HAL open science

Large deviations from a stationary measure for a class of dissipative PDE's with random kicks

Vojkan Jaksic, Vahagn Nersesyan, Claude-Alain Pillet, Armen Shirikyan

To cite this version:

Vojkan Jaksic, Vahagn Nersesyan, Claude-Alain Pillet, Armen Shirikyan. Large deviations from a stationary measure for a class of dissipative PDE's with random kicks. Communications on Pure and Applied Mathematics, 2015, 68 (12), pp.2108-2143. 10.1002/cpa.21568 . hal-00760418v2

HAL Id: hal-00760418

https://hal.science/hal-00760418v2

Submitted on 15 Oct 2014

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Large deviations from a stationary measure for a class of dissipative PDE's with random kicks

V. Jaksic* V. Nersesyan† C.-A. Pillet‡ A. Shirikyan§

Abstract

We study a class of dissipative PDE's perturbed by a bounded random kick force. It is assumed that the random force is non-degenerate, so that the Markov process obtained by the restriction of solutions to integer times has a unique stationary measure. The main result of the paper is a large deviation principle for occupation measures of the Markov process in question. The proof is based on Kifer's large deviation criterion, a coupling argument for Markov processes, and an abstract result on large-time asymptotic for generalised Markov semigroups.

AMS subject classifications: 35Q30, 76D05, 60B12, 60F10

Keywords: Dissipative PDE's, Navier–Stokes system, Ginzburg–Landau equation, large deviations, occupation measures

Contents

0 Introduction **2**

1 The model and the results **6**

1.1 The model 6
1.2 The results 8
1.3 Applications 10
1.4 Scheme of the proof of Theorem 1.3 18
1.5 Scheme of the proof of Theorem 1.4 19
1.6 Uniform large deviations principle 20

2 Large-time asymptotics for generalised Markov semigroups **21**

*Department of Mathematics and Statistics, McGill University, 805 Sherbrooke Street West, Montreal, QC, H3A 2K6 Canada; e-mail: Jaksic@math.mcgill.ca

†Laboratoire de Matématiques, UMR CNRS 8100, Université de Versailles-Saint-Quentin-en-Yvelines, F-78035 Versailles, France; e-mail: Vahagn.Nersesyan@math.uvsq.fr

‡Aix Marseille Université, CNRS, CPT, UMR 7332, Case 907, 13288 Marseille, France; Univeristé de Toulon, CNRS, CPT, UMR 7332, 83957 La Garde, France; e-mail: Pillet@univ-tln.fr

§Department of Mathematics, University of Cergy–Pontoise, CNRS UMR 8088, 2 avenue Adolphe Chauvin, 95302 Cergy–Pontoise, France; e-mail: Armen.Shirikyan@u-cergy.fr

3 The uniform Feller property **25**
3.1 Coupling 25
3.2 Proof of Theorem 3.1 26

4 Proof of Theorem 1.3 **27**

5 Proof of Theorem 1.4 **29**

6 Appendix **31**
6.1 Kifer's sufficient condition for LDP 31
6.2 Exponential equivalence of random probability measures 32
6.3 Dawson–Gärtner theorem 32

Bibliography **34**

0 Introduction

This paper is devoted to the large deviations principle (LDP) for a class of dissipative PDE's perturbed by a smooth random force. The large-time asymptotics of solutions for the problem in question is well understood, and we refer the reader to the articles [FM95, KS00, EMS01, BKL02] for the first results in this direction and to the book [KS12] for further references and a detailed description of the behaviour of solutions as time goes to infinity. In particular, it is known that if the noise is sufficiently non-degenerate, then the Markov process associated with the problem has a unique stationary distribution, which attracts exponentially the law of all solutions. Moreover, the law of iterated logarithm and the central limit theorem hold for Hölder-continuous functionals calculated on trajectories and give some information about fluctuations of their time averages around the mean value. Our aim now is to investigate the probabilities of deviations of order one from the mean value.

Let us describe in more detail the main result of this paper on the example of the Navier–Stokes system. More precisely, we consider the following problem in a bounded domain[1] $D \subset \mathbb{R}^2$ with a C^2-smooth boundary ∂D:

$$\dot{u} - \nu\Delta u + \langle u, \nabla\rangle u + \nabla p = \eta(t,x), \quad \operatorname{div} u = 0, \quad u\big|_{\partial D} = 0, \tag{0.1}$$

$$u(0,x) = u_0(x), \tag{0.2}$$

where $\nu > 0$ is the kinematic viscosity, $u = (u_1(t,x), u_2(t,x))$ is the velocity field of the fluid, $p = p(t,x)$ is the pressure, and η is a random external force. We assume that $\eta(t,x)$ is a random kick force of the form

$$\eta(t,x) = \sum_{k=1}^{+\infty} \delta(t-k)\eta_k(x), \tag{0.3}$$

[1] All the results of this paper remain true for periodic boundary conditions, in which case we assume in addition that the mean values of the velocity field and of the external force are zero.

where δ is the Dirac measure concentrated at zero, and η_k are independent identically distributed (i.i.d.) random variables defined on a probability space $(\Omega, \mathcal{F}, \mathbb{P})$ with range in $L^2(D, \mathbb{R}^2)$ that satisfy

$$\mathbb{P}\{\|\eta_k\|_{L^2} \le b\} = 1 \tag{0.4}$$

for some $b < +\infty$. Problem (0.1), (0.2) is well posed in the space

$$H = \{u \in L^2(D, \mathbb{R}^2) : \operatorname{div} u = 0 \text{ in } D, \langle u, \boldsymbol{n}\rangle = 0 \text{ on } \partial D\}, \tag{0.5}$$

where $\boldsymbol{n}$ stands for the outside unit normal to ∂D. The restrictions of solutions for (0.1), (0.2) to integer times form a Markov chain in H. As is well known (see Chapter 3 of the book [KS12] and the references therein), this process is ergodic under rather general hypotheses on η_k. More precisely, suppose that there exists an increasing sequence of finite-dimensional subspaces $H_N \subset H$ such that the law of the projection of η_k to H_N is absolutely continuous with respect to the Lebesgue measure, and its support contains the origin. Let $\mathcal{P}(H)$ be the set of all Borel probability measures on H endowed with topology of weak convergence. Then the Markov chain in question possesses a unique stationary measure $\mu \in \mathcal{P}(H)$, which is exponentially mixing in the sense that the law of any solution of (0.1) with a deterministic initial condition converges to μ exponentially fast in the Kantorovich–Wasserstein metric. We wish to investigate the probabilities of large deviations of the occupation measures from μ. More precisely, let

$$\zeta_k^\omega = \frac{1}{k}\sum_{j=0}^{k-1} \delta_{v_j}, \quad k \ge 1, \tag{0.6}$$

be a sequence of random probability measures in $\mathcal{P}(H)$, where $\{v_j\}$ denotes a stationary trajectory of the Markov chain. The following theorem is a simplified version of the main result of this paper (see Theorem 1.3).

Theorem A. *Under the above hypotheses, the sequence $\{\zeta_k\}$ satisfies a LDP. More precisely, there is a lower semicontinuous mapping $I : \mathcal{P}(H) \to [0, +\infty]$ which is equal to $+\infty$ outside a compact subset such that*

$$-\inf_{\lambda\in\dot\Gamma} I(\lambda) \le \liminf_{k\to\infty} \frac{1}{k}\log \mathbb{P}\{\zeta_k \in \Gamma\} \le \limsup_{k\to\infty} \frac{1}{k}\log \mathbb{P}\{\zeta_k \in \Gamma\} \le -\inf_{\lambda\in\overline{\Gamma}} I(\lambda), \tag{0.7}$$

where $\Gamma \subset \mathcal{P}(H)$ is an arbitrary Borel subset, and $\dot\Gamma$ and $\overline{\Gamma}$ denote its interior and closure, respectively.

For instance, if $f : H \to \mathbb{R}^m$ is a continuous mapping and $B \subset \mathbb{R}^m$ is a Borel subset, then taking $\Gamma = \{\sigma \in \mathcal{P}(H) : \int_H f \mathrm{d}\sigma \in B\}$ in inequality (0.7), we get (see Section 1.2 for a precise statement)

$$\exp(-c_-\, k) \lesssim \mathbb{P}\Big\{\frac{1}{k}\sum_{j=0}^{k-1} f(v_j) \in B\Big\} \lesssim \exp(-c_+\, k) \quad \text{as } k \to \infty,$$

where $c_\pm = c_\pm(f, B) \geq 0$ are some constants (not depending on k) that can be expressed in terms of the rate function I.

Let us mention that the LDP is well understood for finite-dimensional diffusions and for Markov processes with compact phase space, provided that the randomness is sufficiently non-degenerate and ensures mixing in the total variation norm. This type of results were first obtained by Donsker and Varadhan [DV75] and later extended by many others. A detailed account of the main achievements can be found in the books [FW84, DS89, DZ00].

In the context of randomly forced PDE's, the problem of large deviations was studied in a number of papers. Most of them, however, are devoted to studying PDE's with vanishing random perturbation and provide estimates for the probabilities of deviations from solutions of the limiting deterministic equations. We refer the reader to the papers [Fre88, Sow92a, Sow92b, Cha96, CR04, CR05, SS06, CM10] and the references therein for various results of this type, including the asymptotics of stationary distributions when the amplitude of the perturbation goes to zero. To the best of our knowledge, the only papers devoted to large deviations from a stationary distribution in the case of stochastic PDE's are those by Gourcy [Gou07b, Gou07a]. Using a general result due to Wu [Wu01], he established the LDP for occupation measures of stochastic Burgers and Navier–Stokes equations, provided that the random force is white in time and sufficiently irregular in the space variables. The present paper gives a first result on large deviations from a stationary distribution for PDE's with a *smooth* random perturbation.

Let us note that, in Gourcy's papers, the set of measures is endowed with the *τ-topology* which is generated by the duality with respect to bounded Borel functions (and is much stronger than the weak topology used in our paper). This enables one to apply the LDP to physically relevant observables that are not continuous on the energy space. Under our assumptions, the LDP is not likely to hold for the τ-topology. However, the results established in this paper can be applied to derive the LDP for functionals that are continuous on higher Sobolev spaces. Furthermore, using the Dawson–Gärtner theorem [DG87], we establish the following result on large deviations in the space of trajectories (also called process level LDP). Let us denote by $\boldsymbol{H}$ the space of sequences $\boldsymbol{u} = (u_j, j \geq 0)$ with $u_j \in H$ and endow it with the Tikhonov topology. Given a stationary trajectory $\boldsymbol{v} = \{v_j\}$ for the Markov chain associated with (0.1), we define the sequence of occupation measures

$$\boldsymbol{\zeta}_k^\omega = \frac{1}{k} \sum_{j=0}^{k-1} \delta_{\boldsymbol{v}_j}, \tag{0.8}$$

where we set $\boldsymbol{v}_j = (v_i, i \geq j)$.

Theorem B. *Let us assume that the above-mentioned hypotheses are satisfied. Then the LDP holds for $\boldsymbol{\zeta}_k$ with a rate function $\boldsymbol{I} : \mathcal{P}(\boldsymbol{H}) \to [0, +\infty]$.*

In conclusion, let us mention that the LDP discussed above remains valid in the case of unbounded perturbations; this question will be addressed in a

subsequent publication. We also remark that this paper is a first step of a research program whose aim is to develop a large deviation theory for dissipative PDE's with random perturbation and to justify the Gallavotti–Cohen fluctuation principle for some relevant functionals; cf. [GC95].

The paper is organised as follows. In Section 1 we introduce the model, state our results, describe applications, and outline the schemes of the proofs. Section 2 deals with the large-time asymptotics of generalised Markov semigroups. A central technical part of the proof is the verification of uniform Feller property for a suitable family of semigroups. This verification is based on a coupling argument and is carried out in Section 3. The proofs of the main results are given in Sections 4 and 5. Finally, three auxiliary results used in the main text are recalled in the Appendix.

Acknowledgments

The research of VJ was supported by NSERC. The research of VN was supported by the ANR grants EMAQS (No. ANR 2011 BS01 017 01) and STOSYMAP (No. ANR 2011 BS01 015 01). The research of AS was carried out within the MME-DII Center of Excellence (ANR-11-LABX-0023-01) and supported by the ANR grant STOSYMAP and RSF research project 14-49-00079. The research of CAP was partly supported by ANR grant 09-BLAN-0098.

Notation

Let $\mathbb{Z}$ be the set of integers, let $\mathbb{Z}_+$ be the set of non-negative integers, and let X be a Polish space with a metric $d_X(u,v)$. We denote by X^k the direct product of k copies of X, by $\boldsymbol{X} = X^{\mathbb{Z}_+}$ the space of sequences $(u_k, k \in \mathbb{Z}_+)$ with $u_k \in X$, and by $B_X(a,d)$ the closed ball of radius $d > 0$ centered at $a \in X$. If $a = 0$, we write $B_X(d)$. The distribution of a random variable ξ is denoted by $\mathcal{D}(\xi)$ and the indicator function of a set C by I_C.

$L^p(D)$ and $H^s(D)$ denote the usual Lebesgue and Sobolev spaces in a domain $D \subset \mathbb{R}^n$. We use the same notation for spaces of scalar and vector valued functions. The corresponding norms are denoted by $\|\cdot\|_{L^p}$ and $\|\cdot\|_s$, respectively.

$C_b(X)$ is the space of bounded continuous functions $f : X \to \mathbb{R}$ endowed with the natural norm $\|f\|_\infty = \sup_X |f|$ and $C_+(X)$ is the set of strictly positive functions $f \in C_b(X)$.

$L_b(X)$ stands for the space of functions $f \in C_b(X)$ such that

$$\|f\|_L := \|f\|_\infty + \sup_{0<d_X(u,v)\le 1} \frac{|f(u)-f(v)|}{d_X(u,v)} < \infty.$$

In the case of a compact metric space, we shall drop the subscript b and write $C(X)$ and $L(X)$.

$\mathcal{B}(X)$ denotes the Borel σ-algebra on X, $\mathcal{M}(X)$ the vector space of signed Borel measures on X with finite total mass, $\mathcal{M}_+(X)$ the cone of non-negative measures

$\mu \in \mathcal{M}(X)$, and $\mathcal{P}(X)$ the set of Borel probability measures on X. The vector space $\mathcal{M}(X)$ is endowed with the *total variation norm*

$$\|\mu\|_{\mathrm{var}} := \sup_{\Gamma \in \mathcal{B}(X)} |\mu(\Gamma)| = \frac{1}{2} \sup_{\substack{f \in C_b(X) \\ \|f\|_\infty \le 1}} \left| \int_X f \mathrm{d}\mu \right|.$$

When dealing with $\mathcal{M}_+(X)$, we also use the *Kantorovich–Wasserstein* (also called *dual-Lipschitz*) metric defined by

$$\|\mu_1 - \mu_2\|_L^* := \sup_{\substack{f \in L_b(X) \\ \|f\|_L \le 1}} \left| \int_X f \mathrm{d}\mu_1 - \int_X f \mathrm{d}\mu_2 \right|, \quad \mu_1, \mu_2 \in \mathcal{M}_+(X).$$

The topology defined by the Kantorovich–Wasserstein distance coincides with that of weak convergence. We shall write $\mu_n \rightharpoonup \mu$ to denote the weak convergence of $\{\mu_n\}$ to μ.

For an integrable function $f : X \to \mathbb{R}$ and a measure $\mu \in \mathcal{M}(X)$, we set

$$\langle f, \mu \rangle = \int_X f(u)\, \mu(\mathrm{d}u), \quad \|f\|_\mu = \int_X |f(u)|\, \mu(\mathrm{d}u).$$

Given a function $f : X \to \mathbb{R}$, we denote by f^+ and f^- its positive and negative parts, respectively:

$$f^+ = \frac{1}{2}(|f| + f), \quad f^- = \frac{1}{2}(|f| - f).$$

1 The model and the results

1.1 The model

In this section, we describe a class of discrete-time stochastic systems for which we shall prove the LDP. Let H be a real separable Hilbert space with a scalar product $(\cdot, \cdot)$ and the corresponding norm $\|\cdot\|$ and let $S : H \to H$ be a continuous mapping. We consider the random dynamical system

$$u_k = S(u_{k-1}) + \eta_k, \quad k \ge 1, \tag{1.1}$$

where $\{\eta_k\}$ is a sequence of independent identically distributed (i.i.d.) random variables in H. System (1.1) defines a homogeneous family of Markov chains, and we denote by $P_k(u, \Gamma)$ its transition function and by $\mathfrak{P}_k : C_b(H) \to C_b(H)$ and $\mathfrak{P}_k^* : \mathcal{P}(H) \to \mathcal{P}(H)$ the corresponding Markov operators. We shall assume that S satisfies the following three conditions (which are stronger version of those introduced in [KS00]; see also Section 3.2.1 in [KS12]).

(A) Regularity and stability. *The mapping $S : H \to H$ is continuous. Moreover, for any $R > r > 0$ there are positive constants $C = C(R)$ and $a = a(R, r) < 1$ and an integer $n_0 = n_0(R, r) \geq 1$ such that*

$$\|S(u_1) - S(u_2)\| \leq C(R)\|u_1 - u_2\| \quad \text{for } u_1, u_2 \in B_H(R), \tag{1.2}$$

$$\|S^n(u)\| \leq \max\{a\|u\|, r\} \quad \text{for } u \in B_H(R),\ n \geq n_0, \tag{1.3}$$

where S^n is the nth iteration of S.

Let us denote by $\mathcal{K}$ the support of the law for η_1 and assume that it is a compact subset in H. Given a closed subset $B \subset H$, define the sequence of sets

$$\mathcal{A}_0(B) = B, \quad \mathcal{A}_k(B) = S(\mathcal{A}_{k-1}(B)) + \mathcal{K}, \quad k \geq 1,$$

and denote by $\mathcal{A}(B)$ the closure in H of the union of $\mathcal{A}_k(B)$. We shall call $\mathcal{A}(B)$ the *domain of attainability from B*.

(B) Dissipativity. *There is $\rho > 0$ and a non-decreasing integer-valued function $k_0 = k_0(R)$ such that*

$$\mathcal{A}_k(B_H(R)) \subset B_H(\rho) \quad \text{for } R \geq 0,\ k \geq k_0(R). \tag{1.4}$$

(C) Squeezing. *There is an orthonormal basis $\{e_j\}$ in H such that, for all $R > 0$ and $u_1, u_2 \in B_H(R)$,*

$$\|(I - \mathsf{P}_N)(S(u_1) - S(u_2))\| \leq \gamma_N(R)\|u_1 - u_2\|, \tag{1.5}$$

where $\mathsf{P}_N : H \to H$ denotes the orthogonal projection on the linear span of $e_1, \ldots, e_N$, and $\{\gamma_N(R)\}$ is a decreasing sequence of positive numbers converging to zero as $N \to \infty$.

As for the sequence $\{\eta_k\}$, we assume that it satisfies the following hypothesis:

(D) Structure of the noise. *The random variable η_k has the form*

$$\eta_k = \sum_{j=1}^{\infty} b_j \xi_{jk} e_j, \tag{1.6}$$

where $\{e_j\}$ is the orthonormal basis entering (C)*, $b_j \geq 0$ are constants such that*

$$\mathfrak{B} := \sum_{j=1}^{\infty} b_j^2 < \infty, \tag{1.7}$$

and ξ_{jk} are independent scalar random variables. Moreover, the law of ξ_{jk} is absolutely continuous with respect to the Lebesgue measure, and the corresponding density $p_j(r)$ is a continuously differentiable function such that $p_j(0) > 0$ and $\operatorname{supp} p_j \subset [-1, 1]$.

Recall that a measure $\mu \in \mathcal{P}(H)$ is said to be stationary for (1.1) if $\mathfrak{P}_1^* \mu = \mu$. A proof of the following theorem can be found in Chapter 3 of [KS12].

Theorem 1.1. *Suppose that Conditions* (A)–(D) *are fulfilled and that*[2]

$$b_j \neq 0 \quad \textit{for all } j \geq 1. \tag{1.8}$$

Then there is a unique stationary measure $\mu \in \mathcal{P}(H)$. Moreover, there are constants $C > 0$ and $\alpha > 0$ such that, for any $\lambda \in \mathcal{P}(H)$, we have

$$\|\mathfrak{P}_k^* \lambda - \mu\|_L^* \leq C e^{-\alpha k} \left(1 + \int_H \|u\| \lambda(\mathrm{d}u)\right), \quad k \geq 0. \tag{1.9}$$

We conclude this subsection by a simple remark on the support of the stationary distribution μ. Let us denote by $\mathcal{A} = \mathcal{A}(\{0\})$ the domain of attainability from zero. Since $\mathcal{A}$ is an invariant subset for (1.1), it carries a stationary measure. Since the stationary measure is unique, we must have $\operatorname{supp} \mu \subset \mathcal{A}$. On the other hand, inequality (1.3) and the inclusion $0 \in \operatorname{supp} \mathcal{D}(\eta_1)$ imply that $P_k(u, B_H(r)) > 0$ for any $u \in \mathcal{A}$, $r > 0$, and $k \gg 1$. Combining this fact with the Kolmogorov–Chapman relation, one easily proves that $\operatorname{supp} \mu = \mathcal{A}$.

1.2 The results

Before formulating the main results of this paper, we recall some standard definitions from the theory of large deviations (e.g., see Chapter 6 in [DZ00]). Let X be a Polish space and let $\mathcal{P}(X)$ be the space of probability measures on X endowed with the topology of weak convergence (generated by the Kantorovich–Wasserstein distance). Recall that a *random probability measure* on X is defined as a measurable mapping from a probability space $(\Omega, \mathcal{F}, \mathbb{P})$ to $\mathcal{P}(X)$. A mapping $I : \mathcal{P}(X) \to [0, +\infty]$ is called a *rate function* if it is lower semicontinuous, and a rate function I is said to be *good* if the level set $\{\sigma \in \mathcal{P}(X) : I(\sigma) \leq \alpha\}$ is compact for any $\alpha \in [0, +\infty)$. For a measurable set $\Lambda \subset \mathcal{P}(X)$, we write $I(\Lambda) = \inf_{\sigma \in \Lambda} I(\sigma)$.

Definition 1.2. Let $\{\zeta_k = \zeta_k^\omega, k \geq 1\}$ be a sequence of random probability measures on $\mathcal{A}$. We shall say that $\{\zeta_k\}$ satisfies the LDP with a rate function I if the following two properties are satisfied.

Upper bound. For any closed subset $F \subset \mathcal{P}(X)$, we have

$$\limsup_{k \to \infty} \frac{1}{k} \log \mathbb{P}\{\zeta_k \in F\} \leq -I(F). \tag{1.10}$$

Lower bound. For any open subset $G \subset \mathcal{P}(X)$, we have

$$\liminf_{k \to \infty} \frac{1}{k} \mathbb{P}\{\zeta_k \in G\} \geq -I(G). \tag{1.11}$$

[2] Theorem 1.1 remains valid if finitely many b_j are non-zero. However, the main results of this paper on LDP will be proved under the stronger condition (1.8).

We now consider the family of Markov chains defined by (1.1). To an arbitrary random variable u_0 in H, which we always assume to be independent of $\{\eta_k\}$, one associates a family of *occupation measures* by the formula

$$\zeta_k := \frac{1}{k}\sum_{n=0}^{k-1}\delta_{u_n}, \tag{1.12}$$

where δ_u stands for the Dirac measure concentrated at u. Recall that $\mathcal{A}$ denotes the domain of attainability from zero (see the end of Section 1.1). It is a compact subset of H, invariant under the random dynamics defined by (1.1). Note that if the support of $\mathcal{D}(u_0)$ is contained in $\mathcal{A}$, then ζ_k is also supported by $\mathcal{A}$. The following theorem is the main result of this paper.

Theorem 1.3. *Let Hypotheses* (A)–(D) *and Condition* (1.8) *be fulfilled and let u_0 be an arbitrary random variable in H whose law is supported by $\mathcal{A}$. Then the family $\{\zeta_k, k \geq 1\}$ of random probability measures on $\mathcal{A}$ satisfies the LDP with a good rate function I defined by*

$$I(\sigma) = \sup_{V\in C(\mathcal{A})}\big(\langle V,\sigma\rangle - Q(V)\big), \quad \sigma\in\mathcal{P}(\mathcal{A}), \tag{1.13}$$

where $Q(V)$ is a 1-Lipschitz convex function such that $Q(C) = C$ for any $C \in \mathbb{R}$. Furthermore, the function Q can be written as the limit (1.40)*, which exists for any $V \in C(\mathcal{A})$ and does not depend on the initial point u_0.*

Note that the lower semi-continuity of I is built in its definition, while the fact that I is a good rate function follows from the compactness of $\mathcal{P}(\mathcal{A})$ in the weak topology. Choosing suitable open and closed sets in the LDP for occupation measures, we obtain the asymptotics of the time-averages for various functionals of trajectories of (1.1). For instance, let $f : \mathcal{A} \to \mathbb{R}^m$ be a continuous mapping and let $\Gamma \subset \mathbb{R}^m$ be a Borel set. Define

$$F_\Gamma = \{\sigma\in\mathcal{P}(\mathcal{A}) : \langle f,\sigma\rangle\in\overline{\Gamma}\}, \quad G_\Gamma = \{\sigma\in\mathcal{P}(\mathcal{A}) : \langle f,\sigma\rangle\in\dot{\Gamma}\},$$

where $\overline{\Gamma}$ and $\dot{\Gamma}$ denote the closure and interior of Γ, respectively. In view of the LDP, we have

$$\limsup_{k\to\infty}\frac{1}{k}\log\mathbb{P}\bigg\{\frac{1}{k}\sum_{n=0}^{k-1}f(u_n)\in\Gamma\bigg\} \leq -I(F_\Gamma), \tag{1.14}$$

$$\liminf_{k\to\infty}\frac{1}{k}\log\mathbb{P}\bigg\{\frac{1}{k}\sum_{n=0}^{k-1}f(u_n)\in\Gamma\bigg\} \geq -I(G_\Gamma). \tag{1.15}$$

Theorem 1.3 provides the LDP for the occupation measures (1.12). Some further analysis combined with the Dawson–Gärtner theorem enables one to derive a process level LDP for trajectories of (1.1) issued from $\mathcal{A}$. Namely, denote by $\boldsymbol{H} = H^{\mathbb{Z}_+}$ the direct product of countably many copies of H and, given

a trajectory $\{u_k\}$ for (1.1), define a sequence of random probability measures on $\boldsymbol{H}$ by the relation

$$\boldsymbol{\zeta}_k = \frac{1}{k}\sum_{n=0}^{k-1}\delta_{\boldsymbol{u}_n}, \quad k \geq 1, \tag{1.16}$$

where we set $\boldsymbol{u}_n = (u_k, k \geq n)$.

Theorem 1.4. *Let the hypotheses of Theorem 1.3 be fulfilled and let u_0 be an arbitrary random variable in H whose law is supported by $\mathcal{A}$. Then the family of random probability measures $\{\boldsymbol{\zeta}_k, k \geq 1\}$ satisfies the LDP with a good rate function $\boldsymbol{I} : \mathcal{P}(\boldsymbol{H}) \to [0, +\infty]$, which is equal to $+\infty$ outside a compact subset.*

1.3 Applications

Two-dimensional Navier–Stokes system

Let us consider the Navier–Stokes system (0.1) in which $\eta(t, x)$ is a random kick force of the form (0.3). We assume that the kicks $\{\eta_k\}$ form a sequence of i.i.d. random variables in the space H (see (0.5)). Normalising the solutions of (0.1) to be right continuous and denoting $u_k = u(k, x)$, we see that any solution of (0.1) satisfies relation (1.1) in which S stands for the time-1 shift along trajectories of the homogeneous Navier–Stokes system (e.g., see Section 2.3 in [KS12]). We claim that Theorems 1.3 and 1.4 can be applied to (1.1) with the above choice of S, provided that we restrict our consideration to trajectories lying in $\mathcal{A} = \mathcal{A}(\{0\})$. Indeed, the differentiability of the flow map for the Navier–Stokes system is well known (e.g., see Section 7.5 in [BV92]), and all other properties entering Conditions (A) and (B) are checked in [KS00]. Furthermore, the squeezing property (C) is satisfied for any choice of an orthonormal basis $\{e_j\}$ in H (cf. proof of Proposition 1.6 below). We thus obtain the following result.

Proposition 1.5. *Let the random variables $\{\eta_k\}$ satisfy Condition* (D) *with $b_j \neq 0$ for all $j \geq 1$, let u_0 be an arbitrary H-valued random variable which is independent of $\{\eta_k\}$ and whose law is supported by $\mathcal{A}$, and let $\{u_k\}$ be the corresponding trajectory of* (1.1)*. Then the occupation measures ζ_k and $\boldsymbol{\zeta}_k$ defined by* (1.12) *and* (1.16) *satisfy the LDP with good rate functions.*

In particular, taking for u_0 a random variable distributed according to the stationary measure μ, we obtain Theorems A and B of the Introduction. Furthermore, in view of the discussion after Theorem 1.3, we have an LDP for the time-averages of continuous functionals $f : H \to \mathbb{R}^m$ calculated on trajectories of (1.1). This result is applicable, for instance, to the energy functional $f(u) = \frac{1}{2}\int_D |u(x)|^2 \mathrm{d}x$.

To treat other physically relevant observables, such as the enstrophy or the correlation tensor, we need to change the phase space of the problem, making it more regular. More precisely, let us define the space

$$U = H \cap H_0^1(D) \cap H^2(D)$$

(where $H^s(D)$ denotes the usual Sobolev space of order s) and endow it with the usual scalar product in H^2. Since the flow-map for the Navier–Stokes system preserves the H^2-regularity, system (1.1) can be studied in the space U, provided that the random kicks also belong to U. We have the following result.

Proposition 1.6. *Let $\{\eta_k\}$ be a sequence of i.i.d. random variables in U of the form (1.6), in which $\{e_j\}$ is an orthonormal basis in U, and $\{b_j\}$ and $\{\xi_{jk}\}$ are the same as in Condition* (D)*. Assume that $b_j \neq 0$ for all $j \geq 1$. Then the LDP holds for the occupation measures ζ_k and $\boldsymbol{\zeta}_k$ of any trajectory whose initial state u_0 is a U-valued random variable with range in $\mathcal{A}$.*

For instance, one can take for an initial state any function $u_0 \in U$ or a random variable u_0 distributed according to the stationary measure. Furthermore, relations (1.14) and (1.15) hold for the functional $f : U \to \mathbb{R}^6$ defined by

$$f(u) = \left(\frac{1}{2}\int_D |u(x)|^2 \mathrm{d}x, \frac{1}{2}\int_D |(\nabla \otimes u)(x)|^2 \mathrm{d}x, u^i(x_1)u^j(x_2), 1 \leq i, j \leq 2\right),$$

where $u = (u^1, u^2)$, and $x_1, x_2 \in D$ are given points.

Proof of Proposition 1.6. We shall check that Conditions (A)–(D) of Section 1.1 (with H replaced by U) are fulfilled. The validity of (D) follows from the hypotheses of the proposition. The facts that, for any $T > 0$, the time-T shift along trajectories is uniformly Lipschitz continuous on bounded subsets of U and is continuously differentiable in the Fréchet sense are proved in Chapter 7 of [BV92]. Let us prove (1.3). It is well known that (see the proof of Theorem 6.2 of Chapter 1 in [BV92])

$$\|S(u)\| \leq q\,\|u\|, \quad \|S(u)\|_2 \leq C\,\|u\|,$$

where $q < 1$ and $C > 0$ are some constants, $u \in H$ in the first inequality, and $u \in B_H(1)$ in the second. Combining these two estimates we see that for any $R > 0$ we can find $n_1 = n_1(R) \geq 1$ such that

$$\|S^{n+1}(u)\|_2 \leq Cq^n\|u\| \quad \text{for } u \in B_H(R),\, n \geq n_1.$$

This inequality immediately implies (1.3).

We now establish the dissipativity property (B). We know that this property holds in the space H. Thus, we can find $\rho_1 > 0$ such that, for any $R > 0$ and a sufficiently large integer $k_1(R) \geq 1$, we have

$$\mathcal{A}_k(B_U(R)) \subset B_H(\rho_1) \quad \text{for } k \geq k_1(R).$$

Since the mapping S is continuous from H to U, it follows that

$$\mathcal{A}_{k+1}(B_U(R)) \subset S(B_H(\rho_1)) + \mathcal{K} \quad \text{for } k \geq k_1(R).$$

Choosing $\rho > 0$ such that $S(B_H(\rho_1)) + \mathcal{K} \subset B_U(\rho)$, we obtain (1.4) with $H = U$ and $k_0(R) = k_1(R) + 1$.

It remains to prove the squeezing property (C). Let $u_i(t)$, $i=1,2$, be two solutions of the homogeneous Navier–Stokes system, which we write as a non-local PDE in the space H:

$$\partial_t u + \nu L u + B(u) = 0. \tag{1.17}$$

Here $L = -\Pi\Delta$, $B(u) = B(u,u)$, $B(u,v) = \Pi(\langle u, \nabla\rangle v)$, and Π is the orthogonal projection in $L^2(D,\mathbb{R}^2)$ onto H. We wish to show that, if the initial conditions $u_{i0} = u_i(0)$ belong to the ball $B_U(R)$ and $\{e_j\}$ is an orthonormal basis in U, then

$$\big\|(I-\mathsf{P}_N)\big(u_1(1)-u_2(1)\big)\big\|_2 \le \gamma_N(R)\|u_{10}-u_{20}\|_2,$$

where $\gamma_N(R)$ depends only on the basis $\{e_j\}$ and goes to zero as $N\to\infty$. A simple compactness argument implies that this inequality will hold if we prove that

$$\|u_1(1)-u_2(1)\|_3 \le C(R)\|u_{10}-u_{20}\|_2 \quad \text{for } u_{10}, u_{20} \in B_U(R). \tag{1.18}$$

The proof of this inequality is carried out by standard methods (e.g., see [BV92]), and we confine ourselves to outlining the main steps. We shall denote by $C_i(R)$ unessential positive constants depending only on R.

Step 1. It suffices to prove that

$$\|u_1(1)-u_2(1)\|_2 \le C_1(R)\,\|u_{10}-u_{20}\|_2, \tag{1.19}$$

$$\|\dot u_1(1)-\dot u_2(1)\|_1 \le C_2(R)\,\|u_{10}-u_{20}\|_2 \tag{1.20}$$

for $u_{10}, u_{20} \in B_U(R)$, where $\dot v = \partial_t v$. Set $u = u_1 - u_2$ and note that (1.17) implies

$$\nu L u(1) = -\dot u(1) - B\big(u_1(1), u(1)\big) - B\big(u(1), u_2(1)\big).$$

Since $u_i(1)$, $i=1,2$ are bounded in H^2 (see Theorem 6.2 in [BV92]) and the bilinear mapping $(u,v)\mapsto B(u,v)$ is continuous from U to H^1, we see that

$$\nu\,\|Lu(1)\|_1 \le \|\dot u(1)\|_1 + C_3(R)\,\|u(1)\|_2.$$

Recalling that L^{-1} is continuous from $H^1_0\cap H$ to H^3 and using inequalities (1.19) and (1.20), we obtain the required estimate (1.18).

Step 2. To prove (1.19) and (1.20), we first show that

$$\sup_{0\le t\le 1}\left(\|\dot u(t)\|^2 + \int_0^t \|\dot u(s)\|_1^2 \mathrm{d}s\right) \le C_4(R)\,\|u_{10}-u_{20}\|_2. \tag{1.21}$$

Differentiating (1.17) in time, we derive the following equation for $\dot u = \dot u_1 - \dot u_2$:

$$\partial_t \dot u + \nu L\dot u + B(\dot u_1, u) + B(u_1, \dot u) + B(\dot u, u_2) + B(u, \dot u_2) = 0. \tag{1.22}$$

Taking the L^2-scalar product with $2\dot u$ and performing some standard transformations, we obtain

$$\partial_t\|\dot u\|^2 + \nu\|\dot u\|_1^2 \le C_5\|u_2\|_1^2\|\dot u\|^2 + C_5\|u\|_1\|u\|\,\big(\|\dot u_1\|_1\|\dot u_1\| + \|\dot u_2\|_1\|\dot u_2\|\big).$$

Applying the Gronwall and Cauchy–Schwarz inequalities and using the fact that $\dot{u}_i$ belong to a bounded set in $L^\infty(0,1;H)\cap L^2(0,1;H^1)$, we derive

$$\|\dot{u}(t)\|^2+\nu\int_0^t \|\dot{u}(s)\|_1^2\mathrm{d}s \le C_6(R)\Big(\|\dot{u}(0)\|^2+\|u\|_{L^\infty(J_t;H)}\|u\|_{L^2(J_t;H^1)}\Big), \quad (1.23)$$

where $J_t=(0,t)$. It follows from (1.17) that (cf. Step 1)

$$\|\dot{u}(0)\| \le C_7(R)\,\|u(0)\|_2.$$

Furthermore, it is well known that

$$\sup_{0\le t\le 1}\left(\|u(t)\|^2+t\,\|u(t)\|_1^2+\int_0^t \|u(s)\|_1^2\,\mathrm{d}s\right) \le C_8(R)\,\|u(0)\|^2. \quad (1.24)$$

Combining these two inequalities with (1.23), we obtain (1.21).

Step 3. We now derive (1.19). To this end, note that

$$\nu Lu(1)=g:=-\partial_t u(1)-B(u_1(1),u(1))-B(u(1),u_2(1)). \quad (1.25)$$

Using the continuity properties of B and inequalities (1.21) and (1.24) one easily obtains

$$\|g\| \le \|\partial_t u(1)\|+C_9(R)\,\|u(1)\|_1 \le C_{10}(R)\,\|u(0)\|_2.$$

Combining this with (1.25) and the continuity of L^{-1} from H to H^2, we obtain (1.19).

Step 4. It remains to prove (1.20). To this end, we take the L^2-scalar product of (1.22) with $2tL\dot{u}$. After some standard transformations, we derive

$$\partial_t\big(t\|\dot{u}\|_1^2\big)-\|\dot{u}\|_1^2+2\nu t\,\|\dot{u}\|_2^2=q(t), \quad (1.26)$$

where we set $q(t)=-2t\big(B(\dot{u}_1,u)+B(u_1,\dot{u})+B(\dot{u},u_2)+B(u,\dot{u}_2),L\dot{u}\big)$. Well-known estimates for the bilinear term B imply that

$$|q(t)| \le \nu t\,\|\dot{u}\|_2^2+C_{11}t\,q_1(t), \quad (1.27)$$

where

$$q_1(t)=\|\dot{u}_1\|\,\|\dot{u}_1\|_1\|u\|_1\|u\|_2+\|u_1\|^2\|u_1\|_2^2\|\dot{u}\|^2 \\ +\|u_2\|_1\|u_2\|_2\|\dot{u}\|\,\|\dot{u}\|_1+\|\dot{u}_2\|_1^2\|u\|\,\|u\|_2.$$

Integrating (1.26) in time and using (1.27), we obtain

$$t\,\|\dot{u}\|_1^2+\nu\int_0^t s\,\|\dot{u}\|_2^2\,\mathrm{d}s \le \int_0^t \|\dot{u}\|_1^2\,\mathrm{d}s+C_{11}\int_0^t sq_1(s)\,\mathrm{d}s. \quad (1.28)$$

The first integral on the right-hand side can be estimated with the help of (1.21). We now bound each term of the second integral:

$$\int_0^t s\,\|\dot u_1\|\,\|\dot u_1\|_1\|u\|_1\|u\|_2 \mathrm{d}s \le C_{12}(R)\left(\int_0^t \|u(s)\|_2^2\,\mathrm{d}s\right)^{1/2} \sup_{0\le s\le t}\|u(s)\|_1,$$

$$\int_0^t s\,\|u_1\|^2\|u_1\|_2^2\|\dot u\|^2 \mathrm{d}s \le C_{13}(R)\left(\int_0^t \|\dot u(s)\|^2\,\mathrm{d}s\right)^{1/2},$$

$$\int_0^t s\,\|u_2\|_1\|u_2\|_2\|\dot u\|\,\|\dot u\|_1 \mathrm{d}s \le C_{14}(R)\left(\int_0^t \|\dot u(s)\|_1^2\,\mathrm{d}s\right)^{1/2} \sup_{0\le s\le t}\|\dot u(s)\|,$$

$$\int_0^t s\,\|\dot u_2\|_1^2\|u\|\,\|u\|_2 \mathrm{d}s \le C_{15}(R)\left(\int_0^t \|u(s)\|_2^2\,\mathrm{d}s\right)^{1/2} \sup_{0\le s\le t}\|u(s)\|,$$

where $0 \le t \le 1$, and we used the fact that the functions u_i and $\dot u_i$ belong to bounded sets in the spaces $L^\infty(J_1, H^2)$ and $L^\infty(J_1, L^2) \cap L^2(J_1, H^1)$, respectively. On the other hand, it is well known that

$$\sup_{0\le t\le 1}\left(\|u(t)\|_1^2 + \int_0^t \|u(s)\|_2^2\,\mathrm{d}s\right) \le C_{17}(R)\,\|u(0)\|_1^2.$$

Combining these estimates with (1.28), (1.24), and (1.21), we obtain (1.20). This completes the proof of Proposition 1.6. □

Complex Ginzburg–Landau equation

Let us consider a complex Ginzburg–Landau (CGL) equation perturbed by a random kick force:

$$\partial_t u - (\nu + i)\Delta u + ia|u|^2 u = \eta(t,x), \quad x \in D, \quad u\big|_{\partial D} = 0. \tag{1.29}$$

Here $a > 0$ is a parameter, $D \subset \mathbb{R}^3$ is a bounded domain with C^2-smooth boundary ∂D, $u = u(t,x)$ is a complex-valued unknown function, and η is an external force of the form (0.3), where $\{\eta_k\}$ is a sequence of i.i.d. random variables in the complex space $H_0^1(D)$. It is well known that the Cauchy problem for (1.29) is well posed in $H_0^1(D)$ (see [KS04]), that is, for any $u_0 \in H_0^1(D)$, problem (1.29) has a unique solution such that

$$u(0,x) = u_0(x). \tag{1.30}$$

Let us assume that the random kicks entering (1.29) have the form

$$\eta_k(x) = \sum_{j=1}^{\infty} b_j(\xi_{jk}^1 + i\xi_{jk}^2)e_j,$$

where $\{e_j\}$ is an orthonormal basis in $H_0^1(D)$ consisting of the real eigenfunctions of the Dirichlet Laplacian, $\{b_j\} \subset \mathbb{R}$ is a sequence satisfying (1.7), and ξ_{jk}^i are

independent real-valued random variables whose laws possess the properties stated in Condition (D) of Section 1.1. We denote by $\mathcal{A} \subset H_0^1(D)$ the set of attainability from zero. The following result is an analogue of Propositions 1.5 and 1.6 for the case of the CGL equation.

Proposition 1.7. *In addition to the above hypotheses, assume that $b_j \neq 0$ for all $j \geq 1$. Then the LDP holds for the occupation measures (1.12) and (1.16) of the trajectories whose initial state is an H_0^1-valued random variable with range in $\mathcal{A}$.*

Proof. We endow the space $H_0^1 = H_0^1(D)$ with the scalar product

$$(u_1, u_2)_1 = \mathrm{Re} \int_D \nabla u_1(x) \cdot \overline{\nabla u_2(x)} \, \mathrm{d}x$$

and regard it as a real Hilbert space. Let $S : H_0^1 \to H_0^1$ be the time-1 shift along trajectories of the problem (1.29) with $\eta \equiv 0$. The required results will be established if we check that the stochastic system (1.1) considered in the space $H = H_0^1$ possesses properties (A)–(D). Regularity of the mapping S and its Lipschitz continuity on bounded subsets are standard, and (D) is satisfied in view of the hypotheses of the proposition. Thus, it remains to check (1.3)–(1.5).

Step 1. Let us introduce the following continuous functionals on H_0^1:

$$\mathcal{H}_0(u) = \frac{1}{2} \int_D |u(x)|^2 \mathrm{d}x, \quad \mathcal{H}_1(u) = \int_D \Big(\frac{1}{2} |\nabla u(x)|^2 + \frac{a}{4} |u(x)|^4 \Big) \, \mathrm{d}x.$$

It is well known [3] that if $u(t)$ is a solution of (1.29) and η is a locally integrable function of time with range in H_0^1, then

$$\frac{\mathrm{d}}{\mathrm{d}t} \mathcal{H}_0(u) = -\nu \, \|\nabla u\|^2 + (u, \eta), \tag{1.31}$$

$$\frac{\mathrm{d}}{\mathrm{d}t} \mathcal{H}_1(u) = -\nu \, \|\Delta u\|^2 - 2a\nu(|u|^2, |\nabla u|^2) + a\nu(u^2, (\nabla u)^2) + (-\Delta u + a|u|^2 u, \eta), \tag{1.32}$$

where $(\cdot, \cdot)$ and $\| \cdot \|$ stand for the real L^2-scalar product and the corresponding norm:

$$(u_1, u_2) = \mathrm{Re} \int_D u_1(x) \, \bar{u}_2(x) \, \mathrm{d}x, \quad \|u\|^2 = (u, u).$$

Using the inequalities

$$\|\Delta u\|^2 \geq \alpha_1 \|\nabla u\|^2, \quad |(u^2, (\nabla u)^2)| \leq (|u|^2, |\nabla u|^2), \quad (|u|^2, |\nabla u|^2) \geq c \int_D |u|^4 \mathrm{d}x,$$

where $\alpha_1 > 0$ is the first eigenvalue of the Dirichlet Laplacian, we derive from (1.32) the inequality

$$\frac{\mathrm{d}}{\mathrm{d}t} \mathcal{H}_1(u(t)) \leq -\delta \mathcal{H}_1(u(t)) - \frac{\nu}{2} \|\Delta u\|^2 + (-\Delta u + a|u|^2 u, \eta), \tag{1.33}$$

[3] For instance, see Section 2.2 in [KS04] for the more complicated case of a white noise force.

where $\delta > 0$ depends only on a, ν, and α_1. Taking $\eta \equiv 0$ and applying the Gronwall inequality, we see that

$$\mathcal{H}_1(u(t)) \le e^{-\delta t}\mathcal{H}_1(u(0)), \quad t \ge 0.$$

It follows that if $u_0 \in B_{H_0^1}(R)$, then

$$\|S^n(u_0)\|_1 \le \big(2\mathcal{H}_1(u(n))\big)^{1/2} \le \big(2e^{-\delta n}\mathcal{H}_1(u_0)\big)^{1/2} \le C_1 R\, e^{-\delta n/2}\|u_0\|_1,$$

where we used the inequality

$$\mathcal{H}_1(v) \le C\|v\|_1^4, \quad v \in H_0^1, \tag{1.34}$$

following immediately from the continuity of the embedding of $H_0^1 \subset L^4$. The above estimate for $\|S^n(u_0)\|_1$ implies (1.3).

Step 2. We now prove the dissipativity property (1.4). To this end, we first establish a bound for the L^2-norm. It follows from (1.31) that, for any $\varepsilon > 0$, the function $\varphi_\varepsilon(t) = (\mathcal{H}_0(u(t)) + \varepsilon)^{1/2}$ satisfies the inequality

$$\varphi_\varepsilon'(t) \le -\nu\alpha_1\varphi_\varepsilon(t) + \tfrac{1}{\sqrt{2}}\|\eta(t)\| + \nu\alpha_1\sqrt{\varepsilon}.$$

Applying the Gronwall inequality and passing to the limit $\varepsilon \to 0$, we obtain

$$\|u(t)\| \le e^{-\nu\alpha_1 t}\|u_0\| + \int_0^t e^{-\nu\alpha_1(t-s)}\|\eta(s)\|\,\mathrm{d}s. \tag{1.35}$$

Now note that if $\int_J \|\eta(s)\|\mathrm{d}s \le b_0$ for any interval $J \subset \mathbb{R}_+$ of length 1, then

$$\int_0^t e^{-\nu\alpha_1(t-s)}\|\eta(s)\|\,\mathrm{d}s \le \frac{b_0}{1-e^{-\nu\alpha_1}} \quad \text{for } t \ge 0.$$

Combining this with (1.35), we see that

$$\|u(k)\| \le e^{-\nu\alpha_1 k}\|u_0\| + \frac{b_0}{1-e^{-\nu\alpha_1}}, \quad k \ge 0. \tag{1.36}$$

This inequality, established in the case of locally time-integrable functions $\eta(t)$, remains true for kick forces of the form (0.3) with L^2 bounded functions η_k. In particular, (1.4) holds with $H = L^2(D)$.

We now use the regularising property of the homogeneous CGL equation to prove (1.4) with $H = H_0^1$. Namely, if we show that the mapping $S : u_0 \mapsto u(1)$ from L^2 to H_0^1 is bounded on bounded subsets, then (1.36) will obviously imply the existence of a universal constant $\rho > 0$ satisfying (1.4) with $H = H_0^1$. To prove the boundedness of S, let us fix a solution u of (1.29), define a function $\psi(t) = t\sqrt{\mathcal{H}_1(u(t))}$, and calculate its derivative. It follows from (1.32) and (1.34) that[4]

$$\psi'(t) = \sqrt{\mathcal{H}_1} + \frac{t}{\sqrt{\mathcal{H}_1}}\frac{\mathrm{d}}{\mathrm{d}t}\mathcal{H}_1 \le \sqrt{C}\,\|u(t)\|_1^2.$$

[4] A rigorous derivation of (1.37) can be carried out by the simple argument used to establish (1.35).

Furthermore, integrating relation (1.31) with $\eta \equiv 0$, we see that

$$\|u(t)\|^2 + 2\nu \int_0^t \|u(s)\|_1^2 \mathrm{d}s \le \|u_0\|^2.$$

Combining these two relations, we obtain

$$\|u(t)\|_1 \le \sqrt{2\mathcal{H}_1(u(t))} = \frac{\sqrt{2}}{t}\psi(t) \le \frac{\sqrt{2C}}{t}\int_0^t \|u\|_1^2 \mathrm{d}s \le \frac{C_2}{\nu t}\|u_0\|^2. \tag{1.37}$$

The boundedness of S from L^2 to H_0^1 is a straightforward consequence of this inequality.

Step 3. It remains to establish the squeezing property (1.5), in which P_N is the orthogonal projection in H_0^1 (endowed with the scalar product $(\cdot,\cdot)_1$) to the vector span of $\{e_j, ie_j, j = 1, \dots, N\}$. Let us set $\mathsf{Q}_N = I - \mathsf{P}_N$. By hypothesis, we have $\|e_j\|_1 = 1$, hence it follows that

$$\|e_j\| = \frac{1}{\sqrt{\alpha_j}}, \quad j \ge 1,$$

where α_j is the eigenvalue of the Dirichlet Laplacian corresponding to the eigenfunction e_j. Using this relation, it is straightforward to check that $\{\sqrt{\alpha_j}\, e_j\}$ is an orthonormal basis in L^2 and that the norm of Q_N regarded as an operator in L^2 is equal to 1.

Now let u_1, u_2 be two solutions of (1.29) corresponding to initial data $u_{10}, u_{20} \in B_{H_0^1}(R)$. Applying Q_N to (1.29) and setting $w = \mathsf{Q}_N(u_1 - u_2)$, we derive the equation

$$\dot{w} - (\nu + i)\Delta w + ia\mathsf{Q}_N\big(|u_1|^2 u_1 - |u_2|^2 u_2\big) = 0.$$

It follows that

$$\begin{aligned}\partial_t \|w\|_1^2 &= 2\,\mathrm{Re}\int_D \nabla\dot{w}\cdot\nabla\bar{w}\,\mathrm{d}x = -2\,\mathrm{Re}\int_D \dot{w}\,\Delta\bar{w}\,\mathrm{d}x \\ &= -2\big((\nu+i)\Delta w - ia\mathsf{Q}_N(|u_1|^2u_1 - |u_2|^2u_2), \Delta\bar{w}\big) \\ &\le -2\nu\,\|\Delta w\|^2 + a\,\big\||u_1|^2u_1 - |u_2|^2u_2\big\|\,\|\Delta w\|,\end{aligned} \tag{1.38}$$

where we used the fact that the norm of Q_N is equal to 1. Using the Hölder inequality and the continuity of the embedding $H_0^1 \subset L^6$, we derive

$$\big\||u_1|^2u_1 - |u_2|^2u_2\big\| \le C_3\big(\|u_1\|_1 + \|u_2\|_1\big)^2\,\|w\|_1.$$

Substituting this into (1.38) and using the Poincaré inequality $\|\Delta w\|^2 \ge \alpha_N\|w\|_1^2$, we obtain

$$\partial_t\|w\|_1^2 \le -\big(\nu\alpha_N - C_4(\|u_1\|_1 + \|u_2\|_1)^4\big)\,\|w\|_1^2. \tag{1.39}$$

Since $u_{10}, u_{20} \in B_{H_0^1}(R)$ and the resolving operator for the CGL is bounded on bounded subsets, we can find $C_5(R)$ such that

$$\|u_i(t)\|_1 \le C_5(R) \quad \text{for } 0 \le t \le 1,\ i = 1, 2.$$

Combining this with (1.39) and the Gronwall inequality, we derive

$$\begin{aligned}\|w(t)\|_1^2 &\le \exp\biggl(-\nu\alpha_N t + C_4\int_0^t (\|u_1\|_1+\|u_2\|_1)^4 \mathrm{d}s\biggr)\,\|w(0)\|_1^2\\ &\le \exp\bigl(-\nu\alpha_N t + C_6(R)\,t\bigr)\,\|w(0)\|_1^2.\end{aligned}$$

Since Q_N is an orthogonal projection in H_0^1, we have $\|w(0)\|_1 \le \|u_{10}-u_{20}\|_1$. Substituting this into the above estimate and taking $t=1$, we obtain

$$\|u_1(1)-u_2(1)\|_1 \le \gamma_N(R)\,\|u_{10}-u_{20}\|_1, \quad \gamma_N^2(R) = \exp\bigl(-\nu\alpha_N + C_6(R)\bigr).$$

This completes the proof of (1.5) and Proposition 1.7 follows. □

1.4 Scheme of the proof of Theorem 1.3

Along with ζ_k, let us consider "shifted" occupation measures defined as

$$\hat\zeta_k = \frac{1}{k}\sum_{n=1}^k \delta_{u_n}.$$

The sequences $\{\zeta_k\}$ and $\{\hat\zeta_k\}$ are exponentially equivalent (see Lemma 6.2), and therefore, by Theorem 4.2.13 of [DZ00], it suffices to prove the LDP for $\hat\zeta_k$. The proof of this property is based on an abstract result established by Kifer [Kif90]. For the reader's convenience, its statement is recalled in the Appendix (see Theorem 6.1). We shall prove that the following two properties hold.

Property 1: The existence of a limit. For any $V \in C(\mathcal{A})$, the limit

$$Q(V) = \lim_{k\to+\infty}\frac{1}{k}\log\mathbb{E}\,\exp\biggl(\sum_{n=1}^k V(u_n)\biggr). \tag{1.40}$$

exists and does not depend on the initial condition u_0.

The function $Q(V)$ is convex and 1-Lipschitz, and we denote by $I:\mathcal{M}(\mathcal{A})\to\mathbb{R}_+$ its Legendre transform; see (6.2). It is well known that

$$Q(V) = \sup_{\sigma\in\mathcal{P}(\mathcal{A})}\bigl(\langle V,\sigma\rangle - I(\sigma)\bigr);$$

see Lemma 2.2 in [BD99] and Theorem 2.2.15 in [DS89]. In view of the compactness of $\mathcal{P}(\mathcal{A})$, for any $V\in C(\mathcal{A})$ the supremum in the above relation is attained at some point $\sigma_V \in \mathcal{P}(\mathcal{A})$. Any such point is called an *equilibrium state*.

Property 2: Uniqueness of the equilibrium state. There is a dense vector space $\mathcal{V}\subset C(\mathcal{A})$ such that, for any $V\in\mathcal{V}$, there exists unique σ_V satisfying:

$$Q(V) = \langle V,\sigma_V\rangle - I(\sigma_V). \tag{1.41}$$

According to Kifer's theorem, the first of the above properties implies the LD upper bound for $\hat{\zeta}_k$, while the second is sufficient for the LD lower bound. The proofs of these two properties are related to the large-time behaviour of a generalised Markov semigroup associated with u_k. More precisely, given a function $V \in C(\mathcal{A})$, we consider the semigroup

$$\mathfrak{P}_k^V f(u) := \mathbb{E}_u f(u_k) \exp\bigg(\sum_{n=1}^{k} V(u_n)\bigg), \quad f \in C(\mathcal{A}), \tag{1.42}$$

where the subscript u means that we consider the trajectory of (1.1) starting from $u \in H$. The dual semigroup is denoted by $\mathfrak{P}_k^{V*} : \mathcal{P}(\mathcal{A}) \to \mathcal{P}(\mathcal{A})$. We construct explicitly a dense vector space $\mathcal{V} \subset C(\mathcal{A})$ such that, for any $V \in \mathcal{V}$, the semigroup $\mathfrak{P}_k^V$ is uniformly Feller and uniformly irreducible (see Section 2 for the definition of these concepts). Then, by an abstract result proved in Section 2, there is a number $\lambda_V > 0$, a function $h_V \in C_+(\mathcal{A})$, and a measure $\mu_V \in \mathcal{P}(\mathcal{A})$ satisfying

$$\mathfrak{P}_1^V h_V = \lambda_V h_V, \quad \mathfrak{P}_1^{V*} \mu_V = \lambda_V \mu_V, \tag{1.43}$$

such that for any $f \in C(\mathcal{A})$ and $\nu \in \mathcal{P}(\mathcal{A})$ we have

$$\lambda_V^{-k} \mathfrak{P}_k^V f \to \langle f, \mu_V \rangle h_V \quad \text{in } C(\mathcal{A}) \text{ as } k \to +\infty, \tag{1.44}$$

$$\lambda_V^{-k} \mathfrak{P}_k^{V*} \nu \rightharpoonup \langle h_V, \nu \rangle \mu_V \quad \text{in } \mathcal{M}_+(\mathcal{A}) \text{ as } k \to +\infty. \tag{1.45}$$

Taking $f = 1$ in (1.44), one gets immediately the existence of the limit (1.40) for $V \in \mathcal{V}$ and any initial function u_0 whose law is supported by $\mathcal{A}$. Then, by a simple approximation argument, we prove the existence of the limit for any $V \in C(\mathcal{A})$.

To establish the uniqueness of $\sigma_V \in \mathcal{P}(\mathcal{A})$ satisfying (1.41), we first show that any equilibrium state σ_V is a stationary measure for the dual of the following Markov semigroup:

$$\mathscr{S}_k^V g := \lambda_V^{-k} h_V^{-1} \mathfrak{P}_k^V (g h_V), \quad g \in C(\mathcal{A}). \tag{1.46}$$

We then deduce the uniqueness of stationary measure for $\mathscr{S}_k^V$ from convergence (1.45), showing that $\sigma_V(\mathrm{d}u) = h_V(u)\mu_V(\mathrm{d}u)$.

The crucial point in the realisation of the above scheme is the verification of the uniform Feller property for the semigroup $\{\mathfrak{P}_k^V\}$. This verification is based on a coupling argument and is carried out in Section 3.

1.5 Scheme of the proof of Theorem 1.4

Let $p_m : \boldsymbol{H} \to H^m$ be the projection that maps a sequence $(u_j, j \in \mathbb{Z}_+)$ to the vector $(u_j, 0 \le j \le m-1)$. It is straightforward to check that if $\{u_k, k \ge 0\}$ is a trajectory for (1.1), then the image of $\boldsymbol{\zeta}_k$ (see (1.16)) under p_m coincides with the random probability measure

$$\zeta_k^m = \frac{1}{k} \sum_{n=0}^{k-1} \delta_{\boldsymbol{u}_n^m}, \quad k \ge 1, \tag{1.47}$$

where $\boldsymbol{u}_n^m = (u_n, \dots, u_{n+m-1})$. It follows from the Dawson–Gärtner theorem (see Theorem 6.3 in the Appendix) that to prove Theorem 1.4 it suffices to show that for any integer $m \geq 1$, the LDP holds for ζ_k^m with a good rate function $I_m : \mathcal{P}(H^m) \to [0, +\infty]$. The proof of this fact is very similar to the proof of Theorem 1.3 and the argument is outlined in Section 5. To formulate the result precisely, let $\mathcal{A}^{(m)}$ be the set of vectors $(u_1, \dots, u_m) \in H^m$ such that $u_1 \in \mathcal{A}$ and $u_k = S(u_{k-1}) + \eta_k$ for $2 \leq k \leq m$, where $\eta_k \in \mathcal{K}$. Note that if a trajectory $\{u_k\}$ for (1.1) is such that u_0 is an $\mathcal{A}$-valued random variable, then the measures ζ_k^m are concentrated on $\mathcal{A}^{(m)}$. In Section 5 we prove

Theorem 1.8. *Under the conditions of Theorem 1.3, let u_0 be a random variable in H whose law is supported by $\mathcal{A}$. Then the family $\{\zeta_k^m, k \geq 1\}$ regarded as a sequence of random probability measures on $\mathcal{A}^{(m)}$ satisfies the LDP with a good rate function $I_m : \mathcal{P}(\mathcal{A}^{(m)}) \to [0, +\infty]$. Moreover, I_m can be written as*

$$I_m(\sigma) = \sup_{V \in C(\mathcal{A}^{(m)})} \big(\langle V, \sigma \rangle - Q_m(V)\big), \quad \sigma \in \mathcal{P}(\mathcal{A}^{(m)}), \tag{1.48}$$

where $Q_m : C(\mathcal{A}^{(m)}) \to \mathbb{R}$ is a 1-Lipschitz convex function such that $Q_m(C) = C$ for any $C \in \mathbb{R}$.

This result immediately implies that ζ_k^m, as measures on H^m, satisfy the LDP. To see this, extend the rate function I_m constructed in Theorem 1.8 to the space $\mathcal{P}(H^m)$ by setting $I_m(\sigma) = +\infty$ for any measure $\sigma \in \mathcal{P}(H^m)$ satisfying $\sigma(\mathcal{A}^{(m)}) < 1$. Then, recalling that ζ_k^m are supported on $\mathcal{A}^{(m)}$ if so is the the initial measure $\mathcal{D}(u_0)$, we readily check that the LD upper and lower bounds hold for the family $\{\zeta_k^m, k \geq 1\}$ regarded as random probability measures on H^m.

1.6 Uniform large deviations principle

The arguments of the proofs of Theorems 1.3 and 1.4 enable one to obtain a *uniform LDP* for the families $\{\zeta_k\}$ and $\{\boldsymbol{\zeta}_k\}$, which depend on the initial point. More precisely, let us denote by $\zeta_k(u)$ the occupation measure (1.12) for the trajectory issued from a deterministic point $u \in \mathcal{A}$ and define $\boldsymbol{\zeta}_k(u)$ in a similar way. The definition of the uniform LDP is recalled in the Appendix (see Section 6.3). We have the following result.

Theorem 1.9. *Let Hypotheses* (A)–(D) *and Condition* (1.8) *be satisfied. Then the uniform LDP holds for the families $\{\zeta_k(u), u \in \mathcal{A}\}$ and $\{\boldsymbol{\zeta}_k(u), u \in \mathcal{A}\}$ with the good rate functions I and $\boldsymbol{I}$ defined in Theorems 1.3 and 1.4, respectively.*

Sketch of the proof. Let us define the set $\Theta := \mathbb{N} \times \mathcal{A}$ and introduce an order relation $\prec$ on it by the following rule: if $\theta_i = (k_i, u^i) \in \Theta$ for $i = 1, 2$, then $\theta_1 \prec \theta_2$ if and only if $k_1 \leq k_2$. Then $(\Theta, \prec)$ is a directed set. Defining $r(\theta) = k$, we apply Theorem 6.1 to the family $\zeta_\theta = \zeta_k(u)$ indexed by $\theta = (k, u) \in \Theta$. The scheme of the proof described above for Theorem 1.3 applies equally well in this case, and using the fact that the convergence in (1.40) is uniform with respect

to the deterministic initial condition $u_0 \in \mathcal{A}$, we get the existence of limit (6.1) and uniqueness of equilibrium measure. Thus, we have the LDP

$$-I(\dot{\Gamma}) \leq \liminf_{\theta\in\Theta} \frac{1}{k} \log \mathbb{P}\{\zeta_\theta \in \Gamma\} \leq \limsup_{\theta\in\Theta} \frac{1}{k} \log \mathbb{P}\{\zeta_\theta \in \Gamma\} \leq -I(\overline{\Gamma}). \quad (1.49)$$

Now notice that the middle terms in this inequality can be written as

$$\liminf_{\theta\in\Theta} \frac{1}{k} \log \mathbb{P}\{\zeta_\theta \in \Gamma\} = \liminf_{k\to+\infty} \frac{1}{k} \inf_{u\in\mathcal{A}} \log \mathbb{P}\{\zeta_k(u) \in \Gamma\},$$
$$\limsup_{\theta\in\Theta} \frac{1}{k} \log \mathbb{P}\{\zeta_\theta \in \Gamma\} = \limsup_{k\to+\infty} \frac{1}{k} \sup_{u\in\mathcal{A}} \log \mathbb{P}\{\zeta_k(u) \in \Gamma\}.$$

Substituting these relations into (1.49), we obtain the uniform LDP for $\zeta_k(u)$.

To establish the uniform LDP for $\boldsymbol{\zeta}_k(u)$, we apply Theorem 6.4. We thus need the uniform LDP for the projected measures $\zeta_k^m = \zeta_k^m(u)$ defined in Section 1.5. The latter can be obtained by modifying the proof of Theorem 1.8 exactly in the same way as we did above to get the uniform LDP for $\zeta_k(u)$. □

2 Large-time asymptotics for generalised Markov semigroups

In this section, we prove a general result on the large-time behaviour of trajectories for a class of dual semigroups. This type of results were established earlier for Markov semigroups satisfying a uniform Feller and an irreducibility properties; see [LY94, Sza97, KS00, LS06, KS12]. The main theorem of this section is a generalisation of Theorem 4.2 in [KS00] and has independent interest.

Let X be a compact metric space, let $\mathcal{M}_+(X)$ be the space of non-negative Borel measures on X endowed with the topology of weak convergence, and let $\{P(u,\cdot), u \in X\} \subset \mathcal{M}_+(X)$ be a family satisfying the following condition:

Feller property. The function $u \mapsto P(u,\cdot)$ from X to $\mathcal{M}_+(X)$ is continuous and non-vanishing.

In this case, we shall say that $P(u,\Gamma)$ is a *generalised Markov kernel*. One obvious consequence of the Feller property is the inequality

$$C^{-1} \leq P(u,X) \leq C \quad \text{for all } u \in X.$$

Define the operators

$$\mathfrak{P}f(u) = \int_X P(u,dv)f(v), \quad \mathfrak{P}^*\mu(\Gamma) = \int_X P(u,\Gamma)\mu(du)$$

and denote $\mathfrak{P}_k = \mathfrak{P}^k$ and $\mathfrak{P}_k^* = (\mathfrak{P}^*)^k$. It is easy to see that

$$\mathfrak{P}_k f(u) = \int_X P_k(u,dv)f(v), \quad \mathfrak{P}_k^*\mu(\Gamma) = \int_X P_k(u,\Gamma)\mu(du),$$

where $P_k(u,\Gamma)$ is defined by the relations $P_0(u,\cdot)=\delta_u$, $P_1(u,\cdot)=P(u,\cdot)$, and

$$P_k(u,\cdot)=\int_X P_{k-1}(u,dv)P(v,\cdot),\quad k\geq 2.$$

To simplify the notation, the sup-norm on $C(X)$ is denoted in this section by $\|\cdot\|$. Let $\mathbf{1}$ be the function on X identically equal to 1. Recall that a family $\mathcal{C}\subset C(X)$ is called *determining* if any two measures $\mu,\nu\in\mathcal{M}_+(X)$ satisfying the relation $\langle f,\mu\rangle=\langle f,\nu\rangle$ for all $f\in\mathcal{C}$ coincide. In this section we prove:

Theorem 2.1. *Let $P(u,\Gamma)$ be a generalised Markov kernel satisfying the following conditions.*

Uniform Feller property. *There is a determining family $\mathcal{C}\subset C_+(X)$ of non-zero functions such that $\mathbf{1}\in\mathcal{C}$ and the sequence $\{\|\mathfrak{P}_k f\|^{-1}\mathfrak{P}_k f, k\geq 0\}$ is uniformly equicontinuous for any $f\in\mathcal{C}$.*

Uniform irreducibility. *For any $r>0$ there is an integer $m\geq 1$ and a constant $p>0$ such that*

$$P_m(u,B(\hat{u},r))\geq p\quad \textit{for all } u,\hat{u}\in X. \tag{2.1}$$

Then there is a constant $\lambda>0$, a unique measure $\mu\in\mathcal{P}(X)$ whose support coincides with X, and a unique $h\in C_+(X)$ satisfying $\langle h,\mu\rangle=1$, such that for any $f\in C(X)$ and $\nu\in\mathcal{M}_+(X)$ we have

$$\mathfrak{P}h=\lambda h,\quad \mathfrak{P}^*\mu=\lambda\mu, \tag{2.2}$$

$$\lambda^{-k}\mathfrak{P}_k f\to\langle f,\mu\rangle h\quad \textit{in } C(X) \textit{ as } k\to\infty, \tag{2.3}$$

$$\lambda^{-k}\mathfrak{P}_k^*\nu\rightharpoonup\langle h,\nu\rangle\mu\quad \textit{as } k\to\infty. \tag{2.4}$$

Proof. Note that the uniqueness of h and μ is an immediate consequence of the normalisation and relations (2.2)–(2.4). We split the proof in four steps.

Step 1. We first prove the existence of a measure satisfying the second relation in (2.2). To this end, let $F:\mathcal{P}(X)\to\mathcal{P}(X)$ be a map defined by

$$F(\mu)=(\mathfrak{P}^*\mu(X))^{-1}\mathfrak{P}^*\mu.$$

The Feller property implies that this map is well defined and continuous in the weak topology. Since $\mathcal{P}(X)$ is a convex compact set, by the Leray–Schauder theorem, the mapping F has a fixed point $\mu\in\mathcal{P}(X)$. We thus obtain the second relation in (2.2) with $\lambda=\mathfrak{P}^*\mu(X)$. In what follows, we may assume without loss of generality that $\lambda=1$; otherwise, we can replace $P(u,\Gamma)$ by $\lambda^{-1}P(u,\Gamma)$.

Step 2. Let us prove that, for any $f\in\mathcal{C}$, we have

$$C_f^{-1}\leq\|\mathfrak{P}_k f\|\leq C_f\quad \text{for all } k\geq 1, \tag{2.5}$$

where $C_f>1$ is a constant not depending on k. Indeed, suppose that there is a sequence $k_j\to\infty$ such that

$$\|\mathfrak{P}_{k_j}f\|+\|\mathfrak{P}_{k_j}f\|^{-1}\to+\infty\quad \text{as } j\to\infty. \tag{2.6}$$

In view of the uniform Feller property, we can assume that

$$\|\mathfrak{P}_{k_j} f\|^{-1}\mathfrak{P}_{k_j} f \to g \quad \text{in } C(X) \text{ as } j \to \infty,$$

where $g \in C(X)$ is function whose norm is equal to 1. Integrating with respect to μ and using the invariance of μ, we derive

$$\|\mathfrak{P}_{k_j} f\|^{-1}\langle f, \mu\rangle \to \langle g, \mu\rangle \quad \text{as } j \to \infty. \tag{2.7}$$

The uniform irreducibility implies that for any $\hat{u} \in X$ and $r > 0$ we have

$$\mu\big(B(\hat{u}, r)\big) = \int_X P_m\big(u, B(\hat{u}, r)\big)\mu(du) \geq p\,\mu(X) > 0.$$

Hence, $\operatorname{supp}\mu = X$, and since $f, g \in C_+(X)$ are non-zero functions, we have that $\langle f, \mu\rangle > 0$ and $\langle g, \mu\rangle > 0$. It now follows from (2.7) that the sequence $\|\mathfrak{P}_{k_j} f\|$ has a finite positive limit, and therefore (2.6) cannot hold.

Step 3. Let us prove the existence of $h \in C_+(X)$ satisfying the first relation in (2.2) with $\lambda = 1$. Let $f \in \mathcal{C}$ be an arbitrary function. The uniform Feller property and inequality (2.5) imply that the sequence $\mathfrak{P}_k f$ is uniformly equicontinuous. It follows that so is the sequence

$$f_k = \frac{1}{k}\sum_{l=0}^{k-1} \mathfrak{P}_l f.$$

Let h be a limit point for $\{f_k\}$. It is straightforward to see that $h \geq 0$ and $\mathfrak{P}_1 h = h$. Furthermore, since $\langle f_k, \mu\rangle = \langle f, \mu\rangle > 0$, we see that h is non-zero. Multiplying h by a constant, we can assume that $\langle h, \mu\rangle = 1$. It remains to prove that $h(u) > 0$ for all $u \in X$. Indeed, let $\hat{u} \in X$ be any point at which h is positive. Then there is $r > 0$ such that $h(v) \geq r$ for $v \in B(\hat{u}, r)$. It follows that, for any $u \in X$, we have

$$\begin{aligned} h(u) = \mathfrak{P}_m h(u) &= \int_X P_m(u, dv)h(v) \geq \int_{B(\hat{u},r)} P_m(u, dv)h(v) \\ &\geq r P_m\big(u, B(\hat{u}, r)\big) \geq rp > 0, \end{aligned}$$

where $m \geq 1$ is the integer from (2.1).

Step 4. We now establish convergence (2.3) and (2.4) with $\lambda = 1$. To this end, we first note that (2.4) is an immediate consequence of (2.3). Furthermore, the right-hand inequality in (2.5) with $f = \mathbf{1}$ implies that the norms of the operators $\mathfrak{P}_k$ are bounded by $C_{\mathbf{1}}$ for all $k \geq 1$. Since the linear span of a determining family is dense in $C(X)$, it suffices to establish (2.3) for any $f \in \mathcal{C}$.

Let us fix an arbitrary $f \in \mathcal{C}$ and define the function $g = f - \langle f, \mu\rangle h$. We need to prove that $\mathfrak{P}_k g \to 0$ in $C(X)$. Since $\{\mathfrak{P}_k g, k \geq 0\}$ is uniformly equicontinuous and the norms of $\mathfrak{P}_k$ are bounded, the required assertion will be established if we prove that any sequence of integers $n_i \to \infty$ contains a subsequence $\{k_j\} \subset \{n_i\}$ for which

$$\|g_{k_j}\|_\mu \to 0 \quad \text{as } j \to \infty, \tag{2.8}$$

where we set $g_k = \mathfrak{P}_k g$. Since $\langle g_k, \mu\rangle = 0$ for $k \geq 0$, convergence (2.8) certainly holds for any subsequence $\{k_j\}$ such that $\|g^+_{k_j}\| \to 0$ or $\|g^-_{k_j}\| \to 0$ as $j \to \infty$. Let us assume that there is no subsequence satisfying this property. Then there exist sequences $\{u^\pm_i\} \subset X$ and a constant $\alpha > 0$ such that

$$\tilde{g}^+_i(u^+_i) = \max_{u\in X} \tilde{g}^+_i(u) \geq \alpha, \quad \tilde{g}^-_i(u^-_i) = \max_{u\in X} \tilde{g}^-_i(u) \geq \alpha, \tag{2.9}$$

where we set $\tilde{g}_i = g_{k_i}$. Since $\tilde{g}^\pm_i$ are uniformly equicontinuous, we can find $r > 0$ not depending on i such that

$$\tilde{g}^\pm_i(u) \geq \frac{1}{2}\tilde{g}^\pm_i(u^\pm_i) \quad \text{for } u \in B(u^\pm_i, r). \tag{2.10}$$

Let m and p be the constants arising in the uniform irreducibility condition. Then (2.10) and (2.5) imply

$$\begin{aligned}
\mathfrak{P}_m \tilde{g}^\pm_i(u) &= \int_X P_m(u, dv)\tilde{g}^\pm_i(v) \leq C_{\mathbf{1}}\tilde{g}^\pm_i(u^\pm_i),\\
\mathfrak{P}_m \tilde{g}^\pm_i(u) &\geq \int_{B(u^\pm_i, r)} P_m(u, dv)\tilde{g}^\pm_i(v) \geq p\,\tilde{g}^\pm_i(u^\pm_i)/2,
\end{aligned}$$

and it follows that

$$\sup_{u\in X} \mathfrak{P}_m \tilde{g}^\pm_i(u) \leq A_g \inf_{u\in X} \mathfrak{P}_m \tilde{g}^\pm_i(u), \tag{2.11}$$

where $A_g = 2C_{\mathbf{1}}/p > 1$ (so that $0 < 1 - A_g^{-1} < 1$). In particular, due to the stationarity of μ, we have

$$\mathfrak{P}_m \tilde{g}^\pm_i(u) \geq A_g^{-1}\|\mathfrak{P}_m \tilde{g}^\pm_i\| \geq A_g^{-1}\|\mathfrak{P}_m \tilde{g}^\pm_i\|_\mu = A_g^{-1}\|\tilde{g}^\pm_i\|_\mu.$$

Using this inequality, we now write

$$\begin{aligned}
\|\mathfrak{P}_m \tilde{g}_i\|_\mu &= \int_X |\mathfrak{P}_m(\tilde{g}^+_i - \tilde{g}^-_i)|d\mu\\
&= \int_X \big|(\mathfrak{P}_m \tilde{g}^+_i - A_g^{-1}\|\tilde{g}^+_i\|_\mu) - (\mathfrak{P}_m \tilde{g}^-_i - A_g^{-1}\|\tilde{g}^-_i\|_\mu)\big|d\mu\\
&\leq \int_X \big|\mathfrak{P}_m \tilde{g}^+_i - A_g^{-1}\|\tilde{g}^+_i\|_\mu\big|d\mu + \int_X \big|\mathfrak{P}_m \tilde{g}^-_i - A_g^{-1}\|\tilde{g}^-_i\|_\mu\big|d\mu\\
&= \int_X \mathfrak{P}_m(\tilde{g}^+_i + \tilde{g}^-_i)d\mu - A_g^{-1}\big(\|\tilde{g}^+_i\|_\mu + \|\tilde{g}^-_i\|_\mu\big)\\
&= (1 - A_g^{-1})\|\tilde{g}_i\|_\mu.
\end{aligned} \tag{2.12}$$

Furthermore, for any $f \in C(X)$ and $k \geq 1$, we have

$$\|\mathfrak{P}_k f\|_\mu = \langle|\mathfrak{P}_k f|, \mu\rangle \leq \langle\mathfrak{P}_k|f|, \mu\rangle = \langle|f|, \mu\rangle = \|f\|_\mu.$$

It follows that the sequence $\{\|\mathfrak{P}_k g\|_\mu\}$ is non-increasing. Combining this property with (2.12), we see that if $n_l \geq n_i + m$, then

$$\|g_{n_l}\|_\mu \leq (1 - A_g^{-1})\|g_{n_i}\|_\mu. \tag{2.13}$$

Let us choose a subsequence $\{k_j\} \subset \{n_i\}$ such that $k_{j+1} \geq k_j + m$. Then (2.13) implies that

$$\|g_{k_j}\|_\mu = \|\mathfrak{P}_{k_j} g\|_\mu \leq (1 - A_g^{-1})^j \|g\|_\mu \quad \text{for } j \geq 0,$$

where $k_0 = 0$. This proves convergence (2.8) and completes the proof of the theorem. □

3 The uniform Feller property

We shall use freely the notation introduced in Subsection 1.1 (we recall, in particular, that $\{e_j\}$ is the orthonormal basis introduced in Condition (C), P_N is the orthogonal projection onto $H_N = \operatorname{span}\{e_1, \dots, e_N\}$, and $\mathcal{A} = \mathcal{A}(\{0\})$ is the domain of attainability from zero). Let $\mathcal{V}$ be the set of functions $V \in C(\mathcal{A})$ for which there is an integer $N \geq 1$ and a function $F \in C(H_N)$ such that

$$V(u) = F(\mathsf{P}_N u) \quad \text{for } u \in \mathcal{A}. \tag{3.1}$$

It is easy to see that $\mathcal{V}$ is a dense subspace in $C(\mathcal{A})$ containing the constant functions. In particular, the intersection $\mathcal{C} = \mathcal{V} \cap C_+(\mathcal{A})$ is a determining family for $\mathcal{P}(\mathcal{A})$.

For any $V \in C(\mathcal{A})$, let us consider the following generalised Markov kernel on $\mathcal{A}$:

$$P_1^V(u, \Gamma) = \mathbb{E}_u\big(I_\Gamma(u_1) e^{V(u_1)}\big) = \int_\Gamma P_1(u, dv) e^{V(v)}, \quad u \in \mathcal{A}, \quad \Gamma \in \mathcal{B}(\mathcal{A}). \tag{3.2}$$

The corresponding semigroup of operators is given by (1.42). The goal of this section is to prove:

Theorem 3.1. *Under the hypotheses of Theorem 1.3, for any $V \in \mathcal{V}$ the semigroup $\{\mathfrak{P}_k^V\}$ possesses the uniform Feller property for the determining class $\mathcal{C}$. In other words, for any $V \in \mathcal{V}$ and $f \in \mathcal{C}$ the sequence $\{\|\mathfrak{P}_k^V f\|_\infty^{-1} \mathfrak{P}_k^V f, k \geq 0\}$ is uniformly equicontinuous.*

This theorem will play a key role in the proof of our main results. To prove it, we first recall a coupling construction for the Markov chain associated with (1.1) and then use it to establish Theorem 3.1.

3.1 Coupling

Let us denote by $\boldsymbol{\mathcal{A}} = \mathcal{A}^{\mathbb{Z}_+}$ the direct product of countably many copies of $\mathcal{A}$ endowed with the Tikhonov topology and by $\boldsymbol{P}(v)$ the law of the trajectory $(u_k, k \geq 0)$ for (1.1) issued from $v \in \mathcal{A}$. Thus, $\boldsymbol{P}(v)$ is a probability measure on $\boldsymbol{\mathcal{A}}$. The following result is established in [KS12] (see Section 3.2.2).

Proposition 3.2. *Under the hypotheses of Theorem 1.3, for any sufficiently large integer $N \geq 1$ there is a probability space $(\Omega_N, \mathcal{F}_N, \mathbb{P}_N)$, positive numbers C_N and $\gamma_N < 1$ such that*

$$\gamma_N \to 0 \quad \text{as } N \to \infty, \tag{3.3}$$

and an $\mathcal{A} \times \mathcal{A}$-valued Markov process (u_k, u'_k) on Ω_N parametrised by the initial point $(v, v') \in \mathcal{A} \times \mathcal{A}$ for which the following properties hold.

(a) *The $\mathbb{P}_N$-laws of the sequences $\{u_k, k \geq 0\}$ and $\{u'_k, k \geq 0\}$ coincide with $\boldsymbol{P}(v)$ and $\boldsymbol{P}(v')$, respectively.*

(b) *For any integer $r \geq 1$, we have*[5]

$$\mathbb{P}_N\big\{\mathsf{P}_N u_k = \mathsf{P}_N u'_k \text{ for } 1 \leq k \leq r-1, \mathsf{P}_N u_r \neq \mathsf{P}_N u'_r\big\} \leq C_N \gamma_N^r \|v - v'\|, \tag{3.4}$$

where the trajectory (u_k, u'_k) entering the left-hand side corresponds to the initial point $(v, v') \in \mathcal{A} \times \mathcal{A}$.

Note that, in [KS12], inequality (3.4) is proved with a fixed γ_N. However, since the numbers $\gamma_N(R)$ entering (1.5) go to zero, we can make γ_N arbitrarily small by choosing a large N.

3.2 Proof of Theorem 3.1

Let us fix two functions $V \in \mathcal{V}$ and $f \in \mathcal{C}$. There is no loss of generality in assuming that the integer $N \geq 1$ entering representation (3.1) for V and f is the same. Abusing slightly the notation, we shall write V and f for the corresponding function F in the right-hand side. Let us set

$$g_k(u) = \|\mathfrak{P}_k^V f\|_\infty^{-1} \mathfrak{P}_k^V f(u), \quad k \geq 0.$$

We need to show that $\{g_k, k \geq 0\}$ is uniformly equicontinuous. One can assume that $0 < f \leq 1$ and $\inf_{\mathcal{A}} V = 0$.

Let us fix two points $v, v' \in \mathcal{A}$ and denote by $\{u_k\}$ and $\{u'_k\}$ the sequences constructed in Proposition 3.2. Denoting by $\Xi_V f(z_1, \dots, z_k)$ the expression $\exp(V(z_1) + \dots + V(z_k)) f(z_k)$ and by $A(r)$ the event on the left-hand side of (3.4), we can write

$$\mathfrak{P}_k^V f(v) - \mathfrak{P}_k^V f(v') = \sum_{r=1}^{k} I_k^r(v, v'), \tag{3.5}$$

where we set

$$I_k^r(v, v') = \mathbb{E}_N\big\{I_{A(r)}\big(\Xi_V f(u_1, \dots, u_k) - \Xi_V f(u'_1, \dots, u'_k)\big)\big\},$$

and $\mathbb{E}_N$ stands for the expectation corresponding to $\mathbb{P}_N$. Let us denote by $\mathcal{F}_k^N$ the filtration generated by (u_k, u'_k). Since f is bounded by 1 and separated from

[5] The relation $\mathsf{P}_N u_k = \mathsf{P}_N u'_k$ in (3.4) should be omitted when $r = 1$.

zero by a positive number δ, taking the conditional expectation given $\mathcal{F}_r^N$ and carrying out some simple estimates, we derive

$$\begin{aligned} I_k^r(v,v') &\le \mathbb{E}_N\big\{I_{A(r)}e^{V(u_1)+\dots+V(u_k)}f(u_k)\big\} \\ &\le \mathbb{E}_N\big(I_{A(r)}e^{V(u_1)+\dots+V(u_r)}\,\mathbb{E}_N\big\{\Xi_V\mathbf{1}(u_{r+1},\dots,u_k)\,\big|\,\mathcal{F}_r^N\big\}\big) \\ &\le e^{r\|V\|_\infty}\mathbb{E}_N\big(I_{A(r)}(\mathfrak{P}_{k-r}^V\mathbf{1})(u_r)\big) \le e^{r\|V\|_\infty}\mathbb{E}_N\big(I_{A(r)}\|\mathfrak{P}_k^V\mathbf{1}\|_\infty\big) \\ &\le \delta^{-1}e^{r\|V\|_\infty}\|\mathfrak{P}_k^V f\|_\infty\mathbb{P}_N\big(A(r)\big). \end{aligned} \tag{3.6}$$

Substituting this into (3.5) and using (3.4), we derive

$$|g_k(v)-g_k(v')| \le C_N\delta^{-1}\|v-v'\|\sum_{r=1}^k e^{r\|V\|_\infty}\gamma_N^r.$$

Choosing N so large that $\|V\|_\infty+\log\gamma_N<0$, we obtain

$$|g_k(v)-g_k(v')| \le C_N'\|v-v'\| \quad \text{for all } v,v'\in\mathcal{A},\ k\ge 0.$$

The proof of Theorem 3.1 is complete.

4 Proof of Theorem 1.3

We shall prove Theorem 1.3 by verifying Property 1 (the existence of a limit) and Property 2 (uniqueness of the equilibrium state) of Section 1.4.

Let $P_k^V(u,\Gamma)$, $\{\mathfrak{P}_k^V\}$, $\mathcal{V}$, and $\mathcal{C}$ be as in Theorem 3.1. For any $V\in C(\mathcal{A})$,

$$P_k^V(u,\cdot)\ge e^{-k\|V\|_\infty}P_k(u,\cdot) \quad \text{for any } u\in\mathcal{A}.$$

Since $P_k(u,\Gamma)$ is uniformly irreducible (e.g., see Section 5 of [KS00] for a proof of a similar assertion in a more complicated setting), so is $P_k^V(u,\Gamma)$. By Theorem 3.1, for any $V\in\mathcal{V}$ the semigroup $\{\mathfrak{P}_k^V\}$ possesses the uniform Feller property for the determining class $\mathcal{C}$. Thus, for $V\in\mathcal{V}$, Theorem 2.1 holds for the semigroup $\{\mathfrak{P}_k^V\}$ and the class $\mathcal{C}$.

We now turn to the proof of Property 1 and the existence of the limit (1.40). Theorem 2.1 implies that for any $V\in\mathcal{V}$ there is $h_V\in C_+(\mathcal{A})$ and a constant $\lambda_V>0$ such that

$$\lambda_V^{-k}\mathfrak{P}_k^V\mathbf{1}\to h_V \quad \text{in } C(\mathcal{A}) \text{ as } k\to\infty.$$

It follows that for $V\in\mathcal{V}$

$$Q(V)=\lim_{k\to+\infty}\frac{1}{k}\log(\mathfrak{P}_k^V\mathbf{1})(u)=\log\lambda_V \tag{4.1}$$

uniformly in $u\in\mathcal{A}$. The estimate

$$\begin{aligned}(\mathfrak{P}_k^{V_1}\mathbf{1})(u) &= \mathbb{E}_u\exp\bigg(\sum_{n=1}^k V_1(u_n)\bigg)\le e^{k\|V_1-V_2\|_\infty}\,\mathbb{E}_u\exp\bigg(\sum_{n=1}^k V_2(u_n)\bigg) \\ &= e^{k\|V_1-V_2\|_\infty}\,(\mathfrak{P}_k^{V_2}\mathbf{1})(u),\end{aligned}$$

which holds for any $V_1, V_2 \in C(\mathcal{A})$, implies

$$\left| \frac{1}{k} \log(\mathfrak{P}_k^{V_1} \mathbf{1})(u) - \frac{1}{k} \log(\mathfrak{P}_k^{V_2} \mathbf{1})(u) \right| \leq \|V_1 - V_2\|_\infty \quad \text{for } k \geq 1,\ u \in \mathcal{A}. \tag{4.2}$$

Hence, (4.1) holds for all $V \in C(\mathcal{A})$, the limit is uniform in $u \in \mathcal{A}$, and

$$|Q(V_1) - Q(V_2)| \leq \|V_1 - V_2\|_\infty \quad \text{for } V_1, V_2 \in C(\mathcal{A}). \tag{4.3}$$

The existence of the limit (1.40) for an arbitrary $\mathcal{A}$-valued initial random variable u_0 now follows by integration with respect to the law of u_0. The Hölder inequality implies immediately that Q is a convex function.

Let us prove Property 2. We shall show that, for any $V \in \mathcal{V}$, there is a unique equilibrium state $\sigma_V \in \mathcal{P}(\mathcal{A})$ for $Q(V)$. To this end, we first derive a necessary and sufficient condition for a measure $\sigma \in \mathcal{P}(\mathcal{A})$ to be an equilibrium state.

Recall that $I : \mathcal{M}(\mathcal{A}) \to \mathbb{R}$ is the Legendre transform of the 1-Lipschitz convex function $Q : C(\mathcal{A}) \to \mathbb{R}$. Given a function $V \in C(\mathcal{A})$, introduce a Markov semigroup by (1.46) and denote by $\{\mathscr{S}_k^{V*}\}$ its dual semigroup acting on $\mathcal{P}(\mathcal{A})$. As in the case of $\mathfrak{P}_k$, for any $F \in C(\mathcal{A})$ we can consider a generalised Markov semigroup defined by

$$\mathfrak{S}_1^F f = \mathscr{S}_1^V (e^F f), \quad \mathfrak{S}_k^F = (\mathfrak{S}_1^F)^k.$$

We claim that, for any $F \in C(\mathcal{A})$, the limit

$$Q^V(F) = \lim_{k \to \infty} \frac{1}{k} \log(\mathfrak{S}_k^F \mathbf{1})(u) \tag{4.4}$$

is well defined and does not depend on $u \in \mathcal{A}$. Indeed, it is straightforward to check using (1.46) that

$$\mathfrak{S}_k^F \mathbf{1} = \lambda_V^{-k} h_V^{-1} \mathfrak{P}_k^{V+F} h_V.$$

In view of convergence (1.44), it follows that (4.4) exists for any $F \in \mathcal{V}$ and is equal to

$$Q^V(F) = Q(V + F) - Q(V). \tag{4.5}$$

Repeating the simple approximation argument used in the proof of (4.1), we conclude that the limit (4.4) is well defined for any $F \in C(\mathcal{A})$ and is given by (4.5). It follows that the Legendre transform of Q^V has the form

$$I^V(\sigma) = I(\sigma) + Q(V) - \langle V, \sigma \rangle. \tag{4.6}$$

Thus, a probability measure σ is an equilibrium state for $Q(V)$ if and only if $I^V(\sigma) = 0$. On the other hand, by Lemma 2.5 in [DV75] (see also Lemma 4.1.45 in [DS89]), the latter holds if and only if σ is a stationary measure for $\mathscr{S}_1^{V*}$. Thus, the required property will be established if we prove that $\mathscr{S}_1^{V*}$ has a unique stationary measure.

In view of (1.46), we have

$$\mathscr{S}_k^{V*}\sigma = \lambda_V^{-k} h_V \mathfrak{P}_k^{V*}(h_V^{-1}\sigma). \tag{4.7}$$

It follows that $\sigma_V = h_V \mu_V$ (where μ_V is the measure in (1.43)) is a stationary measure for $\mathscr{S}_1^{V*}$. Moreover, if $\sigma \in \mathcal{P}(\mathcal{A})$ is another stationary measure for $\mathscr{S}_k^{V*}$, then, by (1.45), we have

$$\sigma = \mathscr{S}_k^{V*}\sigma = \lambda_V^{-k} h_V \mathfrak{P}_k^{V*}(h_V^{-1}\sigma) \to h_V \mu_V = \sigma_V.$$

This completes the proof of uniqueness of the equilibrium state for $V \in \mathcal{V}$ and that of Theorem 1.3.

5 Proof of Theorem 1.4

As described in Section 1.5, Theorem 1.4 follows from Theorem 1.8 (which in turn is a generalisation of Theorem 1.3). To establish Theorem 1.8, one follows the general scheme used in the proof of Theorem 1.3, applying it to the Markov chain formed by the segments of trajectories of length m. Namely, let us consider the following family of Markov chains in $\mathcal{A}^{(m)}$:

$$\boldsymbol{u}_k = \boldsymbol{S}(\boldsymbol{u}_{k-1}) + \boldsymbol{\eta}_k, \tag{5.1}$$

where $\boldsymbol{u}_k = (u_k^1, \dots, u_k^m)$, $\boldsymbol{\eta}_k = (0, \dots, 0, \eta_{k+m-1})$, and $\boldsymbol{S} : H^m \to H^m$ is the mapping given by

$$\boldsymbol{S}(v_1, \dots, v_m) = \big(v_2, \dots, v_m, S(v_m)\big), \quad (v_1, \dots, v_m) \in \boldsymbol{H}.$$

It is clear that if u_0 is an $\mathcal{A}$-valued random variable independent of $\{\eta_k\}$ and $\{u_k\}$ is the corresponding trajectory of (1.1), then ζ_k^m is the occupation measure for the trajectory of (5.1) starting from the (random) initial point $(u_0, \dots, u_{m-1})$. Since its law is supported by $\mathcal{A}^{(m)}$, the LDP for ζ_k^m will be established if we prove the LDP for the Markov family (5.1) restricted to the invariant compact set $\mathcal{A}^{(m)}$. By Kifer's theorem and the argument described Section 1.4, the latter result is a consequence of the following two properties (which were described in Section 1.4 for $\{u_k\}$):

Property 1': The existence of a limit. For any function $V \in C(\mathcal{A}^{(m)})$ the limit

$$Q_m(V) = \lim_{k \to +\infty} \frac{1}{k} \log \mathbb{E} \exp\bigg(\sum_{n=1}^{k} V(\boldsymbol{u}_n)\bigg) \tag{5.2}$$

exists and does not depend on the initial point $\boldsymbol{u} = (u^1, \dots, u^m) \in \mathcal{A}^{(m)}$.

Property 2': Uniqueness of equilibrium state. There exists a dense vector space $\mathcal{V}_m \subset C(\mathcal{A}^{(m)})$ such that, for any $V \in \mathcal{V}_m$, there is a unique measure $\sigma_V \in \mathcal{P}(\mathcal{A}^{(m)})$ satisfying the relation

$$Q_m(V) = \sup_{\sigma \in \mathcal{P}(\mathcal{A}^{(m)})} \big(\langle V, \sigma\rangle - I_m(\sigma)\big),$$

where $I_m(\sigma)$ denotes the Legendre transform of Q_m.

To establish these assertions, we introduce a generalised Markov semigroup by the relation (cf. (1.42))

$$\mathfrak{P}_k^V f(\boldsymbol{u}) := \mathbb{E}_{\boldsymbol{u}} f(\boldsymbol{u}_k) \exp\Big(\sum_{n=1}^k V(\boldsymbol{u}_n)\Big), \quad f \in C(\mathcal{A}^{(m)}), \tag{5.3}$$

where $V \in C(\mathcal{A}^{(m)})$ is a given function. If we prove that $\{\mathfrak{P}_k^V\}$ satisfies the uniform Feller and uniform irreducibility properties of Theorem 2.1 for any V belonging to a dense subspace $\mathcal{V}_m$ that contains constant functions, then the required results will following line by line the proof of Theorem 1.3.

To show the uniform irreducibility, note that $\mathcal{A}^{(m)}$ is the domain of attainability from zero for system (5.1). Therefore the required property follows by repeating the proof of a similar property for (1.1).

We now turn to the uniform Feller property. Let $\mathcal{V}_m$ be the space of functions $V \in C(\mathcal{A}^{(m)})$ for which there is an integer $N \geq 1$ and a function $F \in C(H_N^m)$ such that

$$V(\boldsymbol{u}) = F(\mathsf{P}_N \boldsymbol{u}) \quad \text{for } \boldsymbol{u} = (u^1, \dots, u^m) \in \mathcal{A}^{(m)},$$

where $\mathsf{P}_N \boldsymbol{u} = (\mathsf{P}_N u^1, \dots, \mathsf{P}_N u^m)$. Given $\boldsymbol{v}, \boldsymbol{v}' \in \mathcal{A}^{(m)}$, we define

$$\boldsymbol{v}_k = (\boldsymbol{v}, u_1, \dots, u_k), \quad \boldsymbol{v}'_k = (\boldsymbol{v}', u'_1, \dots, u'_k), \quad \boldsymbol{w}_k = (\boldsymbol{v}', u_1, \dots, u_k),$$

where $\{u_k\}$ and $\{u'_k\}$ stand for the trajectories issued from $\boldsymbol{v}$ and $\boldsymbol{v}'$, respectively. Then, for $V, f \in \mathcal{V}_m$ and $k \geq m$, we have (cf. (3.5))

$$\mathfrak{P}_k^V f(\boldsymbol{v}) - \mathfrak{P}_k^V f(\boldsymbol{v}') = I_k(\boldsymbol{v}, \boldsymbol{v}') + \sum_{r=1}^k I_k^r(\boldsymbol{v}, \boldsymbol{v}'),$$

where we set [6]

$$\begin{aligned} I_k(\boldsymbol{v}, \boldsymbol{v}') &= \mathbb{E} f(\boldsymbol{v}_k)\big(e^{V(\boldsymbol{v}_1)+\dots+V(\boldsymbol{v}_k)} - e^{V(\boldsymbol{w}_1)+\dots+V(\boldsymbol{w}_k)}\big), \\ I_k^r(\boldsymbol{v}, \boldsymbol{v}') &= \mathbb{E} I_{A(r)}\big(e^{V(\boldsymbol{w}_1)+\dots+V(\boldsymbol{w}_k)} f(\boldsymbol{v}_k) - e^{V(\boldsymbol{v}'_1)+\dots+V(\boldsymbol{v}'_k)} f(\boldsymbol{v}'_k)\big), \end{aligned}$$

and $A(r)$ is defined in Section 3.2. Since the last k elements of the sequences $\boldsymbol{v}_k$ and $\boldsymbol{w}_k$ are the same, the expression $I_k(\boldsymbol{v}, \boldsymbol{v}')$ can be estimated uniformly in $k \geq m$ by a function of the form $\|\mathfrak{P}_k^V f\|_\infty \, g(\boldsymbol{v}, \boldsymbol{v}')$, where $g(\boldsymbol{v}, \boldsymbol{v}') \to 0$ as $\boldsymbol{v} - \boldsymbol{v}' \to 0$. On the other hand, $I_k^r(\boldsymbol{v}, \boldsymbol{v}')$ can be bounded in exactly the same way as in the proof of Theorem 3.1. This completes the proof of Theorems 1.8 and 1.4.

[6] Given a function $g : \mathcal{A}^{(m)} \to \mathbb{R}$ and any finite sequence $\boldsymbol{z} \subset \mathcal{A}$ of length $\geq m$, with a slight abuse of notation we denote by $g(\boldsymbol{z})$ the value of g calculated on the last m elements of $\boldsymbol{z}$.

6 Appendix

In this section, we recall three results on the large deviation principle (LDP). The first of them was established by Kifer [Kif90] and provides a sufficient condition for the validity of LDP for a family of random probability measures. The second result shows that, when studying the LDP for occupation measures of random processes, one can take the average starting from any non-negative time. The third result due to Dawson and Gärtner [DG87] shows that the process level LDP is a straightforward consequence of the LDP for finite segments of solutions.

6.1 Kifer's sufficient condition for LDP

Let Θ be a directed set, let X be a compact metric space, and let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space. We consider a family $\{\zeta_\theta\} = \{\zeta_\theta^\omega\}$ of random probability measures on X depending on $\theta \in \Theta$ such that the following limit exists for any $V \in C(X)$:

$$Q(V) = \lim_{\theta \in \Theta} \frac{1}{r(\theta)} \log \int_\Omega \exp\left(r(\theta) \int_X V \mathrm{d}\zeta_\theta^\omega \right) \mathrm{d}\mathbb{P}(\omega), \tag{6.1}$$

where $r : \Theta \to \mathbb{R}$ is a given positive function such that $\lim_{\theta \in \Theta} r(\theta) = \infty$. Then $Q : C(X) \to \mathbb{R}$ is a convex 1-Lipschitz functional such that $Q(V) \geq 0$ for any $V \in C_+(X)$ and $Q(C) = C$ for any constant $C \in \mathbb{R}$. Recall that the Legendre transform of Q is defined on the space $\mathcal{M}(X)$ by

$$I(\sigma) = \sup_{V \in C(X)} \big(\langle V, \sigma \rangle - Q(V) \big) \tag{6.2}$$

if $\sigma \in \mathcal{P}(X)$ and $I(\sigma) = \infty$ otherwise. The function $I(\sigma)$ is convex and lower semicontinuous in the weak topology, and Q can be reconstructed by the formula

$$Q(V) = \sup_{\sigma \in \mathcal{P}(X)} \big(\langle V, \sigma \rangle - I(\sigma) \big).$$

Since $\mathcal{P}(X)$ endowed with topology of weak convergence is compact, for any $V \in C(X)$ there is $\sigma_V \in \mathcal{P}(X)$ such that

$$Q(V) = \langle V, \sigma_V \rangle - I(\sigma_V). \tag{6.3}$$

Any measure $\sigma_V \in \mathcal{P}(X)$ satisfying (6.3) is called an *equilibrium state* for V. The following result of Kifer shows that if the equilibrium state is unique for a dense vector subspace of $V \in C(X)$, then the LDP holds for ζ_θ.

Theorem 6.1. *Suppose that limit* (6.1) *exists for any* $V \in C(X)$. *Then the LD upper bound*

$$\limsup_{\theta \in \Theta} \frac{1}{r(\theta)} \log \mathbb{P}\{\zeta_\theta \in F\} \leq -I(F)$$

holds with the rate function I given by (6.2). Furthermore, if there exists a dense vector space $\mathcal{V} \subset C(X)$ such that the equilibrium state $\sigma_V \in \mathcal{P}(X)$ is unique for any $V \in \mathcal{V}$, then the LD lower bound also holds:

$$\liminf_{\theta \in \Theta} \frac{1}{r(\theta)} \log \mathbb{P}\{\zeta_\theta \in G\} \geq -I(G).$$

6.2 Exponential equivalence of random probability measures

Let X be a Polish space and let $\{\mu_k\}$ and $\{\mu'_k\}$ be two sequences of random probability measures on X. Recall that $\{\mu_k\}$ and $\{\mu'_k\}$ are said to be *exponentially equivalent* if

$$\lim_{k \to \infty} \mathbb{P}\big\{\|\mu_k - \mu'_k\|_L^* > \delta\big\}^{1/k} = 0 \quad \text{for any } \delta > 0. \tag{6.4}$$

It is well known that if two sequences of random probability measures are exponentially equivalent, then an LDP with a good rate function for one of them implies the same LDP for the other; see Section 4.2.2 in [DZ00].

Now let $\{u_n\}$ be a random sequence in X. We denote by $\zeta_k^{(m)}$ the occupation measures for $\{u_n\}$ starting at time $m \geq 0$:

$$\mu_k^{(m)} = \frac{1}{k} \sum_{n=m}^{m+k-1} \delta_{u_n}.$$

The following result was used in Sections 1.4.

Lemma 6.2. *The sequences $\mu_k^{(m)}$ and $\mu_k^{(l)}$ are exponentially equivalent for any integers $m, l \geq 0$.*

Proof. Let $f \in L_b(X)$ be such that $\|f\|_L \leq 1$. Then

$$\big|\big(f, \mu_k^{(m)}\big) - \big(f, \mu_k^{(l)}\big)\big| \leq \frac{2|m-l|}{k}.$$

It follows that $\big\|\mu_k^{(m)} - \mu_k^{(l)}\big\|_L^* \leq \frac{2|m-l|}{k}$, whence we see that

$$\mathbb{P}\big\{\|\mu_k^{(m)} - \mu_k^{(l)}\|_L^* > \delta\big\} = 0 \quad \text{for } k > 2\delta^{-1}|m-l|.$$

Hence, condition (6.4) is satisfied for any $\delta > 0$, and the sequences in question are exponentially equivalent. □

6.3 Dawson–Gärtner theorem

For a given Polish space X, we denote by $\boldsymbol{X} = X^{\mathbb{Z}_+}$ the direct product of countably many copies of X, endowed with the Tikhonov topology, and by $p_m : \boldsymbol{X} \to X^m$ the natural projection to the first m components of $\boldsymbol{X}$. Let

$\{\boldsymbol{\zeta}_k\} = \{\boldsymbol{\zeta}_k^\omega\}$ be a sequence of random probability measures on $\mathcal{P}(\boldsymbol{X})$ and let ζ_k^m be the image of $\boldsymbol{\zeta}_k$ under the projection p_m. The following theorem is a particular case of a more general result established in [DG87] (see also Theorem 4.6.1 in [DZ00]).

Theorem 6.3. *Suppose that for any integer $m \geq 1$ the sequence $\{\zeta_k^m\}$ satisfies the LDP with a good rate function $I_m : \mathcal{P}(X^m) \to [0, +\infty]$. Then the LDP holds for $\{\boldsymbol{\zeta}_k\}$ with the good rate function*

$$\boldsymbol{I}(\boldsymbol{\sigma}) = \sup_{m \geq 1} I_m\big(\boldsymbol{\sigma} \circ p_m^{-1}\big). \tag{6.5}$$

Proof. Step 1: Rate function. Let us prove that the function $\boldsymbol{I}$ defined by (6.5) is a good rate function. Indeed, since I_m are good rate functions, for any $\alpha \in \mathbb{R}$ we have

$$\{\boldsymbol{I} \leq \alpha\} = \bigcap_{m=1}^{\infty} \big\{\boldsymbol{\sigma} \in \mathcal{P}(\boldsymbol{X}) : I_m\big(\boldsymbol{\sigma} \circ p_m^{-1}\big) \leq \alpha\big\} = \bigcap_{m=1}^{\infty} \{\boldsymbol{\sigma} \circ p_m^{-1} \in K_\alpha^m\}, \tag{6.6}$$

where K_α^m are compact subsets in $\mathcal{P}(X^m)$. This relation immediately implies that the set $\{\boldsymbol{I} \leq \alpha\}$ is closed and therefore $\boldsymbol{I}$ is lower semicontinuous. Furthermore, since a sequence $\{\boldsymbol{\sigma}_j\} \subset \mathcal{P}(\boldsymbol{X})$ converges if and only if so does $\{\boldsymbol{\sigma}_j \circ p_m^{-1}\}$ for any $m \geq 1$, it follows from (6.6) that the level sets of $\boldsymbol{I}$ are compact.

Step 2: Lower bound. Let $G \subset \mathcal{P}(\boldsymbol{X})$ be an open subset. It suffices to prove that, for any $\boldsymbol{\sigma} \in G$, we have

$$\liminf_{k \to \infty} \frac{1}{k} \log \mathbb{P}\{\boldsymbol{\zeta}_k \in G\} \geq -\boldsymbol{I}(\boldsymbol{\sigma}).$$

Since G is open, for any $\boldsymbol{\sigma} \in G$, one can find an integer $m \geq 1$ and open subset $G_m \subset \mathcal{P}(X^m)$ containing $\boldsymbol{\sigma} \circ p_m^{-1}$ such that $G \supset p_m^{-1}(G_m)$. Since the LDP holds for $\zeta_k^m = \boldsymbol{\zeta}_k \circ p_m^{-1}$, it follows that

$$\begin{aligned}\liminf_{k \to \infty} \frac{1}{k} \log \mathbb{P}\{\boldsymbol{\zeta}_k \in G\} &\geq \liminf_{k \to \infty} \frac{1}{k} \log \mathbb{P}\{\boldsymbol{\zeta}_k \in p_m^{-1}(G_m)\} \\ &= \liminf_{k \to \infty} \frac{1}{k} \log \mathbb{P}\{\zeta_k^m \in G_m\} \geq -I_m(G_m).\end{aligned}$$

It remains to note that $I_m(G_m) \leq I_m(\boldsymbol{\sigma} \circ p_m^{-1}) \leq \boldsymbol{I}(\boldsymbol{\sigma})$.

Step 3: Upper bound. Let $F \subset \mathcal{P}(\boldsymbol{X})$ be a closed subset. It suffices to prove that, if $\alpha < \boldsymbol{I}(F)$, then

$$\liminf_{k \to \infty} \frac{1}{k} \log \mathbb{P}\{\boldsymbol{\zeta}_k \in F\} \leq -\alpha. \tag{6.7}$$

Relation (6.6) implies that

$$\varnothing = F \cap \{\boldsymbol{I} \leq \alpha\} = \bigcap_{m=1}^{\infty} F \cap \big\{I_m\big(\boldsymbol{\sigma} \circ p_m^{-1}\big) \leq \alpha\big\}.$$

Since $F \cap \{\boldsymbol{I} \le \alpha\}$ is a compact set, it follows that one can find an integer $m \ge 1$ such that $F \cap \big\{I_m\big(\boldsymbol{\sigma} \circ p_m^{-1}\big) \le \alpha\big\} = \varnothing$. Denoting by F_m the image of F under the projection p_m, we conclude that $I_m(F_m) > \alpha$. Since $F \subset p_m^{-1}(F_m)$, using the LDP for ζ_k^m, we derive

$$\begin{aligned}\liminf_{k\to\infty} \frac{1}{k} \log \mathbb{P}\{\boldsymbol{\zeta}_k \in F\} &\le \liminf_{k\to\infty} \frac{1}{k} \log \mathbb{P}\{\boldsymbol{\zeta}_k \in p_m^{-1}(F_m)\} \\ &= \liminf_{k\to\infty} \frac{1}{k} \log \mathbb{P}\{\zeta_k^m \in F_m\} \le -I_m(F_m) < -\alpha.\end{aligned}$$

This completes the proof of (6.7) and of the theorem. □

Theorem 6.3 admits a simple generalisation to the case of uniform LDP. Namely, let us assume that we are given a sequence of random probability measures $\{\boldsymbol{\zeta}_k(y)\}$ on $\boldsymbol{X}$ depending on a parameter $y \in Y$, where Y is an arbitrary set. We shall say that $\{\boldsymbol{\zeta}_k(y)\}$ satisfies the *uniform LDP* with a good rate function $\boldsymbol{I} : \mathcal{P}(\boldsymbol{X}) \to [0, +\infty]$ if

$$\begin{aligned}-\boldsymbol{I}(\dot\Gamma) &\le \liminf_{k\to\infty} \frac{1}{k} \log \inf_{y\in Y} \mathbb{P}\{\boldsymbol{\zeta}_k(y) \in \Gamma\} \\ &\le \liminf_{k\to\infty} \frac{1}{k} \log \sup_{y\in Y} \mathbb{P}\{\boldsymbol{\zeta}_k(y) \in \Gamma\} \le -\boldsymbol{I}(\overline{\Gamma}),\end{aligned} \tag{6.8}$$

where $\Gamma \subset \mathcal{P}(\boldsymbol{X})$ is an arbitrary Borel subset. The proof of the following result literally repeats that of Theorem 6.3, and we omit it.

Theorem 6.4. *Suppose that for any integer $m \ge 1$ the sequence $\{\zeta_k^m(y), y \in Y\}$ satisfies the uniform LDP with a good rate function $I_m : \mathcal{P}(X^m) \to [0, +\infty]$. Then the uniform LDP holds for $\{\boldsymbol{\zeta}_k(y), y \in Y\}$ with the good rate function (6.5).*

References

[BD99] T. Bogenschütz and A. Doebler, *Large deviations in expanding random dynamical systems*, Discrete Contin. Dynam. Systems **5** (1999), no. 4, 805–812.

[BKL02] J. Bricmont, A. Kupiainen, and R. Lefevere, *Exponential mixing of the 2D stochastic Navier–Stokes dynamics*, Comm. Math. Phys. **230** (2002), no. 1, 87–132.

[BV92] A. V. Babin and M. I. Vishik, *Attractors of Evolution Equations*, North-Holland Publishing, Amsterdam, 1992.

[Cha96] M.-H. Chang, *Large deviation for Navier–Stokes equations with small stochastic perturbation*, Appl. Math. Comput. **76** (1996), no. 1, 65–93.

[CM10] I. Chueshov and A. Millet, *Stochastic 2D hydrodynamical type systems: Well posedness and large deviations*, Appl. Math. Optim. **61** (2010), no. 3, 379–420.

[CR04] S. Cerrai and M. Röckner, *Large deviations for stochastic reaction–diffusion systems with multiplicative noise and non-Lipschitz reaction term*, Annals Prob. **32** (2004), no. 1B, 1100–1139.

[CR05] ———, *Large deviations for invariant measures of stochastic reaction–diffusion systems with multiplicative noise and non-Lipschitz reaction term*, Ann. Inst. H. Poincaré Probab. Statist. **41** (2005), 69–105.

[DG87] D. A. Dawson and J. Gärtner, *Large deviations from the McKean-Vlasov limit for weakly interacting diffusions*, Stochastics **20** (1987), no. 4, 247–308.

[DS89] J.-D. Deuschel and D. W. Stroock, *Large Deviations*, Academic Press, Boston, 1989.

[DV75] M. D. Donsker and S. R. S. Varadhan, *Asymptotic evaluation of certain Markov process expectations for large time, I-II*, Comm. Pure Appl. Math. **28** (1975), 1–47, 279–301.

[DZ00] A. Dembo and O. Zeitouni, *Large Deviations Techniques and Applications*, Springer–Verlag, Berlin, 2000.

[EMS01] W. E, J. C. Mattingly, and Ya. Sinai, *Gibbsian dynamics and ergodicity for the stochastically forced Navier–Stokes equation*, Comm. Math. Phys. **224** (2001), no. 1, 83–106.

[FM95] F. Flandoli and B. Maslowski, *Ergodicity of the 2D Navier–Stokes equation under random perturbations*, Comm. Math. Phys. **172** (1995), no. 1, 119–141.

[Fre88] M. I. Freidlin, *Random perturbations of reaction–diffusion equations: The quasi deterministic approximation*, Trans. Amer. Math. Soc. **305** (1988), 665–697.

[FW84] M. I. Freidlin and A. D. Wentzell, *Random Perturbations of Dynamical Systems*, Springer, New York–Berlin, 1984.

[GC95] G. Gallavotti and E. G. D. Cohen, *Dynamical ensembles in stationary states*, J. Statist. Phys. **80** (1995), no. 5-6, 931–970.

[Gou07a] M. Gourcy, *A large deviation principle for 2D stochastic Navier–Stokes equation*, Stochastic Process. Appl. **117** (2007), no. 7, 904–927.

[Gou07b] ———, *Large deviation principle of occupation measure for a stochastic Burgers equation*, Ann. Inst. H. Poincaré Probab. Statist. **43** (2007), no. 4, 375–408.

[Kif90] Y. Kifer, *Large deviations in dynamical systems and stochastic processes*, Trans. Amer. Math. Soc. **321** (1990), no. 2, 505–524.

[KS00] S. Kuksin and A. Shirikyan, *Stochastic dissipative PDEs and Gibbs measures*, Comm. Math. Phys. **213** (2000), no. 2, 291–330.

[KS04] _____, *Randomly forced CGL equation: stationary measures and the inviscid limit*, J. Phys. A **37** (2004), no. 12, 3805–3822.

[KS12] _____, *Mathematics of Two-Dimensional Turbulence*, Cambridge University Press, Cambridge, 2012.

[LS06] A. Lasota and T. Szarek, *Lower bound technique in the theory of a stochastic differential equation*, J. Differential Equations **231** (2006), no. 2, 513–533.

[LY94] A. Lasota and J. A. Yorke, *Lower bound technique for Markov operators and iterated function systems*, Random Comput. Dynam. **2** (1994), no. 1, 41–77.

[Sow92a] R. Sowers, *Large deviations for a reaction–diffusion equation with non-Gaussian perturbation*, Ann. Probab. **20** (1992), 504–537.

[Sow92b] _____, *Large deviations for the invariant measure of a reaction–diffusion equation with non-Gaussian perturbations*, Probab. Theory Related Fields **92** (1992), 393–421.

[SS06] S. S. Sritharan and P. Sundar, *Large deviations for the two-dimensional Navier–Stokes equations with multiplicative noise*, Stochastic Process. Appl. **116** (2006), 1636–1659.

[Sza97] T. Szarek, *Markov operators acting on Polish spaces*, Ann. Polon. Math. **67** (1997), no. 3, 247–257.

[Wu01] L. Wu, *Large and moderate deviations and exponential convergence for stochastic damping Hamiltonian systems*, Stochastic Process. Appl. **91** (2001), no. 2, 205–238.