential estimates. This approach, although quite successful, does not lend itself to simple generalization to other norms.

Recently, a partial study of this theorem for Hölder norms was initiated by [1], based on the old results of [4]. In this paper, we propose a different method for the evaluation of the required exponential estimates. Our main tools in the computation are correlation inequalities of the FKG type, and in particular a result of [5]. This technique was already applied, in the context of Onsager-Machlup computations and support evaluation, in the short paper [12] and in [2],[3],[9],[10]. Here, we will take care to make explicit the requirements on the norms involved for the method to be applicable. We will thus obtain the Onsager-Machlup functional for a variety of norms.

The organization of this paper is as follows. In section 2, we present a basic conditional expectation result, under general conditions on the norms involved. We also check that the

[1]AT&T Bell Laboratories, 600 Mountain Av., Murray Hill, N.J.

[2]Dept. of Electrical Engineering, Technion, Haifa 32000, Israel

Hölder norms Lip_α with $\alpha < 1/3$, L^p norms with $p \geq 4$, and the supremum norm all satisfy these conditions. Section 3 deals with the Onsager-Machlup problem.

This paper was motivated by the talk of Prof. Baldi at the stochastic analysis meeting in Sant Feliu de Guixol, where related results, using different methods, were presented. We would like to take this opportunity to thank the organizers, M. Sanz and D. Nualart, for this stimulating conference.

Notations: throughout, v^* denotes the transpose of a vector (matrix) v.

2 Conditional Expectation Theorems

Throughout this paper, d is a given positive integer, and $C_{0,d}[0,T]$ denotes the space of $\mathbb{R}^d$ valued, continuous functions $\phi(t)$ on $[0,T]$, with $\phi(0) = 0$. The components of $\phi(t)$ are denoted by $\phi_i(t) \in C_{0,1}[0,T]$.

A completely convex norm $\|\cdot\|$ is a measurable norm on $C_{0,d}[0,T]$ which satisfies the following condition: For every $i = 1, 2, \ldots, d$, every $\varepsilon > 0$, and every fixed component $\left(\tilde{\phi}_1(\cdot), \ldots, \tilde{\phi}_{i-1}(\cdot),\ \tilde{\phi}_{i+1}(\cdot), \ldots, \tilde{\phi}_d(\cdot)\right) \in C_{0,d-1}[0,T]$, the set

$$A_i = \left\{\phi(\cdot) : \left\| \left(\tilde{\phi}_1(\cdot), \ldots, \tilde{\phi}_{i-1}(\cdot),\ \phi(\cdot),\ \tilde{\phi}_{i+1}(\cdot), \ldots, \tilde{\phi}_d(\cdot)\right) \right\| < \varepsilon \right\}$$

is convex and symmetric in $C_{0,1}[0,T]$.

It is easy to check that all L^p norms (i.e. $||\phi||_p = (\sum_{i=1}^d \int_0^T \phi_i^p(t)dt)^{1/p}$), the supremum norm and all Hölder norms are completely convex norms on $C_{0,d}[0,T]$, as are weighted versions of these norms, while rotations do not preserve the complete convexity property. The reason for our interest in completely convex norms lies in the following.

Theorem 1 *Let $\|\cdot\|$ be a completely convex norm. Let $\mathcal{F}_i$ denote the sub-sigma algebra generated by $\{w_1(t), \ldots, w_{i-1}(t),\ w_{i+1}(t), \ldots, w_d(t), \quad 0 \leq t \leq T\}$. Let $\psi(\cdot)$ be an $\mathcal{F}_i$ adapted function such that, for any c,*

$$E(\exp c^2 \int_0^T \psi^2(t)\, dt \mid \|w\| < \varepsilon) \underset{\varepsilon \to 0}{\longrightarrow} 1 \tag{3}$$

Then, for any ε,

$$E(\exp c \int_0^T \psi(t)\, dw_i(t) \mid \|w\| < \varepsilon) \underset{\varepsilon \to 0}{\longrightarrow} 1 \tag{4}$$

Proof: The proof is similar to the proof of Theorem 1 in [12]. Indeed, note that, by the convexity assumption of the norm, for each path $\{w_1(\cdot), \ldots, w_{i-1}(\cdot),\ w_{i+1}(\cdot), \ldots, w_d(\cdot)\}$, the set $A_i \triangleq \{w_i : \|w\| < \varepsilon\}$ is convex and symmetric and, moreover, the random variable $\int_0^T \psi(t)dw_i(t)$ is conditionally (in $\mathcal{F}_i$) Gaussian of zero mean. Thus, by theorem 1 of [5],

$$\begin{aligned} E(\exp c\int_0^T \psi(t)\,dw_i(t) \mid \|w\| < \varepsilon,\, \mathcal{F}_i) &\le E(\exp |c\int_0^T \psi(t)\,dw_i(t)| \mid \|w\| < \varepsilon,\, \mathcal{F}_i) \\ &= E(\exp |c\int_0^T \psi(t)\,dw_i(t)| \mid \mathcal{F}_i,\, w_i \in A_i) \\ \le E(\exp |c\int_0^T \psi(t)\,dw_i(t)| \mid \mathcal{F}_i) &\le 2\exp c^2\int_0^T \psi^2(t)\,dt \end{aligned}$$

Hence, for any c, it follows from (3) that

$$E(\exp c\int_0^T \psi(t)\,dw_i(t) \mid \|w\| < \varepsilon) \le 2\,E(\exp c^2\int_0^T \psi^2(t)\,dt \mid \|w\| < \varepsilon) \underset{\varepsilon \to 0}{\longrightarrow} 2 \qquad (5)$$

Since (5) holds for any c, one has by substituting c/p for c and applying Jensen's inequality that

$$\limsup_{\varepsilon\to 0} E(\exp c\int_0^T \psi(t)\,dw_i(t) \mid \|w\| < \varepsilon) \le 2^{1/p}$$

Taking $p \to \infty$, one concludes that

$$\limsup_{\varepsilon\to 0} E(\exp c\int_0^T \psi(t)\,dw_i(t) \mid \|w\| < \varepsilon) \le 1$$

Since c is arbitrary, the theorem is a consequence of the following elementary lemma, which will also serve us in the sequel. For a proof, see [7], ch. 6.9. □

Lemma 1 *Let $I_1, I_2, \ldots, I_n$ be random variables defined on the Wiener space. If, for all c,*

$$\limsup_{\varepsilon\to 0} E(\exp c\,I_i \mid \|w\| < \varepsilon) \le 1$$

Then

$$\lim_{\varepsilon\to 0} E(\exp \sum_{i=1}^{n} I_i \mid \|w\| < \varepsilon) = 1$$

Remark: In a recent paper [14], H. Sugita shows that for many nonzero constants a (actually, for a in some dense set of constants $A \subset R$), one may find a norm $||\cdot||^{(a)}$ such that

$$\lim_{\epsilon\to 0} P(|\int_0^1 w_1(t)dw_2(t) - \int_0^1 w_2(t)dw_1(t) - a| > a/2|\, ||w||^{(a)} < \epsilon) = 0,$$

and thus concludes that no intrinsic skeleton of Lévy's stochastic area exists. A consequence of Theorem 1 is that if attention is restricted to the family of completely convex norms, a skeleton does exist.

Another useful classical estimate is based on martingales arguments. We bring below a version suitable for our needs here.

Theorem 2 *Let $g(\cdot)$ be an adapted process and c a given constant. Define $\tau \triangleq \int_0^T g^2(t)dt$. Assume that, for each $\delta > 0$, there exists a function $\delta(y,\varepsilon,\delta)$ such that*

$$\frac{\int_\delta^\infty \delta(y,\epsilon,\delta)^{-1/2} e^{cy} e^{-y^2/2\delta(y,\varepsilon,\delta)} dy}{P(\|w\| < \varepsilon)} \underset{\varepsilon \to 0}{\longrightarrow} 0 \tag{6}$$

and

$$\int_\delta^\infty P(\tau > \delta(y,\varepsilon,\delta) \mid \|w\| < \varepsilon)\, e^{cy} dy \underset{\epsilon \to 0}{\longrightarrow} 0 \tag{7}$$

Then,

$$\limsup_{\varepsilon \to 0} E(\exp c \int_0^T g(t)dw_i(t) \mid \|w\| < \varepsilon) \leq 1$$

Proof: Let $\delta > 0$ be given. Then,

$$\limsup_{\varepsilon \to 0} E(\exp c \int_0^T g(t)dw_i(t) \mid \|w\| < \varepsilon) \leq e^{|c\delta|} + \frac{\int_\delta^\infty e^{|c|y} E(\mathbf{1}_{||w||<\epsilon} \mathbf{1}_{\int_0^T g(t)dw_i(t) > y}) dy}{P(||w|| < \epsilon)}. \tag{8}$$

Note that

$$\begin{aligned} E(\mathbf{1}_{||w||<\epsilon} \mathbf{1}_{\int_0^T g(t)dw_i(t)>y}) &= \\ E(\mathbf{1}_{||w||<\epsilon} \mathbf{1}_{\int_0^T g(t)dw_i(t)>y} \mathbf{1}_{\int_0^T g^2(t)dt \leq \delta(y,\epsilon,\delta)}) &+ E(\mathbf{1}_{||w||<\epsilon} \mathbf{1}_{\int_0^T g(t)dw_i(t)>y} \mathbf{1}_{\int_0^T g^2(t)dt > \delta(y,\epsilon,\delta)}) \\ &\triangleq A_1 + A_2 \end{aligned} \tag{9}$$

Recall that $\int_0^T g(t)dw_i(t)$ is distributed like $\tilde{w}_\tau$, where $\tilde{w}_.$ is a Brownian motion. Thus,

$$A_1 \leq P(\sup_{0 \leq t \leq \delta(y,\epsilon,\delta)} \tilde{w}_t > y) = 2P(\tilde{w}_{\delta(y,\epsilon,\delta)} > y) \leq c\delta(y,\varepsilon,\delta)^{-1/2} e^{-y^2/2\delta(y,\epsilon,\delta)}$$

Thus, using (6),

$$\begin{aligned} \frac{\int_\delta^\infty e^{|c|y} E(\mathbf{1}_{||w||<\epsilon} \mathbf{1}_{\int_0^T g(t)dw_i(t)>y} \mathbf{1}_{\int_0^T g^2(t)dt \leq \delta(y,\epsilon,\delta)}) dy}{P(||w|| < \epsilon)} &= \frac{\int_\delta^\infty e^{|c|y} A_1 dy}{P(||w|| < \epsilon)} \\ \leq 2\frac{\int_\delta^\infty e^{|c|y} P(\tilde{w}_{\delta(y,\epsilon,\delta)} > y) dy}{P(||w|| < \epsilon)} &\underset{\epsilon \to 0}{\longrightarrow} 0. \end{aligned} \tag{10}$$

Similarly, using (7),

$$\begin{aligned} \frac{\int_\delta^\infty e^{|c|y} E(\mathbf{1}_{||w||<\epsilon} \mathbf{1}_{\int_0^T g(t)dw_i(t)>y} \mathbf{1}_{\int_0^T g^2(t)dt > \delta(y,\epsilon,\delta)}) dy}{P(||w|| < \epsilon)} &= \frac{\int_\delta^\infty e^{|c|y} A_2 dy}{P(||w|| < \epsilon)} \\ \leq \int_\delta^\infty e^{|c|y} P(\tau > \delta(y,\epsilon,\delta) \mid ||w|| < \epsilon) dy &\underset{\epsilon \to 0}{\longrightarrow} 0. \end{aligned} \tag{11}$$

Combining (8),(9),(10) and (11), and taking first $\epsilon \to 0$ and then $\delta \to 0$, the theorem follows.

$\square$

Remarks

(1) If the assumptions of Theorem 2 hold for some $g(\cdot)$, we say that $||\cdot||$ is $\underline{g(\cdot) \text{ integrable}}$.

(2) In a typical application, $g(t)$ is such that $\sup_{0\le t\le T}\left|\frac{g(t)}{|w(t)|^2}\right| \le C$ for some C. We check that in this situation, many norms satisfy the assumptions of the theorem. Consider first the case of L^p norms, $p \ge 4$. Take $\delta(y,\epsilon,\delta) = y^{1/2}\epsilon^{2+\alpha}$, some appropriate $\alpha(p)$. Then,

$$P(\int_0^T g^2(t)dt > y^{1/2}\epsilon^{2+\alpha(p)}) \mid ||w||_p < \epsilon) \le P(||w||_4 > C^{1/4}y^{1/8}\epsilon^{1/2+\alpha(p)/4} \mid ||w||_p < \epsilon) \underset{\epsilon\to 0}{\longrightarrow} 0 \tag{12}$$

(actually, for $p \ge 4$, the right hand side of (12) is 0 for all ϵ small enough, however (12) may be true also for values of $p < 4$. The computation of (12) in the latter case remains open). In addition to (12), also

$$\int_\delta^\infty e^{cy}P(||w||_4 > C^{1/4}y^{1/8}\epsilon^{1/2+\alpha(p)/4} \mid ||w||_p < \epsilon)dy \underset{\epsilon\to 0}{\longrightarrow} 0 \tag{13}$$

which implies that (7) holds true. To see (6), use the fact that $P(||w||_p < \epsilon) \ge P(||w||_\infty < \epsilon) \sim e^{-k/\epsilon^2}$, and thus

$$\frac{\int_\delta^\infty e^{cy}e^{-y^2/2y^{1/2}\epsilon^{2+\alpha(p)}}dy}{P(\|w\|_p < \varepsilon)} \le \frac{\int_\delta^\infty e^{cy}e^{-y^{3/2}/2\epsilon^{2+\alpha(p)}}dy}{e^{-k/\epsilon^2}} \underset{\varepsilon\to 0}{\longrightarrow} 0$$

We next consider the case of Hölder norms, i.e. the norm $||\phi||_{Lip_\alpha} \triangleq \sup_{0\le t,s\le T}\frac{|\phi(t)-\phi(s)|}{|t-s|^\alpha}$. By applying a result of [4] (see [1] for details), it follows that for all $\alpha < 1/2$,

$$-\epsilon^{2/(1-2\alpha)}\log P(||w||_{Lip_\alpha} < \epsilon) \underset{\epsilon\to 0}{\longrightarrow} k$$

for some constant k. To satisfy (6), one then takes $\delta(y,\epsilon,\delta) = y^{1/2}\epsilon^{2/(1-2\alpha)+\delta'}$, any $\delta' > 0$. (7) is then satisfied as soon as $\alpha < 1/4$, for

$$P(\tau > \delta(y,\epsilon,\delta) \mid ||w||_{Lip_\alpha} < \epsilon) = P(||w||_4 > \delta^{1/2}\epsilon^{1/2(1-2\alpha)+\delta'/4} \mid ||w||_{Lip_\alpha} < \epsilon) = 0$$

for all ϵ small enough, as soon as $2(1-2\alpha) > 1$.

(3) A similar analysis in the case $\sup_{0\le t\le T}\left|\frac{g(t)}{|w(t)|^3}\right| \le C$ reveals that the assumptions of the theorem hold true for Hölder norms $||\cdot||_{Lip_\alpha}$ as soon as $\alpha < 1/3$. Similarly, for every $\alpha < 1/2$ there exists a $\beta \in (1,\infty)$ such that $\|\cdot\|_{\mathrm{Lip}_\alpha}$ is $g(\cdot)$ integrable as soon as $\sup_{0\le t\le T}\left|\frac{g(t)}{|w(t)|^\beta}\right| \le C$.

3 Onsager-Machlup functionals

Let x_t, $0 \le t \le T$ denote the solution of the SDE

$$dx_t = f(x_t)dt + dw_t,\ x_0 = 0,\ x_t \in \mathbb{R}^d \tag{14}$$

(we discuss extensions to the case of nonconstant diffusion coefficients at the end of this section). Here and throughout, $f(\cdot)$ is a smooth function, i.e. bounded with bounded derivatives of arbitrary order. We compute here the Onsager-Machlup limit

$$J(\phi) = \lim_{\epsilon \to 0} \frac{P(||x-\phi)|| < \epsilon)}{P(||w|| < \epsilon)}. \tag{15}$$

Theorem 3 *Let $||\cdot||$ be a completely convex norm such that $P(||w|| < \epsilon) > 0$. Assume that either*

$$\lim_{||\Psi|| \to 0} \int_0^T |\Psi|^2 dt = 0 \tag{16}$$

$$\forall c,\ \limsup_{\epsilon \to 0} E(\exp c|w(T)| \mid ||w|| < \epsilon) \leq 1 \tag{17}$$

$$||\cdot|| \text{ is } |w|^2 \text{ integrable} \tag{18}$$

or

$$\lim_{||\Psi|| \to 0} \int_0^T |\Psi|^3 dt = 0 \tag{19}$$

$$\forall c,\ \limsup_{\epsilon \to 0} E(\exp c|w(T)|^3 \mid ||w|| < \epsilon) \leq 1 \tag{20}$$

$$||\cdot|| \text{ is } |w|^3 \text{ integrable} \tag{21}$$

(for the definition of $|w|^\beta$ integrable, see the last remark of section 2).

Let $\phi \in H_{1,d}[0,T]$, i.e. $\dot{\phi} \in L_{2,d}[0,T]$ and $\phi_0 = 0$. Then,

$$J(\phi) = \exp -\frac{1}{2}\left(\int_0^T |\dot{\phi}(t) - f(\phi(t))|^2 dt + \int_0^T \nabla f(\phi(t)) dt\right). \tag{22}$$

Proof: The proof follows closely the change of measure argument of [7], except that here the exponential estimates of section 2 are used. Let $z(t) = x(t) - \phi(t)$, then

$$dz(t) = \left(f(z(t) + \phi(t)) - \dot{\phi}(t)\right) dt + dw_t,\ z(0) = 0 \tag{23}$$

An application of Girsanov's theorem yields that

$$\begin{aligned} &\frac{P(||x-\phi|| < \epsilon)}{P(||w|| < \epsilon)} = \frac{P(||z|| < \epsilon)}{P(||w|| < \epsilon)} = \\ &E\left(\exp \int_0^T \left(f(w(t) + \phi(t)) - \dot{\phi}(t)\right)^* dw(t)\right. \\ &\left. -\frac{1}{2}\int_0^T |f(w(t) + \phi(t)) - \dot{\phi}(t)|^2 dt \mid ||w|| < \epsilon\right). \end{aligned} \tag{24}$$

Since $f(\cdot)$ is smooth, a Taylor expansion of it and either (16) or (19) reveal that, uniformly in ω,

$$\int_0^T |f(w(t) + \phi(t)) - \dot{\phi}(t)|^2 dt \underset{||w|| \to 0}{\longrightarrow} \int_0^T |f(\phi(t)) - \dot{\phi}(t)|^2 dt. \tag{25}$$

Next, using either (17) or (20), one may apply theorem 1 of [12] to conclude that, for any c,

$$E(\exp c\int_0^T \dot{\phi}(t)^* dw(t) \mid ||w|| < \epsilon) \underset{\epsilon\to 0}{\longrightarrow} 1 \tag{26}$$

(note that theorem 1 of [12] is based on the same correlation argument which yields theorem 1 in this paper). Finally, by a Taylor expansion of $f(\cdot)$, one obtains that

$$\begin{aligned}
\int_0^T f^*(w(t)+\phi(t))dw(t) &= \\
\int_0^T f^*(\phi(t))dw(t) &+ \sum_{i,j=1}^{d}\int_0^T \frac{\partial f_i}{\partial x_j}(\phi(t))w_j(t)dw_i(t) + \int_0^T g_1^*(t)dw(t) = \\
\int_0^T f^*(\phi(t))dw(t) &+ \sum_{i,j=1}^{d}\int_0^T \frac{\partial f_i}{\partial x_j}(\phi(t))w_j(t)dw_i(t) + \sum_{i,j,k=1}^{d}\int_0^T \frac{\partial^2 f_i}{\partial x_j \partial x_k}(\phi(t))w_j(t)w_k(t)dw_i(t) \\
&+ \int_0^T g_2^*(t)dw(t)
\end{aligned} \tag{27}$$

where $g_i(t)$, $i=1,2$ are $\mathbb{R}^d$ valued adapted processes such that, for some deterministic α,

$$\sup_{0\le t\le T}\left|\frac{g_1(t)}{|w(t)|^2}\right| \le \alpha, \ \sup_{0\le t\le T}\left|\frac{g_2(t)}{|w(t)|^3}\right| \le \alpha.$$

Using integration by parts,

$$\int_0^T f^*(\phi(t))dw(t) = \sum_{i=1}^{d} w_i(T)f_i(\phi(T)) - \sum_{i,j=1}^{d}\int_0^T w_i(t)\frac{\partial f_i}{\partial x_j}(\phi(t))\dot{\phi}_j(t)dt.$$

Thus, by either (16) and (17) or (19) and (20), one concludes that for any c,

$$E(\exp c\int_0^T f^*(\phi(t))dw(t) \mid ||w|| < \epsilon) \underset{\epsilon\to 0}{\longrightarrow} 1. \tag{28}$$

Next, using either (18) (for $i=1$) or (21) (for $i=2$) and applying theorem 2, one concludes that, for every c,

$$E(\exp c\int_0^T g_i(t)dw(t) \mid ||w|| < \epsilon) \underset{\epsilon\to 0}{\longrightarrow} 1. \tag{29}$$

On the other hand, again integrating by parts,

$$\begin{aligned}
E\left(\exp c\left(\sum_{i=1}^{d}\int_0^T \frac{\partial f_i}{\partial x_i}(\phi(t))w_i(t)dw_i(t) + \frac{1}{2}\int_0^T \nabla f(\phi(t))dt\right) \mid ||w|| < \epsilon\right) &= \quad (30)\\
E\left(\exp c\sum_{i=1}^{d}\frac{1}{2}\int_0^T \frac{\partial f_i}{\partial x_i}(\phi(t))dw_i^2(t)) \mid ||w|| < \epsilon\right) &= \\
E\left(\exp(\frac{c}{2}\sum_{i=1}^{d} w_i^2(T)\frac{\partial f_i}{\partial x_i}(\phi(T)) - \frac{c}{2}\sum_{i=1}^{d}\int_0^T w_i^2(t)\frac{\partial}{\partial t}\left(\frac{\partial f_i}{\partial x_i}(\phi(t))\right)dt) \mid ||w|| < \epsilon\right) &\underset{\epsilon\to 0}{\longrightarrow} 1.
\end{aligned}$$

where either (16) and (17) or (19) and (20) were used in the last computation. By theorem 1 and the same assumptions,

$$E\left(\exp c\sum_{i,j=1,i\neq j}^{d}\int_0^T \frac{\partial f_i}{\partial x_j}(\phi(t))w_j(t)dw_i(t) \mid ||w|| < \epsilon\right) \underset{\epsilon\to 0}{\longrightarrow} 1. \tag{31}$$

Combining (27),(28),(29),(30),(31), one obtains (23) under assumptions (16–18). To get the result under (19–21), it remains to check that for any i, j, k,

$$\limsup_{\epsilon\to 0} E(\exp c \int_0^T \frac{\partial^2 f_i}{\partial x_j \partial x_k}(\phi(t)) w_j(t) w_k(t) dw_i(t) \mid ||w|| < \epsilon) \leq 1 \tag{32}$$

If $i \neq j$ and $i \neq k$, (32) follows from theorem 1 and (19). Thus, it is enough to consider the case $k = i$. If $j \neq i$, then

$$\begin{aligned}
\int_0^T \frac{\partial^2 f_i}{\partial x_j \partial x_i}(\phi(t)) w_j(t) w_i(t) dw_i(t) &= \\
\frac{1}{2}\int_0^T \frac{\partial^2 f_i}{\partial x_j \partial x_i}(\phi(t)) w_j(t) dw_i^2(t) - \frac{1}{2}\int_0^T \frac{\partial^2 f_i}{\partial x_j \partial x_i}(\phi(t)) w_j(t) dt &= \\
\frac{1}{2}\frac{\partial^2 f_i}{\partial x_j \partial x_i}(\phi(T)) w_j(T) w_i^2(T) - \frac{1}{2}\int_0^T \frac{\partial^2 f_i}{\partial x_j \partial x_i}(\phi(t)) w_i^2(t) dw_j(t) &- \\
\frac{1}{2}\int_0^T \frac{\partial}{\partial t}\left(\frac{\partial^2 f_i}{\partial x_j \partial x_i}(\phi(t))\right) w_i^2(t) w_j(t) dt - \frac{1}{2}\int_0^T \frac{\partial^2 f_i}{\partial x_j \partial x_i}(\phi(t)) w_j(t) dt &
\end{aligned}$$

Thus, for this case,

$$\limsup_{\epsilon\to 0} E(\exp c \int_0^T \frac{\partial^2 f_i}{\partial x_j \partial x_i}(\phi(t)) w_j(t) w_i(t) dw_i(t) \mid ||w|| < \epsilon) \leq 1.$$

Finally, in the case $i = j = k$, it follows form Ito's lemma that

$$\begin{aligned}
\int_0^T \frac{\partial^2 f_i}{\partial x_i^2}(\phi(t)) w_i^2(t) dw_i(t) &= \frac{1}{3}\int_0^T \frac{\partial^2 f_i}{\partial x_i^2}(\phi(t)) dw_i^3(t) - \frac{1}{3}\int_0^T \frac{\partial^2 f_i}{\partial x_i^2}(\phi(t)) w_i(t) dt \\
&= -\frac{1}{3}\int_0^T \frac{\partial}{\partial t}\left(\frac{\partial^2 f_i}{\partial x_i^2}(\phi(t))\right) w_i^3(t) dt - \frac{1}{3}\int_0^T \frac{\partial^2 f_i}{\partial x_i^2}(\phi(t)) w_i(t) dt \\
&+ \frac{1}{3}\frac{\partial^2 f_i}{\partial x_i^2}(\phi(T)) w_i^3(T)
\end{aligned}$$

Thus, (32) follows in the case $i = j = k$ from (19) and (20). $\square$

Corollary 1 *The Onsager Machlup limit (22) holds true for all L^p norms, $p \geq 4$, and all Hölder norms $||\cdot||_{Lip_\alpha}$, $\alpha < 1/3$.*

Proof: Note first that in view of the remarks following theorem 2 and the fact that all Hölder norms dominate the supremum norm, there is nothing to prove in the case $||\cdot||_{Lip_\alpha}$, $\alpha < 1/3$. Considering the L^p norm case, $p \geq 4$, in view of the same remarks the only thing to check is (17). The latter is obvious for the supremum norm, follows in the case $p = 2$ from corollary 1 in [8], and follows in the general case by writing $w(T) = \int_0^T dw_t = \int_0^T \phi^\delta(t) dw_t + \int_0^T (1 - \phi^\delta(t)) dw_t$, with $\phi^\delta \in C_{1,d}[0,T]$, $\phi^\delta(T) = 0$, and $||1 - \phi^\delta(t)||_2 \to_{\delta\to 0} 0$. Integrating by parts, one has

$$|w(T)| = |\int_0^T (1 - \phi^\delta(t)) dw_t - \int_0^T \dot\phi^\delta(t) w_t dt| \leq c_\delta ||w||_2 + |\int_0^T (1 - \phi^\delta(t)) dw_t|,$$

with $c_\delta \to_{\delta\to 0} 0$. Applying now again Theorem 1 of [5] in exactly the same way as in the proof of Theorem 1, (17) follows. $\Box$

Remarks

(1) In the one dimensional ($d = 1$) case, the results of theorem 3 extend to all Hölder norms $||\cdot||_{Lip_\alpha}$ with $\alpha < 1/2$. This is proved as follows: from the remark following theorem 2, one knows that $||\cdot||_{Lip_\alpha}$, $\alpha < 1/2$ is $|w|^\beta$ integrable for large enough β. Using a Taylor series up to order β in (27) and repeating the proof of theorem 3, the conclusion follows if it can be shown that, for a one dimensional Brownian motion $w(t)$ and any deterministic $\psi \in H_1[0,T]$,

$$\limsup_{\epsilon\to 0} E(\exp c\int_0^T \psi(t)w^\beta(t)dw(t) \mid ||w||_{Lip_\alpha} < \epsilon) < 1.$$

By Ito's lemma,

$$\int_0^T \psi(t)w^\beta(t)dw(t) = \frac{w^{\beta+1}(T)\psi(T)}{\beta+1} - \int_0^T \frac{\psi(t)w^{\beta-1}(t)dt}{2\beta} - \int_0^T \frac{w^{\beta+1}(t)}{\beta+1}\dot{\psi}(t)dt \underset{||w||_{Lip_\alpha}\to 0}{\longrightarrow} 0$$

uniformly in ω, where the last limit follows from the fact that the Hölder norm dominates the supremum norm. The difficulty in extending this argument to the multidimensional case is that the integration by parts yields again stochastic integrals, which can not be handled by either the correlation based bounds or the martingale argument exploited in theorem 2.

(2) The corollary actually holds also for the L^2 norm. However, in that case, the computations following the proof of theorem 2 do not hold, and one has to use different correlation inequalities. An example of such a computation may be found in [8].

(3) In the case of nonconstant diffusion coefficients, the form of the Onsager Machlup limit depends on the norm used, and is known in the case where the norm is related to the Riemannian metric defined by $\sigma(\cdot)\sigma^*(\cdot)$ (c.f. [15],[6]). Since dealing with this problem requires the introduction of differential geometric considerations, we do not pursue it here.

References

[1] Baldi, P., Ben Arous, G. and Kerkyaacharian, G., "Large deviations and Strassen theorem in Hölder norm". Preprint, 1991.

[2] Carmona, R. and Nualart, D., "Traces of random variables on the Wiener space and the Onsager Machlup functional", preprint (1991).

[3] Chaleyat-Maurel, M. and Nualart, D., "The Onsager Machlup functional for a class of anticipating processes", preprint (1991).

[4] Ciesielski, Z., "On the isomorphism of the spaces H_α and m", *Bull. de L'Academie Pol. des Sciences*, Serie des sci. math., astr., phys., VIII(1960), pp. 217-222.

[5] Das Gupta S., Eaton M.L., Olkin I., Perlman M., Savage L.J. and Sobel M., "Inequalities on the probability content of convex regions for elliptically contoured distributions", Sixth Berkeley symposium on Mathematical Statistics and Probability, vol II., 1972, pp. 241-265.

[6] Fujita, J. and Kotani, S., "The Onsager Machlup function for diffusion processes", *J. Math. Kyoto Univ.*, 22 (1982), pp. 115-130.

[7] Ikeda, N. and Watanabe, S., Stochastic Differential Equations and Diffusion Processes, 2nd edition, North-Holland, 1989.

[8] Mayer-Wolf, E. and Zeitouni, O., "The probability of small Gaussian ellipsoids and associated exponential moments", to appear, *Annals of Probability* (1992).

[9] Mayer-Wolf, E. and Zeitouni, O., "Onsager-Machlup functionals for non trace class SPDEs", to appear, *PTRF* (1992).

[10] Millet, A. and Nualart, D., "Theoreme de support pour une classe d'equations differentielles stochastiques anticipantes", C.R. Acad. Sci. Paris, t. 312, Serie I (1991), pp. 743-746.

[11] Onsager, L. and Machlup, S., "Fluctuations and irreversible processes", I,II, *Phys. Rev.*, 91 (1953), pp. 1505-1515.

[12] Shepp, L.A. and Zeitouni, O., "A note on conditional exponential moments and the Onsager Machlup functional", *Annals of Probability*, 20 (1992), pp. 652–654.

[13] Stratonovich, R.L., "On the probability functional of diffusion processes", *Selected Trans. in Math. Stat. Prob.* , 10 (1971), pp. 273-286.

[14] Sugita, H., "Various topologies in the Wiener space and Lévy's stochastic area", *PTRF* 91 (1992), pp. 283–296.

[15] Takahashi, Y. and Watanabe, S., "The probability functional (Onsager–Machlup function) of diffusion processes", *Stochastic Integrals* (D. Williams, Ed.), *Lecture Notes in Math.*, 851 (1981), pp. 433–463. Springer, Berlin – New-York.

[16] Wong, E. and Zakai, M., “On the relation between ordinary and stochastic differential equations”, Int. Jour. Eng. Sc., 3 (1965), pp. 213-229.