Electronic Journal of Probability

Electron. J. Probab. **23** (2018), no. 55, 1–28.
ISSN: 1083-6489 https://doi.org/10.1214/18-EJP177

Non-equilibrium steady states for networks of oscillators

Noé Cuneo* Jean-Pierre Eckmann† Martin Hairer‡
Luc Rey-Bellet§

Abstract

Non-equilibrium steady states for chains of oscillators (masses) connected by harmonic and anharmonic springs and interacting with heat baths at different temperatures have been the subject of several studies. In this paper, we show how some of the results extend to more complicated networks. We establish the existence and uniqueness of the non-equilibrium steady state, and show that the system converges to it at an exponential rate. The arguments are based on controllability and conditions on the potentials at infinity.

Keywords: non-equilibrium statistical mechanics; networks of oscillators; geometric ergodicity; Hörmander's condition; Lyapunov functions.
AMS MSC 2010: 82C05; 60H10; 34C15; 60B10.
Submitted to EJP on January 16, 2018, final version accepted on May 13, 2018.
Supersedes arXiv:1712.09413.

1 Introduction

The aim of this paper is to state and prove an extension of the results of [11, 31, 1] to the multidimensional case. We consider a network of masses connected with springs (interaction potentials), where some of the masses interact with stochastic heat baths which can have different temperatures. We also let each mass interact with a substrate through some pinning potential. We will show that under conditions spelled out in this paper, any such system has a unique *non-equilibrium stationary state* (invariant measure). We show, moreover, that the convergence to the steady state is exponential. The proof follows in principle the ideas of [11, 9], but the controllability argument uses the more general conditions of [14], and the compactness part relies on a Lyapunov function argument similar to [31, 1].

The new aspects of this paper are twofold: First, we deal with networks of springs connecting the masses, and not just with 1-dimensional chains. Second, we correct an

*Département de mathématiques, Université de Cergy-Pontoise, CNRS UMR 8088
†Département de Physique Théorique and Section de Mathématiques, University of Geneva
‡Department of Mathematics, Imperial College London
§Department of Mathematics and Statistics, University of Massachusetts Amherst

oversight of [31, 1] (see Remark 5.12) by a careful analysis of the interplay between the coupling potentials, which hold the system together, and the pinning potentials, which prevent it from "flying away". This will require decomposing the phase space into two regions, depending on whether the pinning forces or the interaction forces dominate. In the process, we also obtain sharper estimates on the rate of energy dissipation (see Remark 5.7).

The conditions on the system come in the following flavors:

C1 : The masses are sufficiently connected to the heat baths.
C2 : The interaction potentials are non-degenerate.
C3 : The potentials are homogeneous at infinity and coercive.
C4 : The limiting interaction forces are locally injective.
C5 : The interaction potentials grow at least as fast as the pinning potentials.

We will make C1–C5 precise in the next section. C1–C2 will be required to show the uniqueness of the steady state, and actually C1 will have to be more specific than what common sense would seem to dictate. C3–C5 will be further required for existence and exponential convergence.

As was shown in [11, 10], it is useful to assume that all potentials are quadratic (at least at infinity). These results have been extended in [9, 30, 1] to potentials of polynomial growth subject to C5.

Without C5, decoupling phenomena (related to "breathers") may lead to subexponential convergence to the invariant measure (and much more difficult proofs). In fact, the existence of the invariant measure when the pinning potentials grow faster than the interaction potentials has only been obtained for a chain of 3 masses so far (see the extensive discussion in [17]), and for some closely related chains of rotors, which correspond to the "infinite pinning" limit in a sense (see [6, 5, 7]).

The paper is organized as follows. In §2 we give the precise definitions of the conditions C1–C5 above and state the main result about existence and uniqueness of the invariant measure, and exponential convergence. The proof relies on two main ingredients: (1) Hörmander's bracket condition, and (2) a Lyapunov condition on the energy. In §3 we prove that these two ingredients lead to the desired result, and there we consider more general thermalized Hamiltonian systems (of which networks of oscillators are a special case). Finally, we check that under C1–C5, networks of oscillators indeed satisfy Hörmander's condition (§4) and the Lyapunov property (§5).

While the discussion in §3 is rather standard, the proofs in §4 and §5 are quite specific to our setup; the main technical difficulty there is that the heat baths do not act on all the oscillators directly, so that propagation within the network has to be carefully studied. This difficulty was already present, to a lesser extent, in the works on chains of oscillators mentioned above.

Finally, although we restrict ourselves to smooth potentials here, we mention that systems of particles with singular interactions (but with heat baths acting on all particles) have attracted significant attention lately (see for example [2, 3, 21, 13]).

2 Setup and results

We consider a finite set $\mathcal{G}$ of masses. We denote by $q_v \in \mathbf{R}^n$ and $p_v \in \mathbf{R}^n$ the position and momentum of each mass $v \in \mathcal{G}$ (we assume $n \geq 1$). The phase space is then $\Omega \equiv \mathbf{R}^{2|\mathcal{G}|n}$, and we write $z = (p, q) = ((p_v)_{v\in\mathcal{G}}, (q_v)_{v\in\mathcal{G}})$.

We then introduce a set $\mathcal{E} \subset \mathcal{G} \times \mathcal{G}$ of edges representing the springs, and consider Hamiltonians of the form

$$H(p, q) = \sum_{v\in\mathcal{G}} \Big(\frac{p_v^2}{2} + U_v(q_v)\Big) + \sum_{e\in\mathcal{E}} V_e(\delta q_e) \ , \tag{2.1}$$

where the functions U_v are *pinning* potentials, the functions V_e are *interaction* potentials, and where for $e = (v, v') \in \mathcal{E}$ we write $\delta q_e = q_{v'} - q_v \in \mathbf{R}^n$.

We view $(\mathcal{G}, \mathcal{E})$ as an undirected graph with no loop (*i.e.*, no edge of the kind (v, v)). Since the edges $e = (v, v')$ and $\bar{e} = (v', v)$ are identified, we also adopt the convention that $V_e(q_{v'} - q_v)$ and $V_{\bar{e}}(q_v - q_{v'})$ are equal and both express just one interaction, which appears only once in (2.1).

We now choose a subset $\mathcal{B} \subset \mathcal{G}$ of vertices where thermal baths act, and for every $b \in \mathcal{B}$ we assume that some temperature $T_b > 0$ and some coupling constant $\gamma_b > 0$ are given. For $v \notin \mathcal{B}$ we set, for convenience, $\gamma_v = T_v = 0$. With this notation, our model is described by the system of stochastic differential equations (one equation per $v \in \mathcal{G}$):

$$dq_v = p_v \, dt \ , \qquad dp_v = -\nabla_{q_v} H(p, q) \, dt - \gamma_v p_v \, dt + \sqrt{2T_v \gamma_v} \, dW_v(t) \ , \tag{2.2}$$

where the W_v are mutually independent standard n-dimensional Wiener processes. Note that for $v \notin \mathcal{B}$, the last two terms in (2.2) are absent. We denote by $z_t = (p_t, q_t)$ the solution of (2.2). For each fixed initial condition $z \in \Omega$, we denote by $\mathbf{P}_z$ the probability distribution of the solutions to (2.2), and by $\mathbf{E}_z$ the corresponding expectation. We also introduce the transition kernels $P_t(z, \cdot\,)$ defined for all $z \in \Omega$, $t \geq 0$, and all Borel sets $A \subset \Omega$ by

$$P_t(z, A) = \mathbf{P}_z\{z_t \in A\} \ . \tag{2.3}$$

The Langevin heat baths used in (2.2) are slightly simpler than those in [11, 9, 31]. There, the oscillators interact with some classical field theories which are initially Gibbs-distributed, and the (linear) coupling between the oscillators and the fields is chosen so that the latter can be integrated out. The resulting dynamics is similar to (2.2), but instead of directly acting on the momenta as in (2.2), the noise and dissipation act on some auxiliary variables which in turn interact with the momenta. The choice of Langevin heat baths (also made in [1]) is only for convenience, and the present analysis is easily transposed to the setup of [11, 9, 31].

We now make C1–C5 precise. We start with C1 in §2.1, which is in particular satisfied if the network is a chain with heat baths at both ends. In §2.2 and §2.3, we discuss C2–C5. An example of potentials satisfying C2–C5 that the reader might want to have in mind is[1] $V_e = (1 + \|\cdot\|^2)^{\ell_i/2}$ and $U_v = (1 + \|\cdot\|^2)^{\ell_p/2}$, where $\ell_i, \ell_p \in \mathbf{R}$ satisfy $\ell_i \geq \ell_p \geq 2$. (If ℓ_i and ℓ_p are even numbers subject to the same condition, then one may also take $V_e = \|\cdot\|^{\ell_i}$ and $U_v = \|\cdot\|^{\ell_p}$.)

2.1 Controllability through the springs

The following definition is useful: Let B be a subset of $\mathcal{G}$. We say that B is *nicely connected* to $v \in \mathcal{G} \setminus B$ if there exists a vertex $b \in B$ and an edge of the form $(b, v) \in \mathcal{E}$, and there is *no other edge* from b to $\mathcal{G} \setminus B$. We define $\mathcal{T}B$ as the union of B with its nicely connected vertices in $\mathcal{G} \setminus B$ (see Figure 1). We denote by $\mathcal{T}^2\mathcal{B}, \mathcal{T}^3\mathcal{B}, \ldots$ the iterates of this construction.

Definition 2.1. *Let $(\mathcal{G}, \mathcal{E})$ be as above. We say that $\mathcal{B} \subset \mathcal{G}$* controls $(\mathcal{G}, \mathcal{E})$ *if there exists $k \geq 1$ such that $\mathcal{T}^k\mathcal{B} = \mathcal{G}$.*

This allows us to make C1 precise as[2]

Condition C1. The graph is connected and $\mathcal{B}$ controls $(\mathcal{G}, \mathcal{E})$.

Remark 2.2. Note that connectedness is a trivial restriction, for if the graph is not connected the results apply to each connected component separately. Chains with heat baths at both ends (or even at just one end) obviously satisfy C1. So do some finite pieces

[1] Throughout the paper, $\|\cdot\|$ denotes the Euclidean norm.

[2] It was brought to our attention that the same condition appears in [8, Section 2.2].

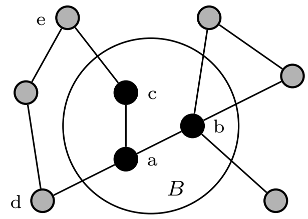

Figure 1: In this network, if $B = \{\mathrm{a}, \mathrm{b}, \mathrm{c}\}$, then $\mathcal{T}B = \{\mathrm{a}, \mathrm{b}, \mathrm{c}, \mathrm{d}, \mathrm{e}\}$.

of regular lattices, see Figure 2. As some examples in Figure 2 illustrate, controllability is, unfortunately, not a monotone property in $\mathcal{E}$: Adding edges, *i.e.*, more springs, will sometimes improve controllability, and sometimes destroy it. On the other hand, given $(\mathcal{G}, \mathcal{E})$, controllability is a monotone property in the set $\mathcal{B}$ of "initially controlled" nodes.

Remark 2.3. One always has the inequality $|\mathcal{T}^{k+1}\mathcal{B}| \le |\mathcal{T}^k\mathcal{B}| + |\mathcal{B}|$. Indeed, let B_k be the set of vertices in $\mathcal{T}^k\mathcal{B}$ that are connected to at least one other vertex in $\mathcal{G} \setminus \mathcal{T}^k\mathcal{B}$. It is then clear from the definition of $\mathcal{T}$ that $|\mathcal{T}^{k+1}\mathcal{B}| \le |\mathcal{T}^k\mathcal{B}| + |B_k|$. On the other hand, it follows from the definition of $\mathcal{T}$ that, for every "newly added" vertex v in $\mathcal{T}^{k+1}\mathcal{B} \setminus \mathcal{T}^k\mathcal{B}$, there must be at least one vertex w in B_k such that v is the only element in $\mathcal{G} \setminus \mathcal{T}^k\mathcal{B}$ that is connected to w. As a consequence, $|B_{k+1}| \le |B_k| \le |\mathcal{B}|$ for every k, from which the claim follows at once.

In a way, this remark says that the system is effectively almost 1-dimensional with respect to the propagation of information. No point in $\mathcal{B}$ and no point in $\mathcal{T}^k\mathcal{B}$ will ever control more than one new point as one iterates from k to $k+1$ above.

Another criterion for the controllability of networks of interacting oscillators was introduced in [4]. While the results in [4] allow in some cases to control networks with more general topologies, in particular some which do not satisfy Condition C1, they only apply to strictly anharmonic polynomial potentials in 1D ($n = 1$).

2.2 Non-degenerate potentials

We now discuss the conditions on the potentials V_e. The attentive reader will note that, in fact, the non-degeneracy conditions below are not necessary on all the links, but only on those which are needed for Condition C1 to hold. This means, for example, that in Figure 2, the potentials associated with the "vertical" springs may be degenerate. We will not deal with this any further, and make the assumptions on all V_e.

Given a multi-index $\alpha = (\alpha_1, \dots, \alpha_n)$ of non-negative integers, we set $|\alpha| = \sum_{i=1}^n \alpha_i$, and define D^α as the differential operator with α_i derivatives in the i^{th} direction of $\mathbf{R}^n$. Given a potential $V\colon \mathbf{R}^n \to \mathbf{R}$, (*i.e.*, any of the V_e) we introduce the following notion of non-degeneracy [31]. The idea is that the V_e do not have "infinitely flat" pieces.

Definition 2.4. *A smooth potential* $V : \mathbf{R}^n \to \mathbf{R}$ *is* non-degenerate *if there exists an* $\ell < \infty$ *such that the set of derivatives*

$$\{D^\alpha \nabla V(x) \ : \ 1 \le |\alpha| \le \ell\}$$

spans $\mathbf{R}^n$ *for* every $x \in \mathbf{R}^n$.

We now have the following precise version of C2:

Condition C2. The interaction potentials V_e are non-degenerate.

Example 2.5. Any potential of the form $V(x) = \|x\|^r$ with $r = 2, 4, 6, \dots$ is non-degenerate. The same is true of $V(x) = (1 + \|x\|^2)^{r/2}$ with any real number $r > 0$. On the contrary, if $\|x\|$ is replaced by $|x_1|$ here, then the resulting potential is degenerate (unless $n = 1$).

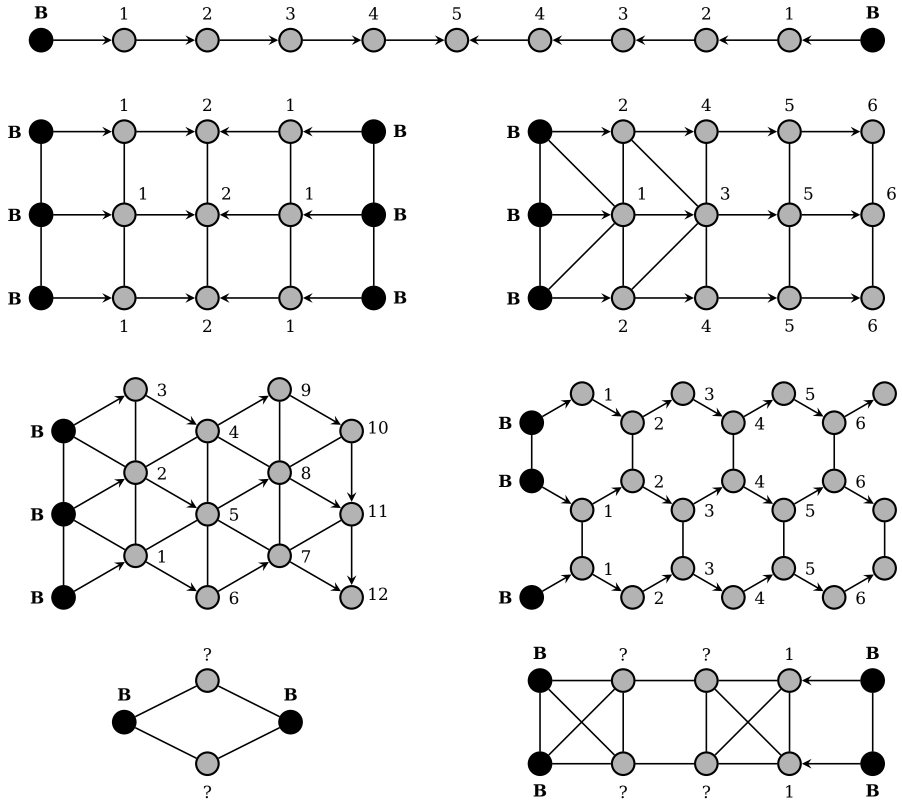

Figure 2: The elements of $\mathcal{B}$ are labeled by “B”. The numerical label k indicates that the vertex is in $\mathcal{T}^k\mathcal{B}$ but not in $\mathcal{T}^{k-1}\mathcal{B}$ (with $\mathcal{T}^0\mathcal{B} = \mathcal{B}$), and the uncontrollable elements are labeled by “?”. The arrows indicate the growth of $\mathcal{T}^k\mathcal{B}$ as a function of k. The top five networks are controlled by $\mathcal{B}$, while the bottom two are not. The example in the lower left corner was used in [12].

Remark 2.6. The condition in Definition 2.4 allows for controllability in the following sense: Consider a *given* continuous trajectory $\bar{q}\colon [0,1] \to \mathbf{R}^n$ and the problem

$$\dot{p}_f(t) = -\nabla V\big(\bar{q}(t) + f(t)\big) \tag{2.4}$$

with $p_f(0) = p_*$. If V is non-degenerate, then the set of solutions $p_f(1)$ of (2.4) at time 1, as f is varied over all smooth functions with $\sup_{t\le 1} |f(t)| \le 1$, contains an open (and in particular “full-dimensional”) set.

2.3 Nearly homogeneous potentials

One of the difficulties with models of the type (2.1), (2.2) is to show the existence of a non-equilibrium steady state. As was demonstrated in [17, 15], this can be highly non-trivial, and even with “nice” potentials, there are situations where the convergence to the steady state can be arbitrarily slow.

For the purpose of proving the existence of the steady state, a convenient class of interactions is given by potentials that behave at infinity like homogeneous functions. We say that a function $\Psi\colon \mathbf{R}^n \to \mathbf{R}$ is homogeneous of degree[3] $r \ge 2$ if $\Psi(\lambda x) = \lambda^r \Psi(x)$

[3] The degree r is not assumed to be an integer. The restriction $r \ge 2$ is required for some of the results below.

for every $\lambda > 0$ and every $x \in \mathbf{R}^n \setminus \{0\}$. With this notion at hand, we give the following definition, which is slightly weaker than the one in [30]:

Definition 2.7. *A smooth function $V : \mathbf{R}^n \to \mathbf{R}$ is said to be* nearly homogeneous *of degree r if there exists a homogeneous (of degree r), differentiable function $V_\infty : \mathbf{R}^n \to \mathbf{R}$ such that ∇V_∞ is locally Lipschitz, and such that for all $0 \le |\alpha| \le 1$,*

$$\lim_{\lambda\to\infty} \sup_{\|x\|=1} \left| \frac{(D^\alpha V)(\lambda x)}{\lambda^{r-|\alpha|}} - D^\alpha V_\infty(x) \right| = 0 \;.$$

Example 2.8. If $V(x) = \|x\|^r$ with $r = 2, 4, 6, \ldots$, then V is nearly homogeneous. Moreover, for any real number $r \ge 2$, the potential $V(x) = (1+\|x\|^2)^{r/2}$ is nearly homogeneous. In both cases, $V_\infty(x) = \|x\|^r$.

Remark 2.9. It is easy to see that nearly homogeneous functions (of degree $r \ge 2$) also satisfy some derived properties, for $0 \le |\alpha| \le 1$:

(i) $\lim_{\|x\|\to\infty} \|x\|^{|\alpha|-r} \left(D^\alpha V(x) - D^\alpha V_\infty(x)\right) = 0$.

(ii) $|D^\alpha V(x)| \le C_V(1 + \|x\|^{r-|\alpha|})$, for some $C_V > 0$.

(iii) $\lim_{\lambda\to\infty} \sup_{x\in K} \left(\lambda^{|\alpha|-r}(D^\alpha V)(\lambda x) - D^\alpha V_\infty(x)\right) = 0$, for every compact set K.

(iv) If $\inf_{\|x\|=1} V_\infty(x) > 0$, then $V(x) \ge C'_V\|x\|^r$, for some $C'_V > 0$ when $\|x\|$ is large enough.

We can now define C3–C5 properly as follows.

Condition C3. The potentials U_v are nearly homogeneous of degree $\ell_p \ge 2$ with limiting functions $U_{v,\infty}$, and the potentials V_e are nearly homogeneous of degree $\ell_i \ge 2$ with limiting functions $V_{e,\infty}$. Moreover, the limiting potentials are coercive, *i.e.*, $\inf_{\|x\|=1} V_{e,\infty}(x) > 0$ and $\inf_{\|x\|=1} U_{v,\infty}(x) > 0$.

Condition C4. The limiting interaction forces $-\nabla V_{e,\infty}$ are locally injective in the sense that for each $e \in \mathcal{E}$ and each $x \in \mathbf{R}^n$, we have $\nabla V_{e,\infty}(x') \ne \nabla V_{e,\infty}(x)$ for all x' in a neighborhood of x.

Condition C5. The interaction and pinning powers satisfy $\ell_i \ge \ell_p$.

Note that Conditions C2 and C4 are not comparable: the former guarantees that the forces $-\nabla V_e$ are locally *surjective* in a sense, and the latter guarantees that the limiting forces $-\nabla V_{e,\infty}$ are locally *injective*.

For example, consider a smooth homogeneous function $V : \mathbf{R}^3 \to \mathbf{R}$ given by $x^4/4 + y^2z^2/2$ on the set $M = \{(x,y,z) \in \mathbf{R}^3 \;:\; z^2 + y^2 \le x^2/10\}$. Then $\nabla V(x,y,z) = (x^3, yz^2, y^2z)$, and obviously $\{D^\alpha \nabla V(x) \;:\; 1 \le |\alpha| \le 3\}$ spans $\mathbf{R}^3$ for all $(x,y,z) \in M$. However, $\nabla V(x,y,0) = (x^3,0,0)$, and thus ∇V is not locally injective.

There are specific systems for which Condition C4 is not actually required, and others for which it is, as we illustrate in Remarks 5.15 and 5.16.

Remark 2.10. Condition C4 holds for example if the $V_{e,\infty}$ are strictly convex. In particular if $n = 1$, then the $V_{e,\infty}$ are automatically strictly convex, since they are homogeneous of degree $\ell_i \ge 2$ and coercive.

Remark 2.11. The requirement that all interaction potentials have the *same* degree ℓ_i is crucial. Indeed, if one of the interactions in the bulk of the network (*i.e.*, involving two oscillators in $\mathcal{G} \setminus \mathcal{B}$) has a higher degree than the others, the system may find itself in a regime where the two corresponding oscillators oscillate in phase opposition and with a frequency much higher than the other natural frequencies of the system, leading to a decoupling phenomenon comparable to the situation in [17]. This is again expected to lead to subgeometric convergence to the invariant measure and much more involved proofs.

Remark 2.12. As will be clear from the proofs in §5, it is actually not necessary for all the limiting pinning potentials $U_{v,\infty}$ to be coercive (or even to be non-zero). In fact, we only need the quantity defined in (5.41) to be coercive.

Without loss of generality, we also assume that the potentials U_v and V_e are non-negative (by the coercivity condition above, this is always achievable by adding a constant).

2.4 Main result

Given the definitions of §2.1–2.3, we can now state the main result. In order to emphasize the role of each assumption, we introduce the following (very weak) auxiliary condition.

Condition CA. The Hamiltonian H has compact level sets (*i.e.*, the set $\{z : H(z) \leq K\}$ is compact for each $K > 0$), and there exists some $\beta > 0$ such that the function $\exp\bigl(-\beta H\bigr)$ is integrable on Ω.

Condition CA follows immediately from Condition C3 (one can choose any $\beta > 0$).

Theorem 2.13. *The following holds.*

1. *Under Conditions C1, C2 and CA, the system (2.2) admits at most one invariant measure, and if it exists, it has a smooth density with respect to Lebesgue measure.*
2. *Under Conditions C1, C3, C4 and C5, the system (2.2) admits a least one invariant measure, and $e^{\vartheta H}$ is integrable with respect to it for all $0 < \vartheta < 1/T_{\max}$, with $T_{\max} = \max\{T_b : b \in \mathcal{B}\}$.*
3. *Finally, assuming Conditions C1–C5, the system (2.2) admits a unique invariant measure $\mu_\star$. Moreover, for all $0 < \vartheta < T_{\max}$, there are constants $C, c > 0$ such that for every initial condition $z = (p, q) \in \Omega$ and all $t \geq 0$,*

$$\sup_{f \in C(\Omega)\,:\,|f| \leq e^{\vartheta H}} \left| \mathbf{E}_z f(z_t) - \int f \, d\mu_\star \right| \leq C e^{\vartheta H(z) - ct} \; . \tag{2.5}$$

This theorem is a special case of Theorem 3.1 below, as we will show.

3 A general result about thermalized Hamiltonian systems

In this section, we prove a version of Theorem 2.13 which applies to more general thermalized Hamiltonian systems subject to two assumptions H1 and H2 (see below). As we show in §4 and §5, these assumptions follow from Conditions C1–C5. Although the material discussed in this section is mostly standard (see for example [25]), we provide a complete exposition relying on the version of Harris' ergodic theorem proved in [18]. We hope that by considering more general Hamiltonian systems and conditions in this section, the proofs will be both easier to read and useful beyond the scope of this paper.

The setup is as in §2, except that we do not assume that the set of masses $\mathcal{G}$ has the structure of a graph and that the Hamiltonian has the form (2.1). More precisely, we study the SDE

$$dq_v = p_v \, dt \; , \qquad dp_v = -\nabla_{q_v} H(p,q) \, dt - \gamma_v p_v \, dt + \sqrt{2 T_v \gamma_v} \, dW_v(t) \; , \tag{3.1}$$

where the friction constants γ_v, the temperatures T_v and the set $\mathcal{B} \subset \mathcal{G}$ are as in §2, and where the Hamiltonian is given by

$$H(p,q) = \sum_{v \in \mathcal{G}} \frac{p_v^2}{2} + U(q) \; ,$$

for some arbitrary smooth, non-negative potential U on $\mathbf{R}^{n|\mathcal{G}|}$.

We also assume throughout this section that Condition CA holds, *i.e.*, that H has compact level sets and that $\exp\bigl(-\beta H\bigr)$ is integrable on Ω for some $\beta > 0$.

We define the semigroup $(\mathcal{P}^t)_{t\geq 0}$ acting on the space of bounded measurable functions on Ω by $\mathcal{P}^t f(z) = \mathbf{E}_z f(z_t) = \int_\Omega f(z')P_t(z, dz')$. We also fix $0 < \vartheta < 1/T_{\max}$, with $T_{\max} = \max\{T_b : b \in \mathcal{B}\}$ as in Theorem 2.13. We let moreover

$$V = e^{\vartheta H} \ .$$

The solutions to (3.1) form a Markov process whose generator $\mathcal{L}$ is

$$\mathcal{L} = X_0 + \sum_{b\in\mathcal{B}} \sum_{i=1,\dots,n} X_{b,i}^2 \ , \tag{3.2}$$

where $X_{b,i} = \sqrt{T_b\gamma_b}\partial_{p_b^i}$ and

$$X_0 = \sum_{v\in\mathcal{G}} \Bigl(p_v \cdot \nabla_{q_v} - \nabla_{q_v} U(q) \cdot \nabla_{p_v} - \gamma_v p_v \cdot \nabla_{p_v}\Bigr) \ .$$

From now on, we will view X_0 and the $X_{b,i}$ interchangeably as first-order differential operators and as vector fields on Ω.

With $C_* = \vartheta \sum_{b\in\mathcal{B}} \gamma_b T_b$, we obtain

$$\mathcal{L}V = \sum_{b\in\mathcal{B}} \vartheta\gamma_b \left([\vartheta T_b - 1]p_b^2 + T_b\right) e^{\vartheta H} \leq C_* V \ . \tag{3.3}$$

Since H, and hence V, have compact level sets by assumption, the process admits strong solutions that are continuous and defined for all $t \geq 0$ (almost surely), the strong Markov property is satisfied, and for all $t \geq 0$ we have

$$\mathcal{P}^t V \leq e^{C_* t} V \tag{3.4}$$

(see for example [20, Theorem 3.5], [29], and [28, Theorem III.3.1] for the strong Markov claim).

We now introduce Hörmander's celebrated "Lie bracket condition" [23]. Define a family of vector fields $\mathcal{A}_0$ by $\mathcal{A}_0 = \{X_{b,i} \,:\, b \in \mathcal{B}, i = 1, \dots, n\}$ and then, recursively,

$$\mathcal{A}_{k+1} = \mathcal{A}_k \cup \bigl\{[X, Y] \,:\, X \in \mathcal{A}_k\, , Y \in \mathcal{A}_0 \cup \{X_0\}\bigr\} \ ,$$

where $[X, Y]$ denotes the Lie bracket (commutator) of X and Y. With this notation at hand, we formulate

Condition H1. The operator $\mathcal{L}$ defined in (3.2) satisfies Hörmander's bracket condition, *i.e.*, for every $z \in \Omega$, there exists an integer $k > 0$ such that the linear span of $\{Y(z) : Y \in \mathcal{A}_k\}$ is all of Ω.

Condition H1 is sufficient (and "almost necessary") for $\partial_t - \mathcal{L}$ to be hypoelliptic, so that the semigroup associated to (3.1) has a smoothing effect (see Proposition 3.2 below). We note that the requirement in Condition H1 is made for all $z \in \Omega$; see for example [27] for an argument which only requires Hörmander's condition to hold at one point, but which is specific to quasi-harmonic systems whose harmonic part is subject to Kalman's controllability condition.

Next, we introduce a Lyapunov condition, which will be crucial in order to obtain the existence of an invariant measure and the exponential convergence (2.5).

Condition H2. There exists $t_* > 0$ and $\varkappa \in (0,1)$ such that[4]

$$\mathcal{P}^{t_*} V \le \varkappa V + c\mathbf{1}_K \;, \tag{3.5}$$

where $c > 0$ is a constant and K is a compact set.

In §4, we show that for the original system (2.2), Conditions C1 and C2 imply Condition H1, and in §5 we show that Conditions C1, C3, C4 and C5 imply Condition H2. With this in mind, Theorem 2.13 is a special case of

Theorem 3.1. *The following holds (recall that Condition CA is assumed throughout this section).*

1. *Under Condition H1, the system (3.1) admits at most one invariant measure, and if it exists, it has a smooth density with respect to Lebesgue measure.*
2. *Assuming Condition H2, the system (3.1) admits a least one invariant measure, and V is integrable with respect to it.*
3. *Finally, assuming Conditions H1 and H2, the system (3.1) admits a unique invariant measure $\mu_\star$, and the exponential convergence in* (2.5) *holds.*

Proof. The three parts of the theorem are proved in Propositions 3.3, 3.7 and 3.8 below. □

3.1 Controllability and uniqueness

The following consequence of Hörmander's condition is well known [23] (see [29, Section 7], [16], and [32, Section 7.4] for introductions).

Proposition 3.2. *Assume Condition H1. Then the transition kernel in* (2.3) *can be written as $P_t(z, dz') = p_t(z, z')dz'$, where the map $(t, z, z') \mapsto p_t(z, z')$ is smooth on $(0,\infty)\times\Omega\times\Omega$. In particular, the process is strong Feller. Finally, every invariant measure has a smooth density with respect to Lebesgue measure on Ω.*

We now prove the following "accessibility" result (see also [6, Section 5.2.1] for another variant of this argument).

Proposition 3.3. *Assume Condition H1. Then the system (3.1) admits at most one invariant measure, and for every non-empty open set $\mathcal{U} \subset \Omega$ and all $z \in \Omega$, we have $\sup_{t>0} P_t(z,\mathcal{U}) > 0$.*

Proof. The argument follows the same lines as the reasoning first given in [14], see also [24]. Take $\beta > 0$ as in Condition CA and consider instead of (3.1) the modified equation

$$dq_v = p_v\, dt \;, \qquad dp_v = -\nabla_{q_v} H(p,q)\, dt - \gamma_v p_v\, dt + \sqrt{2\gamma_v \beta^{-1}}\, dW_v(t) \;. \tag{3.6}$$

The only difference is that all the temperatures have been replaced by $1/\beta$ (we still have $\gamma_v = 0$ for all $v \notin \mathcal{B}$). By the same argument as above, the solutions to (3.1) almost surely exist for all times. It is well known that the measure

$$d\mu_\beta = \frac{1}{Z} e^{-\beta H(p,q)}\, dp\, dq$$

is invariant for (3.6), and by Condition CA, one can choose $Z > 0$ so that μ_β is a probability measure. (The invariance of μ_β can be seen by checking that $\mathcal{L}^* e^{-\beta H} = 0$, where $\mathcal{L}^*$ is the formal adjoint of the generator of (3.6).)

We next show that μ_β is the only invariant probability measure for (3.6). It is easy to show that, as a consequence of Proposition 3.2, the map $z \mapsto \overline{P}_t(z, \cdot\,)$ is continuous

[4] Here and below, $\mathbf{1}_K$ denotes the characteristic function of the set K.

in the total variation topology, where $\overline{P}_t$ denotes the transition probabilities for (3.6). Since distinct ergodic invariant probability measures for (3.6) are mutually singular by Birkhoff's ergodic theorem, this immediately implies that if ν is an ergodic invariant measure for (3.6) and $z \in \operatorname{supp} \nu$, then there exists a neighborhood $\mathcal{U}_z$ of z such that $\mathcal{U}_z \cap \operatorname{supp} \bar{\nu} = \emptyset$ for *every other* ergodic invariant measure $\bar{\nu}$.

As a consequence, let us choose some (there exists at least one) ergodic invariant measure ν of (3.6). Assuming by contradiction that ν is not unique, we have $\operatorname{supp} \nu \neq \Omega$. As a consequence, setting $\mathcal{V} = \bigcup_{z \in \operatorname{supp} \nu} (\mathcal{U}_z \setminus \operatorname{supp} \nu)$, we have constructed a non-empty open set $\mathcal{V}$ such that $\mathcal{V} \cap \operatorname{supp} \bar{\nu} = \emptyset$ for *every* ergodic invariant measure $\bar{\nu}$ of (3.6) and therefore, by the ergodic representation theorem, for *every* invariant measure $\bar{\nu}$. (We must have $\mathcal{V} \neq \emptyset$ for otherwise $\operatorname{supp} \nu$ would be both open and closed, which cannot be.) However, $\operatorname{supp} \mu_\beta = \Omega$, thus yielding a contradiction.

Returning to our main line of argument, since μ_β is the unique invariant probability measure for (3.6), it must be ergodic. Since μ_β has full support, it then follows from Birkhoff's ergodic theorem that for every open set $\mathcal{U}$ and Lebesgue-almost every initial condition $z \in \Omega$, we have $\sup_{t>0} \overline{P}_t(z, \mathcal{U}) > 0$. An easy application of the Chapman-Kolmogorov equation, using the smoothness of the transition probabilities, shows that this actually holds for every $z \in \Omega$.[5] The conclusion of the proposition thus holds for (3.6). We now return to (3.1).

The key is that for each $z \in \Omega$ and $t \geq 0$, the transition probabilities $\overline{P}_t(z, \cdot\,)$ for (3.6) and $P_t(z, \cdot\,)$ for (3.1) are equivalent, since the two stochastic differential equations differ only by the scaling of the Brownian motions. Thus, we indeed have $\sup_{t>0} P_t(z, \mathcal{U}) > 0$ for all $z \in \Omega$ and every non-empty open set $\mathcal{U} \subset \Omega$. Assume now by contradiction that (3.1) admits more than one invariant probability measure. Then by the ergodic decomposition theorem there exist two distinct ergodic measures, which then have distinct supports S_1 and S_2. By smoothness, there exists a non-empty open set $\mathcal{U} \subset S_2$, and by taking $z \in S_1$ we find $\sup_{t>0} P_t(z, \mathcal{U}) = 0$, which is a contradiction. □

Although this will not be needed, we state the following corollary, which follows from the Stroock–Varadhan support theorem (see [33], and [22, Theorem 5.b] for an extension to case of unbounded coefficients).

Corollary 3.4. *Assume Condition H1. Then, for any starting point $z_0 \in \Omega$ and any non-empty open set $\mathcal{U} \subset \Omega$, there exists a time $t > 0$ and smooth controls $u_b \colon [0, t] \to \mathbf{R}^n$ for $b \in \mathcal{B}$ such that the solution at time t to*[6]

$$\dot{q}_v = p_v \;, \qquad \dot{p}_v = -\nabla_{q_v} H(p, q) - \gamma_v p_v + \mathbf{1}_{\mathcal{B}}(v) u_v(t) \;, \qquad v \in \mathcal{G} \;,$$

with initial condition z_0, lies in $\mathcal{U}$.

Remark 3.5. In the case of chains of oscillators, a stronger controllability argument is used in [10]. The argument given above is "softer". As a consequence, it applies to a larger class of Langevin equations, at the expense of having less explicit control. The argument in [10] actually implies that, in the statement of Proposition 3.3, the quantity $P_t(z, \mathcal{U})$ is positive for all $t > 0$.

3.2 Minorization

The next proposition shows that every compact set is *small* in the terminology of [26]. In fact, we show that for each given compact set C, the minorization condition holds for *all* large enough t. In the proof, $p_t(\,\cdot\,, \cdot\,)$ is as in Proposition 3.2.

[5] One can also use that the strong Feller property implies that Birkhoff's ergodic theorem holds for every initial condition in the support of the invariant measure [19, Theorem 4.10].

[6] The same is true without the dissipative terms $-\gamma_v p_v$, since they can be absorbed into the controls u_v (recall that $\gamma_v = 0$ when $v \notin \mathcal{B}$).

Proposition 3.6. *Assume Condition H1. Then, for every compact set C, there exists a time t_C such that for all $t \geq t_C$, there exists a non-negative and non-trivial measure ν (which may depend on t) such that $P_t(z, \cdot\,) \geq \nu$ for all $z \in C$.*

Proof. We start by showing that there exists $z_* \in \Omega$ such that for all $z \in \Omega$, there are $t_\sharp(z)$ and $\delta_z > 0$ satisfying

$$p_t(z', z_*) > 0 \qquad \text{for all } t \geq t_\sharp(z) \text{ and all } z' \in B(z, \delta_z)\ . \tag{3.7}$$

First, pick any $z_0 \in \Omega$. We now fix any z_* such that $p_1(z_0, z_*) > 0$. By continuity, there exists $\delta > 0$ such that $\inf_{z \in B(z_0,\delta)} p_1(z, z_*) > 0$. By Proposition 3.3, there exists for each $z \in \Omega$ some $t_0(z)$ such that $P_{t_0(z)}(z, B(z_0, \delta)) > 0$. It then follows from the semigroup property that $p_{t_1(z)}(z, z_*) > 0$ with $t_1(z) = t_0(z) + 1$. Using continuity again, we can choose $\delta_z > 0$ so that

$$p_{t_1(z)}(z', z_*) > 0 \qquad \text{for all } z' \in B(z, \delta_z)\ . \tag{3.8}$$

We now show that there exists $t_2 > 0$ such that

$$p_t(z_*, z_*) > 0 \qquad \text{for all } t \geq t_2\ . \tag{3.9}$$

Since $p_{t_1(z_*)}(z_*, z_*) > 0$, continuity with respect to time implies that for some $\Delta > 0$ small enough, we have $p_t(z_*, z_*) > 0$ for all $t \in [t_1(z_*), t_1(z_*) + \Delta]$. But then the same holds for all $t \in [nt_1(z_*), nt_1(z_*) + n\Delta]$, $n \in \mathbf{N}$. Thus (3.9) holds with $t_2 = n_* t_1(z_*)$ for any integer $n_* \geq t_1(z_*)/\Delta$. Using (3.8), (3.9) and the semigroup property yields (3.7) with $t_\sharp(z) = t_1(z) + t_2$.

We now prove the main claim. Let C be a compact set. The balls $\{B(z, \delta_z) : z \in C\}$ form an open cover of C, and by compactness we can extract a finite subcover, yielding a maximum time t_C such that $p_t(z, z_*) > 0$ for all $z \in C$ and all $t \geq t_C$. For any such t, since $p_t(\,\cdot\,,\cdot\,)$ is continuous on Ω^2 and C is compact, the result follows with $d\nu = \varepsilon \mathbf{1}_{B(z_*, r)} dz$ for small enough $\varepsilon, r > 0$. □

3.3 Existence of an invariant measure and exponential convergence

As an elementary consequence of Condition H2, we find that

$$\mathcal{P}^{nt_*} V \leq \varkappa^n V + c \sum_{i=0}^{\infty} \varkappa^i \quad \text{for all } n \in \mathbf{N}\ .$$

From this and (3.4), we obtain that

$$\mathcal{P}^t V \leq c_1 + c_2 \rho^t V \quad \text{for all } t \geq 0\ , \tag{3.10}$$

with $c_1, c_2 > 0$ and $\rho = \varkappa^{1/t_*} \in (0, 1)$. In particular, since V has compact level sets, this implies that for any $z \in \Omega$, the family of probability measures $(P_t(z, \cdot\,))_{t \geq 0}$ is tight. Since the process is Feller, the standard Krylov–Bogolyubov construction then implies that for some sequence t_k increasing to infinity, $\frac{1}{t_k} \int_0^{t_k} P_s(z, \cdot\,) ds$ converges weakly to some measure which is invariant, and with respect to which V is integrable. We thus obtain

Proposition 3.7. *Under Condition H2, the process admits an invariant measure $\mu_\star$, and V is integrable with respect to μ_*.*

Assuming in addition Condition H1 implies that $\mu_\star$ is unique (Proposition 3.3), and we now prove exponential convergence.

Proposition 3.8. *Under Conditions H1 and H2, the exponential convergence in* (2.5) *holds.*

Proof. We will apply the main result of [18] to the discrete-time semigroup $(\mathcal{P}^{nt_0})_{n=0,1,2,\ldots}$, for some large enough $t_0 > 0$. Let first $R = 2c_1/(1-\rho)$. Here c_1, c_2 and ρ are as in (3.10). We then define the compact set $C = \{z : V(z) \le R\}$. We choose now $t_0 \ge t_C$ with the t_C from Proposition 3.6, and large enough so that $c_2\rho^{t_0} < \rho$. It follows that $R > 2c_1/(1-c_2\rho^{t_0})$, so that by (3.10) the main result of [18] applies to $(\mathcal{P}^{nt_0})_{n=0,1,2,\ldots}$. We obtain[7] that for some $C_0, c_0 > 0$ and all $z \in \Omega$,

$$\sup_{f\in C(\Omega)\,:\,|f|\le V} \left|\mathbf{E}_z f(z_{nt_0}) - \int f d\mu_\star\right| \le C_0 V(z) e^{-c_0 n t_0} \quad \text{for all } n \in \mathbf{N}\ . \tag{3.11}$$

For $|f| \le V$, we define $g(z,t) = \mathbf{E}_z f(z_t) - \int f d\mu_\star$. Decomposing $t = nt_0 + r$ with $n \in \mathbf{N}$ and $r \in [0, t_0)$, we obtain from the Markov property that

$$|g(z,t)| = |\mathbf{E}_z g(z_r, nt_0)| \le C_0 e^{-c_0 n t_0} \mathbf{E}_z V(z_r) \le C_0 e^{C_* t_0 - c_0 n t_0} V(z)\ ,$$

where we have also used (3.4). This immediately implies (2.5) for some $C, c > 0$, and thus the proof is complete. □

4 Hypoellipticity

In this section, we prove

Proposition 4.1. *Under Conditions C1 and C2, the system (2.2) satisfies Condition H1.*

Proof. For the system (2.2), the vector field X_0 in the decomposition (3.2) reads

$$\begin{aligned} X_0 = \sum_{v\in\mathcal{G}} &\Big(p_v \cdot \nabla_{q_v} - \nabla U_v(q_v)\cdot\nabla_{p_v} - \gamma_v p_v \cdot \nabla_{p_v}\Big) \\ &- \sum_{(u,v)\in\mathcal{E}} \nabla V_{(u,v)}(q_v - q_u)\cdot\big(\nabla_{p_v} - \nabla_{p_u}\big)\ . \end{aligned}$$

We will actually prove the following statement, which implies Condition H1. Let $\bar X_0 = \partial_t - X_0$ and set $\mathcal{M}_0 = \{\bar X_0\} \cup \{X_{b,i} : b \in \mathcal{B}, i = 1,\ldots,n\}$, which we view as a family of smooth vector fields on $\mathbf{R}^{1+2n|\mathcal{G}|}$. Denote by $\mathcal{M}$ the smallest set of vector fields containing $\mathcal{M}_0$ that is closed under Lie brackets and multiplication by smooth functions.

We will show that ∂_t, as well as ∇_{p_v} and ∇_{q_v} for every $v \in \mathcal{G}$, all belong to $\mathcal{M}$. Since $\bar X_0 \in \mathcal{M}$, it is sufficient to prove the claim about the ∇_{p_v} and ∇_{q_v}. (Here and below, what we mean by $\nabla_{p_v} \in \mathcal{M}$ is that $\partial_{p_v^i} \in \mathcal{M}$ for all $i = 1,\ldots,n$, and similarly for ∇_{q_v}.)

Note first that, by the definition of $X_{b,i}$ and $\mathcal{M}_0$, we have $\nabla_{p_b} \in \mathcal{M}$ for all $b \in \mathcal{B}$. Furthermore, since

$$[\partial_{p_v^i}, \bar X_0] = -\partial_{q_v^i} + \gamma_v \partial_{p_v^i}$$

for all $v \in \mathcal{G}$, it follows that one has the implication

$$\nabla_{p_v} \in \mathcal{M} \quad\Rightarrow\quad \nabla_{q_v} \in \mathcal{M}\ .$$

By the definition of the notion of $\mathcal{B}$ controlling $\mathcal{G}$, the claim now follows if we can show that, for any set $\mathcal{B}' \subset \mathcal{G}$, one has the implication

$$\nabla_{q_b}, \nabla_{p_b} \in \mathcal{M} \text{ for all } b \in \mathcal{B}' \quad\Rightarrow\quad \nabla_{p_v} \in \mathcal{M} \text{ for all } v \in \mathcal{T}\mathcal{B}'\ .$$

[7] Another way to obtain (3.11) with $t_0 = t_*$ is to use [26, Theorem 15.0.1]. Indeed, (3.7) implies that the process is aperiodic, Condition H2 provides the required drift condition, and by Proposition 3.6 the compact set K in (3.5) is *small* (and hence *petite*). An alternative proof of convergence using quasi-compactness of the semigroup can be found in [29].

Assume therefore that $\mathcal{B}'$ is such that $\nabla_{q_b}, \nabla_{p_b}$ are in $\mathcal{M}$ for all $b \in \mathcal{B}'$. Note that, for all $i \in \{1, \ldots, n\}$ and every $b \in \mathcal{B}'$,

$$[\partial_{q_b^i}, \bar{X}_0] = \big(\partial_i \nabla U_b\big)(q_b) \cdot \nabla_{p_b} - \sum_{e=(b,v)\in\mathcal{E}_b} \big(\partial_i \nabla V_e\big)(\delta q_e) \cdot \big(\nabla_{p_v} - \nabla_{p_b}\big) \;, \tag{4.1}$$

where we denote by $\mathcal{E}_b$ the subset of those edges in $\mathcal{E}$ that are of the form (b,v) for some $v \in \mathcal{G}$. Fix now $v \in \mathcal{T}\mathcal{B}' \setminus \mathcal{B}'$. By the definition of $\mathcal{T}\mathcal{B}'$, there exists then $b \in \mathcal{B}'$ such that $(b,v) \in \mathcal{E}_b$ and, for every other w for which (b,w) is in $\mathcal{E}_b$, one has $w \in \mathcal{B}'$. For such a $b \in \mathcal{B}'$, we conclude that in (4.1) all the terms but $\big(\partial_i \nabla V_{(b,v)}\big)(q_v - q_b) \cdot \nabla_{p_v}$ are of the form $f_u(z) \cdot \nabla_{p_u}$ for some $u \in \mathcal{B}'$, so that

$$\big(\partial_i \nabla V_{(b,v)}\big)(q_v - q_b) \cdot \nabla_{p_v} \in \mathcal{M} \;. \tag{4.2}$$

By the definition of $\mathcal{T}\mathcal{B}'$, this holds for every $v \in \mathcal{T}\mathcal{B}' \setminus \mathcal{B}'$. We now get rid of the potential term in (4.2). Repeatedly taking Lie brackets with $\partial_{q_b^j}$, (4.2) implies that, for every non-zero multi-index α, we have

$$\big(D^\alpha \nabla V_{(b,v)}\big)(q_v - q_b) \cdot \nabla_{p_v} \in \mathcal{M} \;. \tag{4.3}$$

Let now ℓ be the value appearing in the non-degeneracy assumption for $V_{(b,v)}$ and let M be the $n \times n$ matrix-valued function whose elements are given by

$$M_{ij}(x) = \sum_{1 \le |\alpha| \le \ell} \big(D^\alpha \partial_i V_{(b,v)}\big)(x) \, \big(D^\alpha \partial_j V_{(b,v)}\big)(x) \;, \quad x \in \mathbf{R}^n \;.$$

It follows from the non-degeneracy assumption that M is invertible for every $x \in \mathbf{R}^n$, so that $M_{ij}^{-1}(x)$ is a smooth function. An explicit calculation shows, furthermore, that one has the identity

$$\partial_{p_v^j} = \sum_{i=1}^n \sum_{1 \le |\alpha| \le \ell} M_{ij}^{-1}(q_v - q_b)\big(D^\alpha \partial_i V_{(b,v)}\big)(q_v - q_b) \, \big(D^\alpha \nabla V_{(b,v)}\big)(q_v - q_b) \cdot \nabla_{p_v} \;.$$

From (4.3) and the fact that $M_{ij}^{-1}(q_v - q_b)\big(D^\alpha \partial_i V_{(b,v)}\big)(q_v - q_b)$ is a smooth function, we deduce that we indeed have $\nabla_{p_v} \in \mathcal{M}$, thus completing the proof. □

5 Lyapunov condition

In this section, we show that Conditions C1, C3, C4 and C5 imply that the system (2.2) satisfies Condition H2 above, *i.e.*, that $V = e^{\vartheta H}$ satisfies the Lyapunov property if ϑ is small enough.

The proof follows the lines of the argument that can be found in [31, 1]. Unfortunately, these works both contained a gap in the argument, which we presently correct (see Remark 5.12).

We fix $t_* > 0$ and $\vartheta < 1/T_{\max}$ with $T_{\max} = \max\{T_b : b \in \mathcal{B}\}$. The main result of this section is

Theorem 5.1. *Under Conditions C1, C3, C4 and C5, there is a constant $C_1 > 0$ such that for all z_0 such that $H(z_0)$ is large enough, we have*

$$\mathbf{E}_{z_0} e^{\vartheta H(z_{t_*}) - \vartheta H(z_0)} \le e^{-C_1 H(z_0)} \;. \tag{5.1}$$

Remark 5.2. By the coercivity of H, the theorem above implies that there exist constants $\varkappa \in (0,1)$ and $c > 0$, and a compact set K such that

$$\mathbf{E}_z e^{\vartheta H(z_{t_*})} \le \varkappa e^{\vartheta H(z)} + c \mathbf{1}_K(z) \;,$$

which is the usual Lyapunov condition used in Condition H2.

For the remainder of the paper, we assume that Conditions C1, C3, C4 and C5 are satisfied.

The central role in the proof of Theorem 5.1 will be played by the *dissipation integral*

$$\Gamma(t) = \sum_{b\in\mathcal{B}} \gamma_b \int_0^t p_b^2(s) ds \; . \tag{5.2}$$

In a nutshell, we will prove (5.1) by showing that if $H(z_0)$ is large enough, then with very high probability the main contribution to the energy difference $H(z_{t_*}) - H(z_0)$ comes from (minus) the dissipation integral $\Gamma(t_*)$, which, also with very high probability, scales like $H(z_0)$.

In order to do this, we start by partitioning, for each initial condition $z_0 \in \Omega$, the probability space into the following three events:

$$\begin{aligned} A_1 &= \left\{ H(z_s) \in \left[\frac{H(z_0)}{2}, 2H(z_0)\right] \;\; \forall s \in [0,t_*] \right\} \; , \\ A_2 &= \left\{ \inf_{s\in[0,t_*]} H(z_s) < \frac{H(z_0)}{2} \right\} \; , \\ A_3 &= \left\{ \sup_{s\in[0,t_*]} H(z_s) > 2H(z_0) \right\} \; . \end{aligned}$$

The event A_1 will be the center of most of our analysis. The event A_2 will be of no trouble, since after getting as low as $H(z_0)/2$, it is unlikely that the energy will increase again to a large value. Finally, the event A_3 will be of negligible probability at high energy.

When the event A_1 is realized, we will cut the time interval $[0, t_*]$ into subintervals. The length of each subinterval will depend on the distribution of energy between the *interaction* and *center of mass* degrees of freedom as follows.

We introduce the center of mass coordinates

$$Q = \frac{1}{|\mathcal{G}|} \sum_{v\in\mathcal{G}} q_v \; , \qquad P = \sum_{v\in\mathcal{G}} p_v \; , \tag{5.3}$$

and split the Hamiltonian according to

$$H = H_c + H_i \; , \tag{5.4}$$

where

$$\begin{aligned} H_c &= \frac{P^2}{2|\mathcal{G}|} + \sum_{v\in\mathcal{G}} U_v(q_v) \; , \\ H_i &= \frac{1}{2} \sum_{v\in\mathcal{G}} \left(p_v - \frac{P}{|\mathcal{G}|} \right)^2 + \sum_{e\in\mathcal{E}} V_e(\delta q_e) \; . \end{aligned}$$

We then let

$$\tau(z) = \begin{cases} \lambda H(z)^{\frac{1}{\ell_i} - \frac{1}{2}} & \text{if } H_i(z) \geq H(z)/2 \; , \\ \lambda H(z)^{\frac{1}{\ell_p} - \frac{1}{2}} & \text{if } H_c(z) > H(z)/2 \; , \end{cases} \tag{5.5}$$

where $\lambda > 0$ is arbitrary if $\ell_p > 2$, and subject to the condition $0 < \lambda \leq t_*/2$ if $\ell_p = 2$. Note that $\tau(z)$ is *not* random when z is fixed.

The rationale behind (5.5) is simple: when the system is dominated by the “internal” dynamics, the natural time scale is $H(z)^{1/\ell_i - 1/2}$. In the opposite case, the time scale

$H(z)^{1/\ell_p-1/2}$ of the pinning potentials is relevant. When $\ell_i = \ell_p$, this distinction of time scales obviously vanishes.

The following proposition, which we will prove in §5.1 and §5.2, says that with a very large probability, the average dissipation rate over the time interval $[0, \tau(z_0)]$ is at least some fraction of the initial energy.

Proposition 5.3. *Let*

$$\tilde{A} = \{H(z_s) \le 4H(z_0)\ \forall s \in [0, \tau(z_0)]\} \ .$$

Then there exist $\varepsilon, B > 0$ such that for all z_0 with $H(z_0)$ large enough,

$$\mathbf{P}_{z_0}\left(\tilde{A} \cap \{\Gamma(\tau(z_0)) < \varepsilon H(z_0)\tau(z_0)\}\right) \le e^{-BH(z_0)} \ . \tag{5.6}$$

For the remainder of this section, we assume that ε, B are fixed as in Proposition 5.3. We start with a corollary of Proposition 5.3, which says that one can basically apply Proposition 5.3 to successive time intervals in order to obtain estimates on $\Gamma(t_*)$.

Corollary 5.4. *There exists $B' > 0$ such that for all z_0 with $H(z_0)$ large enough,*

$$\mathbf{P}_{z_0}\left(A_1 \cap \left\{\Gamma(t_*) < \frac{\varepsilon t_*}{4} H(z_0)\right\}\right) \le e^{-B'H(z_0)} \ . \tag{5.7}$$

Proof. Fix z_0 and let $E = H(z_0)$. Consider the sequence of stopping times

$$\tau_0 = 0 \ , \qquad \tau_{j+1} = \tau_j + \tau(z_{\tau_j}) \ , \tag{5.8}$$

with $\tau(z)$ for $z \in \Omega$ as in (5.5). We now introduce the random variable

$$J = \sup\{j : \tau_j \le t_*\} \ .$$

On A_1, we have for all $t \le t_*$ that

$$\lambda(2E)^{\frac{1}{\ell_i}-\frac{1}{2}} \le \tau(z_t) \le \lambda(E/2)^{\frac{1}{\ell_p}-\frac{1}{2}} \ ,$$

and hence that

$$J \le \hat{J} \equiv \lfloor t_*(2E)^{\frac{1}{2}-\frac{1}{\ell_i}}\lambda^{-1} \rfloor \ .$$

Moreover, if E is large enough (and in the case $\ell_p = 2$, using that $\lambda \le t_*/2$), we have on A_1 that $J > 0$ and that

$$\tau_J > t_* - \tau(z_{\tau_J}) \ge \frac{t_*}{2} \ .$$

Consider next the events

$$G_j = \{J > j\} \cap \left\{\sum_{b\in\mathcal{B}} \gamma_b \int_{\tau_j}^{\tau_{j+1}} p_b^2(s) ds < \varepsilon\tau(z_{\tau_j})H(z_{\tau_j})\right\} \ ,$$

$$G = \bigcup_{j\ge 0} G_j = \left\{\exists j < J : \sum_{b\in\mathcal{B}} \gamma_b \int_{\tau_j}^{\tau_{j+1}} p_b^2(s) ds < \varepsilon\tau(z_{\tau_j})H(z_{\tau_j})\right\} \ .$$

We observe that the event $A_1 \cap \{J > j\}$ is a subset of

$$\tilde{A}_j \equiv \left\{H(z_{\tau_j}) \ge \frac{E}{2} \text{ and } H(z_t) \le 4H(z_{\tau_j})\ \forall t \in [\tau_j, \tau_{j+1}]\right\} \ .$$

Thus, if E is large enough, we find by Proposition 5.3 and the strong Markov property that for all $j \ge 0$,

$$\mathbf{P}_{z_0}(A_1 \cap G_j) \le e^{-BE/2} \ ,$$

so that

$$\mathbf{P}_{z_0}(A_1 \cap G) \leq \sum_{j=0}^{\hat{J}-1} \mathbf{P}_{z_0}(A_1 \cap G_j) \leq \hat{J} e^{-BE/2} \leq e^{-B'E} \tag{5.9}$$

if $B' > 0$ is small enough and E large enough.

We observe next that on $A_1 \cap G^c$ and for all E large enough,

$$\begin{aligned}\Gamma(t_*) &\geq \sum_{j=0}^{J-1} \sum_{b\in\mathcal{B}} \gamma_b \int_{\tau_j}^{\tau_{j+1}} p_b^2(s) ds \geq \sum_{j=0}^{J-1} \varepsilon \tau(z_{\tau_j}) H(z_{\tau_j}) \\ &\geq \frac{\varepsilon E}{2} \sum_{j=0}^{J-1} \tau(z_{\tau_j}) = \frac{\varepsilon E}{2} \tau_J \geq \frac{\varepsilon E}{4} t_* \ .\end{aligned} \tag{5.10}$$

Thus, the left-hand side of (5.7) is bounded by $\mathbf{P}_{z_0}(A_1 \cap G)$, which by (5.9) completes the proof. □

Lemma 5.5. *There are constants $\rho, q > 0$ such that for every initial condition z_0, every event A, and all $t > 0$,*

$$\mathbf{E}_{z_0}\left(e^{\vartheta H(z_t) - \vartheta H(z_0)} \mathbf{1}_A\right) \leq e^{C_* t} \left(\mathbf{E}_z(e^{-\rho\Gamma(t)} \mathbf{1}_A)\right)^{\frac{1}{q}} \leq e^{C_* t} \ , \tag{5.11}$$

with again $C_* = \vartheta \sum_{b\in\mathcal{B}} \gamma_b T_b$.

Proof. This proof is as in [31, 1]. By applying the Itô formula to $H(z_t)$, we find

$$\mathbf{E}_{z_0}\left(e^{\vartheta H(z_t) - \vartheta H(z_0)} \mathbf{1}_A\right) = e^{C_* t} \mathbf{E}_{z_0}\left(e^{-\vartheta\Gamma(t) + \vartheta M_t} \mathbf{1}_A\right) \ ,$$

where

$$M_t = \int_0^t \sum_{b\in\mathcal{B}} \sqrt{2\gamma_b T_b} p_b(s) dW_b(s) \ .$$

The quadratic variation of M_t satisfies

$$[M]_t = 2 \int_0^t \sum_{b\in\mathcal{B}} \gamma_b T_b p_b^2(s) ds \leq 2 T_{\max} \Gamma(t) \ . \tag{5.12}$$

Let $p > 1$ be such that $p\vartheta < 1/T_{\max}$ and let q be such that $\frac{1}{q} + \frac{1}{p} = 1$. By Hölder's inequality,

$$\begin{aligned}\mathbf{E}_{z_0}\left(e^{-\vartheta\Gamma(t) + \vartheta M_t} \mathbf{1}_A\right) &= \mathbf{E}_{z_0}\left(e^{-\vartheta\Gamma(t) + \frac{p\vartheta^2}{2}[M]_t} \mathbf{1}_A e^{\vartheta M_t - \frac{p\vartheta^2}{2}[M]_t}\right) \\ &\leq \left(\mathbf{E}_{z_0} e^{-\vartheta q\Gamma(t) + \frac{qp\vartheta^2}{2}[M]_t} \mathbf{1}_A\right)^{\frac{1}{q}} \left(\mathbf{E}_{z_0} e^{p\vartheta M_t - \frac{p^2\vartheta^2}{2}[M]_t}\right)^{\frac{1}{p}} \ .\end{aligned}$$

The expectation in the second bracket in the last line is ≤ 1, since the exponential there is a Doléans–Dade exponential, and thus a supermartingale. Finally, by (5.12) we obtain (5.11) with $\rho = \vartheta q(1 - p\vartheta T_{\max}) > 0$. □

Lemma 5.6. *There exists $c > 0$ such that for all z_0 with $H(z_0)$ large enough,*

$$\mathbf{P}_{z_0}(A_3) \leq e^{-cH(z_0)} \ . \tag{5.13}$$

Proof. This is a classical result (see for example [31] or the proof of Theorem 3.5 in [20]). Observe that by (3.3),

$$(\partial_t + \mathcal{L})(e^{\vartheta H - C_* t}) = (\mathcal{L} - C_*)e^{\vartheta H - C_* t} \le 0 \ .$$

Consider the stopping time $\sigma = \min(t_*, \inf\{t \ge 0 : H(z_t) > 2H(z_0)\})$ (with the convention $\inf \emptyset = +\infty$). Then, σ is a bounded stopping time, and we have by Dynkin's formula

$$\begin{aligned}\mathbf{E}_{z_0} e^{\vartheta H(z_\sigma) - C_* t_*} &\le \mathbf{E}_{z_0} e^{\vartheta H(z_\sigma) - C_* \sigma} \\ &= e^{\vartheta H(z_0)} + \mathbf{E}_{z_0} \int_0^\sigma ((\partial_s + \mathcal{L})(e^{\vartheta H - C_* s}))(z_s) ds \ .\end{aligned}$$

As the expectation in the last line is non-positive, we find $\mathbf{E}_{z_0} e^{\vartheta H(z_\sigma)} \le e^{C_* t_* + \vartheta H(z_0)}$, and thus

$$\mathbf{P}_{z_0}(A_3) = \mathbf{P}_{z_0}\{\sigma < t_*\} \le e^{-2\vartheta H(z_0)} \mathbf{E}_{z_0}\left(e^{\vartheta H(z_\sigma)} \mathbf{1}_{\sigma < t_*}\right) \le e^{C_* t_* - \vartheta H(z_0)} \ ,$$

where the last inequality uses (5.11). Thus, choosing c small enough completes the proof. ☐

We can now give the

Proof of Theorem 5.1. First, we have by Lemma 5.5 and Corollary 5.4 that if $H(z_0)$ is large enough,

$$\begin{aligned}\mathbf{E}_{z_0}\left(e^{\vartheta H(z_{t_*}) - \vartheta H(z_0)} \mathbf{1}_{A_1}\right) &\le e^{C_* t} \left(\mathbf{E}_{z_0}(e^{-\rho \Gamma(t_*)} \mathbf{1}_{A_1})\right)^{\frac{1}{q}} \\ &\le e^{C_* t}\left(e^{-B' H(z_0)} + e^{-\rho \varepsilon t_* H(z_0)/4}\right)^{\frac{1}{q}} \le e^{-cH(z_0)}\end{aligned} \tag{5.14}$$

for some small enough $c > 0$. We next work on A_2. Consider the stopping time $\sigma = \min(t_*, \inf\{t \ge 0 : H(z_t) < H(z_0)/2\})$ (again with $\inf \emptyset = +\infty$). We have $A_2 = \{\sigma < t_*\}$ and

$$\mathbf{E}_{z_0}\left(e^{\vartheta H(z_{t_*})} \mathbf{1}_{A_2}\right) \le e^{\vartheta H(z_0)/2} \mathbf{E}_{z_0}\left(e^{\vartheta H(z_{t_*}) - \vartheta H(z_\sigma)} \mathbf{1}_{A_2}\right) \le e^{\vartheta H(z_0)/2 + C_* t_*} \ ,$$

where we have used the strong Markov property, (5.11), and the fact that $t_* - \sigma \le t_*$. But then,

$$\mathbf{E}_{z_0}\left(e^{\vartheta H(z_{t_*}) - \vartheta H(z_0)} \mathbf{1}_{A_2}\right) \le e^{C_* t_* - \vartheta H(z_0)/2} \le e^{-cH(z_0)} \ , \tag{5.15}$$

if $c > 0$ is small enough and $H(z_0)$ is large enough.

Finally, by Lemma 5.5 and Lemma 5.6, we have

$$\mathbf{E}_{z_0}\left(e^{\vartheta H(z_{t_*}) - \vartheta H(z_0)} \mathbf{1}_{A_3}\right) \le e^{C_* t_*} \left(\mathbf{P}_{z_0}(A_3)\right)^{\frac{1}{q}} \le e^{-cH(z_0)} \ , \tag{5.16}$$

which has the desired form again. Summing (5.14), (5.15) and (5.16) completes the proof. ☐

Remark 5.7. Above, we split the time interval $[0, t_*]$ into many subintervals, and apply Proposition 5.3 to each of them. This is what allows us to obtain (5.1), which is very natural from the dimensional point of view. In comparison, [31, 1] use the same Lyapunov function, but obtain weaker estimates (but still sufficient to obtain exponential convergence in (2.5)): the bound obtained in [31] is $\mathbf{E}_{z_0} e^{\vartheta H(z_{t_*}) - \vartheta H(z_0)} \le e^{-C_1 H^r(z_0)}$ with $r \in (0,1)$, and in [1] it is only shown that $\lim_{\|z_0\| \to \infty} \mathbf{E}_{z_0} e^{\vartheta H(z_{t_*}) - \vartheta H(z_0)} = 0$.

It now remains to prove Proposition 5.3. In order to do so, we start with some technical lemmas.

Lemma 5.8. *Let $r \geq 1$ and let $f : \mathbf{R}^r \to \mathbf{R}^r$ be a locally Lipschitz function. For $T > 0$, let $V \in \mathcal{C}([0,T], \mathbf{R}^r)$ and consider*

$$dx_t = f(x_t)dt + dV(t), \qquad dy_t = f(y_t)dt$$

with initial conditions $x_0 = y_0 \in \mathbf{R}^r$. Then, provided that both x and y exist up to time T,

$$\sup_{t\in[0,T]} \|x_t - y_t\| \leq e^{k_* T} \sup_{t\in[0,T]} \|V(t)\| , \qquad k_* = \sup_{t\in[0,T]} \frac{\|f(x_t) - f(y_t)\|}{\|x_t - y_t\|} ,$$

with the convention $0/0 = 0$.

Proof. Setting $\Delta_s = \|x_s - y_s\|$, we have $\Delta_t \leq \int_0^t k_* \Delta_s ds + \|V(t)\|$ and the result follows from Gronwall's inequality. □

Remark 5.9. We will later use Lemma 5.8 to show that, after adequate rescaling, (2.2) (or a component thereof) converges to a deterministic dynamics at high energy.

As a consequence of the definition of H, Condition C3 and Remark 2.9 (iv), we immediately obtain

Lemma 5.10. *There is a constant $C > 0$ such that for all $z \in \Omega$, $v \in \mathcal{G}$ and $e \in \mathcal{E}$,*

$$\|q_v\| \leq C(1 + H^{\frac{1}{\ell_p}}(z)) , \quad \|\delta q_e\| \leq C(1 + H^{\frac{1}{\ell_i}}(z)) , \quad \|p_v\| \leq C H^{\frac{1}{2}}(z) . \tag{5.17}$$

We are now ready to prove Proposition 5.3. We treat the case where $H_i(z_0) \geq H(z_0)/2$ in §5.1 and the case where $H_c(z_0) > H(z_0)/2$ in §5.2. When $\ell_i = \ell_p$, such a distinction is not necessary and only the analysis in §5.1 is required.

5.1 When the interactions dominate

In this subsection, we make

Assumption 5.11. *If $\ell_i > \ell_p$, we assume that $z_0 \in \Omega$ is such that $H_i(z_0) \geq H(z_0)/2$. (If $\ell_i = \ell_p$, we make no such restriction.)*

We write $E = H(z_0)$. Consider the rescaled time $\sigma = E^{\frac{1}{2}-\frac{1}{\ell_i}} t$ and the variables

$$\begin{aligned} \tilde{p}_v(\sigma) &= E^{-\frac{1}{2}} p_v(E^{\frac{1}{\ell_i}-\frac{1}{2}}\sigma) , \\ \tilde{q}_v(\sigma) &= E^{-\frac{1}{\ell_i}} q_v(E^{\frac{1}{\ell_i}-\frac{1}{2}}\sigma) . \end{aligned} \tag{5.18}$$

We write $\tilde{z} = (\tilde{p}, \tilde{q})$ and $\tilde{z}_0$ for the rescaled initial condition. We consider times $t \in [0, \tau(z_0)] = [0, \lambda E^{1/\ell_i - 1/2}]$, or equivalently $\sigma \in [0, \lambda]$. Observe that in terms of the rescaled time and variables, (5.6) reads

$$\mathbf{P}_{z_0}\left(\tilde{A} \cap \left\{\int_0^\lambda \sum_{b\in\mathcal{B}} \gamma_b \tilde{p}_b^2(\sigma) d\sigma < \varepsilon\lambda\right\}\right) \leq e^{-BE} . \tag{5.19}$$

In the remainder of this section, we show that (5.19) holds provided E is large enough and z_0 satisfies Assumption 5.11.

Introducing

$$\tilde{H}(p,q) = \sum_{v\in\mathcal{G}} \frac{p_v^2}{2} + \sum_{v\in\mathcal{G}} E^{-1} U_v(E^{\frac{1}{\ell_i}} q_v) + \sum_{e\in\mathcal{E}} E^{-1} V_e(E^{\frac{1}{\ell_i}} \delta q_e) , \tag{5.20}$$

it is easy to see that

$$d\tilde{q}_v = \tilde{p}_v d\sigma \;, \\ d\tilde{p}_v = -(\nabla_{q_v}\tilde{H})(\tilde{p},\tilde{q})d\sigma - E^{\frac{1}{\ell_i}-\frac{1}{2}}\gamma_v\tilde{p}_v d + E^{\frac{1}{2\ell_i}-\frac{3}{4}}\sqrt{2T_v\gamma_v}d\tilde{W}_v(\sigma) \;, \tag{5.21}$$

where $\tilde{W}_v(\sigma) = E^{-\frac{1}{2\ell_i}+\frac{1}{4}}W_v(E^{\frac{1}{\ell_i}-\frac{1}{2}}\sigma)$ is again an n-dimensional Brownian motion. Clearly, in (5.21), the stochastic term vanishes in the limit $E \to \infty$, and so does the dissipative term, except when $\ell_i = 2$.

Observe that when $E \to \infty$, the Hamiltonian $\tilde{H}$ converges pointwise to

$$\hat{H}(p,q) = \sum_{v\in\mathcal{G}} \frac{p_v^2}{2} + \delta_{\ell_i,\ell_p}\sum_{v\in\mathcal{G}} U_{v,\infty}(q_v) + \sum_{e\in\mathcal{E}} V_{e,\infty}(\delta q_e) \;, \tag{5.22}$$

where $U_{v,\infty}$ and $V_{e,\infty}$ are defined in Condition C3.

Moreover, by construction,

$$H(z) = E\tilde{H}(\tilde{z}) \;,$$

and in particular,

$$\tilde{H}(\tilde{z}_0) = 1 \;. \tag{5.23}$$

We introduce the set

$$\tilde{K}_E = \{\tilde{z} : \tilde{H}(\tilde{z}) \le 4\} \;.$$

On the event $\tilde{A}$, we have $H(z_t) \le 4E$ for all $t \in [0, \tau(z_0)]$, and hence also

$$\tilde{z}_\sigma \in \tilde{K}_E, \quad 0 \le \sigma \le \lambda \;.$$

By (5.17), there exists $\tilde{C} > 0$ such that if E is large enough, we have that for all $\tilde{z} \in \tilde{K}_E$,

$$\|\tilde{q}_v\| \le \tilde{C}E^{\frac{1}{\ell_p}-\frac{1}{\ell_i}} \;, \quad \|\delta\tilde{q}_e\| \le \tilde{C} \;, \quad \|\tilde{p}_v\| \le \tilde{C} \;. \tag{5.24}$$

Remark 5.12. Note that if $\ell_i > \ell_p$, then $\tilde{q}_v$ may become arbitrarily large when E is large, so that the set $\tilde{K}_E$ is not bounded uniformly in E. In fact, when $\ell_i > \ell_p$, it is *not* true that $\sup_{\tilde{z}\in\tilde{K}_E}|\tilde{H}(\tilde{z}) - \hat{H}(\tilde{z})|$ goes to zero when $E \to \infty$. Indeed, for all E one can find $\tilde{z} \in \tilde{K}_E$ such that all the energy is in the pinning potential, so that $\hat{H}(\tilde{z}) = 0$ but $\tilde{H}(\tilde{z}) = 1$. This explains why we have to restrict ourselves to initial conditions such that $H_i(z_0) \ge H(z_0)/2$ (which will guarantee that $\hat{H}(\tilde{z}_0)$ is not too small), and then treat the opposite case separately in §5.2. This distinction is missing from the proofs in [31, 1].

Lemma 5.13. *For all* $0 \le |\alpha| \le 1$ *and* $e \in \mathcal{E}$*, we have*

$$\lim_{E\to\infty}\sup_{\tilde{z}\in\tilde{K}_E}\left|D^\alpha_{\tilde{z}}\left(E^{-1}V_e(E^{\frac{1}{\ell_i}}\delta\tilde{q}_e) - V_{e,\infty}(\delta\tilde{q}_e)\right)\right| = 0 \;. \tag{5.25}$$

Let $v \in \mathcal{G}$*. If* $\ell_i = \ell_p$ *and* $0 \le |\alpha| \le 1$ *(case 1) or if* $\ell_i > \ell_p$ *and* $|\alpha| = 1$ *(case 2), then:*

$$\lim_{E\to\infty}\sup_{\tilde{z}\in\tilde{K}_E}\left|D^\alpha_{\tilde{z}}\left(E^{-1}U_v(E^{\frac{1}{\ell_i}}\tilde{q}_v) - \delta_{\ell_i,\ell_p}U_{v,\infty}(\tilde{q}_v)\right)\right| = 0 \;, \tag{5.26}$$

$$\lim_{E\to\infty}\sup_{\tilde{z}\in\tilde{K}_E}\left|(D^\alpha\tilde{H})(\tilde{z}) - (D^\alpha\hat{H})(\tilde{z})\right| = 0 \;. \tag{5.27}$$

Proof. The first identity follows immediately from Condition C3, Remark 2.9 (iii) and (5.24). Assume now we are in case 1. By Condition C3 and Remark 2.9 (iii),

$$\lim_{E\to\infty}\sup_{\|x\|\le\tilde{C}}\left|\left(E^{\frac{|\alpha|}{\ell_i}-1}(D^\alpha U_v)(E^{\frac{1}{\ell_i}}x) - (D^\alpha U_{v,\infty})(x)\right)\right| = 0 \;.$$

This together with (5.24) proves (5.26).

Assume now we are in case 2. By (5.24), in order to prove (5.26), it is enough to show that when $|\alpha| = 1$,

$$\lim_{E\to\infty} \sup_{\|x\|\le \tilde{C}E^{\frac{1}{\ell_p}-\frac{1}{\ell_i}}} \left| E^{\frac{|\alpha|}{\ell_i}-1}(D^\alpha U_v)(E^{\frac{1}{\ell_i}}x)\right| = 0 \; . \tag{5.28}$$

By Remark 2.9 (ii), we have $|(D^\alpha U_v)(E^{\frac{1}{\ell_i}}x)| \le c(1+E^{\ell_p/\ell_i-|\alpha|/\ell_i}\|x\|^{\ell_p-|\alpha|})$ for some $c>0$. From this, we obtain that for some $c'>0$, the supremum in (5.28) is bounded above by

$$c'(E^{\frac{|\alpha|}{\ell_i}-1} + E^{\frac{|\alpha|}{\ell_i}-\frac{|\alpha|}{\ell_p}}) \; .$$

Clearly, since $|\alpha| = 1 < 2 \le \ell_p < \ell_i$, the above vanishes when $E \to \infty$, which proves (5.28) and hence (5.26).

Finally, in both cases, combining (5.25) and (5.26) yields (5.27) (recalling that the kinetic parts in $\hat{H}$ and $\tilde{H}$ are identical). ☐

We now observe that for all E large enough,

$$\hat{H}(\tilde{z}_0) \in [1/4, 2] \; . \tag{5.29}$$

Indeed, if $\ell_i = \ell_p$, this follows from (5.23) and (5.27). If $\ell_i > \ell_p$, then Assumption 5.11 ensures that $|\sum_{v\in\mathcal{G}} E^{-1}U_v(E^{1/\ell_i}q_v)| \le 1/2$, so that (5.23) and (5.25) indeed imply (5.29) for E large enough.

Next, (5.21) can be rewritten as

$$\begin{aligned} d\tilde{q}_v &= \tilde{p}_v d\sigma \; , \\ d\tilde{p}_v &= -(\nabla_{q_v}\hat{H})(\tilde{p},\tilde{q})d\sigma + \tilde{R}_v(\tilde{q})d\sigma \\ &\quad - E^{\frac{1}{\ell_i}-\frac{1}{2}}\gamma_v\tilde{p}_v d\sigma + E^{\frac{1}{2\ell_i}-\frac{3}{4}}\sqrt{2T_v\gamma_v}d\tilde{W}_v(\sigma) \; , \end{aligned} \tag{5.30}$$

where $\tilde{R}_v(\tilde{q}) = -\nabla_{\tilde{q}_v}(\tilde{H}(\tilde{z}) - \hat{H}(\tilde{z}))$, which by Lemma 5.13 satisfies, regardless of whether $\ell_i > \ell_p$ or $\ell_i = \ell_p$,

$$\lim_{E\to\infty} \sup_{\tilde{z}\in\tilde{K}_E} \|\tilde{R}_v(\tilde{q}_v)\| \to 0 \; . \tag{5.31}$$

Consider now the deterministic limiting system

$$\begin{aligned} d\hat{q}_v &= \hat{p}_v d\sigma \; , \\ d\hat{p}_v &= -(\nabla_{q_v}\hat{H})(\hat{z})d\sigma - \delta_{\ell_i,2}\gamma_v\hat{p}_v d\sigma \; , \end{aligned} \tag{5.32}$$

with initial condition $\hat{z}_0 = \tilde{z}_0$.

Proposition 5.14. *There is a constant $C > 0$ such that for every initial condition $\hat{z}_0$ such that $\hat{H}(\hat{z}_0) \in [1/4, 2]$, the solution of (5.32) satisfies*

$$\int_0^\lambda \sum_{b\in\mathcal{B}} \gamma_b \hat{p}_b^2(\sigma)d\sigma \ge C \; . \tag{5.33}$$

Proof. We first show that

$$\int_0^\lambda \sum_{b\in\mathcal{B}} \gamma_b \hat{p}_b^2(\sigma)d\sigma > 0 \quad \text{if} \quad \hat{H}(\hat{z}_0) > 0 \; . \tag{5.34}$$

Indeed, assume the left-hand side of (5.34) is zero. Then, for all $b \in \mathcal{B}$, we have $\hat{p}_b(\sigma) \equiv 0$ on $[0,\lambda]$. Take now $v \in \mathcal{TB} \setminus \mathcal{B}$. There exists then $b \in \mathcal{B}$ such that b is linked only to v and possibly some vertices in $\mathcal{B}$. Now, since the masses in $\mathcal{B}$ do not move, all forces among them are constant (this applies to both the interaction forces $-\nabla V_{e,\infty}$ with $e \in \mathcal{B} \times \mathcal{B}$ and, if $\ell_i = \ell_p$, to the pinning forces $-\nabla U_{b',\infty}$ with $b' \in \mathcal{B}$). Thus, since the total force on b is identically zero, we must have that $\nabla V_{(b,v),\infty}(\hat{q}_b(\sigma) - \hat{q}_v(\sigma))$ is constant. But then, by Condition C4, this means that actually $\hat{q}_b(\sigma) - \hat{q}_v(\sigma)$ is constant, and hence that so is $\hat{q}_v(\sigma)$. We have thus shown that $\hat{p}_v(\sigma) \equiv 0$ for all $v \in \mathcal{T}\mathcal{B}$. Proceeding in the same way, we obtain inductively that the same holds for all v in $\mathcal{T}^2\mathcal{B}$, $\mathcal{T}^3\mathcal{B}$, etc. Thus, by Condition C1, we eventually obtain that no mass moves during the time interval $[0, \lambda]$. But then we have $\hat{p}_v(0) = 0$ and $\nabla_{q_v} \hat{H}(\hat{z}_0) = 0$ for all $v \in \mathcal{G}$, which is only possible if $\hat{H}(\hat{z}_0) = 0$, so that (5.34) holds.

We now complete the proof of the proposition using a compactness argument and the fact that the solution of (5.32) depends continuously on the initial condition $\hat{z}_0$. In order to do so, there are two cases to consider.

- $\ell_i = \ell_p$. Then, the set $\{\hat{z} : \hat{H}(\hat{z}) \in [1/4, 2]\}$ is compact, and hence (5.33) holds for some $C > 0$.
- $\ell_i > \ell_p$. Then, the set $\{\hat{z} : \hat{H}(\hat{z}) \in [1/4, 2]\}$ is not compact, since it is invariant under global translations $q_v \mapsto q_v + \rho$, where ρ is any vector in $\mathbf{R}^n$ independent of v. But when $\ell_i > \ell_p$, both the dynamics (5.32) and the left-hand side of (5.33) are invariant under such translations. Since the set $\{\hat{z} : \hat{H}(\hat{z}) \in [1/4, 2]\}$ is compact modulo such translations, we obtain (5.33) for some $C > 0$.

This completes the proof. □

Remark 5.15. Note that Condition C4 is only used to prove (5.34). In fact, there are systems for which (5.34), and hence all the results in the present paper, hold without Condition C4. For example, consider a chain of N oscillators with heat baths at both ends, *i.e.*, $\mathcal{G} = \{1, \dots, N\}$, $\mathcal{B} = \{1, N\}$ and $\mathcal{E} = \{(1, 2), (2, 3), \dots, (N-1, N)\}$. Let $\ell_i > \ell_p$, so that the limiting system only involves the interaction potentials. Assume the left-hand side of (5.34) is zero. Then, on the time interval $[0, \lambda]$, we have $\hat{p}_1(\sigma) \equiv 0$. But then, we must have $\nabla V_{(1,2),\infty}(\hat{q}_2(\sigma) - \hat{q}_1(\sigma)) \equiv 0$ (unlike in the general case, we know here that the constant is zero, since no other force acts on the first oscillator). As a consequence, since the only stationary point of $V_{(1,2),\infty}$ is at the origin (this is true of any coercive, homogeneous function without the need for Condition C4), we must have $\hat{q}_2(\sigma) \equiv \hat{q}_1(\sigma)$. But then we also have $p_2(\sigma) \equiv 0$. Continuing like this along the chain, we eventually obtain that all the masses stand still, and conclude as above that $\hat{H}(\hat{z}_0) = 0$.

Remark 5.16. Condition C4 cannot be waived in general. We give here a counterexample in three[8] dimensions consisting of two oscillators 1 and 2, the first of which is coupled to a heat bath (see Figure 3). We start with both oscillators at rest at position $(0, 1, 0)$ and $(4, 2, 0)$ respectively. We assume that $V_{21}(x, y, z) = \frac{y^4}{4} + \frac{x^2 z^2}{2}$ when $(x, y, z) \approx (4, 1, 0)$, and that $U_1 = U_2 = U$, where $U(x, y, z) = \frac{x^4 + y^4 + z^4}{4}$ when $(x, y, z) \approx (0, 1, 0)$ and $U(x, y, z) = \frac{x^4}{64} - \frac{y^4}{32} + \frac{z^4}{4}$ when $(x, y, z) \approx (4, 2, 0)$. The potentials above can be extended to non-degenerate, coercive, homogeneous functions of degree 4 (note in particular that $\hat{H} = H$). Moreover, Condition C4 is not satisfied, as $\nabla V_{21}(4 + \varepsilon, 1, 0) = (0, 1, 0)$ for all small enough $|\varepsilon|$. In this setup, the initial energy of the system is non-zero, and there exists a finite time interval during which the following happens: oscillator 1 does not move at all (so that (5.34) fails to hold if λ is small enough), and the position of oscillator 2 is $(x(t), 2, 0)$, for some decreasing $x(t)$. The interaction force f_{21} (see Figure 3) remains equal to $(0, 1, 0)$, and the pinning force f_1 acting on oscillator 1 remains equal to $(0, -1, 0)$.

[8] Everything in this example happens in the Oxy-plane. The third dimension is necessary only to ensure that the interaction potential is non-degenerate.

During the same time, the pinning force f_2 acting on oscillator 2 is equal to $(-\frac{x^3(t)}{16}, 1, 0)$, consistently with the motion described above.

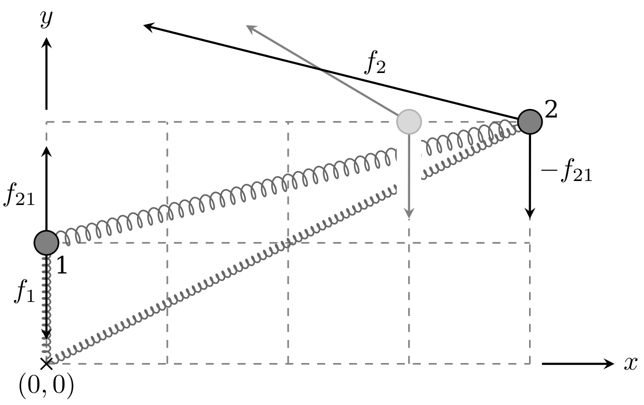

Figure 3: Illustration of the example in Remark 5.16 in the Oxy-plane (in which the motion takes place).

Returning to the proof of (5.19), we compare now the systems $(\tilde{p}, \tilde{q})$ and $(\hat{p}, \hat{q})$.

Lemma 5.17. *There exist a constant $c > 0$, a family of constants $(G_E)_{E>0}$ satisfying $\lim_{E\to\infty} G_E = 0$, and a family of non-negative random variables $(\eta_E)_{E>0}$ satisfying*

$$\mathbf{P}\{\eta_E \geq s\} \leq e^{-\frac{s^2}{2}} \ , \tag{5.35}$$

such that if E is large enough,

$$\mathbf{1}_{\tilde{A}} \sup_{\sigma\in[0,\lambda]} \|\tilde{z}_\sigma - \hat{z}_\sigma\| \leq G_E + cE^{\frac{1}{2\ell_i}-\frac{3}{4}}\eta_E \ .$$

Proof. The result immediately follows from Lemma 5.8 and (5.31), provided we can show that there exists an absolute constant $k > 0$ such that on the event $\tilde{A}$, we have $\|(\nabla\hat{H})(\tilde{z}_\sigma) - (\nabla\hat{H})(\hat{z}_\sigma)\| \leq k\|\tilde{z}_\sigma - \hat{z}_\sigma\|$ for all $0 \leq \sigma \leq \lambda$ (we need not worry about the other terms in (5.32), as they are globally Lipschitz). As mentioned above, on the event $\tilde{A}$, we have $\tilde{z}_\sigma \in \tilde{K}_E$ for all $0 \leq \sigma \leq \lambda$. Moreover, since $\frac{d}{d\sigma}\hat{H}(\hat{z}_\sigma) = -\delta_{\ell_i,2}\sum_{b\in\mathcal{B}}\gamma_b\hat{p}_b^2 \leq 0$, we have by (5.29) that $\hat{H}(\hat{z}_\sigma) \leq 2$ for all $0 \leq \sigma \leq \lambda$. We consider again two cases separately.

- $\ell_i = \ell_p$. Then, there exists $R > 0$ such that for all E large enough, $\|\tilde{z}_\sigma\| \leq R$ and $\|\hat{z}_\sigma\| \leq R$ for all $0 \leq \sigma \leq \lambda$. Since $\nabla\hat{H}$ is locally Lipschitz (by Condition C3), the proof is complete.
- $\ell_i > \ell_p$. Then, one can find $R > 0$ such that $\|\delta\tilde{q}_e(\sigma)\|, \|\delta\hat{q}_e(\sigma)\|, \|\tilde{p}_v(\sigma)\|$ and $\|\hat{p}_v(\sigma)\|$ are bounded by R for all $0 \leq \sigma \leq \lambda$. Since $\nabla\hat{H}$ is locally Lipschitz and depends only the δq_e and p_v, the proof is complete.

Note that by Lemma 5.8, the random variable η_E can be chosen as a constant times $\sup_{\sigma\in[0,\lambda]} \|\tilde{W}(\sigma)\|$, where $\tilde{W} = (\tilde{W}_b)_{b\in\mathcal{B}}$ is an $n|\mathcal{B}|$-dimensional Brownian motion. While $\tilde{W}$ depends on E pathwise, its distribution does not. Moreover, $\tilde{W}$ does not depend on z_0 for a given energy E, and thus the same is true of η_E. □

Using Lemma 5.17, Proposition 5.14 and the inequality $x^2 \geq \frac{y^2}{2} - (x-y)^2$, we obtain that there exist $c, c' > 0$ such that on $\tilde{A}$ and if E is large enough,

$$\int_0^\lambda \sum_{b\in\mathcal{B}} \gamma_b\tilde{p}_b^2(\sigma)d\sigma \geq c - c'(G_E + E^{\frac{1}{2\ell_i}-\frac{3}{4}}\eta_E)^2 \geq c - 2c'G_E^2 - 2c'E^{\frac{1}{\ell_i}-\frac{3}{2}}\eta_E^2 \ .$$

Since $G_E \to 0$, we find for E large enough that

$$\int_0^\lambda \sum_{b\in\mathcal{B}} \gamma_b \tilde{p}_b^2(\sigma)d\sigma \geq \frac{c}{2} - 2c' E^{\frac{1}{\ell_i}-\frac{3}{2}}\eta_E^2 \ ,$$

so that

$$\mathbf{P}_{z_0}\left(\tilde{A} \cap \left\{\int_0^\lambda \sum_{b\in\mathcal{B}} \gamma_b \tilde{p}_b^2(\sigma)d\sigma < \frac{c}{4}\right\}\right) \leq \mathbf{P}\left\{\eta_E > E^{\frac{3}{4}-\frac{1}{2\ell_i}}\sqrt{\frac{c}{8c'}}\right\} \ .$$

Using now (5.35) and the fact that $\frac{1}{\ell_i} - \frac{3}{2} \leq -1$ completes the proof of (5.19) (for an adequate choice of ε and B).

Thus, if $\ell_i = \ell_p$, the proof of Proposition 5.3 is complete. If now $\ell_i > \ell_p$, then because of Assumption 5.11, the conclusion of Proposition 5.3 is proved only in the case where $H_i(z_0) \geq H(z_0)/2$, and the next subsection is required.

5.2 When the pinning dominates

Recalling the decomposition of H introduced in (5.4), we now make the following assumption.

Assumption 5.18. *We assume that $\ell_i > \ell_p$ and that the initial condition $z_0 \in \Omega$ satisfies $H_c(z_0) > H(z_0)/2$.*

We start by rescaling the system in much the same way as in §5.1, except that we now choose the natural scaling of the pinning. More precisely, we introduce the rescaled time $\sigma = E^{1/2-1/\ell_p}t$ and the variables

$$\tilde{p}_v(\sigma) = E^{-\frac{1}{2}}p_v(E^{\frac{1}{\ell_p}-\frac{1}{2}}\sigma) \ ,$$
$$\tilde{q}_v(\sigma) = E^{-\frac{1}{\ell_p}}q_v(E^{\frac{1}{\ell_p}-\frac{1}{2}}\sigma) \ .$$

We consider times $t \in [0, \tau(z_0)] = [0, \lambda E^{\frac{1}{\ell_p}-\frac{1}{2}}]$, or equivalently $\sigma \in [0, \lambda]$. As in §5.1, the analogue of (5.6) in terms of the rescaled variables and time is

$$\mathbf{P}_{z_0}\left(\tilde{A} \cap \left\{\int_0^\lambda \sum_{b\in\mathcal{B}} \gamma_b \tilde{p}_b^2(\sigma)d\sigma < \varepsilon\lambda\right\}\right) \leq e^{-BE} \ . \tag{5.36}$$

We let now

$$\tilde{H}(p,q) = \sum_{v\in\mathcal{G}} \frac{p_v^2}{2} + \sum_{v\in\mathcal{G}} E^{-1}U_v(E^{\frac{1}{\ell_p}}q_v) + \sum_{e\in\mathcal{E}} E^{-1}V_e(E^{\frac{1}{\ell_p}}\delta q_e) \ ,$$

and obtain

$$\begin{aligned} d\tilde{q}_v &= \tilde{p}_v d\sigma \ , \\ d\tilde{p}_v &= -(\nabla_{q_v}\tilde{H})(\tilde{p},\tilde{q})d\sigma - E^{\frac{1}{\ell_p}-\frac{1}{2}}\gamma_v\tilde{p}_v d\sigma + E^{\frac{1}{2\ell_p}-\frac{3}{4}}\sqrt{2T_v\gamma_v}d\tilde{W}_v(\sigma) \ , \end{aligned} \tag{5.37}$$

where $\tilde{W}_v(\sigma) = E^{-\frac{1}{2\ell_p}+\frac{1}{4}}W_v(E^{\frac{1}{\ell_p}-\frac{1}{2}}\sigma)$ is again an n-dimensional Brownian motion. We define, as in §5.1,

$$\tilde{K}_E = \{\tilde{z} : \tilde{H}(\tilde{z}) \leq 4\} \ ,$$

and obtain that on the event $\tilde{A}$, we have $\tilde{z}_\sigma \in \tilde{K}_E$ for all $0 \leq \sigma \leq \lambda$.

By (5.17), there is some $\tilde{C}$ such that if E is large enough, then for all $\tilde{z} \in \tilde{K}_E$,

$$\|\tilde{q}_v\| \leq \tilde{C} \ , \quad \|\delta\tilde{q}_e\| \leq \tilde{C}E^{\frac{1}{\ell_i}-\frac{1}{\ell_p}} \ , \quad \|\tilde{p}_v\| \leq \tilde{C} \ . \tag{5.38}$$

Note that unlike in §5.1, the collection of sets $(\tilde{K}_E)_{E>0}$ is uniformly bounded. In fact, the maximum allowed value of $\delta\tilde{q}_e$ becomes very small at high energy.

Remark 5.19. The difficulty is that the dynamics (5.37) does not converge to a nice limit when E is large. Indeed, we have for any edge $e = (v, v')$ that

$$\nabla_{\tilde{q}_v}(E^{-1}V_e(E^{\frac{1}{\ell_p}}\delta\tilde{q}_e)) \sim E^{\frac{\ell_i}{\ell_p}-1}\|\tilde{\delta}q_e\|^{\ell_i-1} ,$$

which diverges pointwise when $E \to \infty$ if $\delta\tilde{q}_e \neq 0$. The supremum of this quantity over $\tilde{K}_E$ diverges like $E^{1/\ell_p - 1/\ell_i}$ (as can be seen by the scaling in (5.38)). The interpretation is that at high energy and under Assumption 5.18, while the rescaled system behaves like a "tight molecule" with vanishing relative distance $\tilde{\delta}q_e$ between the masses, the dynamics is still dominated by the fast oscillations of the internal degrees of freedom. The way around this is to consider the center of mass coordinates.

The center of mass coordinates in (5.3) are expressed, after rescaling, as

$$\tilde{P}(\sigma) = E^{-\frac{1}{2}}P(E^{\frac{1}{\ell_p}-\frac{1}{2}}\sigma) = E^{-\frac{1}{2}}\sum_{v\in\mathcal{G}} p_v(E^{\frac{1}{\ell_p}-\frac{1}{2}}\sigma) ,$$
$$\tilde{Q}(\sigma) = E^{-\frac{1}{\ell_p}}Q(E^{\frac{1}{\ell_p}-\frac{1}{2}}\sigma) = \frac{1}{|\mathcal{G}|}E^{-\frac{1}{\ell_p}}\sum_{v\in\mathcal{G}} q_v(E^{\frac{1}{\ell_p}-\frac{1}{2}}\sigma) .$$

We denote by $(\tilde{P}_0, \tilde{Q}_0)$ the rescaled initial condition. As the interaction forces cancel out, the dynamics we obtain is

$$\begin{aligned} d\tilde{Q} &= \frac{1}{|\mathcal{G}|}\tilde{P}d\sigma , \\ d\tilde{P} &= -\sum_{v\in\mathcal{G}} E^{\frac{1}{\ell_p}-1}\nabla U_v(E^{\frac{1}{\ell_p}}\tilde{q}_v)d\sigma \\ &\qquad - E^{\frac{1}{\ell_p}-\frac{1}{2}}\sum_{b\in\mathcal{B}}\gamma_b\tilde{p}_b d\sigma + E^{\frac{1}{2\ell_p}-\frac{3}{4}}\sum_{b\in\mathcal{B}}\sqrt{2T_b\gamma_b}d\tilde{W}_b(\sigma) . \end{aligned} \tag{5.39}$$

Moreover, since the graph $(\mathcal{G}, \mathcal{E})$ is connected by Condition C1, we have for all $\tilde{z} \in \tilde{K}_E$ that

$$\max_{v\in\mathcal{G}}\|\tilde{Q} - \tilde{q}_v\| \le \max_{(v,v')\in\mathcal{G}^2}\|\tilde{q}_v - \tilde{q}_{v'}\| \le \sum_{e\in\mathcal{E}}\|\delta\tilde{q}_e\| \le |\mathcal{E}|\tilde{C}E^{\frac{1}{\ell_i}-\frac{1}{\ell_p}} . \tag{5.40}$$

Defining now

$$U_\infty(\tilde{Q}) = \sum_{v\in\mathcal{G}} U_{v,\infty}(\tilde{Q}) , \tag{5.41}$$

we have

Lemma 5.20. *For all* $0 \le |\alpha| \le 1$,

$$\lim_{E\to\infty}\sup_{\tilde{z}\in\tilde{K}_E}\left|\sum_{v\in\mathcal{G}} E^{\frac{|\alpha|}{\ell_p}-1}(D^\alpha U_v)(E^{\frac{1}{\ell_p}}\tilde{q}_v) - (D^\alpha U_\infty)(\tilde{Q})\right| = 0 .$$

Proof. By Condition C3, (5.38) and Remark 2.9 (iii) we have, for all $v \in \mathcal{G}$,

$$\lim_{E\to\infty}\sup_{\tilde{z}\in\tilde{K}_E}\left|E^{\frac{|\alpha|}{\ell_p}-1}(D^\alpha U_v)(E^{\frac{1}{\ell_p}}\tilde{q}_v) - (D^\alpha U_{v,\infty})(\tilde{q}_v)\right| = 0 . \tag{5.42}$$

Moreover, by Condition C3, there exists $c > 0$ such that $D^\alpha U_{v,\infty}$ is c-Lipschitz on the ball $B(0, \tilde{C}) \subset \mathbf{R}^n$, which by (5.38) contains $\tilde{q}_v$ and $\tilde{Q}$ for all $\tilde{z} \in \tilde{K}_E$. Thus,

$$\sup_{\tilde{z}\in\tilde{K}_E}\left|(D^\alpha U_{v,\infty})(\tilde{q}_v) - (D^\alpha U_{v,\infty})(\tilde{Q})\right| \le c\sup_{\tilde{z}\in\tilde{K}_E}\|\tilde{Q} - \tilde{q}_v\| .$$

By (5.40), the right-hand side vanishes when $E \to \infty$. This and (5.42) imply that

$$\lim_{E\to\infty} \sup_{\tilde z \in \tilde K_E} \left| E^{\frac{|\alpha|}{\ell_p}-1} (D^\alpha U_v)(E^{\frac{1}{\ell_p}} \tilde q_v) - (D^\alpha U_{v,\infty})(\tilde Q) \right| = 0 \; .$$

By the definition of U_∞ and the triangle inequality, the proof is complete. □

It is then natural to consider the limiting system

$$\begin{aligned} d\hat Q &= \frac{1}{|\mathcal{G}|} \hat P d\sigma \; , \\ d\hat P &= -\nabla U_\infty(\hat Q) d\sigma \; , \end{aligned} \tag{5.43}$$

which corresponds to the Hamiltonian

$$\hat H(\hat P, \hat Q) = \frac{\hat P^2}{2|\mathcal{G}|} + U_\infty(\hat Q) \; .$$

We can rewrite (5.39) as

$$\begin{aligned} d\tilde Q &= \frac{1}{|\mathcal{G}|} \tilde P d\sigma \; , \\ d\tilde P &= -\nabla U_\infty(\tilde Q) d\sigma + \tilde R(\tilde z) d\sigma - E^{\frac{1}{\ell_p} - \frac12} \sum_{b\in\mathcal{B}} \gamma_b \tilde p_b d\sigma \\ &\quad + E^{\frac{1}{2\ell_p} - \frac34} \sum_{b\in\mathcal{B}} \sqrt{2 T_b \gamma_b} d\tilde W_b(\sigma) \; , \end{aligned} \tag{5.44}$$

where

$$\tilde R(\tilde z) = \nabla U_\infty(\tilde Q) - E^{\frac{1}{\ell_p}-1} \sum_{v\in\mathcal{G}} (\nabla U_v)(E^{\frac{1}{\ell_p}} \tilde q_v) \; ,$$

which by Lemma 5.20 satisfies

$$\lim_{E\to\infty} \sup_{\tilde z \in \tilde K_E} \|\tilde R(\tilde z)\| \to 0 \; . \tag{5.45}$$

Note that the dynamics (5.44) does *not* converge to (5.43) when $\ell_p = 2$, as the dissipative terms in (5.44) remain in the limit. This will complicate the argument slightly (see the proof of Lemma 5.22).

As a consequence of Lemma 5.20, and since $H_c(z_0) > H(z_0)/2$, we have when E is large enough that

$$\hat H(\tilde P_0, \tilde Q_0) \in [1/4, 2] \; .$$

Proposition 5.21. *There is a constant $C > 0$ such that for every initial condition $(\hat P_0, \hat Q_0)$ with $\hat H(\hat P_0, \hat Q_0) \in [1/4, 2]$, the solution of (5.43) satisfies*

$$\sup_{\sigma\in[0,\lambda]} \|\hat Q(\sigma) - \hat Q(0)\| \geq C \; . \tag{5.46}$$

Proof. The left-hand side of (5.46) is obviously strictly positive provided that $\hat H(\hat P_0, \hat Q_0) > 0$. Moreover, the map $(\hat P_0, \hat Q_0) \mapsto \sup_{\sigma\in[0,\lambda]} \|\hat Q(\sigma) - \hat Q(0)\|$ is lower semicontinuous (as the supremum of a family of continuous functions). Thus, since the set $\{(\hat P_0, \hat Q_0) \in \mathbf{R}^{2n} : \hat H(\hat P_0, \hat Q_0) \in [1/4, 2]\}$ is compact, the proof is complete. □

We introduce the random variable

$$X = \sup_{b\in\mathcal{B}} \int_0^\lambda \tilde p_b^2 d\sigma \; .$$

Lemma 5.22. *There exist constants $B, \varepsilon > 0$ such that if E is large enough,*

$$\mathbf{P}_{z_0}\left(\tilde{A} \cap \{X < \varepsilon\}\right) \le e^{-BE} .$$

Proof. In this proof, the constant $c > 0$ may be different each time it appears, and is not allowed to depend on E, provided E is large enough. First, observe that for all $b \in \mathcal{B}$, Hölder's inequality implies that

$$\int_0^\lambda \|\tilde{p}_b\| d\sigma \le c\sqrt{X} . \tag{5.47}$$

Next, by (5.38) and the fact that $\hat{H}$ is conserved by (5.43), we have that on the event $\tilde{A}$, there is some $R > 0$ such that $\|\hat{Q}(\sigma)\|, \|\tilde{Q}(\sigma)\|, \|\hat{P}(\sigma)\|$ and $\|\tilde{P}(\sigma)\|$ are bounded by R for all $0 \le \sigma \le \lambda$. As $\nabla\hat{H}$ is locally Lipschitz by Condition C3, there exists $k > 0$ such that on the event $\tilde{A}$, we have for all $0 \le \sigma \le \lambda$ that

$$\|(\nabla\hat{H})(\hat{P}(\sigma), \hat{Q}(\sigma)) - (\nabla\hat{H})(\tilde{P}(\sigma), \tilde{Q}(\sigma))\| \le k\|(\hat{P}(\sigma), \hat{Q}(\sigma)) - (\tilde{P}(\sigma), \tilde{Q}(\sigma))\| .$$

As a consequence, we can apply Lemma 5.8 to $(\hat{P}, \hat{Q})$ and $(\tilde{P}, \tilde{Q})$ to obtain that

$$\mathbf{1}_{\tilde{A}} \sup_{\sigma\in[0,\lambda]} \|\tilde{Q}(\sigma) - \hat{Q}(\sigma)\| \le cG_E + cE^{\frac{1}{\ell_p} - \frac{1}{2}}\sqrt{X} + cE^{\frac{1}{2\ell_p} - \frac{3}{4}}\eta_E , \tag{5.48}$$

where $\lim_{E\to\infty} G_E = 0$ and where η_E is a non-negative random variable satisfying (5.35).

Pick now any $b \in \mathcal{B}$. By (5.37) and (5.47), we have

$$\sup_{\sigma\in[0,\lambda]} \|\tilde{q}_b(\sigma) - \tilde{q}_b(0)\| \le c\int_0^\lambda \|\tilde{p}_b\| d\sigma \le c\sqrt{X} . \tag{5.49}$$

Moreover, by (5.40), we also find that on $\tilde{A}$,

$$\sup_{\sigma\in[0,\lambda]} \|\tilde{q}_b(\sigma) - \tilde{Q}(\sigma)\| \le cE^{\frac{1}{\ell_i} - \frac{1}{\ell_p}} . \tag{5.50}$$

From (5.49) and (5.50) we deduce that

$$\sup_{\sigma\in[0,\lambda]} \|\tilde{Q}(\sigma) - \tilde{Q}(0)\| \le cE^{\frac{1}{\ell_i} - \frac{1}{\ell_p}} + c\sqrt{X} . \tag{5.51}$$

This together with (5.48) implies that on $\tilde{A}$,

$$\sup_{\sigma\in[0,\lambda]} \|\hat{Q}(\sigma) - \hat{Q}(0)\| \le cE^{\frac{1}{\ell_i} - \frac{1}{\ell_p}} + cG_E + c\sqrt{X}(1 + E^{\frac{1}{\ell_p} - \frac{1}{2}}) + cE^{\frac{1}{2\ell_p} - \frac{3}{4}}\eta_E .$$

But by Proposition 5.21, the left-hand side is bounded below by $C > 0$. Thus,

$$cE^{\frac{1}{2\ell_p} - \frac{3}{4}}\eta_E \ge C - cE^{\frac{1}{\ell_i} - \frac{1}{\ell_p}} - cG_E - c\sqrt{X}(1 + E^{\frac{1}{\ell_p} - \frac{1}{2}}) .$$

We next choose $\varepsilon > 0$ small enough so that for all E large enough,

$$cE^{\frac{1}{\ell_i} - \frac{1}{\ell_p}} + cG_E + c\sqrt{\varepsilon}(1 + E^{\frac{1}{\ell_p} - \frac{1}{2}}) \le \frac{C}{2} .$$

Then, on $\tilde{A}$ and for E large enough, $X < \varepsilon$ implies $cE^{\frac{1}{2\ell_p} - \frac{3}{4}}\eta_E > C/2$, so that

$$\mathbf{P}_{z_0}(\tilde{A} \cap \{X < \varepsilon\}) \le \mathbf{P}_{z_0}\left\{cE^{\frac{1}{2\ell_p} - \frac{3}{4}}\eta_E > C/2\right\} \le e^{-cE^{\frac{3}{2} - \frac{1}{\ell_p}}} .$$

Since $\frac{3}{2} - \frac{1}{\ell_p} \ge 1$, the proof is complete. □

By Lemma 5.22, and since

$$\sum_{b\in\mathcal{B}}\int_0^{\lambda}\gamma_b\tilde{p}_b^2 d\sigma \geq X \inf_{b\in\mathcal{B}}\gamma_b\ ,$$

we obtain (5.36) (for some $\varepsilon > 0$ possibly smaller than that of Lemma 5.22). Thus, the proof of Proposition 5.3 is complete.

References

[1] P. Carmona, *Existence and uniqueness of an invariant measure for a chain of oscillators in contact with two heat baths*, Stochastic Process. Appl. **117** (2007), no. 8, 1076–1092. MR-2340880

[2] F. Conrad and M. Grothaus, *Construction, ergodicity and rate of convergence of N-particle Langevin dynamics with singular potentials*, J. Evol. Equ. **10** (2010), no. 3, 623–662. MR-2674062

[3] B. Cooke, D. P. Herzog, J. C. Mattingly, S. A. McKinley, and S. C. Schmidler, *Geometric ergodicity of two-dimensional Hamiltonian systems with a Lennard-Jones-like repulsive potential*, Commun. Math. Sci. **15** (2017), no. 7, 1987–2025. MR-3717917

[4] N. Cuneo and J.-P. Eckmann, *Controlling general polynomial networks*, Commun. Math. Phys. **328** (2014), no. 3, 1255–1274. MR-3201224

[5] N. Cuneo and J.-P. Eckmann, *Non-equilibrium steady states for chains of four rotors*, Commun. Math. Phys. **345** (2016), no. 1, 185–221. MR-3509013

[6] N. Cuneo, J.-P. Eckmann, and C. Poquet, *Non-equilibrium steady state and subgeometric ergodicity for a chain of three coupled rotors*, Nonlinearity **28** (2015), no. 7, 2397–2421. MR-3366649

[7] N. Cuneo and C. Poquet, *On the relaxation rate of short chains of rotors interacting with Langevin thermostats*, Electron. Commun. Probab. **22** (2017), 8 pp. MR-3666856

[8] A. Dymov, *Asymptotic behaviour of a network of oscillators coupled to thermostats of finite energy*, ArXiv e-prints (2017).

[9] J.-P. Eckmann and M. Hairer, *Non-equilibrium statistical mechanics of strongly anharmonic chains of oscillators*, Commun. Math. Phys. **212** (2000), no. 1, 105–164. MR-1764365

[10] J.-P. Eckmann, C.-A. Pillet, and L. Rey-Bellet, *Entropy production in nonlinear, thermally driven Hamiltonian systems*, J. Stat. Phys. **95** (1999), no. 1-2, 305–331. MR-1705589

[11] J.-P. Eckmann, C.-A. Pillet, and L. Rey-Bellet, *Non-equilibrium statistical mechanics of anharmonic chains coupled to two heat baths at different temperatures*, Commun. Math. Phys. **201** (1999), no. 3, 657–697. MR-1685893

[12] J.-P. Eckmann and E. Zabey, *Strange heat flux in (an)harmonic networks*, J. Stat. Phys. **114** (2004), no. 1-2, 515–523. MR-2032137

[13] M. Grothaus and P. Stilgenbauer, *A hypocoercivity related ergodicity method for singularly distorted non-symmetric diffusions*, Integr. Equ. Oper. Theory **83** (2015), no. 3, 331–379. MR-3413926

[14] M. Hairer, *A probabilistic argument for the controllability of conservative systems*, ArXiv Mathematical Physics e-prints (2005).

[15] M. Hairer, *How hot can a heat bath get?*, Commun. Math. Phys. **292** (2009), no. 1, 131–177. MR-2540073

[16] M. Hairer, *On Malliavin's proof of Hörmander's theorem*, Bull. Sci. Math. **135** (2011), no. 6-7, 650–666. MR-2838095

[17] M. Hairer and J. C. Mattingly, *Slow energy dissipation in anharmonic oscillator chains*, Comm. Pure Appl. Math. **62** (2009), no. 8, 999–1032. MR-2531551

[18] M. Hairer and J. C. Mattingly, *Yet another look at Harris' ergodic theorem for Markov chains*, Seminar on Stochastic Analysis, Random Fields and Applications VI, 2011, pp. 109–117. MR-2857021

[19] M. Hairer, A. M. Stuart, and J. Voss, *Analysis of SPDEs arising in path sampling. II. The nonlinear case*, Ann. Appl. Probab. **17** (2007), no. 5-6, 1657–1706. MR-2358638

[20] R. Z. Has'minskii, *Stochastic stability of differential equations*, Monographs and Textbooks on Mechanics of Solids and Fluids: Mechanics and Analysis, vol. 7, Sijthoff & Noordhoff, Alphen aan den Rijn, 1980, Translated from the Russian by D. Louvish. MR-600653

[21] D. P. Herzog and J. C. Mattingly, *Ergodicity and Lyapunov functions for Langevin dynamics with singular potentials*, ArXiv e-prints (2017).

[22] R. Höpfner, E. Löcherbach, and M. Thieullen, *Ergodicity and limit theorems for degenerate diffusions with time periodic drift. Application to a stochastic Hodgkin-Huxley model*, ESAIM Probab. Stat. **20** (2016), 527–554. MR-3581833

[23] L. Hörmander, *Hypoelliptic second order differential equations*, Acta Math. **119** (1967), 147–171. MR-0222474

[24] B. Leimkuhler, E. Noorizadeh, and F. Theil, *A gentle stochastic thermostat for molecular dynamics*, J. Stat. Phys. **135** (2009), no. 2, 261–277. MR-2505736

[25] J. C. Mattingly, A. M. Stuart, and D. J. Higham, *Ergodicity for SDEs and approximations: locally Lipschitz vector fields and degenerate noise*, Stochastic Process. Appl. **101** (2002), no. 2, 185–232. MR-1931266

[26] S. P. Meyn and R. L. Tweedie, *Markov chains and stochastic stability*, second ed., Cambridge University Press, Cambridge, 2009, With a prologue by Peter W. Glynn. MR-2509253

[27] R. Raquépas, *A note on Harris' ergodic theorem, controllability and perturbations of harmonic networks*, ArXiv e-prints (2018).

[28] D. Revuz and M. Yor, *Continuous martingales and Brownian motion*, third ed., Grundlehren der Mathematischen Wissenschaften, vol. 293, Springer-Verlag, Berlin, 1999. MR-1725357

[29] L. Rey-Bellet, *Ergodic properties of Markov processes*, Open quantum systems. II, Lecture Notes in Math., vol. 1881, Springer, Berlin, 2006, pp. 1–39. MR-2248986

[30] L. Rey-Bellet and L. E. Thomas, *Asymptotic behavior of thermal nonequilibrium steady states for a driven chain of anharmonic oscillators*, Commun. Math. Phys. **215** (2000), no. 1, 1–24. MR-1799873

[31] L. Rey-Bellet and L. E. Thomas, *Exponential convergence to non-equilibrium stationary states in classical statistical mechanics*, Commun. Math. Phys. **225** (2002), no. 2, 305–329. MR-1889227

[32] D. W. Stroock, *Partial differential equations for probabilists*, Cambridge Studies in Advanced Mathematics, vol. 112, Cambridge University Press, Cambridge, 2008. MR-2410225

[33] D. W. Stroock and S. R. S. Varadhan, *On the support of diffusion processes with applications to the strong maximum principle*, Proceedings of the Sixth Berkeley Symposium on Mathematical Statistics and Probability, Vol. III: Probability theory, Univ. California Press, Berkeley, Calif., 1972, pp. 333–359. MR-0400425

Acknowledgments. This work was initiated in the pleasant atmosphere of the Fields Institute, Toronto. Financial support was kindly provided by the ERC Advanced grant 290843:BRIDGES (JPE), the Swiss National Science Foundation Grant 165057 (NC), the National Science Foundation (NSF) grant DMS-1515712 (LRB), the Leverhulme Trust (MH), as well as the ERC Consolidator grant 615897:CRITICAL (MH).