Available online at www.sciencedirect.com

ScienceDirect

Ann. I. H. Poincaré – PR 43 (2007) 441–459

ANNALES DE L'INSTITUT HENRI POINCARÉ
PROBABILITÉS ET STATISTIQUES

www.elsevier.com/locate/anihpb

Large deviation principle of occupation measure for stochastic Burgers equation

Mathieu Gourcy

Laboratoire de Mathématiques, CNRS-UMR 6620, Université Blaise Pascal, 63177 Aubière, France

Received 3 March 2006; accepted 4 July 2006

Available online 26 December 2006

Abstract

In this paper we obtain a Large Deviation Principle for the occupation measure of the solution to a stochastic Burgers equation which describes the exact rate of exponential convergence. This Markov process is strongly Feller and has a unique invariant measure. Moreover, the rate function is explicit: it is the level-2 entropy of Donsker–Varadhan.

Résumé

On obtient un Principe de Grandes Déviations pour la mesure d'occupation associée à la solution d'une équation de Burgers stochastique. Ce résultat décrit convergence exponentielle vers l'unique mesure invariante. La fonction de taux associée est l'entropie de niveau 2 de Donsker–Varadhan.

MSC: 60F10; 60J35; 35Q53

Keywords: Stochastic Burgers equation; Large deviations; Occupation measure

1. Introduction and main results

Let $H = L^2(0,1)$ equipped with its norm $\|\cdot\|_2$. In this paper we are interested in the large time behavior of the solution to the following stochastic Burgers equation:

$$\mathrm{d}X(t) = \left(\Delta X(t) + \frac{1}{2} D_\xi X^2(t)\right)\mathrm{d}t + G\,\mathrm{d}W(t); \quad X(0,\xi) = x_0(\xi) \in H, \tag{1.1}$$

where $G : H \to H$ is a bounded linear operator, $W(t)$ is a standard cylindrical Wiener process on H, and Δ is the Laplacian on $(0,1)$ with the Dirichlet boundary conditions. Indeed, the problem (1.1) is supplemented by:

$$X(t,0) = X(t,1) = 0, \quad t > 0.$$

It is well known that Δ is a negative, self-adjoint, non-bounded operator on H with the domain of definition given by

$$D(\Delta) = \left\{u \in H^2(0,1):\ u(0) = u(1) = 0\right\} = H^2(0,1) \cap H_0^1(0,1),$$

E-mail address: mathieu.gourcy@math.univ-bpclermont.fr.

0246-0203/$ – see front matter
doi:10.1016/j.anihpb.2006.07.003

where

$$H_0^1 = \{x\colon [0,1] \to \mathbb{R};\ x \text{ is abs. continuous}, x(0) = x(1) = 0, \text{ and } \nabla x := D_\xi x \in H\}.$$

We assume that $\operatorname{tr}(GG^*) < \infty$, i.e. the energy injected by the random force is finite, and that, for $Q = GG^*$,

$$\operatorname{Im}\bigl((-\Delta)^{-\delta/2}\bigr) \subset \operatorname{Im}\bigl(Q^{1/2}\bigr) \quad \text{for some } \frac{1}{2} < \delta < 1, \tag{1.2}$$

where $\operatorname{Im}(Q^{1/2})$ is the range of the operator $Q^{1/2}$. The last condition (1.2) means that the noise is not too degenerate. It is equivalent to say that the domain of definition of $(-\Delta)^{\delta/2}$ in H is contained in $\operatorname{Im}(Q^{1/2})$.

The above equation plays an important role in fluid dynamic for understanding of chaotic behavior. This stochastic model has been intensively studied for 10 years, in particular by Da Prato, Debussche, Dermoune, Weinan, Gatarek, Khanin, Mazel, Sinai and Temam among many others (from a chronological point of view, see [5,4,9,2,3,18]). About large deviations, small noise asymptotic was investigated by Cardon-Weber [1]. More recently, Goldys and Maslowski proved the exponential ergodicity [13].

Let $M_1(H)$ (resp. $M_b(H)$) be the space of probability measures (resp. signed σ-additive measures of bounded variation) on H equipped with the Borel σ-field $\mathcal{B}(H)$. The usual duality relation between $\nu \in M_b(H)$ and $f \in b\mathcal{B}(H)$, the set of bounded and measurable functions on H, will be denoted by

$$\nu(f) := \int_H f \,\mathrm{d}\nu.$$

On $M_b(H)$ (or its subspace $M_1(H)$), we will consider the usual weak convergence topology $\sigma(M_b(H), C_b(H))$ and the so called τ-topology $\sigma(M_b(H), b\mathcal{B}(H))$, which is much stronger.

Our aim is to establish the large deviation principle (LDP in short) for the occupation measure L_t of the solution X (or empirical measure of level-2) given by

$$L_t(A) := \frac{1}{t}\int_0^t \delta_{X_s}(A)\,\mathrm{d}s, \quad \forall A \in \mathcal{B}(H)$$

δ_a being the Dirac measure at a. Notice that L_t is an in $M_1(H)$-valued random variable. This is a traditional subject in probability since the pioneering work of Donsker and Varadhan [11]. The main innovation is that we deal about infinite dimensional diffusions for which their assumptions are not satisfied. For an introduction to large deviations we refer to the books of Deuschel and Stroock [10], and Dembo and Zeitouni [8].

Under (1.2), it is known that X_t is a Markov process with a unique invariant measure μ (cf. [7]). So the ergodic theorem says that, almost surely under $\mathbb{P}_\mu$, L_t converges weakly to μ. We establish in this note a much more stronger result:

Theorem 1.1. *Assume that* $\operatorname{tr}(GG^*) < +\infty$ *and* (1.2) (*throughout this paper*). *Let* $0 < \lambda_0 < \frac{\pi^2}{2\|Q\|}$, *where* $\|Q\|$ *is the norm of Q as an operator in H and*

$$\Phi(x) = \mathrm{e}^{\lambda_0\|x\|_2^2}, \qquad \mathcal{M}_{\lambda_0,L} := \left\{\nu \in M_1(H) \;\middle|\; \int_H \Phi(x)\nu(\mathrm{d}x) \leqslant L\right\}. \tag{1.3}$$

The family $\mathbb{P}_\nu(L_T \in \cdot)$ as $T \to +\infty$ satisfies the large deviation principle (*LDP*) *with respect to* (*w.r.t. in short*) *the topology τ, with speed T and the rate function J, uniformly for any initial measure in $\mathcal{M}_{\lambda_0,L}$ where $L > 1$ is any fixed number. Here $J : M_1(H) \to [0,+\infty]$ is the level-2 entropy of Donsker–Varadhan defined by* (3.2) *below.*

More precisely we have:

(i) *J is a good rate function on $M_1(H)$ equipped with the topology τ of the convergence against bounded and Borelian functions, i.e., $[J \leqslant a]$ is τ-compact for every $a \in \mathbb{R}^+$;*

(ii) *for all open set G in $M_1(H)$ with respect to the topology τ,*

$$\liminf_{T\to\infty} \frac{1}{T} \log \inf_{\nu\in\mathcal{M}_{\lambda_0,L}} \mathbb{P}_\nu(L_T \in G) \geqslant -\inf_G J; \tag{1.4}$$

(iii) *for all closed set F in $M_1(H)$ with respect to the topology τ,*

$$\limsup_{T\to\infty}\frac{1}{T}\log\sup_{\nu\in\mathcal{M}_{\lambda_0,L}}\mathbb{P}_\nu(L_T\in F)\leqslant-\inf_F J. \tag{1.5}$$

Furthermore we have

$$J(\nu)<+\infty\quad\Longrightarrow\quad\nu\ll\mu,\ \nu(H_0^1)=1\ \text{and}\int_{H_0^1}\|\nabla x\|_2^2\,\mathrm{d}\nu<+\infty, \tag{1.6}$$

where μ is the unique invariant probability measure of (X_t).

The LDP w.r.t. the topology τ is much stronger than that w.r.t. the usual weak convergence topology as in Donsker and Varadhan [11]. Sometimes considered as a technical detail, the topology τ is crucial here: interesting consequences of this LDP can be deduced for many physical quantities of the system such as $\|x\|_{H^1}=\|\nabla x\|_2$, or more generally $\|x\|_{H^\alpha}:=\|(-\Delta)^{\alpha/2}x\|_2$ for $0\leqslant\alpha\leqslant 1$, which are not continuous on H. In fact, we establish

Corollary 1.2. *Let $\mathbb{B}$ a separable Banach space, and $f:H_0^1\to\mathbb{B}$ a measurable function, bounded on the balls $\{x;\ \|\nabla x\|_2\leqslant R\}$, and satisfying*

$$\lim_{\|\nabla x\|_2\to\infty}\frac{\|f(x)\|_{\mathbb{B}}}{\|\nabla x\|_2^2}=0. \tag{1.7}$$

Then, $\mathbb{P}_\nu(L_T(f)\in\cdot)$ satisfies the LDP on $\mathbb{B}$, with speed T and the rate function I_f given by

$$I_f(z)=\inf\left\{J(\nu);\ J(\nu)<+\infty,\int_{H_0^1}f(x)\,\mathrm{d}\nu(x)=z\right\},\quad\forall z\in\mathbb{B},$$

uniformly over initial distributions ν in $\mathcal{M}_{\lambda_0,L}$ for any fixed $L>1$.

For instance, $f:H_0^1\to\mathbb{B}:=H^\alpha$ with $f(x)=x$ for any $\alpha\in[0,1)$ is allowed, so that the LDP in H^α holds for $\mathbb{P}_\nu(1/T\int_0^T X_t\,\mathrm{d}t\in\cdot)$. An other particular case of the above corollary is the following: for every $p\in(0,2)$,

$$\mathbb{P}_\nu\left(\frac{1}{T}\int_0^T\left\|\nabla X(t)\right\|_2^p\,\mathrm{d}t\in\cdot\right)$$

satisfies the LDP on $\mathbb{R}$ with speed T and the rate function I defined by

$$I(z)=\inf\left\{J(\nu);\ J(\nu)<+\infty,\int_H\|\nabla x\|_2^p\,\mathrm{d}\nu(x)=z\right\},\quad\forall z\in\mathbb{R} \tag{1.8}$$

uniformly over initial distributions ν in $\mathcal{M}_{\lambda_0,L}$ (for any $L>1$).

Finally, we introduce $(e_k)_k$ the complete orthonormal system in $L^2(0,1)$ which diagonalizes Δ on its domain, and by $-\lambda_k$ the corresponding eigenvalues. We have

$$e_k(x)=\sqrt{\frac{2}{\pi}}\sin k\pi x,\quad\lambda_k=\pi^2k^2,\quad k\in\mathbb{N}^*=\{1,2,\ldots\}.$$

Remarks 1.3.

(i) Let us see the meaning of our assumptions: $\mathrm{tr}(Q)<+\infty$ and (1.2). Assume that $Ge_k=\sigma_k e_k$ for every $k\geqslant 1$. Then

$$GW(t)=\sum_{k=1}^{\infty}\sigma_k\beta_k(t)e_k, \tag{1.9}$$

where $(\beta_k)_{k\in\mathbb{N}^*}$ is a family of independent real valued standard Brownian motions. Then $\operatorname{tr}(Q) < +\infty$ and condition (1.2) is satisfied if

$$\frac{c}{k} \leqslant |\sigma_k| \leqslant \frac{C}{k^{1/2+\varepsilon}}$$

for two positive constants c and C and some small $\varepsilon > 0$.

A more general example of noise for which our assumptions hold is

$$G := (-\Delta)^{-\beta} B, \quad \frac{1}{4} < \beta < \frac{1}{2},$$

where B is any linear bounded and invertible operator on H. Indeed $\operatorname{tr}(GG^*) \leqslant \|B\|^2_{H\to H} \operatorname{tr}(\Delta^{-2\beta}) < +\infty$ for $2\beta > 1/2$. Since $\operatorname{Im}(G) = \operatorname{Im}(\Delta^{-\beta})$ and by the polar decomposition, $\operatorname{Im}(G) \subset \operatorname{Im}(\sqrt{GG^*})$, the condition (1.2) is then verified with $\delta = 2\beta$.

(ii) Our approach here is well adapted to the case of a multiplicative (or correlated) random forcing term, that is, the noise $GW(t)$ can be replaced by

$$g\big(X(t,\xi)\big)GW(t),$$

where $g : H \to [\alpha, \beta]$ is Lipschitz continuous, $0 < \alpha < \beta < \infty$, G satisfies (1.2) and $\operatorname{tr}(GG^*) < +\infty$. Indeed, following [4], the strong Feller property and the topological irreducibility hold. All estimates necessary for the LDP in Theorem 1.1 still hold in the actual case, and then all previous results remain valid.

(iii) The class (1.3) of allowed initial distributions for the uniform LDP is sufficiently rich. For example, choosing L large enough, it includes all the Dirac probability measures δ_x with x in any ball of H.

(iv) Our LDP is more precise than the exponential convergence of P_t to the invariant measure μ established in [13]. Indeed the LDP furnishes the exact rate of the exponential convergence in probability of the empirical measures L_T to μ. Moreover by Theorem 6.4 in [21], under the strong Feller and topological irreducibility assumption for (P_t), the LDP in Theorem 1.1 is equivalent to saying that the essential spectral radius in some weighted functions spaces $b_u\mathcal{B}$ is zero.

(v) The assumption (1.2) plays a crucial role for Theorem 1.1: if the noise acts only on a finite number of modes (i.e., $\sigma_k = 0$ for all $k > N$ in (1.9)) as in Kolmogorov's turbulence theory, we believe that the LDP w.r.t. the τ-topology is false. It is a challenging open question for establishing the LDP of L_T w.r.t. the weak convergence topology in the last degenerate noise case.

(vi) For the 2D-stochastic Navier–Stokes equation, we can prove, under suitable conditions, a LDP on some $D(A^\alpha)$, for $\frac{1}{4} < \alpha < \frac{1}{2}$. Here A is also the Laplacian, but regarded as an operator on the subspace of the L^2-vector fields with free divergence. That will be carried out in a future work.

This paper is organized as follows. In Section 2, we recall known results on existence and uniqueness of solution, and existence of an invariant probability measure for Eq. (1.1). In Section 3 we give some general facts about large deviations for strong Feller and irreducible Markov processes and we obtain the uniform lower bound (1.4). Then we prove the convergence of the Galerkin approximations for the considered equation in Section 4. The exponential tightness is investigated in Section 5, and the uniform upper bound (1.5) for the strong τ-topology in Section 6. Finally, the extension to non-bounded functionals on H is discussed in the last Section 7.

2. Solutions of the equation and their properties

Let us specify what we understand by solution. Generally, we are concerned with two ways of giving a rigorous meaning to solutions of stochastic differential equations in infinite dimensional spaces, that is, the variational one [17,15] and the semigroup one [6]. Correspondingly, as in the case of deterministic evolution equations, we have two notions of strong, and "mild" solution. In most situations, one finds that the concept of strong solution is too limited to include important examples. The weaker concept of mild solution seems to be more appropriate. In the sequel, we are working with this concept, that we define more precisely now.

We denote by $S(t)$ the semigroup generated by Δ on $L^2(0,1)$, or from a formal point of view, $S(t) = \mathrm{e}^{t\Delta}$.

Definition 2.1. We say that $X \in C([0,T], L^2(0,1))$ is a "mild" solution of problem (1.1) if $X(t)$ is adapted to $\mathcal{F}_t$, the σ-algebra of the cylindrical Wiener process until time t and for arbitrary $0 \leqslant t$, we have

$$X(t) = S(t)x_0 + \int_0^t S(t-s)\frac{1}{2}D_\xi X^2(s)\,\mathrm{d}s + \int_0^t S(t-s)G\,\mathrm{d}W(s) \tag{2.1}$$

for any $x_0 \in L^2(0,1)$, $\mathbb{P}$ almost surely.

Note that all the terms in (2.1) take sense since the mapping

$$F : u \in C\big([0,T], L^1(0,1)\big) \to \int_0^t S(t-s)\frac{1}{2}D_\xi u(s)\,\mathrm{d}s \in C\big([0,T], L^2(0,1)\big)$$

is well defined (see [7] p. 260) and the stochastic convolution $W_\Delta := \int_0^t S(t-s)G\,\mathrm{d}W(s)$ also (see (4.4) below). Da Prato, Debussche, Temam established in [5] for the first time existence and uniqueness for a stochastic Burgers equation cylindrically perturbed, that is when G is the identity operator. The method they used to obtain local existence in time of a solution consists in considering a fixed path of the noise, to get into a deterministic setting and use a fixed point argument. Then the time of explosion is shown to be infinite, by means of a priori bounds on the solution. The same proof gives in our setting:

Theorem 2.2. *Stochastic Burgers equation* (1.1) *admits a unique mild solution and for all* $T > 0$,

$$X \in C\big([0,T], L^2(0,1)\big) \cap L^2\big([0,T], C[0,1]\big).$$

The solution satisfies Markov and strong Markov properties (see [6]). We can also consider the transition semigroup associated to the dynamics given by

$$P_t\Phi(x) := \mathbb{E}\Phi\big(X(t,x)\big) = \mathbb{E}^x\Phi\big(X(t)\big), \quad \forall \Phi \in b\mathcal{B}(H).$$

As in [5], this semigroup admits an invariant measure. Moreover, under our condition (1.2) on the noise, the following interesting properties hold.

Lemma 2.3.

(i) *The transition semigroup* (P_t) *corresponding to the forced Burgers equation* (1.1) *satisfies the strong Feller property. That is, for any bounded Borelian function* Φ *on* H *and any* $t > 0$, *the function* $P_t\Phi(\cdot)$ *is continuous on* H.

(ii) *For every* $t > 0$, $P_t(x, O) > 0$ *for all* $x \in H$ *and all non-empty open subset* O *of* H. *Hence,* (P_t) *is also topologically irreducible.*

(iii) *In particular, the transition semigroup* (P_t), *corresponding to the forced Burgers equation* (1.1) *admits a unique invariant measure* μ, *which charges all non-empty open subsets of* H.

Part (i) is well known when the cylindrical noise is considered (see [7]). In our case of a finite trace class noise, the non-degeneracy condition (1.2) is essential. More precisely, $\delta < 1$ allows to obtain a bound on the derivative of the semigroup by using the Bismut–Elworthy formula as in [2] or [12]. The condition $\delta > \frac{1}{2}$ is borrowed from the finite trace assumption, crucial in the application of Itô's formula for the exponential tightness.

The point (ii) was proved by Goldys and Maslowski in [13] for our class of noise. We recall that (P_t) is topologically irreducible if, for all non-empty open set Γ in H, and all $x \in H$, we have $P_t(x, \Gamma) > 0$ for some $t > 0$.

According to the general theory [7], we obtain (iii) as first corollary, sometimes called Doob's theorem, of the two preceding points together with the existence of invariant measure. In fact this result gives also the convergence of the transition probabilities to the invariant measure.

Our aim is to complete the study of Eq. (1.1) by giving information on the rare events and the exact rate of exponential convergence by means of a large deviation principle, one of the strongest ergodic behaviors of Markov processes.

3. General results about large deviations

In this section, we introduce some necessary notations and definitions and give general results (essentially following [19]) on large deviations for Markov processes.

3.1. Notations and entropy of Donsker–Varadhan

We first compare the "topological irreducibility" defined above (often called irreducibility in the literature on SPDE) with the probabilistic irreducibility for a Markov process which is the more general assumption under which the large deviations result we use (as Lemma 3.2 below) holds true (see [14,19] for details).

Let ν be a probability measure on H; a transition kernel operator P on H is said ν-irreducible (resp. ν-essentially irreducible) if for all A in H such that $\nu(A) > 0$, and for all x in H (resp. for ν almost all x in H), we can find $n \in \mathbb{N}$ such that $P^n(x, A) > 0$. When ν charges all non-empty open subsets of H, the ν-irreducibility implies the topological irreducibility. But for the strong Feller P, the topological irreducibility implies the ν-irreducibility for all ν such that $\nu \ll \nu P$ (see [19]).

Thus by Lemma 2.3, for the unique invariant measure μ of our model, P_t is μ-irreducible for every $t > 0$. In reality for our model, we have the much stronger property that all the probability measures in the family

$$\{P_t(x, \cdot),\ x \in H,\ t > 0\}$$

are equivalent, and they are also equivalent to μ (see [7, p. 41]).

Consider the $H := L^2(0, 1)$-valued continuous Markov process

$$\big(\Omega, (\mathcal{F}_t)_{t\geqslant 0}, \mathcal{F}, \big(X_t(\omega)\big)_{t\geqslant 0}, (\mathbb{P}_x)_{x\in H}\big)$$

whose semigroup of Markov transitions kernels is denoted by $(P_t(x, \mathrm{d}y))_{t\geqslant 0}$, where

$\Omega = C(\mathbb{R}^+, H)$ is the space of continuous functions from $\mathbb{R}^+$ to H equipped with the compact convergence topology;
$\mathcal{F}_t = \sigma(X_s, 0 \leqslant s \leqslant t)$ for any $t \geqslant 0$ is the natural filtration;
$\mathcal{F} = \sigma(X_s, 0 \leqslant s)$ and $\mathbb{P}_x(X_0 = x) = 1$.

Hence, $\mathbb{P}_x$ is the law of the Markov process with initial state x in H. For any initial measure ν on H, let $\mathbb{P}_\nu(\mathrm{d}\omega) := \int_H \mathbb{P}_x(\mathrm{d}\omega)\nu(\mathrm{d}x)$.

The empirical measure of level-3 (or process level) is

$$R_t := \frac{1}{t}\int_0^t \delta_{\theta_s X}\,\mathrm{d}s,$$

where $(\theta_s X)_t = X_{s+t}$ for all $t, s \geqslant 0$ are the shifts on Ω. Hence R_t is a random element of $M_1(\Omega)$, the space of probability measures on Ω.

The level-3 entropy functional of Donsker and Varadhan $H : M_1(\Omega) \to [0, +\infty]$ is defined by,

$$H(Q) := \begin{cases} \mathbb{E}^{\bar{Q}} h_{\mathcal{F}_1^0}(\bar{Q}_{\omega(-\infty,0]}; \mathbb{P}_{w(0)}), & \text{if } Q \in M_1^s(\Omega), \\ +\infty, & \text{otherwise,} \end{cases} \tag{3.1}$$

where

$M_1^s(\Omega)$ is the space of those elements in $M_1(\Omega)$ which are moreover stationary;
$\bar{Q}$ is the unique stationary extension of $Q \in M_1^s(\Omega)$ to $\bar{\Omega} := C(\mathbb{R}, H)$; $\mathcal{F}_t^s = \sigma(X(u);\ s \leqslant u \leqslant t)$ on $\bar{\Omega}$, $\forall s, t \in \mathbb{R}$, $s \leqslant t$;
$\bar{Q}_{\omega(-\infty,t]}$ is the regular conditional distribution of $\bar{Q}$ knowing $\mathcal{F}_t^{-\infty}$;
$h_{\mathcal{G}}(\nu, \mu)$ is the usual relative entropy or Kullback information of ν with respect to μ restricted on the σ-field $\mathcal{G}$, given by

$$h_{\mathcal{G}}(\nu, \mu) := \begin{cases} \int \frac{\mathrm{d}\nu}{\mathrm{d}\mu}|_{\mathcal{G}} \log(\frac{\mathrm{d}\nu}{\mathrm{d}\mu}|_{\mathcal{G}})\,\mathrm{d}\mu, & \text{if } \nu \ll \mu \text{ on } \mathcal{G}, \\ +\infty, & \text{otherwise.} \end{cases}$$

The level-2 entropy functional $J : M_1(H) \to [0, \infty]$ is defined by

$$J(\beta) = \inf\{H(Q);\ Q \in M_1^s(\Omega) \text{ and } Q_0 = \beta\}, \quad \forall \beta \in M_1(H), \tag{3.2}$$

where $Q_0(\cdot) = Q(X(0) \in \cdot)$ is the marginal law at time $t = 0$.

Lastly introduced in [19], we define the restriction of the Donsker Varadhan entropy to the μ component, by

$$H_\mu(Q) := \begin{cases} H(Q), & \text{if } Q_0 \ll \mu, \\ +\infty, & \text{otherwise} \end{cases}$$

and for the level-2 entropy functional

$$J_\mu(\beta) := \begin{cases} J(\beta), & \text{if } \beta \ll \mu, \\ +\infty, & \text{otherwise.} \end{cases}$$

For our model, let us first establish the

Lemma 3.1. *We have $J(\nu) < +\infty \Rightarrow \nu \ll \mu$. Moreover, $J = J_\mu$ on $M_1(H)$ and $[J = 0] = \{\mu\}$.*

Proof. Consider ν such that $J(\nu) < \infty$. We recall the expression (3.1) of the Level-3 entropy. For $Q \in M_1^s(\Omega)$ such that $Q_0 = \nu$, $H(Q) < \infty$, and for every $t > 0$, noting that the entropy of marginal measure is not larger than the global entropy, we have by Jensen inequality,

$$\begin{aligned} H(Q) &= \mathbb{E}^{\bar{Q}} h_{\mathcal{F}_1}\big(\bar{Q}_{\omega(-\infty,0]}; \mathbb{P}_{w(0)}\big) = \frac{1}{t}\mathbb{E}^{\bar{Q}} h_{\mathcal{F}_t^0}\big(\bar{Q}_{\omega(-\infty,0]}; \mathbb{P}_{w(0)}\big) \\ &\geqslant \frac{1}{t} h_{\mathcal{F}_t^0}(Q; \mathbb{P}_\nu) \geqslant \frac{1}{t} h_{\sigma(w(t))}(Q; \mathbb{P}_\nu) \\ &\geqslant \frac{1}{t} h_{\mathcal{B}(H)}(\nu; \nu P_t). \end{aligned}$$

Taking infinimum over such Q, we get

$$J(\nu) \geqslant \frac{1}{t} h_{\mathcal{B}(H)}(\nu; \nu P_t). \tag{3.3}$$

So the Kullback information of ν with respect to νP_t is finite, which implies by definition that $\nu \ll \nu P_t$. Since all $P_t(x, \mathrm{d}y)$, $t > 0$, $x \in H$ are equivalent to μ ([7]), we have

$$\nu P_t(\cdot) = \int_H P(t, x, \cdot)\nu(\mathrm{d}x) \ll \mu.$$

Thus $\nu \ll \nu P_t \ll \mu$, as desired.

By definition, we have $J \leqslant J_\mu$ and they are equal on

$$\{\nu \in M_1(H) \text{ such that } \nu \ll \mu\}.$$

Since any probability measure ν on H such that $J(\nu) < \infty$ is absolutely continuous with respect to μ, we have $J = J_\mu$ on $M_1(H)$.

At the end, if the probability measure β is such that $J(\beta) = 0$ then $\beta \ll \mu$ and $\beta = \beta P_t$ for every $t > 0$ by (3.3). By the uniqueness in Lemma 2.3, we have $\beta = \mu$ and the proof is finished. □

3.2. The lower bound

Let us first recall the definition of the projective limit τ_p of the strong τ-topology,

$$\tau_p := \sigma\left(M_1(\Omega), \bigcup_{t \geqslant 0} b\mathcal{F}_t^0\right),$$

where $b\mathcal{F}_t^0$ is the set of functions on Ω, that are bounded and measurable for $\mathcal{F}_t^0$.

The following level-3 lower bound of Large Deviations for τ_p was established by Wu (see [19, Theorem B.1]) under more general conditions.

Lemma 3.2. ([19]) *For any open set O in $(M_1(\Omega), \tau_p)$,*

$$\liminf_{t\to\infty} \frac{1}{t} \log \mathbb{P}_x(R_t \in O) \geqslant -\inf_O H_\mu, \quad \mu\text{-a.e. initial state } x \in H.$$

Recall that $H_\mu = H$ by Lemma 3.1 and that a good rate function admits compact level sets (by definition). Our goal here, is to prove the

Proposition 3.3. *If J is a good rate function on $(M_1(H), \tau)$ and the uniform upper bound* (1.5) *is satisfied, then the level-*3 *uniform lower bound holds true: for any measurable open subset O in $(M_1(\Omega), \tau_p)$,*

$$\liminf_{t\to\infty} \frac{1}{t} \log \inf_{\nu\in\mathcal{M}_{\lambda_0,L}} \mathbb{P}_\nu(R_t \in O) \geqslant -\inf_O H.$$

*In particular, the desired Level-*2 *lower bound* (1.4) *holds* (*by the contraction principle*).

Proof. For any $Q \in O$ fixed, we can take a τ_p neighborhood of Q in $M_1(\Omega)$ of form

$$N(Q,\delta) := \left\{ Q' \in M_1(\Omega) \text{ such that } \left| \int F_i \,\mathrm{d}Q' - \int F_i \,\mathrm{d}Q \right| < \delta,\ \forall i = 1, \ldots, d \right\}$$

contained in O, where $\delta > 0$, $1 \leqslant d \in \mathbb{N}$ and $F_i \in b\mathcal{F}_n^0$ for some $n \in \mathbb{N}$. It is sufficient to establish that for every Q in O such that $H(Q) < \infty$

$$\liminf_{t\to\infty} \frac{1}{t} \log \inf_{\nu\in\mathcal{M}_{\lambda_0,L}} \mathbb{P}_\nu\big(R_t \in N(Q,\delta)\big) \geqslant -H_\mu(Q). \tag{3.4}$$

But by Egorov's lemma, Lemma 3.2 implies the existence of a Borelian subset K in H with $\mu(K) > 0$ such that for any $\varepsilon > 0$

$$\inf_{x\in K} \frac{1}{t} \log \mathbb{P}_x\left(R_t \in N\left(Q, \frac{\delta}{2} \right) \right) \geqslant -H_\mu(Q) - \varepsilon \tag{3.5}$$

for all t large enough. Let us fix $a > 0$. For any $0 \leqslant b \leqslant a$, we have

$$\left| \int F_i \,\mathrm{d}(R_t \circ \theta_b - R_t) \right| \leqslant \frac{2(a+1)}{t} \| F_i \|_\infty$$

and then for all $0 \leqslant b \leqslant a$ and for all t large enough (depending on a and δ),

$$\begin{aligned} \mathbb{P}_\nu\big(R_t \in N(Q,\delta)\big) &\geqslant \mathbb{P}_\nu\left(X_b \in K;\ R_t \circ \theta_b \in N\left(Q, \frac{\delta}{2} \right) \right) \\ &\geqslant \mathbb{P}_\nu(X_b \in K) \inf_{x\in K} \mathbb{P}_x\left(R_t \in N\left(Q, \frac{\delta}{2} \right) \right). \end{aligned}$$

Integrating for $0 \leqslant b \leqslant a$, and dividing by a yields

$$\mathbb{P}_\nu\big(R_t \in N(Q,\delta)\big) \geqslant \mathbb{E}^\nu L_a(K) \inf_{x\in K} \mathbb{P}_x\left(R_t \in N\left(Q, \frac{\delta}{2} \right) \right). \tag{3.6}$$

Hence, for proving (3.4), by (3.5) and (3.6), it is enough to establish that for any Borelian subset K with $\mu(K) > 0$, we can find $a > 0$ such that

$$\inf_{\nu\in\mathcal{M}_{\lambda_0,L}} \mathbb{E}^\nu L_a(K) > 0.$$

Notice that

$$\mathbb{E}^\nu L_a(K) \geqslant \frac{\mu(K)}{2} \left(1 - \mathbb{P}_\nu\left(\big| L_a(K) - \mu(K) \big| \geqslant \frac{\mu(K)}{2} \right) \right)$$

and by the assumed level 2 upper bound,

$$\limsup_{a\to+\infty} \frac{1}{a} \log \sup_{\nu\in\mathcal{M}_{\lambda_0,L}} \mathbb{P}_\nu\left(\big| L_a(K) - \mu(K) \big| \geqslant \frac{\mu(K)}{2} \right) \leqslant -\inf_F J(\nu),$$

where

$$F=\left\{\beta\in M_1(H):\ \left|\beta(K)-\mu(K)\right|\geqslant\frac{\mu(K)}{2}\right\}$$

is closed for the τ-topology. So once $\inf_F J>0$, we shall obtain for all a large enough

$$\inf_{\nu\in\mathcal{M}_{\lambda_0,L}}\mathbb{E}^\nu L_a(K)\geqslant\frac{\mu(K)}{2}\left(1-\exp\left(-a\frac{\inf_F J}{2}\right)\right)>0.$$

It remains to prove that $\inf_F J>0$. To this end we may assume that $\inf_F J<+\infty$. In that case, since J is a good rate function (our condition), $\inf_F J$ is attained by some $\beta_0\in F$. But $J(\beta)=0\Leftrightarrow\beta=\mu$ (Lemma 3.1) and $\mu\notin F$, so $\inf_F J=J(\beta_0)>0$ as desired. $\square$

3.3. Cramer functionals and weak upper bound

Let us introduce the uniform upper Cramer functional over a non-empty family of initial measures $\mathcal{A}$ in $M_1(H)$,

$$\Lambda(V|\mathcal{A}):=\limsup_{t\to\infty}\frac{1}{t}\log\sup_{\nu\in\mathcal{A}}\mathbb{E}^\nu\exp\bigl(tL_t(V)\bigr)$$

and several other Cramer functionals,

$$\begin{aligned}
\Lambda(V|x)&:=\Lambda\bigl(V|\{\delta_x\}\bigr)=\limsup_{t\to\infty}\frac{1}{t}\log\mathbb{E}^x\exp\bigl(tL_t(V)\bigr),\\
\Lambda^\infty(V)&:=\Lambda\bigl(V|\{\delta_x;\,x\in H\}\bigr)=\limsup_{t\to\infty}\frac{1}{t}\log\sup_{x\in H}\mathbb{E}^x\exp\bigl(tL_t(V)\bigr),\\
\Lambda^0(V)&:=\sup_{x\in H}\Lambda(V|x)=\sup_{x\in H}\limsup_{t\to\infty}\frac{1}{t}\log\mathbb{E}^x\exp\bigl(tL_t(V)\bigr),
\end{aligned}\tag{3.7}$$

where V is a bounded and Borelian function on H.

The functionals $\Lambda^0(V)$ and $\Lambda^\infty(V)$ are respectively the pointwise and uniform Cramer functionals introduced already in [10]. For $\Lambda:b\mathcal{B}(H)\to\mathbb{R}$ any one of the above functionals, define its Legendre transformation:

$$\begin{aligned}
\Lambda_w^*(\nu)&=\sup_{V\in C_b(H)}\left(\int_H V\,\mathrm{d}\nu-\Lambda(V)\right),\quad\forall\nu\in M_b(H),\\
\Lambda^*(\nu)&=\sup_{V\in b\mathcal{B}(H)}\left(\int_H V\,\mathrm{d}\nu-\Lambda(V)\right),\quad\forall\nu\in M_b(H),
\end{aligned}\tag{3.8}$$

where $M_b(H)$ is the space of all signed σ-additive measures of bounded variation on $(H,\mathcal{B})$.

Remark that $\{\delta_x\}_{x\in H}\subset\bigcup_{L>0}\mathcal{M}_{\lambda_0,L}$, we have for any bounded and measurable function V,

$$\Lambda^0(V)\leqslant\sup_{L>0}\Lambda(V|\mathcal{M}_{\lambda_0,L})\leqslant\Lambda^\infty(V).$$

Since (P_t) is Feller, we have by [19, Proposition B.13]

$$\bigl(\Lambda^0\bigr)^*(\nu)=\bigl(\Lambda^0\bigr)^*_w(\nu)=\bigl(\Lambda^\infty\bigr)^*(\nu)=\bigl(\Lambda^\infty\bigr)^*_w(\nu)=J(\nu),\quad\forall\nu\in M_1(H)$$

which implies the l.s.c. for J and the fact that

$$\sup_{V\in b\mathcal{B}(H)}\left(\int_H V\,\mathrm{d}\nu-\sup_{L>0}\Lambda(V|\mathcal{M}_{\lambda_0,L})\right)=\sup_{V\in C_b(H)}\left(\int_H V\,\mathrm{d}\nu-\sup_{L>0}\Lambda(V|\mathcal{M}_{\lambda_0,L})\right)=J(\nu),\quad\forall\nu\in M_1(H).$$

So by Gärtner and Ellis theorem (see [8]), we have always the following general weak* upper bound

Lemma 3.4. *Let $M_1(H)$ be equipped with the weak convergence topology. For any compact subset K in $M_1(H)$ w.r.t. the weak convergence topology, and for any $\varepsilon>0$, there is a neighborhood $N(K,\varepsilon)$ of K in $M_1(H)$ such that*

$$\limsup_{t\to\infty}\frac{1}{t}\log\sup_{\nu\in\mathcal{M}_{\lambda_0,L}}\mathbb{P}_\nu\big(L_t\in N(K,\varepsilon)\big)\leqslant\begin{cases}-\inf_{\nu\in K}J(\nu)+\varepsilon, & \textit{if } \inf_{\nu\in K}J(\nu)<\infty,\\ -\frac{1}{\varepsilon}, & \textit{otherwise}.\end{cases}$$

Now to obtain the upper bound in Theorem 1.1 w.r.t. the weak convergence topology, we need to prove the exponential tightness of L_t.

4. Convergence of a Galerkin method

Let us introduce the approximation system associated with Eq. (1.1):

$$\mathrm{d}X_n(t)=\left(\Delta X_n(t)+\frac{1}{2}\Pi_n D_\xi(X_n)^2(t)\right)\mathrm{d}t+G_n\,\mathrm{d}W(t);\quad X_n(0)=\Pi_n x,\tag{4.1}$$

where Π_n is the orthogonal projection on H_n, the finite dimensional space spanned by the first n eigenvectors $(e_1,\dots,e_n)$, and $G_n:=\Pi_n G$.

The convergence of a similar approximation but with a non-linearity truncated by the function

$$f_n(x)=\frac{nx^2}{n+x^2}$$

was investigated by Da Prato and Debussche [3]. The aim of this section is to establish some a priori estimates on X_n, and the convergence of the approximation method (4.1).

Theorem 4.1. *The solutions X_n of* (4.1) *converge to the solution X of* (1.1) *in $C([0,T];H)$ and in $L^2([0,T];H_0^1)$ almost surely.*

4.1. A priori estimate for the finite dimensional approximations

From now on, we denote by $\langle\cdot,\cdot\rangle$ the inner product in H. Let us apply Itô's formula to the finite dimensional diffusion X_n. Since $X_n(t)\in H_n$, remark that

$$\begin{aligned}\big\langle X_n(t),\Pi_n D_\xi X_n^2(t)\big\rangle&=\big\langle \Pi_n X_n(t),D_\xi X_n^2(t)\big\rangle=\big\langle X_n(t),D_\xi X_n^2(t)\big\rangle\\&=\int_0^1 X_n(t,\xi)D_\xi X_n^2(t,\xi)\,\mathrm{d}\xi=\left[\frac{X_n^3(t,\xi)}{3}\right]_{\xi=0}^{\xi=1}=0\end{aligned}$$

because of no-slip boundary conditions. So, we obtain:

$$\begin{aligned}\mathrm{d}\big\|X_n(t)\big\|_2^2&=2\big\langle X_n(t),\mathrm{d}X_n(t)\big\rangle+\mathrm{tr}(Q_n)\,\mathrm{d}t\\&=\big[-2\big\|\nabla X_n(t)\big\|_2^2+\mathrm{tr}(Q_n)\big]\,\mathrm{d}t+2\big\langle X_n(t),G_n\,\mathrm{d}W(t)\big\rangle.\end{aligned}$$

In the same spirit, denoting by $d[Y,Y]_t$ the quadratic variation process of a semi-martingale Y, we can also compute with the Itô formula

$$\begin{aligned}\mathrm{d}\mathrm{e}^{\lambda_0\|X_n(t)\|_2^2}&=\mathrm{e}^{\lambda_0\|X_n(t)\|_2^2}\left[\lambda_0\mathrm{d}\big\|X_n(t)\big\|_2^2+\frac{\lambda_0^2}{2}\mathrm{d}\big[\|X_n\|_2^2,\|X_n\|_2^2\big]_t\right]\\&=\mathrm{e}^{\lambda_0\|X_n(t)\|_2^2}\big[-2\lambda_0\big\|\nabla X_n(t)\big\|_2^2+\lambda_0\,\mathrm{tr}(Q_n)+2\lambda_0^2\big\|G_n^*X_n(t)\big\|_2^2\big]\,\mathrm{d}t\\&\quad+2\lambda_0\mathrm{e}^{\lambda_0\|X_n(t)\|_2^2}\big\langle X_n(t),G_n\,\mathrm{d}W(t)\big\rangle.\end{aligned}$$

For any smooth function f on $H_n=\Pi_n(L^2)$, we define $g:=\mathcal{L}_n f$ if

$$f\big(X_n(t)\big)-f\big(X_n(0)\big)-\int_0^t g\big(X_n(s)\big)\,\mathrm{d}s$$

is a local martingale. The following lemma, being a consequence of Itô's formula, is well known to probabilists and it is crucial.

Lemma 4.2. ([16]) *If f is smooth on H_n, and $f \geqslant 1$, then*

$$M_t := \mathrm{e}^{-\int_0^t \frac{\mathcal{L}_n f}{f}(X_n(s))\,\mathrm{d}s} f\big(X_n(t)\big)$$

is a local martingale.

In view of the above definition we have for $f(x) = \mathrm{e}^{\lambda_0 \|x\|_2^2}$ $(x \in H_n)$,

$$\mathcal{L}_n f(x) := f(x)\big[-2\lambda_0 \|\nabla x\|_2^2 + \lambda_0 \operatorname{tr}(Q_n) + 2\lambda_0^2 \big\|G_n^* x\big\|_2^2\big]$$

and

$$\begin{aligned}
-\frac{\mathcal{L}_n f(x)}{f(x)} &= 2\lambda_0 \|\nabla x\|_2^2 - \lambda_0 \operatorname{tr}(Q_n) - 2\lambda_0^2 \big\|G_n^* x\big\|_2^2 \\
&\geqslant 2\lambda_0 \|\nabla x\|_2^2 - \lambda_0 \operatorname{tr}(Q) - 2\lambda_0^2 \|Q\| \|x\|_2^2.
\end{aligned}$$

Moreover, by the Poincaré inequality

$$\|x\|_2 \leqslant \frac{\|\nabla x\|_2}{\pi}, \quad \forall x \in H$$

we obtain for $0 < \lambda_0 \leqslant \frac{\pi^2}{2\|Q\|}$, since $1 - \frac{\lambda_0 \|Q\|}{\pi^2} \geqslant \frac{1}{2}$

$$\begin{aligned}
-\frac{\mathcal{L}_n f(x)}{f(x)} &\geqslant 2\lambda_0 \left(\|\nabla x\|_2^2 \left(1 - \frac{\lambda_0 \|Q\|}{\pi^2}\right) - \frac{\operatorname{tr}(Q)}{2} \right) \\
&\geqslant \lambda_0 \|\nabla x\|_2^2 - \lambda_0 \operatorname{tr}(Q).
\end{aligned}$$

So we conclude by Lemma 4.2 that

$$N_t^n := \exp\left(\lambda_0 \int_0^t \big\|\nabla X_n(s)\big\|_2^2 \,\mathrm{d}s - \lambda_0 \operatorname{tr}(Q) t \right) \mathrm{e}^{\lambda_0 \|X_n(t)\|_2^2} \tag{4.2}$$

is a supermartingale. This proves the following crucial exponential estimate:

Lemma 4.3. *Let $0 < \lambda_0 < \frac{\pi^2}{2\|Q\|}$. For any x in H, we have*

$$\mathbb{E}^x \exp\left(\lambda_0 \int_0^t \big\|\nabla X_n(s)\big\|_2^2 \,\mathrm{d}s \right) \mathrm{e}^{\lambda_0 \|X_n(t)\|_2^2} \leqslant \mathrm{e}^{\lambda_0 \operatorname{tr}(Q) t} \mathrm{e}^{\lambda_0 \|x\|_2^2}, \quad \forall t > 0. \tag{4.3}$$

In particular, we have

$$\sup_{n \in \mathbb{N}} \mathbb{E}^x \mathrm{e}^{\lambda_0 \int_0^t \|\nabla X_n(s)\|_2^2 \,\mathrm{d}s} < \infty$$

so $(X_n)_n$ is uniformly bounded in $L^2(\Omega \times [0, T], H_0^1)$.

This kind of estimates was also investigated by Da Prato and Debussche [2] for proving some properties on derivatives of the transition semigroup.

4.2. Proof of Theorem 4.1

Let us introduce the stochastic convolution, or Ornstein–Uhlenbeck process

$$W_\Delta(t) = \int_0^t S(t-s) G \,\mathrm{d}W(s)$$

which is the mild solution of the linear equation with additive noise

$$\mathrm{d}W_\Delta(t) = \Delta W_\Delta(t)\,\mathrm{d}t + G\,\mathrm{d}W(t); \quad W_\Delta(0) = 0. \tag{4.4}$$

Since $Q = GG^*$ has finite trace, it is known (see [6, p. 99 and p. 148]), that the stochastic integral W_Δ is the limit in $L^2(\Omega, C([0,T]; H))$ and in $L^2(\Omega, L^2([0,T]; H_0^1))$ of its finite dimensional approximation defined by

$$W_\Delta^n(t) = \int_0^t S(t-s)G_n\,\mathrm{d}W(s) = \Pi_n W_\Delta(t).$$

Notice that W_Δ^n is the mild solution of the finite dimensional linear equation with additive noise

$$\mathrm{d}W_\Delta^n(t) = \Delta W_\Delta^n(t)\,\mathrm{d}t + G_n\,\mathrm{d}W(t); \quad W_\Delta(0) = 0. \tag{4.5}$$

Let us prove that the convergences above hold in fact a.s. in $C([0,T]; H)$ and $L^2([0,T]; H_0^1)$. Indeed the a.s. convergence of W_Δ^n to W_Δ in $L^2([0,T]; H_0^1)$ is obvious. For the convergence in $C([0,T], H)$, since for a.e. ω, $t \mapsto W_\Delta(t,\omega)$ is continuous from $[0,T]$ to H, then $K := \{W_\Delta(t,\omega); t \in [0,T]\}$ is compact in H. Notice that if $h \in H$, $\Pi_n h \to h$ in H and that the mappings $h \mapsto \Pi_n h$, $n \geqslant 1$ are equi-continuous on H for $\|\Pi_n\|_{H\to H} = 1$. So the above pointwise convergence is uniform over the compact subset K by Arzela–Ascoli's theorem: as $n \to \infty$,

$$\sup_{t\in[0,T]} \left\| W_\Delta(t,\omega) - \Pi_n W_\Delta(t,\omega) \right\|_H \to 0.$$

Our proof below, as in [7], will be completely deterministic. Fix any $\omega \in \Omega$ such that $W_\Delta^n(\omega) \to W_\Delta(\omega)$ both in $C([0,T], H)$ and $L^2([0,T], H_0^1)$ and we shall remove "ω" in the proof below.

Let us define

$$y := X - W_\Delta = S(t)x + \frac{1}{2}\int_0^t S(t-s)D_\xi(X)^2(s)\,\mathrm{d}s,$$

$$y_n := X_n - W_\Delta^n = S(t)\Pi_n x + \frac{1}{2}\int_0^t S(t-s)\Pi_n D_\xi(X_n)^2(s)\,\mathrm{d}s$$

and

$$z_n := y - y_n = X - X_n - \left(W_\Delta - W_\Delta^n\right).$$

Recall that X is bounded in $L^\infty([0,T], H)$ and also in $L^2([0,T], H_0^1)$ almost surely. Indeed we have the following a-priori estimates (see [7, p. 264]):

$$\left\|y(t)\right\|_2^2 \leqslant \mathrm{e}^{8\int_0^t \|W_\Delta(s)\|_\infty^2\,\mathrm{d}s}\|x\|_2^2 + 2\int_0^t \mathrm{e}^{8\int_r^t \|W_\Delta(s)\|_\infty^2\,\mathrm{d}s}\left\|W_\Delta(r)\right\|_\infty^4\,\mathrm{d}r$$

and

$$\int_0^T \left\|D_\xi y(t)\right\|_2^2\,\mathrm{d}t \leqslant 8\int_0^T \left\|W_\Delta(t)\right\|_\infty^2 \left\|y(t)\right\|_2^4\,\mathrm{d}t + \int_0^T \left\|W_\Delta(t)\right\|_\infty^4\,\mathrm{d}t$$

and the fact that $H_0^1 \subset C[0,1]$ is a compact continuous embedding. The same proof as in [7, p. 264] yields the same estimates for y_n with W_Δ replaced by W_Δ^n, so the sequence $(X_n)_n$ is bounded in $L^\infty([0,T], H)$ and $L^2([0,T], H_0^1)$ almost surely (see also [3]). We can assume without loss of generality that the preceding bounds hold for our "ω".

It remains to show the convergence of z_n to 0 in the desired spaces. Notice that z_n is solution of

$$\frac{\mathrm{d}z_n}{\mathrm{d}t} = \Delta z_n + \frac{1}{2}D_\xi X^2 - \frac{1}{2}\Pi_n D_\xi X_n^2$$

from which we can deduce the a priori estimate:

$$\begin{aligned}\frac{1}{2}\frac{\mathrm{d}}{\mathrm{d}t}\|z_n\|_2^2+\|D_\xi z_n\|_2^2&=\left\langle\frac{1}{2}D_\xi(X)^2-\frac{1}{2}\Pi_nD_\xi(X_n)^2,z_n\right\rangle\\&=\left\langle\frac{1}{2}D_\xi\left(X^2-X_n^2\right),z_n\right\rangle+\left(\frac{1}{2}(I-\Pi_n)D_\xi\left(X^2\right),z_n\right)\\&:=I_1+I_2.\end{aligned}$$

Noting that $X-X_n=z_n+(I-\Pi_n)W_\Delta$, we have

$$\begin{aligned}I_1&=-\frac{1}{2}\left\langle(X_n+X)(X-X_n),D_\xi z_n\right\rangle\\&=-\frac{1}{2}\left\langle(X_n+X)z_n,D_\xi z_n\right\rangle-\frac{1}{2}\left\langle(X_n+X)(I-\Pi_n)W_\Delta,D_\xi z_n\right\rangle\\&:=I_{11}+I_{12}.\end{aligned}$$

We can bound I_{11} as follows

$$\begin{aligned}|I_{11}|&\leqslant\frac{1}{2}\|X_n+X\|_\infty\|z_n\|_2\|D_\xi z_n\|_2\\&\leqslant\frac{1}{4}\|D_\xi z_n\|_2^2+\|X_n+X\|_\infty^2\|z_n\|_2^2\end{aligned}$$

and for the second

$$\begin{aligned}|I_{12}|&\leqslant\frac{1}{2}\|X_n+X\|_\infty\left\|(I-\Pi_n)W_\Delta\right\|_2\|D_\xi z_n\|_2\\&\leqslant\frac{1}{4}\|D_\xi z_n\|_2^2+\|X_n+X\|_\infty^2\left\|(I-\Pi_n)W_\Delta\right\|_2^2.\end{aligned}$$

Similarly, for the remaining term, we have

$$|I_2|\leqslant\left\|(I-\Pi_n)D_\xi(X)\right\|_2^2+\|X\|_\infty^2\|z_n\|_2^2.$$

Hence we obtain the inequality

$$\begin{aligned}\frac{\mathrm{d}}{\mathrm{d}t}\|z_n\|_2^2+\|D_\xi z_n\|_2^2\leqslant{}&2\left(\|X\|_\infty^2+\|X_n+X\|_\infty^2\right)\|z_n\|_2^2+2\left\|(I-\Pi_n)D_\xi(X)\right\|_2^2\\&+2\|X_n+X\|_\infty^2\left\|(I-\Pi_n)W_\Delta\right\|_2^2.\end{aligned}\tag{4.6}$$

By Gronwall's inequality we get

$$\begin{aligned}\left\|z_n(t)\right\|_2^2\leqslant{}&\exp\left(\int_0^t2\left\|X(s)\right\|_\infty^2+2\left\|X_n(s)+X(s)\right\|_\infty^2\mathrm{d}s\right)\left\|(I-\Pi_n)x\right\|_2^2\\&+2\int_0^t\exp\left(\int_s^t2\left\|X(r)\right\|_\infty^2+2\left\|X_n(r)+X(r)\right\|_\infty^2\mathrm{d}r\right)\left\|X_n(s)+X(s)\right\|_\infty^2\left\|(I-\Pi_n)W_\Delta(s)\right\|_2^2\mathrm{d}s\\&+2\int_0^t\exp\left(\int_s^t2\left\|X(r)\right\|_\infty^2+2\left\|X_n(r)+X(r)\right\|_\infty^2\mathrm{d}r\right)\left\|(I-\Pi_n)D_\xi X(s)\right\|_2^2\mathrm{d}s.\end{aligned}$$

In the sequel we denote the norm in the corresponding spaces respectively by

$$|u|_{L^2(0,T,H)}^2:=\int_0^T\left\|u(t)\right\|_2^2\mathrm{d}t,$$

$$|u|_{L^2(0,T,H_0^1)}^2:=\int_0^T\left\|\nabla u(t)\right\|_2^2\mathrm{d}t=\int_0^T\left\|D_\xi u(t)\right\|_2^2\mathrm{d}t,$$

$$|u|_{C(0,T,H)} := \sup_{0\leqslant t\leqslant T} \|u(t)\|_2.$$

Taking the supremum in t, and using again the compact continuous embedding $H_0^1(0,1) \subset C[0,1]$, so that $\|x\|_\infty \leqslant C\|\nabla x\|_2$ for some constant $C > 0$, we obtain

$$\begin{aligned}\sup_{0\leqslant t\leqslant T} \|z_n(t)\|_2^2 &\leqslant \mathrm{e}^{(4C^2|X_n|^2_{L^2(0,T,H_0^1)} + 6C^2|X|^2_{L^2(0,T,H_0^1)})}((1)+(2)+(3)) \\ &\leqslant \mathrm{e}^{M_1}((1)+(2)+(3)) \end{aligned}\tag{4.7}$$

for some number $M_1 > 0$, where

$$\begin{aligned}(1) &= \|(I-\Pi_n)x\|_2^2, \\ (2) &= 4C^2\big(|X_n|^2_{L^2(0,T,H_0^1)} + |X|^2_{L^2(0,T,H_0^1)}\big)\big|(I-\Pi_n)W_\Delta\big|^2_{C(0,T,H)} \\ &\leqslant M_2\big|(I-\Pi_n)W_\Delta\big|^2_{C(0,T,H)}, \\ (3) &= 2\int_0^T \|(I-\Pi_n)D_\xi X(s)\|_2^2\,\mathrm{d}s \end{aligned}$$

for some constant $M_2 > 0$.

Now, (1) $\to 0$ is clear, (2) $\to 0$ is assumed for our "ω", and (3) $\to 0$ by dominated convergence. Consequently, $z_n \to 0$ in $C([0,T],H)$.

Finally, let us integrate (4.6) for t. It gives

$$\begin{aligned}|z_n|^2_{L^2([0,T],H_0^1)} \leqslant{}& \big(4C^2|X_n|^2_{L^2([0,T],H_0^1)} + 6C^2|X|^2_{L^2([0,T],H_0^1)}\big)\sup_{0\leqslant t\leqslant T}\|z_n(t)\|_2^2 + 2\big|(I-\Pi_n)D_\xi(X)\big|^2_{L^2([0,T],H)} \\ &+ \|(I-\Pi_n)x\|_2^2 + 4C^2\big(|X_n|^2_{L^2([0,T],H_0^1)} + |X|^2_{L^2([0,T],H_0^1)}\big)\big|(I-\Pi_n)W_\Delta\big|^2_{C([0,T],H)} \\ \leqslant{}& M_1\sup_{0\leqslant t\leqslant T}\|z_n(t)\|_2^2 + 2\big|(I-\Pi_n)D_\xi(X)\big|^2_{L^2([0,T],H)} \\ &+ \|(I-\Pi_n)x\|_2^2 + M_2\big|(I-\Pi_n)W_\Delta\big|^2_{C([0,T],H)}\end{aligned}$$

which yields $z_n \to 0$ in $L^2(0,T,H_0^1)$ and the proof is finished. □

5. Uniform upper bound for the weak convergence topology: the exponential tightness

In this section, $M_1(H)$ is equipped with $\sigma(M_1(H), C_b(H))$ the weak convergence topology, instead of τ. The aim is to prove the following

Proposition 5.1.

(a) *For any $\varepsilon > 0$, there is some compact subset $K = K_\varepsilon$ in $M_1(H)$ in the weak convergence topology such that*

$$\limsup_{t\to\infty}\frac{1}{t}\log\sup_{\nu\in\mathcal{M}_{\lambda_0,L}}\mathbb{P}_\nu(L_t\notin K) \leqslant -\frac{1}{\varepsilon}.$$

(b) *Consequently for any closed set F in $M_1(H)$ equipped with the weak convergence topology $\sigma(M_1(H), C_b(H))$,*

$$\limsup_{t\to\infty}\frac{1}{t}\log\sup_{\nu\in\mathcal{M}_{\lambda_0,L}}\mathbb{P}_\nu(L_t\in F) \leqslant -\inf_F J.$$

By the weak upper bound in Lemma 3.4 and according to the general theory of large deviations, the upper bound of large deviations in (b) follows from the uniform exponential tightness of the family of $\mathbb{P}_\nu(L_T \in \cdot)$ over $\nu \in \mathcal{M}_{\lambda_0,L}$ stated in part (a).

Before proving it, let us notice the following consequence of our study in Section 4.

Lemma 5.2. *For any fixed* $0 < \lambda_0 < \frac{\pi^2}{2\|Q\|}$,

$$N_t := \exp\left(\lambda_0 \int_0^t \|\nabla X(s)\|_2^2 \, ds - \lambda_0 \operatorname{tr}(Q)t\right) e^{\lambda_0 \|X(t)\|_2^2}$$

is a supermartingale. In particular we have

$$\mathbb{E}^x e^{\lambda_0 \int_0^t \|\nabla X(s)\|_2^2 \, ds} \leqslant e^{\lambda_0 \operatorname{tr}(Q)t} e^{\lambda_0 \|x\|_2^2}, \quad \forall x \in H \tag{5.1}$$

and for any fixed $L > 1$, *and any initial measure in the set* $\mathcal{M}_{\lambda_0,L}$, *the following estimate holds*

$$\mathbb{E}^\nu e^{\lambda_0 \int_0^t \|\nabla X(s)\|_2^2 \, ds} \leqslant e^{\lambda_0 \operatorname{tr}(Q)t} L. \tag{5.2}$$

Proof. By the almost sure convergence in Theorem 4.1 and Fatou's Lemma, (N_t) is a supermartingale by passing to the limit for $n \to \infty$ in (4.2). The estimates (5.1) and (5.2) follow immediately. □

Proof of Proposition 5.1. As said above it is sufficient to prove the uniform exponential tightness of $(\mathbb{P}_\nu(L_t \in \cdot), t \to +\infty)$ over $\nu \in \mathcal{M}_{\lambda_0,L}$ in part (a).

Step 1. Define $\Phi : M_1(H) \to [0, +\infty]$ by

$$\Phi(\beta) = \lambda_0 \int_H \|\nabla x\|_2^2 \, d\beta(x), \quad \text{with } \|\nabla x\|_2 := +\infty \text{ for } \forall x \in H \backslash H_0^1,$$

where λ_0 is a real number such that $0 < \lambda_0 < \frac{\pi^2}{2\|Q\|}$. We claim that this function admits compact level sets.

At first, $x \to \|\nabla x\|_2^2$ is lower semi continuous (l.s.c. in short) on H, as a non-decreasing limit of continuous functions $x \to \|\nabla \Pi_n x\|_2^2$. Thus, Φ is l.s.c. on $M_1(H)$, and for any $a > 0$, the level set $[\Phi \leqslant a]$ is closed in $M_1(H)$.

Now let us show that $[\Phi \leqslant a]$ is tight (so it will be compact in $M_1(H)$ by Prokorov's criterion). Indeed, for any $\delta > 0$ consider

$$A_\delta = \left\{ x \in H_0^1 \text{ s.t. } \|\nabla x\|_2 \leqslant \sqrt{\frac{a}{\lambda_0 \delta}} \right\}.$$

It is compact in H by the compact embedding $H_0^1 \subset H$, and we have

$$\forall \beta \in [\Phi \leqslant a], \quad \beta(A_\delta^c) \leqslant \int_{A_\delta^c} \frac{\lambda_0 \delta \|\nabla x\|_2^2}{a} \, d\beta(x) \leqslant \delta \frac{\Phi(\beta)}{a} \leqslant \delta.$$

Step 2. For any $\varepsilon > 0$, $K := [\Phi \leqslant \lambda_0 \operatorname{tr}(Q) + 1/\varepsilon]$ is a compact subset of $M_1(H)$ by Step 1. For any $\nu \in \mathcal{M}_{\lambda_0,L}$, we have by Chebychev's inequality and Lemma 5.2,

$$\begin{aligned}
\mathbb{P}_\nu(L_t \notin K) &\leqslant \exp\left(-\left[\lambda_0 \operatorname{tr}(Q) + \frac{1}{\varepsilon}\right] t\right) \mathbb{E}^\nu e^{t \Phi(L_t)} \\
&= \exp\left(-\left[\lambda_0 \operatorname{tr}(Q) + \frac{1}{\varepsilon}\right] t\right) \mathbb{E}^\nu \exp\left(\lambda_0 \int_0^t \|\nabla X(s)\|_2^2 \, ds\right) \\
&\leqslant e^{-t/\varepsilon} L,
\end{aligned}$$

the desired uniform exponential tightness. □

6. Uniform upper bound for the τ-topology

Now, we prove the desired upper bound (1.5) for the strong τ-topology. It is based on the following criterion of the so-called hyper-exponential recurrence [20, Theorem 2.1] established by Wu for strong Feller and topologically irreducible Markov processes.

Lemma 6.1. ([20]) *For a subset K in H, let us define $\tau_K := \inf\{t \geqslant 0$ s.t. $X_t \in K\}$ and $\tau_K^{(1)} := \inf\{t \geqslant 1$ s.t $X_t \in K\}$. If for any $\lambda > 0$, there exists a compact subset K in H such that*

$$\sup_{\nu \in \mathcal{M}_{\lambda_0,L}} \mathbb{E}^\nu \mathrm{e}^{\lambda \tau_K} < \infty \tag{6.1}$$

and

$$\sup_{x \in K} \mathbb{E}^x \mathrm{e}^{\lambda \tau_K^{(1)}} < \infty \tag{6.2}$$

then $[J \leqslant a]$ is τ-compact for every $a \in \mathbb{R}^+$, and the upper bound (1.5) *uniform on $\mathcal{M}_{\lambda_0,L}$ for the τ-topology holds true.*

In this section we establish the estimates (6.1) and (6.2) for our model. For the compact subset K of H, we still consider

$$K := \{x \in H_0^1 \text{ s.t. } \|\nabla x\|_2 \leqslant M\}, \tag{6.3}$$

where the real number M will be fixed later. The definition of the occupation measure implies that for $n \geqslant 2$,

$$\mathbb{P}_\nu\big(\tau_K^{(1)} > n\big) \leqslant \mathbb{P}_\nu\left(L_n(K) \leqslant \frac{1}{n}\right) = \mathbb{P}_\nu\left(L_n(K^c) \geqslant 1 - \frac{1}{n}\right).$$

With our choice for K, we remark that

$$L_n\big(K^c\big) \leqslant \frac{1}{M^2} L_n\big(\|\nabla x\|_2^2\big).$$

Hence for any fixed real $0 < \lambda_0 < \frac{\pi^2}{2\|Q\|}$, we have by Chebychev's inequality

$$\begin{aligned}\mathbb{P}_\nu\big(\tau_K^{(1)} > n\big) &\leqslant \mathbb{P}_\nu\left(L_n\big(\|\nabla x\|_2^2\big) > M^2\left(1 - \frac{1}{n}\right)\right) \\ &\leqslant \mathrm{e}^{-n\lambda_0 M^2(1-\frac{1}{n})}\mathbb{E}^\nu \mathrm{e}^{\lambda_0 \int_0^n \|\nabla X_s\|_2^2\,\mathrm{d}s}.\end{aligned}$$

For any initial measure $\nu \in M_1(H)$, integrating (5.1) w.r.t. ν, and using it in the above expression yields

$$\mathbb{P}_\nu\big(\tau_K^{(1)} > n\big) \leqslant \int_H \mathrm{e}^{\lambda_0 \|x\|_2^2} \nu(\mathrm{d}x) \mathrm{e}^{-n\lambda_0 C}, \quad \forall n \geqslant 2,$$

where $C := M^2/2 - \mathrm{tr}(Q)$.

Let $\lambda > 0$ be any fixed real number. By the integration by parts formula, we have

$$\begin{aligned}\mathbb{E}^\nu \mathrm{e}^{\lambda \tau_K^{(1)}} &= 1 + \int_0^{+\infty} \lambda \mathrm{e}^{\lambda t} \mathbb{P}_\nu\big(\tau_K^{(1)} > t\big)\,\mathrm{d}t \\ &\leqslant \mathrm{e}^{2\lambda} + \sum_{n \geqslant 2} \lambda \mathrm{e}^{\lambda(n+1)} \mathbb{P}_\nu\big(\tau_K^{(1)} > n\big) \\ &\leqslant \mathrm{e}^{2\lambda}\left(1 + \lambda \int_H \mathrm{e}^{\lambda_0 \|x\|_2^2} \nu(\mathrm{d}x) \sum_{n \geqslant 2} \mathrm{e}^{-n(\lambda_0 C - \lambda)}\right).\end{aligned}$$

Now, by the definition (6.3) of K, we can choose M such that $\lambda_0 C - \lambda \geqslant 1$. Then, taking the supremum over $\{\nu = \delta_x, x \in K\}$, we get

$$\sup_{x\in K} \mathbb{E}^x \mathrm{e}^{\lambda \tau_K^{(1)}} \leqslant \mathrm{e}^{2\lambda}\left(1 + \lambda \mathrm{e}^{\lambda_0 M^2} \sum_{n\geqslant 2} \mathrm{e}^{-n(\lambda_0 C - \lambda)}\right) < \infty$$

for $\forall x \in K$, $\|x\|_2 \leqslant \|\nabla x\|_2/\pi \leqslant M$. So the bound (6.1) holds true. We obtain (6.2) in the same way: since $\tau_K \leqslant \tau_K^{(1)}$, we have

$$\begin{aligned}\sup_{\nu\in\mathcal{M}_{\lambda_0,L}} \mathbb{E}^\nu \mathrm{e}^{\lambda\tau_K} &\leqslant \sup_{\nu\in\mathcal{M}_{\lambda_0,L}} \mathbb{E}^\nu \mathrm{e}^{\lambda\tau_K^{(1)}} \\ &\leqslant \mathrm{e}^{2\lambda}\left(1 + \lambda L \sum_{n\geqslant 2} \mathrm{e}^{-n(\lambda_0 C - \lambda)}\right) \\ &< \infty.\end{aligned}$$

Proof of Theorem 1.1. At first the good uniform upper bound of large deviations, i.e., parts (i) and (iii) follows by Lemma 6.1 for its conditions (6.1) and (6.2) are verified above.

The uniform lower bound in part (ii) was established in Proposition 3.3.

The first claim in (1.6): "$J(\nu) < +\infty \Rightarrow \nu \ll \mu$" was proven in Lemma 3.1. We conclude the proof with the second claim in (1.6) that for $\nu \in M_1(H)$ with $J(\nu) < \infty$, $\nu(\|\nabla x\|_2^2) < \infty$. Indeed, denoting by $a \wedge b$ the minimum of two real numbers a and b, and for the function $V_n(x) := (\lambda_0 \|\nabla x\|_2^2) \wedge n$ bounded and measurable on H, we have

$$\begin{aligned}\nu(V_n) &\leqslant \left(\Lambda^0\right)^*(\nu) + \Lambda^0(V_n) \\ &\leqslant J(\nu) + \lambda_0 \operatorname{tr}(Q),\end{aligned} \tag{6.4}$$

where we have used the definitions (3.7), (3.8), the crucial estimate (5.1) and the fact that $(\Lambda^0)^* = J$. The conclusion follows by Fatou's lemma. □

7. Extension to some unbounded functionals

In this section we point out the fact that the estimate in Lemma 5.2 is sufficient to extend the LDP of Theorem 1.1, i.e. Corollary 1.2 for unbounded functionals and its consequences.

Proof of Corollary 1.2. For the measurable function $f : H_0^1 \to \mathbb{B}$, let us consider $f_n : H \to \mathbb{B}$ defined by

$$f_n(x) := \begin{cases} f(x), & \text{if } x \in H_0^1,\ \|\nabla x\|_2 \leqslant n, \\ 0, & \text{otherwise} \end{cases} \tag{7.1}$$

which is far from being continuous, but is measurable and bounded on H by our assumptions. Since $\nu \to \nu(f_n) = \int_{\mathbb{B}} z\nu(f_n \in \mathrm{d}z)$ is continuous from $(M_1(H), \tau)$ to $\mathbb{B}$ by [10, Lemma 3.3.8], then $L_T(f_n)$ satisfies the LDP by Theorem 1.1 and the contraction principle.

Now by the approximation lemma in large deviations (see [10, p. 37]), it remains to prove that for any $L > 0$

$$\lim_{n\to\infty} \sup_{\beta:\ J(\beta)\leqslant L} \left\|\beta(f_n) - \beta(f)\right\|_{\mathbb{B}} = 0 \tag{7.2}$$

and for any $\delta > 0$,

$$\lim_{n\to\infty} \limsup_{T\to\infty} \frac{1}{T} \log \sup_{\nu\in\mathcal{M}_{\lambda_0,L}} \mathbb{P}_\nu\left(\left\|L_T(f - f_n)\right\|_{\mathbb{B}} > \delta\right) = -\infty. \tag{7.3}$$

Thanks to our condition (1.7) on f, we can construct a sequence $(\varepsilon(n))_n$ decreasing to 0 such that once $\|\nabla x\|_2 \geqslant n$,

$$\left\|f(x)\right\|_{\mathbb{B}} \leqslant \varepsilon(n)\|\nabla x\|_2^2.$$

Denoting by 1_A the characteristic function of the set A, we have for any β satisfying $J(\beta) \leqslant L$,

$$\begin{aligned}
\left\|\beta(f_n)-\beta(f)\right\|_{\mathbb{B}} &= \left\|\beta(f 1_{\{\|\nabla X(s)\|_2 \geqslant n\}})\right\|_{\mathbb{B}} \\
&\leqslant \beta\big(\varepsilon(n)\|\nabla x\|_2^2 1_{\{\|\nabla X(s)\|_2 \geqslant n\}}\big) \\
&\leqslant \frac{\varepsilon(n)}{\lambda_0}\beta\big(\lambda_0\|\nabla x\|_2^2\big) \\
&\leqslant \frac{\varepsilon(n)}{\lambda_0}\big(L+\lambda_0 \operatorname{tr}(Q)\big)
\end{aligned}$$

by using (6.4). Hence (7.2) follows.

Let us evaluate

$$\begin{aligned}
\mathbb{P}_\nu\big(\left\|L_T(f-f_n)\right\|_{\mathbb{B}} > \delta\big) &= \mathbb{P}_\nu\left(\left\|\frac{1}{T}\int_0^T f(X_s)-f_n(X_s)\,\mathrm{d}s\right\|_{\mathbb{B}} > \delta\right) \\
&\leqslant \mathbb{P}_\nu\left(\frac{1}{T}\int_0^T \varepsilon(n)\left\|\nabla X(s)\right\|_2^2 1_{\{\|\nabla X(s)\|_2 \geqslant n\}}\,\mathrm{d}s > \delta\right) \\
&\leqslant \mathbb{P}_\nu\left(\int_0^T \lambda_0\left\|\nabla X(s)\right\|_2^2 1_{\{\|\nabla X(s)\|_2 \geqslant n\}}\,\mathrm{d}s > \frac{\lambda_0 T\delta}{\varepsilon(n)}\right) \\
&\leqslant \exp\left(-\frac{\lambda_0 T\delta}{\varepsilon(n)}\right)\mathbb{E}^\nu \exp\left(\lambda_0\int_0^T \left\|\nabla X(s)\right\|_2^2\,\mathrm{d}s\right)
\end{aligned}$$

so that (7.3) is consequence of (5.2). □

Remerciements

L'auteur remercie vivement Liming Wu pour les nombreuses discussions enrichissantes et l'attention qu'il a portée à ce travail.

References

[1] C. Cardon-Weber, Large deviations for a Burgers type SPDE, Stochastic Process. Appl. 84 (1999) 53–70.
[2] G. Da Prato, A. Debussche, Differentiability of the transition semigroup of the stochastic Burgers equation, and application to the corresponding Hamilton Jacobi equation, Atti Accad. Naz. Lincei Cl. Sci. Fis. Mat. Natur. Rend. Lincei Mat. Appl. 9 (1998) 267–277.
[3] G. Da Prato, A. Debussche, Dynamic programming for the stochastic Burgers equation, Ann. Mat. Pura Appl. (IV) CLXXVIII (2000) 143–174.
[4] G. Da Prato, D. Gatarek, Stochastic Burgers equation with correlated noise, Stoch. Stoch. Rep. 52 (1995) 29–41.
[5] G. Da Prato, A. Debussche, R. Temam, Stochastic Burgers equation, Nonlinear Anal. 1 (1994) 389–402.
[6] G. Da Prato, J. Zabczyk, Stochastic Equations in Infinite Dimensions, Cambridge University Press, 1992.
[7] G. Da Prato, J. Zabczyk, Ergodicity for Infinite Dimensional Systems, Cambridge University Press, 1996.
[8] A. Dembo, O. Zeitouni, Large Deviations Techniques and Applications, second ed., Springer-Verlag, 1998.
[9] A. Dermoune, Around the stochastic Burgers equation, Stoch. Anal. Appl. 15 (2) (1997) 295–311.
[10] J.D. Deuschel, D.W. Stroock, Large Deviations, Pure Appl. Math., vol. 137, Academic Press, San Diego, 1989.
[11] M.D. Donsker, S.R.S. Varadhan, Asymptotic evaluation of certain Markov process expectations for large time, I–IV, Comm. Pure Appl. Math. 28 (1975) 1–47; 279–301 (1975); 29, 389–461 (1976); 36, 183–212 (1983).
[12] F. Flandoli, B. Maslowski, Ergodicity of the 2D Navier–Stokes equation under random perturbation, Comm. Math. Phys. 171 (1995) 119–141.
[13] B. Goldys, B. Maslowski, Exponential ergodicity for stochastic burgers and 2D Navier–Stokes equation, J. Funct. Anal. 226 (1) (2005) 230–255.
[14] S.P. Meyn, R.L. Tweedie, Markov Chains and Stochastic Stability, Springer-Verlag, 1993.
[15] E. Pardoux, Equations aux dérivées partielles stochastiques non linéaires monotones, Ph.D. thesis. Université Paris XI, 1975.
[16] D. Revuz, M. Yor, Continuous Martingales and Brownian Motion, third ed., Springer-Verlag, 1999.
[17] B.L. Rozovski, Stochastic Evolution Systems: Linear Theory and Application to Non Linear Filtering, Kluver Academic, 1990.
[18] E. Weinan, K. Khanin, A. Mazel, Ya. Sinai, Invariant measures for Burgers equation with stochastic forcing, Ann. of Math. 151 (2000) 877–960.

[19] L. Wu, Uniformly integrable operators and large deviations for Markov processes, J. Funct. Anal. 172 (2000) 301–376.
[20] L. Wu, Large and moderate deviations and exponential convergence for stochastic damping Hamiltonian systems, Stochastic Proc. Appl. 91 (2001) 205–238.
[21] L. Wu, Essential spectral radius for Markov semigroups (I): discrete time case, Probab. Theory Related Fields 128 (2004) 255–321.