Annales de l'Institut Henri Poincaré - Probabilités et Statistiques
2011, Vol. 47, No. 3, 725–747
DOI: 10.1214/10-AIHP382

ANNALES DE L'INSTITUT HENRI POINCARÉ
PROBABILITÉS ET STATISTIQUES
www.imstat.org/aihp

Variational representations for continuous time processes

Amarjit Budhiraja[a,1], Paul Dupuis[b,2] and Vasileios Maroulas[c,3]

[a]*Department of Statistics and Operations Research, University of North Carolina at Chapel Hill, Chapel Hill, NC 27599, USA.*
E-mail: budhiraj@email.unc.edu
[b]*Lefschetz Center for Dynamical Systems, Division of Applied Mathematics, Brown University, Providence, RI 02912, USA.*
E-mail: dupuis@dam.brown.edu
[c]*Institute for Mathematics and its Applications, University of Minnesota, Minneapolis, MN 55455, USA.*
E-mail: maroulas@math.utk.edu

Received 25 June 2009; revised 5 July 2010; accepted 8 July 2010

Abstract. A variational formula for positive functionals of a Poisson random measure and Brownian motion is proved. The formula is based on the relative entropy representation for exponential integrals, and can be used to prove large deviation type estimates. A general large deviation result is proved, and illustrated with an example.

Résumé. Une formule variationnelle pour des fonctionnelles positives d'une mesure de Poisson aléatoire et d'un mouvement brownien est démontrée. Cette formule provient de la représentation des intégrales exponentielles par l'entropie relative, et peut être utilisée pour obtenir des estimées de grandes déviations. Un résultat de grandes déviations général est démontré.

MSC: Primary 60F10; secondary 60G51; 60H15

Keywords: Variational representations; Poisson random measure; Infinite-dimensional Brownian motion; Large deviations; Jump-diffusions

1. Introduction

In this paper we prove a variational representation for positive measurable functionals of a Poisson random measure and an infinite-dimensional Brownian motion. These processes provide the driving noises for a wide range of important process models in continuous time, and thus we also obtain variational representations for these processes when a strong solution exists. The representations have a number of uses, the most important being to prove large deviation estimates.

The theory of large deviations is by now well understood in many settings, but there remain some situations where the topic is not as well developed. These are often settings where technical issues challenge standard approaches, and the problem of finding nearly optimal or even reasonably weak sufficient conditions is hindered as much by technique of proof as any other issue.

Variational representations of the sort developed in this paper have been shown to be particularly useful, when combined with weak convergence methods, for analyzing such systems. For example, Brownian motion representations have been used by [1, 6–8, 9, 10, 17–21, 23–25, 26, 29] in the large deviation analysis of solutions to SPDEs

[1]Supported in part by the National Science Foundation (DMS-1004418), the Army Research Office (W911NF-0-1-0080, W911NF-10-1-0158) and the US-Israel Binational Science Foundation (2008466).
[2]Supported in part by the National Science Foundation (DMS-0706003), the Army Research Office (W911NF-09-1-0155) and the Air Force Office of Scientific Research (FA9550-07-1-0544, FA9550-09-1-0378).
[3]Supported in part by the IMA Industrial Postdoctoral Fellowship.

in the small noise limit, and recently in [5] for interacting particle limits. Other examples occur when there is little regularity associated with the process dynamics, such as non-Lipschitz or even discontinuous coefficients [3,19].

The usefulness of the representations is in part due to the fact that they avoid certain discretization and/or approximation arguments, which can be cumbersome for complex systems. Another reason is that exponential tightness, a property that is often required by other approaches and which often leads to artificial conditions, is replaced by ordinary tightness for controlled processes with uniformly bounded control costs. (Although exponential tightness can a posteriori be obtained as consequence of the large deviation principle (LDP) and properties of the rate function.) What is required for the weak convergence approach, beyond the variational representations, is that basic qualitative properties (existence, uniqueness and law of large number limits) can be demonstrated for certain controlled versions of the original process.

Previous work on variational representations has focused on either discrete time processes [11], functionals of finite-dimensional Brownian motion [2], or various formulations of infinite-dimensional Brownian motion [4,6]. An important class of processes that were not covered are continuous time Markov processes with jumps, e.g., Lévy processes. In this paper we eliminate this gap, and in fact give variational representations for functionals of a fairly general Poisson random measure (PRM) plus an independent infinite-dimensional Brownian motion (BM), thereby covering many continuous time models.

In [28] Zhang has also proved a variational representation for functionals of a PRM. The representation in [28] is given in terms of certain predictable transformations on the canonical Poisson space. Existence of such transformations relies on solvability of certain nonlinear partial differential equations from the theory of mass transportation. This imposes restrictive conditions on the intensity measure (e.g., absolute continuity with respect to Lebesgue measure) of the PRM (see (H1) and (H2) of [28]), and even the very elementary setting of a standard Poisson process is not covered. Additionally, use of such a representation for proving large deviation results for general continuous time models with jumps appears to be unclear.

In contrast, we impose very mild assumptions on the intensity measure (namely, it is a σ-finite measure on a locally compact space), and establish a representation in Theorem 2.1 (see also Theorem 3.1), that is given in terms of a fixed PRM defined on an augmented space. A key question in formulating the representation for PRM is "*what form of controlled PRM is natural for the representation of exponential integrals via relative entropy duality?*" In the Brownian case there is little room for discussion, since control by shifting the mean is obviously very appealing. In [28] the control moves the atoms of the Poisson random measure through a rather complex nonlinear transformation. The fact that atoms are neither created nor destroyed is partly responsible for the fact that the representation does not cover the standard Poisson process. In the representation obtained here the control process enters as a censoring/thinning function in a very concrete fashion, which in turn allows for elementary weak convergence arguments in proofs of large deviation results.

As an application of the representation, we establish a general large deviation principle (LDP) for functionals of a PRM and an infinite-dimensional BM in Section 4. A similar LDP for functionals of an infinite-dimensional BM [4] has been used in recent years by numerous authors to study small noise asymptotics for a variety of infinite-dimensional stochastic dynamical models (e.g., [1, 6–8, 9, 10, 17–21, 23–25, 26, 29]). The LDP obtained in the current paper is expected to be similarly useful in the study of infinite-dimensional stochastic models with jumps (e.g., SPDEs with jumps). We illustrate the use of the LDP in Section 4 via a simple finite-dimensional jump-diffusion model. The goal is to simply show how the approach can be used and no attempt is made to obtain the best possible conditions.

An outline of the paper is as follows. In the next section we state and prove the representation for the case of a PRM. We note that the argument used for the lower bound is much simpler than the corresponding argument used in previous papers [4,6]. A statement of the general representation (i.e., for functionals of both a PRM and an infinite-dimensional BM) is in Section 3 with a sketch of proof given in the Appendix. A general large deviation result and a particular example are given in Section 4, and the paper concludes with an appendix which contains proofs of some auxiliary results.

Notation and a topology

The following notation will be used. For a metric space $\mathbb{S}$ denote by $M_b(\mathbb{S})$, $C(\mathbb{S})$, $C_b(\mathbb{S})$, $C_c(\mathbb{S})$, the spaces of real, bounded Borel measurable functions, continuous functions, continuous bounded functions and continuous functions with compact support, respectively. The Borel σ-field on $\mathbb{S}$ will be denoted as $\mathcal{B}(\mathbb{S})$. For an $\mathbb{S}$-valued measurable map X defined on some probability space (Ω, F, P) we will denote the measure induced by X on $(\mathbb{S}, \mathcal{B}(\mathbb{S}))$ by $P \circ X^{-1}$.

Given $\mathbb{S}$-valued random variables X_n, X, we will write $X_n \Rightarrow X$ to denote the weak convergence of $P \circ X_n^{-1}$ to $P \circ X^{-1}$. For a real bounded measurable map h on a measurable space $(V, \mathcal{V})$, we denote $\sup_{v\in V} |h(v)|$ by $|h|_\infty$. The space of all probability measures on $(V, \mathcal{V})$ is denoted as $P(V, \mathcal{V})$ or merely as $P(V)$, when clear from the context.

For a locally compact Polish space $\mathbb{S}$, we denote by $\mathcal{M}_F(\mathbb{S})$ the space of all measures ν on $(\mathbb{S}, \mathcal{B}(\mathbb{S}))$, satisfying $\nu(K) < \infty$ for every compact $K \subset \mathbb{S}$. We endow $\mathcal{M}_F(\mathbb{S})$ with the weakest topology such that for every $f \in C_c(\mathbb{S})$ the function $\nu \mapsto \langle f, \nu\rangle = \int_{\mathbb{S}} f(u)\nu(\mathrm{d}u)$, $\nu \in \mathcal{M}_F(\mathbb{S})$ is continuous. This topology can be metrized such that $\mathcal{M}_F(\mathbb{S})$ is a Polish space. One metric that is convenient for this purpose is the following. Consider a sequence of open sets $\{O_j, j \in \mathbb{N}\}$ such that $\bar{O}_j \subset O_{j+1}$, each $\bar{O}_j$ is compact, and $\bigcup_{j=1}^\infty O_j = \mathbb{S}$ (cf. Theorem 9.5.21 of [22]). Let $\phi_j(x) = [1 - d(x, O_j)] \vee 0$, where d denotes the metric on $\mathbb{S}$. Given any $\mu \in \mathcal{M}_F(\mathbb{S})$, let $\mu^j \in \mathcal{M}_F(\mathbb{S})$ be defined by $[\mathrm{d}\mu^j/\mathrm{d}\mu](x) = \phi_j(x)$. Given $\mu, \nu \in \mathcal{M}_F(\mathbb{S})$, let

$$\bar{d}(\mu, \nu) = \sum_{j=1}^\infty 2^{-j} \|\mu^j - \nu^j\|_{BL},$$

where $\|\cdot\|_{BL}$ denotes the bounded, Lipschitz norm:

$$\|\mu^j - \nu^j\|_{BL} = \sup\left\{\int_{\mathbb{S}} f\,\mathrm{d}\mu^j - \int_{\mathbb{S}} f\,\mathrm{d}\nu^j \colon |f|_\infty \le 1, \left|f(x) - f(y)\right| \le d(x, y) \text{ for all } x, y \in \mathbb{S}\right\}.$$

It is straightforward to check that $\bar{d}(\mu, \nu)$ defines a metric under which $\mathcal{M}_F(\mathbb{S})$ is a Polish space, and that convergence in this metric is essentially equivalent to weak convergence on each compact subset of $\mathbb{X}$. Specifically, $\bar{d}(\mu_n, \mu) \to 0$ if and only if for each $j \in \mathbb{N}$, $\mu_n^j \to \mu^j$ in the weak topology as finite nonnegative measures, i.e., for all $f \in C_b(\mathbb{X})$

$$\int_{\mathbb{S}} f\,\mathrm{d}\mu_n^j \to \int_{\mathbb{S}} f\,\mathrm{d}\mu^j.$$

Throughout $\mathcal{B}(\mathcal{M}_F(\mathbb{S}))$ will denote the Borel σ-field on $\mathcal{M}_F(\mathbb{S})$, under this topology.

2. Representations for functionals of a PRM

Let $\mathbb{X}$ be a locally compact Polish space. Fix $T \in (0, \infty)$ and let $\mathbb{X}_T = [0, T] \times \mathbb{X}$. Fix a measure $\nu \in \mathcal{M}_F(\mathbb{X})$ and let $\nu_T = \lambda_T \otimes \nu$, where λ_T is Lebesgue measure on $[0, T]$. Let $\mathbb{M} = \mathcal{M}_F(\mathbb{X}_T)$ and denote by $\mathbb{P}$ the unique probability measure on $(\mathbb{M}, \mathcal{B}(\mathbb{M}))$ under which the canonical map, $N : \mathbb{M} \to \mathbb{M}$, $N(m) \doteq m$, is a Poisson random measure with intensity measure ν_T (see [12], Section I.8). With applications to large deviations in mind, we also consider, for $\theta > 0$, the analogous probability measures $\mathbb{P}_\theta$ on $(\mathbb{M}, \mathcal{B}(\mathbb{M}))$ under which N is a Poisson random measure with intensity $\theta\nu_T$. The corresponding expectation operators will be denoted by $\mathbb{E}$ and $\mathbb{E}_\theta$, respectively.

We will obtain variational representations for $-\log \mathbb{E}_\theta(\exp[-F(N)])$, where $F \in M_b(\mathbb{M})$, in terms of a Poisson random measure constructed on a larger space. We now describe this construction.

2.1. *Controlled Poisson random measure*

Let $\mathbb{Y} = \mathbb{X} \times [0, \infty)$ and $\mathbb{Y}_T = [0, T] \times \mathbb{Y}$. Let $\bar{\mathbb{M}} = \mathcal{M}_F(\mathbb{Y}_T)$ and let $\bar{\mathbb{P}}$ be the unique probability measure on $(\bar{\mathbb{M}}, \mathcal{B}(\bar{\mathbb{M}}))$ such that the canonical map, $\bar{N} : \bar{\mathbb{M}} \to \bar{\mathbb{M}}$, $\bar{N}(m) \doteq m$, is a Poisson random measure with intensity measure $\bar{\nu}_T = \lambda_T \otimes \nu \otimes \lambda_\infty$, where λ_∞ is Lebesgue measure on $[0, \infty)$. The corresponding expectation operator will be denoted by $\bar{\mathbb{E}}$. The control will act through this additional component of the underlying point space. Let $\mathcal{G}_t \doteq \sigma\{\bar{N}((0, s] \times A)\colon 0 \le s \le t, A \in \mathcal{B}(\mathbb{Y})\}$, and to facilitate the use of a martingale representation theorem let $\bar{\mathcal{F}}_t$ denote the completion under $\bar{\mathbb{P}}$. We denote by $\bar{\mathcal{P}}$ the predictable σ-field on $[0, T] \times \bar{\mathbb{M}}$ with the filtration $\{\bar{\mathcal{F}}_t\colon 0 \le t \le T\}$ on $(\bar{\mathbb{M}}, \mathcal{B}(\bar{\mathbb{M}}))$. Let $\bar{\mathcal{A}}$ be the class of all $(\bar{\mathcal{P}} \otimes \mathcal{B}(\mathbb{X}))\backslash\mathcal{B}[0, \infty)$ measurable maps $\varphi : \mathbb{X}_T \times \bar{\mathbb{M}} \to [0, \infty)$. Since $\bar{\mathbb{M}}$ is the underlying probability space, following standard convention, we will at times suppress the dependence of $\varphi(t, x, \omega)$ on ω, $(t, x, \omega) \in \mathbb{X}_T \times \bar{\mathbb{M}}$, and merely write $\varphi(t, x)$. For $\varphi \in \bar{\mathcal{A}}$, define a counting process N^φ on $\mathbb{X}_T$ by

$$N^\varphi\big((0, t] \times U\big) = \int_{(0,t]\times U} \int_{(0,\infty)} 1_{[0,\varphi(s,x)]}(r) \bar{N}(\mathrm{d}s\,\mathrm{d}x\,\mathrm{d}r), \quad t \in [0, T], U \in \mathcal{B}(\mathbb{X}). \tag{2.1}$$

N^φ is to be thought of as a controlled random measure, with φ selecting the intensity for the points at location x and time s, in a possibly random but nonanticipating way. When, for some $\theta > 0$, $\varphi(s,x,\omega) = \theta$, for all $(s,x,\omega) \in \mathbb{X}_T \times \bar{\mathbb{M}}$, we will write N^φ as N^θ. Obviously N^θ has the same distribution on $\bar{\mathbb{M}}$ with respect to $\bar{\mathbb{P}}$ as N has on $\mathbb{M}$ with respect to $\mathbb{P}_\theta$. N^θ therefore plays the role of N on $\bar{\mathbb{M}}$. Define $\ell : [0,\infty) \to [0,\infty)$ by

$$\ell(r) = r\log r - r + 1, \quad r \in [0,\infty).$$

As is well known, ℓ is the local rate function for a standard Poisson process and so it is not surprising that it plays a key role in our analysis. For $\varphi \in \bar{\mathcal{A}}$, define a $[0,\infty]$-valued random variable $L_T(\varphi)$ by

$$L_T(\varphi)(\omega) = \int_{\mathbb{X}_T} \ell\big(\varphi(t,x,\omega)\big)\nu_T(\mathrm{d}t\,\mathrm{d}x), \quad \omega \in \bar{\mathbb{M}}. \tag{2.2}$$

The following is the main result of this section. The first equality is elementary.

Theorem 2.1. *Let $F \in M_b(\mathbb{M})$. Then, for $\theta > 0$,*

$$-\log \mathbb{E}_\theta\big(\mathrm{e}^{-F(N)}\big) = -\log \bar{\mathbb{E}}\big(\mathrm{e}^{-F(N^\theta)}\big) = \inf_{\varphi\in\bar{\mathcal{A}}} \bar{\mathbb{E}}\big[\theta L_T(\varphi) + F\big(N^{\theta\varphi}\big)\big].$$

The proof of this theorem follows from an upper bound established in Theorem 2.6 of Section 2.3.1, and a lower bound established in Theorem 2.8 of Section 2.3.2. For notational convenience we will only provide arguments for the case $\theta = 1$. The general case is treated similarly.

Remark 2.2. *In some applications a generalization of this result is useful. Consider a probability space with a complete filtration on which is given a PRM (with respect to this filtration) with intensity measure $\bar{\nu}_T = \lambda_T \otimes \nu \otimes \lambda_\infty$. Then a representation as in Theorem 2.1 holds with expectation computed on this space and the infimum taken over all controls that are predictable with respect to the given filtration (which may be strictly larger than the filtration generated by the PRM). A similar extension is possible for Theorem 3.1. For concreteness, the canonical space and filtration is considered here, but we note that analogous representations, for functionals of Brownian motions, in terms of general filtrations have been established in [4]. It is only in the proof of the upper bound in the representation that additional work is needed, and a remark on how the extension can be proved is given at the end of Section 2.3.1.*

2.2. *A class of nice controls*

Let $\{K_n \subset \mathbb{X}, n = 1,2,\ldots\}$ be an increasing sequence of compact sets such that $\bigcup_{n=1}^\infty K_n = \mathbb{X}$. For each n let

$$\bar{\mathcal{A}}_{b,n} \doteq \big\{\varphi \in \bar{\mathcal{A}} : \text{ for all } (t,\omega) \in [0,T] \times \bar{\mathbb{M}}, n \ge \varphi(t,x,\omega) \ge 1/n \text{ and } \varphi(t,x,\omega) = 1 \text{ if } x \in K_n^c\big\},$$

and let

$$\bar{\mathcal{A}}_b = \bigcup_{n=1}^\infty \bar{\mathcal{A}}_{b,n}.$$

This class of controls is particularly convenient. Let N_c^1 be the compensated version of N^1, which is defined by $N_c^1(A) = N^1(A) - \nu_T(A)$ for all $A \in \mathcal{B}(\mathbb{X}_T)$ such that $\nu_T(A) < \infty$.

Another class of processes, which will be used as test functions, is as follows. Let $\hat{A}_b$ be the set of all of all bounded $(\bar{\mathcal{P}} \otimes \mathcal{B}(\mathbb{Y}))\backslash\mathcal{B}(\mathbb{R})$ measurable maps $\vartheta : \mathbb{Y}_T \times \bar{\mathbb{M}} \to \mathbb{R}$, such that for some for some compact $K \subset \mathbb{Y}$, $\vartheta(s,x,r,\omega) = 0$ whenever $(x,r) \in K^c$. Once again ω will be frequently suppressed from the notation. The following result is standard (see, e.g., Theorem III.3.24 of [14]).

Lemma 2.3. *Let $\varphi \in \bar{\mathcal{A}}_b$. Then*

$$\begin{aligned}\mathcal{E}_t(\varphi) &\doteq \exp\left\{\int_{(0,t]\times\mathbb{X}} \log\big(\varphi(s,x)\big) N_c^1(\mathrm{d}s\,\mathrm{d}x) + \int_{(0,t]\times\mathbb{X}} \big(\log\big(\varphi(s,x)\big) - \varphi(s,x) + 1\big)\nu_T(\mathrm{d}s\,\mathrm{d}x)\right\} \\ &= \exp\left\{\int_{(0,t]\times\mathbb{X}\times[0,1]} \log\big(\varphi(s,x)\big) N(\mathrm{d}s\,\mathrm{d}x\,\mathrm{d}r) + \int_{(0,t]\times\mathbb{X}\times[0,1]} \big(-\varphi(s,x)+1\big)\bar{\nu}_T(\mathrm{d}s\,\mathrm{d}x\,\mathrm{d}r)\right\}\end{aligned}$$

is an $\{\bar{\mathcal{F}}_t\}$-martingale. Define a probability measure $\mathbb{Q}^\varphi$ on $\bar{\mathbb{M}}$ by

$$\mathbb{Q}^\varphi(G) = \int_G \mathcal{E}_T(\varphi)\,\mathrm{d}\bar{\mathbb{P}} \quad \textit{for } G \in \mathcal{B}(\bar{\mathbb{M}}).$$

Then for any $\vartheta \in \hat{A}_b$,

$$\mathbb{E}^{\mathbb{Q}^\varphi}\int \vartheta(s,x,r)\bar{N}(\mathrm{d}s\,\mathrm{d}x\,\mathrm{d}r) = \mathbb{E}^{\mathbb{Q}^\varphi}\int \vartheta(s,x,r)\big[\varphi(s,x)1_{(0,1]}(r) + 1_{(1,\infty)}(r)\big]\bar{\nu}_T(\mathrm{d}s\,\mathrm{d}x\,\mathrm{d}r).$$

The last statement in the lemma says that under $\mathbb{Q}^\varphi$, $\bar{N}$ is a random counting measure with compensator $[\varphi(s,x)1_{(0,1]}(r) + 1_{(1,\infty)}(r)]\bar{\nu}_T(\mathrm{d}s\,\mathrm{d}x\,\mathrm{d}r)$.

Lemma 2.4. *Let $\varphi \in \bar{\mathcal{A}}_{b,n}$. Then there exists a sequence of processes $\varphi_k \in \bar{\mathcal{A}}_{b,n}$ with the following properties*:

(1) *There exist $\ell, n_1, \dots, n_\ell \in \mathbb{N}$ and a partition $0 = t_0 < t_1 < \cdots < t_\ell = T$, $\bar{\mathcal{F}}_{t_{i-1}}$-measurable random variables X_{ij}, $i = 1, \dots, \ell$, $j = 1, \dots, n_\ell$, and for each $i = 1, \dots, \ell$ a disjoint measurable partition E_{ij} of K_n, $j = 1, \dots, n_\ell$, such that $1/n \le X_{ij} \le n$ and*

$$\varphi_k(t,x,\bar{m}) = 1_{\{0\}}(t) + \sum_{i=1}^{\ell}\sum_{j=1}^{n_i} 1_{(t_{i-1},t_i]}(t) X_{ij}(\bar{m}) 1_{E_{ij}}(x) + 1_{K_n^c}(x)1_{(0,T]}(t). \tag{2.3}$$

(2) *N^{φ_k} converges in distribution to N^φ as $k \to \infty$.*
(3) *$\bar{\mathbb{E}}|L_T(\varphi_k) - L_T(\varphi)| \to 0$ and $\bar{\mathbb{E}}|\mathcal{E}_T(\varphi_k) - \mathcal{E}_T(\varphi)| \to 0$, as $k \to \infty$.*

The collection of processes identified in item 1 of the lemma will be denoted $\bar{\mathcal{A}}_{s,n}$. We let $\bar{\mathcal{A}}_s = \bigcup_{n=1}^\infty \bar{\mathcal{A}}_{s,n}$ and refer to elements in $\bar{\mathcal{A}}_s$ as *simple* processes.

Proof of Lemma 2.4. We first construct processes φ_k which satisfy parts (2) and (3) of the lemma but not (necessarily) part (1). For $k \in \mathbb{N}$ define

$$\varphi_k(t,x,\omega) = \frac{k}{n}\left(\frac{1}{k} - t\right)^+ + k\int_{(t-1/k)^+}^t \varphi(s,x,\omega)\,\mathrm{d}s, \quad (t,x,\omega) \in \mathbb{X}_T \times \bar{\mathbb{M}}.$$

An application of Lusin's theorem gives that for $\nu \otimes \bar{\mathbb{P}}$-a.e. (x,ω), as $k \to \infty$

$$\int_{[0,T]} \big|\varphi_k(t,x,\omega) - \varphi(t,x,\omega)\big|\,\mathrm{d}t \to 0, \qquad \int_{[0,T]} \big|\ell\big(\varphi_k(t,x,\omega)\big) - \ell\big(\varphi(t,x,\omega)\big)\big|\,\mathrm{d}t \to 0. \tag{2.4}$$

In particular, $\varphi_k \in \bar{\mathcal{A}}_{b,n}$ for every k and $\bar{\mathbb{E}}|L_T(\varphi_k) - L_T(\varphi)| \to 0$, as $k \to \infty$. Also note that for $f \in C_c(\mathbb{X}_T)$

$$\begin{aligned}\bar{\mathbb{E}}\big|\langle f, N^{\varphi_k}\rangle - \langle f, N^\varphi\rangle\big| &\le \bar{\mathbb{E}}\int_{\mathbb{Y}_T} |f(s,x)|\big|1_{[0,\varphi_k(s,x,\omega)]}(r) - 1_{[0,\varphi(s,x,\omega)]}(r)\big|\bar{\nu}_T(\mathrm{d}s\,\mathrm{d}x\,\mathrm{d}r) \\ &\le c_1\bar{\mathbb{E}}\int_{[0,T]\times K_n} \big|\varphi_k(s,x,\omega) - \varphi(s,x,\omega)\big|\nu_T(\mathrm{d}s\,\mathrm{d}x).\end{aligned}$$

Using (2.4) and $\nu(K_n) < \infty$ shows that the last quantity approaches 0 as $k \to \infty$, and hence $N^{\varphi_k} \Rightarrow N^{\varphi}$.

Next we consider the $L^1(\bar{\mathbb{P}})$ convergence of $\mathcal{E}_T(\varphi_k)$. By Scheffe's lemma it suffices to show that

$$\mathcal{E}_T(\varphi_k) \to \mathcal{E}_T(\varphi) \quad \text{in } \bar{\mathbb{P}}\text{-probability}. \tag{2.5}$$

For this, it is enough to show that

$$\int_{[0,T]\times\mathbb{X}} \bigl(1-\varphi_k(s,x)\bigr)\nu_T(\mathrm{d}s\,\mathrm{d}x) \to \int_{[0,T]\times\mathbb{X}} \bigl(1-\varphi(s,x)\bigr)\nu_T(\mathrm{d}s\,\mathrm{d}x)$$

and

$$\int_{[0,T]\times\mathbb{X}} \log\bigl(\varphi_k(s,x)\bigr)N^1(\mathrm{d}s\,\mathrm{d}x) \to \int_{[0,T]\times\mathbb{X}} \log\bigl(\varphi(s,x)\bigr)N^1(\mathrm{d}s\,\mathrm{d}x)$$

in probability as $k \to \infty$. The first convergence is immediate from (2.4), the uniform bounds on $\varphi_k, \varphi, \nu(K_n) < \infty$, and the fact that $1-\varphi_k(s,x) = 1-\varphi(s,x) = 0$ for $x \notin K_n$. The second convergence follows similarly on noting that

$$\bigl|\log\bigl(\varphi_k(s,x)\bigr) - \log\bigl(\varphi(s,x)\bigr)\bigr| \le n\bigl|\varphi_k(s,x) - \varphi(s,x)\bigr|.$$

This proves (2.5) and so $\bar{\mathbb{E}}|\mathcal{E}_T(\varphi_k) - \mathcal{E}_T(\varphi)| \to 0$ as $k \to \infty$. This completes the construction of φ_k which satisfy parts (2) and (3) of the lemma.

Next we show that part (1) can also be satisfied. Note that by construction, $t \longmapsto \varphi_k(t,x,\omega)$ is continuous for $\nu \otimes \bar{\mathbb{P}}$-a.e. (x,ω). Consider any φ_k as constructed previously, and to simplify the notation drop the k subscript. Two more levels of approximation will be used, and indexed by q and r. Thus for the fixed φ and $q \in \mathbb{N}$ define

$$\varphi_q(t,x,\omega) = \sum_{m=0}^{\lfloor qT \rfloor} \varphi\left(\frac{m}{q}, x, \omega\right) 1_{(m/q,(m+1)/q]}(t), \quad (t,x,\omega) \in \mathbb{X}_T \times \bar{\mathbb{M}}.$$

It is easily checked that (2.4) is satisfied as $q \to \infty$, and so, arguing as above, the sequence $\{\varphi_q\}$ satisfies parts (2) and (3) of the lemma. Note that for fixed q and m, $g(x,\omega) = \varphi(m/q,x,\omega)$ is a $\mathcal{B}(\mathbb{X}) \otimes \bar{\mathcal{F}}_{m/q}$-measurable map with values in $[1/n, n]$ and $g(x,\omega) = 1$ for $x \in K_n^c$. By a standard approximation procedure one can find $\mathcal{B}(\mathbb{X}) \otimes \bar{\mathcal{F}}_{m/q}$-measurable maps $g_r, r \in \mathbb{N}$ with the following properties: $g_r(x,\omega) = \sum_{j=1}^{a(r)} c_j^r(\omega) 1_{E_j^r}(x)$ for $x \in K_n$, where for each r, $\{E_j^r\}_{j=1}^{a(r)}$ is some measurable partition of K_n and for all j, r, $c_j^r(\omega) \in [1/n, n]$ a.s.; $g_r(x,\omega) = 1$ for $x \in K_n^c$; $g_r \to g$ a.s. $\nu \otimes \bar{\mathbb{P}}$. Hence by taking q and r large we can find processes which satisfy all parts (1)–(3) of the lemma. □

A last result is needed before proving the main theorem. This result, which is a key element in the proof of the representation, shows that simple controls under a new measure can always be replicated on the original probability space, and vice versa. The proof of the lemma, which uses an elementary but detailed argument, is in the Appendix.

Lemma 2.5. *For every $\varphi \in \bar{\mathcal{A}}_s$, there is $\tilde{\varphi} \in \bar{\mathcal{A}}_s$ such that $\bar{\mathbb{P}} \circ (N^{\varphi})^{-1} = \mathbb{Q}^{\tilde{\varphi}} \circ (N^1)^{-1}$ and*

$$\mathbb{E}^{\mathbb{Q}^{\tilde{\varphi}}}\bigl[L_T(\tilde{\varphi}) + F\bigl(N^1\bigr)\bigr] = \bar{\mathbb{E}}\bigl[L_T(\varphi) + F\bigl(N^{\varphi}\bigr)\bigr]. \tag{2.6}$$

Conversely, given any $\tilde{\varphi} \in \bar{\mathcal{A}}_s$ there is $\varphi \in \bar{\mathcal{A}}_s$ such that $\bar{\mathbb{P}} \circ (N^{\varphi})^{-1} = \mathbb{Q}^{\tilde{\varphi}} \circ (N^1)^{-1}$ and (2.6) holds.

2.3. *Proof of Theorem* 2.1

The starting point for the proof is the basic relative entropy representation for exponential integrals. For $\mathbb{Q}$ and $\mathbb{P}$ in $P(\bar{\mathbb{M}})$, the relative entropy of $\mathbb{Q}$ with respect to $\mathbb{P}$ is defined by

$$R(\mathbb{Q} \,\|\, \mathbb{P}) = \int_{\bar{\mathbb{M}}} \left(\log \frac{\mathrm{d}\mathbb{Q}}{\mathrm{d}\mathbb{P}}\right) \mathrm{d}\mathbb{Q}$$

whenever $\mathbb{Q}$ is absolutely continuous with respect to $\mathbb{P}$ and $\log(\mathrm{d}\mathbb{Q}/\mathrm{d}\mathbb{P})$ is $\mathbb{Q}$-integrable, and in all other cases $R(\mathbb{Q} \parallel \mathbb{P}) = \infty$.

Let $h : \bar{\mathbb{M}} \to \mathbb{M}$ be defined by

$$h(\bar{m})\big(U \times (0,t]\big) = \int_{U\times(0,t]\times(0,\infty)} 1_{[0,1]}(r)\bar{m}(\mathrm{d}s\,\mathrm{d}x\,\mathrm{d}r), \quad t \in [0,T], U \in \mathcal{B}(\mathbb{X}).$$

Thus $N^1 = h(N)$. By the well-known relative entropy representation for exponential integrals (see, e.g., Proposition 1.4.2 of [11]),

$$\begin{aligned}
-\log \bar{\mathbb{E}}\big(\mathrm{e}^{-F(N^1)}\big) &= -\log \int_{\bar{\mathbb{M}}} \mathrm{e}^{-F(h(\bar{m}))}\bar{\mathbb{P}}(\mathrm{d}\bar{m}) \\
&= \inf_{\mathbb{Q}\in P(\bar{\mathbb{M}})} \left[R(\mathbb{Q} \parallel \bar{\mathbb{P}}) + \int_{\bar{\mathbb{M}}} F\big(h(\bar{m})\big)\mathbb{Q}(\mathrm{d}\bar{m})\right]. \qquad (2.7)
\end{aligned}$$

2.3.1. *Proof of the upper bound*

Theorem 2.6. *For every $F \in M_b(\mathbb{M})$*

$$-\log \bar{\mathbb{E}}\big(\mathrm{e}^{-F(N^1)}\big) \le \inf_{\varphi\in\bar{\mathcal{A}}} \bar{\mathbb{E}}\big[L_T(\varphi) + F\big(N^\varphi\big)\big].$$

Proof. We begin by evaluating $R(\mathbb{Q}^\varphi \parallel \bar{\mathbb{P}})$ for a $\varphi \in \bar{\mathcal{A}}_b$. By Lemma 2.3 $\{\mathcal{E}_t(\varphi)\}$ is an $\{\bar{\mathcal{F}}_t\}$-martingale and under $\mathbb{Q}^\varphi$, $\bar{N}$ is a random counting measure with compensator $[\varphi(s,x)1_{(0,1]}(r) + 1_{(1,\infty)}(r)]\bar{\nu}_T(\mathrm{d}s\,\mathrm{d}x\,\mathrm{d}r)$. It follows from the definition of relative entropy and L_T in (2.2) that

$$\begin{aligned}
R\big(\mathbb{Q}^\varphi \parallel \bar{\mathbb{P}}\big) &= \int\left[\int \log\big(\varphi(s,x)\big)N_c^1(\mathrm{d}s\,\mathrm{d}x) + \int \big(\log\big(\varphi(s,x)\big) - \varphi(s,x) + 1\big)\nu_T(\mathrm{d}s\,\mathrm{d}x)\right]\mathrm{d}\mathbb{Q}^\varphi \\
&= \int\left[\int \log\big(\varphi(s,x)\big)N^1(\mathrm{d}s\,\mathrm{d}x) + \int \big(-\varphi(s,x) + 1\big)\nu_T(\mathrm{d}s\,\mathrm{d}x)\right]\mathrm{d}\mathbb{Q}^\varphi \\
&= \int\left[\int \big(\varphi(s,x)\log\big(\varphi(s,x)\big) - \varphi(s,x) + 1\big)\nu_T(\mathrm{d}s\,\mathrm{d}x)\right]\mathrm{d}\mathbb{Q}^\varphi \\
&= \mathbb{E}^{\mathbb{Q}^\varphi} L_T(\varphi). \qquad (2.8)
\end{aligned}$$

Thus, by (2.7), for $\varphi \in \bar{\mathcal{A}}_b$

$$\begin{aligned}
-\log \bar{\mathbb{E}}\big(\mathrm{e}^{-F(N^1)}\big) &\le \left[R\big(\mathbb{Q}^\varphi \parallel \bar{\mathbb{P}}\big) + \int_{\bar{\mathbb{M}}} F\big(h(\bar{m})\big)\mathbb{Q}^\varphi(\mathrm{d}\bar{m})\right] \\
&= \mathbb{E}^{\mathbb{Q}^\varphi}\big[L_T(\varphi) + F\big(N^1\big)\big]. \qquad (2.9)
\end{aligned}$$

We complete the proof by showing that for any $\varphi \in \bar{\mathcal{A}}$,

$$-\log \bar{\mathbb{E}}\big(\mathrm{e}^{-F(N^1)}\big) \le \bar{\mathbb{E}}\big[L_T(\varphi) + F\big(N^\varphi\big)\big]. \qquad (2.10)$$

Case 1: $\varphi \in \bar{\mathcal{A}}_s$. According to Lemma 2.5 one can find $\tilde{\varphi}$ that is $\bar{\mathcal{F}}_t$-predictable and simple, and such that

$$\mathbb{E}^{\mathbb{Q}^{\tilde{\varphi}}}\big[L_T(\tilde{\varphi}) + F\big(N^1\big)\big] = \bar{\mathbb{E}}\big[L_T(\varphi) + F\big(N^\varphi\big)\big].$$

Thus (2.10) follows directly from (2.9).

Case 2: $\varphi \in \bar{\mathcal{A}}_b$. Given $\varphi \in \bar{\mathcal{A}}_b$, let $\varphi_k \in \bar{\mathcal{A}}_s$ be the sequence constructed as in Lemma 2.4. By case 1, for every $k \in \mathbb{N}$

$$-\log \bar{\mathbb{E}}\big(\mathrm{e}^{-F(N^1)}\big) \le \bar{\mathbb{E}}\big[L_T(\varphi_k) + F\big(N^{\varphi_k}\big)\big]. \qquad (2.11)$$

From Lemma 2.4, under $\bar{\mathbb{P}}$, $N^{\varphi_k} \Rightarrow N^{\varphi}$ and $\bar{\mathbb{E}}[L_T(\varphi_k)] \to \bar{\mathbb{E}}[L_T(\varphi)]$. However, F is not assumed continuous, and so we cannot simply pass to the limit in the last display. Instead, we will apply Lemma 2.8 of [2], which allows F to be just bounded and measurable when there are bounds on certain relative entropies. The first and the last equalities in the following display follow from Lemma 2.5, the second equality is a consequence of (2.8), and the inequality follows from the fact that relative entropy can only decrease when considering measures induced by the same mapping (in this case the random variable N^1):

$$\begin{aligned} R\big(\bar{\mathbb{P}} \circ \big(N^{\varphi_k}\big)^{-1} \parallel \bar{\mathbb{P}} \circ \big(N^1\big)^{-1}\big) &= R\big(\mathbb{Q}^{\tilde{\varphi}_k} \circ \big(N^1\big)^{-1} \parallel \bar{\mathbb{P}} \circ \big(N^1\big)^{-1}\big) \\ &\le R\big(\mathbb{Q}^{\tilde{\varphi}_k} \parallel \bar{\mathbb{P}}\big) \\ &= \mathbb{E}^{\mathbb{Q}^{\tilde{\varphi}_k}}\big[L_T(\tilde{\varphi}_k)\big] \\ &= \bar{\mathbb{E}}\big[L_T(\varphi_k)\big]. \end{aligned} \tag{2.12}$$

Since $\bar{\mathbb{E}}[L_T(\varphi_k)] \to \bar{\mathbb{E}}[L_T(\varphi)] < \infty$ we have $\sup_k R(\bar{\mathbb{P}} \circ (N^{\varphi_k})^{-1} \parallel \bar{\mathbb{P}} \circ (N^1)^{-1}) < \infty$, and so by Lemma 2.8 of [2] we can pass to the limit in (2.11) and obtain (2.10) for all $\varphi \in \bar{\mathcal{A}}_b$. For future use we note that the lower semicontinuity of relative entropy and (2.12) imply $R(\bar{\mathbb{P}} \circ (N^{\varphi})^{-1} \parallel \bar{\mathbb{P}} \circ (N^1)^{-1}) \le \bar{\mathbb{E}}[L_T(\varphi)]$ for $\varphi \in \bar{\mathcal{A}}_b$.

Case 3: $\varphi \in \bar{\mathcal{A}}$. Define

$$\varphi_n(x,t,\omega) = \begin{cases} \big[\varphi(x,t,\omega) \vee (1/n)\big] \wedge n, & x \in K_n, \\ 1, & \text{else.} \end{cases}$$

Note that $\varphi_n \in \bar{\mathcal{A}}_{b,n}$, and so (2.10) holds with φ replaced by φ_n. Since the definition of φ_n implies $\ell(\varphi_n(x,t,\omega))$ is nondecreasing in n, by monotone convergence $\bar{\mathbb{E}}L_T(\varphi_n) \uparrow \bar{\mathbb{E}}L_T(\varphi)$. If $\bar{\mathbb{E}}L_T(\varphi) = \infty$ there is nothing to prove. Assume therefore that

$$\bar{\mathbb{E}}L_T(\varphi) < \infty. \tag{2.13}$$

Then $R(\bar{\mathbb{P}} \circ (N^{\varphi_n})^{-1} \parallel \bar{\mathbb{P}} \circ (N^1)^{-1}) \le \bar{\mathbb{E}}L_T(\varphi_n) \le \bar{\mathbb{E}}L_T(\varphi)$. Since the relative entropies are uniformly bounded and the level sets of the relative entropy function are compact (see Lemma 1.4.3(c) of [11]), at least along a subsequence N^{φ_n} converges in distribution to some random variable N^*. We claim that N^* has same distribution as N^{φ}. If true, we can once again apply Lemma 2.8 of [2], pass to the limit on n, and thereby obtain (2.10) for arbitrary $\varphi \in \bar{\mathcal{A}}$.

To prove the claim and hence complete the proof of Theorem 2.6, it suffices to show that for every $f \in C_c(\mathbb{X}_T)$

$$\langle f, N^{\varphi_n}\rangle \to \langle f, N^{\varphi}\rangle \quad \text{in } \bar{\mathbb{P}}\text{-probability as } n \to \infty. \tag{2.14}$$

Let n_0 be large enough that the support of f is contained in $[0,T] \times K_{n_0}$. Then for all $n \ge n_0$

$$\bar{\mathbb{E}}\big|\langle f, N^{\varphi_k}\rangle - \langle f, N^{\varphi}\rangle\big| \le |f|_\infty \bar{\mathbb{E}} \int_{[0,T]\times K_{n_0}} \left(\frac{1}{n} + \big(\varphi(t,x) - n\big)^+\right) \nu_T(\mathrm{d}s\,\mathrm{d}x).$$

Next note that $\nu_T([0,T] \times K_{n_0}) < \infty$, $(\varphi(t,x) - n)^+ \to 0$ as $n \to \infty$, and that $(\varphi(t,x) - n)^+ \le \ell(\varphi(t,x))$. These observations together with (2.13) show that the right-hand side in the last display approaches 0 as $n \to \infty$. This proves (2.14) and the claim follows. $\square$

Remark 2.7. *Following up on Remark 2.2, we indicate here how one can generalize Theorem 2.1 to allow the infimum in the representation to be over all controls which are predictable with respect to a possibly larger filtration. The only issue is whether the upper bound will continue to hold when a filtration larger than the completion of the one generated by the PRM is used. First note that Lemma 2.4 is valid without change in this more general setting. Letting an asterisk denote objects on the new probability space, it follows that for any bounded control φ^* we can find a sequence of simple controls φ_k^* (all predictable with respect to the given filtration) such that*

$$N^{*,\varphi_k^*} \text{ converges in distribution to } N^{*,\varphi^*} \tag{2.15}$$

and

$$\mathbb{E}^* L_T\big(\varphi_k^*\big) \to \mathbb{E}^* L_T\big(\varphi^*\big). \tag{2.16}$$

Arguing as in case 3 *of Theorem* 2.6, *the same result holds without the boundedness assumption so long as* $\mathbb{E}^* L_T(\varphi^*) < \infty$, *which can be assumed without loss.*

Given any φ_k^*, one can find a sequence of sets $S_j \subset [0,\infty)$, each with finite cardinality, and a sequence of S_j-valued controls $\varphi_{k,j}^*$, such that $\mathbb{E}^* L_T(\varphi_{k,j}^*) \to \mathbb{E}^* L_T(\varphi_k^*)$ and $N^{*,\varphi_{k,j}^*}$ converges in distribution to N^{*,φ_k^*}. Thus in (2.15) and (2.16) we can assume that each φ_k^* takes on only a finite number of values. Finally, using the martingale convergence theorem as in [16], Theorem 10.3.1, and a chattering lemma as in [15], Theorem 3.1, we can further assume that for each k there is $\theta > 0$ and a finite partition $\{\Gamma_k, k = 1,\dots,K\}$ of $\mathbb{Y}$ such that $\varphi_k^*(s)$ depends only on $N^*([0,l\theta],\Gamma_k)$ for l such that $l\theta \le s$ and $k = 1,\dots,K$, where N^* is the PRM with intensity $\bar{\nu}_T = \lambda_T \otimes \nu \otimes \lambda_\infty$ on this space. This exhibits $\varphi_k^*(s)$ as a measurable function of the past values of this PRM, and hence there are corresponding simple controls φ_k on the canonical space such that $\bar{\mathbb{E}} L_T(\varphi_k) \to \mathbb{E}^* L_T(\varphi^*)$ and N^{φ_k} converges in distribution to N^{*,φ^*}. Using the identification of $\bar{\mathbb{E}} L_T(\varphi_k)$ with relative entropies and the convergence in distribution, we can argue as before that $\bar{\mathbb{E}} F(N^{\varphi_k}) \to \mathbb{E}^* F(N^{*,\varphi^*})$. Thus the infimum over all controls with respect to the larger filtration is no smaller than the infimum over all controls on the canonical space.

2.3.2. *Proof of the lower bound*

Theorem 2.8. *For every* $F \in M_b(\mathbb{M})$

$$-\log \bar{\mathbb{E}}\big(\mathrm{e}^{-F(N^1)}\big) \ge \inf_{\varphi \in \bar{\mathcal{A}}} \bar{\mathbb{E}}\big[L_T(\varphi) + F\big(N^\varphi\big)\big]. \tag{2.17}$$

Proof. Following [28], we first consider a class of particularly simple F. Let $F \in M_b(\mathbb{M})$ be of the form $F(m) = g(\langle f_1, m\rangle,\dots,\langle f_k, m\rangle)$, where $k \in \mathbb{N}$, $g \in C_c^\infty(\mathbb{R}^k)$, and $f_i \in C_c(\mathbb{X}_T)$. The class of all such F is denoted by $C_{\mathrm{cyl}}(\mathbb{M})$. Once the result for $F \in C_{\mathrm{cyl}}(\mathbb{M})$ is proved, the general case will follow from an approximation argument based on the fact that $C_{\mathrm{cyl}}(\mathbb{M})$ is dense in $M_b(\mathbb{M})$, namely for an arbitrary $\tilde{F} \in M_b(\mathbb{M})$ one can find a sequence of $F_j \in C_{\mathrm{cyl}}(\mathbb{M})$ such that for all $j \ge 1$, $\|F_j\| \le \|\tilde{F}\| < \infty$ and as $j \to \infty$, $F_j(m) \to \tilde{F}(m)$ for $\bar{\mathbb{P}}$-almost all $m \in \mathbb{M}$. By Proposition 1.4.2 of [11],

$$-\log \bar{\mathbb{E}}\big(\mathrm{e}^{-F(N^1)}\big) = R(\mathbb{Q} \,\|\, \bar{\mathbb{P}}) + \mathbb{E}^{\mathbb{Q}}\big[F\big(h(\bar{N})\big)\big],$$

where $\mathbb{Q}$ is the probability measure defined by

$$\mathbb{Q}(A) = \frac{\int_A \mathrm{e}^{-F(h(\bar{m}))}\, \mathrm{d}\bar{\mathbb{P}}(\bar{m})}{\int_{\bar{\mathbb{M}}} \mathrm{e}^{-F(h(\bar{m}))}\, \mathrm{d}\bar{\mathbb{P}}(\bar{m})}. \tag{2.18}$$

By the martingale representation for Poisson random measures ([14], Theorem III.4.37), there is an $\bar{\mathcal{F}}_t$-predictable process $\tilde{\varphi}$ such that

$$\frac{\mathrm{d}\mathbb{Q}}{\mathrm{d}\bar{\mathbb{P}}} = \mathcal{E}_T(\tilde{\varphi}). \tag{2.19}$$

Owing to the form of F, $\tilde{\varphi} \in \mathcal{A}_{b,n}$ for some $n < \infty$ (see Proposition 4.2 and Eq. (30) of [28]). It follows from (2.8) that

$$-\log \bar{\mathbb{E}}\big(\mathrm{e}^{-F(N^1)}\big) = \mathbb{E}^{\mathbb{Q}^{\tilde{\varphi}}}\big[L_T(\tilde{\varphi}) + F\big(h(\bar{N})\big)\big], \tag{2.20}$$

where we now denote $\mathbb{Q}$ by $\mathbb{Q}^{\tilde{\varphi}}$. For F of this special form, it remains to construct a near minimizer on the original probability space. Given $\varepsilon \in (0,1)$ we will construct $\varphi \in \mathcal{A}_{s,n}$ such that

$$\mathbb{E}^{\mathbb{Q}^{\tilde{\varphi}}}\big[L_T(\tilde{\varphi}) + F\big(h(\bar{N})\big)\big] \ge \bar{\mathbb{E}}\big[L_T(\varphi) + F\big(N^\varphi\big)\big] - \varepsilon. \tag{2.21}$$

Since $\varepsilon > 0$ is arbitrary this will complete the proof of (2.17) for $F \in C_{\mathrm{cyl}}(\mathbb{M})$.

Let $\tilde{\varphi}_k$ be a sequence in $\mathcal{A}_s$ as constructed in Lemma 2.4 for $\tilde{\varphi}$ introduced in (2.19). We claim that as $k \to \infty$

$$\mathbb{E}^{\mathbb{Q}^{\tilde{\varphi}_k}}\left[L_T(\tilde{\varphi}_k) + F\left(h(\bar{N})\right)\right] \to \mathbb{E}^{\mathbb{Q}^{\tilde{\varphi}}}\left[L_T(\tilde{\varphi}) + F\left(h(\bar{N})\right)\right]. \tag{2.22}$$

To see this, rewrite the quantities in (2.22) in terms of the original measure $\bar{\mathbb{P}}$:

$$\mathbb{E}^{\mathbb{Q}^{\tilde{\varphi}_k}}\left[L_T(\tilde{\varphi}_k) + F\left(h(\bar{N})\right)\right] = \bar{\mathbb{E}}\left(\mathcal{E}_T(\tilde{\varphi}_k)\left[L_T(\tilde{\varphi}_k) + F\left(h(\bar{N})\right)\right]\right)$$

and

$$\mathbb{E}^{\mathbb{Q}^{\tilde{\varphi}}}\left[L_T(\tilde{\varphi}) + F\left(h(\bar{N})\right)\right] = \bar{\mathbb{E}}\left(\mathcal{E}_T(\tilde{\varphi})\left[L_T(\tilde{\varphi}) + F\left(h(\bar{N})\right)\right]\right).$$

To show (2.22) it is enough to show that

$$\bar{\mathbb{E}}\left(\left[\mathcal{E}_T(\tilde{\varphi}_k) - \mathcal{E}_T(\tilde{\varphi})\right]\left[L_T(\tilde{\varphi}_k) + F\left(h(\bar{N})\right)\right]\right) \to 0 \tag{2.23}$$

and

$$\bar{\mathbb{E}}\left(\mathcal{E}_T(\tilde{\varphi})\left[L_T(\tilde{\varphi}_k) - L_T(\tilde{\varphi})\right]\right) \to 0. \tag{2.24}$$

However, statements (2.23) and (2.24) follow from part (3) of Lemma 2.4 on noting that F and $\mathcal{E}_T(\tilde{\varphi})$ are bounded, and $L_T(\tilde{\varphi}_k)$ is uniformly bounded for $k \in \mathbb{N}$.

Now fix k large enough that the difference between the expressions on the two sides of (2.22) is bounded by ε. According to the second part of Lemma 2.5 (with $\tilde{\varphi}$ there replaced by $\tilde{\varphi}_k$), we can find φ such that (2.21) holds, which proves the theorem when $F \in C_{\mathrm{cyl}}(\mathbb{M})$.

Next consider an arbitrary $F \in M_b(\mathbb{M})$. Let $F_j \in C_{\mathrm{cyl}}(\mathbb{M})$ be such that $\|F_j\| \le \|F\| < \infty$ and $F_j(m) \to F(m)$ for $\bar{\mathbb{P}}$-almost all $m \in \mathbb{M}$. By dominated convergence

$$-\log \bar{\mathbb{E}}\left(\mathrm{e}^{-F_j(N^1)}\right) \to -\log \bar{\mathbb{E}}\left(\mathrm{e}^{-F(N^1)}\right). \tag{2.25}$$

Let $\tilde{\varphi}_j \in \bar{\mathcal{A}}_{b,n}$ be determined by applying the martingale representation theorem to $\mathrm{d}\mathbb{Q}/\mathrm{d}\bar{\mathbb{P}}$ where $\mathbb{Q}$ is defined by (2.18), but with F replaced by F_j. Let $\varphi_j \in \bar{\mathcal{A}}_{s,n}$ be such that (2.21) holds with $(\tilde{\varphi}_j, \varphi_j, F_j)$ replacing $(\tilde{\varphi}, \varphi, F)$. Then (2.21) along with (2.20) gives $\sup_j \bar{\mathbb{E}}(L_T(\varphi_j)) \le 2|F|_\infty + 1$. As noted in (2.12), $R(\bar{\mathbb{P}} \circ (N^{\varphi_j})^{-1} \,\|\, \bar{\mathbb{P}} \circ (N^1)^{-1}) \le \bar{\mathbb{E}}(L_T(\varphi_j))$. Thus, along a subsequence, N^{φ_j} converges in distribution to some limit N^*. By Lemma 2.8 of [2], along this subsequence

$$\bar{\mathbb{E}}\left[F_j\left(N^{\varphi_j}\right)\right] \to \bar{\mathbb{E}}\left[F\left(N^*\right)\right] \quad \text{and} \quad \bar{\mathbb{E}}\left[F\left(N^{\varphi_j}\right)\right] \to \bar{\mathbb{E}}\left[F\left(N^*\right)\right]. \tag{2.26}$$

Finally, by (2.25), (2.21) (with $(\tilde{\varphi}_j, \varphi_j, F_j)$) and (2.26), for sufficiently large j within the convergent subsequence

$$\begin{aligned}
-\log \bar{\mathbb{E}}\left(\mathrm{e}^{-F(N^1)}\right) &\ge -\log \bar{\mathbb{E}}\left(\mathrm{e}^{-F_j(N^1)}\right) - \varepsilon \\
&= \bar{\mathbb{E}}\left[L_T(\varphi_j) + F_j\left(N^{\varphi_j}\right)\right] - 2\varepsilon \\
&\ge \bar{\mathbb{E}}\left[L_T(\varphi_j) + F\left(N^{\varphi_j}\right)\right] - 3\varepsilon.
\end{aligned}$$

Since $\varepsilon > 0$ is arbitrary, this completes the proof of the lower bound. □

3. Functionals of PRM and BM

In this section we state the representation for functionals of both a PRM and an infinite-dimensional BM. One can obtain, as in [6], representations for related objects, such as Hilbert space-valued BM, Brownian sheet, etc.

We first define the probability space. Denote the product space of countable infinite copies of the real line by $\mathbb{R}^\infty$. Endowed with the topology of coordinate-wise convergence $\mathbb{R}^\infty$ is a Polish space. Recall the definitions of $\mathbb{M}$ and $\bar{\mathbb{M}}$ from the beginning of Section 2. We denote the Polish space $C([0,T]:\ \mathbb{R}^\infty)$ by $\mathbb{W}$ and denote by $\mathbb{V}$ the product space $\mathbb{W}\times\mathbb{M}$. Let $\bar{\mathbb{V}}=\mathbb{W}\times\bar{\mathbb{M}}$. Abusing notation from Section 2, let $N:\mathbb{V}\to\mathbb{M}$ be defined by $N(w,m)=m$, for $(w,m)\in\mathbb{V}$. The map $\bar{N}:\bar{\mathbb{V}}\to\bar{\mathbb{M}}$ is defined analogously. Let $\beta=(\beta_i)_{i=1}^\infty$ be coordinate maps on $\mathbb{V}$ defined as $\beta_i(w,m)=w_i$. Analogous maps on $\bar{\mathbb{V}}$ are denoted again as $\beta=(\beta_i)_{i=1}^\infty$. Define $\mathcal{G}_t\doteq\sigma\{N((0,s]\times A),\beta_i(s):\ 0\le s\le t, A\in\mathcal{B}(\mathbb{X}), i\ge 1\}$. For $\theta>0$, denote by $\mathbb{P}_\theta$ the unique probability measure on $(\mathbb{V},\mathcal{B}(\mathbb{V}))$ such that under $\mathbb{P}_\theta$:

(1) $(\beta_i)_{i=1}^\infty$ is an i.i.d. family of standard Brownian motions.
(2) N is a PRM with intensity measure $\theta\nu_T$.
(3) $\{\beta_i(t), t\in[0,T]\}$, $\{N([0,t]\times A)-\theta t\nu(A), t\in[0,T]\}$ are $\mathcal{G}_t$-martingales for every $i\ge 1$, $A\in\mathcal{B}(\mathbb{X})$ such that $\nu(A)<\infty$.

Define $(\bar{\mathbb{P}},\{\bar{\mathcal{G}}_t\})$ on $(\bar{\mathbb{V}},\mathcal{B}(\bar{\mathbb{V}}))$ analogous to $(\mathbb{P}_\theta,\{\mathcal{G}_t\})$ by replacing $(N,\theta\nu_T)$ with $(\bar{N},\bar{\nu}_T)$ in the above. Throughout we will consider the $\bar{\mathbb{P}}$-completion of the filtration $\{\bar{\mathcal{G}}_t\}$ and denote it by $\{\bar{\mathcal{F}}_t\}$. We denote by $\bar{\mathcal{P}}$ the predictable σ-field on $[0,T]\times\bar{\mathbb{V}}$ with the filtration $\{\bar{\mathcal{F}}_t:\ 0\le t\le T\}$ on $(\bar{\mathbb{V}},\mathcal{B}(\bar{\mathbb{V}}))$. Let $\bar{\mathcal{A}}$ be the class of all $(\bar{\mathcal{P}}\otimes\mathcal{B}(\mathbb{X}))\backslash\mathcal{B}[0,\infty)$ measurable maps $\varphi:\mathbb{X}_T\times\bar{\mathbb{V}}\to[0,\infty)$. For $\varphi\in\bar{\mathcal{A}}$, define $L_T(\varphi)$ and the counting process N^φ on $\mathbb{X}_T$ as in Section 2.

We denote by ℓ_2 the Hilbert space of real sequences $a=(a_i)$ satisfying $\|a\|^2=\sum_{i=1}^\infty a_i^2<\infty$, with the usual inner product. Define

$$\mathcal{P}_2=\left\{\psi=(\psi_i)_{i=1}^\infty:\ \psi_i \text{ is } \bar{\mathcal{P}}\backslash\mathcal{B}(\mathbb{R}) \text{ measurable and } \int_0^T\|\psi(s)\|^2\,\mathrm{d}s<\infty,\ \text{a.s. } \bar{\mathbb{P}}\right\}$$

and set $\mathcal{U}=\mathcal{P}_2\times\bar{\mathcal{A}}$. For $\psi\in\mathcal{P}_2$ define $\tilde{L}_T(\psi)=\frac{1}{2}\int_0^T\|\psi(s)\|^2\,\mathrm{d}s$ and for $u=(\psi,\varphi)\in\mathcal{U}$, set $\bar{L}_T(u)=L_T(\varphi)+\tilde{L}_T(\psi)$. For $\psi\in\mathcal{P}_2$, let $\beta^\psi=(\beta_i^\psi)$ be defined as $\beta_i^\psi(t)=\beta_i(t)+\int_0^t\psi_i(s)\,\mathrm{d}s$, $t\in[0,T]$, $i\in\mathbb{N}$. The following is the main result of this section.

Theorem 3.1. *Let $F\in M_b(\mathbb{V})$. Then for $\theta>0$,*

$$-\log\mathbb{E}_\theta\big(\mathrm{e}^{-F(\beta,N)}\big)=-\log\bar{\mathbb{E}}\big(\mathrm{e}^{-F(\beta,N^\theta)}\big)=\inf_{u=(\psi,\varphi)\in\mathcal{U}}\bar{\mathbb{E}}\big[\theta\bar{L}_T(u)+F\big(\beta^{\sqrt{\theta}\psi},N^{\theta\varphi}\big)\big].$$

The proof of Theorem 3.1 is very similar to that of Theorem 2.1, however for the reader's convenience a sketch is included in the Appendix.

4. Application to large deviations

We now apply the representation obtained in Section 3 to prove a large deviation result. Let $\{\mathcal{G}^\epsilon\}_{\epsilon>0}$, be a family of measurable maps from $\bar{\mathbb{V}}$ to $\mathbb{U}$, where $\mathbb{U}$ is some Polish space. Let $\{Z^\epsilon\}_{\epsilon>0}$ be a collection of $\mathbb{U}$-valued random variables defined on $(\bar{\mathbb{V}},\mathcal{B}(\bar{\mathbb{V}}),\bar{\mathbb{P}})$ by

$$Z^\epsilon=\mathcal{G}^\epsilon\big(\sqrt{\epsilon}\beta,\epsilon N^{\epsilon^{-1}}\big).\tag{4.1}$$

We are interested in a large deviation principle for the family $\{Z^\epsilon\}_{\epsilon>0}$ as $\epsilon\to 0$. We begin with some notation. For $N\in\mathbb{N}$, let

$$\tilde{S}^N=\big\{f\in L^2\big([0,T]:\ \ell_2\big):\ \tilde{L}_T(f)\le N\big\}.$$

With the weak topology on the Hilbert space, $\tilde{S}^N$ is a compact subset of $L^2([0,T]:\ \ell_2)$. We will throughout consider $\tilde{S}^N$ endowed with this topology. Also, let

$$S^N=\big\{g:\ X_T\to[0,\infty):\ L_T(g)\le N\big\}.$$

A function $g \in S^N$ can be identified with a measure $\nu_T^g \in \mathbb{M}$, defined by $\nu_T^g(A) = \int_A g(s,x)\nu_T(\mathrm{d}s\,\mathrm{d}x)$, $A \in \mathcal{B}(\mathbb{X}_T)$. Recalling that convergence in $\mathbb{M}$ is essentially equivalent to weak convergence on compact subsets (see the Introduction), the superlinear growth of ℓ implies that $\{\nu_T^g\colon g \in S^N\}$ is a compact subset of $\mathbb{M}$. Throughout we consider the topology on S^N obtained through this identification which makes S^N a compact space. We let $\bar{S}^N = \tilde{S}^N \times S^N$ with the usual product topology. Let $\mathbb{S} = \bigcup_{N\geq 1} \bar{S}^N$ and let $\mathcal{U}^N$ be the space of $\bar{S}^N$-valued controls:

$$\mathcal{U}^N = \{u = (\psi, \varphi) \in \mathcal{U}\colon u(\omega) \in \bar{S}^N, \bar{\mathbb{P}}\text{-a.e. } \omega\}.$$

The following will be the main assumption in our large deviations result.

Condition 4.1. *There exists a measurable map $\mathcal{G}^0 : \mathbb{V} \to \mathbb{U}$ such that the following hold.*

(1) *For $N \in \mathbb{N}$ let $(f_n, g_n), (f, g) \in \bar{S}^N$ be such that $(f_n, g_n) \to (f, g)$. Then*

$$\mathcal{G}^0\left(\int_0^\cdot f_n(s)\,\mathrm{d}s, \nu_T^{g_n}\right) \to \mathcal{G}^0\left(\int_0^\cdot f(s)\,\mathrm{d}s, \nu_T^{g}\right).$$

(2) *For $N \in \mathbb{N}$ let $u_\epsilon = (\psi_\epsilon, \varphi_\epsilon)$, $u = (\psi, \varphi) \in \mathcal{U}^N$ be such that, as $\epsilon \to 0$, u_ϵ converges in distribution to u. Then*

$$\mathcal{G}^\epsilon\left(\sqrt{\epsilon}\beta + \int_0^\cdot \psi_\epsilon(s)\,\mathrm{d}s, \epsilon N^{\epsilon^{-1}\varphi_\epsilon}\right) \Rightarrow \mathcal{G}^0\left(\int_0^\cdot \psi(s)\,\mathrm{d}s, \nu_T^{\varphi}\right).$$

For $\phi \in \mathbb{U}$, define $\mathbb{S}_\phi = \{(f, g) \in \mathbb{S}\colon \phi = \mathcal{G}^0(\int_0^\cdot f(s)\,\mathrm{d}s, \nu_T^g)\}$. Let $I : \mathbb{U} \to [0, \infty]$ be defined by

$$I(\phi) = \inf_{q=(f,g)\in\mathbb{S}_\phi} \{\bar{L}_T(q)\}. \tag{4.2}$$

The following is the main result of this section. The proof is similar to that of Theorem 5 in [6] (see also [4]) and is therefore relegated to the Appendix.

Theorem 4.2. *For $\epsilon > 0$, let Z^ϵ be defined by (4.1) and suppose that Condition 4.1 holds. Then I is a rate function on $\mathbb{U}$ and the family $\{Z^\epsilon\}_{\epsilon>0}$ satisfies a large deviation principle with rate function I.*

4.1. *An example*: *Finite-dimensional jump diffusions*

As an application of Theorem 4.2, we consider small noise stochastic differential equations (SDE) of the form:

$$\begin{aligned} \mathrm{d}Z^\epsilon(t) &= b(t, Z^\epsilon(t))\,\mathrm{d}t + \sqrt{\epsilon}\sigma(t, Z^\epsilon(t))\,\mathrm{d}\beta(t) + \int_{\mathbb{X}} \gamma(t, Z^\epsilon(t-), x)(\epsilon N^{\epsilon^{-1}}(\mathrm{d}t\,\mathrm{d}x) - \nu_T(\mathrm{d}t\,\mathrm{d}x)), \\ Z^\epsilon(0) &= z \in \mathbb{R}^d. \end{aligned} \tag{4.3}$$

Here β is a d-dimensional standard BM and the coefficients b, σ and γ satisfy the following conditions:

(1) The maps b, σ and γ are bounded and measurable, from $[0, T] \times \mathbb{R}^d$ to $\mathbb{R}^d$, $[0, T] \times \mathbb{R}^d$ to $\mathbb{R}^{d\times d}$ and $[0, T] \times \mathbb{R}^d \times \mathbb{X}$ to $\mathbb{R}^d$, respectively.
(2) For some $L \in (0, \infty)$, we have for all $t \in [0, T]$, $x \in \mathbb{X}$ and $z, z' \in \mathbb{R}^d$

$$|b(t, z) - b(t, z')| + |\sigma(t, z) - \sigma(t, z')| + |\gamma(t, z, x) - \gamma(t, z', x)| \leq L|z - z'|.$$

(3) For some compact $K \subset \mathbb{X}$, $\gamma(t, z, x) = 0$ for all $(t, z, x) \in [0, T] \times \mathbb{R}^d \times K^c$.

Under these conditions, there is a unique strong solution of (4.3) – indeed the conditions can be substantially weakened, see, e.g., Theorem III.2.3.2 of [14].

Let $\mathbb{U} = D([0,T]: \mathbb{R}^d)$, i.e., the space of $\mathbb{R}^d$-valued, right-continuous functions with left limits and the usual Skorohod topology. Then the solution Z^ϵ of (4.3) is a $\mathbb{U}$-valued random variable. We will now prove a large deviation principle for the family $\{Z^\epsilon\}_{\epsilon>0}$ as $\epsilon \to 0$. For $q = (f,g) \in \mathbb{S}$, denote by $\xi = \xi_q \in \mathbb{U}$ the unique solution of the integral equation

$$\xi(t) = z + \int_{[0,t]} \left(b\big(s,\xi(s)\big) + \sigma\big(s,\xi(s)\big) f(s) + \int_{\mathbb{X}} \gamma\big(s,\xi(s),x\big)\big(g(s,x)-1\big)\nu(\mathrm{d}x)\right) \mathrm{d}s.$$

Let $\mathbb{I}: \mathbb{U} \to [0,\infty]$ be defined as

$$\mathbb{I}(\phi) = \inf_{q\in\mathbb{S}:\phi=\xi_q} \bar{L}_T(q).$$

Theorem 4.3. *The map $\mathbb{I}$ is a rate function on $\mathbb{U}$ and $\{Z^\epsilon\}_{\epsilon>0}$ satisfies a large deviation principle on $\mathbb{U}$ with rate function $\mathbb{I}$.*

Proof. Modifying the notation from Section 3, we denote by $\mathbb{V}$ the space $C([0,T]:\mathbb{R}^d) \times \mathbb{M}$. With obvious notational changes, Theorem 4.2 holds for β as in the current section (i.e., a d-dimensional Brownian motion rather than an infinite-dimensional Brownian motion). Since (4.3) has a unique strong solution, there is a measurable map $\mathcal{G}^\epsilon : \mathbb{V} \to \mathbb{U}$ such that $Z^\epsilon = \mathcal{G}^\epsilon(\sqrt{\epsilon}\beta, \epsilon N^{\epsilon^{-1}})$. We will now verify that $\mathcal{G}^\epsilon$ satisfies Condition 4.1. Define $\mathcal{G}^0 : \mathbb{V} \to \mathbb{U}$ as follows. If $(w,m) \in \mathbb{V}$ is of the form $(w,m) = (\int_0^\cdot f(s)\,\mathrm{d}s, \nu_g)$ for some $q = (f,g) \in \mathbb{S}$, we define

$$\mathcal{G}^0(w,m) = \mathcal{G}^0\left(\int_0^\cdot f(s)\,\mathrm{d}s, \nu_g\right) = \xi_q.$$

For all other $(w,m) \in \mathbb{V}$ we set $\mathcal{G}^0(w,m) = 0$. With this definition, $\mathbb{I} = I$, where I is as defined in (4.2). We now show that part (2) of Condition 4.1 holds with this choice of $\mathcal{G}^0$. The proof of part (1) is similar, and hence omitted. Fix $N \in \mathbb{N}$ and $u_\epsilon = (\psi_\epsilon, \varphi_\epsilon)$, $u = (\psi,\varphi) \in \mathcal{U}^N$ such that, as $\epsilon \to 0$, u_ϵ converges in distribution to u. Then $\tilde{Z}^\epsilon = \mathcal{G}^\epsilon(\sqrt{\epsilon}\beta + \int_0^\cdot \psi_\epsilon(s)\,\mathrm{d}s, \epsilon N^{\epsilon^{-1}\varphi_\epsilon})$ is the unique solution of the controlled SDE

$$\begin{aligned} \mathrm{d}\tilde{Z}^\epsilon(t) &= \big(b\big(t,\tilde{Z}^\epsilon(t)\big) + \sigma\big(t,\tilde{Z}^\epsilon(t)\big)\psi_\epsilon(t)\big)\,\mathrm{d}t + \sqrt{\epsilon}\sigma\big(t,\tilde{Z}^\epsilon(t)\big)\,\mathrm{d}\beta(t) \\ &\quad + \int_{\mathbb{X}} \gamma\big(t,\tilde{Z}^\epsilon(t-),x\big)\big(\epsilon N^{\epsilon^{-1}\varphi_\epsilon}(\mathrm{d}t\,\mathrm{d}x) - \nu_T(\mathrm{d}t\,\mathrm{d}x)\big), \\ \tilde{Z}^\epsilon(0) &= z. \end{aligned} \tag{4.4}$$

It is easily checked that $\{\tilde{Z}^\epsilon\}_{\epsilon>0}$ is a tight family of $\mathbb{U}$-valued random variables. Elementary martingale estimates show that

$$\sup_{0\le t\le T} \left| \int_{[0,t]\times\mathbb{X}} \gamma\big(s,\tilde{Z}^\epsilon(s-),x\big)\big(\epsilon N^{\epsilon^{-1}\varphi_\epsilon}(\mathrm{d}t\,\mathrm{d}x) - \varphi_\epsilon(t,x)\nu_T(\mathrm{d}t\,\mathrm{d}x)\big)\right| \to 0$$

and

$$\sup_{0\le t\le T} \sqrt{\epsilon}\left| \int_{[0,t]} \sigma\big(s,\tilde{Z}^\epsilon(s)\big)\,\mathrm{d}\beta(s)\right| \to 0$$

in $\bar{\mathbb{P}}$-probability, as $\epsilon \to 0$. Thus choosing a subsequence along which $(\tilde{Z}^\epsilon, \psi_\epsilon, \varphi_\epsilon)$ converges in distribution (as a sequence of $\mathbb{U} \times \tilde{S}^N \times S^N$-valued random variables) to $(\tilde{Z}, \tilde{\psi}, \tilde{\varphi})$ we have that $(\tilde{\psi}, \tilde{\varphi})$ has the same probability law as (ψ,φ) and, by using conditions (1)–(3) on the coefficients, it can be verified that $\tilde{Z}$ solves

$$\tilde{Z}(t) = z + \int_{[0,t]} \left(b\big(s,\tilde{Z}(s)\big) + \sigma\big(s,\tilde{Z}(s)\big)\tilde{\psi}(s) + \int_{\mathbb{X}} \gamma\big(s,\tilde{Z}(s),x\big)\big(\tilde{\varphi}(s,x)-1\big)\nu(\mathrm{d}x)\right) \mathrm{d}s.$$

This shows that $\tilde{Z} = \mathcal{G}^0(\int_0^{\cdot} \tilde{\psi}(s)\,\mathrm{d}s, \nu_T^{\tilde{\varphi}})$, and proves part (2) of Condition 4.1, i.e.,

$$\mathcal{G}^\epsilon\left(\sqrt{\epsilon}\beta + \int_0^{\cdot} \psi_\epsilon(s)\,\mathrm{d}s, \epsilon N^{\epsilon^{-1}\varphi_\epsilon}\right) \Rightarrow \mathcal{G}^0\left(\int_0^{\cdot} \psi(s)\,\mathrm{d}s, \nu_T^{\varphi}\right).$$

The result follows. □

Appendix

A.1. *Proof of Lemma* 2.5

We need to show that the distribution of N^1 under $\mathbb{Q}^{\tilde{\varphi}}$ is same as that of N^φ under $\bar{\mathbb{P}}$, and that the costs $L_T(\tilde{\varphi})$ under $\mathbb{Q}^{\tilde{\varphi}}$ and $L_T(\varphi)$ under $\bar{\mathbb{P}}$ are the same.

Let $\varphi \in \bar{\mathcal{A}}_s$ be represented as on the right-hand side of (2.3). We will need some notation to describe how measures on $[0,T] \times \mathbb{Y}$ are decomposed into parts on subintervals of the form $(t_{i-1}, t_i]$, and also how after some manipulation such quantities can be recombined. For $i = 1, \dots, \ell$ let $\mathbb{I}_i = (t_{i-1}, t_i]$ and let $\mathbb{Y}_i \doteq \mathbb{I}_i \times \mathbb{Y}$. Denote by $\bar{\mathbb{M}}_i$ the space of nonnegative σ-finite integer valued measures $\bar{m}_i$ on $(\mathbb{Y}_i, \mathcal{B}(\mathbb{Y}_i))$ that satisfy $m_i(\mathbb{I}_i \times K) < \infty$ for all compact $K \subset \mathbb{Y}$. Endow $\mathbb{M}_i$ with the weakest topology making the functions $m \longmapsto \langle f, m\rangle, m \in \mathbb{M}_i$ continuous, for every f in $C(\mathbb{I}_i \times \mathbb{Y})$ vanishing outside some compact subset of $\mathbb{Y}$. Denote by $\mathcal{M}_i$ the corresponding Borel σ-field. Let $\bar{N}_i$ be the $\bar{\mathbb{M}}_i$-valued random variable on $(\bar{\mathbb{M}}, \mathcal{B}(\bar{\mathbb{M}}))$ defined by $\bar{N}_i(A) = \bar{N}(A)$, $A \in \mathcal{B}(\mathbb{Y}_i)$. Also, define $\mathbb{J}_i \doteq [1/n, n]^{n_i}$, and the $\mathbb{J}_i$-valued random variable X_i by $X_i = (X_{i1}, \dots, X_{in_i})$. Let $\hat{\mathbb{M}} = \bar{\mathbb{M}}_1 \times \cdots \times \bar{\mathbb{M}}_\ell$, and define $\varpi : \hat{\mathbb{M}} \to \bar{\mathbb{M}}$ by $\varpi(\hat{m}) = m$ when

$$m(A \times B) = \sum_{i=1}^{q} m_i\big((A \cap \mathbb{I}_i) \times B\big), \quad \hat{m} = (\bar{m}_1, \dots, \bar{m}_\ell), \bar{m}_i \in \bar{\mathbb{M}}_i, B \in \mathcal{B}(\mathbb{Y}), A \in \mathcal{B}[0,T].$$

Thus ϖ concatenates the measures back together: $\varpi((\bar{N}_1, \dots, \bar{N}_\ell)) = \bar{N}$.

From the predictability properties of φ it follows that for $i = 2, \dots, \ell$ there are measurable maps $\xi_i : \bar{\mathbb{M}}_1 \times \cdots \times \bar{\mathbb{M}}_{i-1} \to \mathbb{J}_i$ such that

$$X_{ij}(\bar{m}) = \xi_{ij}\big(\bar{N}_1(\bar{m}), \dots, \bar{N}_{i-1}(\bar{m})\big), \quad \xi_i = (\xi_{i1}, \dots, \xi_{in_i}).$$

Also, for $i = 1$, $X_1 = \xi_1$ a.s.-$\bar{\mathbb{P}}$ for some fixed vector ξ_1 in $\mathbb{J}_1$. The construction of $\tilde{\varphi}$, which takes the same form as φ, is recursive. For $s \in \mathbb{I}_1$ we set $\tilde{\varphi}(s, x, \bar{m}) = \varphi(s, x, \bar{m})$. As we will see, if there were only one time interval we would be done, in that N^φ under $\bar{\mathbb{P}}$ and N^1 under $\mathbb{Q}^{\tilde{\varphi}}$ would have the same distribution, and the costs $L_T(\varphi)$ and $L_T(\tilde{\varphi})$ would obviously be the same. The definition on subsequent intervals will depend on maps $T_i : \bar{\mathbb{M}}_1 \times \cdots \times \bar{\mathbb{M}}_i \to \bar{\mathbb{M}}_i$ for $i = 1, \dots, \ell$, which must also be defined recursively.

Observe that under $\bar{\mathbb{P}}$, $\bar{m}_1$ has intensity $\mathrm{d}s \times \nu(\mathrm{d}x) \times \mathrm{d}r$. Under $\mathbb{Q}^{\tilde{\varphi}}$, regardless of the definition of $\tilde{\varphi}$ on later intervals, $\bar{m}_1$ has intensity

$$\mathrm{d}s \times \nu(\mathrm{d}x) \times \left[\sum_{j=1}^{n_1} \xi_{1j} 1_{E_{1j}}(x) 1_{(0,1]}(r) + 1_{(1,\infty)}(r)\right] \mathrm{d}r.$$

The task of T_1 is to "undo" the effect of the change of measure, so that under $\mathbb{Q}^{\tilde{\varphi}}$, $\tilde{m}_1 = T_1[\bar{m}_1]$ has intensity $\mathrm{d}s \times \nu(\mathrm{d}x) \times \mathrm{d}r$. This can be done by requiring that for any $j \in \{1, \dots, n_1\}$ and any Borel subsets $A \subset \mathbb{I}_1$, $B \subset E_{1j}$, $C_1 \subset [0, \xi_{1j}]$ and $C_2 \subset (\xi_{1j}, \infty)$,

$$\tilde{m}_1\big(A \times B \times [C_1 \cup C_2]\big) = \bar{m}_1\left(A \times B \times \left[\frac{1}{\xi_{1j}} C_1 \cup (C_2 - \xi_{1j} + 1)\right]\right).$$

The mapping T_1 can thus be viewed as a transformation on the underlying space $\mathbb{Y}_1$ on which m_1 is defined. An equivalent characterization of $\tilde{m}_1 = T_1(\bar{m}_1)$ that will be used below is that $\tilde{m}_1$ is the unique measure that satisfies

$$\begin{aligned}&\int_{\mathbb{Y}_1} \psi(s,x,r)\tilde{m}_1(\mathrm{d}s\,\mathrm{d}x\,\mathrm{d}r)\\&\quad=\sum_{j=1}^{n_1}\int_{\mathbb{Y}_1} 1_{E_{1j}}(x)\big(\psi(s,x,\xi_{1j}r)1_{(0,1]}(r)+\psi(s,x,r+\xi_{1j}-1)1_{(1,\infty)}(r)\big)\bar{m}_1(\mathrm{d}s\,\mathrm{d}x\,\mathrm{d}r)\end{aligned}$$

for all nonnegative $\psi \in M_b(\mathbb{Y}_i)$.

With T_1 in hand the definition of $\tilde{\varphi}(s,x,\bar{m})$ for $s \in \mathbb{I}_2$ is straightforward. Indeed, since $\tilde{m}_1$ has the same distribution under $\mathbb{Q}^{\tilde{\varphi}}$ that $\bar{m}_1$ has under $\bar{\mathbb{P}}$, and since each ξ_{2j} is a function only of $\bar{N}_1(\bar{m}) = \bar{m}_1$, with the definition $\tilde{\xi}_{2j}(\bar{m}) = \xi_{2j}(T_1[\bar{m}_1])$, $\tilde{\xi}_{2j}(\bar{m})$ under $\mathbb{Q}^{\tilde{\varphi}}$ has the same distribution as $\xi_{2j}(\bar{m})$ under $\bar{\mathbb{P}}$. The sets E_{2j} are used as in (2.3) to define $\tilde{\varphi}$ on $\mathbb{I}_1 \cup \mathbb{I}_2$, so that $\{\tilde{\varphi}(s,x,\bar{m}), s \in \mathbb{I}_1 \cup \mathbb{I}_2, x \in \mathbb{X}\}$ under $\mathbb{Q}^{\tilde{\varphi}}$ has the same distribution as $\{\varphi(s,x,\bar{m}), s \in \mathbb{I}_1 \cup \mathbb{I}_2, x \in \mathbb{X}\}$ under $\bar{\mathbb{P}}$.

We now proceed recursively, and having defined $T_1,\dots,T_{p-1}$ for some $1 < p \le \ell$, we define T_p by $T_p(\bar{m}_1,\dots,\bar{m}_p) = \tilde{m}_p$, where $\tilde{m}_p$ is the unique measure satisfying, for all nonnegative $\psi \in M_b(\mathbb{Y}_p)$,

$$\begin{aligned}&\int_{\mathbb{Y}_p} \psi(s,x,r)\tilde{m}_p(\mathrm{d}s\,\mathrm{d}x\,\mathrm{d}r)\\&\quad=\sum_{j=1}^{n_p}\int_{\mathbb{Y}_p} 1_{E_{pj}}(x)\big(\psi(s,x,\tilde{\xi}_{pj}r)1_{(0,1]}(r)+\psi(s,x,r+\tilde{\xi}_{pj}-1)1_{(1,\infty)}(r)\big)\bar{m}_p(\mathrm{d}s\,\mathrm{d}x\,\mathrm{d}r),\end{aligned}$$

where $\tilde{\xi}_p = \xi_p(\tilde{m}_1,\dots,\tilde{m}_{p-1})$ and $\tilde{m}_i = T_i(\bar{m}_1,\dots,\bar{m}_i)$. We define the transformation $T : \bar{\mathbb{M}} \to \bar{\mathbb{M}}$ by

$$T(\bar{m}) = \varpi\big(T_1\big(\bar{N}_1(\bar{m})\big),\dots,T_\ell\big(\bar{N}_1(\bar{m}),\dots,\bar{N}_\ell(\bar{m})\big)\big),$$

and define $\tilde{\varphi} \in \bar{\mathcal{A}}_s$ for all times s by replacing X_{ij} with $\tilde{X}_{ij}$ in the right-hand side of (2.3), where

$$\tilde{X}_i(\bar{m}) = X_i\big(T(\bar{m})\big) = \xi_i\big(T_1\big(\bar{N}_1(\bar{m})\big),\dots,T_i\big(\bar{N}_1(\bar{m}),\dots,\bar{N}_i(\bar{m})\big)\big). \tag{A.1}$$

Denoting $T(\bar{N})$ by $\tilde{N}$ we see that for $\vartheta \in \hat{A}_b(\mathbb{Y}_T)$

$$\begin{aligned}&\int \vartheta(s,x,r)\tilde{N}(\mathrm{d}s\,\mathrm{d}x\,\mathrm{d}r)\\&\quad=\int \big(\vartheta\big(s,x,\tilde{\varphi}(s,x)r\big)1_{(0,1]}(r)+\vartheta\big(s,x,r+\tilde{\varphi}(s,x)-1\big)1_{(1,\infty)}(r)\big)\bar{N}(\mathrm{d}s\,\mathrm{d}x\,\mathrm{d}r).\end{aligned} \tag{A.2}$$

Also, let $h_\varphi : \bar{\mathbb{M}} \to \mathbb{M}$ be defined by

$$h_\varphi(\bar{m})(A \times B) = \sum_{i=1}^{\ell}\sum_{j=1}^{n_i} \bar{m}\big((A \cap \mathbb{I}_i) \times (B \cap E_{ij}) \times \big[0, X_{ij}(\bar{m})\big]\big)$$

for $A \times B \in \mathcal{B}(\mathbb{X}_T)$. We want to show that the distribution of N^1 under $\mathbb{Q}^{\tilde{\varphi}}$ is same as that of N^φ under $\bar{\mathbb{P}}$, and that the costs $L_T(\tilde{\varphi})$ under $\mathbb{Q}^{\tilde{\varphi}}$ and $L_T(\varphi)$ under $\bar{\mathbb{P}}$ are the same. To do this, we will prove the following:

(1) The distribution of $T(\bar{N})$ under $\mathbb{Q}^{\tilde{\varphi}}$ is same as that of $\bar{N}$ under $\bar{\mathbb{P}}$.
(2) $h_\varphi(\bar{N}) = N^\varphi$ and $h_\varphi(T(\bar{N})) = h_\varphi(\tilde{N}) = N^1$.
(3) For some measurable map $\Theta : \bar{\mathbb{M}} \to [0,\infty)$, $L_T(\varphi) = \Theta(\bar{N})$ and $L_T(\tilde{\varphi}) = \Theta(T(\bar{N}))$, a.s. $\bar{\mathbb{P}}$.

Item (3) is an immediate consequence of the definition of $\tilde{\varphi}$ via (A.1). We next consider (2). Noting that $\bar{N}(\bar{m}) = \bar{m}$ (and suppressing $\bar{m}$ in notation), we have for $A \times B \in \mathcal{B}(\mathbb{X}_T)$,

$$\begin{aligned}
h_\varphi(\bar{N})(A \times B) &= \sum_{i=1}^{\ell} \sum_{j=1}^{n_i} \bar{N}\big((A \cap \mathbb{I}_i) \times (B \cap E_{ij}) \times [0, X_{ij}]\big) \\
&= \sum_{i=1}^{\ell} \int_{(t_{i-1}, t_i] \times \mathbb{X}} \int_{[0,\infty)} 1_{A \times B}(s,x) 1_{[0,\varphi(s,x)]}(r) \bar{N}(\mathrm{d}s\,\mathrm{d}x\,\mathrm{d}r) \\
&= \int_{\mathbb{Y}_T} 1_{A \times B}(s,x) 1_{[0,\varphi(s,x)]}(r) \bar{N}(\mathrm{d}s\,\mathrm{d}x\,\mathrm{d}r) \\
&= N^\varphi(A \times B).
\end{aligned}$$

This proves the first statement in (2). Next, using (A.1), (A.2), and the observation that $r > 1$ implies $r + \tilde{\varphi}(s,x) - 1 > \tilde{\varphi}(s,x)$,

$$\begin{aligned}
h_\varphi\big(T(\bar{N})\big)(A \times B) &= \int 1_{A \times B}(s,x) 1_{[0,\tilde{\varphi}(s,x)]}(r) \tilde{N}(\mathrm{d}s\,\mathrm{d}x\,\mathrm{d}r) \\
&= \int 1_{A \times B}(s,x) \big(1_{[0,\tilde{\varphi}(s,x)]}\big(\tilde{\varphi}(s,x) r\big) 1_{[0,1]}(r) \\
&\quad + 1_{[0,\tilde{\varphi}(s,x)]}\big(r + \tilde{\varphi}(s,x) - 1\big) 1_{(1,\infty)}(r)\big) \bar{N}(\mathrm{d}s\,\mathrm{d}x\,\mathrm{d}r) \\
&= \int 1_{A \times B}(s,x) 1_{[0,1]}(r) \bar{N}(\mathrm{d}s\,\mathrm{d}x\,\mathrm{d}r) \\
&= N^1(A \times B).
\end{aligned}$$

This proves the second statement in (2). Lastly we prove (1). It suffices to show that for every $\vartheta \in \hat{A}_b$,

$$\mathbb{E}^{\mathbb{Q}^{\tilde{\varphi}}} \int \vartheta(s,x,r) \tilde{N}(\mathrm{d}s\,\mathrm{d}x\,\mathrm{d}r) = \mathbb{E}^{\mathbb{Q}^{\tilde{\varphi}}} \int \vartheta(s,x,r) \bar{\nu}_T(\mathrm{d}s\,\mathrm{d}x\,\mathrm{d}r).$$

Using (A.2) along with Lemma 2.3, we have that

$$\begin{aligned}
&\mathbb{E}^{\mathbb{Q}^{\tilde{\varphi}}} \int \vartheta(s,x,r) \tilde{N}(\mathrm{d}s\,\mathrm{d}x\,\mathrm{d}r) \\
&\quad = \mathbb{E}^{\mathbb{Q}^{\tilde{\varphi}}} \int \big(\vartheta\big(s,x,\tilde{\varphi}(s,x) r\big) \tilde{\varphi}(s,x) 1_{[0,1]}(r) \\
&\qquad + \vartheta\big(s,x,r + \tilde{\varphi}(s,x) - 1\big) 1_{(1,\infty)}(r)\big) \bar{\nu}_T(\mathrm{d}s\,\mathrm{d}x\,\mathrm{d}r) \\
&\quad = \mathbb{E}^{\mathbb{Q}^{\tilde{\varphi}}} \int \big(\vartheta(s,x,r) 1_{[0,\tilde{\varphi}(s,x)]}(r) + \vartheta(s,x,r) 1_{(\tilde{\varphi}(s,x),\infty)}(r)\big) \bar{\nu}_T(\mathrm{d}s\,\mathrm{d}x\,\mathrm{d}r) \\
&\quad = \mathbb{E}^{\mathbb{Q}^{\tilde{\varphi}}} \int \vartheta(s,x,r) \bar{\nu}_T(\mathrm{d}s\,\mathrm{d}x\,\mathrm{d}r),
\end{aligned}$$

which proves (1), and completes the proof of the first part of the lemma.

We now consider the second part. This requires that we start with a simple control $\tilde{\varphi}$ and associated measure $\mathbb{Q}^{\tilde{\varphi}}$, and construct a control φ for use with the original measure $\bar{\mathbb{P}}$. Let $\tilde{\varphi}$ take the form of the right-hand side of (2.3), but with the corresponding tilded quantities $\tilde{X}_{pj}$. As before, let $\tilde{\xi}_{pj}$ indicate the dependence of $\tilde{X}_{pj}$ on points in $\bar{\mathbb{M}}_1 \times \cdots \times \bar{\mathbb{M}}_p$. Define maps $\bar{T}_i : \bar{\mathbb{M}}_1 \times \cdots \times \bar{\mathbb{M}}_i \to \bar{\mathbb{M}}_i$ for $i = 1, \ldots, \ell$, recursively, as follows. Let $\bar{T}_1(\bar{m}_1) = \hat{m}_1$ be defined by

$$\int_{\mathbb{Y}_1} \psi(s,x,r)\hat{m}_1(\mathrm{d}s\,\mathrm{d}x\,\mathrm{d}r) = \sum_{j=1}^{n_1} \int_{\mathbb{Y}_1} 1_{E_{1j}}(x)\left(\psi\left(s,x,\frac{r}{\xi_{1j}}\right)1_{[0,\xi_{1j}]}(r) + \psi(s,x,r-\xi_{1j}+1)1_{(\xi_{1j},\infty)}(r)\right)\bar{m}_1(\mathrm{d}s\,\mathrm{d}x\,\mathrm{d}r)$$

for $\psi \in M_b(\mathbb{Y}_i)$. Having defined $\bar{T}_1,\dots,\bar{T}_{p-1}$ for some $1 < p \le \ell$, we now define $\bar{T}_p$ by $\bar{T}_p(\bar{m}_1,\dots,\bar{m}_p) = \hat{m}_p$ when

$$\int_{\mathbb{Y}_p} \psi(s,x,r)\hat{m}_p(\mathrm{d}s\,\mathrm{d}x\,\mathrm{d}r) = \sum_{j=1}^{n_p} \int_{\mathbb{Y}_p} 1_{E_{pj}}(x)\left(\psi\left(s,x,\frac{r}{\hat{\xi}_{pj}}\right)1_{[0,\hat{\xi}_{pj}]}(r) + \psi(s,x,r-\hat{\xi}_{pj}+1)1_{(\hat{\xi}_{pj},\infty)}(r)\right)\bar{m}_p(\mathrm{d}s\,\mathrm{d}x\,\mathrm{d}r),$$

where $\hat{\xi}_p = \tilde{\xi}_p(\hat{m}_1,\dots,\hat{m}_{p-1})$ and $\hat{m}_i = \bar{T}_i(\bar{m}_1,\dots,\bar{m}_i)$. We now define the transformation $\bar{T}:\bar{\mathbb{M}} \to \bar{\mathbb{M}}$ by the relation

$$\bar{T}(\bar{m}) = \varpi\big(\bar{T}_1\big(\bar{N}_1(\bar{m})\big),\dots,\bar{T}_\ell\big(\bar{N}_1(\bar{m}),\dots,\bar{N}_\ell(\bar{m})\big)\big).$$

Next, define $\varphi \in \bar{\mathcal{A}}_s$ by replacing X_{ij} with $\hat{X}_{ij}$ in the right-hand side of (2.3), where

$$\hat{X}_{ij}(\bar{m}) = \tilde{X}_{ij}\big(\bar{T}(\bar{m})\big) = \tilde{\xi}_i\big(\bar{T}_1\big(\bar{N}_1(\bar{m})\big),\dots,\bar{T}_i\big(\bar{N}_1(\bar{m}),\dots,\bar{N}_i(\bar{m})\big)\big).$$

Denoting $\bar{T}(\bar{N})$ by $\hat{N}$ we see that for $\vartheta \in \hat{\mathcal{A}}_b(\mathbb{Y}_T)$

$$\int \vartheta(s,x,r)\hat{N}(\mathrm{d}s\,\mathrm{d}x\,\mathrm{d}r) = \int\left(\vartheta\left(s,x,\frac{r}{\tilde{\varphi}(s,x)}\right)1_{[0,\tilde{\varphi}(s,x)]}(r) + \psi\big(s,x,r-\tilde{\varphi}(s,x)+1\big)1_{(\tilde{\varphi}(s,x),\infty)}(r)\right)\bar{N}(\mathrm{d}s\,\mathrm{d}x\,\mathrm{d}r).$$

Also, let $h_1:\bar{\mathbb{M}} \to \mathbb{M}$ be defined by

$$h_1(\bar{m})(A\times B) = \sum_{i=1}^{\ell}\sum_{j=1}^{n_i} \bar{m}\big((A\cap \mathbb{I}_i)\times(B\cap E_{ij})\times(0,1]\big)$$

for $A\times B \in \mathcal{B}(\mathbb{X}_T)$. Again, we need to show that the distribution of N^1 under $\mathbb{Q}^{\tilde{\varphi}}$ is same as that of N^{φ} under $\bar{\mathbb{P}}$, and that the costs $L_T(\tilde{\varphi})$ under $\mathbb{Q}^{\tilde{\varphi}}$ and $L_T(\varphi)$ under $\bar{\mathbb{P}}$ are the same. To do this, we show:

(1) The distribution of $\bar{N}$ under $\mathbb{Q}^{\tilde{\varphi}}$ is same as that of $\bar{T}(\bar{N})$ under $\bar{\mathbb{P}}$.
(2) $h_1(\bar{N}) = N^1$ and $h_1(\bar{T}(\bar{N})) = h_1(\hat{N}) = N^{\varphi}$.
(3) For some measurable map $\Theta:\bar{\mathbb{M}} \to [0,\infty)$, $L_T(\tilde{\varphi}) = \Theta(\bar{N})$ and $L_T(\varphi) = \Theta(\bar{T}(\bar{N}))$, a.s. $\bar{\mathbb{P}}$.

The proofs of items (2) and (3) are exactly the same as in the proof of the first part of the lemma. We now prove (1). Once again, following steps similar to the proof of the first part, it is easily seen that for every $\vartheta \in \hat{\mathcal{A}}_b$,

$$\mathbb{E}^{\mathbb{Q}^{\tilde{\varphi}}}\int \vartheta(s,x,r)\bar{N}(\mathrm{d}s\,\mathrm{d}x\,\mathrm{d}r) = \mathbb{E}^{\mathbb{Q}^{\tilde{\varphi}}}\int \vartheta(s,x,r)\big(\tilde{\varphi}(s,x)1_{(0,1]}(r) + 1_{(1,\infty)}(r)\big)\bar{\nu}_T(\mathrm{d}s\,\mathrm{d}x\,\mathrm{d}r), \tag{A.3}$$

and

$$\bar{\mathbb{E}}\int \vartheta(s,x,r)\hat{N}(\mathrm{d}s\,\mathrm{d}x\,\mathrm{d}r) = \bar{\mathbb{E}}\int \vartheta(s,x,r)\big(\varphi(s,x)1_{(0,1]}(r) + 1_{(1,\infty)}(r)\big)\bar{\nu}_T(\mathrm{d}s\,\mathrm{d}x\,\mathrm{d}r). \tag{A.4}$$

Item (1) is essentially a consequence of the above two relations, but we provide additional details. In order to prove (1), it suffices to establish that for every $q = 1, \dots, \ell$,

$$\text{distribution of } (\bar{N}_1, \dots, \bar{N}_q) \text{ under } \mathbb{Q}^{\tilde{\varphi}} \text{ equals that of } (\hat{N}_1, \dots, \hat{N}_q) \text{ under } \bar{\mathbb{P}}, \tag{A.5}$$

where $\hat{N}_q = \bar{T}_q(\bar{N}_1(\bar{m}), \dots, \bar{N}_q(\bar{m}))$ for $q = 1, \dots, \ell$. We proceed recursively. For each $i = 1, \dots, q$ and $y_i \in [1/n, n]^{n_i} \doteq \mathbb{J}_i$, we define a σ-finite measure $\nu_i^{y_i}$ on $(\mathbb{X}, \mathcal{B}(\mathbb{X}))$ as

$$\nu_i^{y_i}(A) = \nu\big(A \cap K_n^c\big) + \sum_{j=1}^{n_i} y_{ij}\nu(A \cap E_{ij}), \quad y_i = (y_{i1}, \dots, y_{in_i}), A \in \mathcal{B}(\mathbb{X}).$$

Define $\bar{\nu}_i^{y_i}$ on $(\mathbb{Y}, \mathcal{B}(\mathbb{Y}))$ as $\bar{\nu}_i^{y_i}(A \times B) = \nu_i^{y_i}(A)\lambda_\infty(B \cap (0,1]) + \lambda_\infty(B \cap (1,\infty))$. Denote by $\mu_i^{y_i}$ the unique probability measure on $(\bar{\mathbb{M}}_i, \mathcal{B}(\bar{\mathbb{M}}_i))$ under which the canonical map $N_i^* : \bar{\mathbb{M}}_i \to \bar{\mathbb{M}}_i$, $N_i^*(m) \doteq m, m \in \mathbb{M}_i$, is a Poisson random measure with intensity measure $\lambda_i \otimes \bar{\nu}_i^{y_i}$, where λ_i is the Lebesgue measure on $\mathbb{I}_i$.

From (A.3) and (A.4) it follows that $\bar{N}_1$ under $\mathbb{Q}^{\tilde{\varphi}}$ and $\hat{N}_1$ under $\bar{\mathbb{P}}$, are both distributed according to $\mu_1^{\xi_1}$. Suppose now that (A.5) holds with q replaced by $q-1$, for some $1 < q \le \ell$. Note that, under $\mathbb{Q}^{\tilde{\varphi}}$, using (A.3), the conditional law of $\bar{N}_q$ given $(\bar{N}_1, \dots, \bar{N}_{q-1})$ equals $\mu_q^{\bar{\xi}_q}$, where $\bar{\xi}_q = \xi_q(\bar{N}_1, \dots, \bar{N}_{q-1})$. Similarly, using (A.4), the conditional law, under $\bar{\mathbb{P}}$, of $\hat{N}_q$ given $(\hat{N}_1, \dots, \hat{N}_{q-1})$ equals $\mu_q^{\hat{\xi}_q}$, where $\hat{\xi}_q = \xi_q(\hat{N}_1, \dots, \hat{N}_{q-1})$. Combining these observations with our assumption that (A.5) holds with q replaced by $q-1$, we have that (A.5) holds with q as well. This completes the proof of (1) and the lemma follows.

A.2. *Sketch of Proof of Theorem* 3.1

As for the proof of Theorem 2.1, we will only consider the case $\theta = 1$. Define $\bar{\mathcal{A}}_b$, $\bar{\mathcal{A}}_s$, $\bar{\mathcal{A}}_{b,n}$, $\bar{\mathcal{A}}_{s,n}$, $\hat{\mathcal{A}}_b$ as in Section 2.2. Denote by $\mathcal{P}_2^{s,n}$ $[\mathcal{P}_2^{b,n}]$, the collection of ℓ_2-valued simple predictable (bounded predictable) processes ψ with $\|\psi(t)\| \le n$, a.s. $\bar{\mathbb{P}}$, for $t \in [0,T]$. Let $\mathcal{P}_2^s = \bigcup_{n\ge1}\mathcal{P}_2^{s,n}$ and $\mathcal{P}_2^b = \bigcup_{n\ge1}\mathcal{P}_2^{b,n}$. Set $\mathcal{U}_{s,n} = \bar{\mathcal{A}}_{s,n} \times \mathcal{P}_2^{s,n}$ and define $\mathcal{U}_s, \mathcal{U}_{b,n}, \mathcal{U}_b$ analogously. For $\varphi \in \bar{\mathcal{A}}_b$, let $\mathcal{E}_t(\varphi)$ be as defined in Lemma 2.3. For $\psi \in \mathcal{P}_2^b$, define the martingale

$$\tilde{\mathcal{E}}_t(\psi) = \exp\left\{\sum_{i=1}^\infty \int_0^t \psi_i(s)\,\mathrm{d}\beta_i(s) - \frac{1}{2}\int_0^t \big\|\psi(s)\big\|^2\,\mathrm{d}s\right\}, \quad t \in [0,T].$$

Finally, for $u = (\psi, \varphi) \in \mathcal{U}_b$, let $\bar{\mathcal{E}}_t(u) = \mathcal{E}_t(\varphi)\tilde{\mathcal{E}}_t(\psi)$, $t \in [0,T]$. The following result is standard, see Theorem III.3.24 of [14].

Lemma A.1. *Let $u \in \mathcal{U}_b$. Then $\{\bar{\mathcal{E}}_t(u)\}$ is an $\{\bar{\mathcal{F}}_t\}$-martingale. Define a probability measure $\mathbb{Q}^u$ on $\bar{\mathbb{V}}$ by*

$$\mathbb{Q}^u(G) = \int_G \bar{\mathcal{E}}_T(u)\,\mathrm{d}\bar{\mathbb{P}} \quad \textit{for } G \in \mathcal{B}(\bar{\mathbb{V}}).$$

Then for any $\vartheta \in \hat{A}_b$,

$$\mathbb{E}^{\mathbb{Q}^u}\int \vartheta(s,x,r)\bar{N}(\mathrm{d}s\,\mathrm{d}x\,\mathrm{d}r) = \mathbb{E}^{\mathbb{Q}^u}\int \vartheta(s,x,r)\big(\varphi(s,x)1_{(0,1]}(r) + 1_{(1,\infty)}(r)\big)\bar{\nu}_T(\mathrm{d}s\,\mathrm{d}x\,\mathrm{d}r).$$

Furthermore, under $\mathbb{Q}^u$, $\{\beta_i(t) - \int_0^t \psi_i(s)\,\mathrm{d}s, 0 \le t \le T\}$ is an i.i.d. sequence of standard Brownian motions.

The next two lemmas are proved in a manner similar to Lemmas 2.4 and 2.5. The proofs are omitted.

Lemma A.2. *Let $u = (\psi, \varphi) \in \mathcal{U}_{b,n}$. Then there exists a sequence of processes $u_k = (\psi_k, \varphi_k) \in \mathcal{U}_{s,n}$ such that, as $k \to \infty$:*

(1) $(\beta^{\psi_k}, N^{\varphi_k})$ *converges in distribution to* $(\beta^{\psi}, N^{\varphi})$.
(2) $\bar{\mathbb{E}}|L_T(\varphi_k) - L_T(\varphi)| \to 0$ *and* $\bar{\mathbb{E}}|\tilde{L}_T(\psi_k) - \tilde{L}_T(\psi)| \to 0$.
(3) $\bar{\mathbb{E}}|\bar{\mathcal{E}}_T(u_k) - \bar{\mathcal{E}}_T(u)| \to 0$, *as* $k \to \infty$.

Lemma A.3. *For every* $u = (\psi, \varphi) \in \mathcal{U}_s$, *there is* $\tilde{u} \in \mathcal{U}_s$ *such that* $\bar{\mathbb{P}} \circ (\beta^{\psi}, N^{\varphi})^{-1} = \mathbb{Q}^{\tilde{u}} \circ (\beta, N^1)^{-1}$ *and*

$$\mathbb{E}^{\mathbb{Q}^{\tilde{u}}}\big[\bar{L}_T(\tilde{u}) + F\big(\beta, N^1\big)\big] = \bar{\mathbb{E}}\big[\bar{L}_T(u) + F\big(\beta^{\psi}, N^{\varphi}\big)\big]. \tag{A.6}$$

Conversely, given any $\tilde{u} \in \mathcal{U}_s$ *there is* $u = (\psi, \varphi) \in \mathcal{U}_s$ *such that* $\bar{\mathbb{P}} \circ (\beta^{\psi}, N^{\varphi})^{-1} = \mathbb{Q}^{\tilde{u}} \circ (\beta, N^1)^{-1}$ *and* (A.6) *holds.*

Using the above three results we can now establish the following upper bound:

Theorem A.4. *For every* $F \in M_b(\mathbb{V})$

$$-\log \mathbb{E}\big(\mathrm{e}^{-F(\beta,N)}\big) \le \inf_{u=(\psi,\varphi)\in\mathcal{U}} \bar{\mathbb{E}}\big[\bar{L}_T(u) + F\big(\beta^{\psi}, N^{\varphi}\big)\big].$$

Sketch of Proof. Similar to the proof of Theorem 2.6, we prove that for any $u = (\psi, \varphi) \in \mathcal{U}$,

$$-\log \bar{\mathbb{E}}\big(\mathrm{e}^{-F(\beta,N^1)}\big) \le \bar{\mathbb{E}}\big[\bar{L}_T(u) + F\big(\beta^{\psi}, N^{\varphi}\big)\big] \tag{A.7}$$

by considering three cases. Proofs for cases (1) and (2) (i.e., $u \in \mathcal{U}_s$ and $u \in \mathcal{U}_b$) follow exactly as for Theorem 2.6 upon using Lemmas A.1–A.3 in place of Lemmas 2.3–2.5. In particular (cf. below (2.12)),

$$R\big(\bar{\mathbb{P}} \circ \big(\beta^{\psi}, N^{\varphi}\big)^{-1} \,\|\, \bar{\mathbb{P}} \circ \big(\beta, N^1\big)^{-1}\big) \le \bar{\mathbb{E}}\big[\bar{L}_T(u)\big] \quad \text{for } u = (\psi, \varphi) \in \mathcal{U}_b. \tag{A.8}$$

We provide additional details for the proof of case (3). Consider now $u = (\psi, \varphi) \in \mathcal{U}$. Without loss of generality we assume $\bar{\mathbb{E}}\bar{L}_T(u) < \infty$. Define φ_n as in case (3) of Theorem 2.6. Then $\varphi_n \in \bar{\mathcal{A}}_{b,n}$, and $\bar{\mathbb{E}}L_T(\varphi_n) \uparrow \bar{\mathbb{E}}L_T(\varphi)$. Next, let

$$\tau_n = \inf\left\{t \in [0, T]\colon\ \int_0^t \big\|\psi(s)\big\|^2 \,\mathrm{d}s \ge n\right\},$$

with the convention that $\tau_n = T$ if $\int_0^T \|\psi(s)\|^2 \,\mathrm{d}s < n$. Let $\psi_n(t) = \psi(t \wedge \tau_n)$. It is easily checked that $\bar{\mathbb{E}}\tilde{L}_T(\psi_n) \uparrow \bar{\mathbb{E}}\tilde{L}_T(\psi)$. Thus, in particular, the relative entropies $R(\bar{\mathbb{P}} \circ (\beta^{\psi_n}, N^{\varphi_n})^{-1} \,\|\, \bar{\mathbb{P}} \circ (\beta, N^1)^{-1})$ are uniformly bounded. Also, noting $\bar{\mathbb{E}}\int_{[0,T]} \|\psi_n(t) - \psi(t)\|^2 \,\mathrm{d}t \to 0$, we see that $\beta^{\psi_n} \to \beta^{\psi}$, in $\bar{\mathbb{P}}$-probability. Combining these observations with calculations similar to those below (2.14), we have that $\bar{\mathbb{E}}[F(\beta^{\psi_n}, N^{\varphi_n})] \to \bar{\mathbb{E}}[F(\beta^{\psi}, N^{\varphi})]$. The result now follows on recalling that (A.7) holds with $u = (\psi, \varphi)$ replaced by $u_n = (\psi_n, \varphi_n)$ and sending $n \to \infty$. □

For the proof of the lower bound, as in Section 2.3.2, we begin by considering a suitable class of "cylindrical" functions. For $\eta = (\eta_i) \in L^2([0, T]\colon\ \ell_2)$, let $\iota(\eta) : \mathbb{V} \to \mathbb{R}$ be defined as $\iota(\eta) = \sum_{i=1}^{\infty} \int_0^T \eta_i(s) \,\mathrm{d}\beta_i(s)$. Let

$$F(w, m) = g\big(\iota\big(\eta^1\big), \dots, \iota\big(\eta^p\big), \langle f_1, m\rangle, \dots, \langle f_k, m\rangle\big), \quad (w, m) \in \mathbb{V},$$

where $p, k \in \mathbb{N}$, $g \in C_c^{\infty}(\mathbb{R}^{p+k})$, $\eta^i \in L^2([0, T]\colon\ \ell_2)$ and $f_i \in C_c(\mathbb{X}_T)$. The class of all such F is denoted as $C_{\mathrm{cyl}}(\mathbb{V})$. Standard approximations show that, for every $F \in M_b(\mathbb{V})$,

$$\text{there is a sequence } F_n \in C_{\mathrm{cyl}}(\mathbb{V}) \text{ such that } |F_n|_{\infty} \le |F|_{\infty} \text{ and } F_n \to F \text{ a.s. } \bar{\mathbb{P}}. \tag{A.9}$$

By Proposition 1.4.2 of [11],

$$-\log \bar{\mathbb{E}}\big(\mathrm{e}^{-F(\beta,N^1)}\big) = R(\mathbb{Q} \,\|\, \bar{\mathbb{P}}) + \bar{\mathbb{E}}^{\mathbb{Q}}\big[F\big(\beta, h(\bar{N})\big)\big],$$

where $\mathbb{Q}$ is the probability measure defined by

$$\mathbb{Q}(A) = \frac{\int_A \mathrm{e}^{-F(w,h(\bar{m}))}\,\mathrm{d}\bar{\mathbb{P}}(w,\bar{m})}{\int_{\bar{\mathbb{V}}} \mathrm{e}^{-F(w,h(\bar{m}))}\,\mathrm{d}\bar{\mathbb{P}}(w,\bar{m})}, \quad A \in \mathcal{B}(\bar{\mathbb{V}}). \tag{A.10}$$

The key ingredient in the proof of the lower bound is the following representation for the Radon–Nikodym derivative $\frac{\mathrm{d}\mathbb{Q}}{\mathrm{d}\bar{\mathbb{P}}}$.

Theorem A.5. *Let $F \in C_{\mathrm{cyl}}(\mathbb{V})$ and $\mathbb{Q}$ be defined by (A.10). Then there is a $u \in \mathcal{U}_b$ such that $\frac{\mathrm{d}\mathbb{Q}}{\mathrm{d}\bar{\mathbb{P}}} = \bar{\mathcal{E}}_T(u)$.*

The proof of the statement that $\frac{\mathrm{d}\mathbb{Q}}{\mathrm{d}\bar{\mathbb{P}}} = \bar{\mathcal{E}}_T(u)$ for some $u \in \mathcal{U}$ follows from classical martingale representation results (see, e.g., Theorem 2 of [13]). The property that u can in fact be chosen to be an element of $\mathcal{U}_b$, when F is a smooth cylindrical function (i.e., F belongs to $C_{\mathrm{cyl}}(\mathbb{V})$), can be deduced using arguments similar to [28] (see Proposition 4.2 and Eq. (30) therein; see also [27], Theorem 3.4). Details are omitted.

Following the proof of (2.20) we have for $F \in C_{\mathrm{cyl}}(\mathbb{V})$ and $u \in \mathcal{U}_b$ as in Theorem A.5, writing $\mathbb{Q} = \mathbb{Q}^u$,

$$-\log \bar{\mathbb{E}}\big(\mathrm{e}^{-F(\beta,N^1)}\big) = \mathbb{E}^{\mathbb{Q}^u}\big[\bar{L}_T(u) + F\big(\beta, h(\bar{N})\big)\big]. \tag{A.11}$$

Let u_k be a sequence in $\mathcal{U}_s$ as constructed in Lemma A.2 for u as above. Then for each k

$$\left[R\big(\mathbb{Q}^{u_k} \,\|\, \bar{\mathbb{P}}\big) + \int_{\bar{\mathbb{V}}} F\big(w, h(\bar{m})\big)\,\mathrm{d}\mathbb{Q}^{u_k}(w,\bar{m})\right] = \mathbb{E}^{\mathbb{Q}^{u_k}}\big[\bar{L}_T(u_k) + F\big(\beta, h(\bar{N})\big)\big].$$

Using Lemma A.3 and Theorem A.5 we now obtain the following lower bound, whose proof is the same as that of Theorem 2.8.

Theorem A.6. *For every $F \in M_b(\mathbb{V})$*

$$-\log \mathbb{E}\big(\mathrm{e}^{-F(\beta,N)}\big) \geq \inf_{u=(\psi,\varphi)\in\mathcal{U}} \bar{\mathbb{E}}\big[\bar{L}_T(u) + F\big(\beta^{\psi}, N^{\varphi}\big)\big].$$

Theorem 3.1 is an immediate consequence of Theorems A.4 and A.6.

A.3. *Proof of Theorem 4.2*

In order to show that I is a rate function, it suffices to prove that for every $M \in (0,\infty)$, the set

$$\Lambda_M = \big\{\phi \in \mathbb{U}\colon\ I(\phi) \leq M\big\}$$

is compact. Part (1) of Condition 4.1 implies that for every $K \in (0,\infty)$ the set

$$\Gamma_K = \left\{\mathcal{G}^0\left(\int_0^{\cdot} f(s)\,\mathrm{d}s, \nu_T^g\right)\colon\ (f,g) \in \bar{S}^K\right\}$$

is compact. Compactness of Λ_M is now an immediate consequence of the identity $\Lambda_M = \bigcap_{n\geq 1} \Gamma_{M+1/n}$. This proves the first part of the theorem. For the second part, it suffices to show that for $F \in C_b(\mathbb{U})$

$$\lim_{\epsilon\to 0} -\epsilon \log \bar{\mathbb{E}}\big(\mathrm{e}^{-\epsilon^{-1}F(Z^\epsilon)}\big) = \inf_{\phi\in\mathbb{U}}\big\{I(\phi) + F(\phi)\big\}. \tag{A.12}$$

Lower bound: We begin by showing that the left-hand side (with lim replaced by liminf) of (A.12) is bounded below by the right-hand side. Note that

$$-\epsilon \log \bar{\mathbb{E}}\big(\mathrm{e}^{-\epsilon^{-1}F(Z^\epsilon)}\big) = -\epsilon \log \bar{\mathbb{E}}\big[\mathrm{e}^{-\epsilon^{-1}F\circ\mathcal{G}^\epsilon(\sqrt{\epsilon}\beta,\epsilon N^{\epsilon^{-1}})}\big].$$

Noting that $N^{\epsilon^{-1}}$ is a PRM with intensity $\epsilon^{-1}\nu_T$, we see from Theorem 3.1 (with $\theta = \epsilon^{-1}$) that the right-hand side above equals

$$\inf_{u=(\psi,\varphi)\in\mathcal{U}} \bar{\mathbb{E}}\left[\bar{L}_T(u) + F \circ \mathcal{G}^\epsilon\left(\sqrt{\epsilon}\beta + \int_0^{\cdot} \psi(s)\,\mathrm{d}s, \epsilon N^{\epsilon^{-1}\varphi}\right)\right]. \tag{A.13}$$

Fix $\delta \in (0,1)$. For each ϵ choose $u_\epsilon = (\psi_\epsilon, \varphi_\epsilon) \in \mathcal{U}$ such that the expression above is bounded below by

$$\bar{\mathbb{E}}\left[\bar{L}_T(u_\epsilon) + F \circ \mathcal{G}^\epsilon\left(\sqrt{\epsilon}\beta + \int_0^{\cdot} \psi_\epsilon(s)\,\mathrm{d}s, \epsilon N^{\epsilon^{-1}\varphi_\epsilon}\right)\right] - \delta. \tag{A.14}$$

Clearly $\bar{\mathbb{E}}(\bar{L}_T(u_\epsilon)) \le 2|F|_\infty + 1$. For $0 \le t \le T$ let

$$L_t(u_\epsilon) = \int_{[0,t]}\left(\|\psi_\epsilon(s)\|^2 + \int_{\mathbb{X}} \ell\big(\varphi_\epsilon(s,x)\big)\nu(\mathrm{d}x)\right)\mathrm{d}s$$

and

$$\tau_M^\epsilon = \inf\big\{t \in [0,T] : \bar{L}_t(u_\epsilon) \ge M\big\} \wedge T.$$

Define

$$\varphi_{\epsilon,M}(t,x) = 1 + \big[\varphi_\epsilon(t,x) - 1\big]1_{[0,\tau_M^\epsilon]}(t),\ \psi_{\epsilon,M}(t) = \psi_\epsilon(t)1_{[0,\tau_M^\epsilon]}(t), \quad t \in [0,T], x \in \mathbb{X}.$$

Note that $u_{\epsilon,M} = (\psi_{\epsilon,M}, \varphi_{\epsilon,M}) \in \mathcal{U}^M$. Also,

$$\bar{\mathbb{P}}(u_\epsilon \ne u_{\epsilon,M}) \le \bar{\mathbb{P}}\big(\bar{L}_T(u_\epsilon) \ge M\big) \le \frac{2|F|_\infty + 1}{M}.$$

Choose M large enough so that the right-hand side above is bounded by $\delta/(2|F|_\infty)$. Then the expression in (A.14) is bounded below by

$$\bar{\mathbb{E}}\left[\bar{L}_T(u_{\epsilon,M}) + F \circ \mathcal{G}^\epsilon\left(\sqrt{\epsilon}\beta + \int_0^{\cdot} \psi_{\epsilon,M}(s)\,\mathrm{d}s, \epsilon N^{\epsilon^{-1}\varphi_{\epsilon,M}}\right)\right] - 2\delta.$$

Note that $\{u_{\epsilon,M}\}_{\epsilon>0}$ is a family of $\bar{S}^M$-valued random variables. Recalling that $\bar{S}^M$ is compact, choose a weakly convergent subsequence and denote by $u = (\psi, \varphi)$ the weak limit point. From part (2) of Condition 4.1 we have that along this subsequence $\mathcal{G}^\epsilon(\sqrt{\epsilon}\beta + \int_0^{\cdot} \psi_{\epsilon,M}(s)\,\mathrm{d}s, \epsilon N^{\epsilon^{-1}\varphi_{\epsilon,M}})$ converges weakly to $\mathcal{G}^0(\int_0^{\cdot} \psi(s)\,\mathrm{d}s, \nu_T^\varphi)$. Thus, using Fatou's lemma and lower semicontinuity properties of $\bar{L}_T$

$$\begin{aligned}
&\liminf_{\epsilon\to 0} -\epsilon \log \bar{\mathbb{E}}\big(\mathrm{e}^{-\epsilon^{-1}F(Z^\epsilon)}\big) \\
&\quad\ge \liminf_{\epsilon\to 0} \bar{\mathbb{E}}\left[\bar{L}_T(u_\epsilon) + F \circ \mathcal{G}^\epsilon\left(\sqrt{\epsilon}\beta + \int_0^{\cdot} \psi_{\epsilon,M}(s)\,\mathrm{d}s, \epsilon N^{\epsilon^{-1}\varphi_{\epsilon,M}}\right)\right] - 2\delta \\
&\quad\ge \bar{\mathbb{E}}\left[\bar{L}_T(u) + F \circ \mathcal{G}^0\left(\int_0^{\cdot} \psi(s)\,\mathrm{d}s, \nu_T^\varphi\right)\right] - 2\delta \\
&\quad\ge \inf_{\phi\in\mathbb{U}} \inf_{q\in\mathbb{S}_\phi} \big(\bar{L}_T(q) + F(\phi)\big) - 2\delta \\
&\quad= \inf_{\phi\in\mathbb{U}} \big(I(\phi) + F(\phi)\big) - 2\delta.
\end{aligned}$$

Since $\delta \in (0,1)$ is arbitrary, this completes the proof of the lower bound.

Upper bound. We now show that the left-hand side in (A.12) (with lim replaced by limsup) is bounded above by the right-hand side. Fix $\delta \in (0,1)$ and $\phi_0 \in \mathbb{U}$ such that

$$I(\phi_0) + F(\phi_0) \le \inf_{\phi \in \mathbb{U}} \bigl(I(\phi) + F(\phi)\bigr) + \delta.$$

Choose $q = (f, g) \in \mathbb{S}_{\phi_0}$ such that $\bar{L}_T(q) \le I(\phi_0) + \delta$. Note that with this choice

$$\phi_0 = \mathcal{G}^0\biggl(\int_0^{\cdot} f(s)\,\mathrm{d}s, \nu_T^g\biggr).$$

Recalling from the proof of the lower bound that $-\epsilon \log \bar{\mathbb{E}}(\exp(-\epsilon^{-1} F(Z^\epsilon)))$ equals the expression in (A.13), we have that

$$\begin{aligned}
&\limsup_{\epsilon \to 0} -\epsilon \log \bar{\mathbb{E}}\bigl(\mathrm{e}^{-\epsilon^{-1} F(Z^\epsilon)}\bigr) \\
&\quad \le \bar{L}_T(q) + \limsup_{\epsilon \to 0} \bar{\mathbb{E}}\biggl[F \circ \mathcal{G}^\epsilon\biggl(\sqrt{\epsilon}\beta + \int_0^{\cdot} f(s)\,\mathrm{d}s, \epsilon N^{\epsilon^{-1} g}\biggr)\biggr] \\
&\quad \le I(\phi_0) + \delta + F \circ \mathcal{G}^0\biggl(\int_0^{\cdot} f(s)\,\mathrm{d}s, \nu_T^g\biggr) \\
&\quad = I(\phi_0) + F(\phi_0) + \delta \\
&\quad \le \inf_{\phi \in \mathbb{U}} \bigl(I(\phi) + F(\phi)\bigr) + 2\delta,
\end{aligned}$$

where the second inequality in the above display makes use of part (2) of Condition 4.1. Since $\delta \in (0,1)$ is arbitrary the desired upper bound follows. This completes the proof of the theorem.

References

[1] H. Bessaih and A. Millet. Large deviation principle and inviscid shell models. *Electron. J. Probab.* (2009) **14** 2551–2579. MR2570011

[2] M. Boue and P. Dupuis. A variational representation for certain functionals of Brownian motion. *Ann. Probab.* **26** (1998) 1641–1659. MR1675051

[3] M. Boué, P. Dupuis and R. S. Ellis. Large deviations for small noise diffusions with discontinuous statistics. *Probab. Theory Related Fields* **116** (2000) 125–149. MR1736592

[4] A. Budhiraja and P. Dupuis. A variational representation for positive functional of infinite dimensional Brownian motions. *Probab. Math. Statist.* **20** (2000) 39–61. MR1785237

[5] A. Budhiraja, P. Dupuis and M. Fischer. Large deviation properties of weakly interacting processes via weak convergence methods. *Ann. Probab.* To appear.

[6] A. Budhiraja, P. Dupuis and V. Maroulas. Large deviations for infinite dimensional stochastic dynamical systems. *Ann. Probab.* **36** (2008) 1390–1420. MR2435853

[7] A. Budhiraja, P. Dupuis and V. Maroulas. Large deviations for stochastic flows of diffeomorphisms. *Bernoulli* **36** (2010) 234–257. MR2648756

[8] I. Chueshov and A. Millet. Stochastic 2D hydrodynamical type systems: Well posedness and large deviations. *Appl. Math. Optim.* **61** (2010) 379–420. MR2609596

[9] A. Du, J. Duan and H. Gao. Small probability events for two-layer geophysical flows under uncertainty. Preprint.

[10] J. Duan and A. Millet. Large deviations for the Boussinesq equations under random influences. *Stochastic Process. Appl.* **119** (2009) 2052–2081. MR2519356

[11] P. Dupuis and R. Ellis. *A Weak Convergence Approach to the Theory of Large Deviations.* Wiley, New York, 1997. MR1431744

[12] N. Ikeda and S. Watanabe. *Stochastic Differential Equations and Diffusion Processes*. North-Holland, Amsterdam, 1981. MR1011252

[13] J. Jacod. A general theorem of representation for martingales. In *Proceedings of Symposia in Pure Mathematics* **31** 37–53. Amer. Math. Soc., Providence, RI, 1977. MR0443074

[14] J. Jacod and A. N. Shiryaev. *Limit Theorems for Stochastic Processes*. Springer, Berlin, 1987. MR0959133

[15] H. J. Kushner. Numerical methods for stochastic control problems in continuous time. *SIAM J. Control Optim.* **28** (1990) 999–1048. MR1064717

[16] H. J. Kushner and P. Dupuis. *Numerical Methods for Stochastic Control Problems in Continuous Time*, 2nd edition. Springer, New York, 2001. MR1800098

[17] W. Liu. Large deviations for stochastic evolution equations with small multiplicative noise. *Appl. Math. Optim.* **61** (2010) 27–56. MR2575313

[18] U. Manna, S. S. Sritharan and P. Sundar. Large deviations for the stochastic shell model of turbulence. *Nonlinear Differential Equations Appl.* **16** (2009) 493–521. MR2525514

[19] J. Ren and X. Zhang. Freidlin–Wentzell's large deviations for homeomorphism flows of non-Lipschitz SDEs. *Bull. Sci. Math.* **129** (2005) 643–655. MR2166732

[20] J. Ren and X. Zhang. Schilder theorem for the Brownian motion on the diffeomorphism group of the circle. *J. Funct. Anal.* **224** (2005) 107–133. MR2139106

[21] M. Rockner, T. Zhang and X. Zhang. Large deviations for stochastic tamed 3D Navier–Stokes equations. *Appl. Math. Optim.* **61** (2010) 267–285. MR2585144

[22] H. L. Royden. *Real Analysis*. Prentice Hall, Englewood Cliffs, NJ, 1988.

[23] S. S. Sritharan and P. Sundar. Large deviations for the two dimensional Navier–Stokes equations with multiplicative noise. *Stochastic Process. Appl.* **116** (2006) 1636–1659. MR2269220

[24] W. Wang and J. Duan. Reductions and deviations for stochastic partial differential equations under fast dynamical boundary conditions. *Stoch. Anal. Appl.* **27** (2009) 431–459. MR2523176

[25] D. Yang and Z. Hou. Large deviations for the stochastic derivative Ginzburg–Landau equation with multiplicative noise. *Phys. D* **237** (2008) 82–91. MR2450925

[26] X. Zhang. Euler schemes and large deviations for stochastic Volterra equations with singular kernels. *J. Differential Equations* **244** (2008) 2226–2250. MR2413840

[27] X. Zhang. A variational representation for random functionals on abstract Wiener spaces. *J. Math. Kyoto Univ.* **9** (2009) 475–490. MR2583599

[28] X. Zhang. Clark–Ocone formula and variational representation for Poisson functionals. *Ann. Probab.* **37** (2009) 506–529. MR2510015

[29] X. Zhang. Stochastic Volterra equations in Banach spaces and stochastic partial differential equations. *J. Funct. Anal.* **258** (2010) 1361–1425. MR2565842