Asymptotic Analysis for the Generalized Langevin Equation with Singular Potentials

Manh Hong Duong[1] · Hung Dang Nguyen[2]

Abstract

We consider a system of interacting particles governed by the generalized Langevin equation (GLE) in the presence of external confining potentials, singular repulsive forces, as well as memory kernels. Using a Mori–Zwanzig approach, we represent the system by a class of Markovian dynamics. Under a general set of conditions on the nonlinearities, we study the large-time asymptotics of the multi-particle Markovian GLEs. We show that the system is always exponentially attractive toward the unique invariant Gibbs probability measure. The proof relies on a novel construction of Lyapunov functions. We then establish the validity of the small-mass approximation for the solutions by an appropriate equation on any finite-time window. Important examples of singular potentials in our results include the Lennard–Jones and Coulomb functions.

Keywords Lyapunov functions · Singular potentials · Small mass limits

Mathematics Subject Classification 60H10

1 Introduction

Given $N \geq 2$, we are interested in the following system of generalized Langevin equations in $\mathbb{R}^d$, $d \geq 1$:

$$\mathrm{d}\, x_i(t) = v_i(t)\mathrm{d}t, \qquad i = 1, \ldots, N,$$

✉ Manh Hong Duong
  h.duong@bham.ac.uk

  Hung Dang Nguyen
  hnguye53@utk.edu

1 School of Mathematics, University of Birmingham, Birmingham, UK

2 Department of Mathematics, University of Tennessee, Knoxville, TN, USA

$$m \, \mathrm{d} v_i(t) = -\gamma v_i(t)\mathrm{d}t - \nabla U(x_i(t))\mathrm{d}t - \sum_{j \neq i} \nabla G\big(x_i(t) - x_j(t)\big)\mathrm{d}t + \sqrt{2\gamma} \, \mathrm{d}W_{i,0}(t)$$

$$+ \int_0^t K_i(t-s)v_i(s)\mathrm{d}s\mathrm{d}t + F_i(t)\mathrm{d}t. \tag{1.1}$$

System (1.1) is introduced to describe the evolution of N interacting micro-particles in a thermally fluctuating viscoelastic medium, see Baczewski and Bond (2013), Bockius et al. (2021), Duong (2015), Duong and Pavliotis (2019), Gomes et al. (2020), Gottwald et al. (2015), Jung et al. (2018, Ness et al. (2015), Wei et al. (2016) and references therein. The bivariate process $(x_i(t), v_i(t))$, $i = 1, \ldots, N$ represents the position and velocity of the ith particle. On the v_i-equation in (1.1), $m > 0$ is the particle's mass, $\gamma > 0$ governs the viscous friction, $\{W_{i,0}\}_{i=1,\ldots,N}$ are independent standard d-dimensional Wiener processes, $U : \mathbb{R}^d \to [0, \infty)$ represents a confining potential satisfying polynomial growth and certain dissipative conditions and $G : \mathbb{R}^d \setminus \{0\} \to \mathbb{R}$ is a singular repulsive potential. Furthermore, the ith particle is subjected to a convolution term involving the convolution kernel $K_i : [0, \infty) \to [0, \infty)$ that characterizes the delayed response of the fluid to the particle's past movement (Kneller 2011; Mason and Weitz 1995; Mori 1965). In accordance with the *fluctuation-dissipation* relationship (Kubo 1966; Zwanzig 2001), the random force $F_i(t)$ is a mean-zero, stationary Gaussian process linked to K_i via the relation

$$\mathbb{E}[F_i(t)F_i(s)] = K_i(|t-s|). \tag{1.2}$$

In the absence of memory effects, that is setting $K_i \equiv 0$ and $F_i \equiv 0$, (1.1) is reduced to the classical underdamped Langevin system modeling Brownian particles driven by repulsive external forces

$$\mathrm{d} x_i(t) = v_i(t)\mathrm{d}t, \quad i = 1, \ldots, N,$$

$$m \, \mathrm{d} v_i(t) = -\gamma v_i(t)\mathrm{d}t - \nabla U(x_i(t))\mathrm{d}t - \sum_{j \neq i} \nabla G(x_i(t) - x_j(t))\mathrm{d}t + \sqrt{2\gamma} \, \mathrm{d}W_{i,0}(t).$$

$$\tag{1.3}$$

In particular, the large-time asymptotic of (1.3) is well-understood. That is under a wide class of polynomial potential U and singular potential G, including the instances of Lennard–Jones and Coulomb functions, system (1.3) admits a unique invariant probability measure which is exponentially attractive and whose formula is given by Conrad and Grothaus (2010), Cooke et al. (2017), Grothaus and Stilgenbauer (2015), Herzog and Mattingly (2019), Lu and Mattingly (2019)

$$\pi(\mathrm{dx}, \mathrm{dv}) = \frac{1}{Z} \exp\left\{ -\left(\frac{1}{2}\sum_{i=1}^N m|v_i|^2 + \sum_{i=1}^N U(x_i) + \sum_{1 \leq i < j \leq N} G(x_i - x_j)\right)\right\}\mathrm{dx}\mathrm{dv}.$$

In the above, Z is the normalization constant, $\mathrm{x} = (x_1, \ldots, x_N)$ and $\mathrm{v} = (v_1, \ldots, v_N)$. However, as pointed out elsewhere in Kubo (1966), Zwanzig (2001), the presence of

elasticity in a viscoelastic medium induces a memory effect between the motion of the particles and the surrounding molecular bombardment. It is thus more physically relevant to consider (1.1). On the other hand, in the absence of singularities in (1.1) ($G \equiv 0$), there is a vast literature in the context of, e.g., large-time behaviors (Glatt-Holtz et al. 2020; Herzog et al. 2023; Ottobre and Pavliotis 2011; Pavliotis 2014; Pavliotis et al. 2021) as well as small-mass limits (Herzog et al. 2016; Hottovy et al. 2015; Lim and Wehr 2019; Lim et al. 2020; Nguyen 2018; Shi and Wang 2021). In contrast, much less is known about the system (1.1) in the presence of both memory kernels and singular potentials for any of those limiting regimes.

The main goal of the present article is thus twofold. Firstly, under a general set of conditions on the nonlinearities and memory kernels, we asymptotically characterize the equilibrium of (1.1) when $t \to \infty$. More specifically, we aim to prove that under these practical assumptions, (1.1) is exponentially attractive toward a unique ergodic probability measure. Secondly, we explore the behaviors of (1.1) in the small-mass regime, i.e., by taking m to zero on the right-hand side of the v_i-equation in (1.1). Due to the singular limit when m is small, the velocity $v(t)$ is oscillating fast, whereas the position $x(t)$ is still moving slow. Hence, we seek to identify a limiting process $q(t)$ such that $x(t)$ can be related to $q(t)$ on any finite-time window. We now provide a more detailed description of the main results.

1.1 Geometric Ergodicity

In general, there is no Markovian dynamics associated with (1.1), owing to the presence of the memory kernels. Nevertheless, it is well known that for kernels which are a sum of exponential functions, we may adopt the Mori–Zwanzig approach to produce a Markovian approximation to (1.1) (Glatt-Holtz et al. 2020; Kubo 1966; Mori 1965; Ottobre and Pavliotis 2011; Pavliotis 2014; Zwanzig 2001). More specifically, when K_i is given by

$$K_i(t) = \sum_{\ell=1}^{k_i} \lambda_{i,\ell}^2 e^{-\alpha_{i,\ell} t}, \quad t \geq 0, \tag{1.4}$$

for some positive constants $\lambda_{i,\ell}, \alpha_{i,\ell}, \ell = 1, \ldots, k_i$, following the framework of Baczewski and Bond (2013), Doob (1942), Duong and Shang (2022), Goychuk (2012), Ottobre and Pavliotis (2011), Pavlioti (2014), we can rewrite (1.1) as the following system

$$\mathrm{d}\, x_i(t) = v_i(t)\mathrm{d}t, \quad i = 1, \ldots, N,$$

$$m\, \mathrm{d}\, v_i(t) = -\gamma v_i(t)\mathrm{d}t - \nabla U(x_i(t))\mathrm{d}t + \sqrt{2\gamma}\, \mathrm{d}W_{i,0}(t)$$

$$- \sum_{j \neq i} \nabla G(x_i(t) - x_j(t))\mathrm{d}t + \sum_{\ell=1}^{k_i} \lambda_{i,\ell} z_{i,\ell}(t)\mathrm{d}t,$$

$$\mathrm{d}\, z_{i,\ell}(t) = -\alpha_{i,\ell} z_{i,\ell}(t)\mathrm{d}t - \lambda_{i,\ell} v_i(t)\mathrm{d}t + \sqrt{2\alpha_{i,\ell}}\, \mathrm{d}W_{i,\ell}(t), \quad \ell = 1, \ldots, k_i. \tag{1.5}$$

See Pavliotis (2014, Proposition 8.1) for a detailed discussion of this formulation.

Denoting $z_i = (z_{i,1}, \ldots, z_{i,k_i}) \in (\mathbb{R}^d)^{k_i}, i = 1, \ldots, N$, we introduce the Hamiltonian function H_N defined as

$$H_N(\mathrm{x}, \mathrm{v}, z_1, \ldots, z_N) = \frac{1}{2}m|\mathrm{v}|^2 + \sum_{i=1}^{N} U(x_i) + \sum_{1 \le i < j \le N} G(x_i - x_j) + \frac{1}{2}\sum_{i=1}^{N} |z_i|^2. \tag{1.6}$$

The corresponding Gibbs measure is given by:

$$\pi_N(\mathrm{x}, \mathrm{v}, z_1, \ldots, z_N) = \frac{1}{Z_N} \exp\left\{-H_N(\mathrm{x}, \mathrm{v}, z_1, \ldots, z_N)\right\} \mathrm{dxdvdz}_1 \ldots \mathrm{dz}_N, \tag{1.7}$$

where Z_N is the normalization constant. Under suitable assumptions on the potentials, cf. Assumptions 2.1 and 2.3, one can rely on the Hamiltonian structure to show that (1.5) is always well-posed. That is a strong solution of (1.5) exists and is unique for all finite time. Furthermore, by a routine computation (Pavliotis 2014, Proposition 8.2), it is not difficult to see that the Gibbs measure π_N as in (1.7) is an invariant probability measure of (1.5) and that there is dissipation toward such a measure. On the other hand, as observed elsewhere in Conrad and Grothaus (2010), Cooke et al. (2017), Grothaus and Stilgenbauer (2015), Herzog and Mattingly (2019), Lu and Mattingly (2019), in the presence of the nonlinearities, the Hamiltonian (1.6) does not produce an energy estimate of the form

$$\frac{\mathrm{d}}{\mathrm{d}t}\mathbb{E}[V(t)] \le -c\,\mathbb{E}[V(t)] + C, \quad t \ge 0,$$

which is needed to obtain geometric ergodicity. In this paper, we tackle the problem by exploiting the technique of Lyapunov function, cf. Definition 3.1, and successfully establish the uniqueness of π_N as well as an exponential convergent rate toward π_N. We note that our result covers important examples of singular potentials such as Lennard–Jones functions and Coulomb functions. We refer the reader to Theorem 2.8 for a precise statement of this result and to Sect. 3 for its proof.

Historically, in the absence of repulsive forces, (1.5) is reduced to the following single-particle GLE

$$\mathrm{d}\,x(t) = v(t)\mathrm{d}t,$$

$$m\,\mathrm{d}\,v(t) = -\gamma v(t)\mathrm{d}t - \nabla U(x(t))\mathrm{d}t + \sum_{i=1}^{k} \lambda_i z_i(t)\mathrm{d}t + \sqrt{2\gamma}\,\mathrm{d}W_0(t),$$

$$\mathrm{d}\,z_i(t) = -\alpha_i z_i(t)\mathrm{d}t - \lambda_i v(t)\mathrm{d}t + \sqrt{2\alpha_i}\,\mathrm{d}W_i(t), \quad i = 1, \ldots, k, \tag{1.8}$$

whose large-time asymptotic has been studied extensively (Glatt-Holtz et al. 2020; Herzog et al. 2023; Ottobre and Pavliotis 2011; Pavliotis 2014). Particularly, mixing rates for the kernel instances of finite sum of exponentials were established in Ottobre and Pavliotis (2011, Pavliotis (2014) via the weak Harris Theorem (Hairer and

Mattingly 2011; Meyn and Tweedie 2012). Analogously, kernel instances as an infinite sum of exponentials were explored in Glatt-Holtz et al. (2020). In this work, the uniqueness of invariant probability measures was obtained by employing the so-called asymptotic coupling argument. Recently in Herzog et al. (2023), a Gibbsian approach was adopted from Liu (2002), Mattingly and Sinai (2001), Mattingly (2002) to study more general kernels that need not be a sum of exponentials.

Turning back to our ergodicity results for singular potentials, the proof of Theorem 2.8 relies on two ingredients: an irreducibility condition and a suitable Lyapunov function (Hairer and Mattingly 2011; Mattingly et al. 2002; Meyn and Tweedie 2012). Whereas the irreducibility is relatively standard and can be addressed by adapting to the argument in, e.g., Herzog and Mattingly (2019, Corollary 5.12) and Mattingly et al. (2002, Proposition 2.5), the construction of Lyapunov functions is highly nontrivial requiring a deeper understanding of the dynamics. Notably, in case the viscous drag is positive ($\gamma > 0$), we draw upon the ideas developed in Herzog and Mattingly (2019), Lu and Mattingly (2019) tailored to our settings, cf. Lemma 3.5. On the other hand, when $\gamma = 0$, it is worth pointing out that the method therein is unfortunately not applicable owing to a lack of dissipation in v. To circumvent the issue, we effectively build up a novel Lyapunov function specifically designed for (1.5), cf. Lemma 3.6. The main idea of the construction is to realize the dominating effects at high energy ($H_N \gg 1$). We do so by employing a heuristic asymptotic scaling allowing for determining the leading order terms in "bad" regions, i.e., when $|x| \to 0$ and $|v|, |z| \to \infty$. In turn, this will be crucial in the derivation of our Lyapunov functions that ultimately will be invoked to conclude ergodicity when $\gamma = 0$. The heuristic argument will be presented out in Sect. 3.1.1, whereas the proofs of Lemmas 3.5 and 3.6 are supplied in Sect. 3.2.

1.2 Small-Mass Limit

In the second main topic of the paper, we investigate the small-mass limit for the process $x(t) = x_m(t)$ in (1.5). Namely, by taking m to zero on the right-hand side of the v_i-equation in (1.5), we aim to derive a process $q(t)$ taking values in $(\mathbb{R}^d)^N$ such that $x_m(t)$ can be well-approximated by $q(t)$ on any finite-time window, i.e.,

$$\sup_{t \in [0,T]} |x_m(t) - q(t)| \to 0, \quad \text{as } m \to 0, \tag{1.9}$$

where the limit holds in an appropriate sense. In the literature, such a statement is also known as the Smoluchowski–Kramer approximation (Freidlin 2004; Kramers 1940; Smoluchowski 1916).

We note that in the absence of the singularities, there is a vast literature on limits of the form (1.9) for various settings of the single-particle GLE as well as other second-order systems. For examples, numerical simulations were performed in Hottovy et al. (2012a, b). Rigorous results in this direction for state-dependent drift terms appear in the work of Cerrai et al. (2020), Hottovy et al. (2015, 2012a), Pardoux and Veretennikov (2003). Similar results are also established for Langevin dynamics (Herzog et al. 2016; Duong et al. 2017), finite-dimensional single-particle GLE (Lim and Wehr 2019; Lim

et al. 2020), as well as infinite-dimensional single-particle GLE (Nguyen 2018; Shi and Wang 2021). Analogous study for the stochastic wave equation was central in the work of Cerrai and Freidlin (2006a, 2006b), Cerrai et al. (2017, Cerrai and Glatt-Holtz (2020, 2014), Cerrai and Salins (2016), Nguyen (2022). On the other hand, small-mass limits in the context of repulsive forces are much less studied, but see the recent paper (Choi and Tse 2022) for a quantification of the small-mass limit of the kinetic Vlasov–Fokker–Planck equations with singularities. In our work, we investigate this problem under a general set of conditions on the nonlinear potentials. More specifically, let us introduce the following system

$$\gamma dq_i(t) = -\nabla U(q_i(t))dt - \sum_{j \neq i} \nabla G(q_i(t) - q_j(t))dt + \sqrt{2\gamma}dW_{i,0}(t)$$

$$- \sum_{\ell=1}^{k_i} \lambda_{i,\ell}^2 q_i(t)dt + \sum_{\ell=1}^{k_i} \lambda_{i,\ell} f_{i,\ell}(t)dt, \qquad i = 1, \ldots, N,$$

$$df_{i,\ell}(t) = -\alpha_{i,\ell} f_{i,\ell}(t) + \lambda_{i,\ell} \alpha_{i,\ell} q_i(t)dt + \sqrt{2\alpha_{i,\ell}}dW_{i,\ell}(t), \quad \ell = 1, \ldots, k_i. \tag{1.10}$$

Our second main result states that the process $q(t)$ satisfies the following limit in probability for all $\xi, T > 0$

$$\mathbb{P}\left(\sup_{t \in [0,T]} |x_m(t) - q(t)| > \xi \right) \to 0, \quad \text{as } m \to 0. \tag{1.11}$$

The precise statement of (1.11) is provided in Theorem 2.10, while its detailed proof is supplied in Sect. 4.

In order to derive the limiting system (1.10), we will adopt the framework developed in Herzog et al. (2016), Nguyen (2018) dealing with the same issue in the absence of singular potentials. This involves exploiting the structure of the z_i-equation in (1.5) while making use of Duhamel's formula and an integration by parts. In turn, this allows for completely decoupling z_i from v_i, ultimately arriving at (1.10). See Sect. 4.1 for a further discussion of this point. The proof of (1.11) draws upon the argument in Herzog et al. (2016), Nguyen (2018), Ottobre and Pavliotis (2011) tailored to our settings. Namely, we first reduce the general problem to the special case when the nonlinearities are assumed to be Lipschitz. This requires a careful analysis on the auxiliary memory variables $z_i(t), i = 1, \ldots, N$, as well as the velocity process $v_m(t)$. We then proceed to remove the Lipschitz constraint by making use of crucial moment bounds on the limiting process $q(t)$. In turn, this relies on a delicate estimate on (1.10) via suitable Lyapunov functions, which are also of independent interest. Similarly to the ergodicity results, we note that the limit (1.11) is applicable to a wide range of repulsive forces, e.g., the Lennard–Jones and Coulomb functions. To the best of the authors' knowledge, the type of limit (1.11) that we establish in this work seems to be the first in this direction for stochastic systems with memory and singularities. The explicit argument for (1.11) will be carried out while making use of a series of auxiliary results in Sect. 4. The proof of Theorem 2.10 will be presented in Sect. 4.4.

Finally, we remark that a crucial property that is leveraged in this work is the choice of memory kernels as a finite sum of exponentials, cf. (1.4). This allows for the convenience of employing Markovian framework to study asymptotic analysis for (1.1) (Mori 1965; Ottobre and Pavliotis 2011; Pavliotis 2014; Zwanzig 2001) under the impact of interacting repulsive forces. For SDEs with non-exponentially decaying kernels, e.g., sub-exponential or power-law, with smooth nonlinearities though, we refer the reader to Baeumer et al. (2015, Desch and Londen (2011, Glatt-Holtz et al. (2020), Herzog et al. (2023, Nguyen (2018).

1.3 Organization of the Paper

The rest of the paper is organized as follows: in Sect. 2, we introduce the notations as well as the assumptions that we make on the nonlinearities. We also state the main results of the paper, including Theorem 2.8 on geometric ergodicity of (1.5) and Theorem 2.10 on the validity of the approximation of (1.5) by (1.10) in the small-mass regime. In Sect. 3, we address the construction of Lyapunov functions for (1.5) and prove Theorem 2.8. In Sect. 4, we detail a series of auxiliary results that we employ to prove the convergence of (1.5) toward (1.10). We also conclude Theorem 2.10 in this section. In Appendix A, we provide useful estimates on singular potentials that are exploited to establish the main results.

2 Assumptions and Main Results

Throughout, we let $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t \geq 0}, \mathbb{P})$ be a filtered probability space satisfying the usual conditions (Karatzas and Shreve 2012) and $(W_{i,j}(t))$, $i = 1, \ldots, N$, $j = 0, \ldots, k_i$, be i.i.d standard d-dimensional Brownian Motions on $(\Omega, \mathcal{F}, \mathbb{P})$ adapted to the filtration $(\mathcal{F}_t)_{t \geq 0}$.

In Sect. 2.1, we detail sufficient conditions on the nonlinearities U and G that we will employ throughout the analysis. We also formulate the well-posedness through Proposition 2.7. In Sect. 2.2, we state the first main result through Theorem 2.8 giving the uniqueness of the invariant Gibbs measure π_N defined in (1.7), as well as the exponential convergent rate toward π_N in suitable Wasserstein distances. In Sect. 2.3, we provide our second main result through Theorem 2.10 concerning the validity of the small mass limit of (1.5) on any finite-time window.

2.1 Main Assumptions

For notational convenience, we denote the inner product and the norm in $\mathbb{R}^d$ by $\langle \cdot, \cdot \rangle$ and $|\cdot|$, respectively. Concerning the potential U, we will impose the following condition (Ottobre and Pavliotis 2011; Pavliotis 2014).

Assumption 2.1 (i) $U \in C^\infty(\mathbb{R}^d; [1, \infty))$ satisfies

$$|U(x)| \leq a_1(1 + |x|^{\lambda+1}), \quad |\nabla U(x)| \leq a_1(1 + |x|^\lambda), \quad x \in \mathbb{R}^d, \qquad (2.1)$$

and

$$|\nabla^2 U(x)| \leq a_1(1 + |x|^{\lambda-1}), \tag{2.2}$$

for some constants $a_1 > 0$ and $\lambda \geq 1$.

(ii) Furthermore, there exist positive constants a_2, a_3 such that

$$\langle \nabla U(x), x \rangle \geq a_2|x|^{\lambda+1} - a_3, \quad x \in \mathbb{R}^d. \tag{2.3}$$

(iii) If $\gamma = 0$, then $\lambda = 1$.

Remark 2.2 The first two conditions (i) and (ii) in Assumption 2.1 are quite popular and can be found in many previous literature for the Langevin dynamics (Glatt-Holtz et al. 2020; Mattingly et al. 2002; Nguyen 2018). In particular, $U(x)$ essentially behaves like $|x|^{\lambda+1}$ at infinity. On the other hand, for the generalized Langevin counterpart, in the absence of the viscous drag, i.e., $\gamma = 0$ in $v-$equation in (1.5), we have to impose condition (iii) Ottobre and Pavliotis (2011) requiring U be essentially a quadratic potential. We note, however, that this condition is unnecessary for the well-posedness. Rather, it is to guarantee the existence of suitable Lyapunov functions, so as to ensure geometric ergodicity of (1.5). See the proofs of Lemmas 3.4 and 3.6 for a further discussion of this point.

Concerning the singular potential G, we will make the following condition (Bolley et al. 2018; Herzog and Mattingly 2019; Lu and Mattingly 2019).

Assumption 2.3 (i) $G \in C^\infty(\mathbb{R}^d \backslash \{0\}; \mathbb{R})$ satisfies $G(x) \rightarrow \infty$ as $|x| \rightarrow 0$. Furthermore, there exists a positive constant $\beta_1 \geq 1$ such that for all $x \in \mathbb{R}^d \backslash \{0\}$

$$|G(x)| \leq a_1\left(1 + |x| + \frac{1}{|x|^{\beta_1}}\right), \tag{2.4}$$

$$|\nabla G(x)| \leq a_1\left(1 + \frac{1}{|x|^{\beta_1}}\right), \tag{2.5}$$

$$\text{and} \quad |\nabla^2 G(x)| \leq a_1\left(1 + \frac{1}{|x|^{\beta_1+1}}\right), \tag{2.6}$$

where a_1 is the constant as in Assumption 2.1.

(ii) There exist positive constants $\beta_2 \in [0, \beta_1)$, a_4, a_5 and a_6 such that

$$\left| \nabla G(x) + a_4 \frac{x}{|x|^{\beta_1+1}} \right| \leq \frac{a_5}{|x|^{\beta_2}} + a_6, \quad x \in \mathbb{R}^d \backslash \{0\}. \tag{2.7}$$

Remark 2.4 (i) Although (2.7) is slightly odd looking, this condition essentially states that ∇G can be expressed as:

$$\nabla G(x) = -c\frac{x}{|x|^{\beta_1+1}} + \text{lower-order terms},$$

hence the requirement $\beta_2 < \beta_1$. Also, ∇G has the order of $|x|^{-\beta_1}$ as $|x| \to 0$, i.e.,

$$\frac{c}{|x|^{\beta_1}} - C \leq |\nabla G(x)| \leq \frac{C}{|x|^{\beta_1}} + C,$$

which need not imply (2.7). Condition (2.7) will be employed in the Lyapunov proofs in Sect. 3.

(ii) A routine calculation shows that both the Coulomb functions

$$G(x) = \begin{cases} -\log|x|, & d = 2, \\ \frac{1}{|x|^{d-2}}, & d \geq 3, \end{cases} \tag{2.8}$$

and the Lennard–Jones functions

$$G(x) = \frac{c_0}{|x|^{12}} - \frac{c_1}{|x|^6}, \tag{2.9}$$

satisfy (2.7). Particularly, in case of (2.9), we have

$$\nabla G(x) = -12c_0 \frac{x}{|x|^{14}} + 6c_1 \frac{x}{|x|^8},$$

which verifies (2.7) with $\beta_1 = 13$ and $\beta_2 = 7$. Another well-known example for the case $\beta_1 = 1$ is the log function $G(x) = -\log|x|$, whereas the case $\beta_1 > 1$ includes the instance $G(x) = |x|^{-\beta_1+1}$.

(iiI) Without loss of generality, we may assume that

$$\sum_{i=1}^{N} U(x_i) + \sum_{1 \leq i < j \leq N} G(x_i - x_j) \geq 0. \tag{2.10}$$

Otherwise, we may replace $U(x)$ by $U(x) + c$ for some sufficiently large constant c, which does not affect (1.5).

Under the above two assumptions, we are able to establish the geometric ergodicity of (1.5), cf. Theorem 2.8, in any dimension $d \geq 1$. They are also sufficient for the purpose of investigating the small mass limits when either $d \geq 2$ or $\beta_1 > 1$. On the other hand, when $d = 1$ and $\beta_1 = 1$ (e.g., log potentials), we will impose the following extra condition on G.

Assumption 2.5 Let G and β_1 be as in Assumption 2.3. In dimension $d = 1$, if $\beta_1 = 1$, then

$$a_4 > \frac{1}{2}, \tag{2.11}$$

where a_4 is the positive constant from condition (2.7).

Remark 2.6 As it turns out, in dimension $d = 1$, log potentials ($\beta_1 = 1$) induce further difficulty for the small-mass limit. To circumvent the issue, we have to impose Assumption 2.5, which requires that the repulsive force be strong enough so as to establish suitable energy estimates. See the proof of Lemma 4.3 for a detailed explanation of this point. It is also worth mentioning that the threshold $a_4 > 1/2$ is a manifestation of Lemma A.2 and is perhaps far from optimality. Thus, the small-mass limit in the case $d = 1$ and $\beta_1 = 1$, cf. Theorem 2.10, remains an open problem for all $a_4 \in (0, 1/2]$.

Having introduced sufficient conditions on the potentials, we turn to the issue of well-posedness for (1.5). The domain where the displacement process $x_m(t)$ evolves on is denoted by $\mathcal{D}$ and is defined as Bolley et al. (2018, Herzog and Mattingly (2019), Lu and Mattingly (2019)

$$\mathcal{D} = \{x = (x_1, \ldots, x_N) \in (\mathbb{R}^d)^N : x_i \neq x_j \text{ if } i \neq j\}. \tag{2.12}$$

Then, we define the phase space for the solution of (1.5) as follows:

$$\mathbf{X} = \mathcal{D} \times (\mathbb{R}^d)^N \times \prod_{i=1}^{N} (\mathbb{R}^d)^{k_i}. \tag{2.13}$$

The first result of this paper is the following well-posedness result ensuring the existence and uniqueness of strong solutions to system (1.5).

Proposition 2.7 *Under Assumptions 2.1 and 2.3, for every initial condition* $X_0 = (x(0), v(0), z_1(0), \ldots, z_N(0)) \in \mathbf{X}$, *system* (1.5) *admits a unique strong solution* $X_m(t; X_0) = \big(x_m(t), v_m(t), z_{1,m}(t), \ldots, z_{N,m}(t)\big) \in \mathbf{X}$.

The proof of Proposition 2.7 is a consequence of the existence of Lyapunov functions below in Lemmas 3.5 and 3.6. The argument can be adapted from the well-posedness proof of Glatt-Holtz et al. (2020, Section 3) tailored to our settings. Alternatively, the result can also be proved by using the Hamiltonian as in (1.6) to establish suitable moment bounds. It, however, should be noted that Proposition 2.7 itself is highly nontrivial as well-posedness does not automatically follow a standard argument for SDEs (Khasminskii 2011) based on locally Lipschitz continuity, owing to the presence of singular potentials. Nevertheless, the issue can be tackled by constructing appropriate Lyapunov functions.

As a consequence of the well-posedness, we can thus introduce the Markov transition probabilities of the solution $X_m(t)$ by

$$P_t^m(X_0, A) := \mathbb{P}(X_m(t; X_0) \in A),$$

which are well-defined for $t \geq 0$, initial condition $X_0 \in \mathbf{X}$ and Borel sets $A \subset \mathbf{X}$. Letting $\mathcal{B}_b(\mathbf{X})$ denote the set of bounded Borel measurable functions $f : \mathbf{X} \to \mathbb{R}$, the associated Markov semigroup $P_t^m : \mathcal{B}_b(\mathbf{X}) \to \mathcal{B}_b(\mathbf{X})$ is defined and denoted by

$$P_t^m f(X_0) = \mathbb{E}[f(X_m(t; X_0))], \quad f \in \mathcal{B}_b(\mathbf{X}).$$

2.2 Geometric Ergodicity

We now turn to the topic of large-time properties of equation (1.5). Recall that a probability measure μ on Borel subsets of $\mathbf{X}$ is called *invariant* for the semigroup P_t^m if for every $f \in \mathcal{B}_b(\mathbf{X})$

$$\int_{\mathbf{X}} f(X)(P_t^m)^* \mu(\mathrm{d}X) = \int_{\mathbf{X}} f(X)\mu(\mathrm{d}X),$$

where $(P_t^m)^*\mu$ is defined as in Hairer and Mattingly (2011)

$$(P_t^m)^*\mu(A) = \int_{\mathbf{X}} P_t^m(X, A)\mu(\mathrm{d}X),$$

for all Borel sets $A \subset \mathbf{X}$. Next, we denote by $\mathcal{L}_{m,\gamma}^N$ the generator associated with (1.5). One defines $\mathcal{L}_{m,\gamma}^N$ for any $\varphi \in C^2(\mathbf{X}; \mathbb{R})$ by

$$\mathcal{L}_{m,\gamma}^N \varphi = \sum_{i=1}^{N} \left[\langle \partial_{x_i}\varphi, v_i \rangle + \frac{1}{m}\Big\langle \partial_{v_i}\varphi, -\gamma v_i - \nabla U(x_i) - \sum_{j \neq i} \nabla G(x_i - x_j) + \sum_{\ell=1}^{k_i} \lambda_{i,\ell} z_{i,\ell} \Big\rangle \right.$$
$$\left. + \frac{\gamma}{m^2}\Delta_{v_i}\varphi + \sum_{\ell=1}^{k_i} \langle \partial_{z_{i,\ell}}\varphi, -\alpha_{i,\ell} z_{i,\ell} - \lambda_{i,\ell} v_i \rangle + \sum_{\ell=1}^{k_i} \alpha_{i,\ell} \Delta_{z_{i,\ell}}\varphi \right]. \tag{2.14}$$

In order to show that $\pi_N(X)\mathrm{d}X$ defined in (1.7) is invariant for (1.5), it suffices to show that $\pi_N(X)$ is a solution of the stationary Fokker–Planck equation

$$(\mathcal{L}_{m,\gamma}^N)^* \pi_N(X) = 0, \tag{2.15}$$

where $(\mathcal{L}_{m,\gamma}^N)^*$ is the dual of $\mathcal{L}_{m,\gamma}^N$, i.e.,

$$\int_{\mathbf{X}} \mathcal{L}_{m,\gamma}^N f_1(X) \cdot f_2(X)\mathrm{d}X = \int_{\mathbf{X}} f_1(X) \cdot (\mathcal{L}_{m,\gamma}^N)^* f_2(X)\mathrm{d}X,$$

for any $f_1, f_2 \in C_c^2(\mathbf{X}; \mathbb{R})$. In the absence of the singular potential G, this approach was previously employed in Pavliotis (2014) for a finite-dimensional GLE and in Glatt-Holtz et al. (2020) for an infinite-dimensional GLE. With regard to (2.15), we may simply adapt to the proof of Pavliotis (2014, Proposition 8.2) tailored to our setting with the appearance of G.

Concerning the unique ergodicity of π_N, we will work with suitable Wasserstein distances allowing for the convenience of measuring the convergent rate toward equilibrium. For a measurable function $V : \mathbf{X} \to (0, \infty)$, we introduce the following weighted supremum norm

$$\|\varphi\|_V := \sup_{X \in \mathbf{X}} \frac{|\varphi(X)|}{1 + V(X)}.$$

We denote by $\mathcal{M}_V$ the collection of probability measures μ on Borel subsets of $\mathbf{X}$ such that

$$\int_{\mathbf{X}} V(X)\mu(\mathrm{d}X) < \infty.$$

Let $\mathcal{W}_V$ be the corresponding weighted total variation distance in $\mathcal{M}_V$ associated with $\|\cdot\|_V$, given by

$$\mathcal{W}_V(\mu_1, \mu_2) = \sup_{\|\varphi\|_V \leq 1} \left| \int_{\mathbf{X}} \varphi(X)\mu_1(\mathrm{d}X) - \int_{\mathbf{X}} \varphi(X)\mu_2(\mathrm{d}X) \right|.$$

We remark that $\mathcal{W}_V$ is a Wasserstein distance. Indeed, by the dual Kantorovich theorem

$$\mathcal{W}_V(\mu_1, \mu_2) = \inf \mathbb{E}\Big[\big((1 + V(X_1)) + (1 + V(X_2))\big)\mathbf{1}\{X_1 \neq X_2\}\Big],$$

where the infimum runs over all pairs of random variables (X_1, X_2) such that $X_1 \sim \mu_1$ and $X_2 \sim \mu_2$. We refer the reader to the monograph (Villani 2021) for a detailed account of Wasserstein distances and optimal transport problems. With this setup, we now state the first main result of the paper, establishing the unique ergodicity of π_N defined in (1.7) as well as the exponential convergent rate toward π_N.

Theorem 2.8 *Under Assumptions 2.1 and 2.3, for every $m > 0$ and $\gamma \geq 0$, the probability measure π_N defined in (1.7) is the unique invariant probability measure for (1.5). Furthermore, there exists a function $V \in C^2(\mathbf{X}; [1, \infty))$ such that for all $\mu \in \mathcal{M}_V$,*

$$\mathcal{W}_V\big((P_t^m)^*\mu, \pi_N\big) \leq Ce^{-ct}\mathcal{W}_V(\mu, \pi_N), \quad t \geq 0, \tag{2.16}$$

for some positive constants c and C independent of μ and t.

In order to establish Theorem 2.8, we will draw upon the framework of Bolley et al. (2018), Hairer and Mattingly (2011), Herzog and Mattingly (2019), Lu and Mattingly (2019), Mattingly et al. (2002), Meyn and Tweedie (2012) tailored to our settings. The argument relies on two crucial ingredients: a suitable Lyapunov function, cf. Definition 3.1, and a minorization condition, cf. Definition 3.2. Since it is not difficult to see that the system (1.5) satisfies hypoellipticity (Ottobre and Pavliotis 2011; Pavliotis 2014), we may employ a relatively standard argument (Mattingly et al. 2002; Pavliotis 2014) to establish the minorization. On the other hand, constructing Lyapunov functions is quite involved requiring a deeper understanding of the dynamics in the presence of the singular potentials. Particularly, while the construction in the case $\gamma > 0$ can be adapted to those in the previous work of Herzog and Mattingly (2019), Lu and Mattingly (2019), the absence of the viscous drag $\gamma = 0$ induces further difficulty owing to the interaction between the singular potentials, the velocity v and the auxiliary variables $\{z_i\}_{i=1,\ldots,N}$. To overcome this issue, we will follow (Athreya et al. 2012; Cooke et al. 2017; Herzog and Mattingly 2015a, b, 2019; Lu and Mattingly 2019) and

perform the technique of *asymptotic scaling* to determine the leading-order terms in the dynamics at large energy states. All of this will be carried out in Sect. 3. The proof of Theorem 2.8 will be given in Sect. 3.3.

2.3 Small-Mass Limit

We now consider the topic of small-mass limit and rigorously compare the solution of (1.5) with that of (1.10) as $m \to 0$.

As mentioned in Sect. 1.2, the derivation of (1.10) follows the framework of Cerrai et al. (2020), Herzog et al. (2016), Lim and Wehr (2019, Lim et al. (2020) for finite-dimensional second-order systems as well as Nguyen (2018, Shi and Wang (2021) for infinite-dimensional systems. The trick employed involves an integration by parts on the z-equation of (1.5), so as to decouple the velocity from the other processes. As a result, it induces extra drift terms appearing on the right-hand side of the q-equation in (1.10). Also, the initial conditions of (1.10) are closely related to those of (1.5). For the sake of clarity, we defer the explanation in detail to Sect. 4.1.

Due to the presence of singular potentials, the first issue arising from (1.10) is the well-posedness. To this end, we introduce the new phase space for the solutions of (1.10) defined as

$$\mathbf{Q} = \mathcal{D} \times \prod_{i=1}^{N} (\mathbb{R}^d)^{k_i},$$

where we recall $\mathcal{D}$ being the state space for the displacement as in (2.12). The existence and uniqueness of a strong solution in $\mathbf{Q}$ for (1.10) are guaranteed in the following auxiliary result.

Proposition 2.9 *Under Assumptions* 2.1 *and* 2.3, *for every initial condition* $Q_0 = (q(0), f_1(0), \ldots, f_N(0)) \in \mathbf{Q}$, *system* (1.10) *admits a unique strong solution* $Q(t; Q_0) = (q(t), f_1(t), \ldots, f_N(t)) \in \mathbf{Q}$.

Similar to the proof of Proposition 2.7, the argument of Proposition 2.9 follows from the energy estimate established in Lemma 4.3 in Sect. 4.3. In turn, the result in Lemma 4.3 relies on suitable Lyapunov function specifically designed for (1.10). It will also be employed to study the small-mass limit.

We now state the second main result of the paper giving the validity of the approximation of (1.5) by (1.10) in the small-mass regime.

Theorem 2.10 *Suppose that Assumptions* 2.1, 2.3 *and* 2.5 *hold. For every* $(x(0), v(0), z_1(0), \ldots, z_N(0)) \in \mathbf{X}$, *let* $X_m(t) = (x_m(t), v_m(t), z_{1,m}(t), \ldots, z_{N,m}(t))$ *be the solution of* (1.5) *and let* $Q(t) = (q(t), f_1(t), \ldots, f_N(t))$ *be the solution of* (1.10) *with the following initial condition*

$$q_i(0) = x_i(0), \quad f_{i,\ell}(0) = z_{i,\ell}(0) + \lambda_{i,\ell} x_i(0), \quad i = 1, \ldots, N, \quad \ell = 1, \ldots, k_i. \tag{2.17}$$

Then, for every T, $\xi > 0$, it holds that

$$\mathbb{P}\left\{ \sup_{t \in [0,T]} |x_m(t) - q(t)| > \xi \right\} \to 0, \quad m \to 0. \tag{2.18}$$

In order to establish Theorem 2.10, we will adapt to the approach in Herzog et al. (2016), Nguyen (2018) tailored to our settings. The argument can be summarized as follows: we first truncate the nonlinearities in (1.5) and (1.10) while making use of conditions (2.2) and (2.6). This results in Lipschitz systems, thus allowing for proving the small-mass limit of the truncated systems. We then exploit the moment estimates on (1.10), cf. Lemma 4.3, to remove the Lipschitz assumption. The explicit argument will be carried out in a series of results in Sect. 4.

Finally, we remark that the convergence in (2.18) only holds in probability, but not in L^p. Following Higham et al. (2002, Theorem 2.2), the latter is a consequence of an L^p estimate on $x_m(t)$ that is uniform with respect to the mass, i.e.,

$$\limsup_{m \to 0} \sup_{t \in [0,T]} |x_m(t)|^p < \infty. \tag{2.19}$$

Due to the singularities, (2.19) is not available in this work and would require further insight. Therefore, the small-mass convergence in L^p remains an open problem.

3 Geometric Ergodicity

Throughout the rest of the paper, c and C denote generic positive constants that may change from line to line. The main parameters that they depend on will appear between parenthesis, e.g., $c(T, q)$ is a function of T and q. In this section, since we do not take $m \to 0$, we will drop the subscript m in x_m, v_m and elsewhere for notational convenience.

In this section, we establish the unique ergodicity and the exponential convergent rate toward the Gibbs measure π_N defined in (1.7) for (1.5). The argument will rely on the construction of suitable Lyapunov functions while making use of a standard irreducibility condition. For the reader's convenience, we first recall the definitions of these notions below.

Definition 3.1 A function $V \in C^2(\mathbf{X}; [1, \infty))$ is called a *Lyapunov* function for (1.5) if the followings hold:
  (i) $V(X) \to \infty$ whenever $|X| + \sum_{1 \le i < j \le N} |x_i - x_j|^{-1} \to \infty$ in $\mathbf{X}$; and
  (ii) for all $X \in \mathbf{X}$,

$$\mathcal{L}_{m,\gamma}^N V(X) \le -c \mathcal{L}_{m,\gamma}^N V(X) + D, \tag{3.1}$$

for some constants $c > 0$ and $D \ge 0$ independent of X.

Definition 3.2 Let V be a Lyapunov function as in Definition 3.1. Denote

$$\mathbf{X}_R = \left\{ X \in \mathbf{X} : V(X) \le R \right\}.$$

The system (1.5) is said to satisfy a *minorization* condition if for all R sufficiently large, there exist positive constants t_R, c_R, a probability measure ν_R such that $\nu_R(\mathbf{X}_R) = 1$ and for every $X \in \mathbf{X}_R$ and any Borel set $A \subset \mathbf{X}$,

$$P_{t_R}^m(X, A) \geq c_R \nu_R(A). \tag{3.2}$$

In Sect. 3.1, we provide a heuristic argument on the construction of a Lyapunov function for the instance of single-particle GLE. In particular, we show that our construction works well for this simpler setting, both in the case $\gamma > 0$ and $\gamma = 0$ through Lemmas 3.3 and 3.4, respectively. In Sect. 3.2, we adapt the construction in the single-particle to the general N-particle system (1.5), and establish suitable Lyapunov function so as to achieve the desired dissipative estimates. Together with an auxiliary result on minorization, cf. Lemma 3.7, we will conclude the proof of Theorem 2.8 in Sect. 3.3.

3.1 Single-Particle GLE

Following the framework developed in Herzog and Mattingly (2019), Lu and Mattingly (2019), we will build up intuition for the construction of Lyapunov functions for the full system (1.5) by investigating a simpler equation in the absence of interacting forces. More specifically, we introduce the following single-particle GLE ($N = 1$)

$$\mathrm{d}\,x(t) = v(t)\mathrm{d}t,$$

$$m\,\mathrm{d}\,v(t) = -\gamma v(t)\mathrm{d}t - \nabla U(x(t))\mathrm{d}t - \nabla G(x(t))\mathrm{d}t + \sum_{i=1}^{k} \lambda_i z_i(t)\mathrm{d}t + \sqrt{2\gamma}\,\mathrm{d}W_0(t),$$

$$\mathrm{d}\,z_i(t) = -\alpha_i z_i(t)\mathrm{d}t - \lambda_i v(t)\mathrm{d}t + \sqrt{2\alpha_i}\,\mathrm{d}W_i(t), \quad i = 1, \ldots, k. \tag{3.3}$$

Let $\mathcal{L}_{m,\gamma}$ be the generator associated with (3.3). That is for $\gamma \geq 0$,

$$\mathcal{L}_{m,\gamma}\varphi = \langle \partial_x \varphi, v \rangle + \frac{1}{m}\left\langle \partial_v \varphi, -\gamma v - \nabla U(x) - \nabla G(x) + \sum_{i=1}^{k} \lambda_i z_i \right\rangle$$

$$+ \frac{\gamma}{m^2}\triangle_v \varphi + \sum_{i=1}^{k}\langle \partial_{z_i}\varphi, -\alpha_i z_i - \lambda_i v \rangle + \sum_{i=1}^{k}\alpha_i \triangle_{z_i}\varphi, \tag{3.4}$$

where $\varphi = \varphi(x, v, z_1, \ldots, z_k) \in C^2(\mathbb{R}^{(k+2)d})$. Denote H

$$H\big(x, v, z_1, \ldots, z_k\big) = U(x) + G(x) + \frac{1}{2}m|v|^2 + \frac{1}{2}\sum_{i=1}^{k}|z_i|^2. \tag{3.5}$$

Note that $\mathcal{L}_{m,\gamma}\varphi$ can be written as

$$\mathcal{L}_{m,\gamma}\varphi = J\nabla H \cdot \nabla\varphi - \sigma\nabla H \cdot \nabla\varphi + \mathrm{div}(\sigma\nabla\varphi), \tag{3.6}$$

where

$$\nabla H = \begin{pmatrix} \partial_x H \\ \partial_v H \\ \partial_{z_1} H \\ \vdots \\ \partial_{z_k} H \end{pmatrix} = \begin{pmatrix} \nabla U(x) + \nabla G(x) \\ mv \\ z_1 \\ \vdots \\ z_k \end{pmatrix}, \quad \nabla \varphi = \begin{pmatrix} \partial_x \varphi \\ \partial_v \varphi \\ \partial_{z_1} \varphi \\ \vdots \\ \partial_{z_k} \varphi \end{pmatrix},$$

$$J := \frac{1}{m} \begin{pmatrix} 0 & I & 0 & \cdots & 0 \\ -I & 0 & \lambda_1 & \cdots & \lambda_k \\ 0 & -\lambda_1 & 0 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & -\lambda_k & 0 & \cdots & 0 \end{pmatrix}, \quad \text{and} \quad \sigma = \sigma_\gamma := \begin{pmatrix} 0 & 0 & 0 & \cdots & 0 \\ 0 & \frac{\gamma}{m^2} & 0 & \cdots & 0 \\ 0 & 0 & \alpha_1 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & \alpha_k \end{pmatrix}.$$

We notice that J is anti-symmetric while σ is positive semi-definite. This formulation will be more convenient in the subsequent computations. In particular, if φ is linear in v and z then the last term in (3.6) vanishes.

Now, there are two cases to be considered depending on the value of the viscous constant γ, of which, the first case is when $\gamma > 0$.

3.1.1 Positive Viscous Constant $\gamma > 0$

In this case, we observe that system (3.3) is almost the same as the following single-particle Langevin equation without the auxiliary memory variables

$$\mathrm{d}\,x(t) = v(t)\mathrm{d}t,$$
$$m\,\mathrm{d}\,v(t) = -\gamma v(t)\mathrm{d}t - \nabla U(x(t))\mathrm{d}t - \nabla G(x(t))\mathrm{d}t + \sqrt{2\gamma}\,\mathrm{d}W_0(t). \tag{3.7}$$

Before discussing the Lyapunov construction for (3.3), it is illuminating to recapitulate the heuristic argument for (3.7) from Herzog and Mattingly (2019), Lu and Mattingly (2019, Mattingly et al. 2002). As a first ansatz, one can facilitate the Hamiltonian of (3.7) given by

$$\tilde{H}(x, v) = \frac{1}{2}m|v|^2 + U(x) + G(x).$$

Denoting by $\tilde{\mathcal{L}}$ the generator associated with (3.7), i.e.,

$$\tilde{\mathcal{L}}\varphi = \langle \partial_x \varphi, v \rangle + \frac{1}{m}\langle \partial_v \varphi, -\gamma v - \nabla U(x) - \nabla G(x) \rangle + \frac{\gamma}{m^2}\triangle_v \varphi, \tag{3.8}$$

a routine computation shows that

$$\tilde{\mathcal{L}}\tilde{H} \leq -c|v|^2 + D.$$

While this is sufficient for the well-posedness of (3.7), it does not produce the dissipative effect on x when $|x| \to \infty$ and $|x| \to 0$, so as to establish geometric ergodicity. In the absence of the singular potential G, one can exploit the trick in Mattingly et al. (2002) by considering a perturbation of the form

$$\tilde{H}(x, v) + \varepsilon \langle x, v \rangle, \tag{3.9}$$

where ε is sufficiently small. Applying $\tilde{\mathcal{L}}$ to the above function while making use of condition (2.3), it is not difficult to see that

$$\tilde{\mathcal{L}}(\tilde{H}(x, v) + \varepsilon \langle x, v \rangle) \leq -c(|v|^2 + U(x)) + C.$$

On the other hand, in the work of Herzog and Mattingly (2019), the following function was introduced

$$\exp \left\{ b(\tilde{H}(x, v) + \psi(x, v)) \right\},$$

where the perturbation $\psi(x, v)$ satisfies

$$\tilde{\mathcal{L}} \psi(x, v) \leq -C,$$

for a sufficiently large constant C. In order to construct ψ, it is crucial to determine the leading order terms in $\tilde{\mathcal{L}}$ when $U(x) + G(x)$ is large while v is being fixed. Following the idea previously presented in Herzog and Mattingly (2019, Section 3.2), in this situation, $\tilde{\mathcal{L}}$ is approximated by

$$\tilde{\mathcal{L}} \approx \mathcal{A}_1 = -\nabla(U(x) + G(x)) \cdot \nabla_v.$$

In turn, this suggests ψ satisfies

$$\mathcal{A}_1 \psi \leq -C.$$

A candidate for the above inequality is given by Herzog and Mattingly (2019)

$$\psi(x, v) \propto \frac{\langle v, \nabla(U(x) + G(x)) \rangle}{|\nabla(U(x) + G(x))|^2}.$$

We refer the reader to Herzog and Mattingly (2019) for the derivation of ψ in more detail. While this choice of ψ works well for Lennard–Jones and Riesz potentials, it is not applicable to the class of log functions, of which, the Coulomb potential in dimension $d = 2$ is a well-known example.

With regard to the specific instance of Coulomb potentials (2.8), in the work of Lu and Mattingly (2019), the authors tackle the Lyapunov issue by employing the scaling transformation

$$x = \kappa^{-1} \hat{x}, \quad v = \hat{v},$$

for $\kappa > 0$. In dimension $d = 2$ (with $G(x) = -\log|x|$), $\tilde{\mathcal{L}}$ as in (3.8) can be recast as

$$\tilde{\mathcal{L}} = v \cdot \nabla_x + \frac{1}{m}\left(-\gamma v - \nabla U(x) + \frac{x}{|x|^2}\right) \cdot \nabla_v + \frac{\gamma}{m^2}\triangle_v$$

$$= \kappa\,\hat{v} \cdot \nabla_{\hat{x}} + \frac{1}{m}\left(-\gamma\hat{v} - \nabla U(\kappa^{-1}\hat{x}) + \kappa\frac{\hat{x}}{|\hat{x}|^2}\right) \cdot \nabla_{\hat{v}} + \frac{\gamma}{m^2}\triangle_{\hat{v}}.$$

Recalling condition (2.1), since $|\nabla U(\kappa^{-1}\hat{x})| \approx \kappa^{-\lambda}|\hat{x}|^\lambda$, which is negligible as κ is large, we observe that

$$\tilde{\mathcal{L}} \approx \kappa\,\hat{v} \cdot \nabla_{\hat{x}} + \frac{\kappa}{m}\frac{\hat{x}}{|\hat{x}|^2} \cdot \nabla_{\hat{v}}, \quad \kappa \to \infty.$$

Likewise, in dimension $d \geq 3$ (with $G(x) = |x|^{2-d}$),

$$\tilde{\mathcal{L}} = \kappa\,\hat{v} \cdot \nabla_{\hat{x}} + \frac{1}{m}\left(-\gamma\hat{v} - \nabla U(\kappa^{-1}\hat{x}) + 2\kappa^{d-1}\frac{\hat{x}}{|\hat{x}|^d}\right) \cdot \nabla_{\hat{v}} + \frac{\gamma}{m^2}\triangle_{\hat{v}}$$

$$\approx \frac{2}{m}\kappa^{d-1}\frac{\hat{x}}{|\hat{x}|^2} \cdot \nabla_{\hat{v}}, \quad k \to \infty.$$

So, taking κ to infinity (i.e., taking $|x| \to 0$) indicates

$$\tilde{\mathcal{L}} \approx \mathcal{A}_2 = \begin{cases} v \cdot \nabla_x + \frac{x}{|x|^2} \cdot \nabla_v, & d = 2, \\ \frac{x}{|x|^d} \cdot \nabla_v, & d \geq 3. \end{cases}$$

A typical choice of ψ satisfying $\mathcal{A}_2\psi \leq -C$ is given by

$$\psi(x, v) \propto -\frac{\langle x, v\rangle}{|x|}.$$

Together with (3.9), we deduce that a Lyapunov function V for (3.7) with Coulomb potentials has the following form Lu and Mattingly (2019)

$$V \propto \tilde{H}(x, v) + \varepsilon_1\langle x, v\rangle - \varepsilon_2\frac{\langle x, v\rangle}{|x|},$$

for some positive constants ε_1 and ε_2 sufficiently small.

Turning back to (3.3) in the case $\gamma > 0$, motivated by the above discussion, for $\varepsilon > 0$, we introduce the function V_1 given by

$$V_1(x, v, z_1, \ldots, z_k) := H(x, v, z_1, \ldots, z_k) + m\varepsilon\langle x, v\rangle - m\varepsilon\frac{\langle x, v\rangle}{|x|}, \qquad (3.10)$$

where H is defined in (3.5). Since (3.3) only differs from (3.7) by the appearance of the linear memory variables z_i's, it turns out that V_1 is indeed a Lyapunov function for (3.3). This is summarized in the following lemma.

Lemma 3.3 *Under Assumptions* 2.1 *and* 2.3*, let V_1 be defined as in* (3.10)*. For each $\gamma > 0$ and $m > 0$, there exists a positive constant ε sufficiently small such that V_1 is a Lyapunov function for* (3.3)*.*

Proof Letting $\mathcal{L}_{m,\gamma}$ and H, respectively, be defined in (3.4) and (3.5), Itô's formula yields

$$\mathcal{L}_{m,\gamma} H = -\sigma \nabla H \cdot \nabla H + \operatorname{div}(\sigma \nabla H)$$
$$= -\gamma |v|^2 - \sum_{i=1}^{k} \alpha_i |z_i|^2 + \frac{1}{m} \cdot \frac{d}{2}\gamma + \frac{d}{2}\sum_{i=1}^{k} \alpha_i. \qquad (3.11)$$

Similarly, we compute

$$\mathcal{L}_{m,\gamma}\big(m\langle x, v\rangle\big) = m|v|^2 - \gamma\langle x, v\rangle - \langle \nabla U(x) + \nabla G(x), x\rangle + \sum_{i=1}^{k} \lambda_i \langle z_i, x\rangle. \quad (3.12)$$

To estimate the right-hand side above, we recall from condition (2.3) that

$$-\langle \nabla U(x), x\rangle \le -a_2|x|^{\lambda+1} + a_3.$$

With regard to ∇G, condition (2.5) implies that

$$|\langle \nabla G(x), x\rangle| \le \frac{a_1}{|x|^{\beta_1 - 1}} + a_1|x|.$$

Concerning the cross terms $\langle x, v\rangle$ and $\langle z_i, x\rangle$, for $\varepsilon \in (0, 1)$, we employ Cauchy–Schwarz inequality to deduce

$$-\gamma\varepsilon\langle x, v\rangle + \varepsilon\sum_{i=1}^{k}\lambda_i\langle z_i, x\rangle \le \gamma\varepsilon^{1/2}|v|^2 + \varepsilon^{1/2}\sum_{i=1}^{k}\lambda_i^2|z_i|^2 + (k+\gamma)\varepsilon^{3/2}|x|^2.$$

Together with (3.12), we find

$$\mathcal{L}_{m,\gamma}\big(\varepsilon m\langle x, v\rangle\big) \le (\varepsilon m + \gamma\varepsilon^{1/2})|v|^2 + \varepsilon^{1/2}\sum_{i=1}^{k}\lambda_i^2|z_i|^2 + (k+\gamma)\varepsilon^{3/2}|x|^2$$
$$- a_2\varepsilon|x|^{\lambda+1} + a_3\varepsilon + \varepsilon\frac{a_1}{|x|^{\beta_1-1}} + \varepsilon a_1|x|, \qquad (3.13)$$

whence,

$$\mathcal{L}_{m,\gamma}\big(\varepsilon m\langle x, v\rangle\big) \le C\left(\varepsilon^{1/2}|v|^2 + \varepsilon^{1/2}\sum_{i=1}^{k}|z_i|^2 + \varepsilon^{3/2}|x|^2 + \varepsilon\frac{1}{|x|^{\beta_1-1}} + 1\right)$$
$$- a_2\varepsilon|x|^{\lambda+1}, \qquad (3.14)$$

for some positive constant C independent of ε.

Turning to $-\langle x, v \rangle / |x|$, it holds that

$$
\begin{aligned}
\mathcal{L}_{m,\gamma} & \left(-m \frac{\langle x, v \rangle}{|x|} \right) \\
= & -m \frac{|v|^2}{|x|} + m \frac{|\langle x, v \rangle|^2}{|x|^3} + \gamma \frac{\langle x, v \rangle}{|x|} + \frac{\langle \nabla U(x), x \rangle}{|x|} + \frac{\langle \nabla G(x), x \rangle}{|x|} \\
& - \sum_{i=1}^{k} \lambda_i \frac{\langle z_i, x \rangle}{|x|}.
\end{aligned}
\tag{3.15}
$$

It is clear that

$$
-m \frac{|v|^2}{|x|} + m \frac{|\langle x, v \rangle|^2}{|x|^3} \leq 0.
$$

From condition (2.1), we readily have

$$
\frac{\langle \nabla U(x), x \rangle}{|x|} \leq |\nabla U(x)| \leq a_1 (1 + |x|^\lambda).
$$

Also, recalling (2.7),

$$
\begin{aligned}
\frac{\langle \nabla G(x), x \rangle}{|x|} = -\frac{a_4}{|x|^{\beta_1}} + \frac{\langle \nabla G(x) + a_4 \frac{x}{|x|^{\beta_1+1}}, x \rangle}{|x|} & \leq -\frac{a_4}{|x|^{\beta_1}} + \frac{a_5}{|x|^{\beta_2}} + a_6 \\
& \leq -\frac{a_4}{2|x|^{\beta_1}} + C.
\end{aligned}
\tag{3.16}
$$

In the last estimate above, we subsumed $|x|^{-\beta_2}$ into $-|x|^{-\beta_1}$ thanks to the fact that $\beta_2 \in [0, \beta_1)$ by virtue of the condition (2.7). Altogether, we deduce that

$$
\begin{aligned}
\mathcal{L}_{m,\gamma} \left(-m\varepsilon \frac{\langle x, v \rangle}{|x|} \right) & \leq C\varepsilon \left(|v| + \sum_{i=1}^{k} |z_i| + |x|^\lambda + 1 \right) - \varepsilon \frac{a_4}{2|x|^{\beta_1}} \\
& \leq C \left(\varepsilon^2 |v|^2 + \varepsilon^2 \sum_{i=1}^{k} |z_i|^2 + \varepsilon |x|^\lambda + 1 \right) - \varepsilon \frac{a_4}{2|x|^{\beta_1}}.
\end{aligned}
\tag{3.17}
$$

Now, we combine estimates (3.14) and (3.17) together with identities (3.10) and (3.11) to infer

$$
\begin{aligned}
\mathcal{L}_{m,\gamma} V_1 = \mathcal{L}_{m,\gamma} & \left(H + m\varepsilon \langle x, v \rangle - m\varepsilon \frac{\langle x, v \rangle}{|x|} \right) \\
& \leq -c|v|^2 - c \sum_{i=1}^{k} |z_i|^2 - c\varepsilon |x|^{\lambda+1} - \varepsilon \frac{c}{|x|^{\beta_1}} + C
\end{aligned}
$$

$$+ C\left(\varepsilon^{1/2}|v|^2 + \varepsilon^{1/2}\sum_{i=1}^k |z_i|^2 + \varepsilon^{3/2}|x|^2 + \varepsilon|x|^\lambda + \varepsilon\frac{1}{|x|^{\beta_1-1}}\right),$$

for some positive constants c, C independent of ε. By taking ε sufficiently small, we observe that the positive non-constant terms on the above right-hand side are dominated by the negative terms. In particular, since $\lambda \geq 1$, $\varepsilon^{3/2}|x|^2$ can be subsumed into $-\varepsilon|x|^{\lambda+1}$. As a consequence, we arrive at

$$\mathcal{L}_{m,\gamma} V_1 \leq -c|v|^2 - c\sum_{i=1}^k |z_i|^2 - c\varepsilon|x|^{\lambda+1} - \varepsilon\frac{c}{|x|^{\beta_1}} + C. \qquad (3.18)$$

Since U and G are bounded by $|x|^{\lambda+1} + |x|^{-\beta_1}$, (3.18) produces the desired Lyapunov property of V_1 for (3.3). The proof is thus finished. $\qquad\square$

3.1.2 Zero Viscous Constant $\gamma = 0$

We now turn to the instance $\gamma = 0$. In this case, since there is no viscous drag on the right-hand side of the v-equation in (3.3), the function V_1 defined in (3.10) does not produce the dissipation in v for large v. To circumnavigate this issue, we note that the z_i-equation in (3.3) still depends on v. So, we may exploit this fact to transfer the dissipation from, say z_1 to v. More specifically, let us consider adding a small perturbation to V_1 as follows:

$$V_1 + m\varepsilon\langle v, z_1\rangle.$$

Denote by $\mathcal{L}_{m,0}$ the generator associated with (3.3) when $\gamma = 0$. That is, from (3.4), we have

$$\mathcal{L}_{m,0}\varphi = \langle \partial_x\varphi, v\rangle + \frac{1}{m}\left\langle \partial_v\varphi, -\nabla U(x) - \nabla G(x) + \sum_{i=1}^k \lambda_i z_i\right\rangle$$
$$+ \sum_{i=1}^k \langle \partial_{z_i}\varphi, -\alpha_i z_i - \lambda_i v\rangle + \sum_{i=1}^k \alpha_i \triangle_{z_i}\varphi. \qquad (3.19)$$

Applying $\mathcal{L}_{m,0}$ to the cross term $\langle v, z_1\rangle$, although $\mathcal{L}_{m,0}\langle v, z_1\rangle$ provides the required dissipative effect in v, it also induces a cross product $\langle \nabla G(x), z_1\rangle$, which has the order of $|z_1|/|x|^{-\beta_1}$ by virtue of condition (2.5). That is

$$\mathcal{L}_{m,0}\langle v, z_1\rangle \leq -c|v|^2 + \frac{|z_1|}{|x|^{\beta_1}}. \qquad (3.20)$$

How to annihilate the effect caused by this extra term is the main difficulty that we face in the case $\gamma = 0$.

To circumvent the issue, from the proof of Lemma 3.3, particularly the estimate (3.17), we see that

$$\mathcal{L}_{m,0}\Big(- \frac{\langle x, v\rangle}{|x|}\Big) \propto -\frac{1}{|x|^{\beta_1}} + \varphi(x, v, z_1, \ldots, z_k),$$

where $\varphi(x, v, z_1, \ldots, z_k)$ consists of lower-order terms. This suggests that we look for a perturbation of the form

$$-\frac{\langle x, v\rangle}{|x|}\psi(x, v, z_1),$$

where ψ satisfies

$$\psi \geq c|z_1|, \quad \text{and} \quad |\mathcal{L}_{m,0}\psi| = O(|v| + |z|). \tag{3.21}$$

From (3.20) and (3.21), the terms $|z_1|/|x|^{\beta_1}$ and $O(|v| + |z|)$ suggest that the issue is where $|v|$, $|z|$ are large and $|x|$ is small. To derive ψ, it is important to understand the dynamics in $\mathcal{L}_{m,0}$ in this "bad region". So, we introduce the following scaling transformation

$$(x, v, z_1, \ldots, z_k) = (\kappa^{-a}\hat{x}, \kappa\hat{v}, \kappa\hat{z}_1, \kappa\hat{z}_2, \ldots, \kappa\hat{z}_k),$$

for some positive constant $a > 1$. Recalling $\mathcal{L}_{m,0}$ as in (3.19), under this scaling, we find that

$$
\begin{aligned}
\mathcal{L}_{m,0} &= v \cdot \nabla_x + \frac{1}{m}\Big(- \nabla U(x) - \nabla G(x) + \sum_{i=1}^{k}\lambda_i z_i\Big) \cdot \nabla_v \\
&\quad + \sum_{i=1}^{k}\big(- \alpha_i z_i - \lambda_i v\big) \cdot \nabla_{z_i} + \sum_{i=1}^{k}\alpha_i \triangle_{z_i} \\
&= \kappa^{a+1}\hat{v} \cdot \nabla_{\hat{x}} + \frac{1}{m\kappa}\Big(- \nabla U(\kappa^{-a}\hat{x}) - \nabla G(\kappa^{-a}\hat{x}) + \kappa\sum_{i=1}^{k}\lambda_i\hat{z}_i\Big) \cdot \nabla_{\hat{v}} \\
&\quad + \sum_{i=1}^{k}\big(- \alpha_i\hat{z}_i - \lambda_i\hat{v}\big) \cdot \nabla_{\hat{z}_i} + \kappa^{-2}\sum_{i=1}^{k}\alpha_i \triangle_{\hat{z}_i}.
\end{aligned}
$$

Recalling condition (2.1) and condition (2.5), suppose heuristically that

$$-\nabla U(\kappa^{-a}\hat{x}) - \nabla G(\kappa^{-a}\hat{x}) \approx -\kappa^{-a\lambda}\nabla U(\hat{x}) - \kappa^{a\beta_1}\nabla G(\hat{x}),$$

implying,

$$\mathcal{L}_{m,0} \approx \kappa^{a+1}\hat{v} \cdot \nabla_{\hat{x}} - \kappa^{-a\lambda-1}\nabla U(\hat{x}) \cdot \nabla_{\hat{v}} - \kappa^{a\beta_1-1}\nabla G(\hat{x}) \cdot \nabla_{\hat{v}} + \sum_{i=1}^{k}\hat{z}_i \cdot \nabla_{\hat{v}}$$

$$-\sum_{i=1}^{k}(\hat{z}_i + \hat{v}) \cdot \nabla_{\hat{z}_i} + \kappa^{-2}\sum_{i=1}^{k}\Delta_{\hat{z}_i}.$$

Observe that

$$\mathcal{L}_{m,0} \approx \kappa^{a+1}\hat{v} \cdot \nabla_{\hat{x}} - \kappa^{a\beta_1 - 1}\nabla G(\hat{x}) \cdot \nabla_{\hat{v}}, \quad \kappa \to \infty.$$

In other words, when κ is large, the dominant balance of terms in the above transformation is contained in

$$\mathcal{A}_{m,0} = v \cdot \nabla_x - \nabla G(x) \cdot \nabla_v.$$

Together with the requirement (3.21), a typical choice for ψ is given by

$$\psi = \sqrt{a_1 |z_1|^2 + \tfrac{1}{2}m|v|^2 + G(x) + U(x) + a_2},$$

for some positive constants a_1, a_2 to be chosen later. In the above, we note that the appearance of $U(x)$ is to ensure the expression under the square root is positive. In summary, the candidate Lyapunov function for (3.3) looks like

$$V_1 + m\varepsilon\langle v, z_1\rangle - m\varepsilon\frac{\langle x, v\rangle}{|x|}\psi.$$

In Lemma 3.4, we will see that by picking a_1, a_2 carefully, we will achieve the Lyapunov effect for (3.3). We finish this discussion by introducing the following function V_2 defined for $\varepsilon \in (0, 1)$, $R > 1$,

$$
\begin{aligned}
&V_2\big(x, v, z_1, \ldots, z_k\big) \\
&:= H\big(x, v, z_1, \ldots, z_k\big) + \varepsilon R m\langle x, v\rangle + \varepsilon R^2 m\langle v, z_1\rangle - \varepsilon m\frac{\langle x, v\rangle}{|x|}\sqrt{Q_R}, \quad (3.22)
\end{aligned}
$$

where

$$Q_R = R^4|z_1|^2 + m|v|^2 + 2U(x) + 2G(x) + R. \qquad (3.23)$$

Lemma 3.4 *Under Assumptions 2.1 and 2.3, let V_2 be defined as in (3.22). For $\gamma = 0$ and every $m > 0$, there exist positive constants ε small and R large enough such that V_2 is a Lyapunov function for (3.3).*

Proof Firstly, we note that when $\gamma = 0$, by virtue of Assumption 2.1, $\lambda = 1$. From the estimate (3.13), we immediately obtain

$$\mathcal{L}_{m,0}\Big(\varepsilon R m\langle x, v\rangle\Big) \leq CR\Big(\varepsilon|v|^2 + \varepsilon^{1/2}\sum_{i=1}^{k}|z_i|^2 + \varepsilon^{3/2}|x|^2 + \varepsilon\frac{1}{|x|^{\beta_1 - 1}} + 1\Big)$$

$$- a_2 \varepsilon R |x|^2.$$

Also, (3.11) is reduced to

$$\mathcal{L}_{m,0} H = -\sum_{i=1}^{k} \alpha_i |z_i|^2 + \frac{d}{2} \sum_{i=1}^{k} \alpha_i.$$

As a consequence, we obtain

$$\mathcal{L}_{m,0}\big(H + \varepsilon Rm \langle x, v \rangle\big) \leq CR\left(\varepsilon |v|^2 + \varepsilon^{1/2} \sum_{i=1}^{k} |z_i|^2 + \varepsilon^{3/2} |x|^2 + \varepsilon \frac{1}{|x|^{\beta_1 - 1}} + 1 \right)$$
$$- a_2 \varepsilon R |x|^2 - \sum_{i=1}^{k} \alpha_i |z_i|^2 + \frac{d}{2} \sum_{i=1}^{k} \alpha_i.$$

By taking ε sufficiently small, it follows that

$$\mathcal{L}_{m,0}\big(H + m\varepsilon \langle x, v \rangle\big)$$
$$\leq -c\varepsilon R |x|^2 - c \sum_{i=1}^{k} |z_i|^2 + C\varepsilon Rm |v|^2 + C\varepsilon R \frac{1}{|x|^{\beta_1 - 1}} + CR, \qquad (3.24)$$

for some positive constants c, C independent of ε and R.

Next, we consider the cross term $\langle v, z_1 \rangle$ on the right-hand side of (3.22). Itô's formula yields (recalling $\gamma = 0$)

$$\mathcal{L}_{m,0}\big(\varepsilon m \langle v, z_1 \rangle\big)$$
$$= -\varepsilon \langle \nabla U(x) + \nabla G(x), z_1 \rangle + \varepsilon \sum_{i=1}^{k} \lambda_i \langle z_i, z_1 \rangle - \varepsilon \alpha_1 m \langle v, z_1 \rangle - \lambda_1 \varepsilon m |v|^2. \quad (3.25)$$

We invoke condition (2.1) with Cauchy–Schwarz inequality to infer

$$-\varepsilon \langle \nabla U(x), z_1 \rangle \leq \varepsilon a_1 (1 + |x|) |z_1| \leq a_1 \varepsilon^{3/2} |x|^2 + a_1 \varepsilon^{1/2} |z_1|^2 + a_1 \varepsilon |z_1|.$$

Similarly,

$$\varepsilon \sum_{i=1}^{k} \lambda_i \langle z_i, z_1 \rangle - \varepsilon \alpha_1 m \langle v, z_1 \rangle \leq \varepsilon \sum_{i=1}^{k} \lambda_i^2 |z_i|^2 + \varepsilon k |z_1|^2 + \varepsilon^{1/2} \alpha_1^2 m |z_1|^2 + \varepsilon^{3/2} m |v|^2.$$

Concerning the cross term $\langle \nabla G(x), z_1 \rangle$, recall from (2.5) that

$$|\nabla G(x)| \leq \frac{a_1}{|x|^{\beta_1}} + a_1.$$

As a consequence,

$$-\varepsilon\langle\nabla G(x), z_1\rangle \leq \varepsilon a_1 \frac{|z_1|}{|x|^{\beta_1}} + \varepsilon a_1|z_1| \leq \varepsilon a_1 \frac{|z_1|}{|x|^{\beta_1}} + \frac{1}{2}\varepsilon(|z_1|^2 + a_1^2).$$

Altogether, we deduce that the bound

$$\mathcal{L}_{m,0}\big(R^2\varepsilon m\langle v, z_1\rangle\big)$$

$$\leq -\lambda_1\varepsilon R^2 m|v|^2 + C\varepsilon R^2 \frac{|z_1|}{|x|^{\beta_1}} + C\varepsilon^{3/2} R^2\big(|v|^2 + |x|^2\big) + C\varepsilon^{1/2} R^2 \sum_{i=1}^{k} |z_i|^2 + C\varepsilon R^2, \tag{3.26}$$

holds for some positive constant C independent of R and ε.

Turning to the last term on the right-hand side of (3.22), a routine calculation together with (3.15) gives

$$\mathcal{L}_{m,0}\left(-\varepsilon m \frac{\langle x, v\rangle}{|x|}\sqrt{Q_R}\right)$$

$$= \varepsilon\left(-m\frac{|v|^2|x|^2 - |\langle x, v\rangle|^2}{|x|^3} + \frac{\langle\nabla U(x) - \sum_{i=1}^{k}\lambda_i z_i, x\rangle}{|x|} + \frac{\langle\nabla G(x), x\rangle}{|x|}\right)\sqrt{Q_R}$$

$$- \varepsilon m\frac{\langle x, v\rangle}{|x|} \cdot \frac{\langle v, \sum_{i=1}^{k}\lambda_i z_i\rangle}{\sqrt{Q_R}} + \varepsilon R^6 m\frac{\langle x, v\rangle}{|x|} \cdot \frac{\alpha_1|z_1|^2 + \lambda_1\langle v, z_1\rangle}{\sqrt{Q_R}}$$

$$+ \frac{1}{2}\alpha_1\varepsilon R^6 m\frac{\langle x, v\rangle}{|x|} \cdot \frac{1}{\sqrt{Q_R}}\left(d - \frac{R^6|z_1|^2}{Q_R}\right)$$

$$= I_1 - I_2 + I_3 + I_4. \tag{3.27}$$

Concerning I_4, we recall from (3.23) that

$$Q_R \geq R^6|z_1|^2 + m|v|^2, \tag{3.28}$$

whence

$$I_4 \leq \frac{1}{2}\alpha_1\varepsilon R^6 d\sqrt{m}.$$

With regard to I_1, we invoke conditions (2.1) and (2.4) to see that

$$Q_R \leq R^6|z_1|^2 + m|v|^2 + a_1\left(1 + |x|^2 + \frac{1}{|x|^{\beta_1}} + |x|\right) + R^2.$$

As a consequence,

$$\varepsilon\frac{\langle\nabla U(x) - \sum_{i=1}^{k}\lambda_i z_i, x\rangle}{|x|}\sqrt{Q_R} \leq \varepsilon\left[a_1(1 + |x|) + \sum_{i=1}^{k}\lambda_i|z_i|\right]\sqrt{Q_R}$$

$$\leq \frac{1}{2}\varepsilon \left[a_1(1 + |x|) + \sum_{i=1}^{k} \lambda_i |z_i| \right]^2 + \frac{1}{2}\varepsilon Q_R$$

$$\leq C\varepsilon \left(R^6 \sum_{i=1}^{k} |z_i|^2 + |v|^2 + |x|^2 + \frac{1}{|x|^{\beta_1}} + R^2 \right).$$

On the other hand, estimate (3.16) implies the bound

$$\varepsilon \frac{\langle \nabla G(x), x \rangle}{|x|} \sqrt{Q_R} \leq \varepsilon \left(-\frac{a_4}{2|x|^{\beta_1}} + C \right) \sqrt{Q_R}$$

$$= -\varepsilon \frac{a_4}{2} \cdot \frac{1}{|x|^{\beta_1}} \sqrt{Q_R} + C\varepsilon \sqrt{Q_R}$$

$$\leq -c\varepsilon \frac{R^3 |z_1| + R}{|x|^{\beta_1}} + C\varepsilon \left(R^6 |z_1|^2 + |v|^2 + |x|^2 + \frac{1}{|x|^{\beta_1}} + R^2 \right).$$

It follows that

$$I_1 \leq -c\varepsilon \frac{R^3 |z_1| + R}{|x|^{\beta_1}} + C\varepsilon \left(R^6 \sum_{i=1}^{k} |z_i|^2 + |v|^2 + |x|^2 + \frac{1}{|x|^{\beta_1}} + R^2 \right).$$

With regard to I_2 on the right-hand side of (3.27), since $Q_R \geq m|v|^2$, we find

$$-I_2 = -\varepsilon m \frac{\langle x, v \rangle}{|x|} \cdot \frac{\langle v, \sum_{i=1}^{k} \lambda_i z_i \rangle}{\sqrt{Q_R}} \leq \varepsilon \sqrt{m} |v| \sum_{i=1}^{k} \lambda_i |z_i| \leq C\varepsilon \left(|v|^2 + \sum_{i=1}^{k} |z_i|^2 \right).$$

Turning to I_3, we estimate as follows:

$$I_3 = \varepsilon R^6 m \frac{\langle x, v \rangle}{|x|} \cdot \left(\frac{\alpha_1 |z_1|^2}{\sqrt{Q_R}} + \frac{\lambda_1 \langle v, z_1 \rangle}{\sqrt{Q_R}} \right)$$

$$\leq \varepsilon R^3 m |v| \cdot \alpha_1 |z_1| + \varepsilon R^6 \sqrt{m} |v| \cdot \lambda_1 |z_1|$$

$$\leq C\varepsilon (|v|^2 + R^{12} |z_1|^2),$$

where in the first inequality we have used $\sqrt{Q_R} \geq R^3 |z_1|$ (which follows from (3.28)) and $\sqrt{Q_R} \geq \sqrt{m} |v|$. Now, we collect the estimates on I_j, $j = 1, \ldots, 4$, together with expression (3.27) to infer (recalling $\varepsilon < 1 < R$)

$$\mathcal{L}_{m,0} \left(-\varepsilon m \frac{\langle x, v \rangle}{|x|} \sqrt{Q_R} \right)$$

$$\leq -c\varepsilon \frac{R^3 |z_1| + R}{|x|^{\beta_1}} + C\varepsilon \left(R^{12} \sum_{i=1}^{k} |z_i|^2 + |v|^2 + |x|^2 + \frac{1}{|x|^{\beta_1}} + R^6 \right)$$

$$\leq -c\varepsilon \frac{R^3|z_1| + R}{|x|^{\beta_1}} + C\varepsilon \left(R^{12} \sum_{i=1}^{k} |z_i|^2 + |v|^2 + |x|^2 + R^6 \right). \qquad (3.29)$$

In the last implication above, we subsumed $C\varepsilon |x|^{-\beta_1}$ into $-c\varepsilon R|x|^{-\beta_1}$, by taking R large enough.

Turning back to V_2 given by (3.22), we combine (3.24), (3.26) and (3.29) to arrive at the estimate

$$\mathcal{L}_{m,0} V_2 \leq -c\varepsilon R|x|^2 - c \sum_{i=1}^{k} |z_i|^2 + C\varepsilon R|v|^2 + C\varepsilon R \frac{1}{|x|^{\beta_1 - 1}} + CR$$

$$- c\varepsilon R^2 |v|^2 + C\varepsilon R^2 \frac{|z_1|}{|x|^{\beta_1}} + C\varepsilon^{3/2} R^2 \big(|v|^2 + |x|^2 \big) + C\varepsilon^{1/2} R^2 \sum_{i=1}^{k} |z_i|^2$$

$$+ C\varepsilon R^2$$

$$- c\varepsilon \frac{R^3|z_1| + R}{|x|^{\beta_1}} + C\varepsilon \left(R^{12} \sum_{i=1}^{k} |z_i|^2 + |v|^2 + |x|^2 + R^6 \right),$$

i.e.,

$$\mathcal{L}_{m,0} V_2 \leq -(c\varepsilon R^2 - C\varepsilon R - C\varepsilon^{3/2} R^2 - C\varepsilon)|v|^2 - (c - C\varepsilon^{1/2}R^2 - C\varepsilon R^{12}) \sum_{i=1}^{k} |z_i|^2$$

$$- (c\varepsilon R - C\varepsilon^{3/2}R^2 - C\varepsilon)|x|^2 - c\varepsilon R \frac{1}{|x|^{\beta_1}} - (c\varepsilon R^3 - c\varepsilon R^2)\frac{|z_1|}{|x|^{\beta_1}}$$

$$+ C\varepsilon R \frac{1}{|x|^{\beta_1 - 1}} + CR + C\varepsilon R^6.$$

Since c, C are independent of ε and R, we may take R sufficiently large and then shrink ε to zero while making use of the fact that $\frac{1}{|x|^{\beta_1 - 1}}$ can be subsumed into $|x|^2 + |x|^{\beta_1}$. It follows that

$$\mathcal{L}_{m,0} V_2 \leq -c\varepsilon R^2 |v|^2 - c \sum_{i=1}^{k} |z_i|^2 - c\varepsilon R|x|^2 - c\varepsilon R \frac{1}{|x|^{\beta_1}} + C\varepsilon R^6. \qquad (3.30)$$

This produces the Lyapunov property of V_2 for (3.3) in the case $\gamma = 0$, thereby finishing the proof. □

3.2 N-particle GLEs

We now turn our attention to the full system (1.5) and construct Lyapunov functions for (1.5) based on the discussion of the single-particle GLE in Sect. 3.1.

We start with the case $\gamma > 0$ and observe that there is a natural generalization of the function V_1 defined in (3.10) to an arbitrary number of particles $N \geq 2$ in an

arbitrary number of dimensions $d \geq 1$. More specifically, for $\varepsilon > 0$, we introduce the following function V_N^1 given by

$$
\begin{aligned}
V_N^1(\mathrm{x}, \mathrm{v}, z_1, \ldots, z_N) \\
= H_N(\mathrm{x}, \mathrm{v}, z_1, \ldots, z_N) + \varepsilon m \langle \mathrm{x}, \mathrm{v} \rangle - \varepsilon m \sum_{i=1}^{N} \left\langle v_i, \sum_{j \neq i} \frac{x_i - x_j}{|x_i - x_j|} \right\rangle,
\end{aligned}
\tag{3.31}
$$

where H_N is the Hamiltonian as in (1.6). In Lemma 3.5, stated and proven next, we assert that one may pick ε sufficiently small to ensure the Lyapunov property of V_N^1.

Lemma 3.5 *Under Assumptions 2.1 and 2.3, let V_N^1 be the function defined as (3.31). For all $\gamma > 0$ and $m > 0$, there exists a positive constant ε sufficiently small such that V_N^1 is a Lyapunov function for (1.5).*

Proof We first consider the Hamiltonian H_N given by (1.6). Applying $\mathcal{L}_{m,\gamma}^N$ as in (2.14) to H_N gives

$$
\mathcal{L}_{m,\gamma}^N H_N = -\gamma |\mathrm{v}|^2 - \sum_{i=1}^{N} \sum_{\ell=1}^{k_i} \alpha_{i,\ell} |z_{i,\ell}|^2 + \frac{1}{2m} \gamma N d + \frac{1}{2} \sum_{i=1}^{N} \sum_{\ell=1}^{k_i} \alpha_{i,\ell}.
\tag{3.32}
$$

Next, a routine calculation on the cross term $\langle \mathrm{x}, \mathrm{v} \rangle$ gives

$$
\begin{aligned}
\mathcal{L}_{m,\gamma}^N \big(\varepsilon m \langle \mathrm{x}, \mathrm{v} \rangle \big) = \varepsilon m |\mathrm{v}|^2 - \varepsilon \gamma \langle \mathrm{x}, \mathrm{v} \rangle + \varepsilon \sum_{i=1}^{N} \left\langle x_i, \sum_{\ell=1}^{k_i} z_{i,\ell} \right\rangle \\
- \varepsilon \sum_{i=1}^{N} \langle x_i, \nabla U(x_i) \rangle - \varepsilon \sum_{1 \leq i < j \leq N} \langle x_i - x_j, \nabla G(x_i - x_j) \rangle.
\end{aligned}
\tag{3.33}
$$

Recalling (2.3), we readily have

$$
-\sum_{i=1}^{N} \langle x_i, \nabla U(x_i) \rangle \leq -a_2 \sum_{i=1}^{N} |x_i|^{\lambda+1} + N a_3.
$$

Also, from (2.5), it holds that

$$
-\sum_{1 \leq i < j \leq N} \langle x_i - x_j, \nabla G(x_i - x_j) \rangle \leq a_1 \left(N^2 + \sum_{1 \leq i < j \leq N} \frac{1}{|x_i - x_j|^{\beta_1 - 1}} \right).
$$

Concerning the other cross terms on the right-hand side of (3.33) we invoke Cauchy–Schwarz inequality to infer

$$-\varepsilon\gamma\langle x, v\rangle + \varepsilon \sum_{i=1}^{N}\Big\langle x_i, \sum_{\ell=1}^{k_i} z_{i,\ell}\Big\rangle \leq C\Big(\varepsilon^{1/2}|v|^2 + \varepsilon^{1/2}\sum_{i=1}^{N}|z_i|^2 + \varepsilon^{3/2}|x|^2\Big)$$

for some positive constant C independent of ε. We now collect the above estimates together with expression (3.33) to obtain

$$\mathcal{L}_{m,\gamma}^{N}\big(\varepsilon m\langle x, v\rangle\big) \leq C\Big(\varepsilon^{1/2}|v|^2 + \varepsilon^{1/2}\sum_{i=1}^{N}|z_i|^2 + \varepsilon^{3/2}|x|^2 + \varepsilon \sum_{1\leq i<j\leq N}\frac{1}{|x_i - x_j|^{\beta_1-1}}\Big)$$

$$- a_2\varepsilon \sum_{i=1}^{N}|x_i|^{\lambda+1} + C. \tag{3.34}$$

Next, we turn to the last term on the right-hand side of (3.31). A routine calculation produces

$$\mathcal{L}_{m,\gamma}^{N}\Big(-m\sum_{i=1}^{N}\Big\langle v_i, \sum_{j\neq i}\frac{x_i - x_j}{|x_i - x_j|}\Big\rangle\Big)$$

$$= -m\sum_{1\leq i<j\leq N}\frac{|v_i - v_j|^2}{|x_i - x_j|} + m\sum_{1\leq i<j\leq N}\frac{|\langle v_i - v_j, x_i - x_j\rangle|^2}{|x_i - x_j|^3}$$

$$+ \gamma\sum_{1\leq i<j\leq N}\frac{\langle v_i - v_j, x_i - x_j\rangle}{|x_i - x_j|} + \sum_{1\leq i<j\leq N}\frac{\langle\nabla U(x_i) - \nabla U(x_j), x_i - x_j\rangle}{|x_i - x_j|}$$

$$+ \sum_{i=1}^{N}\Big\langle\sum_{j\neq i}\frac{x_i - x_j}{|x_i - x_j|}, \sum_{\ell\neq i}\nabla G(x_i - x_\ell)\Big\rangle - \sum_{i=1}^{N}\Big\langle\sum_{j\neq i}\frac{x_i - x_j}{|x_i - x_j|}, \sum_{\ell=1}^{k_i}\lambda_{i,\ell} z_{i,\ell}\Big\rangle. \tag{3.35}$$

We proceed to estimate the above right-hand side while making use of Cauchy–Schwarz inequality. It is clear that

$$-m\sum_{1\leq i<j\leq N}\frac{|v_i - v_j|^2}{|x_i - x_j|} + m\sum_{1\leq i<j\leq N}\frac{|\langle v_i - v_j, x_i - x_j\rangle|^2}{|x_i - x_j|^3} \leq 0,$$

which is negligible. Also,

$$\gamma\sum_{1\leq i<j\leq N}\frac{\langle v_i - v_j, x_i - x_j\rangle}{|x_i - x_j|} - \sum_{i=1}^{N}\Big\langle\sum_{j\neq i}\frac{x_i - x_j}{|x_i - x_j|}, \sum_{\ell=1}^{k_i}\lambda_{i,\ell} z_{i,\ell}\Big\rangle$$

$$\leq \gamma(N-1)\sum_{i=1}^{N}|v_i| + (N-1)\sum_{i=1}^{N}\sum_{\ell=1}^{k_i}\lambda_{i,\ell}|z_{i,\ell}|.$$

Concerning the cross terms involving ∇U, we invoke condition (2.1) and obtain

$$\sum_{1\leq i<j\leq N}\frac{\langle\nabla U(x_i)-\nabla U(x_j), x_i-x_j\rangle}{|x_i-x_j|} \leq (N-1)a_1\Big(N+\sum_{i=1}^{N}|x_i|^{\lambda}\Big).$$

With regard to the cross terms involving ∇G, we recast them as follows:

$$\sum_{i=1}^{N}\Big\langle\sum_{j\neq i}\frac{x_i-x_j}{|x_i-x_j|}, \sum_{\ell\neq i}\nabla G(x_i-x_\ell)\Big\rangle = -a_4\sum_{i=1}^{N}\Big\langle\sum_{j\neq i}\frac{x_i-x_j}{|x_i-x_j|}, \sum_{\ell\neq i}\frac{x_i-x_\ell}{|x_i-x_\ell|^{\beta_1+1}}\Big\rangle$$

$$+\sum_{i=1}^{N}\Big\langle\sum_{j\neq i}\frac{x_i-x_j}{|x_i-x_j|}, \sum_{\ell\neq i}\nabla G(x_i-x_\ell)$$

$$+a_4\frac{x_i-x_\ell}{|x_i-x_\ell|^{\beta_1+1}}\Big\rangle.$$

In view of Lemma A.1, we readily have

$$-a_4\sum_{i=1}^{N}\Big\langle\sum_{j\neq i}\frac{x_i-x_j}{|x_i-x_j|}, \sum_{\ell\neq i}\frac{x_i-x_\ell}{|x_i-x_\ell|^{\beta_1+1}}\Big\rangle \leq -2a_4\sum_{1\leq i<j\leq N}\frac{1}{|x_i-x_j|^{\beta_1}}.$$

On the other hand, condition (2.7) implies the bound

$$\sum_{i=1}^{N}\Big\langle\sum_{j\neq i}\frac{x_i-x_j}{|x_i-x_j|}, \sum_{\ell\neq i}\nabla G(x_i-x_\ell)+a_4\frac{x_i-x_\ell}{|x_i-x_\ell|^{\beta_1+1}}\Big\rangle$$

$$\leq (N-1)\sum_{i=1}^{N}\sum_{\ell\neq i}\Big|\nabla G(x_i-x_\ell)+a_4\frac{x_i-x_\ell}{|x_i-x_\ell|^{\beta_1+1}}\Big|$$

$$\leq (N-1)\Big[2\sum_{1\leq i<\ell\leq N}\frac{a_5}{|x_i-x_\ell|^{\beta_2}}+N(N-1)a_6\Big].$$

In the above, a_4, a_5 and a_6 are the constants as in condition (2.7). Since $\beta_2 \in [0, \beta_1)$, we observe that $|x_i-x_\ell|^{-\beta_2}$ can be subsumed into $-|x_i-x_\ell|^{-\beta_1}$. It follows that

$$\sum_{i=1}^{N}\Big\langle\sum_{j\neq i}\frac{x_i-x_j}{|x_i-x_j|}, \sum_{\ell\neq i}\nabla G(x_i-x_\ell)\Big\rangle \leq -a_4\sum_{1\leq i<j\leq N}\frac{1}{|x_i-x_j|^{\beta_1}}+C. \quad (3.36)$$

From the identity (3.35), we infer the estimate

$$\mathcal{L}_{m,\gamma}^N \left(-\varepsilon m \sum_{i=1}^N \left\langle v_i, \sum_{j\neq i} \frac{x_i - x_j}{|x_i - x_j|} \right\rangle \right)$$

$$\leq -a_4 \varepsilon \sum_{1 \leq i < j \leq N} \frac{1}{|x_i - x_j|^{\beta_1}} + C\varepsilon \left(1 + \sum_{i=1}^N |v_i| + \sum_{i=1}^N |x_i|^\lambda + \sum_{i=1}^N \sum_{\ell=1}^{k_i} |z_{i,\ell}| \right).$$

$$(3.37)$$

Now, we collect (3.32), (3.34), (3.35) together with the expression (3.31) of V_N^1 and deduce

$$\mathcal{L}_{m,\gamma}^N V_N^1 \leq -\gamma |v|^2 - \sum_{i=1}^N \sum_{\ell=1}^{k_i} \alpha_{i,\ell} |z_{i,\ell}|^2 - a_2 \varepsilon \sum_{i=1}^N |x_i|^{\lambda+1}$$

$$- a_4 \varepsilon \sum_{1 \leq i < j \leq N} \frac{1}{|x_i - x_j|^{\beta_1}} + C$$

$$+ C \left(\varepsilon^{1/2} |v|^2 + \varepsilon^{1/2} \sum_{i=1}^N |z_i|^2 + \varepsilon^{3/2} |x|^2 + \varepsilon \sum_{1 \leq i < j \leq N} \frac{1}{|x_i - x_j|^{\beta_1 - 1}} \right)$$

$$+ C\varepsilon \left(\sum_{i=1}^N |v_i| + \sum_{i=1}^N |x_i|^\lambda + \sum_{i=1}^N \sum_{\ell=1}^{k_i} |z_{i,\ell}| \right).$$

$$(3.38)$$

In the above, we emphasize that C is a positive constant independent of ε. Finally, by taking ε sufficiently small, we may infer

$$\mathcal{L}_{m,\gamma}^N V_N^1 \leq -\frac{1}{2} \left(\gamma |v|^2 + \sum_{i=1}^N \sum_{\ell=1}^{k_i} \alpha_{i,\ell} |z_{i,\ell}|^2 + a_2 \varepsilon \sum_{i=1}^N |x_i|^{\lambda+1} \right.$$

$$\left. + a_4 \varepsilon \sum_{1 \leq i < j \leq N} \frac{1}{|x_i - x_j|^{\beta_1}} \right) + C.$$

This produces the desired Lyapunov property of V_N^1 for system (1.5) in the case $\gamma > 0$, as claimed. $\qquad\square$

Turning to the case $\gamma = 0$, analogous to the function V_2 defined in (3.22) for the single-particle system (3.3), for $\varepsilon \in (0, 1)$ and $R > 1$, we introduce the function V_N^2 given by

$$V_N^2(x, v, z_1, \ldots, z_N) = H_N(x, v, z_1, \ldots, z_N) + \varepsilon R m \langle x, v \rangle + \varepsilon R^2 m \sum_{i=1}^N \langle v_i, z_{i,1} \rangle$$

$$-\varepsilon\left(\sum_{i=1}^{N}m\left\langle v_i, \sum_{j\neq i}\frac{x_i - x_j}{|x_i - x_j|}\right\rangle\right)\sqrt{Q_R^N}, \tag{3.39}$$

where

$$Q_R^N(\mathrm{x}, \mathrm{v}, z_1, \ldots, z_N) = R^6 \sum_{i=1}^{N}|z_{i,1}|^2 + m|\mathrm{v}|^2 + 2\sum_{i=1}^{N}U(x_i)$$

$$+ 2\sum_{1\leq i<j\leq N}G(x_i - x_j) + R^2. \tag{3.40}$$

In Lemma 3.6, we prove that V_N^2 is indeed a Lyapunov function for (1.5) in the case $\gamma = 0$.

Lemma 3.6 *Under Assumptions* 2.1 *and* 2.3, *let* V_N^2 *be the function defined in* (3.39). *In the case* $\gamma = 0$, *for every* $m > 0$, *there exist positive constants* ε *small and* R *large enough such that* V_N^2 *is a Lyapunov function for* (1.5).

Proof We first consider the Hamiltonian H_N as in (1.6). Since $\gamma = 0$, the identity (3.32) is reduced to

$$\mathcal{L}_{m,0}^N H_N = -\sum_{i=1}^{N}\sum_{\ell=1}^{k_i}\alpha_{i,\ell}|z_{i,\ell}|^2 + \frac{1}{2}\sum_{i=1}^{N}\sum_{\ell=1}^{k_i}\alpha_{i,\ell}. \tag{3.41}$$

With regard to the cross term $\langle \mathrm{x}, \mathrm{v}\rangle$, we note that since $\gamma = 0$, by virtue of Assumption 2.1, $\lambda = 1$. We employ an argument similarly to that of the estimate (3.34) and obtain the bound

$$\mathcal{L}_{m,0}^N\left(\varepsilon Rm\langle \mathrm{x}, \mathrm{v}\rangle\right) \leq -a_2\varepsilon R|\mathrm{x}|^2 + CR\left(\varepsilon|\mathrm{v}|^2 + \varepsilon^{1/2}\sum_{i=1}^{N}|z_i|^2 + \varepsilon^{3/2}|\mathrm{x}|^2 + 1\right)$$

$$+ C\varepsilon R\sum_{1\leq i<j\leq N}\frac{1}{|x_i - x_j|^{\beta_1-1}}. \tag{3.42}$$

In the above, C is a positive constant independent of ε and R.

Concerning the cross terms $\langle v_i, z_{i,1}\rangle$, $i = 1, \ldots, N$, on the right-hand side of (3.39), applying Itô's formula gives

$$\mathcal{L}_{m,0}^N\left(m\langle v_i, z_{i,1}\rangle\right) = \left\langle -\nabla U(x_i) - \sum_{j\neq i}\nabla G(x_i - x_j), z_{i,1}\right\rangle + \sum_{\ell=1}^{k_i}\lambda_{i,\ell}\langle z_{i,\ell}, z_{i,1}\rangle$$

$$- \alpha_{i,1}m\langle v_i, z_{i,1}\rangle - \lambda_{i,1}m|v_i|^2.$$

From the condition (2.1) ($\lambda = 1$), we have

$$-\varepsilon\langle \nabla U(x_i), z_{i,1}\rangle \leq a_1\varepsilon(|x| + 1)|z_{i,1}| \leq C(\varepsilon^{3/2}|x_i|^2 + \varepsilon^{1/2}|z_{i,1}|^2 + \varepsilon).$$

In the last estimate above, we employed Cauchy–Schwarz inequality. Likewise,

$$\varepsilon \sum_{\ell=1}^{k_i} \lambda_{i,\ell} \langle z_{i,\ell}, z_{i,1} \rangle - \varepsilon \alpha_{i,1} m \langle v_i, z_{i,1} \rangle \leq C \varepsilon^{1/2} |z_i|^2 + C \varepsilon^{3/2} |v_i|^2.$$

Also, condition (2.5) implies

$$-\sum_{i=1}^{N} \left\langle \sum_{j \neq i} \nabla G(x_i - x_j), z_{i,1} \right\rangle \leq a_1 \sum_{i=1}^{N} |z_{i,1}| \left(2 \sum_{1 \leq i < j \leq N} \frac{1}{|x_i - x_j|^{\beta_1}} + N^2 \right)$$

$$\leq C \sum_{i=1}^{N} |z_{i,1}| \sum_{1 \leq i < j \leq N} \frac{1}{|x_i - x_j|^{\beta_1}}$$

$$+ C \sum_{i=1}^{N} |z_i|^2 + C.$$

It follows that for $\varepsilon \in (0, 1)$ and $R > 1$,

$$\mathcal{L}_{m,0}^N \left(\varepsilon R^2 m \sum_{i=1}^{N} \langle v_i, z_{i,1} \rangle \right) \leq -\varepsilon R^2 m \sum_{i=1}^{N} \lambda_{i,1} |v_i|^2$$

$$+ C R^2 \left(\varepsilon^{3/2} (|\mathbf{x}|^2 + |\mathbf{v}|^2) + \varepsilon^{1/2} \sum_{i=1}^{N} |z_i|^2 + \varepsilon \right)$$

$$+ C \varepsilon R^2 \left(\sum_{i=1}^{N} |z_i| \sum_{1 \leq i < j \leq N} \frac{1}{|x_i - x_j|^{\beta_1}} + \sum_{i=1}^{N} |z_i|^2 + 1 \right). \tag{3.43}$$

In the above, we emphasize again that the positive constant C does not depend on ε and R.

Next, we consider the last term on the right-hand side of (3.39). Observe that

$$\sum_{i=1}^{N} \left\langle v_i, \sum_{j \neq i} \frac{x_i - x_j}{|x_i - x_j|} \right\rangle = \sum_{1 \leq i < j \leq N} \frac{\langle v_i - v_j, x_i - x_j \rangle}{|x_i - x_j|}.$$

So,

$$\mathcal{L}_{m,0}^N \left(-\varepsilon \sum_{i=1}^{N} m \left\langle v_i, \sum_{j \neq i} \frac{x_i - x_j}{|x_i - x_j|} \right\rangle \right) \sqrt{Q_R^N}$$

$$= -\varepsilon \sqrt{Q_R^N} \, \mathcal{L}_{m,0}^N \left(\sum_{i=1}^{N} m \left\langle v_i, \sum_{j \neq i} \frac{x_i - x_j}{|x_i - x_j|} \right\rangle \right)$$

$$- \varepsilon m \sum_{1 \leq i < j \leq N} \frac{\langle v_i - v_j, x_i - x_j \rangle}{|x_i - x_j|} \mathcal{L}_{m,0}^N \sqrt{Q_R^N}$$

$$= I_1 + I_2. \tag{3.44}$$

Concerning I_1, from the estimate (3.37) (with $\lambda = 1$), we have

$$I_1 = \sqrt{Q_R^N} \, \mathcal{L}_{m,0}^N \left(- \varepsilon \sum_{i=1}^N m \left\langle v_i, \sum_{j \neq i} \frac{x_i - x_j}{|x_i - x_j|} \right\rangle \right)$$

$$\leq -a_4 \varepsilon \sqrt{Q_R^N} \sum_{1 \leq i < j \leq N} \frac{1}{|x_i - x_j|^{\beta_1}} + C \varepsilon \left(1 + |v| + |x| + \sum_{i=1}^N \sum_{\ell=1}^{k_i} |z_{i,\ell}| \right) \sqrt{Q_R^N}.$$

Recalling Q_R^N given by (3.40)

$$Q_R^N = R^6 \sum_{i=1}^N |z_{i,1}|^2 + m|v|^2 + 2 \sum_{i=1}^N U(x_i) + 2 \sum_{1 \leq i < j \leq N} G(x_i - x_j) + R^2,$$

it is clear that

$$\sqrt{Q_R^N} \geq \left(R^6 \sum_{i=1}^N |z_{i,1}|^2 + R^2 \right)^{1/2} \geq c R^3 \sum_{i=1}^N |z_{i,1}| + c R,$$

whence

$$- a_4 \varepsilon \sqrt{Q_R^N} \sum_{1 \leq i < j \leq N} \frac{1}{|x_i - x_j|^{\beta_1}} \leq$$

$$- c \varepsilon R^3 \sum_{i=1}^N |z_{i,1}| \sum_{1 \leq i < j \leq N} \frac{1}{|x_i - x_j|^{\beta_1}}$$

$$- c \varepsilon R \sum_{1 \leq i < j \leq N} \frac{1}{|x_i - x_j|^{\beta_1}}.$$

On the other hand, in view of conditions (2.1) and (2.4), it holds that

$$\sqrt{Q_R^N} \leq C \left(R^3 \sum_{i=1}^N |z_{i,1}| + \sqrt{m}|v| + |x| + \left(\sum_{1 \leq i < j \leq N} \frac{1}{|x_i - x_j|^{\beta_1}} \right)^{1/2} + R \right).$$

It follows that

$$\varepsilon \left(1 + |v| + |x| + \sum_{i=1}^N \sum_{\ell=1}^{k_i} |z_{i,\ell}| \right) \sqrt{Q_R^N}$$

$$\leq C\,\varepsilon\left(R^6\sum_{i=1}^{N}|z_i|^2 + |v|^2 + |x|^2 + \sum_{1\leq i<j\leq N}\frac{1}{|x_i-x_j|^{\beta_1}} + R^2\right).$$

As a consequence,

$$I_1 \leq -a_4\varepsilon\sqrt{Q_R^N}\sum_{1\leq i<j\leq N}\frac{1}{|x_i-x_j|^{\beta_1}} + C\varepsilon\left(1 + |v| + |x| + \sum_{i=1}^{N}\sum_{\ell=1}^{k_i}|z_{i,\ell}|\right)\sqrt{Q_R^N}$$

$$\leq -c\varepsilon R^3\sum_{i=1}^{N}|z_{i,1}|\sum_{1\leq i<j\leq N}\frac{1}{|x_i-x_j|^{\beta_1}} - c\varepsilon R\sum_{1\leq i<j\leq N}\frac{1}{|x_i-x_j|^{\beta_1}}$$

$$+ C\,\varepsilon\left(R^6\sum_{i=1}^{N}|z_i|^2 + |v|^2 + |x|^2 + \sum_{1\leq i<j\leq N}\frac{1}{|x_i-x_j|^{\beta_1}} + R^2\right), \qquad (3.45)$$

holds for some positive constants c, C independent of ε and R.

With regard to I_2 on the right-hand side of (3.44), Itô's formula yields the identity

$$\mathcal{L}_{m,0}^{N}\sqrt{Q_R^N} = \frac{1}{\sqrt{Q_R^N}}\sum_{i=1}^{N}\left[\left\langle v_i, \sum_{\ell=1}^{k_i}\lambda_{i,\ell}z_{i,\ell}\right\rangle - R^6\alpha_{i,1}|z_{i,1}|^2 - R^6\langle z_{i,\ell}, v_i\rangle\right.$$

$$\left. + \tfrac{1}{2}\alpha_{i,1}d - \frac{1}{2Q_R^N}\alpha_{i,1}|z_{i,1}|^2\right].$$

Since

$$\sum_{i=1}^{N}U(x_i) + \sum_{1\leq i<j\leq N}G(x_i-x_j) \geq 0,$$

we deduce the bound

$$|\mathcal{L}_{m,0}^{N}\sqrt{Q_R^N}| \leq CR^6\frac{|v|\sum_{i=1}^{N}|z_i| + \sum_{i=1}^{N}|z_i|^2 + 1}{(R^6\sum_{i=1}^{N}|z_i|^2 + |v|^2 + 1)^{1/2}} \leq CR^6\left(\sum_{i=1}^{N}|z_i| + 1\right).$$

It follows that I_2 satisfies

$$I_2 = -\varepsilon m\sum_{1\leq i<j\leq N}\frac{\langle v_i-v_j, x_i-x_j\rangle}{|x_i-x_j|}\mathcal{L}_{m,0}^{N}\sqrt{Q_R^N}$$

$$\leq C\varepsilon\sum_{1\leq i<j\leq N}|v_i-v_j|R^6\left(\sum_{i=1}^{N}|z_i| + 1\right)$$

$$\leq C\varepsilon R^6\sum_{i=1}^{N}|v_i|\left(\sum_{i=1}^{N}|z_i| + 1\right),$$

whence

$$I_2 \le C\varepsilon\left(|v|^2 + R^{12}\sum_{i=1}^N |z_i|^2 + R^{12} \right). \tag{3.46}$$

Now, we collect (3.45), (3.46) together with (3.44) to arrive at the bound

$$\mathcal{L}_{m,0}^N\left(-\varepsilon\sum_{i=1}^N m\left\langle v_i, \sum_{j\ne i}\frac{x_i - x_j}{|x_i - x_j|}\right\rangle \right)\sqrt{Q_R^N}\right)$$

$$\le -c\varepsilon R^3 \sum_{i=1}^N |z_{i,1}| \sum_{1\le i<j\le N}\frac{1}{|x_i - x_j|^{\beta_1}} - c\varepsilon R \sum_{1\le i<j\le N}\frac{1}{|x_i - x_j|^{\beta_1}}$$

$$+ C\varepsilon\left(R^6 \sum_{i=1}^N |z_i|^2 + |v|^2 + |x|^2 + \sum_{1\le i<j\le N}\frac{1}{|x_i - x_j|^{\beta_1}} + R^2 \right)$$

$$+ C\varepsilon\left(|v|^2 + R^{12}\sum_{i=1}^N |z_i|^2 + R^{12} \right),$$

whence

$$\mathcal{L}_{m,0}^N\left(-\varepsilon\sum_{i=1}^N m\left\langle v_i, \sum_{j\ne i}\frac{x_i - x_j}{|x_i - x_j|}\right\rangle \right)\sqrt{Q_R^N}\right)$$

$$\le -c\varepsilon R^3 \sum_{i=1}^N |z_{i,1}| \sum_{1\le i<j\le N}\frac{1}{|x_i - x_j|^{\beta_1}} - c\varepsilon R \sum_{1\le i<j\le N}\frac{1}{|x_i - x_j|^{\beta_1}}$$

$$+ C\varepsilon\left(R^{12}\sum_{i=1}^N |z_i|^2 + |v|^2 + |x|^2 + \sum_{1\le i<j\le N}\frac{1}{|x_i - x_j|^{\beta_1}} + R^{12} \right). \tag{3.47}$$

In the above, we emphasize that c, C are independent of ε and R.

Turning back to V_N^2 given by (3.39), from the estimates (3.41), (3.42), (3.43) and (3.47), we obtain

$$\mathcal{L}_{m,0}^N V_N^2 \le -I_3 + I_4,$$

where

$$I_3 = -\sum_{i=1}^N \sum_{\ell=1}^{k_i} \alpha_{i,\ell}|z_{i,\ell}|^2 - a_2\varepsilon R|x|^2 - \varepsilon R^2 m \sum_{i=1}^N \lambda_{i,1}|v_i|^2$$

$$- c\varepsilon R^3 \sum_{i=1}^N |z_{i,1}| \sum_{1\le i<j\le N}\frac{1}{|x_i - x_j|^{\beta_1}} - c\varepsilon R \sum_{1\le i<j\le N}\frac{1}{|x_i - x_j|^{\beta_1}},$$

and

$$I_4 = \frac{1}{2}\sum_{i=1}^{N}\sum_{\ell=1}^{k_i}\alpha_{i,\ell} + CR\left(\varepsilon|v|^2 + \varepsilon^{1/2}\sum_{i=1}^{N}|z_i|^2 + \varepsilon^{3/2}|x|^2 + 1\right)$$

$$+ C\varepsilon R^2\left(\sum_{i=1}^{N}|z_{i,1}| \sum_{1\le i<j\le N}\frac{1}{|x_i-x_j|^{\beta_1}} + \sum_{i=1}^{N}|z_i|^2 + 1\right)$$

$$+ C\varepsilon R \sum_{1\le i<j\le N}\frac{1}{|x_i-x_j|^{\beta_1-1}} + CR^2\left(\varepsilon^{3/2}(|x|^2+|v|^2) + \varepsilon^{1/2}\sum_{i=1}^{N}|z_i|^2 + \varepsilon\right)$$

$$+ C\varepsilon\left(R^{12}\sum_{i=1}^{N}|z_i|^2 + |v|^2 + |x|^2 + \sum_{1\le i<j\le N}\frac{1}{|x_i-x_j|^{\beta_1}} + R^{12}\right).$$

Since the constants c, C are independent of ε, R, we may infer

$$\mathcal{L}_{m,0}^{N}V_N^2 \le -c\sum_{i=1}^{N}|z_i|^2 - c\varepsilon R|x|^2 - c\varepsilon R^2|v|^2$$

$$- c\varepsilon R^3\sum_{i=1}^{N}|z_{i,1}| \sum_{1\le i<j\le N}\frac{1}{|x_i-x_j|^{\beta_1}} - c\varepsilon R \sum_{1\le i<j\le N}\frac{1}{|x_i-x_j|^{\beta_1}}$$

$$+ C\varepsilon R(1 + R\varepsilon^{1/2})|v|^2 + C\varepsilon(1 + R^2\varepsilon^{1/2})|x|^2 + C\varepsilon^{1/2}R^{12}\sum_{i=1}^{N}|z_i|^2$$

$$+ C\varepsilon R^2\sum_{i=1}^{N}|z_{i,1}| \sum_{1\le i<j\le N}\frac{1}{|x_i-x_j|^{\beta_1}} + C\varepsilon \sum_{1\le i<j\le N}\frac{1}{|x_i-x_j|^{\beta_1}} + CR^{12}.$$

Now, by first taking R sufficiently large and then shrinking ε small enough, we observe that all the positive (non-constant) terms on the above right-hand side are dominated by the negative terms. That is, the following holds

$$\mathcal{L}_{m,0}^{N}V_N^2 \le -c\sum_{i=1}^{N}|z_i|^2 - c\varepsilon R|x|^2 - c\varepsilon R^2|v|^2 - c\varepsilon R \sum_{1\le i<j\le N}\frac{1}{|x_i-x_j|^{\beta_1}} + CR^{12}.$$

This produces the desired Lyapunov property of V_N^2 for (1.5) in the case $\gamma = 0$. The proof is thus finished.   □

3.3 Proof of Theorem 2.8

The proof of Theorem 2.8 is based on the Lyapunov functions constructed in Sect. 3.2 and a local minorization, cf. Definition 3.2, on the transition probabilities $P_t^m(X, \cdot)$. The latter property is summarized in the following auxiliary result.

Lemma 3.7 *Under Assumption 2.1 and Assumption 2.3, system* (1.5) *satisfies the minorization condition as in Definition 3.2.*

The proof of Lemma 3.7 is relatively standard and can be found in literature for the Langevin dynamics (Herzog and Mattingly 2019; Ottobre and Pavliotis 2011). For the sake of completeness, we briefly sketch the argument without going into details.

Sketch of the proof of Lemma 3.7 First of all, by verifying a Hormander's condition, see Ottobre and Pavliotis (2011, page 1639), we note that the operator $\partial_t + \mathcal{L}^N_{m,\gamma}$ ($\gamma \geq 0$) is hypoelliptic (Hörmander 1967). This implies that the transition probabilities P^m_t has a smooth density in **X**. Furthermore, we may adapt the proof of Herzog and Mattingly (2019, Proposition 2.5) to study the control problem for (1.5). In particular, by the Stroock-Varadhan support Theorem, it can be shown that $\mathbb{P}^m_t(X, A) > 0$ for $t > 0$, $X \in \mathbf{X}$ and every open set $A \subset \mathbf{X}$. We then employ the same argument as those in Herzog and Mattingly (2019,Corollary 5.12) and Mattingly et al. (2002, Lemma 2.3) to conclude the inequality (3.2), thereby concluding the minorization.   □

We are now in a position to conclude Theorem 2.8. The proof employs the weak-Harris theorem proved in (Hairer and Mattingly 2011, Theorem 1.2).

Proof of Theorem 2.8 First of all, it is clear that π_N defined in (1.7) is an invariant measure for (1.5). Next, we observe that from Lemma 3.5 and Lemma 3.6, the functions V^1_N and V^2_N, respectively, defined in (3.31) and (3.39) are Lyapunov functions for (1.5) in the case $\gamma > 0$ and $\gamma = 0$. It follows that Hairer and Mattingly (2011, Assumption 1) holds. On the other hand, Lemma 3.7 verifies Hairer and Mattingly (2011, Assumption 2). In view of Hairer and Mattingly (2011, Theorem 1.2), we conclude the uniqueness of π_N as well as the exponential convergent rate (2.16).   □

4 Small-Mass Limit

We turn to the topic of the small-mass limit for (1.5). In Sect. 4.1, we provide a heuristic argument on how we derive the limiting system (1.10) as $m \to 0$. We also outline the main steps of the proof of Theorem 2.10 in this section. In Sect. 4.2, we prove a partial result on the small-mass limit assuming the nonlinearities are globally Lipschitz. In Sect. 4.3, we establish useful moment estimates on the limiting system (1.10). Lastly, in Sect. 4.4, we establish Theorem 2.10 while making use of the auxiliary results from Sect. 4.2 and Sect. 4.3.

4.1 Heuristic Argument for the Limiting System (1.10)

In this subsection, we provide a heuristic argument detailing how we derive the limiting system (1.10) as well as its initial conditions from the original system (1.5). The argument draws upon recent works in Herzog et al. (2016), Nguyen (2018) where similar issues were dealt with in the absence of singular potentials. We formally set $m = 0$ on the left-hand side of the v_i−equation in (1.5) while substituting the $v_i(t)$

term on the right-hand side by $\mathrm{d}x_i(t)$ to obtain

$$\gamma \mathrm{d}x_i(t) = \left[-\nabla U(x_i(t)) - \sum_{j \neq i} \nabla G(x_i(t) - x_j(t)) + \sum_{\ell=1}^{k_i} \lambda_{i,\ell} z_{i,\ell}(t) \right] \mathrm{d}t$$
$$+ \sqrt{2\gamma} d W_0(t).$$

Next, considering the $z_{i,\ell}$−equation in (1.5), by Duhamel's formula, $z_{i,\ell}(t)$ may be written as:

$$z_{i,\ell}(t) = e^{-\alpha_{i,\ell} t} z_{i,\ell}(0) - \lambda_{i,\ell} \int_0^t e^{-\alpha_{i,\ell}(t-r)} v_i(r) \mathrm{d}r + \sqrt{2\alpha_{i,\ell}} \int_0^t e^{-\alpha_{i,\ell}(t-r)} \mathrm{d}W_{i,\ell}(r). \tag{4.1}$$

We note that the above expression still depends on $v_i(t)$. Nevertheless, this can be circumvented by employing an integration by parts as follows:

$$\int_0^t e^{-\alpha_{i,\ell}(t-r)} v_i(r) \mathrm{d}r = x_i(t) - e^{-\alpha_{i,\ell} t} x_i(0) + \alpha_{i,\ell} \int_0^t e^{-\alpha_{i,\ell}(t-r)} x_i(r) \mathrm{d}r.$$

Alternatively, we note that the above identity can be derived by applying Itô's formula to $e^{\alpha_{i,\ell} t} x_i(t)$. Plugging back into (4.1), we find

$$z_{i,\ell}(t) = e^{-\alpha_{i,\ell} t} (z_{i,\ell}(0) + \lambda_{i,\ell} x_i(0)) - \lambda_{i,\ell} x_i(t) - \lambda_{i,\ell} \alpha_{i,\ell} \int_0^t e^{-\alpha_{i,\ell}(t-r)} x_i(r) \mathrm{d}r$$
$$+ \sqrt{2\alpha_{i,\ell}} \int_0^t e^{-\alpha_{i,\ell}(t-r)} \mathrm{d}W_{i,\ell}(r), \tag{4.2}$$

whence

$$z_{i,\ell}(t) + \lambda_{i,\ell} x_i(t) = e^{-\alpha_{i,\ell} t} (z_{i,\ell}(0) + \lambda_{i,\ell} x_i(0)) - \lambda_{i,\ell} \alpha_{i,\ell} \int_0^t e^{-\alpha_{i,\ell}(t-r)} x_i(r) \mathrm{d}r$$
$$+ \sqrt{2\alpha_{i,\ell}} \int_0^t e^{-\alpha_{i,\ell}(t-r)} \mathrm{d}W_{i,\ell}(r).$$

Setting $f_{i,\ell}(t) := z_{i,\ell}(t) + \lambda_{i,\ell} x_i(t)$, we observe that (using Duhamel's formula again)

$$\mathrm{d}f_{i,\ell}(t) = -\alpha_{i,\ell} f_{i,\ell}(t) + \lambda_{i,\ell} \alpha_{i,\ell} x_i(t) \mathrm{d}t + \sqrt{2\alpha_{i,\ell}} \mathrm{d}W_{i,\ell}(t), \quad \ell = 1, \ldots, k_i,$$
$$f_{i,\ell}(0) = z_{i,\ell}(0) + \lambda_{i,\ell} x_i(0).$$

This together with setting $q_i(t) := x_i(t)$ deduces the limiting system (1.10) as well as the corresponding shifted initial conditions as in Theorem 2.10.

Next, for the reader's convenience, we summarize the idea of the proof of Theorem 2.10. The argument essentially consists of three steps as follows (Herzog et al. 2016; Lim and Wehr 2019; Lim et al. 2020).

Step 1: We first truncate the nonlinear potentials in (1.5) and (1.10) obtaining a truncated system whose coefficients are globally Lipschitz, and establish a convergence in probability in the small mass limit. This result appears in Proposition 4.1 found in Sect. 4.2.

Step 2: Next, we establish an exponential moment bound on any finite-time window for the limiting system (1.10). This is discussed in details in Sect. 4.3, cf. Lemma 4.3.

Step 3: We prove Theorem 2.10 by removing the Lipschitz constraint from Proposition 4.1 while making use of the energy estimates in Lemma 4.3.

4.2 Truncating (1.5) and (1.10)

For $R > 2$, let $\theta_R : [0, \infty) \to \mathbb{R}$ be a smooth function satisfying

$$\theta_R(t) = \begin{cases} 1, & 0 \le t \le R, \\ \text{decreasing}, & R \le t \le R+1, \\ 0, & t \ge R+1. \end{cases} \tag{4.3}$$

With the above cut-off θ_R, we consider a truncating version of (1.5) given by

$$\mathrm{d}\,x_i(t) = v_i(t)\mathrm{d}t, \qquad i = 1, \ldots, N,$$
$$m\mathrm{d}\,v_i(t) = -\gamma v_i(t)\mathrm{d}t - \theta_R(|x_i(t)|)\nabla U(x_i(t))\mathrm{d}t + \sqrt{2\gamma}\,\mathrm{d}W_{i,0}(t)$$
$$- \sum_{j \ne i} \theta_R\big(|x_i(t) - x_j(t)|^{-1}\big)\nabla G\big(x_i(t) - x_j(t)\big)\mathrm{d}t + \sum_{\ell=1}^{k_i} \lambda_{i,\ell} z_{i,\ell}(t)\mathrm{d}t,$$
$$\mathrm{d}\,z_{i,\ell}(t) = -\alpha_{i,\ell} z_{i,\ell}(t)\mathrm{d}t - \lambda_{i,\ell} v_i(t)\mathrm{d}t + \sqrt{2\alpha_{i,\ell}}\,\mathrm{d}W_{i,\ell}(t), \quad \ell = 1, \ldots, k_i, \tag{4.4}$$

as well as the following truncated version of (1.10)

$$\gamma \mathrm{d}q_i(t) = -\theta_R(|q_i(t)|)\nabla U(q_i(t))\mathrm{d}t - \sum_{j \ne i} \theta_R(|q_i(t) - q_j(t)|^{-1})\nabla G\big(q_i(t) - q_j(t)\big)\mathrm{d}t$$
$$- \sum_{i=1}^{k_i} \lambda_{i,\ell}^2 q_i(t)\mathrm{d}t + \sum_{\ell=1}^{k_i} \lambda_{i,\ell} f_{i,\ell}(t)\mathrm{d}t + \sqrt{2\gamma}\mathrm{d}W_{i,0}(t),$$
$$\mathrm{d}f_{i,\ell}(t) = -\alpha_{i,\ell} f_{i,\ell}(t) + \lambda_{i,\ell} \alpha_{i,\ell} q_i(t)\mathrm{d}t + \sqrt{2\alpha_{i,\ell}}\mathrm{d}W_{i,\ell}(t), \quad \ell = 1, \ldots, k_i,$$
$$q_i(0) = x_i(0), \quad f_{i,\ell}(0) = z_{i,\ell}(0) + \lambda_{i,\ell} x_i(0). \tag{4.5}$$

We now show that system (4.4) can be approximated by (4.5) on any finite-time window in the small-mass regime.

Proposition 4.1 *Under Assumptions 2.1 and 2.3, given any initial condition* $(\mathrm{x}(0), \mathrm{v}(0), \mathrm{z}_1(0), \ldots, \mathrm{z}_N(0)) \in \mathbf{X}$ *and* $R > 2$, *let* $\big(\mathrm{x}_m^R(t), \mathrm{v}_m^R(t), \mathrm{z}_{1,m}^R(t), \ldots, \mathrm{z}_{N,m}^R(t)\big)$ *and* $\big(\mathrm{q}^R(t), \mathrm{f}_1^R(t), \ldots, \mathrm{f}_N^R(t)\big)$ *respectively solve (4.4) and (4.5). Then, for every* $T > 0$, *the following holds*

$$\mathbb{E}\left[\sup_{0\le t\le T} \left|x_m^R(t) - q^R(t)\right|^4\right] \le m \cdot C, \quad as \quad m \to 0, \tag{4.6}$$

for some positive constant $C = C(T, R)$ independent of m.

In order to establish Proposition 4.1, it is crucial to derive useful moment bounds on the velocity process $v_m^R(t)$. More precisely, we have the following result.

Lemma 4.2 *Under Assumptions 2.1 and 2.3, given any initial condition $(x(0), v(0), z_1(0), \ldots, z_N(0)) \in \mathbf{X}$ and $R > 2$, let $\left(x_m^R(t), v_m^R(t), z_{1,m}^R(t), \ldots, z_{N,m}^R(t)\right)$ be the solution of (4.4). Then, for every $T > 0$, $n > 1$ and $\varepsilon > 0$, it holds that*

$$m^n \mathbb{E}\left[\sup_{0\le t\le T} \left|v_m^R(t)\right|^n\right] \le m^{\frac{n}{2}-\varepsilon} C, \quad as \quad m \to 0, \tag{4.7}$$

for some positive constant $C = C(T, n, R, \varepsilon)$ independent of m.

For the sake of clarity, the proof of Lemma 4.2 will be deferred to the end of this subsection. In what follows, we will assume Lemma 4.2 holds and prove Proposition 4.1. The argument is adapted from the proof of (Nguyen 2018, Proposition 9) tailored to our settings.

Proof of Proposition 4.1 From the (x_i, v_i)−equations in (4.4), we find

$$m dv_i^R(t) + \gamma dx_i^R(t) = -\theta_R(x_i^R(t))\nabla U(x_i^R(t))dt + \sqrt{2\gamma}\, dW_{i,0}(t)$$
$$- \sum_{j\ne i} \theta_R\left(|x_i^R(t) - x_j^R(t)|^{-1}\right)\nabla G\left(x_i^R(t) - x_j^R(t)\right)dt$$
$$+ \sum_{\ell=1}^{k_i} \lambda_{i,\ell} z_{i,\ell}^R(t)dt.$$

Substituting $z_{i,\ell}^R(t)$ by the expression (4.2) into the above equation produces

$$m dv_i^R(t) + \gamma dx_i^R(t)$$
$$= -\theta_R(|x_i^R(t)|)\nabla U(x_i^R(t))dt$$
$$+ \sqrt{2\gamma}\, dW_{i,0}(t) - \sum_{j\ne i} \theta_R\left(|x_i^R(t) - x_j^R(t)|^{-1}\right)\nabla G\left(x_i^R(t) - x_j^R(t)\right)dt$$
$$+ \sum_{\ell=1}^{k_i} \lambda_{i,\ell} e^{-\alpha_{i,\ell} t}\left[z_{i,\ell}(0) + \lambda_{i,\ell} x_i(0)\right]$$
$$- \left(\sum_{\ell=1}^{k_i} \lambda_{i,\ell}^2\right) x_i^R(t)$$
$$- \sum_{\ell=1}^{k_i} \lambda_{i,\ell}^2 \alpha_{i,\ell} \int_0^t e^{-\alpha_{i,\ell}(t-r)} x_i^R(r)dr$$

$$+ \sum_{\ell=1}^{k_i} \lambda_{i,\ell} \sqrt{2\alpha_{i,\ell}} \int_0^t e^{-\alpha_{i,\ell}(t-r)} \mathrm{d}W_{i,\ell}(r)\mathrm{d}t. \tag{4.8}$$

Similarly, from the $f_{i,\ell}-$equation of system (4.5), we have

$$f_{i,\ell}^R(t) = e^{-\alpha_{i,\ell}t}\left(z_{i,\ell}(0)+\lambda_{i,\ell}x_i(0)\right) - \lambda_{i,\ell}\alpha_{i,\ell}\int_0^t e^{-\alpha_{i,\ell}(t-r)}q_i^R(r)\mathrm{d}r$$
$$+ \sqrt{2\alpha_{i,\ell}}\int_0^t e^{-\alpha_{i,\ell}(t-r)}\mathrm{d}W_{i,\ell}(r).$$

Plugging into q_i-equation of (4.5) yields

$$\gamma \,\mathrm{d}q_i^R(t)$$
$$= -\theta_R(|q_i^R(t)|)\nabla U(q_i^R(t))\mathrm{d}t$$
$$+ \sqrt{2\gamma}\,\mathrm{d}W_{i,0}(t - \sum_{j\neq i}\theta_R(|q_i^R(t)-q_j^R(t)|^{-1})\nabla G(q_i^R(t)-q_j^R(t))\mathrm{d}t$$
$$+ \sum_{\ell=1}^{k_i}\lambda_{i,\ell}e^{-\alpha_{i,\ell}t}\left[z_{i,\ell}(0)+\lambda_{i,\ell}x_i(0)\right] - \left(\sum_{\ell=1}^{k_i}\lambda_{i,\ell}^2\right)q_i^R(t)$$
$$- \sum_{\ell=1}^{k_i}\lambda_{i,\ell}^2\alpha_{i,\ell}\int_0^t e^{-\alpha_{i,\ell}(t-r)}x_i^R(r)\mathrm{d}r$$
$$+ \sum_{\ell=1}^{k_i}\lambda_{i,\ell}\sqrt{2\alpha_{i,\ell}}\int_0^t e^{-\alpha_{i,\ell}(t-r)}\mathrm{d}W_{i,\ell}(r)\mathrm{d}t. \tag{4.9}$$

Setting $\bar{x}_i^R = x_i^R - q_i^R$, we subtract (4.9) from (4.8) to obtain the identity

$$m\mathrm{d}v_i^R(t) + \gamma\mathrm{d}\bar{x}_i^R(t) = -\left[\theta_R(|x_i^R(t)|)\nabla U(x_i^R(t)) - \theta_R(|q_i^R(t)|)\nabla U(q_i^R(t))\right]\mathrm{d}t$$
$$- \sum_{j\neq i}\left[\theta_R(|x_i^R(t)-x_j^R(t)|^{-1})\nabla G(x_i^R(t)-x_j^R(t))\right.$$
$$\left. - \theta_R(|q_i^R(t)-q_j^R(t)|^{-1})\nabla G(q_i^R(t)-q_j^R(t))\right]\mathrm{d}t$$
$$- \left(\sum_{\ell=1}^{k_i}\lambda_{i,\ell}^2\right)\bar{x}_i^R(t) - \sum_{\ell=1}^{k_i}\lambda_{i,\ell}^2\alpha_{i,\ell}\int_0^t e^{-\alpha_{i,\ell}(t-r)}\bar{x}_i^R(r)\mathrm{d}r.$$
$$\tag{4.10}$$

By the choice of θ_R as in (4.3) and conditions (2.2) and (2.6), we invoke the mean value theorem to infer

$$|\theta_R(|x|)\nabla U(x) - \theta_R(|y|)\nabla U(y)| \le C|x-y|, \quad x,y \in \mathbb{R}^d,$$

and

$$|\theta_R(|x|^{-1})\nabla G(x) - \theta_R(|y|^{-1})\nabla G(y)| \leq C|x - y|, \quad x, y \in \mathbb{R}^d \setminus \{0\},$$

for some positive constant $C = C(R)$. From (4.10), we arrive at the a.s. bound

$$|\bar{x}_i^R(t)|^n \leq Cm^n|v_i^R(t) - v(0)|^n + C\int_0^t \sum_{j=1}^N |\bar{x}_j^R(r)|^n \mathrm{d}r,$$

whence

$$|\bar{\mathbf{x}}^R(t)|^n \leq Cm^n|\mathbf{v}^R(t) - \mathbf{v}(0)|^n + C\int_0^t \sum_{j=1}^N |\bar{\mathbf{x}}^R(r)|^n \mathrm{d}r.$$

Gronwall's inequality implies that

$$\mathbb{E}\sup_{t\in[0,T]}|\bar{\mathbf{x}}^R(t)|^n \leq Cm^n\mathbb{E}\sup_{t\in[0,T]}|\mathbf{v}^R(t) - \mathbf{v}(0)|^n.$$

Setting $n = 4$, by virtue of Lemma 4.2, this produces the small-mass limit (4.6), as claimed. □

We now turn to the proof of Lemma 4.2.

Proof of Lemma 4.2 From the v_i−equation in (4.4), variation constant formula yields

$$mv_i^R(t) = me^{-\frac{\gamma}{m}t}v_i(0) - \int_0^t e^{-\frac{\gamma}{m}(t-r)}\theta_R(|x_i^R(r)|)\nabla U(x_i^R(r))\mathrm{d}r$$

$$- \int_0^t e^{-\frac{\gamma}{m}(t-r)}\sum_{j\neq i}\left[\theta_R(|x_i^R(r) - x_j^R(r)|^{-1})\nabla G(x_i^R(r) - x_j^R(r))\right]\mathrm{d}r$$

$$+ \sum_{\ell=1}^{k_i}\lambda_{i,\ell}\int_0^t e^{-\frac{\gamma}{m}(t-r)}z_{i,\ell}^R(r)\mathrm{d}r + \sqrt{2\gamma}\int_0^t e^{-\frac{\gamma}{m}(t-r)}\mathrm{d}W_{i,0}(r). \quad (4.11)$$

With regard to the integral involving $z_{i,\ell}$, since $z_{i,\ell}^R$ satisfies the third equation in (4.4), we have

$$z_{i,\ell}^R(t) = e^{-\alpha_{i\ell}t}z_{i,\ell}(0) - \lambda_{i,\ell}\int_0^t e^{-\alpha_{i,\ell}(t-r)}v_i^R(r)\mathrm{d}r + \sqrt{2\alpha_{i,\ell}}\int_0^t e^{-\alpha_{i,\ell}(t-r)}\mathrm{d}W_{i,\ell}(r).$$

Plugging back into (4.11), we obtain the identity

$$mv_i^R(t) = me^{-\frac{\gamma}{m}t}v_i(0) - \int_0^t e^{-\frac{\gamma}{m}(t-r)}\theta_R(|x_i^R(r)|)\nabla U(x_i^R(r))\mathrm{d}r$$

$$- \int_0^t e^{-\frac{\gamma}{m}(t-r)} \sum_{j \neq i} \left[\theta_R \left(|x_i^R(r) - x_j^R(r)|^{-1} \right) \nabla G(x_i^R(r) - x_j^R(r)) \right] dr$$

$$+ \sum_{\ell=1}^{k_i} \lambda_{i,\ell} \int_0^t e^{-\frac{\gamma}{m}(t-r)} e^{-\alpha_{i,\ell} r} z_{i,\ell}(0) dr$$

$$- \sum_{\ell=1}^{k_i} \lambda_{i,\ell}^2 \int_0^t e^{-\frac{\gamma}{m}(t-r)} \int_0^r e^{-\alpha_{i,\ell}(r-s)} v_i^R(s) ds dr$$

$$+ \sum_{\ell=1}^{k_i} \lambda_{i,\ell} \sqrt{2\alpha_{i,\ell}} \int_0^t e^{-\frac{\gamma}{m}(t-r)} \int_0^r e^{-\alpha_{i,\ell}(r-s)} dW_{i,\ell}(s) dr$$

$$+ \sqrt{2\gamma} \int_0^t e^{-\frac{\gamma}{m}(t-r)} dW_{i,0}(r)$$

$$= m e^{-\frac{\gamma}{m}t} v_i(0) - I_1 - I_2 + I_3 - I_4 + I_5 + I_6. \tag{4.12}$$

Concerning I_1, we invoke condition (2.1) while making use of the choice of θ_R as in (4.3) to obtain

$$|I_1| \leq C \int_0^t e^{-\frac{\gamma}{m}(t-r)} dr = mC,$$

for some positive constant $C = C(R)$ independent of m. Likewise, we employ condition (2.5) to infer

$$|I_2| \leq C \int_0^t e^{-\frac{\gamma}{m}(t-r)} dr = mC.$$

Similarly, it is clear that

$$|I_3| \leq m \sum_{\ell=1}^{k_i} \lambda_{i,\ell} |z_{i,\ell}(0)|.$$

Concerning I_4, observe that

$$|I_4| \leq \sum_{\ell=1}^{k_i} \lambda_{i,\ell}^2 \int_0^t e^{-\frac{\gamma}{m}(t-r)} dr \int_0^t \sup_{s \in [0,r]} |v_i^R(s)| dr \leq m \sum_{\ell=1}^{k_i} \lambda_{i,\ell}^2 \int_0^t \sup_{s \in [0,r]} |v_i^R(s)| dr.$$

With regard to the noise term I_5, note that

$$\left| \int_0^t e^{-\frac{\gamma}{m}(t-r)} \int_0^r e^{-\alpha_{i,\ell}(r-s)} dW_{i,\ell}(s) dr \right|^n$$

$$\leq \left| \int_0^t e^{-\frac{\gamma}{m}(t-r)} dr \right|^n \sup_{r \in [0,T]} \left| \int_0^r e^{-\alpha_{i,\ell}(r-s)} dW_{i,\ell}(s) \right|^n$$

$$\leq \frac{m^n}{\gamma^n} \sup_{r \in [0,T]} \left| \int_0^r e^{-\alpha_{i,\ell}(r-s)} \mathrm{d}W_{i,\ell}(s) \right|^n.$$

We employ Burkholder's inequality to infer

$$\mathbb{E} \sup_{t \in [0,T]} \left| \int_0^t e^{-\frac{\gamma}{m}(t-r)} \int_0^r e^{-\alpha_{i,\ell}(r-s)} \mathrm{d}W_{i,\ell}(s) \mathrm{d}r \right|^n$$

$$\leq \frac{m^n}{\gamma^n} \mathbb{E} \sup_{t \in [0,T]} \left| \int_0^t e^{\alpha_{i,\ell} s} \mathrm{d}W_{i,\ell}(s) \right|^n \leq m^n C(T,n).$$

Turning to I_6, we set

$$Y_{i,0}(t) := \int_0^t e^{-\beta(t-r)} \mathrm{d}W_{i,0}(r), \quad \beta := \frac{\gamma}{m},$$

and let $Z_{i,0} \sim N(0,1)$ be a random variable independent of $W_{i,0}(t)$. Then, the process

$$X_{i,0}(t) := Z_{i,0} e^{-\beta t} + \sqrt{2\beta} Y_{i,0}(t)$$

is a stationary solution to

$$\mathrm{d}X_{i,0}(t) = -\beta X_{i,0}(t) \mathrm{d}t + \sqrt{2\beta} \mathrm{d}W_{i,0}(t), \quad X_{i,0}(0) = Z.$$

For $n \geq 1$, it holds by the definition of $X_{i,0}(t)$ that

$$\mathbb{E} \sup_{t \in [0,T]} |Y_{i,0}(t)|^n \leq C\beta^{-\frac{n}{2}} \left(1 + \mathbb{E} \sup_{t \in [0,T]} |X_{i,0}(t)|^n \right).$$

By Pavliotis et al. (2022, Lemma B.1), it holds for all $n > 1$ that

$$\mathbb{E} \sup_{t \in [0,T]} |X_{i,0}(t)|^n \leq C\big(1 + \log(1 + \beta T)\big)^{n/2}.$$

As a consequence, for all $\varepsilon > 0$

$$\beta^{\frac{n}{2} - \varepsilon} \mathbb{E} \sup_{t \in [0,T]} |Y_{i,0}(t)|^n < C(T,n).$$

In other words,

$$\mathbb{E} \sup_{t \in [0,T]} |Y_{i,0}(t)|^n < m^{\frac{n}{2} - \varepsilon} C(T,n).$$

Now we collect the above estimates together with expression (4.12) to deduce that for all $n > 1$ and $\varepsilon > 0$

$$m^n \mathbb{E} \sup_{t \in [0,T]} |\mathrm{v}^R(t)|^n$$

$$\leq m^n C\Big(|v(0)|^n + \sum_{i=1}^{N} |z_i^R(0)|^n + 1\Big) + m^{\frac{n}{2}-\varepsilon} C + Cm^n \int_0^T \mathbb{E} \sup_{r\in[0,t]} |v^R(r)|^n dt,$$

holds for some positive constant $C = C(T, n, R, \varepsilon)$ independent of m. In view of Gronwall's inequality, for all $n > 1$, we arrive at (4.7), as claimed. $\square$

4.3 Estimates on (1.10)

In this subsection, we provide several energy estimates on the limiting system (1.10) on any finite-time window. More precisely, we have the following result.

Lemma 4.3 *Under Assumptions 2.1, 2.3 and 2.5, for all* $(x(0), z_1(0),..., z_N(0)) \in \mathcal{D} \times (\mathbb{R}^d)^N$, *let* $Q(t) = \big(q(t), f_1(t),\ldots, f_N(t)\big)$ *be the solution of* (1.10) *and* β_1 *be the constant as in Assumptions 2.3 and 2.5. For all* $\varepsilon, \kappa > 0$ *sufficiently small and* $T > 0$, *the followings hold:*
   (a) If $\beta_1 > 1$,

$$\mathbb{E}\Bigg[\exp\Bigg\{ \sup_{t\in[0,T]} \Big(\kappa|q(t)|^2 + \kappa\varepsilon \sum_{1\leq i < j \leq N} \frac{1}{|q_i(t) - q_j(t)|^{\beta_1 - 1}} \Big) \Bigg\} \Bigg] \leq C, \quad (4.13)$$

for some positive constant $C = C(\kappa, \varepsilon, T, x(0), z_1(0), \ldots, z_N(0))$.
   (b) Otherwise, if $\beta_1 = 1$,

$$\mathbb{E}\Bigg[\exp\Bigg\{ \sup_{t\in[0,T]} \Big(\kappa|q(t)|^2 - \kappa\varepsilon \sum_{1\leq i < j \leq N} \log|q_i(t) - q_j(t)| \Big) \Bigg\} \Bigg] \leq C. \quad (4.14)$$

The proof of Lemma 4.3 relies on two ingredients: the choice of Lyapunov functions specifically designed for (1.10) and the exponential Martingale inequality. Later in Sect. 4.4, we will particularly exploit Lemma 4.3 to remove the Lipschitz constraint in Proposition 4.1 so as to conclude the main Theorem 2.10.

Proof of Lemma 4.3 (a) Suppose that $\beta_1 > 1$. For $\varepsilon > 0$, we consider the functions Γ_1 given by

$$\Gamma_1(q, f_1, \ldots, f_N) = \frac{1}{2}\gamma|q|^2 + \frac{1}{2}\sum_{i=1}^{N}\sum_{\ell=1}^{k_i} \frac{1}{\alpha_{i,\ell}} |f_{i,\ell}|^2 + \varepsilon\gamma \sum_{1\leq i<j\leq N} \frac{1}{|q_i - q_j|^{\beta_1-1}}. \quad (4.15)$$

We aim to show that $\Gamma_1(t)$ satisfies suitable energy estimate, allowing for establishing the moment bound in sup norm (4.13).
   From (1.10), we employ Itô's formula to obtain

$$d\Bigg(\frac{1}{2}\gamma|q(t)|^2 + \frac{1}{2}\sum_{i=1}^{N}\sum_{\ell=1}^{k_i} \frac{1}{\alpha_{i,\ell}} |f_{i,\ell}(t)|^2 \Bigg)$$

$$= -\sum_{i=1}^{N}\left(\sum_{\ell=1}^{k_i}\lambda_{i,\ell}^2\right)|q_i(t)|^2\mathrm{d}t - \sum_{i=1}^{N}\langle\nabla U(q_i(t)), q_i(t)\rangle\mathrm{d}t + N\gamma^2 d\,\mathrm{d}t$$

$$-\sum_{1\le i<j\le N}\langle\nabla G(q_i(t)-q_j(t)), q_i(t)-q_j(t)\rangle\mathrm{d}t + \sum_{i=1}^{N}\sqrt{2\gamma}\langle q_i(t),\mathrm{d}W_{i,0}(t)\rangle$$

$$-\sum_{i=1}^{N}\sum_{\ell=1}^{k_i}|f_{i,\ell}|^2\mathrm{d}t + \alpha_{i,\ell}^2 d\,\mathrm{d}t + \sqrt{2\alpha_{i,\ell}}\langle f_{i,\ell}(t),\mathrm{d}W_{i,\ell}(t)\rangle.$$

Recalling condition (2.3), we readily have

$$-\sum_{i=1}^{N}\langle\nabla U(q_i(t)), q_i(t)\rangle \le -Na_2\sum_{i=1}^{N}|q_i(t)|^{\lambda+1} + Na_3.$$

To bound the cross terms involving G, we invoke (2.5) and obtain

$$-\sum_{1\le i<j\le N}\langle\nabla G(q_i(t)-q_j(t)), q_i(t)-q_j(t)\rangle$$

$$\le a_1\sum_{1\le i<j\le N}\left(|q_i(t)-q_j(t)| + \frac{1}{|q_i(t)-q_j(t)|^{\beta_1-1}}\right).$$

Since $\sum_{1\le i<j\le N}|q_i - q_j|$ can be subsumed into $-|\mathbf{q}|^2$, we infer from the above estimates that

$$\mathrm{d}\left(\frac{1}{2}\gamma|\mathbf{q}(t)|^2 + \frac{1}{2}\sum_{i=1}^{N}\sum_{\ell=1}^{k_i}\frac{1}{\alpha_{i,\ell}}|f_{i,\ell}(t)|^2\right)$$

$$\le -c|\mathbf{q}(t)|^2\mathrm{d}t - c|\mathbf{q}(t)|^{\lambda+1}\mathrm{d}t - \sum_{i=1}^{N}|f_i|^2 + C\mathrm{d}t + a_1\sum_{1\le i<j\le N}\frac{1}{|q_i-q_j|^{\beta_1-1}}\mathrm{d}t$$

$$+\sum_{i=1}^{N}\sqrt{2\gamma}\langle q_i(t),\mathrm{d}W_{i,0}(t)\rangle + \sum_{i=1}^{N}\sum_{\ell=1}^{k_i}\sqrt{2\alpha_{i,\ell}}\langle f_{i,\ell}(t),\mathrm{d}W_{i,\ell}(t)\rangle, \qquad (4.16)$$

for some positive constant c, C independent of t.

Turning to the last term on the right-hand side of (4.15), a routine computation gives

$$\mathrm{d}\left(\varepsilon\gamma\sum_{1\le i<j\le N}\frac{1}{|q_i(t)-q_j(t)|^{\beta_1-1}}\right)$$

$$= -\varepsilon(\beta_1-1)\sum_{i=1}^{N}\left\langle\sum_{j\ne i}\frac{q_i(t)-q_j(t)}{|q_i(t)-q_j(t)|^{\beta_1+1}}, -\nabla U(q_i(t))\mathrm{d}t\right.$$

$$- \sum_{\ell \neq i} \nabla G \big(q_i(t) - q_\ell(t) \big) dt$$

$$- \sum_{i=1}^{k_i} \lambda_{i,\ell}^2 q_i(t) dt + \sum_{\ell=1}^{k_i} \lambda_{i,\ell} f_{i,\ell}(t) dt + \sqrt{2\gamma} dW_{i,0}(t) \Big\rangle$$

$$- \varepsilon(\beta_1 - 1)(\beta_1 + 1 - d) \sum_{1 \leq i < j \leq N} \frac{1}{|q_i(t) - q_j(t)|^{\beta_1 + 1}} dt. \qquad (4.17)$$

Using Cauchy–Schwarz inequality, it is clear that

$$- \varepsilon(\beta_1 - 1) \sum_{i=1}^{N} \Big\langle \sum_{j \neq i} \frac{q_i(t) - q_j(t)}{|q_i(t) - q_j(t)|^{\beta_1 + 1}}, - \sum_{i=1}^{k_i} \lambda_{i,\ell}^2 q_i(t) + \sum_{\ell=1}^{k_i} \lambda_{i,\ell} f_{i,\ell}(t) \Big\rangle$$

$$\leq C\varepsilon^{1/2} |q(t)|^2 + C\varepsilon^{1/2} \sum_{i=1}^{N} |f_i(t)|^2 + C\varepsilon^{3/2} \sum_{1 \leq i < j \leq N} \frac{1}{|q_i(t) - q_j(t)|^{2\beta_1}}.$$

Concerning the cross terms involving U on the right-hand side of (4.17), we recall that ∇U satisfies (2.2). In light of the mean value theorem, we infer

$$- \varepsilon(\beta_1 - 1) \sum_{i=1}^{N} \Big\langle \sum_{j \neq i} \frac{q_i(t) - q_j(t)}{|q_i(t) - q_j(t)|^{\beta_1 + 1}}, - \nabla U(q_i(t)) \Big\rangle$$

$$= \varepsilon(\beta_1 - 1) \sum_{1 \leq i < j \leq N} \frac{\langle q_i(t) - q_j(t), \nabla U(q_i(t)) - \nabla U(q_j(t)) \rangle}{|q_i(t) - q_j(t)|^2}$$

$$\leq C\varepsilon \sum_{1 \leq i < j \leq N} \big(|q_i(t)|^{\lambda - 1} + |q_j(t)|^{\lambda - 1} \big) \leq C\varepsilon |q(t)|^{\lambda - 1}.$$

Turning to the cross terms involving G on the right-hand side of (4.25), we recast them as follows:

$$- \varepsilon(\beta_1 - 1) \sum_{i=1}^{N} \Big\langle \sum_{j \neq i} \frac{q_i(t) - q_j(t)}{|q_i(t) - q_j(t)|^{\beta_1 + 1}}, - \sum_{\ell \neq i} \nabla G \big(q_i(t) - q_\ell(t) \big) \Big\rangle$$

$$= -\varepsilon(\beta_1 - 1) \sum_{i=1}^{N} \Big\langle \sum_{j \neq i} \frac{q_i(t) - q_j(t)}{|q_i(t) - q_j(t)|^{\beta_1 + 1}}, a_4 \sum_{\ell \neq i} \frac{q_i(t) - q_\ell(t)}{|q_i(t) - q_\ell(t)|^{\beta_1 + 1}} \Big\rangle$$

$$+ \varepsilon(\beta_1 - 1) \sum_{i=1}^{N} \Big\langle \sum_{j \neq i} \frac{q_i(t) - q_j(t)}{|q_i(t) - q_j(t)|^{\beta_1 + 1}}, \sum_{\ell \neq i} \nabla G \big(q_i(t) - q_\ell(t) \big)$$

$$+ a_4 \frac{q_i(t) - q_\ell(t)}{|q_i(t) - q_\ell(t)|^{\beta_1 + 1}} \Big\rangle.$$

In view of Lemma A.2, cf. (A.2), we find

$$
-\varepsilon(\beta_1 - 1) \sum_{i=1}^{N} \left\langle \sum_{j \neq i} \frac{q_i(t) - q_j(t)}{|q_i(t) - q_j(t)|^{\beta_1+1}}, a_4 \sum_{\ell \neq i} \frac{q_i(t) - q_\ell(t)}{|q_i(t) - q_\ell(t)|^{\beta_1+1}} \right\rangle
$$
$$
\leq -\varepsilon(\beta_1 - 1) a_4 \cdot \frac{4}{N(N-1)^2} \sum_{1 \leq i < j \leq N} \frac{1}{|q_i(t) - q_j(t)|^{2\beta_1}}.
$$

On the other hand, the condition (2.7) implies

$$
\varepsilon(\beta_1 - 1) \sum_{i=1}^{N} \left\langle \sum_{j \neq i} \frac{q_i(t) - q_j(t)}{|q_i(t) - q_j(t)|^{\beta_1+1}}, \sum_{\ell \neq i} \nabla G\big(q_i(t) - q_\ell(t)\big) \right.
$$
$$
\left. + a_4 \frac{q_i(t) - q_\ell(t)}{|q_i(t) - q_\ell(t)|^{\beta_1+1}} \right\rangle
$$
$$
\leq \varepsilon(\beta_1 - 1) \sum_{i=1}^{N} \left\langle \sum_{j \neq i} \frac{1}{|q_i(t) - q_j(t)|^{\beta_1}}, \sum_{\ell \neq i} \left(a_5 \frac{1}{|q_i(t) - q_\ell(t)|^{\beta_2}} + a_6 \right) \right\rangle
$$
$$
\leq \varepsilon^{3/2} C \sum_{1 \leq i < j \leq N} \frac{1}{|q_i(t) - q_j(t)|^{2\beta_1}} + \varepsilon^{1/2} C \sum_{1 \leq i < j \leq N} \frac{1}{|q_i(t) - q_j(t)|^{2\beta_2}} + C,
$$

for some positive constant C independent of t and ε. Since $\beta_2 < \beta_1$, cf. (2.7), taking ε small enough produces the bound

$$
-\varepsilon(\beta_1 - 1) \sum_{i=1}^{N} \left\langle \sum_{j \neq i} \frac{q_i(t) - q_j(t)}{|q_i(t) - q_j(t)|^{\beta_1+1}}, -\sum_{\ell \neq i} \nabla G\big(q_i(t) - q_\ell(t)\big) \right\rangle
$$
$$
\leq -\varepsilon c \sum_{1 \leq i < j \leq N} \frac{1}{|q_i(t) - q_j(t)|^{2\beta_1}}. \tag{4.18}
$$

Similarly, regarding the last term on the right-hand side of (4.17), since in Case 1, $\beta_1 > 1$, it is clear that $\varepsilon|q_i - q_j|^{-\beta_1-1}$ can also be subsumed into $-\varepsilon|q_i - q_j|^{-2\beta_1}$ as in (4.18). Altogether with the expression (4.17), we arrive at the estimate

$$
d\left(\varepsilon \gamma \sum_{1 \leq i < j \leq N} \frac{1}{|q_i(t) - q_j(t)|^{\beta_1-1}} \right)
$$
$$
\leq \varepsilon^{1/2} C |q(t)|^2 dt + \varepsilon^{1/2} C |q(t)|^{\lambda-1} dt + \varepsilon^{1/2} C \sum_{i=1}^{N} |f_i(t)|^2 dt + C dt
$$
$$
- \varepsilon c \sum_{1 \leq i < j \leq N} \frac{1}{|q_i(t) - q_j(t)|^{2\beta_1}} dt
$$

$$- \varepsilon \sum_{i=1}^{N} \left\langle \sum_{j \neq i} \frac{q_i(t) - q_j(t)}{|q_i(t) - q_j(t)|^{\beta_1+1}}, \sqrt{2\gamma} dW_{i,0}(t) \right\rangle. \tag{4.19}$$

Next, from the expression (4.15) of Γ_1 and the estimates (4.16) and (4.19), we obtain (by taking ε small enough)

$$\begin{aligned}
d\Gamma_1(t) \leq &-c|q(t)|^2 dt - c|q(t)|^{\lambda+1} dt - c \sum_{i=1}^{N} |f_i|^2 \\
&- \varepsilon c \sum_{1 \leq i < j \leq N} \frac{1}{|q_i(t) - q_j(t)|^{2\beta_1}} dt + C dt \\
&+ \sum_{i=1}^{N} \sqrt{2\gamma} \langle q_i(t), dW_{i,0}(t) \rangle + \sum_{i=1}^{N} \sum_{\ell=1}^{k_i} \sqrt{2\alpha_{i,\ell}} \langle f_{i,\ell}(t), dW_{i,\ell}(t) \rangle \\
&- \varepsilon \sum_{i=1}^{N} \left\langle \sum_{j \neq i} \frac{q_i(t) - q_j(t)}{|q_i(t) - q_j(t)|^{\beta_1+1}}, \sqrt{2\gamma} dW_{i,0}(t) \right\rangle.
\end{aligned} \tag{4.20}$$

We emphasize that in (4.20), $c > 0$ is independent of ε, whereas $C > 0$ may still depend on ε.

Now, to establish (4.13), we aim to employ the well-known exponential Martingale inequality applied to (4.20). The argument below is similarly to that found in Hairer and Mattingly (2008, Lemma 5.1). See also Glatt-Holtz et al. (2021, 2022).

For $\kappa \in (0, 1)$ to be chosen later, from (4.20), we observe that

$$\begin{aligned}
\kappa d\Gamma_1(t) \leq &-c\kappa|q(t)|^2 dt - c\kappa|q(t)|^{\lambda+1} dt - c\kappa \sum_{i=1}^{N} |f_i|^2 + C\kappa dt \\
&- \kappa \varepsilon c \sum_{1 \leq i < j \leq N} \frac{1}{|q_i(t) - q_j(t)|^{2\beta_1}} dt + dM_1(t).
\end{aligned} \tag{4.21}$$

In the above, the semi-Martingale term $M_1(t)$ is defined as

$$\begin{aligned}
dM_1(t) = &\kappa \sum_{i=1}^{N} \sqrt{2\gamma} \langle q_i(t), dW_{i,0}(t) \rangle + \kappa \sum_{i=1}^{N} \sum_{\ell=1}^{k_i} \sqrt{2\alpha_{i,\ell}} \langle f_{i,\ell}(t), dW_{i,\ell}(t) \rangle \\
&- \kappa \varepsilon \sum_{i=1}^{N} \left\langle \sum_{j \neq i} \frac{q_i(t) - q_j(t)}{|q_i(t) - q_j(t)|^{\beta_1+1}}, \sqrt{2\gamma} dW_{i,0}(t) \right\rangle,
\end{aligned}$$

whose variation process $\langle M_1 \rangle(t)$ is given by

$$d\langle M_1 \rangle(t) = 2\gamma\kappa^2 \left| \sum_{i=1}^{N} \left(q_i(t) - \varepsilon \sum_{j \neq i} \frac{q_i(t) - q_j(t)}{|q_i(t) - q_j(t)|^{\beta_1+1}} \right) \right|^2 dt$$

$$+ \kappa^2 \sum_{i=1}^{N} \sum_{\ell=1}^{k_i} 2\alpha_{i,\ell} |f_{i,\ell}(t)|^2 \mathrm{d}t.$$

Using Cauchy–Schwarz inequality, it is clear that

$$\mathrm{d}\langle M_1 \rangle(t) \leq -\tilde{c}\kappa^2 |\mathrm{q}(t)|^2 \mathrm{d}t - \tilde{c}\kappa^2 \sum_{i=1}^{N} |f_i|^2 \mathrm{d}t - \kappa^2 \varepsilon \tilde{c} \sum_{1 \leq i < j \leq N} \frac{1}{|q_i(t) - q_j(t)|^{2\beta_1}} \mathrm{d}t,$$

for some positive constant $\tilde{c}$ independent of both κ and ε. It follows from (4.21) that

$$\kappa \mathrm{d}\Gamma_1(t) \leq C\kappa \mathrm{d}t - \frac{c}{\kappa} \langle M_1 \rangle(t) \mathrm{d}t + \mathrm{d}M_1(t).$$

Recalling the exponential Martingale inequality applying to $M_1(t)$,

$$\mathbb{P}\left(\sup_{t \geq 0} \left[M_1(t) - \frac{c}{\kappa} \langle M_1 \rangle(t) \right] > r \right) \leq e^{-\frac{2c}{\kappa} r}, \quad r \geq 0, \tag{4.22}$$

we deduce that

$$\mathbb{P}\left(\sup_{t \in [0,T]} \left[\kappa \Gamma_1(t) - \kappa \Gamma_1(0) - \kappa C t \right] > r \right) \leq e^{-\frac{2c}{\kappa} r}.$$

In particular, by choosing κ sufficiently small, the above inequality implies

$$\mathbb{E} \exp \left\{ \sup_{t \in [0,T]} \kappa \Gamma_1(t) \right\} \leq C, \tag{4.23}$$

for some positive constant $C = C(T, \kappa, \varepsilon, \mathrm{x}(0), z_1(0), \ldots, z_N(0))$. Recalling Γ_1 as in (4.15), the estimate (4.23) produces (4.13). Hence, part (a) is established for $\beta_1 > 1$.

(b) Considering $\beta_1 = 1$, in this case, we introduce the function Γ_2 defined as

$$\Gamma_2(\mathrm{q}, f_1, \ldots, f_N) = \frac{1}{2}\gamma|\mathrm{q}|^2 + \frac{1}{2} \sum_{i=1}^{N} \sum_{\ell=1}^{k_i} \frac{1}{\alpha_{i,\ell}} |f_{i,\ell}|^2 - \varepsilon \gamma \sum_{1 \leq i < j \leq N} \log |q_i - q_j|. \tag{4.24}$$

With regard to the log term on the above right-hand side, the following identity holds

$$\mathrm{d}\left(-\varepsilon \gamma \sum_{1 \leq i < j \leq N} \log |q_i(t) - q_j(t)| \right)$$

$$= -\varepsilon \sum_{i=1}^{N} \left\langle \sum_{j \neq i} \frac{q_i(t) - q_j(t)}{|q_i(t) - q_j(t)|^2}, -\nabla U(q_i(t)) \mathrm{d}t - \sum_{\ell \neq i} \nabla G\big(q_i(t) - q_\ell(t)\big) \mathrm{d}t \right.$$

$$
- \sum_{i=1}^{k_i} \lambda_{i,\ell}^2 q_i(t) \mathrm{d}t + \sum_{\ell=1}^{k_i} \lambda_{i,\ell} f_{i,\ell}(t) \mathrm{d}t + \sqrt{2\gamma} \mathrm{d}W_{i,0}(t) \Big\rangle
$$

$$
- \varepsilon (d-2) \sum_{1 \le i < j \le N} \frac{1}{|q_i(t) - q_j(t)|^2} \mathrm{d}t. \tag{4.25}
$$

Similarly to the estimates in Case 1, we readily have

$$
- \varepsilon \sum_{i=1}^{N} \Big\langle \sum_{j \ne i} \frac{q_i(t) - q_j(t)}{|q_i(t) - q_j(t)|^2}, -\nabla U(q_i(t)) - \sum_{i=1}^{k_i} \lambda_{i,\ell}^2 q_i(t) + \sum_{\ell=1}^{k_i} \lambda_{i,\ell} f_{i,\ell}(t) \Big\rangle
$$

$$
\le C \varepsilon^{1/2} |q(t)|^2 + C \varepsilon |q(t)|^{\lambda-1} + C \varepsilon^{1/2} \sum_{i=1}^{N} |f_i(t)|^2
$$

$$
+ C \varepsilon^{3/2} \sum_{1 \le i < j \le N} \frac{1}{|q_i(t) - q_j(t)|^2}. \tag{4.26}
$$

Concerning the cross terms involving G on the right-hand side of (4.25). We employ an argument from Case 1 while making use of condition (2.7) and the estimate (A.3) to arrive at

$$
- \varepsilon \sum_{i=1}^{N} \Big\langle \sum_{j \ne i} \frac{q_i(t) - q_j(t)}{|q_i(t) - q_j(t)|^2}, - \sum_{\ell \ne i} \nabla G\big(q_i(t) - q_\ell(t)\big) \Big\rangle
$$

$$
\le -2a_4 \varepsilon \sum_{1 \le i < j \le N} \frac{1}{|q_i(t) - q_j(t)|^2} + C \varepsilon^{3/2} \sum_{1 \le i < j \le N} \frac{1}{|q_i(t) - q_j(t)|^2} + \tilde{C}. \tag{4.27}
$$

In the above, we emphasize that C is independent of ε even though $\tilde{C}$ may still depend on ε. Turning to the last term on the right-hand side of (4.25), i.e.,

$$
-\varepsilon(d-2) \sum_{1 \le i < j \le N} \frac{1}{|q_i(t) - q_j(t)|^2} \mathrm{d}t,
$$

there are two cases to be considered, depending on the dimension d. In dimension $d \ge 2$, it is clear that the above expression is negative and thus is negligible. On the other hand, in dimension $d = 1$, it is reduced to

$$
\varepsilon \sum_{1 \le i < j \le N} \frac{1}{|q_i(t) - q_j(t)|^2} \mathrm{d}t.
$$

In view of Assumption 2.5, we combine with (4.27) to obtain

$$
- \varepsilon \sum_{i=1}^{N} \Big\langle \sum_{j \ne i} \frac{q_i(t) - q_j(t)}{|q_i(t) - q_j(t)|^2}, - \sum_{\ell \ne i} \nabla G\big(q_i(t) - q_\ell(t)\big) \Big\rangle + \varepsilon \sum_{1 \le i < j \le N} \frac{1}{|q_i(t) - q_j(t)|^2}
$$

$$\leq -(2a_4 - 1)\varepsilon \sum_{1 \leq i < j \leq N} \frac{1}{|q_i(t) - q_j(t)|^2} + C\varepsilon^{3/2} \sum_{1 \leq i < j \leq N} \frac{1}{|q_i(t) - q_j(t)|^2} + \tilde{C},$$

whence (by taking ε sufficiently small)

$$-\varepsilon \sum_{i=1}^{N} \left\langle \sum_{j \neq i} \frac{q_i(t) - q_j(t)}{|q_i(t) - q_j(t)|^2}, -\sum_{\ell \neq i} \nabla G\big(q_i(t) - q_\ell(t)\big) \right\rangle + \varepsilon \sum_{1 \leq i < j \leq N} \frac{1}{|q_i(t) - q_j(t)|^2}$$

$$\leq -c\varepsilon \sum_{1 \leq i < j \leq N} \frac{1}{|q_i(t) - q_j(t)|^2} + \tilde{C},$$

From (4.25) and (4.26), we find

$$\mathrm{d}\left(-\sum_{1 \leq i < j \leq N} \log |q_i(t) - q_j(t)| \right)$$

$$\leq -c\varepsilon \sum_{1 \leq i < j \leq N} \frac{1}{|q_i(t) - q_j(t)|^2} \mathrm{d}t + C\varepsilon^{1/2}|q(t)|^2 \mathrm{d}t + C\varepsilon|q(t)|^{\lambda-1}\mathrm{d}t$$

$$+ C\varepsilon^{1/2} \sum_{i=1}^{N} |f_i(t)|^2 \mathrm{d}t - \varepsilon \sum_{i=1}^{N} \left\langle \sum_{j \neq i} \frac{q_i(t) - q_j(t)}{|q_i(t) - q_j(t)|^2}, \sqrt{2\gamma} \mathrm{d}W_{i,0}(t) \right\rangle + \tilde{C}\mathrm{d}t. \tag{4.28}$$

Next, we combine estimates (4.16), (4.28) with expression (4.24) of Γ_2 to infer

$$\mathrm{d}\Gamma_2(t) \leq -c|q(t)|^2\mathrm{d}t - c|q(t)|^{\lambda+1}\mathrm{d}t - c\sum_{i=1}^{N}|f_i|^2$$

$$- \varepsilon c \sum_{1 \leq i < j \leq N} \frac{1}{|q_i(t) - q_j(t)|^2}\mathrm{d}t + C\mathrm{d}t$$

$$+ \sum_{i=1}^{N} \sqrt{2\gamma}\langle q_i(t), \mathrm{d}W_{i,0}(t)\rangle + \sum_{i=1}^{N}\sum_{\ell=1}^{k_i} \sqrt{2\alpha_{i,\ell}}\langle f_{i,\ell}(t), \mathrm{d}W_{i,\ell}(t)\rangle$$

$$- \varepsilon \sum_{i=1}^{N} \left\langle \sum_{j \neq i} \frac{q_i(t) - q_j(t)}{|q_i(t) - q_j(t)|^2}, \sqrt{2\gamma}\mathrm{d}W_{i,0}(t) \right\rangle. \tag{4.29}$$

We may now employ an argument similarly to the exponential Martingale approach as in Case 1 to deduce the bound for all κ sufficiently small

$$\mathbb{E} \exp \left\{ \sup_{t \in [0,T]} \kappa \Gamma_2(t) \right\} \leq C,$$

for some positive constant $C = C(T, \kappa, \varepsilon, \mathrm{x}(0), \mathrm{z}_1(0), \ldots, \mathrm{z}_N(0))$. Recalling Γ_2 defined in (4.24), this produces the estimate (4.14). The proof is thus finished. $\qquad \square$

4.4 Proof of Theorem 2.10

We are now in a position to conclude Theorem 2.10. The argument follows along the lines of the proof of Nguyen (2018, Theorem 4) adapted to our setting. See also Herzog et al. (2016, Theorem 2.4). The key observation is that instead of controlling the exiting time of the process $x_m(t)$ as $m \to 0$, we are able to control $q(t)$ since $q(t)$ is independent of m.

Proof of Theorem 2.10 Let $\big(x_m(t), v_m(t), z_{1,m}(t), \ldots, z_{N,m}(t)\big)$ and $\big(q(t), f_1(t), \ldots, f_N(t)\big)$, respectively, solve (1.5) and (1.10). For $R, m > 0$, define the following stopping times

$$\sigma^R = \inf_{t \geq 0} \left\{ |q(t)| + \sum_{1 \leq i < j \leq N} |q_i(t) - q_j(t)|^{-1} \geq R \right\}, \qquad (4.30)$$

and

$$\sigma_m^R = \inf_{t \geq 0} \left\{ |x_m(t)| + \sum_{1 \leq i < j \leq N} |x_{i,m}(t) - x_{j,m}(t)|^{-1} \geq R \right\}.$$

Fixing $T, \xi > 0$, observe that

$$\mathbb{P}\left(\sup_{t \in [0,T]} |x_m(t) - q(t)| > \xi \right) \leq \mathbb{P}\left(\sup_{t \in [0,T]} |x_m(t) - q(t)| > \xi, \sigma^R \wedge \sigma_m^R \geq T \right)$$
$$+ \mathbb{P}\big(\sigma^R \wedge \sigma_m^R < T \big). \qquad (4.31)$$

To control the first term on the above right-hand side, observe that

$$\mathbb{P}\big(0 \leq t \leq \sigma^R \wedge \sigma_m^R, \, q(t) = q^R(t), \, x_m(t) = x_m^R(t) \big) = 1,$$

where $q^R(t)$ and $x_m^R(t)$ are the first components of the solutions of (4.5) and (4.4), respectively. As a consequence,

$$\mathbb{P}\left(\sup_{0 \leq t \leq T} |x_m(t) - q(t)| > \xi, \sigma^R \wedge \sigma_m^R \geq T \right)$$
$$\leq \mathbb{P}\left(\sup_{0 \leq t \leq T} |x_m^R(t) - q^R(t)| > \xi \right) \leq \frac{m}{\xi^4} \cdot C(T, R). \qquad (4.32)$$

In the last estimate above, we employed Proposition 4.1 while making use of Markov's inequality.

Turning to $\mathbb{P}(\sigma^R \wedge \sigma_m^R < T)$, we note that

$$\mathbb{P}\big(\sigma^R \wedge \sigma_m^R < T \big)$$

$$\leq \mathbb{P}\left(\sup_{t \in [0,T]} |\mathrm{x}_m^R(t) - \mathrm{q}^R(t)| \leq \frac{\xi}{R}, \sigma^R \wedge \sigma_m^R < T \right)$$

$$+ \mathbb{P}\left(\sup_{t \in [0,T]} |\mathrm{x}_m^R(t) - \mathrm{q}^R(t)| > \frac{\xi}{R} \right)$$

$$\leq \mathbb{P}\left(\sup_{t \in [0,T]} |\mathrm{x}_m^R(t) - \mathrm{q}^R(t)| \leq \frac{\xi}{R}, \sigma_m^R < T \leq \sigma^R \right) + \mathbb{P}(\sigma^R < T)$$

$$+ \mathbb{P}\left(\sup_{t \in [0,T]} |\mathrm{x}_m^R(t) - \mathrm{q}^R(t)| > \frac{\xi}{R} \right)$$

$$= I_1 + I_2 + I_3. \tag{4.33}$$

Concerning I_3, the same argument as in (4.32) produces the bound

$$I_3 = \mathbb{P}\left(\sup_{0 \leq t \leq T} |\mathrm{x}_m^R(t) - \mathrm{q}^R(t)| > \frac{\xi}{R} \right) \leq \frac{m}{\xi^4} \cdot C(T, R). \tag{4.34}$$

Next, considering I_2, from (4.30), observe that for all ε small and R large enough,

$$\{\sigma^R < T\} = \left\{ \sup_{t \in [0,T]} |\mathrm{q}(t)| + \sum_{1 \leq i < j \leq N} |q_i(t) - q_j(t)|^{-1} \geq R \right\}$$

$$\subseteq \left\{ \sup_{t \in [0,T]} |\mathrm{q}(t)| \geq \frac{R}{N^2} \right\} \bigcup_{1 \leq i < j \leq N} \left\{ -\varepsilon \log |q_i(t) - q_j(t)| \geq \varepsilon \log\left(\frac{R}{N^2} \right) \right\}$$

$$\subseteq \left\{ \sup_{t \in [0,T]} |\mathrm{q}(t)|^2 - \varepsilon \sum_{1 \leq i < j \leq N} \log |q_i(t) - q_j(t)| \geq \frac{R}{N^2} \right\}$$

$$\bigcup_{1 \leq i < j \leq N} \left\{ \sup_{t \in [0,T]} |\mathrm{q}(t)|^2 - \varepsilon \sum_{1 \leq i < j \leq N} \log |q_i(t) - q_j(t)| \geq \varepsilon \log\left(\frac{R}{N^2} \right) \right\},$$

whence,

$$\{\sigma^R < T\} \subseteq \left\{ \sup_{t \in [0,T]} |\mathrm{q}(t)|^2 - \varepsilon \sum_{1 \leq i < j \leq N} \log |q_i(t) - q_j(t)| \geq \varepsilon \log\left(\frac{R}{N^2} \right) \right\}.$$

We note that Proposition 4.1 implies the estimate

$$\mathbb{E}\left[\sup_{t \in [0,T]} \left(|\mathrm{q}(t)|^2 - \varepsilon \sum_{1 \leq i < j \leq N} \log |q_i(t) - q_j(t)| \right) \right] \leq C, \tag{4.35}$$

for some positive constant $C = C(T, \varepsilon)$. Together with Markov's inequality, we infer the bound for R large enough

$$I_2 = \mathbb{P}(\sigma^R < T) \leq \frac{C(T)}{\varepsilon \log(R/N^2)} \leq \frac{C(T)}{\varepsilon \log R}. \tag{4.36}$$

Turning to I_1 on the right-hand side of (4.33), for R large enough and $\xi \in (0, 1)$, a chain of event implications is derived as follows:

$$\left\{ \sup_{t \in [0,T]} |\mathrm{x}_m^R(t) - \mathrm{q}^R(t)| \le \frac{\xi}{R}, \sigma_m^R < T \le \sigma^R \right\}$$

$$= \left\{ \sup_{t \in [0,T]} |\mathrm{x}_m^R(t) - \mathrm{q}(t)| \le \frac{\xi}{R}, \sup_{t \in [0,T]} \left(|\mathrm{x}_m^R(t)| + \sum_{1 \le i < j \le N} |x_{i,m}^R(t) - x_{j,m}^R(t)|^{-1} \right) \right.$$

$$\left. \ge R, \sigma_m^R < T \le \sigma^R \right\}$$

$$\subseteq \left\{ \sup_{t \in [0,T]} |\mathrm{x}_m^R(t) - \mathrm{q}(t)| \le \frac{\xi}{R}, \sup_{t \in [0,T]} |\mathrm{x}_m^R(t)| \ge \frac{R}{N^2} \right\}$$

$$\bigcup_{1 \le i < j \le N} \left\{ \sup_{t \in [0,T]} |\mathrm{x}_m^R(t) - \mathrm{q}(t)| \le \frac{\xi}{R}, \sup_{t \in [0,T]} |x_{i,m}^R(t) - x_{j,m}^R(t)|^{-1} \ge \frac{R}{N^2} \right\}$$

$$= \left\{ \sup_{t \in [0,T]} |\mathrm{x}_m^R(t) - \mathrm{q}(t)| \le \frac{\xi}{R}, \sup_{t \in [0,T]} |\mathrm{x}_m^R(t)| \ge \frac{R}{N^2} \right\}$$

$$\bigcup_{1 \le i < j \le N} \left\{ \sup_{t \in [0,T]} |\mathrm{x}_m^R(t) - \mathrm{q}(t)| \le \frac{\xi}{R}, \inf_{t \in [0,T]} |x_{i,m}^R(t) - x_{j,m}^R(t)| \le \frac{N^2}{R} \right\}.$$

Since ξ and N are fixed, for R large enough, say, $\frac{R}{N^2} - \frac{\xi}{R} \ge \sqrt{R}$, we have

$$\left\{ \sup_{t \in [0,T]} |\mathrm{x}_m^R(t) - \mathrm{q}(t)| \le \frac{\xi}{R}, \sup_{t \in [0,T]} |\mathrm{x}_m^R(t)| \ge \frac{R}{N^2} \right\}$$

$$\subseteq \left\{ \sup_{t \in [0,T]} |\mathrm{q}(t)| \ge \frac{R}{N^2} - \frac{\xi}{R} \right\} \subseteq \left\{ \sup_{t \in [0,T]} |\mathrm{q}(t)| \ge \sqrt{R} \right\}$$

$$\subseteq \left\{ \sup_{t \in [0,T]} |\mathrm{q}(t)|^2 - \varepsilon \sum_{1 \le i < j \le N} \log |q_i(t) - q_j(t)| \ge \sqrt{R} \right\}.$$

On the other hand, by triangle inequality,

$$\inf_{t \in [0,T]} |q_i(t) - q_j(t)| \le 2 \sup_{t \in [0,T]} |\mathrm{x}_m^R(t) - \mathrm{q}(t)| + \inf_{t \in [0,T]} |x_{i,m}^R(t) - x_{j,m}^R(t)|,$$

implying

$$\left\{ \sup_{t \in [0,T]} |\mathrm{x}_m^R(t) - \mathrm{q}(t)| \le \frac{\xi}{R}, \inf_{t \in [0,T]} |x_{i,m}^R(t) - x_{j,m}^R(t)| \le \frac{N^2}{R} \right\}$$

$$\subseteq \left\{ \inf_{t \in [0,T]} |q_i(t) - q_j(t)| \le \frac{2\xi + N^2}{R} \right\} \subseteq \left\{ \inf_{t \in [0,T]} |q_i(t) - q_j(t)| \le \frac{1}{\sqrt{R}} \right\}$$

$$= \left\{ -\varepsilon \sup_{t\in[0,T]} \log|q_i(t) - q_j(t)| \geq \frac{1}{2}\varepsilon \log R \right\}.$$

It follows that for ε small and R large enough

$$\left\{ \sup_{t\in[0,T]} |\mathrm{x}_m^R(t) - \mathrm{q}(t)| \leq \frac{\xi}{R}, \inf_{t\in[0,T]} |x_{i,m}^R(t) - x_{j,m}^R(t)| \leq \frac{N^2}{R} \right\}$$
$$\subseteq \left\{ \sup_{t\in[0,T]} |\mathrm{q}(t)|^2 - \varepsilon \sum_{1\leq i<j\leq N} \log|q_i(t) - q_j(t)| \geq \frac{1}{2}\varepsilon \log R \right\}.$$

As a consequence, the following holds

$$\left\{ \sup_{t\in[0,T]} |\mathrm{x}_m^R(t) - \mathrm{q}^R(t)| \leq \frac{\xi}{R}, \sigma_m^R < T \leq \sigma^R \right\}$$
$$\subseteq \left\{ \sup_{t\in[0,T]} |\mathrm{q}(t)|^2 - \varepsilon \sum_{1\leq i<j\leq N} \log|q_i(t) - q_j(t)| \geq \frac{1}{2}\varepsilon \log R \right\}.$$

We employ (4.35) and Markov's inequality to infer

$$I_1 = \mathbb{P}\left(\sup_{t\in[0,T]} |\mathrm{x}_m^R(t) - \mathrm{q}^R(t)| \leq \frac{\xi}{R}, \sigma_m^R < T \leq \sigma^R \right) \leq \frac{C(T)}{\varepsilon \log R}. \qquad (4.37)$$

Turning back to (4.33), we collect (4.34), (4.36) and (4.37) to arrive at the bound

$$\mathbb{P}\left(\sigma^R \wedge \sigma_m^R < T \right) \leq \frac{m}{\xi^4} \cdot C(T, R) + \frac{C(T)}{\varepsilon \log R}. \qquad (4.38)$$

We emphasize that in the above estimate $C(T, R)$ and $C(R)$ are independent of m.

Now, putting everything together, from (4.31), (4.32) and (4.38), we obtain the estimate

$$\mathbb{P}\left(\sup_{t\in[0,T]} |\mathrm{x}_m(t) - \mathrm{q}(t)| > \xi \right) \leq \frac{m}{\xi^4} \cdot C(T, R) + \frac{C(T)}{\varepsilon \log R},$$

for all ε small and R large enough. By sending R to infinity and then shrinking m further to zero, this produces the small-mass limit (2.18), thereby completing the proof. $\square$

Remark 4.4 For the underdamped Langevin dynamics, it is well known that the small-mass limit $m \to 0$ and the high-friction limit $\gamma \to +\infty$ (under an appropriate time (or noise) rescaling) both lead to the same limiting system, which is the overdamped Langevin dynamics, see, e.g., (Lelièvre et al. 2010, Section 2.2.4) and also (Duong et al. 2017) for the Vlasov–Fokker–Planck system. Similarly, one can also derive the underdamped Langevin dynamics from the generalized Langevin dynamics (GLE)

under the white-noise limit, by rescaling the friction coefficients (λ_i) and the strength of the noises $(\alpha)_i$ appropriately, see, e.g., (Ottobre and Pavliotis 2011; Nguyen 2018) for a rigorous analysis for the non-interacting GLE and (Duong and Pavliotis 2019) for a formal derivation for the interacting-GLE with regular interactions. This white-noise limit is different from the small-mass limit studied in this paper. It would be interesting to study the white-noise limit for the generalized Langevin dynamics with irregular interactions, which we leave for future work.

Appendix A. Auxiliary Estimates on Singular Potentials

In this section, we collect some useful estimates on singular potentials, cf. Lemma A.1 and Lemma A.2. In particular, Lemma A.1 was employed to construct Lyapunov functions in Sect. 3. Lemma A.2 appears in the proof of Lemma 4.3 which was invoked to prove the small-mass result of Theorem 2.10.

Lemma A.1 *For all $s \geq 0$ and* $\mathrm{x} = (x_1, \ldots, x_N) \in \mathcal{D}$*, the following holds:*

$$\sum_{i=1}^{N} \left\langle \sum_{j \neq i} \frac{x_i - x_j}{|x_i - x_j|^{s+1}}, \sum_{\ell \neq i} \frac{x_i - x_\ell}{|x_i - x_\ell|} \right\rangle \geq 2 \sum_{1 \leq i < j \leq N} \frac{1}{|x_i - x_j|^s}. \tag{A.1}$$

Proof The proof is the same as that of (Lu and Mattingly 2019, Lemma 4.2). See also (Bolley et al. 2018, Inequality (5.1)). □

Lemma A.2 *For all* $\mathrm{x} = (x_1, \ldots, x_N) \in \mathcal{D}$*, the followings hold:*
(a) *For all $s \geq 0$,*

$$\sum_{i=1}^{N} \left\langle \sum_{j \neq i} \frac{x_i - x_j}{|x_i - x_j|^{s+1}}, \sum_{\ell \neq i} \frac{x_i - x_\ell}{|x_i - x_\ell|^{s+1}} \right\rangle \geq \frac{4}{N(N-1)^2} \sum_{1 \leq i < j \leq N} \frac{1}{|x_i - x_j|^{2s}}. \tag{A.2}$$

(b) *Furthermore, for $s \in [0, 1]$,*

$$\sum_{i=1}^{N} \left\langle \sum_{j \neq i} \frac{x_i - x_j}{|x_i - x_j|^{s+1}}, \sum_{\ell \neq i} \frac{x_i - x_\ell}{|x_i - x_\ell|^{s+1}} \right\rangle \geq 2 \sum_{1 \leq i < j \leq N} \frac{1}{|x_i - x_j|^{2s}}. \tag{A.3}$$

Proof (a) With regard to (A.2), we employ Cauchy-Schwarz inequality to infer

$$\left| \sum_{i=1}^{N} \left\langle \sum_{j \neq i} \frac{x_i - x_j}{|x_i - x_j|^{s+1}}, \sum_{\ell \neq i} \frac{x_i - x_\ell}{|x_i - x_\ell|} \right\rangle \right|^2$$

$$\leq \sum_{i=1}^{N} \left| \sum_{j \neq i} \frac{x_i - x_j}{|x_i - x_j|^{s+1}} \right|^2 \cdot \sum_{i=1}^{N} \left| \sum_{\ell \neq i} \frac{x_i - x_\ell}{|x_i - x_\ell|} \right|^2$$

$$\leq \sum_{i=1}^{N} \left| \sum_{j\neq i} \frac{x_i - x_j}{|x_i - x_j|^{s+1}} \right|^2 \cdot N(N-1)^2.$$

In view of Lemma A.1, the following holds

$$\left| \sum_{i=1}^{N} \left\langle \sum_{j\neq i} \frac{x_i - x_j}{|x_i - x_j|^{s+1}}, \sum_{\ell\neq i} \frac{x_i - x_\ell}{|x_i - x_\ell|} \right\rangle \right|^2 \geq 4 \left(\sum_{1\leq i<j\leq N} \frac{1}{|x_i - x_j|^s} \right)^2.$$

Therefore, we obtain

$$\sum_{i=1}^{N} \left| \sum_{j\neq i} \frac{x_i - x_j}{|x_i - x_j|^{s+1}} \right|^2 \geq \frac{4}{N(N-1)^2} \left(\sum_{1\leq i<j\leq N} \frac{1}{|x_i - x_j|^s} \right)^2.$$

This produces (A.2), thus establishing part (a).

(b) Turning to (A.3), we will follow the proof of (Lu and Mattingly 2019, Lemma 4.2) adapted to our setting. A routine computation gives

$$\sum_{i=1}^{N} \left\langle \sum_{j\neq i} \frac{x_i - x_j}{|x_i - x_j|^{s+1}}, \sum_{\ell\neq i} \frac{x_i - x_\ell}{|x_i - x_\ell|^{s+1}} \right\rangle$$

$$= 2 \sum_{1\leq i<j\leq N} \frac{1}{|x_i - x_j|^{2s}} + 2 \sum_{1\leq i<j<\ell\leq N} \left[\frac{\langle x_i - x_j, x_i - x_\ell \rangle}{|x_i - x_j|^{s+1} |x_i - x_\ell|^{s+1}} \right.$$

$$\left. + \frac{\langle x_j - x_i, x_j - x_\ell \rangle}{|x_j - x_i|^{s+1} |x_j - x_\ell|^{s+1}} + \frac{\langle x_\ell - x_i, x_\ell - x_j \rangle}{|x_\ell - x_i|^{s+1} |x_\ell - x_j|^{s+1}} \right].$$

It suffices to prove that the second sum on the above right-hand side is nonnegative. To this end, we claim that for $i < j < \ell$,

$$\frac{\langle x_i - x_j, x_i - x_\ell \rangle}{|x_i - x_j|^{s+1} |x_i - x_\ell|^{s+1}} + \frac{\langle x_j - x_i, x_j - x_\ell \rangle}{|x_j - x_i|^{s+1} |x_j - x_\ell|^{s+1}} + \frac{\langle x_\ell - x_i, x_\ell - x_j \rangle}{|x_\ell - x_i|^{s+1} |x_\ell - x_j|^{s+1}} \geq 0.$$

Denote $\theta_i, \theta_j, \theta_\ell$ to be the angles formed by these points and whose vertices are respectively x_i, x_j, x_ℓ. Observe that the above inequality is equivalent to

$$\cos(\theta_i)|x_j - x_\ell|^s + \cos(\theta_j)|x_i - x_\ell|^s + \cos(\theta_\ell)|x_i - x_j|^s \geq 0. \qquad \text{(A.4)}$$

Now, there are two cases to be considered depending on the triangle formed by x_i, x_j, x_ℓ in $\mathbb{R}^d$.

Case 1: the triangle is an acute or right triangle. In this case, it is clear that $\cos(\theta_i)$, $\cos(\theta_j)$ and $\cos(\theta_\ell)$ are both non negative. This immediately produces (A.4).

Case 2: The triangle is an obtuse triangle. Without loss of generality, suppose that the angle $\theta_j \in (\pi/2, \pi]$, and thus $\theta_i + \theta_\ell \in [0, \pi/2)$. Observe that

$$\min\{\cos(\theta_i), \cos(\theta_\ell)\} \geq \cos(\theta_i + \theta_\ell) = -\cos(\theta_j) \geq 0.$$

As a consequence, since $s \in [0, 1]$,

$$\cos(\theta_i)|x_j - x_\ell|^s + \cos(\theta_j)|x_i - x_\ell|^s + \cos(\theta_\ell)|x_i - x_j|^s$$
$$\geq |\cos(\theta_j)|\left(|x_j - x_\ell|^s - |x_i - x_\ell|^s + |x_i - x_j|^s\right) \geq 0.$$

This establishes (A.4), thereby finishing the proof of part (b).   □

Author Contributions M.H.D and H.N wrote the main manuscript text. All authors reviewed the manuscript.

Declarations

Conflict of interest The authors have no conflicts of interest to declare that are relevant to the content of this article.

References

Athreya, A., Kolba, T., Mattingly, J.C.: Propagating Lyapunov functions to prove noise-induced stabilization. Electron. J. Probab. **17**(96), 1–38 (2012)

Baczewski, A.D., Bond, S.D.: Numerical integration of the extended variable generalized Langevin equation with a positive Prony representable memory kernel. J. Chem. Phys. **139**(4), 044107 (2013)

Baeumer, B., Geissert, M., Kovács, M.: Existence, uniqueness and regularity for a class of semilinear stochastic Volterra equations with multiplicative noise. J. Differ. Equ. **258**(2), 535–554 (2015)

Bockius, N., Shea, J., Jung, G., Schmid, F., Hanke, M.: Model reduction techniques for the computation of extended Markov parameterizations for generalized Langevin equations. J. Phys. Condens. Matter **33**(21), 214003 (2021)

Bolley, F., Chafaï, D., Fontbona, J.: Dynamics of a planar Coulomb gas. Ann. Appl. Probab. **28**(5), 3152–3183 (2018)

Cerrai, S., Freidlin, M.: On the Smoluchowski-Kramers approximation for a system with an infinite number of degrees of freedom. Probab. Theory Relat. Fields **135**(3), 363–394 (2006)

Cerrai, S., Freidlin, M.: Smoluchowski-Kramers approximation for a general class of SPDEs. J. Evol. Equ. **6**(4), 657–689 (2006)

Cerrai, S., Freidlin, M., Salins, M.: On the Smoluchowski-Kramers approximation for SPDEs and its interplay with large deviations and long time behavior. Discrete Contin. Dyn. Syst. A **37**(1), 33 (2017)

Cerrai, S., Glatt-Holtz, N.: On the convergence of stationary solutions in the Smoluchowski-Kramers approximation of infinite dimensional systems. J. Funct. Anal. **278**(8), 108421 (2020)

Cerrai, S., Salins, M.: Smoluchowski-Kramers approximation and large deviations for infinite dimensional gradient systems. Asymptot. Anal. **88**(4), 201–215 (2014)

Cerrai, S., Salins, M.: Smoluchowski-Kramers approximation and large deviations for infinite-dimensional nongradient systems with applications to the exit problem. Ann. Probab. **44**(4), 2591–2642 (2016)

Cerrai, S., Wehr, J., Zhu, Y.: An averaging approach to the Smoluchowski-Kramers approximation in the presence of a varying magnetic field. J. Stat. Phys. **181**(1), 132–148 (2020)

Choi, Y.-P., Tse, O.: Quantified overdamped limit for kinetic Vlasov-Fokker-Planck equations with singular interaction forces. J. Diff. Equ. **330**, 150–207 (2022)

Conrad, F., Grothaus, M.: Construction, ergodicity and rate of convergence of N-particle Langevin dynamics with singular potentials. J. Evol. Equ. **10**(3), 623–662 (2010)

Cooke, B., Herzog, D.P., Mattingly, J.C., McKinley, S.A., Schmidler, S.C.: Geometric ergodicity of two-dimensional Hamiltonian systems with a Lennard-Jones-like repulsive potential. Commun. Math. Sci. **15**(7), 1987–2025 (2017)

Desch, W., Londen, S.-O.: An L p-theory for stochastic integral equations. J. Evol. Equ. **11**(2), 287–317 (2011)

Doob, J.L.: The Brownian movement and stochastic equations. Ann. Math. **43**, 351–369 (1942)

Duong, M.H.: Long time behaviour and particle approximation of a generalised Vlasov dynamic. Nonlinear Anal. Theory Methods Appl. **127**, 1–16 (2015)

Duong, M.H., Lamacz, A., Peletier, M.A., Sharma, U.: Variational approach to coarse-graining of generalized gradient flows. Calc. Var. Partial. Differ. Equ. **56**, 1–65 (2017)

Duong, M.H., Pavliotis, G.A.: Mean field limits for non-Markovian interacting particles: convergence to equilibrium, GENERIC formalism, asymptotic limits and phase transitions. Commun. Math. Sci. **16**(8), 2199–2230 (2019)

Duong, M.H., Shang, X.: Accurate and robust splitting methods for the generalized Langevin equation with a positive Prony series memory kernel. J. Comput. Phys. **464**, 111332 (2022)

Liu, D.: Gibbsian dynamics and invariant measures for stochastic dissipative PDEs. J. Stat. Phys. **108**(5–6), 1125–1156 (2002)

Mattingly, J.C., Sinai, Y.: Gibbsian dynamics and ergodicity for the stochastically forced Navier-Stokes equation. Commun. Math. Phys. **224**(1), 83–106 (2001)

Freidlin, M.: Some remarks on the Smoluchowski-Kramers approximation. J. Stat. Phys. **117**(3–4), 617–634 (2004)

Glatt-Holtz, N., Martinez, V.R., Richards, G.H.: On the long-time statistical behavior of smooth solutions of the weakly damped, stochastically-driven KdV equation. arXiv preprint arXiv:2103.12942 (2021)

Glatt-Holtz, N.E., Herzog, D.P., McKinley, S.A., Nguyen, H.D.: The generalized Langevin equation with power-law memory in a nonlinear potential well. Nonlinearity **33**(6), 2820 (2020)

Glatt-Holtz, N.E., Martinez, V.R., Nguyen, H.D.: The short memory limit for long time statistics in a stochastic Coleman-Gurtin model of heat conduction. arXiv preprint arXiv:2212.05646 (2022)

Gomes, S.N., Pavliotis, G.A., Vaes, U.: Mean field limits for interacting diffusions with colored noise: phase transitions and spectral numerical methods. Multiscale Model. Simul. **18**(3), 1343–1370 (2020)

Gottwald, F., Karsten, S., Ivanov, S.D., Kühn, O.: Parametrizing linear generalized Langevin dynamics from explicit molecular dynamics simulations. J. Chem. Phys. **142**(24), 244110 (2015)

Goychuk, I.: Viscoelastic subdiffusion: generalized Langevin equation approach. Adv. Chem. Phys. **150**, 187 (2012)

Grothaus, M., Stilgenbauer, P.: A hypocoercivity related ergodicity method with rate of convergence for singularly distorted degenerate Kolmogorov equations and applications. Integr. Equ. Oper. Theory **83**(3), 331–379 (2015)

Hairer, M., Mattingly, J.C.: Spectral gaps in Wasserstein distances and the 2D stochastic Navier-Stokes equations. Ann. Prob. **36**(6), 2050–2091 (2008)

Hairer, M., Mattingly, J.C.: Yet another look at Harris' ergodic theorem for Markov chains. In: Seminar on Stochastic Analysis, Random Fields and Applications VI, pp. 109–117 (2011)

Herzog, D.P., Hottovy, S., Volpe, G.: The small-mass limit for Langevin dynamics with unbounded coefficients and positive friction. J. Stat. Phys. **163**(3), 659–673 (2016)

Herzog, D.P., Mattingly, J.C.: Noise-induced stabilization of planar flows I. Electron. J. Probab. **20**(111), 1–43 (2015)

Herzog, D.P., Mattingly, J.C.: Noise-induced stabilization of planar flows II. Electron. J. Probab. **20**(113), 1–37 (2015)

Herzog, D.P., Mattingly, J.C.: Ergodicity and Lyapunov functions for Langevin dynamics with singular potentials. Commun. Pure Appl. Math. **72**(10), 2231–2255 (2019)

Herzog, D.P., Mattingly, J.C., Nguyen, H.D.: Gibbsian dynamics and the generalized Langevin equation. Electron. J. Probab. **28**, 1–29 (2023)

Higham, D.J., Mao, X., Stuart, A.M.: Strong convergence of Euler-type methods for nonlinear stochastic differential equations. SIAM J. Numer. Anal. **40**(3), 1041–1063 (2002)

Hörmander, L.: Hypoelliptic second order differential equations. Acta Math. **119**(1), 147–171 (1967)

Hottovy, S., McDaniel, A., Volpe, G., Wehr, J.: The Smoluchowski-Kramers limit of stochastic differential equations with arbitrary state-dependent friction. Commun. Math. Phys. **336**(3), 1259–1283 (2015)

Hottovy, S., Volpe, G., Wehr, J.: Noise-induced drift in stochastic differential equations with arbitrary friction and diffusion in the Smoluchowski-Kramers limit. J. Stat. Phys. **146**, 762–773 (2012)

Hottovy, S., Volpe, G., Wehr, J.: Thermophoresis of Brownian particles driven by coloured noise. Europhys. Lett. **99**(6), 60002 (2012)

Jung, G., Hanke, M., Schmid, F.: Generalized Langevin dynamics: construction and numerical integration of non-Markovian particle-based models. Soft Matter **14**(46), 9368–9382 (2018)

Karatzas, I., Shreve, S.: Brownian Motion and Stochastic Calculus, vol. 113. Springer Science & Business Media, Berlin (2012)

Khasminskii, R.: Stochastic Stability of Differential Equations, vol. 66. Springer Science & Business Media, Berlin (2011)

Kneller, G.: Generalized Kubo relations and conditions for anomalous diffusion: physical insights from a mathematical theorem. J. Chem. Phys. **134**(22), 224106 (2011)

Kramers, H.A.: Brownian motion in a field of force and the diffusion model of chemical reactions. Physica **7**(4), 284–304 (1940)

Kubo, R.: The fluctuation-dissipation theorem. Rep. Prog. Phys. **29**(1), 255 (1966)

Lelièvre, T., Rousset, M., Stoltz, G.: Free Energy Computations: A Mathematical Perspective. World Scientific, Singapore (2010)

Lim, S.H., Wehr, J.: Homogenization for a class of generalized Langevin equations with an application to thermophoresis. J. Stat. Phys. **174**, 656–691 (2019)

Lim, S.H., Wehr, J., Lewenstein, M.: Homogenization for generalized Langevin equations with applications to anomalous diffusion. In: Annales Henri Poincaé, pp. 1–59 (2020)

Lu, Y., Mattingly, J.C.: Geometric ergodicity of Langevin dynamics with Coulomb interactions. Nonlinearity **33**(2), 675 (2019)

Mason, T.G., Weitz, D.A.: Optical measurements of frequency-dependent linear viscoelastic moduli of complex fluids. Phys. Rev. Lett. **74**(7), 1250 (1995)

Mattingly, J.C.: Exponential convergence for the stochastically forced Navier-Stokes equations and other partially dissipative dynamics. Commun. Math. Phys. **230**(3), 421–462 (2002)

Mattingly, J.C., Stuart, A.M., Higham, D.J.: Ergodicity for SDEs and approximations: locally Lipschitz vector fields and degenerate noise. Stoch. Process. Their Appl. **101**(2), 185–232 (2002)

Meyn, S.P., Tweedie, R.L.: Markov Chains and Stochastic Stability. Springer Science & Business Media, Berlin (2012)

Mori, H.: A continued-fraction representation of the time-correlation functions. Prog. Theor. Phys. **34**(3), 399–416 (1965)

Ness, H., Stella, L., Lorenz, C., Kantorovich, L.: Applications of the generalized Langevin equation: towards a realistic description of the baths. Phys. Rev. B **91**(1), 014301 (2015)

Nguyen, H.D.: The small-mass limit and white-noise limit of an infinite dimensional generalized Langevin equation. J. Stat. Phys. **173**(2), 411–437 (2018)

Nguyen, H.D.: The small mass limit for long time statistics of a stochastic nonlinear damped wave equation. J. Differ. Equ. **371**, 481–548 (2023)

Ottobre, M., Pavliotis, G.A.: Asymptotic analysis for the generalized Langevin equation. Nonlinearity **24**(5), 1629 (2011)

Pardoux, E., Veretennikov, A.Y.: On Poisson equation and diffusion approximation 2. Ann. Probab. **31**(3), 1166–1192 (2003)

Pavliotis, G.A.: Stochastic Processes and Applications: Diffusion Processes, the Fokker-Planck and Langevin Equations, vol. 60. Springer, Berlin (2014)

Pavliotis, G.A., Stoltz, G., Vaes, U.: Scaling limits for the generalized Langevin equation. J. Nonlinear Sci. **31**(1), 8 (2021)

Pavliotis, G.A., Stuart, A.M., Vaes, U.: Derivative-free Bayesian inversion using multiscale dynamics. SIAM J. Appl. Dyn. Syst. **21**(1), 284–326 (2022)

Shi, C., Wang, W.: Small mass limit and diffusion approximation for a generalized Langevin equation with infinite number degrees of freedom. J. Differ. Equ. **286**, 645–675 (2021)

Smoluchowski, M.V.: Drei vortrage uber diffusion, brownsche bewegung und koagulation von kolloidteilchen. Z. Phys. **17**, 557–585 (1916)

Villani, C.: Topics in Optimal Transportation, vol. 58. American Mathematical Soc., Providence (2021)

Wei, X., Zhang, T., Luo, T.: Chain conformation-dependent thermal conductivity of amorphous polymer blends: the impact of inter-and intra-chain interactions. Phys. Chem. Chem. Phys. **18**(47), 32146–32154 (2016)

Zwanzig, R.: Nonequilibrium Statistical Mechanics. Oxford University Press, Oxford (2001)