# DETERMINISTIC EQUIVALENTS OF ADDITIVE FUNCTIONALS OF RECURRENT DIFFUSIONS AND DRIFT ESTIMATION

D. LOUKIANOVA AND O. LOUKIANOV

Abstract. Let $X_t$ be a one-dimensional Harris recurrent diffusion, with a drift depending on an unknown parameter $\theta$ belonging to some metric compact $\Theta$. We firstly show that all integrable additive functionals of $X_t$ are asymptotically equivalent in probability to some deterministic process $v_t$. Then we use this result to study the behavior of the maximum likelihood estimator for the parameter $\theta$. Under mild regularity assumptions, we find an upper rate of its convergence as a function of $v_t$, extending some recent results for ergodic diffusions.

## Introduction

The present paper addresses the problem of maximum likelihood estimation (MLE) for continuously observed recurrent diffusions. Consider a stochastic process $X = (X_t)_{t \geq 0}$ in $\mathbb{R}$ given by

$$dX_t = \sigma(X_t)\, dB_t + b(\theta_0, X_t)\, dt,$$

where $(B_t)_{t \geq 0}$ is a Brownian motion and $\theta_0$ is an unknown parameter belonging to some compact metric space $\Theta$. A MLE $\hat{\theta}_t$ of $\theta_0$ is an element of $\Theta$ such that the trajectory $X_s$ observed up to time $t$ has "the highest probability density".

It is well-known that $\hat{\theta}_t$ can be written as

$$\hat{\theta}_t = \arg\sup_{\theta \in \Theta} \left( M_t(\theta) - \frac{1}{2} A_t(\theta) \right),$$

where $M(\theta)$ are martingales under the law of $X$ and $A(\theta) = \langle M(\theta) \rangle$ are their quadratic variations. This representation of $\hat{\theta}_t$ explains why the methods used in MLE are frequently based on the knowledge of asymptotic behavior of $A_t(\theta)$ and $M_t(\theta)$.

There is an important framework where numerous results in this direction can be obtained: that of Harris recurrent diffusions. Recall that this notion means that $X$ admits an invariant measure $\mu$ such that $\mu(A) > 0$ implies $\int_0^\infty \mathbf{1}_A(X_t)\, dt = \infty$ a.s. for every measurable $A$. If $\mu(\mathbb{R})$ is finite then $X$ is called positively recurrent (or ergodic), and null-recurrent otherwise. All diffusions considered in this paper are assumed to be Harris recurrent.

An important point about recurrent diffusions is the Chacon-Ornstein ratio-limit theorem: all integrable additive functionals (IAF) are asymptotically equivalent. In

1991 *Mathematics Subject Classification.* 60G17, 60F10, 92B20, 68T10.

*Key words and phrases.* Regularity of martingale families, Harris diffusion, maximum likelihood estimator, rate of convergence.

ergodic case it yields that every IAF of $X_t$ (in particular $A_t(\theta)$) a.s. grows asymptotically like $t$, providing a natural *deterministic* normalizer for MLE. Moreover, the central limit theorem for martingales and IAF holds here [10]. This makes it possible to find not only the rate of convergence but also the limit distribution of normalized MLE. The results for ergodic diffusions are abundant (see the works of Basawa, Prakasa Rao [1, 17], Feigin [4], Kutoyants [11], Lanska [12], van Zanten [23, 24], Yoshida [26] to mention just a few); an extensive survey can be found in the recent book [10] by Kutoyants.

The situation is different for null-recurrent diffusions, where the behavior of $M_t$ and $A_t$ is more complicated. To the best of our knowledge, there is no general result on MLE here, and the existing ones concern particular models. A notable exception is the class of *regularly varying* null-recurrent diffusions, thoroughly studied by Khasminskii [9, 8, 5], Touati [19], Höpfner and Löcherbach [7] and many other authors. For such diffusions the couples $(M_t/\sqrt{t^\alpha}, A_t/t^\alpha)$ converge to a non-degenerated limit law as $t \to \infty$ for some $\alpha \leq 1$. Using this fact, Höpfner and Kutoyants have presented in [6] one of the first examples concerning the rate of convergence and the limit distribution of MLE for null-recurrent diffusions. Probably, their method can be extended to the whole class of regular variation, but it is not clear how to proceed in the general case when there is no information about *a priori* random behavior of martingales and IAF.

An important attempt to take this randomness into consideration was made by van Zanten in [25]. His approach is based on a natural choice of *random* metrics $d(\theta, \theta')$ and he is able to formulate fairy general results. Unfortunately, all concrete examples he gives in our setting (Harris diffusions) concern the ergodic case. Actually, he uses some kind of uniform CLT theorem to establish the existence of an asymptotic *deterministic* equivalent of $d(\theta, \theta')$, but this theorem is proved only for ergodic diffusions (see [21]).

Trying to find out what can be done in the general Harris recurrent setting, one could hardly expect to obtain a limit distribution of normed MLE beyond the regular variation case, at least by classical technique. Indeed, this class is known to be the largest one where deterministically normed pairs $(M, A)$ converge to some law. Nevertheless, it turns out that for questions like rate of convergence there is no need for such a strong property, but a much weaker one (asymptotic tightness) suffices.

The aim of our paper is to provide a unified treatment for (upper) rates of convergence of MLE for general (both positively and null) recurrent diffusions. Denote as above by $d$ the metrics of $\Theta$, by $\theta_0$ the true value of parameter, by $\hat{\theta}_t$ a MLE of $\theta_0$, and let $\mathbf{P}$ be the law of $X$. Any process $r_t$ satisfying

$$\lim_{M\to\infty} \limsup_{t\to\infty} \mathbf{P}\left(r_t d(\hat{\theta}_t, \theta_0) > M\right) = 0$$

will be called (*probabilistic*[1]) upper rate of convergence. Naturally, one is looking for a "maximal" possible rate.

In a preliminary version of this paper we have obtained a theorem expressing $r_t$ as a function of an IAF $V_\tau$ stopped at some random moment $\tau(t)$. However, its formulation and proof were greatly simplified if we supposed that $V_t$ behaves deterministically. Namely, we needed to show that for some deterministic function

[1]See [14] for the *almost sure* counterpart of the present paper.

$v_t$, any integrable additive functional $V_t$ satisfies

$$\lim_{M\to\infty}\liminf_{t\to\infty}\mathbf{P}\left(\frac{1}{M}\le\frac{V_t}{v_t}\le M\right)=1. \tag{T}$$

This property of asymptotic tightness is obviously true for ergodic or regularly varying diffusions, with $v_t=t^\alpha$. Surprisingly enough, it does not seem to be well known in the general case, since the references we have found in the literature mainly concern discrete settings. The property (T) itself for recurrent Markov chains is proved in the Chen's work [3].

Our first section is thus devoted to the proof of (T) for recurrent diffusions. We follow the technique developed by Chen for Markov chains. A natural candidate for the norming function $v_t$ is $\mathbb{E}_\pi A_t$ for some IAF $A_t$ and initial distribution $\pi$. Clearly, the order of growth of such $v_t$ should be independent of $A_t$ and $\pi$, which leads to the use of *special* additive functionals. In this way we firstly prove (T) for local times (lemma 1.1), and then extend it to all IAF by Chacon-Ornstein theorem.

Finally, we use (T) in section 2 to obtain a statistical result (theorem 4) that we briefly describe now. Recall that we consider a recurrent diffusion $X_t$ given by

$$dX_t=\sigma(X_t)\,dB_t+b(\theta_0,X_t)\,dt,\quad \theta_0\in\Theta,$$

and estimate $\theta_0$ by a MLE $\hat\theta_t$ found from

$$\hat\theta_t=\arg\sup_{\theta\in\Theta}\left(M_t(\theta)-\frac{1}{2}A_t(\theta)\right),$$

where $M_t(\theta)$ is a martingale and $A(\theta)$ is its quadratic variation.

Suppose that there exist two continuous increasing processes $U_t$ and $V_t\to\infty$ such that $\liminf_{t\to\infty}U_t/V_t>0$ and for $t$ great enough

$$\forall\theta\in\Theta,\ d^{2\kappa}(\theta,\theta_0)U_t\le\langle M(\theta)-M(\theta_0)\rangle_t$$
$$\text{and}\quad\forall(\theta,\theta')\in\Theta^2,\ \langle M(\theta)-M(\theta')\rangle_t\le d^{2\delta}(\theta,\theta')V_t \tag{$*$}$$

for some constants $\delta\le 1$ and $\kappa\ge\delta$. Denote by $N(\Theta,d,\varepsilon)$ the $\varepsilon$-covering number of $\Theta$, i.e. the smallest number of balls of $d$-radius $\varepsilon$ covering $\Theta$. Further, suppose that for some function $\phi(x)$ such that $\phi(x)/x^p$ is decreasing for some $p<2$, the following entropy inequality holds:

$$\int_0^\eta\sqrt{\log\left(1+N(\Theta_\eta,d^\delta,\varepsilon)\right)}\,d\varepsilon\le\phi(\eta),$$

where $\Theta_\eta=\{\theta\in\Theta:\ d^\delta(\theta,\theta_0)\le\eta\}$. For example, one can take $\phi(\eta)=\eta$ if $\Theta\subset\mathbb{R}^m$, but some infinite-dimensional compacts $\Theta$ satisfy this condition too.

Let $v_t$ be a deterministic equivalent of IAF of $X_t$. Then we show that any positive process $r_t$ satisfying

$$r_t^{2\kappa}\phi(1/r_t^\delta)\le\sqrt{v_t}$$

is an upper rate of convergence of MLE, i.e.

$$\lim_{M\to\infty}\limsup_{t\to\infty}\mathbf{P}\left(r_t d(\hat\theta_t,\theta_0)>M\right)=0.$$

Naturally, the concrete applications are based on the fact that $\langle M(\theta)-M(\theta')\rangle$ can be expressed in terms of $b$ and $\sigma$. To simplify the formulations, suppose that $\sigma(x)=1$, then $\langle M(\theta)-M(\theta')\rangle_t=\int_0^t(b(\theta,X_s)-b(\theta',X_s))^2\,ds$. Hence an easy way to fulfill the necessary conditions above is to suppose that the function $b(\theta, x)$ satisfies

$$K(x)d^{\kappa}(\theta,\theta') \leq |b(\theta,x) - b(\theta',x)| \leq C(x)d^{\delta}(\theta,\theta') \qquad (**)$$

for some $C(x) > 0$ and $K(x) \geq 0$. Then one can take $U_t = \int_0^t K^2(X_s)\,ds$ and $V_t = \int_0^t C^2(X_s)\,ds$ to satisfy $(*)$. The recurrence of $X_t$ implies that

$$\lim_{t\to\infty} V_t = \infty \quad \text{and} \quad \liminf_{t\to\infty} \frac{U_t}{V_t} \to \frac{\mu(K^2)}{\mu(C^2)} > 0,$$

as soon as $0 < \mu(K^2), \mu(C^2) < \infty$, where $\mu$ is the invariant measure of $X_t$. Thus our result applies and generalizes, for $\Theta \subset \mathbb{R}^m$, the well-known rate $r_t = \sqrt{t^{1/(2\kappa-\delta)}}$ for ergodic diffusions (see [24, 25]).

The $\mu$-integrability condition is not really restrictive even for null-recurrent diffusions, but the assumption $(**)$ sometimes is. For example, if $b(\theta, x)$ has a single "cusp" of order $\kappa$ at $\theta_0$, it gives a rate of convergence which is not optimal. Moreover, if $b(\theta, x)$ is discontinuous or $\Theta$ is infinite-dimensional, $(**)$ does not hold at all. Nevertheless, one can apply theorem 4 in these cases too, proceeding by a different technique, based on some kind of "uniform Chacon-Ornstein theorem". We are able, for example, to treat an equation

$$dX_t = \theta(X_t)\,dt + dB_t$$

provided that the entropy of $\Theta$ is not too great. Some concrete examples are given at the end of the paper.

## 1. Deterministic equivalents of additive functionals

Let $X_t$ be a one-dimensional diffusion given by

$$dX_t = b(X_t)dt + \sigma(X_t)dW_t,$$

whose state space $E$ is an interval $(l, r)$ of $\mathbb{R}$, which may be open, closed or semi-closed, bounded or unbounded. Denote by $\mathcal{E}$ the borelian sigma field on $E$. We assume that the death time of $X_t$ is a.s. infinite and that $X_t$ is Harris recurrent. Recall that it means that $X_t$ admits a unique (up to constant multiples) invariant measure $\mu$ on $E$, such that $\mu(A) > 0$ implies $\int_0^\infty \mathbf{1}_A(X_s)ds = \infty$ a.s. for every measurable $A \subset E$. This condition is equivalent to the following, which shows that BM, OU process, Bessel process with index less or equal to 2 and many linear diffusions are Harris recurrents:

> $X$ has an invariant measure and for all open set $\mathcal{O} \subset E$ one has $\limsup_{t\to\infty} \mathbf{1}_{\mathcal{O}}(X_t) = 1$ a.s.

According to the case $\mu(E) < \infty$ or $\mu(E) = \infty$ the diffusion is said ergodic or null-recurrent.

In this section we will show that there exist a deterministic function $v_t$ associated with the process $X_t$, such that any integrable additive functional is equivalent to $v_t$ in probability when $t \to \infty$.

As usually, we denote by $(\theta_t)_{t\geq 0}$ the family of shift operators. Recall that

**Definition 1.1.** An *additive functional* (abbreviated AF) of the process $X_t$ is a $\bar{\mathbb{R}}_+$-valued, adapted process $A = \{A_t,\ t \geq 0\}$ such that

- it is a.s. non-decreasing, right continuous, vanishing at zero;
- for each pair $(s, t)$, $A_{s+t} = A_t + A_s \circ \theta_t$ a.s.

A *continuous additive functional* (CAF) is an additive functional such that the map $t \to A_t$ is continuous. The most important examples of CAF are $\int_0^t f(X_s)ds$ where $f \geq 0$, and the local time of diffusion at level $x$.

To each AF we can associate a measure $\nu_A$ on the state space by

$$\nu_A(B) = \mathbb{E}_\mu \int_0^1 \mathbf{1}_B(X_s)dA_s, \quad B \in \mathcal{E}.$$

If $\|\nu_A\| =: \nu_A(E) = \mathbb{E}_\mu(A_1)$ is finite, $A$ is said integrable. For example, the measure associated with $A_t = \int_0^t f(X_s)ds$ is $fd\mu$, and if $f \in L_1(\mu)$, $A$ is integrable.

The measure associated with Tanaka-Meyer diffusion local time $L_t^x$ is the Dirac measure in $x$, pondered by $1/s'(x)$, where $s(x)$ is a scale function of $X$, so $L_t^x$ is integrable.

To define an equivalent $v_t$ of IAF, we need a concept of special (for $X_t$) function. We denote by $\mathcal{E}_+$ (resp. $b\mathcal{E}_+$) the space of positive borelian (resp. bounded) functions. If $h \in b\mathcal{E}_+$ , we can define the $U^h$ potential operator associated to $h$ by

$$\forall f \in \mathcal{E}_+, \quad \forall x \in \mathbb{R}, \quad U^h f(x) = \mathbb{E}_x \int_0^\infty \exp\left[ - \int_0^t h(X_s)ds \right] f(X_t)dt.$$

**Definition 1.2.** Following Neveu [15] the function $f \in \mathcal{E}_+$ is said special if for all $h \in b\mathcal{E}_+$ such that $\mu(h) > 0$ the function $U^h f(x)$ is bounded.

This notion can be extended to additive functionals (Brancovan [2]):

**Definition 1.3.** An CAF A is said special (abbreviated SAF), if for all $h \in b\mathcal{E}_+$ such that $\mu(h) > 0$, the function

$$U_A^h \mathbf{1}(x) = \mathbb{E}_x \int_0^\infty \exp\left[ - \int_0^t h(X_s)ds \right] dA_t$$

is bounded.

It is well known that the Harris recurrent diffusion satisfies the Chacon-Ornstein limit quotient theorem:

**Theorem 1.** *For two integrable AF $A$, $B$ of $X$ with $\|\nu_B\| > 0$,*

$$\lim_{t\to\infty} \frac{A_t}{B_t} = \frac{\|\nu_A\|}{\|\nu_B\|} \quad \mathbf{P}_x - a.s. \quad \forall x.$$

*and*

$$\lim_{t\to\infty} \frac{\mathbb{E}_x A_t}{\mathbb{E}_x B_t} = \frac{\|\nu_A\|}{\|\nu_B\|} \quad \forall x.$$

The most important application of special functions and functionals is the following "strong" version of limit-quotient theorem:

**Theorem 2.** *Let $A$ and $B$ be two SAF such that $0 < \|\nu_B\|$. For all initial probabilities $\pi_1$ and $\pi_2$ it holds*

$$\lim_{t\to\infty} \frac{\mathbb{E}_{\pi_1} A_t}{\mathbb{E}_{\pi_2} B_t} = \frac{\|\nu_A\|}{\|\nu_B\|}.$$

*Remark* 1.1. SAF is always integrable.

It follows from Brancovan [2] that every function $f \in \mathcal{E}_+$ with compact support such that $\mu(f) > 0$, is special for one dimensional diffusion.

Now to construct an equivalent of additive functionals we will take some special function $g$, with $\mu(g) = 1$, and put

$$v_t = \mathbb{E}_\pi \int_0^t g(X_s)ds,$$

where $\pi$ is some probability on $E$.

Clearly, $v_t$ is non-negative and non-decreasing. In view of strong limit-quotient theorem the asymptotic order of $v_t$ depends only on the process $X_t$. The following theorem claims that each IAF of $X$ is equivalent to $v_t$ in probability, as $t \to \infty$.

**Theorem 3.** *Let $(X_t)$ be a Harris recurrent scalar diffusion. Let $v_t$ be associated to $(X_t)$ as previously. Then for every IAF $A_t$ such that $\|\nu_A\| > 0$ and every initial distribution $\pi$, it holds:*

$$\lim_{M\to\infty} \liminf_{t\to\infty} \mathbf{P}_\pi(1/M \le A_t/v_t \le M) = 1.$$

This result was proved by Chen for Harris Markov chains in [3]. Here we extend it on continuous setting, following Chen for the proof. Namely, we firstly prove an analogous result for local times, which is based on the following lemma.

**Lemma 1.1.** *Let $(X_t)$ be one dimensional Harris diffusion with a state space $E$. For all $x \in E^\circ$ and $s > 0$, for any initial probability $\pi$ the following inequality holds:*

$$\mathbb{E}_\pi(L_t^x \mathbf{1}_{\{L_t^x \ge s\}}) \le (s + \mathbb{E}_x L_t^x)\mathbf{P}_\pi(L_t^x \ge s)$$

*Proof.* Denote $\tau_s = \inf\{t \ge 0;\ L_t^x = s\}$.

$$\begin{aligned}\mathbb{E}_\pi(L_t^x \mathbf{1}_{\{L_t^x \ge s\}}) &= \mathbb{E}_\pi(L_t^x \mathbf{1}_{\{\tau_s < t\}}) \le \mathbb{E}_\pi(L_{t+\tau_s}^x \mathbf{1}_{\{\tau_s < t\}}) = \\ &= \mathbb{E}_\pi\left([L_{t+\tau_s}^x - L_{\tau_s}^x + s]\mathbf{1}_{\{\tau_s < t\}}\right) = \mathbb{E}_\pi\left([L_t^x \circ \theta_{\tau_s} + s]\mathbf{1}_{\{\tau_s < t\}}\right) = \\ &= \mathbb{E}_\pi\left([\mathbb{E}_{X_{\tau_s}} L_t^x + s]\mathbf{1}_{\{\tau_s < t\}}\right) = (\mathbb{E}_x L_t^x + s)\mathbf{P}_\pi(L_t^x > s),\end{aligned}$$

where in the last line we have used the strong Markov property and the fact that $X_{\tau_s} = x$. ☐

Now we can prove theorem 3.

*Proof of theorem 3.* We only have to prove the theorem in reduced case where $A_t = L_t^x$ with some $x \in E^\circ$. The general case can be easily deduced from this one by Chacon-Ornstein theorem.

In reduced case we have to prove

$$\lim_{M\to\infty} \limsup_{t\to\infty} \mathbf{P}_\pi(L_t^x/v_t \ge M) = 0$$

and

$$\lim_{M\to\infty} \liminf_{t\to\infty} \mathbf{P}_\pi(L_t^x/v_t \ge 1/M) = 1.$$

The first limit easily follows from Chebychev inequality and the fact that $L_t^x$ is special. To prove the second limit we use the lemma 1.1 with $s = v_t/M$ and we get

$$\mathbf{P}_\pi(L_t^x \ge v_t/M) \ge \frac{\mathbb{E}_\pi L_t^x/v_t - 1/M}{\mathbb{E}_x L_t^x/v_t + 1/M}$$

By the strong Chacon-Ornstein theorem

$$\liminf_{t\to\infty} \mathbf{P}_\pi(L_t^x/v_t \geq 1/M) \geq \frac{\|L_t^x\| - 1/M}{\|L_t^x\| + 1/M}$$

whence the theorem follows. □

## 2. Rate of convergence of MLE for Harris diffusions

Let $x \in \mathbb{R}$ and $(\Theta, d)$ be a compact metric space. Consider a family of one-dimensional SDE given by

$$dX_t = \sigma(X_t)dB_t + b(X_t, \theta)dt, \quad X_0 = x, \quad \theta \in \Theta, \tag{1}$$

where $B_t$ is a one-dimensional Brownian motion. We suppose that for each $\theta \in \Theta$ the functions $b$ and $\sigma$ satisfy the usual assumptions for the existence of a weak solution on $[0, +\infty[$; we also suppose $\sigma^2 > 0$.

If the true value of the parameter $\theta$ is unknown and one observes a trajectory $(X_s,\ s \leq t)$, one can approximate $\theta$ by a maximum likelihood estimator (MLE). In this section we give its upper rate of convergence, provided that the observed diffusion is recurrent.

Take $\Omega = C([0, \infty[\to \mathbb{R})$ and let $\mathcal{F}$ be its borelian $\sigma$-field, and $(\mathcal{F}_t)$ its natural filtration. Denote $\mathbf{P}_\theta$ the law of the solution of (1) on $\Omega$. The measures $\mathbf{P}_\theta$ are locally absolutely continuous w.r.t. each other. Let $\theta_0$ be the (unknown) true value of the parameter and $\mathcal{L}_t(\theta)$ be the local density of $\mathbf{P}_\theta$ w.r.t. $\mathbf{P} = \mathbf{P}_{\theta_0}$:

$$\mathcal{L}_t(\theta) = \exp\left(M_t(\theta) - \frac{1}{2}A_t(\theta)\right),$$

$$M_t(\theta) = \int_0^t \left(\frac{b_\theta - b_{\theta_0}}{\sigma}\right)(X_s)\, dW_s$$

and

$$A_t(\theta) = \langle M(\theta)\rangle_t = \int_0^t \left(\frac{b_\theta - b_{\theta_0}}{\sigma}\right)^2 (X_s)\, ds$$

where $b_\theta = b(\theta, \cdot\,)$ and $W_t$ is a Brownian motion under $\mathbf{P}$. Denote also

$$A_t(\theta, \theta') = \langle M(\theta) - M(\theta')\rangle_t = \int_0^t \left(\frac{b_\theta - b_{\theta'}}{\sigma}\right)^2 (X_s)\, ds. \tag{2}$$

Any maximizer $\hat{\theta}_t$ of the random map $\theta \to \mathcal{L}_t(\theta)$ (provided one exists) is called a maximum likelihood estimator of $\theta_0$. In what follows we are interested in finding an (upper) rate of convergence of $\hat{\theta}_t$. Our result concerns the case when $A_t(\theta, \theta')$ satisfies some kind of Hölder property w.r.t. $\theta$ and some kind of non degeneracy property at $\theta_0$. Before we formulate the theorem, let us recall one inequality of Nishiyama [16] which we will need in the proof.

Let $(\Theta, \rho)$ be a compact metric space. Denote by $N = N(\Theta, \rho, \varepsilon)$ its $\varepsilon$-covering number, i.e. the smallest number of closed balls with radius $\varepsilon > 0$ which cover $\Theta$. Let $M = \{M_t(\theta);\ \theta \in \Theta\}$ be a $\rho$-separable family of continuous local martingales indexed by $\Theta$, such that $M_0(\theta) = 0$. Following Nishiyama denote by $\|M_\rho\|_t$ the quadratic $\rho$-modulus of the family $\left\{M_t^\theta\right\}$:

$$\|M_\rho\|_t = \sup_{\theta, \theta'} \frac{\sqrt{\langle M(\theta) - M(\theta')\rangle_t}}{\rho(\theta, \theta')}. \tag{3}$$

**Theorem** (Nishiyama [16]). *Let $M$ be a martingale family as above, and let $\tau$ be a finite stopping time. Then, for any choice of quadratic $\rho$ modulus $\|M_\rho\|_t$ of $M$, it holds that for every $\eta, \kappa$,*

$$\mathbb{E} \sup_{t \le \tau} \sup_{\rho(\theta,\theta') \le \eta} |M_t(\theta) - M_t(\theta')| \mathbf{1}_{\|M_\rho\|_\tau \le \kappa} \le C\kappa \int_0^\eta \sqrt{\log[1 + N(\Theta, \rho, \varepsilon)]}\, d\epsilon, \tag{4}$$

*where $C$ is a universal constant.*

Now we are ready to prove the main statistical result of this paper.

**Theorem 4.** *Consider a family of diffusions given by* (1)*. Suppose that $X_t$ corresponding to $\theta_0$ is Harris recurrent, and let $v_t$ be a deterministic equivalent of integrable additive functionals of $X_t$.*

(1) *Suppose that there are two positive IAF $V_t$ and $U_t$, two constants $\delta \in ]0, 1]$ and $\kappa \ge \delta$ and a random variable $\tau = \tau(\omega)$ such that $\mathbf{P}$-a.s.*

$$\forall t > \tau, \quad \forall (\theta, \theta') \in \Theta^2, \quad A_t(\theta, \theta') \le d^{2\delta}(\theta, \theta') V_t \tag{5}$$

*and*

$$\forall t > \tau, \quad \forall \theta \in \Theta, \quad d^{2\kappa}(\theta, \theta_0) U_t \le A_t(\theta, \theta_0) \tag{6}$$

(2) *Suppose that $\phi$ is a function such that $\phi(\eta)/\eta^p$ is decreasing for some $p < 2$, and that the following entropy inequality holds:*

$$\int_0^\eta \sqrt{\log[1 + N(\Theta_\eta, d^\delta, \varepsilon)]}\, d\epsilon \le \phi(\eta),$$

*where $\Theta_\eta = \{\theta \in \Theta; d^\delta(\theta, \theta_0) \le \eta\}$.*

*Then $\mathbf{P}$-a.s. a MLE $\hat\theta_t$ exists for $t$ large enough, and if $r_t$ is a deterministic function increasing to infinity such that*

$$r_t^{2\kappa} \phi(1/r_t^\delta) \le \sqrt{v_t},$$

*it holds that $r_t$ is an upper rate of convergence, i.e.*

$$\lim_{M \to \infty} \limsup_{t \to \infty} \mathbf{P}\left( r_t d(\hat\theta_t, \theta_0) > M \right) = 0.$$

*Proof.* Recall that we have under $\mathbf{P}$

$$\mathcal{L}_t(\theta) = \exp\left( M_t(\theta) - \frac{1}{2} A_t(\theta) \right).$$

It follows from [13] that $M_t(\theta)$ and $A_t(\theta)$ have continuous in $\theta$ modifications for $t > T_1(\omega)$ great enough. For this modification, $\mathcal{L}_t(\theta)$ is continuous on $\Theta$ and since $\Theta$ is compact, a MLE $\hat\theta_t$ exists and satisfies

$$\hat\theta_t = \arg\sup_{\theta \in \Theta} \left( M_t(\theta) - \frac{1}{2} A_t(\theta) \right).$$

Denote

$$S_{t,j} = \left\{ 2^j < r_t d(\theta, \theta_0) \le 2^{j+1} \right\} \subset \Theta.$$

Let $M \in \mathbb{N}$. We have for fixed $t$ and any $\hat{\theta}_t$ such that $\mathcal{L}_t(\hat{\theta}_t) \geq \mathcal{L}_t(\theta_0) = 0$

$$\left\{ r_t d(\hat{\theta}_t, \theta_0) > 2^M \right\} = \bigcup_{j=M}^{\infty} \left\{ 2^j < r_t d(\hat{\theta}_t, \theta_0) \leq 2^{j+1} \right\} \subseteq$$
$$\bigcup_{j=M}^{\infty} \left\{ \sup_{S_{t,j}} \left( M_t(\theta) - \frac{1}{2} A_t(\theta) \right) \geq 0 \right\} \subseteq \bigcup_{j=M}^{\infty} \left\{ \sup_{S_{t,j}} M_t(\theta) \geq \inf_{S_{t,j}} \frac{1}{2} A_t(\theta) \right\}.$$

If necessary, redefine $\tau$ as $\max(\tau, T_1)$ and denote by

$$\Omega_t(M) = \left\{ t > \tau(\omega); \quad \frac{1}{M} \leq \frac{V_t}{v_t} \leq M; \quad \frac{1}{M} \leq \frac{U_t}{v_t} \leq M \right\}$$

and

$$B_t(M) = \Omega_t(M) \cap \left\{ r_t d(\hat{\theta}_t, \theta_0) > 2^M \right\}.$$

Then

$$\mathbf{P}^* \left\{ r_t d(\hat{\theta}_t, \theta_0) > 2^M \right\} \leq \mathbf{P}^*(\Omega_t^c(M)) + \mathbf{P}^*(B_t(M))$$

and by the theorem 3

$$\lim_{M \to \infty} \limsup_{t \to \infty} \mathbf{P}^*(\Omega_t^c(M)) = 0.$$

Now let us estimate the second term:

$$\mathbf{P}^*(B_t(M)) \leq \sum_{j \geq M} \mathbf{P}^* \left\{ \sup_{S_{t,j}} M_t(\theta) \geq \frac{1}{2} \inf_{S_{t,j}} A_t(\theta);\ \Omega_t(M) \right\} \leq$$
$$\sum_{j \geq M} \mathbf{P}^* \left\{ \sup_{S_{t,j}} M_t(\theta) \geq \frac{1}{2} \inf_{S_{t,j}} d^{2\kappa}(\theta, \theta_0) U_t;\ \Omega_t(M) \right\} \leq$$
$$\sum_{j \geq M} \mathbf{P}^* \left\{ \sup_{S_{t,j}} |M_t(\theta)| \cdot \mathbf{1}_{\Omega_t(M)} \geq \frac{2^{2j\kappa} v_t}{r_t^{2\kappa} M} \cdot \frac{1}{4} \right\} \leq$$
$$4M \frac{r_t^{2\kappa}}{v_t} \sum_{j \geq M} 2^{-2j\kappa} \mathbb{E}^* \left( \sup_{d(\theta, \theta_0) \leq \frac{2^{j+1}}{r_t}} |M_t(\theta) - M_t(\theta_0)| \cdot \mathbf{1}_{\Omega_t(M)} \right), \quad (7)$$

where in the first inequality we have used the assumption (6) and in the last inequality the fact that $M_t(\theta_0) = 0$.

Denote

$$\rho(\theta, \theta') = d^{\delta}(\theta, \theta'), \quad \rho_j = \left( \frac{2^{j+1}}{r_t} \right)^{\delta}$$

and

$$\Psi_j = \left\{ \theta : d(\theta, \theta_0) \leq 2^{j+1}/r_t \right\} = \{ \theta : \rho(\theta, \theta_0) \leq \rho_j \}.$$

The assumption (5) implies on $\Omega_t(M)$ for the quadratic $\rho$ modulus $\|M_\rho\|_t$ given by (3) the inequality

$$\|M_\rho\|_t \leq \sqrt{V_t}.$$

The Nishiyama martingale inequality (4) implies that the last term of (7) is bounded by

$$4M\frac{r_t^{2\kappa}}{v_t}\sum_{j\geq M}2^{-2j\kappa}\mathbb{E}^*\sup_{\substack{(\theta,\theta')\in\Psi_j^2\\ \rho(\theta,\theta')\leq\rho_j}}|M_t(\theta)-M_t(\theta')|\cdot\mathbf{1}_{\|M_\rho\|_t\leq\sqrt{Mv_t}}\leq$$

$$4M\frac{r_t^{2\kappa}}{v_t}\sum_{j\geq M}2^{-2j\kappa}C\sqrt{Mv_t}\int_0^{\rho_j}\sqrt{\log(1+N(\Psi_j,\rho,\varepsilon))}d\varepsilon, \quad (8)$$

provided the last integral is finite. In this case by the second assumption the integral is bounded by $\phi(\rho_j)$ and by the hypothesis on $\phi$

$$\phi(\rho_j)=\phi\left(\left(\frac{2^{j+1}}{r_t}\right)^\delta\right)\leq 2^{(j+1)\delta p}\phi\left(\frac{1}{r_t^\delta}\right). \quad (9)$$

From (7), (8) and (9) we deduce the bound

$$\mathbf{P}^*(B_t(M))\leq 4M\frac{r_t^{2\kappa}}{v_t}\sum_{j\geq M}2^{-2j\kappa}C\sqrt{Mv_t}2^{(j+1)\delta p}\phi\left(\frac{1}{r_t^\delta}\right)\leq$$

$$CM\sqrt{M}\frac{r_t^{2\kappa}}{\sqrt{v_t}}\phi\left(\frac{1}{r_t^\delta}\right)\sum_{j\geq M}2^{-j(2\kappa-\delta p)}\leq CM\sqrt{M}\sum_{j\geq M}2^{-j(2\kappa-\delta p)}, \quad (10)$$

where we have used the last assumption of the theorem and denoted by $C$ an universal constant. Hence the sum (10) is equivalent to $M\sqrt{M}\sum_{j\geq M}2^{-j(2\kappa-\delta p)}$, which converges to 0 as $M\to\infty$, and we get the requested result. □

*Remark* 2.1. It is easy to see that equivalence in probability is a transitive relation, so any $r'_t$ equivalent in probability to $r_t$ will be an upper rate too. For each particular case we will find some function $f$ such that $r_t=f(v_t)$. According to the theorem 3, we can replace $v_t$ by $A_t$, where $A_t$ is an IAF of $X$, and check that $f(A_t)$ will be an upper rate too. Though we don't know the true point $\theta_0$ and can not calculate $v_t$, taking $A_t=\int_0^t g(X_s)ds$ where $g$ is bounded with compact support on $E$ we obtain an IAF. Hence the (random) rate $f(A_t)$ can be evaluated from the observation of trajectory $(X_s;\ s\leq t)$.

Before proceeding to examples, let us recall a simple criterion of recurrence. Consider a diffusion

$$dX_t=b(X_t)\,dt+\sigma(X_t)\,dB_t$$

and put

$$s(x)=\exp\left(-\int_0^x\frac{2b(y)}{s^2(y)}\,dy\right).$$

The diffusion $X_t$ is recurrent if and only if the function $S(x)=\int_0^x s(y)\,dy$ is a space transformation of $\mathbb{R}$: $\lim_{x\to\pm\infty}S(x)=\pm\infty$ (see [18]). In this case, its invariant measure is

$$\mu(dx)=\frac{2\,dx}{\sigma^2(x)s(x)}.$$

**Example 2.1.** Suppose that $\Theta\subset\mathbb{R}^d$. Let $X_t$ be Harris with invariant measure $\mu$. Suppose that

- the drift coefficient has the Holder property with respect to $\theta$:
$$\forall x \in \mathbb{R},\ \forall(\theta,\theta') \in \Theta^2, \quad |b(\theta,x)-b(\theta',x)| \le C(x)d^{\delta}(\theta,\theta').$$
- for some $K(x) > 0$ and $\kappa \ge \delta$ the following condition holds:
$$\forall x \in \mathbb{R},\ \forall\theta \in \Theta, \quad |b(\theta,x)-b(\theta_0,x)| \ge K(x)d^{\kappa}(\theta,\theta_0);$$
- the functions $K(x)/\sigma(x)$ and $C(x)/\sigma(x)$ belong to $L^2(\mu)$.

It is easy to see that in this case two first assumption of the theorem 4 are satisfied with
$$V_t = \int_0^t \frac{C^2}{\sigma^2}(X_s)\,ds \quad \text{and} \quad U_t = \int_0^t \frac{K^2}{\sigma^2}(X_s)\,ds.$$
It is easy to see also that $N(\Psi_j,\rho,\varepsilon) \sim (\rho_j/\varepsilon)^{d/\delta}$, so
$$\int_0^{\rho_j} \sqrt{\log(1+N(\Psi_j,\rho,\varepsilon))}d\varepsilon \sim \int_0^{\rho_j} \sqrt{\log(1+(\rho_j/\varepsilon)^{d/\delta})}d\varepsilon \sim \rho_j$$
and $\phi(x) = x$ in this case. The relation
$$r_t^{2\kappa}\phi(1/r_t^{\delta}) \le \sqrt{v_t}$$
yields the rate $r_t = v_t^{\frac{1}{2(2\kappa-\delta)}}$ usual in this case (see for instance Van Zanten [24]). Hence this example generalizes Van Zanten's result on the general recurrent case.

The assumption of $\mu$-integrability can seem restrictive for null-recurrent diffusions, since $\mu$ is infinite, but in fact it often holds. For example, consider the following SDE:
$$dX_t = \frac{\theta X_t}{1+\theta^2+X_t^2}\,dt + dB_t.$$
It is easy to see that its solution is recurrent if and only if $\theta \le 1/2$ and null-recurrent for $|\theta| \le 1/2$. A simple calculation shows that $\delta = \kappa = 1$ and $C(x) \sim K(x) \sim (1+|x|)^{-1}$ for $|x|$ great enough. If $\theta_0 \in ]-\infty,1/2[$, then $\mu(x) \sim (1+|x|)^{2\theta_0}\,dx$, so the integrability assumptions are fulfilled and theorem 4 applies. Unfortunately, it is not clear what happens in the "border" case $\theta_0 = 1/2$.

**Example 2.2.** We borrow this example from [10, p. 270]. Consider a diffusion with a switching drift
$$dX_t = -\mathrm{sgn}\,(X_t-\theta)\,dt + dW_t, \quad \theta \in [a,b] \subset \mathbb{R}.$$
It is easy to see that $X_t$ is ergodic with invariant measure $\mu(dx) = \exp(-2|x-\theta|)\,dx$. By the ergodic theorem
$$\frac{1}{t}A_t(\theta,\theta') = \frac{4}{t}\int_0^t \mathbf{1}_{[\theta,\theta']}(X_s)\,ds \to \int \mathbf{1}_{[\theta,\theta']}(x)\mu(dx) = f(\theta,\theta').$$
The function $f(\theta,\theta')$ satisfies $K|\theta-\theta'| \le f(\theta,\theta') \le C|\theta-\theta'|$ for some constants $0 < K < C$. Hence, if we prove that the convergence above is uniform in $(\theta,\theta')$, then the assumptions of theorem 4 will hold with $\kappa = \delta = 1/2$, $U_t = Kt/2$ and $V_t = 2t$. The proof follows immediately from the uniform ergodic lemmas 4.1 and 4.2 in [22], which claim that
$$\sup_{\theta}\left|\frac{1}{t}\int_0^t \mathbf{1}_{]-\infty,\theta]}(X_s)\,dW_s - \int \mathbf{1}_{]-\infty,\theta]}(x)\mu(dx)\right| \to 0 \quad \text{as } t \to \infty.$$
As above, we can take $\phi(x) = x$, which gives the well-known upper rate $r_t \sim t$.

**Example 2.3.** The last example concerns infinite-dimensional parametric spaces. Consider a diffusion

$$dX_t = \theta(X_t)dt + dW_t,$$

where $\theta(x)$ is an unknown function to be estimated. Suppose that any candidate $\theta(x)$ is supported by $[0,1]$ and satisfies for some $\alpha \in ]1/2, 1]$

$$\sup_{0\le s\le 1} |\theta(s)| + \sup_{0\le s<t\le 1} \frac{|\theta(t)-\theta(s)|}{|t-s|^\alpha} \le 1,$$

Let $\Theta$ be the set of all such $\theta$, then $X_t$ is null-recurrent for any $\theta \in \Theta$. Remark also that the invariant measures $\mu_\theta$ satisfy

$$0 < m_\theta \le d\mu_\theta/dx \le M_\theta$$

for some constants $m_\theta$ and $M_\theta$.

Denote by $\mu$ the invariant measure corresponding to the true parameter $\theta_0$. The metric on $\Theta$ we will consider is $d(\theta_1, \theta_2) = \|\theta_1 - \theta_2\|_{L^2(\mu)}$. The set $\Theta$ is compact in $L^2(\mu)$. Moreover, since $\|\cdot\|_{L^2(\mu)} \le c\,\|\cdot\|_\infty$, we have (see [20])

$$N(\Theta, \varepsilon, d) \le N(\Theta, c\varepsilon, \|\cdot\|_\infty) \sim \exp(A\varepsilon^{-1/\alpha}),$$

and the entropy condition of theorem 4 is satisfied with $\phi(x) = x^{1-1/2\alpha}$.

Now let us find two IAF $V_t$ and $U_t$, such that the first assumptions of the theorem 4 hold. In our case

$$A_t(\theta_1, \theta_2) = \int_0^t (\theta_1 - \theta_2)^2(X_s)ds.$$

Put

$$Z_t = \int_0^t \mathbf{1}_{[0,1]}(X_s)ds,$$

then $A_t(\theta_1, \theta_2)$ and $Z_t$ are IAF of $X_t$, hence by Chacon-Ornstein theorem

$$\forall(\theta_1,\theta_2), \quad \frac{A_t(\theta_1,\theta_2)}{Z_t} \to \frac{d^2(\theta_1,\theta_2)}{\mu([0,1])} \quad \text{a.s.}$$

We will show that this convergence is uniform on $\Theta^2$. Choose a dense countable subset $\Theta'$ of $\Theta$ such that the convergence above almost surely holds for all $(\theta_1, \theta_2) \in \Theta' \times \Theta'$. We have

$$\begin{aligned}\frac{|A_t(\theta_1,\theta_2) - A_t(\psi_1,\psi_2)|}{Z_t} &= \frac{1}{Z_t}\left|\int_0^t ((\theta_1-\theta_2)^2 - (\psi_1-\psi_2)^2)(X_s)\,ds\right| \le \\ &\frac{1}{Z_t}\int_0^t |\theta_1-\theta_2-\psi_1+\psi_2|\cdot|\theta_1-\theta_2+\psi_1-\psi_2|(X_s)\,ds \le \\ &(\|\theta_1-\psi_1\|_\infty + \|\theta_2-\psi_2\|_\infty)\frac{\int_0^t 4\cdot\mathbf{1}_{[0,1]}(X_s)\,ds}{Z_t} = \\ &4(\|\theta_1-\psi_1\|_\infty + \|\theta_2-\psi_2\|_\infty)\end{aligned}$$

A.s. the family $A_t(\theta_1, \theta_2)/Z_t$ (indexed by $t$) being equicontinuous and convergent on dense subset $\Theta' \times \Theta'$ to a continuous limit $d^2(\theta_1, \theta_2)/\mu([0,1])$, it converges uniformly on whole $\Theta^2$. Hence a.s. there exists some $\tau(\omega)$ such that

$$\forall t > \tau,\ \forall(\theta_1,\theta_2) \in \Theta^2, \quad \frac{d^2(\theta_1,\theta_2)}{2\mu([0,1])} \le \frac{A_t(\theta_1,\theta_2)}{Z_t} \le \frac{2d^2(\theta_1,\theta_2)}{\mu([0,1])}.$$

Taking $U_t = Z_t/2\mu([0,1])$ and $V_t = 2Z_t/\mu([0,1])$ we satisfy all the assumptions of theorem 4 and obtain

$$r_t \sim (v_t)^{\frac{1}{2+1/\alpha}} ,$$

where by $v_t$ we denote the deterministic equivalent of the IAF of $X_t$. According to the remark 2.1, $r_t = Z_t^{\frac{1}{2+1/\alpha}}$ is an upper rate too. This one is random and depends only on observations of unknown processes.

## Acknowledgments

The authors thank Sylvain Delattre for pointing them out the result of Chen which permitted to simplify the earlier version of this manuscript.

## References

[1] I.V. Basawa, B.L.S. Prakasa Rao (1980), *Statistical Inference for Stochastic Processes*, Academic Press, London.
[2] M. Brancovan (1979), Fonctionnelles additives spéciales des processus récurrents au sens de Harris *Wahrscheinlichkeitstheorie* **47**, 163–194.
[3] X. Chen (1999), How often does a Harris recurrent Markov chain recur? *Ann. Prob.* **3**, 1324–1346.
[4] P. D. Feigin (1976), Maximum likelihood estimation for continuous-time stochastic processes, *Adv. Appl. Prob.* **8**, 712–736.
[5] R. Z. Has'minskii (1980), *Stochastic stability of differential equations.*, Aalphen: Sijthoff and Noordoff.
[6] R. Höpfner, Yu. Kutoyants (2003), On a problem of statistical inference in null recurrent diffusions, *Stat. Inference Stoch. Process.* **6**(1) 25–42.
[7] R. Höpfner, E. Löcherbach (2003), Limit Theorems for Null recurrent Markov Processes. *Memoirs of the AMS*, **768**, Providence, Rhode Island.
[8] R. Khasminskii (2000), Asymptotic behavior of parabolic equation arising from one-dimensional null-recurrent diffusions. *J. Diff. Equations* **161**, 154–173.
[9] R. Khasminskii (2001), Limit distributions of some integral functionals for null-recurrent diffusions. *Stochastic Process. Appl.* **92**, 1–9.
[10] Yu. Kutoyants (2004), *Statistical Inference for Ergodic Diffusion Processes*, Springer-Verlag.
[11] Yu. Kutoyants (1977), Estimation of the trend parameter of a diffusion process in the smooth case, *Theor. Probab. Appl.*, **22**, 399–405.
[12] V. Lanska (1979), Minimum contrast estimation in diffusion process. *J. Appl. Probab.* **16**, 65–75.
[13] D. Loukianova, O. Loukianov (preprint), Uniform law of large numbers and consistency of estimators for Harris diffusions, to appear in *Statist. Probab. Letters.*
[14] D. Loukianova, O. Loukianov (preprint), Almost sure rate of convergence of maximum likelihood estimators for multidimensional diffusions, submitted.
[15] J. Neveu (1972), Potentiel markovien récurrent des chaînes de Harris, *Ann. Inst. Fourier* **22**, 85–130.
[16] Nishiyama (1999), A maximum inequality for continuous martingales and M-estimation in gaussian white noise model, *Ann. Statist.* **27**(2), 675–696.
[17] B.L.S. Prakasa Rao (1999), *Statistical Inference for Diffusion Type Processes.* London: Arnold.
[18] D. Revuz, M. Yor (1994), *Continuous Martingales and Brownian Motion*, Second Edition, Springer-Verlag, Berlin, Heidelberg.
[19] A. Touati (1988), Théorèmes limites pour les processus de Markov récurrents, *unpublished paper*. See also C.R.A.S. Paris Série I **305**, 841-844 (1987)
[20] A.W. van der Vaart, J.A. Wellner (1996), *Weak Convergence and Empirical Processes*, Springer-Verlag, New York.
[21] A.W. van der Vaart, H. van Zanten (preprint), Donsker theorems for diffusions: necessary and sufficient conditions, to appear in *Ann. Prob.*

[22] H. van Zanten (1999), On the Uniform Convergence of Local Time and the Uniform Consistency of Density Estimators for Ergodic Diffusions, *Probability, Networks and Algorithms* PNA-R9909.

[23] H. van Zanten (2003), On uniform laws of large numbers for ergodic diffusions and consistency of estimators, *Stat. Inference Stoch. Process.* **6**(2), 199-213.

[24] H. van Zanten (2003), On empirical processes for ergodic diffusions and rates of convergence of $M$-estimators, *Scand. J. Statist.* **30**(3), 443–458.

[25] H. van Zanten (preprint), On the rate of convergence of the MLE in brownian semimartingale models, to appear in *Bernoulli*.

[26] N. Yoshida (1990), Asymptotic behavior of $M$-estimators and related random field for diffusion process, *Ann. Inst. Statist. Math.* **42** (2), 221–251.

Département Mathématique, Université d'Evry, France
*E-mail address*: dasha@maths.univ-evry.fr

Département Informatique, IUT de Fontainebleau, France
*E-mail address*: oleg@iut-fbleau.fr