Chapter 9
Abstract Sufficient Conditions for Large and Moderate Deviations in the Small Noise Limit

In this chapter we use the representations derived in Chap. 8 to study large and moderate deviations for stochastic systems driven by Brownian and/or Poisson noise, and consider a "small noise" limit, as in Sects. 3.2 and 3.3. We will prove general abstract large deviation principles, and in later chapters apply these to models in which the noise enters the system in an additive and independent manner.[1] For these systems, one can view the mapping that takes the noise into the state of the system as "nearly" continuous, and it is this property that allows a unified and relatively straightforward treatment. In contrast, for the corresponding discrete time processes of Chap. 4, the noise entered in a possibly nonadditive way, and a more involved analysis was required. If we had restricted our attention in Chap. 4 to recursive models of the form

$$X^n_{i+1} = X^n_i + \frac{1}{n} b(X^n_i) + \frac{1}{n}\sigma(X^n_i)\theta_i, \quad X^n_0 = x_0,$$

with $\{\theta_i\}_{i\in\mathbb{N}}$ an iid sequence (the discrete time analogues of the models in this chapter), then the analysis of Chap. 4 would have been much simpler. If one were to generalize within the continuous time framework to systems in which the noise enters in a more complicated manner, as in for example processes with multiple time or space scales (e.g., [111]), then the mapping from noise to state becomes more complex, as do the formulation of large deviation results and the methods of proof.

The main results of this chapter are Theorems 9.2 and 9.9 on uniform Laplace principles for a sequence of measurable functions of a Brownian motion and a PRM. Theorem 9.2 is well suited for proving large deviation results for small noise systems, whereas Theorem 9.9 is motivated by applications to moderate deviations. The proof of Theorem 9.2 is given in Sect. 9.3, and that of Theorem 9.9 is in Sect. 9.4. Theorems 9.2 and 9.9 are applied in Chap. 10 to develop large and moderate deviation approximations for certain finite dimensional systems. Infinite dimensional systems

[1] In our terminology, this includes systems with multiplicative noise, namely settings in which the noise term is multiplied by a state-dependent coefficient.

A. Budhiraja and P. Dupuis, *Analysis and Approximation of Rare Events*, Probability Theory and Stochastic Modelling 94,
https://doi.org/10.1007/978-1-4939-9579-0_9

are considered in Chap. 11, with the case of reaction–diffusion equations being developed in some detail, and in Chap. 12, where stochastic flows of diffeomorphisms are considered.

9.1 Definitions and Notation

The noise processes that drive the stochastic dynamical systems of this chapter were introduced in Chap. 8, and we adopt the notation used there. All stochastic processes are on the time horizon $[0, T]$, for some $T \in (0, \infty)$. We recall that Λ is a symmetric strictly positive trace class operator on the real separable Hilbert space $(\mathscr{H}, \langle \cdot, \cdot \rangle)$, $\mathscr{H}_0 \doteq \Lambda^{1/2}\mathscr{H}$, and $\mathbb{W} \doteq \mathscr{C}([0, T] : \mathscr{H}_0)$. Recall also that for a locally compact Polish space $\mathscr{S}$, $\Sigma(\mathscr{S})$ is the space of all measures ν on $(\mathscr{S}, \mathscr{B}(\mathscr{S}))$ satisfying $\nu(K) < \infty$ for every compact $K \subset \mathscr{S}$.

Throughout this chapter we deal simultaneously with Brownian and Poisson noise models. Because of this, we slightly modify the notation from Chap. 8 to make associations clear. Given a locally compact Polish space $\mathscr{X}$ that models the different types of jumps under the PRM, define $\mathscr{X}_T \doteq [0, T] \times \mathscr{X}$, the augmented space $\mathscr{Y} \doteq \mathscr{X} \times \mathbb{R}_+$ and the time-dependent version $\mathscr{Y}_T \doteq [0, T] \times \mathscr{Y}$, and canonical spaces $\mathbb{M} \doteq \Sigma(\mathscr{X}_T)$, $\bar{\mathbb{M}} \doteq \Sigma(\mathscr{Y}_T)$, $\mathbb{V} \doteq \mathbb{W} \times \mathbb{M}$, and $\bar{\mathbb{V}} \doteq \mathbb{W} \times \bar{\mathbb{M}}$. Let $\bar{N}$ and W be the maps from $\bar{\mathbb{V}}$ to $\bar{\mathbb{M}}$ and $\bar{\mathbb{V}}$ to $\mathbb{W}$ such that

$$\bar{N}(w, m) = m, \quad W(w, m) = w, \text{ for } (w, m) \in \bar{\mathbb{V}}.$$

Define

$$\mathscr{F}_t^0 \doteq \sigma \left\{ \bar{N}((0, s] \times A), W(s) : 0 \le s \le t, A \in \mathscr{B}(\mathscr{Y}) \right\}.$$

Assume $\nu \in \Sigma(\mathscr{X})$, and define $\bar{\nu} \doteq \nu \times \lambda_\infty$ and $\bar{\nu}_T \doteq \lambda_T \times \bar{\nu}$, where λ_T and λ_∞ are Lebesgue measure on $[0, T]$ and $[0, \infty)$, respectively. Let P denote the unique probability measure on $(\bar{\mathbb{V}}, \mathscr{B}(\bar{\mathbb{V}}))$ such that under P:

(a) W is a Λ-Wiener process with respect to $\mathscr{F}_t^0$;
(b) $\bar{N}$ is an $\mathscr{F}_t^0$-PRM with intensity measure $\bar{\nu}_T$;
(c) for all $0 \le s \le t < \infty$, $(\bar{N}((s, t] \times \cdot), W(t) - W(s))$ is independent of $\mathscr{F}_s^0$.

It follows that W and $\bar{N}$ are independent under P [167, Lemma 13.6]. Throughout this chapter we use $\{\mathscr{F}_t\}$, the augmentation of the filtration $\{\mathscr{F}_t^0\}$ with all P-null sets in $\mathscr{B}(\bar{\mathbb{V}})$. Recall the collections of controls $\bar{\mathscr{A}}^W$, $\bar{\mathscr{A}}_b^W$, $\bar{\mathscr{A}}_{b,n}^W$, $\bar{\mathscr{A}}^N$, $\bar{\mathscr{A}}_b^N$, and $\bar{\mathscr{A}}_{b,n}^N$ introduced in Sect. 8.3, where a subscript b, n means that costs are w.p.1 bounded by n, a b denotes the union over finite n of such controls, and the overbar indicates that the filtration used in defining these spaces is $\{\mathscr{F}_t\}$. We also have the definitions $\bar{\mathscr{A}}_{b,n} \doteq \bar{\mathscr{A}}_{b,n}^W \times \bar{\mathscr{A}}_{b,n}^N$, and $\bar{\mathscr{A}}_b \doteq \cup_{n \in \mathbb{N}} \bar{\mathscr{A}}_{b,n}$. For $u = (\psi, \varphi) \in \bar{\mathscr{A}}_b$, define the costs

$$L_T^W(\psi) \doteq \frac{1}{2} \int_0^T \|\psi(s)\|_0^2 \, ds \text{ and } L_T^N(\varphi) \doteq \int_{\mathscr{X}_T} \ell(\varphi(t, x)) \, \nu_T(dt \times dx) \qquad (9.1)$$

as in Sect. 8.3, and let $\bar{L}_T(u) \doteq L_T^W(\psi) + L_T^N(\varphi)$. The controlled PRM N^φ is also defined as in that section.

Let S_n^W denote the subset of $\mathscr{L}^2([0,T]:\mathscr{H}_0)$ defined as in (8.1), and recall that this is a compact space with the weak topology on $\mathscr{L}^2([0,T]:\mathscr{H}_0)$. For $n \in \mathbb{N}$, define the analogous space

$$S_n^N \doteq \left\{ g : \mathscr{X}_T \to [0,\infty) : L_T^N(g) \le n \right\}.$$

A function $g \in S_n^N$ can be identified with a measure $\nu_T^g \in \mathbb{M}$ according to $\nu_T^g(A) = \int_A g(s,x)\,\nu_T(ds \times dx)$, $A \in \mathscr{B}(\mathscr{X}_T)$. Since convergence in $\mathbb{M}$ is essentially equivalent to weak convergence on compact subsets, the superlinear growth of ℓ implies that $\left\{\nu_T^g : g \in S_n^N\right\}$ is a compact subset of $\mathbb{M}$. The proof of this fact is given in Appendix A.4.3. We equip S_n^N with the topology obtained through this identification, which makes S_n^N a compact space. We then let $\bar{S}_n \doteq S_n^W \times S_n^N$ with the usual product topology, with respect to which it is also a compact space. An element $u \in \bar{\mathscr{A}}_{b,n}$ is regarded as a random variable with values in the compact space $\bar{S}_n$. Finally, let $\bar{S} \doteq \cup_{n\in\mathbb{N}} \bar{S}_n$.

9.2 Abstract Sufficient Conditions for LDP and MDP

Recall from Chap. 1 that various normalizations or scaling sequences are possible when one is formulating an LDP. In this section we formulate sufficient conditions for a Laplace principle to hold for general measurable functions of $(\sqrt{\varepsilon}W, \varepsilon N^{1/\varepsilon})$ and with two different scaling sequences. The first sufficient condition will be used in Chaps. 10, 11 and 12 to study large deviation principles for small noise stochastic dynamical systems. The second sufficient condition is for a moderate deviation principle. The condition is applied to finite dimensional models in Chap. 10, and for an example of its use in an infinite dimensional setting we refer to [41]. The results that we prove in fact give more, namely uniform Laplace principles in the sense of Definition 1.11. The uniformity is with respect to a parameter z (typically an initial condition), which takes values in some compact subset of a Polish space $\mathscr{Z}$.

The definition of a uniform Laplace principle was given in Chap. 1. The statement there considered the scale sequence $\varepsilon = 1/n$, and the analogous definition for a general scale function $\varkappa(\varepsilon)$ is as follows. Let $\{I_z, z \in \mathscr{Z}\}$ be a family of rate functions on $\mathscr{X}$ parametrized by z in a Polish space $\mathscr{Z}$ and assume that this family has compact level sets on compacts, namely, for each compact subset K of $\mathscr{Z}$ and each $M < \infty$, $\cup_{z\in K}\{x \in \mathscr{X} : I_z(x) \le M\}$ is a compact subset of $\mathscr{X}$. Let $\{X^\varepsilon\}$ be a collection of $\mathscr{X}$-valued random variables with distributions that depend on $z \in \mathscr{Z}$ and denote the corresponding expectation operator by E_z. The collection $\{X^\varepsilon\}$ is said to satisfy the Laplace principle on $\mathscr{X}$ with scale function $\varkappa(\varepsilon)$ and rate function I_z, uniformly on compacts, if for all compact subsets K of $\mathscr{Z}$ and all bounded continuous functions h mapping $\mathscr{X}$ into $\mathbb{R}$,

$$\lim_{\varepsilon\to 0}\sup_{z\in K}\left|\varkappa(\varepsilon)\log E_z \exp\{-\varkappa(\varepsilon)^{-1}h(X^\varepsilon)\} - F(z,h)\right| = 0,$$

where $F(z,h) \doteq -\inf_{x\in\mathscr{X}}[h(x)+I_z(x)]$.

In this chapter it will be convenient to work with a common probability measure (instead of a collection parametrized by $z\in\mathscr{Z}$) and instead note the dependence on z in the collection of random variables, i.e., we write X_z^ε instead of X^ε.

9.2.1 An Abstract Large Deviation Result

In this section we present a sufficient condition for a uniform Laplace principle with scale function $\varkappa(\varepsilon)=\varepsilon$ to hold for measurable functions of $(\sqrt{\varepsilon}W, \varepsilon N^{1/\varepsilon})$. Recall that $N^{1/\varepsilon}$ is the PRM defined through (8.16) with $\varphi\equiv 1/\varepsilon$. It is defined on $\bar{\mathbb{V}}$, takes values in $\mathbb{M}$, and has an intensity measure that is scaled by $1/\varepsilon$. Let $\{\mathscr{G}^\varepsilon\}_{\varepsilon>0}$, be a family of measurable maps from $\mathscr{Z}\times\mathbb{V}$ to $\mathbb{U}$, where $\mathscr{Z}$ and $\mathbb{U}$ are some Polish spaces. Let $\left\{Z_z^\varepsilon\right\}_{\varepsilon>0,z\in\mathscr{Z}}$ be the collection of $\mathbb{U}$-valued random variables on $(\bar{\mathbb{V}},\mathscr{B}(\bar{\mathbb{V}}),P)$ defined by

$$Z_z^\varepsilon \doteq \mathscr{G}^\varepsilon(z,\sqrt{\varepsilon}W,\varepsilon N^{1/\varepsilon}). \tag{9.2}$$

We are interested in a uniform large deviation principle for the family $\left\{Z_z^\varepsilon\right\}$ as $\varepsilon\to 0$. We recall from Proposition 1.14 that a uniform large deviation principle is implied by a uniform Laplace principle.

A control $u=(\psi,\varphi)\in\bar{\mathscr{A}}_{b,n}$ will be regarded as a random variable with values in the compact metric space S_n. The following is a sufficient condition for a large deviation property. As noted in Sect. 9.1, $\bar{L}_T(u)=L_T^W(\psi)+L_T^N(\varphi)$. Recall also the notation $W^\psi(\cdot)=W(\cdot)+\int_0^\cdot\psi(s)ds$.

Condition 9.1 *There exists a measurable map $\mathscr{G}^0:\mathscr{Z}\times\mathbb{V}\to\mathbb{U}$ such that the following hold.*

(a) *For $n\in\mathbb{N}$ and compact $K\subset\mathscr{Z}$, the set*

$$\Gamma_{n,K}\doteq\left\{\mathscr{G}^0\left(z,\int_0^\cdot f(s)ds,\nu_T^g\right):q=(f,g)\in S,\bar{L}_T(q)\le n, z\in K\right\} \tag{9.3}$$

is a compact subset of $\mathbb{U}$.

(b) *For $n\in\mathbb{N}$, let $u^\varepsilon=(\psi^\varepsilon,\varphi^\varepsilon)\in\bar{\mathscr{A}}_{b,n}$, $u=(\psi,\varphi)\in\bar{\mathscr{A}}_{b,n}$ be such that u^ε converges in distribution to u as $\varepsilon\to 0$. Also, let $\{z^\varepsilon\}\subset\mathscr{Z}$ be such that $z^\varepsilon\to z$ as $\varepsilon\to 0$. Then*

$$\mathscr{G}^\varepsilon\left(z^\varepsilon,\sqrt{\varepsilon}W^{\psi^\varepsilon/\sqrt{\varepsilon}},\ \varepsilon N^{\varphi^\varepsilon/\varepsilon}\right)\Rightarrow\mathscr{G}^0\left(z,\int_0^\cdot\psi(s)ds,\nu_T^\varphi\right).$$

For $\phi\in\mathbb{U}$ and $z\in\mathscr{Z}$, define $S_{z,\phi}^{\mathscr{G}}\doteq\left\{(f,g)\in S:\phi=\mathscr{G}^0(z,\int_0^\cdot f(s)ds,\nu_T^g)\right\}$. These are the controls that produce the output ϕ. For $z\in\mathscr{Z}$, let $I_z:\mathbb{U}\to[0,\infty]$ be defined by

$$I_z(\phi) \doteq \inf_{q=(f,g)\in S^{\mathscr{G}}_{z,\phi}} \bar{L}_T(q). \tag{9.4}$$

Theorem 9.2 *Suppose that $\mathscr{G}^\varepsilon$ and $\mathscr{G}^0$ satisfy Condition 9.1. Suppose also that for all $\phi \in \mathbb{U}$, $z \mapsto I_z(\phi)$ is a lower semicontinuous mapping from $\mathscr{Z}$ to $[0,\infty]$. Then for all $z \in \mathscr{Z}$, I_z defined in (9.4) is a rate function on $\mathbb{U}$, the family $\{I_z, z \in \mathscr{Z}\}$ of rate functions has compact level sets on compacts, and $\left\{Z^\varepsilon_z\right\}$ satisfies a Laplace principle with scale function ε and rate function I_z, uniformly on compact subsets of $\mathscr{Z}$.*

Remark 9.3 Note that the lower semicontinuity of $z \mapsto I_z(\phi)$ is typically automatic in the situation in which z is an initial condition for a stochastic process defined by the mapping $\mathscr{G}^\varepsilon$, since in this case $I_z(\phi) < \infty$ only when $z = \phi(0)$.

The proof of Theorem 9.2 is given in Sect. 9.3. Two examples were given in Chap. 3 to illustrate the role of $\mathscr{G}^\varepsilon$ and the scaling. For convenience, we recall the diffusion example, and hence take $\mathbb{V} = \mathscr{C}([0,1]:\mathbb{R}^k)$, $\mathbb{U} = \mathscr{C}([0,1]:\mathbb{R}^d)$, and $\mathscr{Z} = \mathbb{R}^d$.

Example 9.4 Suppose $b:\mathbb{R}^d \to \mathbb{R}^d$ and $\sigma:\mathbb{R}^d \to \mathbb{R}^{d\times k}$ satisfy

$$\|b(x)-b(y)\| + \|\sigma(x)-\sigma(y)\| \le C\,\|x-y\|$$

for all $x, y \in \mathbb{R}^d$, with $C \in (0,\infty)$. Let W be a standard k-dimensional Brownian motion, and for fixed $z \in \mathbb{R}^d$ and $\varepsilon > 0$, let $X^\varepsilon_z = \left\{X^\varepsilon_z(t)\right\}_{0\le t\le 1}$ be the strong solution of the SDE

$$dX^\varepsilon_z(t) = b(X^\varepsilon_z(t))dt + \sqrt{\varepsilon}\sigma(X^\varepsilon_z(t))dW(t),\;\; X^\varepsilon_z(0) = z. \tag{9.5}$$

From the unique pathwise solvability of this SDE (see [172, Definition 5.3.2 and Corollary 5.3.23]), it follows that for each $\varepsilon > 0$, there is a measurable map $\mathscr{G}^\varepsilon : \mathbb{R}^d \times \mathscr{C}([0,1]:\mathbb{R}^k) \to \mathscr{C}([0,1]:\mathbb{R}^d)$ such that $X^\varepsilon_z = \mathscr{G}^\varepsilon(z, \sqrt{\varepsilon}W)$ is the solution to (9.5). The corresponding map $\mathscr{G}^0$ can be defined by $\mathscr{G}^0\left(z, \int_0^\cdot f(s)ds\right) = \varphi$ if for $z \in \mathbb{R}^d$ and $f \in \mathscr{L}^2([0,1]:\mathbb{R}^k)$,

$$\varphi(t) = z + \int_0^t b(\varphi(s))ds + \int_0^t \sigma(\varphi(s))f(s)ds, \quad t \in [0,1],$$

and $\mathscr{G}^0(z,\gamma) \equiv 0$ for all other $(z,\gamma) \in \mathbb{R}^d \times \mathscr{C}([0,1]:\mathbb{R}^d)$. Along the lines of the discussion in Chap. 3, it is easily checked that Condition 9.1 is valid, and in particular, part (b) is simply a restatement of the LLN limit $\bar{X}^\varepsilon_{z^\varepsilon} \Rightarrow \bar{X}_z$, where $\bar{X}^\varepsilon_{z^\varepsilon}$ and $\bar{X}_z$ are the solutions to

$$d\bar{X}^\varepsilon_{z^\varepsilon}(t) = b(\bar{X}^\varepsilon_{z^\varepsilon}(t))dt + \sigma(\bar{X}^\varepsilon_{z^\varepsilon}(t))\psi^\varepsilon(t)dt + \sqrt{\varepsilon}\sigma(\bar{X}^\varepsilon_{z^\varepsilon}(t))dW(t),$$

$\bar{X}^{\varepsilon}_{z^{\varepsilon}}(0) = z_{\varepsilon}$, and

$$d\bar{X}_z(t) = b(\bar{X}_z(t))dt + \sigma(\bar{X}_z(t))\psi(t)dt, \ \ \bar{X}_z(0) = z.$$

In Chaps. 10, 11, and 12 we consider other applications of Theorem 9.2.

Remark 9.5 The discussion of Example 9.4 shows that the main additional work needed to prove a uniform LDP instead of the ordinary LDP is to prove, instead of the convergence $\bar{X}^{\varepsilon}_z \Rightarrow \bar{X}_z$, the stronger convergence property $\bar{X}^{\varepsilon}_{z^{\varepsilon}} \Rightarrow \bar{X}_z$ whenever $z^{\varepsilon} \to z$. However, the proof of this stronger convergence property, at least in most situations of interest, requires the same analysis as that used for the convergence with a fixed initial condition. Thus in some uses later of Theorems 9.2 and 9.9 we present the argument for a fixed initial condition, and leave it to the reader to check that the same arguments could be used for converging initial conditions and thereby yield the uniform Laplace principle. Two exceptions are the reaction–diffusion example of Chap. 11 and the serve-the-longest queueing example of Chap. 13. For the latter example, as with many models in queueing, the discrete nature of the state space for the prelimit models requires initial conditions that depend on the scaling parameter.

9.2.2 An Abstract Moderate Deviation Result

Let $\{\mathscr{K}^{\varepsilon}\}_{\varepsilon>0}$ be a family of measurable maps from $\mathscr{Z} \times \mathbb{V}$ to $\mathbb{U}$. Let $a : (0, \infty) \to (0, \infty)$ be such that as $\varepsilon \to 0$,

$$a(\varepsilon) \to 0 \text{ and } \varkappa(\varepsilon) \doteq \frac{\varepsilon}{a^2(\varepsilon)} \to 0. \tag{9.6}$$

For $\varepsilon > 0$ and $z \in \mathscr{Z}$, let $Y^{\varepsilon}_z \doteq \mathscr{K}^{\varepsilon}(z, \sqrt{\varepsilon}W, \varepsilon N^{1/\varepsilon})$. In this section we formulate a sufficient condition for the collection $\left\{Y^{\varepsilon}_z\right\}_{\varepsilon>0}$ to satisfy a uniform Laplace principle with scale function $\varkappa(\varepsilon)$ and a rate function that is given through a suitable quadratic form.

While the large and moderate deviation assumptions and arguments are very similar when one is considering a diffusion model, a significant difference occurs when a PRM driving noise is included. This is very similar to the situation encountered in the discrete time analogue presented in Chap. 5. In particular, the Poisson cost must be replaced by an appropriate quadratic functional in the limit $\varepsilon \to 0$, and the dynamics are also adjusted to make analysis of the LLN limits easier. This is done by centering the controls on $\varphi \equiv 1$ and rescaling. The following inequalities will be used to translate bounds on controls φ^{ε} into bounds on this quadratic approximation. Recall the function $\ell(r) \doteq r \log r - r + 1$. The following properties can be easily shown. Part (a) has been used many times already, but is included here for convenience.

Lemma 9.6 (a) *For* $a, b \in (0, \infty)$ *and* $\sigma \in [1, \infty)$, $ab \le e^{\sigma a} + \frac{1}{\sigma}\ell(b)$.

(b) *For every* $\beta > 0$, *there exist* $\kappa_1(\beta), \bar{\kappa}_1(\beta) \in (0, \infty)$ *such that* $\kappa_1(\beta)$ *and* $\bar{\kappa}_1(\beta)$ *converge to* 0 *as* $\beta \to \infty$, *and for* $r \ge 0$,

$$|r-1| \le \kappa_1(\beta)\ell(r) \text{ if } |r-1| \ge \beta, \text{ and } r \le \bar{\kappa}_1(\beta)\ell(r) \text{ if } r \ge \beta > 1.$$

(c) *There is a nondecreasing function* $\kappa_2 : (0, \infty) \to (0, \infty)$ *such that for each* $\beta > 0$,

$$|r-1|^2 \le \kappa_2(\beta)\ell(r) \quad \text{for } |r-1| \le \beta, r \ge 0.$$

(d) *There exists* $\kappa_3 \in (0, \infty)$ *such that*

$$\ell(r) \le \kappa_3|r-1|^2, \quad |\ell(r) - (r-1)^2/2| \le \kappa_3|r-1|^3 \text{ for all } r \ge 0.$$

Recall $L_T^N(g) \doteq \int_{\mathscr{X}_T} \ell(g(s,y))\nu_T(ds \times dy)$ and $L_T^W(\psi) \doteq \frac{1}{2}\int_0^T \|\psi(s)\|_0^2\, ds$. For $\varepsilon > 0$ and $n \in \mathbb{N}$, define the spaces

$$\begin{aligned} S_{n,+}^{N,\varepsilon} &\doteq \{g : \mathscr{X}_T \to \mathbb{R}_+ \text{ such that } L_T^N(g) \le na^2(\varepsilon)\} \\ S_n^{N,\varepsilon} &\doteq \{f : \mathscr{X}_T \to \mathbb{R} \text{ such that } f = (g-1)/a(\varepsilon), \text{ with } g \in S_{n,+}^{N,\varepsilon}\}. \end{aligned} \tag{9.7}$$

Thus $S_n^{N,\varepsilon}$ are the centered and rescaled versions of the nonnegative functions appearing in $S_{n,+}^{N,\varepsilon}$. The following result is immediate from Lemma 9.6.

Lemma 9.7 *Suppose* $g \in S_{n,+}^{N,\varepsilon}$ *for some* $n < \infty$ *and let* $f = (g-1)/a(\varepsilon)$. *Then:*

(a) $\displaystyle\int_{\mathscr{X}_T} |f(s,y)| 1_{\{|f(s,y)| \ge \beta/a(\varepsilon)\}} \nu_T(ds \times dy) \le na(\varepsilon)\kappa_1(\beta)$ *for all* $\beta > 0$;

(b) $\displaystyle\int_{\mathscr{X}_T} g(s,y) 1_{\{g(s,y) \ge \beta\}} \nu_T(ds \times dy) \le na^2(\varepsilon)\bar{\kappa}_1(\beta)$ *for all* $\beta > 1$;

(c) $\displaystyle\int_{\mathscr{X}_T} |f(s,y)|^2 1_{\{|f(s,y)| \le \beta/a(\varepsilon)\}} \nu_T(ds \times dy) \le n\kappa_2(\beta)$ *for all* $\beta > 0$,

where $\kappa_1, \bar{\kappa}_1$ *and* κ_2 *are as in Lemma 9.6.*

Let

$$\mathscr{U}_{n,+}^{\varepsilon} \doteq \left\{(u_W, u_N) \in \bar{\mathscr{A}}_b : u_W(\cdot, \omega) \in S_{na(\varepsilon)^2}^W, u_N(\cdot,\cdot,\omega) \in S_{n,+}^{N,\varepsilon},\ \bar{P}\text{-a.s.}\right\}. \tag{9.8}$$

Thus by (9.7), $\mathscr{U}_{n,+}^{\varepsilon}$ is the class of controls for both types of noise for which the cost scales proportionally with $a(\varepsilon)^2$. Owing to the moderate deviation scaling, one can assume without loss that the control appearing in the representation can be restricted to a class of this form, with n depending on the function F. However, as $\varepsilon \to 0$ we will need to use the centered and rescaled analogues, which are related to a diffusion approximation to the original process. This requires additional notation and definitions.

The norm in the Hilbert space $\mathscr{L}^2(\nu_T)$ is denoted by $\|\cdot\|_{N,2}$, and the norm in $\mathscr{L}^2([0,T]:\mathscr{H}_0)$ by $\|\cdot\|_{W,2}$. Let $\mathscr{L}^2 \doteq \mathscr{L}^2([0,T]:\mathscr{H}_0)\times\mathscr{L}^2(\nu_T)$ and recall that $\mathscr{P}\mathscr{F}$ is the predictable σ-field on $[0,T]\times\bar{\mathbb{V}}$ with the filtration $\{\mathscr{F}_t\}$ on $(\bar{\mathbb{V}},\mathscr{B}(\bar{\mathbb{V}}))$.

Given a map $\mathscr{K}^0:\mathscr{Z}\times\mathscr{L}^2\to\mathbb{U}$, $z\in\mathscr{Z}$, and $\eta\in\mathbb{U}$, let

$$S^{\mathscr{K}}_{z,\eta}\doteq\{q=(f_1,f_2)\in\mathscr{L}^2:\eta=\mathscr{K}^0(z,q)\},$$

and define I_z for $z\in\mathscr{Z}$ by

$$I_z(\eta)\doteq\inf_{q=(f_1,f_2)\in S^{\mathscr{K}}_{z,\eta}}\left[\frac{1}{2}\left(\|f_1\|^2_{W,2}+\|f_2\|^2_{N,2}\right)\right]. \tag{9.9}$$

Here $S^{\mathscr{K}}_{z,\eta}$ identifies the $\mathscr{L}^2$ spaces that lead to the outcome q under the map $\mathscr{K}^0$, which can be associated with the map $\mathscr{K}^\varepsilon$ linearized about the LLN limit. As always, we follow the convention that the infimum over an empty set is $+\infty$.

We now introduce a sufficient condition that ensures that I_z is a rate function for every $z\in\mathscr{Z}$, the collection $\{I_z\}_{z\in\mathscr{Z}}$ has compact level sets on compacts, and the collection $\{Y^\varepsilon_z\}$ satisfies a Laplace principle with scale function $\varkappa(\varepsilon)$ and rate function I_z as $\varepsilon\to 0$. Let

$$\hat{S}_n\doteq\{(f_1,f_2)\in\mathscr{L}^2:\|f_1\|^2_{W,2}+\|f_2\|^2_{N,2}\le n\}. \tag{9.10}$$

Condition 9.8 *For some measurable map $\mathscr{K}^0:\mathscr{Z}\times\mathscr{L}^2\to\mathbb{U}$, the following two conditions hold.*

(a) *For every $n\in\mathbb{N}$ and compact $K\subset\mathscr{Z}$, the set*

$$\Gamma_{n,K}\doteq\left\{\mathscr{K}^0(z,q):z\in K,q\in\hat{S}_n\right\}$$

is a compact subset of $\mathbb{U}$.

(b) *Given $n\in\mathbb{N}$ and $\varepsilon>0$, let $(\psi^\varepsilon,\varphi^\varepsilon)\in\mathscr{U}^\varepsilon_{n,+}$ [defined in (9.8)]. Let $\theta^\varepsilon=\psi^\varepsilon/a(\varepsilon)$ and $\zeta^\varepsilon=(\varphi^\varepsilon-1)/a(\varepsilon)$. Suppose that for some $\beta\in(0,1]$, there is $m\in\mathbb{N}$ such that $(\theta^\varepsilon,\zeta^\varepsilon 1_{\{|\zeta^\varepsilon|\le\beta/a(\varepsilon)\}})\Rightarrow(\theta,\zeta)$ in $\hat{S}_m$. Also, let $\{z^\varepsilon\}\subset\mathscr{Z}$ be such that $z^\varepsilon\to z$ as $\varepsilon\to 0$. Then*

$$\mathscr{K}^\varepsilon\left(z^\varepsilon,\sqrt{\varepsilon}W^{\psi^\varepsilon/\sqrt{\varepsilon}},\ \varepsilon N^{\varphi^\varepsilon/\varepsilon}\right)\Rightarrow\mathscr{K}^0(z,\theta,\zeta).$$

Note that from Lemma 9.7, $(\theta^\varepsilon,\zeta^\varepsilon 1_{\{|\zeta^\varepsilon|\le\beta/a(\varepsilon)\}})$, as in part (b) of Condition 9.8, takes values in $\hat{S}_m$ with $m=n(1+\kappa_2(\beta))$. Thus if the condition holds, one can take $m=n(1+\kappa_2(\beta))$ without loss of generality. The following is the analogue of Theorem 9.2 for the moderate deviation scaling.

Theorem 9.9 *Suppose that $\mathscr{K}^\varepsilon$ and $\mathscr{K}^0$ satisfy Condition 9.8. Suppose also that for all $\phi\in\mathbb{U}$, $z\mapsto I_z(\phi)$ is a lower semicontinuous mapping from $\mathscr{Z}$ to $[0,\infty]$.*

Then for $z \in \mathscr{Z}$, I_z defined in (9.9) is a rate function on $\mathbb{U}$, the family $\{I_z, z \in \mathscr{Z}\}$ of rate functions has compact level sets on compacts, and $\{Y_z^\varepsilon\}$ satisfies a uniform Laplace principle with scale function $\varkappa(\varepsilon)$ and rate function I_z as $\varepsilon \to 0$.

The proof of Theorem 9.9 is in Sect. 9.4. As with Theorem 9.2, the assumption of lower semicontinuity of $z \mapsto I_z(\phi)$ is often vacuous when z plays the role of an initial condition (see Remark 9.3).

Example 9.10 Let X_z^ε be as in Example 9.4. In addition to the Lipschitz condition on the coefficients b, σ, assume that b is continuously differentiable. Let X_z^0 be the solution of the ODE $\dot{X}_z^0(t) = b(X_z^0(t))$, $X_z^0(0) = z$, and let $Y_z^\varepsilon = (X_z^\varepsilon - X_z^0)/a(\varepsilon)$. Then for each $\varepsilon > 0$, there is a measurable map $\mathscr{K}^\varepsilon$ from $\mathbb{R}^d \times \mathscr{C}([0,1] : \mathbb{R}^k)$ to $\mathscr{C}([0,1] : \mathbb{R}^d)$ such that $Y_z^\varepsilon = \mathscr{K}^\varepsilon(z, \sqrt{\varepsilon}W)$. The corresponding map

$$\mathscr{K}^0 : \mathbb{R}^d \times \mathscr{L}^2([0,1] : \mathbb{R}^k) \to \mathscr{C}([0,1] : \mathbb{R}^d)$$

is defined by $\mathscr{K}^0(z, f) = \eta_{z,f}$, where $\eta_{z,f}$ is the unique solution of the equation

$$\eta_{z,f}(t) = \int_0^t [Db(X_z^0(s))](\eta_{z,f}(s))ds + \int_0^t \sigma(X_z^0(s)) f(s)ds,$$

where $Db(x)$ is the matrix $(\partial b_i(x)/\partial x_j)_{ij}$. In Chap. 10, under the assumption that Db is Lipschitz continuous, it will be shown using Theorem 9.9 that for each fixed z, Y_z^ε satisfies a Laplace principle with scale function $\varkappa(\varepsilon)$. The proof of the uniform Laplace principle can be given similarly by considering arbitrary $z_\varepsilon \to z$ as $\varepsilon \to 0$. This result can be viewed as a moderate deviation principle for the diffusion process X_z^ε. Chapter 10 will treat the more general setting of d-dimensional jump-diffusions, and we refer the reader to [41] for analogous results in an infinite dimensional setting.

9.3 Proof of the Large Deviation Principle

In this section we prove Theorem 9.2. We first argue that for all compact $K \subset \mathscr{Z}$ and each $M < \infty$,

$$\Lambda_{M,K} \doteq \cup_{z \in K} \{\phi \in \mathbb{U} : I_z(\phi) \le M\} \tag{9.11}$$

is a compact subset of $\mathbb{U}$. Note that this will show that for each $z \in \mathscr{Z}$, I_z is a rate function and the collection $\{I_z, z \in \mathscr{Z}\}$ has compact level sets on compacts. To establish this, we will show that $\Lambda_{M,K}$ equals $\cap_{\delta \in (0,1)} \Gamma_{M+\delta,K}$, where $\Gamma_{M,K}$ is as in (9.3). In view of part (a) of Condition 9.1, the compactness of $\Lambda_{M,K}$ will then follow. Let $\phi \in \Lambda_{M,K}$. Then there exists $z \in K$ such that $I_z(\phi) \le M$. We can now find, for each $\delta \in (0,1)$, $q_\delta = (f_\delta, g_\delta) \in S_{z,\phi}^{\mathscr{G}}$, i.e., $\phi = \mathscr{G}^0(z, \int_0^\cdot f_\delta(s)ds, \nu_T^{g_\delta})$, such that $\bar{L}_T(q_\delta) \le M + \delta$. In particular, $\phi \in \Gamma_{M+\delta,K}$. Since $\delta \in (0,1)$ is arbitrary, we have $\Lambda_{M,K} \subset \cap_{\delta \in (0,1)} \Gamma_{M+\delta,K}$. Conversely, suppose $\phi \in \Gamma_{M+\delta,K}$ for all $\delta \in (0,1)$.

Then for each $\delta \in (0, 1)$, there exists $z_\delta \in K$, $q_\delta = (f_\delta, g_\delta) \in S$, $\bar{L}_T(q_\delta) \le M + \delta$, such that $\phi = \mathscr{G}^0(z_\delta, \int_0^\cdot f_\delta(s)ds, \nu_T^{g_\delta})$. In particular, we have

$$\inf_{z\in K} I_z(\phi) \le I_{z_\delta}(\phi) \le \bar{L}_T(q_\delta) \le M + \delta.$$

Sending $\delta \to 0$ gives $\inf_{z\in K} I_z(\phi) \le M$. Since the map $z \mapsto I_z(\phi)$ is lower semicontinuous, $\phi \in \Lambda_{M,K}$, and the inclusion $\cap_{\delta\in(0,1)} \Gamma_{M+\delta,K} \subset \Lambda_{M,K}$ follows. This proves the compactness of $\Lambda_{M,K}$ and finishes the first part of the theorem.

We next prove the second statement in the theorem. Fix $z \in \mathscr{Z}$ and let $\{z^\varepsilon\}_{\varepsilon>0} \subset \mathscr{Z}$ be such that $z^\varepsilon \to z$ as $\varepsilon \to 0$. Fix $F \in \mathscr{C}_b(\mathbb{U})$. In view of Proposition 1.12, it suffices to prove the Laplace upper bound

$$\limsup_{\varepsilon\to 0} \varepsilon \log E \exp\left\{-\frac{1}{\varepsilon} F(Z^\varepsilon_{z^\varepsilon})\right\} \le -\inf_{\phi\in\mathbb{U}} [I_z(\phi) + F(\phi)] \tag{9.12}$$

and lower bound

$$\liminf_{\varepsilon\to 0} \varepsilon \log E \exp\left\{-\frac{1}{\varepsilon} F(Z^\varepsilon_{z^\varepsilon})\right\} \ge -\inf_{\phi\in\mathbb{U}} [I_z(\phi) + F(\phi)]. \tag{9.13}$$

Proof of the Laplace upper bound. Fix $\delta > 0$ and recall from Sect. 8.3 that W^ψ denotes $W + \int_0^\cdot \psi(s)ds$. Using Theorem 8.19 and the definition (9.2) of $Z^\varepsilon_{z^\varepsilon}$, there exists $M < \infty$ such that for each $\varepsilon \in (0, 1)$, one can find $u^\varepsilon = (\psi^\varepsilon, \varphi^\varepsilon) \in \bar{\mathscr{A}}_{b,M}$ with

$$\begin{aligned}
&-\varepsilon \log E \exp\left\{-\frac{1}{\varepsilon} F(Z^\varepsilon_{z^\varepsilon})\right\} \\
&\quad = -\varepsilon \log E \exp\left\{-\frac{1}{\varepsilon} F \circ \mathscr{G}^\varepsilon(z^\varepsilon, \sqrt{\varepsilon} W, \varepsilon N^{1/\varepsilon})\right\} \\
&\quad \ge E\left[\bar{L}_T(u^\varepsilon) + F \circ \mathscr{G}^\varepsilon\left(z^\varepsilon, \sqrt{\varepsilon} W^{\psi^\varepsilon/\sqrt{\varepsilon}}, \varepsilon N^{\varphi^\varepsilon/\varepsilon}\right)\right] - \delta.
\end{aligned}$$

Using the compactness of S_M, we can find a subsequence $\{\varepsilon_k\}$ along which u^{ε_k} converges in distribution to some $u = (\psi, \varphi)$ that takes values in S_M a.s. By a standard subsequential argument, it is enough to demonstrate the lower bound for this subsequence, which for simplicity we label as ε. From part (b) of Condition 9.1, it follows that

$$\begin{aligned}
&\liminf_{\varepsilon\to 0} E\left[\bar{L}_T(u^\varepsilon) + F \circ \mathscr{G}^\varepsilon\left(z^\varepsilon, \sqrt{\varepsilon} W^{\psi^\varepsilon/\sqrt{\varepsilon}}, \varepsilon N^{\varphi^\varepsilon/\varepsilon}\right)\right] \\
&\quad \ge E\left[\bar{L}_T(u) + F \circ \mathscr{G}^0\left(z, \int_0^\cdot \psi(s)ds, \nu_T^\varphi\right)\right] \\
&\quad \ge \inf_{\{(\phi,q)\in\mathbb{U}\times S_M : q\in S^{\mathscr{G}}_{z,\phi}\}} \left[\bar{L}_T(q) + F(\phi)\right] \\
&\quad = \inf_{\phi\in\mathbb{U}} [I_z(\phi) + F(\phi)],
\end{aligned}$$

where the first inequality is a consequence of Fatou's lemma and the lower semicontinuity of $q \mapsto \bar{L}_T(q)$ on S_M, and the second inequality follows from the definition of $S^{\mathscr{G}}_{z,\phi}$, and the equality is due to the definition of $I_z(\phi)$ in (9.4). Since $\delta > 0$ is arbitrary, this completes the proof of the upper bound (9.12). □

Proof of the Laplace lower bound. We need to prove the inequality in (9.13). Without loss of generality we can assume that $\inf_{\phi\in\mathbb{U}}[I_z(\phi) + F(\phi)] < \infty$. Let $\delta > 0$ be arbitrary, and let $\phi_0 \in \mathbb{U}$ be such that

$$I_z(\phi_0) + F(\phi_0) \le \inf_{\phi\in\mathbb{U}}[I_z(\phi) + F(\phi)] + \frac{\delta}{2}. \tag{9.14}$$

Choose $q_0 = (f_0, g_0) \in S^{\mathscr{G}}_{z,\phi_0}$ such that

$$\bar{L}_T(q_0) \le I_z(\phi_0) + \frac{\delta}{2}. \tag{9.15}$$

Note that $\phi_0 = \mathscr{G}^0\left(z, \int_0^{\cdot} f_0(s)ds, \nu_T^{g_0}\right)$. Using the representation in Theorem 8.19, we obtain

$$\begin{aligned}
&\limsup_{\varepsilon\to 0} -\varepsilon \log E \exp\left\{-\frac{1}{\varepsilon}F(Z^{\varepsilon}_{z^{\varepsilon}})\right\} \\
&\quad= \limsup_{\varepsilon\to 0} \inf_{u=(\psi,\varphi)\in\bar{\mathscr{A}}} E\left[\bar{L}_T(u) + F\circ\mathscr{G}^{\varepsilon}\left(z^{\varepsilon}, \sqrt{\varepsilon}W^{\psi/\sqrt{\varepsilon}}, \varepsilon N^{\varphi/\varepsilon}\right)\right] \\
&\quad\le \limsup_{\varepsilon\to 0} E\left[\bar{L}_T(q_0) + F\circ\mathscr{G}^{\varepsilon}\left(z^{\varepsilon}, \sqrt{\varepsilon}W^{f_0/\sqrt{\varepsilon}}, \varepsilon N^{g_0/\varepsilon}\right)\right] \\
&\quad= \bar{L}_T(q_0) + \limsup_{\varepsilon\to 0} E\left[F\circ\mathscr{G}^{\varepsilon}\left(z^{\varepsilon}, \sqrt{\varepsilon}W^{f_0/\sqrt{\varepsilon}}, \varepsilon N^{g_0/\varepsilon}\right)\right].
\end{aligned}$$

By part (b) of Condition 9.1, we have

$$\begin{aligned}
\lim_{\varepsilon\to 0} E\left[F\circ\mathscr{G}^{\varepsilon}\left(z^{\varepsilon}, \sqrt{\varepsilon}W^{f_0/\sqrt{\varepsilon}}, \varepsilon N^{g_0/\varepsilon}\right)\right] &= F\circ\mathscr{G}^0\left(z, \int_0^{\cdot} f_0(s)ds, \nu_T^{g_0}\right) \\
&= F(\phi_0).
\end{aligned}$$

In view of (9.14) and (9.15), the left side of (9.13) can be at most $\inf_{\phi\in\mathbb{U}}[I_z(\phi) + F(\phi)] + \delta$. Since δ is arbitrary, the proof of the Laplace lower bound is complete. □

9.4 Proof of the Moderate Deviation Principle

In this section we prove Theorem 9.9. In order to show that I_z defined in (9.9) is a rate function on $\mathbb{U}$ and the family $\{I_z, z \in \mathscr{Z}\}$ of rate functions has compact level sets on compacts, we need to show that for every compact $K \subset \mathscr{Z}$ and $M < \infty$, the set $\Lambda_{M,K}$ defined in (9.11), but with I_z as in (9.9), is compact. The proof of this property is exactly the same as that in the proof of Theorem 9.2, except we make use of part (a) of Condition 9.8 instead of the corresponding part of Condition 9.1. We omit the details.

We next prove the second statement in the theorem. Fix $z \in \mathscr{Z}$ and let $\{z^\varepsilon\}_{\varepsilon>0} \subset \mathscr{Z}$ be such that $z^\varepsilon \to z$ as $\varepsilon \to 0$. Fix $F \in \mathscr{C}_b(\mathbb{U})$. It suffices to prove the Laplace upper and lower bounds:

$$\limsup_{\varepsilon\to 0} \varkappa(\varepsilon) \log E \exp\left\{-\frac{1}{\varkappa(\varepsilon)} F(Y^\varepsilon_{z^\varepsilon})\right\} \le -\inf_{\phi\in\mathbb{U}}[I_z(\phi) + F(\phi)], \tag{9.16}$$

$$\liminf_{\varepsilon\to 0} \varkappa(\varepsilon) \log E \exp\left\{-\frac{1}{\varkappa(\varepsilon)} F(Y^\varepsilon_{z^\varepsilon})\right\} \ge -\inf_{\phi\in\mathbb{U}}[I_z(\phi) + F(\phi)]. \tag{9.17}$$

Proof of the Laplace upper bound. Since $Y^{\varepsilon,z^\varepsilon} \doteq \mathscr{K}^\varepsilon(z^\varepsilon, \sqrt{\varepsilon}W, \varepsilon N^{\varepsilon^{-1}})$, Theorem 8.19 implies

$$\begin{aligned} &-\varkappa(\varepsilon) \log E \exp\left\{-\frac{1}{\varkappa(\varepsilon)} F(Y^\varepsilon_{z^\varepsilon})\right\} \\ &\quad = \inf_{u=(\psi,\varphi)\in\bar{\mathscr{A}}_b} E\left[\frac{1}{a^2(\varepsilon)}\bar{L}_T(u) + F\circ\mathscr{K}^\varepsilon\left(z^\varepsilon, \sqrt{\varepsilon}W^{\psi/\sqrt{\varepsilon}}, \varepsilon N^{\varphi/\varepsilon}\right)\right]. \end{aligned} \tag{9.18}$$

For later use, recall that by Theorem 8.3, $\bar{\mathscr{A}}_b$ in the representation can be replaced by $\bar{\mathscr{A}}$. Choose $\tilde{u}^\varepsilon = (\tilde{\psi}^\varepsilon, \tilde{\varphi}^\varepsilon) \in \bar{\mathscr{A}}_b$ such that

$$\begin{aligned} &-\varkappa(\varepsilon) \log E \exp\left\{-\frac{1}{\varkappa(\varepsilon)} F(Y^\varepsilon_{z^\varepsilon})\right\} + \varepsilon \\ &\ge E\left[\frac{1}{a^2(\varepsilon)}\left[L^W_T(\tilde{\psi}^\varepsilon) + L^N_T(\tilde{\varphi}^\varepsilon)\right] + F\circ\mathscr{K}^\varepsilon\left(z^\varepsilon, \sqrt{\varepsilon}W^{\tilde{\psi}^\varepsilon/\sqrt{\varepsilon}}, \varepsilon N^{\tilde{\varphi}^\varepsilon/\varepsilon}\right)\right]. \end{aligned} \tag{9.19}$$

Note that for $\varepsilon \in (0,1)$,

$$\frac{1}{a^2(\varepsilon)} E\left[L^W_T(\tilde{\psi}^\varepsilon) + L^N_T(\tilde{\varphi}^\varepsilon)\right] \le \tilde{M} \doteq (2\|F\|_\infty + 1).$$

We would like to argue as in Theorem 8.4 that one can consider controls that are in a certain sense bounded, but in this case the bound should depend on ε. Fix $\delta > 0$ and define

$$\tau^\varepsilon \doteq \inf\left\{t \in [0,T] : L^W_t(\tilde{\psi}^\varepsilon) \ge a^2(\varepsilon)2M \text{ or } L^N_t(\tilde{\varphi}^\varepsilon) \ge a^2(\varepsilon)2M\right\} \wedge T,$$

where $M \doteq \tilde{M}\|F\|_\infty/\delta$. Let

$$\varphi^\varepsilon(s,y) \doteq \tilde{\varphi}^\varepsilon(y,s)1_{\{s\le\tau^\varepsilon\}} + 1_{\{s>\tau^\varepsilon\}}, \quad \psi^\varepsilon(s) \doteq \tilde{\psi}^\varepsilon(s)1_{\{s\le\tau^\varepsilon\}}$$

for $(s, y) \in \mathscr{X}_T$. Then $u^\varepsilon \doteq (\psi^\varepsilon, \varphi^\varepsilon) \in \bar{\mathscr{A}}_b$,

$$L_T^N(\varphi^\varepsilon) \le a^2(\varepsilon)2M, \quad L_T^W(\psi^\varepsilon) \le a^2(\varepsilon)2M,$$

and

$$P\{\varphi^\varepsilon \neq \tilde{\varphi}^\varepsilon \text{ or } \psi^\varepsilon \neq \tilde{\psi}^\varepsilon\} \le \frac{1}{a^2(\varepsilon)2M} E\left[L_T^W(\tilde{\psi}^\varepsilon) + L_T^N(\tilde{\varphi}^\varepsilon)\right] \le \frac{\delta}{2\|F\|_\infty}.$$

For $(s, y) \in \mathscr{X}_T$, define the rescaled and (for φ^ε) centered controls

$$\zeta^\varepsilon(s, y) \doteq \frac{\varphi^\varepsilon(s, y) - 1}{a(\varepsilon)}, \quad \theta^\varepsilon(s) \doteq \frac{\psi^\varepsilon(s)}{a(\varepsilon)}.$$

Fix any $\beta \in (0, 1]$. Applying part (d) of Lemma 9.6 yields

$$\begin{aligned} E\left[\frac{1}{a^2(\varepsilon)} \int_{\mathscr{X}_T} \ell(\tilde{\varphi}^\varepsilon) d\nu_T\right] &\ge E\left[\frac{1}{a^2(\varepsilon)} \int_{\mathscr{X}_T} \ell(\varphi^\varepsilon) 1_{\{|\zeta^\varepsilon| \le \beta/a(\varepsilon)\}} d\nu_T\right] \\ &\ge E\left[\int_{\mathscr{X}_T} \left(\frac{1}{2}(\zeta^\varepsilon)^2 - \kappa_3 a(\varepsilon)|\zeta^\varepsilon|^3\right) 1_{\{|\zeta^\varepsilon| \le \beta/a(\varepsilon)\}} d\nu_T\right] \\ &\ge \left(\frac{1}{2} - \kappa_3\beta\right) E\left[\int_{\mathscr{X}_T} (\zeta^\varepsilon)^2 1_{\{|\zeta^\varepsilon| \le \beta/a(\varepsilon)\}} d\nu_T\right]. \end{aligned} \tag{9.20}$$

Also, from the definition of τ^ε, it follows that

$$\begin{aligned} &E\left[F \circ \mathscr{K}^\varepsilon\left(z^\varepsilon, \sqrt{\varepsilon} W^{\tilde{\psi}^\varepsilon/\sqrt{\varepsilon}}, \varepsilon N^{\tilde{\varphi}^\varepsilon/\varepsilon}\right) - F \circ \mathscr{K}^\varepsilon\left(z^\varepsilon, \sqrt{\varepsilon} W^{\psi^\varepsilon/\sqrt{\varepsilon}}, \varepsilon N^{\varphi^\varepsilon/\varepsilon}\right)\right] \\ &\le 2\|F\|_\infty P\{\varphi^\varepsilon \neq \tilde{\varphi}^\varepsilon \text{ or } \psi^\varepsilon \neq \tilde{\psi}^\varepsilon\} \\ &\le \delta. \end{aligned}$$

The definition of τ^ε implies $\varphi^\varepsilon \in S_{2M,+}^{N,\varepsilon}$, which was defined in (9.7), and thus part (c) of Lemma 9.7 implies an upper bound of $2M\kappa_2(\beta)$ on the expected value in (9.20). Using the last two displays, (9.19), and $\kappa_2(1) \ge \kappa_2(\beta)$, we have

$$\begin{aligned} &- \varkappa(\varepsilon) \log E \exp\left\{-\frac{1}{\varkappa(\varepsilon)} F(Y_{z^\varepsilon}^\varepsilon)\right\} \\ &\ge E\left[\frac{1}{2} \int_{\mathscr{X}_T} (\zeta^\varepsilon)^2 1_{\{|\zeta^\varepsilon| \le \beta/a(\varepsilon)\}} d\nu_T + L_T^W(\theta^\varepsilon)\right] \\ &\quad + E\left[F \circ \mathscr{K}^\varepsilon\left(z^\varepsilon, \sqrt{\varepsilon} W^{\psi^\varepsilon/\sqrt{\varepsilon}}, \varepsilon N^{\varphi^\varepsilon/\varepsilon}\right)\right] - \delta - \varepsilon - 2\beta\kappa_3 M\kappa_2(1). \end{aligned} \tag{9.21}$$

Recall $\hat{S}_n$ defined in (9.10), and note that $\{\theta^\varepsilon, \zeta^\varepsilon 1_{\{|\zeta^\varepsilon| \le \beta/a(\varepsilon)\}}\}$ is a sequence in the compact set $\hat{S}_K$ for sufficiently large but finite K, and is therefore automatically tight. Let (θ, ζ) be a limit point along a subsequence that we index once more by

ε. By a standard argument by contradiction, it suffices to prove (9.16) along this subsequence. Using part (b) of Condition 9.8, we have that along this subsequence,

$$\mathscr{K}^{\varepsilon}\left(z^{\varepsilon}, \sqrt{\varepsilon}W^{\psi^{\varepsilon}/\sqrt{\varepsilon}}, \varepsilon N^{\varphi^{\varepsilon}/\varepsilon}\right) \Rightarrow \mathscr{K}^{0}(z,\theta,\zeta) \doteq \eta.$$

Hence taking limits in (9.21) along this subsequence yields

$$\begin{aligned}
&\liminf_{\varepsilon\to 0} -\varkappa(\varepsilon)\log E\exp\left\{-\frac{1}{\varkappa(\varepsilon)}F(Y^{\varepsilon}_{z^{\varepsilon}})\right\}\\
&\geq E\left[\frac{1}{2}(\|\theta\|^2_{W,2}+\|\zeta\|^2_{N,2})+F(\eta)\right]-\delta-\beta\kappa_3 M\kappa_2(1)\\
&\geq E\left[I_z(\eta)+F(\eta)\right]-\delta-\beta\kappa_3 M\kappa_2(1)\\
&\geq \inf_{\eta\in\mathbb{U}}\left[I_z(\eta)+F(\eta)\right]-\delta-\beta\kappa_3\kappa_2(1)\frac{\tilde{M}\|F\|_\infty}{\delta},
\end{aligned}$$

where the first line is from Fatou's lemma, and the second uses the definition of I_z in (9.9). Sending first β to 0 and then δ to 0 gives (9.16). □

Proof of the Laplace lower bound. For $\delta > 0$, there exists $\eta \in \mathbb{U}$ such that

$$I_z(\eta)+F(\eta) \leq \inf_{\eta\in\mathbb{U}}[I_z(\eta)+F(\eta)]+\delta/2. \tag{9.22}$$

Choose $(\theta,\zeta)\in S^{\mathscr{K}}_{z,\eta}$ such that

$$\frac{1}{2}(\|\theta\|^2_{W,2}+\|\zeta\|^2_{N,2}) \leq I_z(\eta)+\delta/2. \tag{9.23}$$

For $\beta \in (0,1]$, define

$$\zeta^{\varepsilon} \doteq \zeta 1_{\{|\zeta|\leq\beta/a(\varepsilon)\}},\quad \varphi^{\varepsilon} \doteq 1+a(\varepsilon)\zeta^{\varepsilon},\quad \psi^{\varepsilon} \doteq a(\varepsilon)\theta.$$

For every $\varepsilon > 0$, using $\zeta^{\varepsilon} = (\varphi^{\varepsilon}-1)/a(\varepsilon)$ and part (d) of Lemma 9.6, we have

$$\int_{\mathscr{X}_T}\ell(\varphi^{\varepsilon})d\nu_T \leq \kappa_3\int_{\mathscr{X}_T}(\varphi^{\varepsilon}-1)^2 d\nu_T = a^2(\varepsilon)\kappa_3\int_{\mathscr{X}_T}|\zeta^{\varepsilon}|^2 d\nu_T \leq a^2(\varepsilon)M,$$

where $M \doteq \kappa_3\int_{\mathscr{X}_T}|\zeta|^2 d\nu_T$. Thus $\varphi^{\varepsilon}\in\mathscr{U}^{\varepsilon}_{M,+}$, with this space defined in (9.8), for all $\varepsilon > 0$. Also

$$\zeta^{\varepsilon}1_{\{|\zeta^{\varepsilon}|\leq\beta/a(\varepsilon)\}} = \zeta 1_{\{|\zeta|\leq\beta/a(\varepsilon)\}},$$

which converges to ζ in $L^2(\nu_T)$ as $\varepsilon\to 0$. Thus by part (b) of Condition 9.8,

$$\mathscr{K}^{\varepsilon}\left(z^{\varepsilon}, \sqrt{\varepsilon}W^{\psi^{\varepsilon}/\sqrt{\varepsilon}},\ \varepsilon N^{\varphi^{\varepsilon}/\varepsilon}\right) \Rightarrow \mathscr{K}^{0}(z,\theta,\zeta) = \eta. \tag{9.24}$$

Using part (d) of Lemma 9.6 and $\varkappa(\varepsilon)\varepsilon^{-1} = 1/a(\varepsilon)^2$, we obtain

$$\begin{aligned}\varkappa(\varepsilon)\varepsilon^{-1} L_T^N(\varphi^\varepsilon) &\le \frac{1}{2}\int_{\mathscr{X}_T} |\zeta^\varepsilon|^2\, d\nu_T + \kappa_3 \int_{\mathscr{X}_T} a(\varepsilon)|\zeta^\varepsilon|^3\, d\nu_T \\ &\le \frac{1}{2}(1+2\kappa_3\beta)\int_{\mathscr{X}_T} |\zeta|^2\, d\nu_T.\end{aligned}$$

For φ^ε as defined in terms of ζ, there is no guarantee that $\varphi^\varepsilon \in \bar{\mathscr{A}}_b^N$. However, as noted previously, the variational representation (9.18) holds with $\bar{\mathscr{A}}_b$ replaced by $\bar{\mathscr{A}}$. Hence by the last display,

$$\begin{aligned}&-\varkappa(\varepsilon)\log E\exp\left\{-\frac{1}{\varkappa(\varepsilon)}F(Y^\varepsilon_{z^\varepsilon})\right\} \\ &\le \frac{1}{a^2(\varepsilon)}\left[L_T^W(\psi^\varepsilon) + L_T^N(\varphi^\varepsilon)\right] + E\left[F\circ\mathscr{K}^\varepsilon\left(z^\varepsilon, \sqrt{\varepsilon}W^{\psi^\varepsilon/\sqrt{\varepsilon}}, \varepsilon N^{\varphi^\varepsilon/\varepsilon}\right)\right] \\ &\le \frac{1}{2}(\|\theta\|^2_{W,2} + \|\zeta\|^2_{N,2}) + E\left[F\circ\mathscr{K}^\varepsilon\left(z^\varepsilon, \sqrt{\varepsilon}W^{\psi^\varepsilon/\sqrt{\varepsilon}}, \varepsilon N^{\varphi^\varepsilon/\varepsilon}\right)\right] \\ &\quad + \kappa_3\beta\int_{\mathscr{X}_T}|\zeta|^2\, d\nu_T.\end{aligned}$$

Taking the limit as $\varepsilon \to 0$ and using (9.24) yields

$$\begin{aligned}&\limsup_{\varepsilon\to 0} -\varkappa(\varepsilon)\log E\exp\left\{-\frac{1}{\varkappa(\varepsilon)}F(Y^\varepsilon_{z^\varepsilon})\right\} \\ &\quad\le \frac{1}{2}(\|\theta\|^2_{W,2} + \|\zeta\|^2_{N,2}) + F(\eta) + \kappa_3\beta\int|\zeta|^2 d\nu_T.\end{aligned}$$

Finally, sending $\beta \to 0$ gives

$$\begin{aligned}\limsup_{\varepsilon\to 0} -\varkappa(\varepsilon)\log E\exp\left\{-\frac{1}{\varkappa(\varepsilon)}F(Y^\varepsilon_{z^\varepsilon})\right\} &\le \frac{1}{2}(\|\theta\|^2_{W,2} + \|\zeta\|^2_{N,2}) + F(\eta) \\ &\le I_z(\eta) + F(\eta) + \delta/2 \\ &\le \inf_{\eta\in\mathbb{U}}[I_z(\eta) + F(\eta)] + \delta,\end{aligned}$$

where the second inequality is from (9.23) and the last inequality follows from (9.22). Since $\delta > 0$ is arbitrary, this completes the proof of (9.17) and consequently the proof of Theorem 9.9. □

9.5 Notes

The sufficient condition for a Laplace principle given in Theorem 9.2, in the case that there is no Poisson noise, was established in [39], and the general case was treated in

[45], where its application to the study of a small noise LDP for finite dimensional jump-diffusions was studied as well. The sufficient condition given in Theorem 9.9 for the case in which the driving noise does not have a Gaussian component was given in [41]. This work also gave applications of Theorem 9.9 to the study of moderate deviation principles for finite and infinite dimensional stochastic dynamical systems driven by Poisson random measures.

The sufficient conditions given in this chapter have found applications in many different problems. Some of the works that have used the sufficient condition for Brownian motion functional given in [39] include [20–22, 37, 43, 44, 63, 64, 71, 91, 142, 143, 156, 191, 192, 196, 203, 207, 212–216, 218, 236, 244, 250, 254–256, 263, 268, 270]. Sufficient conditions given in this chapter for functionals of PRM and BM have been used in [12, 38, 47, 55, 70, 75, 251, 253, 258, 262, 264, 267, 272]. MDP sufficient conditions have found applications in [48, 57, 186–188, 194, 257, 265, 273].