The Annals of Probability
1993, Vol. 21, No. 2, 1015–1044

LARGE DEVIATIONS FOR MARKOV PROCESSES CORRESPONDING TO PDE SYSTEMS

BY ALEXANDER EIZENBERG[1] AND MARK FREIDLIN[2]

Hebrew University and University of Maryland

We continue the study of the asymptotic behavior of Markov processes $(X^\varepsilon(t), \nu^\varepsilon(t))$ corresponding to systems of elliptic PDE with a small parameter $\varepsilon > 0$. In the present paper we consider the case where the process $(X^\varepsilon(t), \nu^\varepsilon(t))$ can leave a given domain D only due to large deviations from the degenerate process $(X^0(t), \nu^0(t))$. In this way we study the limit behavior of solutions of corresponding Dirichlet problems.

1. Introduction. Consider the Dirichlet problem for the PDE system

$$(1.1)\qquad \begin{aligned} &L_k^\varepsilon u_k^\varepsilon(x) + \sum_{j=1}^n d_{kj}(x)\big(u_j^\varepsilon(x) - u_k^\varepsilon(x)\big) = 0, \quad x \in D \subset \mathbb{R}^r,\\ &u_k^\varepsilon|\partial D = \psi_k, \qquad u_k^\varepsilon \in C^2(D) \cap C(\overline{D}), \qquad k = 1, \dots, n. \end{aligned}$$

Here D is a connected bounded domain with a C^2-class smooth boundary ∂D. The differential operators L_k^ε are given by the formula

$$(1.2)\qquad L_k^\varepsilon u = \frac{\varepsilon}{2} \sum_{i,j=1}^r a_k^{ij}(x) \frac{\partial^2 u}{\partial x^i \partial x^j} + \sum_{i=1}^r b_k^i(x) \frac{\partial u}{\partial x^i}$$

for any $u \in C^2(D)$, $\psi_k \in C(\partial D)$; $(a_k^{ij}(x)) = a_k(x) \in C^2(\mathbb{R}^r)$ are positive definite matrices, $k = 1, \dots, n$; $d_{kj}(x) > 0$, $\varepsilon > 0$ and $b_k^j, d_{kj} \in C^1(\mathbb{R}^r)$. Without loss of generality, we assume that the coefficients as well as their first derivatives remain bounded in $\mathbb{R}^r$. In the present paper we continue the study of the asymptotic behavior of u_k^ε when $\varepsilon \downarrow 0$ that was started in Eizenberg and Freidlin (1990). Recall that the system (1.1) can be associated with a right continuous Markov process $(X^\varepsilon(t), \nu^\varepsilon(t))$ with the phase space $\mathbb{R}^r \times \{1, \dots, n\}$. The first component $X^\varepsilon(t)$ satisfies the stochastic differential equation

$$(1.3)\qquad dX^\varepsilon(t) = \sqrt{\varepsilon}\, \sigma_{\nu^\varepsilon(t)}(X^\varepsilon(t))\, dw(t) + b_{\nu^\varepsilon(t)}(X^\varepsilon(t))\, dt,$$

where $\sigma_k(x)\sigma_k^*(x) = a_k(x)$, $b_k(x) = (b_k^1(x), \dots, b_k^r(x))$, $k = 1, \dots, n$, and $w(t)$ is a Brownian motion in $\mathbb{R}^r$. The second component $\nu^\varepsilon(t)$ is a random process

---

Received June 1990; revised October 1991.

[1]The essential part of the preparation of this work was completed when the first author was Visiting Assistant Professor at University of Maryland, College Park, Maryland.

[2]Research partially supported by NSF Grant DMS-87-21440.

AMS 1991 *subject classifications*. 60F10, 35B25, 35J55.

Key words and phrases. Large deviations, small random perturbations, PDE systems, singular perturbations.

with the states $\{1, \dots, n\}$ such that

$$(1.4)\qquad P\{\nu^\varepsilon(t+\Delta) = j | \nu^\varepsilon(t) = i,\ X^\varepsilon(t) = x\} = d_{ij}(x)\Delta + o(\Delta) \quad \text{uniformly in } \mathbb{R}^r,$$

provided $\Delta \downarrow 0$, $1 \le j, i \le n$, $i \ne j$. Denote

$$(1.5)\qquad \tau^\varepsilon = \min\{t\colon X^\varepsilon(t) \in \partial D\}.$$

The solution $\{u_1^\varepsilon(x), \dots, u_n^\varepsilon(x)\}$ has the stochastic representation [see Eizenberg and Freidlin (1990). Theorem 3]

$$(1.6)\qquad u_k^\varepsilon(x) = E_{x,k}\psi_{\nu^\varepsilon(\tau^\varepsilon)}(X^\varepsilon(\tau^\varepsilon)).$$

[Here, as usual, the subscript x, k means the initial condition $X^\varepsilon(0) = x$, $\nu^\varepsilon(0) = k$.] It was pointed out in Eizenberg and Freidlin (1990) that in contrast to the case of a single equation, in the case of systems the degenerate process $(X^0(t), \nu^0(t))$ is not a dynamical system, but a random process. In the case of a single equation the asymptotic behavior of solutions depends, as is known, on the properties of trajectories of corresponding dynamical systems. Similarly, in the case of systems we can describe several typical situations according to the behavior of the degenerate process $(X^0(t), \nu^0(t))$. Let $\tau^0 = \min\{t\colon X^0(t) \in \partial D\}$ be the exit time of $X^0(t)$ from D. In Eizenberg and Freidlin (1990) we considered the situation where $P_{x,k}\{\tau^0 < \infty\} = 1$. The last assumption is analogous to the Levinson conditions for a single equation. It was shown that in such a case the solution $\{u_1^\varepsilon(x), \dots, u_n^\varepsilon(x)\}$ of (1.1) tends to the solution $\{u_1^0(x), \dots, u_n^0(x)\}$ of the corresponding degenerate system with the same boundary conditions on the regular part of the boundary, as $\varepsilon \downarrow 0$.

In the present paper we will discuss the situations where the trajectories of the process $X^0(t)$ starting from some point $x_0 \in D$ remain in D with a positive probability; that is,

$$(1.7)\qquad P_{x_0,k_0}\{\tau^0 = \infty\} > 0 \quad \text{for some } x_0 \in D,\ 1 \le k_0 \le n.$$

More precisely, we will concentrate our attention on the case where the degenerate process hinders the exit from D and the exit occurs only due to the large deviations of the process $(X^\varepsilon(t), \nu^\varepsilon(t))$ from the degenerate process $(X^0(t), \nu^0(t))$.

This case is a generalization of the situation considered in Wentzell and Freidlin (1970) for the Dirichlet problem with a single equation

$$(1.8)\qquad \frac{\varepsilon}{2}\sum_{i,j}^{r} a^{ij}(x)\frac{\partial^2 u^\varepsilon(x)}{\partial x^i\,\partial x^j} + \sum_{i=1}^{r} b^i(x)\frac{\partial u^\varepsilon(x)}{\partial x^i} = 0, \qquad u^\varepsilon|\partial D = \psi$$

[see also, Freidlin and Wentzell (1984) and Freidlin (1985)], where the vector field $b(x) = (b^1(x), \dots, b^r(x))$ satisfies the condition

$$(1.9)\qquad (n(x), b(x)) < 0 \quad \text{for } x \in \partial D \quad [\text{here } n(x) \text{ is the outward normal to } \partial D]$$

and the corresponding diffusion process $X^\varepsilon(t)$ exits D only as a result of large deviations from the dynamical system $X^0(t)$ defined by

$$\frac{dX^0(t)}{dt} = b(X^0(t)). \tag{1.10}$$

It was shown in Wentzell and Freidlin (1970) that the rough asymptotic properties of the probabilities of large deviations in the case of a single equation are defined by the functional

$$I_T(\varphi) = \tfrac{1}{2}\int_0^T L(\varphi(s), \dot{\varphi}(s))\, ds \tag{1.11}$$

for any $\varphi \in C_{0,T}(\mathbb{R}^r)$ absolutely continuous [otherwise one sets $I_T(\varphi) \equiv \infty$], where $L(x, v) = (a^{-1}(x)(v - b(x)), v - b(x))$ for any $x, v \in \mathbb{R}^r$. More precisely, $(1/\varepsilon)I_T(\varphi)$ is the action functional for the family of measures μ^ε induced by the diffusion processes $X^\varepsilon(t)$, $0 \le t \le T$, on $C_{0,T}(\mathbb{R}^r)$ as $\varepsilon \downarrow 0$. One of the most important results concerning the cases where (1.9) is satisfied can be formulated as follows.

Denote

$$V(x,y) = \inf_{\substack{\varphi:\,[0,T]\to\mathbb{R}^r\\ \varphi(0)=x,\\ \varphi(T)=y\\ T\ge 0}} I_T(\varphi).$$

Suppose that there exists a compact $K \subset D$ that contains all the limit points of (1.10) and

$$V(x,y) = 0 \quad \text{for } x, y \in K.$$

If there exists $y_0 \in \partial D$ such that

$$V(K, y_0) < V(K, y) \quad \text{for } y \in \partial D,\ y \ne y_0$$
$$[\text{here } V(K,y) \equiv V(x,y) \text{ for any } x \in K],$$

then

$$\lim_{\varepsilon \downarrow 0} u^\varepsilon(x) = \psi(y_0) \quad \text{for any} \quad x \in D,$$

where $u^\varepsilon(x)$ is the solution of (1.8).

In the present paper we will develop a similar approach for the case of systems and, in particular, we will study the large deviations of the diffusion component $X^\varepsilon(t)$ from the degenerate random process $X^0(t)$. Such a class of large deviations problems was discussed by Bezuidenhout (1987) for the particular case when the discrete component was a Markov process itself or, in other words, the functions $d_{ij}(x)$ were independent of x.

In the present work we consider the general case of the Markov process $(X^\varepsilon(t), \nu^\varepsilon(t))$ given by (1.3) and (1.4). Similarly to Bezuidenhout (1987), we will assume throughout this paper that $a_k^{ij}(x) = \delta_{ij}$. It seems that the case of general diffusion matrices $a_k(x)$ requires a more sophisticated approach. We will give a simple and explicit representation for the lower semicontinuous version of the action functional for the family $X^\varepsilon(t)$ when $\varepsilon \downarrow 0$.

One can see, however, from the formula (1.6) that it is necessary to know not only the action functional for $X^\varepsilon(t)$, but also the limit behavior of $\nu^\varepsilon(\tau^\varepsilon)$ to get the complete picture concerning the asymptotic behavior of $u_k^\varepsilon(x)$.

We describe the limit behavior for the second component in Section 3, together with the precise formulation of the main results of this paper.

2. The general approach. Through all this work we will assume that the second order terms of the operators L_k^ε are equal to the Laplacian, that is, $a_k^{ij}(x) = \delta_{ij}$ for any $k = 1, \dots, n$. Thus, the stochastic differential equation (1.3) takes the form

$$dX^\varepsilon(t) = \sqrt{\varepsilon}\, dw(t) + b_{\nu^\varepsilon(t)}(X^\varepsilon(t))\, dt. \tag{2.1}$$

Our purpose in this section is to derive uniform large deviation estimates for the process X^ε. Using absolute continuity we reduce the problem to large deviations for the simpler process Z^ε defined in (2.2) and (2.3) and utilize the fact that Z^ε is a continuous functional of the independent processes εw and ν. We accomplish this in three steps. Our first step is to formulate some class of rather abstract Assumptions A–D concerning continuous functionals acting on products of measure spaces and to show that those assumptions are satisfied for the process Z^ε. Next, in Theorem 1 we prove that the assumptions imply some large deviation estimates for the measures induced by such functionals. Finally, in Theorem 2, we extend these estimates to the process X^ε.

Now we pass to details. As in Eizenberg and Freidlin (1990), we will use the fact that the measure $\mu_1^{T,\varepsilon}$ induced by $(X^\varepsilon(t), \nu^\varepsilon(t))$ in the space of trajectories for $0 \le t \le T$ is absolutely continuous with respect to the measure $\mu_2^{T,\varepsilon}$ induced by the Markov process $(Z^\varepsilon(t), \nu(t))$, which can be defined in the following way: $\nu(t)$ is a discrete Markov process on a phase space $\{1, \dots, n\}$ with the transition probabilities

$$P\{\nu(t+\Delta) = k \mid \nu(t) = i\} = \Delta + o(\Delta), \qquad \Delta \downarrow 0 \tag{2.2}$$

for any $i \neq k$, $1 \le i, k \le n$, and the process $Z^\varepsilon(t)$ is the solution of the stochastic differential equation

$$dZ^\varepsilon(t) = \sqrt{\varepsilon}\, dw(t) + b_{\nu(t)}(Z^\varepsilon(t))\, dt, \tag{2.3}$$

where w and ν are assumed to be independent. As was pointed out in Eizenberg and Freidlin [(1990), Section 3], the density is given by the formula

$$\begin{aligned} p_T^\varepsilon(Z^\varepsilon(\cdot), \nu(\cdot)) &= \frac{d\mu_1^{T,\varepsilon}}{d\mu_2^{T,\varepsilon}}(Z^\varepsilon(\cdot), \nu(\cdot)) \\ &= \prod_{i=0}^{n(T)-1} d_{\nu(\eta_i)\nu(\eta_{i+1})}(Z^\varepsilon(\eta_{i+1})) \\ &\quad \times \exp\left(-\sum_{i=0}^{n(T)} \int_{\eta_i}^{\eta_{i+1}\wedge T} \big(d_{\nu(\eta_i)}(Z^\varepsilon(s)) - n + 1\big)\, ds \right), \end{aligned} \tag{2.4}$$

where

$$d_k(x) = \sum_{\substack{l=1 \\ l \neq k}}^{n} d_{kl}(x), \tag{2.5}$$

η_i is the sequence of Markov times defined by

$$\eta_0 = 0, \qquad \eta_{i+1} = \inf\{s > \eta_i : \nu(s) \neq \nu(\eta_i)\} \tag{2.6}$$

and

$$n(T) = \max\{i | \eta_i \leq T\}. \tag{2.7}$$

[If $n(T) = 0$, $p_T^\varepsilon(Z^\varepsilon(\cdot), \nu(\cdot)) = \exp(-\int_0^T (d_{\nu(0)}(Z^\varepsilon(s)) - n + 1)\, ds)$.] Similarly to Eizenberg and Freidlin (1990), at first we will consider the properties of the auxiliary Markov process $(Z^\varepsilon(t), \nu(t))$, and then we will extend these results onto the process $(X^\varepsilon(t), \nu^\varepsilon(t))$.

The Markov processes of type $(Z^\varepsilon(t), \nu(t))$ were studied by Bezuidenhout (1987), where they were considered as a particular case of a more general class of Markov processes $(X^\varepsilon(t), \xi(t))$ such that $\xi(t)$ is a Markov random process itself and $X^\varepsilon(t)$ satisfies the stochastic differential equation

$$dX^\varepsilon(t) = \varepsilon\, dw(t) + b(X^\varepsilon(t), \xi(t))\, dt. \tag{2.8}$$

Here the Brownian motion $w(t)$ is supposed to be independent of $\xi(t)$, the bounded function b: $\mathbb{R}^r \times \mathbb{R}^d \to \mathbb{R}^r$ satisfies Lipschitz conditions and $\mathbb{R}^d$, $d \geq 1$, is a phase space of $\xi(t)$. [The other example of such processes was considered in Freidlin and Gartner (1978), where $\xi(t)$ was assumed to be a Markov process in $\mathbb{R}^d$ with continuous trajectories and $b(x, y)$ was assumed to be continuously differentiable.] Since only part of the estimates that we need can be obtained directly from the results of Bezuidenhout (1987), we prefer to develop a slightly different approach to the processes of form (2.8). This approach is similar to the method of Bezuidenhout (1987); however, it will allow us to apply the results concerning the processes given by (2.2) and (2.3) to a general case of processes $(X^\varepsilon(t), \nu^\varepsilon(t))$.

We will consider the following general situation.

Assumption A. Let Ω be a metric space with a metric ρ. Suppose that the family of probability measures μ^ε is defined on Borel subsets of Ω and that a function S: $\Omega \to [0, \infty]$ is the action functional for μ^ε with respect to the parameter $\varepsilon \downarrow 0$, that is [see Freidlin and Wentzell (1984), Chapter 3]:

(H0) For any $s \geq 0$, the set $\Phi(s) = \{\omega: S(\omega) \leq s\}$ is compact.

(H1) For any $\delta > 0$, $\gamma > 0$, $\omega \in \Omega$, there exists $\varepsilon_1 = \varepsilon_1(\delta, \gamma, \omega) > 0$ such that for $\varepsilon < \varepsilon_1$,

$$\mu^\varepsilon\{\omega' : \rho(\omega', \omega) < \delta\} \geq \exp\left(-\frac{1}{\varepsilon}(S(\omega) + \gamma)\right).$$

(H2) For any $\gamma, \delta > 0$, $s > 0$ there exists $\varepsilon_2 = \varepsilon_2(\gamma, \delta, s)$ such that for $\varepsilon < \varepsilon_2$,

$$\mu^\varepsilon\{\omega' : \rho(\omega', \Phi(s)) \geq \delta\} \leq \exp\left(-\frac{1}{\varepsilon}(s - \gamma)\right).$$

ASSUMPTION B. Let F and G be two metric spaces with the metrics ρ_F and ρ_G, respectively, and X be some set of parameters x. Suppose that m is some probability measure on Borel subsets of F. Let $\mathscr{L}_x$: $\Omega \times F \to G_x \subset G$ be a family of operators Lipschitz continuous uniformly with respect to $x \in X$:

$$\rho_G(\mathscr{L}_x(\omega_1, f_1), \mathscr{L}_x(\omega_2, f_2)) \leq C_0 \max(\rho_F(f_1, f_2), \rho(\omega_2, \omega_2)) \tag{2.9}$$

for any $f_1, f_2 \in F$, $\omega_1, \omega_2 \in \Omega$, $x \in X$. Let $\tilde{\mu}^\varepsilon$ be the family of measures on $\Omega \times F$ defined by $\tilde{\mu}^\varepsilon = \mu^\varepsilon \times m$. Then one can define a family of the measures P_x^ε on $G \times F$, $\varepsilon > 0$, $x \in X$ using the operators $\mathscr{L}_x$: for any Borel sets $A \subset G$, $B \subset F$ and for any $x \in X$, $\varepsilon > 0$,

$$P_x^\varepsilon(A \times B) = \tilde{\mu}^\varepsilon\{(\omega', f') : \mathscr{L}_x(\omega', f') \in A,\ f' \in B\}. \tag{2.10}$$

Moreover we assume that $G = \bigcup_{x \in X} G_x$ and G_x are disjoint for different x. [Clearly, $P_x^\varepsilon(G_x \times F) = 1$.]

To make the situation clear, let us show the connection between Assumptions A and B and the processes of form $(Z^\varepsilon(t), \nu(t))$. Notice, that although we are interested in obtaining our estimates uniformly with respect to initial conditions (x, i), we should not worry about the second component because it admits only a finite number of possible initial values. Thus, without loss of generality, we can fix $\nu(0)$, for instance $\nu(0) = 1$. For a given $T > 0$ choose F to be the space of trajectories of $\nu(t)$, that is, the set of right continuous functions ν: $[0, T] \to \{1, \dots, n\}$, $\nu(0) = 1$, with the metric ρ_F given by the L_1 norm on F. Set

$$\Omega = \{\psi(\cdot) \in C_{0,T}(\mathbb{R}^r) : \psi(0) = 0\}$$

with the family of measures μ^ε induced by the process $\varepsilon w(t)$ $[w(0) = 0]$ and with the metric $\rho = \rho_G$ given by the uniform norm in $C_{0,T}(\mathbb{R}^r)$. Next, let $G = C_{0,T}(\mathbb{R}^r)$, $X = \mathbb{R}^r$ and $G_x = \{\varphi(\cdot) \in G : \varphi(0) = x\}$. Finally, we define the measure m on F as the measure induced by the trajectories of the process $\nu(t)$ under the initial condition $\nu(0) = 1$.

Each trajectory $\varphi(t)$ of the process $Z^\varepsilon(t)$ with the initial condition $\varphi(0) = x \in X = \mathbb{R}^r$ which corresponds to given trajectories $\psi(t)$ and $\nu(t)$ of the processes $\varepsilon w(t)$ and $\nu(t)$, respectively [we will designate the process $\nu(t)$ and its trajectories in the same way], is the unique solution of the integral equation

$$\varphi(t) = x + \int_0^t b_{\nu(s)}(\varphi(s))\, ds + \psi(t), \qquad 0 \leq t \leq T. \tag{2.11}$$

The equation (2.11) defines the operator

$$\mathscr{L}_x : \Omega \times F \to G_x$$

as follows:

$$\mathscr{L}_x(\psi(\cdot), \nu(\cdot)) = \varphi(\cdot). \tag{2.12}$$

The measures P_x^ε defined in (2.10) are exactly the measures induced by the process $(Z^\varepsilon(t), \nu(t))$ on $C_{0,T}(\mathbb{R}^r) \times F$, corresponding to initial conditions $Z^\varepsilon(0) = x$. Notice that the family of measures μ^ε in our case satisfies the conditions (H0)–(H2) with respect to the parameter $\varepsilon \downarrow 0$, where

$$S(\psi(\cdot)) = \tfrac{1}{2}\int_0^T |\dot{\psi}(s)|^2\, ds \tag{2.13}$$

[see Freidlin and Wentzell (1984)].

Finally, using the Gronwall inequality and the fact that b_i are smooth, one can show immediately in the standard way that (2.9) holds, where the Lipschitz constant C_0 certainly depends on $T > 0$.

Another example where Assumptions A and B are satisfied is the case of Freidlin and Gartner (1978) mentioned above.

To obtain the main results of this section we will need some additional assumptions concerning our general model.

ASSUMPTION C. Suppose that for any $f \in F$, $g \in G_x$, $x \in X$ there exists a unique $\omega = \omega_x(f, g) \in \Omega$ such that

$$\mathscr{L}_x(\omega_x(f, g), f) = g. \tag{2.14}$$

Denote for any $g \in G$ and $f \in F$,

$$S(f, g) \equiv S(\omega_x(f, g)), \tag{2.15}$$

where x is defined by $g \in G_x$ and

$$\tilde{S}(g) = \inf_{f \in F} S(f, g). \tag{2.16}$$

We will suppose here that for any given g such that $\tilde{S}(g) < \infty$, $S(f, g)$ is continuous with respect to f.

Particularly, in our special case of the process $(Z^\varepsilon(t), \nu(t))$, one has for any sample paths $\varphi(\cdot)$ and $\nu(\cdot)$ of $Z^\varepsilon(t)$ and $\nu(t)$, respectively,

$$\omega_x(\nu(\cdot), \varphi(\cdot)) = \psi(\cdot) \in C_{0,T}(\mathbb{R}^r), \tag{2.17}$$

where $\psi(\cdot)$ is defined by

$$\psi(t) = \varphi(t) - x - \int_0^t b_{\nu(s)}(\varphi(s))\, ds, \tag{2.18}$$

as follows from (2.11). Thus, by (2.17), (2.18) and (2.13) one has

$$S(\nu(\cdot), \varphi(\cdot)) = \tfrac{1}{2}\int_0^T |\dot{\varphi}(s) - b_{\nu(s)}(\varphi(s))|^2\, ds \tag{2.19}$$

and

$$\tilde{S}(\varphi(\cdot)) = I_T(\varphi) \equiv \inf_{\nu(\cdot)} \tfrac{1}{2}\int_0^T |\dot{\varphi}(t) - b_{\nu(t)}(\varphi(t))|^2\,dt. \tag{2.20}$$

[We will use the more conventional designation $I_T(\varphi)$ when we formulate the main results for our process $(Z^\varepsilon(t), \nu(t))$.] It is clear that if $\tilde{S}(\varphi(\cdot)) < \infty$, then

$$\int_0^T |\dot{\varphi}(t)|^2\,dt < \infty \tag{2.21}$$

and one can easily show using the Cauchy–Schwarz inequality that $S(\nu(\cdot), \varphi(\cdot))$ is continuous with respect to $\nu(\cdot)$.

Next, we will formulate our last assumption.

Assumption D. Suppose that there exists a sequence of Borel subsets $F_k \subseteq F$, $F_k \subset F_{k+1}$, $k \geq 1$, such that for any $k, \alpha > 0$, $f_0 \in F_k$,

$$m(f' \in F_k\colon \rho_F(f', f_0) < \alpha) \geq d(k, \alpha) > 0, \tag{2.22a}$$

where $d(k, \alpha)$ is independent of f_0. Moreover, for any $\delta, s > 0$ there exists $k = k(\delta, s)$ such that for any $f \in F$, $g \in G_x$, $x \in X$ satisfying $S(f, g) \leq s$ one can find $\tilde{f} \in F_k$ and $\tilde{g} \in G_x$ such that

$$\rho_G(g, \tilde{g}) \leq \delta, \tag{2.22b}$$

$$S(\tilde{f}, \tilde{g}) \leq S(f, g) + \delta. \tag{2.22c}$$

We will concentrate our results concerning the connection between the process $(Z^\varepsilon(t), \nu(t))$ and our general approach in the following proposition.

Proposition 1. *Let F be the set of right continuous functions ν: $[0, T] \to \{1, \dots, n\}$ with the metric ρ_F defined by the L_1 norm and with the measure m induced by the process $\nu(t), \nu(0) = 1$ and choose $F_k = \{\nu(\cdot) \in F\colon N(T, \nu(\cdot)) \leq k\}$, where*

$$N(T, \nu(\cdot)) = \max\{i\colon \eta_i \leq T\} \tag{2.23}$$

and $\eta_i = \eta_i(\nu(\cdot))$ are the "jumps" of the trajectory $\nu(\cdot)$ defined as in (2.6).

Set $\Omega = \{\psi(\cdot) \in C_{0T}(\mathbb{R}^r)\colon \psi(0) = 0\}$, $X = \mathbb{R}^r$, $G = C_{0,T}(\mathbb{R}^r)$ and $G_x = \{\varphi \in G\colon \varphi(0) = x\}$.

Let ρ and ρ_G be the uniform metrics on Ω and G, respectively, and μ^ε be the family of measures induced by the processes $\varepsilon w(t)$ on Ω. Define the family of operators $\mathscr{L}_x$ as in (2.11) *and* (2.12).

Then the measures P_x^ε defined in (2.10) *are exactly the measures induced by $(Z^\varepsilon(t), \nu(t))$ on the space of trajectories, Assumptions A–D are satisfied, and the functionals $S(w)$, $S(f, g)$ and $\tilde{S}(g)$ have the forms* (2.13), (2.19) *and* (2.20), *respectively.*

Proof. In light of our discussion just following Assumptions B and C we need to prove only that Assumption D is satisfied. First, we will prove (2.22a).

For a given $k \geq 1$ and $\nu_0(\cdot) \in F_k$ denote $U(\beta) = \bigcup_{i=1}^{N(T,\nu_0(\cdot))}[\eta_i - \beta, \eta_i + \beta]$, where $N(T, \nu_0(\cdot))$ was defined in (2.23). In general, the segments $[\eta_i - \beta, \eta_i + \beta]$ are not disjoint. It can be seen immediately that one can represent $U(\beta)$ in the form

$$U(\beta) = \bigcup_{i=1}^{j_0} A_i, \tag{2.24}$$

where A_i are pairwise disjoint segments and each one of them is a union of some segments $[\eta_i - \beta, \eta_i + \beta]$. Therefore, $A_i = [a_i, a'_i]$, such that $a_{i+1} > a'_i$, $1 \leq i \leq j_0 - 1$, $1 \leq j_0 \leq k$ and

$$|A_i| > 2\beta, \qquad \sum_{i=1}^{j_0} |A_i| \leq 2\beta k, \qquad 1 \leq i \leq j_0, \tag{2.25}$$

where $|A_i|$ is the length of the segment A_i.

Denote the segments A_i such that $\nu_0(a_i) \neq \nu_0(a'_i)$ by $B_1, \dots, B_l$, where $B_i = [b_i, b'_i]$ and $b_{i+1} > b'_i$.

Define the function $\tilde{\nu}_0(t)$ on $[0, T] \setminus \bigcup_{i=1}^{l} B_i$ in the following way:

$$\tilde{\nu}_0(t) = \begin{cases} \nu_0(t), & t \in [0, T] \setminus U(\beta), \\ \nu_0(a_i), & t \in A_i \text{ such that } \nu_0(a_i) = \nu_0(a'_i). \end{cases} \tag{2.26}$$

Notice that the points a_i, a'_i, $1 \leq i \leq j_0$, cannot coincide with any of jump points η_i.

Therefore, it is clear that $[0, T] \setminus \bigcup_{j=1}^{l} B_i$ is a union of some disjoint intervals and semiintervals, and $\tilde{\nu}_0(t)$ is constant on each one of them. Then according to elementary properties of the process $\nu(t)$ and (2.25), one has

$$P\{\mathscr{A}(\nu_0)\} \geq e^{-T}(\beta)^k \tag{2.27}$$

provided $0 < \beta \leq \beta_0$, where $\mathscr{A}(\nu_0)$ is the event that the process $\nu(t)$ is equal to $\tilde{\nu}_0(t)$ on $[0, T] \setminus \bigcup_{i=1}^{l} B_i$ and that all the jumps of $\nu(t)$ can occur on the segments B_i only, not more than one jump on each segment B_i.

On the other hand,

$$\begin{aligned} &m\{\nu(\cdot) \in F_k : \rho_F(\nu(\cdot), \nu_0(\cdot)) < \alpha\} \\ &\quad = P\left\{N(T, \nu(\cdot)) \leq k, \int_0^T |\nu(s) - \nu_0(s)|\, ds < \alpha\right\} \\ &\quad \geq P\{\mathscr{A}(\nu_0)\}, \end{aligned} \tag{2.28}$$

provided $\beta = \alpha/2kn$, which together with (2.27) proves (2.22a).

Now we will prove (2.22b) and (2.22c). Let $s > 0$. Suppose that

$$\nu(\cdot) \in F \quad \text{and} \quad \varphi(\cdot) \in C_{0,T}(\mathbb{R}^r), \qquad \varphi(0) = x$$

are such that

$$S(\nu(\cdot), \varphi(\cdot)) \leq s.$$

It follows immediately from (2.19) and the last inequality that

$$\int_0^T |\dot{\varphi}(t)|^2\,dt \le 4s + 4T \sup_{\substack{1\le i\le n\\ y\in\mathbb{R}^r}} |b_i(y)|^2 \tag{2.29}$$

and thus, $\varphi(t)$ satisfies the Hölder conditions of order $1/2$ [see, e.g., Freidlin and Wentzell (1984), proof of Lemma 2.1b, Chapter 3]. More precisely,

$$|\varphi(t_1) - \varphi(t_2)| \le C(s)|t_1 - t_2|^{1/2} \tag{2.30}$$

for $0 \le t_1, t_2 \le T$, where $C(s)$ depends on T, but is independent of $\varphi(\cdot)$ and $\nu(\cdot)$. For any given $m > 0$ let us break $[0, T]$ into m segments of length T/m. On each segment $jT/m \le t \le (j+1)T/m$ we have

$$\left| b_i(\varphi(t)) - b_i\left(\varphi\left(\frac{jT}{k}\right)\right)\right| \le C_1(s)\left(\frac{T}{m}\right)^{1/2}, \tag{2.31}$$

where

$$C_1(s) = C_2 C(s), \qquad C_2 = \sup_{\substack{x\in\mathbb{R}^r\\ 1\le i\le n}} \left|\frac{\partial b_i}{\partial x}\right|, \qquad 0 \le j \le m-1,\ 1 \le i \le n.$$

Thus, by (2.19) and (2.30),

$$\begin{aligned}
&\left| S(\nu(\cdot), \varphi(\cdot)) - \frac{1}{2}\sum_{j=0}^{m-1} \int_{jT/m}^{(j+1)T/m} \left|\dot{\varphi}(t) - b_{\nu(t)}\left(\varphi\left(\frac{jT}{m}\right)\right)\right|^2 dt \right| \\
&\qquad \le \frac{1}{2} C_1(s)\left(\frac{T}{m}\right)^{1/2}\left(\int_0^T |\dot{\varphi}(t)|\,dt + 2T \max_{\substack{y\in\mathbb{R}^r\\ 1\le i\le n}} |b_i(y)|\right) \\
&\qquad \le C_2(s) m^{-1/2},
\end{aligned} \tag{2.32}$$

where $C_2(s)$ depends on s and T only. [In the right inequality we also use (2.29) and the Cauchy–Schwarz inequality.]

Let us introduce the following functions: for $1 \le i \le n$,

$$\chi^i(t) = \begin{cases} 1, & \nu(t) = i, \\ 0, & \text{otherwise}, \end{cases} \tag{2.33}$$

and denote

$$\gamma_i^j = \int_{jT/m}^{(j+1)T/m} \chi^i(t)\,dt \quad \text{for } 0 \le j \le m-1,\ 1 \le i \le n. \tag{2.34}$$

Define $\nu_m\colon [0, T] \to \{1, \dots, n\}$ in the following way:

$$\nu_m(t) = i \quad \text{for } t \in \Delta_i^j,\ 0 \le j \le m-1,\ 1 \le i \le n, \tag{2.35}$$

where

$$\Delta_i^j = \left[\frac{jT}{m} + \sum_{k=0}^{i-1}\gamma_k^j, \frac{jT}{m} + \sum_{k=0}^{i}\gamma_k^j\right) \tag{2.36}$$
$$\left(\bigcup_{i=1}^{n}\Delta_i^j = \left[\frac{jT}{m}, \frac{(j+1)T}{m}\right) \quad \text{since } \sum_{i=1}^{n}\gamma_i^j = \frac{T}{m}, 0 \le j \le m-1\right);$$

Clearly, $\nu_m(t)$ is defined for any $t \in [0,1]$. Notice that some Δ_i^j can be empty, but

$$\nu_m(0) = 1.$$

Next, we define φ_m by

$$\dot{\varphi}_m(t) = \dot{\varphi}(t) + b_{\nu_m(t)}\left(\varphi\left(\frac{jT}{m}\right)\right) - b_{\nu(t)}\left(\varphi\left(\frac{jT}{m}\right)\right), \tag{2.37}$$
$$\text{for } t \in \left[\frac{jT}{m}, \frac{(j+1)T}{m}\right), \varphi_m(0) = \varphi(0).$$

One can verify directly that

$$\varphi_m\left(\frac{jT}{m}\right) = \varphi\left(\frac{jT}{m}\right) \quad \text{for any } 0 \le j \le m-1. \tag{2.38}$$

Moreover,

$$\max_{0 \le t \le T}|\varphi_m(t) - \varphi(t)| \le M_0T/m, \tag{2.39}$$

where by M_i we will denote constants depending on the fields b_i only.

It follows from (2.37) and (2.30) that the functions φ_m satisfy the Hölder condition with the same constant of the form $C_3(s) = C(s)(1 + M_1T^{1/2})$. Therefore, similarly to (2.32) one has

$$\left|S(\nu_m(\cdot), \varphi(\cdot)) - \frac{1}{2}\sum_{j=0}^{m-1}\int_{jT/m}^{(j+1)T/m}\left|\dot{\varphi}_m(t) - b_{\nu_m(t)}\left(\varphi_m\left(\frac{jT}{m}\right)\right)\right|^2 dt\right| \le C_4(s)m^{-1/2}. \tag{2.40}$$

However, by (2.38) and (2.37), for $0 \le t \le T$, $0 \le j \le m-1$,

$$\dot{\varphi}_m(t) - b_{\nu_m(t)}\left(\varphi_m\left(\frac{jT}{m}\right)\right) = \dot{\varphi}(t) - b_{\nu(t)}\left(\varphi\left(\frac{jT}{m}\right)\right).$$

The last equality together with (2.40) and (2.32) implies

$$|S(\nu(\cdot), \varphi(\cdot)) - S(\nu_m(\cdot), \varphi_m(\cdot))| \le C_5(s)m^{-1/2}. \tag{2.41}$$

Let $\delta > 0$. Choose $m = m(s, \delta)$ large enough such that $\max(C_5(s)m^{-1/2},$

$M_0Tm^{-1}) \le \delta$. Set $\tilde{\nu} = \nu_m \in F_{mn}$ and $\tilde{\varphi} = \varphi_m$. Then by (2.39) and (2.41) we obtain a much stricter statement than (2.22b) and (2.22c) setting $k(s,\delta) = nm(s,\delta)$. □

Now we will formulate the main results of this section for our general model. Denote $S^*(g) = \liminf_{g' \to g} \tilde{S}(g') \wedge \tilde{S}(g)$. The functional $S^*(g)$ is lower semicontinuous.

Set $\Phi^*(s) = \{g: S^*(g) \le s\}$, $\Phi_x^*(s) = \Phi^*(s) \cap G_x$.

THEOREM 1. (a) *For any* $\delta, \gamma, s > 0$ *there exist* $\varepsilon_1 = \varepsilon_1(s,\delta,\gamma)$ *and* $M = M(s,\gamma,\delta)$ *such that*

$$(2.42)\quad P_x^\varepsilon\{(g', f'): \rho_G(g', g) \le \delta \text{ and } f' \in F_M\} \ge \exp\left(-\frac{S^*(g)+\gamma}{\varepsilon}\right)$$

provided $g \in \Phi_x^*(s)$, $x \in X$, $\varepsilon \le \varepsilon_1$.

(b) *For any* $\delta > 0$, $\gamma > 0$, $s > 0$, *there exists* $\varepsilon_2 = \varepsilon_2(\delta,\gamma,s)$ *such that*

$$(2.43)\qquad P_x^\varepsilon\{g: \rho_G(g, \Phi^*(s)) \ge \delta\} \le \exp\left(-\frac{s-\gamma}{\varepsilon}\right)$$

for $\varepsilon \le \varepsilon_2$, $x \in X$.

PROOF. (a) By the definitions of $\tilde{S}$ and S^*, for any β: $0 < \beta < 1$ and for any $g \in \Phi_x(s)$ there exist $g_1 = g_1(g)$ and $f \in F$ such that

$$(2.44)\qquad S = S(f, g_1) \le \tilde{S}(g_1) + \frac{\beta}{2} \le S^*(g) + \beta \le s + \beta$$

and

$$(2.45)\qquad \rho_G(g_1, g) \le \beta.$$

On the other hand, by (2.22) there exist $\tilde{g}_1 \in \Phi_x(s)$ and $\tilde{f} \in F_k$, $k = k(\beta, s)$, such that $\rho(\tilde{g}_1, g_1) < \beta$ and

$$(2.46)\qquad S\big(w_x(\tilde{f}, \tilde{g}_1)\big) = S\big(\tilde{f}, \tilde{g}_1\big) \le S(f, g_1) + \beta.$$

By (H1) there exists $\varepsilon_3 = \varepsilon_3(\alpha, \beta, s)$ such that for $\varepsilon \le \varepsilon_3$,

$$(2.47)\qquad \begin{aligned} \mu^\varepsilon\{\omega': \rho(\omega', \omega_x(f, g_1)) \le \alpha\} &\ge \exp\left(-\frac{S(\omega_x(f,g_1)) + \beta}{\varepsilon}\right) \\ &\ge \exp\left(-\frac{S^*(g) + 3\beta}{\varepsilon}\right). \end{aligned}$$

[Notice that ε_3 is independent of x, g and f by Theorem 3.2, Chapter 3 of Freidlin and Wentzell (1984).] The right inequality in (2.47) follows from (2.44) and (2.46).

Choose $\beta = \min(\delta, \gamma)/4$. Thus, by (2.10), (2.14) and (2.9), the definition of $\tilde{\mu}^{\varepsilon}$, (2.22) and (2.44)–(2.47),

$$\begin{aligned}
&P_x^{\varepsilon}\{(g', f'): \rho_G(g', g) \le \delta,\ f' \in F_k\} \\
&\quad \ge P_x^{\varepsilon}\left\{(g', f'): \rho_G(g, \tilde{g}_1) \le \frac{\delta}{3},\ f' \in F_k\right\} \\
&\quad = \tilde{\mu}^{\varepsilon}\left\{(\omega', f'): \rho_G\Big(\mathscr{L}_x(\omega', f'), \mathscr{L}_x\big(\omega_x(\tilde{f}, \tilde{g}_1), \tilde{f}\big)\Big) \le \frac{\delta}{3},\ f' \in F_k\right\} \\
&\quad \ge \tilde{\mu}^{\varepsilon}\left\{(\omega', f'): \rho\big(\omega', \omega_x(\tilde{f}, \tilde{g}_1)\big) \le \frac{\delta}{3C_0},\ \rho(f', \tilde{f}') \le \frac{\delta}{3C_0},\ f' \in F_k\right\} \\
&\quad = \mu^{\varepsilon}\left(\omega': \rho(\omega', \omega_x(f, \tilde{g}_1)) \le \frac{\delta}{3C_0}\right) m\left\{f': \rho(f', \tilde{f}) \le \frac{\delta}{3C_0},\ f' \in F_k\right\} \\
&\quad \ge \exp\left(-\frac{S^*(g) + 3\beta}{\varepsilon}\right) d\left(k, \frac{\delta}{3C_0}\right) \\
&\quad \ge \exp\left(-\frac{S^*(g) + \gamma}{\varepsilon}\right)
\end{aligned} \tag{2.48}$$

provided $\varepsilon \le \varepsilon_4(\delta, \gamma, s) = \min(\varepsilon_3(\alpha, \beta, s), \beta|\ln d(k, \delta/3C_0)|)$. [Here we set $\alpha = \delta/3C_0$ in (2.47).]

Setting $M(s, \gamma, \delta) = k(\beta, s)$, we complete the proof.

(b) Denote $\tilde{\Phi}(s) = \{g \in G: \tilde{S}(g) \le s\}$. It is clear that

$$\tilde{\Phi}(s) \subset \Phi^*(s)$$

and thus

$$P_x^{\varepsilon}\{g: \rho_G(g, \Phi^*(s)) \ge \delta\} \le P_x^{\varepsilon}\{g: \rho_G(g, \tilde{\Phi}(s)) \ge \delta\}. \tag{2.49}$$

Next, for any $\omega \in \Phi(s)$, $f \in F$ and $x \in X$ one has, according to definitions (2.14) and (2.15),

$$\omega = \omega_x(f, \mathscr{L}_x(\omega, f)) \tag{2.50}$$

and thus

$$S(f, \mathscr{L}_x(\omega, f)) = S(\omega) \le s. \tag{2.51}$$

Therefore, by (2.16), $\mathscr{L}_x(\omega, f) \in \tilde{\Phi}(s)$ for any $x \in X$. Then, if $g' \in G_x$ is such that $\rho_G(g', \tilde{\Phi}_x(s)) \ge \delta$, we have in particular that

$$\rho_G(g', \mathscr{L}_x(\omega, f)) \ge \delta \tag{2.52}$$

for any $x \in X$, $f \in F$ and $\omega \in \Phi(s)$.

On the other hand, $g' = \mathscr{L}_x(\omega_x(f, g'), f)$ and then, by (2.9),

$$\rho_G(g', \mathscr{L}_x(\omega, f)) \le C_0 \rho(\omega_x(f, g'), \omega) \tag{2.53}$$

for any $\omega \in \Phi(s)$. Thus, by (2.50)–(2.52) and (2.41), taking into account (2.10), (2.14) and (H2),

$$\begin{aligned}
&P_x^\varepsilon\left\{(g', f'): \rho_G\left(g', \tilde{\Phi}_x(s)\right) \geq \delta\right\} \\
&\quad \leq P_x^\varepsilon\left\{(g', f'): \rho\left(\omega_x(g', f'), \Phi(s)\right) \geq \frac{\delta}{C_0}\right\} \\
&\quad = \tilde{\mu}^\varepsilon\left\{(\omega', f'): \rho\left(\omega_x\left(\mathscr{L}_x(\omega', f'), f'\right), \Phi(s)\right) \geq \frac{\delta}{C_0}\right\} \\
&\quad = \tilde{\mu}^\varepsilon\left((\omega', f'): \rho(\omega', \Phi(s)) \geq \frac{\delta}{C_0}\right) \\
&\quad = \mu^\varepsilon\left(\omega': \rho(\omega', \Phi(s)) \geq \frac{\delta}{C_0}\right) \leq \exp\left(-\frac{s-\gamma}{\varepsilon}\right)
\end{aligned}$$

for ε small enough, which together with (2.49) proves (2.43). □

Before we formulate the consequences of the last theorem for our special case, let us find the explicit representation of the lower semicontinuous modification $S^*(\varphi(\cdot)) \equiv I_T^*(\varphi)$ for $\tilde{S}(\varphi(\cdot)) = I_T(\varphi)$ defined in (2.20).

LEMMA 1. *The functional*

$$I_T^*(\varphi) = \begin{cases} \frac{1}{2}\int_0^T R(\dot{\varphi}(t), \varphi(t))\, dt, & \text{if } \varphi \text{ is absolutely continuous}, \\ \infty, & \text{otherwise}, \end{cases} \tag{2.54}$$

where

$$R(v, x) = \min_{\substack{\alpha_1, \dots, \alpha_n \geq 0 \\ \sum_{i=1}^n \alpha_n = 1}} \left| v - \sum_{i=1}^n \alpha_i b_i(x) \right|^2,$$

for any $x, v \in \mathbb{R}^r$, is the lower semicontinuous modification of $I_T(\varphi)$ for $T > 0$, that is,

$$I_T^*(\varphi) = \liminf_{\varphi' \to \varphi} I_T(\varphi') \wedge I_T(\varphi). \tag{2.55}$$

PROOF. It is clear that

$$I_T^*(\varphi) \leq I_T(\varphi) \quad \text{for any } \varphi \in C_{0,T}(\mathbb{R}^r). \tag{2.56}$$

Let us prove that

$$I_T^*(\varphi) \geq \liminf_{\varphi' \to \varphi} I_T(\varphi'). \tag{2.57}$$

The last inequality trivially holds if φ is not absolutely continuous or if $\dot\varphi \notin L_2([0,T])$. Now we are going to show that for any absolutely continuous φ with L_2-derivative one can find a sequence $\varphi_m \in C_{0,T}(\mathbb{R}^r)$ such that $\varphi_m \to \varphi$ uniformly as $m \to \infty$ and

$$I_T^*(\varphi) \ge \liminf_{m\to\infty} I_T(\varphi_m). \tag{2.58}$$

Since $C^2_{0,T}(\mathbb{R}^r)$ is dense in $\{\varphi \in C_{0,T}(\mathbb{R}^r)\colon \dot\varphi \in L_2\}$ with respect to the norm $|||\cdot|||$, defined by

$$|||\varphi||| = \|\varphi\|_\infty + \|\dot\varphi\|_2,$$

and $I_T^*(\varphi)$ is obviously continuous with respect to $|||\cdot|||$, it is enough to prove the existence of $\{\varphi_m\}_{n=1}^\infty$ satisfying (2.58) for $\varphi \in C^2_{0,T}(\mathbb{R}^r)$. Suppose, therefore, that $\varphi \in C^2_{0,T}(\mathbb{R}^r)$. One can easily see that the function $R(\dot\varphi(t), \varphi(t))$ satisfies Lipschitz conditions with constant

$$C_1 = (2\gamma_1(\varphi) + C_3)(\gamma_2(\varphi) + C_3\gamma_1(\varphi)) \tag{2.59}$$

where

$$C_3 = 2\max_{\substack{x\in\mathbb{R}^r\\ 1\le i\le n}}(|\nabla b_i|, |b_i|), \qquad \gamma_1(\varphi) = \max_{0\le t\le T}|\dot\varphi|, \qquad \gamma_2(\varphi) = \max_{0\le t\le T}|\ddot\varphi|.$$

By compactness reasons, for any $t \in [0,T]$ there exist $\alpha_i = \alpha_i(t)$, $i = 1,\dots,n$, $\sum_{i=1}^n \alpha_i = 1$ such that

$$R(\dot\varphi(t), \varphi(t)) = \left|\dot\varphi(t) - \sum_{i=1}^n \alpha_i b_i(\varphi(t))\right|^2. \tag{2.60}$$

Then, breaking $[0,T]$ into m equal segments for a given $m \ge 1$, we obtain

$$\left| I_T^*(\varphi) - \frac{1}{2}\sum_{j=0}^m \left|\dot\varphi\left(\frac{jT}{m}\right) - \sum_{i=1}^n \alpha_i^j b_i\left(\varphi\left(\frac{jT}{m}\right)\right)\right|^2 \frac{T}{m}\right| \le \frac{C_1 T^2}{2m}, \tag{2.61}$$

where $\alpha_i^j = \alpha_i(jT/m)$, $0 \le j \le m-1$, $1 \le i \le n$.

Define $\nu_m\colon [0,T] \to \{1,\dots,n\}$ in the following way:

$$\nu_m(t) = i \quad \text{for} \quad t \in \Delta_i^j,\ 0 \le j \le m-1,\ 1 \le i \le n, \tag{2.62}$$

where

$$\Delta_i^j = \left[\frac{jT}{m} + \frac{T}{m}\sum_{k=0}^{i-1}\alpha_k^j, \frac{jT}{m} + \frac{T}{m}\sum_{k=0}^{i}\alpha_k^j\right) \tag{2.63}$$

$$\left(\bigcup_{i=1}^n \Delta_i^j = \left[\frac{jT}{m}, \frac{(j+1)T}{m}\right) \quad \text{since } \sum_{i=1}^n \alpha_i^j = 1, 0 \le j \le m-1\right),$$

and thus, $\nu_m(t)$ is defined for any $t \in [0,1]$).

Now we can define $\varphi_m(t)$. Set

$$\dot{\varphi}_m(t) = \dot{\varphi}(t) + b_{\nu_m(t)}\left(\varphi\left(\frac{jT}{m}\right)\right) - \sum_{i=1}^{n} \alpha_i^j b_i\left(\varphi\left(\frac{jT}{m}\right)\right)$$

$$\text{for } t \in \left[\frac{jT}{m}, \frac{(j+1)T}{m}\right), \varphi_m(0) = \varphi(0). \tag{2.64}$$

One can verify directly that

$$\varphi_m\left(\frac{jT}{m}\right) = \varphi\left(\frac{jT}{m}\right) \quad \text{for any } 1 \le j \le m-1$$

and that

$$\max_{0 \le t \le T} |\varphi_m(t) - \varphi(t)| \le \frac{C_3 T}{m}, \quad \text{where } C_3 \text{ is defined in (2.59).} \tag{2.65}$$

Further,

$$\begin{aligned} I_T(\varphi_m) &\le \frac{1}{2}\int_0^T |\dot{\varphi}_m(t) - b_{\nu_m(t)}(\varphi_m(t))|^2\, dt \\ &= \frac{1}{2}\sum_{j=0}^{m-1}\int_{jT/m}^{((j+1)T)/m} \left| \dot{\varphi}(t) - \sum_{i=1}^{n} \alpha_i^j b_i\left(\varphi\left(\frac{jT}{m}\right)\right) \right. \\ &\qquad \left. - b_{\nu_m(t)}(\varphi_m(t)) + b_{\nu_m(t)}\left(\varphi\left(\frac{jT}{m}\right)\right)\right|^2 dt \\ &\le \frac{1}{2}\sum_{j=0}^{m-1}\int_{jT/m}^{((j+1)T)/m} \left| \dot{\varphi}(t) - \sum_{i=1}^{n} \alpha_i^j b_i\left(\varphi\left(\frac{jT}{m}\right)\right)\right|^2 dt + \frac{C_4}{m} \end{aligned} \tag{2.66}$$

for some $C_4 > 0$ depending on φ [here we take into account the smoothness of $b(x)$ and $\varphi(t)$].

Finally, by (2.66) and (2.61),

$$\begin{aligned} I_T(\varphi_m) &\le \frac{1}{2}\sum_{j=0}^{m-1}\left|\dot{\varphi}\left(\frac{jT}{m}\right) - \sum_{i=1}^{n} \alpha_i^j b_i\left(\varphi\left(\frac{jT}{m}\right)\right)\right|^2 \frac{T}{m} \\ &\quad + \frac{1}{2}\sum_{j=1}^{m-1}\int_{jT/m}^{((j+1)T)/m}\left|\dot{\varphi}(t) - \dot{\varphi}\left(\frac{jT}{m}\right)\right| \\ &\quad \times \left|\dot{\varphi}\left(\frac{jT}{m}\right) + \dot{\varphi}(t) + 2\sum_{i=1}^{n} \alpha_i^j b_i\left(\varphi\left(\frac{jT}{m}\right)\right)\right| dt + \frac{C_4}{m} \\ &\le I_T^*(\varphi) + \frac{C_5}{m} \end{aligned} \tag{2.67}$$

for some $C_5 = C_5(\varphi)$, which together with (2.65) proves (2.58), and therefore (2.57). Thus

$$\liminf_{\varphi' \to \varphi} I_T(\varphi') \le I_T^*(\varphi) \le I_T(\varphi). \tag{2.68}$$

To complete the proof of the lemma, we need to show that I_T^* is lower semicontinuous. According to Ioffe and Tikhomirov (1978), it is enough to prove that $R(v, x)$ is convex with respect to v. Indeed, for any $\alpha_i, \alpha_i' \ge 0$ such that $\sum_{i=1}^n \alpha_i = \sum_{i=1}^n \alpha_i' = 1$ and for any β: $0 \le \beta \le 1$, $x, v_1, v_2 \in \mathbb{R}^r$, one has

$$\begin{aligned} &R(\beta v_1 + (1-\beta) v_2, x) \\ &\quad \le \left| \beta v_1 + (1-\beta) v_2 - \sum_{i=1}^n (\alpha_i \beta + \alpha_i'(1-\beta)) b_i(x) \right|^2 \\ &\quad = \left| \beta \left(v_1 - \sum_{i=1}^n \alpha_i b_i(x) \right) + (1-\beta) \left(v_2 - \sum_{i=1}^n \alpha_i' b_i(x) \right) \right|^2 \\ &\quad \le \beta \left| v_1 - \sum_{i=1}^n \alpha_i b_i(x) \right|^2 + (1-\beta) \left| v_2 - \sum_{i=1}^n \alpha_i' b_i(x) \right|^2. \end{aligned} \tag{2.69}$$

Since α_i and α_i' are arbitrary, we can take the minimum over all possible $\{\alpha_i\}_{i=1}^n$ and $\{\alpha_i'\}_{i=1}^n$, which proves that $R(v, x)$ is really convex, that is, $I_T^*(\varphi)$ is lower semicontinuous. □

REMARK. Generally speaking, $I_T^*(\varphi) \ne I_T(\varphi)$, as follows from the example given in Bezuidenhout (1987).

Now we can formulate our results concerning large deviations for the process $(X^\varepsilon(t), \nu^\varepsilon(t))$.

THEOREM 2. (i) *The set* $\Phi_x^*(s) = \{I^*(\varphi) \le s\}$ *is compact for any* $s > 0$, $x \in \mathbb{R}^r$.

(ii) *For any* $\delta > 0$, $\gamma > 0$, $s > 0$ *there exists* $\varepsilon_5 = \varepsilon_5(s, \delta, \gamma)$ *and* $M = M(s, \gamma, \delta)$ *such that*

$$P_{x,i}\left\{ \sup_{0 \le t \le T} |X^\varepsilon(t) - \varphi(t)| \le \delta \text{ and } \tilde{n}(T) \le M(s, \gamma, \delta) \right\} \ge \exp\left(-\frac{I^*(\varphi) + \gamma}{\varepsilon} \right)$$

for any $\varphi \in \Phi_x^*(s)$, $\varepsilon \le \varepsilon_5$, $x \in \mathbb{R}^r$, $1 \le i \le n$, *where* $\tilde{n}(T)$ *is the number of jumps of* $\nu^\varepsilon(t)$ *defined as in* (2.6) *and* (2.7).

(iii) *For any* $\delta, \gamma, s > 0$ *there exists* $\varepsilon = \varepsilon_6(s, \delta, \gamma)$ *such that*

$$P_{x,i}\{\rho_{0T}(X^\varepsilon(\cdot), \Phi^*(s)) \ge \delta\} \le \exp\left(-\frac{s-\gamma}{\varepsilon} \right)$$

for $\varepsilon \le \varepsilon_6$, $x \in \mathbb{R}^r$, $1 \le i \le n$, *where* $\rho_{0,T}$ *is the uniform metric on* $C_{0,T}(\mathbb{R}^r)$.

PROOF. (i) One can prove the compactness of $\Phi_x^*(s)$ in a completely similar way to the case of a simple equation, using the facts that I_T^* is lower semicontinuous and that $\int_0^T |\dot{\varphi}|^2\,dt$ is bounded on the set $\Phi(s)$ and the Arzela theorem.

(ii) Without loss of generality we can assume that

$$0 < \min_{\substack{1\le i,j\le n\\ x\in\mathbb{R}^r}} d_{ij} \le \max_{\substack{1\le i,j\le n\\ x\in\mathbb{R}^r}} d_{ij} \le 1 \text{ in } \mathbb{R}^r. \tag{2.70}$$

Denote

$$Q = \left\{ \sup_{0\le t\le T} |Z^\varepsilon(t) - \varphi(t)| \le \delta \text{ and } n(T) \le M(s,\gamma,\delta) \right\},$$

where $n(T)$ is the number of jumps of $\nu(t)$. On Q it holds by (2.4) and (2.70) that

$$p_T^\varepsilon(Z^\varepsilon(\cdot),\nu(\cdot)) \ge \left(\min_{\substack{1\le i,j\le n\\ x\in\mathbb{R}^r}} d_{ij} \right)^{M(s,\gamma,\delta)} \equiv C(s,\gamma,\delta) > 0. \tag{2.71}$$

Thus, by (2.61),

$$\begin{aligned} P_{x,i}&\left\{ \sup_{0\le t\le T} |X^\varepsilon(t) - \varphi(t)| \le \delta \text{ and } \tilde{n}(T) \le M(s,\gamma,\delta) \right\} \\ &= E_{x,i}\chi_Q p_T^\varepsilon(Z^\varepsilon(\cdot),\nu(\cdot)) \\ &\ge C(s,\gamma,\delta) P_{x,i}\{Q\}. \end{aligned} \tag{2.72}$$

Here χ_Q means the indicator of Q.

However, by Theorem 1(a) and Proposition 1, one can immediately see that

$$P_{x,i}\{Q\} \ge \exp\left(-\frac{I_T^*(\varphi) + \gamma/2}{\varepsilon} \right) \tag{2.73}$$

for ε small enough and for any $\varphi \in \Phi^*(s)$, which is compact according to (i). The estimates (2.72) and (2.73) together prove the statement (ii).

(iii) Under assumption (2.70), p_T^ε is bounded from above. Thus

$$\begin{aligned} P_{x,i}&\{\rho_{0,T}(X^\varepsilon(\cdot),\Phi_x^*(s)) \ge \delta\} \\ &= Ep_T^\varepsilon(Z^\varepsilon(\cdot),\nu(\cdot))\chi_{\{\rho_{0T}(Z^\varepsilon(\cdot),\Phi^*(s))\ge\delta\}} \\ &\le C_6 P\{\rho_{0T}(Z^\varepsilon(\cdot),\Phi_x{}^*(s))\} \le \exp\left(-\frac{s-\gamma}{\varepsilon} \right) \end{aligned} \tag{2.74}$$

for ε small enough [by Proposition 1 and Theorem 1b]. The fact that the estimates are uniform with respect to $1 \le i \le n$ is trivial. □

3. The main results. Denote for given $x, y \in D$,

$$V(x,y) = \inf_{\varphi,T} I_T^*(\varphi), \tag{3.1}$$

where the infinum is taken over all

$$\varphi \in C_{0,T}(\mathbb{R}^r), \qquad T > 0: \varphi(0) = x, \qquad \varphi(T) = y.$$

In this section we suppose that:

ASSUMPTION 1. There exists a compact $\mathscr{K}_0 \subset D$ containing all the limit points of the dynamical systems

$$\frac{dS_i^t x}{dt} = b_i(S_i^t x), \qquad S_i^0 x = x, \qquad 1 \le i \le n, \tag{3.2}$$

such that

$$V(x,y) = V(y,x) = 0 \tag{3.3}$$

for any $x, y \in \mathscr{K}_0$.

ASSUMPTION 2. $(n(x), b_i(x)) < 0$ for $x \in \partial D$, $1 \le i \le n$, where $n(x)$ is the outward normal at $x \in \partial D$.

It is clear that, as in the case of a single equation, for any $x, y, z \in \mathbb{R}^r$ it holds that

$$V(x,y) \le V(x,z) + V(z,y). \tag{3.4}$$

Thus, for any $y \in \mathbb{R}^r$ one can define

$$V(\mathscr{K}_0, y) = V(x,y) \quad \text{for any } x \in \mathscr{K}_0. \tag{3.5}$$

It follows from (3.4) that $V(x,y)$ is continuous with respect to x, y.
Our next assumption is:

ASSUMPTION 3. There exists $x_0 \in \partial D$ such that

$$V(\mathscr{K}_0, x_0) < V(\mathscr{K}_0, x) \quad \text{for any } x \in \partial D,\ x \ne x_0. \tag{3.6}$$

In the second part of our main theorem we will need the following assumption.

ASSUMPTION 4. There exists i_0: $1 \le i_0 \le n$ such that

$$g_{i_0}(x_0) < g_i(x_0) \quad \text{for } 1 \le i \le n,\ i \ne i_0, \tag{3.7}$$

where x_0 has been defined in Assumption 3 and

$$g_i(x) \equiv -(b_i(x), n(x)) > 0, \qquad x \in \partial D,\ 1 \le i \le n. \tag{3.8}$$

Now we can formulate our main results.

THEOREM 3. (i) *Under Assumptions* 1–3,

$$\lim_{\varepsilon \downarrow 0} P_{x,k}\{|X^\varepsilon(\tau^\varepsilon) - x_0| > \delta\} = 0 \tag{3.9}$$

for any $\delta > 0$, $1 \le k \le n$, *uniformly on compact* $K \subset D$.

(ii) *If, in addition, Assumption* 4 *is satisfied, then*

$$\lim_{\varepsilon \downarrow 0} P_{x,k}\{\nu^\varepsilon(\tau^\varepsilon) = i_0\} = 1 \tag{3.10}$$

for $1 \le i \le n$ *uniformly on compact* $K \subset D$.

COROLLARY . *Let Assumptions* 1–4 *hold. Then*

$$\lim_{\varepsilon \downarrow 0} u_k^\varepsilon = \psi_{i_0}(x_0), \qquad 1 \le k \le n, \tag{3.11}$$

uniformly on compact $K \subset D$, *where* $u_k^\varepsilon(x)$ *are the solutions of the Dirichlet problem* (1.1) *with* $a_k^{ij}(x) \equiv \delta_{ij}$.

REMARK. Similar results also can be proved for general a_k^{ij}; however, one cannot find such a convenient representation of lower-semicontinuous modification for the action functional as (2.43).

PROOF OF THEOREM 3. The proof of (i) is based on Theorem 2 in a completely standard way, similar to the case of single equations [see Wentzell and Freidlin (1970) and Freidlin (1985)] and we will leave it to the reader. Notice, that although X^ε is not, in general, a Markov process, its hitting times are stopping times with respect to the Markov process $(X^\varepsilon, \nu^\varepsilon)$. Therefore we can use the standard technique of hitting times sequences. We use a similar technique in a much more complicated situation in the proof of (ii) just below. For example, the formula (3.16) illustrates the use of the strong Markov property in our circumstances.

(ii) Without loss of generality we can set $i_0 = 1$ in Assumption 4.

Choose $\delta_0 > 0$ such that

$$\min_{\substack{x \in \partial D: |x - x_0| \le \delta_0 \\ 2 \le i \le n}} g_i(x) - g_1(x) \equiv a_0 > 0. \tag{3.12}$$

Let $\gamma = \gamma(\varepsilon)$ be such that $\gamma = \exp(-a_0\gamma/\varepsilon)$ (one can easily see that such γ exists). Denote

$$\Gamma_1^\varepsilon = \{x \in D: \operatorname{dist}(x, \partial D) = \gamma\},$$

$$\Gamma_2^\varepsilon = \{x \in D: \operatorname{dist}(x, \partial D) = 2\gamma\},$$

$$K_1^\varepsilon = \Gamma_1^\varepsilon \cap \left\{x: |x - x_0| \le \frac{\delta_0}{2}\right\} \ne \phi, \quad \text{for } \varepsilon \text{ small enough},$$

$$K_2^\varepsilon = \Gamma_1^\varepsilon \setminus K_1^\varepsilon.$$

Define $\tilde{\tau}^\varepsilon = \inf\{t > 0: X^\varepsilon(t) \in K_2^\varepsilon\} \wedge \tau^\varepsilon$. Then for any $i \in \{1, \dots, n\}$, $x \in K \subset D$,

$$P_{x,i}\{\nu^\varepsilon(\tau^\varepsilon) \ne 1\} \le P_{x,i}\{\tilde{\tau}^\varepsilon < \tau^\varepsilon\} + P_{x,i}\{\nu(\tau^\varepsilon) \ne 1 \text{ and } \tilde{\tau}^\varepsilon = \tau^\varepsilon\}. \tag{3.13}$$

Notice, that for $\varepsilon > 0$ it follows from Assumption 3 and from the continuity of $V(x, y)$ that

$$V(\mathscr{K}_0, x_0) = \min_{x \in \partial D} V(\mathscr{K}_0, x) < \inf_{x \in K_2^\varepsilon} V(\mathscr{K}_0, x) \quad \text{for } \varepsilon \text{ small enough.} \tag{3.14}$$

Thus, it can be proved, again by the standard way, that

$$\lim_{\varepsilon \downarrow 0} P_{x,i}\{\tilde{\tau}^\varepsilon < \tau^\varepsilon\} = 0 \quad \text{uniformly on compacts in } D. \tag{3.15}$$

Thus, our main problem is to estimate the second term of (3.13).

Define sequences of Markov times

$$\begin{aligned} \sigma_1^\varepsilon &= \inf\{t \geq 0 \colon X^\varepsilon(t) \in \Gamma_1^\varepsilon\} \wedge \tau^\varepsilon, \\ \tau_m^\varepsilon &= \inf\{t \geq \sigma_m^\varepsilon \colon X^\varepsilon(t) \in \partial D \cup \Gamma_2^\varepsilon\}, \\ \sigma_{m+1}^\varepsilon &= \inf\{t \geq \tau_m^\varepsilon \colon X^\varepsilon(t) \in \Gamma_1^\varepsilon\} \wedge \tau^\varepsilon. \end{aligned}$$

Set $m_0^\varepsilon = \min\{m \colon \tau_m^\varepsilon = \tau^\varepsilon\}$. Clearly, $m_0^\varepsilon < \infty$ a.e. and

$$\tau_m^\varepsilon = \sigma_{m+1}^\varepsilon = \tau^\varepsilon \quad \text{for } m \geq m_0^\varepsilon.$$

Thus

$$\begin{aligned} &P_{x,i}\{\nu^\varepsilon(\tau^\varepsilon) \neq 1 \text{ and } \tilde{\tau}^\varepsilon = \tau^\varepsilon\} \\ &\quad = \sum_{m=1}^{\infty} P_{x,i}\{\nu^\varepsilon(\tau^\varepsilon) \neq 1, \tilde{\tau}^\varepsilon = \tau^\varepsilon, m_0^\varepsilon = m\} \\ &\quad \leq \sum_{m=1}^{\infty} P_{x,i}\{\nu^\varepsilon(\tau^\varepsilon) \neq 1, X^\varepsilon(\tau_m^\varepsilon) \in \partial D \text{ and } X^\varepsilon(\sigma_m^\varepsilon) \in K_1^\varepsilon\} \\ &\quad = E_{x,i} \sum_{m=1}^{\infty} \chi_{\{X^\varepsilon(\sigma_m^\varepsilon) \in K_1^\varepsilon\}} P_{X^\varepsilon(\sigma_m^\varepsilon), \nu^\varepsilon(\sigma_m^\varepsilon)}\{X^\varepsilon(\tau_1^\varepsilon) \in \partial D, \nu^\varepsilon(\tau_1^\varepsilon) \neq 1\}. \end{aligned} \tag{3.16}$$

Similarly,

$$\begin{aligned} 1 &= \sum_{m=1}^{\infty} P_{x,i}\{m_0^\varepsilon = m\} \\ &= E_{x,i} \sum_{m=1}^{\infty} \chi_{\{X^\varepsilon(\sigma_m^\varepsilon) \in \Gamma_1^\varepsilon\}} P_{X^\varepsilon(\sigma_m^\varepsilon), \nu^\varepsilon(\sigma_m^\varepsilon)}\{X^\varepsilon(\tau_1^\varepsilon) \in \partial D\}. \end{aligned} \tag{3.17}$$

(Remember that χ_A is the indicator of A.) According to (3.13), (3.15), (3.16) and (3.17) it is enough to show

$$\sup_{\substack{x \in K_1^\varepsilon \\ 1 \leq j \leq n}} P_{x,j}\{\nu^\varepsilon(\tau_1^\varepsilon) \neq 1 | X^\varepsilon(\tau_1^\varepsilon) \in \partial D\} \to 0 \tag{3.18}$$

as $\varepsilon \downarrow 0$. To prove (3.18) we will reduce the problem to the one-dimensional case. It is a well known fact that the function

$$\tilde{\rho}(x) = \begin{cases} \operatorname{dist}(x, \partial D), & x \in \partial D \cup D, \\ -\operatorname{dist}(x, \partial D), & x \notin D, \end{cases}$$

is smooth for x: $\mathrm{dist}(x, \partial D) \le \rho_0$ for some $\rho_0 > 0$ small enough, and that

$$\nabla\tilde{\rho}(x) = -n(x) \quad \text{for} \quad x \in \partial D. \tag{3.19}$$

Clearly, one can introduce a function $\rho(x)$ smooth and bounded with bounded first and second derivatives in $\mathbb{R}^r$ such that

$$\rho(x) = \tilde{\rho}(x) \quad \text{for } x\text{: } \mathrm{dist}(x, \partial D) < \frac{\rho_0}{2} \tag{3.20}$$

and

$$|\rho(x)| > \frac{\rho_0}{2} \quad \text{for } x\text{: } \mathrm{dist}(x, \partial D) > \frac{\rho_0}{2}.$$

Then $\{\rho(x) = \alpha\} = \{\tilde{\rho}(x) = \alpha\}$ for α: $|\alpha| \le \rho_0/2$. Let us consider the random processes $F^\varepsilon(t)$ and $\psi^\varepsilon(t)$ defined by

$$F^\varepsilon(t) = \gamma^{-1}\rho(Z^\varepsilon(\gamma t)) \tag{3.21}$$

and

$$\psi^\varepsilon(t) = \nu(\gamma t), \tag{3.22}$$

where the process $(Z^\varepsilon(t), \nu(t))$ is defined as in (2.2) and (2.3). Using the Itô formula one can easily verify that for any $t > 0$,

$$\begin{aligned} F^\varepsilon(t) = F^\varepsilon(0) &+ \varepsilon^{1/2}\gamma^{-1/2}\int_0^t (\nabla\rho(Z^\varepsilon(\gamma s)), d\tilde{w}(s)) \\ &+ (\varepsilon/2)\int_0^t \Delta\rho(Z^\varepsilon(\gamma s))\, ds \\ &+ \int_0^t \big(\nabla\rho(Z^\varepsilon(\gamma s)), b_{\psi^\varepsilon(s)}(Z^\varepsilon(\gamma s))\big)\, ds, \end{aligned} \tag{3.23}$$

where $\tilde{w}(t) = \gamma^{-1/2}w(\gamma t)$ is a Brownian motion in $\mathbb{R}^r$. Notice that by the definition of γ,

$$\frac{\varepsilon}{\gamma} \sim \frac{1}{|\ln \varepsilon|}. \tag{3.24}$$

The transition probabilities for $\psi^\varepsilon(t)$ are

$$P\{\psi^\varepsilon(t+h) = i | \psi^\varepsilon(t) = j\} = \gamma h + o(\gamma h) \quad \text{for } i \ne j, t, h \ge 0. \tag{3.25}$$

Together with the process $F^\varepsilon(t)$ we will consider the one-dimensional Markov random process

$$Y_y^\varepsilon(t) = Y_y^\varepsilon(0) - \int_0^t \big(b_{\psi^\varepsilon(s)}(y), n(y)\big)\, ds + \varepsilon^{1/2}\gamma^{-1/2}w_y(t), \tag{3.26}$$

where $y \in \partial D$ and the one-dimensional Brownian motion $w_y(t)$ is defined by

$$w_y(t) = -(n(y), \tilde{w}(t)). \tag{3.27}$$

Now we will stop the proof of Theorem 3 for a moment and will prove the following.

PROPOSITION 2. *If* $Y_y^\varepsilon(0) = F^\varepsilon(0)$ *and* $|Z^\varepsilon(0) - y| \le \gamma$ *for some* $y \in \partial D$, *then*

$$P\Big\{\sup_{0 \le s \le t} |Y_y^\varepsilon(s) - F^\varepsilon(s)| \ge \delta\Big\} \le \exp\big(-C_{10}(t)\delta\gamma^{-1/2}\big) \tag{3.28}$$

for $\varepsilon \le \varepsilon(\delta, t)$, *where* $\varepsilon(\delta, t)$ *is independent of* y.

PROOF. By (3.19), (3.23) and (3.26), for any $t > 0$,

$$\begin{aligned} F^\varepsilon(s) - Y_y^\varepsilon(s) &= (\varepsilon/2)\int_0^t \Delta\rho(Z^\varepsilon(\gamma s))\, ds + \varepsilon^{1/2}\gamma^{-1/2}\int_0^t \big(V_y(Z^\varepsilon(\gamma s)), d\tilde{w}(s)\big) \\ &\quad + \int_0^t \big(b_{\psi^\varepsilon(s)}(Z^\varepsilon(\gamma s)), \nabla\rho(Z^\varepsilon(\gamma s))\big) - \big(b_{\psi^\varepsilon(s)}(y), \nabla\rho(y)\big)\, ds, \end{aligned} \tag{3.29}$$

where

$$V_y(x) \equiv \nabla\rho(x) - \nabla\rho(y). \tag{3.30}$$

Thus

$$\begin{aligned} \sup_{0 \le s \le t} |Y_y^\varepsilon(s) - F^\varepsilon(s)| &\le (\varepsilon/2)t \max_{x \in \mathbb{R}^r}|\Delta\rho(z)| \\ &\quad + \sup_{0 \le s \le t} \varepsilon^{1/2}\gamma^{-1/2}\left|\int_0^s \big(V_y(Z^\varepsilon(\gamma u)), d\tilde{w}(u)\big)\right| \\ &\quad + C_7 t \sup_{0 \le s \le t} |Z^\varepsilon(\gamma s) - y|, \end{aligned} \tag{3.31}$$

where $C_7 > 0$ depends only on the upper bounds of the first and second derivatives of $b_i(x)$ and $\rho(x)$. According to the assumption of the proposition, one has by definition of $Z^\varepsilon(\gamma u)$,

$$\begin{aligned} \sup_{0 \le u \le t} |Z^\varepsilon(\gamma u) - y| &\le |y - Z^\varepsilon(0)| + \sup_{0 \le u \le t} |Z^\varepsilon(\gamma u) - Z^\varepsilon(0)| \\ &\le \gamma + \gamma t \max_{\substack{1 \le i \le n \\ x \in \mathbb{R}^r}} |b_i(x)| + \varepsilon^{1/2}\gamma^{1/2} \sup_{0 \le u \le t} |\tilde{w}(u)|. \end{aligned} \tag{3.32}$$

By the standard martingale estimates for Brownian motion we derive from (3.32),

$$P\left\{\sup_{0\le s\le t}|Z^\varepsilon(\gamma s)-y|\ge\frac{\delta}{3C_7t}\right\}$$
(3.33)
$$\le\exp\left(-\frac{\delta^2C_8(t)}{\varepsilon\gamma}\right)\quad\text{for }\varepsilon\text{ small enough.}$$

Now we will estimate the second term of (3.31). By the exponential martingale inequality [e.g., Friedman (1976), Theorem 7.5], one has for any $\lambda>0$,

$$P\left\{\varepsilon^{1/2}\gamma^{-1/2}\sup_{0\le s\le t}\left|\int_0^s(V_y(Z^\varepsilon(\gamma u)),d\tilde w(u))\right|\ge\frac{\delta}{3}\right\}$$
$$\le P\left\{\sup_{0\le s\le t}\left|\int_0^s(V_y(Z^\varepsilon(\gamma u)),d\tilde w(u))-\frac{\lambda}{2}\int_0^s|V_y(Z^\varepsilon(\gamma u))|^2\,du\right|\right.$$
(3.34)
$$\left.\ge\frac{1}{6}\varepsilon^{-1/2}\gamma^{1/2}\delta\right\}$$
$$+P\left\{\sup_{0\le s\le t}\left|\int_0^s|V_y(Z^\varepsilon(\gamma u))|^2\,du\right|\ge\frac{1}{3}\varepsilon^{-1/2}\gamma^{1/2}\lambda^{-1}\delta\right\}$$
$$\le\exp\left(-\frac{1}{6}\lambda\varepsilon^{-1/2}\gamma^{1/2}\delta\right)+P\left\{t\sup_{0\le u\le t}|V_y(Z^\varepsilon(\gamma u))|^2\ge\frac{1}{3}\varepsilon^{-1/2}\gamma^{1/2}\lambda^{-1}\delta\right\}.$$

However, by (3.30), the smoothness of $\nabla\rho$ and (3.32), using the standard martingale estimates for $\tilde w(t)$ and setting

$$\lambda=\varepsilon^{-1/2},\tag{3.35}$$

one has for some $C_9, C_{10}(t)>0$ and ε small enough:

$$P\left\{\sup_{0\le u\le t}|V_y(Z^\varepsilon(\gamma u))|^2\ge\tfrac{1}{3}t^{-1}\varepsilon^{-1/2}\gamma^{1/2}\lambda^{-1}\delta\right\}$$
(3.36)
$$\le P\left\{\sup_{0\le u\le t}|\tilde w(u)|^2\ge C_9t^{-1}\gamma^{-1/2}\varepsilon^{-1}\delta\right\}$$
$$\le\exp\left(-C_{10}(t)\varepsilon^{-1}\gamma^{-1/2}\delta\right).$$

Now, by (3.31), (3.33), (3.34), (3.35) and (3.36) we obtain the statement of the proposition. □

[The estimate (3.28) is sufficient for our purposes, but, certainly, one can get a stronger estimate.]

We will need the following simple fact concerning the process $Y_y^\varepsilon(s)$: For any $A > 0$ there exists $T = T(A)$ large enough such that for any $y \in \partial D$: $|y - x_0| \le \delta_0$, the initial condition $Y_y^\varepsilon(0) \in [0, 2]$ and ε small enough it holds that

$$P\Big\{\sup_{0 \le s \le T} Y_y^\varepsilon(s) \le 3\Big\} \le \exp(-A\gamma\varepsilon^{-1}). \tag{3.37}$$

Indeed, for any trajectory of $\psi^\varepsilon(t)$, one has

$$Y_y^\varepsilon(0) - \int_0^t \big(b_{\psi^\varepsilon(s)}(y), n(y)\big)\, ds \ge a_1 t, \tag{3.38}$$

where $a_1 = \min_{1 \le i \le n,\, y: |y - x_0| \le \delta_0} |(b_i(y), n(y))|$. Thus, for T large enough and ε small enough,

$$\begin{aligned} P\Big\{\sup_{0 \le s \le T} Y_y^\varepsilon(s) \le 3\Big\} &\le P\big\{P|w_y(T)| \ge \gamma^{1/2}\varepsilon^{-1/2}(a_1 T - 3)\big\} \\ &\le \exp\Big(-\gamma\varepsilon^{-1}\big((a_1 T - 3)^2 (2T)^{-1} - 1\big)\Big) \\ &\le \exp(-A\gamma\varepsilon^{-1}) \end{aligned} \tag{3.39}$$

according to the standard martingale estimates, which proves (3.37).

We will use also processes of the form

$$x_{a,y}^\varepsilon(t) = \varepsilon^{1/2}\gamma^{-1/2} w_y(t) + \beta + at \qquad \text{for given } a > 0,\ \beta \in \mathbb{R},\ y \in \partial D, \tag{3.40}$$

where the Brownian motion $w_y(t)$ was defined in (3.27). The processes $x_{a,y}^\varepsilon(t)$ are an almost trivial particular case of diffusion processes for single equations with a small parameter, and we can apply to them the well known large deviations estimates of Wentzell and Freidlin (1970). The action functional for $x_{a,y}^\varepsilon(t)$ has the form

$$I_T^a(\varphi) = \tfrac{1}{2}\int_0^T |\dot\varphi(t) - a|^2\, dt \quad \text{for } \varphi: [0, T] \to \mathbb{R} \tag{3.41}$$

(if φ is not absolutely continuous, then, as usual, $I_T^a(\varphi) = \infty$). Simple analysis of the proof of Theorem 3.1 of Freidlin and Wentzell [(1984), Chapter 4] for this particular case shows that for any $\beta > \beta_1$ it holds that

$$V^a(\beta, \beta_1, T) \equiv \inf_{\substack{\varphi \in C_{0,T}(\mathbb{R}) \\ \varphi(0) = \beta,\ \varphi(t) = \beta_1 \\ \text{for some } t \in [0, T]}} I_T^a(\varphi) = 2a(\beta - \beta_1), \tag{3.42}$$

provided $T \ge (\beta - \beta_1)/a$ [the extremal is $\varphi_0(t) = -at + \beta,\ 0 \le t \le (\beta - \beta_1/a)$]. Moreover, for any $\beta \in (\beta_1, \beta_2)$, $T \ge 0$, $\beta_2 > \beta_1$,

$$\inf_{\varphi \in \Phi} I_T^a(\varphi) = V^a(\beta, \beta_1, T), \tag{3.43}$$

where

$$\Phi = \{\varphi \in C_{0,T}(\mathbb{R}): \varphi(0) = \beta; \exists\, s \le T \text{ such that } \varphi(s) = \beta_1 \text{ and } \varphi(u) \ne \beta_2, \text{ for } 0 \le u \le s\} \tag{3.44}$$

(since the extremal $\varphi_0 \in \Phi$). Thus, under the initial condition $x^{\varepsilon}_{a,y}(0) = \beta$ one has by the standard Wentzell and Freidlin estimates,

$$\lim_{\varepsilon \downarrow 0} -\varepsilon\gamma^{-1} \ln P\left\{x^{\varepsilon}_{a,y}\left(\tau^{\varepsilon}_{a,y}(\beta_1, \beta_2)\right) = \beta_1 \text{ and } \tau^{\varepsilon}_1 \le T\right\} = 2a(\beta - \beta_1), \tag{3.45}$$

where $\tau^{\varepsilon}_{a,y}(\beta_1, \beta_2)$ is the exit time of $x^{\varepsilon}_{a,y}(t)$ from $[\beta_1, \beta_2]$ and T is large enough.

Now we can return to the proof of the theorem. Our goal is to prove (3.18), that is, to estimate $P_{x,j}\{\nu^{\varepsilon}(\tau^{\varepsilon}_1) \ne 1 | X^{\varepsilon}(\tau^{\varepsilon}_1) \in \partial D\}$ for different $x \in K^{\varepsilon}_1$ and j. For a given $x \in K^{\varepsilon}_1$ one can find $y \in \partial D$ such that $|x - y| \le \gamma$. Then by (3.12) for ε small enough it holds that

$$g_i(y) - g_1(y) \ge a_0 \quad \text{for } i \ge 2. \tag{3.46}$$

Denote by σ^{ε} the exit time of $F^{\varepsilon}(t)$ from $[0,2]$ and by $\tau^{\varepsilon}(y; \beta_1, \beta_2)$ the exit times of $Y^{\varepsilon}_y(s)$ from $[\beta_1, \beta_2]$, $\beta_2 > \beta_1$. Notice that if $\psi^{\varepsilon}(t) = i$ for each $t \in [t_1, t_2]$ and $Y^{\varepsilon}_y(t_1) = x^{\varepsilon}_{g_i(y),y}(t_1)$, then

$$Y^{\varepsilon}_y(t) = x^{\varepsilon}_{g_i(y),y}(t), \quad \text{provided } t \in [t_1, t_2]. \tag{3.47}$$

Under the initial conditions $X^{\varepsilon}(0) = x$, $\nu^{\varepsilon}(0) = 1$, it holds by (2.4), (2.70), (3.21), (3.28), (3.45), (3.47), and (3.25) and the fact that $w_y(t)$ and $\psi^{\varepsilon}(t)$ are independent that

$$\begin{aligned}
&P\{X^{\varepsilon}(\tau^{\varepsilon}_1) \in \partial D\} \\
&\quad \ge E p^{\varepsilon}_{T\gamma}(Z^{\varepsilon}(\cdot), \nu(\cdot)) \chi_{\{Z^{\varepsilon}(\tilde{\tau}^{\varepsilon}_1) \in \partial D,\, \tilde{\tau}^{\varepsilon}_1 \le T\gamma;\, \nu(t) = 1 \text{ for each } t \in [0, \gamma T]\}} \\
&\quad \ge \exp\left(-\gamma T\left(\max_{z \in \mathbb{R}^r}(d_1(z)) - n + 1\right)\right) \\
&\qquad \times P\{F^{\varepsilon}(\sigma^{\varepsilon}) = 0, \sigma^{\varepsilon} \le T; \psi^{\varepsilon}(t) = 1 \text{ for each } t \in [0, T]\} \\
&\quad \ge P\{Y^{\varepsilon}_y(\tau^{\varepsilon}(y; -\alpha, 2 - \alpha)) = -\alpha; \tau^{\varepsilon}(y; -\alpha, 2 - \alpha) \le T; \psi^{\varepsilon}(t) = 1 \\
&\qquad\qquad \text{for each } t \in [0, T]\} \\
&\qquad - P\left\{\sup_{0 \le a \le T} |F^{\varepsilon}(s) - Y^{\varepsilon}_y(s)| \ge \alpha/2\right\} \\
&\quad \ge P\left\{x^{\varepsilon}_{g_1(y),y}\left(\tau^{\varepsilon}_{g_1(y),y}(-\alpha, 2 - \alpha)\right) = \alpha; \tau^{\varepsilon}_{g_1(y),y}(-\alpha, 2 - \alpha) \le T\right\} \\
&\qquad \times P\{\psi^{\varepsilon}(t) = 1 \text{ for each } t \in [0, T]\} - \exp(-(1/2)\alpha\gamma^{-1/2} C_{10}(T)) \\
&\quad \ge \exp(-(2g_1(y) + \delta))\varepsilon^{-1}\gamma
\end{aligned} \tag{3.48}$$

for a given $\delta > 0$, T large enough, $\alpha = \alpha(\delta)$ and $\varepsilon \le \varepsilon(\delta)$, where $\tilde{\tau}^{\varepsilon}_1 = \inf\{t \ge 0: Z^{\varepsilon}(t) \in \partial D \cup \Gamma^{\varepsilon}_2\}$. [Here and later $x^{\varepsilon}_{g_1(y),y}(0) = F^{\varepsilon}(0) = Y^{\varepsilon}_y(0) = 1$.]

Next, by (2.4), (2.70), (3.28) and (3.37) for any $A > 0$ there exists $T = T(A)$ large enough such that

$$
\begin{aligned}
P\{\tau_1^\varepsilon \ge \gamma T\} &= E\chi_{\{\tilde\tau_1^\varepsilon \ge \gamma T\}} p_{\gamma T}^\varepsilon(Z^\varepsilon(\cdot), \nu(\cdot)) \\
&\le \exp(\gamma T(n-1)) P\{\sigma^\varepsilon \ge T\} \\
&\le 2P\Big\{\sup_{1 \le t \le T} Y_y^\varepsilon \le 3\Big\} + 2P\Big\{\sup_{0 \le t \le T} |Y_y^\varepsilon(t) - F^\varepsilon(t)| \ge 1/2\Big\} \\
&\le \exp(-A\gamma\varepsilon^{-1})
\end{aligned}
\tag{3.49}
$$

provided $\varepsilon \le \varepsilon(A)$.

If $\nu^\varepsilon(0) = 1$, $X^\varepsilon(0) \in K_1^\varepsilon$, $|y - x| \le \delta_0/2$, then according to the fact that $x_{g_1(y),y}^\varepsilon(s) \le Y_y^\varepsilon(s)$ everywhere for any $s \ge 0$, taking into account (3.28), (2.4), (2.70), and (3.45), one has for a given $T > 0$, $\delta > 0$,

$$
\begin{aligned}
&P\{X^\varepsilon(\tau_1^\varepsilon) \in \partial D, \tau_1^\varepsilon \le \gamma T \text{ and } \nu^\varepsilon(\tau_1^\varepsilon) \ne 1\} \\
&\quad \le \exp(\gamma T(n-1)) P\{Z^\varepsilon(\tilde\tau_1^\varepsilon) \in \partial D, \tilde\tau_1^\varepsilon \le \gamma T, \nu(\tilde\tau_1^\varepsilon) \ne 1\} \\
&\quad \le 2P\{F^\varepsilon(\sigma^\varepsilon) = 0, \sigma^\varepsilon \le T, \tilde n(T) \ge 1\} \\
&\quad \le 2P\{Y_y^\varepsilon(\tau^\varepsilon(y;\alpha,2+\alpha)) = \alpha; \tau^\varepsilon(y,\alpha,2+\alpha) \le T; \tilde n(T) \ge 1\} \\
&\qquad + 2P\Big\{\sup_{0 \le s \le T} |Y_y^\varepsilon(s) - F^\varepsilon(s)| \ge \alpha/2\Big\} \\
&\quad \le 2P\Big\{x_{g_1(y),y}^\varepsilon\big(\tau_{g_1(y),y}^\varepsilon(\alpha, 2+\alpha)\big) = \alpha, \tau_{g_1(y),y}^\varepsilon(\alpha,2+\alpha) \le T\Big\} P\{\tilde n(T) \ge 1\} \\
&\qquad + 2\exp\big(-(1/2)C_{10}(T)\alpha\gamma^{-1/2}\big) \\
&\quad \le \gamma \exp\big(-2(g_1(y) - \delta)\gamma\varepsilon^{-1}\big)
\end{aligned}
\tag{3.50}
$$

for $\alpha = \alpha(\delta) > 0$, provided $\varepsilon \le \varepsilon(\delta, T)$, where $\tilde n(T)$ is the number of jumps of $\psi^\varepsilon(t)$ on the time segment $[0, T]$. Then, choosing A and, therefore, T in (3.49) large enough, we derive from (3.49) and (3.50),

$$
P\{X^\varepsilon(\tau_1^\varepsilon) \in \partial D \text{ and } \nu^\varepsilon(\tau_1^\varepsilon) \ne 1\} \le \gamma \exp\big(-\gamma\varepsilon^{-1}(2g_1(y) - 3\delta)\big) \quad \text{for } \varepsilon \le \varepsilon(\delta). \tag{3.51}
$$

Therefore, by (3.48), (3.51), (3.24) and the definition of γ, one has

$$
\begin{aligned}
&\sup_{x \in K_1^\varepsilon} P_{x,1}\{X^\varepsilon(\tau_1^\varepsilon) \ne 1 | X^\varepsilon(\tau_1^\varepsilon) \in \partial D\} \\
&\quad \le \gamma \exp(4\gamma\varepsilon^{-1}\delta) = \exp\big(-\gamma\varepsilon^{-1}(a_0 - 4\delta)\big) \le \varepsilon^{a_2}
\end{aligned}
\tag{3.52}
$$

for $\delta \le a_0/8$ and $\varepsilon \le \varepsilon(\delta)$.

Suppose now that $\nu^\varepsilon(0) = i \ne 1$, $x = X^\varepsilon(0) \in K_1^\varepsilon$ and $|x - y| \le \delta_0/2$ for some $y \in \partial D$. Notice that for any given $\alpha > 0$ one can find $h = h(\alpha)$ such that if $\psi^\varepsilon(t) = 1$ for $h \le t \le T$ and $x^\varepsilon_{g_1(y),y}(0) = Y^\varepsilon_y(0)$, then

$$|Y^\varepsilon_y(t) - x^\varepsilon_{g_1(y),y}(t)| \le \frac{\alpha}{3} \quad \text{for } 0 \le t \le T \tag{3.53}$$

by (3.26), (3.40) and (3.8), since b_i are bounded.

Thus, by (2.4), (2.60), (3.53), (3.28) and (3.45) one has

$$\begin{aligned}
P_{x,i}\{X^\varepsilon(\tau_1^\varepsilon) \in \partial D\} &\ge P_{x,i}\{X^\varepsilon(\tau_1^\varepsilon) \in \partial D,\, n(\gamma T) = 1,\, \tau_1^\varepsilon \le \gamma T\} \\
&= E_{x,i} p^\varepsilon_{\gamma T}(Z^\varepsilon(\cdot), \nu(\cdot)) X_{\{Z^\varepsilon(\tilde\tau_1^\varepsilon) \in \partial D,\, n(\gamma T) = 1,\, \tilde\tau_1^\varepsilon \le \gamma T\}} \\
&\ge \min_{\substack{1 \le i, j \le n \\ x \in \overline{D}}} d_{ij}(x) P\{F^\varepsilon(\sigma^\varepsilon) = 0,\, \tilde n(T) = 1,\, \psi^\varepsilon(t) = 1 \\
&\qquad\qquad \text{for } h \le t \le T,\, \sigma^\varepsilon \le T\} \\
&\ge C_{12} P\{Y^\varepsilon_y(\tau^\varepsilon(y; -\alpha, 2-\alpha)) = -\alpha; \\
&\qquad \tau^\varepsilon(y; -\alpha, 2-\alpha) \le T,\, \tilde n(T) = 1,\, \psi^\varepsilon(t) = 1, \\
&\qquad\qquad \text{for } h \le t \le T\} \\
&\qquad - P\Big\{\sup_{0 \le s \le T} |F^\varepsilon(s) - Y^\varepsilon_y(s)| \ge \alpha/2\Big\} \qquad (3.54) \\
&\ge C_{12} P\Big\{x^\varepsilon_{g_1(y),y}\big(\tau^\varepsilon_{g_1(y),y}(-2\alpha, 2-2\alpha)\big) \\
&\qquad = -2\alpha,\, \tau^\varepsilon_{g_1(y),y}(-2\alpha, 2-2\alpha) \le T, \\
&\qquad \tilde n(T) = 1,\, \psi^\varepsilon(t) = 1 \text{ for each } t \in [h, T]\Big\} \\
&\qquad - \exp\big(-C_{10}(T)(\alpha/2)\gamma^{-1/2}\big) \\
&\ge h\gamma \exp\big(-\varepsilon^{-1}\gamma(2g_1(y) + \delta)\big) \\
&\qquad - \exp\big(-C_{10}(T)(\alpha/2)\gamma^{-1/2}\big) \\
&\ge \gamma \exp\big(-2\varepsilon^{-1}\gamma(g_1(y) + \delta)\big),
\end{aligned}$$

setting $x^\varepsilon_{g_1(y),y}(0) = Y^\varepsilon_y(0) = 1$, choosing T large enough, $\alpha \le \alpha(\delta)$ and $\varepsilon \le \varepsilon(\delta)$ and taking into account the independence of $x^\varepsilon_{g_1(y),y}(t)$ and $\psi^\varepsilon(t)$.

Next, notice that if $\nu(0) = i \ne 1$, $\nu(\tilde\tau_1) \ne 1$ and $\tilde\tau_1^\varepsilon \le \gamma T$ for some $T > 0$, then only two possibilities exist:

(i) $\nu(t) \ne 1$ for any $t \in [0, \gamma T]$;
(ii) there was more than one jump of ν prior to γT.

Clearly, the possibilities (i) and (ii) can intersect. If (i) holds, then

$$Y^\varepsilon_y(t) \ge x^\varepsilon_{g_0(y),y}(t) \quad \text{for any } t \in [0, T], \tag{3.55}$$

where $g_0(y) = \min_{2 \le i \le n} g_i(y)$, $Y_y^\varepsilon(0) = x^\varepsilon_{g_0(y),y}(0)$. Thus, for T large enough,

$$(3.56)\quad \begin{aligned} &P_{i,x}\{Z^\varepsilon(\tilde\tau_1^\varepsilon) \in \partial D, \tilde\tau_1^\varepsilon \le T\gamma \text{ and (i) holds}\} \\ &\quad = P\{F^\varepsilon(\sigma^\varepsilon) = 0, \sigma^\varepsilon \le T \text{ and } \psi^\varepsilon(t) \ne 1 \text{ for } t \in [0,T]\} \\ &\quad \le P\{Y_y^\varepsilon(\tau^\varepsilon(y;\alpha, 2+\alpha)) = \alpha, \tau^\varepsilon(y,\alpha,2+\alpha) \le T \\ &\qquad\qquad \text{and } \psi(t) \ne 1 \text{ for } t \in [0,T]\} \\ &\qquad + P\Big\{\sup_{0 \le s \le T} |Y_y^\varepsilon(s) - F^\varepsilon(s)| > \alpha/2\Big\} \\ &\quad \le P\Big\{x^\varepsilon_{g_0(y),y}\big(\tau_{g_0(y),y}(\alpha, 2+\alpha)\big) = \alpha, \tau^\varepsilon_{g_0(y),y}(\alpha,2+\alpha) \le T\Big\} \\ &\qquad + \exp\big(-C_{10}(T)\alpha(2\gamma)^{-1}\big) \\ &\quad \le \exp\big(-2(g_0(y) - \delta)\gamma\varepsilon^{-1}\big) \end{aligned}$$

for $\alpha = \alpha(\delta)$, $\varepsilon \le \varepsilon(\delta, T)$, by (3.55), (3.45) and (3.28).

Suppose now that (ii) holds. By the elementary properties of Poisson processes one has

$$(3.57)\quad P\{\tilde n(T) \ge 2\} \le C_{13}(T)\gamma^2,$$

where we recall that $\tilde n(T)$ is the number of jumps of $\psi^\varepsilon(t)$ prior to T.

Notice that for any $t \ge 0$ and y satisfying (3.46) one has

$$Y_y^\varepsilon(t) \ge x^\varepsilon_{g_1(y),y}(t),$$

provided $Y_y^\varepsilon(0) = x^\varepsilon_{g_1(y),y}(0)$.

Thus, one has by the standard arguments of this section that

$$(3.58)\quad \begin{aligned} &P_{x,i}\{Z^\varepsilon(\tilde\tau_1^\varepsilon) = \partial D, \tilde\tau_1^\varepsilon \le T\gamma, \tilde n(T) \ge 2\} \\ &\quad \le P_{x,i}\Big\{\sup_{0 \le t \le T} |Y_y^\varepsilon(s) - F^\varepsilon(s)| \ge \alpha/2\Big\} \\ &\qquad + P\Big\{x^\varepsilon_{g_1(y),y}\big(\tau_{g_1(y),y}(\alpha, 2+\alpha)\big) = \alpha, \tau_{g_1(y),y}(\alpha, 2+\alpha) \le T\Big\} \\ &\qquad \times P\{\tilde n(T) \ge 2\} \\ &\quad \le \exp\big(-\alpha C_{10}(T)(2\gamma)^{-1}\big) + \gamma^2 \exp\big(-2(g_1(y) - \delta)\gamma\varepsilon^{-1}\big) \\ &\quad \le \gamma^2 \exp\big(-2(g_1(y) - 2\delta)\gamma\varepsilon^{-1}\big), \end{aligned}$$

where we use the independence of $\psi^\varepsilon(t)$ and $x^\varepsilon_{g_1(y),y}(t)$ and assume $x \in K_1^\varepsilon$, $F^\varepsilon(0) = Y_y^\varepsilon(0) = x^\varepsilon_{g_1(y),y}(0) = 1$ and $\varepsilon \le \varepsilon(\delta, T)$. Choosing A and $T(A)$ large enough in (3.49) and taking into account our assumption that p_T^ε is bounded, we obtain for $x \in K_1^\varepsilon$, y: $|y - x| < \delta/2$, $\nu^\varepsilon(0) = i \ne 1$, by (3.56) and (3.58),

$$(3.59)\quad \begin{aligned} &P_{x,i}\{X^\varepsilon(\tau_1^\varepsilon) \in \partial D, \nu^\varepsilon(\tau_1^\varepsilon) \ne 1\} \\ &\quad \le \gamma^2 \exp\big(-2(g_1(y) - 3\delta)\gamma\varepsilon^{-1}\big) \\ &\qquad + 2\exp\big(-2(g_0(y) - \delta)\gamma\varepsilon^{-1}\big). \end{aligned}$$

Finally, by (3.54), (3.59), (3.12) and the definition of $\gamma = \gamma(\varepsilon)$ one has

$$\sup_{\substack{x\in K_1^\varepsilon \\ 2\le i\le n}} P_{x,i}\{\nu^\varepsilon(\tau_1^\varepsilon) \neq 1 | X^\varepsilon(\tau_1^\varepsilon) \in \partial D\}$$

$$\leq \frac{1}{\gamma} \sup_{|y-y_0|\le\delta_0} \exp\big(-2\gamma\varepsilon^{-1}(g_0(y) - g_1(y) - 2\delta)\big) + \gamma \exp\big(8\delta\gamma\varepsilon^{-1}\big) \tag{3.60}$$

$$\leq \gamma^{1/2},$$

provided δ is chosen to be small enough and $\varepsilon \leq \varepsilon_0$.

The estimates (3.60) and (3.52) prove (3.18) and this completes the proof of the theorem. □

REFERENCES

BEZUIDENHOUT, C. (1987). A large deviations principle for small perturbations of random evolution equations. *Ann. Probab.* **15** 646–658.

EIZENBERG, A. and FREIDLIN, M. (1990). On the Dirichlet problems for a class of second order PDE systems with small parameter. *Stochastics Stochastic Rep.* **33** 111–148.

FREIDLIN, M. I. and WENTZELL, A. D. (1984). *Random Perturbations of Dynamical Systems.* Springer, Berlin.

FREIDLIN, M. (1985). *Functional Integration and Partial Differential Equations.* Princeton Univ. Press.

FREIDLIN, M. and GARTNER, J. (1978). A new contribution to the question of large deviations for random processes. *Vestnik Mosk. Univ. Ser. I Mat. Meh.* **5** 52–59.

FRIEDMAN, A. (1976). *Stochastic Differential Equations and Applications.* Academic, New York.

IOFFE, A. D. and TIKHOMIROV, V. M. (1978). Theory of extremal problems. In *Studies in Math. and Its Application* **6**. North-Holland, Amsterdam.

WENTZELL, A. D. and FREIDLIN, M. I. (1970). On small random perturbations of dynamical systems. *Russian Math. Surveys* **25** 1–55.

INSTITUTE OF MATHEMATICS
HEBREW UNIVERSITY
JERUSALEM
ISRAEL

DEPARTMENT OF MATHEMATICS
UNIVERSITY OF MARYLAND
COLLEGE PARK, MARYLAND 20742