ALEA, Lat. Am. J. Probab. Math. Stat. 22, 415–449 (2025)
DOI: 10.30757/ALEA.v22-15

Central limit theorem and moderate deviations for a class of semilinear stochastic partial differential equations in any space dimension

Dane Johnson and Leila Setayeshgar

Department of Mathematics and Statistics, Elon University, 100 Campus Drive, Elon, NC 27244, USA
E-mail address: djohnson52@elon.edu

Department of Mathematics, The College of Charleston, 175 Calhoun St., Charleston, SC 29466, USA
E-mail address: setayeshgarl@cofc.edu

Abstract. We prove a functional central limit theorem and a uniform moderate deviation principle for the law of the solutions to a class of parabolic semilinear stochastic partial differential equations, where the driving noise is a finite dimensional Wiener process. The space variable is of any dimension and the uniformity is with respect to initial conditions that are bounded and do not necessarily belong to a compact set. Our proof is based on the weak convergence method.

1. Introduction

We consider a family of parabolic, semilinear stochastic partial differential equations (SPDEs) indexed by $0 < \epsilon \leq 1$. The equation reads

$$\frac{\partial}{\partial t}u^{\epsilon,\xi}(t,x) = \left[\frac{\partial}{\partial x_i}\left(b_{ij}(x)\frac{\partial}{\partial x_j}u^{\epsilon,\xi}(t,x) + g_i\big(t,x,u^{\epsilon,\xi}(t,x)\big)\right) + f\big(t,x,u^{\epsilon,\xi}(t,x)\big)\right] + \sqrt{\epsilon}\sigma_j\big(t,x,u^{\epsilon,\xi}(t,x)\big)\frac{d}{dt}B_j,\ t \geq 0,\ x \in D, \tag{1.1}$$

supplemented with Dirichlet boundary conditions

$u^{\epsilon,\xi}(t,x) = 0,\ t \geq 0,\ x \in \partial D,$

and the initial condition

$u^{\epsilon,\xi}(0,x) = \xi(x),\ x \in D.$

Here $D \subset \mathbb{R}^d$ with $d \geq 1$ is a bounded convex domain with smooth boundary, ∂D, the driving noise $B_j := \{B_j(t), t \geq 0, j = 1, 2, \cdots, k\}$ is a k-dimensional Wiener process, and the space variable is of any dimension. The compromise for the space variable being of any dimension is the finite-dimensionality of the noise. In other words, it is the finite-dimensionality of the noise that allows for the space variable to be of any dimension. The initial condition, ξ, belongs to $L^\rho(D)$ for ρ

Received by the editors April 10th, 2024; accepted January 14th, 2025.
2010 *Mathematics Subject Classification.* Primary 60H15, 60H10, 60F05, 60F10; Secondary 37L55.
Key words and phrases. Moderate deviation principle, infinite-dimensional dynamical systems, weak convergence method, stochastic partial differential equations.

sufficiently large, and the small parameter $\epsilon \in (0,1]$ denotes the noise intensity. The functions $g_i := g_i(t,x,r)$, $i = 1,2,\cdots d$ are locally Lipschitz continuous in the third variable and have at most quadratic growth in $r \in \mathbb{R}$. The functions $f := f(t,x,r)$ and $\sigma_i := \sigma_i(t,x,r)$, $i = 1,2,\cdots k$ are locally Lipschitz continuous in the third variable and have linear growth in r. Additionally, we assume that $f := f(t,x,r)$ and $g_i := g_i(t,x,r)$, $i = 1,2,\cdots d$ are differentiable in the third variable, and that their derivatives are uniformly Lipschitz continuous in r. The existence and uniqueness of a probabilistically strong solution to Eq. (1.1) was proven by Gyöngy and Rovira (2000) via an approximation method. In Setayeshgar (2019), a large deviation principle (LDP) for the law of the solutions to Eq. (1.1) was proven, which was uniform over initial conditions on compact sets. In Salins (2019), a definition for an equicontinuous uniform Laplace principle (EULP) was introduced. It was then shown that the EULP is equivalent to the Freidlin and Wentzell's definition of the uniform large deviation principle (FWULDP). A sufficient condition under which a measurable function of an infinite-dimensional Wiener process satisfies an EULP was also provided with uniformity being over initial conditions that belong to bounded and not necessarily compact sets. In Salins and Setayeshgar (2023), the authors extended the work in Setayeshgar (2019), and relaxed the restriction on initial conditions over which the large deviations was uniform in that they established a uniform large deviation principle for the law of the solutions to Eq. (1.1), where uniformity was over bounded and not necessarily compact sets of initial conditions. Essentially, that work characterized the asymptotic behavior of probabilities of large deviations of the solution to the SPDE (i.e., $u^{\epsilon,\xi}$) away from its law of large numbers limit (i.e., $u^{0,\xi}$) as $\epsilon \to 0$.

In this work, we first consider the scaled and centered process, $v^{\epsilon,\xi} := \frac{1}{\sqrt{\epsilon}}(u^{\epsilon,\xi} - u^{0,\xi})$, and show that, $\{v^{\epsilon,\xi}\}_{\epsilon\in(0,1]}$ converges to the solution of a related SPDE as $\epsilon \to 0$, and prove the first result of our paper: a Functional Central Limit Theorem (FCLT). We then consider, $z^{\epsilon,\xi} := \frac{1}{\sqrt{\epsilon}\gamma(\epsilon)}(u^{\epsilon,\xi} - u^{0,\xi})$, where $\gamma(\epsilon) \to \infty$ and $\gamma(\epsilon)\sqrt{\epsilon} \to 0$ as $\epsilon \to 0$, and prove the second result of our paper: a Uniform Moderate Deviation Principle (UMDP). We do so by establishing that, $\{z^{\epsilon,\xi}\}_{\epsilon\in(0,1]}$ converges to 0 and satisfies a ULDP. Note that the ULDP that we prove is essentially a UMDP due to the scaling which is under.

Among the applications of large deviations is the design and analysis of accelerated Monte Carlo techniques such as importance sampling. Importance sampling algorithms employ a change of measure, which attempts to reduce the variance of the estimator. There exists a vast amount of literature exploring how the large deviation principle rate function may suggest effective changes of measure (see e.g. Dupuis and Wang, 2004; Vanden-Eijnden and Weare, 2012; Dupuis et al., 2015; Wang et al., 2015). As discussed in Dupuis and Johnson (2017), changes of measure based on the moderate deviation principle (MDP) rather than the large deviation principle are often easier to obtain due to the simplified form of the MDP rate function and may be effective when one is interested in trajectories that fall between the central limit theorem approximation and that of the LDP. Although most of the research currently present in the literature is concerned with finite-dimensional settings, there are some interesting infinite-dimensional examples. In Salins and Spiliopoulos (2017), the authors prove desirable asymptotic as well as pre-asymptotic performance under the large deviations scaling for importance sampling schemes used to estimate probabilities of escape from an attractor for solutions to a linear parabolic SPDE. The paper provides a useful discussion on the additional challenges posed by an infinite dimensional setting and the approach presented relies on the existence of a spectral gap, which allows the scheme to focus on a lower dimensional manifold, where escape is most likely to occur. In Gasteratos et al. (2024), the authors apply a related approach and achieve similar results for a semilinear SPDE, where they instead work with a moderate deviations scaling, which provides a linear approximation of the more complicated nonlinear dynamics in a neighborhood of the equilibrium. There are other examples of proving an MDP for small noise solutions to SPDEs (see e.g., Wang and Zhang, 2015; Gasteratos et al., 2023; Hu et al., 2020; Dong et al., 2017; Xiong and Zhang, 2021; Belfadli et al., 2019). Among them, our work is most similar to Hu et al. (2020), who also prove both an FCLT and an MDP using the weak convergence approach. However, there are significant differences between Hu et al. (2020) and what is carried out here. The method of proof in Hu et al. (2020) uses stopping times to bound the expected value of the stopped process, while here we avoid that approach and work with pathwise inequalities instead. These pathwise inequalities allow us to prove a sufficient condition under which a uniform moderate deviation principle with uniformity being over bounded sets of initial conditions that do not necessarily belong to compact sets, holds.

Lastly, we provide the motivation for our work as follows. As already outlined, large deviations analysis has many applications including but not limited to, the analysis of Monte Carlo and fast simulation techniques such as importance sampling. To the best of the authors' knowledge, construction of efficient importance sampling schemes for infinite dimensional models has not been carried out extensively in the literature and is thought to be expensive even for a single sample. This cost suggests that fast simulation techniques may be effective even for deviations that are only *moderately* away from the law of large numbers limit. For this reason, moderate deviations analysis – which is essentially large deviations but with a different deviation scaling – may be useful in constructing efficient importance sampling schemes for analyzing rare events arising in infinite dimensional systems.

The main contribution of this paper is twofold: a functional central theorem and a uniform moderate deviation principle for the law of the solutions to Eq. (1.1), where the uniformity is with respect to initial conditions that are bounded and do not necessarily belong to a compact set. Note that our proof to the latter verifies the sufficient condition of an EULP as opposed to that of Budhiraja et al. (2008), which cannot be used in this setting. Note that we also prove a uniform Laplace principle (ULP) with uniformity being over initial conditions on compact sets. We reiterate that our work has potential for constructing efficient importance sampling schemes capable of identifying the probability of rare events occurring in infinite dimensional settings. Finally, we mention that it would be interesting to study Eq. (1.1) with the white-in-time noise replaced by that of white-in-time, colored-in-space, where the injection of coloring will allow the process to gain better regularity. In such case, first it must be proven that under appropriate conditions on the noise correlation structure, a unique solution to Eq. (1.1) exists, that is strong in the sense of probability theory and further depends continuously on initial data. The existence and uniqueness of solutions often depends upon the conditions on the noise covariance function, e.g., decay and smoothness properties. The existence and uniqueness of a solution to Eq. (1.1) with a white-in-time, colored-in-space noise will be a prerequisite for carrying out other future work, such as a uniform large deviation principle, among others.

1.1. *Outline of the paper.* In Section 2 we state the assumptions and preliminaries that are necessary for the formulation of the problem. The existence and uniqueness results for the family of semilinear SPDEs is also asserted in this section. In section 3 we state and prove the first main result of our paper: a functional central theorem. Section 4 states the definition of the uniform large deviation principle as well as a sufficient condition under which the principle is said to hold. Section 5 introduces the moderate deviations scaling as well as the controlled and skeleton equations. The existence and uniqueness of solutions to the aforementioned equations is also established in this section. Section 6 states the second main result of our paper: a uniform moderate deviation principle, where uniformity is over bounded and not necessarily compact sets of initial conditions. Proving the UMDP requires a set of preliminary results, which are provided in Section 7. Finally, the proof of the UMDP is carried out in Section 8. A collection of results used in the proofs of our main theorems is presented in Appendix A.

1.2. *Notation.* For a spatial domain $D \subset \mathbb{R}^d$ and $\rho \in [1, \infty)$, let $L^\rho(D)$ be the set of functions $g : D \to \mathbb{R}$, such that $\int_D |g(x)|^\rho dx < +\infty$ endowed with the norm

$$|g|_\rho := \left(\int_D |g(x)|^\rho dx \right)^{\frac{1}{\rho}}. \tag{1.2}$$

For any $T > 0$ and $\rho \in [1, \infty)$, let $C\big([0,T] : L^\rho(D)\big)$ denote the space of functions $g : [0,T] \times D$, such that $t \mapsto g(t, \cdot)$ is continuous in the $L^\rho(D)$ norm. We endow this space with the norm

$$\big|g\big|_{C([0,T]:L^\rho(D))} := \sup_{t \in [0,T]} \big|g(t, \cdot)\big|_\rho.$$

Finally, for any metric space $(\mathcal{E}, \tau)$, we define the distance between an element $x \in E$ and a set $B \subset E$ by

$$\mathrm{dist}_{\mathcal{E}}(x, B) := \inf_{y \in B} \tau(x, y). \tag{1.3}$$

Unless otherwise noted, we adopt the following notation throughout the paper. The summation convention is in place. The notation := means by definition. C denotes a free positive constant, which may take on different values and depend upon other parameters. When the dependence of the constant C on a parameter is of significance, we denote it with a subscript, i.e., C_k will represent that C depends upon the parameter k. If $\theta \in \mathbb{R}^n$ for some $n \in \mathbb{N}$, then θ_i denotes the i^{th} component of that vector and $|\theta| := \big(\sum_{i=1}^n |\theta_i|^2\big)^{\frac{1}{2}}$. We use the notation $|h(t, \cdot)|_p := |h(t)|_p$ to denote the $L^p(D)$-norm of a function $h := h(t, x)$ with respect to the variable $x \in D$. Throughout the manuscript, we will work with inequalities and properties related to random fields that hold almost surely. To avoid repeatedly mentioning the foregoing, we state that these results are only implied to hold almost surely rather than everywhere. If $z := z(t, x)$ is a random field and E is a function space, then saying that z is almost surely in E means that z has a stochastic modification, which is in E almost surely. For ease of notation, we typically do not distinguish between a random field and a modification, which has the desired properties. For instance, if we say that $h(t, x) \in C\big([0,T] : L^\rho(D)\big)$, this means that there is a modification of $h(t, x)$, which is almost surely in $C\big([0,T] : L^\rho(D)\big)$. Occasionally, we will be more explicit about this in situations where we desire to exercise care. Depending on the situation, we may indicate the (t, x)-dependence in different ways to make the notation more convenient. For example, at times we write $f\big(t, x, u^{\epsilon,\xi}(t, x)\big)$ as $f\big(t, u^{\epsilon,\xi}(t)\big)(x)$ or $f\big(u^{\epsilon,\xi}\big)(t, x)$, and $z^\epsilon(t, x) \cdot f\big(t, x, u^{\epsilon,\xi}(t, x)\big)$ as $\Big(z^\epsilon f\big(u^{\epsilon,\xi}\big)\Big)(t, x)$.

2. Preliminaries

In this section we introduce a set of assumptions and preliminaries that are necessary for the formulation of the problem. Let $(\Omega, \mathcal{F}, \{\mathcal{F}_t\}_{t \in [0,T]}, P)$ be a filtered probability space or stochastic basis carrying a k-dimensional Brownian motion, $\{B_j(t), t \geq 0, 1 \leq j \leq k\}$. The following assumptions will be in effect throughout the paper.

(A1) The domain $D \subset \mathbb{R}^d,\ d \geq 1$ is a bounded convex set with smooth boundary.

(A2) The matrix $b_{ij}(x) \in C^2(\bar{D})$ is symmetric for every $x \in D$, and satisfies the uniform ellipticity condition, i.e., there exists a constant $\kappa > 0$, such that

$$\frac{1}{\kappa}|\theta|^2 \geq b_{ij}(x)\theta_i\theta_j \geq \kappa|\theta|^2,$$

for all $\theta \in \mathbb{R}^d$ and $x \in D$.

(A3) The functions g_i are of the form $g_i(t,x,r) := g_{i1}(x,t,r) + g_{i2}(t,r)$, where g_{i1} and g_{i2} are Borel functions of $(t,x,r) \in [0,T] \times D \times \mathbb{R}$ and of $(t,r) \in [0,T] \times \mathbb{R}$, respectively. Moreover, there exists a constant $K > 0$, such that

$$\big|g_{i1}(t,x,r)\big| \le K(1+|r|),\ \big|g_{i2}(t,r)\big| \le K(1+|r|^{\nu}),\ i \in \{1,...,d\},$$

for all $t \in [0,T]$, $x \in D$, $r \in \mathbb{R}$, with some $\nu \ge 1$.

(A4) The functions $f := f(t,x,r)$ and $\sigma_j := \sigma_j(t,x,r)$ are Borel functions and have linear growth in r, i.e., there exists a constant $L > 0$, such that

$$\sum_j \big|\sigma_j(t,x,r)\big| \le L(|r|+1),\ j \in \{1,...,k\},$$
$$\big|f(t,x,r)\big| \le L(|r|+1),$$

for all $t \in [0,T]$, $x \in D$, and $r \in \mathbb{R}$.

(A5) There exists a constant $L > 0$, such that

$$\sum_j \big|\sigma_j(t,x,r) - \sigma_j(t,x,s)\big| \le L(|r-s|),\ j \in \{1,...,k\},$$
$$\big|f(t,x,r) - f(t,x,s)\big| \le L|r-s|,$$
$$\big|g_i(t,x,r) - g_i(t,x,s)\big| \le L\big(1+|r|^{\nu-1}+|s|^{\nu-1}\big)|r-s|,\ i \in \{1,...,d\},$$

for all $t \in [0,T]$, $x \in D$, $r,s \in \mathbb{R}$.

(A6) The functions $f(x,t,r)$ and $g_i(x,t,r)$ are differentiable with respect to the third variable and their derivatives $f'(x,t,r)$ and $g_i'(x,t,r)$, are uniformly Lipschitz with respect to the last variable. That is, there exists a constant $K > 0$, such that

$$\big|f'(x,t,r) - f'(x,t,s)\big| \le K|r-s|,$$
$$\big|g_i'(x,t,r) - g_i'(x,t,s)\big| \le K|r-s|,\ i \in \{1,...,d\},$$

for all $t \in [0,T]$, $x \in D$, and $r,s \in \mathbb{R}$.
Observe that the addition of (A6) implies $\nu \le 2$ and $\big|f'(x,t,r)\big| \le L$ for all $t \in [0,T]$, $x \in D$, and $r \in \mathbb{R}$.

Definition 2.1 (Mild Solution)**.** A random field $u_\xi^\epsilon := \big\{u_\xi^\epsilon(t,x) : t \in [0,T], x \in D\big\}$ is called a mild solution of Eq. (1.1) with initial condition $\xi \in L^\rho(D)$, if $u_\xi^\epsilon(\cdot,\cdot)$ is $C\big([0,T] : L^\rho(D)\big)$-valued a.s., and $u_\xi^\epsilon(t,x)$ is $\{\mathcal{F}_t\}$-measurable for any $t \in [0,T]$, and $x \in D$, and if

$$\begin{aligned} u^{\epsilon,\xi}(t,x) &= \int_D G_t(x,y)\xi(y)dy + \sqrt{\epsilon}\int_0^t\int_D G_{t-s}(x,y)\sigma_j\big(s,u^{\epsilon,\xi}(s)\big)(y)dydB_j(s) \\ &\quad - \int_0^t\int_D \partial_{y_i}G_{t-s}(x,y)g_i\big(s,u^{\epsilon,\xi}(s)\big)(y)dyds \\ &\quad + \int_0^t\int_D G_{t-s}(x,y)f\big(s,u^{\epsilon,\xi}(s)\big)(y)dyds. \end{aligned} \tag{2.1}$$

The function $G_t(x,y),\ t \ge 0,\ x,y \in D$ is the Green kernel associated with the following linear equation

$$\frac{\partial}{\partial t}u(t,x) = \frac{\partial}{\partial x_i}\bigg(b_{ij}(x)\frac{\partial}{\partial x_j}u(t,x)\bigg),$$

with Dirichlet boundary condition

$$u(t,x) = 0,\ t \ge 0,\ x \in \partial D,$$

where $b_{ij} \in C^2(\bar{D})$, and ∂D is Lipschitz. We now state some estimates on the Green kernel that result from Proposition 3.5 of Gyöngy and Rovira (2000).

2.1. *Estimates on the Green kernel.* The following estimates on the Green kernel are results that follow from Proposition 3.5 of Gyöngy and Rovira (2000). The proof of that proposition uses a key estimate by Ladyženskaja et al. (1968). The below estimates on the kernel are used in proving the main results of the paper.

There exist Borel functions a, b, c and for $p \geq 1$, a corresponding nonnegative constant K_p, such that for all $0 \leq s < t \leq T$, $x, y \in D$, and $i \in \{1, ...d\}$, the following hold:

(E1) $\left|G_{t-s}(x,y)\right| \leq a(t-s, x-y),\ \left|a(t,\cdot)\right|_p \leq K_p t^{-\frac{d}{2}(\frac{p-1}{p})}$,

(E2) $\left|\frac{\partial}{\partial y_i} G_{t-s}(x,y)\right| \leq b(t-s, x-y),\ \left|b(t,\cdot)\right|_p \leq K_p t^{-\frac{d}{2}(\frac{p-1}{p})-\frac{1}{2}}$,

(E3) $\left|\frac{\partial}{\partial s} G_{t-s}(x,y)\right| \leq c(t-s, x-y),\ \left|c(t,\cdot)\right|_p \leq K_p t^{-\frac{d}{2}(\frac{p-1}{p})-1}$.

The following theorem (Gyöngy and Rovira, 2000, Theorem 2.1) asserts the existence of a unique solution to Eq. (1.1).

Theorem 2.2 (Existence & uniqueness of solution mapping)**.** *Assume the set of Hypotheses (A1)-(A5). Then, there exists $\rho_0(d,\nu)$, such that for every $\rho > \rho_0(d,\nu)$, Eq. (1.1) has a unique solution in $C\big([0,T], L^\rho(D)\big)$, provided ξ is an $\mathcal{F}_0$-measurable, $L^\rho(D)$-valued random element. Moreover, if ξ has a continuous stochastic modification, then $u^\epsilon_\xi(t,x)$ has a stochastic modification, which is continuous in $(t,x) \in [0,\infty) \times D$.*

Hereon, we assume $\rho < \infty$ is fixed and satisfies $\rho > \max\{\rho_0(d,\nu), d+2\}$. We further collect related results used in the remainder of the paper in Appendix A.

3. Functional Central Limit Theorem

In Setayeshgar (2019), a large deviation principle for the law of the solutions to Eq. (1.1) was proven, which was uniform over initial conditions on compact sets. In Salins and Setayeshgar (2023), the authors extended that work and relaxed the restriction on initial conditions over which the large deviations was uniform in that they proved a uniform large deviation principle for the law of the solutions to Eq. (1.1), where uniformity was over bounded and not necessarily compact sets of initial conditions. In this work, we center $u^{\epsilon,\xi}$ around its law of large numbers limit and consider the following scaling

$$v^{\epsilon,\xi} := \frac{1}{\sqrt{\epsilon}}(u^{\epsilon,\xi} - u^{0,\xi}), \tag{3.1}$$

to prove the first result of our paper: a functional central limit theorem. Note that for a.e. $\omega \in \Omega$, and for all $t \in [0,T]$, we have

$$\begin{aligned} v^{\epsilon,\xi}(x,t) = &- \int_0^t \int_D \partial_{y_i} G_{t-s}(x,y) \frac{1}{\sqrt{\epsilon}} \Big(g_i\big(u^{0,\xi} + \sqrt{\epsilon} v^{\epsilon,\xi}\big) - g_i\big(u^{0,\xi}\big)\Big)(s,y)dyds \\ &+ \int_0^t \int_D G_{t-s}(x,y) \frac{1}{\sqrt{\epsilon}} \Big(f\big(u^{0,\xi} + \sqrt{\epsilon} v^{\epsilon,\xi}\big) - f\big(u^{0,\xi}\big)\Big)(s,y)dyds \\ &+ \int_0^t \int_D G_{t-s}(x,y) \sigma_j\big(u^{0,\xi} + \sqrt{\epsilon} v^{\epsilon,\xi}\big)(s,y)dydB_j(s), \end{aligned} \tag{3.2}$$

for a.e. $x \in D$. The first result of our paper – a functional central limit theorem – shows that for any $\xi \in L^\rho(D)$, $\{v^{\epsilon,\xi}\}_{\epsilon\in(0,1]}$ converges in probability as $\epsilon \to 0$ to an $\mathcal{F}_t$-adapted random process,

$\hat{v}^\xi \in C([0,T] : L^\rho(D))$, which satisfies for a.e. $\omega \in \Omega$ and for all $t \in [0,T]$

$$\begin{aligned}\hat{v}^\xi(t,x) = &- \int_0^t \int_D \partial_{y_i} G_{t-s}(x,y)\big(g_i(u^{0,\xi})'\hat{v}^\xi\big)(s,y)dyds \\ &+ \int_0^t \int_D G_{t-s}(x,y)\big(f(u^{0,\xi})'\hat{v}^\xi\big)(s,y)dyds \\ &+ \int_0^t \int_D G_{t-s}(x,y)\sigma_j\big(u^{0,\xi}\big)(s,y)dydB_j(s), \end{aligned} \tag{3.3}$$

for a.e. $x \in D$. Note that because $\sup_{t\in[0,T],x\in D} |u^{0,\xi}(t,x)|$ is not necessarily finite, existence and uniqueness of solutions to (3.3) do not follow directly from Theorem 2.2. We are now in a position to state the functional central limit theorem.

Theorem 3.1 (Functional Central Limit Theorem)**.** *For any $\xi \in L^\rho(D)$, there exists a unique $\mathcal{F}_t$-adapted random process, $\hat{v}^\xi \in C([0,T] : L^\rho(D))$, satisfying (3.3) and $\{v^{\epsilon,\xi}\}_{\epsilon\in(0,1]}$ converges to $\hat{v}^\xi$ in probability in $C\big([0,T] : L^\rho(D)\big)$ as $\epsilon \to 0$.*

Next we state and prove two propositions below, which will subsequently be used in the proof of Theorem 3.1.

3.1. *Preliminary results required to prove Theorem 3.1.* We first establish that the sequence $\big\{\sup_{t\in[0,T]} |v^{\epsilon,\xi}(t)|_\rho\big\}_{\epsilon\in(0,1]}$ is bounded in probability, uniformly in $\epsilon \in (0,1]$.

Proposition 3.2. *For any $\xi \in L^\rho(D)$, the sequence $\Big\{\sup\limits_{t\in[0,T]} |v^{\epsilon,\xi}(t)|_\rho\Big\}_{\epsilon\in(0,1]}$ is bounded in probability, uniformly in $\epsilon \in (0,1]$.*

Proof: Let $\epsilon \in (0,1]$ be arbitrary. We have

$$\begin{aligned}\big|u^{\epsilon,\xi}(t) - u^{0,\xi}(t)\big|_\rho \le &\left| \int_0^t \int_D \partial_{y_i} G_{t-s}(\cdot,y)\Big[g_i\big(s,u^{\epsilon,\xi}(s)\big)(y) - g_i\big(s,u^{0,\xi}(s)\big)(y)\Big]dyds\right|_\rho \\ &+ \left|\int_0^t \int_D G_{t-s}(\cdot,y)\Big[f\big(s,u^{\epsilon,\xi}(s)\big)(y) - f\big(s,u^{0,\xi}(s)\big)(y)\Big]dyds\right|_\rho \\ &+ \sqrt{\epsilon}\left|\int_0^t \int_D G_{t-s}(\cdot,y)\sigma_j\big(s,u^{\epsilon,\xi}(s)\big)(y)dydB_j(s)\right|_\rho \\ &:= J_1^\epsilon(t) + J_2^\epsilon(t) + \sqrt{\epsilon}J_3^\epsilon(t),\end{aligned}$$

for all $t \in [0,T]$. Next define

$$\zeta_1^\epsilon := 1 + \sup_{t\in[0,T]} \big|u^{0,\xi}(s)\big|_\rho + \sup_{t\in[0,T]} \big|u^{\epsilon,\xi}(s)\big|_\rho,$$

and

$$\zeta_2^\epsilon := \sup_{t\in[0,T]} J_3^\epsilon(t).$$

Due to Lemma A.6, $\{\zeta_1^\epsilon\}_{\epsilon\in(0,1]}$ is bounded in probability, uniformly in $\epsilon \in (0,1]$, and due to Lemma A.5, combined with Lemma A.6, $\{\zeta_2^\epsilon\}_{\epsilon\in(0,1]}$ enjoys the same property.

Consider $J_1^\epsilon(t)$. Due to (A6), for all $i \in \{1, ..., d\}$, we have

$$\begin{aligned}
&\Big|g_i\big(s, u^{\epsilon,\xi}(s)\big)(y) - g_i\big(s, u^{0,\xi}(s)\big)(y)\Big| \\
&\leq \big|u^{\epsilon,\xi}(s,y) - u^{0,\xi}(s,y)\big| \int_0^1 \Big|g_i'\big(s, u^{0,\xi}(s)\big) + r\big(u^{\epsilon,\xi}(s) - u^{0,\xi}(s)\big)(y)\Big| dr \\
&\leq C\Bigg(1 + \big|u^{0,\xi}(s,y)\big| + \big|u^{\epsilon,\xi}(s,y)\big|\Bigg)\big|u^{\epsilon,\xi}(s,y) - u^{0,\xi}(s,y)\big|. \qquad (3.4)
\end{aligned}$$

Consequently

$$\begin{aligned}
J_1^\epsilon(t) &\leq \int_0^t \Bigg| \int_D \big|\partial_{y_i} G_{t-s}(\cdot, y)\big| \Big|g_i\big(s, u^{\epsilon,\xi}(s)\big)(y) - g_i\big(s, u^{0,\xi}(s)\big)(y)\Big| dy \Bigg|_\rho ds \\
&\leq \int_0^t \Big|b(t-s,\cdot)\Big|_{\frac{\rho}{\rho-1}} \Big|g_i\big(s, u^{\epsilon,\xi}(s)\big)(\cdot) - g_i\big(s, u^{0,\xi}(s)\big)(\cdot)\Big|_{\frac{\rho}{2}} ds \\
&\leq \sum_{i=1}^d C \int_0^t (t-s)^{-\frac{d}{2\rho}-\frac{1}{2}} \Big|g_i\big(s, u^{\epsilon,\xi}(s)\big)(\cdot) - g_i\big(s, u^{0,\xi}(s)\big)(\cdot)\Big|_{\frac{\rho}{2}} ds \\
&\leq C \int_0^t (t-s)^{-\frac{d}{2\rho}-\frac{1}{2}} \Bigg(1 + |u^{\epsilon,\xi}(s)|_\rho + |u^{0,\xi}(s)|_\rho\Bigg) \big|u^{\epsilon,\xi}(s) - u^{0,\xi}(s)\big|_\rho ds \\
&\leq C\zeta_1^\epsilon \int_0^t (t-s)^{-\frac{d}{2\rho}-\frac{1}{2}} \big|u^{\epsilon,\xi}(s) - u^{0,\xi}(s)\big|_\rho ds,
\end{aligned}$$

where line 1 used Minkowski's inequality, line 2 used (E2) and Young's inequality for convolutions with $p = \rho/(\rho-1)$, $q = \rho/2$, and $r = \rho$, where $1/p + 1/q = 1 + 1/r$ and $p, q, r \in \mathbb{R}_{\geq 1}$, line 3 used (E2), and line 4 used Hölder's inequality and (3.4).

Similarly

$$\begin{aligned}
J_2^\epsilon(t) &\leq \int_0^t \Bigg| \int_D G_{t-s}(\cdot, y)\Big[f\big(s, u^{\epsilon,\xi}(s)\big)(y) - f\big(s, u^{0,\xi}(s)\big)(y)\Big] dy \Bigg|_\rho ds \\
&\leq C \int_0^t \Bigg| \int_D \big|G_{t-s}(\cdot, y)\big| \Big|u^{\epsilon,\xi}(s,y) - u^{0,\xi}(s,y)\Big| dy \Bigg|_\rho ds \\
&\leq C \int_0^t \big|a(t-s,\cdot)\big|_1 \big|u^{\epsilon,\xi}(s) - u^{0,\xi}(s)\big|_\rho ds \\
&\leq C \int_0^t \big|u^{\epsilon,\xi}(s) - u^{0,\xi}(s)\big|_\rho ds,
\end{aligned}$$

where line 2 used (A5), line 3 used Young's inequality for convolutions with exponents $p = 1$, $q = \rho$, and $r = \rho$, and line 4 used (E1). Combining the foregoing implies that for all $t \in [0, T]$, we have

$$\big|u^{\epsilon,\xi}(t) - u^{0,\xi}(t)\big|_\rho \leq \sqrt{\epsilon}\zeta_2^\epsilon + C\int_0^t \Big((t-s)^{-\frac{d}{2\rho}-\frac{1}{2}}\zeta_1^\epsilon + 1\Big)\big|u^{\epsilon,\xi}(s) - u^{0,\xi}(s)\big|_\rho ds,$$

from which it follows by Grönwall's inequality that

$$\sup_{t\in[0,T]} \big|u^{\epsilon,\xi}(t) - u^{0,\xi}(t)\big|_\rho \leq \sqrt{\epsilon}\zeta_2^\epsilon e^{C\int_0^T \big((T-s)^{-\frac{d}{2\rho}-\frac{1}{2}}\zeta_1^\epsilon + 1\big) ds}.$$

Since $\epsilon > 0$ was arbitrary and both $\{\zeta_1^\epsilon\}_{\epsilon\in(0,1]}$ and $\{\zeta_2^\epsilon\}_{\epsilon\in[0,1]}$ are bounded in probability, uniformly in $\epsilon \in (0, 1]$, the result follows from (3.1). This completes the proof. □

The following Lemma will allow us to control the stochastic integral term in (3.2), which is also needed in the proof of Theorem 3.1.

Proposition 3.3. *For any* $\xi \in L^\rho(D)$

$$\sup_{t\in[0,T]}\left|\int_0^t\int_D G_{t-s}(\cdot,y)\Big(\sigma_j\big(s,u^{\epsilon,\xi}(s)\big)(y)-\sigma_j\big(s,u^{0,\xi}(s)\big)(y)\Big)dydB_j(s)\right|_\rho \to 0,$$

in probability as $\epsilon \to 0$.

Proof: Due to Lemma A.6 and (A5), we have

$$\sup_{\epsilon\in(0,1]}\left\{E\left[\sum_{j=1}^k \sup_{t\in[0,T]}\left|\sigma_j\big(t,u^{\epsilon,\xi}(t)\big)-\sigma_j\big(t,u^{0,\xi}(t)\big)\right|_\rho^\rho\right]\right\} \le C\big(1+|\xi|_\rho^\rho\big).$$

Therefore, Lemma A.5 implies that

$$\int_0^t\int_D G_{t-s}(\cdot,y)\Big(\sigma_j\big(s,u^{\epsilon,\xi}(s)\big)(y)-\sigma_j\big(s,u^{0,\xi}(s)\big)(y)\Big)dydB_j(s),$$

is tight in $C\big([0,T]:L^\rho(D)\big)$. Consequently, given any subsequence we can choose a further subsequence indexed by ϵ^*, that converges weakly in $C\big([0,T]:L^\rho(D)\big)$ to some limit $X^*(t)$. Since $\rho > d+2$, we can choose $p \in (d+2,\rho)$ and since D is bounded, this subsequence also converges weakly to $X^*(t)$ in $C\big([0,T]:L^p(D)\big)$. Therefore

$$\begin{aligned}\left(\sup_{t\in[0,T]}\left|u^{\epsilon^*,\xi}(t)-u^{0,\xi}(t)\right|_p^p\right)^{\frac{\rho}{p}} &\le \left(\big|D\big|^{\frac{\rho-p}{\rho}}\sup_{t\in[0,T]}\left|u^{\epsilon^*,\xi}(t)-u^{0,\xi}(t)\right|_\rho^p\right)^{\frac{\rho}{p}}\\ &\le C\left(\sup_{t\in[0,T]}\left|u^{\epsilon^*,\xi}(t)\right|_\rho^\rho+\sup_{t\in[0,T]}|u^{0,\xi}(t)|_\rho^\rho\right),\end{aligned}$$

where we have employed Hölder's inequality with exponents $\rho/(\rho-p)$ and ρ/p and Minkowski's inequality. This result combined with Lemma A.6, implies

$$\sup_{\epsilon^*}E\left[\left(\sup_{t\in[0,T]}\left|u^{\epsilon^*,\xi}(t)-u^{0,\xi}(t)\right|_p^p\right)^{\frac{\rho}{p}}\right] \le C\big(1+|\xi|_\rho^\rho\big).$$

Consequently, for any $K>0$, we have

$$\begin{aligned}&K^{\frac{\rho}{p}}\sup_{\epsilon^*}E\left[\sup_{t\in[0,T]}\left|u^{\epsilon^*,\xi}(t)-u^{0,\xi}(t)\right|_p^p\mathcal{I}_{\left\{\sup_{t\in[0,T]}|u^{\epsilon^*,\xi}(t)-u^{0,\xi}(t)|_p^p\ge K\right\}}\right]\\ &\le \sup_{\epsilon^*}E\left[\left(\sup_{t\in[0,T]}\left|u^{\epsilon^*,\xi}(t)-u^{0,\xi}(t)\right|_p^p\right)^{\frac{\rho}{p}}\right] \le C\big(1+|\xi|_\rho^\rho\big),\end{aligned}$$

which implies that

$$\left\{\sup_{t\in[0,T]}\left|u^{\epsilon^*,\xi}(t)-u^{0,\xi}(t)\right|_p^p\right\}_{\epsilon^*},$$

is uniformly integrable. This combined with (A5), implies

$$\lim_{\epsilon^*\to 0}E\left[\sum_{j=1}^k\sup_{t\in[0,T]}\left|\sigma_j\big(t,u^{\epsilon^*,\xi}(t)\big)-\sigma_j\big(t,u^{0,\xi}(t)\big)\right|_p^p\right]=0.$$

Combining this with Lemma A.2 part (a), gives

$$\lim_{\epsilon^*\to 0}E\left[\left|\int_0^t\int_D G_{t-s}(\cdot,y)\Big(\sigma_j\big(s,u^{\epsilon^*,\xi}(s)\big)(y)-\sigma_j\big(s,u^{0,\xi}(s)\big)(y)\Big)dydB_j(s)\right|_p^p\right]=0,$$

for all $t \in [0,T]$. It follows that $\sup_{t\in[0,T]} |X^*(t)|_p = 0$ for a.e. $\omega \in \Omega$ and consequently $\sup_{t\in[0,T]} |X^*(t)|_\rho = 0$ for a.e. $\omega \in \Omega$. Since the initial subsequence was arbitrary, it follows that

$$\sup_{t\in[0,T]} \left| \int_0^t \int_D G_{t-s}(\cdot,y)\Big(\sigma_j\big(s,u^{\epsilon,\xi}(s)\big)(y) - \sigma_j\big(s,u^{0,\xi}(s)\big)(y)\Big) dy dB_j(s) \right|_\rho \to 0,$$

in probability as $\epsilon \to 0$. This completes the proof. □

3.2. *Proof of Theorem 3.1*. For any $\mathcal{F}_t$-adapted random process $z \in C\big([0,T] : L^\rho(D)\big)$, define

$$L_1^\epsilon(z,t,x) := -\int_0^t \int_D \partial_{y_i} G_{t-s}(x,y) \frac{1}{\sqrt{\epsilon}} \Big(g_i\big(u^{0,\xi} + \sqrt{\epsilon} z\big) - g_i\big(u^{0,\xi}\big)\Big)(s,y) dy ds,$$

$$\tilde{L}_1(z,t,x) := -\int_0^t \int_D \partial_{y_i} G_{t-s}(x,y) \Big(z g_i'\big(u^{0,\xi}\big)\Big)(s,y) dy ds,$$

$$L_2^\epsilon(z,t,x) := \int_0^t \int_D G_{t-s}(x,y) \frac{1}{\sqrt{\epsilon}} \Big(f\big(u^{0,\xi} + \sqrt{\epsilon} z\big) - f\big(u^{0,\xi}\big)\Big)(s,y) dy ds,$$

$$\tilde{L}_2(z,t,x) := \int_0^t \int_D G_{t-s}(x,y) \Big(z f'\big(u^{0,\xi}\big)\Big)(s,y) dy ds,$$

and

$$\tilde{L}_3(t,x) := \int_0^t \int_D G_{t-s}(x,y) \Big(\sigma_j\big(u^{0,\xi}\big)\Big)(s,y) dy dB_j(s).$$

Due to (A6) and the Fundamental Theorem of Calculus, for all $x \in D$, $t \in [0,T]$, and $i \in \{1,...,d\}$, we have

$$\begin{aligned}
&\left| \left(\frac{1}{\sqrt{\epsilon}} \Big(g_i\big(u^{0,\xi} + \sqrt{\epsilon} z\big) - g_i\big(u^{0,\xi}\big)\Big) - z g_i'\big(u^{0,\xi}\big) \right)(t,x) \right| \\
&= \left| \left(\left(z \frac{g_i\big(u^{0,\xi} + r\sqrt{\epsilon} z\big)}{\sqrt{\epsilon} z} \Big|_0^1 \right) - z g_i'\big(u^{0,\xi}\big) \right)(t,x) \right| \\
&= \left| \left(z \int_0^1 g_i'\big(u^{0,\xi} + r\sqrt{\epsilon} z\big) dr - z g_i'\big(u^{0,\xi}\big) \right)(t,x) \right| \\
&\le C\sqrt{\epsilon} |z(t,x)|^2.
\end{aligned}$$

Similarly

$$\left| \left(\frac{1}{\sqrt{\epsilon}} \Big(f\big(u^{0,\xi} + \sqrt{\epsilon} z\big) - f\big(u^{0,\xi}\big)\Big) - z f'\big(u^{0,\xi}\big) \right)(t,x) \right| \le C\sqrt{\epsilon} \big|z(t,x)\big|^2.$$

Consequently

$$\Big| L_1^\epsilon(z,t,x) - \tilde{L}_1(z,t,x) \Big| \le C\sqrt{\epsilon} \int_0^t \sum_{i=1}^d \int_D \big|\partial_{y_i} G_{t-s}(x,y)\big| \big|z(s,y)\big|^2 dy ds,$$

and

$$\Big| L_2^\epsilon(z,t,x) - \tilde{L}_2(z,t,x) \Big| \le C\sqrt{\epsilon} \int_0^t \int_D \big|G_{t-s}(x,y)\big| \big|z(s,y)\big|^2 dy ds.$$

Using Minkowski's inequality, Young's inequality with (E2), and Hölder's inequality gives

$$\begin{aligned}\Big|L_1^\epsilon(z,t)-\tilde{L}_1(z,t)\Big|_\rho &\le C\sqrt{\epsilon}\sum_{i=1}^d\int_0^t\Big|\int_D\big|\partial_{y_i}G_{t-s}(\cdot,y)\big|\big|z(s,y)\big|^2dy\Big|_\rho ds\\ &\le C\sqrt{\epsilon}\int_0^t(t-s)^{-\frac{d}{2\rho}-\frac{1}{2}}\big|z(s)\big|_\rho^2ds\\ &\le C\sqrt{\epsilon}\sup_{t\in[0,T]}\big|z(t)\big|_\rho^2,\end{aligned}$$

and

$$\begin{aligned}\big|\tilde{L}_1(z,t)\big|_\rho &\le C\int_0^t\Big|\int_D\Big|\partial_{y_i}G_{t-s}(\cdot,y)\Big|\Big|\big(zg_i'(u^{0,\xi})\big)(s,y)\Big|dy\Big|_\rho ds\\ &\le C\sum_{i=1}^d\int_0^t(t-s)^{-\frac{d}{2\rho}-\frac{1}{2}}\Big|\big(zg_i'(u^{0,\xi})\big)(s,\cdot)\Big|_{\frac{\rho}{2}}dyds\\ &\le C\int_0^t(t-s)^{-\frac{d}{2\rho}-\frac{1}{2}}\Big(1+|u^{0,\xi}(s)|_\rho\Big)\big|z(s)\big|_\rho ds\\ &\le C\Big(1+|\xi|_\rho^\rho\Big)\int_0^t(t-s)^{-\frac{d}{2\rho}-\frac{1}{2}}\big|z(s)\big|_\rho ds,\end{aligned}$$

where the second to last line used (A6) and the last line used Lemma A.6. Similarly, using Minkowski's inequality, Young's inequality with (E1), (A6), Hölder's inequality, and Lemma A.6 gives

$$\begin{aligned}\Big|L_2^\epsilon(z,t)-\tilde{L}_2(z,t)\Big|_\rho &\le C\sqrt{\epsilon}\int_0^t(t-s)^{-\frac{d}{2\rho}}\big|z(s)\big|_\rho^2ds\\ &\le C\sqrt{\epsilon}\sup_{t\in[0,T]}|z(t)|_\rho^2,\end{aligned}$$

and

$$\big|\tilde{L}_2(z,t)\big|_\rho \le C\Big(1+|\xi|_\rho^\rho\Big)\int_0^t(t-s)^{-\frac{d}{2\rho}}\big|z(s)\big|_\rho ds.$$

For notational convenience, for all $\epsilon\in(0,1]$, define

$$\zeta_1^\epsilon := \sup_{t\in[0,T]}\left|\int_0^t\int_D G_{t-s}(\cdot,y)\Big(\sigma_j\big(u^{\epsilon,\xi}\big)-\sigma_j\big(u^{0,\xi}\big)\Big)(s,y)dydB_j(s)\right|_\rho,$$

and

$$\zeta_2^\epsilon := \sqrt{\epsilon}\sup_{t\in[0,T]}\Big|v^{\epsilon,\xi}(t)\Big|_\rho^2.$$

Note that due to Proposition 3.3, $\lim_{\epsilon\to 0}\zeta_1^\epsilon = 0$ in probability and due to Proposition 3.2, $\lim_{\epsilon\to0}\zeta_2^\epsilon = 0$ in probability. Recall (3.2) and note that for all $\epsilon_1,\epsilon_2\in(0,1]$ and $t\in[0,T]$, we have

$$\begin{aligned}
&\Big|v^{\epsilon_1,\xi}(t)-v^{\epsilon_2,\xi}(t)\Big|_\rho\\
&\le \Big|\tilde{L}_1(v^{\epsilon_1,\xi}-v^{\epsilon_2,\xi},t)\Big|_\rho+\Big|\tilde{L}_2\big(v^{\epsilon_1,\xi}-v^{\epsilon_2,\xi},t\big)\Big|_\rho+\Big|L_1^{\epsilon_1}\big(v^{\epsilon_1,\xi},t\big)-\tilde{L}_1\big(v^{\epsilon_1,\xi},t\big)\Big|_\rho+\zeta_1^{\epsilon_1}+\zeta_1^{\epsilon_2}+\\
&+\Big|L_1^{\epsilon_2,\xi}\big(v^{\epsilon_2,\xi},t\big)-\tilde{L}_1\big(v^{\epsilon_2,\xi},t\big)\Big|_\rho+\Big|L_2^{\epsilon_1}\big(v^{\epsilon_1,\xi},t\big)-\tilde{L}_2\big(v^{\epsilon_1,\xi},t\big)\Big|_\rho+\Big|L_2^{\epsilon_2}(v^{\epsilon_2,\xi},t)-\tilde{L}_2(v^{\epsilon_2,\xi},t)\Big|_\rho\\
&\le C\Big(1+|\xi|_\rho^\rho\Big)\int_0^t\Big((t-s)^{-\frac{d}{2\rho}-\frac{1}{2}}+(t-s)^{-\frac{d}{2\rho}}\Big)\Big|v^{\epsilon_1,\xi}(s)-v^{\epsilon_2,\xi}(s)\Big|_\rho ds\\
&+C\Big(\zeta_2^{\epsilon_1}+\zeta_2^{\epsilon_2}\Big)+\zeta_1^{\epsilon_1}+\zeta_1^{\epsilon_2}.
\end{aligned}$$

Consequently, Grönwall's inequality yields

$$\sup_{t\in[0,T]}\Big|v^{\epsilon_1,\xi}(t)-v^{\epsilon_2,\xi}(t)\Big|_\rho\le C\Big(\zeta_2^{\epsilon_1}+\zeta_2^{\epsilon_2}+\zeta_1^{\epsilon_1}+\zeta_1^{\epsilon_2}\Big)e^{C(1+|\xi|_\rho^\rho)}. \tag{3.5}$$

Next, define the sets

$$\Theta_n^\epsilon:=\Big\{C(\zeta_2^\epsilon+\zeta_1^\epsilon)e^{C(1+|\xi|_\rho^\rho)}\le 2^{-n-1}\Big\}. \tag{3.6}$$

Since both $\zeta_1^\epsilon\to 0$ and $\zeta_2^\epsilon\to 0$ in probability for any subsequence, we can choose a further subsequence indexed by $\{\bar{\epsilon}_n\}_{n\in\mathbb{N}}$, such that $\bar{\epsilon}_n\downarrow 0$, which satisfies

$$P\Big(\big(\Theta_n^{\bar{\epsilon}_n}\big)^c\Big)\le 2^{-n-1}, \tag{3.7}$$

for all $n\in\mathbb{N}$. For any $N\in\mathbb{N}$, note that on the set $\cap_{n\ge N}\Theta_n^{\bar{\epsilon}_n}$ due to (3.5), (3.6), and (3.7) for all $n,m\ge N$, we have

$$\sup_{t\in[0,T]}\Big|v^{\bar{\epsilon}_n,\xi}(t)-v^{\bar{\epsilon}_m,\xi}(t)\Big|_\rho\le 2^{-n-1}+2^{-m-1}\le 2^{-\min\{n,m\}}.$$

Since $C\big([0,T]:L^\rho(D)\big)$ is complete, this implies that $\{v^{\bar{\epsilon}_n,\xi}\}_{n\in\mathbb{N}}$ converges to a limit on the set $\cup_{N\in\mathbb{N}}\cap_{n\ge N}\Theta_n^{\bar{\epsilon}_n}$ and due to the Borel-Cantelli lemma $P\big(\cup_{N\in\mathbb{N}}\cap_{n\ge N}\Theta_n^{\bar{\epsilon}_n}\big)=1$. Define $\hat{v}^\xi:=\lim_{n\to\infty}v^{\bar{\epsilon}_n,\xi}$ in $C\big([0,T]:L^\rho(D)\big)$ and note that since $\{v^{\bar{\epsilon}_n,\xi}\}_{n\in\mathbb{N}}$ are $\mathcal{F}_t$-adapted, so is $\hat{v}^\xi$. We have for all $t\in[0,T]$ and $n\in\mathbb{N}$

$$\begin{aligned}
&\Big|\hat{v}^\xi(t)-\big(\tilde{L}_1(\hat{v}^\xi,t)+\tilde{L}_2(\hat{v}^\xi,t)+\tilde{L}_3(t)\big)\Big|_\rho\\
&\le\Big|\hat{v}^\xi(t)-v^{\bar{\epsilon}_n,\xi}(t)\Big|_\rho+\Big|L_1^{\bar{\epsilon}_n}\big(v^{\bar{\epsilon}_n,\xi},t\big)-\tilde{L}_1\big(v^{\bar{\epsilon}_n,\xi},t\big)\Big|_\rho+\Big|L_2^{\bar{\epsilon}_n}\big(v^{\bar{\epsilon}_n,\xi},t\big)-\tilde{L}_2\big(v^{\bar{\epsilon}_n,\xi},t\big)\Big|_\rho\\
&+\zeta_1^{\bar{\epsilon}_n}+\Big|\tilde{L}_1(v^{\bar{\epsilon}_n,\xi}-v,t)\Big|_\rho+\Big|\tilde{L}_2\big(v^{\bar{\epsilon}_n,\xi}-\hat{v}^\xi,t\big)\Big|_\rho\\
&\le C\big(\zeta_1^{\bar{\epsilon}_n}+\zeta_2^{\bar{\epsilon}_n}\big)+C\big(1+|\xi|_\rho^\rho\big)\sup_{t\in[0,T]}\Big|v^{\bar{\epsilon}_n,\xi}(t)-\hat{v}^\xi(t)\Big|_\rho.
\end{aligned}$$

Since with probability 1

$$\lim_{n\to\infty}\left(C\big(\zeta_1^{\bar{\epsilon}_n}+\zeta_2^{\bar{\epsilon}_n}\big)+C\big(1+|\xi|_\rho^\rho\big)\sup_{t\in[0,T]}\Big|v^{\bar{\epsilon}_n,\xi}(t)-\hat{v}^\xi(t)\Big|_\rho\right)=0,$$

it follows that for a.e. $\omega\in\Omega$

$$\sup_{t\in[0,T]}\left|\hat{v}^\xi(t)-\Big(\tilde{L}_1\big(\hat{v}^\xi,t\big)+\tilde{L}_2(\hat{v}^\xi,t)+\tilde{L}_3(t)\Big)\right|_\rho=0,$$

meaning $\hat{v}^\xi$ satisfies (3.3). Recall that the original subsequence from which we chose the further subsequence $\{v^{\bar{\epsilon}_n,\xi}\}_{n\in\mathbb{N}}$ that converged to $\hat{v}^\xi$ a.s. was arbitrary. Consequently, it follows that

$\{v^{\epsilon,\xi}\}_{\epsilon\in(0,1]}$ converges to $\hat{v}^\xi$ in probability in $C\big([0,T] : L^\rho(D)\big)$ as $\epsilon \to 0$. To show uniqueness, assume an $\mathcal{F}_t$-adapted random process, $\tilde{v}^\xi \in C\big([0,T] : L^\rho(D)\big)$, satisfies

$$\sup_{t\in[0,T]} \left|\tilde{v}^\xi(t) - \tilde{L}_1\big(\tilde{v}^\xi, t\big) + \tilde{L}_2(\tilde{v}^\xi, t) + \tilde{L}_3(t)\right|_\rho = 0,$$

for a.e. $\omega \in \Omega$. Then with probability 1, for all $t \in [0,T]$, we have

$$\begin{aligned}\left|\hat{v}^\xi(t) - \tilde{v}^\xi(t)\right|_\rho &= \left|\tilde{L}_1\big(\hat{v}^\xi - \tilde{v}^\xi, t\big)\right|_\rho + \left|\tilde{L}_2\big(\hat{v}^\xi - \tilde{v}^\xi, t\big)\right|_\rho \\ &\le C\big(1 + |\xi|_\rho^\rho\big) \int_0^t \left((t-s)^{-\frac{d}{2\rho}-\frac{1}{2}} + (t-s)^{-\frac{d}{2\rho}}\right)\left|\hat{v}^\xi(s) - \tilde{v}^\xi(s)\right|_\rho ds.\end{aligned}$$

Therefore, Grönwall's inequality implies that for a.e. $\omega \in \Omega$

$$\sup_{t\in[0,T]} \left|\hat{v}^\xi(t) - \tilde{v}^\xi(t)\right|_\rho = 0.$$

This completes the proof. □

4. Uniform Large Deviation Principle

In this section we recall Freidlin and Wentzell's definition of the uniform large deviation principle Freidlin and Wentzell (2012) as well as that of an equicontinuous uniform Laplace principle Salins (2019). We further assert that they are equivalent. We then provide a sufficient condition under which an EULP and therefore an FWULDP holds. Let $(\mathcal{E}, \tau)$ be a Polish space and $\mathcal{E}_0$ a set. For any $\xi \in \mathcal{E}_0$, let $I_\xi : \mathcal{E} \to [0, +\infty]$ be a so-called rate function, meaning that it is lower-semicontinuous. Let $\Gamma_\xi(s)$ denote the level set $\Gamma_\xi(s) := \{\varphi \in \mathcal{E} : I_\xi(\varphi) \le s\}$, and let $\mathscr{L}$ be a collection of subsets of $\mathcal{E}_0$. Further, let $\alpha(\varepsilon)$ be a function converging to zero as ε approaches zero itself.

Definition 4.1 (FWULDP)**.** A family of $\mathcal{E}$-valued random variables, $\{X_\xi^\varepsilon\}_{\varepsilon\in(0,1]}^{\xi\in\mathcal{E}_0}$, is said to satisfy an FWULDP with speed $\alpha(\varepsilon)$ and rate function I_ξ, uniformly over $\mathscr{L}$, if

(1) For any $L \in \mathscr{L}$, $s_0 \ge 0$, and $\delta > 0$

$$\liminf_{\varepsilon\to 0} \inf_{\xi\in L} \inf_{\varphi\in\Gamma_\xi(s_0)} \left(\alpha(\varepsilon) \log P\Big(\tau(X_\xi^\varepsilon, \varphi) < \delta\Big) + I_\xi(\varphi)\right) \ge 0.$$

(2) For any $L \in \mathscr{L}$, $s_0 \ge 0$, and $\delta > 0$

$$\limsup_{\varepsilon\to 0} \sup_{\xi\in L} \sup_{s\in[0,s_0]} \left(\alpha(\varepsilon) \log P\Big(\operatorname{dist}_{\mathcal{E}}\big(X_\xi^\varepsilon, \Gamma_\xi(s)\big) \ge \delta\Big) + s\right) \le 0.$$

Recall that a family $Q \subset C_b(\mathcal{E})$ of functions $z : \mathcal{E} \to \mathbb{R}$ is equibounded and equicontinuous if

$$\sup_{z\in Q} \sup_{\varphi\in\mathcal{E}} \big|z(\varphi)\big| < \infty \quad \text{and} \quad \lim_{\delta\to 0} \sup_{z\in Q} \sup_{\substack{\varphi,\eta\in\mathcal{E} \\ \tau(\varphi,\eta)<\delta}} \big|z(\varphi) - z(\eta)\big| = 0.$$

Definition 4.2 (EULP)**.** A family of $\mathcal{E}$-valued random variables, $\{X_\xi^\varepsilon\}_{\varepsilon\in(0,1]}^{\xi\in\mathcal{E}_0}$, is said to satisfy an EULP with speed $\alpha(\varepsilon)$ and rate function I_ξ, uniformly over $\mathscr{L}$, if for any $L \in \mathscr{L}$ and any equicontinuous and equibounded family $Q \subset C_b(\mathcal{E})$

$$\lim_{\varepsilon\to 0} \sup_{\xi\in L} \sup_{z\in Q} \left|\alpha(\varepsilon) \log \mathbb{E} \exp\left(-\frac{z(X_\xi^\varepsilon)}{\alpha(\varepsilon)}\right) + \inf_{\varphi\in\mathcal{E}} \big\{z(\varphi) + I_\xi(\varphi)\big\}\right| = 0.$$

Theorem 4.3 (Salins, 2019, Theorem 2.10)**.** *FWULDP and EULP are equivalent.*

We now present a sufficient condition under which a family of random variables satisfies an EULP and therefore also an FWULDP. To this end, for any $N > 0$ and $T > 0$, let $\mathcal{A}_2^N$ be the collection of $\mathcal{F}_t$-adapted square integrable $\mathbb{R}^k$-valued processes, such that

$$P\Big(|\beta|_{L^2([0,T]:\mathbb{R}^k)} \le N\Big) = 1.$$

Assumption 4.4. Assume that there exists a family of measurable maps $\mathscr{G}_\xi^\varepsilon : C\big([0,T] \to \mathbb{R}^k\big) \to \mathcal{E}$ indexed by $\varepsilon \in [0,1]$ and $\xi \in \mathcal{E}_0$. Let B be a k-dimensional Brownian motion, $X_\xi^{\varepsilon,\beta} := \mathscr{G}_\xi^\varepsilon\Big(\sqrt{\alpha(\varepsilon)}B + \int_0^\cdot \beta(s)ds\Big)$, and $\mathscr{L}$ be a collection of subsets of $\mathcal{E}_0$. Assume that for any $\delta > 0$, $L \in \mathscr{L}$ and $N > 0$

$$\lim_{\varepsilon\to 0}\sup_{\xi\in L}\sup_{\beta\in\mathcal{A}_2^N} P\bigg(\tau\bigg(\mathscr{G}_\xi^\varepsilon\bigg(\sqrt{\alpha(\varepsilon)}B + \int_0^\cdot \beta(s)ds\bigg), \mathscr{G}_\xi^0\bigg(\int_0^\cdot \beta(s)ds\bigg)\bigg) > \delta\bigg) = 0.$$

Theorem 4.5 (Salins, 2019, Theorem 2.13)**.** *If Assumption 4.4 holds, then the family* $X_\xi^{\varepsilon,\beta} := \mathscr{G}_\xi^\varepsilon\Big(\sqrt{\alpha(\varepsilon)}B + \int_0^\cdot \beta(s)ds\Big)$ *satisfies an EULP, uniformly over* $\mathscr{L}$ *with speed* $\alpha(\varepsilon)$ *and rate function*

$$I_\xi(\varphi) := \inf_{\big\{\beta\in L^2([0,T]:\mathbb{R}^k):\varphi:=\mathscr{G}_\xi^0(\int_0^\cdot \beta(s)ds)\big\}} \bigg\{\frac{1}{2}\int_0^T |\beta(s)|^2 ds\bigg\},$$

with the convention that $\inf(\emptyset) = +\infty$.

5. Controlled and Skeleton Equations

In this section we introduce the scaling, which is used in the second of our two main results: the uniform moderate deviation principle. Let $\{\gamma(\epsilon)\}_{\epsilon\in(0,1]}$ be a sequence satisfying

$$\lim_{\epsilon\to 0}\gamma(\epsilon) = \infty \quad \text{and} \quad \lim_{\epsilon\to 0}\gamma(\epsilon)\sqrt{\epsilon} = 0, \tag{5.1}$$

and define

$$z^{\epsilon,\xi} := \frac{1}{\gamma(\epsilon)\sqrt{\epsilon}}(u^{\epsilon,\xi} - u^{0,\xi}) = \frac{1}{\gamma(\epsilon)}v^{\epsilon,\xi}. \tag{5.2}$$

Note that for a.e. $\omega \in \Omega$ and for all $t \in [0,T]$, we have

$$\begin{aligned} z^{\epsilon,\xi}(t,x) = &-\int_0^t\int_D \partial_{y_i}G_{t-s}(x,y)\frac{1}{\gamma(\epsilon)\sqrt{\epsilon}}\Big(g_i\big(u^{0,\xi} + \gamma(\epsilon)\sqrt{\epsilon}z^{\epsilon,\xi}\big) - g_i\big(u^{0,\xi}\big)\Big)(s,y)dyds \\ &+ \int_0^t\int_D G_{t-s}(x,y)\frac{1}{\gamma(\epsilon)\sqrt{\epsilon}}\Big(f\big(u^{0,\xi} + \gamma(\epsilon)\sqrt{\epsilon}z^{\epsilon,\xi}\big) - f\big(u^{0,\xi}\big)\Big)(s,y)dyds \\ &+ \frac{1}{\gamma(\epsilon)}\int_0^t\int_D G_{t-s}(x,y)\sigma_j\big(u^{0,\xi} + \gamma(\epsilon)\sqrt{\epsilon}z^{\epsilon,\xi}\big)(s,y)dydB_j(s), \end{aligned} \tag{5.3}$$

for a.e. $x \in D$. Observe that due to Theorem 3.1 and (5.1), $z^{\epsilon,\xi}$ converges to zero in probability in $C\big([0,T] : L^\rho(D)\big)$ as ϵ approaches zero. Further, note that the unique mild solution to Eq. (2.1) is probabilistically strong. The same holds for the solution to (5.3). For this reason, for any $\epsilon > 0$ and initial condition $\xi \in L^\rho(D)$, there exists a measurable map $\mathscr{H}_\xi^\epsilon : C([0,T] : \mathbb{R}^k) \to C([0,T] : L^\rho(D))$, such that $z^{\epsilon,\xi} := \mathscr{H}_\xi^\epsilon\Big(\frac{1}{\gamma(\epsilon)}B\Big)$.

Next we provide the following definition [Eq. (4.1) in Budhiraja et al. (2008)], which is used in the proof of our main results.

Definition 5.1. For any $N > 0$, define

$$\Lambda^N := \left\{ \beta \in L^2\big([0,T] : \mathbb{R}^k\big) : \int_0^T |\beta(s)|^2 ds \le N \right\},$$

and endow Λ^N with the metric

$$d_1(f,g) := \sum_{i=1}^{\infty} \frac{1}{2^i} \left| \int_0^T \big(f_j(s) - g_j(s)\big) e_j^i(s) ds \right|,$$

where $\{e^i\}_{i\in\mathbb{N}}$ is an orthonormal basis of $L^2([0,T] : \mathbb{R}^k)$ under which Λ^N is a compact metric space. Note that when we say convergence in Λ^N, we will be referring to convergence under the d_1 metric (which is consistent with the weak $L^2\big([0,T] : \mathbb{R}^k\big)$ topology) rather than convergence under the $L^2\big([0,T] : \mathbb{R}^k\big)$ norm.

Let $\mathcal{A}_2$ be the set of $\mathcal{F}_t$-adapted controls $\beta : [0,T] \to \mathbb{R}^k$, such that $P\big(|\beta|_{L^2([0,T]:\mathbb{R}^k)} < \infty\big) = 1$. Further, as introduced in Section 4, for any $N > 0$, let $\mathcal{A}_2^N \subset \mathcal{A}_2$ be the set of admissible controls such that

$$\mathcal{A}_2^N := \Big\{ \beta \in \mathcal{A}_2 : \beta(\omega) \in \Lambda^N \ a.s. \Big\},$$

where Λ^N is given by Definition 5.1.

We now define the following change of measure, which will allow us to introduce the controlled version of $z^{\epsilon,\xi}(t,x)$.

Definition 5.2 (Change of measure)**.** For any $N > 0$ and $\beta \in \mathcal{A}_2^N$ define the measure $Q^{\epsilon,\beta}$ as follows

$$\frac{dQ^{\epsilon,\beta}}{dP} := \exp\left\{ \gamma(\epsilon) \int_0^T \beta_j(s) dB^j(s) - \frac{\gamma(\epsilon)^2}{2} \int_0^T |\beta(s)|^2 ds \right\}.$$

Note that Girsanov's theorem implies that under the measure, $Q^{\epsilon,\beta}$, the process

$$\bar{B}(t) := B(t) - \gamma(\epsilon) \int_0^t \beta(s) ds,$$

is a k-dimensional Wiener process.

Let $z^{\epsilon,\xi,\beta}(t,x) := \mathscr{H}_\xi^\epsilon\Big(\frac{1}{\gamma(\epsilon)} B + \int_0^{\cdot} \beta(s)ds\Big)$. Note that, $z^{\epsilon,\xi,\beta}(t,x)$ is the so-called controlled equation, and has the same distribution under the measure P, that $z^{\epsilon,\xi}(t,x) := \mathscr{H}_\xi^\epsilon\Big(\frac{1}{\gamma(\epsilon)} B\Big)$ has under the measure $Q^{\epsilon,\beta}$. It solves

$$\begin{aligned} z^{\epsilon,\xi,\beta}(t,x) = &- \int_0^t \int_D \partial_{y_i} G_{t-s}(x,y) \frac{1}{\gamma(\epsilon)\sqrt{\epsilon}} \Big(g_i\big(u^{0,\xi} + \gamma(\epsilon)\sqrt{\epsilon} z^{\epsilon,\xi,\beta}\big) - g_i\big(u^{0,\xi}\big)\Big)(s,y) dy ds \\ &+ \int_0^t \int_D G_{t-s}(x,y) \frac{1}{\gamma(\epsilon)\sqrt{\epsilon}} \Big(f\big(u^{0,\xi} + \gamma(\epsilon)\sqrt{\epsilon} z^{\epsilon,\xi,\beta}\big) - f\big(u^{0,\xi}\big)\Big)(s,y) dy ds \\ &+ \int_0^t \int_D G_{t-s}(x,y) \sigma_j\big(u^{\epsilon,\xi,\beta}\big)(s,y) dy \beta_j(s) ds \\ &+ \frac{1}{\gamma(\epsilon)} \int_0^t \int_D G_{t-s}(x,y) \sigma_j\big(u^{\epsilon,\xi,\beta}\big)(s,y) dy dB_j(s). \end{aligned} \tag{5.4}$$

The following theorem asserts the existence and uniqueness of solutions to (5.4). Its proof is almost verbatim to that of Budhiraja et al. (2008, Theorem 10) with Girsanov's theorem being the main ingredient.

Theorem 5.3 (Existence & uniqueness of controlled process)**.** *Let $\mathscr{H}^{\epsilon}_{\xi}$ denote the solution mapping, and let $\beta \in \mathcal{A}^N_2$ for some $N \in \mathbb{N}$. Then, for all $\epsilon \in (0,1]$ and $\xi \in L^{\rho}(D)$*

$$z^{\epsilon,\beta,\xi} := \mathscr{H}^{\epsilon}_{\xi}\Big(\frac{1}{\gamma(\epsilon)}B + \int_0^{\cdot} \beta(s)ds\Big),$$

is the unique solution to (5.4)*.*

When $\epsilon = 0$, $\mathscr{H}^{0}_{\xi}\Big(\int_0^{\cdot} \beta(s)ds\Big) := z^{0,\xi,\beta}(t,x)$, will represent the solution to the so-called skeleton equation. It solves

$$\begin{aligned} z^{0,\xi,\beta}(t,x) = &- \int_0^t \int_D \partial_{y_i} G_{t-s}(x,y)\Big(g_i'(u^{0,\xi})z^{0,\xi,\beta}\Big)(s,y)dyds \\ &+ \int_0^t \int_D G_{t-s}(x,y)\Big(f'(u^{0,\xi})z^{0,\xi,\beta}\Big)(s,y)dyds \\ &+ \int_0^t \int_D G_{t-s}(x,y)\sigma_j\big(u^{0,\xi}\big)(s,y)dy\beta_j(s)ds. \end{aligned} \tag{5.5}$$

The following lemma asserts the existence of a unique solution to the skeleton equation. It is essentially a deterministic version of Theorem 3.1. Therefore, for the sake of brevity, and because the arguments employed in the proof of this lemma are indeed similar to those used in the proof of Theorem 3.1, we will only provide a sketch, which is located in Section A.2.

Theorem 5.4 (Existence & uniqueness of skeleton)**.** *For any $N > 0$ and $\beta \in \Lambda^N$, there exists a unique solution to* (5.5) *in $C\big([0,T] : L^{\rho}(D)\big)$.*

We are now in a position to provide a definition for the large deviations rate function.

Definition 5.5 (Large deviations rate function)**.** Let $\xi \in L^{\rho}(D)$ and $\psi \in C\big([0,T] : L^{\rho}(D)\big)$. We define the large deviations rate function or action functional, $I_{\xi} : C\big([0,T] : L^{\rho}(D)\big) \to [0,+\infty]$ by

$$I_{\xi}(\psi) := \inf_{\big\{\beta \in L^2([0,T]:\mathbb{R}^k) : \psi := \mathscr{H}^0_{\xi}(\int_0^{\cdot} \beta(s)ds)\big\}} \Big\{\frac{1}{2}\int_0^T |\beta(s)|^2 ds\Big\}, \tag{5.6}$$

with the convention that $\inf(\emptyset) = +\infty$. We say that I_{ξ} is a rate function if for any fixed $\xi \in L^{\rho}(D)$, the mapping $I_{\xi} : C\big([0,T] : L^{\rho}(D)\big) \to [0,+\infty]$ is lower-semicontinuous.

6. Uniform Moderate Deviation Principle in $C\big([0,T] : L^{\rho}(D)\big)$

In this section we state the uniform moderate deviation principle – the second main result of our paper – which is indeed a uniform large deviation principle but with the scaling introduced in Section 5 .

Theorem 6.1 (Uniform Moderate Deviation Principle in $C\big([0,T] : L^{\rho}(D)\big)$ topology)**.** *The processes, $\{z^{\epsilon,\xi}\}^{\xi \in L^{\rho}(D)}_{\epsilon \in (0,1]}$ satisfy a uniform moderate deviation principle on $C\big([0,T] : L^{\rho}(D)\big)$, where the uniformity is over $L^{\rho}(D)$-bounded sets of initial conditions with rate function I_{ξ}, given by* (5.6)*.*

(1) For any $R > 0$, $s_0 \geq 0$, and $\delta > 0$

$$\liminf_{\epsilon \to 0} \inf_{|\xi|_{\rho} \leq R} \inf_{\psi \in \Gamma_{\xi}(s_0)} \bigg(\frac{1}{(\gamma(\epsilon))^2}\log P\Big(\big|z^{\epsilon,\xi} - \psi\big|_{C([0,T]:L^{\rho}(D))} < \delta\Big) + I_{\xi}(\psi)\bigg) \geq 0.$$

(2) For any $R > 0$, $s_0 \geq 0$, and $\delta > 0$

$$\limsup_{\epsilon \to 0} \sup_{|\xi|_{\rho} \leq R} \sup_{s \in [0,s_0]} \bigg(\frac{1}{(\gamma(\epsilon))^2}\log P\Big(dist_{C([0,T]:L^{\rho}(D))}\big(z^{\epsilon,\xi}, \Gamma_{\xi}(s)\big)\Big) \geq \delta\Big) + s\bigg) \leq 0,$$

where

$$\Gamma_\xi(s) := \Big\{\psi \in C\big([0,T] : L^\rho(D)\big) : I_\xi(\psi) \le s\Big\}.$$

The proof of Theorem 6.1 is in Section 8.

Corollary 6.2 (ULP)**.** *Let I_ξ be given by Definition 5.6 and $\{z^{\epsilon,\xi}\}_{\epsilon\in(0,1]}^{\xi\in L^\rho(D)}$ be given by (5.3). Then, for any bounded, continuous function $f : C\big([0,T] : L^\rho(D)\big) \to \mathbb{R}$ and compact subset $K \subset L^\rho(D)$, we have*

$$\limsup_{\epsilon\to 0}\sup_{\xi\in K}\left|\frac{1}{\big(\gamma(\epsilon)\big)^2}\log E\Big[e^{-(\gamma(\epsilon))^2 f(z^{\epsilon,\xi})}\Big] + \inf_{\psi\in C([0,T]:L^\rho(D))}\big\{I_\xi(\psi)+f(\psi)\big\}\right| = 0.$$

The proof of Corollary 6.2 is in Section 8.

The following corollary follows immediately from Corollary 6.2, as for Polish space random elements the Laplace principle implies the large deviation principle. (see Dupuis and Ellis, 1997, Theorem 1.2.1). Due to the scaling that the LDP is under, we will refer to this corollary as an MDP.

Corollary 6.3 (MDP)**.** *Let I_ξ be given by Definition 5.6 and $\{z^{\epsilon,\xi}\}_{\epsilon\in(0,1]}^{\xi\in L^\rho(D)}$ be given by (5.3). Then, for any $\xi \in L^\rho(D)$, the following hold:*

(a) For any closed $F \subset C\big([0,T] : L^\rho(D)\big)$

$$\limsup_{\epsilon\to 0}\frac{1}{\big(\gamma(\epsilon)\big)^2}\log P\big(z^{\epsilon,\xi}\in F\big) \le -\inf_{\psi\in F}\big\{I_\xi(\psi)\big\}.$$

(b) For any open $G \subset C\big([0,T] : L^\rho(D)\big)$

$$\liminf_{\epsilon\to 0}\frac{1}{\big(\gamma(\epsilon)\big)^2}\log P\big(z^{\epsilon,\xi}\in G\big) \ge -\inf_{\psi\in G}\Big\{I_\xi(\psi)\Big\}.$$

7. Preliminary Results Required to Prove Theorem 6.1

The following lemma demonstrates that the law of large numbers limit is continuous as a function of the initial condition, which is needed for the proof of Theorem 7.3.

Lemma 7.1. *If $\xi^n \to \xi$ in $L^\rho(D)$ as $n\to\infty$, then $u^{0,\xi^n} \to u^{0,\xi}$ in $C\big([0,T] : L^\rho(D)\big)$.*

Proof: We have for all $t\in[0,T]$

$$\begin{aligned}\Big|u^{0,\xi^n}(t) - u^{0,\xi}(t)\Big|_\rho &\le \left|\int_0^t\int_D \partial_{y_i}G_{t-s}(\cdot,y)\Big(g_i(u^{0,\xi^n}) - g_i(u^{0,\xi})\Big)(s,y)dyds\right|_\rho\\ &+\left|\int_0^t\int_D G_{t-s}(\cdot,y)\Big(f(u^{0,\xi^n}) - f(u^{0,\xi})\Big)(s,y)dyds\right|_\rho\\ &+\left|\int_D G_t(\cdot,y)\big|\xi^n(y)-\xi(y)\big|dy\right|_\rho\\ &:= J_1^n(t)+J_2^n(t)+J_3^n(t).\end{aligned}$$

Using an argument similar to that used in the proof of Proposition 3.2 gives

$$J_2^n(t) \le C\int_0^t\big|u^{0,\xi^n}(s)-u^{0,\xi}(s)\big|_\rho ds,$$

and

$$\begin{aligned} J_1^n(t) &\le C\Big(1+\sup_{s\in[0,T]}\big|u^{0,\xi^n}(s)\big|_\rho+\sup_{s\in[0,T]}\big|u^{0,\xi^n}(s)\big|_\rho\Big)\int_0^t (t-s)^{-\frac{d}{2\rho}-\frac{1}{2}}\big|u^{0,\xi^n}(s)-u^{0,\xi}(s)\big|_\rho ds\\ &\le C\Big(1+\big|\xi^n\big|_\rho^\rho+\big|\xi\big|_\rho^\rho\Big)\int_0^t (t-s)^{-\frac{d}{2\rho}-\frac{1}{2}}\big|u^{0,\xi^n}(s)-u^{0,\xi}(s)\big|_\rho ds, \end{aligned}$$

where the second inequality used Lemma A.6. In addition, Young's inequality and (E1) give

$$J_3^n(t)\le C\big|\xi^n(s)-\xi(s)\big|_\rho.$$

Therefore, for all $t\in[0,T]$ we have

$$\begin{aligned} \big|u^{0,\xi^n}(t)-u^{0,\xi}(t)\big|_\rho &\le C\big|\xi^n(s)-\xi(s)\big|_\rho\\ &\quad + C\int_0^t\Big(\big(1+\big|\xi^n\big|_\rho^\rho+\big|\xi\big|_\rho^\rho\big)(t-s)^{-\frac{d}{2\rho}-\frac{1}{2}}+1\Big)\big|u^{0,\xi^n}(s)-u^{0,\xi}(s)\big|_\rho ds. \end{aligned}\tag{7.1}$$

Finally, Grönwall's inequality implies

$$\sup_{t\in[0,T]}\big|u^{0,\xi^n}(t)-u^{0,\xi}(t)\big|_\rho\le C\big|\xi^n(s)-\xi(s)\big|_\rho e^{C\left(1+|\xi^n|_\rho^\rho+|\xi|_\rho^\rho\right)}.$$

Now tending n to ∞, taking the limit of both sides, and further using that $\xi^n\to\xi$ in $L^\rho(D)$ as $n\to\infty$, yields the result. This completes the proof. □

The proof of Theorem 7.3 also requires the following lemma, which gives a bound on the L^ρ-norm of the skeleton equation in terms of the initial condition and the control.

Lemma 7.2. *There exists a constant $C>0$, such that for all $\xi\in L^\rho(D)$, $N>0$, and $\beta\in\Lambda^N$, we have*

$$\sup_{t\in[0,T]}\big|z^{0,\xi,\beta}(t)\big|_\rho\le CN\Big(1+\big|\xi\big|_\rho^\rho\Big)e^{C(1+|\xi|_\rho^\rho)}.$$

Proof: For all $t\in[0,T]$, we have

$$\begin{aligned} \big|z^{0,\xi,\beta}(t)\big|_\rho &\le \bigg|\int_0^t\int_D \partial_{y_i}G_{t-s}(\cdot,y)\Big(g_i'(u^{0,\xi})z^{0,\xi,\beta}\Big)(s,y)dyds\bigg|_\rho\\ &\quad+\bigg|\int_0^t\int_D G_{t-s}(\cdot,y)\Big(f'(u^{0,\xi})z^{0,\xi,\beta}\Big)(s,y)dyds\bigg|_\rho\\ &\quad+\bigg|\int_0^t\int_D G_{t-s}(\cdot,y)\sigma_j\big(u^{0,\xi}\big)(s,y)dy\beta_j(s)ds\bigg|_\rho\\ &:= J_1(t)+J_2(t)+J_3(t). \end{aligned}$$

Using a similar argument to that applied to the $\tilde{L}_1^\epsilon$ and $\tilde{L}_2^\epsilon$ terms in the proof of Theorem 3.1, gives

$$J_1(t)\le C\big(1+\big|\xi\big|_\rho^\rho\big)\int_0^t (t-s)^{-\frac{d}{2\rho}-\frac{1}{2}}\big|z^{0,\xi,\beta}(s)\big|_\rho ds,$$

and

$$J_2(t)\le C\big(1+\big|\xi\big|_\rho^\rho\big)\int_0^t (t-s)^{-\frac{d}{2\rho}}\big|z^{0,\xi,\beta}(s)\big|_\rho ds.$$

In addition, Minkowski's inequality, Young's inequality with (E1) and (A4), Lemma A.6, and the fact that $\beta \in \Lambda^N$, give

$$J_3(t) \le C \int_0^t \Big(1 + \big|u^{0,\xi}(s)\big|_\rho\Big)\big|\beta(s)\big| ds \le C\Big(1 + \sup_{t\in[0,T]} \big|u^{0,\xi}(t)\big|_\rho\Big)\Big(\int_0^t \big|\beta(s)\big|^2 ds\Big)^{\frac{1}{2}} \le CN\Big(1 + |\xi|_\rho^\rho\Big).$$

Therefore

$$\big|z^{0,\beta,\xi}(t)\big|_\rho \le CN\big(1 + |\xi|_\rho^\rho\big) + C\big(1 + |\xi|_\rho^\rho\big) \int_0^t \Big((t-s)^{-\frac{d}{2\rho}} + (t-s)^{-\frac{d}{2\rho}-\frac{1}{2}}\Big)\big|z^{0,\xi,\beta}(s)\big|_\rho ds,$$

for all $t \in [0,T]$ and Grönwall's inequality yields

$$\sup_{t\in[0,T]} \big|z^{0,\xi,\beta}(t)\big|_\rho \le CN\big(1 + |\xi|_\rho^\rho\big) e^{C\left(1+|\xi|_\rho^\rho\right)}.$$

This completes the proof. □

The next theorem demonstrates the continuity of the skeleton equation and is used in the proof Corollary 6.2.

Theorem 7.3 (Continuity of skeleton equation)**.** *For any $N > 0$, the mapping*

$$(\xi, \beta) \to \mathscr{H}_\xi^0\Big(\int_0^\cdot \beta(s) ds\Big) \text{ from } L^\rho(D) \times \Lambda^N \text{ to } C\big([0,T] : L^\rho(D)\big)$$

is continuous.

Proof: Assume that $(\xi^n, \beta^n) \to (\xi, \beta)$ in $L^\rho(D) \times \Lambda^N$ (recall the metric on Λ^N given in Definition 5.1). We have

$$\sup_{t\in[0,T]} \Big|z^{0,\xi^n,\beta^n}(t) - z^{0,\xi,\beta,}(t)\Big|_\rho \le \sup_{t\in[0,T]} \Big|z^{0,\xi^n,\beta^n}(t) - z^{0,\xi,\beta^n}(t)\Big|_\rho + \sup_{t\in[0,T]} \Big|z^{0,\xi,\beta^n}(t) - z^{0,\xi,\beta}(t)\Big|_\rho. \quad (7.2)$$

First, consider $\sup_{t\in[0,T]} \big|z^{0,\xi^n,\beta^n}(t) - z^{0,\xi,\beta^n}(t)\big|_\rho$ and note that for all $t \in [0,T]$

$$\begin{aligned}\Big|z^{0,\xi^n,\beta^n}(t) - z^{0,\xi,\beta^n}(t)\Big|_\rho \le& \left|\int_0^t \int_D \partial_{y_i} G_{t-s}(\cdot,y)\Big(g_i'(u^{0,\xi^n}) z^{0,\xi^n,\beta^n} - g_i'(u^{0,\xi}) z^{0,\xi,\beta^n}\Big)(s,y) dy ds\right|_\rho \\ &+ \left|\int_0^t \int_D G_{t-s}(\cdot,y)\Big(f'(u^{0,\xi^n}) z^{0,\xi^n,\beta^n} - f'(u^{0,\xi}) z^{0,\xi,\beta^n}\Big)(s,y) dy ds\right|_\rho \\ &+ \left|\int_0^t \int_D G_{t-s}(\cdot,y)\Big(\sigma_j(u^{0,\xi^n}) - \sigma_j(u^{0,\xi})\Big)(s,y) dy \beta_j^n(s) ds\right|_\rho \\ :=& J_1^n(t) + J_2^n(t) + J_3^n(t).\end{aligned}$$

Using Minkowski's inequality, we have

$$\begin{aligned}J_1^n \le& \int_0^t \left|\int_D \partial_{y_i} G_{t-s}(\cdot,y)\Big(g_i'(u^{0,\xi^n})\big(z^{0,\xi^n,\beta^n} - z^{0,\xi,\beta^n}\big)\Big)(s,y) dy\right|_\rho ds \\ &+ \int_0^t \left|\int_D \partial_{y_i} G_{t-s}(\cdot,y)\Big(\big(g_i'(u^{0,\xi^n}) - g_i'(u^{0,\xi})\big) z^{0,\xi,\beta^n}\Big)(s,y) dy\right|_\rho ds \\ :=& J_{11}^n(t) + J_{12}^n(t),\end{aligned}$$

and

$$J_2^n \leq \int_0^t \Big| \int_D G_{t-s}(\cdot,y)\Big(f'(u^{0,\xi^n})(z^{0,\xi^n,\beta^n} - z^{0,\xi,\beta^n})\Big)(s,y)dy\Big|_\rho ds$$
$$\int_0^t \Big| \int_D G_{t-s}(\cdot,y)\Big(\big(f'(u^{0,\xi^n}) - f'(u^{0,\xi})\big)z^{0,\xi,\beta^n}\Big)(s,y)dy\Big|_\rho ds$$
$$:= J_{21}^n(t) + J_{22}^n(t).$$

A similar argument to that applied to the $\tilde{L}_1^\epsilon$ and $\tilde{L}_2^\epsilon$ terms in the proof of Theorem 3.1, gives

$$J_{11}^n(t) \leq C\big(1 + |\xi^n|_\rho^\rho\big) \int_0^t (t-s)^{-\frac{d}{2\rho}-\frac{1}{2}} \Big|z^{0,\xi^n,\beta^n}(s) - z^{0,\xi,\beta^n}(s)\Big|_\rho ds,$$

and

$$J_{21}^n(t) \leq C\big(1 + |\xi^n|_\rho^\rho\big) \int_0^t (t-s)^{-\frac{d}{2\rho}} \Big|z^{0,\xi^n,\beta^n}(s) - z^{0,\xi,\beta^n}(s)\Big|_\rho ds.$$

In addition, using Young's inequality combined with (E2), Hölder's inequality, and (A6), gives

$$\begin{aligned} J_{12}^n(t) &\leq C \sum_{i=1}^d \int_0^t (t-s)^{-\frac{d}{2\rho}-\frac{1}{2}} \Big| \Big(\big(g_i'(u^{0,\xi^n}) - g_i'(u^{0,\xi})\big)z^{0,\xi,\beta^n}\Big)(s,\cdot)\Big|_{\frac{\rho}{2}} ds \\ &\leq C \sum_{i=1}^d \int_0^t (t-s)^{-\frac{d}{2\rho}-\frac{1}{2}} \Big|g_i'(u^{0,\xi^n})(s) - g_i'(u^{0,\xi})(s)\Big|_\rho \big|z^{0,\xi,\beta^n}(s)\big|_\rho ds \\ &\leq C \sup_{t\in[0,T]} \big|z^{0,\xi,\beta^n}(t)\big|_\rho \sup_{t\in[0,T]} \Big|u^{0,\xi^n}(t) - u^{0,\xi}(t)\Big|_\rho. \end{aligned}$$

Similarly, using Young's inequality combined with (E1), Hölder's inequality, and (A6), gives

$$J_{22}^n(t) \leq C \sup_{t\in[0,T]} \big|z^{0,\xi,\beta^n}(t)\big|_\rho \sup_{t\in[0,T]} \Big|u^{0,\xi^n}(t) - u^{0,\xi}(t)\Big|_\rho.$$

Again using Minkowski's inequality, Young's inequality with (E1), (A5), and that $\beta^n \in \Lambda^N$, gives

$$\begin{aligned} J_3^n(t) &\leq \int_0^t \Big| \int_D G_{t-s}(\cdot,y)\Big(\sigma_j(u^{0,\xi^n}) - \sigma_j(u^{0,\xi})\Big)(s,y)dy\beta_j^n(s)\Big|_\rho ds \\ &\leq C \int_0^t \Big|u^{0,\xi^n}(s) - u^{0,\xi}(s)\Big|_\rho \big|\beta^n(s)\big| ds \leq C \sup_{t\in[0,T]} \Big|u^{0,\xi^n}(t) - u^{0,\xi}(t)\Big|_\rho. \end{aligned}$$

Therefore

$$\begin{aligned} \Big|z^{0,\xi^n,\beta^n}(t) - z^{0,\xi,\beta^n}(t)\Big|_\rho &\leq C\Big(1 + \sup_{t\in[0,T]} \big|z^{0,\xi,\beta^n}(t)\big|_\rho\Big) \sup_{t\in[0,T]} \Big|u^{0,\xi^n}(t) - u^{0,\xi}(t)\Big|_\rho \\ &\quad + C\Big(1 + |\xi^n|_\rho^\rho\Big) \int_0^t c_t(s) \Big|z^{0,\xi^n,\beta^n}(s) - z^{0,\xi,\beta^n}(s)\Big|_\rho ds, \end{aligned}$$

for all $t \in [0,T]$, where

$$c_t(s) := (t-s)^{-\frac{d}{2\rho}-\frac{1}{2}} + (t-s)^{-\frac{d}{2\rho}}.$$

Consequently, Grönwall's inequality implies

$$\sup_{t\in[0,T]} \Big|z^{0,\xi^n,\beta^n}(t) - z^{0,\xi,\beta^n}(t)\Big|_\rho \leq C\Big(1 + \sup_{t\in[0,T]} \big|z^{0,\xi,\beta^n}(t)\big|_\rho\Big) \sup_{t\in[0,T]} \Big|u^{0,\xi^n}(t) - u^{0,\xi}(t)\Big|_\rho e^{C(1+|\xi^n|_\rho^\rho)},$$

and due to Lemma 7.1, Lemma 7.2, and the fact that $\xi^n \to \xi$ in $L^\rho(D)$ as $n \to \infty$, we have

$$\lim_{n\to\infty} \sup_{t\in[0,T]} \Big|z^{0,\xi^n,\beta^n}(t) - z^{0,\xi,\beta^n}(t)\Big|_\rho = 0. \tag{7.3}$$

Now consider $\sup_{t\in[0,T]}\big|z^{0,\xi,\beta}(t)-z^{0,\xi,\beta^n}(t)\big|_\rho$. For all $t\in[0,T]$, we have

$$\begin{aligned}z^{0,\xi,\beta}(t,x)-z^{0,\xi,\beta^n}(t,x)&=\int_0^t\int_D\partial_{y_i}G_{t-s}(x,y)\Big(g_i'(u^{0,\xi})\big(z^{0,\xi,\beta}-z^{0,\xi,\beta^n}\big)\Big)(s,y)dyds\\&\quad+\int_0^t\int_D G_{t-s}(x,y)\Big(f'(u^{0,\xi})\big(z^{0,\xi,\beta}-z^{0,\xi,\beta^n}\big)\Big)(s,y)dyds\\&\quad+\int_0^t\int_D G_{t-s}(x,y)\Big(\sigma_j(u^{0,\xi})\Big)(s,y)dy\big(\beta_j(s)-\beta_j^n(s)\big)ds\\&:=L_1^n(t,x)+L_2^n(t,x)+L_3^n(t,x),\end{aligned}$$

for a.e. $x\in D$. Therefore, for all $t\in[0,T]$

$$\Big|z^{0,\xi,\beta}(t)-z^{0,\xi,\beta^n}(t)\Big|_\rho\le\big|L_1^n(t,\cdot)\big|_\rho+\big|L_2^n(t,\cdot)\big|_\rho+\big|L_3^n(t,\cdot)\big|_\rho.$$

Similar to J_{11}^n and J_{21}^n above, it follows that

$$\big|L_1^n(t,\cdot)\big|_\rho\le C\big(1+|\xi|_\rho^\rho\big)\int_0^t(t-s)^{-\frac{d}{2\rho}-\frac12}\Big|z^{0,\xi,\beta}(s)-z^{0,\xi,\beta^n}(s)\Big|_\rho ds,$$

and

$$\big|L_2^n(t,\cdot)\big|_\rho\le C\big(1+|\xi|_\rho^\rho\big)\int_0^t(t-s)^{-\frac{d}{2\rho}}\Big|z^{0,\xi,\beta}(s)-z^{0,\xi,\beta^n}(s)\Big|_\rho ds.$$

Therefore, for all $t\in[0,T]$

$$\Big|z^{0,\xi,\beta}(t)-z^{0,\xi,\beta^n}(t)\Big|_\rho\le\sup_{s\in[0,T]}\big|L_3^n(s)\big|_\rho+C\big(1+|\xi|_\rho^\rho\big)\int_0^t c_t(s)\Big|z^{0,\xi,\beta}(s)-z^{0,\xi,\beta^n}(s)\Big|_\rho ds,$$

where

$$c_t(s):=(t-s)^{-\frac{d}{2\rho}-\frac12}+(t-s)^{-\frac{d}{2\rho}},$$

and Grönwall's inequality gives

$$\sup_{t\in[0,T]}\Big|z^{0,\xi,\beta}(t)-z^{0,\xi,\beta^n}(t)\Big|_\rho\le\sup_{t\in[0,T]}\big|L_3^n(t)\big|_\rho e^{C(1+|\xi|_\rho^\rho)}.\tag{7.4}$$

Now consider $L_3^n(t,x)$. Note that for all $(t,x)\in[0,T]\times D$ and $j\in\{1,...,k\}$, Hölder's inequality combined with (E1), (A4), and Lemma A.6, gives

$$\begin{aligned}&\int_0^t\bigg(\int_D G_{t-s}(x,y)\sigma_j\big(u^{0,\xi}\big)(s,y)dy\bigg)^2ds\\&\le C\int_0^t(t-s)^{-\frac{d}{\rho}}\big|\sigma_j\big(u^{0,\xi}\big)(s,\cdot)\big|_\rho^2ds\le C\bigg(1+\sup_{s\in[0,T]}|u^{0,\xi}(s)|_\rho^2\bigg)ds\le C\big(1+|\xi|_\rho^\rho\big).\end{aligned}$$

Consequently, for all $(t,x)\in[0,T]\times D$, we have $\mathcal{I}_{\{s<t\}}\int_D G_{t-s}(x,y)\sigma_j\big(u^{0,\xi}\big)(s,y)dy\in L^2\big([0,T]:\mathbb{R}^k\big)$ so the convergence of $\beta^n\to\beta$ in Λ^N $\big($which implies $\beta^n\to\beta$ weakly in $L^2\big([0,T]:\mathbb{R}^k\big)\big)$ gives

$$\lim_{n\to\infty}L_3^n(t,x)\to0,\tag{7.5}$$

for all $(t,x)\in[0,T]\times D$. Due to (A4), Lemma A.6, and that $\{\beta^n\}_{n\in\mathbb{N}},\beta\in\Lambda^N$, we have for all $j\in\{1,...,k\}$

$$\begin{aligned}\bigg(\int_0^T\Big|\sigma_j\big(u^{0,\xi}\big)(s,\cdot)\big(\beta_j^n(s)-\beta_j(s)\big)\Big|_\rho^2ds\bigg)^{\frac12}&\le C\bigg(1+\sup_{r\in[0,T]}\big|u^{0,\xi}(r)\big|_\rho\bigg)\bigg(\int_0^T\Big|\beta_j^n(s)-\beta_j(s)\Big|^2ds\bigg)^{\frac12}\\&\le C\big(1+|\xi|_\rho^\rho\big).\end{aligned}$$

Consequently, Lemma A.3 implies that there exists a constant $C > 0$, such that for all $n \in \mathbb{N}$

$$\Big|L_3^n(t,\cdot) - L_3^n(s,\cdot)\Big|_\rho \le C\big(1 + |\xi|_\rho^\rho\big)|t-s|^{\frac{1}{4}},$$

for all $s, t \in [0,T]$ and

$$\Big|L_3^n(t,x) - L_3^n(t,y)\Big| \le C\big(1 + |\xi|_\rho^\rho\big)|x-y|^{\frac{\rho-d}{2\rho}},$$

for all $x, y \in D$, and $t \in [0,T]$. It follows from a deterministic version of the proof of Lemma A.5 that $\{L_3^n\}_{n\in\mathbb{N}}$ is tight in $C\big([0,T] : L^\rho(D)\big)$. Therefore, any subsequence of $\{L_3^n\}_{n\in\mathbb{N}}$ has a further subsequence that converges in $C\big([0,T] : L^\rho(D)\big)$ and due to (7.5), this limit must be $0 \in C\big([0,T] : L^\rho(D)\big)$. From this it follows that

$$\lim_{n\to\infty} \sup_{t\in[0,T]} \big|L_3^n(t)\big|_\rho = 0.$$

Plugging this into (7.4), gives

$$\lim_{n\to\infty} \sup_{t\in[0,T]} \Big|z^{0,\xi,\beta}(t) - z^{0,\xi,\beta^n}(t)\Big|_\rho = 0,$$

which combined with (7.3) and (7.2), yields

$$\lim_{n\to\infty} \sup_{t\in[0,T]} \Big|z^{0,\xi^n,\beta^n}(t) - z^{0,\xi,\beta}(t)\Big|_\rho = 0.$$

This completes the proof. □

The following lemma, which can be thought of as an analogue of Lemma A.6 for $u^{\epsilon,\xi,\beta}$ is used to prove Lemma 7.5.

Lemma 7.4. *Let $N > 0$ be arbitrary. Then, there exists a constant $C_N > 0$, such that for any $\beta \in \mathcal{A}_2^N$, $\xi \in L^\rho(D)$, and $\epsilon \in (0,1]$, we have*

$$E\Big[\sup_{t\in[0,T]} \big|u^{\epsilon,\beta,\xi}(t)\big|_\rho^\rho\Big] \le C_N\big(1 + |\xi|_\rho^\rho\big).$$

Proof: This proof is essentially an application of Girsanov's theorem to what was already done in Gyöngy and Rovira (2000). For this reason, we only provide a sketch here and refer the reader to that paper for additional details.

Recall that $\lim_{\epsilon\to 0}\sqrt{\epsilon}\gamma(\epsilon) = 0$. For this reason, $\sup_{\epsilon\in(0,1]}\sqrt{\epsilon}\gamma(\epsilon) < \infty$. Let $\epsilon \in (0,1]$ be arbitrary. As in the proofs of Theorem 2.1 and Proposition 4.4 in Gyöngy and Rovira (2000), we can consider a sequence of $\mathcal{F}_t$-adapted random processes $\{\hat{u}_n^\epsilon\}_{n\in\mathbb{N}} \in C\big([0,T] : L^\rho(D)\big)$, which satisfy the following:

(a) $\lim_{n\to\infty} \hat{u}_n^\epsilon = u^{\epsilon,\xi}$ a.s. in $C\big([0,T] : L^\rho(D)\big)$.

(b) For all $q \in [1,\infty)$ and $n \in \mathbb{N}$

$$E\Big[\sup_{t\in[0,T]} \big|\hat{u}_n^\epsilon(t)\big|_q^q\Big] < \infty.$$

(c) There exists a constant $K_1 > 0$, which does not depend on n or ϵ, such that for all $n \in \mathbb{N}$, and for a.e. $\omega \in \Omega$

$$\begin{aligned}\big|\hat{u}_n^\epsilon(t)\big|_\rho^\rho &\le |\xi|_\rho^\rho + K_1 + K_1\int_0^t \big|\hat{u}_n^\epsilon(s)\big|_\rho^\rho ds \\ &\quad + \sqrt{\epsilon}\rho\int_0^t\int_D \big|\hat{u}_n^\epsilon(s,x)\big|^{\rho-2}\hat{u}_n^\epsilon(s,x)\sigma_{jn}\big(s,x,\hat{u}_n^\epsilon(s,x)\big)dx dB_j(s),\end{aligned}$$

where σ_{jn} is a truncated version of σ_j, which satisfies (A4) with the same $L > 0$.

Recall the measure $Q^{\epsilon,\beta}$ given in Definition 5.2 and that $\bar{B}(t) := B(t) - \gamma(\epsilon)\int_0^t \beta(s)ds$ is a k-dimensional Brownian motion under $Q^{\epsilon,\beta}$. Consequently, under the measure $Q^{\epsilon,\beta}$, $u^{\epsilon,\xi}$ has the same distribution as $u^{\epsilon,\xi,\beta}$. For a.e. $\omega \in \Omega$

$$\begin{aligned}\big|\hat{u}_n^\epsilon(t)\big|_\rho^\rho &\le \big|\xi\big|_\rho^\rho + K_1 + K_1\int_0^t \big|\hat{u}_n^\epsilon(s)\big|_\rho^\rho ds \\ &+ \sqrt{\epsilon}\gamma(\epsilon)\rho\int_0^t\int_D |\hat{u}_n^\epsilon(s,x)|^{\rho-2}\hat{u}_n^\epsilon(s,x)\sigma_{jn}\big(s,x,\hat{u}_n^\epsilon(s,x)\big)dx\beta_j^\epsilon(s)ds \\ &+ \sqrt{\epsilon}\rho\int_0^t\int_D \big|\hat{u}_n^\epsilon(s,x)\big|^{\rho-2}\hat{u}_n^\epsilon(s,x)\sigma_{jn}\big(s,x,\hat{u}_n^\epsilon(s,x)\big)dxd\bar{B}_j(s), \end{aligned} \tag{7.6}$$

for all $t \in [0,T]$. Since $\beta \in \mathcal{A}_2^N$, it follows that $E\Big[\Big(\frac{dQ^{\epsilon,\beta}}{dP}\Big)^2\Big] < \infty$. Consequently for all $q < \infty$

$$E^{Q^{\epsilon,\beta}}\Big[\sup_{t\in[0,T]}\big|\hat{u}_n^\epsilon(t)\big|_q^q\Big] \le E\Big[\sup_{t\in[0,T]}\big|\hat{u}_n^\epsilon(t)\big|_{2q}^{2q}\Big]^{\frac{1}{2}} E\Big[\Big(\frac{dQ^{\epsilon,\beta}}{dP}\Big)^2\Big]^{\frac{1}{2}} < \infty.$$

The Burkholder-Davis-Gundy inequality gives

$$\begin{aligned}&E^{Q^{\epsilon,\beta}}\Big[\sup_{r\in[0,t]}\Big|\sqrt{\epsilon}\rho\int_0^r\int_D \big|\hat{u}_n^\epsilon(s,x)\big|^{\rho-2}\hat{u}_n^\epsilon(s,x)\sigma_{jn}\big(s,x,\hat{u}_n^\epsilon(s,x)\big)dxd\bar{B}_j(s)\Big|\Big] \\ &\le C\sum_{j=1}^k E^{Q^{\epsilon,\beta}}\Big[\Big(\int_0^t\Big(\int_D \big|\hat{u}_n^\epsilon(s,x)\big|^{\rho-2}\hat{u}_n^\epsilon(s,x)\sigma_{jn}\big(s,x,\hat{u}_n^\epsilon(s,x)\big)dx\Big)^2 ds\Big)^{\frac{1}{2}}\Big] \\ &\le CE^{Q^{\epsilon,\beta}}\Big[\Big(\int_0^t (1+\big|\hat{u}_n^\epsilon(s)\big|_\rho^\rho)^2 ds\Big)^{\frac{1}{2}}\Big] \le C\Big(1 + E^{Q^{\epsilon,\beta}}\Big[\Big(\sup_{r\in[0,t]}\big|\hat{u}_n^\epsilon(r)\big|_\rho^\rho\int_0^t \big|\hat{u}_n^\epsilon(s)\big|_\rho^\rho ds\Big)^{\frac{1}{2}}\Big]\Big) \\ &\le \frac{1}{4}E^{Q^{\epsilon,\beta}}\Big[\sup_{r\in[0,t]}\big|\hat{u}_n^\epsilon(r)\big|_\rho^\rho\Big] + K_2\Big(1 + E^{Q^{\epsilon,\beta}}\Big[\int_0^t \big|\hat{u}_n^\epsilon(s)\big|_\rho^\rho ds\Big]\Big),\end{aligned}$$

for all $t \in [0,T]$, where the constant $K_2 > 0$ neither depends on n nor on ϵ. Similarly, since $\sup_{\epsilon\in(0,1]}\sqrt{\epsilon}\gamma(\epsilon) < \infty$, for all $t \in [0,T]$, we have

$$\begin{aligned}&\Big|\sqrt{\epsilon}\gamma(\epsilon)\rho\int_0^t\int_D \big|\hat{u}_n^\epsilon(s,x)\big|^{\rho-2}\hat{u}_n^\epsilon(s,x)\sigma_{jn}\big(s,x,\hat{u}_n^\epsilon(s,x)\big)dx\beta_j^\epsilon(s)ds\Big| \\ &\le C\int_0^t \Big(1+\big|\hat{u}_n^\epsilon(s)\big|_\rho^\rho\Big)\big|\beta^\epsilon(s)\big|ds \le C\Big(\int_0^t \Big(1+\big|\hat{u}_n^\epsilon(s)\big|_\rho^\rho\Big)^2 ds\Big)^{\frac{1}{2}}\Big(\int_0^t \big|\beta^\epsilon(s)\big|^2 ds\Big)^{\frac{1}{2}} \\ &\le CN\Big(1+\Big(\sup_{r\in[0,t]}\big|\hat{u}_n^\epsilon(r)\big|_\rho^\rho\int_0^t \big|\hat{u}_n^\epsilon(s)\big|_\rho^\rho ds\Big)^{\frac{1}{2}}\Big) \le \frac{1}{4}\sup_{r\in[0,t]}\big|\hat{u}_n^\epsilon(r)\big|_\rho^\rho + K_{3,N}\Big(1+\int_0^t \big|\hat{u}_n^\epsilon(s)\big|_\rho^\rho ds\Big).\end{aligned}$$

Combining the foregoing with (7.6) and defining $K_{4,N} := K_1 + K_2 + K_{3,N}$, gives

$$\frac{1}{2}E^{Q^{\epsilon,\beta}}\Big[\sup_{s\in[0,t]}\big|\hat{u}_n^\epsilon(s)\big|_\rho^\rho\Big] \le \big|\xi\big|_\rho^\rho + K_{4,N} + K_{4,N}\int_0^t E^{Q^{\epsilon,\beta}}\Big[\sup_{r\in[0,s]}\big|\hat{u}_n^\epsilon(r)\big|_\rho^\rho\Big]ds.$$

Since $t \to E^{Q^{\epsilon,\beta}}\Big[\sup_{s\in[0,t]}\big|\hat{u}_n^\epsilon(s)\big|_\rho^\rho\Big]$ is continuous (this is due to $\hat{u}_n^\epsilon \in C\big([0,T]:L^\rho(D)\big)$ Q-a.s., and the monotone convergence theorem), Grönwall's inequality gives

$$E^{Q^{\epsilon,\beta}}\Big[\sup_{t\in[0,T]}\big|\hat{u}_n^\epsilon(t)\big|_\rho^\rho\Big] \le 2\Big(|\xi|_\rho^\rho + K_{4,N}\Big)e^{2K_{4,N}},$$

for all $n\in\mathbb{N}$. Since $\lim_{n\to\infty}\hat{u}_n^\epsilon = u^{\epsilon,\xi}$ in $C\big([0,T]:L^\rho(D)\big)$ $Q^{\epsilon,\beta}$-a.s., Fatou's Lemma gives

$$E^{Q^{\epsilon,\beta}}\Big[\sup_{t\in[0,T]}\big|u^{\epsilon,\xi}(t)\big|_\rho^\rho\Big] \le \liminf_{n\to\infty} E^{Q^{\epsilon,\beta}}\Big[\sup_{t\in[0,T]}\big|\hat{u}_n^\epsilon(t)\big|_\rho^\rho\Big] \le 2\Big(\big|\xi\big|_\rho^\rho + K_{4,N}\Big)e^{2K_{4,N}}.$$

Since the distribution of $u^{\epsilon,\xi}$ under $Q^{\epsilon,\beta}$ is the same as $u^{\epsilon,\xi,\beta}$ and ϵ was arbitrary, this completes the proof. □

Lemma 7.5. *For any $N>0$ and $R>0$*

$$\Big\{\sup_{t\in[0,T]}\big|z^{\epsilon,\xi,\beta}(t)\big|_\rho\Big\}_{\epsilon\in(0,1]}^{|\xi|_\rho\le R,\beta\in\mathcal{A}_2^N},$$

is bounded in probability, uniformly in $\epsilon\in(0,1]$, $|\xi|_\rho\le R$ and $\beta\in\mathcal{A}_2^N$.

Proof: Throughout this proof, the free positive constant C will neither depend on R nor on N. For convenience, define

$$\zeta_1^{\epsilon,\xi,\beta} := \sup_{t\in[0,T]}\big|u^{\epsilon,\xi,\beta}(s)\big|_\rho + \sup_{t\in[0,T]}\big|u^{0,\xi}(s)\big|_\rho.$$

Due to Lemma 7.4 and Lemma A.6, there exists a constant $K_{R,N}>0$, such that

$$\sup_{\epsilon\in(0,1]}\sup_{|\xi|_\rho\le R}\sup_{\beta\in\mathcal{A}_2^N} E\Big[\big(\zeta_1^{\epsilon,\xi,\beta}\big)^\rho\Big] \le K_{R,N}.$$

For a.e. $\omega\in\Omega$, we have for all $t\in[0,T]$

$$\begin{aligned} z^{\epsilon,\beta,\xi}(t,x) &= \int_0^t\int_D \partial_{y_i}G_{t-s}(x,y)\frac{1}{\gamma(\epsilon)\sqrt{\epsilon}}\Big(g_i\big(u^{0,\xi}+\gamma(\epsilon)\sqrt{\epsilon}z^{\epsilon,\beta,\xi}\big) - g_i\big(u^{0,\xi}\big)\Big)(s,y)dyds \\ &+ \int_0^t\int_D G_{t-s}(x,y)\frac{1}{\gamma(\epsilon)\sqrt{\epsilon}}\Big(f\big(u^{0,\xi}+\gamma(\epsilon)\sqrt{\epsilon}z^{\epsilon,\beta,\xi}\big) - f\big(u^{0,\xi}\big)\Big)(s,y)dyds \\ &+ \int_0^t\int_D G_{t-s}(x,y)\sigma_j\big(u^{\epsilon,\beta,\xi}\big)(s,y)dy\beta_j(s)ds \\ &+ \frac{1}{\gamma(\epsilon)}\int_0^t\int_D G_{t-s}(x,y)\sigma_j\big(u^{\epsilon,\beta,\xi}\big)(s,y)dydB_j(s) \\ &:= L_1^{\epsilon,\xi,\beta}(t,x) + L_2^{\epsilon,\xi,\beta}(t,x) + L_3^{\epsilon,\xi,\beta}(t,x) + \frac{1}{\gamma(\epsilon)}L_4^{\epsilon,\xi,\beta}(t,x), \end{aligned} \tag{7.7}$$

for a.e. $x\in D$. Next define

$$\tilde{L}_1^{\epsilon,\xi,\beta}(t,x) := \int_0^t\int_D \partial_{y_i}G_{t-s}(x,y)\Big(z^{\epsilon,\beta,\xi}g_i'\big(u^{0,\xi}\big)\Big)(s,y)dyds.$$

As in the proof of Theorem 3.1, it follows that

$$\begin{aligned} \Big|L_1^{\epsilon,\xi,\beta}(t,\cdot) - \tilde{L}_1^{\epsilon,\xi,\beta}(t,\cdot)\Big|_\rho &\le C\gamma(\epsilon)\sqrt{\epsilon}\int_0^t (t-s)^{-\frac{d}{2\rho}-\frac12}\big|z^{\epsilon,\beta,\xi}(s)\big|_\rho^2 ds \\ &\le C\zeta_1^{\epsilon,\xi,\beta}\int_0^t (t-s)^{-\frac{d}{2\rho}-\frac12}\big|z^{\epsilon,\beta,\xi}(s)\big|_\rho ds, \end{aligned}$$

and

$$\big|\tilde{L}_1^{\epsilon,\xi,\beta}(t,\cdot)\big|_\rho \le C\big(1+\zeta_1^{\epsilon,\xi,\beta}\big)\int_0^t (t-s)^{-\frac{d}{2\rho}-\frac{1}{2}}\big|z^{\epsilon,\beta,\xi}(s)\big|_\rho ds,$$

which gives

$$\big|L_1^{\epsilon,\xi,\beta}(t,\cdot)\big|_\rho \le C\big(1+\zeta_1^{\epsilon,\xi,\beta}\big)\int_0^t (t-s)^{-\frac{d}{2\rho}-\frac{1}{2}}\big|z^{\epsilon,\beta,\xi}(s)\big|_\rho ds.$$

A similar argument gives

$$\big|L_2^{\epsilon,\xi,\beta}(t,\cdot)\big|_\rho \le C\big(1+\zeta_1^{\epsilon,\xi,\beta}\big)\int_0^t (t-s)^{-\frac{d}{2\rho}}\big|z^{\epsilon,\beta,\xi}(s)\big|_\rho ds.$$

Similar to the treatment of the $J_3(t)$ term in the proof of Lemma 7.2, we have

$$\big|L_3^{\epsilon,\xi,\beta}(t,\cdot)\big|_\rho \le C\Big(1+\sup_{s\in[0,T]}\big|u^{\epsilon,\beta,\xi}(s)\big|_\rho\Big)\Big(\int_0^t \big|\beta(s)\big|^2 ds\Big)^{\frac{1}{2}} \le CN\big(1+\zeta_1^{\epsilon,\xi,\beta}\big),$$

where the second inequality used that $\beta\in\mathcal{A}_2^N$. Finally, consider $L_4^{\epsilon,\xi,\beta}(t,x)$. Due to Lemma 7.4, (A4), and Lemma A.5, it follows that $\Big\{\big|L_4^{\epsilon,\xi,\beta}(t)\big|_\rho\Big\}_{\epsilon\in(0,1]}^{|\xi|_\rho\le R,\beta\in\mathcal{A}_2^N}$ is tight in $C([0,T]:L^\rho(D))$, and consequently $\Big\{\sup_{t\in[0,T]}\big|L_4^{\epsilon,\xi,\beta}(t)\big|_\rho\Big\}_{\epsilon\in(0,1]}^{|\xi|_\rho\le R,\beta\in\mathcal{A}_2^N}$ is bounded in probability, uniformly in $\epsilon\in(0,1]$, $|\xi|_\rho\le R$ and $\beta\in\mathcal{A}_2^N$. For convenience, let

$$\zeta_2^{\epsilon,\xi,\beta} := \sup_{t\in[0,T]}\big|L_4^{\epsilon,\xi,\beta}(t)\big|_\rho.$$

Therefore

$$\begin{aligned}\big|z^{\epsilon,\beta,\xi}(t)\big|_\rho &\le \big|L_1^{\epsilon,\xi,\beta}(t)\big|_\rho+\big|L_2^{\epsilon,\xi,\beta}(t)\big|_\rho+\big|L_3^{\epsilon,\xi,\beta}(t)\big|_\rho+\frac{1}{\gamma(\epsilon)}\sup_{t\in[0,T]}\big|L_4^{\epsilon,\xi,\beta}(t)\big|_\rho\\ &\le CN\Big(1+\zeta_1^{\epsilon,\xi,\beta}+\frac{1}{\gamma(\epsilon)}\zeta_2^{\epsilon,\xi,\beta}\Big)+C\big(1+\zeta_1^{\epsilon,\xi,\beta}\big)\int_0^t c_t(s)\big|z^{\epsilon,\beta,\xi}(s)\big|_\rho ds,\end{aligned}$$

where

$$c_t(s) := (t-s)^{-\frac{d}{2\rho}}+(t-s)^{-\frac{d}{2\rho}-\frac{1}{2}}.$$

Consequently, Grönwall's inequality implies

$$\sup_{t\in[0,T]}|z^{\epsilon,\beta,\xi}(t)|_\rho \le CN\left(1+\zeta_1^{\epsilon,\xi,\beta}+\frac{1}{\gamma(\epsilon)}\zeta_2^{\epsilon,\xi,\beta}\right)e^{C(1+\zeta_1^{\epsilon,\xi,\beta})}.$$

Since $\lim_{\epsilon\to 0}\dfrac{1}{\gamma(\epsilon)}=0$ and both $\{\zeta_1^{\epsilon,\xi,\beta}\}_{\epsilon\in(0,1]}^{|\xi|_\rho\le R,\beta\in\mathcal{A}_2^N}$ and $\{\zeta_2^{\epsilon,\xi,\beta}\}_{\epsilon\in(0,1]}^{|\xi|_\rho\le R,\beta\in\mathcal{A}_2^N}$ are bounded in probability, uniformly in $\epsilon\in(0,1]$, $|\xi|_\rho\le R$ and $\beta\in\mathcal{A}_2^N$, the result follows and the proof is complete. □

8. Proof of Theorem 6.1

Theorem 4.3 provides that FWULDP and EULP are equivalent. To this effect, in order to prove an FWULDP, we will prove Theorem 4.5, which is an EULP, whose proof hinges upon showing that Assumption 4.4 holds in the $C\big([0,T]:L^\rho(D)\big)$ topology. The next theorem proves the holding of that assumption in the said topology.

Theorem 8.1 (Uniform Convergence in probability)**.** *For any $\lambda>0$, $R>0$, and $N>0$*

$$\lim_{\epsilon\to 0}\sup_{|\xi|_\rho\le R}\sup_{\beta\in\mathcal{A}_2^N} P\Big(\big|z^{\epsilon,\xi,\beta}-z^{0,\xi,\beta}\big|_{C([0,T]:L^\rho(D))}>\lambda\Big)=0.$$

Proof: Throughout this proof, C denotes a free positive constant, which does not depend on R or N. Note that for all $t \in [0,T]$ and a.e. $x \in D$,

$$\begin{aligned} z^{\epsilon,\xi,\beta}(t,x) - z^{0,\xi,\beta}(t,x) &= J_1^{\epsilon,\xi,\beta}(t,x) - \tilde{J}_1^{\xi,\beta}(t,x) + J_2^{\epsilon,\xi,\beta}(t,x) - \tilde{J}_2^{\xi,\beta}(t,x) \\ &\quad + J_3^{\epsilon,\xi,\beta}(t,x) - \tilde{J}_3^{\xi,\beta}(t,x) + \frac{1}{\gamma(\epsilon)} J_4^{\epsilon,\xi,\beta}(t,x), \end{aligned}$$

where

$$J_1^{\epsilon,\xi,\beta}(t,x) := -\int_0^t \int_D \partial_{y_i} G_{t-s}(x,y) \Big(\frac{1}{\gamma(\epsilon)\sqrt{\epsilon}} \Big(g_i(u^{0,\xi} + \gamma(\epsilon)\sqrt{\epsilon} z^{\epsilon,\xi,\beta}) - g_i(u^{0,\xi}) \Big) \Big)(s,y) dy ds,$$

$$\tilde{J}_1^{\xi,\beta}(t,x) := -\int_0^t \int_D \partial_{y_i} G_{t-s}(x,y) \Big(z^{0,\xi,\beta} g_i'(u^{0,\xi}) \Big)(s,y) dy ds,$$

$$J_2^{\epsilon,\xi,\beta}(t,x) := \int_0^t \int_D G_{t-s}(x,y) \Big(\frac{1}{\gamma(\epsilon)\sqrt{\epsilon}} \Big(f(u^{0,\xi} + \gamma(\epsilon)\sqrt{\epsilon} z^{\epsilon,\xi,\beta}) - f(u^{0,\xi}) \Big) \Big)(s,y) dy ds,$$

$$\tilde{J}_2^{\xi,\beta}(t,x) := \int_0^t \int_D G_{t-s}(x,y) \Big(z^{0,\xi,\beta} f'(u^{0,\xi}) \Big)(s,y) dy ds,$$

$$J_3^{\epsilon,\xi,\beta}(t,x) := \int_0^t \int_D G_{t-s}(x,y) \sigma_j \big(u^{0,\xi} + \gamma(\epsilon)\sqrt{\epsilon} z^{\epsilon,\xi,\beta} \big)(s,y) dy \beta_j(s) ds,$$

$$\tilde{J}_3^{\xi,\beta}(x,t) := \int_0^t \int_D G_{t-s}(x,y) \sigma_j \big(u^{0,\xi} \big)(s,y) dy \beta_j(s) ds,$$

and

$$J_4^{\epsilon,\xi,\beta}(t,x) := \int_0^t \int_D G_{t-s}(x,y) \sigma_j \big(u^{0,\xi} + \gamma(\epsilon)\sqrt{\epsilon} z^{\epsilon,\xi,\beta} \big)(s,y) dy dB_j(s).$$

Note that for all $t \in [0,T]$, Minkowski's inequality implies

$$\begin{aligned} \big| z^{\epsilon,\xi,\beta}(t) - z^{0,\xi,\beta}(t) \big|_\rho &\le \big| J_1^{\epsilon,\xi,\beta}(t) - \tilde{J}_1^{\xi,\beta}(t) \big|_\rho + \big| J_2^{\epsilon,\xi,\beta}(t) - \tilde{J}_2^{\xi,\beta}(t) \big|_\rho \\ &\quad + \big| J_3^{\epsilon,\xi,\beta}(t) - \tilde{J}_3^{\xi,\beta}(t) \big|_\rho + \frac{1}{\gamma(\epsilon)} \big| J_4^{\epsilon,\xi,\beta}(t) \big|_\rho. \end{aligned} \tag{8.1}$$

We reiterate that, as a matter of notation, any display even if not made explicit will hold true with probability one. We now aim to find desired bounds on the RHS of (8.1). To this end, we first attend to $\big| J_1^{\epsilon,\xi,\beta}(t) - \tilde{J}_1^{\xi,\beta}(t) \big|_\rho + \big| J_2^{\epsilon,\xi,\beta}(t) - \tilde{J}_2^{\xi,\beta}(t) \big|_\rho$. For all $t \in [0,T]$ and $x \in D$, define

$$L_1^{\epsilon,\xi,\beta}(t,x) := \int_0^t \int_D \partial_{y_i} G_{t-s}(x,y) \Big(z^{\epsilon,\xi,\beta} g_i'(u^{0,\xi}) \Big)(s,y) dy ds,$$

and observe that

$$\big| J_1^{\epsilon,\xi,\beta}(t) - \tilde{J}_1^{\xi,\beta}(t) \big|_\rho \le \big| J_1^{\epsilon,\xi,\beta}(t) - L_1^{\epsilon,\xi,\beta}(t) \big|_\rho + \big| L_1^{\epsilon,\xi,\beta}(t) - \tilde{J}_1^{\xi,\beta}(t) \big|_\rho. \tag{8.2}$$

Applying the same argument as in the proof of Theorem 3.1 to bound $\big| L_1^{\epsilon}(z,t) - \tilde{L}_1(z,t) \big|_\rho$, gives

$$\big| J_1^{\epsilon,\xi,\beta}(t) - L_1^{\epsilon,\xi,\beta}(t) \big|_\rho \le C\gamma(\epsilon)\sqrt{\epsilon} \sup_{s\in[0,t]} \big| z^{\epsilon,\xi,\beta}(s) \big|_\rho^2. \tag{8.3}$$

Next we use the boundedness of the initial condition and the same argument as in the proof of Theorem 3.1 to bound $|\tilde{L}_1(z,t)|_\rho$, which yields

$$\big| L_1^{\epsilon,\xi,\beta}(t) - \tilde{J}_1^{\xi,\beta}(t) \big|_\rho \le C\big(1 + R^\rho\big) \int_0^t (t-s)^{\frac{d}{2\rho} - \frac{1}{2}} \big| z^{\epsilon,\xi,\beta}(s) - z^{0,\xi,\beta}(s) \big|_\rho ds. \tag{8.4}$$

Next define

$$c_t(s) := (t-s)^{-\frac{d}{2\rho}-\frac{1}{2}} + (t-s)^{-\frac{d}{2\rho}}, \quad \forall\ 0 \le s < t \le T,$$

and

$$\zeta_1^{\epsilon,\xi,\beta} := \Big(1 + \sup_{t\in[0,T]} \big|z^{\epsilon,\xi,\beta}(t)\big|_\rho\Big)^2.$$

Note that due to Lemma 7.5, $\{\zeta_1^{\epsilon,\xi,\beta}\}_{\epsilon\in(0,1]}^{|\xi|_\rho\le R,\beta\in\mathcal{A}_2^N}$ is bounded in probability, uniformly in $\epsilon \in (0,1]$, $|\xi|_\rho \le R$ and $\beta \in \mathcal{A}_2^N$. In view of (8.2), and combining (8.3) and (8.4)

$$\big|J_1^{\epsilon,\xi,\beta}(t) - \tilde{J}_1^{\xi,\beta}(t)\big|_\rho \le C\gamma(\epsilon)\sqrt{\epsilon}\zeta_1^{\epsilon,\xi,\beta} + C\big(1+R^\rho\big)\int_0^t c_t(s)\big|z^{\epsilon,\xi,\beta}(s) - z^{0,\xi,\beta}(s)\big|_\rho ds. \tag{8.5}$$

Similarly, we get

$$\big|J_2^{\epsilon,\xi,\beta}(t) - \tilde{J}_2^{\xi,\beta}(t)\big|_\rho \le C\gamma(\epsilon)\sqrt{\epsilon}\zeta_1^{\epsilon,\xi,\beta} + C\big(1+R^\rho\big)\int_0^t c_t(s)\big|z^{\epsilon,\xi,\beta}(s) - z^{0,\xi,\beta}(s)\big|_\rho ds. \tag{8.6}$$

It remains to bound $\big|J_3^{\epsilon,\xi,\beta}(t) - \tilde{J}_3^{\xi,\beta}(t)\big|_\rho + \dfrac{1}{\gamma(\epsilon)}\big|J_4^{\epsilon,\xi,\beta}(t)\big|_\rho$. We first attend to $\big|J_3^{\epsilon,\xi,\beta}(t) - \tilde{J}_3^{\xi,\beta}(t)\big|_\rho$. Applying the same argument as in the proof of Theorem 7.3 to bound $J_3^n(t)$, and that $\beta \in \Lambda^N$ a.s., we have

$$\big|J_3^{\epsilon,\xi,\beta}(t) - \tilde{J}_3^{\xi,\beta}(t)\big|_\rho \le CN\gamma(\epsilon)\sqrt{\epsilon}\sup_{s\in[0,t]}\big|z^{\epsilon,\xi,\beta}(s)\big|_\rho \le CN\gamma(\epsilon)\sqrt{\epsilon}\zeta_1^{\epsilon,\xi,\beta}. \tag{8.7}$$

Note that due to Lemma 7.4 and Lemma A.5, $\big\{J_4^{\epsilon,\xi,\beta}(t,x)\big\}_{\epsilon\in(0,1]}^{|\xi|_\rho\le R,\beta\in\mathcal{A}_2^N}$ is tight in $C\big([0,T] : L^\rho(D)\big)$. Define

$$\zeta_2^{\epsilon,\xi,\beta} := \sup_{t\in[0,T]} |J_4^{\epsilon,\xi,\beta}(t)|_\rho.$$

Consequently, $\big\{\zeta_2^{\epsilon,\xi,\beta}\big\}_{\epsilon\in(0,1]}^{|\xi|_\rho\le R,\beta\in\mathcal{A}_2^N}$ is bounded in probability, uniformly in $\epsilon \in (0,1]$, $|\xi|_\rho \le R$ and $\beta \in \mathcal{A}_2^N$. Finally, observe that in view of (8.1), and by combining (8.5), (8.6) and (8.7) we have

$$\begin{aligned}\big|z^{\epsilon,\xi,\beta}(t) - z^{0,\xi,\beta}(t)\big|_\rho \le C(1+N)\Big(\gamma(\epsilon)\sqrt{\epsilon}\zeta_1^{\epsilon,\xi,\beta} + \frac{1}{\gamma(\epsilon)}\zeta_2^{\epsilon,\xi,\beta}\Big)& \\ + C\big(1+R^\rho\big)\int_0^t c_t(s)\big|z^{\epsilon,\xi,\beta}(s) - z^{0,\xi,\beta}(s)\big|_\rho ds.&\end{aligned}$$

Consequently, Grönwall's inequality implies

$$\sup_{t\in[0,T]}\big|z^{\epsilon,\xi,\beta}(t) - z^{0,\xi,\beta}(t)\big|_\rho \le C(1+N)\Big(\gamma(\epsilon)\sqrt{\epsilon}\zeta_1^{\epsilon,\xi,\beta} + \frac{1}{\gamma(\epsilon)}\zeta_2^{\epsilon,\xi,\beta}\Big)e^{C\left(1+R^\rho\right)}.$$

Therefore, for any $\lambda > 0$

$$\begin{aligned}\sup_{|\xi|_\rho\le R}\sup_{\beta\in\mathcal{A}_2^N} P\Big(\sup_{t\in[0,T]}\big|z^{\epsilon,\xi,\beta}(t) - z^{0,\xi,\beta}(t)\big|_\rho > \lambda\Big) \le \sup_{|\xi|_\rho\le R}\sup_{\beta\in\mathcal{A}_2^N} P\Big(\zeta_1^{\epsilon,\xi,\beta} > \frac{\lambda'}{\gamma(\epsilon)\sqrt{\epsilon}}\Big)& \\ + \sup_{|\xi|_\rho\le R}\sup_{\beta\in\mathcal{A}_2^N} P\Big(\zeta_2^{\epsilon,\xi,\beta} > \lambda'\gamma(\epsilon)\Big),&\end{aligned}$$

where $\lambda' := \dfrac{\lambda}{2C(1+N)e^{C(1+R^\rho)}}$. Note that by Lemma 7.5, $\{\zeta_1^{\epsilon,\xi,\beta}\}_{\epsilon\in(0,1]}^{|\xi|_\rho\le R,\beta\in\mathcal{A}_2^N}$ is bounded in probability, uniformly in $\epsilon\in(0,1]$, $|\xi|_\rho\le R, \beta\in\mathcal{A}_2^N$. Lemma 7.4 and Lemma A.5 imply the same for $\{\zeta_2^{\epsilon,\xi,\beta}\}_{\epsilon\in(0,1]}^{|\xi|_\rho\le R,\beta\in\mathcal{A}_2^N}$. Therefore, since $\gamma(\epsilon)\sqrt{\epsilon}\to 0$ and $\dfrac{1}{\gamma(\epsilon)}\to 0$ as $\epsilon\to 0$, we have

$$\lim_{\epsilon\to 0}\sup_{|\xi|_\rho\le R}\sup_{\beta\in\mathcal{A}_2^N} P\left(\sup_{t\in[0,T]}\left|z^{\epsilon,\xi,\beta}(t)-z^{0,\xi,\beta}(t)\right|_\rho>\lambda\right)=0.$$

This completes the proof.

□

Proof of Theorem 6.1. Theorem 6.1 is a direct consequence of Theorem 4.5 and Theorem 8.1.

Proof of Corollary 6.2. Due to Theorem 6.1 and Salins (2019, Theorem 2.5), it suffices to show that for any compact set $K\subset L^\rho(D)$ and $s\ge 0$

$$\cup_{\xi\in K}\{\psi : I_\xi(\psi)\le s\},$$

is a pre-compact subset of $C([0,T]:L^\rho(D))$. This follows from Theorem 7.3.

Appendix A.

A.1. *Related Results.* The following definitions and results are used in proving the main theorems of the paper.

For notational convenience, the following definition is introduced.

Definition A.1. For $p>d+2$, $j\in\{1,...,k\}$, and an $\mathcal{F}_t$-adapted random process $h(t)\in C([0,T]: L^p(D))$, define

$$H_j(h)(t,x) := \int_0^t\int_D G_{t-s}(x,y)h(s,y)dydB_j(s),$$

for all $t\in[0,T]$ and $x\in D$.

It follows from Hölder's inequality and (E1) that $H_j(h)(t,x)$ is well defined (see Gyöngy and Rovira, 2000, Lemma 3.3). The next two lemmas will be imperative when proving tightness and precompactness and are versions of Gyöngy and Rovira (2000, Lemma 3.3) and Gyöngy and Rovira (2000, Lemma 3.1) that have been slightly adjusted to address what is needed here.

Lemma A.2. *Let $p>d+2$, $j\in\{1,...,k\}$, and $h(t)$ be an $\mathcal{F}_t$-adapted random process in $C([0,T]: L^p(D))$, which satisfies $E\Big[\sup_{t\in[0,T]}|h(t)|_p^p\Big]<\infty$. Let $H_j(h)$ be given by Definition A.1. Then, the following hold:*

(a) There exists a constant $C>0$, such that

$$E\left[\left|H_j(h)(t,x)\right|^p\right]\le Ct^{\frac{1}{4}(p-d)}E\Big[\sup_{r\in[0,T]}|h(r)|_p^p\Big],$$

for all $t\in[0,T]$ and $x\in D$.

(b) For any $\tau\in\left(0,\frac{1}{2}(p-d-2)\right)$, there exists a constant $C>0$, such that

$$E\left[\left|H_j(h)(t,x)-H_j(h)(s,x)\right|^p\right]\le C|t-s|^{1+\tau}E\Big[\sup_{r\in[0,T]}|h(r)|_p^p\Big],$$

for all $s,t\in[0,T]$ and $x\in D$.

(c) For any $\alpha \in (0, p-d)$, there exists a constant $C > 0$, such that

$$E\left[\left|H_j(h)(t,x) - H_j(h)(t,y)\right|^p\right] \le C\left|x-y\right|^\alpha E\Big[\sup_{r\in[0,T]}\left|h(r)\right|_p^p\Big],$$

for all $x, y \in D$ and $t \in [0,T]$.

Proof: Since we assumed $p > d+2$, all claims follow directly from Lemma 3.3 in Gyöngy and Rovira (2000) by (using their notiation) setting $m = p$, choosing γ sufficiently large, and observing that $\kappa = 1 - \frac{d}{2p}$, $\theta = \frac{1}{2}$, and $\epsilon = 1$ (note that there is a typo in part (1) of the statement of Lemma 3.3 in Gyöngy and Rovira (2000) and it can be verified by checking the proof that the exponent on t should be $\kappa_p - \frac{1}{2} - \frac{1}{2\gamma}$ rather than $\kappa_p - \frac{1}{2} - \frac{1}{\gamma}$). □

Note that Lemma A.2 implies there is a modification of $H_j(h)(t,x)$, which is an $L^p(D)$-valued, $\mathcal{F}_t$-adapted random process.

Lemma A.3. *For $p > d+2$ and $h(s) \in L^2\big([0,T] : L^p(D)\big)$, define*

$$J(h)(t,x) := \int_0^t \int_D G_{t-s}(x,y)h(s,y)dyds,$$

for all $x \in D$ and $t \in [0,T]$. Then $J(h)(t,x)$ satisfies the following:

(a) For any $\tau \in (0, \frac{1}{2})$, there exists a constant $C > 0$, such that

$$\left|J(h)(t) - J(h)(s)\right|_p \le C\left|t-s\right|^\tau\left(\int_0^T |h(r)|_p^2 dr\right)^{\frac{1}{2}},$$

for all $s, t \in [0,T]$.

(b) For any $\alpha \in \big(0, \dfrac{p-d}{p}\big)$, there exists a constant $C > 0$, such that

$$\left|J(h)(t,x) - J(h)(t,y)\right| \le C\left|x-y\right|^\alpha\left(\int_0^T |h(r)|_p^2 dr\right)^{\frac{1}{2}},$$

for all $x, y \in D$ and $t \in [0,T]$.

Proof: Part (a) follows from Gyöngy and Rovira (2000, Lemma 3.1) by setting both q and ρ in their notation equal to our p, which results in $\kappa = 1$ and $\epsilon = 1$ in their notation. To prove part (b) note that for all $x, y, z \in D$

$$\begin{aligned}
&\Big|G_{t-s}(x,z) - G_{t-s}(y,z)\Big| \\
&\le \bigg(\left|G_{t-s}(x,z)\right|^{1-\alpha} + \left|G_{t-s}(y,z)\right|^{1-\alpha}\bigg)\Big|G_{t-s}(x,z) - G_{t-s}(y,z)\Big|^\alpha \\
&\le \bigg(\left|G_{t-s}(x,z)\right|^{1-\alpha} + \left|G_{t-s}(y,z)\right|^{1-\alpha}\bigg)\left|x-y\right|^\alpha\bigg(\int_0^1 b\Big(t-s, y+\lambda(x-y)-z\Big)d\lambda\bigg)^\alpha \\
&\le \left|x-y\right|^\alpha\bigg(a(t-s,x-z)^{1-\alpha} + a(t-s,y-z)^{1-\alpha}\bigg)\bigg(\int_0^1 b\Big(t-s, y+\lambda(x-y)-z\Big)d\lambda\bigg)^\alpha, \quad \text{(A.1)}
\end{aligned}$$

where the second line used the inequality $|a-b| \le \big(|a|^{1-\alpha} + |b|^{1-\alpha}\big)|a-b|^\alpha$ for any $a, b \in \mathbb{R}$ and $\alpha \in [0,1]$, the third line used (E2), and the fourth line used (E1). In addition

$$\left|\Big(a(t-s,x-\cdot)^{1-\alpha}+a(t-s,y-\cdot)^{1-\alpha}\Big)\bigg(\int_0^1 b\Big(t-s,y+\lambda(x-y)+\cdot\Big)d\lambda\bigg)^{\alpha}\right|_{\frac{p}{p-1}}$$
$$\leq\bigg(\Big|(a(t-s,x-\cdot)\Big|^{1-\alpha}_{\frac{p}{p-1}}+\Big|a(t-s,y-\cdot)\Big|^{1-\alpha}_{\frac{p}{p-1}}\bigg)\bigg(\int_0^1\Big|b(t-s,y+\lambda(x-y)-\cdot)\Big|_{\frac{p}{p-1}}d\lambda\bigg)^{\alpha}$$
$$\leq C\big(t-s\big)^{-\frac{d}{2p}(1-\alpha)+(-\frac{d}{2p}-\frac{1}{2})\alpha}=C\big(t-s\big)^{-\frac{d}{2p}-\frac{\alpha}{2}}, \tag{A.2}$$

where Hölder's inequality, Minkowski's inequality, (E1), and (E3) were used. Therefore

$$\Big|J(h)(t,x)-J(h)(t,y)\Big|$$
$$\leq\int_0^t\Big|\big(G_{t-s}(x,\cdot)-G_{t-s}(y,\cdot)\big)h(s,\cdot)\Big|_1 ds\leq\int_0^t\Big|G_{t-s}(x,\cdot)-G_{t-s}(y,\cdot)\Big|_{\frac{p}{p-1}}\big|h(s)\big|_p ds$$
$$\leq C\big|x-y\big|^{\alpha}\int_0^t\big(t-s\big)^{-\frac{d}{2p}-\frac{\alpha}{2}}\big|h(s)\big|_\rho ds\leq C\big|x-y\big|^{\alpha}\bigg(\int_0^T\big|h(s)\big|_p^2 ds\bigg)^{\frac{1}{2}}.$$

This proves (b), where Hölder's inequality was used twice and the third inequality used (A.1) and (A.2). ☐

We will rely on the following version of the Kolmogorov continuity theorem.

Theorem A.4 (Kolmogorov Continuity Theorem)**.** *Let $h(t)$ be an $L^p(D)$-valued, $\mathcal{F}_t$-adapted stochastic process and assume there exist constants $\alpha,\tau>0$ and $C>0$, such that for all $s,t\in[0,T]$ we have*

$$E\Big[\big|h(t)-h(s)\big|_p^{\alpha}\Big]\leq C\big|t-s\big|^{1+\tau}. \tag{A.3}$$

Then, for any $\lambda\in(0,\dfrac{\tau}{\alpha})$, there exists a λ-Hölder continuous, $\mathcal{F}_t$-adapted modification $\tilde{h}(t)$. In addition, if we let Υ_λ represent the corresponding (random) λ-Hölder constant of $\tilde{h}(t)$, meaning with probability 1 for all $t,s\in[0,T]$, we have

$$\big|\tilde{h}(t)-\tilde{h}(s)\big|_p\leq\Upsilon_\lambda\big|t-s\big|^{\lambda},$$

then, there exists a (deterministic) constant $K_{T,\alpha,\tau,\lambda}$ (which does not depend on C), such that

$$E\Big[\big(\Upsilon_\lambda\big)^{\alpha}\Big]\leq K_{T,\alpha,\tau,\lambda}C.$$

Proof: That there exists a λ-Hölder continuous, $\mathcal{F}_t$-adapted modfication comes form the standard Kolmogorov continuity theorem (see Da Prato and Zabczyk, 2014, Theorem 3.3). It remains to address the λ-Hölder constant. Using a version of the Garsia, Rademich, and Rumsey lemma (see Garsia et al., 1970/71, Lemma 1.1), there exists a constant $\tilde{K}_{T,\alpha,\tau,\lambda}$, such that with probability 1 (specifically, on the set where $\tilde{h}(t)$ is continuous)

$$\Upsilon_\lambda=\sup_{s,t\in[0,T]}\frac{\big|\tilde{h}(t)-\tilde{h}(s)\big|_p}{|t-s|^{\lambda}}\leq\tilde{K}_{T,\alpha,\tau,\lambda}\bigg(\int_0^T\int_0^T\frac{\big|h(t)-h(s)\big|_p^{\alpha}}{|t-s|^{2+\alpha\lambda}}dtds\bigg)^{\frac{1}{\alpha}}.$$

Raising both sides to the power of α, taking expectations, and recalling (A.3), gives

$$E\big[(\Upsilon_\lambda)^{\alpha}\big]\leq K_{T,\alpha,\tau,\lambda}C.$$

☐

The next lemma will be used to prove tightness of the stochastic integral terms in both the proof of the FCLT and the proof of the UMDP.

Lemma A.5. *Let $p > d+2$ and $\{h^{\alpha}(t)\}_{\alpha\in A}$ be a collection of $\mathcal{F}_t$-adapted random processes in $C([0,T]:L^p(D))$ indexed by some $\alpha \in A$, which satisfy*

$$\sup_{\alpha\in A}\left\{E\left[\sup_{t\in[0,T]}|h^{\alpha}(t)|_p^p\right]\right\} < \infty. \tag{A.4}$$

For arbitrary $j \in \{1,...,k\}$, let $H_j(h^{\alpha})(t,x)$ be given by Definition A.1 for all $\alpha \in A$. Then, $\{H_j(h^{\alpha})(t)\}_{\alpha\in A}$ is a collection of $\mathcal{F}_t$-adapted random processes in $C\big([0,T]:L^p(D)\big)$, which is tight in $C\big([0,T]:L^p(D)\big)$.

Proof: Note that due to Lemma A.2, $\{H_j(h^{\alpha})(t)\}_{\alpha\in A}$ is a collection of $\mathcal{F}_t$-adapted, $L^p(D)$-valued random processes. Using Lemma A.2 part (b), combined with (A.4) implies that for any $\tau \in \big(0,\frac{1}{2}(p-d-2)\big)$, there exists some $C>0$, such that

$$\sup_{\alpha\in A}\left\{E\left[\Big|H_j(h^{\alpha})(t)-H_j(h^{\alpha})(s)\Big|_p^p\right]\right\} \le C\big|t-s\big|^{1+\tau},$$

for all $s,t\in[0,T]$. This, combined with Theorem A.4 implies that for an arbitrary $\lambda\in(0,\frac{\tau}{p})$ and all $\alpha\in A$, $H_j(h^{\alpha})(t)\in C^{\lambda}\big([0,T]:L^p(D)\big)$ and

$$\sup_{\alpha\in A}\left\{E\big[(\Upsilon_{\lambda}^{\alpha})^p\big]\right\} < \infty, \tag{A.5}$$

where for all $\alpha\in A$, $\Upsilon_{\lambda}^{\alpha}$ is the (random) λ-Hölder constant, which for a.e. $\omega\in\Omega$ satisfies

$$\Big|H_j(h^{\alpha})(t)-H_j(h^{\alpha})(s)\Big|_p \le \Upsilon_{\lambda}^{\alpha}\big|t-s\big|^{\lambda},$$

for all $s,t\in[0,T]$.

Recall that for any $\lambda^* < \lambda$, a set $\mathcal{G}\subset C^{\lambda^*}\big([0,T]:L^p(D)\big)$ is precompact if there exists a constant $C>0$, such that for all $g\in\mathcal{G}$ and $s,t\in[0,T]$, we have $\big|g(s)-g(t)\big|_p \le C|t-s|^{\lambda}$ and there is a countable, dense subset $\mathcal{D}\subset[0,T]$, such that for all $t\in\mathcal{D}$

$$\mathcal{G}(t) := \Big\{g(t) : g\in\mathcal{G}\Big\},$$

is precompact in $L^p(D)$.

Due to the Kolmogorov-Riesz compactness theorem, $\mathcal{H}\subset L^p(D)$ is precompact if

(i) $\sup_{h\in\mathcal{H}}\big|h\big|_p < \infty$,

(ii) $\limsup_{|y|\to 0}\sup_{h\in\mathcal{H}}\big|\tilde{h}(\cdot+y)-\tilde{h}(\cdot)\big|_p = 0$, where $\tilde{h}(x) := \mathcal{I}_D(x)h(x)$ and in the above equation the L^p-norm is on $\mathbb{R}^d$ rather than on D.

We will use the following sufficient condition for (ii). Define the $\mathcal{C}_c^{\infty}(\mathbb{R}^d)$ mollifier

$$\eta(x) := \begin{cases} C_{\eta}\exp\left(\frac{1}{|x|^2-1}\right) & \text{if } |x|\le 1\\ 0 & \text{if } |x|>1,\end{cases}$$

where the constant C_{η} is chosen so that $\int_{\mathbb{R}^d}\eta(x)dx = 1$. For all $r>0$ and $f\in L^p(\mathbb{R}^d)$, let

$$\eta_r(x) := \frac{1}{r^d}\eta\Big(\frac{x}{r}\Big), \text{ and } f^r(x) := \int_{\mathbb{R}^d} f(x-z)\eta_r(z)dz.$$

Then (i) in combination with

$$\lim_{r\to 0}\sup_{h\in\mathcal{H}}\big|\tilde{h}^r-\tilde{h}\big|_p = 0, \tag{A.6}$$

implies (ii) and that $\mathcal{H}$ is precompact in $L^p(D)$.
To see this, assume (A.6) and (i) so that $\sup_{h\in\mathcal{H}}|h|_p = C_{\mathcal{H}} < \infty$. Let $\epsilon > 0$ be arbitrary and note that, we can choose $r > 0$ sufficiently small, such that

$$\sup_{h\in\mathcal{H}}\left|\tilde{h}^r - \tilde{h}\right|_p \le \frac{\epsilon}{3},$$

and since $\eta_r \in \mathcal{C}_c^\infty(\mathbb{R}^d)$, we can choose $\delta > 0$ sufficiently small, such that

$$\sup_{x\in\mathbb{R}^d, |y|\le\delta}\left|\eta_r(x) - \eta_r(x+y)\right| \le \frac{\epsilon}{3C_{\mathcal{H}}|B_{r+\delta}|}, \tag{A.7}$$

where $B_{r+\delta}$ is the ball with radius $r+\delta$ centered at the origin and $|B_{r+\delta}| := \int_{B_{r+\delta}} dx$. Then, for any $|y| \le \delta$ and $h \in \mathcal{H}$, we have

$$\begin{aligned}
\left|\tilde{h}(\cdot) - \tilde{h}(\cdot+y)\right|_p &\le \left|\tilde{h}(\cdot) - \tilde{h}^r(\cdot)\right|_p + |\tilde{h}(\cdot+y) - \tilde{h}^r(\cdot+y)|_p + \left|\tilde{h}^r(\cdot) - \tilde{h}^r(\cdot+y)\right|_p \\
&\le \frac{2}{3}\epsilon + \left(\int_{\mathbb{R}^d}\left|\int_{B_{r+\delta}}\tilde{h}(x-z)\big(\eta_r(z) - \eta_r(z+y)\big)dz\right|^p dx\right)^{\frac{1}{p}} \\
&\le \frac{2}{3}\epsilon + \frac{\epsilon}{3C_{\mathcal{H}}|B_{r+\delta}|}\left(\int_{\mathbb{R}^d}\left|\int_{B_{r+\delta}}\left|\tilde{h}(x-z)\right|dz\right|^p dx\right)^{\frac{1}{p}} \\
&\le \frac{2}{3}\epsilon + \frac{\epsilon}{3C_{\mathcal{H}}|B_{r+\delta}|}|B_{r+\delta}|^{1-\frac{1}{p}}\left(\int_{\mathbb{R}^d}\int_{B_{r+\delta}}\left|\tilde{h}(x-z)\right|^p dz dx\right)^{\frac{1}{p}} \le \epsilon,
\end{aligned}$$

where the third line used (A.7) and the fourth used Hölder's inequality.

Consequently, the tightness of $\left\{H_j(h^\alpha)(t,x)\right\}_{\alpha\in A}$ in $C^{\lambda^*}\big([0,T] : L^p(D)\big)$ for an arbitrary $\lambda^* < \lambda$ will follow from Prokhorov's theorem if we can show that for any $\delta > 0$, the following hold:

(a) There exists a constant $C > 0$, such that for all $\alpha \in A$

$$P\Big(\left|H_j(h^\alpha)(s) - H_j(h^\alpha)(t)\right|_p \le C|t-s|^\lambda \text{ for all } s,t \in [0,T]\Big) \ge 1-\delta.$$

(b) For all $t \in [0,T]$, there exists a constant $C > 0$, such that for all $\alpha \in A$

$$P\Big(\left|H_j(h^\alpha)(t)\right|_p \le C\Big) \ge 1-\delta.$$

(c) For all $t \in [0,T]$ and $n \in \mathbb{N}$, there exists an $r > 0$, such that for all $\alpha \in A$

$$P\left(\int_{\mathbb{R}^d}\left|\mathcal{I}_D(x)H_j(h^\alpha)(t,x) - \int_{B_r}\mathcal{I}_D(x-z)H_j(h^\alpha)(t,x-z)\eta_r(z)dz\right|^p dx \le \frac{1}{n}\right) \ge 1-\delta.$$

Note that (a) follows directly from (A.5) and (b) follows from Lemma A.2 part (a), combined with (A.4). For (c), note that due to Lemma A.2 parts (a) and (c), combined with (A.4) and the fact that $p-d > 2$, there exists a constant $C > 0$, such that for all $\alpha \in A$, $x, y \in D$, and $t \in [0,T]$, we have

$$E\Big[\left|H_j(h^\alpha)(t,x)\right|^p\Big] \le C \text{ and } E\Big[\left|H_j(h^\alpha)(t,x) - H_j(h^\alpha)(t,y)\right|^p\Big] \le C|x-y|^2. \tag{A.8}$$

Let

$$\tilde{D}_r^i := \{x \in D : \operatorname{dist}(x,\partial D) > r\}, \;\; |\tilde{D}_r^i| := \int_{\tilde{D}_r^i} dx, \;\; \tilde{D}_r^b := \{x \in \mathbb{R}^d : \operatorname{dist}(x,\partial D) \le r\}, \;\; |D| := \int_D dx,$$

and observe that $\lim_{r\to 0}|\tilde{D}_r^b| = 0$, so we can choose $r > 0$ sufficiently small, such that

$$Cr^2|D| + 2^{p+1}C|\tilde{D}_r^b| \le \frac{\delta}{n}.$$

Then, we have

$$
\begin{aligned}
&E\left[\int_{\mathbb{R}^d}\left|\mathcal{I}_D(x)H_j(h^\alpha)(t,x)-\int_{B_r}\mathcal{I}_D(x-z)H_j(h^\alpha)(t,x-z)\eta_r(z)dz\right|^p dx\right]\\
&\le E\left[\int_{\tilde{D}_r^i}\left|\int_{B_r}\Big(H_j(h^\alpha)(t,x)-H_j(h^\alpha)(t,x-z)\Big)\eta_r(z)dz\right|^p dx\right]\\
&+2^pE\left[\int_{\tilde{D}_r^b}\left|\mathcal{I}_D(x)H_j(h^\alpha)(t,x)\right|^p dx\right]+2^pE\left[\int_{\tilde{D}_r^b}\left|\int_{B_r}\mathcal{I}_D(x-z)H_j(h^\alpha)(t,x-z)\eta_r(z)dz\right|^p dx\right]\\
&\le \int_{B_r}\int_{\tilde{D}_r^i}E\left[\left|H_j(h^\alpha)(t,x)-H_j(h^\alpha)(t,x-z)\right|^p dx\right]\eta_r(z)dx+2^p\int_{\tilde{D}_r^b}E\left[\left|\mathcal{I}_D(x)H_j(h^\alpha)(t,x)\right|^p\right]dx\\
&+2^p\int_{B_r}\int_{\tilde{D}_r^b}E\left[\left|\mathcal{I}_D(x-z)H_j(h^\alpha)(t,x-z)\right|^p\right]dx\eta_r(z)dz\le Cr^2|D|+2^{p+1}C|\tilde{D}_r^b|\le\frac{\delta}{n},
\end{aligned}
$$

where the second inequality used Jensen's inequality and the third used (A.8).

Markov's inequality then gives

$$
P\left(\int_{\mathbb{R}^d}\left|\mathcal{I}_D(x)H_j(h^\alpha)(t,x)-\int_{B_r}\mathcal{I}_D(x-z)H_j(h^\alpha)(t,x-z)\eta_r(z)dz\right|^p dx>\frac{1}{n}\right)\le n\frac{\delta}{n}=\delta,
$$

which gives (c) and completes the proof that $\{H_j(h^\alpha)(t,x)\}_{\alpha\in A}$ is tight in $C^{\lambda^*}\big([0,T]:L^p(D)\big)$ and consequently also tight in $C\big([0,T]:L^p(D)\big)$. □

The following lemma is a direct consequence of Gyöngy and Rovira (2000, Proposition 4.4) and will allow us to use Lemmas A.2, A.3, and A.5.

Lemma A.6. *There exists a constant $C>0$, such that for all $\xi\in L^\rho(D)$, we have*

$$
\sup_{\epsilon\in[0,1]}\left\{E\left[\sup_{t\in[0,T]}|u^{\epsilon,\xi}(t)|_\rho^\rho\right]\right\}\le C\big(1+|\xi|_\rho^\rho\big).
$$

Proof: This is given by the proof of Gyöngy and Rovira (2000, Theorem 2.1) and a careful reading of the proof shows that the same constant $C>0$ applies for all $\epsilon\in[0,1]$. Note that, the $\exp^{-|u_0|_\rho}$ weighting used in Gyöngy and Rovira (2000) is not necessary here because we are working with deterministic initial conditions. □

A.2. *Proof of Theorem 5.4*. Theorem 5.4 is essentially a deterministic version of Theorem 3.1. Therefore, for the sake of brevity, and because the arguments used in the proof of this lemma are indeed similar to those used in the proof of Theorem 3.1, we will only provide a sketch.

For all $m\in\mathbb{N}$, define $\beta^m\in\Lambda^N$ by

$$
\beta_j^m(t):=\begin{cases} m & \text{if } \beta_j(t)>m\\ -m & \text{if } \beta_j(t)<-m\\ \beta_j(t) & \text{otherwise,}\end{cases}
$$

for all $j \in \{1, ..., k\}$ and $t \in [0, T]$. Due to Theorem 2.2, we know for all $\epsilon \in (0, 1]$, there is a unique $\hat{u}^{\epsilon,\xi,\beta^m}(t, x) \in C\big([0, T] : L^\rho(D)\big)$, such that for all $t \in [0, T]$

$$\begin{aligned}\hat{u}^{\epsilon,\xi,\beta^m}(t,x) &= \int_D G_t(x,y)\xi(y)dy + \sqrt{\epsilon}\int_0^t \int_D G_{t-s}(x,y)\sigma_j\big(\hat{u}^{\epsilon,\xi,\beta^m}\big)(s,y)dy\beta_j^m(s)ds \\ &- \int_0^t \int_D \partial_{y_i} G_{t-s}(x,y) g_i\big(\hat{u}^{\epsilon,\xi,\beta^m}\big)(s,y)dyds \\ &+ \int_0^t \int_D G_{t-s}(x,y) f\big(\hat{u}^{\epsilon,\xi,\beta^m}\big)(s,y)dyds,\end{aligned}$$

for a.e. $x \in D$ and due to Lemma A.6

$$\sup_{\epsilon\in[0,1]}\left\{\sup_{t\in[0,T]}\big|\tilde{u}^{\epsilon,\xi,\beta^m}(t)\big|_\rho^\rho\right\} \le C_m\big(1 + |\xi|_\rho^\rho\big).$$

Applying the same arguments used to treat the J_3 term in the proof of Lemma 7.2 and the J_1^ϵ and J_2^ϵ terms in the proof of Proposition 3.2, combined with Grönwall's inequality gives

$$\sup_{\epsilon\in(0,1]}\left\{\epsilon^{-\frac{1}{2}}\sup_{t\in[0,T]}\big|\hat{u}^{\epsilon,\xi,\beta^m}(t) - u^{0,\xi}(t)\big|_\rho\right\} \le C_m.$$

Define

$$\hat{z}^{\epsilon,\xi,\beta^m} := \frac{1}{\sqrt{\epsilon}}\big(\hat{u}^{\epsilon,\xi,\beta^m} - u^{0,\xi}\big).$$

A deterministic version of the proof of Theorem 3.1, combined with the argument used to treat the J_3^n term in the proof of Theorem 7.3, shows that $\{\hat{z}^{\epsilon,\xi,\beta^m}\}_{\epsilon\in(0,1]}$ converges in $C\big([0,T] : L^\rho(D)\big)$ to a limit z^{0,ξ,β^m}, which satisfies for all $t \in [0, T]$

$$\begin{aligned}z^{0,\xi,\beta^m}(t) &= \sqrt{\epsilon}\int_0^t \int_D G_{t-s}(x,y)\sigma_j\big(u^{0,\xi}\big)(s,y)dy\beta_j^m(s)ds \\ &- \int_0^t \int_D \partial_{y_i} G_{t-s}(x,y)\Big(z^{0,\xi,\beta^m} g_i'\big(u^{0,\xi}\big)\Big)(s,y)dyds \\ &+ \int_0^t \int_D G_{t-s}(x,y)\Big(z^{0,\xi,\beta^m} f'\big(u^{0,\xi}\big)\Big)(s,y)dyds,\end{aligned}$$

for a.e. $x \in D$. A similar approach, combined with the argument used to treat the J_3 term in the proof of Lemma 7.2, shows that $\{z^{0,\xi,\beta^m}\}_{m\in\mathbb{N}}$ converges in $C\big([0,T] : L^\rho(D)\big)$ to a limit $z^{0,\xi,\beta}$, which satisfies (5.5). Uniqueness follows as in the proof of Theorem 3.1.

□

References

Belfadli, R., Boulanba, L., and Mellouk, M. Moderate deviations for a stochastic Burgers equation. *Mod. Stoch. Theory Appl.*, **6** (2), 167–193 (2019). MR3963877.

Budhiraja, A., Dupuis, P., and Maroulas, V. Large deviations for infinite dimensional stochastic dynamical systems. *Ann. Probab.*, **36** (4), 1390–1420 (2008). MR2435853.

Da Prato, G. and Zabczyk, J. *Stochastic equations in infinite dimensions*, volume 152 of *Encyclopedia of Mathematics and its Applications*. Cambridge University Press, Cambridge, second edition (2014). ISBN 978-1-107-05584-1. MR3236753.

Dong, Z., Xiong, J., Zhai, J., and Zhang, T. A moderate deviation principle for 2-D stochastic Navier-Stokes equations driven by multiplicative Lévy noises. *J. Funct. Anal.*, **272** (1), 227–254 (2017). MR3567505.

Dupuis, P. and Ellis, R. S. *A weak convergence approach to the theory of large deviations*. Wiley Series in Probability and Statistics: Probability and Statistics. John Wiley & Sons, Inc., New York (1997). ISBN 0-471-07672-4. MR1431744.

Dupuis, P. and Johnson, D. Moderate deviations-based importance sampling for stochastic recursive equations. *Adv. in Appl. Probab.*, **49** (4), 981–1010 (2017). MR3732184.

Dupuis, P., Spiliopoulos, K., and Zhou, X. Escaping from an attractor: importance sampling and rest points I. *Ann. Appl. Probab.*, **25** (5), 2909–2958 (2015). MR3375891.

Dupuis, P. and Wang, H. Importance sampling, large deviations, and differential games. *Stoch. Stoch. Rep.*, **76** (6), 481–508 (2004). MR2100018.

Freidlin, M. I. and Wentzell, A. D. *Random perturbations of dynamical systems*, volume 260 of *Grundlehren der mathematischen Wissenschaften*. Springer, Heidelberg, third edition (2012). ISBN 978-3-642-25846-6. MR2953753.

Garsia, A. M., Rodemich, E., and Rumsey, H., Jr. A real variable lemma and the continuity of paths of some Gaussian processes. *Indiana Univ. Math. J.*, **20**, 565–578 (1970/71). MR267632.

Gasteratos, I., Salins, M., and Spiliopoulos, K. Moderate deviations for systems of slow-fast stochastic reaction-diffusion equations. *Stoch. Partial Differ. Equ. Anal. Comput.*, **11** (2), 503–598 (2023). MR4588617.

Gasteratos, I., Salins, M., and Spiliopoulos, K. Importance sampling for stochastic reaction–diffusion equations in the moderate deviation regime. *Stoch. Partial Differ. Equ. Anal. Comput.*, **12**, 2081–2150 (2024). DOI: 10.1007/s40072-023-00320-x.

Gyöngy, I. and Rovira, C. On L^p-solutions of semilinear stochastic partial differential equations. *Stochastic Process. Appl.*, **90** (1), 83–108 (2000). MR1787126.

Hu, S., Li, R., and Wang, X. Central limit theorem and moderate deviations for a class of semilinear stochastic partial differential equations. *Acta Math. Sci. Ser. B (Engl. Ed.)*, **40** (5), 1477–1494 (2020). MR4143604.

Ladyženskaja, O. A., Solonnikov, V. A., and Ural'ceva, N. N. *Linear and quasilinear equations of parabolic type*, volume 23 of *Translations of Mathematical Monographs*. American Mathematical Society, Providence, RI (1968). MR241822.

Salins, M. Equivalences and counterexamples between several definitions of the uniform large deviations principle. *Probab. Surv.*, **16**, 99–142 (2019). MR3960292.

Salins, M. and Setayeshgar, L. Uniform large deviations for a class of Burgers-type stochastic partial differential equations in any space dimension. *Potential Anal.*, **58** (1), 181–201 (2023). MR4535922.

Salins, M. and Spiliopoulos, K. Rare event simulation via importance sampling for linear SPDE's. *Stoch. Partial Differ. Equ. Anal. Comput.*, **5** (4), 652–690 (2017). MR3736657.

Setayeshgar, L. Large deviations for a class of parabolic semilinear stochastic partial differential equations in any space dimension. *Markov Process. Related Fields*, **25** (5), 899–914 (2019). MR4246020.

Vanden-Eijnden, E. and Weare, J. Rare event simulation of small noise diffusions. *Comm. Pure Appl. Math.*, **65** (12), 1770–1803 (2012). MR2982641.

Wang, R., Zhai, J., and Zhang, T. A moderate deviation principle for 2-D stochastic Navier-Stokes equations. *J. Differential Equations*, **258** (10), 3363–3390 (2015). MR3319423.

Wang, R. and Zhang, T. Moderate deviations for stochastic reaction-diffusion equations with multiplicative noise. *Potential Anal.*, **42** (1), 99–113 (2015). MR3297988.

Xiong, J. and Zhang, R. Semilinear stochastic partial differential equations: central limit theorem and moderate deviations. *Math. Methods Appl. Sci.*, **44** (8), 6808–6838 (2021). MR4258396.