

# Chapter 10

## Large and Moderate Deviations for Finite Dimensional Systems



In this chapter we use the abstract sufficient conditions from Chap. 9 to prove large and moderate deviation principles for small noise finite dimensional jump-diffusions. We will consider only Laplace principles rather than uniform Laplace principles, since, as was noted in Chap. 9, the extension from the nonuniform to the uniform case is straightforward. The first general results on large deviation principles for jump-diffusions of the form considered in this chapter are due to Wentzell [245–248] and Freidlin and Wentzell [140]. The conditions for an LDP identified in the current chapter relax some of the assumptions made in these works. Results on moderate deviation principles in this chapter are based on the recent work [41]. We do not aim for maximal generality, and from the proofs it is clear that many other models (e.g., time inhomogeneous jump diffusions, SDEs with delay) can be treated in an analogous fashion.

The perspective here is different from that of the discrete time model of Chap. 4. In particular, the emphasis is on viewing the stochastic model as a relatively well behaved mapping on a fixed noise space consisting of a Brownian motion and one or more Poisson random measures. This has important consequences for models with degeneracy, i.e., systems for which the noise does not push the state in all directions. Processes of this sort in the general discrete time setting required complicated assumptions such as Condition 4.8 and a delicate mollification argument as in Sect. 4.8. In contrast, the degeneracy is essentially irrelevant when the process of interest can be viewed as a nice mapping (at least asymptotically) on a fixed noise space. This distinction becomes even more significant for the infinite dimensional models of Chap. 11, where the analogous degeneracy is ubiquitous. In this chapter we use Lipschitz continuity assumptions on the coefficients to guarantee that the mapping is well behaved. However, this is not necessary, especially with regard to Poisson noise, and for one such weakening we refer to Sect. 13.3.

The chapter is organized as follows. Section 10.1 introduces the basic stochastic process model that will be considered. Conditions under which the stochastic equation and its deterministic analogue have unique solutions are given. In Sect. 10.2 we use Theorem 9.2 to establish an LDP for the solution under additional integrability

conditions. Finally, in Sect. 10.3 we apply Theorem 9.9 to prove a moderate deviations result. For this result the integrability conditions we require are somewhat weaker, though additional smoothness conditions on the coefficients are assumed so that one can easily expand around the LLN limit, and the proof of tightness, as with the discrete time model of Chap. 5, is more involved than for the large deviation counterpart. We use the notation from Chap. 9, except that in this chapter,  $W$  is a finite dimensional standard Brownian motion, i.e.,  $\mathcal{H}_0 = \mathcal{H} = \mathbb{R}^d$ ,  $\mathbb{W} = \mathcal{C}([0, T] : \mathbb{R}^d)$ , and  $\Lambda$  is the identity operator.

## 10.1 Small Noise Jump-Diffusion

We consider small noise stochastic differential equations (SDEs) of the form

$$\begin{aligned} X^\varepsilon(t) = x_0 + \int_0^t b(X^\varepsilon(s))ds + \sqrt{\varepsilon} \int_0^t \sigma(X^\varepsilon(s))dW(s) \\ + \varepsilon \int_{\mathcal{X}_i} G(X^\varepsilon(s-), y)N^{1/\varepsilon}(ds \times dy), \end{aligned} \quad (10.1)$$

where  $W$  is a standard  $d$ -dimensional Wiener process,  $N^{1/\varepsilon}$  is a PRM with intensity measure  $\lambda_T \times \nu$  (see Definition 8.11) constructed from  $\bar{N}$  as in (8.16) with  $\varphi = 1/\varepsilon$ , and  $W, \bar{N}$  satisfy (a)–(c) in Sect. 9.1. The coefficients are assumed to satisfy the following condition.

**Condition 10.1** *The functions  $b : \mathbb{R}^d \rightarrow \mathbb{R}^d$ ,  $\sigma : \mathbb{R}^d \rightarrow \mathbb{R}^{d \times d}$ , and  $G : \mathbb{R}^d \times \mathcal{X} \rightarrow \mathbb{R}^d$  satisfy*

(a) *for some  $L_b \in (0, \infty)$ ,*

$$\|b(x) - b(\bar{x})\| \leq L_b \|x - \bar{x}\|, \quad x, \bar{x} \in \mathbb{R}^d;$$

(b) *for some  $L_\sigma \in (0, \infty)$ ,*

$$\|\sigma(x) - \sigma(\bar{x})\| \leq L_\sigma \|x - \bar{x}\|, \quad x, \bar{x} \in \mathbb{R}^d;$$

(c) *for some  $L_G \in \mathcal{L}^1(\nu)$ ,*

$$\|G(x, y) - G(\bar{x}, y)\| \leq L_G(y) \|x - \bar{x}\|, \quad x, \bar{x} \in \mathbb{R}^d, \quad y \in \mathcal{X};$$

(d) *for some  $M_G \in \mathcal{L}^1(\nu)$ ,*

$$\|G(x, y)\| \leq M_G(y)(1 + \|x\|), \quad x \in \mathbb{R}^d, \quad y \in \mathcal{X}.$$

The following result follows by standard arguments (see Theorem IV.9.1 of [159]). It says that under Condition 10.1, Eq. (10.1) has a unique pathwise solution. In

applying results from Chap. 9, we take  $\mathbb{U} = \mathcal{D}([0, T] : \mathbb{R}^d)$ , i.e., the space of  $\mathbb{R}^d$ -valued right-continuous functions with left limits and the usual Skorokhod topology [24, Chap. 3, Sect. 12].

**Theorem 10.2** Fix  $x_0 \in \mathbb{R}^d$ , and assume Condition 10.1. Then for each  $\varepsilon > 0$ , there is a measurable map  $\mathcal{G}^\varepsilon : \mathbb{V} \rightarrow \mathcal{D}([0, T] : \mathbb{R}^d)$  such that for every probability space  $(\tilde{\Omega}, \tilde{\mathcal{F}}, \tilde{P})$  on which are given a  $d$ -dimensional Brownian motion  $\tilde{W}$  and an independent Poisson random measure  $\tilde{N}_\varepsilon$  on  $\mathcal{X}_T$  with intensity measure  $\varepsilon^{-1} \nu_T$ ,  $\tilde{X}^\varepsilon \doteq \mathcal{G}^\varepsilon(\sqrt{\varepsilon} \tilde{W}, \varepsilon \tilde{N}_\varepsilon)$  is an  $\tilde{\mathcal{F}}_t \doteq \sigma\{\tilde{W}(s), \tilde{N}_\varepsilon(B \times [0, s]), s \leq t, B \in \mathcal{B}(\mathcal{X}), \nu(B) < \infty\}$  adapted process that is the unique solution of the stochastic integral equation

$$\begin{aligned} \tilde{X}^\varepsilon(t) = x_0 &+ \int_0^t b(\tilde{X}^\varepsilon(s)) ds + \sqrt{\varepsilon} \int_0^t \sigma(\tilde{X}^\varepsilon(s)) d\tilde{W}(s) \\ &+ \varepsilon \int_{\mathcal{X}_t} G(\tilde{X}^\varepsilon(s-), y) \tilde{N}_\varepsilon(ds \times dy), \end{aligned} \tag{10.2}$$

for  $t \in [0, T]$ . In particular,  $X^\varepsilon = \mathcal{G}^\varepsilon(\sqrt{\varepsilon} W, \varepsilon N^{1/\varepsilon})$  is the unique solution of (10.1).

## 10.2 An LDP for Small Noise Jump-Diffusions

The solution  $X^\varepsilon$  of (10.1) is a  $\mathcal{D}([0, T] : \mathbb{R}^d)$ -valued random variable. To prove a large deviation principle for  $\{X^\varepsilon\}_{\varepsilon > 0}$  as  $\varepsilon \rightarrow 0$ , we will assume the following additional condition on the coefficient function  $G$ . For  $\rho \in (0, \infty)$ , let  $\mathcal{L}_{\text{exp}}^\rho$  be the collection of all measurable  $\theta : \mathcal{X} \rightarrow \mathbb{R}_+$  such that whenever  $A \in \mathcal{B}(\mathcal{X})$  satisfies  $\nu(A) < \infty$ ,

$$\int_A e^{\rho\theta(y)} \nu(dy) < \infty. \tag{10.3}$$

Let  $\mathcal{L}_{\text{exp}} \doteq \bigcap_{\rho \in (0, \infty)} \mathcal{L}_{\text{exp}}^\rho$ .

**Condition 10.3**  $M_G \in \mathcal{L}_{\text{exp}}$  and  $L_G \in \mathcal{L}_{\text{exp}}^\rho$  for some  $\rho > 0$ .

*Remark 10.4* This exponential integrability condition on jump distributions is a natural requirement for the model; it should be compared with Condition 4.3, assumed in the study of small noise discrete time Markov recursive systems. Consider, for example, the case in which  $\mathcal{X} = \mathbb{R}^d$ ,  $\nu \in \mathcal{P}(\mathbb{R}^d)$  satisfies  $\int_{\mathbb{R}^d} e^{\langle \alpha, y \rangle} \nu(dy) < \infty$  for all  $\alpha \in \mathbb{R}^d$ , and for some  $d \times d$  matrix  $A$  and a vector  $w \in \mathbb{R}^d$ ,

$$G(x, y) = Ay + y \langle x, w \rangle.$$

Then  $G$  satisfies parts (c) and (d) of Condition 10.1, as well as Condition 10.3. In the general case (in which  $\nu$  need not be a probability measure), note that the local rate function corresponding to just the jump part of the SDE (10.2) would be

$$L(x, \beta) = \inf \left[ \int_{\mathcal{X}} \ell(g(y))\nu(dy) : \int_{\mathcal{X}} G(x, y)g(y)\nu(dy) = \beta \right]. \tag{10.4}$$

Using convex duality and that  $\ell(b)$  is dual to  $(e^a - 1)$ , for this to be superlinear in  $\beta$  [a condition needed for the rate function on path space to have compact level sets in the usual topology of  $\mathcal{D}([0, T] : \mathbb{R}^d)$ ], one needs

$$\int_{\mathcal{X}} [e^{\langle \alpha, G(x, y) \rangle} - 1] \nu(dy) < \infty \text{ for all } \alpha \in \mathbb{R}^d. \tag{10.5}$$

However, this follows under Conditions 10.1 and 10.3. One can break the integral in (10.5) according to  $\{y : M_G(y) > 1\}$  and  $\{y : M_G(y) \leq 1\}$ . Then  $M_G \in \mathcal{L}^1(\nu)$  implies  $\nu(\{y : M_G(y) > 1\}) < \infty$ , and the integral over  $\{y : M_G(y) > 1\}$  is finite due to Condition 10.3. The mean value theorem gives the bound  $\|\alpha\| e^{\|\alpha\|} M_G(y)$  for the integrand on  $\{y : M_G(y) \leq 1\}$ , and finiteness for the corresponding integral follows from  $M_G \in \mathcal{L}^1(\nu)$ .

We recall that  $S \doteq \cup_{n \in \mathbb{N}} S_n$ , with each  $S_n \doteq S_n^W \times S_n^N$  compact in the appropriate topology. The proof of the following theorem proceeds using a standard argument and is given after Lemma 10.8.

**Theorem 10.5** *Fix  $x_0 \in \mathbb{R}^d$ , and assume Conditions 10.1 and 10.3. Then for each  $q = (f, g) \in S$ , there is a unique  $\xi = \xi_q \in \mathcal{C}([0, T] : \mathbb{R}^d)$  such that for all  $t \in [0, T]$ ,*

$$\begin{aligned} \xi(t) = & x_0 + \int_0^t b(\xi(s))ds + \int_0^t \sigma(\xi(s))f(s)ds \\ & + \int_{\mathcal{X}_t} G(\xi(s), y)g(s, y)\nu(dy)ds. \end{aligned} \tag{10.6}$$

For  $q = (f, g) \in S$ , let  $\xi = \xi_q$  denote the solution of (10.6). Let  $I : \mathcal{D}([0, T] : \mathbb{R}^d) \rightarrow [0, \infty]$  be defined by

$$I(\phi) \doteq \inf_{q \in S : \phi = \xi_q} \bar{L}_T(q), \tag{10.7}$$

where  $\bar{L}_T(q) \doteq L_T^W(f) + L_T^N(g)$ , with the individual costs defined as in (9.1).

**Theorem 10.6** *Assume Conditions 10.1 and 10.3. Then  $I$  is a rate function on  $\mathcal{D}([0, T] : \mathbb{R}^d)$  and  $\{X^\varepsilon\}_{\varepsilon > 0}$  satisfies a large deviation principle on  $\mathcal{D}([0, T] : \mathbb{R}^d)$  with rate function  $I$ .*

Following our standard convention, the proof is given for  $T = 1$ . Before proceeding with the proof, we present two lemmas. The first will be used to prove tightness. Recall that  $g \in S_n^N$  means that  $\int_{[0, T] \times \mathcal{X}} \ell(g(u, y))\nu(dy)du \leq n$ . For a function  $f : [0, 1] \rightarrow \mathbb{R}^k$ , define  $\|f\|_{\infty, t} \doteq \sup_{0 \leq s \leq t} \|f(s)\|$  for  $t \in [0, 1]$ . Note that

the constant  $c(\delta, n)$  appearing in the lemma may also depend on the function  $\theta$ . However, when the lemma is used, this will be a fixed quantity, such as  $M_G(y)$ , that is associated with a particular process model under consideration.

**Lemma 10.7** *Let  $\theta \in \mathcal{L}_{exp}$  and suppose that  $v(\{\theta > 1\}) < \infty$ . Then for every  $\delta > 0$  and  $n \in \mathbb{N}$ , there exists  $c(\delta, n) \in (1, \infty)$  such that for all  $\tilde{\theta} : \mathcal{X} \rightarrow \mathbb{R}_+$  satisfying  $\tilde{\theta} \leq \theta$ , every measurable map  $f : [0, 1] \rightarrow \mathbb{R}_+$ , and all  $0 \leq s \leq t \leq 1$ ,*

$$\begin{aligned} & \sup_{g \in S_n^N} \int_{(s,t] \times \mathcal{X}} f(u) \tilde{\theta}(y) g(u, y) v(dy) du \\ & \leq c(\delta, n) \left( \int_{\mathcal{X}} \tilde{\theta}(y) v(dy) \right) \left( \int_s^t f(u) du \right) + \delta \|f\|_{\infty,1}. \end{aligned} \tag{10.8}$$

*Proof* Let  $f : [0, 1] \rightarrow \mathbb{R}_+$ ,  $g \in S_n^N$ , and  $\delta > 0$  be given. Then for each  $m \in (0, \infty)$ ,

$$\int_{(s,t] \times \mathcal{X}} f(u) \tilde{\theta}(y) g(u, y) v(dy) du = T_1(m) + T_2(m), \tag{10.9}$$

where

$$T_1(m) \doteq \int_{(s,t] \times \{\theta \leq m\}} f(u) \tilde{\theta}(y) g(u, y) v(dy) du,$$

and

$$T_2(m) \doteq \int_{(s,t] \times \{\theta > m\}} f(u) \tilde{\theta}(y) g(u, y) v(dy) du.$$

Using part (a) of Lemma 9.6 with  $\sigma = k$ ,  $a = \theta(y)$  and  $b = g(u, y)$ , for each  $k \in \mathbb{N}$  we have the bound

$$T_2(m) \leq \|f\|_{\infty,1} \left( \int_{\{\theta > m\}} e^{k\theta(y)} v(dy) + \frac{n}{k} \right).$$

Also, for each  $\beta \in (1, \infty)$ ,  $T_1(m)$  can be bounded by

$$T_1(m) \leq T_3(m, \beta) + T_4(m, \beta),$$

where

$$\begin{aligned} T_3(m, \beta) & \doteq \int_{E_1(m, \beta)} f(u) \tilde{\theta}(y) g(u, y) v(dy) du, \\ T_4(m, \beta) & \doteq \int_{E_2(m, \beta)} f(u) \tilde{\theta}(y) g(u, y) v(dy) du, \end{aligned}$$

and

$$E_1(m, \beta) \doteq \{(u, y) \in (s, t] \times \mathcal{X} : \theta(y) \leq m \text{ and } g(u, y) \leq \beta\},$$

$$E_2(m, \beta) \doteq \{(u, y) \in (s, t] \times \mathcal{X} : \theta(y) \leq m \text{ and } g(u, y) > \beta\}.$$

Using part (b) of Lemma 9.6 and that  $g \in S_n^N$ , we obtain

$$T_3(m, \beta) + T_4(m, \beta) \leq \beta \left( \int_{\mathcal{X}} \tilde{\theta}(y) \nu(dy) \right) \left( \int_s^t f(u) du \right) + \bar{\kappa}_1(\beta) mn \|f\|_{\infty,1}. \tag{10.10}$$

Thus the left side of (10.8) can be bounded by

$$\beta \left( \int_{\mathcal{X}} \tilde{\theta}(y) \nu(dy) \right) \left( \int_s^t f(u) du \right) + \|f\|_{\infty,1} \left( \bar{\kappa}_1(\beta) mn + \int_{\{\theta > m\}} e^{k\theta(y)} \nu(dy) + \frac{n}{k} \right).$$

Given  $\delta > 0$ , choose  $k \in \mathbb{N}$  such that  $n/k < \delta/3$ . Then use that  $\theta \in \mathcal{L}_{\text{exp}}$  to choose  $m \in (0, \infty)$  such that  $\int_{\{\theta > m\}} e^{k\theta(y)} \nu(dy) < \delta/3$ . This is possible, since  $\nu(\{\theta > 1\}) < \infty$ . Finally, choose  $\beta \in (1, \infty)$  such that  $\bar{\kappa}_1(\beta) mn < \delta/3$ . The result now follows by taking  $c(\delta, n) = \beta$ .  $\square$

The following lemma is proved similarly to Lemma 10.7 and therefore only a sketch is provided. Recall that  $\mathcal{X}_1 \doteq [0, 1] \times \mathcal{X}$  and  $\nu_1 \in \mathcal{P}(\mathcal{X}_1)$  is the product measure  $\nu_1(ds \times dy) = \nu(dy)ds$ .

**Lemma 10.8** *Let  $\theta \in \mathcal{L}_{\text{exp}}^\rho \cap \mathcal{L}^1(\nu)$  for some  $\rho \in (0, \infty)$ . Then for every  $n \in \mathbb{N}$ ,*

$$\sup_{g \in S_n^N} \int_{\mathcal{X}_1} \theta(y) g(u, y) \nu(dy) du < \infty.$$

*Proof* Consider the equality in (10.9) with  $m = 1, s = 0, t = 1, f = 1$ , and  $\tilde{\theta} = \theta$ . Then as in the proof of Lemma 10.7,

$$T_2(1) \leq \int_{\{\theta > 1\}} e^{\rho\theta(y)} \nu(dy) + \frac{n}{\rho}.$$

Also, as with the proof of (10.10),

$$T_1(1) \leq \int_{\mathcal{X}} \theta(y) \nu(dy) + \bar{\kappa}_1(1)n.$$

The result follows by combining the two estimates.  $\square$

*Proof of Theorem 10.5.* Fix  $q = (f, g) \in S$  and let  $k \in \mathbb{N}$  be such that  $q \in S_k$ . We first prove the existence of a solution to (10.6). Consider a sequence  $\{\phi_n\}$  in  $\mathcal{C}([0, 1] : \mathbb{R}^d)$

constructed recursively as follows. Define  $\phi_1(t) \doteq x_0$  for all  $t \in [0, 1]$ , and for  $n \in \mathbb{N}$ , let

$$\begin{aligned} \phi_{n+1}(t) &\doteq x_0 + \int_0^t b(\phi_n(s))ds + \int_0^t \sigma(\phi_n(s))f(s)ds \\ &\quad + \int_{\mathcal{X}_t} G(\phi_n(s), y)g(s, y)v(dy)ds, \quad t \in [0, 1]. \end{aligned} \quad (10.11)$$

Using the growth conditions on  $b$ ,  $\sigma$ , and  $G$ , we have that there is a  $c_1 \in (0, \infty)$  such that for all  $n \in \mathbb{N}$  and  $t \in [0, 1]$ ,

$$\begin{aligned} \|\phi_{n+1}\|_{\infty, t} &\leq \|x_0\| + c_1 \int_0^t (1 + \|\phi_n\|_{\infty, s})(1 + |f(s)|)ds \\ &\quad + c_1 \int_{\mathcal{X}_t} M_G(y)(1 + \|\phi_n\|_{\infty, s})g(s, y)v(dy)ds. \end{aligned}$$

Thus, with  $h(s) \doteq (1 + |f(s)| + \int_{\mathcal{X}} M_G(y)g(s, y)v(dy))$ , for some  $c_2 \in (0, \infty)$ , we have

$$\|\phi_{n+1}\|_{\infty, t} \leq c_2 \left( 1 + \int_0^t \|\phi_n\|_{\infty, s} h(s) ds \right), \quad t \in [0, 1], n \in \mathbb{N}.$$

From Lemma 10.8, we obtain  $\int_{[0, 1]} h(s)ds < \infty$ . A standard recursive argument now shows that for all  $n$ ,  $\|\phi_n\|_{\infty, 1} \leq c_2 \exp \int_0^1 h(s)ds < \infty$ .

Using Lemma 10.7, it is easily seen that for each  $n \in \mathbb{N}$ ,  $\phi_n \in \mathcal{C}([0, 1] : \mathbb{R}^d)$ . Indeed, the continuity of the last term in (10.11) follows on observing that for every  $\delta > 0$  and  $0 \leq s \leq t \leq 1$ ,

$$\begin{aligned} &\int_{(s, t] \times \mathcal{X}} \|G(\phi_n(u), y)\|g(u, y)v(dy)du \\ &\leq (1 + \|\phi_n\|_{\infty, 1}) \left( c(\delta, k)(t - s) \int_{\mathcal{X}} M_G(y)v(dy) + \delta \right). \end{aligned}$$

For  $n \in \mathbb{N}$  and  $t \in [0, 1]$ , let  $a_n(t) \doteq \|\phi_{n+1} - \phi_n\|_{\infty, t}$ . Then there exists  $c_3 \in (0, \infty)$  such that for all  $n \geq 2$  and  $t \in [0, 1]$ ,

$$\begin{aligned} a_n(t) &\leq c_3 \int_0^t a_{n-1}(s)ds + c_3 \int_0^t a_{n-1}(s)|f(s)|ds \\ &\quad + \int_0^t a_{n-1}(s) \left( \int_{\mathcal{X}} L_G(y)g(s, y)v(dy) \right) ds. \end{aligned}$$

Thus, with  $m(s) \doteq c_3(1 + |f(s)|) + \int_{\mathcal{X}} L_G(y)g(s, y)v(dy)$ , we have for all  $t \in [0, 1]$  and  $n \geq 2$  that

$$a_n(t) \leq \int_0^t a_{n-1}(s)m(s)ds.$$

This shows that for all  $n \in \mathbb{N}$ ,

$$a_{n+1}(1) \leq a_1(1) \frac{\left(\int_0^1 m(s)ds\right)^n}{n!}.$$

Lemma 10.8 implies  $\int_{\mathcal{X}_1} L_G(y)g(s, y)v(dy)ds < \infty$ , and thus  $\int_0^1 m(s)ds < \infty$ . From this it follows that  $\{\phi_n\}$  is a Cauchy sequence in  $\mathcal{C}([0, 1] : \mathbb{R}^d)$  and therefore must converge to some  $\phi \in \mathcal{C}([0, 1] : \mathbb{R}^d)$ . From the continuity of  $b$  and  $\sigma$  it follows that for every  $t \in [0, 1]$ ,

$$\int_0^t b(\phi_n(s))ds + \int_0^t \sigma(\phi_n(s))f(s)ds \rightarrow \int_0^t b(\phi(s))ds + \int_0^t \sigma(\phi(s))f(s)ds.$$

Also,

$$\begin{aligned} & \int_{\mathcal{X}_1} \|G(\phi_n(s), y) - G(\phi(s), y)\|g(s, y)v(dy)ds \\ & \leq \|\phi_n - \phi\|_{\infty, 1} \int_{\mathcal{X}_1} L_G(y)g(s, y)v(dy)ds. \end{aligned}$$

Since  $\int_{\mathcal{X}_1} L_G(y)g(s, y)v(dy)ds < \infty$ , the right-hand side in the last display converges to 0 as  $n \rightarrow \infty$ . Combining these observations, we have that  $\phi$  solves (10.6), proving the existence of solutions.

We now consider uniqueness. Suppose that  $\phi_1, \phi_2$  are two solutions of (10.6) in  $\mathcal{C}([0, 1] : \mathbb{R}^d)$ . Then using the Lipschitz property of  $b, \sigma$ , and  $G$ , we have that for some  $c_4 \in (0, \infty)$  and all  $t \in [0, 1]$ ,

$$\begin{aligned} \|\phi_1 - \phi_2\|_{\infty, t} & \leq c_4 \int_0^t \|\phi_1 - \phi_2\|_{\infty, s} ds + c_4 \int_0^t \|\phi_1 - \phi_2\|_{\infty, s} |f(s)| ds \\ & \quad + \int_0^t \|\phi_1 - \phi_2\|_{\infty, s} \left( \int_{\mathcal{X}} L_G(y)g(s, y)v(dy) \right) ds. \end{aligned}$$

Thus

$$\|\phi_1 - \phi_2\|_{\infty, t} \leq \int_0^t \|\phi_1 - \phi_2\|_{\infty, s} \left( c_4 + c_4 |f(s)| + \int_{\mathcal{X}} L_G(y)g(s, y)v(dy) \right) ds.$$

Recalling that  $\int_{\mathcal{X}_1} L_G(y)g(s, y)v(dy) < \infty$ , an application of Gronwall's lemma implies  $\phi_1 = \phi_2$ .  $\square$

The following lemma is useful in characterizing the limit points of weakly converging controlled jump processes. It will be used in the proof of Theorem 10.6.



Recall that as discussed in Sect. 9.1,  $g_k \rightarrow g$  in the topology of  $S_n^N$  if  $\int_{\mathcal{X}_1} f g_k d\nu_1 \rightarrow \int_{\mathcal{X}_1} f g d\nu_1$  for bounded continuous  $f$  with compact support.

**Lemma 10.9** Fix  $n \in \mathbb{N}$  and let  $K$  be a compact subset of  $\mathcal{X}_1$ . Let  $g, g_k \in S_n^N, k \in \mathbb{N}$  be such that  $g_k \rightarrow g$ . Also, let  $\gamma : \mathcal{X}_1 \rightarrow \mathbb{R}$  be a bounded measurable function. Then as  $k \rightarrow \infty$ ,

$$\int_K \gamma(s, y) g_k(s, y) \nu(dy) ds \rightarrow \int_K \gamma(s, y) g(s, y) \nu(dy) ds. \tag{10.12}$$

*Proof* We assume that  $\nu_1(K) \neq 0$ , since otherwise, the result is trivially true. By replacing, if needed,  $g_k$  and  $g$  with  $g_k + 1_K$  and  $g + 1_K$  respectively, we can assume without loss of generality that  $\int_K g(s, y) \nu_1(ds \times dy) \neq 0$  and  $\int_K g_k(s, y) \nu_1(ds \times dy) \neq 0$  for all  $k \in \mathbb{N}$ . If (10.12) holds with  $K$  replaced by  $K_1$ , where  $K \subset K_1$  for some compact  $K_1$ , then by taking  $\tilde{\gamma} = \gamma 1_{K_1}$ , we see that (10.12) also holds with the compact set  $K$ . Also, since  $\nu$  is finite on every compact set, we can always find a compact set  $K_1 \supset K$  such that  $\nu_1(\partial K_1) = 0$ . Hence in proving the lemma, we can assume without loss of generality that  $\nu_1(\partial K) = 0$ . Recall from Chap. 9 that for  $g \in S_n^N$ ,  $\nu_1^g$  is defined by setting  $\nu_1^g(A) = \int_A g(s, y) \nu_1(ds \times dy)$  for  $A \in \mathcal{B}(\mathcal{X}_1)$ . Define probability measures  $\tilde{\nu}^k$  and  $\tilde{\nu}$  as follows:

$$\tilde{\nu}^k(\cdot) \doteq \frac{\nu_1^{g_k}(\cdot \cap K)}{m_k}, \quad \tilde{\nu}(\cdot) \doteq \frac{\nu_1^g(\cdot \cap K)}{m},$$

where  $m \doteq \nu_1^g(K)$  and  $m_k \doteq \nu_1^{g_k}(K)$ . Let  $\theta(\cdot) \doteq \nu_1(\cdot \cap K) / \nu_1(K)$ . Then

$$\begin{aligned} R(\tilde{\nu}^k \parallel \theta) &= \int_K \log \left( \frac{\nu_1(K)}{m_k} g_k(s, y) \right) \frac{1}{m_k} g_k(s, y) \nu_1(ds \times dy) \\ &= \frac{1}{m_k} \int_K [\ell(g_k(s, y)) + g_k(s, y) - 1] \nu_1(ds \times dy) + \log \frac{\nu_1(K)}{m_k} \\ &\leq \frac{n}{m_k} + 1 - \frac{\nu_1(K)}{m_k} + \log \frac{\nu_1(K)}{m_k}. \end{aligned}$$

Since  $g_k \rightarrow g$  and  $\nu_1(\partial K) = 0$ , it follows that  $m_k \rightarrow m$ , and therefore the last display implies  $\sup_{k \in \mathbb{N}} R(\tilde{\nu}^k \parallel \theta) < \infty$ . Also note that  $\tilde{\nu}^k$  converges weakly to  $\tilde{\nu}$ . From Lemma 2.5, it follows that

$$\frac{1}{m_k} \int_K \gamma(s, y) g_k(s, y) \nu_1(ds \times dy) \rightarrow \frac{1}{m} \int_K \gamma(s, y) g(s, y) \nu_1(ds \times dy),$$

which proves (10.12). □

### 10.2.1 Proof of the Large Deviation Principle

In this section we prove Theorem 10.6. Theorem 10.2 shows that the solution to the SDE (10.2) can be expressed as a measurable mapping on the input noises:  $X^\varepsilon = \mathcal{G}^\varepsilon(\sqrt{\varepsilon}W, \varepsilon N^{1/\varepsilon})$ . We now verify that  $\mathcal{G}^\varepsilon$  satisfies Condition 9.1, which by Theorem 9.2 will complete the proof of the LDP. The notation used is that of Sect. 9.2. In particular, the spaces  $\mathbb{V}$  and  $\bar{\mathbb{V}}$  are the canonical spaces for a Brownian motion and PRM on  $\mathcal{X}_1$  and  $\mathcal{Y}_1$ , respectively.

Define  $\mathcal{G}^0 : \mathbb{V} \rightarrow \mathcal{D}([0, T] : \mathbb{R}^d)$  as follows. If  $(w, m) \in \mathbb{V}$  is of the form  $(w, m) = (\int_0^\cdot f(s)ds, \nu_1^g)$  for some  $q = (f, g) \in \mathcal{S}$ , set

$$\mathcal{G}^0(w, m) = \mathcal{G}^0\left(\int_0^\cdot f(s)ds, \nu_1^g\right) = \xi_q,$$

where  $\xi_q$  is the unique solution of (10.6). For all other  $(w, m) \in \mathbb{V}$  set  $\mathcal{G}^0(w, m) = 0$ . Since  $\bar{L}_T(q) = \infty$  for such  $q$ , with this definition,  $I$  defined in (10.7) is the same as the function  $I$  defined in (9.4).

We will show that part (b) of Condition 9.1, which is the weak convergence of the controlled processes as  $\varepsilon \rightarrow 0$ , holds with this choice of  $\mathcal{G}^0$ . Part (a) of the condition follows if we prove continuity of  $q \mapsto \mathcal{G}^0(q)$  for  $q$  such that  $\bar{L}_T(q) \leq n$  (recall that the initial condition has been assumed fixed). This is in fact an easier deterministic analogue of the proof of part (b), and hence omitted (see, for example, the proof of Theorem 11.25). Fix  $n \in \mathbb{N}$  and let  $u^\varepsilon = (\psi^\varepsilon, \varphi^\varepsilon) \in \bar{\mathcal{A}}_{b,n}$ ,  $u = (\psi, \varphi) \in \mathcal{A}_{b,n}$  be such that  $u^\varepsilon$  converges in distribution to  $u$  as  $\varepsilon \rightarrow 0$ . We recall that this implies the a.s. bounds

$$\int_0^1 \|\psi^\varepsilon(s)\|^2 ds \leq n \text{ and } \int_{\mathcal{X}_1} \ell(\varphi^\varepsilon(s, x)) \nu_1(ds \times dx) \leq n. \quad (10.13)$$

Furthermore, almost surely  $\varphi^\varepsilon(s, x)$  has upper and lower bounds of the form  $1/\delta$  and  $\delta$  for all  $x$  in some compact set  $K$ , and  $\varphi^\varepsilon(s, x) = 1$  for  $x \notin K$  (where  $\delta > 0$  and  $K$  depend on  $\varphi^\varepsilon$ ). The analogous statements also hold for  $(\psi, \varphi)$ , and as we will see, the ability to restrict to controls with such nice properties greatly simplifies the arguments. Almost all of the difficulties in the proof are due to the jump term [for comparison, one can consider the proof of the analogous diffusion model in Theorem 3.19].

Let  $\tilde{\varphi}^\varepsilon = 1/\varphi^\varepsilon$  and for  $t \in [0, 1]$  define

$$\begin{aligned} \mathcal{E}_1^\varepsilon(t) \doteq & \exp \left[ \int_{\mathcal{X}_T \times [0, \infty)} 1_{[0, \varphi^\varepsilon(s, y)/\varepsilon]}(r) \log(\tilde{\varphi}^\varepsilon(s, y)) \bar{N}(ds \times dy \times dr) \right. \\ & \left. + \int_{\mathcal{X}_T \times [0, \infty)} 1_{[0, \varphi^\varepsilon(s, y)/\varepsilon]}(r) (-\tilde{\varphi}^\varepsilon(s, y) + 1) \bar{\nu}_1(ds \times dy \times dr) \right] \end{aligned}$$

and

$$\mathcal{E}_2^\varepsilon(t) \doteq \exp \left[ -\frac{1}{\sqrt{\varepsilon}} \int_0^t \psi^\varepsilon(s) dW(s) - \frac{1}{2\varepsilon} \int_0^t \|\psi^\varepsilon(s)\|^2 ds \right].$$

Let  $\mathcal{E}^\varepsilon(t) = \mathcal{E}_1^\varepsilon(t)\mathcal{E}_2^\varepsilon(t)$ . Then using the independence between the Brownian and Poisson noises, it follows that  $\{\mathcal{E}^\varepsilon(t)\}_{0 \leq t \leq 1}$  is a  $\mathcal{F}_t$ -martingale, and consequently

$$\bar{Q}^\varepsilon(A) = \int_A \mathcal{E}^\varepsilon(1) dP, \quad A \in \mathcal{B}(\bar{\mathbb{V}})$$

defines a probability measure on  $\bar{\mathbb{V}}$ . The bounds (10.13) on  $\psi^\varepsilon$  and  $\varphi^\varepsilon$  along with the properties of  $\varphi^\varepsilon$  in terms of some compact set  $K$  noted below (10.13) imply that  $P$  and  $\bar{Q}^\varepsilon$  are mutually absolutely continuous, and by Girsanov's theorem [see Theorem D.3],

$$\left( \sqrt{\varepsilon} W^{\psi^\varepsilon/\sqrt{\varepsilon}}, \varepsilon N^{\varphi^\varepsilon/\varepsilon} \right)$$

under  $\bar{Q}^\varepsilon$  has the same probability law as  $(\sqrt{\varepsilon} W, \varepsilon N^{1/\varepsilon})$  under  $P$ , where we recall  $W^{\psi/\sqrt{\varepsilon}} \doteq W + \int_0^\cdot \psi(s) ds / \sqrt{\varepsilon}$ . Thus it follows that  $\bar{X}^\varepsilon = \mathcal{G}^\varepsilon(\sqrt{\varepsilon} W^{\psi^\varepsilon/\sqrt{\varepsilon}}, \varepsilon N^{\varphi^\varepsilon/\varepsilon})$  is the unique solution, under both  $\bar{Q}^\varepsilon$  and  $P$ , of the controlled SDE given by  $\bar{X}^\varepsilon(0) = x_0$  and

$$\begin{aligned} d\bar{X}^\varepsilon(t) &= [b(\bar{X}^\varepsilon(t)) + \sigma(\bar{X}^\varepsilon(t))\psi^\varepsilon(t)] dt + \sqrt{\varepsilon}\sigma(\bar{X}^\varepsilon(t))dW(t) \\ &\quad + \varepsilon \int_{\mathcal{X}} G(\bar{X}^\varepsilon(t-), x) N^{\varphi^\varepsilon/\varepsilon}(dt \times dx). \end{aligned} \tag{10.14}$$

We end this section by stating martingale bounds that will be useful in the sequel. Recall the definition of  $\mathcal{A}^N$  from Sect. 8.3. Also recall that for  $\varphi \in \mathcal{A}^N$ ,  $N_c^\varphi$  denotes the compensated form of  $N^\varphi$ .

**Lemma 10.10** *Let  $\varphi \in \mathcal{A}^N$ , and assume that  $\psi : [0, 1] \times \Omega \times \mathcal{X} \rightarrow \mathbb{R}$  is  $\mathcal{P}\mathcal{F} \otimes \mathcal{B}(\mathcal{X})/\mathcal{B}(\mathbb{R})$ -measurable and*

$$E \int_{\mathcal{X}_1} (|\psi(s, x)| \vee |\psi(s, x)|^2) \varphi(s, x) \nu_1(ds \times dx) < \infty.$$

*Then there exists  $C \in (0, \infty)$  such that for all  $t \in [0, T]$*

$$E \left[ \sup_{0 \leq s \leq t} \left| \int_{\mathcal{X}_t} \psi(s, x) N_c^\varphi(ds \times dx) \right| \right] \leq CE \left[ \int_{\mathcal{X}_t} \psi(s, x)^2 \varphi(s, x) \nu_1(ds \times dx) \right]^{\frac{1}{2}},$$

*and there is also the bound*

$$E \left[ \sup_{0 \leq s \leq t} \left| \int_{\mathcal{X}_t} \psi(s, x) N_c^\varphi(ds \times dx) \right|^2 \right] \leq 4E \left[ \int_{\mathcal{X}_t} \psi(s, x)^2 \varphi(s, x) \nu_1(ds \times dx) \right].$$

*Proof* The bounds follow from Doob's maximal inequality and the Lenglart–Lepingle–Pratelli inequality [(D.2) and (D.4) of Appendix D], and from expressions for the quantities on the right-hand sides that are stated in Sect. D.2.2. In particular, in applying (D.2), we use that the expected quadratic variation of the stochastic integral in the last display is

$$E \int_{\mathcal{X}_t} \psi(s, x)^2 N^\varphi(ds \times dx) = E \int_{\mathcal{X}_t} \psi(s, x)^2 \varphi(s, x) \nu_1(ds \times dx).$$

□

### 10.2.1.1 Tightness

**Lemma 10.11** *Assume Conditions 10.1 and 10.3. Given controls  $(\psi^\varepsilon, \varphi^\varepsilon) \in \bar{\mathcal{A}}_{b,n}$ , let  $\bar{X}^\varepsilon$  be the corresponding unique solution to (10.14). Then  $\{\bar{X}^\varepsilon\}$  is a tight family of  $\mathcal{D}([0, 1] : \mathbb{R}^d)$ -valued random variables.*

*Proof* We begin with an estimate on the supremum of  $\bar{X}^\varepsilon(t)$ . Recalling  $\|x\|_{\infty,t} \doteq \sup_{0 \leq s \leq t} \|x(s)\|$ , by Condition 10.1 we have that for suitable  $c_1 < \infty$ ,

$$\begin{aligned} \|\bar{X}^\varepsilon\|_{\infty,t} &\leq \|x_0\| + c_1 \int_0^t (1 + \|\bar{X}^\varepsilon\|_{\infty,s})(1 + \|\psi^\varepsilon(s)\|) ds \\ &\quad + \sqrt{\varepsilon} \left\| \int_0^\cdot \sigma(\bar{X}^\varepsilon(s)) dW(s) \right\|_{\infty,t} \\ &\quad + \varepsilon \left\| \int_{\mathcal{X}_\cdot} M_G(y)(1 + \|\bar{X}^\varepsilon(s-)\|) N_c^{\varphi^\varepsilon/\varepsilon}(ds \times dy) \right\|_{\infty,t} \\ &\quad + \int_{\mathcal{X}_t} M_G(y)(1 + \|\bar{X}^\varepsilon\|_{\infty,s}) \varphi^\varepsilon(s, y) \nu(dy) ds, \end{aligned}$$

where  $N_c^{\varphi^\varepsilon/\varepsilon}(ds \times dy) = N^{\varphi^\varepsilon/\varepsilon}(ds \times dy) - \varepsilon^{-1} \varphi^\varepsilon(s, y) \nu_1(ds \times dy)$ . Let

$$\begin{aligned} \mathcal{R}_t^\varepsilon &\doteq \sqrt{\varepsilon} \left\| \int_0^\cdot \sigma(\bar{X}^\varepsilon(s)) dW(s) \right\|_{\infty,t} \\ &\quad + \varepsilon \left\| \int_{\mathcal{X}_\cdot} M_G(y)(1 + \|\bar{X}^\varepsilon(s-)\|) N_c^{\varphi^\varepsilon/\varepsilon}(ds \times dy) \right\|_{\infty,t}. \end{aligned} \tag{10.15}$$

Using the bound (10.13) for  $\psi^\varepsilon$  and Hölder's inequality, Gronwall's inequality [Lemma E.2] gives

$$\begin{aligned} 1 + \|\bar{X}^\varepsilon\|_{\infty,1} &\leq (1 + \|x_0\| + \mathcal{R}_1^\varepsilon) \exp \left\{ c_1(1 + \sqrt{n}) + \int_{\mathcal{X}_1} M_G(y) \varphi^\varepsilon(s, y) \nu(dy) ds \right\} \\ &\leq c_2(1 + \|x_0\| + \mathcal{R}_1^\varepsilon), \end{aligned} \tag{10.16}$$

where, using Lemma 10.7 with  $\delta = 1$ ,  $f \equiv 1$  and  $g = \varphi^\varepsilon$  and the fact that  $M_G \in \mathcal{L}^1(\nu) \cap \mathcal{L}_{\text{exp}}$ , we obtain

$$c_2 \doteq \exp \left\{ c_1(1 + \sqrt{n}) + c(1, n) \int_{\mathcal{X}} M_G(y) \nu(dy) + 1 \right\} < \infty.$$

Also, Condition 10.1 implies that for some  $c_3 < \infty$ ,

$$\sqrt{\varepsilon} E \left\| \int_0^\cdot \sigma(\bar{X}^\varepsilon(s)) dW(s) \right\|_{\infty, 1} \leq c_3 \sqrt{\varepsilon} (E \|\bar{X}^\varepsilon\|_{\infty, 1} + 1).$$

Let  $m \in (0, \infty)$ . Then the expectation of the second term in the definition of  $\mathcal{R}_t^\varepsilon$  can be bounded by the sum of

$$T_1^\varepsilon \doteq \varepsilon E \left\| \int_{\mathcal{X}_t} M_G(y) 1_{\{M_G \leq m\}} (1 + \|\bar{X}^\varepsilon(s-)\|) N_c^{\varphi^\varepsilon/\varepsilon}(ds \times dy) \right\|_{\infty, 1} \quad (10.17)$$

and

$$T_2^\varepsilon \doteq 2E \int_{\mathcal{X}_t} M_G(y) 1_{\{M_G \geq m\}} (1 + \|\bar{X}^\varepsilon(s)\|) \varphi^\varepsilon(s, y) \nu(dy) ds. \quad (10.18)$$

For (10.18), we use the representation  $N_c^{\varphi^\varepsilon/\varepsilon} = N^{\varphi^\varepsilon/\varepsilon} - \varepsilon^{-1} \varphi^\varepsilon \nu_1$ , that the corresponding integrals are almost surely nondecreasing in  $t$ , and the identity

$$\begin{aligned} \varepsilon E \int_{\mathcal{X}_t} M_G(y) 1_{\{M_G \geq m\}} (1 + \|\bar{X}^\varepsilon(s-)\|) N_c^{\varphi^\varepsilon/\varepsilon}(ds \times dy) \\ = E \int_{\mathcal{X}_t} M_G(y) 1_{\{M_G \geq m\}} (1 + \|\bar{X}^\varepsilon(s)\|) \varphi^\varepsilon(s, y) \nu(dy) ds. \end{aligned}$$

An application of Lemmas 10.7 and 10.10 as used before yield that for some  $c_4 < \infty$ ,

$$\begin{aligned} T_1^\varepsilon &\leq c_4 \sqrt{\varepsilon m} E \left[ (1 + \|\bar{X}^\varepsilon\|_{\infty, 1}) \int_{\mathcal{X}_t} M_G(y) \varphi^\varepsilon(s, y) \nu(dy) ds \right]^{1/2} \\ &\leq c_4 \sqrt{\varepsilon m} (1 + E \|\bar{X}^\varepsilon\|_{\infty, 1}) \left( c(1, n) \int_{\mathcal{X}} M_G(y) \nu(dy) + 1 \right)^{1/2}. \end{aligned}$$

Also, for every  $\delta > 0$ ,

$$T_2^\varepsilon \leq 2(1 + E \|\bar{X}^\varepsilon\|_{\infty, 1}) \left( c(\delta, n) \int_{\mathcal{X}} M_G(y) 1_{\{M_G \geq m\}} \nu(dy) + \delta \right).$$

Choosing  $\delta > 0$  sufficiently small and then  $m < \infty$  sufficiently large, there is  $\varepsilon_0 > 0$  such that for all  $\varepsilon \leq \varepsilon_0$ ,

$$E\mathcal{R}_1^\varepsilon = c_3\sqrt{\varepsilon} (E\|\bar{X}^\varepsilon\|_{\infty,1} + 1) + T_1^\varepsilon + T_2^\varepsilon \leq \frac{1}{2c_2} (E\|\bar{X}^\varepsilon\|_{\infty,1} + 1).$$

Using this in (10.16) then gives that for  $\varepsilon \leq \varepsilon_0$ ,

$$\bar{E}\|\bar{X}^\varepsilon\|_{\infty,1} \leq 2c_2(\|x_0\| + 1),$$

and therefore

$$\sup_{\varepsilon \leq \varepsilon_0} \bar{E}\|\bar{X}^\varepsilon\|_{\infty,1} < \infty. \tag{10.19}$$

Henceforth we consider only  $\varepsilon < \varepsilon_0$ . We next argue that  $\mathcal{R}_1^\varepsilon$  defined in (10.15) converges to 0 in probability. The term with the Brownian motion is easy. Using Condition 10.1, the estimate in (10.19), and the Burkholder–Davis–Gundy inequality, it follows that

$$\sqrt{\varepsilon} \left\| \int_{[0, \cdot]} \sigma(\bar{X}^\varepsilon(s)) dW(s) \right\|_{\infty,1} \rightarrow 0 \text{ in probability as } \varepsilon \rightarrow 0. \tag{10.20}$$

Next we consider the Poisson term, and write

$$\begin{aligned} & \varepsilon \int_{\mathcal{X}_t} G(\bar{X}^\varepsilon(s-), y) N^{\varphi^\varepsilon/\varepsilon}(ds \times dy) \\ &= \varepsilon \int_{\mathcal{X}_t} G(\bar{X}^\varepsilon(s-), y) N_c^{\varphi^\varepsilon/\varepsilon}(ds \times dy) + \int_{\mathcal{X}_t} G(\bar{X}^\varepsilon(s), y) \varphi^\varepsilon(s, y) \nu(dy) ds. \end{aligned} \tag{10.21}$$

Consider the first term on the right side of (10.21). For  $\alpha \in (0, \infty)$ , define the stopping time  $\tau_\alpha^\varepsilon \doteq \inf\{s : \|\bar{X}^\varepsilon(s)\| > \alpha\}$ . We first show that

$$T^{\alpha,\varepsilon} \doteq \varepsilon \left\| \int_{(0, \tau_\alpha^\varepsilon \wedge \cdot] \times \mathcal{X}} G(\bar{X}^\varepsilon(s-), y) N_c^{\varphi^\varepsilon/\varepsilon}(ds \times dy) \right\|_{\infty,1}$$

converges to zero in probability as  $\varepsilon \rightarrow 0$ . To do this, note that for every  $r \in (0, \infty)$ ,

$$T^{\alpha,\varepsilon} \leq T_{\leq r}^{\alpha,\varepsilon} + T_{> r}^{\alpha,\varepsilon},$$

where

$$T_{\leq r}^{\alpha,\varepsilon} \doteq \varepsilon \left\| \int_{(0, \tau_\alpha^\varepsilon \wedge \cdot] \times \{M_G(y) \leq r\}} G(\bar{X}^\varepsilon(s-), y) N_c^{\varphi^\varepsilon/\varepsilon}(ds \times dy) \right\|_{\infty,1},$$

$$T_{>r}^{\alpha,\varepsilon} \doteq \varepsilon \left\| \int_{(0, \tau_\alpha^\varepsilon \wedge 1] \times \{M_G(y) > r\}} G(\bar{X}^\varepsilon(s-), y) N_c^{\varphi^\varepsilon/\varepsilon}(ds \times dy) \right\|_{\infty,1}.$$

By Lemma 10.10, there is  $c_5 \in (0, \infty)$  such that

$$E(T_{\leq r}^{\alpha,\varepsilon})^2 \leq \varepsilon c_5 E \int_{(0, \tau_\alpha^\varepsilon \wedge 1] \times \{M_G(y) \leq r\}} M_G^2(y) (1 + \|\bar{X}^\varepsilon(s)\|^2) \varphi^\varepsilon(s, y) \nu(dy) ds.$$

We then use that  $\|\bar{X}^\varepsilon(s)\| \leq \alpha$  for  $s \in (0, \tau_\alpha^\varepsilon \wedge 1)$  and write  $(0, \tau_\alpha^\varepsilon \wedge 1] \times \{M_G(y) \leq r\}$  as the disjoint union of two sets, the first for which  $\varphi^\varepsilon(s, y) < \beta$  and the second for which  $\varphi^\varepsilon(s, y) \geq \beta$ , and finally apply part (b) of Lemma 9.6 and use (10.13) to get

$$E(T_{\leq r}^{\alpha,\varepsilon})^2 \leq \varepsilon c_5 (1 + \alpha^2) (\beta r \|M_G\|_1 + r^2 n \bar{\kappa}_1(\beta)).$$

To deal with the term  $T_{>r}^{\alpha,\varepsilon}$ , note that for every  $k \in \mathbb{N}$ ,

$$\begin{aligned} ET_{>r}^{\alpha,\varepsilon} &\leq 2E \int_{(0, \tau_\alpha^\varepsilon \wedge 1] \times \{M_G(y) > r\}} \|G(\bar{X}^\varepsilon(s-), y)\| \varphi^\varepsilon(s, y) \nu(dy) ds \\ &\leq 2(1 + \alpha) E \int_{[0,1] \times \{M_G(y) > r\}} M_G(y) \varphi^\varepsilon(s, y) \nu(dy) ds \\ &\leq 2(1 + \alpha) \left( \int_{\{M_G(y) > r\}} e^{kM_G(y)} \nu(dy) + \frac{n}{k} \right), \end{aligned}$$

where the second inequality follows on using the growth bound stated in part (d) of Condition 10.1 and recalling the definition of  $\tau_\alpha^\varepsilon$ , and the last inequality is a consequence of part (a) of Lemma 9.6 with  $\sigma = k$ ,  $a = M_G(y)$  and  $b = \varphi^\varepsilon(s, y)$  and again the fact that  $\varphi^\varepsilon$  takes values in  $S_n^N$ . Since  $M_G \in \mathcal{L}^1(\nu) \cap \mathcal{L}_{\text{exp}}$ , for every  $k \in \mathbb{N}$ ,  $\int_{\{M_G(y) > r\}} e^{kM_G(y)} \nu(dy) \rightarrow 0$  as  $r \rightarrow \infty$ . Combining these two bounds and sending  $\varepsilon \rightarrow 0$ ,  $r \rightarrow \infty$ ,  $k \rightarrow \infty$  in that order shows that for each  $\alpha \in (0, \infty)$ ,

$$T^{\alpha,\varepsilon} \rightarrow 0 \text{ in probability as } \varepsilon \rightarrow 0. \tag{10.22}$$

Next let

$$T^\varepsilon \doteq \varepsilon \left\| \int_{(0, \cdot] \times \mathcal{X}} G(\bar{X}^\varepsilon(s-), y) N_c^{\varphi^\varepsilon/\varepsilon}(ds \times dy) \right\|_{\infty,1},$$

where the restriction on the time variable in  $T^{\alpha,\varepsilon}$  has been dropped. Defining  $A_\alpha \doteq \{\|\bar{X}^\varepsilon\|_{\infty,1} < \alpha\}$ , for all  $\eta > 0$ ,

$$\begin{aligned} P(T^\varepsilon > \eta) &= P(\{T^\varepsilon > \eta\} \cap A_\alpha) + P(\{T^\varepsilon > \eta\} \cap A_\alpha^c) \\ &\leq P(T_1^{\alpha,\varepsilon} > \eta) + P(A_\alpha^c). \end{aligned}$$

Combining this last bound with (10.22) and (10.19), we see that  $T^\varepsilon \rightarrow 0$  in probability as  $\varepsilon \rightarrow 0$ . Together with (10.20), this shows that

$$\mathcal{R}_1^\varepsilon \rightarrow 0 \text{ in probability as } \varepsilon \rightarrow 0.$$

Thus

$$\begin{aligned} \bar{X}^\varepsilon(t) &= x_0 + \int_0^t b(\bar{X}^\varepsilon(s))ds + \int_0^t \sigma(\bar{X}^\varepsilon(s))\psi^\varepsilon(s)ds \\ &\quad + \int_{\mathcal{X}_t} G(\bar{X}^\varepsilon(s), y)\varphi^\varepsilon(s, y)\nu(dy) ds + \bar{\mathcal{R}}^\varepsilon(t), \end{aligned} \tag{10.23}$$

where  $\|\bar{\mathcal{R}}^\varepsilon\|_{\infty,1} \leq \mathcal{R}_1^\varepsilon$  converges to 0 in probability as  $\varepsilon \rightarrow 0$ . Tightness of the terms in (10.23) that involve  $b$  or  $\sigma$  follows from standard estimates, and are the same as calculations used for the small noise diffusion model in Chap. 3. Thus in order to prove tightness of  $\{\bar{X}^\varepsilon\}$ , it suffices to argue that

$$\xi^\varepsilon(t) = \int_{\mathcal{X}_t} G(\bar{X}^\varepsilon(s), y)\varphi^\varepsilon(s, y)\nu(dy)ds, \quad t \in [0, 1]$$

is tight in  $\mathbb{U} = \mathcal{D}([0, 1] : \mathbb{R}^d)$ . By Lemma 10.7, for every  $\delta > 0$  and  $0 \leq s \leq t \leq 1$ ,

$$\begin{aligned} \|\xi^\varepsilon(t) - \xi^\varepsilon(s)\| &\leq \int_{[s,t] \times \mathcal{X}} (1 + \|\bar{X}^\varepsilon\|_{\infty,u})M_G(y)\varphi^\varepsilon(u, y)\nu(dy)du \\ &\leq (1 + \|\bar{X}^\varepsilon\|_{\infty,1})c(\delta, n)(t - s)\|M_G\|_1 + \delta(1 + \|\bar{X}^\varepsilon\|_{\infty,1}), \end{aligned}$$

where as before,  $\|\cdot\|_1$  is the norm in  $\mathcal{L}^1(\nu)$ . Tightness of  $\xi^\varepsilon$  is now a consequence of (10.19). Thus we have shown that  $\{\bar{X}^\varepsilon\}$  is tight in  $\mathcal{D}([0, 1] : \mathbb{R}^d)$ .  $\square$

### 10.2.1.2 Identification of Limits

The following lemma completes the verification of part (b) of Condition 9.1, and hence the proof of Theorem 10.6.

**Lemma 10.12** *Assume Conditions 10.1 and 10.3. Given controls  $(\psi^\varepsilon, \varphi^\varepsilon) \in \bar{\mathcal{A}}_{b,n}$ , let  $\bar{X}^\varepsilon$  be the corresponding unique solution to (10.14). Assume that  $(\psi^\varepsilon, \varphi^\varepsilon)$  converges in distribution to  $(\psi, \varphi)$ . Then  $\bar{X}^\varepsilon$  converges in distribution to the unique solution to (10.6) with  $(f, g) = (\psi, \varphi)$ .*

*Proof* From Lemma 10.11 it follows that if for some fixed  $n$ , the controls  $(\psi^\varepsilon, \varphi^\varepsilon)$  are in  $\bar{\mathcal{A}}_{b,n}$  for every  $\varepsilon > 0$ , then  $\{\bar{X}^\varepsilon\}_{\varepsilon>0}$  is a tight collection of  $\mathcal{D}([0, 1] : \mathbb{R}^d)$ -valued random variables. It was also shown in the proof of the lemma that  $\|\bar{\mathcal{R}}^\varepsilon\|_{\infty,1}$  appearing in (10.23) converges to 0 in probability as  $\varepsilon \rightarrow 0$ . It follows from this last property and (10.23) that  $\bar{X}$  has continuous sample paths a.s. By appealing to the



Skorohod representation theorem, we assume without loss of generality the almost sure convergence  $(\bar{X}^\varepsilon, \psi^\varepsilon, \varphi^\varepsilon, \bar{\mathcal{P}}^\varepsilon) \rightarrow (\bar{X}, \psi, \varphi, 0)$ . Using the assumed conditions on  $b$  and  $\sigma$ , it is straightforward [see, for example, the proof of Lemma 3.21] using Hölder's inequality and the dominated convergence theorem to show that for every  $t$ , the sum of the first three terms on the right side of (10.23) converges a.s. to

$$x_0 + \int_0^t b(\bar{X}(s))ds + \int_0^t \sigma(\bar{X}(s))\psi(s)ds.$$

In view of the unique solvability of (10.6), to complete the verification that  $\mathcal{G}^\varepsilon$  satisfies part (b) of Condition 9.1, it then suffices to show that for all  $t \in [0, 1]$ ,

$$\int_{\mathcal{X}_t} G(\bar{X}^\varepsilon(s), y)\varphi^\varepsilon(s, y)\nu(dy)ds - \int_{\mathcal{X}_t} G(\bar{X}(s), y)\varphi(s, y)\nu(dy)ds \rightarrow 0 \tag{10.24}$$

as  $\varepsilon \rightarrow 0$ .

We write the expression in (10.24) as  $T_3^\varepsilon(t) + T_4^\varepsilon(t)$ , where

$$\begin{aligned} T_3^\varepsilon(t) &\doteq \int_{\mathcal{X}_t} (G(\bar{X}^\varepsilon(s), y) - G(\bar{X}(s), y))\varphi^\varepsilon(s, y)\nu(dy)ds, \\ T_4^\varepsilon(t) &\doteq \int_{\mathcal{X}_t} G(\bar{X}(s), y)(\varphi^\varepsilon(s, y) - \varphi(s, y))\nu(dy)ds. \end{aligned}$$

Using Condition 10.1, we obtain

$$T_3^\varepsilon(t) \leq \|\bar{X}^\varepsilon - \bar{X}\|_{\infty, 1} \int_{\mathcal{X}_1} L_G(y)\varphi^\varepsilon(s, y)\nu(dy)ds.$$

Since  $L_G \in \mathcal{L}^1(\nu) \cap \mathcal{L}_{\text{exp}}^\rho$ , we see from Lemma 10.8 that  $T_3^\varepsilon(t) \rightarrow 0$  a.s. as  $\varepsilon \rightarrow 0$ . Let  $\{K_r\}_{r \in \mathbb{N}}$  be a sequence of compact subsets of  $\mathcal{X}$  such that  $K_r \uparrow \mathcal{X}$  as  $r \rightarrow \infty$ , and let  $E_r \doteq K_r \cap \{M_G \leq r\}$ . Write  $T_4^\varepsilon(t) = T_{4,r \leq}^\varepsilon(t) + T_{4,r >}^\varepsilon(t)$ , where

$$\begin{aligned} T_{4,r \leq}^\varepsilon(t) &\doteq \int_{\mathcal{X}_t} G(\bar{X}(s), y)1_{E_r}(y)(\varphi^\varepsilon(s, y) - \varphi(s, y))\nu(dy)ds \\ T_{4,r >}^\varepsilon(t) &\doteq \int_{\mathcal{X}_t} G(\bar{X}(s), y)1_{E_r^c}(y)(\varphi^\varepsilon(s, y) - \varphi(s, y))\nu(dy)ds. \end{aligned}$$

Using Lemma 10.9, for every  $r \in (0, \infty)$ ,  $T_{4,r \leq}^\varepsilon(t) \rightarrow 0$  as  $\varepsilon \rightarrow 0$ . Also, using Lemma 10.7 again, for every  $\delta > 0$ ,

$$\begin{aligned} T_{4,r>}^\varepsilon(t) &\leq (1 + \|\bar{X}\|_{\infty,1}) \int_{\mathcal{X}_r} M_G(y) 1_{E_r^c}(y) (\varphi^\varepsilon(s, y) + \varphi(s, y)) \nu(dy) ds \\ &\leq 2(1 + \|\bar{X}\|_{\infty,1}) \left( c(\delta, n)t \int_{\mathcal{X}} M_G(y) 1_{E_r^c}(y) \nu(dy) + \delta \right). \end{aligned}$$

Since  $M_G \in L^1(\nu)$ , it follows that  $\sup_{\varepsilon \in (0, \varepsilon_0)} T_{4,r>}^\varepsilon(t) \rightarrow 0$  if we first send  $r \rightarrow \infty$  and then  $\delta \rightarrow 0$ . Combining these two estimates, we have that  $T_4^\varepsilon(t)$  converges to 0 as  $\varepsilon \rightarrow 0$ . Thus we have proved (10.24), which completes the proof of the lemma and therefore, as noted previously, also the proof of Theorem 10.6.  $\square$

### 10.3 An MDP for Small Noise Jump-Diffusions

Throughout this section we assume Condition 10.1, which implies Lipschitz continuity and linear growth conditions on  $b$ ,  $\sigma$ , and  $G$  in the  $x$  variable. Let  $X^0 \in \mathcal{C}([0, T] : \mathbb{R}^d)$  be the unique solution of the equation

$$X^0(t) = x_0 + \int_0^t b(X^0(s)) ds + \int_{\mathcal{X}_t} G(X^0(s), y) \nu(dy) ds, \quad t \in [0, T].$$

We now establish a Laplace principle for  $\{Y^\varepsilon\}$  with scaling function  $\varkappa(\varepsilon)$ , where

$$Y^\varepsilon = \frac{1}{a(\varepsilon)} (X^\varepsilon - X^0),$$

and as in (9.6),  $a(\varepsilon)$  satisfies  $a(\varepsilon) \rightarrow 0$  and  $\varkappa(\varepsilon) \doteq \varepsilon/a^2(\varepsilon) \rightarrow 0$ . For the MDP we assume some additional smoothness on the coefficients. For a differentiable function  $f : \mathbb{R}^d \rightarrow \mathbb{R}^d$  let  $Df(x) = (\partial f_i(x)/\partial x_j)_{i,j}$ . Following our convention, for matrices we use the operator norm, so that  $\|Df(x)\| \doteq \sup_{w \in \mathbb{R}^d : \|w\|=1} \|Df(x)w\|$ . Similarly, if  $g : \mathbb{R}^d \times \mathcal{X} \rightarrow \mathbb{R}^d$  is differentiable in  $x$  for each fixed  $y \in \mathcal{X}$  and  $D_x g(x, y) = (\partial g_i(x, y)/\partial x_j)_{i,j}$ , then  $\|D_x g(x, y)\|$  denotes the norm of this matrix.

For the MDP, the integrability assumption on  $M_G$  in Condition 10.3 can be weakened, analogous to the corresponding weakening in going from the LDP to MDP in the setting of discrete time models (Chaps. 4 and 5). The following is the only assumption besides Condition 10.1 needed for the MDP.

**Condition 10.13** (a) *The functions  $L_G$  and  $M_G$  are in  $\mathcal{L}^1(\nu) \cap \mathcal{L}_{exp}^\rho$  for some  $\rho \in (0, \infty)$ .*

(b) *For every  $y \in \mathcal{X}$ , the maps  $x \mapsto b(x)$  and  $x \mapsto G(x, y)$  are differentiable. For some  $L_{Db} \in (0, \infty)$ ,*

$$\|Db(x) - Db(\bar{x})\| \leq L_{Db} \|x - \bar{x}\|, \quad x, \bar{x} \in \mathbb{R}^d;$$

for some  $L_{DG} \in \mathcal{L}^1(\nu)$ ,

$$\|D_x G(x, y) - D_x G(\bar{x}, y)\| \leq L_{DG}(y) \|x - \bar{x}\|, \quad x, \bar{x} \in \mathbb{R}^d, \quad y \in \mathcal{X};$$

and lastly,

$$\sup_{\{x \in \mathbb{R}^d : \|x\| \leq \|X^0\|_{\infty, T}\}} \int_{\mathcal{X}} \|D_x G(x, y)\| \nu(dy) < \infty.$$

Recall from Chap. 9 that we define  $\mathcal{L}^2 \doteq \mathcal{L}^2([0, T] : \mathcal{H}_0) \times \mathcal{L}^2(\nu_T)$ , and that in this chapter,  $\mathcal{H}_0 = \mathbb{R}^d$ . For  $q = (f_1, f_2) \in \mathcal{L}^2$ , consider the equation

$$\begin{aligned} \eta(t) &= \int_0^t [Db(X^0(s))] \eta(s) ds + \int_{\mathcal{X}_t} [D_x G(X^0(s), y)] \eta(s) \nu(dy) ds \\ &\quad + \int_0^t \sigma(X^0(s)) f_1(s) ds + \int_{\mathcal{X}_t} G(X^0(s), y) f_2(s, y) \nu(dy) ds. \end{aligned} \quad (10.25)$$

Since  $M_G \in \mathcal{L}^1(\nu) \cap \mathcal{L}_{\text{exp}}^p \subset \mathcal{L}^2(\nu)$ , the last integral on the right side is finite by Hölder's inequality, and so under Condition 10.13, (10.25) has a unique solution  $\eta_q \in \mathcal{C}([0, T] : \mathbb{R}^d)$ . For  $\eta \in \mathcal{D}([0, T] : \mathbb{R}^d)$ , let

$$\bar{I}(\eta) \doteq \inf_{q=(f_1, f_2) \in \mathcal{L}^2 : \eta = \eta_q} \left[ \frac{1}{2} (\|f_1\|_{W,2}^2 + \|f_2\|_{N,2}^2) \right].$$

In particular,  $\bar{I}(\eta) = \infty$  for all  $\eta \in \mathcal{D}([0, T] : \mathbb{R}^d) \setminus \mathcal{C}([0, T] : \mathbb{R}^d)$ .

**Theorem 10.14** *Assume Conditions 10.1 and 10.13. Then  $\{Y^\varepsilon\}_{\varepsilon>0}$  satisfies the Laplace principle in  $\mathcal{D}([0, T] : \mathbb{R}^d)$  with scaling function  $\varkappa(\varepsilon)$  and rate function  $\bar{I}$ .*

The following theorem gives an alternative expression for the rate function. From part (d) of Condition 10.1 and part (a) of Condition 10.13, it follows that  $y \mapsto G_i(X^0(s), y)$  is in  $\mathcal{L}^2(\nu)$  for all  $s \in [0, T]$  and  $i = 1, \dots, d$ , where  $G = (G_1, \dots, G_d)^T$ . For  $i = 1, \dots, d$ , let  $e_i : \mathcal{X}_T \rightarrow \mathbb{R}$  be measurable functions such that for each  $s \in [0, T]$ ,  $\{e_i(s, \cdot)\}_{i=1}^d$  is an orthonormal collection in  $\mathcal{L}^2(\nu)$  and the linear span of the collection contains that of  $\{G_i(X^0(s), \cdot)\}_{i=1}^d$ . Define  $\bar{b}(x) \doteq \int_{\mathcal{X}} D_x G(x, y) \nu(dy)$ ,  $x \in \mathbb{R}^d$ , and define also  $A : [0, T] \rightarrow \mathbb{R}^{d \times d}$  by

$$A_{ij}(s) \doteq \langle G_i(X^0(s), \cdot), e_j(s, \cdot) \rangle_{\mathcal{L}^2(\nu)}, \quad i, j = 1, \dots, d, \quad s \in [0, T], \quad (10.26)$$

where  $\langle \cdot, \cdot \rangle_{\mathcal{L}^2(\nu)}$  is the inner product in  $\mathcal{L}^2(\nu)$ .

For  $\eta \in \mathcal{D}([0, T] : \mathbb{R}^d)$ , let

$$I(\eta) = \inf_{\tilde{q}=(\tilde{f}_1, \tilde{f}_2)} \left[ \frac{1}{2} (\|\tilde{f}_1\|_2^2 + \|\tilde{f}_2\|_2^2) \right],$$

where the infimum is taken over all  $\tilde{q} = (\tilde{f}_1, \tilde{f}_2)$ ,  $\tilde{f}_1, \tilde{f}_2 \in \mathcal{L}^2([0, T] : \mathbb{R}^d)$  such that for  $t \in [0, T]$ ,

$$\begin{aligned} \eta(t) &= \int_0^t [Db(X^0(s)) + \bar{b}(X^0(s))]\eta(s)ds + \int_0^t \sigma(X^0(s))\tilde{f}_1(s)ds \\ &\quad + \int_0^t A(s)\tilde{f}_2(s)ds. \end{aligned} \quad (10.27)$$

Here  $\|\cdot\|_2$  is the usual norm on  $\mathcal{L}^2([0, T] : \mathbb{R}^d)$ , and thus the same as  $\|\cdot\|_{W,2}$ . The proof of the following theorem is given in Sect. 10.3.3.

**Theorem 10.15** *Under the conditions of Theorem 10.14,  $I = \bar{I}$ .*

*Remark 10.16* Theorem 10.15 in particular says that the rate function for  $\{Y^\varepsilon\}$  is the same as that appearing in the large deviation principle with scaling function  $\varepsilon$  for the Gaussian process

$$dZ^\varepsilon(t) = B(t)Z^\varepsilon(t)dt + \sqrt{\varepsilon}A(t)dW_1(t) + \sqrt{\varepsilon}\sigma(X^0(t))dW_2(t), \quad Z^\varepsilon(0) = 0,$$

where  $W_1, W_2$  are independent standard  $d$ -dimensional Brownian motions and  $B(t) = Db(X^0(t)) + \bar{b}(X^0(t))$ .

### 10.3.1 Some Preparatory Results

Following our standard convention, the proof is given for  $T = 1$ , and thus  $\mathbb{U} = \mathcal{D}([0, 1] : \mathbb{R}^d)$ . From Theorem 10.2, it follows that there exists a measurable map  $\mathcal{G}^\varepsilon : \mathbb{V} \rightarrow \mathbb{U}$  such that  $X^\varepsilon = \mathcal{G}^\varepsilon(\sqrt{\varepsilon}W, \varepsilon N^{1/\varepsilon})$ . Using  $Y^\varepsilon = (X^\varepsilon - X^0)/a(\varepsilon)$ , there is measurable  $\mathcal{H}^\varepsilon$  such that  $Y^\varepsilon = \mathcal{H}^\varepsilon(\sqrt{\varepsilon}W, \varepsilon N^{1/\varepsilon})$ . Define  $\mathcal{H}^0 : \mathcal{L}^2 \rightarrow \mathbb{U}$  by  $\mathcal{H}^0(q) = \eta$  if  $\eta$  solves (10.25) for  $q = (f_1, f_2) \in \mathcal{L}^2$ . In order to prove Theorem 10.14, we will verify that Condition 9.8 holds with these choices of  $\mathcal{H}^\varepsilon$  and  $\mathcal{H}^0$ .

The following lemma verifies a continuity property of  $\mathcal{H}^0$ . Recall the space  $\hat{S}_n \doteq \{(f_1, f_2) \in \mathcal{L}^2 : \|f_1\|_{W,2}^2 + \|f_2\|_{N,2}^2 \leq n\}$  introduced above Condition 9.8. This is viewed as a subset of the Hilbert space  $\mathcal{L}^2$  defined there, and with respect to the topology of weak convergence in  $\mathcal{L}^2$  is a compact Polish space. Together with the continuity established in Lemma 10.17, the compactness of  $\hat{S}_n$  implies part (a) of Condition 9.8.

**Lemma 10.17** *Suppose Condition 10.1 holds and  $M_G \in \mathcal{L}^2(v)$ . Fix  $n \in (0, \infty)$  and let  $q^k, q \in \hat{S}_n, k \in \mathbb{N}$  be such that  $q^k \rightarrow q$ . Let  $\mathcal{H}^0(q) = \eta$ , where  $\eta$  solves (10.25). Then  $\mathcal{H}^0(q^k) \rightarrow \mathcal{H}^0(q)$ .*

*Proof* Note that from part (d) of Condition 10.1 and since  $M_G \in \mathcal{L}^2(v)$ , the map  $(s, y) \mapsto G(X^0(s), y)1_{[0,t]}(s)$  is in  $\mathcal{L}^2(v_1)$ . Let  $q^k = (f_1^k, f_2^k)$  and  $q = (f_1, f_2)$ . Since  $f_2^k \rightarrow f_2$ , we have for every  $t \in [0, 1]$  that

$$\int_{\mathcal{X}_t} f_2^k(s, y)G(X^0(s), y)v(dy)ds \rightarrow \int_{\mathcal{X}_t} f_2(s, y)G(X^0(s), y)v(dy)ds. \quad (10.28)$$

We argue that the convergence is in fact uniform in  $t$ . Note that for  $0 \leq s \leq t \leq 1$ ,

$$\begin{aligned} & \left\| \int_{[s,t] \times \mathcal{X}} f_2^k(u, y) G(X^0(u), y) v(dy) du \right\| \\ & \leq (1 + \|X^0\|_{\infty,1}) \int_{[s,t] \times \mathcal{X}} M_G(y) |f_2^k(u, y)| v(dy) du \\ & \leq (1 + \|X^0\|_{\infty,1}) |t - s|^{1/2} \sqrt{n} \|M_G\|_{\mathcal{L}^2(v)}, \end{aligned} \tag{10.29}$$

where  $\|\cdot\|_{\mathcal{L}^2(v)}$  denotes the norm in  $\mathcal{L}^2(v)$ . This implies equicontinuity, and hence the convergence in (10.28) is uniform in  $t \in [0, 1]$ .

Next, since  $f_1^k \rightarrow f_1$ , and since Condition 10.1 implies that  $\sigma(\cdot)$  is continuous, it follows that for every  $t \in [0, 1]$ ,

$$\int_0^t \sigma(X^0(s)) f_1^k(s) ds \rightarrow \int_0^t \sigma(X^0(s)) f_1(s) ds.$$

Once again an equicontinuity estimate similar to (10.29) shows that the convergence is uniform. The conclusion of the lemma now follows from Gronwall's inequality.  $\square$

In order to verify part (b) of Condition 9.8, we first prove some a priori estimates. Recall the spaces  $\mathcal{L}_{\text{exp}}^\rho$  introduced in (10.3) and  $S_{n,+}^{N,\varepsilon}$  and  $S_n^{N,\varepsilon}$  in (9.7). Here  $S_{n,+}^{N,\varepsilon}$  are controls for the Poisson noise with cost bounded by  $na^2(\varepsilon)$ , the scaling that is appropriate for an MDP, and  $S_n^{N,\varepsilon}$  are the centered and rescaled versions of elements of  $S_{n,+}^{N,\varepsilon}$ .

**Lemma 10.18** *Let  $h \in \mathcal{L}^1(v) \cap \mathcal{L}_{\text{exp}}^\rho$  for some  $\rho > 0$  and let  $I$  be a measurable subset of  $[0, 1]$ . Let  $n \in (0, \infty)$ . Then there exist maps  $\vartheta, \xi, \zeta$  from  $(0, \infty)$  to  $(0, \infty)$  such that  $\vartheta(u) \rightarrow 0$  as  $u \rightarrow \infty$  and  $\xi(u) \rightarrow 0$  as  $u \rightarrow 0$ , and for all  $\varepsilon, \beta \in (0, \infty)$ ,*

$$\sup_{f \in S_n^{N,\varepsilon}} \int_{I \times \mathcal{X}} h(y) |f(s, y)| 1_{\{|f| \geq \beta/a(\varepsilon)\}} v(dy) ds \leq \sqrt{a(\varepsilon)} \vartheta(\beta) + (1 + \lambda_1(I)) \xi(\varepsilon)$$

and

$$\sup_{f \in S_n^{N,\varepsilon}} \int_{I \times \mathcal{X}} h(y) |f(s, y)| v(dy) ds \leq \zeta(\beta) \lambda_1(I)^{1/2} + \sqrt{a(\varepsilon)} \vartheta(\beta) + (1 + \lambda_1(I)) \xi(\varepsilon),$$

where  $\lambda_1(I)$  denotes the Lebesgue measure of  $I$ .

*Proof* Let  $f \in S_n^{N,\varepsilon}$  and  $\beta \in (0, \infty)$ . Then

$$\begin{aligned} \int_{I \times \mathcal{X}} h(y)|f(s, y)|\nu(dy)ds &\leq \int_{I \times \mathcal{X}} h(y)|f(s, y)|1_{\{|f| \leq \beta/a(\varepsilon)\}}\nu(dy)ds \quad (10.30) \\ &\quad + \int_{I \times \mathcal{X}} h(y)|f(s, y)|1_{\{|f| \geq \beta/a(\varepsilon)\}}\nu(dy)ds. \end{aligned}$$

Recall that  $\mathcal{L}^1(\nu) \cap \mathcal{L}_{\text{exp}}^\rho \subset \mathcal{L}^p(\nu)$  for all  $p \geq 1$ . By the Cauchy-Schwarz inequality and part (c) of Lemma 9.7, we have

$$\begin{aligned} &\int_{I \times \mathcal{X}} h(y)|f(s, y)|1_{\{|f| \leq \beta/a(\varepsilon)\}}\nu(dy)ds \quad (10.31) \\ &\leq \left( \lambda_1(I) \|h\|_2^2 \int_{\mathcal{X}_1} f(s, y)^2 1_{\{|f| \leq \beta/a(\varepsilon)\}}\nu(dy)ds \right)^{1/2} \\ &\leq \|h\|_2 (n\kappa_2(\beta))^{1/2} \lambda_1(I)^{1/2}. \end{aligned}$$

We now consider the second term on the right side of (10.30). We decompose  $h$  as  $h1_{\{h \leq 1/a(\varepsilon)^{1/2}\}} + h1_{\{h > 1/a(\varepsilon)^{1/2}\}}$ . Then using part (a) of Lemma 9.7, we obtain

$$\begin{aligned} \int_{I \times \mathcal{X}} h(y)1_{\{h \leq 1/a(\varepsilon)^{1/2}\}}|f(s, y)|1_{\{|f| \geq \beta/a(\varepsilon)\}}\nu(dy) ds &\leq \frac{1}{\sqrt{a(\varepsilon)}} na(\varepsilon)\kappa_1(\beta) \\ &= n\sqrt{a(\varepsilon)}\kappa_1(\beta). \end{aligned}$$

Also, letting  $g = a(\varepsilon)f + 1$  and noting that the definition of  $S_n^{N,\varepsilon}$  implies  $g \geq 0$ , we have

$$\begin{aligned} &\int_{I \times \mathcal{X}} h(y)1_{\{h > 1/a(\varepsilon)^{1/2}\}}|f(s, y)|1_{\{|f| \geq \beta/a(\varepsilon)\}}\nu(dy) ds \\ &\leq \frac{\lambda_1(I)}{a(\varepsilon)} \int_{\mathcal{X}} h(y)1_{\{h > 1/a(\varepsilon)^{1/2}\}}\nu(dy) \\ &\quad + \frac{1}{a(\varepsilon)} \int_{I \times \mathcal{X}} h(y)1_{\{h > 1/a(\varepsilon)^{1/2}\}}g(s, y)\nu(dy) ds. \quad (10.32) \end{aligned}$$

The first term on the right side can be bounded by

$$\lambda_1(I)C_1(\varepsilon) \doteq \lambda_1(I) \int_{\mathcal{X}} h(y)^3 1_{\{h > 1/a(\varepsilon)^{1/2}\}}\nu(dy),$$

where  $h \in \mathcal{L}_{\text{exp}}^\rho \cap \mathcal{L}^1(\nu)$  implies  $C_1(\varepsilon) \rightarrow 0$  as  $\varepsilon \rightarrow 0$ . The second term on the right side of (10.32) can be bounded, using part (a) of Lemma 9.6 with  $a = \rho h(y)/2$ ,  $b = g(s, y)$ ,  $\sigma = 1$ , and also using that  $g \in S_{n,+}^{N,\varepsilon}$ , by

$$\begin{aligned} & \frac{2\lambda_1(I)}{\rho a(\varepsilon)} \int_{\mathcal{X}} e^{\rho h(y)/2} 1_{\{h > 1/a(\varepsilon)^{1/2}\}} v(dy) + \frac{2}{\rho a(\varepsilon)} na^2(\varepsilon) \\ & \leq \frac{2\lambda_1(I)}{\rho} \int_{\mathcal{X}} h(y)^2 e^{\rho h(y)/2} 1_{\{h > 1/a(\varepsilon)^{1/2}\}} v(dy) + \frac{2na(\varepsilon)}{\rho} \\ & = \lambda_1(I)C_2(\varepsilon) + \frac{2na(\varepsilon)}{\rho}, \end{aligned}$$

where  $C_2(\varepsilon)$  converges to 0 as  $\varepsilon \rightarrow 0$ . Thus the second term on the right side of (10.30) can be bounded by

$$\sqrt{a(\varepsilon)}\vartheta(\beta) + (1 + \lambda_1(I))\xi(\varepsilon),$$

where

$$\vartheta(\beta) = n\kappa_1(\beta), \quad \xi(\varepsilon) = C_1(\varepsilon) + C_2(\varepsilon) + \frac{2na(\varepsilon)}{\rho}.$$

This gives the first bound of the lemma. The second bound also follows with these choices of  $\xi$ ,  $\vartheta$ , and  $\zeta(\beta) = \|h\|_2(n\kappa_2(\beta))^{1/2}$  using (10.31).  $\square$

The following lemma is proved in a fashion similar to that of Lemma 10.7, and so only a sketch is given.

**Lemma 10.19** *Let  $\theta \in \mathcal{L}_{exp}^\rho$  for some  $\rho > 0$  and suppose that  $v(\{\theta > 1\}) < \infty$ . Then for every  $\delta > 0$  and  $n \in \mathbb{N}$ , there exists  $\tilde{c}(\delta, n) \in (0, \infty)$  such that for all measurable maps  $\tilde{\theta} : \mathcal{X} \rightarrow \mathbb{R}_+$  satisfying  $\tilde{\theta} \leq \theta$ , any measurable  $f : [0, 1] \rightarrow \mathbb{R}_+$ , and all  $0 \leq s \leq t \leq 1$ ,*

$$\begin{aligned} & \sup_{g \in S_{n,+}^{N,\varepsilon}} \int_{(s,t] \times \mathcal{X}} f(u)\tilde{\theta}(y)g(u, y)v(dy)du \\ & \leq \tilde{c}(\delta, n) \left( \int_{\mathcal{X}} \tilde{\theta}(y)v(dy) \right) \left( \int_s^t f(u)du \right) + (\delta + n\rho^{-1}a^2(\varepsilon))\|f\|_{\infty,1}. \end{aligned}$$

*Proof* Let  $f : [0, 1] \rightarrow \mathbb{R}_+$  and  $g \in S_{n,+}^{N,\varepsilon}$ . For  $m \in (0, \infty)$ , let  $T_i(m)$ ,  $i = 1, 2$ , be as in the proof of Lemma 10.7. Then using part (a) of Lemma 9.6 with  $a = \rho\theta(y)$ ,  $b = g(u, y)$ , and  $\sigma = 1$ , we can bound  $T_2(m)$  as

$$T_2(m) \leq \frac{\|f\|_{\infty,1}}{\rho} \left( \int_{\{\theta > m\}} e^{\rho\theta(y)} v(dy) + na^2(\varepsilon) \right).$$

Also, as in the proof of Lemma 10.7,

$$T_1(m) \leq \beta \left( \int_{\mathcal{X}} \tilde{\theta}(y)v(dy) \right) \left( \int_s^t f(u)du \right) + \bar{\kappa}_1(\beta)mna^2(\varepsilon)\|f\|_{\infty,1},$$

where  $\bar{\kappa}_1(\beta) \rightarrow 0$  as  $\beta \rightarrow \infty$ . The result now follows on recalling (10.9) and choosing first  $m$  sufficiently large and then  $\beta$  sufficiently large.  $\square$

Recall the map  $\mathcal{H}^\varepsilon$  introduced at the beginning of Sect. 10.3.1 and the definition of  $\mathcal{U}_{n,+}^\varepsilon$  in (9.8), and recall that by (9.7),  $\mathcal{U}_{n,+}^\varepsilon$  is the class of controls for both types of noise for which the cost scales proportionally with  $a^2(\varepsilon)$ . Since  $\mathcal{U}_{n,+}^\varepsilon$  is contained in  $\tilde{\mathcal{A}}_{b,\bar{n}}$  for some  $\bar{n} \in \mathbb{N}$ , it follows from Sect. 10.2.1 that for  $n \in \mathbb{N}$  and  $u = (\zeta, \varphi) \in \mathcal{U}_{n,+}^\varepsilon$ , the equation

$$d\bar{X}^\varepsilon(t) = [b(\bar{X}^\varepsilon(t)) + \sigma(\bar{X}^\varepsilon(t))\zeta(t)] dt + \sqrt{\varepsilon}\sigma(\bar{X}^\varepsilon(t))dW(t) + \varepsilon \int_{\mathcal{X}} G(\bar{X}^\varepsilon(t-), x)N^{\varphi/\varepsilon}(dt \times dx),$$

$\bar{X}^\varepsilon(0) = x_0$  has a unique solution.

Define  $\bar{Y}^\varepsilon = \mathcal{H}^\varepsilon(\sqrt{\varepsilon}W^{\zeta/\sqrt{\varepsilon}}, \varepsilon N^{\varphi/\varepsilon})$ , and note that using Girsanov’s theorem [Theorem D.3] as in the proof of Theorem 3.19 yields

$$\bar{Y}^\varepsilon = \frac{1}{a(\varepsilon)}(\bar{X}^\varepsilon - X^0). \tag{10.33}$$

The following moment bound on  $\bar{X}^\varepsilon$  follows along the lines of the proof of (10.19).

**Lemma 10.20** *Assume Conditions 10.1 and 10.13. For every  $n \in \mathbb{N}$ , there exists an  $\varepsilon_0 \in (0, 1)$  such that*

$$\sup_{\varepsilon \in (0, \varepsilon_0)} \sup_{u=(\zeta, \varphi) \in \mathcal{U}_{n,+}^\varepsilon} E \|\bar{X}^\varepsilon\|_{\infty,1} < \infty.$$

*Proof* Using the same argument as that used to establish (10.16), for all  $\varepsilon \in (0, 1)$ , we have

$$\|\bar{X}^\varepsilon\|_{\infty,1} \leq (1 + \|x_0\| + \tilde{\mathcal{H}}_1^\varepsilon) \exp \left\{ c_1(1 + \sqrt{n}a_1) + \int_{\mathcal{X}_1} M_G(y)\varphi(s, y)v(dy) ds \right\}, \tag{10.34}$$

where  $a_1 \doteq \sup_{\varepsilon \in (0,1)} a(\varepsilon)$  and

$$\begin{aligned} \tilde{\mathcal{H}}_1^\varepsilon &\doteq \sqrt{\varepsilon} \left\| \int_0^\cdot \sigma(\bar{X}^\varepsilon(s))dW(s) \right\|_{\infty,1} \\ &\quad + \varepsilon \left\| \int_{\mathcal{X}^\cdot} M_G(y)(1 + \|\bar{X}^\varepsilon(s-)\|)N_c^{\varphi/\varepsilon}(ds \times dy) \right\|_{\infty,1}. \end{aligned}$$

Using Lemma 10.19 with  $\delta = 1$ ,  $f \equiv 1$  and recalling that  $\varphi \in S_{n,+}^{N,\varepsilon}$  and the fact that  $M_G \in \mathcal{L}^1(v) \cap \mathcal{L}_{\text{exp}}$ , we obtain



$$\|\bar{X}^\varepsilon\|_{\infty,1} \leq c_2(1 + \|x_0\| + \tilde{\mathcal{R}}_1^\varepsilon), \tag{10.35}$$

where

$$c_2 \doteq \exp \left\{ c_1(1 + \sqrt{n}a_1) + \tilde{c}(1, n) \int_{\mathcal{X}} M_G(y)\nu(dy) + 1 + \frac{n}{\rho}a_1^2 \right\} < \infty.$$

We split the expected value of the second term in the definition of  $\tilde{\mathcal{R}}_1^\varepsilon$  as  $T_1^\varepsilon + T_2^\varepsilon$ , where  $T_i^\varepsilon$  are just as in (10.17)–(10.18) in the proof of the corresponding LDP, and then follow the same procedure as that below (10.17)–(10.18) to bound the two terms. In this case, however, we use Lemma 10.19 rather than Lemma 10.7, and find that for some  $c_3 \in (0, \infty)$ ,

$$T_1^\varepsilon \leq c_3\sqrt{\varepsilon}m(1 + E\|\bar{X}^\varepsilon\|_{\infty,1}) \left( \tilde{c}(1, n) \int_{\mathcal{X}} M_G(y)\nu(dy) + 1 + \frac{n}{\rho}a_1^2 \right)^{1/2},$$

and for every  $\delta > 0$  and  $\varepsilon \in (0, 1)$ ,

$$T_2^\varepsilon \leq 2(1 + E\|\bar{X}^\varepsilon\|_{\infty,1}) \left( \tilde{c}(\delta, n) \int_{\mathcal{X}} M_G(y)1_{\{M_G \geq m\}}\nu(dy) + \delta + \frac{n}{\rho}a^2(\varepsilon) \right).$$

Choosing first  $\delta$  sufficiently small, next  $m$  sufficiently large, and finally  $\varepsilon_0$  sufficiently small, we have for all  $\varepsilon \leq \varepsilon_0$  that

$$E\tilde{\mathcal{R}}_1^\varepsilon \leq \frac{1}{2c_2} (E\|\bar{X}^\varepsilon\|_{\infty,1} + 1).$$

The result now follows on using this estimate in (10.35). □

The following tightness property plays a key role in the proof of Theorem 10.14.

**Lemma 10.21** *Suppose Conditions 10.1 and 10.13 hold, and define  $\bar{Y}^\varepsilon$  by (10.33). For every  $n \in \mathbb{N}$ , there exists an  $\varepsilon_1 \in (0, 1)$  such that*

$$\{\|\bar{Y}^\varepsilon\|_{\infty,1}, u \in \mathcal{U}_{n,+}^\varepsilon, \varepsilon \in (0, \varepsilon_1)\}$$

*is a tight collection of  $\mathbb{R}_+$ -valued random variables.*

*Proof* Let  $u = (\zeta, \varphi) \in \mathcal{U}_{n,+}^\varepsilon$  and let  $\psi \doteq (\varphi - 1)/a(\varepsilon)$ . Then

$$\begin{aligned} \bar{X}^\varepsilon(t) - X^0(t) &= \int_0^t (b(\bar{X}^\varepsilon(s)) - b(X^0(s))) ds + \sqrt{\varepsilon} \int_0^t \sigma(\bar{X}^\varepsilon(s)) dW(s) \\ &\quad + \int_{\mathcal{X}_t} \varepsilon G(\bar{X}^\varepsilon(s-), y) N_c^{\varphi/\varepsilon}(ds \times dy) \\ &\quad + \int_{\mathcal{X}_t} (G(\bar{X}^\varepsilon(s), y) - G(X^0(s), y)) \varphi(s, y) \nu(dy) ds \end{aligned}$$

$$\begin{aligned}
 &+ \int_{\mathcal{X}_t} G(X^0(s), y)(\varphi(s, y) - 1)v(dy)ds \\
 &+ \int_0^t \sigma(\bar{X}^\varepsilon(s))\zeta(s)ds.
 \end{aligned}$$

Write  $\bar{Y}^\varepsilon = (\bar{X}^\varepsilon - X^0)/a(\varepsilon)$  as

$$\bar{Y}^\varepsilon = M^\varepsilon + A^\varepsilon + B^\varepsilon + \mathcal{E}^\varepsilon + C^\varepsilon, \tag{10.36}$$

where for  $t \in [0, 1]$ ,

$$\begin{aligned}
 M^\varepsilon(t) &\doteq \frac{\varepsilon}{a(\varepsilon)} \int_{\mathcal{X}_t} G(\bar{X}^\varepsilon(s-), y)N_c^{\varphi/\varepsilon}(ds \times dy) + \left(\frac{\varepsilon}{a(\varepsilon)}\right)^{1/2} \int_0^t \sigma(\bar{X}^\varepsilon(s))dW(s) \\
 A^\varepsilon(t) &\doteq \frac{1}{a(\varepsilon)} \int_0^t (b(\bar{X}^\varepsilon(s)) - b(X^0(s))) ds, \\
 B^\varepsilon(t) &\doteq \frac{1}{a(\varepsilon)} \int_{\mathcal{X}_t} (G(\bar{X}^\varepsilon(s), y) - G(X^0(s), y)) v(dy)ds, \\
 \mathcal{E}^\varepsilon(t) &\doteq \int_{\mathcal{X}_t} (G(\bar{X}^\varepsilon(s), y) - G(X^0(s), y)) \psi(s, y)v(dy)ds, \\
 C^\varepsilon(t) &\doteq \int_{\mathcal{X}_t} G(X^0(s), y)\psi(s, y)v(dy)ds + \frac{1}{a(\varepsilon)} \int_0^t \sigma(\bar{X}^\varepsilon(s))\zeta(s)ds.
 \end{aligned}$$

With  $\varepsilon_0$  as in Lemma 10.20, we have from the Burkholder–Davis–Gundy inequality (see Sect. D.1) that

$$\left\{ \left\| \int_0^\cdot \sigma(\bar{X}^\varepsilon(s))dW(s) \right\|_{\infty,1} \right\}_{\varepsilon \leq \varepsilon_0}$$

is tight. Also, as in the proof of Lemma 10.20, for some  $c_1 \in (0, \infty)$ , we have

$$\begin{aligned}
 E \left\| \int_{\mathcal{X}} G(\bar{X}^\varepsilon(s-), y)N_c^{\varphi/\varepsilon}(ds \times dy) \right\|_{\infty,1} &\leq (1 + E\|\bar{X}^\varepsilon\|_{\infty,1}) \\
 &\times \left( c_1\sqrt{\varepsilon}m + 2\tilde{c}(\delta, n) \int_{\mathcal{X}} M_G(y)1_{\{M_G \geq m\}}v(dy) + 2\delta + \frac{2n}{\rho}a^2(\varepsilon) \right)
 \end{aligned}$$

for every  $\delta > 0$ ,  $\varepsilon \in (0, \varepsilon_0)$ , and  $u \in \mathcal{U}_{n,+}^\varepsilon$ . Combining these two estimates, we see that  $\{\|M^\varepsilon\|_{\infty,1}\}_{\varepsilon \leq \varepsilon_0}$  is tight, and since  $\varepsilon/a(\varepsilon) \rightarrow 0$

$$\|M^\varepsilon\|_{\infty,1} \rightarrow 0 \text{ in probability} \tag{10.37}$$

as  $\varepsilon \rightarrow 0$ .

In the rest of the proof, we will show upper bounds of the form  $c, c \int_0^t \|\bar{Y}^\varepsilon\|_{\infty,s} ds$  or  $ca(\varepsilon)\|\bar{Y}^\varepsilon\|_{\infty,1}$  for each of the remaining terms in (10.36). By the Lipschitz condition

on  $G$  [part (c) of Condition 10.1 and part (a) of Condition 10.13], there is  $c_2 \in (0, \infty)$  such that for all  $t \in [0, 1]$ ,  $u \in \mathcal{U}_{n,+}^\varepsilon$ , we have

$$\begin{aligned} \|\mathcal{E}^\varepsilon\|_{\infty,t} &\leq a(\varepsilon) \int_{\mathcal{X}_t} L_G(y) \|\bar{Y}^\varepsilon(s)\| |\psi(s, y)| \nu(dy) ds \\ &\leq a(\varepsilon) \|\bar{Y}^\varepsilon\|_{\infty,t} \int_{\mathcal{X}_t} L_G(y) |\psi(s, y)| \nu(dy) ds \\ &\leq c_2 a(\varepsilon) \|\bar{Y}^\varepsilon\|_{\infty,t}, \end{aligned} \quad (10.38)$$

where the last inequality follows from the second bound in Lemma 10.18. Again using the Lipschitz condition on  $G$ , we have for all  $t \in [0, 1]$  that

$$\|B^\varepsilon\|_{\infty,t} \leq \|L_G\|_1 \int_0^t \|\bar{Y}^\varepsilon(s)\| ds.$$

Similarly, the Lipschitz condition on  $b$  gives

$$\|A^\varepsilon\|_{\infty,t} \leq L_b \int_0^t \|\bar{Y}^\varepsilon(s)\| ds.$$

Finally, we come to the term  $C^\varepsilon$ . Again using the second bound in Lemma 10.18, we have that for some  $c_3 \in (0, \infty)$  and all  $u \in \mathcal{U}_{n,+}^\varepsilon$ , with  $u = (\zeta, \varphi)$ ,  $\psi = (\varphi - 1)/a(\varepsilon)$ ,

$$\left\| \int_{\mathcal{X}} G(X^0(s), y) \psi(s, y) \nu(dy) ds \right\|_{\infty,1} \leq c_3.$$

Since  $(\zeta, \varphi) \in \mathcal{U}_{n,+}^\varepsilon$ , the bound  $\int_0^1 \|\zeta(s)\|^2 ds \leq na^2(\varepsilon)$  applies. Thus for  $t \in [0, 1]$ ,

$$\frac{1}{a(\varepsilon)} \int_0^t \sigma(\bar{X}^\varepsilon(s)) \zeta(s) ds = \frac{1}{a(\varepsilon)} \int_0^t \sigma(X^0(s)) \zeta(s) ds + \frac{1}{a(\varepsilon)} \mathcal{R}_1^\varepsilon(t),$$

where

$$\|\mathcal{R}_1^\varepsilon\|_{\infty,1} \leq a(\varepsilon) L_\sigma \|\bar{Y}^\varepsilon\|_{\infty,1} \sqrt{n}, \quad (10.39)$$

and

$$\left\| \frac{1}{a(\varepsilon)} \int_0^t \sigma(X^0(s)) \zeta(s) ds \right\|_{\infty,1} \leq (\|X^0\|_{\infty,1} L_\sigma + \|\sigma(0)\|) \sqrt{n}.$$

Bringing terms in (10.36) of the form  $ca(\varepsilon) \|\bar{Y}^\varepsilon\|_{\infty,1}$  to the left side and renormalizing for a coefficient of one, we have, for some  $c_4 \in (0, \infty)$ ,  $\tilde{\varepsilon}_0 \in (0, \varepsilon_0)$ , and all  $u \in \mathcal{U}_{n,+}^\varepsilon$ ,  $t \in [0, 1]$ , and  $\varepsilon \leq \tilde{\varepsilon}_0$ , that

$$\|\bar{Y}^\varepsilon\|_{\infty,t} \leq c_4 \left( 1 + \int_0^t \|\bar{Y}^\varepsilon\|_{\infty,s} ds \right) + Z^\varepsilon,$$

where  $\{Z^\varepsilon\}_{\varepsilon \leq \varepsilon_0}$  is tight. The result now follows by an application of Gronwall's inequality.  $\square$

The next two lemmas will be needed in the weak convergence arguments of Lemma 10.24. The first is an immediate consequence of Lemma 10.18.

**Lemma 10.22** *Let  $h \in \mathcal{L}^1(\nu) \cap \mathcal{L}_{exp}^\rho$  for some  $\rho > 0$ . Then for all  $n \in \mathbb{N}$ , and  $\beta \in (0, \infty)$ ,*

$$\sup_{f \in S_n^{N,\varepsilon}} \int_{\mathcal{X}_1} h(y) |f(s, y)| 1_{\{|f| \geq \beta/a(\varepsilon)\}} \nu(dy) ds \rightarrow 0 \text{ as } \varepsilon \rightarrow 0.$$

For  $n \in (0, \infty)$ , let  $\hat{S}_n^N \doteq \{f \in \mathcal{L}^2(\nu_1) : \|f\|_{N,2}^2 \leq n\}$ .

**Lemma 10.23** *Let  $n \in \mathbb{N}$ ,  $\varepsilon > 0$ , and  $f^\varepsilon \in S_n^{N,\varepsilon}$ . Let  $\eta : \mathcal{X}_1 \rightarrow \mathbb{R}^d$  be a measurable function such that*

$$|\eta(s, y)| \leq h(y) \text{ for } y \in \mathcal{X}, s \in [0, 1],$$

where  $h \in \mathcal{L}^1(\nu) \cap \mathcal{L}_{exp}^\rho$  for some  $\rho > 0$ . Suppose there is  $\beta \in (0, 1]$  such that  $f^\varepsilon 1_{\{|f^\varepsilon| \leq \beta/a(\varepsilon)\}}$  converges in  $\hat{S}_{nk_2(1)}^N$  to  $f$ . Then for all  $t \in [0, 1]$ ,

$$\int_{\mathcal{X}_t} \eta(s, y) f^\varepsilon(s, y) \nu_1(ds \times dy) \rightarrow \int_{\mathcal{X}_t} \eta(s, y) f(s, y) \nu_1(ds \times dy).$$

*Proof* It follows from Lemma 10.22 that

$$\int_{\mathcal{X}_1} |\eta(s, y) f^\varepsilon(s, y)| 1_{\{|f^\varepsilon| \geq \beta/a(\varepsilon)\}} \nu_1(ds \times dy) \rightarrow 0 \text{ as } \varepsilon \rightarrow 0.$$

Also, since  $\eta 1_{[0,t]} \in \mathcal{L}^2(\nu_1)$  for all  $t \in [0, 1]$  and  $f^\varepsilon 1_{\{|f^\varepsilon| \leq \beta/a(\varepsilon)\}} \rightarrow f$ , we have

$$\int_{\mathcal{X}_t} \eta(s, y) f^\varepsilon(s, y) 1_{\{|f^\varepsilon| \leq \beta/a(\varepsilon)\}} \nu_1(ds \times dy) \rightarrow \int_{\mathcal{X}_t} \eta(s, y) f(s, y) \nu_1(ds \times dy).$$

The result follows on combining the last two displays.  $\square$

### 10.3.2 Proof of the Moderate Deviation Principle

The following is the key result needed in the proof of Theorem 10.14. It gives tightness of the joint distribution of controls and controlled processes, and indicates how limits of these two quantities are related. Recall  $\hat{S}_n^N \doteq \{f \in \mathcal{L}^2(\nu_1) : \|f\|_{N,2}^2 \leq n\}$ .

**Lemma 10.24** *Suppose Conditions 10.1 and 10.13 hold. Let  $n \in \mathbb{N}$ ,  $\varepsilon > 0$ , and  $u^\varepsilon = (\zeta^\varepsilon, \varphi^\varepsilon) \in \mathcal{U}_{n,+}^\varepsilon$ . Let  $\psi^\varepsilon \doteq (\varphi^\varepsilon - 1)/a(\varepsilon)$ , suppose  $\beta \in (0, 1]$ , and set  $\bar{Y}^\varepsilon \doteq \mathcal{H}^\varepsilon(\sqrt{\varepsilon}W^{\zeta^\varepsilon}/\sqrt{\varepsilon}, \varepsilon N^{\varphi^\varepsilon/\varepsilon})$ . Then  $\{(\bar{Y}^\varepsilon, \psi^\varepsilon 1_{\{|\psi^\varepsilon| \leq \beta/a(\varepsilon)\}}, \zeta^\varepsilon/a(\varepsilon))\}_{\varepsilon>0}$  is tight in*

$$\mathcal{D}([0, 1] : \mathbb{R}^d) \times \hat{S}_{n(\kappa_2(1)+1)},$$

and any limit point  $(\bar{Y}, \psi, \zeta)$  satisfies (10.25) a.s., with  $\eta$  replaced by  $\bar{Y}$  and  $(f_1, f_2)$  replaced by  $(\psi, \zeta)$ .

*Proof* We use the notation from the proof of Lemma 10.21 but replace  $(\zeta, \varphi)$  throughout by  $(\zeta^\varepsilon, \varphi^\varepsilon)$ . Assume without loss of generality that  $\varepsilon \leq \varepsilon_0$ . From (10.37) we have that  $\|M^\varepsilon\|_{\infty,1} \rightarrow 0$  in probability as  $\varepsilon \rightarrow 0$ . Also, since from Lemma 10.21  $\{\|\bar{Y}^\varepsilon\|_{\infty,1}\}_{\varepsilon \leq \varepsilon_0}$  is tight, (10.38) implies that  $\|\mathcal{E}^\varepsilon\|_{\infty,1} \rightarrow 0$  in probability.

Noting that  $\bar{X}^\varepsilon(t) = X^0(t) + a(\varepsilon)\bar{Y}^\varepsilon(t)$ , we have by Taylor’s formula,

$$G(\bar{X}^\varepsilon(s), y) - G(X^0(s), y) = a(\varepsilon)D_x G(X^0(s), y)\bar{Y}^\varepsilon(t) + R^\varepsilon(s, y),$$

where

$$\|R^\varepsilon(s, y)\| \leq L_{DG}(y)a^2(\varepsilon)\|\bar{Y}^\varepsilon(s)\|^2.$$

Hence

$$B^\varepsilon(t) = \int_{\mathcal{X}_t} D_x G(X^0(s), y)\bar{Y}^\varepsilon(s)v(dy) ds + T_1^\varepsilon(t),$$

where

$$\|T_1^\varepsilon\|_{\infty,1} \leq \|L_{DG}\|_1 a(\varepsilon) \int_0^1 \|\bar{Y}^\varepsilon(s)\|^2 ds.$$

Thus using Lemma 10.21 again, we have that  $\|T_1^\varepsilon\|_{\infty,1} \rightarrow 0$  in probability. Similarly,

$$A^\varepsilon(t) = \int_0^t Db(X^0(s))\bar{Y}^\varepsilon(s)ds + T_2^\varepsilon(t),$$

where  $\|T_2^\varepsilon\|_{\infty,1} \rightarrow 0$  in probability. Also, from (10.39)  $\|\mathcal{R}_1^\varepsilon\|_{\infty,1} \rightarrow 0$  in probability.

Putting these estimates together, we have from (10.36) that

$$\begin{aligned} \bar{Y}^\varepsilon(t) &= T_3^\varepsilon(t) + \int_0^t Db(X^0(s))\bar{Y}^\varepsilon(s)ds + \frac{1}{a(\varepsilon)} \int_0^t \sigma(X^0(s))\zeta^\varepsilon(s)ds \quad (10.40) \\ &\quad + \int_{\mathcal{X}_t} D_x G(X^0(s), y)\bar{Y}^\varepsilon(s)v(dy)ds + \int_{\mathcal{X}_t} G(X^0(s), y)\psi^\varepsilon(s, y)v(dy)ds, \end{aligned}$$

where

$$T_3^\varepsilon \doteq M^\varepsilon + \mathcal{E}^\varepsilon + T_1^\varepsilon + T_2^\varepsilon + \mathcal{R}_1^\varepsilon \Rightarrow 0.$$

We now prove tightness of

$$\begin{aligned}\tilde{A}^\varepsilon(\cdot) &\doteq \int_0^\cdot Db(X^0(s))\bar{Y}^\varepsilon(s)ds, \quad \tilde{B}^\varepsilon(\cdot) \doteq \int_{\mathcal{X}} D_x G(X^0(s), y)\bar{Y}^\varepsilon(s)v(dy)ds, \\ \tilde{C}^\varepsilon(\cdot) &\doteq \int_{\mathcal{X}} G(X^0(s), y)\psi^\varepsilon(s, y)v(dy)ds, \quad \tilde{D}^\varepsilon(\cdot) \doteq \frac{1}{a(\varepsilon)} \int_0^\cdot \sigma(X^0(s))\zeta^\varepsilon(s)ds.\end{aligned}$$

Applying Lemma 10.18 with  $h = M_G$ , for every  $\beta \in (0, 1]$  and  $\delta \in (0, 1)$ , we have

$$\begin{aligned}\|\tilde{C}^\varepsilon(t + \delta) - \tilde{C}^\varepsilon(t)\| &= \int_{[t, t+\delta] \times \mathcal{X}} \|G(X^0(s), y)\| |\psi^\varepsilon(s, y)| v(dy) ds \\ &\leq (1 + \|X^0\|_{\infty, 1}) \int_{[t, t+\delta] \times \mathcal{X}} M_G(y) |\psi^\varepsilon(s, y)| v(dy) ds \\ &\leq (1 + \|X^0\|_{\infty, 1}) (\zeta(\beta)\delta^{1/2} + \sqrt{a(\varepsilon)}\vartheta(\beta) + 2\xi(\varepsilon)).\end{aligned}$$

Tightness of  $\{\tilde{C}^\varepsilon\}_{\varepsilon>0}$  in  $\mathcal{C}([0, 1] : \mathbb{R}^d)$  is now immediate from the properties of  $\vartheta$  and  $\xi$ .

Next we argue the tightness of  $\tilde{B}^\varepsilon$ . For  $0 \leq t \leq t + \delta \leq 1$ , we have

$$\begin{aligned}\|\tilde{B}^\varepsilon(t + \delta) - \tilde{B}^\varepsilon(t)\| &\leq \int_{[t, t+\delta] \times \mathcal{X}} \|D_x G(X^0(s), y)\bar{Y}^\varepsilon(s)\| v(dy) ds \\ &\leq \left( \sup_{\|x\| \leq \|X^0\|_{\infty, 1}} \int_{\mathcal{X}} \|D_x G(x, y)\| v(dy) \right) \int_{[t, t+\delta]} \|\bar{Y}^\varepsilon(s)\| ds \\ &\leq c_1 \|\bar{Y}^\varepsilon\|_{\infty, 1} \delta,\end{aligned}$$

where  $c_1 \doteq \sup_{\|x\| \leq \|X^0\|_{\infty, 1}} \int_{\mathcal{X}} \|D_x G(x, y)\| v(dy)$  is finite by part (b) of Condition 10.13. Tightness of  $\{\tilde{B}^\varepsilon\}_{\varepsilon>0}$  in  $\mathcal{C}([0, 1] : \mathbb{R}^d)$  now follows as a consequence of Lemma 10.21. Similarly, it can be seen that  $\tilde{A}^\varepsilon$  is tight in  $\mathcal{C}([0, 1] : \mathbb{R}^d)$ . Finally, since  $\zeta^\varepsilon \in S_{na^2(\varepsilon)}^W$  implies the bound  $na^2(\varepsilon)$  for  $\int_0^1 \|\zeta^\varepsilon\|^2 ds$ , it follows that for  $0 \leq t \leq t + \delta \leq 1$ , we have

$$\begin{aligned}\|\tilde{D}^\varepsilon(t + \delta) - \tilde{D}^\varepsilon(t)\| &\leq \frac{1}{a(\varepsilon)} \int_t^{t+\delta} \|\sigma(X^0(s))\zeta^\varepsilon(s)\| ds \\ &\leq \sqrt{\delta} (\|X^0\|_{\infty, 1} L_\sigma + \|\sigma(0)\|) \sqrt{n}.\end{aligned}$$

Tightness of  $\{\tilde{D}^\varepsilon\}_{\varepsilon>0}$  in  $\mathcal{C}([0, 1] : \mathbb{R}^d)$  is now immediate. Since each of these terms is tight,  $\{\bar{Y}^\varepsilon\}_{\varepsilon>0}$  is tight in  $\mathcal{D}([0, 1] : \mathbb{R}^d)$ . Also, from part (c) of Lemma 9.7,

$$\left( \psi^\varepsilon 1_{\{|\psi^\varepsilon| \leq \beta/a(\varepsilon)\}}, \frac{1}{a(\varepsilon)} \zeta^\varepsilon \right)$$

takes values in the compact space  $\hat{S}_{n(\kappa_2(1)+1)}$  for all  $\varepsilon > 0$  and is therefore automatically tight. This completes the proof of the first part of the lemma.

Suppose now that

$$\left( \bar{Y}^\varepsilon, \psi^\varepsilon \mathbf{1}_{\{|\psi^\varepsilon| \leq \beta/a(\varepsilon)\}}, \frac{1}{a(\varepsilon)} \zeta^\varepsilon \right)$$

converges in distribution along a subsequence to  $(\bar{Y}, \psi, \zeta)$ . From Lemma 10.23 and the tightness of  $\tilde{C}^\varepsilon, \tilde{D}^\varepsilon$  established above, we have that

$$\left( \bar{Y}^\varepsilon, \int_{\mathcal{X}} G(X^0(s), y) \psi^\varepsilon(s, y) \nu(dy) ds, \frac{1}{a(\varepsilon)} \int_0^\cdot \sigma(X^0(s)) \zeta^\varepsilon(s) ds \right)$$

converges in distribution, in  $\mathcal{D}([0, 1] : \mathbb{R}^{3d})$ , to

$$\left( \bar{Y}, \int_{\mathcal{X}} G(X^0(s), y) \psi(s, y) \nu(dy) ds, \int_0^\cdot \sigma(X^0(s)) \zeta(s) ds \right).$$

The result now follows on using this convergence in (10.40) and recalling that  $T_3^\varepsilon \Rightarrow 0$ .  $\square$

We now complete the proof of the moderate deviation principle.

*Proof (of Theorem 10.14)* It suffices to show that Condition 9.8 holds with  $\mathcal{H}^\varepsilon$  and  $\mathcal{H}^0$  defined as at the beginning of Sect. 10.3.1. Part (a) of the condition was verified in Lemma 10.17. Consider now part (b). Fix  $n \in \mathbb{N}$  and  $\beta \in (0, 1]$ . Let  $(\zeta^\varepsilon, \varphi^\varepsilon) \in \mathcal{U}_{n,+}^\varepsilon$  and  $\psi^\varepsilon = (\varphi^\varepsilon - 1)/a(\varepsilon)$ . Suppose that for some  $\beta \in (0, 1]$ ,  $(\psi^\varepsilon \mathbf{1}_{\{|\psi^\varepsilon| \leq \beta/a(\varepsilon)\}}, \zeta^\varepsilon/a(\varepsilon)) \Rightarrow (\psi, \zeta)$ . To complete the proof, we need to show that

$$\mathcal{H}^\varepsilon \left( \sqrt{\varepsilon} W^{\zeta^\varepsilon/\sqrt{\varepsilon}}, \varepsilon N^{\varphi^\varepsilon/\varepsilon} \right) \Rightarrow \mathcal{H}^0(\zeta, \psi). \tag{10.41}$$

Recall that the left side of (10.41) equals  $\bar{Y}^\varepsilon$  defined in (10.33). From Lemma 10.24,  $\{(\bar{Y}^\varepsilon, \psi^\varepsilon \mathbf{1}_{\{|\psi^\varepsilon| \leq \beta/a(\varepsilon)\}}, \zeta^\varepsilon/a(\varepsilon))\}$  is tight in  $\mathcal{D}([0, 1] : \mathbb{R}^d) \times \hat{S}_{n(\kappa_2(1)+1)}$ , and every limit point  $(\bar{Y}, \psi, \zeta)$  satisfies (10.25) a.s., with  $\eta$  replaced by  $\bar{Y}$  and  $(f_1, f_2)$  with  $(\bar{\psi}, \zeta)$ . Since (10.25) has a unique solution,  $\mathcal{H}^0(f_1, f_2)$  for every  $f = (f_1, f_2) \in \mathcal{L}^2$  [recall  $\mathcal{L}^2 \doteq \mathcal{L}^2([0, 1] : \mathbb{R}^d) \times \mathcal{L}^2(\nu_T)$ ],  $(\zeta, \psi)$  has the same law as  $(\bar{\zeta}, \bar{\psi})$ , and every limit point of  $\bar{Y}^\varepsilon$  must have the same distribution as  $\mathcal{H}^0(\zeta, \psi)$ . The result follows.  $\square$

### 10.3.3 Equivalence of Two Rate Functions

In this section we present the proof of Theorem 10.15. To simplify notation, suppose without loss that  $T = 1$ . Fix  $\eta \in \mathcal{C}([0, 1] : \mathbb{R}^d)$  and  $\delta > 0$ . Let  $\tilde{f} = (\tilde{f}_1, \tilde{f}_2)$ , where  $\tilde{f}_i \in \mathcal{L}^2([0, 1] : \mathbb{R}^d)$ ,  $i = 1, 2$ , be such that

$$\frac{1}{2} \int_0^1 \left( \|\tilde{f}_1(s)\|^2 + \|\tilde{f}_2(s)\|^2 \right) ds \leq I(\eta) + \delta$$

and  $(\eta, \tilde{f})$  satisfy (10.27). Define  $f_2 : \mathcal{X}_1 \rightarrow \mathbb{R}$  by

$$f_2(s, y) \doteq \sum_{i=1}^d \tilde{f}_{2,i}(s) e_i(s, y), \quad (s, y) \in \mathcal{X}_1, \quad (10.42)$$

where the  $e_i$  were introduced just before the statement of Theorem 10.15. From the orthonormality of  $e_i(s, \cdot)$ , it follows that

$$\frac{1}{2} \int_{\mathcal{X}_1} |f_2(s, y)|^2 \nu_1(ds \times dy) = \frac{1}{2} \int_0^1 \|\tilde{f}_2(s)\|^2 ds. \quad (10.43)$$

Also, with  $A(s)$  defined as in (10.26), we have

$$\begin{aligned} [A(s)\tilde{f}_2(s)]_i &= \sum_{j=1}^d \langle G_i(X^0(s), \cdot), e_j(s, \cdot) \rangle_{\mathcal{L}^2(\nu)} \tilde{f}_{2,j}(s) \\ &= \left\langle G_i(X^0(s), \cdot), \sum_{j=1}^d e_j(s, \cdot) \tilde{f}_{2,j}(s) \right\rangle_{\mathcal{L}^2(\nu)} \\ &= \langle G_i(X^0(s), \cdot), f_2(s, \cdot) \rangle_{\mathcal{L}^2(\nu)}, \end{aligned}$$

so that  $A(s)\tilde{f}_2(s) = \int_{\mathcal{X}} f_2(s, y) G(X^0(s), y) \nu(dy)$ . Consequently,  $\eta$  satisfies (10.25) with  $f_2$  as in (10.42) and  $f_1 = \tilde{f}_1$ . Combining this with (10.43), we have

$$\begin{aligned} \bar{I}(\eta) &\leq \frac{1}{2} \int_0^1 \|\tilde{f}_1(s)\|^2 ds + \frac{1}{2} \int_{\mathcal{X}_1} |f_2(s, y)|^2 \nu_2(ds \times dy) \\ &= \frac{1}{2} \int_0^1 \left( \|\tilde{f}_1(s)\|^2 + \|\tilde{f}_2(s)\|^2 \right) ds \\ &\leq I(\eta) + \delta. \end{aligned}$$

Since  $\delta > 0$  is arbitrary, we have  $\bar{I}(\eta) \leq I(\eta)$ .

Conversely, suppose  $\delta > 0$  and  $q = (f_1, f_2) \in \mathcal{L}^2$  is such that

$$\frac{1}{2} \int_{\mathcal{X}_1} |f_2(s, y)|^2 \nu_1(ds \times dy) + \frac{1}{2} \int_0^1 \|f_1(s)\|^2 ds \leq \bar{I}(\eta) + \delta$$

and (10.25) holds. For  $i = 1, \dots, d$ , define  $\tilde{f}_{2,i} : [0, 1] \rightarrow \mathbb{R}$  by

$$\tilde{f}_{2,i}(s) = \langle f_2(s, \cdot), e_i(s, \cdot) \rangle_{\mathcal{L}^2(\nu)}.$$



For  $s \in [0, 1]$ , let  $\{e_j(s, \cdot)\}_{j=d+1}^\infty$  be defined in such a manner that  $\{e_j(s, \cdot)\}_{j=1}^\infty$  is a complete orthonormal system in  $\mathcal{L}^2(\nu)$ . Then for every  $s \in [0, 1]$ ,  $i = 1, \dots, d$ ,

$$\begin{aligned} [A(s)\tilde{f}_2(s)]_i &= \sum_{j=1}^d \langle G_i(X^0(s), \cdot), e_j(s, \cdot) \rangle_{\mathcal{L}^2(\nu)} \langle f_2(s, \cdot), e_j(\cdot, s) \rangle_{\mathcal{L}^2(\nu)} \\ &= \sum_{j=1}^\infty \langle G_i(X^0(s), \cdot), e_j(\cdot, s) \rangle_{\mathcal{L}^2(\nu)} \langle f_2(s, \cdot), e_j(s, \cdot) \rangle_{\mathcal{L}^2(\nu)} \\ &= \langle G_i(X^0(s), \cdot), f_2(s, \cdot) \rangle_{\mathcal{L}^2(\nu)}, \end{aligned}$$

where the second equality follows on observing that  $G_i(X^0(s), \cdot)$  is in the linear span of  $\{e_j(s, \cdot)\}_{j=1}^d$  for  $i = 1, \dots, d$ . Thus  $A(s)\tilde{f}_2(s) = \int_{\mathcal{X}} f_2(s, y)G(X^0(s), y)\nu(dy)$ , and therefore  $(\eta, \tilde{f})$  satisfy (10.27) with  $\tilde{f}_2$  defined as above and  $\tilde{f}_1 = f_1$ . Note that with  $\tilde{f}_2 = (\tilde{f}_{2,1}, \dots, \tilde{f}_{2,d})$ ,

$$\begin{aligned} \frac{1}{2} \int_0^1 \|\tilde{f}_2(s)\|^2 ds &= \frac{1}{2} \int_0^1 \sum_{j=1}^d \langle f_2(s, \cdot), e_j(s, \cdot) \rangle_{\mathcal{L}^2(\nu)}^2 ds \\ &\leq \frac{1}{2} \int_0^1 \int_{\mathcal{X}} f_2^2(s, y)\nu(dy) ds. \end{aligned}$$

Thus

$$\begin{aligned} I(\eta) &\leq \frac{1}{2} \int_0^1 \left( \|\tilde{f}_1(s)\|^2 + \|\tilde{f}_2(s)\|^2 \right) ds \\ &\leq \frac{1}{2} \int_0^1 \|f_1(s)\|^2 ds + \frac{1}{2} \int_0^1 \int_{\mathcal{X}} f_2^2(s, y)\nu(dy) ds \\ &\leq \bar{I}(\eta) + \delta. \end{aligned}$$

Since  $\delta > 0$  is arbitrary,  $I(\eta) \leq \bar{I}(\eta)$ , which completes the proof. □

### 10.4 Notes

The first results for a general class of continuous time small noise Markov processes appear to be those of Wentzell [245–248]. The class covered includes both Gaussian and Poisson driving noises, and the proof uses approximation by discrete time processes.

Large deviation principles for small noise infinite dimensional stochastic differential equations driven by PRM are considered in [38]. Although this paper considers a considerably more complex setting than the one studied in the current chapter, it

assumed a somewhat stronger condition on the coefficient function that plays the role of  $G$  in this chapter. Specifically, it required the functions  $M_G$  and  $L_G$  to satisfy a more stringent integrability condition than the one used here (Condition 10.3). Some of the results used to deal with these weaker integrability conditions are from [47]. Another distinction between this chapter and [38] is that here we consider systems driven by both Gaussian and Poisson noise, and consider both large and moderate deviations. A moderate deviation principle of the form given in Sect. 10.3, applicable to both finite and infinite dimensional SDE, was presented in [41] for the setting in which the driving noise is only Poisson. It is worth noting that most of the technical details in this chapter arise from the treatment of the PRM term.

There is an important distinction between the types of processes one can represent using PRMs and their large deviation analysis. For one class, which includes the models considered in this chapter as well as the example in Sect. 3.3, different points in the point space of the underlying PRM are used to model different types of jumps in the solution to the SDE. The conditions placed on the coefficient that modulates the impact of the noise on the state ( $G$  in this chapter) tend to be continuity-type conditions, analogous to those one places on the diffusion coefficient of an SDE driven by Brownian motion, though stated in terms of integration over the space  $\mathcal{X}$ . These continuity properties are used to establish uniqueness of the map that takes the controls into the state for the limit dynamics, under the assumption that the cost of the controls is bounded. With the second class, there are only finitely many different types of jumps of the state, and the role of  $G$  is simply to “thin” the PRM to produce state-dependent jump rates. Examples of this type include the process models of Chap. 13 as well as those in [23, 42]. Owing to its role in thinning,  $G$  is typically not continuous, and one does not expect uniqueness of the limiting deterministic map that takes controls to the state process. However, as we will see, it is in fact sufficient to prove the following restricted uniqueness: given a state trajectory for which there is a corresponding control with finite cost, find a control that produces the same state trajectory with nearly the same cost and for which there is uniqueness. These points are illustrated in the analysis carried out in Chap. 13.