Exponential Stability in Discrete Time Filtering for Non-Ergodic Signals

A. Budhiraja*
Department of Mathematics
University of Notre Dame
Notre Dame, IN 46656
budhiraja.1@nd.edu

D. Ocone
Department of Mathematics
Rutgers University
110 Frelinghuysen Rd.
New Brunswick, NJ 08854-8019
ocone@math.rutgers.edu

December 11, 2000

Abstract

In this paper we prove exponential asymptotic stability for discrete time filters for signals arising as solutions of d-dimensional stochastic difference equations. The observation process is the signal corrupted by an additive white noise of sufficiently small variance. The model for the signal admits non-ergodic processes. We show that almost surely, the total variation distance between the optimal filter and an incorrectly initialized filter converges to 0 exponentially fast as time approaches ∞.

Key Words: nonlinear filtering, asymptotic stability, measure valued processes.

*Research suppored by the NSF grant DMI 9812857.

1 Introduction

The central problem of nonlinear filtering is to study the conditional distribution of a signal process at any time instant given noisy observations on the signal available up until that time. If the signal–observation pair is Markov, the conditional distribution process, referred to hereafter as the *optimal filter*, is determined completely by the observation process, the transitional probability function of the pair, and its initial distribution. In practice, the model parameters, i.e. the initial distribution and the transition probability function, are rarely known exactly, and so one constructs sub-optimal filters by replacing the unknown parameters with suitable approximations. Thus it is of interest to study the sensitivity of the filter to errors in the model parameters, especially over large time intervals. The simplest problem assumes that the transition probability function is known exactly, so that the only error in the filter comes from use of the wrong initial distribution. A filter computed with the wrong initial distribution is called an *incorrectly initialized filter* and is suboptimal. We say that the filter is asymptotically stable if the distance (appropriately measured) between the optimal filter and the incorrectly initialized filter converges to 0 as time approaches ∞. Thus if the filter is asymptotically stable the errors in the initial conditions do not significantly influence the long term performance of the filter.

In recent years there has been significant progress in the study of asymptotic stability of filters for models in which a Markov signal is observed in independent, additive, usually Gaussian, white noise. The best general results have been obtained in the case of signal dynamics which admit ergodic solutions. Here, the pioneering paper is that of Kunita[10] who showed that if the signal is Feller-Markov with a compact state space then the filter process is also Feller-Markov, and, furthermore, if the signal admits a unique invariant measure so also does the filter, provided appropriate technical conditions are satisfied. This result was extended to locally compact state spaces in Stettner[16] and Kunita[11]. These papers suggest that under appropriate conditions the ergodicity of the signal should lead to the asymptotic stability of filters. Indeed, Ocone and Pardoux[15] showed that if the Kunita-Stettner conditions on the signal are satisfied and the signal "forgets its initial conditions" then so does the filter. The connection between ergodic signals and asymptotic stability of filters has been greatly clarified by a recent series of papers: Delyon and Zeitouni[9], Atar and Zeitouni[4, 3, 2], Atar[1], LeGland and Mevel[12], Budhiraja and Kushner [5], and Malliavin, Da Prato, and Fuhrmann[8].

Conditions under which asymptotic stability holds in absence of signal ergodicity is an interesting and challenging problem. The known results support the intuition that even for non-ergodic signals "sufficiently good" observations should exert a correcting influence on an incorrectly initialized filter. For Kalman filters, it is a classical result that asymptotic stability does not require signal ergodicity, but only detectability and observability assumprions on the signal–observation system. Asymptotic stability for scalar Beneš filtering models, whose signal processes are generically transient. is established in Ocone[13].

In Budhiraja and Ocone[6] it is shown that for one dimensional stochastic difference equations which are observed in bounded observation noise, asymptotic stability holds in the sense that the total variation distance between the optimal and an incorrectly initialized filter converges to 0 exponentially fast as $t \to \infty$, under appropriate smoothness of the signal process coefficients. General information inequalities relating optimal and incorrectly initialized filters, are given in Ocone[14] and Clark, Ocone and Coumarbatch[7].

The object of this paper is to prove an asymptotic stability result for discrete-time systems in which the assumption of bounded noise made in [6] is dropped, yet the signal is allowed to be non-ergodic. The main result, stated precisely in Theorem 2.4, establishes exponential asymptotic stability in the total variation norm of the corresponding filters. In order to clearly bring out the key points we first study the case where the observation noise and the noise in the signal dynamics are both Gaussian. This is done in Theorem 2.1. The method of proof is quite different from the proof of the bounded observation case studied in[6], which, like [3], used Hilbert's projective metric. This metric is not useful for studying measures on a non-compact space, as required in the case of noise with unbounded support. Rather, in this paper, we start from the crucial observation, made in [3], that if $\rho^{(0)}, \rho^{(1)}$ are nonnegative integrable functions on $I\!R^d$ and $p^{(i)}(x) \doteq \int_{I\!R^d} \rho^{(i)}(x)dx; i = 0, 1$ then

$$||p^{(0)} - p^{(1)}||_1 \leq \frac{||\rho^{(0)} \wedge \rho^{(1)}||_1}{||\rho^{(0)}||_1 ||\rho^{(1)}||_1},$$

where $\rho^{(0)} \wedge \rho^{(1)}$ is an element of $L^1(I\!R^{2d})$ defined as

$$\rho^{(0)} \wedge \rho^{(1)}(x, y) \doteq \rho^{(0)}(x)\rho^{(1)}(y) - \rho^{(0)}(y)\rho^{(1)}(x)$$

and we have denoted the natural norm on $L^1(I\!R^d)$ and $L^1(I\!R^{2d})$ by the same symbol: $|| \cdot ||_1$. This inequality is used in the proof of the theorem with $\rho^{(0)} \equiv \rho_n^{(0)}$ and $\rho^{(1)} \equiv \rho_n^{(1)}$ where $\rho_n^{(0)}$ and $\rho_n^{(1)}$ are the unnormalized filtering densities corresponding to the optimal and the incorrectly initialized filter respectively. The advantage of considering these unnormalized densities is that the map $\rho_{n-1}^{(0)} \wedge \rho_{n-1}^{(1)} \to \rho_n^{(0)} \wedge \rho_n^{(1)}$ is a linear operator on $L^1(I\!R^{2d})$ and Lemma 2.3 shows that almost everywhere this operator is a strict contraction for a small enough observation noise variance: σ^2. Furthermore as $\sigma \to 0$ the contraction coefficient approaches 0. This shows that $||\rho_n^{(0)} \wedge \rho_n^{(1)}||_1$ decays exponentially fast with an arbitrarily large exponential rate if σ is appropriately small. The remaining work is to show that $\frac{1}{||\rho_n^{(0)}||_1 ||\rho_n^{(1)}||_1}$ grows at an at most exponential rate which is bounded as σ approaches 0. This is done in Lemma 2.2. Lemmas 2.2 and 2.3 are stated and proved for Gaussian noises but as is seen in Theorem 2.4 they hold more generally. Theorem 2.1, follows as a direct consequence of Lemmas 2.2 and 2.3. Finally in Theorem 2.4 we consider the case of non-Gaussian noises.

2 The main result:

The filtering model that we consider is as follows. Let $(\Omega, \mathcal{F}, P)$ be some probability space. The signal and the observation processes are given as follows.

$$X_n = a(X_{n-1}) + b(X_{n-1})\xi_n; \;\; n \geq 1,$$

and

$$Y_n = X_n + \sigma\nu_n; \;\; n \geq 1,$$

where X_0 is a $\mathbb{R}^d$ valued random variable with distribution μ_0, $a : \mathbb{R}^d \to \mathbb{R}^d$ and $b : \mathbb{R}^d \to \mathbb{R}^{d\times d}$ are measurable maps satisfying conditions (A1) and (A2) below.

(A1) There exists a finite positive constant a_{lip} such that for all $x, y \in \mathbb{R}^d$

$$|a(x) - a(y)| \leq a_{\text{lip}}|x - y|.$$

(A2) There exist finite positive constant $0 < \underline{b} < \bar{b} < \infty$ such that $\forall u \in \mathbb{R}^d$,

$$\underline{b}|u|^2 \leq \langle u, b(x)u\rangle \leq \bar{b}|u|^2,$$

where $\langle\cdot,\cdot\rangle$ denotes the usual inner product in $\mathbb{R}^d$.

We assume that $\{\xi_n\}_{n\geq 1}$ is a sequence of i.i.d random variables with probability density function q, $\{\nu_n\}_{n\geq 1}$ is another sequence of i.i.d random variables, which is independent of $\{X_0, \{\xi_n; n \geq 1\}\}$, with density r and $\sigma > 0$ is a fixed constant. For notational simplicity denote $\frac{1}{\det(b(x))}\phi\left(b^{-1}(x)(y - a(x))\right)$ by $G(x, y)$. Note that $\{X_n\}_{n\geq 1}$ is a Markov chain with initial distribution μ_0 and transition probability density $G(x, y)$.

Henceforth if a measure on $\mathbb{R}^d$ admits a density with respect to the Lebesgue measure, we will denote the density by the same symbol as the measure. The optimal nonlinear filter is obtained as follows. Define a sequence of finite measures $\{\rho_n^{(0)}\}_{n\geq 1}$ recursively as follows.

$$\begin{aligned} \rho_n^{(0)}(x) &\doteq \frac{1}{\sigma^d} r(\frac{1}{\sigma}(Y_n - x)) \int_{\mathbb{R}^d} \rho_{(n-1)}^{(0)}(y) G(y, x) dy; \;\; x \in \mathbb{R}^d \\ \rho_0^{(0)} &\doteq \mu_0. \end{aligned} \tag{2.1}$$

Finally let $p_n^{(0)}(x) \doteq \frac{\rho_n^{(0)}(x)}{||\rho_n^{(0)}(\cdot)||_1}$, $n \geq 1$ and $p_0^{(0)} \doteq \mu_0$, where for an integrable function g we denote $\int_{\mathbb{R}^d} |g(x)|dx$ by $||g||_1$. The function $p_n^{(0)}$ is the optimal filter, i.e. it is the conditional density of X_n given $Y_1, \cdots, Y_n$.

Now let μ_1 be an arbitrary probability measure on $\mathbb{R}^d$. Define $\rho_n^{(1)}(x)$ and $p_n^{(1)}(x)$ in a similar manner as above by replacing μ_0 with μ_1. Define for $f, g \in L^1(\mathbb{R}^d)$, $f \wedge g \in L^1(\mathbb{R}^{2d})$ as

$$(f \wedge g)(x, y) \doteq f(x)g(y) - f(y)g(x); \;\; x, y \in \mathbb{R}^d.$$

Denote the natural norm on $L^1(I\!R^{2d})$ by $||\cdot||_1$ as well. A straightforward calculation shows (cf. [3]) that

$$||p_n^{(0)} - p_n^{(1)}||_1 \leq \frac{||\rho_n^{(0)} \wedge \rho_n^{(1)}||_1}{||\rho_n^{(0)}||_1 ||\rho_n^{(1)}||_1}. \tag{2.2}$$

We will now show that $||p_n^{(0)} - p_n^{(1)}||_1$ converges to 0 exponentialy fast for sufficiently small σ if q and r satisfy appropriate conditions. Before presenting the general result we will consider the case where $q = r = \phi$ where ϕ denotes the d-dimensional standard normal density. This case contains all the ideas required to prove the general case and is notationally simpler to state.

Theorem 2.1 *Assume that $r = q = \phi$ and $\int_{I\!R^d} |a(z)|^2 \mu_i(dz) < \infty$; $i = 0, 1$. There exists $0 < \sigma_0 < \infty$ such that for all $\sigma < \sigma_0$;*

$$\limsup_{n\to\infty} \frac{1}{n} \log ||p_n^{(0)} - p_n^{(1)}||_1 < 0, \tag{2.3}$$

a.s. P.

Before proceeding to the proof of the theorem we will present two lemmas first of which considers the denominator and the second the numerator of (2.2) .

Lemma 2.2 *Assume that $q = r = \phi$, then there is a finite positive constant K depending only on $a_{lip}, d, \underline{b}, \overline{b}$ such that*

$$\limsup_{n\to\infty} \sup_{0<\sigma<1} \left(-\frac{1}{n} \log ||\rho_n^{(i)}(\cdot)||_1 \right) \leq K \ \ a.s., \tag{2.4}$$

for $i = 0, 1$.

Remark *The constant K in the statement of the lemma is non-random and is obtained by applying the strong law of large numbers to the i.i.d sequences $\{\xi_n\}, \{\nu_n\}$, as will be seen in the proof of the lemma.*

Proof of the lemma: For $i = 0, 1$,

$$\rho_n^{(i)}(x_n) = \int_{(I\!R^d)^n} \prod_{j=1}^n \left(\frac{1}{\sigma^d} \phi(\frac{1}{\sigma}(Y_j - x_j)) G(x_{j-1}, x_j) \right) \mu_i(dx_0) dx_1 \cdots dx_{n-1}.$$

Hence,

$$||\rho_n^{(i)}(\cdot)||_1 = \int_{(I\!R^d)^{n+1}} \prod_{j=1}^n \left(\frac{1}{\sigma^d} \phi(\frac{1}{\sigma}(Y_j - x_j)) G(x_{j-1}, x_j) \right) \mu_i(dx_0) dx_1 \cdots dx_n.$$

Substituting $x_j = Y_j + z_j, j = 0, \cdots, n$, where $Y_0 \doteq 0$, in the above equation we have

$$||\rho_n^{(i)}(\cdot)||_1 = \int_{(I\!R^d)^{n+1}} \prod_{j=1}^n \left(\frac{1}{\sigma^d} \phi(\frac{1}{\sigma} z_j) G(Y_{j-1} + z_{j-1}, Y_j + z_j) \right) \mu_i(dz_0) dz_1 \cdots dz_n.$$

Applying Jensen's inequality to the function $x \to \log(x)$ with respect to the probability measure: $\prod_{j=1}^{n} \frac{1}{\sigma^d} \phi(\frac{1}{\sigma} z_j) \mu_i(dz_0) dz_1 \cdots dz_n$ we have that

$$
\begin{aligned}
\log(||\rho_n^{(i)}(\cdot)||_1) &\geq \frac{1}{\sigma^{nd}} \int_{(\mathbb{R}^d)^{n+1}} \log \left(\prod_{j=1}^{n} G(Y_{j-1}+z_{j-1}, Y_j+z_j) \right) \left(\prod_{j=1}^{n} \phi(\frac{1}{\sigma} z_j) dz_j \right) \mu_i(dz_0) \\
&= \sum_{j=1}^{n} \frac{1}{\sigma^{nd}} \int_{(\mathbb{R}^d)^{n+1}} \log \left(G(Y_{j-1}+z_{j-1}, Y_j+z_j) \right) \left(\prod_{j=1}^{n} \phi(\frac{1}{\sigma} z_j) dz_j \right) \mu_i(dz_0) \\
&= \sum_{j=2}^{n} \frac{1}{\sigma^{2d}} \int_{(\mathbb{R}^d)^{2}} \log \left(G(Y_{j-1}+z_{j-1}, Y_j+z_j) \right) \phi(\frac{1}{\sigma} z_{j-1}) \phi(\frac{1}{\sigma} z_j) dz_{j-1} dz_j \\
&+ \frac{1}{\sigma^{d}} \int_{(\mathbb{R}^d)^{2}} \log \left(G(z_0, Y_1+z_1) \right) \phi(\frac{1}{\sigma} z_1) dz_1 \mu_i(dz_0). \qquad (2.5)
\end{aligned}
$$

In view of (A2) we have that there exist finite positive constants $\underline{B}$ and $\overline{B}$ such that

$$\underline{B} \leq \min(\det b(x), |b(x)|) \leq \max(\det b(x), |b(x)|) \leq \overline{B}.$$

Observe next that for $x, y \in \mathbb{R}^d$,

$$
\begin{aligned}
G(x,y) &= \frac{1}{\det(b(x))} \phi(b^{-1}(x)(y - a(x))) \\
&\geq \frac{1}{(2\pi)^{d/2} \overline{B}} \exp\left(- \frac{|y-a(x)|^2}{\underline{B}^2} \right).
\end{aligned}
$$

Hence

$$\log G(x,y) \geq - \frac{|y-a(x)|^2}{(\underline{B})^2} - \frac{d}{2} \log(2\pi) - \log(\overline{B}) \qquad (2.6)$$

Next note that for $j = 2, \cdots, n$

$$
\begin{aligned}
|Y_j + z_j - a(Y_{j-1}+z_{j-1})| &= |X_j + \sigma\nu_j + z_j - a(Y_{j-1}+z_{j-1})| \\
&= |a(X_{j-1}) + b(X_{j-1})\xi_j + \sigma\nu_j + z_j - a(Y_{j-1}+z_{j-1})| \\
&\leq \overline{B}|\xi_j| + a_{\text{lip}}\sigma|\nu_{j-1}| + \sigma\nu_j + |z_j| + a_{\text{lip}}|z_{j-1}|. \qquad (2.7)
\end{aligned}
$$

Combining the inequalities (2.6), (2.7) we have that

$$
\begin{aligned}
\log G(Y_{j-1}+z_{j-1}, Y_j+z_j) &\geq -2\frac{\overline{B}^2}{\underline{B}^2}|\xi_j|^2 - 2\frac{a_{\text{lip}}^2 \sigma^2 |\nu_{j-1}|^2}{\underline{B}^2} - 2\frac{\sigma^2|\nu_j|^2}{\underline{B}^2} - 2\frac{|z_j|^2}{\underline{B}^2} \\
&- 2\frac{a_{\text{lip}}^2 |z_{j-1}|^2}{\underline{B}^2} - \frac{d}{2}\log(2\pi) - \log(\overline{B}) \qquad (2.8)
\end{aligned}
$$

In a similar fashion we have the inequality

$$\log G(z_0, Y_1+z_1) \geq -2\frac{\overline{B}^2}{\underline{B}^2}|\xi_1|^2 - 2\frac{\sigma^2|\nu_1|^2}{\underline{B}^2} - 2\frac{|z_1|^2}{\underline{B}^2} - 2\frac{|a(X_0)|^2}{\underline{B}^2} - 2\frac{|a(z_0)|^2}{\underline{B}^2} - \frac{d}{2}\log(2\pi) - \log(\overline{B}) \qquad (2.9)$$

Using the inequalities (2.8), (2.9) in (2.5) and observing that $\frac{1}{\sigma^d}\int_{I\!R^d}|z_j|^2\phi(\frac{1}{\sigma}z_j)dz_j = d\sigma^2$ we have

$$\begin{aligned}
\log||\rho_n^{(i)}(\cdot)||_1 &\geq -2\frac{\overline{B}^2}{\underline{B}^2}\sum_{j=1}^n|\xi_j|^2 - 2\frac{(a_{\text{lip}}^2+1)\sigma^2}{\underline{B}^2}\sum_{j=1}^n|\nu_j|^2 - 2\frac{d\sigma^2 n}{\underline{B}^2} - 2\frac{a_{\text{lip}}^2 d\sigma^2(n-1)}{\underline{B}^2} \\
&- 2\frac{|a(X_0)|^2}{\underline{B}^2} - \frac{2}{\underline{B}^2}\int_{I\!R^d}|a(z_0)|^2\mu_i(dz_0) - \frac{nd}{2}\log(2\pi) - n\log(\overline{B})
\end{aligned}$$

Finally an application of the strong law of large number gives that for $i = 0, 1$,

$$\limsup_{n\to\infty}\sup_{0<\sigma<1} -\frac{1}{n}\log||\rho_n^{(i)}(\cdot)||_1 \leq K,$$

where

$$K \doteq 2\frac{\overline{B}^2}{\underline{B}^2} + 2\frac{(a_{\text{lip}}^2+1)}{\underline{B}^2} + 2\frac{d}{\underline{B}^2} + \frac{d}{2}\log(2\pi) + \log(\overline{B}).$$

■

We now consider the numerator in (2.2) in the following lemma.

Lemma 2.3 *Assume that $q = r = \phi$. Then*

$$\lim_{\sigma\to 0}\limsup_{n\to\infty}\frac{1}{n}\log||\rho_n^{(0)}\wedge\rho_n^{(1)}||_1 = -\infty.$$

Proof: Observe that

$$\begin{aligned}
\rho_n^{(0)}\wedge\rho_n^{(1)}(x,y) &= \rho_n^{(0)}(x)\rho_n^{(1)}(y) - \rho_n^{(1)}(x)\rho_n^{(0)}(y) \\
&= \frac{1}{\sigma^{2d}}\phi(\frac{1}{\sigma}(Y_n - x))\phi(\frac{1}{\sigma}(Y_n - y)) \\
&\quad \int_{I\!R^{2d}}\left[\rho_{n-1}^{(0)}(u)G(u,x)\rho_{n-1}^{(1)}(v)G(v,y) - \rho_{n-1}^{(0)}(v)G(v,y)\rho_{n-1}^{(1)}(u)G(u,x)\right]dudv \\
&= \frac{1}{\sigma^{2d}}\phi(\frac{1}{\sigma}(Y_n - x))\phi(\frac{1}{\sigma}(Y_n - y))\int_{I\!R^{2d}}\rho_{n-1}^{(0)}\wedge\rho_{n-1}^{(1)}(u,v)G(u,x)G(v,y)dudv,
\end{aligned} \tag{2.10}$$

where the second equality follows on using (2.1). Now denote (suppressing x and y in the notation) $\rho_{n-1}^{(0)}\wedge\rho_{n-1}^{(1)}(u,v)G(u,x)G(v,y)$ by $M(u,v)$. Also write the vectors u and v as $(u_1, u^{(1)})$ and $(v_1, v^{(1)})$ respectively, where $u_1, v_1 \in I\!R$ and $u^{(1)}, v^{(1)} \in I\!R^{d-1}$. It is easy to check that $\int_{I\!R^{2d}}|M(u,v)|dudv < \infty$ thus we can freely interchange the orders of integration which we will do without any further comment. Next note that

$$\begin{aligned}
\int_{I\!R^{2d}}M(u,v)dudv &= \int_{I\!R^{2(d-1)}}\left(\int_{-\infty}^{\infty}\left(\int_{-\infty}^{\infty}M(u,v)du_1\right)dv_1\right)du^{(1)}dv^{(1)} \\
&= \int_{I\!R^{2(d-1)}}\left(\int_{-\infty}^{\infty}\left(\int_{-\infty}^{v_1}M(u,v)du_1 + \int_{v_1}^{\infty}M(u,v)du_1\right)dv_1\right)du^{(1)}dv^{(1)}
\end{aligned}$$

$$
\begin{aligned}
&= \int_{I\!R^{2(d-1)}} \left(\int_{-\infty}^{\infty} \left(\int_{-\infty}^{v_1} M(u,v) du_1 \right) dv_1 \right) du^{(1)} dv^{(1)} \\
&+ \int_{I\!R^{2(d-1)}} \left(\int_{-\infty}^{\infty} \left(\int_{u_1}^{\infty} M(v,u) dv_1 \right) du_1 \right) dv^{(1)} du^{(1)} \\
&= \int_{I\!R^{2(d-1)}} \left(\int_{-\infty}^{\infty} \left(\int_{-\infty}^{v_1} M(u,v) du_1 \right) dv_1 \right) du^{(1)} dv^{(1)} \\
&+ \int_{I\!R^{2(d-1)}} \left(\int_{-\infty}^{\infty} \left(\int_{-\infty}^{v_1} M(v,u) du_1 \right) dv_1 \right) du^{(1)} dv^{(1)} \\
&= \int_{I\!R^{2(d-1)}} \left(\int_{-\infty}^{\infty} \left(\int_{-\infty}^{v_1} (M(u,v)+M(v,u)) du_1 \right) dv_1 \right) du^{(1)} dv^{(1)},
\end{aligned}
$$

where the third equality follows by renaming (u, v) as (v, u) in the second integral and the fourth equality follows by changing the order of the two innermost integrals in the second expression. Consider now,

$$
\begin{aligned}
M(u,v) + M(v,u) &= \rho_{n-1}^{(0)} \wedge \rho_{n-1}^{(1)}(u,v) G(u,x) G(v,y) + \rho_{n-1}^{(0)} \wedge \rho_{n-1}^{(1)}(v,u) G(v,x) G(u,y) \\
&= \rho_{n-1}^{(0)} \wedge \rho_{n-1}^{(1)}(u,v) \left[G(u,x) G(v,y) - G(v,x) G(u,y) \right] \\
&= \left(\rho_{n-1}^{(0)} \wedge \rho_{n-1}^{(1)}(u,v) \right) \left(G(\cdot,x) \wedge G(\cdot,y)(u,v) \right).
\end{aligned}
$$

Hence

$$
\left| \int_{I\!R^{2d}} M(u,v) du dv \right| \leq \int_{I\!R^{2d}} |\rho_{n-1}^{(0)} \wedge \rho_{n-1}^{(1)}(u,v)| |G(\cdot,x) \wedge G(\cdot,y)|(u,v) du dv. \tag{2.11}
$$

Using (2.11) in (2.10) we have that

$$
\begin{aligned}
||\rho_n^{(0)} \wedge \rho_n^{(1)}||_1 &\doteq \int_{I\!R^{2d}} |\rho_n^{(0)} \wedge \rho_n^{(1)}(x,y)| dx dy \\
&\leq \int_{I\!R^{2d}} \Big(\int_{I\!R^{2d}} \frac{1}{\sigma^{2d}} \phi(\frac{1}{\sigma}(Y_n - x)) \phi(\frac{1}{\sigma}(Y_n - y)) |\rho_{n-1}^{(0)} \wedge \rho_{n-1}^{(1)}(u,v)| \\
&\quad |G(\cdot,x) \wedge G(\cdot,y)|(u,v) dx dy \Big) du dv
\end{aligned}
$$

Define,

$$
K(x,y,u,v) \doteq \frac{1}{\sigma^{2d}} \phi(\frac{1}{\sigma}(Y_n - x)) \phi(\frac{1}{\sigma}(Y_n - y)) |\rho_n^{(0)} \wedge \rho_n^{(1)}(u,v)| |G(\cdot,x) \wedge G(\cdot,y)|(u,v).
$$

Let $\epsilon > 0$ be arbitrary, then

$$
\begin{aligned}
||\rho_n^{(0)} \wedge \rho_n^{(1)}||_1 &= \int_{I\!R^{2d}} \left(\int_{(x,y):|x-y|>\epsilon} K(x,y,u,v) dx dy \right) du dv \\
&+ \int_{I\!R^{2d}} \left(\int_{(x,y):|x-y|\leq\epsilon} K(x,y,u,v) dx dy \right) du dv. \qquad (2.12)
\end{aligned}
$$

Now consider the first integral in (2.12). If $|x-y| > \epsilon$ then $|Y_n - x|^2 + |Y_n - y|^2 > \frac{\epsilon^2}{2}$. This implies that $\exp\left(\frac{-1}{2\sigma^2}\left(|Y_n - x|^2 + |Y_n - y|^2\right)\right) \leq e^{\frac{-\epsilon^2}{4\sigma^2}}$. Thus observing that $\int_{\mathbb{R}^d} G(u,x)dx = \int_{\mathbb{R}^d} G(v,y)dy = 1$ for all $u, v \in \mathbb{R}^d$ we have that

$$\begin{aligned}
\int_{\mathbb{R}^{2d}} \left(\int_{(x,y):|x-y|>\epsilon} K(x,y,u,v)dxdy \right) dudv &\leq \frac{2}{\sigma^{2d}} e^{\frac{-\epsilon^2}{4\sigma^2}} \int_{\mathbb{R}^{2d}} |\rho^{(0)}_{n-1} \wedge \rho^{(1)}_{n-1}|(u,v)dudv \\
&= \frac{2}{\sigma^{2d}} e^{\frac{-\epsilon^2}{4\sigma^2}} ||\rho^{(0)}_{n-1} \wedge \rho^{(1)}_{n-1}||_1. \qquad (2.13)
\end{aligned}$$

Next we consider the second term in (2.12). Let $|x - y| \leq \epsilon$. A straightforward calculation shows that

$$|G(\cdot, x) \wedge G(\cdot, y)|(u,v) \leq G(u,x)|G(v,y) - G(v,x)| + G(v,x)|G(u,x) - G(u,y)| \qquad (2.14)$$

Now

$$\begin{aligned}
|G(u,x) - G(u,y)| &= \frac{1}{\det(b(u))} \left| \phi(b^{-1}(u)(x - a(u))) - \phi(b^{-1}(u)(y - a(u))) \right| \\
&\leq \frac{1}{\underline{B}(2\pi)^{d/2}} \left| e^{-\frac{1}{2}|b^{-1}(u)(x-a(u))|^2} - e^{-\frac{1}{2}|b^{-1}(u)(y-a(u))|^2} \right| \\
&\leq \frac{1}{(\underline{B})^2 (2\pi)^{d/2}} ||x - y||. \qquad (2.15)
\end{aligned}$$

Thus using (2.15) and the observation that $||G||_\infty \doteq \sup_{u,x} |G(u,x)| < \infty$ in (2.14) we get that

$$|G(\cdot, x) \wedge G(\cdot, y)|(u,v) \leq \frac{2||G||_\infty}{(\underline{B})^2 (2\pi)^{d/2}} ||x - y||. \qquad (2.16)$$

Therefore we can now conclude that

$$\begin{aligned}
\int_{\mathbb{R}^{2d}} \left(\int_{(x,y):|x-y|\leq\epsilon} K(x,y,u,v)dxdy \right) dudv &\leq \frac{2||G||_\infty}{(\underline{B})^2 (2\pi)^{d/2}} \epsilon \int_{\mathbb{R}^{2d}} |\rho^{(0)}_{n-1} \wedge \rho^{(1)}_{n-1}|(u,v) \\
&\quad \left(\int_{\mathbb{R}^{2d}} \frac{1}{\sigma^{2d}} \phi(\frac{1}{\sigma}(Y_n - x)) \phi(\frac{1}{\sigma}(Y_n - y)) dxdy \right) dudv \\
&\leq \frac{2||G||_\infty}{(\underline{B})^2 (2\pi)^{d/2}} \epsilon ||\rho^{(0)}_{n-1} \wedge \rho^{(1)}_{n-1}||_1.
\end{aligned}$$

Hence combining the above inequality with (2.12) and (2.13) we have that

$$||\rho^{(0)}_n \wedge \rho^{(1)}_n||_1 \leq \left(\frac{2}{\sigma^{2d}} e^{\frac{-\epsilon^2}{4\sigma^2}} + \frac{2||G||_\infty}{(\underline{B})^2 (2\pi)^{d/2}} \epsilon \right) ||\rho^{(0)}_{n-1} \wedge \rho^{(1)}_{n-1}||_1.$$

Thus

$$\begin{aligned}
\limsup_{\sigma \to 0} \limsup_{n \to \infty} \frac{1}{n} \log \left(||\rho^{(0)}_n \wedge \rho^{(1)}_n||_1 \right) &\leq \limsup_{\epsilon \to 0} \limsup_{\sigma \to 0} \log \left(\frac{2}{\sigma^{2d}} e^{\frac{-\epsilon^2}{4\sigma^2}} + \frac{2||G||_\infty}{(\underline{B})^2 (2\pi)^{d/2}} \epsilon \right) \\
&= \limsup_{\epsilon \to 0} \log \left(\frac{2||G||_\infty}{(\underline{B})^2 (2\pi)^{d/2}} \epsilon \right) \\
&= -\infty
\end{aligned}$$

$\blacksquare$

Proof of Theorem 2.1: Assume without loss of generality that $\sigma < 1$. Observe that

$$\begin{aligned}\limsup_{n\to\infty} \frac{1}{n} \log ||p_n^{(0)} - p_n^{(1)}||_1 &= \limsup_{n\to\infty} \left(\frac{1}{n} \log ||\rho_n^{(0)} \wedge \rho_n^{(1)}||_1 - \frac{1}{n} \log ||\rho_n^{(0)}||_1 - \frac{1}{n} \log ||\rho_n^{(1)}||_1 \right) \\ &\leq \limsup_{n\to\infty} \left(\frac{1}{n} \log ||\rho_n^{(0)} \wedge \rho_n^{(1)}||_1 \right) + 2K\end{aligned}$$

where the inequality follows from Lemma 2.2. Now the theorem follows on applying Lemma 2.3. $\blacksquare$

We now proceed to the general case which relaxes the Gaussian assumption on the densities of ξ_n and ν_n. We will impose the following conditions on q and r:

(A3) There exists a measurable function η, from $\mathbb{R}_+ \cup \{0\} \to \mathbb{R}$ satisfying:

- (a) η is a decreasing function on $\mathbb{R}_+ \cup \{0\}$.
- (b) $\log q(u) \geq \eta(|u|)$ for all $u \in \mathbb{R}^d$.
- (c) For every $c > 0$ there exists a constant $\kappa(c) > -\infty$ such that
  1. $\int_{\mathbb{R}^d} \eta(c|u|) r(u) du > \kappa(c)$.
  2. $\int_{\mathbb{R}^d} \eta(c|u|) q(u) du > \kappa(c)$.
  3. $\int_{\mathbb{R}^d} \eta(c|a(u)|) \mu_i(du) > \kappa(c)$ for $i = 1, 2$.

(A4) There exists a measurable function γ from $\mathbb{R}_+ \cup \{0\} \to \mathbb{R}_+$ satisfying:

- (a) γ is a decreasing function on $\mathbb{R}_+ \cup \{0\}$.
- (b) $\limsup_{t\to\infty} \frac{\gamma(t)}{t^{2d}} = 0$.
- (c) For all $u, v \in \mathbb{R}^d$, $r(u)r(v) \leq \gamma(|u - v|)$.

(A5) The density function q satisfies:

- (a) $||q||_\infty \doteq \sup_{x\in\mathbb{R}^d} |q(x)| < \infty$.
- (b) There exists a finite constant q_{lip} such that for all $x, y \in \mathbb{R}^d$, $|q(x) - q(y)| \leq q_{\text{lip}} |x - y|$.

We now have the following result.

Theorem 2.4 *Assume that assumptions (A1) through (A5) hold. Then there exists $0 < \sigma_0 < \infty$ such that for all $\sigma < \sigma_0$;*

$$\limsup_{n\to\infty} \frac{1}{n} \log ||p_n^{(0)} - p_n^{(1)}||_1 < 0,$$

a.s. P.

Sketch of the Proof: The idea of the proof is to show once more that (2.4) holds for $i = 0, 1$ and then show that

$$\lim_{\sigma\to 0}\limsup_{n\to\infty}\frac{1}{n}\log||\rho_n^{(0)}\wedge\rho_n^{(1)}||_1 = -\infty. \tag{2.17}$$

In order to see that (2.4) holds, observe initially that as in Lemma 2.2 we have the inequality (2.5) with ϕ replaced by r. Observe now that from (A3) ((a) and (b)),

$$\begin{aligned} G(x,y) &= \frac{1}{\det(b(x))}q(b^{-1}(x)(y-a(x))) \\ &\geq \frac{1}{\overline{B}}\eta(|b^{-1}(x)(y-a(x))|) \\ &\geq \frac{1}{\overline{B}}\eta(\frac{1}{\underline{B}}|y-a(x)|). \end{aligned}$$

Using (2.7) and (A3) ((a)), we see that

$$\begin{aligned} \log G(Y_{j-1}+z_{j-1},Y_j+z_j) &\geq \eta\left(\frac{1}{\underline{B}}\left(\overline{B}|\xi_j|+a_{\text{lip}}\sigma|\nu_{j-1}|+\sigma\nu_j+|z_j|+a_{\text{lip}}|z_{j-1}|\right)\right)-\log(\overline{B}) \\ &\geq \eta(\frac{5\overline{B}}{\underline{B}}|\xi_j|)+\eta(\frac{5a_{\text{lip}}}{\underline{B}}|\nu_{j-1}|)+\eta(\frac{5}{\underline{B}}|\nu_j|) \\ &+ \eta(\frac{5}{\underline{B}}|z_j|)+\eta(\frac{5a_{\text{lip}}}{\underline{B}}|z_{j-1}|)-4\eta(0)-\log(\overline{B}), \end{aligned}$$

where the last inequality follows on observing that since η is decreasing, $\eta\left(\sum_{j=1}^k u_j\right)\geq \sum_{j=1}^k\eta(ku_j)-(k-1)\eta(0)$.

In a similar fashion we have that

$$\log G(z_0,Y_1+z_1)\geq\eta(\frac{5\overline{B}}{\underline{B}}|\xi_1|)+\eta(\frac{5}{\underline{B}}|\nu_1|)+\eta(\frac{5}{\underline{B}}|z_1|)+\eta(\frac{5|a(X_0)|}{\underline{B}}|)+\eta(\frac{5|a(z_0)|}{\underline{B}}|)-4\eta(0)-\log(\overline{B}).$$

Observing that,

$$\begin{aligned} \frac{1}{\sigma^d}\int_{I\!R^d}\eta(c|z_j|)r(\frac{z_j}{\sigma})dz_j &= \int_{I\!R^d}\eta(c\sigma|z_j|)r(z_j)dz_j \\ &\geq \int_{I\!R^d}\eta(c|z_j|)r(z_j)dz_j\geq\kappa(c) \end{aligned}$$

we have that

$$\begin{aligned} \log\left(||\rho_n^{(1)}||_1\right) &\geq \sum_{j=1}^n\eta\left(\frac{5\overline{B}}{\underline{B}}|\xi_j|\right)+\sum_{j=1}^n\eta\left(\frac{5}{\underline{B}}|\nu_j|\right)+\sum_{j=1}^{n-1}\eta\left(\frac{5a_{\text{lip}}}{\underline{B}}|\nu_j|\right)+n\kappa(\frac{5}{\underline{B}}) \\ &+ (n-1)\kappa(\frac{5a_{\text{lip}}}{\underline{B}})+\eta\left(\frac{5}{\underline{B}}|a(X_0)|\right)+\int_{I\!R^d}\eta\left(\frac{5}{\underline{B}}|a(z_0)|\right)\mu_i(dz_0)-4n\eta(0)-n\log(\overline{B}). \end{aligned}$$

Now applying the strong law of large numbers and using (A3) ((c); 1, 2, 3) we have that

$$\limsup_{n\to\infty} \sup_{0<\sigma<1} -\frac{1}{n} \log ||\rho_n^{(i)}(\cdot)||_1 \le K,$$

where K is a finite constant.

We now outline the proof of (2.17). As before we have that (2.12) holds where in the definition of $K(\cdot,\cdot,\cdot,\cdot)$, ϕ is replaced by r. Applying (A4) ((a) and (c)) with $u = \frac{Y_n - x}{\sigma}$ and $v = \frac{Y_n - y}{\sigma}$ we have that

$$r\left(\frac{Y_n - x}{\sigma}\right) r\left(\frac{Y_n - y}{\sigma}\right) \le \gamma(\frac{|x-y|}{\sigma}) \le \gamma(\frac{\epsilon}{\sigma}).$$

Therefore as in (2.13),

$$\int_{\mathbb{R}^{2d}} \left(\int_{(x,y):|x-y|>\epsilon} K(x,y,u,v)dxdy \right) dudv \le \frac{2}{\sigma^{2d}} \gamma(\frac{\epsilon}{\sigma}) ||\rho_{n-1}^{(0)} \wedge \rho_{n-1}^{(1)}||_1. \tag{2.18}$$

Next using (2.14) and (A5) ((a) and (b)) we have as in the proof of (2.15) that

$$|G(\cdot,x) \wedge G(\cdot,y)|(u,v) \le \frac{2||q||_\infty q_{\text{lip}}}{\underline{B}^2} ||x-y||. \tag{2.19}$$

This implies that

$$\begin{aligned}\int_{\mathbb{R}^{2d}} \left(\int_{(x,y):|x-y|\le\epsilon} K(x,y,u,v)dxdy \right) dudv \quad &\le \quad \frac{2||q||_\infty q_{\text{lip}}}{\underline{B}^2} \epsilon \int_{\mathbb{R}^{2d}} |\rho_{n-1}^{(0)} \wedge \rho_{n-1}^{(1)}|(u,v) \\ &= \quad \frac{2||q||_\infty q_{\text{lip}}}{\underline{B}^2} \epsilon ||\rho_{n-1}^{(0)} \wedge \rho_{n-1}^{(1)}||_1.\end{aligned}$$

Combining the above inequality with (2.18) we have that

$$||\rho_n^{(0)} \wedge \rho_n^{(1)}||_1 \le \left(\frac{1}{\sigma^{2d}} \gamma(\frac{\epsilon}{\sigma}) + \frac{2||q||_\infty q_{\text{lip}}}{\underline{B}^2} \epsilon \right) ||\rho_{n-1}^{(0)} \wedge \rho_{n-1}^{(1)}||_1.$$

The proof is now completed on observing that

$$\begin{aligned}\limsup_{\sigma\to 0} \limsup_{n\to\infty} \frac{1}{n} \log \left(||\rho_n^{(0)} \wedge \rho_n^{(1)}||_1 \right) \quad &\le \quad \limsup_{\epsilon\to 0} \limsup_{\sigma\to 0} \log \left(\frac{1}{\sigma^{2d}} \gamma(\frac{\epsilon}{\sigma}) + \frac{2||q||_\infty q_{\text{lip}}}{\underline{B}^2} \epsilon \right) \\ &= \quad \limsup_{\epsilon\to 0} \log \left(\frac{2||q||_\infty q_{\text{lip}}}{\underline{B}^2} \epsilon \right) \\ &= \quad -\infty,\end{aligned}$$

where the first equality follows from (A4) ((b) and (a)).■

Remark: From the calculations of the proof one can find lower bound on σ_0 in terms of $\kappa(c)$, a_{lip}, q_{lip}, $||q||_\infty$, $\underline{B}$, $\overline{B}$, and γ. Similarly, one can derive a refined lower bound on σ_0 in the case of Gaussian noise from the proof of Theorem 2.1.

References

[1] R. Atar. Exponential stability for nonlinear filtering of diffusion processes in non-compact domain. *To appear in Annals of Probability.*

[2] R. Atar, F. Viens, and O. Zeitouni. Robustness of Zakai's equations via Feynman-Kac representations. *Stochastic analysis, Control, Optimization and Applications: A Volume in Honor of W. H. Fleming, edited by W. M. McEneaney, G. Yin and Q. Zhang; Birkhauser*, 1997.

[3] R. Atar and O. Zeitouni. Exponential stability for nonlinear filtering. *Annales de l'Institut H. Poincare Probabilites et Statistique*, 33:697–726, 1997.

[4] R. Atar and O. Zeitouni. Lyapunov exponents for finite state nonlinear filtering. *SIAM Journal of Control and Optimization*, 35:36–55, 1997.

[5] A. Budhiraja and H.J. Kushner. Robustness of nonlinear filters over the infinite time interval. *SIAM J. on Control and Optimization*, 36:1618–1637, 1998.

[6] A. Budhiraja and D. Ocone. Exponential stability of discrete time filters for non-ergodic signals. 30:185–193, 1997.

[7] J.M.C Clark, D.L. Ocone, and C. Coumarbatch. Relative entropy and error bounds for filtering of Markov processes. *To appear in Math. Signals Systems & Control*, 1997.

[8] G. Da Prato, M. Fuhrman, and P. Malliavin. Asymptotic ergodicity for the Zakai filtering equation. *C.R. Acad. Sci. Paris, t. 321, Serie I*, pages 613–616, 1995.

[9] B. Delyon and O. Zeitouni. Lyapunov exponents for filtering problems. In M.H.A. Davis and R.J. Elliot, editors, *Applied Stochastic Analysis*, pages 511–521, 1991.

[10] H. Kunita. Asymptotic behavior of the nonlinear filtering errors of Markov processes. *Journal of Multivariate Analysis*, 1:365–393, 1971.

[11] H. Kunita. Ergodic properties of nonlinear filtering processes. In K.C. Alexander and J.C. Watkins, editors, *Spatial Stochastic Processes*, 1991.

[12] F. Le Gland and L. Mevel. Geometric ergodicity in hidden markov models. *Preprint.*

[13] D. Ocone. Asymptotic stability of Benes filters. *To appear in Stochastic Analysis and Applications*, 1997.

[14] D. Ocone. Entropy inequalities and entropy dynamics in nonlinear filtering of diffusion processes. *Stochastic analysis, Control, Optimization and Applications: A Volume in Honor of W. H. Fleming, edited by W. M. McEneaney, G. Yin and Q. Zhang; Birkhauser*, 1997.

[15] D. Ocone and E. Pardoux. Asymptotic stability of the optimal filter with respect to its initial condition. *SIAM Journal of Control and Optimization*, 34:226–243, 1996.

[16] L. Stettner. On invariant measures of filtering processes. In K. Helmes, N. Christopeit, and M. Kohlmann, editors, *Stochastic differential systems, Proc. 4th Bad Honnef Conf.,1988, Lecture Notes in Control and Inform Sci.*, pages 279–292, 1989.