HYPOTHESIS TESTING FOR SIGNAL DETECTION PROBLEM AND LARGE DEVIATIONS

TAIZO CHIYONOBU

Abstract. We consider a signal detection problem for the continuous-time stationary diffusion processes. The optimal decision region is given by Neyman-Pearson's lemma. We establish certain large deviation estimates, and with the help of it we show that the error probability of the second kind of the signal detection tends to zero or one exponentially fast, depending on the fixed exponent of the decay of the error probability of the first kind, as the observation time goes to infinity.

§1. Introduction

We consider the signal detection problem for the continuous time channel; namely, the problem of hypothesis testing to decide, after observing from time 0 to time T, whether x_t or y_t is the actual outcome. Let R be the probability law of the actual outcome process. The two hypothesis we form are;

$$H_0: \ R = P$$

and

$$H_1: \ R = Q,$$

where P and Q are the laws of two distinct stationary diffusion processes on the one-dimensional flat torus $\mathbf{T}$.

The process of hypothesis testing is based on a partition of Ω_T of the space of the outcome processes on $[0, T]$ into two regions (decision regions) B_T and its compliment B_T^c. Suppose that we are given a outcome $z. \in \Omega_T$, and if $z. \in B_T^c$, then we decide H_0 is true and otherwise, we decide that H_1 is true. The probability $e_1(T)$ of accepting H_1 when H_0 is actually true is called the error probability of the 1-st kind. On the other hand, the probability of accepting H_0 when H_1 is actually true is called the error

probability of the 2-nd kind. That is;

$$e_1(T) = P(B_T)$$

and

$$e_2(T) = Q(B_T^c).$$

Given a maximal value $e_1(T)^*$ of $e_1(T)$, the way of determining the optional decision region B_T so as to minimize $e_2(T)$ subject to the constraint

$$e_1(T) \leq e_1(T)^*$$

is given by the following well-known result due to Neyman-Pearson:

LEMMA. (Neyman-Pearson) *For any* $\xi \in \mathbf{R}$, *let*

$$B_T(\xi) \equiv \left\{ z. \in \Omega_T \; ; \; \frac{1}{T} \log \frac{dQ}{dP}\bigg|_{\mathcal{F}_{[0,T]}} (z.) \geq \xi \right\},$$

and let us define
(1.1) $$e_1(T, \xi) \equiv P(B_T(\xi))$$
and
(1.2) $$e_2(T, \xi) \equiv Q(B_T(\xi)^c).$$

Then for the error probabilities $e_1(T)$ *and* $e_2(T)$ *with the choice of any decision region,*

$$e_2(T) \geq e_2(T, \xi)$$

whenever

$$e_1(T) \leq e_1(T, \xi).$$

In view of this result, we are interested in the asymptotic behavior of the error probabilities $e_1(T, \xi)$ and $e_2(T, \xi)$ for the stationary processes P and Q. The purpose of this paper is to show, in the case P and Q are in a class of ergodic stationary diffusion processes which satisfy the large deviation principle of Donsker-Varadhan's type, that when the rate α of the exponential decay of $e_1(T)$ is specified, the second error probability $e_2(T)$ converges to 0 or 1 exponentially, as the observation time goes to infinity, depending on the value of α and the rate of the decay is calculated in terms of the relative entropy. We will be content with the situation that P and Q are the stationary processes on the torus $\mathbf{T}$, since in that case the whole

argument can be simplified considerably and we can illustrate more clearly how the large deviation technique can be used to study the signal detection problem.

Let us give the notations we will use in this article. Let $\Omega = \mathcal{C}([0, \infty) \to \mathbf{T})$ be the space of $\mathbf{T}$-valued continuous trajectories x_t on $\mathbf{T}$. Let θ_t, $t \in \mathbf{R}$ be the shift map on Ω, i.e., $(\theta_t x)_s = x_{t+s}$. We denote by $\mathcal{F}_I$ the σ-field in Ω generated by x_s, $s \in I$ for any interval I. We denote by $\mathcal{M}_s(\Omega)$ the set of shift invariant measures on Ω.

We denote by $\mathcal{M}_1(\mathbf{T})$ the space of probability measures on $\mathbf{T}$. Let $\mathcal{K}$ be the collection of $U \in \mathcal{C}^\infty(\mathbf{T})$ such that $m_U(dx) = e^{-U(x)} dx \in \mathcal{M}_1(\mathbf{T})$. For each $U \in \mathcal{K}$, let $L^U \phi(x) = \frac{1}{2}(\phi''(x) - U'(x)\phi'(x))$ and let $\{P_x^U\}_{x \in \mathbf{T}}$ be the diffusion measure generated by L^U. Note that, since $(L^U)^* m_U = 0$, m_U is the invariant measure of the L^U-diffusion. See, for example [IW, Proposition 4.5], and [DS, Section 6.2]. Since $m_U(dx) \in \mathcal{M}_1(\mathbf{T})$, we see that $P^U = \int_{\mathbf{T}} P_x^U m_U(dx) \in \mathcal{M}_s(\Omega)$.

Now we assume that $P \in \mathcal{M}_s(\Omega)$ and $Q \in \mathcal{M}_s(\Omega)$ are given by the following.

(**A**) $\qquad P = P^a$ and $Q = P^b$ for some distinct $a \in \mathcal{K}$ and $b \in \mathcal{K}$.

Note that P and Q have $m_a(dx) = e^{-a(x)} dx$ and $m_b(dx) = e^{-b(x)} dx$ as their marginal distributions respectively. We also notice, and later give the precise statement, that the diffusion P and Q satisfy the large deviation principle and the relative entropy plays the crucial role. The relative entropy of $R \in \mathcal{M}_s(\Omega)$ with respect to $P \in \mathcal{M}_s(\Omega)$ is given by

$$
H(R|P) \equiv \begin{cases} \lim_{T \to \infty} \dfrac{1}{T} \int \log \left. \dfrac{dR}{dP} \right|_{\mathcal{F}_{[0,T]}} dR, \\ \qquad\qquad \text{if } dR|_{\mathcal{F}_{[0,T]}} \ll dP|_{\mathcal{F}_{[0,T]}} \text{ for each } T, \\ \infty, \qquad\qquad \text{otherwise.} \end{cases}
$$

Now we state our main result. We say that the map $I : \mathcal{M}_1(\mathbf{T}) \to [0, \infty)$ is a rate function if it is lower semi-continuous and it is a good rate function if it has a compact level set, i.e., $\{\mu \in \mathcal{M}_1(\mathbf{R}) ; I(\mu) \leq L\}$ is compact in $\mathcal{M}_1(\mathbf{R})$ for all $L \geq 0$.

MAIN THEOREM 1.3. (i) *For each $\alpha > 0$, there is a $\xi(\alpha) \in \mathbf{R}$, such that for any small $\epsilon > 0$,*

$$
\liminf_{T \to \infty} \frac{1}{T} \log P\left(\frac{1}{T} \log \left. \frac{dQ}{dP} \right|_{\mathcal{F}_{[0,T]}} > \xi(\alpha) - \epsilon \right) \geq -\alpha
$$

and

$$\limsup_{T\to\infty} \frac{1}{T} \log P\left(\frac{1}{T} \log \frac{dQ}{dP}\bigg|_{\mathcal{F}_{[0,T]}} \geq \xi(\alpha) + \epsilon \right) \leq -\alpha.$$

(ii) *There is a good, convex rate function* $I_Q : \mathbf{R} \to [0, \infty]$ *such that* $I_Q(\xi) = 0$ *if and only if* $\xi = H(Q|P)$ *for which the following holds*: *If* $0 < \alpha < H(Q|P)$, *then*

$$(1.4) \qquad \liminf_{T\to\infty} \frac{1}{T} \log Q\left(\frac{1}{T} \log \frac{dQ}{dP}\bigg|_{\mathcal{F}_{[0,T]}} \leq \xi(\alpha) \right) \leq -I_Q(\xi(\alpha)),$$

and if $\alpha > H(Q|P)$, *then*

$$(1.5) \quad \liminf_{T\to\infty} \frac{1}{T} \log \left(1 - Q\left(\frac{1}{T} \log \frac{dQ}{dP}\bigg|_{\mathcal{F}_{[0,T]}} < \xi(\alpha) \right) \right) \leq -I_Q(\xi(\alpha)).$$

In other words, if we fix the exponent α of the decay of the probability of the first kind, then the probability of the error of the second kind approaches to 0 or 1 according to whether $\alpha > H(Q|P)$ or $\alpha < H(Q|P)$.

We will give some remarks. The signal detection problem has been considered by many authors. For the recent result, we refer to [HK], [NK1] and [NK2]. In those papers, they treat the discrete time channel which are i.i.d. sequence or the Markov chain. More recently, in [IS], they showed the same result for the continuous Gaussian channel, i.e., y_t is given by Brownian motion plus the mean 0 Gaussian process independent of the Brownian motion. They are able to do so with the help of the spectral analysis for the Gaussian process.

The structure of the paper is as follows: In Section 2, we state some preliminary facts which we will use in the coming sections. In Section 3, we give the ergodic theorem for the Radon-Nikodym density which appears in the Neyman-Pearson Lemma. We also prove the variational formula associated with it. In Section 4, we prove the large deviation principle for the Radon-Nikodym density and as a corollary of them we prove our main result.

The author is grateful to Prof. Ihara of Nagoya University for the stimulating discussions we shared on the material.

§2. Preliminaries

Since

$$(2.1) \qquad\qquad L^U = \frac{e^U}{2} \frac{d}{dx}\left(e^{-U} \frac{d}{dx} \right),$$

for each $U \in \mathcal{K}$, it has a non-negative self-adjoint extension $\bar{L}^U$ on $L^2(m_U)$. Define the Dirichlet form $\mathcal{E}^U$ on $L^2(m_U)$ by

$$\mathcal{E}^U(\phi, \phi) = \int_{[0,\infty)} \lambda \, d(E_\lambda \phi, \phi)_{L^2(m_U)},$$

where $\{E_\lambda, \lambda \in [0,\infty)\}$ is the spectral resolution of the identity for $\bar{L}^U$. Then, by (2.1),

$$(2.2) \qquad \mathcal{E}^U(\phi, \phi) = \frac{1}{2} \int_{\mathbf{T}} |\phi'|^2(x) m_U(dx)$$

for all $\phi \in L^2(m_U) \cap C^1(\mathbf{T})$. By virtue of (2.2), we easily see the following. The detailed proof can be found in [DS, Theorem 6.2.9].

LEMMA 2.3. *If $\phi \in L^2(m_U)$ and $\mathcal{E}^U(\phi, \phi) = 0$, then ϕ is constant.*

As a result, P^U is ergodic, and so by the individual ergodic theorem, we obtain the following lemma.

LEMMA 2.4. *For all $U \in \mathcal{K}$,*

$$\frac{1}{T} \int_0^T \phi(x_s) \, ds \to \int_{\mathbf{T}} \phi \, dm_U, \quad P^U\text{-a.s.}$$

as $T \to \infty$, for all $\phi \in L^1(m_U)$.

Let $J^U : \mathcal{M}_1(\mathbf{T}) \to [0, \infty]$ be given by

$$J^U(\mu) \equiv \begin{cases} \mathcal{E}^U(f^{1/2}, f^{1/2}), & \text{if } \mu \ll m_U \text{ and } f = d\mu/dm_U, \\ \infty, & \text{otherwise.} \end{cases}$$

Recall that for $U \in \mathcal{K}$, $m_U \in \mathcal{M}_1(\mathbf{T})$ is given by $m_U(dx) = e^{-U(x)}dx$ and $H(R|P)$ is the relative entropy of $R \in \mathcal{M}_s(\Omega)$ with respect to $P \in \mathcal{M}_s(\Omega)$.

LEMMA 2.5. *For all $U \in \mathcal{K}$ and $V \in \mathcal{K}$,*

$$(2.6) \qquad J^U(m_V) = \frac{1}{8} \int_{\mathbf{T}} |V' - U'|^2(x) m_V(dx) = H(P^V | P^U).$$

Proof. By (2.2), the first equality can easily be checked by direct computation.

Before we prove the second equality, we remark that the diffusion P_x^U can be constructed in the following way: Let $\bar{U} \in C_b^\infty(\mathbf{R})$ be defined by

$\bar{U}(\bar{x}) = U([\bar{x}])$ for all $\bar{x} \in \mathbf{R}$, where $[\,\cdot\,]$ is the Gauss symbol, and let $\bar{x}. = (\bar{x}_t, t \in [0,\infty))$ be the diffusion process on $\mathbf{R}$ given by the solution of the stochastic differential equation:

$$(2.7) \qquad d\bar{x}_t = dw_t + \frac{1}{2}\bar{U}'(\bar{x}_t)\,dt$$

with $\bar{x}_0 = \bar{x}$, where $w. = (w_t)_{t \in [0,\infty)}$ is the standard Brownian motion. Then, $x_t = [\bar{x}_t]$ is the diffusion process on $\mathbf{T}$, the law of which governed by P_x^U.

Observe that for any $f. \in C([0,\infty) \to \mathbf{T})$ such that $f_0 \in [0,1)$, there is a $\bar{f}. \in C([0,\infty) \to \mathbf{R})$ for which $[\bar{f}_t] = f_t$ for all $t \in [0,\infty)$. Thus, we can and will identify the diffusion $x.$ on $\mathbf{T}$ with the diffusion $\bar{x}.$ on $\mathbf{R}$. Now we see that, by Girsanov formula (See [IW, Section 4.4], See also [D, Section 2.12 and Section 5.5]),

$$(2.8) \quad \left.\frac{dP^V}{dP^U}\right|_{\mathcal{F}_{[0,T]}}(x.)$$

$$= \phi(x_0)\phi(x_T)\exp\left(\frac{1}{2}\int_0^T (\bar{V}' - \bar{U}')(\bar{x}_t)\,d\bar{x}_t\right.$$

$$\left. - \frac{1}{8}\int_0^T (|\bar{V}'|^2 - |\bar{U}'|^2)(\bar{x}_t)\,dt\right)$$

$$= \phi(x_0)\phi(x_T)\exp\left(\frac{1}{2}\int_0^T (\bar{V}' - \bar{U}')(\bar{x}_t)\left(d\bar{x}_t - \frac{1}{2}\bar{U}'(\bar{x}_t)\,dt\right)\right.$$

$$\left. - \frac{1}{8}\int_0^T |\bar{V}' - \bar{U}'|^2(\bar{x}_t)\,dt\right)$$

where

$$\phi^2(x) = \left.\frac{dP^V}{dP^U}\right|_{\mathcal{F}_{\{0\}}}(x).$$

In fact, by using Ito formula and noting that $m_U(x) \equiv e^{-U(x)}$ satisfies $(L^U)^*m_U = 0$ for all $U \in \mathcal{K}$, one can check that the RHS of (2.8) is a local martingale.

Now we denote by $z.$ and $\bar{z}.$ the diffusion processes on $\mathbf{T}$ and $\mathbf{R}$ corresponding to P^V. Then, it satisfies the stochastic differential equation

$$d\bar{z}_t = dw_t + \frac{1}{2}\bar{V}'(\bar{z}_t)\,dt$$

and thus, by Lemma 2.4,

$$
\begin{aligned}
H(P^V|P^U) &= \lim_{T\to\infty} \frac{1}{T} E^{P^V}\left[\frac{1}{2}\int_0^T (\bar{V}'-\bar{U}')(\bar{z}_t)\Big(dz_t - \frac{1}{2}\bar{U}'(\bar{z}_t)\,dt\Big)\right. \\
&\qquad\qquad\qquad\qquad\qquad \left. -\frac{1}{8}\int_0^T |\bar{V}'-\bar{U}'|^2(z_t)\,dt\right] \\
&= \lim_{T\to\infty} \frac{1}{T} E^{P^V}\left[\frac{1}{4}\int_0^T (\bar{V}'-\bar{U}')(\bar{z}_t)(\bar{V}'-\bar{U}')(\bar{z}_t)\,dt\right. \\
&\qquad\qquad\qquad\qquad\qquad \left. -\frac{1}{8}\int_0^T |\bar{V}'-\bar{U}'|^2(\bar{z}_t)\,dt\right] \\
&= \frac{1}{8}\lim_{T\to\infty} \frac{1}{T} E^{P^V}\left[\int_0^T |V'-U'|^2(z_t)\,dt\right] \\
&= \frac{1}{8}\int_{\mathbf{T}} |V'-U'|^2(x)m_V(dx).
\end{aligned}
$$

$\square$

For eahc $T>0$, let $\mathbf{L}^T:\Omega\to\mathcal{M}_1(\mathbf{T})$ be the empirical measure, i.e.,

$$
\mathbf{L}^T(w) = \frac{1}{T}\int_0^T \delta_{w_s}\,ds.
$$

The following large deviation estimate is the main ingredient for further discussions. The proof of it can be found in [DS, Lemma 6.3.7],

LEMMA 2.9. *For all $U\in\mathcal{K}$, J^U is a good convex rate function and for all measurable $\Gamma\subset M_1(\mathbf{T})$,*

$$
\begin{aligned}
-\inf_{\mu\in\Gamma^\circ} J^U(\mu) &\le \liminf_{T\to\infty} \frac{1}{T}\log P^U(w:\mathbf{L}^T(w)\in\Gamma) \\
&\le \limsup_{T\to\infty} \frac{1}{T}\log P^U(w:\mathbf{L}^T(w)\in\Gamma) \le -\inf_{\mu\in\bar{\Gamma}} J^U(\mu).
\end{aligned}
$$

§3. The ergodic theorem and the variational principle

We first state the ergodic theorem which is the starting point of the discussion about the asymptotic behavior of (1.1) and (1.2).

PROPOSITION 3.1. *We have*

(3.2)
$$
\lim_{T\to\infty} \frac{1}{T}\log \frac{dQ}{dP}\bigg|_{\mathcal{F}_{[0,T]}} = -\frac{1}{8}\int_{\mathbf{T}} |b'-a'|(x)^2 m_a(dx) = -H(P|Q), \quad P\text{-}a.s.
$$

and

(3.3)
$$\lim_{T\to\infty} \frac{1}{T} \log \frac{dQ}{dP}\bigg|_{\mathcal{F}_{[0,T]}} = \frac{1}{8} \int_{\mathbf{T}} |b'-a'|(x)^2 m_b(dx) = H(Q|P), \quad Q\text{-a.s.}$$

  Proof. Let
$$\phi_1^2(x) = \frac{dQ}{dP}\bigg|_{\mathcal{F}_{\{0\}}}(x) = e^{-(b-a)(x)}.$$

We use the same notations as the ones we used in the proof of Lemma 2.5. By the Girsanov formula,

$$(3.4) \quad \frac{dQ}{dP}\bigg|_{\mathcal{F}_{[0,T]}} = \phi_1(x_0)\phi_1(x_T)\exp\bigg(\frac{1}{2}\int_0^T (\bar{b}'-\bar{a}')(\bar{x}_t)\,dw_t$$
$$- \frac{1}{8}\int_0^T |\bar{b}'-\bar{a}'|^2(\bar{x}_t)\,dt\bigg)$$

and thus

$$(3.5) \quad \frac{1}{T}\log\frac{dQ}{dP}\bigg|_{\mathcal{F}_{[0,T]}} = \frac{1}{T}\log\phi_1(x_0)\phi_1(x_T)$$
$$+ \frac{1}{2T}\int_0^T (\bar{b}'-\bar{a}')(\bar{x}_t)\,dw_t - \frac{1}{8T}\int_0^T |\bar{b}'-\bar{a}'|^2(\bar{x}_t)\,dt.$$

The first term in (3.5) disappears as T goes to infinity. The second term is of the form M_T/T where M_T is the martingale with the quadratic variation process

$$\langle M\rangle_T = \frac{1}{4}\int_0^T |\bar{b}'-\bar{a}'|^2(\bar{x}_t)\,dt.$$

By the moment inequality for martingales,

$$\frac{1}{T}M_T \to 0, \quad \text{a.s.}$$

as $T\to\infty$. The third term converges, due to the above inequality, to

$$-\frac{1}{8}\int_{\mathbf{T}} |b'-a'|^2(x)m_a(dx), \quad P\text{-a.s.}$$

by Lemma 2.4. Hence we obtain (3.2). By replacing P and Q in (3.2), we easily obtain (3.3). □

Let, for any $\mu \in \mathcal{M}_1$,

$$\mathcal{H}(\mu) \equiv J^a(\mu) - J^b(\mu)$$

and for any $\xi \in \mathbf{R}$,

$$\Sigma_\xi \equiv \{\mu \in \mathcal{M}_1(\mathbf{T}), \mathcal{H}(\mu) = \xi\}.$$

By Lemma 2.5, for all $m_U \in \mathcal{M}_1(\mathbf{T})$ where $U \in \mathcal{K}$,

$$(3.6) \quad \mathcal{H}(m_U) = \int_{\mathbf{T}} \left\{ \frac{1}{4}(b' - a') \cdot (U' - a') - \frac{1}{8}|b' - a'|^2 \right\}(x) m_U(dx).$$

Now let $I_P : \mathbf{R} \to [0, \infty]$ and $I_Q : \mathbf{R} \to [0, \infty]$ be given by

$$(3.7) \qquad I_P(\xi) \equiv \begin{cases} \inf\{J^a(\mu) \, ; \, \mu \in \Sigma_\xi\}, & \text{if } \Sigma_\xi \neq \phi, \\ \infty, & \text{otherwise,} \end{cases}$$

and

$$(3.8) \qquad I_Q(\xi) \equiv \begin{cases} \inf\{J^b(\mu) \, ; \, \mu \in \Sigma_\xi\}, & \text{if } \Sigma_\xi \neq \phi, \\ \infty, & \text{otherwise.} \end{cases}$$

Then, we have the following simple relation between I_P and I_Q.

LEMMA 3.9.   (i)

$$(3.10) \qquad\qquad I_P(\xi) = 0 \quad \text{if and only if} \quad \xi = -H(P|Q)$$

and

$$(3.11) \qquad\qquad I_Q(\xi) = 0 \quad \text{if and only if} \quad \xi = H(Q|P).$$

(ii) *For all* $\xi \in \mathbf{R}$,

$$(3.12) \qquad\qquad I_Q(\xi) = I_P(\xi) - \xi.$$

Proof.   The relation (3.12) comes directly from the definitions of I_P and I_Q. If $I_P(\xi) = 0$, then, by the definition of I_P, there is a $\mu \in \mathcal{M}_1(\mathbf{T})$ such that $J^a(\mu) = 0$ and $J^b(\mu) = -\xi$, but by Lemma 2.3, this implies that the $\mu = m_a$ and thus $\xi = -J^b(m_a) = -H(P|Q)$ by Lemma 2.5. "Only if" part also follows by taking $\mu = m_a$. (3.11) follows by totally the same argument.   ☐

LEMMA 3.13.   *For each* $\xi \in \mathbf{R}$ *for which* $I_P(\xi) < \infty$, *there is a* $\mu_\xi \in \mathcal{M}_1(\mathbf{T})$ *such that*

$$(3.14) \qquad\qquad \mathcal{H}(\mu_\xi) = \xi \quad \text{and} \quad I_P(\xi) = J^a(\mu_\xi).$$

Proof. The assertion follows directly from the definition of $\mathcal{H}$ and I_P. See, e.g., [DS, Lemma 2.1.2]. $\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad$ ⬜

LEMMA 3.15. *I_P and I_Q are good, convex rate functions.*

Proof. We will only prove the statement for I_P. The lower semicontinuity follows applying [DS, Lemma 2.1.4]. To show that I_P is convex, it suffices to show that

$$I_P(\bar{\xi}) \leq \frac{1}{2}(I_P(\xi_1) + I_P(\xi_2))$$

for and $\xi_1 \in \mathbf{R}$ and $\xi_2 \in \mathbf{R}$, where $\bar{\xi} = (\xi_1 + \xi_2)/2$.

There is nothing to be done in the case $I_P(\xi_1) = \infty$ or $I_P(\xi_2) = \infty$ and so we assume both $I_P(\xi_1) < \infty$ and $I_P(\xi_2) < \infty$. By Lemma 3.13, there are $\mu_{\xi_1} \in \mathcal{M}_1(\mathbf{T})$ and $\mu_{\xi_2} \in \mathcal{M}_1(\mathbf{T})$ such that (3.14) holds. In this case we may assume that there are $U_1 \in \mathcal{K}$ and $U_2 \in \mathcal{K}$ such that $\mu_{\xi_1} = m_{U_1}$ and $\mu_{\xi_2} = m_{U_2}$, since we can approximate both $\mathcal{H}(\mu)$ and $J^a(\mu)$ by $\mathcal{H}(m_{U_n})$ and $J^a(m_{U_n})$ for some $\{U_n\} \subset \mathcal{K}$. (See e.g., [S, Lemma 8.20].)

Let $U_0 \in \mathcal{K}$ be given by

$$e^{-U_0(x)} = \frac{1}{2}\left(e^{-U_1(x)} + e^{-U_2(x)}\right).$$

Then, noting

$$U_0'(x)m_{U_0}(dx) = \frac{1}{2}\left(U_1'(x)m_{U_1}(dx) + U_2'(x)m_{U_2}(dx)\right),$$

we have, by (3.6) and (3.14),

$$
\begin{aligned}
\mathcal{H}(m_{U_0}) &= \int_{\mathbf{T}}\left[\frac{1}{4}(b'-a')(U_0'-a') - \frac{1}{8}|b'-a'|^2\right](x)m_{U_0}(dx) \\
&= \frac{1}{2}\int_{\mathbf{T}}\left[\frac{1}{4}(b'-a')(U_1'-a') - \frac{1}{8}|b'-a'|^2\right](x)m_{U_1}(dx) \\
&\quad + \frac{1}{2}\int_{\mathbf{T}}\left[\frac{1}{4}(b'-a')(U_2'-a') - \frac{1}{8}|b'-a'|^2\right](x)m_{U_2}(dx) \\
&= \frac{1}{2}(\mathcal{H}(m_{U_1}) + \mathcal{H}(m_{U_2})) = \bar{\xi}.
\end{aligned}
$$

On the other hand, since $\mu \mapsto J^a(\mu)$ is convex,

$$\frac{1}{2}(I_P(\xi_1) + I_P(\xi_2)) = \frac{1}{2}(J^a(m_{U_1}) + J^a(m_{U_2})) \geq J^a(m_{U_0}).$$

and thus, by the definition of I_P, we obtain

$$\frac{1}{2}(I_P(\xi_1) + I_P(\xi_2)) \geq I_P(\bar{\xi}).$$

Finally, we show that I_P has a compact level set. For this we only have to show I_P has a bounded level set, since it is lower semi-continuous. For this purpose we first observe that for each $L_2 > 0$, there is a $L_1 > 0$ such that $\{\mu \in \mathcal{M}_1(\mathbf{T}) \,;\, J^a(\mu) \leq L_2\} \subset \{\mu \in \mathcal{M}_1(\mathbf{T}) \,;\, J^b(\mu) \leq L_1\}$. To see this, we observe that, for $\mu(dx) = e^{-U(x)}dx \in \mathcal{M}_1(\mathbf{T})$,

$$
\begin{aligned}
J^b(\mu) &= \frac{1}{8}\int_{\mathbf{T}} |U' - b'|^2 m_b(dx) \\
&\leq \frac{1}{4}\int_{\mathbf{T}} |U' - a'|^2 m_b(dx) + \frac{1}{4}\int_{\mathbf{T}} |a' - b'|^2 m_b(dx) \\
&\leq 2C_1 J^a(\mu) + C_2,
\end{aligned}
$$

where $C_1 = \sup_{x \in \mathbf{T}} e^{-(b-a)}(x)$ and $C_2 = \frac{1}{4}\int_{\mathbf{T}} |a' - b'|^2 m_b(dx)$. Thus,

$$\mathcal{H}(\mu) = J^a(\mu) - J^b(\mu) \geq -J^b(\mu) \geq -C_1 J^a(\mu) - C_2,$$

and so

$$J^a(\mu) \geq -c_1 \mathcal{H}(\mu) - c_2$$

for some $c_1 > 0$ and $c_2 > 0$. Therefore,

$$I_P(\xi) \geq -c_1 \xi - c_2.$$

On the other hand, since $J^a(\mu) \geq \xi$ for μ for which $\mathcal{H}(\mu) = \xi$, we see that

$$I_P(\xi) \leq \xi.$$

By these, we have showed that I_P has a bounded level set. □

LEMMA 3.16.   *For every* $v \in [0,1]$, $\lim_{T \to \infty}(1/T) \log E^P\big[\exp\big(v \log \frac{dQ}{dP}\big|_{\mathcal{F}_{[0,T]}}\big)\big]$ *exists and if we define* $\lambda_P : \mathbf{R} \to [0,\infty]$ *by*

(3.17) $\quad \lambda_P(v) \equiv \begin{cases} \infty, & \text{if } v < 0, \\[2mm] \displaystyle\lim_{T \to \infty} \frac{1}{T} \log E^P\Big[\exp\Big(v \log \frac{dQ}{dP}\Big|_{\mathcal{F}_{[0,T]}}\Big)\Big], & \text{if } 0 \leq v \leq 1, \\[2mm] \infty, & \text{if } v > 1, \end{cases}$

then

(3.18) $$\lambda_P(v) = \sup_{\xi \in \mathbf{R}}(v \cdot \xi - I_P(\xi))$$

for all $v \in \mathbf{R}$.

Remark. By (3.18) and the Lemma 3.15, we obtain

$$(3.19) \qquad I_P(\xi) = \sup_{v \in \mathbf{R}} (v \cdot \xi - \lambda_P(v)).$$

applying the general result on the convex analysis (See [DS, Theorem 2.2.15]).

Proof. First, note that

$$\sup_{\xi \in \mathbf{R}} (v \cdot \xi - I_P(\xi)) = \sup_{\xi \in \mathbf{R},\, \mu \in \mathcal{M}_1(\mathbf{T})} \{v \cdot \xi - J^a(\mu),\, J^a(\mu) - J^b(\mu) = \xi\}$$

$$= \sup_{\mu \in \mathcal{M}_1(\mathbf{T})} \{v(J^a(\mu) - J^b(\mu)) - J^a(\mu)\}$$

$$= - \inf_{\mu \in \mathcal{M}_1(\mathbf{T})} \{(1 - v)J^a(\mu) + vJ^b(\mu)\},$$

and thus for $v < 0$ or for $v > 1$,

$$\sup_{\xi \in \mathbf{R}} (v \cdot \xi - I_P(\xi)) = \infty$$

since $\sup_{\mu \in \mathcal{M}_1(\mathbf{T})} J^a(\mu) = \sup_{\mu \in \mathcal{M}_1(\mathbf{T})} J^b(\mu) = \infty$.

For each $v \in [0,1]$, let $P^v = P^{a+v(b-a)-c}$ where the constant c is given by $e^c = \int_{\mathbf{T}} e^{-(a+v(b-a))(x)} dx$. By the Girsanov formula, adopting the notation as the one used in the proof of Lemma 2.5,

$$\frac{dP^v}{dP}\bigg|_{\mathcal{F}_{[0,T]}} (x.) = \phi_v(x_0)\phi_v(x_T) \exp\left(\frac{v}{2} \int_0^T (\bar{b}' - \bar{a}')(\bar{x}_t)\, dw_t\right.$$

$$\left. - \frac{v^2}{8} \int_0^T |\bar{b}' - \bar{a}'|^2(\bar{x}_t)\, dt\right)$$

where

$$\phi_v^2(x) = \frac{dP^v}{dP}\bigg|_{\mathcal{F}_{\{0\}}} (x),$$

and thus, by (3.4),

$$E^P\left[\exp\left(v \log \frac{dQ}{dP}\bigg|_{\mathcal{F}_{[0,T]}}\right)\right]$$

$$= E^P\left[\phi_1(x_0)^v \phi(x_T)^v \exp\left(\frac{v}{2} \int_0^T (\bar{b}' - \bar{a}')(\bar{x}_t)\, dw_t\right.\right.$$

$$- \frac{v}{8} \int_0^T |\bar{b}' - \bar{a}'|^2 (\bar{x}_t) \, dt \Bigg) \Bigg]$$

$$= E^{P^v} \Bigg[\phi_1(x_0)^v \phi_v(x_0)^{-1} \phi_1(x_T)^v \phi_v(x_T)^{-1}$$

$$\times \exp \Bigg(\frac{1}{8} (v^2 - v) \int_0^T |b' - a'|^2 (x_t) dt \Bigg) \Bigg],$$

and so

$$\lambda_P(v) = \lim_{T \to \infty} \frac{1}{T} \log E^{P^v} \Bigg[\exp \Bigg(\frac{1}{8} (v^2 - v) \int_0^T |b' - a'|^2 (x_t) \, dt \Bigg) \Bigg]$$

$$= \lim_{T \to \infty} \frac{1}{T} \log E^{P^v} \Bigg[\exp \Bigg(\frac{1}{8} (v^2 - v) \int |b' - a'|^2 \, d\mathbf{L}_T(x.) \Bigg) \Bigg].$$

Thus, as a consequence of Lemma 2.9, by the Laplace-type asymptotic formula (See [DS, Theorem 2.1.10]), denoting $J^v = J^{a+v(b-a)-c}$,

$$\lambda_P(v) = \sup_{\mu \in \mathcal{M}_1(\mathbf{T})} \Bigg\{ \frac{1}{8} (v^2 - v) \int_{\mathbf{T}} |b' - a'|^2 (x) \mu(dx) - J^v(\mu) \Bigg\}$$

$$= \sup_{U \in \mathcal{K}} \Bigg\{ \frac{1}{8} \int_{\mathbf{T}} \{ (v^2 - v)|b' - a'|^2 - |U' - a' - v(b' - a')|^2 \}(x) m_U(dx) \Bigg\}$$

$$= \sup_{U \in \mathcal{K}} \Bigg\{ v \int \Big(\frac{1}{4} (b' - a')(U' - a') - \frac{1}{8} |b' - a'|^2 \Big)(x) m_U(dx)$$

$$- \frac{1}{8} \int |U' - a'|(x)^2 m_U(dx) \Bigg\}$$

$$= \sup_{\xi \in \mathbf{R}} (v \cdot \xi - I_P(\xi)).$$

$$\qquad \qquad \qquad \qquad \qquad \qquad \qquad \qquad \qquad \qquad \qquad \square$$

§4. Large deviations and the proof of the main theorem

In this section we give the large deviation estimate for the error probabilities (1.1) and (1.2) and then give the proof of the main Theorem.

THEOREM 4.1.   (i) *For any Borel set* $\Gamma \subset \mathbf{R}$,

$$(4.2) \quad - \inf\{ I_P(\xi) \, ; \, \xi \in \Gamma^\circ \} \leq \liminf_{T \to \infty} \frac{1}{T} \log P \Bigg(\frac{1}{T} \log \frac{dQ}{dP} \bigg|_{\mathcal{F}_{[0,T]}} \in \Gamma \Bigg)$$

$$\leq \limsup_{T \to \infty} \frac{1}{T} \log P \Bigg(\frac{1}{T} \log \frac{dQ}{dP} \bigg|_{\mathcal{F}_{[0,T]}} \in \Gamma \Bigg)$$

$$\leq - \inf\{ I_P(\xi) \, ; \, \xi \in \bar{\Gamma} \}.$$

(ii) *For any Borel set $\Gamma \subset \mathbf{R}$,*

$$(4.3) \quad -\inf\{I_Q(\xi)\,;\,\xi \in \Gamma^\circ\} \leq \liminf_{T \to \infty} \frac{1}{T} \log Q\left(\frac{1}{T} \log \frac{dQ}{dP}\bigg|_{\mathcal{F}_{[0,T]}} \in \Gamma\right)$$

$$\leq \limsup_{T \,to\infty} \frac{1}{T} \log Q\left(\frac{1}{T} \log \frac{dQ}{dP}\bigg|_{\mathcal{F}_{[0,T]}} \in \Gamma\right)$$

$$\leq -\inf\{I_Q(\xi)\,;\,\xi \in \bar{\Gamma}\}.$$

The following Propositions give the proof of (4.2). The proof of (4.3) goes totally in the same way.

PROPOSITION 4.4. *For any closed set $F \subset \mathbf{R}$,*

$$(4.5) \quad \limsup_{T \to \infty} \frac{1}{T} \log P\left(\frac{1}{T} \log \frac{dQ}{dP}\bigg|_{\mathcal{F}_{[0,T]}} \in F\right) \leq -\inf\{I_P(\xi)\,;\,\xi \in F\}.$$

Proof. Let $\kappa = J^a(m_b) = H(Q|P)$. First, we assume that $F \subset (\kappa, \infty)$ and set $\xi_* = \inf\{\xi\,;\,\xi \in F\}$. Then

$$P\left(\frac{1}{T} \log \frac{dQ}{dP}\bigg|_{\mathcal{F}_{[0,T]}} \in F\right) \leq P\left(\frac{1}{T} \log \frac{dQ}{dP}\bigg|_{\mathcal{F}_{[0,T]}} \in [\xi_*, \infty)\right)$$

$$\leq e^{-T \cdot v \cdot \xi_*} E^P\left[\exp\left(v \log \frac{dQ}{dP}\bigg|_{\mathcal{F}_{[0,T]}}\right)\right],$$

and thus

$$\frac{1}{T} \log P\left(\frac{1}{T} \log \frac{dQ}{dP}\bigg|_{\mathcal{F}_{[0,T]}} \in F\right) \leq -v \cdot \xi_* + \frac{1}{T} \log E^P\left[\exp\left(v \log \frac{dQ}{dP}\bigg|_{\mathcal{F}_{[0,T]}}\right)\right]$$

for all $0 \leq v \leq 1$. Here, note that $E^P\left[\exp\left(v \log \frac{dQ}{dP}\big|_{\mathcal{F}_{[0,T]}}\right)\right]$ exists for all $v \in [0,1]$. Hence, noting (3.17),

$$\limsup_{T \to \infty} \frac{1}{T} \log P\left(\frac{1}{T} \log \frac{dQ}{dP}\bigg|_{\mathcal{F}_{[0,T]}} \in F\right) \leq -\inf_{0 \leq v \leq 1} (v \cdot \xi_* - \lambda_P(v))$$

$$= -\inf_{v \in \mathbf{R}} (v \cdot \xi_* - \lambda_P(v))$$

and thus, by (3.19), we get Proposition 4.4 in this case.

Secondly, in the case $F \subset (-\infty, \kappa)$, by the same argument, we get (4.5). Finally in the case that neither $F_1 = F \cap (-\infty, \kappa)$ nor $F_2 = F \cap (\kappa, \infty)$ is empty, we only have to notice that

$$\limsup_{T \to \infty} \frac{1}{T} \log P \left(\frac{1}{T} \log \frac{dQ}{dP} \bigg|_{\mathcal{F}_{[0,T]}} \in F \right)$$

$$\leq \limsup_{T \to \infty} \frac{1}{T} \log \left[P \left(\frac{1}{T} \log \frac{dQ}{dP} \bigg|_{\mathcal{F}_{[0,T]}} \in F_1 \right) \vee P \left(\frac{1}{T} \log \frac{dQ}{dP} \bigg|_{\mathcal{F}_{[0,T]}} \in F_2 \right) \right]$$

$$\leq -(I_P(\xi_1) \wedge I_P(\xi_2)),$$

where $\xi_2 = \sup\{\xi \,;\, \xi \in F_1\}$ and $\xi_2 = \inf\{\xi \,;\, \xi \in F_2\}$ since, by Lemma 3.9), I_P is nonincreasing on $(-\infty, \kappa)$ and nondecreasing on (κ, ∞). □

PROPOSITION 4.6. *For any open set $G \subset \mathbf{R}$,*

$$(4.7) \quad \liminf_{T \to \infty} \frac{1}{T} \log P \left(\frac{1}{T} \log \frac{dQ}{dP} \bigg|_{\mathcal{F}_{[0,T]}} \in G \right) \geq -\inf\{I_P(\xi) \,;\, \xi \in G\}.$$

Proof. It is sufficient to prove that for any $\xi \in \mathbf{R}$ such that $I_P(\xi) < \infty$ and any neighborhood N of ξ,

$$\liminf_{T \to \infty} \frac{1}{T} \log P(A_T) \geq -I_P(\xi).$$

for $A_T = \{(1/T) \log \frac{dQ}{dP}\big|_{\mathcal{F}_{[0,T]}} \in N\}$. Since $I_P(\xi) < \infty$, by Lemma 3.13, there is a $\mu_\xi \in \mathcal{M}_1(\mathbf{T})$ such that (3.14) holds. Again, we only have to show in the case $\mu_\xi = m_{U_\xi}$ for some $U_\xi \in \mathcal{K}$.

Since $\frac{dQ}{dP}\big|_{\mathcal{F}_{[0,T]}}$ is given by (3.4),

$$(4.8) \quad \frac{1}{T} E^{P^{U_\xi}} \left[\log \frac{dQ}{dP} \bigg|_{\mathcal{F}_{[0,T]}} \right]$$

$$= \frac{1}{T} E^{P^{U_\xi}} \left[\frac{1}{2} \int_0^T (\bar{b}' - \bar{a}')(\bar{z}_t) \, dw_t - \frac{1}{8} \int_0^T |\bar{b}' - \bar{a}'|^2(\bar{z}_t) \, dt \right] + \mathcal{R}$$

where $\mathcal{R}$ is the term which decays as $T \to \infty$. Since P^{U_ξ} is the law of the solution of the stochastic differential equation

$$d\bar{z}_t = dw_t + \frac{1}{2} \bar{U}'_\xi(\bar{z}_t) \, dt,$$

by (4.8) and (3.14),

$$\lim_{T\to\infty} \frac{1}{T} E^{P^{U_\xi}} \left[\log \frac{dQ}{dP} \Big|_{\mathcal{F}_{[0,T]}} \right]$$

$$= \int_{\mathbf{T}} \left[\frac{1}{4}(b'-a')\cdot(U'_\xi-a') - \frac{1}{8}|b'-a'|^2 \right](x)m_{U_\xi}(dx)$$

$$= \xi.$$

Therefore, by the ergodic theorem,

$$(4.9) \qquad\qquad P^{U_\xi}(A_T) \to 1 \quad \text{as} \quad T \to \infty.$$

By Jensen's inequality,

$$P(A_T) = E^{P^{U_\xi}} \left[\frac{dP}{dP^{U_\xi}} \Big|_{\mathcal{F}_{[0,T]}} ; A_T \right]$$

$$= E^{P^{U_\xi}} \left[\frac{dP}{dP^{U_\xi}} \Big|_{\mathcal{F}_{[0,T]}} \Big| A_T \right] P^{U_\xi}(A_T)$$

$$\geq \exp\left(E^{P^{U_\xi}} \left[\log \frac{dP}{dP^{U_\xi}} \Big|_{\mathcal{F}_{[0,T]}} \Big| A_T \right] \right) P^{U_\xi}(A_T)$$

$$= \exp\left(E^{P^{U_\xi}} \left[\log \frac{dP}{dP^{U_\xi}} \Big|_{\mathcal{F}_{[0,T]}}, A_T \right] P^{U_\xi}(A_T)^{-1} \right) P^{U_\xi}(A_T).$$

Thus, by (4.9),

$$\liminf_{T\to\infty} \frac{1}{T} \log P(A_T) \geq \lim_{T\to\infty} E^{P^{U_\xi}} \left[\frac{1}{T} \log \frac{dP}{dP^{U_\xi}} \Big|_{\mathcal{F}_{[0,T]}} \right]$$

$$= -H(P^{U_\xi}|P),$$

and the Proposition follows noting Lemma 2.5 and (3.14). □

Proof of the Main Theorem. Since the rate function I_P is increasing on $(-H(P|Q), \infty)$, for $\alpha > 0$, there is a $\xi(\alpha) > -H(P|Q)$ such that $\inf_{t\in[\xi(\alpha),\infty)} I_P(t) \leq \alpha \leq \inf_{t\in(\xi(\alpha),\infty)} I_P(t)$. Hence, (i) follows from Theorem 4.1. It is also easy to see, by (3.11) that $\xi(\alpha)$ is increasing in $\alpha > 0$ and $\xi(\alpha) \geq -H(P|Q)$. By (3.11), $I_Q(\xi_0) = 0$ for $\xi_0 = H(Q|P)$ and thus by (3.12), $I_P(\xi_0) = \xi_0$, which concludes that $\xi(\xi_0) = \xi_0$. Hence for $0 \leq \alpha \leq H(Q|P)$, $\xi(\alpha) \leq H(Q|P)$ and so

$$I_Q(\xi(\alpha)) = \inf\{I_Q(\xi) ; \xi \leq \xi(\alpha)\}$$

since $I_Q(\xi)$ is non-increasing for $\xi \leq H(Q|P)$. For $\alpha \geq H(Q|P)$, $\xi(\alpha) \geq H(Q|P)$ and so

$$I_Q(\xi(\alpha)) = \inf\{I_Q(\xi) \,;\, \xi \geq \xi(\alpha)\},$$

since $I_Q(\xi)$ is non-decreasing in ξ for $\xi \geq H(Q|P)$. These conclude (1.4) and (1.5). □

References

[D]   R. Durrett, Stochastic Calculus, A Practical Introduction, CRC Press, New York, 1996.

[DS]   J-D. Deuschel and D. W. Stroock, Large Deviations, Academic Press, San Diego, 1989.

[HK]   T. S. Han and K. Kobayashi, *The strong converse theorem in hypothesis testing*, IEEE Trans. Inform. Theory, IT-**35** (1989), 178–180.

[IS]   S. Ihara and Y. Sakuma, *Signal Detection in White Gaussian Channel*, Proc. Seventh Japan-Russian Sym. Prob. Theory and Math. Stat., World Scientific, Singapore (1996), 147–156.

[IW]   N. Ikeda and S. Watanabe, Stochastic Differential Equations and Diffusion Processes, Second Edition, North Holland, Amsterdam, 1989.

[NK1]   K. Nakagawa and F. Kanaya, *On the converse theorem in statistical hypothesis testing*, IEEE Trans. Inform. Theory, IT-**39** (1993), 623–628.

[NK2]   K. Nakagawa and F. Kanaya, *On the converse theorem in statistical hypothesis testing for Markov chains*, IEEE Trans. Inform. Theory, IT-**39** (1993), 629–633.

[S]   D. W. Stroock, An Introduction to the Theory of Large Deviations, Springer, New York, 1984.

Grauduate School of Mathematics
Nagoya University
Nagoya, 464-8602
Japan