Probab. Theory Relat. Fields 116, 181–220 (2000)

Mark Freidlin · Shuenn-Jyi Sheu

Diffusion processes on graphs: stochastic differential equations, large deviation principle

Received: 12 February 1997 / Revised version: 3 March 1999

Abstract. Ito's rule is established for the diffusion processes on the graphs. We also consider a family of diffusions processes with small noise on a graph. Large deviation principle is proved for these diffusion processes and their local times at the vertices.

1. Introduction

Let Γ be a connected graph consisting of vertices $O_1, \ldots, O_M$ and edges $I_1, \ldots, I_N$ connecting the vertices. Write $I_i \sim O_k$ if O_k is one of the end points of I_i. Some of the edges may just have only one end point. We assume that the graph is imbedded in an Euclidian space so that any two edges can only have intersection at a vertex. A coordinate y_i is chosen in I_i. In terms of this coordinate, we assume that a second order elliptic operator L_i is defined on I_i,

$$L_i = \frac{\sigma_i^2(y_i)}{2}\frac{d^2}{dy_i^2} + b_i(y_i)\frac{d}{dy_i} \quad .$$

It is well known, under some general conditions, L_i defines uniquely (distribution-wise) a diffusion process $X^{(i)}(t)$ in I_i up to the first time it reaches the end points of I_i (see [24]). To define the process on Γ, which coincides with $X^{(i)}(t)$ in I_i, one should prescribe the behavior of the process after it reaches the vertices. In [18, 19], all possible continuous Markov processes on Γ with Feller property governed by L_i inside I_i were described. It is shown that there exists a one to one correspondence between such processes and the sets of nonnegative constants α_k, α_{ki}, for each O_k and $I_i \sim O_k$, such that

$$\sum_{I_i \sim O_k} \alpha_{ki} + \alpha_k > 0 \quad .$$

Namely, given a set of such constants, an operator A can be defined on a subset of continuous functions F on Γ satisfying

M. Freidlin: Department of Mathematics, University of Maryland, College Park, Maryland 20742, USA

S.-J. Sheu: Institute of Mathematics, Academia Sinica, Nankang, Taipei, Taiwan 11529, ROC. e-mail: sheusj@math.sinica.edu.tw

Mathematics Subject Classification (1991): Primary 60J60; Secondary 60H10, 60J55, 60F10

Key words and phrases: Diffusions on graphs – Local time – Small random perturbation – Large deviations

$$\alpha_k AF(O_k) = \sum_{I_i \sim O_k} \alpha_{ki} D_i F(O_k), \quad k = 1, \dots, M \ ,$$

such that AF is continuous on Γ and equal to $L_i F_i$ inside I_i. Here F_i is the restriction of F on I_i and

$$D_i F(O_k) \equiv \lim_{X \in I_i^0 : X \to O_k} \frac{F(X) - F(O_k)}{|y_i(X) - y_i(O_k)|} \ .$$

Then the corresponding process is the Markov process generated by A.

Skew Brownian motion ([25]) and Walsh's Process ([38]) are interesting examples for the processes introduced above. See also [3] and an extensive references there. Diffusion processes on graphs arise naturally in models of physical processes such as electrical network, nerve impulse propagation, etc. ([15, 20, 21, 30, 32]). Such processes also appear in a number of limiting theorems for classical processes ([7, 16, 18, 19]). Many related interesting mathematical problems has been studied in the literature ([4, 5, 22, 28, 30, 33]).

In this paper, we consider the particular case that $\alpha_k = 0$ and $\alpha_{ki} > 0$ if $I_i \sim O_k$. We shall develop stochastic analysis for such processes. We also give a useful description for them. Of course, inside each I_i, the process is governed by a classical stochastic differential equation, but at the vertices we face some new effects. We derive Ito's differential rule for the processes. These are given in Sections 2 and 3. In Section 4, we consider a family of diffusion processes on graph with small noise and study the large deviation properties of these processes and their local times at the vertices. Some applications are given in Section 5 which in particular implies some of the results obtained in [8, 26]. Finally, we remark that the analysis presented in this paper allows us to consider, for example, the Dirichlet problem for second order elliptic operators on graphs with a small second order coefficients, as well as wave front propagation for reaction-diffusion on graphs. The last problem will be studied in detail elsewhere.

2. Stochastic calculus, stochastic equation for the process

In this section, we shall introduce notations, terminology which will be used for what follows.

Let Γ be a connected graph, as described in Section 1, consisting of vertices, $O_1, \dots, O_M$, and edges, $I_1, \dots, I_N$. A coordinate y_i is chosen on I_i such that I_i is homeomorphic to an interval of real line using this coordinate; I_i^0 is the interior of I_i. We say $I_i \sim O_k$ if I_i has O_k as one of its boundary (end points); I_i has only one end if and only if I_i is homeomorphic to $[0, \infty)$.

Let $C_b(\Gamma)$ be the space of bounded continuous functions on Γ. We say that a continuous function F defined on Γ is in $C_b^\infty(\Gamma)$ if it is bounded and continuous on I_i^0 together with all its derivatives, which have natural extension to the ends of I_i if I_i is homeomorphic to a bounded interval.

Let L_i be an operator given by

$$L_i f(X) = \frac{1}{2}\sigma_i^2(y_i)\frac{d^2 f}{dy_i^2}(y_i) + b_i(y_i)\frac{df}{dy_i}(y_i), \quad X \in I_i, \quad y_i(X) = y_i \ . \tag{2.1}$$

We assume that σ_i and b_i, $i = 1, \ldots, N$, are functions in $C_b^\infty(I_i^0)$. We also assume that σ_i, $i = 1, \ldots, N$, are uniformly strictly positive.

Let $I_i \sim O_k$. For a function F defined on a neighborhood of O_k, we denote

$$D_i F(O_k) = \lim_{X \in I_i^0, X \to O_k} \frac{F(X) - F(O_k)}{|y_i(X) - y_i(O_k)|}$$

if the limit exists.

Let $\alpha_{ki} > 0$ be given for $I_i \sim O_k$. We define a linear operator A on $C_b(\Gamma)$ by

$$AF(X) = L_i F_i(X), \quad X \in I_i \ , \tag{2.2}$$

for F in the domain $D(A)$ which consists of functions F in $C_b^\infty(\Gamma)$ satisfying

$$\rho(F)(O_k) \equiv \frac{1}{\sum_{i:I_i \sim O_k} \alpha_{ki}} \sum_{i:I_i \sim O_k} \alpha_{ki} D_i F(O_k) = 0 \tag{2.3}$$

for each O_k and $AF \in C_b(\Gamma)$. Here F_i is the restriction of F on I_i. We shall call (2.3) the gluing condition in the sequel. The following result taken from [18] gives the existence and uniqueness of a Markov process on Γ with the generator A . We mention that the distribution for Walsh's process is given in [3].

Theorem 2.1. *The operator A generates a Feller Markov process on Γ with continuous sample paths. The operator L_i and the gluing condition at the vertices define such a process in a unique way* (*in the sense of distribution*).

We know that this Markov process inside I_i is the diffusion process generated by L_i. In the following, we shall try to describe the behavior of this process in the neighborhood of O_k. This is a local property. Therefore, in the rest of this section, we shall assume that Γ has only one vertex O and is given by

$$\Gamma = \{(x, i); \quad x \geq 0, \ i = 1, \ldots, N\} \ . \tag{2.4}$$

Here we identify $(0, i), i = 1, \ldots, N$, and call it O. We shall use F to denote a function on Γ and define

$$F_i(x) = F(x, i), \quad x \geq 0 \ .$$

We say $F \in C_b^\infty(\Gamma)$ if $F_i \in C_b^\infty((0, \infty))$ for each i and $F_i(0)$, $i = 1, \ldots, N$, have the same value.

The operator L_i can be considered for functions on $[0, \infty)$ and is given by

$$L_i f(x) = \frac{1}{2}\sigma_i^2(x)\frac{d^2 f}{dx^2}(x) + b_i(x)\frac{df}{dx}(x), \quad x > 0 \tag{2.5}$$

for f defined on $[0, \infty)$. Here σ_i, b_i are in $C_b^\infty((0,\infty))$ for each i. We assume that there is a $c_0 > 0$ such that

$$\sigma_i(x) > c_0 \text{ for all } i \text{ and } x \geq 0 \ .$$

Let the constants α_i, $i = 1, \dots, N$, be positive. We may assume that

$$\sum_{i=1}^N \alpha_i = 1 \ .$$

For $F \in C_b^\infty(\Gamma)$, define

$$\rho(F) = \sum \alpha_i \frac{dF_i}{dx}(0) \ . \tag{2.6}$$

The operator A, defined according to (2.2) (2.3), generates a Markov process and is denoted by $X(t) = (x(t), i(t))$, $t \geq 0$. Note that if $x(t) = 0$, $i(t)$ can be chosen in a nonunique way. Let, for the definition, put $i(t) = 1$ in this case. Since all $(0, i)$ are identified, the choice of $i(t)$ is not essential when $x(t) = 0$.

Lemma 2.2. *There is a* 1*-dim Wiener process* $W(t)$ *and a continuous increasing process* $\ell(t)$, *which are measurable with respect to* $\mathscr{F}_t$, *the* σ*-field generated by* $X(s)$, $s \leq t$, *such that*

$$dx(t) = \sigma_{i(t)}(x(t))dW(t) + b_{i(t)}(x(t))dt + d\ell(t) \ . \tag{2.7}$$

Here $\ell(\cdot)$ *increases only when* $x(t) = 0$.

Proof. For each $\delta > 0$, let us define the stopping times $\tau_n^\delta, \theta_n^\delta$, $n = 0, 1, 2, \dots$, as follows.

$$\begin{aligned}
\theta_0^\delta &= 0 \\
\tau_0^\delta &= \inf\{t > 0; \quad x(t) = 0\} \\
\theta_1^\delta &= \inf\{t > \tau_0^\delta; \quad x(t) = \delta\} \\
&\vdots \\
\tau_n^\delta &= \inf\{t > \theta_n^\delta; \quad x(t) = 0\} \\
\theta_{n+1}^\delta &= \inf\{t > \tau_n^\delta; \quad x(t) = \delta\} \ .
\end{aligned}$$

We define a function H on Γ by

$$H(x, i) = H_i(x) = \int_0^x \frac{1}{\sigma_i(u)} du, \quad x > 0 \ . \tag{2.8}$$

We consider the following decomposition,

$$\begin{aligned}
H(X(t)) = H(X(0)) &+ \sum \left(H\left(X\left(\tau_n^\delta \wedge t\right)\right) - H\left(X\left(\theta_n^\delta \wedge t\right)\right)\right) \\
&+ \sum \left(H\left(X\left(\theta_{n+1}^\delta \wedge t\right)\right) - H\left(X\left(\tau_n^\delta \wedge t\right)\right)\right) \ .
\end{aligned} \tag{2.9}$$

Denote

$$W^\delta(t) = \sum \xi_n^\delta(t)$$

with

$$\xi_n^\delta(t) = H\left(X\left(\tau_n^\delta \wedge t\right)\right) - H\left(X\left(\theta_n^\delta \wedge t\right)\right)$$

$$-\int_{\theta_n^\delta \wedge t}^{\tau_n^\delta \wedge t} \left(\tfrac{1}{2}\sigma_{i(s)}^2(x(s))\frac{d^2 H_{i(s)}}{d^2 x}(x(s)) + b_{i(s)}(x(s))\frac{dH_{i(s)}}{dx}(x(s))\right) ds \quad .$$

It is easy to see from the Ito's formula that $W^\delta(t)$ is a martingale with the quadratic variation

$$\sum \left(\tau_n^\delta \wedge t - \theta_n^\delta \wedge t\right)$$

which converges to t as δ tends to 0. Here we use the property

$$\lim_{\delta \to 0} E\left\{\int_0^t \chi_{\{x(s)\le\delta\}} ds\right\} = 0 \tag{2.10}$$

and the relation

$$\sum_{\theta_n^\delta \le t} \left(\theta_n^\delta - \tau_{n-1}^\delta\right) \le \int_0^t \chi_{\{x(s)\le\delta\}} ds \quad .$$

We remark that (2.10) can be proved using the relation

$$E\left\{\int_0^t \chi_{\{x(s)\le\delta\}} ds\right\} \le E\left\{F^\delta(X(t)) - F^\delta(X(0))\right\} \quad ,$$

where $F^\delta \in C_b^\infty(\Gamma)$ satisfies

$$L_i F_i^\delta(x) = \alpha(x),$$

$$F_i^\delta(0) = \frac{dF_i^\delta}{dx}(0) = 0$$

with $\alpha(x)$ a function such that $0 \le \alpha(x) \le 1, \alpha(x) = 1$ on $[0, \delta]$ and 0 if $x \ge 2\delta$. Indeed, solving the equation we get

$$F_i^\delta(x) = \int_0^x \exp\left(-\int_0^z \frac{2b_i(u)}{\sigma_i^2(u)} du\right) \int_0^z \frac{2\alpha(y)}{\sigma_i^2(y)} \exp\left(\int_0^y \frac{2b_i(v)}{\sigma_i^2(v)} dv\right) dy\, dz \quad .$$

By the assumption on $\alpha(\cdot)$, we can show

$$\left|F_i^\delta(x)\right| \le c\delta x$$

for some $c > 0$. The assertion follows from this.

Therefore, W^δ converges to a Wiener process $W(t)$ as δ tends to 0. By using (2.10), we can also prove

$$\sum \int_{\theta_n^\delta \wedge t}^{\tau_n^\delta \wedge t} \left(\frac{1}{2}\sigma_{i(s)}^2(x(s)) \frac{d^2 H_{i(s)}}{d^2 x}(x(s)) + b_{i(s)}(x(s)) \frac{d H_{i(s)}}{dx}(x(s)) \right) ds$$

$$\to \int_0^t \left(\frac{1}{2}\sigma_{i(s)}^2(x(s)) \frac{d^2 H_{i(s)}}{d^2 x}(x(s)) + b_{i(s)}(x(s)) \frac{d H_{i(s)}}{dx}(x(s)) \right) ds, \quad \text{as } \delta \to 0 \ .$$

From these results,

$$\sum (H(X(\theta_{n+1}^\delta \wedge t)) - H(X(\tau_n^\delta \wedge t)))$$

also converges to a process $\tilde{\ell}(t)$ as δ tends to 0. But since

$$\sum_n (H(X(\theta_{n+1}^\delta \wedge t)) - H(X(\tau_n^\delta \wedge t))) - \sum_{n,i} H(\delta, i) \chi_{\left\{\theta_{n+1}^\delta \le t, i(\theta_{n+1}^\delta) = i\right\}}$$

and

$$\sum_{n,i} H(\delta, i) \chi_{\{\theta_{n+1}^\delta \le t, i(\theta_{n+1}^\delta) = i\}} - \sum_{n,i} \frac{1}{\sigma_i(0)} \delta \chi_{\left\{\theta_{n+1}^\delta \le t, i(\theta_{n+1}^\delta) = i\right\}}$$

both tend to 0 in mean as $\delta \to 0$, $\tilde{\ell}(t)$ should be an increasing process and it increases only when $x(t) = 0$.

Now, consider

$$x(t) = x(0) + \sum (x(\tau_n^\delta \wedge t) - x(\theta_n^\delta \wedge t)) + \sum (x(\theta_{n+1}^\delta \wedge t) - x(\tau_n^\delta \wedge t)) \ .$$

We observe that

$$\sum (x(\tau_n^\delta \wedge t) - x(\theta_n^\delta \wedge t)) = \int_0^t \sigma_{i(s)}(x(s)) dW^\delta(s) + \sum \int_{\theta_n^\delta \wedge t}^{\tau_n^\delta \wedge t} b_{i(s)}(x(s)) ds$$

which converges to

$$\int_0^t \sigma_{i(s)}(x(s)) dW(s) + \int_0^t b_{i(s)}(x(s)) ds$$

as $\delta \to 0$ by using (2.10). On the other hand,

$$\sum (x(\theta_{n+1}^\delta \wedge t) - x(\tau_n^\delta \wedge t)) - \sum \delta \chi_{\{\theta_{n+1}^\delta \le t\}}$$

tends to 0 in mean as $\delta \to 0$ and $\sum \delta \chi_{\{\theta_{n+1}^\delta \le t\}}$ is an increasing process. Therefore,

$$\sum (x(\theta_{n+1}^\delta \wedge t) - x(\tau_n^\delta \wedge t))$$

also converges to an increasing process $\ell(t)$ satisfying the desired property. This completes the proof.

Lemma 2.3. (Ito's formula) *Let $W(t), \ell(t)$ be as in Lemma* 2.2. *Assume $F \in C_b^\infty(\Gamma)$. Then*

$$F(X(t)) = F(X(0)) + \int_0^t \sigma_{i(s)}(x(s))\frac{dF_{i(s)}}{dx}(x(s))dW_s + \int_0^t AF(X(s))ds + \rho(F)\ell(t) \ .$$

Proof. For each $\delta > 0$, let the stopping times $\tau_n^\delta, \theta_n^\delta$ be defined as in the proof of Lemma 2.2. We use a decomposition for $F(X(t))$ similar to (2.9) and a similar arguments. To show the result we only need to show that

$$\sum_{n,i} \frac{dF_i}{dx}(0)\delta\chi_{\{\theta_{n+1}^\delta \le t, i(\theta_{n+1}^\delta)=i\}} \tag{2.11}$$

converges to $\rho(F)\ell(t)$ as $\delta \to 0$, by noting that

$$\sum_n (F(X(\theta_{n+1}^\delta \wedge t)) - F(X(\tau_n^\delta \wedge t))) - \sum_{n,i} \frac{dF_i}{dx}(0)\delta\chi_{\{\theta_{n+1}^\delta \le t, i(\theta_{n+1}^\delta)=i\}} \to 0$$

as $\delta \to 0$.

By the strong Markovian property of the process we see that $i(\theta_n^\delta)$, $n = 0, 1, \ldots$, are $i.i.d.$ with distribution

$$P\{i(\theta_n^\delta) = k\} = \beta_k^\delta \ .$$

Let β_k be a limit of β_k^δ, $k = 1, \ldots, N$, for a subsequence of δ that tends to 0. Then, by the independence of $i(\theta_n^\delta)$, $n = 0, 1, \ldots$, we can show that (2.11) converges to $\tilde{\rho}(F)\ell(t)$, where

$$\tilde{\rho}(F) = \sum \beta_i \frac{dF_i}{dx}(0) \ .$$

Therefore,

$$F(X(t)) = F(X(0)) + \int_0^t \sigma_{i(s)}(x(s))\frac{dF_{i(s)}}{dx}(x(s))dW_s + \int_0^t AF(X(s))ds + \tilde{\rho}(F)\ell(t) \ . \tag{2.12}$$

By the Markovian property of the process $X(t)$, it is known that

$$F(X(t)) - \int_0^t AF(X(s))ds$$

is a martingale if $F \in D(A)$. See [14]. This together with (2.12) imply that $\tilde{\rho}(F) = 0$ if $\rho(F) = 0$. Then we can deduce that $\beta_k = \alpha_k$, $k = 1, \ldots, N$. Therefore, β_k^δ converges to α_k, $k = 1, \ldots, N$, as δ tends to 0. Also, the equation

$$F(X(t)) = F(X(0)) + \int_0^t \sigma_{i(s)}(x(s)) \frac{dF_{i(s)}}{dx}(x(s))dW_s + \int_0^t AF(X(s))ds + \rho(F)\ell(t) \ .$$

holds as claimed.

The above argument also proves the following result.

Corollary 2.4. *Let* θ^δ, $\delta > 0$, *be the stopping time defined by*

$$\theta^\delta = \inf\{t > 0; x(t) = \delta\} \ .$$

Then

$$\lim_{\delta \to 0} P_O\{i(\theta^\delta) = i\} = \alpha_i \quad \textit{for all } i \ .$$

Remark 2.5. Let $\delta > 0$. Define

$$F_i(x) = \begin{cases} \int_0^x \exp\left(\int_z^\delta \frac{2b_i(u)}{\sigma_i^2(u)} du\right) \int_z^\delta \frac{2}{\sigma_i^2(y)} \exp\left(-\int_y^\delta \frac{2b_i(u)}{\sigma_i^2(u)} du\right) dy\, dz & \text{if } x \le \delta \\ 0 \quad \text{otherwise} \end{cases}$$

Applying Lemma 2.3 to this function and letting $\delta \to 0$ we can prove that

$$\sum \alpha_i \frac{2}{\sigma_i^2(0)} \ell(t) = \lim_{\delta \to 0} \frac{1}{\delta} \int_0^t \chi_{\{x(s) \le \delta\}} ds$$

holds in mean. That is, $\ell(\cdot)$ is the occupation density of $x(\cdot)$ ($X(\cdot)$) at 0 (O) up to a constant factor.

3. Some basic transformations

In this section, we assume that our graph Γ has one vertex and is given by (2.4). Let $X(t) = (x(t), i(t))$ be the process on Γ generated by the operator (2.5) with gluing condition $\rho(F) = 0$. Here $\rho(F)$ is defined by (2.6). By Lemma 2.2, $X(t)$ satisfies the equation (2.7). However, simple example, such as the case with $\sigma_i(\cdot) = 1$, and $b_i(\cdot) = 0$ for all i, shows that (2.7) alone is not enough to determine $X(\cdot)$. It is interesting to see if a complete set of equations exists, including (2.7), which can be used to describe the process. Basically, what we need is a random selection of $i(t)$ on each excursion from O. In the case $n = 2$, [23] provides another equation to describe the process. It is not known if this idea can be generalized to $n > 2$. However, it is shown in [35] that for Walsh's process, the filtration generated by the process is not the same as the filtration generated by the Brownian motion in Lemma 2.2. In the following we shall introduce some basic transformations, including space transformations and the transformations by the change of probability measure resembling the Girsanov transformation. We think that these are useful for studying the problems related to the process. In Section 4, we shall apply these to the study of large deviations of small perturbed diffusions on graphs.

For the space transformation, let $c_1(x), \ldots, c_N(x)$ be bounded, strictly positive, smooth functions on $(0, \infty)$ with bounded derivatives. Define

$$G_i(x) = \int_0^x c_i(u)du \ ,$$

and

$$y(t) = G(X(t)) = G_{i(t)}(x(t)) \ .$$

We denote by $Y(t)$ the process $(y(t), i(t))$ which is a continuous process on Γ.

Lemma 3.1. *The process $Y(t)$ defined above is the diffusion generated by $\hat{A}$ with the gluing condition $\hat{\rho}(F) = 0$. Here*

$$\hat{A}F(y,i) = \frac{1}{2}\hat{\sigma}_i^2(y)\frac{d^2F_i}{dy^2}(y) + \hat{b}_i(y)\frac{dF_i}{dy}(y), \quad y > 0 \ ,$$

$$\hat{\sigma}_i(y) = \sigma_i(G_i^{-1}(y)) \cdot c_i(G_i^{-1}(y)) \ ,$$

$$\hat{b}_i(y) = AG(G_i^{-1}(y), i)$$

$$= \frac{1}{2}\sigma_i^2(G_i^{-1}(y))\frac{dc_i}{dx}(G_i^{-1}(y)) + b_i(G_i^{-1}(y))c_i(G_i^{-1}(y)) \ ,$$

$$\hat{\rho}(F) = \frac{1}{\sum \alpha_i c_i(0)} \sum \alpha_i c_i(0)\frac{dF_i}{dy}(0) \ ,$$

where G_i^{-1} is the inverse of G_i. Let $\ell, \hat{\ell}$ be the local time of $X(t), Y(t)$ at 0. Then

$$\hat{\ell}(t) = \sum \alpha_i c_i(0)\ell(t) = \rho(G)\ell(t).$$

Proof. Let F be a function in $C_b^\infty(\Gamma)$. We may apply Lemma 2.3 to $F(G(X(t)))$ to get

$$F(Y(t)) = F(Y(0)) + \int_0^t \hat{\sigma}_{i(t)}(y(t))\frac{dF_{i(t)}}{dy}(y(t))dW(t) + \int_0^t \hat{A}F(Y(t))dt + \hat{\rho}(F)\hat{\ell}(t) \ .$$

This implies that

$$F(Y(t)) - \int_0^t \hat{A}F(Y(t))dt$$

is a martingale if $\hat{\rho}(F) = 0$. That is, the distribution generated by $Y(\cdot)$ is a solution of martingale problem. Then by applying Theorem 2.1 to $\hat{A}F$, $\hat{\rho}(F)$ together with the result in [14] (Chapter 4, Theorem 4.1), we know that the solution of the martingale problem is unique. Therefore, $Y(\cdot)$ is the process generated by $\hat{A}$ with gluing condition $\hat{\rho}(F) = 0$. The rest is easy and the detail is omitted.

Now we consider the Girsanov type transformation. Let $e_i(x), i = 1, \ldots, N$, defined on $(0, \infty)$, be bounded smooth functions with bounded derivatives.

Lemma 3.2. *Assume that the process $X(t)$ is defined on a probability space $(\Omega, \mathscr{F}, P)$, $\mathscr{F}_t$ is the σ-field generated by $X(s), s \le t$, and $W(t)$, given in Lemma 2.2., is a Wiener process adapted to $\mathscr{F}_t$. We define a probability measure Q on $(\Omega, \mathscr{F})$ such that on $(\Omega, \mathscr{F}_T)$ we have*

$$\frac{dQ}{dP} = \exp\left(\int_0^T e_{i(t)}(x(t))dW(t) - \frac{1}{2}\int_0^T |e_{i(t)}(x(t))|^2 dt\right) .$$

Then under Q, $X(t)$ is the diffusion process generated by the operator

$$\hat{A}F(x,i) = AF(x,i) + \sigma_i(x)e_i(x)\frac{dF_i}{dx}(x)$$

with the gluing condition $\rho(F) = 0$.

Proof. We know that under Q,

$$\hat{W}(t) = W(t) - \int_0^t e_{i(s)}(x(s))ds$$

is a Wiener process. Therefore under Q,

$$F(X(t)) - \int_0^t \hat{A}F(X(s))ds$$

is a martingale if $\rho(F) = 0$. The rest follows by using the same argument as in the proof of Lemma 3.1.

Remark 3.3. By Lemma 3.1, it is not difficult to see that $\bar{\ell}(t)$, $\bar{\rho}(F)$ defined as

$$\bar{\ell}(t) = \sum \alpha_i \frac{1}{\sigma_i(0)}\ell(t), \quad \bar{\rho}(F) = \frac{1}{\sum \alpha_i \frac{1}{\sigma_i(0)}}\rho(F)$$

are independent of the local coordinate $\{y_i\}$ chosen on Γ. It seems nature to replace $\ell(t)$ by $\bar{\ell}(\cdot)$. Then the Ito's formula in Lemma 2.3 still holds if we replace $\ell(t)$ and $\rho(F)$ by $\bar{\ell}(t)$ and $\bar{\rho}(F)$.

4. Small perturbed diffusion processes on Γ and their large deviations

In this section, we shall consider a family of diffusion processes $X^\varepsilon(\cdot)$ on the graph satisfying stochastic differential equations with small noise depending on ε. We will study their large deviation properties. The large deviation properties for a family of stochastic processes or probability distributions has been well studied. We shall briefly review the definition in the following. We refer to [17, 37] for more details and some basic properties.

Let $\mathcal{X}$ be a metric space with metric ρ and Z^h denote an $\mathcal{X}$-valued random variable for each $h > 0$. Let $\lambda(h)$ be a positive real-valued function and tends to ∞ as h tends to 0. And let $S(x)$ be a function on $\mathcal{X}$ assuming values in $[0, \infty]$. The sequence Z^h as $h \to 0$ is said to satisfy the large deviation principle with action functional $\lambda(h)S(x)$ if the following assertions hold:

(i) the set $\Phi(s) = \{x \in \mathcal{X}; S(x) \leq s\}$ is compact for every $s \geq 0$;

(ii) for any $\delta > 0$, $\gamma > 0$ and $x \in \mathcal{X}$ there is $h_0 > 0$ such that

$$P\{\rho(Z^h, x) < \delta\} \geq \exp(-\lambda(h)(S(x) + \gamma))$$

for all $h \leq h_0$;

(iii) for any $\delta > 0$, $\gamma > 0$ and $s > 0$ there is an $h_0 > 0$ such that

$$P\{\rho(Z^h, \Phi(s)) \geq \delta\} \leq \exp(-\lambda(h)(s - \gamma))$$

for $h \leq h_0$.

The functions $S(x)$ and $\lambda(h)$ are called the normalized action functional and normalizing coefficient.

The following is an equivalent definition which is frequently used:

(i) the set $\Phi(s) = \{x \in \mathcal{X}; S(x) \leq s\}$ is compact for every $s \geq 0$;

(ii) $\liminf_{h\to 0} \lambda(h)^{-1} \log P\{Z^h \in A\} \geq -\inf_{x\in A^0} S(x)$;

(iii) $\limsup_{h\to 0} \lambda(h)^{-1} \log P\{Z^h \in A\} \leq -\inf_{x\in \bar{A}} S(x)$;

Here A is any Borel subset of $\mathcal{X}$, $\bar{A}$ and A^0 are the closure and interior of A.

The following is another equivalent definition which we shall adopt in this paper.

(i) the set $\Phi(s) = \{x \in \mathcal{X}; S(x) \leq s\}$ is compact for every $s \geq 0$;

(ii) for any $r > 0$ there is a compact set K of $\mathcal{X}$ such that for any $\delta > 0$ there is an $h_0 > 0$ such that

$$P\{\rho(Z^h, K) \geq \delta\} \leq \exp(-\lambda(h)r)$$

for $h \leq h_0$;

(iii) for any $x \in \mathcal{X}$,

$$-S(x) = \lim_{\delta\to 0} \liminf_{h\to 0} \lambda(h)^{-1} \log P\{\rho(Z^h, x) < \delta\}$$

$$= \lim_{\delta\to 0} \limsup_{h\to 0} \lambda(h)^{-1} \log P\{\rho(Z^h, x) < \delta\} .$$

We remark that (i) and (ii) are easier to verify. The main effort is to prove (iii)

In the rest, we shall adopt the notations in Section 2. Let Γ be a connected graph with set of vertices, $O_1, \dots, O_M$, and edges, $I_1, \dots I_N$. On each I_i, a proper coordinate y_i is chosen. The process $X^\varepsilon(\cdot)$ has generator A^ε which is given by (2.2) and (2.3) with L_i replaced by L_i^ε,

$$L_i^\varepsilon f(X) = \frac{\varepsilon}{2}\sigma_i^2(y_i)\frac{d^2 f}{dy_i^2}(y_i) + b_i(y_i)\frac{df}{dy_i}(y_i), \quad X \in I_i, \quad y_i(X) = y_i \ . \quad (4.1)$$

for each I_i. Here σ_i and b_i satisfy the conditions in Section 2. We note that it is also interesting to consider the case where α_{ki} depends on ε for all i and k, $I_i \sim O_k$. However, here we assume α_{ki} to be independent of ε. Denote $\ell_k^\varepsilon(\cdot)$ the local time of $X^\varepsilon(\cdot)$ at O_k defined in Lemma 2.2.

For any two points X_1, X_2 in Γ, the distance $|X_1 - X_2|$ is defined as $|y_i(X_1) - y_i(X_2)|$ if X_1, X_2 are in I_i. Otherewise, it is defined as the minimum of $\sum_{i=1}^K |Y_{i+1} - Y_i|$ over all those Y_i satisfying that $Y_1 = X_1$, $Y_{K+1} = X_2$ and Y_i, Y_{i+1} for each i are in the same I_j for some j. For any two continuous functions $\Phi_1, \Phi_2 : [0, T] \to \Gamma$, we define

$$\|\Phi_1 - \Phi_2\|_T = \sup_{0 \le t \le T} \{|\Phi_1(t) - \Phi_2(t)|\} \ . \quad (4.2)$$

The limiting behavior of $X^\varepsilon(t)$ as $\varepsilon \to 0$ in general is not obvious. However, in this section we are able to establish the large deviation properties for these processes. This in turn can be used to obtain the limiting behavior of these processes. Indeed, because of some applications such as the one which will be mentioned in Section 5, we shall study the large deviation properties for the coupled processes $(X^\varepsilon(t), \ell_k^\varepsilon(t), k = 1, \dots, M)$, considered as a family of continuous processes on $[0, T]$ for a fixed $T > 0$ with sample paths in $\mathscr{C}$. Here

$$\begin{aligned}\mathscr{C} = \{&(\Phi, \eta_1, \dots, \eta_M); \Phi : [0, T] \to \Gamma,\ \eta_k : [0, T] \to R \text{ are continuous} \\ &\text{such that } \eta_k(0) = 0 \text{ and } \eta_k(\cdot) \text{ is nondecreasing for } k = 1, \dots, M.\}\end{aligned} \quad (4.3)$$

We equip $\mathscr{C}$ with uniform topology as described above.

It will become clear later (see the calculation in the proof of Theorem 4.5.) that the normalized action functional for the couple $(X^\varepsilon(\cdot), \ell_k^\varepsilon(\cdot), k = 1, \dots, M)$ can be calculated involving the large time behavior of some diffusion described by (2.4), (2.5) and (2.6), maybe with different σ_i, b_i and α_i. For this reason, we shall first present some results concerning the large time asymptotics for such processes.

Let Γ be given by (2.4) and $\alpha_i > 0$, $c_i > 0$, $i = 1, \dots, N$, be fixed constants. We also assume

$$\sum_{i=1}^N \alpha_i = 1 \ .$$

We consider the diffusion process $X(t) = (x(t), i(t))$ on Γ with generator

$$AF(x,i) = \frac{1}{2}\frac{d^2F_i}{dx^2}(x) - c_i\frac{dF_i}{dx}(x), \ \ x > 0, \ \ i = 1, \cdots, N \ \ , \tag{4.4}$$

satisfying the gluing condition $\rho(F) = 0$ with $\rho(F)$ given by (2.6).

Lemma 4.1. *Let the process $X(t)$ be defined as above. Then $X(t)$ has the invariant density*

$$p_i(x) = a_i \exp(-2c_i x), \quad x > 0, \quad i = 1, \ldots, N \ \ ,$$

with

$$a_i = 2\frac{1}{\sum\frac{\alpha_i}{c_i}}\alpha_i \ \ .$$

Moreover,

$$\frac{1}{T}\int_0^T \chi_{\{i(t)=i\}}dt \to \frac{a_i}{2c_i}$$

in probability as $T \to \infty$ for all i.

In particular, when $c_i = 1$ for all i, define

$$\tau = \inf\{t > 0; x(t) = 0\} \ \ .$$

Then,

$$P_x\{\tau > t\} \le cxe^x\frac{1}{t^{\frac{3}{2}}}e^{-\frac{1}{2}t} \tag{4.5}$$

for some constant c. The process $x(t)$ is also a one dimensional diffusion process. Let $p_t(x, y)$ be the transition density of the process $x(t)$. We have,

$$\begin{aligned} p_t(x,y) = {} & e^{-2y}(\frac{2}{\pi t})^{\frac{1}{2}} e^{-\frac{1}{2t}(x+y-t)^2} + e^{-2y}(\frac{2}{\pi t})^{\frac{1}{2}}\int_x^\infty e^{-\frac{1}{2t}(b+y-t)^2}db \\ & +(\frac{1}{2\pi t})^{\frac{1}{2}}e^{-\frac{1}{2}t}e^{x-y}[e^{-\frac{1}{2t}(x-y)^2} - e^{-\frac{1}{2t}(x+y)^2}] \ \ . \end{aligned} \tag{4.6}$$

Let $X = (x, i)$, $Y = (y, j)$. Then, $P_t(X, Y)$, the transition density of the process $X(t)$, is given by

$$P_t(X,Y) = \alpha_j p_t(x,y) - \alpha_j\left(\frac{1}{2\pi t}\right)^{\frac{1}{2}} e^{-\frac{1}{2}t}e^{x-y}\left[e^{-\frac{(x-y)^2}{2t}} - e^{-\frac{(x+y)^2}{2t}}\right]$$

if $i \neq j$ and

$$P_t(X,Y) = \alpha_j p_t(x,y) + (1-\alpha_j)\left(\frac{1}{2\pi t}\right)^{\frac{1}{2}} e^{-\frac{1}{2}t}e^{x-y}\left[e^{-\frac{(x-y)^2}{2t}} - e^{-\frac{(x+y)^2}{2t}}\right]$$

if $i = j$.

We postpone the proof of these results to Appendix 2.

Lemma 4.2. *Let $X(t) = (x(t), i(t))$ be the process in Lemma 4.1. with $c_i = 1$ for all i. Then, for each $V = (v_1, \cdots, v_N)$, we have*

$$\lim_{\delta\to 0} \limsup_{T\to\infty} \frac{1}{T} \log E_O \left\{ \exp\left(-\int_0^T v_{i(t)} dt \right) ; |W(T)| \le \delta T, x(T) \le \delta T \right\}$$

$$= \lim_{\delta\to 0} \liminf_{T\to\infty} \frac{1}{T} \log E_O \left\{ \exp\left(-\int_0^T v_{i(t)} dt \right) ; |W(T)| \le \delta T, x(T) \le \delta T \right\}$$

This limit depends on V and is denoted by $-\hat{\Lambda}(V)$. We have the following expression for $\hat{\Lambda}(V)$,

$$\hat{\Lambda}(V) = \inf \left\{ \sum v_i \beta_i + J(\beta) \right\} \ , \tag{4.7}$$

$$J(\beta) = \frac{1}{2} \left(\sum \frac{\alpha_i^2}{\beta_i} - 1 \right) = \frac{1}{2} \sum \beta_i (\frac{\alpha_i}{\beta_i} - 1)^2 \ ,$$

where the inf *is taken over all $\beta = (\beta_1, \ldots, \beta_N)$ satisfying*

$$\beta_i > 0, \quad i = 1, \ldots, N; \quad \sum \beta_i = 1 \ . \tag{4.8}$$

Proof. Denote

$$\mu_i(T) = \frac{1}{T} \int_0^T \chi_{\{i(t)=i\}} dt, \quad i = 1, \ldots, N, \ \mu(T) = (\mu_1(T), \ldots, \mu_N(T)) \ .$$

It is enough to prove the following result,

$$\lim_{\delta\to 0} \limsup_{T\to\infty} \tfrac{1}{T} \log E_O\{|\mu(T) - \beta| \le \delta, |W(T)| \le \delta T, x(T) \le \delta T\}$$

$$= \lim_{\delta\to 0} \liminf_{T\to\infty} \tfrac{1}{T} \log E_O\{|\mu(T) - \beta| \le \delta, |W(T)| \le \delta T, x(T) \le \delta T\} \tag{4.9}$$

$$= -J(\beta)$$

for each $\beta = (\beta_1, \ldots, \beta_N)$ satisfying (4.8), since $\int_0^T v_{i(t)} dt = T \sum v_i \mu_i(T)$.

To prove this, we use the idea of changing the probability measure. We consider a new Markov process $\hat{X}(t) = (\hat{x}(t), \hat{i}(t))$ with generator (4.4) and satisfying the gluing condition $\rho(F) = 0$. Here $c_i, i = 1, \ldots, N$, are positive and shall be determined later. By Lemma 2.2, there is a one dimensional Wiener process $\hat{W}(t)$ such that

$$d\hat{x}(t) = -c_{\hat{i}(t)} dt + d\hat{W}(t) + d\hat{\ell}(t) \ , \tag{4.10}$$

$\hat{\ell}(t)$ is the local time of $\hat{X}(t)$ at O.

Denote,

$$\hat{\mu}_i(T) = \frac{1}{T}\int_0^T \chi_{\{\hat{i}(t)=i\}} dt, \quad \hat{\mu}(T) = (\hat{\mu}_1(T), \ldots, \hat{\mu}_N(T)) \ ,$$

and

$$\xi(t) = \int_0^t (c_{\hat{i}(s)} - 1) ds \ .$$

Then, by using Lemma 3.2, we have

$$\begin{aligned} &E_O\{|\mu(T) - \beta| \le \delta, |W(T)| \le \delta T, x(T) \le \delta T\} \\ &= E_O\left\{\exp\left(\int_0^T (c_{\hat{i}(t)} - 1) d\hat{W}(t) - \frac{1}{2}\int_0^T (c_{\hat{i}(t)} - 1)^2 dt\right); \right. \qquad (4.11) \\ &\qquad \left. |\hat{\mu}(T) - \beta| \le \delta, |\hat{W}(T) - \xi(T)| \le \delta T, \hat{x}(T) \le \delta T\right\} \ . \end{aligned}$$

We claim that (4.11) has the following as an upper and lower bound respectively,

$$\begin{aligned} &\exp(\pm c\delta)\exp\left(-T\left(\frac{1}{2}\sum (c_i - 1)^2\beta_i - \sum c_i(c_i - 1)\beta_i\right)\right) \\ &E_O\{\exp(F(\hat{X}(T)) - \rho(F)\hat{\ell}(T)); |\hat{\mu}(T) - \beta| \le \delta, \qquad (4.12) \\ &|\hat{W}(T) + \xi(T)| \le \delta T, \hat{x}(T) \le \delta T\} \ , \end{aligned}$$

c is a constant and F is given by

$$F(x, i) = (c_i - 1)x, \quad x > 0, \ i = 1, \ldots, N \ .$$

To prove (4.12), by Lemma 2.3,

$$F(\hat{X}(T)) = \int_0^T (c_{\hat{i}(t)} - 1) d\hat{W}(t) - \int_0^T c_{\hat{i}(t)}(c_{\hat{i}(t)} - 1) dt + \rho(F)\hat{\ell}(T) \ ,$$

where $\rho(F) = \sum \alpha_i(c_i - 1)$. Now assume $|\hat{\mu}(T) - \beta| \le \delta$ and δ is small. Then using

$$\int_0^T (c_{\hat{i}(t)} - 1)^2 dt = T\sum (c_i - 1)^2 \hat{\mu}_i(T)$$

we have

$$\left|\sum (c_i - 1)^2 \hat{\mu}_i(T) - \sum (c_i - 1)^2 \beta_i\right| \le c\delta \ ,$$

Similarly, under the same condition, we have

$$\left|\int_0^T c_{\hat{i}(t)}(c_{\hat{i}(t)} - 1) dt - T\sum c_i(c_i - 1)\beta_i\right| \le c\delta T$$

and

$$|\xi(T) - T(\sum(c_i - 1)\beta_i)| \le c\delta T \ .$$

These imply (4.12). Next we show how to choose c_i, $i = 1, \dots, N$.

The choice of c_i is to ensure that the following properties hold,

$$\begin{aligned} &(1)\ \sum(c_i - 1)\beta_i = 0, \\ &(2)\ \hat{\mu}(T) \to \beta, \quad \text{in probability as } T \to \infty \ . \end{aligned} \tag{4.13}$$

To see this, we assume (4.13), then

$$|\hat{\mu}(T) - \beta| \le \delta, \quad |\hat{W}(T) + \xi(T)| \le \delta T, \ \hat{x}(T) \le \delta T$$

imply

$$|\hat{\mu}(T) - \beta| \le \delta, \quad |\hat{W}(T)| \le c\delta T, \ \hat{x}(T) \le \delta T$$

Here, and also in the following, we use c to denote a positive constant which maybe different from place to place. Then together with (4.10), we have

$$\left|\hat{\ell}(T) - T\sum c_i\beta_i\right| \le c\delta T \ .$$

This implies that the following are upper and lower bounds for (4.11),

$$\begin{aligned} \exp(\pm c\delta T)\exp\Bigg(-T\Bigg(&\frac{1}{2}\sum(c_i - 1)^2\beta_i - \sum c_i(c_i - 1)\beta_i \\ &+ \sum(c_i - 1)\alpha_i \cdot \sum c_i\beta_i\Bigg)\Bigg) \ . \end{aligned} \tag{4.14}$$

Here we use the fact that

$$P_O\{|\hat{\mu}(T) - \beta| \le \delta, \quad |\hat{W}(T)| \le \delta T, \quad \hat{x}(T) \le \delta T\} \to 1$$

which can be proved by using (4.13) and Lemma 4.1. In the following we shall show that the only $\{c_i\}_{i=1}^N$ satisfying (4.13) is given by

$$c_i = \frac{\alpha_i}{\beta_i}, \quad i = 1, \dots, N \ . \tag{4.15}$$

Then with this particular choice, (4.14) is easily shown to be equal to $\exp(-TJ(\beta) \pm c\delta T)$, with $J(\beta)$ given by (4.7). This completes the proof of (4.9). It remains to prove that (4.15) implies (4.13).

By Lemma 4.1, the invariant density for the process $\hat{X}(t)$ is given by

$$p_i(x) = a_i \exp(-2c_i x) \ ,$$

$$a_i = 2\frac{1}{\sum \frac{\alpha_i}{c_i}}\alpha_i \tag{4.16}$$

for $x > 0, i = 1, \ldots, N$. We also have,

$$\hat{\mu}_i(T) \to \frac{a_i}{2c_i}, \quad i = 1, \ldots, N \ ,$$

in probability as $T \to 0$. Therefore, (4.13)(2) implies $\beta_i = \frac{a_i}{2c_i}$. Together with (4.16), we have

$$c_i = c\frac{\alpha_i}{\beta_i}, \quad i = 1, \ldots, N \ ,$$

for some $c > 0$. Furthermore, using the relation (4.13)(1) we can easily deduce that $c = 1$. This gives precisely (4.15).

Corollary 4.3. *Let $X(t) = (x(t), i(t))$ be the process in Lemma* 4.1. *with $c_i = 1$ for all i. We define the function $\hat{\Lambda}(V)$ for $V = (v_1, \ldots, v_N)$ as in Lemma* 4.2. *Then we have,*

$$\lim_{\delta\to 0} \limsup_{T\to\infty} \frac{1}{T} \log \sup_{|X|\leq\delta T} E_X \left\{ \exp\left(- \int_0^T v_{i(t)} dt \right); |W(T)| \leq \delta T, \right.$$
$$\left. x(T) \leq \delta T \right\} = -\hat{\Lambda}(V) \ .$$

Proof. We use exactly the same argument as in the proof of Lemma 4.2. The detail is omitted.

The following result, although will not be used in the rest, is of independent interest. It can be proved mimicing the above argument. We shall omit the detail.

Lemma 4.4. *Let $X(t) = (x(t), i(t))$ be the process as defined in Lemma* 4.2. *Then*

$$\lim_{T\to\infty} \frac{1}{T} \log E_O \left\{ \exp\left(- \int_0^T v_{i(t)} dt \right) \right\}$$

exists for $V = (v_1, \ldots, v_N)$. Let denote $-\tilde{\Lambda}(V)$ for the limit. Then

$$\tilde{\Lambda}(V) = \inf \left\{ \sum v_i \beta_i + \tilde{J}(\beta) \right\} \ ,$$

$$\tilde{J}(\beta) = \frac{1}{2} \left(1 - \frac{1}{\sum \frac{\alpha_i^2}{\beta_i}} \right)$$

where the inf *is taken over all $\beta = (\beta_1, \ldots, \beta_N)$ satisfying*

$$\beta_i > 0, \ i = 1, \ldots, N, \ \sum \beta_i = 1 \ .$$

We begin our discussion on the large deviation properties for the processes determined by (4.1) with $\varepsilon > 0$. We shall first consider the special case that Γ is given by (2.4). The operator A^ε is defined by (2.2) and (2.3) with L_i being replaced by L_i^ε,

$$L_i^\varepsilon f(x) = \frac{\varepsilon}{2}\frac{d^2}{dx^2} f(x) + b_i(x)\frac{d}{dx} f(x), \quad x > 0 \ . \tag{4.1$'$}$$

We again assume

$$\sum \alpha_i = 1 \ . \tag{4.17}$$

Let $X^\varepsilon(t) = (x^\varepsilon(t), i^\varepsilon(t))$ be the diffusion on Γ generated by A^ε and $\ell^\varepsilon(t)$ be the local time of this process at $O \equiv (0, i)$. The coupled process $(X^\varepsilon(t), \ell^\varepsilon(t))$ restricted to $[0, T]$ has the sample paths in the following space,

$$\begin{aligned}\mathscr{C} = \{(\Phi, \eta);\ \Phi : [0, T] \to \Gamma,\ \ \eta : [0, T] \to R \ \text{ are continuous}\\ \text{such that } \eta(0) = 0 \text{ and } \eta(\cdot) \text{ is nondecreasing}\} \ .\end{aligned} \tag{4.3$'$}$$

We shall first describe the action functional for the large deviation properties of the coupled processes $(X^\varepsilon(\cdot), \ell^\varepsilon(\cdot))$.

Given $X_0 = (x_0, i_0) \in \Gamma$. Let $\Phi(t) = (\phi(t), j(t))$, $0 \le t \le T$, be a continuous function with value on Γ such that $\Phi(0) = X_0$. Let $\eta(t)$, $0 \le t \le T$, be continuous and nondecreasing, $\eta(0) = 0$, $\eta(\cdot)$ be constant on intervals where $\phi(\cdot)$ is away from 0. We also assume that $\phi(\cdot), \eta(\cdot)$ are absolutely continuous such that

$$\int_0^T \left|\frac{d}{dt}\phi(t)\right|^2 dt < \infty, \quad \int_0^T \left|\frac{d}{dt}\eta(t)\right|^2 dt < \infty \ . \tag{4.18}$$

The functional $I_T(\Phi, \eta)$ is defined by

$$\begin{aligned}I_T(\Phi, \eta) = \frac{1}{2}\int_0^T \left|\frac{d\phi}{dt}(t) - b_{j(t)}(\phi(t))\right|^2 \chi_{\{\Phi(t)\neq O\}} dt\\ + \int_0^T \Lambda\left(\frac{d\eta}{dt}(t)\right) \chi_{\{\Phi(t)=O\}} dt \ ,\end{aligned} \tag{4.19}$$

where

$$\Lambda(\theta) = \frac{1}{2}\inf\left\{\sum \beta_i\left(\theta\frac{\alpha_i}{\beta_i} + b_i(0)\right)^2;\ 0 < \beta_i < 1,\ \sum \beta_i = 1\right\} \ . \tag{4.20}$$

For other (Φ, η), we define $I_T(\Phi, \eta) = \infty$.

Now we can state our main result.

Theorem 4.5. *Let $X^\varepsilon(\cdot), \ell^\varepsilon(\cdot)$ be the processes defined as above. For each $T > 0$, the family of continuous processes*

$$\left(X^\varepsilon(t), \ell^\varepsilon(t)\right), \quad 0 \le t \le T, \ \ \varepsilon > 0 \ ,$$

satisfies the large deviation principle in the uniform topology with action functional $\varepsilon^{-1} I_T(\cdot, \cdot)$.

Remark 4.6. Let define

$$L^{(i)}(x,p) = \frac{1}{2}|p - b_i(x)|^2, \quad x \ge 0, \ i = 1, \dots, N \ .$$

It is well known that

$$\varepsilon^{-1} \cdot \int_0^T L^{(i)}\left(\phi(t), \frac{d\phi}{dt}(t)\right) dt$$

is the large deviation action functional for the processes $X^\varepsilon(\cdot)$ on $[0,T]$ before exit from $\{(x,i); x > 0\}$. Then $\Lambda(\theta)$ can be expressed in terms of $L^{(i)}(x,p)$ as follows.

$$\Lambda(\theta) = \inf \left\{ \sum \beta_i L^{(i)}(0,p_i);\ 0 < \beta_i < 1, \right. \\ \left. \sum \beta_i = 1,\ p_i = -\theta \frac{\alpha_i}{\beta_i},\ i = 1, \dots, N \right\} \ .$$

Moreover, by the contraction principle (see [17, 37]), the processes $X^\varepsilon(t), t \in [0,T], \varepsilon > 0$, satisfy large deviation properties with the action functional $\varepsilon^{-1} I_T^0(\Phi)$,

$$I_T^0(\Phi) = \int_0^T L\left(\Phi(t), \frac{d}{dt}\phi(t)\right) dt \ ,$$

where $\Phi(t) = (\phi(t), j(t))$,

$$L((x,i),p) = L^{(i)}(x,p), \qquad \text{if } x \neq 0 \ ,$$

and

$$L(O,0) = \inf \left\{ \sum \beta_i L^{(i)}(0,p_i);\ \theta \ge 0,\ 0 < \beta_i < 1, \right. \\ \left. \sum \beta_i = 1,\ p_i = -\theta \frac{\alpha_i}{\beta_i}, i = 1, \dots, N \right\} \ .$$

We remark that the expression of this normalized action functional is closely related to the one given in [11] for a family of stochastic processes with a feature similar to ours. See also [9, 13].

The following gives the basic properties of $I_T(\cdot,\cdot)$ as a normalized large deviation action functional. Its proof will be postponed to Appendix 1.

Lemma 4.7.

(i) $I_T(\cdot,\cdot)$ *is lower semicontinuous in the uniform topology.*
(ii) *For any* $r > 0$, $\{(\Phi,\eta) \in \mathscr{C};\ I_T(\Phi,\eta) \le r, \Phi(0)$ *belongs to a compact set*$\}$ *is compact.*

Proof of Theorem 4.5. We mention first the main steps of the proof.

We shall prove the following "local" large deviation properties, (4.21) and (4.22). Here $(\Phi(\cdot), \eta(\cdot))$ is given and satisfies the condition (4.18). For any $\delta > 0$,

$$\liminf_{\varepsilon \to 0} \varepsilon \log P_{X_0} \left\{ \|X^\varepsilon - \Phi\|_T \le \delta, \|\ell^\varepsilon - \eta\|_T \le \delta \right\} \ge -I_T(\Phi,\eta) \ . \qquad (4.21)$$

And for any $\rho > 0$, there is $\delta > 0$ such that

$$\limsup_{\varepsilon\to 0} \varepsilon \log P_{X_0} \left\{ \|X^\varepsilon - \Phi\|_T \le \delta, \|\ell^\varepsilon - \eta\|_T \le \delta \right\} \le -I_T(\Phi,\eta) + \rho \quad (4.22)$$

We shall prove that for any $r > 0$ there is a compact set K of $\mathscr{C}$,

$$\limsup_{\varepsilon\to 0} \varepsilon \log P_{X_0} \left\{ (X^\varepsilon(\cdot), \ell^\varepsilon(\cdot)) \notin K \right\} \le -r \ . \quad (4.23)$$

By the results in [17, Chapter 3] or [37], we know that (4.21), (4.22) and (4.23) imply the large deviation properties for $(X^\varepsilon(\cdot), \ell^\varepsilon(\cdot))$.

To prove (4.23), we take

$$K = \bigcap_{n=1}^{\infty} \left\{ (\Phi,\eta) \in \mathscr{C}; |\Phi(t) - \Phi(s)| \le \delta_n,\ |\eta(t) - \eta(s)| \le \delta_n, \right.$$
$$\left. \text{for } |t-s| \le \frac{1}{n},\ 0 \le s, t \le T \right\}$$

for some $\delta_n \to 0$ as $n \to \infty$. For example, it is sufficient to take $\delta_n = Rn^{-1/3}$ with R large enough. To see this, we first consider the special case that

$$b_i(x) = 0, \quad i = 1, \ldots, N,\ x > 0 \ . \quad (4.24)$$

We denote $\bar{X}^\varepsilon(t)$ and $\bar{\ell}^\varepsilon(t)$ the corresponding process and its local time. Then

$$P_{X_0} \left\{ (\bar{X}^\varepsilon(\cdot), \bar{\ell}^\varepsilon(\cdot)) \notin K \right\}$$
$$\le \sum P_{X_0} \left\{ \sup \left\{ |\bar{x}^\varepsilon(t) - \bar{x}^\varepsilon(s)|; |t-s| \le \frac{1}{n}, 0 \le s, t \le T \right\} > \frac{1}{2}\delta_n \right\}$$
$$+ \sum P_{X_0} \left\{ \sup \left\{ |\bar{\ell}^\varepsilon(t) - \bar{\ell}^\varepsilon(s)|; |t-s| \le \frac{1}{n}, 0 \le s, t \le T \right\} > \frac{1}{2}\delta_n \right\}$$

We know that $\bar{x}^\varepsilon(t) = |x_0 + \sqrt{\varepsilon} w(t)|$ for a 1-dim Wiener process $w(t)$. Then

$$P_{X_0} \left\{ \sup \left\{ \left|\bar{x}^\varepsilon(t) - \bar{x}^\varepsilon(s)\right|; |t-s| \le \frac{1}{n}, 0 \le s, t \le T \right\} > \frac{1}{2}\delta_n \right\}$$
$$\le P_{X_0} \left\{ \sqrt{\varepsilon} \sup \left\{ \left| w(t) - w\left(\frac{k}{n}\right) \right|; \frac{k}{n} \le t \le \frac{k+1}{n},\ k = 1, \ldots, K \right\} > \frac{1}{8}\delta_n \right\}$$
$$\le \sum\nolimits_{k=1}^{K} P_{X_0} \left\{ \sqrt{\varepsilon} \sup \left\{ \left| w(t) - w\left(\frac{k}{n}\right) \right|; \frac{k}{n} \le t \le \frac{k+1}{n} \right\} > \frac{1}{8}\delta_n \right\} \ ,$$

where K is equal to the largest integer smaller than nT. We need the following estimate,

$$P_{X_0} \left\{ \sqrt{\varepsilon} \sup \left\{ \left| w(t) - w\left(\frac{k}{n}\right) \right|; \frac{k}{n} \le t \le \frac{k+1}{n} \right\} > \frac{1}{8}\delta_n \right\}$$

$$\leq 2\exp\left(-\frac{1}{2\frac{1}{n}}\frac{\delta_n^2}{64\varepsilon}\right) \tag{4.25}$$

$$\leq 2\exp\left(-cR^2 n^{1/3}\frac{1}{\varepsilon}\right)$$

for some $c > 0$. See [34, Theorem 4.2.1] for a proof. Then we can deduce

$$P_{X_0}\left\{\sup_{|t-s|\leq\frac{1}{n},0\leq s,t\leq T}\left|\bar{x}^{\varepsilon}(t)-\bar{x}^{\varepsilon}(s)\right| > \frac{1}{2}\delta_n\right\} \leq 4nT\exp\left(-cR^2 n^{1/3}\frac{1}{\varepsilon}\right) \ . \tag{4.26}$$

On the other hand, by Lemma 2.2, there is a 1-dim Wiener process $W(t)$ such that

$$\bar{x}^{\varepsilon}(t) = x_0 + \sqrt{\varepsilon}W(t) + \bar{\ell}^{\varepsilon}(t) \ .$$

Then we use (4.25) and (4.26) to deduce

$$P_{X_0}\left\{\sup_{|t-s|\leq\frac{1}{n},0\leq s,t\leq T}\left|\bar{\ell}^{\varepsilon}(t)-\bar{\ell}^{\varepsilon}(s)\right| > \frac{1}{2}\delta_n\right\} \leq 8nT\exp\left(-cR^2 n^{1/3}\frac{1}{\varepsilon}\right) \ .$$

It is easy to see that (4.23) follows from this and (4.26) if we choose R to be large enough.

We now prove (4.23) for the general cases. We remark that by Lemma 3.2,

$$\begin{aligned}
P_{X_0}\left\{\left(X^{\varepsilon}(\cdot),\ell^{\varepsilon}(\cdot)\right)\notin K\right\} &= E_{X_0}\left\{\exp\left(\frac{1}{\sqrt{\varepsilon}}\int_0^T b_{\bar{i}^{\varepsilon}(t)}\left(\bar{x}^{\varepsilon}(t)\right)dW(t)\right.\right.\\
&\qquad\left.\left.-\frac{1}{2\varepsilon}\int_0^T\left|b_{\bar{i}^{\varepsilon}(t)}\left(\bar{x}^{\varepsilon}(t)\right)\right|^2dt\right);\left(\bar{X}^{\varepsilon}(\cdot),\bar{\ell}^{\varepsilon}(\cdot)\right)\notin K\right\}\\
&\leq\left(E_{X_0}\left\{\exp\left(2\frac{1}{\sqrt{\varepsilon}}\int_0^T b_{\bar{i}^{\varepsilon}(t)}\left(\bar{x}^{\varepsilon}(t)\right)dW(t)\right.\right.\right.\\
&\qquad\left.\left.\left.-2\frac{1}{2\varepsilon}\int_0^T\left|b_{\bar{i}^{\varepsilon}(t)}\left(\bar{x}^{\varepsilon}(t)\right)\right|^2dt\right)\right\}\right)^{\frac{1}{2}}\\
&\qquad\times\left(P_{X_0}\left\{\left(\bar{X}^{\varepsilon}(\cdot),\bar{\ell}^{\varepsilon}(\cdot)\right)\notin K\right\}\right)^{\frac{1}{2}}
\end{aligned}$$

by applying Schwarz inequality in the last step. But

$$\begin{aligned}
&E_{X_0}\left\{\exp\left(2\frac{1}{\sqrt{\varepsilon}}\int_0^T, b_{\bar{i}^{\varepsilon}(t)}\left(\bar{x}^{\varepsilon}(t)\right)dW(t)-2\frac{1}{2\varepsilon}\int_0^T\left|b_{\bar{i}^{\varepsilon}(t)}\left(\bar{x}^{\varepsilon}(t)\right)\right|^2dt\right)\right\}\\
&\leq\exp\left(\frac{\|b\|}{\varepsilon}T\right) \ ,
\end{aligned}$$

since

$$E_{X_0}\left\{\exp\left(2\frac{1}{\sqrt{\varepsilon}}\int_0^T b_{\bar{i}^\varepsilon(t)}\left(\bar{x}^\varepsilon(t)\right)dW(t)-4\frac{1}{2\varepsilon}\int_0^T |b_{\bar{i}^\varepsilon(t)}\left(\bar{x}^\varepsilon(t)\right)|^2 dt\right)\right\}=1$$

Here $\|b\| = \sup\{|b_i(x)|; i = 1, \ldots, N, x > 0\}$. Combining this with the result just proved for the cases satisfying (4.24), (4.23) follows.

We now consider (4.21).

We remark that for any δ there are $\hat{\Phi}(\cdot) = \left(\hat{\phi}(\cdot), \hat{j}(\cdot)\right)$ and $\hat{\eta}(\cdot)$ such that the following properties hold: $\hat{\phi}(\cdot)$, $\hat{\eta}(\cdot)$ are piecewise linear;

$$\left\|\Phi - \hat{\Phi}\right\|_T \le \frac{\delta}{2},\ \ \|\eta - \hat{\eta}\|_T \le \frac{\delta}{2} \tag{4.27}$$

$$I_T\left(\hat{\Phi}, \hat{\eta}\right) \le I_T(\Phi, \eta) + \delta\ .$$

We shall give a sketch for how to choose $\hat{\Phi}(\cdot)$ and $\hat{\eta}(\cdot)$.

We first choose a set $U \in [0, T]$ which is a disjoint union of finitely many closed intervals,

$$U = \cup_{i=1}^{N_0}[a_i, b_i]\ ,$$

such that

$$U \supseteq \{t \in [0, T]; \Phi(t) = O\}$$

and

$$0 \le \phi(t) \le \frac{\delta}{2} \quad \text{for } t \in U,$$
$$\Lambda(0)|U \setminus \{t \in [0, T]; \Phi(t) = O\}| \le \frac{\delta}{4}\ .$$

Here $|\cdot|$ is the Lebesgue measure of a set. We may choose a_i, b_i in such a way that $\Phi(a_i) = \Phi(b_i) = O$. Otherewise, we can replace a_i, b_i by

$$\tilde{a}_i = \inf\{t \ge a_i; \Phi(t) = O\}, \quad \tilde{b}_i = \sup\{t \le b_i; \Phi(t) = O\}\ .$$

Denote

$$[0, T] \setminus U = \cup_{i=1}^{N_1}(c_i, d_i)\ ,$$

a decomposition by disjoint open intervals. We choose $\hat{\phi}$, a piecewise linear function on each (c_i, d_i), such that $\hat{\phi}(c_i) = \hat{\phi}(d_i) = 0$,

$$\left\|\phi - \hat{\phi}\right\|_{(c_i, d_i)} \le \frac{\delta}{2}\ ,$$

$$\frac{1}{2}\int_{c_i}^{d_i}\left|b_{j(t)}\left(\hat{\phi}(t)\right)-\frac{d}{dt}\hat{\phi}(t)\right|^2 dt \le \frac{1}{2}\int_{c_i}^{d_i}\left|b_{j(t)}\left(\phi(t)\right)-\frac{d}{dt}\phi(t)\right|^2 dt+\frac{\delta}{4N_1} \ .$$

We define $\hat{j}(t)=j(t)$ on (c_i,d_i) and $\hat{\phi}(t)=0$ on U.

To define $\hat{\eta}$, let $\frac{d}{dt}\hat{\eta}=0$ on $[0,T]\setminus U$. On $[a_i,b_i]$, we choose $\hat{\eta}$ such that

$$\hat{\eta}(a_i)=\eta(a_i), \quad \hat{\eta}(b_i)=\eta(b_i), \quad \left\|\hat{\eta}-\eta\right\|_{[a_i,b_i]}\le\frac{\delta}{2},$$

$$\int_{a_i}^{b_i}\Lambda\left(\frac{d\hat{\eta}}{dt}(t)\right)dt \le \int_{a_i}^{b_i}\Lambda\left(\frac{d\eta}{dt}(t)\right)dt+\frac{\delta}{4N_0} \ .$$

Then $\left(\hat{\Phi},\hat{\eta}\right)$ so defined satisfies the required properties.

This argument can also be applied to construct $\left(\Phi^{(n)}(\cdot),\ \eta^{(n)}(\cdot)\right), n=1,2,3,\ldots$ satisfying the following properties.

(i) For each n, there are $0=t_0^{(n)}<t_1^{(n)}<\cdots<t_{K_n+1}^{(n)}=T$ such that

$$\Phi^{(n)}\left(t_i^{(n)}\right)=\Phi\left(t_i^{(n)}\right), \quad \eta^{(n)}\left(t_i^{(n)}\right)=\eta\left(t_i^{(n)}\right) \ ,$$

and $\phi^{(n)}(\cdot),\eta^{(n)}(\cdot)$ are linear on $[t_i^{(n)},t_{i+1}^{(n)}]$ for each i. Here

$$\Phi^{(n)}(t)=\left(\phi^{(n)}(t),j^{(n)}(t)\right) \ .$$

Moreover,

$$\{t\in[0,T];\Phi(t)=0\}\subseteq\cup\left\{\left[t_i^{(n)},t_{i+1}^{(n)}\right];\Phi\left(t_i^{(n)}\right)=\Phi\left(t_{i+1}^{(n)}\right)=0\right\}; \tag{4.28}$$

(ii) As n tends to infinity, $\left(\Phi^{(n)},\eta^{(n)}\right)$ converges to (Φ,η) uniformly on $[0,T]$ and

$$\lim_{n\to\infty}I_T\left(\Phi^{(n)},\eta^{(n)}\right)=I_T(\Phi,\eta)$$

holds.

This result will be used in the proof of (4.22).

In the following, we fix a small $\rho\ge 0$. Let choose $\hat{\Phi},\hat{\eta}$ satisfying (4.27). Since

$$P_{X_0}\left\{\left\|X^\varepsilon-\Phi\right\|_T\le\delta,\left\|\ell^\varepsilon-\eta\right\|_T\le\delta\right\}\ge P_{X_0}\left\{\left\|X^\varepsilon-\hat{\Phi}\right\|_T\le\frac{\delta}{2},\left\|\ell^\varepsilon-\hat{\eta}\right\|_T\le\frac{\delta}{2}\right\} \ ,$$

we may assume that $\phi(\cdot),\eta(\cdot)$ are piecewise linear. By the Girsanov theorem,

$$P_{X_0}\left\{\left\|X^\varepsilon-\Phi\right\|_T\le\delta,\left\|\ell^\varepsilon-\eta\right\|_T\le\delta\right\}$$

is equal to

$$E_{X_0}\left\{\exp\left(\frac{1}{\sqrt{\varepsilon}}\int_0^T b_{i_0^\varepsilon(t)}\left(x_0^\varepsilon(t)\right)dW(t)-\frac{1}{2\varepsilon}\int_0^T\left|b_{i_0^\varepsilon(t)}\left(x_0^\varepsilon(t)\right)\right|^2 dt\right);\right.$$

$$\left\|X_0^\varepsilon - \Phi\right\|_T \le \delta, \left\|\ell_0^\varepsilon - \eta\right\|_T \le \delta\Big\} \ .$$

Here $X_0^\varepsilon(t) = \left(x_0^\varepsilon(t), i_0^\varepsilon(t)\right)$ and $\ell_0^\varepsilon(t)$ satisfy

$$x_0^\varepsilon(t) = x_0 + \sqrt{\varepsilon} W(t) + \ell_0^\varepsilon(t) \ , \tag{4.29}$$

and $X_0^\varepsilon(t)$ is the process generated by A_0^ε with the gluing condition $\rho(F) = 0$, where

$$A_0^\varepsilon F(x, i) = \frac{\varepsilon}{2} \frac{d^2 F_i}{dx^2}(x) \ .$$

We define $F(\cdot)$ on Γ by

$$F(x, i) = \int_0^x b_i(u)\, du \ . \tag{4.30}$$

By Ito's rule in Lemma 2.3,

$$F\left(X_0^\varepsilon(t)\right) = F(X_0) + \sqrt{\varepsilon} \int_0^T b_{i_0^\varepsilon(t)}\left(x_0^\varepsilon(t)\right) dW(t) + \frac{\varepsilon}{2} \int_0^T \frac{d}{dy} b_{i_0^\varepsilon(t)}\left(x_0^\varepsilon(t)\right) dt + \sum b_i(0) \alpha_i \ell_0^\varepsilon(T) \ .$$

The above expectation is bounded below by

$$E_{X_0}\left\{\exp\left(-\frac{1}{2\varepsilon} \int_0^T \left|b_{i_0^\varepsilon(t)}\left(x_0^\varepsilon(t)\right)\right|^2 dt\right); \left\|X_0^\varepsilon - \Phi\right\|_T \le \delta, \left\|\ell_0^\varepsilon - \eta\right\|_T \le \delta\right\} \times \exp\left(\frac{1}{\varepsilon}\left(F(\Phi(t)) - F(\Phi(0)) - \sum b_i(0)\alpha_i \eta(T)\right) - \frac{c\delta}{\varepsilon}\right)$$

for some $c > 0$. This in turn has the lower bound

$$E_{X_0}\left\{\exp\left(-\frac{1}{2\varepsilon} \int_0^T \left|b_{i_0^\varepsilon(t)}(\phi(t))\right|^2 dt\right); \left\|x_0 + \sqrt{\varepsilon} W - (\phi - \eta)\right\|_T \le \frac{\delta}{2}, \left\|X_0^\varepsilon - \Phi\right\|_T \le \frac{\delta}{2}\right\} \exp\left(\frac{1}{\varepsilon}\Big(F(\Phi(T)) - F(\Phi(0)) - \sum b_i(0)\alpha_i \eta(T)\Big) - \frac{c\delta}{\varepsilon}\right) \tag{4.31}$$

by (4.29).

Since ϕ, η are piecewise linear, there are $0 = t_0 < t_1 < \cdots < t_{K+1} = T$ such that $\phi(\cdot), \eta(\cdot)$ are linear on $[t_i, t_{i+1}]$ for each i. Denote $\Delta_i = t_{i+1} - t_i$. The expectation in (4.31) can be bounded from below by

$$\prod_{i=1}^K \inf_{Y=(y,j):|Y-\Phi(t_i)|\le\delta'} E_Y\left\{\exp\left(-\frac{1}{2\varepsilon} \int_0^{\Delta_i} \left|b_{i_0^\varepsilon(t)}(\phi_i(t))\right|^2 dt\right); \left\|X_0^\varepsilon - \Phi_i\right\|_{\Delta_i} \le \delta', \ \left\|y + \sqrt{\varepsilon} W - (\phi_i - \eta_i)\right\|_{\Delta_i} \le \delta'\right\} \tag{4.32}$$

for $\delta' < \frac{\delta}{2}$ small enough. Here

$$\Phi_i(t) = \Phi(t_i + t), \ \phi_i(t) = \phi(t_i + t), \ \eta_i(t) = \eta(t_i + t) - \eta(t_i) \quad .$$

To estimate

$$E_Y \left\{ \exp\left(-\frac{1}{2\varepsilon} \int_0^{\Delta_i} \left| b_{i_0^\varepsilon(t)}(\phi_i(t)) \right|^2 dt \right); \left\| X_0^\varepsilon - \Phi_i \right\|_{\Delta_i} \le \delta', \right.$$
$$\left. \left\| y + \sqrt{\varepsilon} W - (\phi_i - \eta_i) \right\|_{\Delta_i} \le \delta' \right\} \tag{4.33}$$

for $|Y - \Phi(t_i)| \le \delta'$, we consider different cases separately.

Assume that (t_i, t_{i+1}) is in $\{0 \le t \le T; \Phi(t) \ne O\}$, then $\eta_i = 0$ on $[0, \Delta_i]$. Using the large deviation properties for small perturbed diffusion processes, see [17, Chapter 3], it is easy to see that (4.33) has a lower bound

$$\exp\left(-\frac{1}{2\varepsilon} \left(\int_{t_i}^{t_{i+1}} \left(\left| \frac{d}{dt}\phi(t) \right|^2 + \left| b_{j(t)}(\phi(t)) \right|^2 \right) dt + \rho \right) \right) \tag{4.34}$$

for ε small. Here $\rho > 0$ is given.

Now assume that (t_i, t_{i+1}) is included in $\{0 \le t \le T; \Phi(t) = O\}$. Either η is constant or has constant positive derivative in this interval. We may assume $Y = O$ for both cases. If η is constant, i.e., $\eta_i = 0$, then we can take a function ψ on $[0, \Delta_i]$ such that $\psi(0) = 0$ and

$$0 < \psi(t) < \frac{\delta'}{2}, \ t \in [0, \Delta_i] \quad .$$

Then for small enough $\delta'' < \delta'$, $\|\sqrt{\varepsilon} W - \psi\|_{\Delta_i} < \delta''$ implies $\|x_0^\varepsilon\|_{\Delta_i} < \delta'$ and $\|\sqrt{\varepsilon} W\| < \delta'$. Here we use the fact that

$$\ell_0^\varepsilon(t) = -\inf\left\{ x_0 + \sqrt{\varepsilon} W(s); 0 \le s \le t \right\} \wedge 0 \quad .$$

See [24, Theorem 4.2,Chapter 3]. Therefore, $x_0^\varepsilon(\cdot)$, $\ell_0^\varepsilon(\cdot)$ can be considered as continuous functionals of $W(\cdot)$.

Using this, (4.33) has a lower bound

$$E_O \left\{ \exp\left(-\frac{1}{2\varepsilon} \int_0^{\Delta_i} \left| b_{i_0^\varepsilon(t)}(0) \right|^2 dt; \left\| \sqrt{\varepsilon} W - \psi \right\|_{\Delta_i} \le \delta'' \right) \right\}$$

which in turn has a lower bound

$$\exp\left(-\frac{1}{2\varepsilon} \min_j \left| b_j(0) \right|^2 \Delta_i - \frac{1}{2\varepsilon} \int_0^{\Delta_i} \left| \frac{d}{dt}\psi(t) \right|^2 dt - \frac{\rho}{2\varepsilon} \right)$$
$$\ge \exp\left(-\frac{1}{\varepsilon} \Lambda(0) \Delta_i - \frac{\rho}{\varepsilon} \right) \tag{4.35}$$

if we choose δ' small enough such that $\frac{1}{2\varepsilon} \int_0^{\Delta_i} \left| \frac{d}{dt}\psi(t) \right|^2 dt < \frac{\rho}{2}$.

We now assume that $\eta_i(t) = \alpha t$ for some $\alpha > 0$ on (t_i, t_{i+1}) . Note that here $\phi_i(t) = 0$ on (t_i, t_{i+1}). By the change of probability measure using the Girsanov theorem, (4.33) becomes

$$E_O\left\{\exp\left(-\frac{1}{2\varepsilon}\int_0^{\Delta_i}\left|b_{i^\varepsilon_\alpha(t)}(0)\right|^2 dt\right)\exp\left(-\frac{\alpha}{\sqrt{\varepsilon}}W(\Delta_i) - \frac{1}{2\varepsilon}\|\alpha\|^2\Delta_i\right);\right.$$
$$\left.\left\|x^\varepsilon_\alpha\right\|_{\Delta_i} \le \delta',\ \left\|\sqrt{\varepsilon}W\right\|_{\Delta_i} \le \delta'\right\}$$

which in turn is bounded from below by

$$E_O\left\{\exp\left(-\frac{1}{2\varepsilon}\int_0^{\Delta_i}\left|b_{i^\varepsilon_\alpha(t)}(0)\right|^2 dt\right);\ \left\|x^\varepsilon_\alpha\right\|_{\Delta_i} \le \delta',\ \ \left\|\sqrt{\varepsilon}W\right\|_{\Delta_i} \le \delta'\right\}$$
$$\times\exp\left(-\frac{1}{2\varepsilon}|\alpha|^2\Delta_i - \frac{\rho}{\varepsilon}\right) \tag{4.36}$$

if δ' is small enough. Here $X^\varepsilon_\alpha(t) = \left(x^\varepsilon_\alpha(t), i^\varepsilon_\alpha(t)\right)$ and $\ell^\varepsilon_\alpha(t)$ satisfy

$$x^\varepsilon_\alpha(t) = \sqrt{\varepsilon}W(t) - \alpha t + \ell^\varepsilon_\alpha(t)\ \ , \tag{4.37}$$

X^ε_α is the process generated by A^ε_α with gluing condition $\rho(F) = 0$, where

$$A^\varepsilon_\alpha F(x, i) = \frac{\varepsilon}{2}\frac{d^2F_i}{dx^2}(x) - \alpha\frac{dF_i}{dx}(x)\ \ .$$

It remains to get a lower bound for the expectation in (4.36). The following argument ensures that it is enough to study the following expectation:

$$E_O\left\{\exp\left(-\frac{1}{2\varepsilon}\int_0^{\Delta_i}\left|b_{i^\varepsilon_0(t)}(0)\right|^2 dt\right);\ \left|\sqrt{\varepsilon}W(\Delta_i)\right| \le \delta'',\ x^\varepsilon(\Delta_i) \le \delta''\right\} \tag{4.38}$$

for δ'' small enough. For a fixed positive integer n, we take $s_k = \frac{k}{n}\Delta_i, k = 1, \ldots, n$. By a routine argument using the estimate in [34, Theorem 4.2.1], we can show that, for any $R > 0$, we can choose n large enough,

$$P_O\left\{\|x^\varepsilon_0\|_\Delta \le \delta',\ \|\sqrt{\varepsilon}W\|_{\Delta_i} \le \delta'\right\}$$
$$\ge P_O\left\{|x^\varepsilon_0(s_k)| \le \frac{\delta'}{2},\ |\sqrt{\varepsilon}W(s_k)| \le \frac{\delta'}{2},\ k = 1, \ldots, n\right\} - \exp\left(-\frac{R}{\varepsilon}\right)$$

Our assertion follows from this and an argument involving conditioning.

To estimate (4.38), we change the scales, $t \to \varepsilon t, x \to \frac{x}{\varepsilon}$, which gives the processes $X_\alpha(t) = (x_\alpha(t), i_\alpha(t))$ and $\ell_\alpha(t)$ satisfying

$$x_\alpha(t) = W(t) - \alpha t + \ell_\alpha(t)\ \ ,$$

and $X_\alpha(t)$ is the process generated by A_α with gluing condition $\rho(F)=0$, where

$$A_\alpha F(x,i)=\frac{1}{2}\frac{d^2F_i}{dx^2}(x)-\alpha\frac{dF_i}{dx}(x) \quad .$$

Then (4.38) becomes

$$E_O\left\{\exp\left(-\frac{1}{2}\int_0^{\frac{\Delta_i}{\varepsilon}}|b_{i_\alpha(t)}(0)|^2dt\right);\ \left|W\left(\frac{\Delta_i}{\varepsilon}\right)\right|\le\frac{\delta''}{\varepsilon},\ x_0\left(\frac{\Delta_i}{\varepsilon}\right)\le\frac{\delta''}{\varepsilon}\right\} \quad .$$

By Lemma 4.2 applying to the process $x(t)=\alpha x_\alpha\left(\frac{t}{\alpha^2}\right)$, the limit

$$\lim_{\delta''\to 0}\lim_{\varepsilon\to 0}\varepsilon\log E_O\left\{\exp\left(-\frac{1}{2}\int_0^{\frac{\Delta_i}{\varepsilon}}\left|b_{i_0(t)}(0)\right|^2dt\right);\right.$$

$$\left.\left|W\left(\frac{\Delta_i}{\varepsilon}\right)\right|\le\frac{\delta''}{\varepsilon},\ x_0\left(\frac{\Delta_i}{\varepsilon}\right)\le\frac{\delta''}{\varepsilon}\right\} \tag{4.39}$$

exists and is equal to $-\alpha^2\Delta_i\hat{\Lambda}(V)$. Here $V=\frac{1}{2\alpha^2}\left(|b_1(0)|^2,\dots,|b_n(0)|^2\right)$. Note that

$$\alpha^2\hat{\Lambda}(V)=\inf\left\{\frac{1}{2}\sum\beta_i\left|b_i(0)+\alpha\frac{\alpha_i}{\beta_i}\right|^2-\alpha\sum\alpha_ib_i(0)-\frac{1}{2}\alpha^2\right\} \quad ,$$

where the inf is taken over all $\{\beta_i\}$ satisfying (4.8). Then (4.21) follows from (4.31) $\sim$ (4.39).

We now prove (4.22). Let $\Phi^{(n)},\eta^{(n)},0=t_0^{(n)}<t_1^{(n)}<\dots<t_{K_n+1}=T$ be chosen as in (4.28). We fix n and a small $\rho\ge 0$. We denote

$$K=K_n,\ t_i=t_i^{(n)},\ i=1,\dots,K+1 \quad .$$

Then as (4.32),

$$P_{X_0}\left\{\|X^\varepsilon-\Phi\|_T\le\delta,\|\ell^\varepsilon-\eta\|_T\le\delta\right\}$$

has an upper bound

$$\prod_{i=1}^K\sup_{Y=(y,j):|Y-\Phi(t_i)|\le\delta}E_Y\left\{\exp\left(-\frac{1}{2\varepsilon}\int_0^{\Delta_i}|b_{i_0^\varepsilon(t)}\left(\phi_i^{(n)}(t)\right)|^2dt\right)\right);$$

$$\left.\|X_0^\varepsilon-\Phi_i^{(n)}\|_{\Delta_i}\le 2\delta,\ \|y+\sqrt{\varepsilon}W-\left(\phi_i^{(n)}-\eta_i^{(n)}\right)\|_{\Delta_i}\le 2\delta\right\}$$

$$\exp\left(\frac{1}{\varepsilon}\left(F\left(\Phi^{(n)}(T)\right)-F\left(\Phi^{(n)}(0)\right)-\sum b_i(0)\alpha_i\eta^{(n)}(T)\right)\right)\exp\left(\frac{\rho}{\varepsilon}\right) \tag{4.40}$$

for small ρ if δ is small and n is large enough. Here

$$\Phi_i^{(n)}(t)=\Phi^{(n)}(t_i+t),\ \phi_i^{(n)}(t)=\phi^{(n)}(t_i+t),\ \eta_i^{(n)}(t)=\eta^{(n)}(t_i+t)-\eta^{(n)}(t_i)\quad .$$

The processes $X_0^\varepsilon(t)=\left(x_0^\varepsilon(t), i_0^\varepsilon(t)\right)$ and $\ell_0^\varepsilon(t)$ satisfy (4.29).

To evaluate

$$E_Y\left\{\exp\left(-\frac{1}{2\varepsilon}\int_0^{\Delta_i}|b_{i_0^\varepsilon(t)}\left(\phi_i^{(n)}(t)\right)|^2dt\right);\ \|X_0^\varepsilon-\Phi_i^{(n)}\|_{\Delta_i}\le 2\delta,\right.$$
$$\left.\|y+\sqrt{\varepsilon}W-\left(\phi_i^{(n)}-\eta_i^{(n)}\right)\|_{\Delta_i}\le 2\delta\right\}\ , \tag{4.41}$$

we consider the following cases separately:

(i) (t_i, t_{i+1}) is contained in $\left\{t;\ \Phi^{(n)}(t)\neq 0\right\}$;
(ii) $\Phi^{(n)}(t_i)=\Phi^{(n)}(t_{i+1})=0,\ \eta(t_i)=\eta(t_{i+1})$;
(iii) $\Phi^{(n)}(t_i)=\Phi^{(n)}(t_{i+1})=0,\ \eta(t_i)\neq\eta(t_{i+1})$;

For the cases (i)(ii), we get an upper bound similar to (4.34)(4.35). For the case (iii), assuming $\frac{d}{dt}\eta(t)=\alpha$ on (t_i, t_{i+1}), similar to (4.36), (4.41) has an upper bound

$$E_Y\left\{\exp\left(-\frac{1}{2\varepsilon}\int_0^{\Delta_i}|b_{i_\alpha^\varepsilon(t)}\left(\phi_i^{(n)}(t)\right)|^2dt\right);\ |X_\alpha^\varepsilon(\Delta_i)|\le 2\delta,\right.$$
$$\left.|\sqrt{\varepsilon}W(\Delta_i)|\le 3\delta\right\}\exp\left(-\frac{1}{2}\alpha^2\Delta_i+\frac{\rho}{\varepsilon}\right)\ ,$$

since $Y=(y,j)$ satisfying $y\le\delta$. Using the previous argument for relating (4.36) and (4.38), then applying Corollary 4.3., letting δ small enough, (4.41) has an upper bound

$$\exp\left(-\frac{\Delta_i}{\varepsilon}\left(\Lambda(\alpha)-\alpha\sum\alpha_i b_i(0)\right)+\frac{1}{K\varepsilon}\rho\right)\ .$$

From these and collecting the terms, we can show that the left hand side of (4.22) is bounded from above by $I_T\left(\Phi^{(n)},\eta^{(n)}\right)+\rho$ for large n. Then we get the relation (4.22) by letting $n\to\infty$. This completes the proof of the theorem.

Corollary 4.8. *Let $X^\varepsilon(t)$, $\ell^\varepsilon(t)$ be the same processes as in Theorem* 4.5. *but with L_i^ε given by* (4.1). *Then, for fixed $T>0$, the family of continuous processes,*

$$\left(X^\varepsilon(t),\ell^\varepsilon(t)\right)_{0\le t\le T},\quad \varepsilon>0\ ,$$

satisfies the large deviation principle in the uniform topology with action functional $\varepsilon^{-1}I_T(\cdot)$. Here, for $(\Phi,\eta)\in\mathscr{C}$, $\Phi(t)=(\phi(t), j(t))$, such that $\phi(\cdot)$, $\eta(\cdot)$ are absolutely continuous and satisfy (4.18),

$$I_T(\Phi,\eta)=\frac{1}{2}\int_0^T\sigma_{j(t)}^{-2}(\phi(t))\,|\frac{d\phi}{dt}(t)-b_{j(t)}(\phi(t))|^2\chi_{\{\phi(t)\neq 0\}}dt$$
$$+\int_0^T\Lambda\left(\frac{d\eta}{dt}(t)\right)\chi_{\{\phi(t)=0\}}dt. \tag{4.42}$$

Here

$$\Lambda(\theta) = \inf\left\{\frac{1}{2}\sum \beta_i \frac{1}{\sigma_i^2(0)}|b_i(0) + \theta\frac{\alpha_i}{\beta_i}|^2; 0 < \beta_i < 1, \sum \beta_i = 1\right\} \quad ,$$

For other (Φ, η), $I_T(\Phi, \eta) = \infty$.

Proof. We apply Lemma 3.1 with $c_i(x) = \sigma_i(x)^{-1}$. Denote $Y^\varepsilon(t) = G(X^\varepsilon(t))$. Here

$$G(x, i) = G_i(x) = \int_0^x c_i(u)\, du \quad .$$

Then $Y^\varepsilon(t)$ is the diffusion process generated by $\hat{A}^\varepsilon$ as in Theorem 4.5. with L_i^ε replaced by

$$\hat{L}_i^\varepsilon f(y) = \frac{1}{2}\varepsilon \frac{d^2 f}{dy^2}(y) + \hat{b}_i^\varepsilon(y)\frac{df}{dy}(y) \quad ,$$

and α_i replaced by $\hat{\alpha}_i = \left(\sum \alpha_j \sigma_j(0)^{-1}\right)^{-1} \alpha_i \sigma_i(0)^{-1}$. Here,

$$\hat{b}_i^\varepsilon(y) = b_i\left(G_i^{-1}(y)\right)\frac{1}{\sigma_i\left(G_i^{-1}(y)\right)} - \frac{1}{2}\varepsilon\frac{d\sigma_i}{dx}\left(G_i^{-1}(y)\right) \quad .$$

The local time of $Y^\varepsilon(\cdot)$ at O is given by

$$\hat{\ell}^\varepsilon(t) = \sum \alpha_i \frac{1}{\sigma_i(0)}\ell^\varepsilon(t) \quad .$$

Now consider another diffusion process $\bar{Y}^\varepsilon$ generated by $\bar{A}^\varepsilon$ as in Theorem 4.5 with L_i^ε replaced by

$$\bar{L}_i^\varepsilon f(y) = \frac{1}{2}\varepsilon \frac{d^2 f}{dy^2}(y) + \hat{b}_i(y)\frac{df}{dy}(y) \quad ,$$

and α_i replaced by $\hat{\alpha}_i$ defined above. Here $\hat{b}_i(y) = b_i\left(G_i^{-1}(y)\right)\sigma_i\left(G_i^{-1}(y)\right)^{-1}$. Then by Lemma 3.2.,

$$E\left[\left(Y^\varepsilon, \hat{\ell}^\varepsilon\right) \in B\right] = E\left[\exp\left(\int_0^T e_{\bar{i}^\varepsilon(t)}\left(\bar{y}^\varepsilon(t)\right) dW(t) - \frac{1}{2}\int_0^T |e_{\bar{i}^\varepsilon(t)}\left(\bar{y}^\varepsilon(t)\right)|^2 dt\right);\right.$$
$$\left.\left(\bar{Y}^\varepsilon, \hat{\ell}^\varepsilon\right) \in B\right]$$

for any B, a Borel set in $\mathscr{C}$. Here $e_i(y) = -\frac{1}{2}\frac{d\sigma_i}{dx}(G_i^{-1}(y))$. Using this, we can show that $(Y^\varepsilon, \hat{\ell}^\varepsilon)$ satisfies the large deviation principle if and only if $(\bar{Y}^\varepsilon, \hat{\ell}^\varepsilon)$ satisfies the large deviation principle, with the same action functional. By Theorem 4.5, $(\bar{Y}^\varepsilon, \hat{\ell}^\varepsilon)$ satisfies the large deviation principle. Therefore, $(Y^\varepsilon, \hat{\ell}^\varepsilon)$, and $(X^\varepsilon, \ell^\varepsilon)$ too, also satisfies the large deviation principle. A simple calculation also shows that (4.42) gives the normalized action functional for $(X^\varepsilon, \ell^\varepsilon)$. This completes the proof.

In the rest we consider general Γ and state the large deviation properties for the diffusion processes described by the operators in (4.1). The results can be proved using Corollary 4.8., an argument involving conditioning after suitably dividing the time interval. We shall skip the details of the proof. To describe the action functional, let $(\Phi, \eta_1, \eta_2, \dots, \eta_M) \in \mathscr{C}$, $\Phi(t) = (\phi(t), j(t))$, satisfying the following properties: $\eta_k(t)$, $y_{j(t)}(\Phi(t))$ are absolutely continuous on $\{t; \Phi(t) \neq O_k, k = 1, \dots, M\}$; η_k is constant on the intervals where $\Phi(t) \neq O_k$; the following integrals are finite:

$$\int_0^T \left| \frac{dy_{j(t)}\,(\Phi(t))}{dt} \right|^2 \chi_{\{\Phi(t)\neq O_k, k=1,\dots M\}} dt, \quad \int_0^T \left| \frac{d\eta_k}{dt}(t) \right|^2 dt, \ k = 1, \dots, M \ .$$

Then we define

$$\begin{aligned} & I_T\,(\Phi, \eta_1, \dots, \eta_M) \\ & = \frac{1}{2} \int_0^T \sigma^{-2}_{j(t)} \left(y_{j(t)}\,(\Phi(t))\right) \left| \frac{dy_{j(t)}\,(\Phi(t))}{dt} - b_{j(t)} \left(y_{j(t)}\,(\Phi(t))\right) \right|^2 \\ & \quad \times \chi_{\{\Phi(t)\neq O_k, k=1,\dots,M\}} dt + \sum_k \int_0^T \Lambda^{(k)} \left(\frac{d\eta_k(t)}{dt} \right) \chi_{\{\Phi(t)=O_k\}} dt \ . \end{aligned}$$

Here

$$\Lambda^{(k)}\,(\theta) = \inf \left\{ \frac{1}{2} \sum_{I_i \sim O_k} \beta_i \frac{1}{\sigma_i^2\,(y_i\,(O_k))} \left| b_i\,(y_i\,(O_k)) + \theta \frac{\alpha_{ik}}{\sum_{I_i\sim O_k} \alpha_{ik}} \frac{1}{\beta_i} \right|^2 \right\} \ ,$$

where the inf is taken over all $\{\beta_i\}$ satisfying $0 < \beta_i < 1$, $\sum_{I_i \sim O_k} \beta_i = 1$.

We define $I_T(\Phi, \eta_1, \dots, \eta_M)$ to be infinite if $(\Phi, \eta_1, \dots, \eta_M)$ fails to satisfy the above conditions.

Theorem 4.9. *For each $T > 0$, the family of coupled processes*

$$\left(X^\varepsilon(t), \ell_k^\varepsilon(t), k = 1, \dots, M\right), \ 0 \le t \le T, \ \varepsilon > 0 \ ,$$

satisfy the large deviation properties in the uniform topology with action functional $\varepsilon^{-1} I_T(\cdot)$.

5. Some applications

In this section, we present two examples to show the application of our results obtained in Section 4. The first example is concerning the diffusion processes with discontinuous drift and their large deviation properties. Such processes have been studied before in different context.See [1, 2, 6, 8–13]. The second example is concerning the small diffusion asymptotics for Dirichlet problem in a domain on a graph. Such problem is of interest when study the diffusion processes in narrow channels. See [16].

For the first example, let $b(\cdot)$ be a function defined on R. We assume that the restriction of $b(\cdot)$ on $(0, \infty)$ and $(-\infty, 0)$, denoted by $b_+(\cdot)$ and $b_-(\cdot)$ respectively, are smooth with bounded derivatives of any order. We allow $b(\cdot)$ to be discontinuous at 0. It has right and left limits at 0 which are denoted as $b_+(0)$ and $b_-(0)$. Let $z^\varepsilon(t)$ be the diffusion solving the following equation,

$$dz^\varepsilon(t) = b\left(z^\varepsilon(t)\right) dt + \sqrt{\varepsilon} dW(t) \ . \tag{5.1}$$

Here $W(t)$ is a one dimensional Wiener process. We note by [31] that (5.1) has a unique solution. The large deviation properties of this process are studied in [8, 13, 26, 27] with different conditions on $b(\cdot)$. Here we shall derive these results by using Theorem 4.5.

By the following consideration, $z^\varepsilon(t)$ can be viewed as a particular example studied in Section 4. Let $X^\varepsilon(t) = (x^\varepsilon(t), i^\varepsilon(t))$ be defined by ,

$$x^\varepsilon(t) = |z^\varepsilon(t)|,$$

$$i^\varepsilon(t) = \begin{cases} 1, & \text{if } z^\varepsilon(t) > 0 \\ 2, & \text{otherwise} \end{cases}$$

This gives a representation of the process $z^\varepsilon(t)$. The new process is a Markov process on Γ,

$$\Gamma = \{(x, i)\,, x \geq 0, i = 1, 2\} \ \ ,$$

and is generated by

$$A^\varepsilon F\,(x, i) = \frac{\varepsilon}{2}\frac{d^2 F_i}{dx^2}(x) + b_i(x)\frac{dF_i}{dx}(x) \ \ ,$$

with gluing condition $\rho\,(F) = 0$, where

$$\rho\,(F) = \frac{1}{2}\left(\frac{dF_1}{dx}\,(0+) + \frac{dF_2}{dx}\,(0+)\right) \ \ ,$$

and

$$b_1(x) = b_+(x),\ b_2(x) = -b_-\,(-x)\,,\ \ x > 0 \ \ .$$

We note that $\alpha_i = \frac{1}{2}$ by using our terminology. In the following, we denote

$$\dot{\phi}(t) = \frac{d}{dt}\phi(t)$$

for a function $\phi : [0, T] \to R$. Then we have the following result.

Theorem 5.1. *Let $z \in R$ and $\phi : [0, T] \to R$ a smooth function with $\phi(0) = z$. Then*

$$\lim_{\delta\to 0} \liminf_{\varepsilon\to 0} \varepsilon \log P_z \left\{\|z^\varepsilon - \phi\|_T < \delta\right\}$$
$$= \lim_{\delta\to 0} \limsup_{\varepsilon\to 0} \varepsilon \log P_z \left\{\|z^\varepsilon - \phi\|_T < \delta\right\} \quad . \tag{5.2}$$

and is equal to

$$-\frac{1}{2}\left(\int_0^T |\dot\phi(t) - b(\phi(t))|^2 \chi_{\{\phi(t)\neq 0\}} dt + \int_0^T \min_{0\le\beta\le 1} \left\{|\beta b_+(0) + (1-\beta) b_-(0)|^2\right\} \chi_{\{\phi(t)=0\}} dt\right)$$

if $b_+(0) < b_-(0)$ *and is equal to*

$$-\frac{1}{2}\left(\int_0^T |\dot\phi(t) - b(\phi(t))|^2 \chi_{\{\phi(t)\neq 0\}} dt + \int_0^T \min\left\{|b_+(0)|^2, |b_-(0)|^2\right\} \chi_{\{\phi(t)=0\}} dt\right)$$

if $b_+(0) \ge b_-(0)$.

Proof. By Theorem 4.5. and the contraction principle ([17, Theorem 3.1., Chapter 3] or [37]) applying to the mapping

$$\left(X^\varepsilon(\cdot), \ell^\varepsilon(\cdot)\right) \to X^\varepsilon(\cdot) \quad ,$$

(5.2) is equal to

$$-\frac{1}{2}\int_0^T |\dot\phi(t) - b(\phi(t))|^2 \chi_{\{\phi(t)\neq 0\}} dt - \inf\left\{\int_0^T \Lambda(\dot\eta)\, \chi_{\{\phi(t)=0\}} dt\right\}$$

where inf is taken over all absolutely continuous η such that $\dot\eta(\cdot) \ge 0$ and $\dot\eta(\cdot) = 0$ where $\phi(t) \neq 0$. The above is equal to

$$-\frac{1}{2}\int_0^T |\dot\phi(t) - b(\phi(t))|^2 \chi_{\{\phi(t)\neq 0\}} dt - \inf_{\theta\ge 0}\{\Lambda(\theta)\} \int_0^T \chi_{\{\phi(t)=0\}} dt \quad . \tag{5.3}$$

It remains to calculate

$$\inf_{\theta\ge 0}\{\Lambda(\theta)\} \quad . \tag{5.4}$$

We note that

$$\Lambda(\theta) = \inf_{0<\beta<1} \left\{ \frac{1}{2}\beta b_{+}(0)^2 + \frac{1}{2}(1-\beta) b_{-}(0)^2 + \frac{1}{2}(b_{+}(0) - b_{-}(0))\theta + \frac{1}{8}\theta^2 \left(\frac{1}{\beta} + \frac{1}{1-\beta} \right) \right\} .$$

Therefore, (5.4) is equal to

$$\inf_{\beta} \left\{ \frac{1}{2} b_{+}(0)^2 \beta + \frac{1}{2} b_{-}(0)^2 (1-\beta) + \inf_{\theta} \left\{ \frac{1}{2}(b_{+}(0) - b_{-}(0))\theta + \frac{1}{8}\theta^2 \left(\frac{1}{\beta} + \frac{1}{1-\beta} \right) \right\} \right\} .$$

If $b_{+}(0) - b_{-}(0) > 0$, the *inf* over θ attains when $\theta = 0$. The above is equal to $\frac{1}{2} min \left\{ b_{+}(0)^2, b_{-}(0)^2 \right\}$. On the other hand , if $b_{+}(0) - b_{-}(0) < 0$, then it is equal to

$$\inf_{\beta} \left\{ \frac{1}{2} \left(b_{+}(0)^2 \beta + b_{-}(0)^2 (1-\beta) \right) - \frac{1}{2}\beta(1-\beta)(b_{+}(0) - b_{-}(0))^2 \right\}$$

$$= \inf_{\beta} \left\{ \frac{1}{2} |\beta b_{+}(0) + (1-\beta) b_{-}(0)|^2 \right\}$$

which gives the value of (5.4) in this case. This completes the proof.

We note that the coefficient in $\int_0^T \chi_{\{\phi(t)=0\}} dt$ is also consistent with the one using a formula suggested by the work of [11] which can be written as:

$$\inf \left\{ \beta L_{+}(\theta_{+}) + (1-\beta) L_{-}(\theta_{-}) ; \beta\theta_{+} + (1-\beta)\theta_{-} = 0, 0 < \beta < 1, \theta_{+} < 0 \right\}$$

where

$$L_{+}(\theta) = \frac{1}{2} |\theta - b_{+}(0)|^2$$

$$L_{-}(\theta) = \frac{1}{2} |\theta - b_{-}(0)|^2 .$$

It is also interesting to study the large deviation properties for the process $(z^{\varepsilon}(t), \eta^{\varepsilon}(t))$, where

$$\eta^{\varepsilon}(t) = \int_0^t \chi_{\{z^{\varepsilon}(s)>0\}} ds .$$

This is important when we study the large deviational properties for two dimensional (or high dimensional) diffusion processes with discontinuous statistics. See [9–13, 26, 27].

Now we give another application. Let Γ be the graph described in Section 1 consisting of vertices $O_1, \ldots, O_M$ and edges $I_1, \ldots, I_N$. The operator A^{ε} and the function $\rho(F)(O_k)$, $k = 1, \ldots, M$, are the same as in Section 4. They are given by (2.2) (2.3) with L_i replaced by L_i^ε in (4.1). Here we shall consider a function $F_t^\varepsilon(x),\ t \geq 0,\ X \in \Gamma$ satisfying the following equation,

$$
\begin{aligned}
&\frac{d}{dt} F_t^\varepsilon = A^\varepsilon F_t^\varepsilon + \frac{V}{\varepsilon} F_t^\varepsilon,\ t > 0 \\
&F_0^\varepsilon(X) = \exp\left(-\frac{U(X)}{\varepsilon}\right),\ \ X \in \Gamma, \qquad (5.5)\\
&\rho\left(F_t^\varepsilon\right)(O_k) = \frac{c_k}{\varepsilon} F_t^\varepsilon\,(O_k)\,,\ k = 1, \ldots, M,\ t > 0\ .
\end{aligned}
$$

Here $U(\cdot)$, $V(\cdot)$ are bounded continuous on Γ. The solvability of the equation (5.5) is not obvious and shall be discussed eleswhere. Here we assume that such F_t^ε exists in $C_b^\infty(\Gamma)$. We wish to study the asymptotic behavior of these functions. The following is our main result. We recall that $\mathscr{C}$ is defined in (4.3).

Theorem 5.2. *Let F_t^ε be in $C_b^\infty(\Gamma)$ and satisfy the equation (5.5). Then the following holds:*

$$
\begin{aligned}
\lim_{\varepsilon \to 0} \varepsilon \log F_T^\varepsilon(X) = \sup \Bigg\{ &-U\left(\Phi(T)\right) + \int_0^T V\left(\Phi(t)\right) dt \\
&- \sum_{k=1}^M c_k \eta_k(t) - I_T\left(\Phi, \eta_1, \ldots, \eta_M\right) \Bigg\} \qquad (5.6)
\end{aligned}
$$

for any $X \in \Gamma$, where the inf *is taken over all $(\Phi, \eta_1, \ldots, \eta_M) \in \mathscr{C}$ with $\Phi(0) = X$.*

Proof. Let $X^\varepsilon(t)$ be the process generated by A^ε. We denote by $\ell_1^\varepsilon, \ldots, \ell_M^\varepsilon$ the local times of the process at $O_1, \ldots, O_M$. By Ito's formula applying to F_{T-t}^ε, we have

$$
dF_{T-t}^\varepsilon\left(X_t^\varepsilon\right) = -\frac{V\left(X^\varepsilon(t)\right)}{\varepsilon} F_{T-t}^\varepsilon\left(X_t^\varepsilon\right) dt + \sum_{k=1}^M \frac{c_k}{\varepsilon} F_{T-t}^\varepsilon\,(O_k)\, d\ell_k^\varepsilon(t) + dM^\varepsilon(t)\ ,
$$

where $M^\varepsilon(t)$ is a martingale. From this, we see that

$$
F_{T-t}^\varepsilon\left(X_t^\varepsilon\right) \exp\left(\frac{1}{\varepsilon} \int_0^t V\left(X^\varepsilon(s)\right) ds - \frac{1}{\varepsilon} \sum_{k=1}^M c_k \ell_k^\varepsilon(t)\right)
$$

is a martingale. Therefore,

$$
F_T^\varepsilon(X) = E_X\left\{F_0^\varepsilon\left(X_T^\varepsilon\right) \exp\left(\frac{1}{\varepsilon} \int_0^T V\left(X^\varepsilon(s)\right) ds - \frac{1}{\varepsilon} \sum_{k=1}^M c_k \ell_k^\varepsilon(T)\right)\right\}\ .
$$

The result (5.6) follows from this by applying the large deviation principle in Theorem 4.9. and the Lapace method. See [37, Theorem 2.2.].

Appendix 1

In this appendix, we shall prove Lemma 4.7. First, we note that Lemma 4.7. (ii) is easy since there are $c_1, c_2 > 0$ such that

$$\int_0^T \Lambda\left(\frac{d\eta}{dt}(t)\right)\chi_{\{\Phi(t)=O\}}dt \geq c_1\int_0^T \left|\frac{d\eta}{dt}(t)\right|^2 dt - c_2 T \quad .$$

Therefore, for any $r > 0$, the set

$$\{(\Phi,\eta) \in \mathscr{C};\, I_T(\Phi,\eta) \leq r,\, \Phi(0) \text{ belongs to a compact set}\} \quad ,$$

is contained in

$$\begin{aligned}\Bigg\{(\Phi,\eta) \in \mathscr{C};\, &\int_0^T \left|\frac{d}{dt}y_{j(t)}(\Phi(t))\right|^2 dt \\ &+ \int_0^T \left|\frac{d\eta}{dt}(t)\right|^2 dt \leq r',\, \Phi(0) \text{ is in a compact set}\Bigg\}\end{aligned}$$

for some $r' > 0$, and is precompact. It remains to prove that $I_T(\cdot)$ is lower semi-continuous in the uniform topology.

Let $(\Phi^{(n)}, \eta^n) \in \mathscr{C}$ converges to $(\Phi, \eta) \in \mathscr{C}$ uniformly on $[0, T]$ as $n \to \infty$. We assume that $I_T(\Phi^{(n)}, \eta^n)$, $I_T(\Phi, \eta)$ are finite. Denote

$$\Phi^{(n)}(t) = \left(\phi^{(n)}(t), j^{(n)}(t)\right),\; \Phi(t) = (\phi(t), j(t)) \quad .$$

Let fix a $\delta > 0$. We choose $U \subset [0, T]$ which is a union of finitely many open intervals such that

$$U = \cup_{i=1}^K (a_i, b_i),\; \{t; \Phi(t) = 0\} \subset U,\; |U \setminus \{t; \Phi(t) = 0\}| < \delta$$

and

$$\int_{U\setminus\{t;\Phi(t)=0\}} \frac{1}{2}\left|\frac{d\phi}{dt}(t) - b_{j(t)}(\phi(t))\right|^2 dt < \delta \quad ,$$

By a modification of the argument in the proof of [17, Chapter 3, Lemma 2.1], we can show

$$\begin{aligned}\liminf_{n\to\infty} &\int_{U^c} \frac{1}{2}\left|\frac{d\phi^{(n)}}{dt}(t) - b_{j^{(n)}(t)}\left(\phi^{(n)}(t)\right)\right|^2 dt \\ &\geq \int_{U^c} \frac{1}{2}\left|\frac{d\phi}{dt}(t) - b_{j(t)}(\phi(t))\right|^2 dt \quad . \qquad (A1.1)\end{aligned}$$

On the other hand,

$$\int_{U\backslash\{t;\Phi^{(n)}(t)=0\}} \frac{1}{2}|\frac{d\phi^{(n)}}{dt}(t) - b_{j^{(n)}(t)}\left(\phi^{(n)}(t)\right)|^2 dt$$
$$= \int_{U\backslash\{t;\Phi^{(n)}(t)=0\}} \frac{1}{2}\left(|\frac{d\phi^{(n)}}{dt}(t)|^2 + |b_{j^{(n)}(t)}\left(\phi^{(n)}(t)\right)|^2\right) dt$$
$$+ \sum_{i=1}^{K}\left[B_{j^{(n)}(b_i)}\left(\phi^{(n)}(b_i)\right) - B_{j^{(n)}(a_i)}\left(\phi^{(n)}(a_i)\right)\right] .$$

Here

$$B_j(x) = -\int_0^x b_j(u)\,du .$$

By the definition of $\Lambda(\cdot)$, it is easy to see

$$\int_{U\backslash\{t;\Phi^{(n)}(t)=0\}} \frac{1}{2}\left|b_{j^{(n)}(t)}\left(\phi^{(n)}(t)\right)\right|^2 dt \geq \int_{U\backslash\{t;\Phi^{(n)}(t)=0\}} \Lambda\left(\frac{d\eta^{(n)}}{dt}(t)\right) dt . \tag{A1.2}$$

Therefore,

$$\int_{U\backslash\{t;\Phi^{(n)}(t)=0\}} \frac{1}{2}\left|\frac{d\phi^{(n)}}{dt}(t) - b_{j^{(n)}(t)}\left(\phi^{(n)}(t)\right)\right|^2 dt + \int_{\{t;\Phi^{(n)}(t)=0\}} \Lambda\left(\frac{d\eta^{(n)}}{dt}(t)\right) dt$$
$$\geq \int_U \Lambda\left(\frac{d\eta^{(n)}}{dt}(t)\right) dt + \sum_{i=1}^{K}\left[B_{j^{(n)}(b_i)}\left(\phi^{(n)}(b_i)\right) - B_{j^{(n)}(a_i)}\left(\phi^{(n)}(a_i)\right)\right] .$$

We can mimic the proof of [17, Chapter 3, Lemma 2.1] to obtain

$$\liminf_{n\to\infty} \int_U \Lambda\left(\frac{d\eta^{(n)}}{dt}(t)\right) dt \geq \int_U \Lambda\left(\frac{d\eta}{dt}(t)\right) dt . \tag{A1.3}$$

Moreover,

$$\sum_{i=1}^{K}\left[B_{j^{(n)}(b_i)}\left(\phi^{(n)}(b_i)\right) - B_{j^{(n)}(a_i)}\left(\phi^{(n)}(a_i)\right)\right]$$

converges to

$$\sum_{i=1}^{K}[B_{j(b_i)}(\phi(b_i)) - B_{j(a_i)}(\phi(a_i))] .$$

The last quantity is smaller than $c|U \setminus \{t; \Phi(t) = O\}|$ for some constant c. This combining with (A1.1) $\sim$ (A1.3) and by letting $\delta \to 0$ give the desired result that

$$\liminf_{n\to\infty} I\left(\Phi^{(n)}, \eta^{(n)}\right) \geq I(\Phi, \eta) .$$

The proof of Lemma 4.7. is complete.

Appendix 2

In this appendix we prove Lemma 4.1.

First, we know that the invariant density, $P_i(x)$, satisfies the relation

$$\sum_i \int_0^\infty \left(\frac{1}{2}\frac{d^2}{dx^2}F_i(x) - c_i\frac{d}{dx}F_i(x)\right)P_i(x)dx = 0$$

if F_i defined on $(0, \infty)$ is smooth with bounded derivatives and satisfies

$$\sum \beta_i \frac{dF_i}{dx}(0+) = 0 \ .$$

By using suitable test functions, $F_i(x)$, it is easy to prove that

$$P_i(x) = a_i \exp(-2c_i x)$$

with

$$a_i = 2\frac{1}{\sum \frac{\alpha_i}{c_i}}\alpha_i \ .$$

The result

$$\lim_{T\to\infty}\frac{1}{T}\int_0^T \chi_{\{i(t)=i\}}dt = \frac{a_i}{2c_i}, \ \ i = 1, \ldots, N$$

in probability can be proved using the argument in [36, Sections 31,32].

To prove (4.5), we first note

$$P\{\tau_0 \geq t\} = \left(\frac{2}{\pi}\right)^{\frac{1}{2}} \int_0^{\frac{x}{\sqrt{t}}} e^{-\frac{1}{2}y^2}dy \ ,$$

where $\tau_0 = \inf\{t \geq 0; x + W(t) = 0\}$ if $x \geq 0$. This can be proved by using the observation that the desired quantity satisfies the equation

$$\frac{d}{dt}f(t,x) = \frac{1}{2}\frac{d^2}{dx^2}f(t,x),$$

$$f(0,t) = 0, \ \ t > 0,$$

$$f(x,0) = 1, \ \ x > 0 \ ,$$

and $f(t,x) = f(1, \frac{x}{\sqrt{t}})$ holds by the scaling property of the Brownian motion. After this, (4.5) follows by the change of measures using Girsanov theorem:

$$\begin{aligned} P_x\{\tau \geq t\} &= E\left\{e^{-W(t)-\frac{1}{2}t}; \tau_0 \geq t\right\} \\ &= E\left\{e^{-W(\tau_0)-\frac{1}{2}\tau_0}; \tau_0 \geq t\right\} \qquad (A2.1) \\ &= \left(\frac{2}{\pi}\right)^{\frac{1}{2}} e^x \int_t^\infty e^{-\frac{1}{2}s}\frac{x}{2s^{\frac{3}{2}}}e^{-\frac{x^2}{2s}}ds \ . \end{aligned}$$

Now we prove (4.6). Let

$$x_0(t) = x + W(t) - \min_{0\le s\le t}(x + W(s)) \wedge 0 \ .$$

By [25, p.27], the joint distribution of $W(t),\ \max_{0\le s\le t} W(s)$ is given by

$$P\left\{W(t) \in da,\ \max_{0\le s\le t} W(s) \in db\right\}$$
$$= \left(\frac{2}{\pi t^3}\right)^{\frac{1}{2}} (2b-a)\exp\left(-\frac{(2b-a)^2}{2t}\right),\ \ b\ge 0,\ \ b\ge a \ .$$

By changing the probability measure using Girsanov theorem, we have

$$E_x\{f(x(t))\}\exp\left(\frac{1}{2}t\right)$$
$$= E\{f(x_0(t))\exp(-W(t))\}$$
$$= E\left\{f\left(W(t) - \min_{0\le s\le t} W(s)\right)\exp(-W(t)),\, x + \min_{0\le s\le t} W(s) \le 0\right\}$$
$$+ E\left\{f(x + W(t))\exp(-W(t)),\, x + \min_{0\le s\le t} W(s) \ge 0\right\}$$
$$= E\left\{f\left(-W(t) + \max_{0\le s\le t} W(s)\right)\exp(W(t)),\, \max_{0\le s\le t} W(s) \ge x\right\}$$
$$+ E\left\{f(x - W(t))\exp(W(t)),\, \max\nolimits_{0\le s\le t} W(s) \le x\right\} \ .$$

In the last step we use the property that $W(\cdot)$ and $-W(\cdot)$ have the same distribution. Then (4.6) follows from this by an easy calculation.

Acknowledgements. We wish to thank referees for the useful suggestions.

References

1. Alanyali, M. and Hajek, B.: On large deviations of Markov processes with discontinuous statistics. Ann. Appl. Probab. **8**, 45–66, (1998)
2. Alanyali, M. and Hajek, B.: On large deviations in load sharing network. Ann. Appl. Probab. **8**, 67–97, (1998)
3. Barlow, M., Pitman, J. and Yor, M.: On Walsh's Brownian Motion. LN in Math. **No. 1373**, 275–293, (1989)
4. Baxter, J.R. and Chacon, R.V.: The equivalence of diffusions on networks to Brownian motion. Cont. Math. **26**, 33–49, (1984)
5. Below, J.v.: Classical solvability of linear parabolic equations on networks. J. Diff. Equation **72**, 316–337 (1988)

6. Blinovskii, V.M. and Dobrushin, R.L.: Process level large deviations for a class of piecewise homogeneous random walks. In *Dynkin Festschrift: Markov Processes and Their Applications* (1994) 1–59. Birkhauser, Boston
7. Borodin, A.N. and Freidlin, M.: Fast oscillating random perturbations of dynamical systems with conservation laws. Ann. Inst. H. Poincaré (to appear)
8. Chiang, T.Z. and Sheu, S.J.: Large deviation of small perturbation of some unstable systems. Stoch. Analy. Appl. **15**, 31–50, (1997)
9. Chiang, T.Z. and Sheu, S.J.: Large deviations of diffusion processes with discontinuous drift and their occupation times. Ann. Probab. (to appear)
10. Dupuis, P., Ellis, R.S. and Weiss, A.: Large deviations for Markov processes with discontinuous statistics. I: general upper bounds, Ann. Probab. **19**, 1280–1297, (1991)
11. Dupuis, P., Ellis, R.S.: Large deviations for Markov processes with discontinuous statistics. II: random walks, Probab. Theory Relat. Fields **91**, 153–194, (1992)
12. Dupuis, P., Ellis, R.S.: The large deviation principle for a general class of queueing systems I. Trans. AMS. **347**, 2689–2751, (1995)
13. Boue, M., Dupuis, P. and Ellis, R.S.: Large deviations for diffusions with discontinuous statistics. Preprint (1997)
14. Ethier, S.N. and Kurtz, T.G.: *Markov Processes: Characterization and Convergence.* (1986), John Wiley & Sons
15. Frank, D. and Durham, S.: Random motion on binary trees. J. Appl. Prob. **21**, 58–69, (1984)
16. Freidlin, M.: *Markov Processes and Differential Equation: Asymptotic Problems.* (1996), Birkhauser
17. Freidlin, M.I. and Wentzell, A.D.: *Random Perturbations of Dynamical Systems.* (1984), Spring-Verlag
18. Freidlin, M.I. and Wentzell, A.D.: Diffusion processes on graphs and the averaging principle. Ann. Probab. **21**, 2215–2245, (1993)
19. Freidlin, M.I. and Wentzell, A.D.: *Random Perturbations of Hamiltonian Systems.* Memoirs AMS No 523 (1994)
20. Freidlin, M. and Pavlopoulos, H.: On a stochastic model for moisture budget in an Eulerian atmospheric column. Preprint (1996)
21. Friedman, A. and Huang, C.: Diffusion in network. J. Math. Analy. Appl. **183**, 352–384, (1994)
22. Gaveau, B. and Okada, M.: Differential forms and heat diffusion on one dimensional singular varieties. Bull. Sc. Math. **115**, 61–80, (1991)
23. Harrison, J.M. and Shepp, L.A.: On skew Brownian motion. Ann. Probab. **9**, 309–313, (1981)
24. Ikeda, N. and Watanabe, S.: *Stochastic Differential Equations and Diffusion Processes.* (1981), North-Holland
25. Ito, K. and Mckean, H.P.: *Diffusion Processes and Their Sample Paths.* (1965), Springer-Verlag
26. Korostelev, A.P. and Leonov, S.L.: Action functional for diffusion process with discontinuous drift. Theory Probab. Appl. **37**, 543–550, (1992)
27. Korostelev, A.P. and Leonov, S.L.: Action functional for diffusions in discontinuous media. Probab. Theory Relat. Fields **94**, 317–333, (1992)
28. Lumer, G.: Connecting of local operators and evolution equations on networks. LN in Math. **787**, 219–234, (1980)
29. Mandl, P.: *Analytic Treatment of One Dimensional Markov Processes.* (1968), Springer-Verlag
30. Ali Mehmeti, F.: A charaterization of a generalized C^∞-notation on nets. Integral Equations and Operator Theory **9**, 753–766, (1986)
31. Nakao, S.: On the pathwise uniqueness of solutions of one-dimensional stochastic differential equations. Osaka J. Math. **9**, 513–518, (1972)
32. Nicaise, S.: Some results on spectral theory over networks, applied to nerve impulse transmission. LN in Math. **1171**, 532–541, (1985)

33. Okada, T.: Asymptotic behavior of skew conditional heat kernels on graph networks. Canad. J. Math. **45**, 863–878, (1993)
34. Stroock, D.W. and Varadhan, S.R.S.: *Multidimensional Diffusion Processes.* (1979), Springer-Verlag
35. Tsirelson, B.: Walsh process filtration is not Brownian. Preprint (1996)
36. Varadhan, S.R.S.: *Lectures on Diffusion Problems and Partial Differential Equations.* Lectures of Tata Institute (1980), Spinger-Verlag
37. Varadhan, S.R.S.: *Large Deviations and Applications.* CBMS-NSF Regional Conference Series in Applied Mathematics (1984)
38. Walsh, J.: A diffusion with a discontinuous local time. Temps locaux, Asterisque **52-53**, 37–45, (1978)