Scaling limits for the generalized Langevin equation

G.A. Pavliotis* G. Stoltz† U. Vaes‡

January 4, 2021

Abstract

In this paper, we study the diffusive limit of solutions to the *generalized Langevin equation* (GLE) in a periodic potential. Under the assumption of quasi-Markovianity, we obtain sharp longtime equilibration estimates for the GLE using techniques from the theory of hypocoercivity. We then show asymptotic results for the effective diffusion coefficient in the small correlation time regime, as well as in the overdamped and underdamped limits. Finally, we employ a recently developed numerical method [66] to calculate the effective diffusion coefficient for a wide range of (effective) friction coefficients, confirming our asymptotic results.

Keywords: Generalized Langevin equation, Quasi-Markovian models, Longtime behavior, Hypocoercivity, Effective diffusion coefficient, Overdamped and underdamped limits, Fourier/Hermite spectral methods.

AMS subject classifications: 35B40, 35Q84, 46N30, 82M22, 60H10.

1 Introduction

The *generalized Langevin equation* (GLE) was originally proposed in the context of molecular dynamics and nonequilibrium statistical mechanics, in order to describe the motion of a particle interacting with a heat bath at equilibrium [53, 52, 77]; see also [41, 63, 57] for a rigorous derivation of the equation from a simple model of an open system, consisting of a small Hamiltonian system coupled to an infinite-dimensional, Hamiltonian heat reservoir modeled by the linear wave equation. The GLE has applications in many areas of science and engineering, ranging from atom/solid-surface scattering [19] to polymer dynamics [70], sampling in molecular dynamics [12, 13, 11], and global optimization with simulated annealing [26, 14].

The GLE is closely related, in a sense made precise below, to the simpler *Langevin* (also known as *underdamped Langevin*) equation, which itself reduces to the *overdamped Langevin* equation in the large friction limit. Arranged from the simplest to the most general, and written

*Department of Mathematics, Imperial College London (g.pavliotis@imperial.ac.uk)

†CERMICS, École des Ponts, France & MATHERIALS, Inria Paris (gabriel.stoltz@enpc.fr)

‡Department of Mathematics, Imperial College London (until October 2020) and MATHERIALS, Inria Paris (since November 2020) (urbain.vaes@inria.fr)

in one dimension for simplicity, these three standard models are the following:

$$\dot{q} = -V'(q) + \sqrt{2\,\beta^{-1}}\,\dot{W}, \tag{1.1a}$$

$$\ddot{q} = -V'(q) - \gamma\,\dot{q} + \sqrt{2\,\gamma\,\beta^{-1}}\,\dot{W}, \tag{1.1b}$$

$$\ddot{q} = -V'(q) - \int_0^t \gamma(t-s)\,\dot{q}(s)\,\mathrm{d}s + F(t), \quad \text{where} \quad \mathbf{E}(F(t_1)F(t_2)) = \beta^{-1}\,\gamma(|t_1 - t_2|). \tag{1.1c}$$

Here V is an external potential, F is a stationary Gaussian stochastic forcing, β is the inverse temperature, the parameter γ in (1.1b) is the *friction coefficient*, and the function $\gamma(\cdot)$ in (1.1c) is a *memory kernel*. The constraint $\mathbf{E}(F(t_1)F(t_2)) = \beta^{-1}\,\gamma(|t_1 - t_2|)$ is known as the *fluctuation/dissipation* relation, and it guarantees that the canonical measure at temperature β^{-1} is a stationary distribution of (1.1c); see Section 2. Since the force field in (1.1c) is conservative – it derives from the potential V – and the fluctuation/dissipation relation is assumed to hold, equation (1.1c) is sometimes called an *equilibrium GLE* [44].

The GLE (1.1c) is a non-Markovian stochastic integro-differential equation which, in general, is less amenable to analysis than the Langevin (1.1b) and overdamped Langevin (1.1a) equations. Instead of studying the GLE in its full generality, we will restrict our attention to the case where the GLE is equivalent to a finite-dimensional system of Markovian stochastic differential equations (SDEs). This assumption is known as the quasi-Markovian approximation, and it is employed in many mathematical works on the GLE. It is possible to show that it is satisfied when the Laplace transform of the memory kernel is a finite continued fraction [52] or, relatedly, when the spectral density of the memory kernel is rational, in the sense of [64, 63]; see also [57] for more details on the quasi-Markovian approximation. In this paper, we will study two particular quasi-Markovian GLEs, corresponding to the cases where $\gamma(\cdot)$ is the autocorrelation function of scalar Ornstein–Uhlenbeck (OU) noise and harmonic noise; see Section 2 for precise definitions.

For quasi-Markovian GLEs, it is possible to rigorously prove the passage to (1.1b) in the so-called *white noise limit*. This was done in [55] by leveraging recent developments in multiscale analysis [58]. More precisely, in [55] the authors showed that, with appropriate scalings, the solution of the quasi-Markovian GLE with OU noise converges, in the sense of weak convergence of probability measures on the space of continuous functions, to that of the Langevin equation (1.1b) when the autocorrelation function of the noise converges to a Dirac delta measure.

Our objectives in this paper are twofold: to study the longtime behavior of solutions to a simple quasi-Markovian GLE under quite general assumptions on the potential V and, based on this analysis, to study scaling limits of the effective diffusion coefficient associated with the dynamics in the particular case where V is periodic.

Longtime behavior. The longtime behavior of quasi-Markovian GLEs was studied in several settings in the literature. The exponential convergence of the corresponding semigroup to equilibrium was proved in [64]. In this paper, which is part of a series of papers studying a model consisting of a chain of anharmonic oscillators coupled to Hamiltonian heat reservoirs [23, 22, 21], the authors proved the convergence in an appropriately weighted L^∞ norm, by relying on Lyapunov-based techniques for Markov chains. We also mention [49, 50, 62, 29, 45]

as useful references on the Lyapunov-based approach. Later, in [55], the exponential convergence to equilibrium was proved for the GLE driven by OU noise using Villani's hypocoercivity framework. The authors showed the exponential convergence of the Markov semigroup both in relative entropy and in a weighted H^1 space. More recently, exponential convergence results in an appropriately weighted L^∞ norm were obtained in [44] for a more general class of quasi-Markovian GLEs than had been considered previously, allowing non-conservative forces and position-dependent noise.

Roughly speaking, the first aim of this paper is to obtain, for quasi-Markovian GLEs, H^1 and L^2 convergence estimates similar to those of [55] but valid *uniformly over the space of parameters* that enter the equations, i.e. the parameters of the noise process driving the dynamics. This turns out to be crucial for proving the validity of asymptotic expansions for the effective diffusion coefficient in several limits of interest – our second goal.

Effective diffusion in a periodic potential. The behavior of a Brownian particle in a periodic potential has applications in many areas of science, including electronics [71, 72], biology [61], surface diffusion [27] and Josephson tunneling [5]. For the Langevin (1.1b) and overdamped Langevin (1.1a) equations, as well as for all finite-dimensional approximations of the GLE (1.1c), a functional central limit theorem (FCLT) holds under appropriate assumptions on the initial condition (e.g. stationarity, see [55, Theorem 2.5]): applying the diffusive rescaling, the position process q converges as $\varepsilon \to 0$, in the sense of weak convergence of probability measures on $C([0,T],\mathbf{R})$, to a Brownian motion:

$$\left\{\varepsilon\, q(t/\varepsilon^2)\right\}_{t\in[0,T]} \Rightarrow \left\{\sqrt{2D}\, W(t)\right\}_{t\in[0,T]}, \tag{1.2}$$

where the *effective diffusion coefficient* D depends on the model and its parameters. This is shown formally in [59] for Langevin dynamics, and was proved more recently in [55, Theorem 2.5] for finite-dimensional approximations of the GLE.

In spatial dimension one, the behavior of the effective diffusion coefficient associated with Langevin dynamics (1.1b) is well understood; see, for example, [30] for a theoretical treatment and [59] for numerical experiments. The scaling of the effective diffusion coefficient with respect to the friction coefficient for Langevin dynamics has also been studied extensively in the physics literature. Whereas in the large friction limit a universal bound scaling as $\frac{1}{\gamma}$ holds for the diffusion coefficient in arbitrary dimensions, such a bound is true in the underdamped limit only in one dimension [30]. Claims that underdamped Brownian motion in periodic and random potentials in dimensions higher than one can lead to anomalous diffusion have been made [69, 42] but seem hard to justify rigorously.

The case of non-Markovian Brownian motion in a periodic potential has received less attention, even in one dimension. Early quantitative results were obtained in [36] by means of numerical experiments using the matrix-continued fraction method (see, e.g., [65, Section 9.1.2]), and verified in [35] by analog simulation. In these papers, the authors studied the dependence of the diffusion coefficient on the memory of the noise, and they were also able to calculate the velocity autocorrelation function and to study its dependence on the type of noise, i.e. OU or harmonic noise. Given that few authors have investigated the problem quantitatively since

then, and in light of the increased computational power available today, there is now scope for a more in-depth numerical study of the problem.

Our contributions. Our contributions in this paper are the following:

- We obtain sharp parameter-dependent estimates for the rate of convergence of the GLE to equilibrium in the particular cases of scalar OU and harmonic noises, thereby complementing previous results in [55]. Our approach is an explicit version of the standard hypocoercivity method [76, 32] and uses ideas from [46, 34] for the definition of an appropriate auxiliary norm.

- We show rigorously that the diffusive and white noise limits commute for quasi–Markovian approximations of the GLE. In other words, assuming that the memory of the noise in the GLE is encoded by a small parameter ν, and denoting by D_γ and $D_{\gamma,\nu}$ the effective diffusion coefficients associated with (1.1b) and (1.1c), respectively, we prove that

$$\lim_{\nu\to 0} D_{\gamma,\nu} = D_\gamma.$$

- For the case of OU noise, we study the influence on the effective diffusion coefficient of the friction coefficient that appears in the $\nu \to 0$ limiting Langevin equation, a coefficient that we will also refer to as the *friction coefficient* by a slight abuse of terminology. We show in particular, both by rigorous asymptotics and by numerical experiments, that the diffusive limit commutes with the overdamped limit $\gamma \to \infty$.

- We corroborate most of our theoretical analysis by careful numerical experiments, thereby complementing the results of the early studies [35, 36]. In these studies, because of the hardware limitations at the time, only about 15 basis functions per dimension could be used in 3 dimensions (3D) – position, momentum, and one auxiliary variable – and very few simulations could be achieved in 4 dimensions (4D). With today's hardware and the availability of high-quality mathematical software libraries, we were able to run accurate simulations in both 3D and 4D over a wide range of friction coefficients, including the underdamped limit $\gamma \to 0$.

The rest of the paper is organized as follows. In Section 2, we present the finite-dimensional Markovian models of the GLE that we focus on throughout the paper and we summarize our main results. In Section 3, we obtain an explicit estimate for the rate of convergence to equilibrium of the solution to the GLE. In Section 4, we carry out a multiscale analysis with respect to the correlation time of the noise, and we also study the overdamped and underdamped limits of the effective diffusion coefficient of the GLE. Section 5 is reserved for conclusions and perspectives for future work.

In the appendices, we present a few auxiliary results: in Appendix A, we assess the sharpness of the convergence rate found in Section 3, in the particular case of a quadratic potential; in Appendix B, we present a convergence estimate for harmonic noise; in Appendix C, we derive the technical results used in Section 4.3.

2 Model and main results

The model and the results we present are all stated in a one-dimensional setting. This allows to simplify the presentation and reduces the number of parameters to be considered: the mass of the system is set to 1 (instead of considering a general symmetric positive definite mass matrix) and the friction $\gamma(t)$ is scalar valued (whereas in general it would be a function with values in the set of symmetric positive matrices). The extension of our analysis to higher dimensional cases poses however no difficulties for most of the arguments – with the notable exception of the underdamped limit in Section 4.3.

2.1 Model

Throughout this paper, we assume that V is a smooth one-dimensional potential that is either confining (in particular, $\mathrm{e}^{-\beta V} \in L^1(\mathbf{R})$) or periodic with period 2π. The configuration of the system is described by its position $q \in \mathcal{X}$ and the associated momentum $p \in \mathbf{R}$. Positions are either in $\mathcal{X} = \mathbf{R}$ for confining potentials, or in the torus $\mathcal{X} = \mathbf{T} = \mathbf{R}/2\pi\mathbf{Z}$ for periodic potentials.

General structure of the colored noise. Let us first consider a memory kernel of the form

$$\gamma(t) = \mathbb{1}_{t \geqslant 0}\, \boldsymbol{\lambda}^{\mathsf{T}}\, \mathrm{e}^{-t\mathbf{A}}\, \boldsymbol{\lambda}, \tag{2.1}$$

for a (possibly nonsymmetric) matrix $\mathbf{A} \in \mathbf{R}^{n\times n}$ with eigenvalues with positive real parts, and a vector $\boldsymbol{\lambda} \in \mathbf{R}^n$. It is well-known that the GLE associated with (2.1) is quasi-Markovian [57, Proposition 8.1]: it is equivalent to a Markovian system of stochastic differential equations (SDEs),

$$\mathrm{d}q_t = p_t\, \mathrm{d}t, \tag{2.2a}$$

$$\mathrm{d}p_t = -V'(q_t)\, \mathrm{d}t + \boldsymbol{\lambda}^{\mathsf{T}} \mathbf{z}_t\, \mathrm{d}t, \tag{2.2b}$$

$$\mathrm{d}\mathbf{z}_t = -p_t\, \boldsymbol{\lambda}\, \mathrm{d}t - \mathbf{A}\mathbf{z}_t\, \mathrm{d}t + \boldsymbol{\Sigma}\, \mathrm{d}\mathbf{W}_t, \qquad \mathbf{z}_0 \sim \mathcal{N}(0, \beta^{-1}\mathbf{I}_n), \tag{2.2c}$$

where $\Sigma \in \mathbf{R}^{n\times n}$ is related to $\mathbf{A}$ by the fluctuation/dissipation theorem:

$$\boldsymbol{\Sigma}\boldsymbol{\Sigma}^{\mathsf{T}} = \beta^{-1} \left(\mathbf{A} + \mathbf{A}^{\mathsf{T}}\right).$$

The equivalence comes from the fact that (2.2c) can be integrated as

$$\mathbf{z}_t = -\int_0^t \mathrm{e}^{-(t-s)\mathbf{A}}\, \boldsymbol{\lambda}\, p_s\, \mathrm{d}s + \mathcal{N}_t, \qquad \mathcal{N}_t = \mathrm{e}^{-t\mathbf{A}}\, \mathbf{z}_0 + \int_0^t \mathrm{e}^{-(t-s)\mathbf{A}}\, \boldsymbol{\Sigma}\, \mathrm{d}\mathbf{W}_s,$$

with $\mathbf{E}(\mathcal{N}_t) = 0$ and, by the Itô isometry,

$$\begin{aligned}
\mathbf{E}\left(\mathcal{N}_{t_1}\mathcal{N}_{t_2}^\mathsf{T}\right) &= \mathrm{e}^{-t_1\mathbf{A}}\,\mathbf{E}\left(\mathbf{z}_0\mathbf{z}_0^\mathsf{T}\right)\mathrm{e}^{-t_2\mathbf{A}^\mathsf{T}} + \frac{1}{\beta}\int_0^{\min(t_1,t_2)} \mathrm{e}^{-(t_1-s)\mathbf{A}}\left(\mathbf{A}+\mathbf{A}^\mathsf{T}\right)\mathrm{e}^{-(t_2-s)\mathbf{A}^\mathsf{T}}\,\mathrm{d}s \\
&= \mathrm{e}^{-t_1\mathbf{A}}\left(\frac{1}{\beta}\mathbf{I}_n\right)\mathrm{e}^{-t_2\mathbf{A}^\mathsf{T}} + \frac{1}{\beta}\,\mathrm{e}^{-t_1\mathbf{A}}\left(\int_0^{\min(t_1,t_2)} \frac{\mathrm{d}}{\mathrm{d}u}\left(\mathrm{e}^{u\mathbf{A}}\,\mathrm{e}^{u\mathbf{A}^\mathsf{T}}\right)\Big|_{u=s}\mathrm{d}s\right)\mathrm{e}^{-t_2\,\mathbf{A}^\mathsf{T}} \\
&= \frac{1}{\beta}\,\mathrm{e}^{-(t_1-\min(t_1,t_2))\mathbf{A}}\,\mathrm{e}^{-(t_2-\min(t_1,t_2))\mathbf{A}^\mathsf{T}},
\end{aligned}$$

so $\mathbf{E}\left[\left(\boldsymbol{\lambda}^\mathsf{T}\mathcal{N}_{t_1}\right)\left(\boldsymbol{\lambda}^\mathsf{T}\mathcal{N}_{t_2}\right)\right] = \beta^{-1}\gamma\left(|t_1-t_2|\right)$.

The dynamics (2.2) is ergodic with respect to the probability measure

$$\mu(\mathrm{d}q\,\mathrm{d}p\,\mathrm{d}\mathbf{z}) = Z^{-1}\exp\left(-\beta\left(H(q,p)+\frac{|\mathbf{z}|^2}{2}\right)\right)\mathrm{d}q\,\mathrm{d}p\,\mathrm{d}\mathbf{z}, \qquad H(q,p) = V(q)+\frac{p^2}{2}, \tag{2.3}$$

with Z the normalization constant. Note that the invariant measure is independent of the parameters of the noise $\boldsymbol{\lambda}$ and $\mathbf{A}$. The generator of the Markov semigroup associated with the dynamics is given by

$$\mathcal{L} = \mathcal{L}_{\mathrm{anti}} + \mathcal{L}_{\mathrm{sym}},$$

where the symmetric part $\mathcal{L}_{\mathrm{sym}}$ of the generator, considered as an operator on $L^2(\mu)$, is related to the fluctuation and dissipation terms in (2.2c); while the antisymmetric part $\mathcal{L}_{\mathrm{anti}}$ corresponds to the Hamiltonian part of the dynamics (with Hamiltonian subdynamics for the couples (q,p) and $(p,\boldsymbol{\lambda}^\mathsf{T}\mathbf{z})$) and an additional evolution in the $\mathbf{z}$ degrees of freedom associated with the antisymmetric part of the matrix $\mathbf{A}$:

$$\begin{aligned}
\mathcal{L}_{\mathrm{anti}} &= p\,\frac{\partial}{\partial q} - V'(q)\,\frac{\partial}{\partial p} + \boldsymbol{\lambda}^\mathsf{T}\mathbf{z}\,\frac{\partial}{\partial p} - p\,\boldsymbol{\lambda}^\mathsf{T}\nabla_{\mathbf{z}} + \mathbf{z}^\mathsf{T}\mathbf{A}_{\mathrm{a}}\nabla_{\mathbf{z}}, \\
\mathcal{L}_{\mathrm{sym}} &= -\mathbf{z}^\mathsf{T}\mathbf{A}_{\mathrm{s}}\nabla_{\mathbf{z}} + \beta^{-1}\,\mathbf{A}_{\mathrm{s}} : \nabla_{\mathbf{z}}^2,
\end{aligned}$$

with $\mathbf{A}_{\mathrm{s}} = (\mathbf{A}+\mathbf{A}^\mathsf{T})/2$ and $\mathbf{A}_{\mathrm{a}} = (\mathbf{A}-\mathbf{A}^\mathsf{T})/2$ the symmetric and antisymmetric parts of $\mathbf{A}$, respectively, $\nabla_{\mathbf{z}}^2$ the Hessian operator and : the Frobenius inner product.

Specific models for the colored noise. In this study, we consider the two following models for the process $\mathbf{z}$:

GL1 The noise is modeled by a scalar OU process ($n = 1$), so $\boldsymbol{\lambda}$, $\mathbf{A}$, $\boldsymbol{\Sigma}$ and $\mathbf{z}$ are scalar quantities. We employ the parametrization

$$\boldsymbol{\lambda} = \frac{\sqrt{\gamma}}{\nu}, \qquad \mathbf{A} = \frac{1}{\nu^2},$$

for two positive parameters ν and γ. The associated memory kernel is

$$\gamma(t) = \frac{\gamma}{\nu^2}\,\mathrm{e}^{-t/\nu^2}\,\mathbb{1}_{t\geqslant 0}.$$

Note that

$$\int_0^{+\infty} \gamma(t)\,\mathrm{d}t = \gamma,$$

which motivates the abuse of notation between the constant γ and the function $\gamma(t)$ (the meaning of the object γ under consideration should however be clear from the context). Moreover,

$$\gamma(t) \xrightarrow[\nu\to 0]{} \gamma\delta_0, \tag{2.4}$$

which corresponds to a memoryless, Markovian limit.

GL2 The noise is modeled by a generalized version of harmonic noise:

$$\boldsymbol{\lambda} = \frac{1}{\nu}\begin{pmatrix}\sqrt{\gamma}\\ 0\end{pmatrix}, \qquad \mathbf{A} = \frac{1}{\nu^2}\begin{pmatrix}0 & -\alpha\\ \alpha & \alpha^2\end{pmatrix}, \qquad \boldsymbol{\Sigma} = \sqrt{\frac{2\alpha^2}{\beta\nu^2}}\begin{pmatrix}0 & 0\\ 0 & 1\end{pmatrix}.$$

The associated memory kernel is given by $\frac{\gamma}{\nu^2}\,\mathrm{e}^{-2t/\nu^2}\,\mathbb{1}_{t\geqslant 0}$ when $\alpha = 2$ and otherwise by

$$\gamma(t) = \frac{\gamma}{\nu^2}\exp\left(-\frac{\alpha^2 t}{2\,\nu^2}\right)\left[\frac{\alpha}{\sqrt{|4-\alpha^2|}}\,s_\alpha\left(\frac{\sqrt{|4-\alpha^2|}}{2}\,\frac{\alpha t}{\nu^2}\right) + c_\alpha\left(\frac{\sqrt{|4-\alpha^2|}}{2}\,\frac{\alpha t}{\nu^2}\right)\right]\mathbb{1}_{t\geqslant 0}, \tag{2.5}$$

where (s_α, c_α) are the functions $(\sin, \cos)$ when $\alpha < 2$ and $(\sinh, \cosh)$ when $\alpha > 2$. The latter expression can be found by computing the eigenvalues of $\mathbf{A}$, writing the solution as a sum of exponentials of these eigenvalues multiplied by the time t, and adjusting the coefficients in the linear combination so that $\gamma(0) = \boldsymbol{\lambda}^\mathsf{T}\boldsymbol{\lambda} = \gamma/\nu^2$ and $\gamma'(0^+) = -\boldsymbol{\lambda}^\mathsf{T}\mathbf{A}\boldsymbol{\lambda}$. In particular, since $\alpha^2 - \alpha\sqrt{\alpha^2-4} \to 2$ as $\alpha \to \infty$, we obtain that

$$\gamma(t) \xrightarrow[\alpha\to\infty]{} \frac{\gamma}{\nu^2}\,\mathrm{e}^{-t/\nu^2}\,\mathbb{1}_{t\geqslant 0}$$

for any $t \geqslant 0$, which is the autocorrelation function of the noise in the model GL1. The limit $\alpha \to \infty$ corresponds to an overdamped limit of the noise since (2.2c) reads, in the absence of the forcing term $-p_t\boldsymbol{\lambda}\,\mathrm{d}t$ and with the notation $\mathbf{z} = (z_1, z_2)$,

$$\begin{aligned}\mathrm{d}z_{1,t} &= \frac{\alpha}{\nu^2}z_{2,t}\,\mathrm{d}t,\\ \mathrm{d}z_{2,t} &= -\frac{\alpha}{\nu^2}z_{1,t}\,\mathrm{d}t - \frac{\alpha^2}{\nu^2}z_{2,t}\,\mathrm{d}t + \sqrt{\frac{2\alpha^2}{\beta\nu^2}}\,\mathrm{d}W_t,\end{aligned}$$

which, after rescaling time by a factor α/ν^2, corresponds to a Langevin dynamics with friction α for the $\mathbf{z}$ variable.

In both models GL1 and GL2, the parameters γ and ν (or, rather, ν^2) enter as scalings in the autocorrelation function, with ν being essentially the square root of the correlation time of the noise. In the model GL2, the parameter α encodes the shape of the function. Examples of memory kernels are illustrated in Fig. 1 for the two models under consideration and various values of α.

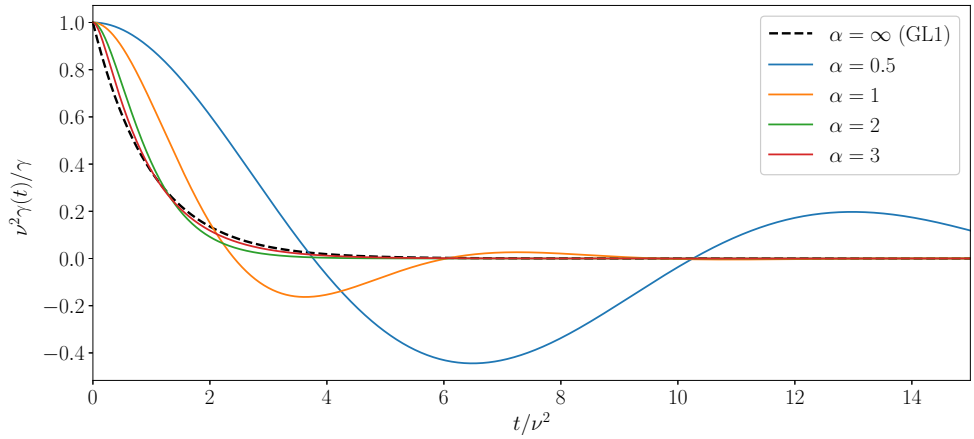


Figure 1: Memory kernels associated with models GL1 and GL2 (with time rescaled by a factor ν^{-2} and memory kernel rescaled by a factor ν^2/γ in order to have shapes independent of the choices of γ, ν).

Effective diffusion coefficient. We consider here the case when $\mathcal{X} = \mathbf{T}$, since the definition of an effective diffusion does not make sense for confining potentials. The derivation of the effective diffusion coefficient for systems of SDEs of the type (2.2) in a periodic potential is well understood; see for example [59, 58] for a formal derivation and [55] for a rigorous proof for GLE. The diffusion coefficient can be expressed in terms of the solution to a Poisson equation:

$$D = \int_{\mathbf{T}\times\mathbf{R}\times\mathbf{R}^n} \phi\, p \,\mathrm{d}\mu, \tag{2.6a}$$

$$-\mathcal{L}\phi = p. \tag{2.6b}$$

The Poisson equation (2.6b) is equipped with periodic boundary conditions in $[-\pi, \pi]$ and the condition that ϕ is square integrable with respect to the Gibbs measure, *i.e.* $\phi \in L^2(\mu)$. In fact, in order to guarantee the uniqueness of the solution to (2.6b), one should further assume that ϕ has average 0 with respect to μ. Since $\mathcal{L}_{\mathrm{FD}}$ and $\mathcal{L}_{\mathrm{ham}}$ are symmetric and antisymmetric operators on $L^2(\mu)$, respectively, and since $\mathcal{L}_{\mathrm{FD}} = \beta^{-1}\, \mathrm{e}^{\beta\frac{|\mathbf{z}|^2}{2}}\, \nabla_{\mathbf{z}} \cdot (\mathrm{e}^{-\beta\frac{|\mathbf{z}|^2}{2}}\, \mathbf{A}_{\mathrm{s}} \nabla_{\mathbf{z}})$, the effective diffusion coefficient can also be rewritten as

$$D = \beta^{-1} \int_{\mathbf{T}\times\mathbf{R}\times\mathbf{R}^n} \nabla_{\mathbf{z}}\phi^{\mathsf{T}} \mathbf{A}_{\mathrm{s}} \nabla_{\mathbf{z}}\phi \,\mathrm{d}\mu. \tag{2.7}$$

Note finally that definitions similar to (2.6a) and (2.6b) hold also for the Langevin dynamics, provided that $\mathcal{L}$ is defined as the corresponding generator (see (4.2)).

2.2 Main results

Before stating our main results, let us introduce some notation. In this paper, $H^1(\mu)$ denotes the subspace of $L^2(\mu)$ of functions whose gradient is in $L^2(\mu)$, equipped with the usual weighted Sobolev inner product. The spaces $L^2_0(\mu)$ and $H^1_0(\mu)$ are the subspaces of $L^2(\mu)$ and $H^1(\mu)$ of functions with mean 0 with respect to μ, respectively, and $\mathcal{B}(E)$ is the space of bounded linear

operators on a Banach space E, equipped with the operator norm

$$\|\mathcal{A}\|_{\mathcal{B}(E)} = \sup_{f \in E \setminus \{0\}} \frac{\|\mathcal{A} f\|_E}{\|f\|_E}.$$

Exponential decay and resolvent estimates. Our first result concerns the convergence to equilibrium for the model GL1. In line with the standard approach in molecular dynamics, we state our results for the semigroup $\mathrm{e}^{t\mathcal{L}}$, which describes the time evolution of average properties, but we note that our estimates apply equally to the adjoint semigroup $\mathrm{e}^{t\mathcal{L}^*}$, where $\mathcal{L}^*$ denotes the $L^2(\mu)$ adjoint of $\mathcal{L}$. This can be shown by duality and, as mentioned in [45, 57], can also be understood from the fact that $\mathcal{L}^* = -\mathcal{L}_{\mathrm{ham}} + \mathcal{L}_{\mathrm{FD}}$ coincides with $\mathcal{L}$ up to the sign of the Hamiltonian part. In turn, convergence estimates for $\mathrm{e}^{t\mathcal{L}^*}$ can be translated into estimates for $\mathrm{e}^{t\mathcal{L}^\dagger}$, where $\mathcal{L}^\dagger$ denotes the Fokker–Planck operator associated with the dynamics. This is because $\mathrm{e}^{t\mathcal{L}^\dagger}\,\psi = \mu\,\mathrm{e}^{t\mathcal{L}^*}(\mu^{-1}\psi)$ for any test function $\psi \in L^2(\mu^{-1})$ where, by a slight abuse of notation, μ denotes in this context the Lebesgue density of the measure (2.3).

Although our results on the effective diffusion apply only to the case of a periodic potential, Theorems 2.1 and 2.2 below apply also when V is a confining potential, provided that the following assumption holds.

Assumption 2.1. *The potential V is smooth. When the position space is $\mathbf{R}$ we moreover assume that the following conditions are satisfied:*

(i) $\mathrm{e}^{-\beta V} \in L^1(\mathbf{R})$,

(ii) $\left\|V''\right\|_\infty := \sup_{q \in \mathbf{R}} |V''(q)| < \infty$,

(iii) the following Poincaré inequality holds true for some constant $R_\beta > 0$:

$$\forall \varphi \in H^1(\mathrm{e}^{-\beta V}), \qquad \|\varphi - \bar{\varphi}\|^2_{L^2(\mathrm{e}^{-\beta V})} \leqslant \frac{1}{R_\beta}\,\|\nabla \varphi\|^2_{L^2(\mathrm{e}^{-\beta V})}\,, \qquad \bar{\varphi} := \frac{\displaystyle\int_{\mathcal{X}} \varphi\,\mathrm{e}^{-\beta V}}{\displaystyle\int_{\mathcal{X}} \mathrm{e}^{-\beta V}}. \tag{2.8}$$

Note that, for smooth periodic potentials, the Poincaré inequality (2.8) holds true without any additional condition on V (see for instance the discussion in [45, Section 2.2.1]). The above assumption allows us to prove the following results.

Theorem 2.1 (Hypoelliptic regularization)**.** *Let $\mathcal{L}$ denote the generator associated with the model GL1 and suppose that Assumption 2.1 holds. Then for any $h \in L^2_0(\mu)$ and any parameters $\gamma > 0$ and $\nu > 0$, it holds $\mathrm{e}^{t\mathcal{L}}\,h \in H^1_0(\mu)$ for all $t > 0$ and there is an inner product $(\!(\,\cdot\,,\,\cdot\,)\!)_{\gamma,\nu}$ equivalent to the usual $H^1(\mu)$ inner product such that*

$$\forall h \in L^2_0(\mu), \qquad \left(\!\left(\mathrm{e}^{\mathcal{L}}\,h, \mathrm{e}^{\mathcal{L}}\,h\right)\!\right)_{\gamma,\nu} \leqslant \|h\|^2_{L^2(\mu)}\,. \tag{2.9}$$

We next state a convergence result in $H^1_0(\mu)$, which relies on the hypocoercive framework described in [76].

Theorem 2.2 ($H^1(\mu)$ hypocoercivity)**.** *Suppose that Assumption 2.1 holds, and consider the inner product* $(\!(\cdot\,,\cdot)\!)_{\gamma,\nu}$ *constructed in the proof of Theorem 2.1. Then there exists a constant* $C_1 \in \mathbf{R}_+$ *such that, for any* $t \geqslant 0$ *and any parameters* $\nu > 0$ *and* $\gamma > 0$*, it holds*

$$\forall f \in H^1_0(\mu), \qquad (\!(\mathrm{e}^{t\mathcal{L}} f, \mathrm{e}^{t\mathcal{L}} f)\!)_{\gamma,\nu} \leqslant \exp\left(-C_1 \min\left(\gamma, \frac{1}{\gamma}, \frac{\gamma}{\nu^4}\right) t\right) (\!(f,f)\!)_{\gamma,\nu}\,. \tag{2.10}$$

In particular, there exists $C_2(\gamma,\nu) \in \mathbf{R}_+$ *such that*

$$\left\|\mathrm{e}^{t\mathcal{L}}\right\|_{\mathcal{B}(H^1_0(\mu))} \leqslant C_2(\gamma,\nu)\, \exp\left(-C_1 \min\left(\gamma, \frac{1}{\gamma}, \frac{\gamma}{\nu^4}\right) t\right).$$

Theorems 2.1 and 2.2 can be combined to obtain a decay estimate in $L^2(\mu)$. More precisely, employing the fact that $\|\cdot\|^2 \leqslant (\!(\cdot\,,\cdot)\!)_{\gamma,\nu}$ (see the definition (3.4) and (3.5)) together with (2.9) and (2.10), we obtain, for $h \in L^2_0(\mu)$,

$$\begin{aligned}
\left\|\mathrm{e}^{t\mathcal{L}} h\right\|^2_{L^2(\mu)} \leqslant (\!(\mathrm{e}^{t\mathcal{L}} h, \mathrm{e}^{t\mathcal{L}} h)\!)_{\nu,\gamma} &\leqslant \exp\left(-2\,C_1\, \min\left(\gamma, \frac{1}{\gamma}, \frac{\gamma}{\nu^4}\right)(t-1)\right)\, (\!(\mathrm{e}^{\mathcal{L}} h, \mathrm{e}^{\mathcal{L}} h)\!)_{\nu,\gamma} \\
&\leqslant \exp\left(-2\,C_1\, \min\left(\gamma, \frac{1}{\gamma}, \frac{\gamma}{\nu^4}\right)(t-1)\right)\, \|h\|^2_{L^2(\mu)} \\
&\leqslant \mathrm{e}^{2C_1} \exp\left(-2\,C_1\, \min\left(\gamma, \frac{1}{\gamma}, \frac{\gamma}{\nu^4}\right) t\right)\, \|h\|^2_{L^2(\mu)}
\end{aligned} \tag{2.11}$$

for any $t \geqslant 1$. This inequality also holds for $0 \leqslant t \leqslant 1$ in view of the trivial bound $\left\|\mathrm{e}^{t\mathcal{L}}\right\|_{\mathcal{B}(L^2_0(\mu))} \leqslant 1$, so the following resolvent bound holds [45, Proposition 2.1].

Corollary 2.1. *Under the same assumptions as in Theorem 2.2, there exist a constant* $C \in \mathbf{R}_+$ *independent of* γ *and* ν *such that*

$$\left\|\mathcal{L}^{-1}\right\|_{\mathcal{B}(L^2_0(\mu))} \leqslant C \max\left(\gamma, \frac{1}{\gamma}, \frac{\nu^4}{\gamma}\right). \tag{2.12}$$

In fact, $C = \mathrm{e}^{2C_1}/2C_1$. The dependence of the exponential decay rate in (2.11) with respect to the parameters ν, γ is illustrated in Fig. 2. It is worth comparing the scaling of the exponential decay rate to the scalings obtained for Langevin dynamics, for which $\lambda_{\mathrm{Lang}}(\gamma)$ scales as $\min(\gamma, \gamma^{-1})$, as proved in [16, 28, 34] (see also the discussion at the beginning of Section 3). The rates obtained for GLE are therefore in line with these rates in the limit $\nu \to 0$, which is precisely the limit (2.4) in which GLE reduces to Langevin dynamics. The additional term γ/ν^4 in the scaling factor for the exponential decay rate of GLE is important only in the limit $\nu \to \infty$ (with the additional condition $\nu^2 \gg \gamma$ if the limits $\gamma, \nu \to \infty$ are taken simultaneously).

Sharpness of the bounds on the exponential decay rate. In the particular case where V is the quadratic potential $k\frac{q^2}{2}$ with $k > 0$, the scaling of the exponential growth bound with respect to γ, ν in all the limits of interest can be calculated explicitly. Indeed, in this case (2.2) can be written in the general form

$$\dot{\mathbf{X}} = \mathbf{D}\mathbf{X}\,\mathrm{d}t + \boldsymbol{\sigma}\,\mathrm{d}\mathbf{W}, \tag{2.13}$$

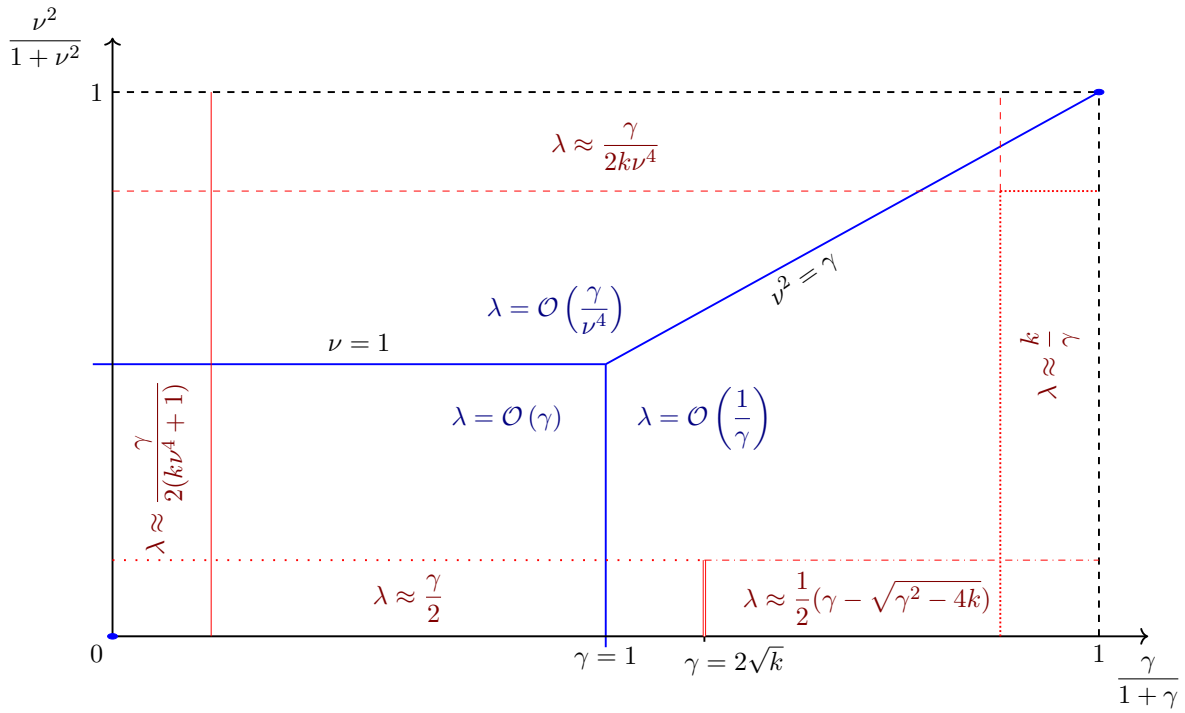


Figure 2: Schematic illustration of the scaling of the exponential decay rate λ in $L^2_0(\mu)$. The general estimate, implied by (2.11) and valid under the general Assumption 2.1, is illustrated in blue. The behavior of the exponential growth bound in the particular case of the quadratic potential $V(q) = k\frac{q^2}{2}$ is depicted in red. In this context, the symbol $\approx$ means that the relative error is arbitrarily close to zero in the corresponding limit.

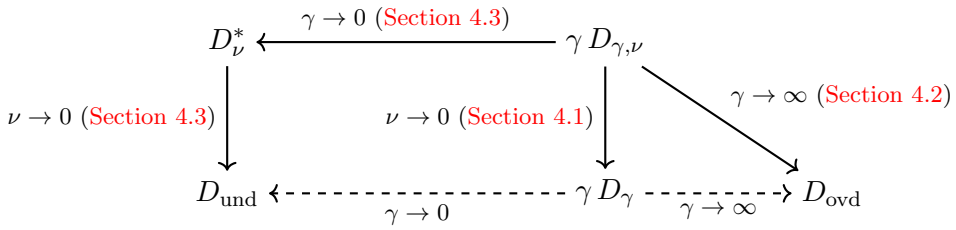
where $\mathbf{X}^\mathsf{T} = (q, p, \mathbf{z}^\mathsf{T})$, $\mathbf{D}$ and $\boldsymbol{\sigma}$ are constant matrices, and $\mathbf{W}$ is a standard Brownian motion on $\mathbf{R}^{2+n}$. It is known by a result from Metafune, Pallara and Priola [51] that the corresponding generator $\mathcal{L}$ generates a strongly continuous and compact semigroup in $L^p(\mu)$ for any $p \in (1, \infty)$, and that the associated spectrum can be obtained explicitly by a linear combination of the eigenvalues of the drift matrix $\mathbf{D}$ in (2.2):

$$\sigma(\mathcal{L}) = \left\{ \sum_{\lambda \in \sigma(\mathbf{D})} \lambda\, k_\lambda, \quad k_\lambda \in \mathbf{N}_{\geqslant 0} \right\}; \tag{2.14}$$

see also [48, Section 9.3] and [56]. By [2, Theorem 5.3], the spectral bound of the generator, *i.e.* the eigenvalue of $\mathcal{L}$ with the largest real part, coincides up to a sign change with the exponential decay rate of the semigroup (and in fact the norms of the propagators $\mathrm{e}^{t\mathbf{D}}$ and $\mathrm{e}^{t\mathcal{L}}$ coincide, as made precise in [4]), so estimating this growth bound in the quadratic case amounts to calculating the eigenvalues of the matrix $\mathbf{D}$. This can be achieved either numerically or analytically in the limiting regimes where the parameters go to either 0 or ∞, based on rigorous asymptotics for the associated characteristic polynomial (as made precise in Appendix A). The behavior of the spectral bound in the limiting regimes is indicated in Fig. 2.

Scaling limits for the effective diffusion coefficient. In Sections 4.1 to 4.3, we establish, either formally or rigorously, the limits in solid arrows in the following diagram (the limits in

dashed arrows are already known results):



Here D_{ovd} denotes the effective diffusion coefficient associated with the overdamped Langevin dynamics (1.1a), while D_{und} and D_γ are diffusion coefficients associated with the Langevin dynamics (1.1b). Let us emphasize that D_ν^* depends only on ν and that D_{und} is a constant, independent of the parameters of the noise, that can be calculated using the approach outlined in [59].

More precisely, we establish rigorously in Section 4.1 that, for fixed γ and for quite general Markovian approximations of the noise, $D_{\gamma,\nu} = D_\gamma + \mathcal{O}\left(\nu^2\right)$ as $\nu \to 0$. Our proof is based on an asymptotic expansion of the solution to the Poisson equation (2.6b) and on the resolvent estimate (2.12). Using similar techniques, we show in Section 4.2 that $D_{\gamma,\nu} \to D_{\mathrm{ovd}}$ in the limit $\gamma \to \infty$ for fixed $\nu > 0$ and in the particular case of model GL1. Finally, in Section 4.3 we motivate, with a formal asymptotic expansion similar to that employed in [59], that $\gamma D_{\gamma,\nu} \to D_\nu^*$ in the underdamped limit $\gamma \to 0$.

In principle, we could also study the limit $\nu \to \infty$. However, we refrain from doing so here because, first, this limit is less relevant from a physical viewpoint than the other limits considered and, second, this limit is technically more difficult. The main difficulty originates from the fact that the leading-order part of the generator is $p\,\partial_q - V'\,\partial_p$, so the terms in the asymptotic expansion of the solution to (2.6b) are not explicit. In addition, the operator norm of the resolvent scales as ν^4, so a large number of terms are required for proving a rigorous result.

2.3 Numerical experiments

Here we verify numerically the limits $\gamma D_{\gamma,\nu} \to D_{\mathrm{ovd}}$ as $\gamma \to \infty$ and $D_{\gamma,\nu} \to D_\gamma$ as $\nu \to 0$, with a spectral method to approximate the solution to the Poisson equation (2.6b). We come back to the case when $\mathcal{X} = \mathbf{T}$. We employ a Galerkin method that is in general non-conformal, in the sense that the finite-dimensional approximation space, which we denote by V_N, does not necessarily contain only mean-zero functions with respect to μ. Following the ideas developed in [66], we use a saddle point formulation to obtain an approximation of the solution to (2.6b):

$$\begin{cases} -\,\Pi_N\,\mathcal{L}\,\Pi_N\Phi_N + \alpha_N u_N = \Pi_N p, \\ \Phi_N^{\mathsf{T}} u_N = 0, \end{cases} \tag{2.15}$$

where Π_N is the $L^2(\mu)$ projection operator on V_N, $u_N = \Pi_N 1 / \|\Pi_N 1\| \in V_N$ and α_N is a Lagrange multiplier. As above, $\langle \cdot, \cdot \rangle$ and $\|\cdot\|$ denote respectively the standard scalar product and norm of $L^2(\mu)$. We choose V_N to be the subspace of $L^2(\mu)$ spanned by tensor products of appropriate one-dimensional functions constructed from trigonometric functions (in the q

direction) and Hermite polynomials (in the p and $\mathbf{z}$ directions). In the case of OU noise, for example, we use the basis functions

$$e_{i,j,k} = Z^{1/2}\, \mathrm{e}^{\frac{\beta}{2}\left(H(q,p)+\frac{z^2}{2}\right)}\, G_i(q)\, H_j(p)\, H_k(z), \qquad 0 \leqslant i,j,k \leqslant N,$$

where G_i are trigonometric functions,

$$G_i(q) = \begin{cases} (2\pi)^{-1/2}, & \text{if } i = 0, \\ \pi^{-1/2} \sin\left(\dfrac{i+1}{2} q\right), & \text{if } i \text{ is odd}, \\ \pi^{-1/2} \cos\left(\dfrac{i}{2} q\right), & \text{if } i \text{ is even}, i > 0. \end{cases} \tag{2.16}$$

and H_j are rescaled normalized Hermite functions,

$$H_j(p) = \frac{1}{\sqrt{\sigma}}\, \psi_j\left(\frac{p}{\sigma}\right), \qquad \psi_j(p) := (2\pi)^{-\frac{1}{4}} \frac{(-1)^j}{\sqrt{j!}}\, \mathrm{e}^{\frac{p^2}{4}}\, \frac{\mathrm{d}^j}{\mathrm{d}p^j}\left(\mathrm{e}^{-\frac{p^2}{2}}\right). \tag{2.17}$$

The functions $(H_j)_{j\in\mathbf{N}}$ are orthonormal in $L^2(\mathbf{R})$ regardless of the value of σ, and σ is a scaling parameter that can be adjusted to better resolve Φ_N. We use the same number N of basis functions in every direction because, although Hermite series converge much slower than Fourier series as $N \to \infty$ when σ is fixed, their spatial resolution is comparable to that of Fourier series when σ is chosen appropriately, as demonstrated in [74].

To solve the linear system associated with (2.15) and the basis functions (2.16), we use either the *SciPy* [37] function `scipy.sparse.linalg.spsolve`, which implements a direct method, or, when the time or memory required to solve (2.15) with a direct method is prohibitive, the function `scipy.sparse.linalg.gmres`, which implements the generalized minimal residual method (GMRES) [68].

The numerical results presented below are for the one-dimensional periodic cosine potential $V(q) = \frac{1}{2}(1 - \cos q)$, and they were all obtained with $\beta = 1$. We examine the variation of the diffusion coefficient with respect to γ for fixed $\nu = 1$ in Fig. 3. The parameters used in the simulations are presented in Table 1. The effective diffusion coefficient was computed for 100 values of γ evenly spaced on a logarithmic scale, and for each value of γ the numerical error was approximated by carrying out the computation with half the number of basis functions in each direction. When the relative error estimated in this manner was over 1%, which occurred roughly when $\gamma \leqslant 10^{-2}$ for the model GL1 and $\gamma \leqslant 10^{-1}$ for the model GL2, the corresponding data points were considered inaccurate and were removed.

Model	Method when $\gamma < 1$	Method when $\gamma > 1$
L	Direct ($N = 250$, $\sigma^{-2} = 16$)	Direct ($N = 250$, $\sigma^{-2} = 16$)
GL1	GMRES ($N = 100$, $\sigma^{-2} = 9$, tol $= 10^{-3}$)	Direct ($N = 40$, $\sigma^{-2} = 3$)
GL2	GMRES ($N = 40$, $\sigma^{-2} = 6$, tol $= 10^{-3}$)	Direct ($N = 16$, $\sigma^{-2} = 2$)

Table 1: Numerical parameters used to generate the data presented in Fig. 3. We employed the *SciPy* function `scipy.sparse.linalg.spsolve` for the direct method, and the function `scipy.sparse.linalg.gmres` for GMRES.

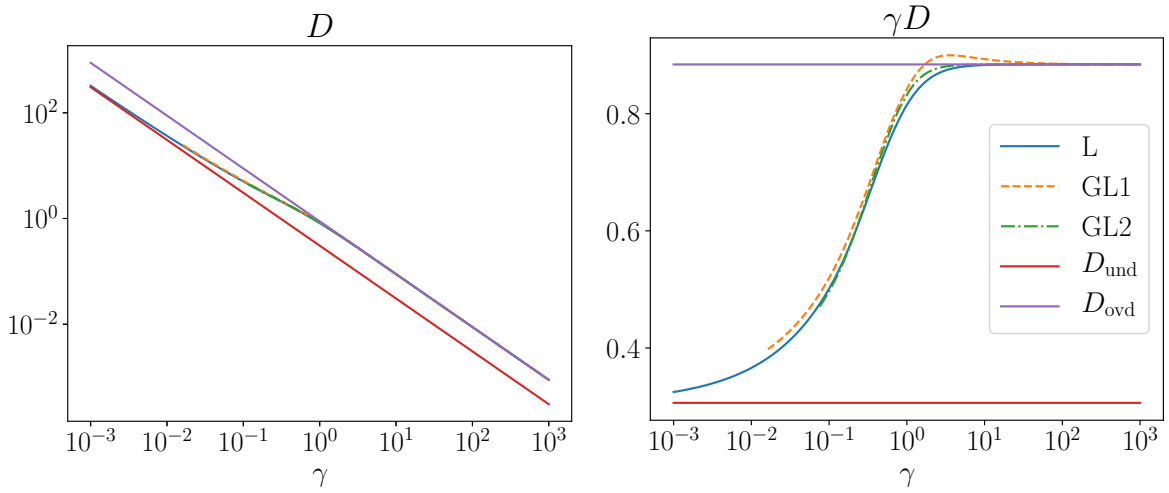


Figure 3: Diffusion coefficient as a function of γ, for the parameters $\nu = 1$ (for the models GL1 and GL2) and $\alpha = 1$ (for GL2). We observe that, for values of γ in the range $[1, 10]$, the GL1 diffusion coefficient is slightly larger than D_{ovd}/γ.

We observe from Fig. 3 that the effective diffusion coefficient is of the same order of magnitude for the three models across the whole range of γ, and that $\gamma D \to D_{\text{ovd}}$ for all models in the limit as $\gamma \to \infty$. We also notice that the inequality $\gamma D \leqslant D_{\text{ovd}}$, which was proved to hold for the Langevin dynamics in [30], is not satisfied for all values of γ in the case of the GLE; indeed it is clear from the figure that, for γ close to 2, the effective diffusion coefficient for the model GL1 is strictly greater than D_{ovd}/γ.

To conclude this section, we verify numerically that $D_{\gamma,\nu} \to D_\gamma$ in the limit as $\nu \to 0$ for fixed γ. Figure 4 presents the dependence on ν of the diffusion coefficient for the models GL1 and GL2 for various values of α. As expected, we recover the effective diffusion coefficient corresponding to the model GL1 as $\alpha \to \infty$, and that of the Langevin dynamics as $\nu \to 0$. For $\alpha = 1$, the convergence to the limit as $\nu \to 0$ appears to be faster than for the other values of α. In fact, it is possible to show that the deviation from the limiting effective diffusion coefficient is of order ν^4 in this case; see [75]. The convergence is illustrated in Fig. 5 in a log-log scale, which confirms the expected rates. When $\alpha = 1$, round off errors appear for small ν, which explains the deviation from the theoretical scaling ν^4.

3 Longtime behavior for model GL1

There are many results on the longtime convergence of the evolution semigroup $e^{t\mathcal{L}}$ of Langevin-like operators, as reviewed for instance in [8] (see also the recent review [33]). Among the approaches allowing us to quantify the scaling of the convergence rate as a function of the parameters of the dynamics, one can quote:

- $H^1(\mu)$ hypocoercivity, pioneered in [73] and [54], was later abstracted in [76]. The application of this theory to Langevin dynamics allows us to quantify the convergence rates in terms of the parameters of the dynamics; see for instance [30] for the Hamiltonian limit and [43, 45] for the overdamped limit. Moreover, the exponential convergence can be

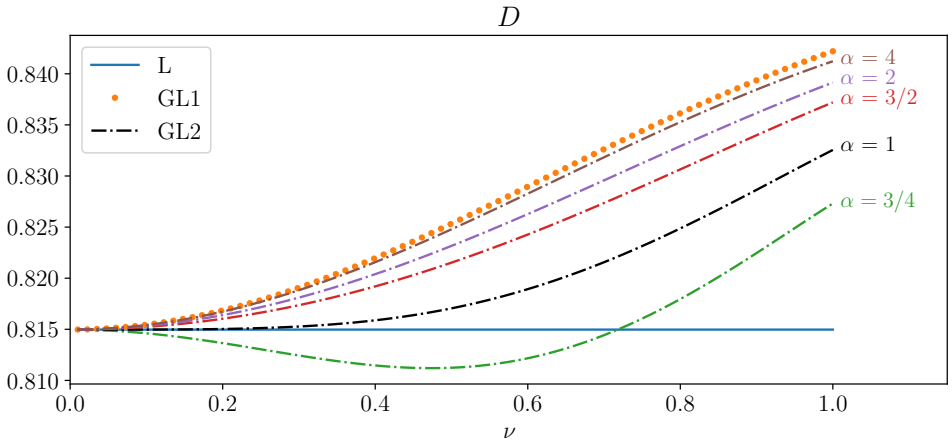


Figure 4: Effective diffusion coefficient against ν, whose square encodes the characteristic time of the autocorrelation function of the noise, for fixed values $\beta = \gamma = 1$.

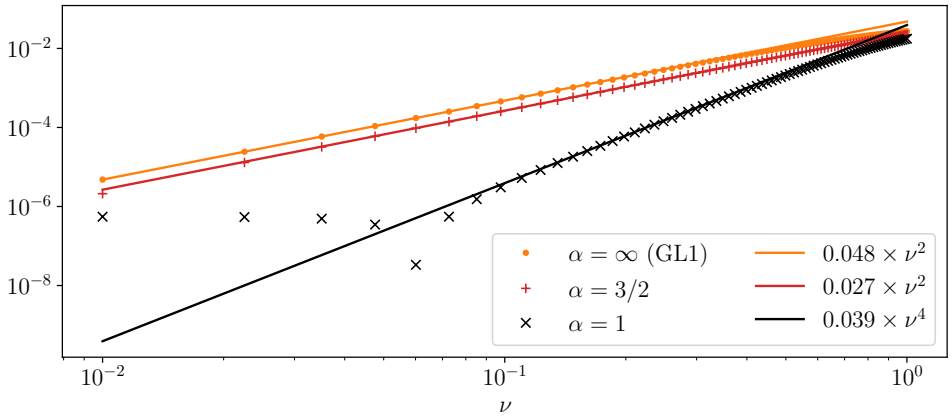


Figure 5: Deviation of the effective diffusion coefficient from its limiting value as $\nu \to 0$. For $\alpha = 1$, we observe that the data is not aligned with the straight line for small values of ν, which we attribute to round off errors.

transferred to $L^2(\mu)$ by hypoelliptic regularization [32].

- Entropic estimates, starting with [15], have been abstracted in [76], under conditions stronger than the ones for $H^1(\mu)$ hypocoercivity. Recently, it was shown how to remove the assumption that the Hessian of the potential is bounded [10].

- A more direct route to prove the convergence in $L^2(\mu)$ was first proposed in [31], then extended in [17, 18], and revisited in [28] where domain issues of the operators at play are addressed. It is based on a modification of the $L^2(\mu)$ scalar product with some regularization operator. This more direct approach makes it even easier to quantify convergence rates; see [16, 28, 66] for studies on the dependence of parameters such as the friction coefficient in Langevin dynamics, as well as [1] for sharp estimates for equilibrium Langevin dynamics and a harmonic potential energy function.

- Fully probabilistic techniques, based on clever coupling strategies, can also be used to obtain the exponential convergence of the law of Langevin processes to their stationary

state [20]. One interest of this approach is that the drift needs not be gradient, in contrast to standard analytical approaches for which the analytical expression of the invariant measure should be known in order to separate the symmetric and antisymmetric parts of the generator under consideration.

- Finally, it was recently shown how to directly obtain $L^2(\mu)$ estimates without changing the scalar product, relying on a space-time Poincaré inequality to conclude to an exponential convergence in time of the evolution semigroup [3, 9].

Our focus in this work is on functional analytic estimates, in $L^2(\mu)$ (where μ defined in (2.3) is the invariant measure of the dynamics), which is a natural framework for giving a meaning to quantities such as effective diffusion coefficients (which have the same form as asymptotic variances in central limit theorems for time averages). We were not able to work directly in $L^2(\mu)$ by generalizing the approach from [17, 18], because of the hierarchical structure of the dynamics, where the noise in $\mathbf{z}$ is first transferred to p and then to q. It is not so easy to construct a modified $L^2(\mu)$ scalar product in this case. On the other hand, the $H^1(\mu)$ framework of [76] can be used directly, as already done in [55]. Our contribution, compared to the latter work, is to carefully track the dependence of the convergence rate on the parameters of the dynamics.

In the calculations below we consider the model GL1 for simplicity, but similar calculations can be carried out for other quasi-Markovian models. Our results apply to both the periodic and confining settings. Throughout this section, all operators are considered by default on the functional space $L^2(\mu)$, the adjoint of a closed unbounded operator T on this space being denoted by T^*.

We will prove the main results in an order different from that in which they are stated in Section 2.2. This is both more usual and more natural because, as will become clear later, it is simpler to construct an inner product $(\!(\cdot\,,\cdot)\!)_{\gamma,\nu}$ such that (2.10) holds than to construct one such that (2.9) holds.

Remark 3.1. *The approach taken in this section can be applied in particular to the model GL2, as discussed in Appendix B. The computations are however algebraically more cumbersome, so the scalings we obtain for the resolvent bound appear not to be sharp, at least in the limit $\gamma \to \infty$.*

3.1 Proof of Theorem 2.2: decay in $H^1(\mu)$

We first introduce the adjoint operator $\partial_z^* = \beta z - \partial_z$ and rewrite the generator of the dynamics for the model GL1 in the standard form of the $H^1(\mu)$ coercivity framework [76]:

$$-\mathcal{L} = A^*A + B, \qquad A = \nu^{-1}\beta^{-1/2}\partial_z, \qquad B = \sqrt{\gamma}\,\nu^{-1}(p\,\partial_z - z\,\partial_p) + (V'(q)\,\partial_p - p\,\partial_q). \quad (3.1)$$

The relevant operators for the study of hypocoercivity are obtained from (iterated) commutators of B with A:

$$C_0 := \nu\,\beta^{1/2}\,A = \partial_z,$$
$$C_1 := \nu\,\gamma^{-1/2}\,[C_0, B] = -\partial_p, \tag{3.2a}$$
$$C_2 := [C_1, B] + \nu^{-1}\,\gamma^{1/2}\,C_0 = \partial_q, \tag{3.2b}$$
$$[C_2, B] + V''(q)\,C_1 = 0. \tag{3.2c}$$

If we were interested only in showing that $-\mathcal{L}$ is hypocoercive, it would be sufficient to invoke at this stage [76, Theorem 24], as done in [55]. Here, however, we are interested not only in whether the dynamics converge to equilibrium but also in the scaling of the rate of convergence with respect to ν and γ, so a careful analysis is required. Recalling that $\langle\cdot,\cdot\rangle$ and $\|\cdot\|$ denote respectively the standard scalar product and norm of $L^2(\mu)$, we denote by $(\!(\cdot,\cdot)\!)$ the inner product defined by polarization from the norm constructed with the operators C_0, C_1, C_2 defined above:

$$(\!(h,h)\!) = \|h\|^2 + a_0\,\|\partial_z h\|^2 + a_1\,\|\partial_p h\|^2 + a_2\,\|\partial_q h\|^2 - 2b_0\,\langle\partial_z h, \partial_p h\rangle - 2b_1\,\langle\partial_p h, \partial_q h\rangle\,, \tag{3.3}$$

that is

$$\begin{aligned}(\!(h_1,h_2)\!) = \langle h_1,h_2\rangle &+ a_0\,\langle\partial_z h_1, \partial_z h_2\rangle + a_1\,\langle\partial_p h_1, \partial_p h_2\rangle + a_2\,\langle\partial_q h_1, \partial_q h_2\rangle \\ &- b_0\,\langle\partial_z h_1, \partial_p h_2\rangle - b_0\,\langle\partial_z h_2, \partial_p h_1\rangle - b_1\,\langle\partial_p h_1, \partial_q h_2\rangle - b_1\,\langle\partial_p h_2, \partial_q h_1\rangle\,.\end{aligned} \tag{3.4}$$

In these expressions, the coefficients a_0, a_1, a_2 are positive, while b_0, b_1 are nonnegative (this sign convention is motivated by the computations performed in this section).

To prove hypocoercivity for the norm of $H^1(\mu)$, we must show that it is possible to find coefficients $a_0, a_1, a_2 > 0$ and $b_0, b_1 \geqslant 0$ such that:

(i) the standard weighted Sobolev norm $\|\cdot\|_{H^1(\mu)}$ (which corresponds to $a_0 = a_1 = a_2 = 1$ and $b_0 = b_1 = 0$ in (3.3)) is equivalent to the norm (3.3). Let us however mention that degenerate norms not equivalent to (3.3) can be considered, as initially done in [73], and recently used in [6, 46, 34];

(ii) coercivity holds for this modified norm, *i.e.* there exists $\lambda > 0$ such that $-(\!(h, \mathcal{L}h)\!) \geqslant \lambda\,(\!(h,h)\!)$ for all smooth, compactly supported and mean-zero $h \in C_c^\infty \cap L_0^2(\mu)$.

In order to be complete, we should in principle show that it is indeed sufficient to prove the coercivity inequality (ii) only for $h \in C_c^\infty \cap L_0^2(\mu)$. As discussed in the proof of [76, Theorem A.8] in a slightly different context, this requires an approximation argument, which we will omit here because this argument is both standard and somewhat technical. For the same reason, we will not worry about the technical justification of the calculations in Section 3.2.

By the Cauchy–Schwarz inequality and since $a_0, a_1, a_2, b_0, b_1 \geqslant 0$, it is clear that

$$((h,h)) - \|h\|^2 \geqslant \begin{pmatrix} \|\partial_z h\| \\ \|\partial_p h\| \\ \|\partial_q h\| \end{pmatrix}^{\mathsf{T}} \mathbf{M}_1 \begin{pmatrix} \|\partial_z h\| \\ \|\partial_p h\| \\ \|\partial_q h\| \end{pmatrix}, \qquad \mathbf{M}_1 = \begin{pmatrix} a_0 & -b_0 & 0 \\ -b_0 & a_1 & -b_1 \\ 0 & -b_1 & a_2 \end{pmatrix}, \tag{3.5}$$

so it is sufficient that $\mathbf{M}_1$ be positive definite in order to meet the first condition. For the second condition, we rely on the following auxiliary result.

Lemma 3.1. *Suppose that Assumption 2.1 holds. Then*

$$-((h,\mathcal{L}h)) \geqslant \frac{1}{\nu^2\beta} \begin{pmatrix} \|C_0\, C_0\, h\| \\ \|C_0\, C_1\, h\| \\ \|C_0\, C_2\, h\| \end{pmatrix}^{\mathsf{T}} \mathbf{M}_1 \begin{pmatrix} \|C_0\, C_0\, h\| \\ \|C_0\, C_1\, h\| \\ \|C_0\, C_2\, h\| \end{pmatrix} + \begin{pmatrix} \|C_0\, h\| \\ \|C_1\, h\| \\ \|C_2\, h\| \end{pmatrix}^{\mathsf{T}} \widetilde{\mathbf{M}}_2 \begin{pmatrix} \|C_0\, h\| \\ \|C_1\, h\| \\ \|C_2\, h\| \end{pmatrix}, \tag{3.6}$$

where $\mathbf{M}_1$ *is the same matrix as in (3.5) and* $\widetilde{\mathbf{M}}_2$ *is given by*

$$\widetilde{\mathbf{M}}_2 = \begin{pmatrix} \frac{1}{\beta\nu^2} + \frac{a_0}{\nu^2} - \frac{b_0\sqrt{\gamma}}{\nu} & -\left|(a_0 - a_1)\frac{\sqrt{\gamma}}{\nu} + \frac{b_0}{\nu^2}\right| & -\left|b_0 - \frac{b_1\sqrt{\gamma}}{\nu}\right| \\ 0 & \frac{b_0\sqrt{\gamma}}{\nu} - b_1 \left\|V''\right\|_\infty & -a_1 - a_2 \left\|V''\right\|_\infty \\ 0 & 0 & b_1 \end{pmatrix}. \tag{3.7}$$

It would be desirable to relax the condition $\|V''\|_\infty < \infty$ in Assumption 2.1 by following the approach presented in [76, Section 7], but this is not possible as such here; see Remark 3.2 for further precisions.

Proof. We calculate the action of the symmetric part of the generator on the terms multiplying a_0, a_1, a_2 in (3.3):

$$\langle C_0 h, C_0 (A^* A) h\rangle = \nu^{-2}\,\beta^{-1}\,\left\|C_0^2\, h\right\|^2 + \nu^{-2}\,\|C_0 h\|^2\,, \tag{3.8a}$$
$$\langle C_1 h, C_1 (A^* A) h\rangle = \nu^{-2}\,\beta^{-1}\,\|C_0\, C_1\, h\|^2\,, \tag{3.8b}$$
$$\langle C_2 h, C_2 (A^* A) h\rangle = \nu^{-2}\,\beta^{-1}\,\|C_0\, C_2\, h\|^2\,, \tag{3.8c}$$

where we took into account that C_0 commutes with C_1, C_2, while $[C_0, C_0^*] = \beta$. The action of the antisymmetric part of the generator B on the the same terms is, in view of the commutator relations (3.2a)-(3.2c),

$$\langle C_0 h, C_0 B h\rangle = \langle C_0 h, [C_0, B] h\rangle = \gamma^{1/2}\,\nu^{-1}\,\langle C_0 h, C_1 h\rangle\,, \tag{3.9a}$$
$$\langle C_1 h, C_1 B h\rangle = \langle C_1 h, [C_1, B] h\rangle = \langle C_1 h, C_2 h\rangle - \gamma^{1/2}\,\nu^{-1}\,\langle C_1 h, C_0 h\rangle\,, \tag{3.9b}$$
$$\langle C_2 h, C_2 B h\rangle = \langle C_2 h, [C_2, B] h\rangle = -\left\langle C_2 h, V''(q)\, C_1\, h\right\rangle. \tag{3.9c}$$

For the terms multiplying b_0, b_1 in (3.3), we have

$$\langle C_0(A^*A)h, C_1h\rangle + \langle C_0h, C_1(A^*A)h\rangle = \nu^{-2}\left(\langle C_0h, C_1h\rangle + 2\,\beta^{-1}\left\langle C_0^2h, C_0C_1h\right\rangle\right), \tag{3.10a}$$

$$\langle C_1(A^*A)h, C_2h\rangle + \langle C_1h, C_2(A^*A)h\rangle = 2\,\nu^{-2}\,\beta^{-1}\,\langle C_0C_1h, C_0C_2h\rangle\,, \tag{3.10b}$$

and

$$\begin{aligned}\langle C_0Bh, C_1h\rangle + \langle C_0h, C_1Bh\rangle &= \langle [C_0, B]h, C_1h\rangle + \langle C_0h, [C_1, B]h\rangle \\ &= \gamma^{1/2}\,\nu^{-1}\,\|C_1h\|^2 + \langle C_0h, C_2h\rangle - \gamma^{1/2}\,\nu^{-1}\,\|C_0h\|^2\,, \qquad (3.11a)\\ \langle C_1Bh, C_2h\rangle + \langle C_1h, C_2Bh\rangle &= \langle [C_1, B]h, C_2h\rangle + \langle C_1h, [C_2, B]h\rangle \\ &= \|C_2h\|^2 - \gamma^{1/2}\,\nu^{-1}\,\langle C_0h, C_2h\rangle - \left\langle V''(q)C_1h, C_1h\right\rangle. \qquad (3.11b)\end{aligned}$$

The inequality (3.6) then follows by combining (3.8) to (3.11) and using Assumption 2.1 as well as the Cauchy–Schwarz inequality. □

Remark 3.2. *For underdamped Langevin dynamics, it is possible to relax the condition* $\|V''\|_\infty < \infty$ *in Assumption 2.1 by following the approach of [76, Section 7]; see also the presentation in the proof of [45, Theorem 2.15], which relies on an estimate provided by [76, Lemma A.24]. The latter result states that, if* $V \in C^2(\mathbf{R}^n)$ *satisfies the inequality*

$$\forall q \in \mathbf{R}^n, \qquad \left|\nabla^2 V(q)\right| \leqslant c(1 + |\nabla V(q)|) \tag{3.12}$$

for some constant $c > 0$*, then there exist nonnegative constants* A_V *and* B_V *such that*

$$\forall h \in H^1(\mathrm{e}^{-\beta V}), \qquad \left\|h\,\nabla^2 V\right\|_{L^2(\mathrm{e}^{-\beta V})} \leqslant A_V\,\|h\|_{L^2(\mathrm{e}^{-\beta V})} + B_V\,\|\nabla h\|_{L^2(\mathrm{e}^{-\beta V})}\,.$$

Unfortunately, this approach does not enable to replace the condition of bounded Hessian by the weaker condition (3.12) in the case of model GL1. In particular, it seems difficult to control the term on the right-hand side of (3.9c). Indeed, quantities such as $|\langle C_1h, V''(q)C_2h\rangle|$ *would be bounded by factors such as* $\|C_2^2h\|$ *or* $\|C_1C_2h\|$*, which cannot be controlled with the first term on the right-hand side of (3.6). A similar issue arises with the last term on the right-hand side of (3.11b).*

Lemma 3.1 shows that the coercivity of $-\mathcal{L}$ for the modified norm is ensured if we can find parameters a_0, a_1, a_2, b_0, b_1 such that the matrix $\widetilde{\mathbf{M}}_2$ in (3.7) is positive definite. This is made precise in the following result (Proposition 3.1), which is a weaker version of Proposition 3.2 proved in the next section. In order to state it, we introduce the following notation: for matrices $\mathbf{X}, \mathbf{Y} \in \mathbf{R}^{d\times d}$,

$$\mathbf{X} \succcurlyeq_+ \mathbf{Y} \qquad \text{if} \qquad \mathbf{v}^\mathsf{T}\mathbf{X}\mathbf{v} \geqslant \mathbf{v}^\mathsf{T}\mathbf{Y}\mathbf{v} \qquad \forall \mathbf{v} \geqslant 0,$$

where the notation $\mathbf{v} \geqslant 0$ for $\mathbf{v} = (v_1, \dots, v_d) \in \mathbf{R}^d$ means that $v_i \geqslant 0$ for all $1 \leqslant i \leqslant d$. We also define the minimum of the Rayleigh quotient under a positivity constraint:

$$\lambda_{\min}^+(\mathbf{X}) := \min_{\mathbf{v}\neq 0, \mathbf{v}\geqslant 0} \frac{\mathbf{v}^\mathsf{T}\mathbf{X}\mathbf{v}}{\mathbf{v}^\mathsf{T}\mathbf{v}}. \tag{3.13}$$

Note that $\mathbf{X} \succcurlyeq_+ \mathbf{Y}$ implies that $\lambda_{\min}^+(\mathbf{X}) \geqslant \lambda_{\min}^+(\mathbf{Y})$.

Remark 3.3. *The inequality $\mathbf{X} \succcurlyeq \mathbf{Y}$ for two symmetric matrices implies $\mathbf{X} \succcurlyeq_+ \mathbf{Y}$, but not conversely: consider e.g. the 2-by-2 matrices with entries δ_{ij} and $(-1)^{i+j}$ for $1 \leqslant i, j \leqslant 2$. Remark also that, for a matrix $\mathbf{X}$ with nonpositive off-diagonal entries (such as $\widetilde{\mathbf{M}}_2$), it is equivalent to define $\lambda_{\min}^+(\mathbf{X})$ as the smallest eigenvalue of the symmetrized matrix $(\mathbf{X}+\mathbf{X}^\mathsf{T})/2$, since the minimum of $\mathbf{v}^\mathsf{T}\mathbf{X}\mathbf{v}$ on the sphere $|\mathbf{v}| = 1$ is achieved for some $\mathbf{v}$ with nonnegative elements.*

We easily deduce from (3.7) that

$$\widetilde{\mathbf{M}}_2 \succcurlyeq_+ \begin{pmatrix} (\beta^{-1}+a_0)\nu^{-2} - b_0 r & -a_0 r - a_1 r - b_0 \nu^{-2} & -b_0 - b_1 r \\ 0 & b_0 r - b_1 \left\|V''\right\|_\infty & -a_1 - a_2 \left\|V''\right\|_\infty \\ 0 & 0 & b_1 \end{pmatrix}, \tag{3.14}$$

where

$$r := \gamma^{1/2}\,\nu^{-1}.$$

It is therefore sufficient to work with the matrix on the right-hand side of (3.14) to derive a lower bound for $\lambda_{\min}^+\left(\widetilde{\mathbf{M}}_2\right)$, and thus for the rate of convergence.

Proposition 3.1. *There exist parameters $(a_0, a_1, a_2, b_0, b_1)$, as well as a constant $C > 0$ (independent of γ, ν) and $\alpha(\gamma, \nu) > 0$, such that $\alpha(\gamma,\nu)\mathbf{I}_3 \preccurlyeq \mathbf{M}_1 \preccurlyeq \mathbf{I}_3$ and*

$$\widetilde{\mathbf{M}}_2 \succcurlyeq_+ C \min\left(r^2\nu^2, \frac{1}{r^2\nu^2}, \frac{r^2}{\nu^2}\right)\mathbf{I}_3 = C\min\left(\gamma, \frac{1}{\gamma}, \frac{\gamma}{\nu^4}\right)\mathbf{I}_3.$$

We obtain from the previous proposition and (3.6) that

$$-\left(\!\left(h, \mathcal{L}h\right)\!\right) \geqslant \lambda_{\min}^+\left(\widetilde{\mathbf{M}}_2\right)\sum_{i=0}^{2}\left\|C_i h\right\|^2.$$

We rely at this stage on Poincaré's inequality to control $\|h\|^2$ with the right-hand side of the above inequality. Since μ is a product of probability measures that satisfy a Poincaré's inequality (the marginals in p, z being Gaussian distributions of variance β^{-1}), it itself satisfies a Poincaré inequality, see for instance [45, Proposition 2.6]:

$$\forall h \in H_0^1(\mu), \qquad \|h\|^2 \leqslant \frac{1}{\min(R_\beta, \beta)}\sum_{i=0}^{2}\left\|C_i h\right\|^2.$$

where R_β is the Poincaré constant from Assumption 2.1. Denoting by K_1 is the largest eigenvalue of $\mathbf{M}_1$ and by $\kappa = \max(R_\beta^{-1}, \beta^{-1})$, this implies that, for any $\zeta \in (0,1)$ and any $h \in H_0^1(\mu)$,

$$-\left(\!\left(h, \mathcal{L}h\right)\!\right) \geqslant \lambda_{\min}^+\left(\widetilde{\mathbf{M}}_2\right)\left(\zeta\kappa\left\|h\right\|^2 + (1-\zeta)\sum_{i=0}^{2}\left\|C_i h\right\|^2\right) \geqslant \lambda_{\min}^+\left(\widetilde{\mathbf{M}}_2\right)\min\left(\zeta\kappa, \frac{1-\zeta}{K_1}\right)\left(\!\left(h, h\right)\!\right).$$

The optimal choice for ζ is $\zeta = (1+K_1\kappa)^{-1}$, which leads to the following inequality for $h \in H_0^1(\mu)$

(noting that $\mathrm{e}^{t\mathcal{L}}\, h \in H_0^1(\mu)$ for all $t \geqslant 0$ by Theorem 2.1):

$$\frac{1}{2}\,\frac{\mathrm{d}}{\mathrm{d}t}\left((\mathrm{e}^{t\mathcal{L}}\, h, \mathrm{e}^{t\mathcal{L}}\, h)\right) = \left((\mathcal{L}\,\mathrm{e}^{t\mathcal{L}}\, h, \mathrm{e}^{t\mathcal{L}}\, h)\right) \leqslant -\frac{\lambda_{\min}^{+}\left(\widetilde{\mathbf{M}}_2\right)}{K_1 + \kappa^{-1}}\left((\mathrm{e}^{t\mathcal{L}}\, h, \mathrm{e}^{t\mathcal{L}}\, h)\right).$$

Theorem 2.2 then follows from Gronwall's inequality.

3.2 Proof of Theorem 2.1: hypoelliptic regularization

In this section, we prove the hypoelliptic regularization estimate (2.9). For any $h \in L_0^2(\mu)$, we define, analogously to [32, 30, 55],

$$\begin{aligned} N_h(t) = \left\|\mathrm{e}^{t\mathcal{L}}\, h\right\|^2 + a_0 t\left\|\partial_z\,\mathrm{e}^{t\mathcal{L}}\, h\right\|^2 + a_1 t^3\left\|\partial_p\,\mathrm{e}^{t\mathcal{L}}\, h\right\|^2 + a_2 t^5\left\|\partial_q\,\mathrm{e}^{t\mathcal{L}}\, h\right\|^2 \\ - 2b_0 t^2\left\langle \partial_z\,\mathrm{e}^{t\mathcal{L}}\, h, \partial_p\,\mathrm{e}^{t\mathcal{L}}\, h\right\rangle - 2b_1 t^4\left\langle \partial_p\,\mathrm{e}^{t\mathcal{L}}\, h, \partial_q\,\mathrm{e}^{t\mathcal{L}}\, h\right\rangle, \end{aligned} \tag{3.15}$$

where $a_0, a_1, a_2 > 0$ and $b_0, b_1 \geqslant 0$ are small parameters. We calculate, by computations similar to the ones performed in the proof of Lemma 3.1,

$$\frac{1}{2}\,N_h'(t) \leqslant -\frac{1}{\beta\nu^2}\begin{pmatrix}\|C_0\, C_0\, h\|\\ \|C_0\, C_1\, h\|\\ \|C_0\, C_2\, h\|\end{pmatrix}^{\mathsf{T}}\widetilde{\mathbf{M}}_1(t)\begin{pmatrix}\|C_0\, C_0\, h\|\\ \|C_0\, C_1\, h\|\\ \|C_0\, C_2\, h\|\end{pmatrix} - \begin{pmatrix}\|C_0\, h\|\\ \|C_1\, h\|\\ \|C_2\, h\|\end{pmatrix}^{\mathsf{T}}\mathbf{M}_2(t)\begin{pmatrix}\|C_0\, h\|\\ \|C_1\, h\|\\ \|C_2\, h\|\end{pmatrix}, \tag{3.16}$$

where

$$\widetilde{\mathbf{M}}_1(t) = \begin{pmatrix} a_0\, t & -b_0\, t^2 & 0\\ -b_0\, t^2 & a_1\, t^3 & -b_1\, t^4\\ 0 & -b_1\, t^4 & a_2\, t^5\end{pmatrix},$$

and, employing again the notation $r = \gamma^{1/2}\,\nu^{-1}$,

$$\begin{aligned}\mathbf{M}_2(t) =& \begin{pmatrix}\left(\beta^{-1} + a_0 t\right)\nu^{-2} - b_0 r t^2 & -a_0 r t - a_1 r t^3 - b_0\nu^{-2} t^2 & -b_0 t^2 - b_1 r t^4\\ 0 & b_0 r t^2 - b_1 t^4\,\|V''\|_\infty & -a_1 t^3 - a_2 t^5\,\|V''\|_\infty\\ 0 & 0 & b_1 t^4\end{pmatrix}\\ &+\begin{pmatrix}-\dfrac{a_0}{2} & -2b_0 t & 0\\ 0 & -\dfrac{3a_1}{2}t^2 & -4b_1 t^3\\ 0 & 0 & -\dfrac{5a_2}{2}t^4\end{pmatrix}\end{aligned} \tag{3.17}$$

Note that $\widetilde{\mathbf{M}}_1(1) = \mathbf{M}_1$ and, for $t = 1$ also, the first matrix on the right-hand side of (3.17) coincides with the matrix on the right-hand side of (3.14). The matrix $\widetilde{\mathbf{M}}_1(t)$ is clearly positive semidefinite for any $t \in [0,1]$ if $\mathbf{M}_1$ is positive semidefinite, which can be viewed from the factorization

$$\widetilde{\mathbf{M}}_1(t) = \begin{pmatrix} t^{1/2} & 0 & 0\\ 0 & t^{3/2} & 0\\ 0 & 0 & t^{5/2}\end{pmatrix}\mathbf{M}_1\begin{pmatrix} t^{1/2} & 0 & 0\\ 0 & t^{3/2} & 0\\ 0 & 0 & t^{5/2}\end{pmatrix}.$$

We also notice that, for any $t \in [0,1]$,

$$\begin{pmatrix} 1 & 0 & 0 \\ 0 & t^{-1} & 0 \\ 0 & 0 & t^{-2} \end{pmatrix} \mathbf{M}_2(t) \begin{pmatrix} 1 & 0 & 0 \\ 0 & t^{-1} & 0 \\ 0 & 0 & t^{-2} \end{pmatrix} \succcurlyeq_+ \mathbf{M}_2,$$

with

$$\mathbf{M}_2 = \begin{pmatrix} \frac{1}{\beta\nu^2} - b_0 r - \frac{a_0}{2} & -a_0 r - a_1 r - \left(2+\nu^{-2}\right) b_0 & -b_0 - b_1 r \\ 0 & b_0 r - \frac{3a_1}{2} - b_1 \left\|V''\right\|_\infty & -4b_1 - a_1 - a_2 \left\|V''\right\|_\infty \\ 0 & 0 & b_1 - \frac{5a_2}{2} \end{pmatrix}. \tag{3.18}$$

The following key result shows that, for an appropriate choice of the parameters $(a_0, a_1, a_2, b_0, b_1)$, the matrix $\mathbf{M}_2$ is bounded from below, in the sense of $\succcurlyeq_+$, by the identity matrix multiplied by a positive prefactor.

Proposition 3.2. *There exist parameters $(a_0, a_1, a_2, b_0, b_1)$, as well as a constant $C > 0$ (independent of γ and ν) and $\alpha(\gamma, \nu) > 0$, such that $\alpha(\gamma,\nu)\mathbf{I}_3 \preccurlyeq \mathbf{M}_1 \preccurlyeq \mathbf{I}_3$ and*

$$\mathbf{M}_2 \succcurlyeq_+ C \min\left(r^2\nu^2, \frac{1}{r^2\nu^2}, \frac{r^2}{\nu^2}\right) \mathbf{I}_3 = C \min\left(\gamma, \frac{1}{\gamma}, \frac{\gamma}{\nu^4}\right) \mathbf{I}_3.$$

Observe that $\widetilde{\mathbf{M}}_2 \succcurlyeq_+ \mathbf{M}_2$, where $\widetilde{\mathbf{M}}_2$ is the matrix defined in (3.7), so the lower bound on $\mathbf{M}_2$ in Proposition 3.2 implies Proposition 3.1 as a byproduct. Proposition 3.2 also implies that $N_h'(t) \leqslant 0$ for any $t \in [0,1]$, by (3.16). This leads to the inequality

$$N_h(1) \leqslant N_h(0),$$

which is precisely the hypoelliptic regularization inequality (2.9).

Proof. Inspecting the entries of $\mathbf{M}_2$, we notice that, since $a_0, a_1, a_2 \geqslant 0$ always appear with a negative sign,

$$\mathbf{M}_2^+ := \begin{pmatrix} \frac{1}{\beta\nu^2} & -\left(2+\nu^{-2}\right) b_0 & -b_0 - b_1 r \\ 0 & b_0 r & -4b_1 \\ 0 & 0 & b_1 \end{pmatrix} \succcurlyeq_+ \mathbf{M}_2. \tag{3.19}$$

Therefore, any bound from below for $\mathbf{M}_2$ is necessarily a bound from below also for $\mathbf{M}_2^+$. By examining the latter matrix, we obtain tentative scalings for the coefficients b_0 and b_1. In a second step, we show that these scalings are in fact also suitable for $\mathbf{M}_2$.

Step 1: bound from below on $\mathbf{M}_2^+$. In order to obtain a bound from below on $\lambda_{\min}^+\left(\mathbf{M}_2^+\right)$, we consider vectors $\mathbf{v}$ in (3.13) which have two non-zero elements. A necessary condition for the positivity of $\lambda_{\min}^+\left(\mathbf{M}_2^+\right)$ is that the determinants of the following 2×2 symmetrized submatrices

are positive:

$$\mathbf{M}_2^{+,i,j} := \begin{pmatrix} \left[M_2^+\right]_{i,i} & \frac{1}{2}\left[M_2^+\right]_{i,j} \\ \frac{1}{2}\left[M_2^+\right]_{j,i} & \left[M_2^+\right]_{j,j} \end{pmatrix}, \qquad 1 \leqslant i < j \leqslant 3.$$

This leads to the following conditions:

$$\frac{b_0 r}{\beta\nu^2} - b_0^2\left(1 + \frac{1}{2\nu^2}\right)^2 > 0 \quad \Longrightarrow \quad b_0 < \frac{r}{\beta}\min(4\nu^2, \nu^{-2}), \tag{3.20a}$$

$$\frac{b_1}{\beta\nu^2} - \frac{1}{4}\left(b_0 + b_1 r\right)^2 > 0 \quad \Longrightarrow \quad \begin{cases} b_1 < \dfrac{1}{\beta r^2 \nu^2}, \\ b_0^2 < \dfrac{4b_1}{\beta\nu^2}, \end{cases} \tag{3.20b}$$

$$b_0 b_1 r - 4b_1^2 > 0 \quad \Longleftrightarrow \quad 0 < b_1 < \frac{b_0 r}{4}. \tag{3.20c}$$

Equation (3.20a) shows that b_0 is at most of order $\min(r\nu^2, r\nu^{-2})$, so that, from (3.20b) and with (3.20c), b_1 is at most of order

$$m(r,\nu) := \min(r^2\nu^2, r^{-2}\nu^{-2}, r^2\nu^{-2}) = \min\left(\gamma, \frac{1}{\gamma}, \frac{\gamma}{\nu^4}\right).$$

Note the following inequalities, which will prove useful later on:

$$m(r,\nu) \leqslant \min(r^2\nu^2, r^{-2}\nu^{-2}) \leqslant 1, \tag{3.21a}$$

$$m(r,\nu) \leqslant \min(r^2\nu^2, r^2\nu^{-2}) \leqslant r^2, \tag{3.21b}$$

$$m(r,\nu) \leqslant \min(r^{-2}\nu^{-2}, r^2\nu^{-2}) \leqslant \nu^{-2}. \tag{3.21c}$$

Condition (3.20c) suggests that b_0 is of order $r^{-1}m(r,\nu)$. We therefore consider the choice

$$b_0 = Ar^{-1}m(r,\nu), \qquad b_1 = Bm(r,\nu), \tag{3.22}$$

with $A, B > 0$ yet to be chosen. The matrix $\mathbf{M}_2^+$ then reads

$$\mathbf{M}_2^+ = m(r,\nu)\begin{pmatrix} \beta^{-1}\nu^{-2}m(r,\nu)^{-1} & -\left(2+\nu^{-2}\right)Ar^{-1} & -Ar^{-1} - Br \\ 0 & A & -4B \\ 0 & 0 & B \end{pmatrix}.$$

Now let

$$\mathbf{U} := \begin{pmatrix} \beta^{1/2}\nu & 0 & 0 \\ 0 & A^{-1/2}m^{-1/2} & 0 \\ 0 & 0 & B^{-1/2}m^{-1/2} \end{pmatrix} \tag{3.23}$$

and observe that

$$\mathbf{U}\mathbf{M}_2^+\mathbf{U} = \begin{pmatrix} 1 & -\left(2\nu + \nu^{-1}\right)r^{-1}\sqrt{\beta m}\sqrt{A} & -\sqrt{\beta}\left[AB^{-1/2}(m^{1/2}r^{-1}\nu) + B^{1/2}(m^{1/2}r\nu)\right] \\ 0 & 1 & -4A^{-1/2}B^{1/2} \\ 0 & 0 & 1 \end{pmatrix},$$

where, to simplify the notation, we omitted the dependence of m on r, ν. By definition of $m(r,\nu)$, it holds $m^{1/2}r^{-1}\nu \leqslant 1$, $m^{1/2}r\nu \leqslant 1$ and $\left(2\nu+\nu^{-1}\right)r^{-1}\sqrt{m} \leqslant 3$. We moreover choose $B = A^\delta$, which leads to

$$\mathbf{U}\mathbf{M}_2^+\mathbf{U} \succcurlyeq_+ \begin{pmatrix} 1 & -3\sqrt{\beta A} & -\sqrt{\beta}\left[A^{1-\delta/2}+A^{\delta/2}\right] \\ 0 & 1 & -4A^{(\delta-1)/2} \\ 0 & 0 & 1 \end{pmatrix}.$$

This shows that it is possible to choose $1 < \delta < 2$ and $A \in (0,1]$ sufficiently small so that $\lambda_{\min}^+\left(\mathbf{U}\mathbf{M}_2^+\mathbf{U}\right) \geqslant 1/2$. To conclude this step, notice that, for any $\mathbf{x} \geqslant 0$,

$$\mathbf{x}^\mathsf{T}\mathbf{M}_2^+\mathbf{x} = \mathbf{x}^\mathsf{T}\mathbf{U}^{-1}(\mathbf{U}\mathbf{M}_2^+\mathbf{U})\mathbf{U}^{-1}\mathbf{x} \geqslant \frac{1}{2}\left\|\mathbf{U}^{-1}\mathbf{x}\right\|^2 \geqslant \frac{A^\delta}{2}m(r,\nu)\left\|\mathbf{x}\right\|^2.$$

which shows that $\mathbf{M}_2^+ \succcurlyeq_+ \frac{A^\delta}{2}m(r,\nu)\mathbf{I}_3$.

Step 2: Bound from below on $\mathbf{M}_2$. We now consider the matrix $\mathbf{M}_1$ in (3.5) to be of the form

$$\mathbf{M}_1 = x\begin{pmatrix} 2y^{-1} & -1 & 0 \\ -1 & y & 0 \\ 0 & 0 & 0 \end{pmatrix} + w\begin{pmatrix} 0 & 0 & 0 \\ 0 & 2z^{-1} & -1 \\ 0 & -1 & z \end{pmatrix},$$

where x, y, w, z are new positive parameters. This corresponds to setting $b_0 = x$, $b_1 = w$ and $a_0 = 2x/y$, $a_1 = xy + 2w/z$, $a_2 = zw$. The interest of this parametrization is to bring the number of parameters down from 5 to 4 and to directly ensure that $\mathbf{M}_1$ is positive definite. Motivated by the scalings (3.22), we choose x to be of order $r^{-1}\,m(r,\nu)$ and w to be of order $m(r,\nu)$. To guess the scalings of y and z with respect to r and ν, we rely on the following observations:

- to ensure that the $(2,2)$ entry of $\mathbf{M}_2$, which reads $b_0 r - \frac{3a_1}{2} - b_1\left\|V''\right\|_\infty$, is positive for all values of r and ν, it is necessary that a_1 is at most of order $b_0\,r$, which suggests that y scales as r;
- the coefficient a_2 appears only in matrix entries where b_1 is also present and, in these entries, both coefficients appear with prefactors that scale identically with respect to r and ν. This suggests that z is of order 1.

Guided by these observations, we consider the following choice:

$$x = A\,r^{-1}\,m(r,\nu), \qquad y = A^\eta\,r, \qquad w = A^\delta\,m(r,\nu), \qquad z = A^\rho,$$

where η, δ, ρ are exponents independent of r and ν yet to be determined, while $A > 0$ is a small parameter. With the same matrix $\mathbf{U}$ as in (3.23) with $B = A^\delta$, we obtain $\mathbf{U}\mathbf{M}_2\mathbf{U} \succcurlyeq_+ \mathbf{I}_3 + \mathbf{R}$,

where $\mathbf{R}$ is an upper diagonal matrix with entries

$$
\begin{aligned}
[\mathbf{R}]_{1,1} &= -\beta \left(A + A^{1-\eta}\right), \\
[\mathbf{R}]_{1,2} &= -\sqrt{\beta} \left(2A^{1/2-\eta} + A^{1/2+\eta} + 2A^{\delta-\rho-1/2} + 3A^{1/2}\right), \\
[\mathbf{R}]_{1,3} &= -\sqrt{\beta} \left(A^{1-\delta/2} + A^{\delta/2}\right), \\
[\mathbf{R}]_{2,2} &= -\frac{3A^{\eta}}{2} - 3A^{\delta-\rho-1} - A^{\delta-1} \left\|V''\right\|_{\infty}, \\
[\mathbf{R}]_{2,3} &= -4A^{(\delta-1)/2} - A^{\eta+(1-\delta)/2} - 2A^{(\delta-1)/2-\rho} - A^{\rho+(\delta-1)/2} \left\|V''\right\|_{\infty}, \\
[\mathbf{R}]_{3,3} &= -\frac{5A^{\rho}}{2}.
\end{aligned}
$$

In order for all entries of $\mathbf{R}$ to converge to 0 as $A \to 0$, we require $0 < \eta < 1/2$, $1 < \delta < 1 + 2\eta$ and $0 < \rho < (\delta - 1)/2$. By the same reasoning at the one allowing to conclude Step 1, we can show that there exist $A > 0$ sufficiently small and a constant $C > 0$ (which depends on A) for which $\mathbf{M}_2 \succcurlyeq_+ Cm(r, \nu)\mathbf{I}_3$.

Finally, it is easy to see that the smallest eigenvalue $\alpha(\gamma, \nu)$ of the real symmetric matrix $\mathbf{M}_1$ is positive and scales as $\min(1, r^{-2})m(r, \nu)$. Upon decreasing A if necessary, we can further ensure that $\mathbf{M}_1 \preccurlyeq \mathbf{I}_3$. □

Remark 3.4. *Since it holds $\widetilde{\mathbf{M}}_2 \succcurlyeq_+ \mathbf{M}_2$, one might wonder whether a sharper lower bound in Proposition 3.1 could have been obtained by working directly with $\widetilde{\mathbf{M}}_2$. While it is possible that a larger constant C on the right-hand side could have been obtained, the dependence of the lower bound on γ and ν in Proposition 3.1 is in fact sharp. Indeed, better scalings with respect to γ and ν in Proposition 3.1 would lead to better scalings in Theorem 2.2 and therefore also in (2.11), but we show in Appendix A that the scalings of the decay rate in (2.11) are optimal.*

4 Scaling limits of the effective diffusion coefficient

We study in this section various limits for the effective diffusion coefficient (2.6a) of GLE, namely the short memory limit $\nu \to 0$ in Section 4.1 (for which we expect to recover the behavior of standard Langevin dynamics), the overdamped limit $\gamma \to \infty$ in Section 4.2 (for which we expect to recover the behavior of overdamped Langevin dynamics), and finally the underdamped limit $\gamma \to 0$ in Section 4.3 (for which we expect the effective diffusion coefficient to scale as γ^{-1}, as for standard Langevin dynamics in the same limit).

All the analysis in this section is done for GLE in a *periodic* potential on the domain $\mathcal{X} = \mathbf{T}$. The general strategy is the following:

(i) we first formally approximate the solution to the Poisson equation (2.6b) by some function $\widehat{\phi}$ obtained by an asymptotic analysis where the solution ϕ is expanded in powers of a small parameter;

(ii) we next rely on the the resolvent estimates (2.12) to provide bounds on $\left\|\phi - \widehat{\phi}\right\|$;

(iii) we finally deduce the leading order behavior of the diffusion coefficient by replacing ϕ by $\widehat{\phi}$ in (2.6a).

Let us also already emphasize that, while the results presented in Sections 4.1 and 4.2 are mathematically rigorous, the discussion in Section 4.3 is only formal since the asymptotic analysis is quite cumbersome in the underdamped setting, where the leading part of the dynamics is a Hamiltonian evolution.

4.1 The short memory limit

In this section, we show rigorously that, in the limit as $\nu \to 0$ and for

$$\gamma = \lambda^{\mathsf{T}} \mathbf{A}^{-1} \boldsymbol{\lambda} = \int_0^\infty \gamma(t)\, \mathrm{d}t > 0 \tag{4.1}$$

fixed, the effective diffusion coefficient $D_{\gamma,\nu}$ associated with the GLE converges to that associated with the Langevin dynamics for the same value of γ, denoted by D_γ. The latter diffusion coefficient is defined in terms of the solution ϕ_{Lang} of the Poisson equation $-\mathcal{L}_{\mathrm{Lang}}\phi_{\mathrm{Lang}} = p$, where

$$\mathcal{L}_{\mathrm{Lang}} = p\frac{\partial}{\partial q} - V'(q)\frac{\partial}{\partial p} - \gamma p \frac{\partial}{\partial p} + \frac{\gamma}{\beta}\frac{\partial^2}{\partial p},$$

the generator of the Langevin dynamics, acts on functions of (q,p). More precisely, denoting by μ_{Lang} the marginal of the invariant probability measure μ in the (q,p) variables,

$$D_\gamma = \int_{\mathbf{T}\times\mathbf{R}} \phi_{\mathrm{Lang}}(q,p)\, p\, \mu_{\mathrm{Lang}}(\mathrm{d}q\, \mathrm{d}p), \qquad \mu_{\mathrm{Lang}}(\mathrm{d}q\, \mathrm{d}p) = \frac{1}{Z_\beta}\, \mathrm{e}^{-\beta H(q,p)}\, \mathrm{d}q\, \mathrm{d}p, \tag{4.2}$$

where $Z_\beta = \int_{\mathbf{T}\times\mathbf{R}} \mathrm{e}^{-\beta H(q,p)}\, \mathrm{d}q\, \mathrm{d}p$ is the normalization constant. Let us recall that, by the results of [73, 40], the solution ϕ_{Lang} is a smooth function which, together with all its derivatives, grows at most polynomially as $|p| \to \infty$. It particular it belongs to $L^2(\mu_{\mathrm{Lang}})$, so that D_γ is well defined by a Cauchy–Schwarz inequality.

We present the analysis for a general quasi-Markovian approximation of the noise of the form (2.2), with parameters $\boldsymbol{\lambda}$ and $\mathbf{A}$ rescaled in such a way that the correlation time of the noise appears explicitly as a parameter, while keeping γ fixed. More precisely, we rewrite the generator of GLE as (indicating explicitly the dependence on ν)

$$\mathcal{L}_\nu = \mathcal{A}_0 + \frac{1}{\nu}\mathcal{A}_1 + \frac{1}{\nu^2}\mathcal{A}_2,$$

with

$$\mathcal{A}_0 = p\frac{\partial}{\partial q} - V'(q)\frac{\partial}{\partial p}, \qquad \mathcal{A}_1 = \boldsymbol{\lambda}^{\mathsf{T}}\mathbf{z}\frac{\partial}{\partial p} - p\boldsymbol{\lambda}^{\mathsf{T}}\nabla_{\mathbf{z}}, \qquad \mathcal{A}_2 = -\mathbf{z}^{\mathsf{T}}\mathbf{A}^{\mathsf{T}}\nabla_{\mathbf{z}} + \frac{1}{\beta}\mathbf{A} : \nabla_{\mathbf{z}}^2.$$

Note that models GL1 and GL2 are already in this rescaled form. Recall also that, denoting by $(q_t^\nu, p_t^\nu, \mathbf{z}_t^\nu)$ the solution to (2.2), (q_t^ν, p_t^ν) converges in the short memory limit $\nu \to 0$, in the sense of weak convergence of probability measures on $C([0,T], \mathbf{T}\times\mathbf{R})$ for some fixed final time $T > 0$, to the solution of the Langevin equation (1.1b) with friction coefficient (4.1); see for example [55, Theorem 2.6] or [57, Result 8.4].

We have the following result, which can be obtained constructively by formal asymptotics; see [75] for details.

Lemma 4.1. *The function*

$$\widehat{\phi}(q,p,\mathbf{z}) := \phi_{\mathrm{Lang}}(q,p) + \nu \boldsymbol{\lambda}^{\mathsf{T}}\mathbf{A}^{-1}\mathbf{z}\,\frac{\partial \phi_{\mathrm{Lang}}}{\partial p}(q,p) + \nu^2\,\psi_2(q,p,\mathbf{z}) + \nu^3\,\psi_3(q,p,\mathbf{z}),$$

where

$$\begin{aligned}
\psi_2(q,p,\mathbf{z}) &= \frac{1}{2}\left(\left(\boldsymbol{\lambda}^{\mathsf{T}}\mathbf{A}^{-1}\mathbf{z}\right)^2 - \frac{1}{\beta}\boldsymbol{\lambda}^{\mathsf{T}}\mathbf{A}^{-1}\mathbf{A}^{-\mathsf{T}}\boldsymbol{\lambda}\right)\frac{\partial^2\phi_{\mathrm{Lang}}}{\partial p^2}(q,p),\\
\psi_3(q,p,\mathbf{z}) &= \left(\frac{1}{6}\left(\boldsymbol{\lambda}^{\mathsf{T}}\mathbf{A}^{-1}\mathbf{z}\right)^3 + \frac{\gamma}{\beta}\boldsymbol{\lambda}^{\mathsf{T}}\mathbf{A}^{-2}\mathbf{z} - \frac{1}{2\beta}(\boldsymbol{\lambda}^{\mathsf{T}}\mathbf{A}^{-1}\mathbf{z})(\boldsymbol{\lambda}^{\mathsf{T}}\mathbf{A}^{-1}\mathbf{A}^{-\mathsf{T}}\boldsymbol{\lambda})\right)\frac{\partial^3\phi_{\mathrm{Lang}}}{\partial p^3}(q,p)\\
&\quad - \left(\gamma p + V'(q)\right)\boldsymbol{\lambda}^{\mathsf{T}}\mathbf{A}^{-2}\mathbf{z}\frac{\partial^2\phi_{\mathrm{Lang}}}{\partial p^2}(q,p) + p\boldsymbol{\lambda}^{\mathsf{T}}\mathbf{A}^{-2}\mathbf{z}\frac{\partial^2\phi_{\mathrm{Lang}}}{\partial q\partial p}(q,p),
\end{aligned}$$

belongs to $L^2_0(\mu)$ *and satisfies*

$$-\mathcal{L}_\nu\widehat{\phi} = p - \nu^2(\mathcal{A}_0\psi_2 + \mathcal{A}_1\psi_3) - \nu^3\mathcal{A}_0\psi_3. \tag{4.3}$$

Proof. We first compute $\mathcal{A}_2\left(\boldsymbol{\lambda}^{\mathsf{T}}\mathbf{A}^{-1}\mathbf{z}\right) = -\boldsymbol{\lambda}^{\mathsf{T}}\mathbf{z}$ and $\mathcal{A}_2\left(\boldsymbol{\lambda}^{\mathsf{T}}\mathbf{A}^{-2}\mathbf{z}\right) = -\boldsymbol{\lambda}^{\mathsf{T}}\mathbf{A}^{-1}\mathbf{z}$, as well as, for $\alpha = 2, 3$,

$$\mathcal{A}_2\left[\left(\boldsymbol{\lambda}^{\mathsf{T}}\mathbf{A}^{-1}\mathbf{z}\right)^\alpha\right] = -\alpha\left(\boldsymbol{\lambda}^{\mathsf{T}}\mathbf{A}^{-1}\mathbf{z}\right)^{\alpha-1}\boldsymbol{\lambda}^{\mathsf{T}}\mathbf{z} + \frac{\alpha(\alpha-1)\gamma}{\beta}\left(\boldsymbol{\lambda}^{\mathsf{T}}\mathbf{A}^{-1}\mathbf{z}\right)^{\alpha-2}.$$

The result can then be verified directly by calculating $-\mathcal{L}_\nu\widehat{\phi}$ and gathering terms with the same powers of ν. □

We can then provide the convergence result on the effective diffusion coefficient.

Proposition 4.1. *Fix* $\gamma > 0$ *and assume that there exists* $C > 0$ *such that*

$$\forall \nu \in (0,1], \qquad \left\|\mathcal{L}_\nu^{-1}\right\|_{\mathcal{B}(L^2_0(\mu))} \leqslant C. \tag{4.4}$$

Then there exists $R > 0$ *such that, for any* $0 < \nu \leqslant 1$,

$$|D_{\gamma,\nu} - D_\gamma| \leqslant R\nu^2.$$

Note that the condition (4.4) follows for GL1 from the resolvent estimate (2.12), and for GL2 from the resolvent estimate (B.1) in Appendix B. It would be possible to weaken this condition by allowing some power law growth with respect to ν on the right-hand side of (4.4) upon further continuing the asymptotic expansion of Lemma 4.1 in order to have higher order terms in (4.3).

Proof. By the result from [40] mentioned previously, ϕ_{Lang} and all its derivatives are smooth and grow at most polynomially as $|p| \to \infty$. Given the definitions of $\mathcal{A}_1$ and $\mathcal{A}_2$, this implies that the coefficients of ν^2 and ν^3 on the right-hand side of (4.3) are smooth functions in $L^2(\mu)$. Since these functions are independent of ν, there exists $K > 0$ such that

$$\left\|\mathcal{L}_\nu\left(\phi_\nu - \widehat{\phi}\right)\right\| \leqslant K\nu^2,$$

where ϕ_ν denotes the solution to the Poisson equation $-\mathcal{L}_\nu \phi_\nu = p$. We therefore obtain with (4.4) that $\left\|\phi_\nu - \widehat{\phi}\right\| \leqslant CK\nu^2$ for all $\nu \leqslant 1$. Since the function $(q,p,\mathbf{z}) \mapsto \boldsymbol{\lambda}^{\mathsf{T}} \mathbf{A}^{-1} \mathbf{z}\, \partial_p \phi_{\mathrm{Lang}}(q,p)$ has average 0 with respect to μ, the desired estimate follows by substituting ϕ_ν by $\widehat{\phi}$ in (2.6a), integrating over $\mathbf{z}$, and comparing with (4.2). □

Remark 4.1. *For the special case of the model GL1, if we had wanted to show only that $|D_{\gamma,\nu} - D_\gamma| \to 0$ in the limit as $\nu \to 0$ without making precise a convergence rate, we could have proceeded more directly from [55, Theorem 2.6], which can be leveraged by using a reasoning similar to that in the proof of [30, Proposition 3.3]. With the same notation as above, [55, Theorem 2.6] implies that the random variable (q_t^ν, p_t^ν) converges weakly, in the limit as $\nu \to 0$, to the solution of (1.1b) evaluated at t, for any $t \geqslant 0$ and any initial condition with finite moments of all orders. Consequently, for any bounded and continuous function $f(q,p)$, it holds as $\nu \to 0$*

$$\forall (q,p,z) \in \mathbf{T} \times \mathbf{R} \times \mathbf{R}, \qquad \mathrm{e}^{t\mathcal{L}_\nu} f(q,p,z) \to \mathrm{e}^{t\mathcal{L}_{\mathrm{Lang}}} f(q,p),$$

and so also $\mathrm{e}^{t\mathcal{L}_\nu} f \to \mathrm{e}^{t\mathcal{L}_{\mathrm{Lang}}} f$ in $L^2(\mu)$ by dominated convergence. By density of the bounded and continuous functions in $L^2(\mu_{\mathrm{Lang}})$ and the continuity of the propagators on $L^2(\mu_{\mathrm{Lang}})$, this limit holds in fact for any $f \in L^2(\mu_{\mathrm{Lang}})$. Therefore, for any $f \in L^2_0(\mu_{\mathrm{Lang}})$, it holds, as $\nu \to 0$,

$$\begin{aligned}\left\|\mathcal{L}_\nu^{-1} f - \mathcal{L}_{\mathrm{Lang}}^{-1} f\right\| &= \left\|\int_0^\infty \mathrm{e}^{t\mathcal{L}_\nu} f \,\mathrm{d}t - \int_0^\infty \mathrm{e}^{t\mathcal{L}_{\mathrm{Lang}}} f \,\mathrm{d}t\right\| \\ &\leqslant \int_0^\infty \left\|\mathrm{e}^{t\mathcal{L}_\nu} f - \mathrm{e}^{t\mathcal{L}_{\mathrm{Lang}}} f\right\| \mathrm{d}t \xrightarrow[\nu\to 0]{} 0.\end{aligned}$$

The last limit is justified by dominated convergence because, by (2.11) and the corresponding result for the Langevin equation, we have the bound $\left\|\mathrm{e}^{t\mathcal{L}_\nu} f - \mathrm{e}^{t\mathcal{L}_{\mathrm{Lang}}} f\right\| \leqslant \left\|\mathrm{e}^{t\mathcal{L}_\nu} f\right\| + \left\|\mathrm{e}^{t\mathcal{L}_{\mathrm{Lang}}} f\right\| \leqslant C\,\mathrm{e}^{-\lambda t}$, for some positive constants C and λ independent of ν. This concludes the proof since

$$D_{\gamma,\nu} = \left\langle -\mathcal{L}_\nu^{-1} p, p \right\rangle \xrightarrow[\nu\to 0]{} \left\langle -\mathcal{L}_{\mathrm{Lang}}^{-1} p, p \right\rangle = D_\gamma.$$

4.2 The overdamped limit

We prove in this section that the effective diffusion coefficient $D_{\gamma,\nu}$ associated with the model GL1 converges, as the effective friction (4.1) goes to infinity, to the effective diffusion coefficient D_{ovd} associated with the overdamped Langevin dynamics (1.1a). It is certainly possible to extend our analysis to more general models of noise than GL1, but the algebra involved in the asymptotic analysis of Lemma 4.2 below becomes more cumbersome, so we refrain from doing so.

Denoting by $\mu_{\mathrm{ovd}}(\mathrm{d}q)$ the marginal of $\mu(\mathrm{d}q\,\mathrm{d}p\,\mathrm{d}z)$ (which has a density proportional to $\mathrm{e}^{-\beta V(q)}\,\mathrm{d}q$), the effective diffusion coefficient for overdamped dynamics is defined from the unique solution $\phi_{\mathrm{ovd}} \in L^2(\mu_{\mathrm{ovd}})$ to the Poisson equation

$$-\mathcal{L}_{\mathrm{ovd}} \phi_{\mathrm{ovd}} = -V', \qquad \int_{\mathbf{T}} \phi_{\mathrm{ovd}} \,\mathrm{d}\mu_{\mathrm{ovd}} = 0, \tag{4.5}$$

where $\mathcal{L}_{\mathrm{ovd}}$ acts on functions of q as

$$\mathcal{L}_{\mathrm{ovd}} = -V'(q)\frac{\mathrm{d}}{\mathrm{d}q} + \beta^{-1}\frac{\mathrm{d}^2}{\mathrm{d}q^2}.$$

By elliptic regularity, the solution ϕ_{ovd} belongs to $C^\infty(\mathbf{T})$. It can then be shown, using the tools from [7, Chapter 3] (see for instance [24, Section 1.2] and [58, Chapter 13]) that

$$D_{\mathrm{ovd}} = \beta^{-1} + \int_{\mathbf{T}} \phi_{\mathrm{ovd}} V' \, \mathrm{d}\mu_{\mathrm{ovd}}. \tag{4.6}$$

The overdamped limit of the effective diffusion coefficient (4.2) for the Langevin dynamics was already studied in [30] (see also [43, Section 3.1.1]), where it is shown that $D_\gamma = D_{\mathrm{ovd}}/\gamma + \mathcal{O}\left(\gamma^{-3}\right)$. We provide the counterpart of this estimate for GL1 in the following result.

Proposition 4.2. *Consider the model GL1 and recall the definition (4.1) of the effective friction γ. There exists $R > 0$ such that*

$$\forall \gamma \geqslant 1, \qquad \left| D_{\gamma,\nu} - \frac{1}{\gamma} D_{\mathrm{ovd}} \right| \leqslant \frac{R}{\gamma^{3/2}}.$$

In fact, using a more involved asymptotic analysis (see Remark 4.2), it would be possible to show that the difference $D_{\gamma,\nu} - D_{\mathrm{ovd}}/\gamma$ is of order γ^{-2}. In order to perform the asymptotic analysis, we rewrite the generator of the GL1 model as (indicating explicitly the dependence on the friction $\gamma > 0$)

$$\mathcal{L}_\gamma = \sqrt{\gamma}\mathcal{A}_0 + \mathcal{A}_1, \qquad \mathcal{A}_0 = \frac{1}{\nu}\left(z\frac{\partial}{\partial p} - p\frac{\partial}{\partial z}\right), \qquad \mathcal{A}_1 = p\frac{\partial}{\partial q} - V'(q)\frac{\partial}{\partial p} + \frac{1}{\nu^2}\left(-z\frac{\partial}{\partial z} + \frac{1}{\beta}\frac{\partial^2}{\partial z^2}\right).$$

The resolvent estimates (2.12) suggest that the solution to the Poisson equation $-\mathcal{L}_\gamma \phi_\gamma = p$ is of order γ as $\gamma \to \infty$. However, by analogy with asymptotic calculations for the Langevin equation [59] where the leading order term in the series expansion in inverse powers of γ of the solution to the Poisson equation $-\mathcal{L}_{\mathrm{Lang}}\phi_{\mathrm{Lang}} = p$ scales as $\mathcal{O}(1)$ (which can also be seen from the expressions of $\mathcal{L}_{\mathrm{Lang}}$ provided in [43, Theorem 2.5] and [8]), we formally expand ϕ_γ as

$$\phi_\gamma = \overline{\phi}_0 + \gamma^{-1/2}\overline{\phi}_1 + \gamma^{-1}\overline{\phi}_2 + \cdots.$$

The correctness of this assumption can be checked *a posteriori*, using formal asymptotics (the details of which are presented in [75]) to calculate the functions $\overline{\phi}_i$ in a systematic way. This is made precise in the following technical result, where $L^2_0(\mu_{\mathrm{ovd}})$ is the subset of functions of $L^2(\mu_{\mathrm{ovd}})$ with mean 0 with respect to $\mu_{\mathrm{ovd}}(\mathrm{d}q)$. For simplicity of notation we assume that $\nu = 1$.

Lemma 4.2. *Denote by ϕ_γ the solution to the Poisson equation (2.6b) for GL1, and by ψ the unique solution in $L^2_0(\mu_{\mathrm{ovd}})$ of the Poisson equation*

$$\mathcal{L}_{\mathrm{ovd}}\psi = \left|V'\right|^2 \phi''_{\mathrm{ovd}} - \frac{3}{2\beta} V' \phi^{(3)}_{\mathrm{ovd}} - \frac{1}{\beta} V'' \phi''_{\mathrm{ovd}} + \frac{1}{2\beta^2}\phi^{(4)}_{\mathrm{ovd}}. \tag{4.7}$$

Define the function

$$\widehat{\phi} = \phi_{\mathrm{ovd}} + \gamma^{-\frac{1}{2}}\, z(\phi'_{\mathrm{ovd}} + 1) + \gamma^{-1}\left(\frac{z^2}{2}\,\phi''_{\mathrm{ovd}} + p\,(\phi'_{\mathrm{ovd}} + 1) + \psi\right) + \gamma^{-\frac{3}{2}}\Phi_{3/2} + \gamma^{-2}\Phi_2 + \gamma^{-\frac{5}{2}}\Phi_{5/2},$$

with

$$\begin{aligned}
\Phi_{3/2} &= pz\,\phi''_{\mathrm{ovd}} + \frac{z^3}{6}\phi^{(3)}_{\mathrm{ovd}} + z\psi', \\
\Phi_2 &= \left(pV' + \frac{p^2}{2} + \frac{p^2}{2}V'' + \frac{z^2}{2}V''\right)\phi''_{\mathrm{ovd}} + \left(\frac{pz^2}{2} - \frac{p}{\beta} + \frac{p^2}{4}V' + \frac{z^2}{4}V'\right)\phi^{(3)}_{\mathrm{ovd}} \\
&\quad + \left(\frac{z^4}{24} - \frac{p^2}{4\beta} - \frac{z^2}{4\beta}\right)\phi^{(4)}_{\mathrm{ovd}} + p\psi' + \frac{z^2}{2}\psi'',
\end{aligned}$$

and

$$\begin{aligned}
\Phi_{5/2} &= \left(pz\,\psi'' + \frac{z^3}{6}\psi^{(3)}\right) + \left(\frac{p^2 z}{2}\,V^{(3)} + pzV'' + \frac{z^3}{2}\,V^{(3)} - z\,V'V'' - z\,V'\right)\phi''_{\mathrm{ovd}} \\
&\quad + \left(\frac{3p^2 z}{4}\,V'' + \frac{p^2 z}{2} + pz\,V' + \frac{3z^3}{4}V'' - \frac{z}{2}\,\big|V'\big|^2 + \frac{z}{\beta}\right)\phi^{(3)}_{\mathrm{ovd}} \\
&\quad + \left(\frac{p^2 z}{4}\,V' + \frac{pz^3}{6} - \frac{pz}{\beta} + \frac{z^3}{4}\,V' + \frac{zV'}{2\beta}\right)\phi^{(4)}_{\mathrm{ovd}} + \left(-\frac{p^2 z}{4\beta} + \frac{z^5}{120} - \frac{z^3}{4\beta}\right)\phi^{(5)}_{\mathrm{ovd}}.
\end{aligned}$$

Let also $\widehat{\phi}_0 := \widehat{\phi} - \int_{\mathbf{T}\times\mathbf{R}\times\mathbf{R}} \widehat{\phi}\,\mathrm{d}\mu$, *so that* $\widehat{\phi}_0 \in L^2_0(\mu)$. *Then* $\mathcal{R} = \gamma^{\frac{5}{2}}\mathcal{L}_\gamma\left(\widehat{\phi}_0 - \phi_\gamma\right)$ *is a well defined function of* $L^2_0(\mu)$ *which is independent of* $\gamma > 0$.

Including the term multiplying $\gamma^{\frac{5}{2}}$ adds a significant amount of complexity, but this is required for rigorously proving the convergence of $\gamma D_{\gamma,\nu}$ to D_{ovd} in the limit as $\gamma \to \infty$.

Proof. Note first that (4.7) admits a solution because the right-hand side belongs to $L^2_0(\mu_{\mathrm{ovd}})$, which can be shown by using integration by part on the last term:

$$\int_{\mathbf{T}} \frac{\phi^{(4)}_{\mathrm{ovd}}}{2\beta^2}\,\mathrm{d}\mu_{\mathrm{ovd}} = \int_{\mathbf{T}} \frac{V'\,\phi^{(3)}_{\mathrm{ovd}}}{2\beta}\,\mathrm{d}\mu_{\mathrm{ovd}} = \int_{\mathbf{T}} \frac{3}{2\beta}\,V'\,\phi^{(3)}_{\mathrm{ovd}}\,\mathrm{d}\mu_{\mathrm{ovd}} - \int_{\mathbf{T}} \left(\big|V'\big|^2 - \frac{V''}{\beta}\right)\phi''_{\mathrm{ovd}}\,\mathrm{d}\mu_{\mathrm{ovd}}.$$

A straightforward computation shows that $\mathcal{L}_\gamma\left(\widehat{\phi}_0 - \phi_\gamma\right) = \gamma^{-5/2}\mathcal{A}_1\Phi_{5/2}$. The function $\mathcal{R}$ is in $L^2(\mu)$ by direct inspection, since ϕ_{ovd} and ψ are smooth and defined on a compact domain. Finally, $\mathcal{R}$ has average 0 with respect to μ since it is in the image of $\mathcal{A}_1$, which is the sum of two generators of stochastic dynamics leaving μ invariant. □

We can conclude this section with the proof of Proposition 4.2.

Proof of Proposition 4.2. Note first that, by an integration by parts in (4.6),

$$D_{\mathrm{ovd}} = \beta^{-1}\left(1 + \int \phi'_{\mathrm{ovd}}\,\mathrm{d}\mu\right). \tag{4.8}$$

The resolvent estimate (2.12) and Lemma 4.2 next imply that

$$\left\|\widehat{\phi}_0 - \phi_\gamma\right\| \leqslant C\,\|\mathcal{R}\|\,\gamma^{-\frac{3}{2}}. \tag{4.9}$$

Moreover, using that ϕ_{ovd} and ψ have average 0 with respect to μ, and that functions with odd powers of p and z also have vanishing averages with respect to μ,

$$\begin{aligned}\int_{\mathbf{T}\times\mathbf{R}\times\mathbf{R}} \widehat{\phi}_0 p \,\mathrm{d}\mu &= \int_{\mathbf{T}\times\mathbf{R}\times\mathbf{R}} \widehat{\phi} p \,\mathrm{d}\mu \\ &= \frac{1}{\beta\gamma}\int_{\mathbf{T}\times\mathbf{R}\times\mathbf{R}} \left(1+\phi'_{\mathrm{ovd}}\right)\mathrm{d}\mu + \frac{1}{\gamma^2}\int_{\mathbf{T}\times\mathbf{R}\times\mathbf{R}} \left(\Phi_2 + \gamma^{-1/2}\Phi_{5/2}\right) p \,\mathrm{d}\mu \end{aligned} \tag{4.10}$$

The result then follows by combining the previous equality with (4.8) and (4.9). □

Remark 4.2. *Note that the term of order $\gamma^{-3/2}$ vanishes in the effective diffusion coefficient (4.10). We therefore expect that the bound in Proposition 4.2 can be improved to γ^{-2}, as for Langevin dynamics. There is no conceptual obstruction to this end, but this would require going to an extra order in Lemma 4.2 by making explicit the term $\gamma^{-3}\Phi_3$ and changing $\Phi_{5/2}$ in the definition of $\widehat{\phi}$, which is algebraically cumbersome.*

4.3 The underdamped limit

We consider in this section the underdamped limit $\gamma \to 0$ for the GL1 model. We will assume without loss of generality that $\min_{q\in[-\pi,\pi]} V(q) = 0$. The underdamped limit of Langevin dynamics in one dimension was carefully studied in [30], see also [59], where it is shown that the effective diffusion coefficient associated with the Langevin dynamics (1.1b), multiplied by the factor γ, converges in the limit as $\gamma \to 0$ to

$$D_{\mathrm{und}} = \frac{8\pi^2}{\beta Z_\beta}\int_{E_0}^{\infty} \frac{\mathrm{e}^{-\beta E}}{S_{\mathrm{und}}(E)}\,\mathrm{d}E, \qquad Z_\beta = \int_{\mathbf{T}\times\mathbf{R}} \mathrm{e}^{-\beta H} = \sqrt{\frac{2\pi}{\beta}}\int_{\mathbf{T}} \mathrm{e}^{-\beta V}, \tag{4.11}$$

where $E_0 = \max_{q\in[-\pi,\pi]} V(q)$, and

$$S_{\mathrm{und}}(E) = \int_{-\pi}^{\pi} P(q,E)\,\mathrm{d}q, \qquad P(q,E) := \sqrt{2(E-V(q))}. \tag{4.12}$$

In this section, we show that a similar result holds for the underdamped limit of GL1. In particular, we motivate the following result.

Result 4.1. *In the limit as $\gamma \to 0$, the effective diffusion coefficient $D_{\gamma,\nu}$ for GL1 scales as D^*_ν/γ in the limit $\gamma \to 0$, for some limiting coefficient D^*_ν. More precisely, it holds*

$$|\gamma D_{\gamma,\nu} - D^*_\nu| \xrightarrow[\nu\to 0]{} 0.$$

The effective diffusion coefficient is given by

$$D^*_\nu = \frac{8\pi^2}{\beta Z_\beta}\int_{E_0}^{\infty} \frac{\nu^2\,\mathrm{e}^{-\beta\mathcal{E}}}{S_\nu(\mathcal{E})}\,\mathrm{d}\mathcal{E}. \tag{4.13}$$

Here

$$S_\nu(E) = \int_{\mathbf{T}} s_\nu(q,E)\,\mathrm{d}q, \tag{4.14}$$

with $s_\nu(\cdot\,,E)$, for $E > E_0$, the unique smooth periodic solution to the following first order

differential equation (see Lemma C.1 in Appendix C):

$$\forall q \in \mathbf{T}, \qquad P(q,E)\frac{\partial s_\nu}{\partial q}(q,E) - \frac{1}{\nu^2} s_\nu(q,E) = -P(q,E). \tag{4.15}$$

Remark 4.3. *We use here the environment* Result*, as opposed to* Proposition *or* Theorem*, because our proof of the result relies on formal asymptotics, and additional work would be required to turn these asymptotics into a rigorous proof.*

The derivation of this result using formal asymptotics is presented at the end of this section. Note that the integral on the right-hand side of (4.13) is well defined since

$$\forall q \in \mathbf{T}, \qquad \nu^2\sqrt{2(E-E_0)} \leqslant \nu^2 \inf_{q\in\mathbf{T}} P(q,E) \leqslant s_\nu(q,E) \leqslant \nu^2 \sup_{q\in\mathbf{T}} P(q,E) = \nu^2\sqrt{2E}. \tag{4.16}$$

The latter inequalities are obtained from the ordinary differential equation (ODE) (4.15) satisfied by s_ν, since $\partial_q s_\nu(q,E)$ vanishes at the extrema of $q \mapsto s_\nu(q,E)$. The relationship between D_ν^* and the diffusion coefficient D_{und} given by (4.11) is made precise in the following result.

Proposition 4.3. *There exists $C > 0$ depending only on V such that*

$$\forall \nu > 0, \qquad |D_\nu^* - D_{\mathrm{und}}| \leqslant C\nu^4(1+\nu^2). \tag{4.17}$$

Proof. If the potential V is constant, then $s_\nu(q,E) = \nu^2 P(q,E)$, so $\nu^{-2}S_\nu(E) = S_{\mathrm{und}}(E)$ and, consequently, $D_\nu^* = D_{\mathrm{und}}$. If V is not constant, then by (C.2) below, it holds

$$\begin{aligned}\left\|\nu^{-2}s_\nu(\,\cdot\,,E) - P(\,\cdot\,,E) - \nu^2\,P(\,\cdot\,,E)\partial_q P(\,\cdot\,,E)\right\|_\infty \\ \leqslant \nu^4\left\|P(\,\cdot\,,E)\partial_q\left[P(\,\cdot\,,E)\partial_q P(\,\cdot\,,E)\right]\right\|_\infty = \nu^4\left\|P(\,\cdot\,,E)V''\right\|_\infty.\end{aligned}$$

Since $P(\,\cdot\,,E)\partial_q P(\,\cdot\,,E)$ integrates to zero over $\mathbf{T}$ and $0 \leqslant P(\,\cdot\,,E) \leqslant \sqrt{2E}$ by (4.12), there exists $K > 0$ independent of ν and E such that

$$\forall E > E_0, \quad \forall \nu > 0, \qquad \left|\nu^{-2}S_\nu(E) - S_{\mathrm{und}}(E)\right| \leqslant K\nu^4\sqrt{E},$$

in view of the definition (4.12) of S_{und}. By (C.4) in the appendix (where we use the assumption that V is not constant) and the fact that $S_{\mathrm{und}}(E)$ is an increasing function of E with $S_{\mathrm{und}}(E_0) > 0$, it holds

$$\begin{aligned}\forall E > E_0, \quad \forall \nu > 0, \qquad \left|\frac{\nu^2}{S_\nu(E)} - \frac{1}{S_{\mathrm{und}}(E)}\right| &= \left|\frac{\nu^2}{S_\nu(E)\,S_{\mathrm{und}}(E)}\right| \left|\nu^{-2}S_\nu(E) - S_{\mathrm{und}}(E)\right| \\ &\leqslant \frac{1+M\nu^2}{\left|S_{\mathrm{und}}(E_0)\right|^2} K\nu^4\sqrt{E},\end{aligned}$$

which, given the definitions (4.11) and (4.13), shows (4.17). □

We conclude this section by first presenting some numerical results confirming that (4.17) holds and then motivating Result 4.1.

4.3.1 Numerical illustration

In order to estimate D_ν^* for $\nu > 0$, the last integral in (4.13) can be truncated and approximated by numerical quadrature. This requires to numerically approximate the integrand, in particular the term $S_\nu(E)$, for discrete values of E. To this end, we employ the function `solve_bvp` from the *SciPy* module `scipy.integrate`. This function implements a solver for boundary value problems (BVP) using the approach proposed in [39], and we employ it in order to calculate an approximation $\hat{\mathbf{y}} = (\hat{y}_1, \hat{y}_2)^\mathsf{T}$ of the solution to the following first order system of ODEs:

$$\mathbf{y}'(q) = \begin{pmatrix} \dfrac{y_1(q)}{\nu^2\, P(q,E)} - 1 \\ y_1(q) \end{pmatrix} =: \mathbf{f}(q), \qquad -\pi \leqslant q \leqslant \pi,$$

subject to the boundary condition

$$\mathbf{g}(\mathbf{y}(-\pi), \mathbf{y}(\pi)) := \begin{pmatrix} y_1(-\pi) - y_1(\pi) \\ y_2(-\pi) \end{pmatrix} = \mathbf{0}.$$

The first line in the ODE for $\mathbf{y}$ is (4.15) divided by P, while the second one corresponds to (4.14). Since, given $E > E_0$, the unique exact solution of this problem is given by

$$\mathbf{y}(q) = \begin{pmatrix} s_\nu(q,E) \\ \displaystyle\int_{-\pi}^{q} s_\nu(Q,E)\,\mathrm{d}Q \end{pmatrix},$$

an approximation of $S_\nu(E)$ is obtained by simply evaluating $\hat{y}_2(\pi)$.

Results from the numerical simulation are illustrated in Fig. 6 for the case of the cosine potential $V(q) = \frac{1}{2}(1 - \cos q)$, which was also employed in Section 2.3. Recall that it is difficult to compare D_ν^* with (2.6a) since the spectral method we use in Section 2.3 cannot tackle values of γ smaller than 0.01, and such values are not sufficiently small to see the asymptotic regime $\gamma \to 0$ (as evidenced by Figure 3, Right).

To generate the results in Fig. 6, the integral in (4.13) was truncated at $E = 25$ and approximated using the *SciPy* function `scipy.integrate.quad` with a relative tolerance equal to 10^{-12}. The tolerance used in `scipy.integrate.solve_bvp` was equal to 10^{-11}. The limiting coefficient D_{und} was computed by truncating and approximating the integral in (4.11) with the same parameters, and by using the explicit expression of $S(E)$ derived in [59].

We observe that, although D_ν^* does vary with ν, the relative variation is very small ($< 5\%$ over the interval $\nu \in [0, 1]$). We also notice that $D_\nu^* \to D_{\text{und}}$ as $\nu \to 0$, as expected from (4.17). In fact, the difference $D_\nu^* - D_{\text{und}}$ clearly scales as ν^4 in the limit $\nu \to 0$, consistently with (4.17).

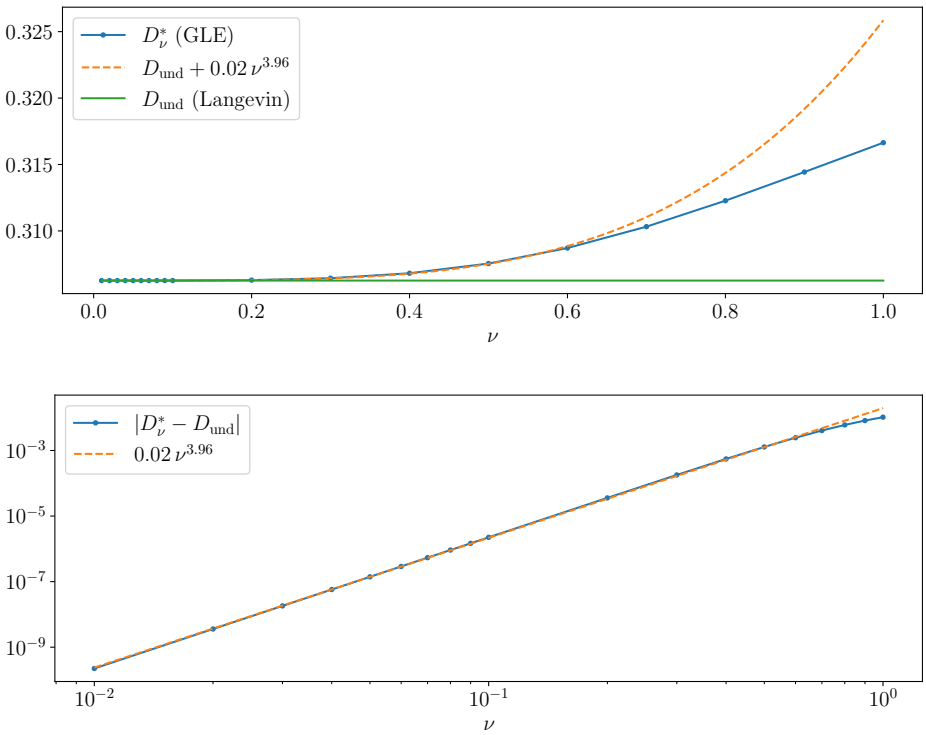


Figure 6: Comparison between the effective diffusion coefficients (multiplied by γ) for the Langevin and the generalized Langevin dynamics in the underdamped limit, in linear (top) and logarithmic (bottom) scales.

4.3.2 Formal derivation of Result 4.1

In order to formally obtain Result 4.1, we rewrite the generator as $\mathcal{L}_\gamma = \mathcal{A}_0 + \sqrt{\gamma}\mathcal{A}_1$ (indicating explicitly the dependence on the friction $\gamma > 0$), with

$$\mathcal{A}_0 = p\frac{\partial}{\partial q} - V'(q)\frac{\partial}{\partial p} + \frac{1}{\nu^2}\left(-z\frac{\partial}{\partial z} + \frac{1}{\beta}\frac{\partial^2}{\partial z^2}\right), \qquad \mathcal{A}_1 = \frac{1}{\nu}\left(z\frac{\partial}{\partial p} - p\frac{\partial}{\partial z}\right).$$

We next consider the following ansatz on the solution ϕ_γ to the Poisson equation (2.6b), motivated by the fact that the leading order of the resolvent is γ^{-1} by (2.12):

$$\phi_\gamma = \frac{1}{\gamma}\overline{\phi}_0 + \frac{1}{\sqrt{\gamma}}\overline{\phi}_1 + \overline{\phi}_2 + \cdots$$

By substituting into the Poisson equation (2.6b) and successively identifying terms of order $\gamma^{-1}, \gamma^{-1/2}, 1, \cdots$, we obtain

$$\mathcal{A}_0\overline{\phi}_0 = 0, \tag{4.18a}$$

$$\mathcal{A}_0\overline{\phi}_1 + \mathcal{A}_1\overline{\phi}_0 = 0, \tag{4.18b}$$

$$\mathcal{A}_0\overline{\phi}_2 + \mathcal{A}_1\overline{\phi}_1 = -p; \tag{4.18c}$$

while $\mathcal{A}_0\overline{\phi}_{i+1} + \mathcal{A}_1\overline{\phi}_i = 0$ for $i \geqslant 2$. We motivate below that

$$\overline{\phi}_0(q,p,z) = \operatorname{sign}(p)\,\psi_0(H(q,p)), \tag{4.19}$$

where $\operatorname{sign}(p) = \mathbb{1}_{[0,\infty)}(p) - \mathbb{1}_{(-\infty,0]}(p)$ and

$$\psi_0(E) = \mathbb{1}_{[E_0,\infty)}(E)\,2\pi\nu^2 \int_{E_0}^{E} \frac{1}{S_\nu(\mathcal{E})}\,\mathrm{d}\mathcal{E}. \tag{4.20}$$

Unfortunately, ψ_0' is discontinuous at $E = E_0$, so $\mathcal{L}_\gamma\overline{\phi}_1$ does not make sense as a function. This invalidates a posteriori the assumed asymptotic expansion (4.16). Despite the breakdown of the naive expansion, which was also noted in [59] and [30] for the Langevin equation in the underdamped regime, we assume that $\gamma^{-1}\overline{\phi}_0$ still captures ϕ_γ at dominant order. For the Langevin equation, the correctness of the dominant term in the naive expansion can be shown rigorously based on results by Freidlin and Weber [25], but showing this for GL1 is an open problem that would probably require considerable additional work.

From the above discussion, the effective diffusion coefficient $D_{\gamma,\nu}$ should scale at dominant order as $\gamma^{-1}D_\nu^*$ with

$$\begin{aligned}
D_\nu^* &= \int_{\mathbf{T}\times\mathbf{R}\times\mathbf{R}} \overline{\phi}_0\, p\,\mathrm{d}\mu = \frac{1}{Z_\beta}\int_{\mathbf{T}\times\mathbf{R}} \overline{\phi}_0(q,p) p\, \mathrm{e}^{-\beta H(q,p)}\,\mathrm{d}q\,\mathrm{d}p = \frac{2}{Z_\beta}\int_{E_0}^{\infty}\int_{\mathbf{T}} \psi_0(E)\,\mathrm{e}^{-\beta E}\,\mathrm{d}q\,\mathrm{d}E \\
&= \frac{4\pi}{Z_\beta}\int_{E_0}^{\infty} \psi_0(E)\,\mathrm{e}^{-\beta E}\,\mathrm{d}E = \frac{8\pi^2}{Z_\beta}\int_{E_0}^{\infty}\int_{E_0}^{E} \frac{\nu^2\,\mathrm{e}^{-\beta E}}{S_\nu(\mathcal{E})}\,\mathrm{d}\mathcal{E}\,\mathrm{d}E \\
&= \frac{8\pi^2}{Z_\beta}\int_{E_0}^{\infty} \frac{\nu^2}{S_\nu(\mathcal{E})}\left(\int_{\mathcal{E}}^{\infty} \mathrm{e}^{-\beta E}\,\mathrm{d}E\right)\mathrm{d}\mathcal{E} = \frac{8\pi^2}{\beta Z_\beta}\int_{E_0}^{\infty} \frac{\nu^2\,\mathrm{e}^{-\beta\mathcal{E}}}{S_\nu(\mathcal{E})}\,\mathrm{d}\mathcal{E},
\end{aligned}$$

where we used in the first line that the change of variables $(q,p) \mapsto (q,E)$ has Jacobian $P(q,E)$. This concludes the derivation of Result 4.1.

Motivation for (4.19). We assume for simplicity, in addition to Assumption 2.1, that $V(q)$ is an even function around $q = 0$. This is not required to obtain the final result, but it leads to a simplified ansatz for $\overline{\phi}_0$, which for general potentials can be justified only *a posteriori*. Under this additional assumption, one can check by substitution that, if ϕ_γ denotes the solution to the Poisson equation $-\mathcal{L}_\gamma\phi_\gamma = p$, then $\psi_\gamma(q,p,z) = -\phi_\gamma(-q,-p,-z)$ is also a solution to the equation, so $\psi_\gamma = \phi_\gamma$ by uniqueness of the solution. Therefore, ϕ_γ and all the summands $\{\overline{\phi}_i\}_{i=0,1,\ldots}$ must satisfy the symmetry relation

$$u(q,p,z) = -u(-q,-p,-z). \tag{4.21}$$

Multiplying (4.18a) by $\overline{\phi}_0$, integrating with respect to μ, and taking into account that the contribution of the antisymmetric part of $\mathcal{A}_0$ vanishes, we obtain

$$\frac{1}{\beta}\int_{\mathbf{T}\times\mathbf{R}\times\mathbf{R}} \left(\frac{\partial\overline{\phi}_0}{\partial z}\right)^2 \mu(\mathrm{d}q\,\mathrm{d}p\,\mathrm{d}z) = 0, \tag{4.22}$$

which shows that $\overline{\phi}_0 = \overline{\phi}_0(q,p)$. Substituting again in (4.18a), we deduce that $\overline{\phi}_0$ must lie

in the kernel of the operator $p\,\partial_q - V'(q)\,\partial_p$, which consists of functions that are constant on the contour lines of the Hamiltonian $H(q,p)$. Together with the symmetry relation (4.21), and distinguishing closed and open orbits (corresponding respectively to $H(q,p) < E_0$ and $H(q,p) > E_0$), this motivates that

$$\overline{\phi}_0(q,p) = \begin{cases} \psi_0(H(q,p)), & \text{if } H(q,p) \geqslant E_0 \text{ and } p \geqslant 0, \\ 0 & \text{if } H(q,p) < E_0, \\ -\psi_0(H(q,p)), & \text{if } H(q,p) \geqslant E_0 \text{ and } p \leqslant 0, \end{cases}$$

for some function ψ_0 still to be determined.

For the next order we use the ansatz $\overline{\phi}_1(q,p,z) = z\,\psi_1(\mathrm{sign}(p)\,q, H(q,p))$. We remark that any function in the kernel of $\mathcal{A}_0$, *i.e.* any function of only (q,p) that is constant on the contour lines of the Hamiltonian, could in principle be added to this ansatz. However, this will not be necessary for our purposes, because any constant-in-z part of $\overline{\phi}_1$ cancels out in the equation (4.26) for ψ_0 derived below. Restricting our attention first to the region where $H(q,p) \geqslant E_0$ and $p \geqslant 0$, we obtain the following equation for the function $(q,E) \mapsto \psi_1(q,E)$ from (4.18b):

$$p\frac{\partial \psi_1}{\partial q}\left(q, V(q) + \frac{p^2}{2}\right) - \frac{1}{\nu^2}\psi_1\left(q, V(q) + \frac{p^2}{2}\right) + \frac{1}{\nu} p\,\psi_0'\left(V(q) + \frac{p^2}{2}\right) = 0.$$

This equation is satisfied pointwise provided

$$\forall (q,E) \in \mathbf{T} \times (E_0, +\infty), \qquad P(q,E)\,\frac{\partial \psi_1}{\partial q}(q,E) - \frac{1}{\nu^2}\psi_1(q,E) + \frac{1}{\nu}\,P(q,E)\,\psi_0'(E) = 0. \quad (4.23)$$

In view of the definition (4.15) of $s_\nu(q,E)$, it holds

$$\psi_1(q,E) = \frac{1}{\nu} s_\nu(q,E)\,\psi_0'(E). \quad (4.24)$$

In the region $E < E_0$, equation (4.18b) simplifies to $\mathcal{A}_0\overline{\phi}_1 = 0$. We can follow the treatment of (4.22), by taking into account the domain in the (q,p) variables. This is done by multiplying (4.24) by $\overline{\phi}_1$, integrating over the set $A := \{(q,p,z) \in \mathbf{T} \times \mathbf{R} \times \mathbf{R} : H(q,p) \leqslant E_0\}$ with respect to the Gaussian weight $g(z) := \sqrt{\frac{\beta}{2\pi}}\,\mathrm{e}^{-\beta z^2/2}$, and employing the formal antisymmetry of $p\,\partial_q - V'(q)\,\partial_p$, which leads to

$$\frac{1}{\beta}\int_A \left(\frac{\partial \overline{\phi}_1}{\partial z}\right)^2 g(z)\,\mathrm{d}q\,\mathrm{d}p\,\mathrm{d}z = 0,$$

so necessarily $\overline{\phi_1} = \overline{\phi_1}(q,p)$ in that region. Consequently, $\overline{\phi_1}(q,p)$ must lie in the kernel of the operator $p\,\partial_q - V'(q)\,\partial_p$, so it is a function of $H(q,p)$ only. By the symmetry relation (4.21), we obtain that $\psi_1(q,E) = 0$ in that region.

Substituting the expression of $\overline{\phi_1}$ in (4.18c) and integrating in the z direction with respect to the Gaussian weight $g(z)$, we obtain

$$\left(p\frac{\partial}{\partial q} - V'(q)\,\frac{\partial}{\partial p}\right)\int_{\mathbf{R}} \overline{\phi}_2(q,p,z)\,g(z)\,\mathrm{d}z + \frac{p}{\nu}\left[\left(\frac{1}{\beta}\frac{\partial}{\partial E} - 1\right)\psi_1\right](q, H(q,p)) = -p, \quad (4.25)$$

which can be viewed as a PDE on $\mathbf{T} \times \mathbf{R}$ for $(q,p) \mapsto \int_{\mathbf{R}} \overline{\phi}_2(q,p,z)\, g(z)\, \mathrm{d}z$. Since the operator acting on this function is formally antisymmetric in $L^2(B)$, with $B := \{(q,p) \in \mathbf{T} \times \mathbf{R} : p \geqslant 0 \text{ and } H(q,p) > E_0\}$, the solvability condition associated with this equation is

$$\int_B \left(\frac{p}{\nu} \left[\left(\frac{1}{\beta} \frac{\partial}{\partial E} - 1 \right) \psi_1 \right] (q, H(q,p)) + p \right) f(H(q,p))\, \mathrm{d}q\, \mathrm{d}p = 0$$

for any smooth function $E \mapsto f(E)$ with compact support in (E_0, ∞). Using again the change of variables $(q,p) \mapsto (q,E)$, the latter equation reads

$$\int_{\mathbf{T} \times (E_0, \infty)} \left(\frac{1}{\nu} \left[\left(\frac{1}{\beta} \frac{\partial}{\partial E} - 1 \right) \psi_1 \right] (q,E) + 1 \right) f(E)\, \mathrm{d}q\, \mathrm{d}E = 0.$$

In order for this equation to be satisfied for any choice of f, it is necessary that

$$\forall E > E_0, \qquad \int_{\mathbf{T}} \left[\frac{1}{\nu} \left(\frac{1}{\beta} \frac{\partial \psi_1}{\partial E}(q,E) - \psi_1(q,E) \right) + 1 \right] \mathrm{d}q = 0,$$

which, by substituting the expression of ψ_1 given by (4.24), gives

$$\begin{aligned} 0 &= \int_{\mathbf{T}} \left[\frac{1}{\nu^2} \left(\frac{1}{\beta} \frac{\partial}{\partial E} \left(s_\nu(q,E)\, \psi_0'(E) \right) - s_\nu(q,E)\, \psi_0'(E) \right) + 1 \right] \mathrm{d}q \\ &= \frac{1}{\nu^2} \left(\frac{1}{\beta} \frac{\mathrm{d}}{\mathrm{d}E} \left(\psi_0'(E) S_\nu(E) \right) - \psi_0'(E) S_\nu(E) \right) + 2\pi. \end{aligned} \tag{4.26}$$

The latter equation is similar to that obtained for the Langevin dynamics in [59]. Viewed as an ODE for $\psi_0'(E) S_\nu(E)$, equation (4.26) admits the general solution

$$\psi_0'(E) S_\nu(E) = 2\pi\nu^2 + C\, \mathrm{e}^{\beta E}, \tag{4.27}$$

for any constant C. A necessary condition for $\overline{\phi}_0$ to belong to $H^1(\mu)$ is that $C = 0$ and that $\overline{\phi}_0$ be continuous at the homoclinic orbit $H(q,p) = E_0$, which leads to (4.20). It is in fact possible to prove that $\overline{\phi}_0$ is in $H^1(\mu)$, see Appendix C.

Remark 4.4. *The above formal calculations, which could be replicated at the level of the backward Kolmogorov equation associated with the GL1 dynamics, suggest that the Hamiltonian of the rescaled process $H(q^\gamma(t/\gamma), p^\gamma(t/\gamma))$, where $(q^\gamma, p^\gamma, z^\gamma)$ is the solution to (2.2), converges weakly to a Markov process on a graph, with a generator similar to that in the case of underdamped Langevin dynamics; see [25, 30, 57].*

5 Conclusions

In this work, we studied quasi-Markovian approximations of the GLE, and we scrutinized in particular two finite-dimensional models of the noise: the scalar Ornstein–Uhlenbeck noise and the harmonic noise. For the former model, we obtained decay estimates with explicit scalings with respect to the parameters, and we investigated the asymptotic behavior of the associated effective diffusion coefficient in several limits of physical relevance. We also employed

an efficient Fourier/Hermite spectral method to verify most of our findings numerically, thereby complementing previous numerical works [36, 35] on the subject.

Exciting questions remain open for future work.

- On the theoretical front, it is not clear whether a direct $L^2(\mu)$ hypocoercivity approach of the type introduced in [17, 18] can be applied to the GLE. If this was the case, we are hopeful that the approach could be replicated at the discrete level to obtain a hypocoercivity estimate for the Fourier/Hermite numerical method, which would enable the calculation of bounds on the numerical error, as done in [66] for Langevin dynamics. It would also be interesting to investigate whether the approach of [8], based on Schur complements, could be generalized in order to more directly obtain the resolvent bounds (2.12) and (B.1). Finally, it would be interesting to investigate the longtime behavior of more general generalized Langevin equations and, in particular, the application of hypocoercivity techniques to the case of arbitrary stationary Gaussian noise processes. In principle, this would require taking into account an infinite number of auxiliary Ornstein-Uhlenbeck processes and it is related to the problem of stochastic realization theory [47].

- On the numerical front, it would be interesting to investigate carefully the underdamped limit of systems in dimension greater or equal to 2 with Monte Carlo simulations, which could be made more efficient with the variance reduction technique based on control variates recently developed in [67].

Acknowledgements. The authors are grateful to the anonymous referees for their careful reading of our work and their very useful suggestions. The work of G.S. was partially funded by the European Research Council (ERC) under the European Union's Horizon 2020 research and innovation programme (grant agreement No 810367), and by the Agence Nationale de la Recherche under grant ANR-19-CE40-0010-01 (QuAMProcs). The work of G.P. and U.V. was partially funded by EPSRC through grants number EP/P031587/1, EP/L024926/1, EP/L020564/1 and EP/K034154/1. The work of U.V. was partially funded by the Fondation Sciences Mathématique de Paris (FSMP), through a postdoctoral fellowship in the "mathematical interactions" program. The work of G.P. was partially funded by JPMorgan Chase & Co. Any views or opinions expressed herein are solely those of the authors listed, and may differ from the views and opinions expressed by JPMorgan Chase & Co. or its affiliates. This material is not a product of the Research Department of J.P. Morgan Securities LLC. This material does not constitute a solicitation or offer in any jurisdiction.

A Confirmation of the rate of convergence in the quadratic case

To assess the sharpness of the lower bounds on the convergence rate provided by Theorem 2.2, we consider the generalized Langevin dynamics confined by the quadratic potential $V(q) = kq^2/2$

(with $k > 0$) in $\mathcal{X} = \mathbf{R}$. In this case, (2.2) can be written as (2.13) with

$$\mathbf{D} = \begin{pmatrix} 0 & 1 & 0 \\ -k & 0 & \frac{\sqrt{\gamma}}{\nu} \\ 0 & -\frac{\sqrt{\gamma}}{\nu} & -\frac{1}{\nu^2} \end{pmatrix} = \mathbf{A} + \frac{\sqrt{\gamma}}{\nu}\mathbf{B} + \frac{1}{\nu^2}\mathbf{C}, \tag{A.1}$$

with

$$\mathbf{A} = \begin{pmatrix} 0 & 1 & 0 \\ -k & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}, \qquad \mathbf{B} = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & -1 & 0 \end{pmatrix}, \qquad \mathbf{C} = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & -1 \end{pmatrix}.$$

In order to determine the elements in (2.14), it suffices to compute the three eigenvalues of $\mathbf{D}$, and combine them linearly with positive coefficients. This means that the spectral bound corresponds to the smallest absolute value of the real parts of the eigenvalues of $\mathbf{D}$. For the model GL1, the eigenvalues x_j (for $1 \leqslant j \leqslant 3$) are the roots of the polynomial $p(x) = \det(x\mathbf{I} - \mathbf{D})$, which reads

$$p(x) = x^3 + \frac{x^2}{\nu^2} + \left(\frac{\gamma}{\nu^2} + k\right) x + \frac{k}{\nu^2}. \tag{A.2}$$

The spectral gap is easily seen to be a continuous function of $(\gamma, \nu) \in (0, +\infty) \times (0, +\infty)$, with positive values. To obtain the scaling of the spectral gap as a function of the parameters, it therefore suffices to consider the various limiting regimes where at least one of the parameters goes to 0 or $+\infty$. The expression (A.1) suggests that the eigenvalues can be expanded in series of (inverse or fractional) powers of γ and ν. The leading order term depends on the asymptotic regime which is considered. We start by investigating regimes where only one of the parameters goes to 0 or $+\infty$, and then discuss regimes where both parameters are varied. The organization of this appendix is illustrated in Fig. 7.

(i) Limit $\nu \to +\infty$ with γ fixed. We denote by $\varepsilon = \nu^{-1}$ and rewrite (A.1) as $\mathbf{A} + \varepsilon\sqrt{\gamma}\mathbf{B} + \varepsilon^2\mathbf{C}$. The matrix $\mathbf{A}$ is diagonalizable and its eigenvalues 0 and $\pm \mathrm{i}\sqrt{k}$ are isolated and non-degenerate. Perturbation theory [38, Chapter II] then shows that the eigenvalues $x_j(\varepsilon)$ are analytic functions of ε for ε sufficiently small. We write

$$x_j(\varepsilon) = x_j^0 + x_j^1\varepsilon + x_j^2\varepsilon^2 + \dots \tag{A.3}$$

To identify the coefficients x_j^k, we plug the above expansion in the characteristic polynomial, which reads

$$p(x) = x^3 + \varepsilon^2 x^2 + \left(\gamma\varepsilon^2 + k\right) x + k\varepsilon^2, \tag{A.4}$$

and identify terms with the same powers of ε in $p(x_j(\varepsilon)) = 0$. Some straightforward computations show that $x_j^{2n+1} = 0$ (in fact, the expansion in (A.3) could be immediately restricted to even powers of ε by [60, Theorem XII.2] since the coefficients of the polynomial are analytic in ε^2), and

$$x_j^2 = -\frac{k + \gamma x_j^0 + (x_j^0)^2}{3(x_j^0)^2 + k}, \qquad x_j^4 = -\frac{(\gamma + 2x_j^0)x_j^2 + 3x_j^0(x_j^2)^2}{3(x_j^0)^2 + k},$$

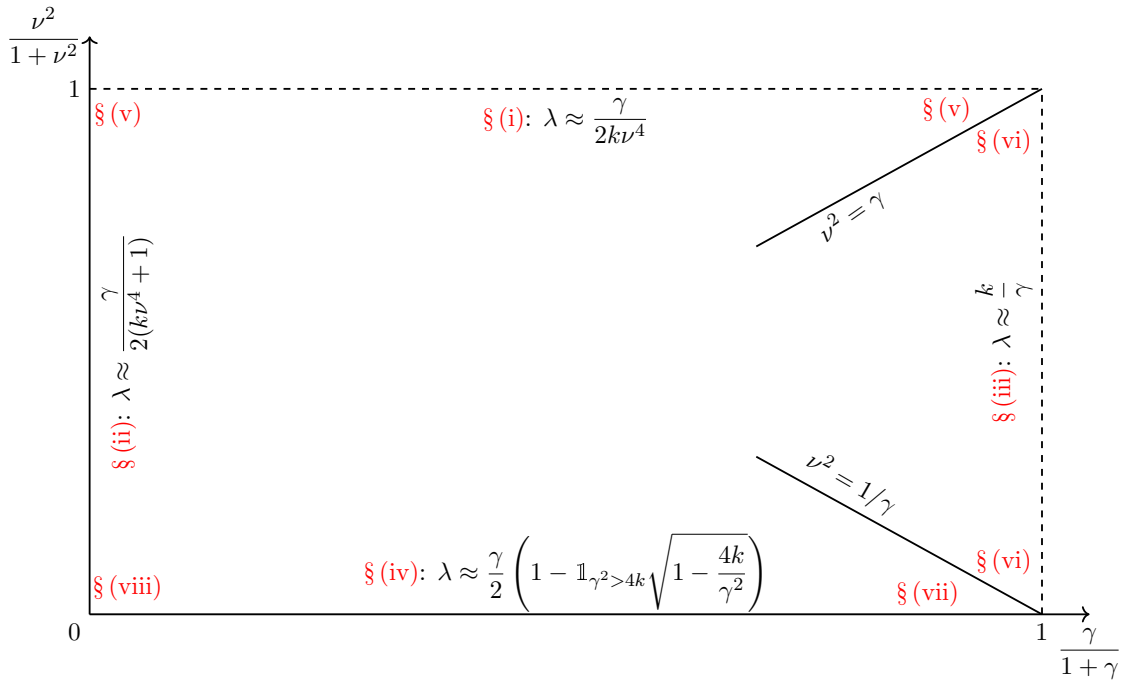


Figure 7: Illustration of the different limits considered in this appendix. The joint limit $\gamma \to \infty$ and $\nu \to \infty$ with γ/ν^2 bounded from above and below, as well as the joint limit $\gamma \to \infty$ and $\nu \to 0$ with $\gamma\nu^2$ bounded from above and below, are omitted for conciseness.

so

$$x_1(\varepsilon) = -\varepsilon^2 + \mathcal{O}\left(\varepsilon^4\right), \qquad x_2(\varepsilon) = \mathrm{i}\sqrt{k}\left(1 + \frac{\gamma}{2k}\varepsilon^2 - \frac{\gamma^2}{8k^2}\varepsilon^4\right) - \frac{\gamma}{2k}\varepsilon^4 + \mathcal{O}\left(\varepsilon^6\right), \tag{A.5}$$

and a similar expression for $x_3(\varepsilon)$ with an imaginary part of opposite sign. This shows that the spectral gap λ is $\gamma/(2k\nu^4) + \mathcal{O}\left(\nu^{-6}\right)$ as $\nu \to +\infty$ with $\gamma > 0$ fixed.

(ii) Limit $\gamma \to 0$ with ν fixed. We denote by $\varepsilon = \sqrt{\gamma}$ and rewrite (A.1) as $\mathbf{A} + \nu^{-2}\mathbf{C} + \varepsilon\nu^{-1}\mathbf{B}$. The dominant part $\mathbf{A} + \nu^{-2}\mathbf{C}$ is diagonalizable with isolated and non-degenerate eigenvalues $x_1^0 = -\nu^{-2}$, $x_2^0 = \mathrm{i}\sqrt{k}$ and $x_3^0 = -\mathrm{i}\sqrt{k}$. The eigenvalues $x_j(\varepsilon)$ of (A.1) are therefore analytic in ε for ε small enough, and an expansion of the form (A.3) also holds in this case. By an analysis similar to the one leading to (A.5), we obtain $x_j^1 = x_j^3 = 0$ and $x_j^2 = -x_j^0/(3\nu^2(x_j^0)^2 + 2x_j^0 + k\nu^2)$, so

$$x_1(\varepsilon) = -\frac{1}{\nu^2} + \mathcal{O}\left(\gamma\right), \qquad x_2(\varepsilon) = \mathrm{i}\sqrt{k} - \frac{1 - \mathrm{i}\sqrt{k}\nu^2}{2(1+k\nu^4)}\gamma + \mathcal{O}\left(\gamma^2\right), \tag{A.6}$$

and a similar expression for $x_3(\varepsilon)$ with an imaginary part of opposite sign. This shows that the spectral gap λ is $\gamma/(2(1+k\nu^4)) + \mathcal{O}\left(\gamma^2\right)$ as $\gamma \to 0$ with $\nu > 0$ fixed.

(iii) Limit $\gamma \to +\infty$ for $\nu > 0$ fixed. In order to treat the situation when γ diverges, we introduce $\varepsilon = \gamma^{-1/2}$ and rewrite (A.1) as $\varepsilon^{-1}\left[\nu^{-1}\mathbf{B} + \varepsilon\left(\mathbf{A} + \nu^{-2}\mathbf{C}\right)\right]$. The eigenvalues $x_j(\varepsilon)$ of $\mathbf{D}$ are then obtained by rescaling the eigenvalues $y_j(\varepsilon)$ of $\nu^{-1}\mathbf{B} + \varepsilon\left(\mathbf{A} + \nu^{-2}\mathbf{C}\right)$ as $x_j(\varepsilon) =$

$y_j(\varepsilon)/\varepsilon$. The eigenproblem associated with $\nu^{-1}\mathbf{B} + \varepsilon\left(\mathbf{A} + \nu^{-2}\mathbf{C}\right)$ has the form of a standard perturbation problem, with associated characteristic polynomial

$$P(y) = y^3 + \varepsilon\nu^{-2}y^2 + (\nu^{-2} + k\varepsilon^2)y + k\nu^{-2}\varepsilon^3.$$

The dominant part $\nu^{-1}\mathbf{B}$ is diagonalizable with isolated and non-degenerate eigenvalues $y_1^0 = 0$, $y_2^0 = \mathrm{i}/\nu$ and $y_3^0 = -\mathrm{i}/\nu$. The eigenvalues $y_j(\varepsilon)$ of $\nu^{-1}\mathbf{B} + \varepsilon\left(\mathbf{A} + \nu^{-2}\mathbf{C}\right)$ are therefore again analytic in ε for ε small enough, and an expansion of the form (A.3) still holds, with x replaced by y. By gathering terms with the same powers of ε in $P(y_j(\varepsilon)) = 0$, we find after some computations that

$$y_1(\varepsilon) = -k\gamma^{-3/2} + \mathcal{O}\left(\frac{1}{\gamma^{5/2}}\right), \qquad y_2(\varepsilon) = \frac{\mathrm{i}}{\nu} - \frac{1}{2\nu^2\sqrt{\gamma}} + \mathcal{O}\left(\frac{1}{\gamma}\right),$$

and a similar expression for $y_3(\varepsilon)$ with an imaginary part of opposite sign. This shows that the spectral gap λ is $k/\gamma + \mathcal{O}\left(\gamma^{-2}\right)$ as $\gamma \to +\infty$ with $\nu > 0$ fixed.

(iv) Limit $\nu \to 0$ for γ fixed. We introduce here $\varepsilon = \nu$ and rewrite (A.1) as $\varepsilon^{-2}\left[\mathbf{C} + \varepsilon\sqrt{\gamma}\mathbf{B} + \varepsilon^2\mathbf{A}\right]$. As in the previous situation, we look for the eigenvalues $y_j(\varepsilon)$ of $\mathbf{C} + \varepsilon\sqrt{\gamma}\mathbf{B} + \varepsilon^2\mathbf{A}$, whose characteristic polynomial reads

$$P(y) = y^3 + y^2 + (\gamma\varepsilon^2 + k\varepsilon^4)y + k\varepsilon^4. \tag{A.7}$$

The dominant part $\mathbf{C}$ of the eigenvalue problem is diagonalizable with eigenvalues $y_1^0 = -1$ and $y_2^0 = y_3^0 = 0$. Since the latter eigenvalue is doubly degenerate, the results of [38, Chapter II] show, by reducing the eigenvalue problem to the subspace associated with $y_2(\varepsilon)$ and $y_3(\varepsilon)$, that $y_1(\varepsilon)$ is analytic in ε for ε sufficiently small, while $y_2(\varepsilon)$ and $y_3(\varepsilon)$ are analytic in $\sqrt{\varepsilon}$:

$$y_2(\varepsilon) = y_2^{1/2}\sqrt{\varepsilon} + y_2^1\varepsilon + y_2^{3/2}\varepsilon^{3/2} + \dots, \tag{A.8}$$

and a similar expansion for $y_3(\varepsilon)$. Note that $y_1(\varepsilon) = -1 + \mathcal{O}(\varepsilon)$, so the spectral gap is determined by the lowest order terms in the expansions of $\lambda_2(\varepsilon)$ and $\lambda_3(\varepsilon)$. By plugging (A.8) into (A.7) we find that the first non-zero terms in the expansions satisfy

$$(y_j^2)^2 + \gamma y_j^2 + k = 0, \qquad y_j^{5/2}(2y_j^2 + \gamma) = 0.$$

We need to distinguish three situations at this stage:

(i) When $\gamma^2 > 4k$, one finds $y_2^2 = -\frac{\gamma}{2}\left(1 + \sqrt{1 - \frac{4k}{\gamma^2}}\right)$ and $y_3^2 = -\frac{\gamma}{2}\left(1 - \sqrt{1 - \frac{4k}{\gamma^2}}\right)$, so $y_j^{5/2} = 0$. In this case the spectral gap of the rescaled matrix $\mathbf{C} + \varepsilon\sqrt{\gamma}\mathbf{B} + \varepsilon^2\mathbf{A}$ scales as $\frac{\gamma}{2}\left(1 - \sqrt{1 - \frac{4k}{\gamma^2}}\right) + \mathcal{O}(\nu)$ as $\nu \to 0$ with $\gamma > 0$ fixed.

(ii) When $\gamma^2 < 4k$, one finds $y_2^2 = -\frac{\gamma}{2}\left(1 - \mathrm{i}\sqrt{\frac{4k}{\gamma^2} - 1}\right)$ and $y_3^2 = -\frac{\gamma}{2}\left(1 + \mathrm{i}\sqrt{\frac{4k}{\gamma^2} - 1}\right)$, so $y_j^{5/2} = 0$. In this case the spectral gap of the rescaled matrix scales as $\frac{\gamma}{2} + \mathcal{O}(\nu)$ as $\nu \to 0$ with $\gamma > 0$ fixed.

(iii) When $\gamma^2 = 4k$, one finds $y_2^2 = y_3^2 = -\frac{\gamma}{2}$, and actually $y_j^{5/2} = 0$ (because of the next order condition in ε which reads $(y_j^{5/2})^2 + (\gamma + 2y_j^2)y_j^3 = 0$). The spectral gap of the rescaled matrix scales as in the previous case as $\frac{\gamma}{2} + \mathcal{O}(\nu)$ as $\nu \to 0$.

In conclusion, the spectral gap λ of $\mathbf{D}$ scales as $\frac{\gamma}{2}\left(1 - \mathbb{1}_{\gamma^2>4k}\sqrt{1 - \frac{4k}{\gamma^2}}\right) + \mathcal{O}(\nu^3)$.

(v) Joint limits $\gamma \to 0$ and $\nu \to +\infty$, or $\nu \to +\infty$ with $\gamma/\nu^2 \to 0$. We denote by $\eta = \frac{\sqrt{\gamma}}{\nu}$ and $\varepsilon = \frac{1}{\nu^2}$. These two parameters are small in the asymptotic regime which is considered. The eigenvalues of the leading order matrix $\mathbf{A}$ in (A.1) are isolated, so, by perturbation theory (upon writing the spectral projector using contour integrals and expanding the resolvents which appear), we can write the eigenvalues of $\mathbf{D} = \mathbf{A} + \eta\mathbf{B} + \varepsilon\mathbf{C}$ as

$$x_j(\eta, \varepsilon) = x_j^0 + x_j^{1,0}\eta + x_j^{0,1}\varepsilon + x_j^{2,0}\eta^2 + x_j^{1,1}\eta\varepsilon + x_j^{0,2}\varepsilon^2 + \dots, \tag{A.9}$$

with $x_1^0 = 0$, $x_2^0 = \mathrm{i}\sqrt{k}$ and $x_3^0 = -\mathrm{i}\sqrt{k}$. We next identify terms with the same powers of η, ε in $p(x_j(\eta, \varepsilon)) = 0$ with $p(x) = x^3 + \varepsilon x^2 + (\eta^2 + k)x + k\varepsilon$. Some straightforward computations show that

$$x_1(\eta, \varepsilon) = -\varepsilon + \frac{1}{k}\eta^2\varepsilon + \mathcal{O}\left(\eta^2\left(\varepsilon^2 + \eta^2\right)\right), \qquad x_2(\eta, \varepsilon) = \mathrm{i}\left(\sqrt{k} + \frac{\eta^2}{2\sqrt{k}}\right) - \frac{1}{2k}\eta^2\varepsilon + \mathcal{O}\left(\varepsilon^4 + \eta^4\right),$$

and a similar expansion for $x_3(\eta, \varepsilon)$ with imaginary parts of opposite signs. Note that the expressions of x_1, x_2 coincide with (A.5), as well with (A.6) in the limit $\nu \to +\infty$. It can be shown that $x_j^{2m,0} \in \mathrm{i}\mathbb{R}$ and $x_j^{2m+1,0} = 0$ for any $m \geqslant 1$ and $1 \leqslant j \leqslant 3$ (it suffices in fact to set $\varepsilon = 0$ to identify these coefficients, which amounts to considering the polynomial $x(x^2+\eta^2+k)$, whose roots are 0 and $\pm\mathrm{i}\sqrt{k+\eta^2}$). Similarly, $x_j^{0,m} = 0$ for $m \geqslant 2$ and $1 \leqslant j \leqslant 3$, as seen by setting $\eta = 0$ and factorizing $p(x)$ as $(x + \varepsilon)(x^2 + k)$. The spectral gap is therefore $\frac{\gamma}{2k\nu^4} + \mathcal{O}\left(\varepsilon^2\eta^2\right)$.

(vi) Joint limits $\gamma \to +\infty$ with $\gamma/\nu^2 \to +\infty$ and $\gamma\nu^2 \to +\infty$. We denote by $\eta = \nu/\sqrt{\gamma}$ and $\varepsilon = 1/(\nu\sqrt{\gamma})$ the two small parameters in the asymptotic regime considered here. We write $\mathbf{D}$ as $\eta^{-1}(\mathbf{B} + \eta\mathbf{A} + \varepsilon\mathbf{C})$, the characteristic polynomial associated with $\mathbf{B} + \eta\mathbf{A} + \varepsilon\mathbf{C}$ being $P(y) = y^3 + \varepsilon y^2 + (k\eta^2 + 1)y + k\varepsilon\eta^2$. We can use an argument similar to the one used to write (A.9) with $y_1^0 = 0$, $y_2^0 = \mathrm{i}$ and $y_3^0 = -\mathrm{i}$ to obtain the following expansion for the zeros of the polynomial $P(y)$:

$$y_1(\eta, \varepsilon) = -k\eta^2\varepsilon + \mathcal{O}\left(\eta^4 + \eta^4\right), \qquad y_2(\eta, \varepsilon) = \mathrm{i} - \frac{\varepsilon}{2} + \mathcal{O}\left(\varepsilon^2 + \eta^2\right),$$

and a similar expansion for $y_3(\eta, \varepsilon)$ with imaginary parts of opposite signs. Note that it can be shown that the remainders for $y_1(\eta, \varepsilon)$ involve only powers of η^2 (since the polynomial itself is analytic in η^2, see [60, Theorem XII.2]) and there are no remainders of the form ε^n or η^n (as seen by setting respectively $\eta = 0$ and $\varepsilon = 0$ in the expression of $P(y)$). For y_2, it can similarly be shown that $y_2^{2m,0} \in \mathrm{i}\mathbb{R}$ and $y_2^{2m+1,0} = 0$ for any $m \geqslant 1$. The spectral gap is therefore $k\eta\varepsilon + \mathcal{O}\left(\eta\varepsilon^2\right)$ with $k\eta\varepsilon = k/\gamma$, which coincides with the limit obtained for $\gamma \to +\infty$ with ν fixed.

(vii) Joint limits $\nu \to 0$ and $\gamma \to +\infty$ with $\gamma\nu^2 \to 0$. It is difficult to rely on a perturbative approach here. Indeed, denoting by $\eta = \nu\sqrt{\gamma}$ and $\varepsilon = \nu^2$ the two small parameters in the asymptotic regime considered here, we write $\mathbf{D}$ as $\varepsilon^{-1}(\mathbf{C} + \eta\mathbf{B} + \varepsilon\mathbf{A})$. The dominant term $\mathbf{C}$ however has degenerate eigenvalues, so it is not clear whether the eigenvalues can be expanded as in (A.9) (in fact, it is in general not true that this is the case, see [38, Section II.5.7]). We use for this case another argument, based on a localization of the zeros of the characteristic polynomial $P(y) = y^3 + y^2 + (k\varepsilon^2 + \eta^2)y + k\varepsilon^2$ associated with $\mathbf{C} + \eta\mathbf{B} + \varepsilon\mathbf{A}$. We first note that the discriminant of the third order polynomial P is

$$\left(k\varepsilon^2 + \eta^2\right)^2 \left[1 - 4\left(k\varepsilon^2 + \eta^2\right)\right] - \left(4 + 9k\varepsilon^2 - 18\eta^2\right) k\varepsilon^2 = -4k\varepsilon^2 + \eta^4 + \mathcal{O}\left(\varepsilon^4 + \eta^2\varepsilon^2 + \eta^6\right).$$

Since $\eta^4/\varepsilon^2 = \gamma^2 \to +\infty$, the discriminant is positive in the limiting regime we consider, so that P admits three real zeros. The polynomial p in (A.4) therefore also admits three real zeros. We next compute

$$p\left(-\frac{1}{\nu^2} + \delta\gamma\right) = \frac{(\delta - 1)\gamma}{\nu^4} + \delta\gamma\left(\frac{(1 - 2\delta)\gamma}{\nu^2} + \delta^2\gamma^2 + k\right).$$

When $\delta = 1$, the right-hand side of the above equality is $-\gamma^2/\nu^2 < 0$ at dominant order, while, for $\delta = 1 + \varepsilon$ with a fixed small parameter $\varepsilon > 0$, the right-hand side scales as $\varepsilon\gamma/\nu^4 > 0$. This shows that one of the roots is in the interval $[-\nu^{-2} + \gamma, -\nu^{-2} + (1 + \varepsilon)\gamma]$ for γ large enough and $\nu, \gamma\nu^2$ sufficiently small. To localize the second root, we consider

$$\frac{\nu^2}{\gamma^2} p\left(-(1 + \delta)\gamma\right) = \delta(1 + \delta) - \gamma\nu^2(1 + \delta)^3 - \frac{k\nu^2}{\gamma}(1 + \delta) + \frac{k}{\gamma^2}.$$

The right-hand side converges to $\delta(1 + \delta)$ as $\gamma \to +\infty$ and $\nu \to 0$ with $\gamma\nu^2 \to 0$, which shows that, for any $\varepsilon \in (0, 1]$, the second root is in the interval $[-(1 + \varepsilon)\gamma, -(1 - \varepsilon)\gamma]$ for γ large enough and $\nu, \gamma\nu^2$ sufficiently small . Finally,

$$\nu^2 p\left(-\frac{k}{\gamma} - \frac{(1 + \delta)k^2}{\gamma^3}\right) = -\delta\frac{k^2}{\gamma^2} - \frac{k\nu^2}{\gamma}\left(k + \frac{(1 + \delta)k^2}{\gamma^2}\right) - \frac{\nu^2}{\gamma^3}\left(k + \frac{(1 + \delta)k^2}{\gamma^2}\right)^3 + \frac{(1 + \delta)k^3}{\gamma^4}\left(2 + \frac{(1 + \delta)k}{\gamma^2}\right).$$

The right-hand side scales as $-\delta k^2/\gamma^2$ as $\gamma \to +\infty$ and $\nu \to 0$ with $\gamma\nu^2 \to 0$, which shows that, for any $\varepsilon \in (0, 1]$, the third root is in the interval $[-k/\gamma - (1 + \varepsilon)k^2/\gamma^2, -k/\gamma - (1 - \varepsilon)k^2/\gamma^2]$ for γ large enough and $\nu, \gamma\nu^2$ sufficiently small. The spectral gap is dictated by the location of this eigenvalue, and scales therefore as k/γ in the limit which is considered here.

(viii) Limit $\gamma \to 0$ and $\nu \to 0$. Here also, it is difficult to rely on a perturbative approach, so we use an alternative method. Employing the same reasoning as in § (vii), it is possible to show that the characteristic polynomial p admits only one real root in this limit, denoted by x_1, and that this root is, for $\varepsilon \in (0, 1]$ fixed, in the interval $[-\nu^{-2} + (1 - \varepsilon)\gamma, -\nu^{-2} + (1 + \varepsilon)\gamma]$ provided that $\gamma\nu^2$ is sufficiently small. Inspired by (A.6), we show that the other two, complex conjugate roots, which we denote by $x^{\pm}$, with the superscript indicating the sign of the imaginary part,

are close to $\hat{x}^{\pm} := -\gamma/2 \pm \mathrm{i}\sqrt{k}$. To this end, we calculate

$$p(\hat{x}^{\pm}) = k\gamma \pm \frac{3}{4}\mathrm{i}\sqrt{k}\gamma^2 - \frac{\gamma^3}{8} - \frac{\gamma^2}{4\nu^2}.$$

Given the factorization $p(x) = (x - x_1)(x - x^+)(x - x^-)$, this implies

$$(\hat{x}^{\pm} - x^+)(\hat{x}^{\pm} - x^-) = \frac{k\gamma \pm \frac{3}{4}\mathrm{i}\sqrt{k}\gamma^2 - \frac{\gamma^3}{8} - \frac{\gamma^2}{4\nu^2}}{\hat{x}^{\pm} - x_1}. \tag{A.10}$$

Since $|\hat{x}^{\pm} - x_1|$ scales as ν^{-2}, the modulus of the right-hand side scales as $k\gamma\nu^2 - \frac{\gamma^2}{4}$ in the limit as $\gamma \to 0$ and $\nu \to 0$. This implies

$$|\hat{x}^{\pm} - x^{\pm}|^2 \leqslant |\hat{x}^{\pm} - x^+|\,|\hat{x}^{\pm} - x^-| = \mathcal{O}\left(\gamma\nu^2 + \gamma^2\right),$$

so $|\hat{x}^{\pm} - x^{\pm}| = \mathcal{O}\left(\sqrt{\gamma}\nu + \gamma\right)$. Therefore, $|\hat{x}^{\pm} - x^{\mp}|$ converges to $2\sqrt{k}$ in the limit $\gamma, \nu \to 0$, so (A.10) implies that in fact $|\hat{x}^{\pm} - x^{\pm}| = \mathcal{O}\left(\gamma\nu^2 + \gamma^2\right)$, which is negligible in front of $\gamma/2 = -\mathrm{Re}(\hat{x}^{\pm})$. In conclusion, the spectral gap scales as $\gamma/2$ in the limit considered here.

Remark A.1. *The scaling of the exponential decay rate can also be obtained from explicit expressions for the roots of the characteristic polynomial, as we illustrate below in the particular case of the limit $\nu \to \infty$ with fixed γ. For simplicity, we consider $k = \gamma = 1$. In this case, the characteristic polynomial is*

$$p(x) = \frac{1}{\nu^2}\left(\nu^2 x^3 + x^2 + \left(1 + \nu^2\right)x + 1\right).$$

Letting $\lambda = \nu^2$, we define

$$q(x) = \lambda x^3 + x^2 + (1 + \lambda)\,x + 1 =: ax^3 + bx^2 + cx + d.$$

Using the standard notation for expressing the roots of a cubic polynomial, we define

$$\begin{aligned}\Delta_0 &= b^2 - 3ac = 1 - 3\lambda(1 + \lambda) = -3\lambda^2 - 3\lambda + 1,\\ \Delta_1 &= 2b^3 - 9abc + 27a^2d = 2 - 9\lambda(1 + \lambda) + 27\lambda^2 = 18\lambda^2 - 9\lambda + 2,\end{aligned}$$

and

$$C = \sqrt[3]{\frac{\Delta_1 + \sqrt{\Delta_1^2 - 4\Delta_0^3}}{2}} = \sqrt[3]{\frac{18\lambda^2 - 9\lambda + 2 + \sqrt{108\lambda^6 + 324\lambda^5 + 540\lambda^4 - 432\lambda^3 + 81\lambda^2}}{2}}.$$

By a Taylor expansion, obtained by symbolic calculation, we calculate

$$\begin{aligned}\frac{C}{\lambda} &= \sqrt{3} + \left(\frac{\sqrt{3}}{2} + 1\right)\frac{1}{\lambda} - \left(3 + \frac{\sqrt{3}}{4}\right)\frac{1}{2\lambda^2} + \mathcal{O}\left(\frac{1}{\lambda^3}\right)\\ &=: \hat{C}(\lambda) + \mathcal{O}\left(\frac{1}{\lambda^3}\right).\end{aligned}$$

The roots of q (and therefore also of p) are given by

$$x_k = -\frac{1}{3a}\left(b + \xi^k C + \frac{\Delta_0}{\xi^k C}\right),$$
$$= -\frac{1}{3}\left(\frac{1}{\lambda} + \xi^k \hat{C} + \frac{\Delta_0}{\xi^k \hat{C}\lambda^2} + \mathcal{O}\left(\frac{1}{\lambda^3}\right)\right), \qquad \xi = \frac{-1+i\sqrt{3}}{2}, \qquad k \in \{0,1,2\}.$$

Carrying out the calculation of the leading order terms using symbolic calculations, we obtain

$$\mathrm{Re}(x_0) = -\frac{1}{\lambda} + \mathcal{O}\left(\frac{1}{\lambda^2}\right), \quad \mathrm{Re}(x_1) = -\frac{1}{2\lambda^2} + \mathcal{O}\left(\frac{1}{\lambda^3}\right), \quad \mathrm{Re}(x_2) = -\frac{1}{2\lambda^2} + \mathcal{O}\left(\frac{1}{\lambda^3}\right),$$

so the decay rate scales as $-1/2\nu^4$ in the limit, which coincides with the result found in § (i). This example shows that, while feasible, obtaining the scaling of the spectral gap from the explicit expressions for the roots of the characteristic polynomial presents a level of difficulty similar to the one encountered above. One advantage of our approach is that it can easily be extended to higher dimensions, which is relevant when more auxiliary processes are considered in the GLE.

B Longtime behavior for model GL2

We present here elements on how to obtain results for GL2 similar to the ones proved for GL1 in Section 3, with choices of coefficients providing a modified $H^1(\mu)$ inner product and guaranteeing hypocoercivity and hypoelliptic regularization. We omit the details of the calculations, which are more cumbersome than for GL1. Defining the generator associated with GL2 as

$$-\mathcal{L} = \nu^{-2}\,\alpha^2\,\beta^{-1}\,\partial^*_{z_2}\,\partial_{z_2}$$
$$+ \nu^{-2}\,\alpha(z_1\,\partial_{z_2} - z_2\,\partial_{z_1}) + \sqrt{\gamma}\,\nu^{-1}(p\,\partial_{z_1} - z_1\,\partial_p) + (V'\,\partial_p - p\,\partial_q),$$

where $\partial^*_{z_2} := \beta z_2 - \partial_{z_2}$, the approach outlined below leads to the resolvent bound

$$\left\|\mathcal{L}^{-1}\right\|_{\mathcal{B}(L^2_0(\mu))} \leqslant C \max\left(\gamma,\ \frac{1}{\gamma},\ \frac{\nu^4}{\gamma},\ \frac{\nu^8}{\alpha^4\gamma},\ \frac{\gamma}{\alpha^2},\ \frac{\gamma^2\nu^2}{\alpha^4}\right). \tag{B.1}$$

In contrast with our observations in the case of model GL1, the scaling on the right-hand side of this equation appears not to be sharp. Indeed, an explicit calculation of the scaling of the spectral bound in the case of a quadratic potential, which is possible by using the same approach as in Appendix A, shows that the resolvent bound scales in fact as γ in the limit as $\gamma \to \infty$ for fixed α and ν, and not as γ^2 as suggested by (B.1).

Let us now make precise how we obtain (B.1). Define the coefficients

$$
\begin{aligned}
a_0 &= 2A^4 \min\left(1, \alpha, \frac{\alpha^2}{\nu^2}, \frac{\alpha^2}{\sqrt{\gamma}\nu}\right), \\
a_1 &= 2A^{10} \min\left(1, \alpha, \frac{1}{\nu^2}, \frac{\alpha^4}{\nu^6}, \frac{1}{\sqrt{\gamma}\nu}, \frac{\alpha^4}{\gamma^{\frac{3}{2}}\nu^3}\right), \\
a_2 &= 2A^{14} \min\left(1, \alpha, \gamma, \frac{\gamma}{\nu^4}, \frac{\alpha^4\gamma}{\nu^8}, \frac{1}{\sqrt{\gamma}\nu}, \frac{\alpha^4}{\gamma^{\frac{3}{2}}\nu^3}\right), \\
a_3 &= 2A^{16} \min\left(\gamma, \frac{1}{\gamma}, \frac{\gamma}{\nu^4}, \frac{\alpha^2}{\gamma}, \frac{\alpha^4\gamma}{\nu^8}, \frac{\alpha^4}{\gamma^2\nu^2}\right),
\end{aligned}
$$

and

$$
\begin{aligned}
b_0 &= A^7 \min\left(\alpha, \frac{1}{\alpha}, \frac{\alpha^3}{\nu^4}, \frac{\alpha^3}{\gamma\nu^2}\right), \\
b_1 &= A^{12} \min\left(\sqrt{\gamma}\nu, \frac{1}{\sqrt{\gamma}\nu}, \frac{\sqrt{\gamma}}{\nu^3}, \frac{\alpha^4\sqrt{\gamma}}{\nu^7}, \frac{\alpha^4}{\gamma^{\frac{3}{2}}\nu^3}\right), \\
b_2 &= A^{15} \min\left(\gamma, \frac{1}{\gamma}, \frac{\gamma}{\nu^4}, \frac{\alpha^2}{\gamma}, \frac{\alpha^4\gamma}{\nu^8}, \frac{\alpha^4}{\gamma^2\nu^2}\right).
\end{aligned}
$$

It is possible to show that, for any sufficiently small A and for fixed β, the following conditions are satisfied for any values of the parameters γ, ν and α:

- The following matrix is positive definite:

$$
\begin{pmatrix}
a_0 & -b_0 & 0 & 0 \\
-b_0 & a_1 & -b_1 & 0 \\
0 & -b_1 & a_2 & -b_2 \\
0 & 0 & -b_2 & a_3
\end{pmatrix}
$$

  Consequently, the inner product defined by polarization from the norm

$$
\begin{aligned}
(\!(h,h)\!) := \|h\|^2 + a_0 \|\partial_{z_2} h\|^2 + a_1 \|\partial_{z_2} h\|^2 + a_2 \|\partial_p h\|^2 + a_3 \|\partial_q h\|^2 \\
- 2b_0 \langle \partial_{z_2} h, \partial_{z_1} h \rangle - 2b_1 \langle \partial_{z_1} h, \partial_p h \rangle - 2b_2 \langle \partial_p h, \partial_q h \rangle ,
\end{aligned}
$$

  is equivalent to the usual $H^1(\mu)$ inner product, by an inequality similar to (3.5).

- The operator $-\mathcal{L}$ is coercive in $H_0^1(\mu)$ endowed with the inner product $(\!(\cdot,\cdot)\!)$. More precisely, it holds for any $h \in H_0^1(\mu)$ that

$$
-(\!(h, \mathcal{L}h)\!) \geqslant C(A)\lambda(\mu,\nu,\alpha)\,(\!(h,h)\!),
$$

  where

$$
\lambda(\mu,\nu,\alpha) = \min\left(\gamma, \frac{1}{\gamma}, \frac{\gamma}{\nu^4}, \frac{\alpha^4\gamma}{\nu^8}, \frac{\alpha^2}{\gamma}, \frac{\alpha^4}{\gamma^2\nu^2}\right). \tag{B.2}
$$

- For any $h \in L^2_0(\mu)$, it holds $N_h'(t) \leqslant 0$ for $t \in [0,1]$, where $N_h(t)$ is defined by

$$N_h(t) = \left\| \mathrm{e}^{t\mathcal{L}} h \right\|^2 + a_0 t \left\| \partial_{z_2} \mathrm{e}^{t\mathcal{L}} h \right\|^2 + a_1 t^3 \left\| \partial_{z_1} \mathrm{e}^{t\mathcal{L}} h \right\|^2 + a_2 t^5 \left\| \partial_p \mathrm{e}^{t\mathcal{L}} h \right\|^2 + a_3 t^7 \left\| \partial_q \mathrm{e}^{t\mathcal{L}} h \right\|^2 - 2b_0 t^2 \left\langle \partial_{z_2} \mathrm{e}^{t\mathcal{L}} h, \partial_{z_1} \mathrm{e}^{t\mathcal{L}} h \right\rangle - 2b_1 t^4 \left\langle \partial_{z_1} \mathrm{e}^{t\mathcal{L}} h, \partial_p \mathrm{e}^{t\mathcal{L}} h \right\rangle - 2b_2 t^6 \left\langle \partial_p \mathrm{e}^{t\mathcal{L}} h, \partial_q \mathrm{e}^{t\mathcal{L}} h \right\rangle .$$

This leads finally to (B.1).

C Technical results used in Section 4.3

In this section, we present technical results that are used in Section 4.3. The following estimate is useful in motivating (4.19).

Lemma C.1. *Assume that $c > 0$ is a constant, that h is a smooth function on the torus $\mathbf{T}$ and that g is a smooth, everywhere positive function on $\mathbf{T}$. Then there exists a unique smooth solution f to the equation*

$$g(q) f'(q) - c\, f(q) = h(q), \qquad q \in \mathbf{T}. \tag{C.1}$$

In addition,

$$\max_{q \in \mathbf{T}} \left| f(q) + \frac{h(q)}{c} + \frac{g(q) h'(q)}{c^2} \right| \leqslant \frac{1}{c^3} \max_{q \in \mathbf{T}} \left| g(q)\, (g\, h')'(q) \right| . \tag{C.2}$$

Proof. Let $f_a(q)$ denote the solution to

$$\forall q \in [-\pi, \pi], \quad g(q) f_a'(q) - c\, f_a(q) = h(q), \qquad f_a(-\pi) = a.$$

Since g is smooth and everywhere positive, this equation admits a unique smooth solution on the interval $[-\pi, \pi]$, by the standard theory of ordinary differential equations. Rewriting the equation for f_a as

$$g(q) \exp\left(\int_{-\pi}^{q} \frac{c}{g(x)} \, \mathrm{d}x \right) \frac{\mathrm{d}}{\mathrm{d}q} \left[f_a(q) \, \exp\left(- \int_{-\pi}^{q} \frac{c}{g(x)} \, \mathrm{d}x \right) \right] = h(q),$$

it is clear that

$$f_a(q) = \exp\left(\int_{-\pi}^{q} \frac{c}{g(x)} \, \mathrm{d}x \right) \left[a + \int_{-\pi}^{q} \exp\left(- \int_{-\pi}^{y} \frac{c}{g(x)} \, \mathrm{d}x \right) \frac{h(y)}{g(y)} \, \mathrm{d}y \right] . \tag{C.3}$$

The unique value a^* for which the function is periodic (namely $f_{a^*}(\pi) = a^*$) is

$$a^* = - \left[1 - \exp\left(- \int_{-\pi}^{\pi} \frac{c}{g(x)} \, \mathrm{d}x \right) \right]^{-1} \int_{-\pi}^{\pi} \exp\left(- \int_{-\pi}^{y} \frac{c}{g(x)} \, \mathrm{d}x \right) \frac{h(y)}{g(y)} \, \mathrm{d}y.$$

It is easily checked that all derivatives of f_{a^*} are also continuous, so that $f_{a^*} \in C^\infty(\mathbf{T})$.

To obtain the estimate (C.2), we introduce $r(q) := f(q) + c^{-1} h(q) + c^{-2} g(q)\, h'(q)$ and note by that, by (C.1),

$$g(q)\, r'(q) - c\, r(q) = \frac{1}{c^2}\, g(q)\, (g\, h')'(q).$$

The term $r'(q)$ vanishes at the extrema of $r(q)$, which leads then to (C.2). □

We next prove another technical result, which we then employ to show that the leading order term $\overline{\phi}_0$ in the underdamped limit, defined in (4.20), indeed belongs to $H^1(\mu)$ when V is not constant.

Lemma C.2. *If $q \mapsto V(q)$ is not constant, then there exists $M > 0$ such that*

$$\forall E \in (E_0, \infty), \qquad S_\nu(E) \geqslant \frac{\nu^2}{1+\nu^2 M} \int_{\mathbf{T}} \sqrt{2(E_0 - V(q))}\, \mathrm{d}q. \tag{C.4}$$

In particular, $S_\nu(E)$ is bounded below by a positive constant uniformly in $E \in (E_0, \infty)$.

Proof. Fix $E > E_0$. We integrate (4.15) over $\mathbf{T}$ and use integration by parts for the first term, which gives

$$\int_{\mathbf{T}} \left(\frac{1}{\nu^2} + \frac{\partial P}{\partial q}(q, E) \right) s_\nu(q, E)\, \mathrm{d}q = \int_{\mathbf{T}} P(q, E)\, \mathrm{d}q. \tag{C.5}$$

We now show that that there exists a positive constant M such that

$$\forall (q, E) \in \mathbf{T} \times (E_0, \infty), \qquad \left| \frac{\partial P}{\partial q}(q, E) \right| = \frac{|V'(q)|}{\sqrt{2(E - V(q))}} \leqslant M. \tag{C.6}$$

To this end, notice that $|\partial_q P(q, E)|$ is a decreasing function of E for fixed q, so it is sufficient to show that

$$\sup_{q \in \mathbf{T}} \left(\lim_{E \to E_0^+} \left| \frac{\partial P}{\partial q}(q, E) \right| \right) < \infty.$$

Since V is smooth, it holds that $V'(q) = 0$ for any q such that $V(q) = E_0$, so

$$L(q) := \lim_{E \to E_0^+} \left| \frac{\partial P}{\partial q}(q, E) \right| = \begin{cases} 0 & \text{if } V(q) = E_0, \\ \dfrac{|V'(q)|}{\sqrt{2(E_0 - V(q))}} & \text{otherwise.} \end{cases}$$

By using L'Hôpital's rule, we notice that, for values of q in a neighborhood of any $q^* \in V^{-1}\{E_0\}$ with $V'(q) \neq 0$,

$$\lim_{q \to q^*} \frac{|V'(q)|}{\sqrt{2(E_0 - V(q))}} = \sqrt{\lim_{q \to q^*} \frac{|V'(q)|^2}{2(E_0 - V(q))}} = \sqrt{- \lim_{q \to q^*} \frac{V'(q)\, V''(q)}{V'(q)}} = \sqrt{- \lim_{q \to q^*} V''(q)},$$

so $L(q)$ is bounded uniformly from above, implying that (C.6) holds. This concludes the proof because, by (C.5), and recalling that $s_\nu \geqslant 0$ by (4.16),

$$\begin{aligned} \int_{\mathbf{T}} s_\nu(q, E)\, \mathrm{d}q &\geqslant \int_{\mathbf{T}} \frac{\nu^{-2} + |\partial_q P(q, E)|}{\nu^{-2} + M} s_\nu(q, E)\, \mathrm{d}q \\ &\geqslant \frac{\nu^2}{1+\nu^2 M} \int_{\mathbf{T}} \left(\frac{1}{\nu^2} + \frac{\partial P}{\partial q}(q, E) \right) s_\nu(q, E)\, \mathrm{d}q = \frac{\nu^2}{1+\nu^2 M} \int_{\mathbf{T}} P(q, E)\, \mathrm{d}q \\ &\geqslant \frac{\nu^2}{1+\nu^2 M} \int_{\mathbf{T}} P(q, E_0)\, \mathrm{d}q, \end{aligned}$$

and the last integral is positive because V is not constant. □

Lemma C.3. *If $q \mapsto V(q)$ is not constant, then the function $\overline{\phi}_0$ defined in (4.19) belongs to $H^1(\mu)$.*

Proof. It is easy to see that the function $\overline{\phi}_0$ defined in (4.19) belongs to $L^2(\mu)$. We next consider the distributional derivatives $\partial_y \overline{\phi}_0$ for $y \in \{q, p\}$, since derivatives in z vanish: for $H(q,p) < E_0$, these derivatives vanish; while for $H(q,p) > E_0$, it holds $\partial_q \overline{\phi}_0 = \operatorname{sign}(p) \psi_0'(H(q,p)) V'(q)$ and $\partial_p \overline{\phi}_0 = \psi_0'(H(q,p))|p|$. Note that all these derivatives belong to $L^\infty_{\mathrm{loc}}(\mathbf{T} \times \mathbf{R} \times \mathbf{R})$ in view of the lower bound provided by (C.4). Moreover, $\left|\nabla \overline{\phi}_0(q,p,z)\right|$ grows sufficiently slowly as $|p| \to \infty$. This allows to conclude that $\overline{\phi}_0 \in H^1(\mu)$, as claimed. □

References

[1] F. Achleitner, A. Arnold, and D. Stürzer. Large-time behavior in non-symmetric Fokker-Planck equations. *Riv. Math. Univ. Parma*, 6(1):1–68, 2015.

[2] W. Arendt, C. J. K. Batty, M. Hieber, and F. Neubrander. *Vector-valued Laplace transforms and Cauchy problems*, volume 96 of *Monographs in Mathematics*. Birkhäuser Verlag, Basel, 2001.

[3] S. Armstrong and J.-C. Mourrat. Variational methods for the kinetic Fokker-Planck equation. *arXiv preprint*, 1902.04037, 2019.

[4] A. Arnold, C. Schmeiser, and B. Signorello. Propagator norm and sharp decay estimates for Fokker–Planck equations with linear drift. *arXiv preprint*, 2003.01405, 2020.

[5] A. Barone and G. Paterno. *Physics and Applications of the Josephson Effect*. Wiley, New York, 1982.

[6] F. Baudoin. Bakry-Émery meet Villani. *J. Funct. Anal.*, 273(7):2275–2291, 2017.

[7] A. Bensoussan, J.-L. Lions, and G. Papanicolaou. *Asymptotic Analysis for Periodic Structures*. AMS Chelsea Publishing, 2011.

[8] E. Bernard, M. Fathi, A. Levitt, and G. Stoltz. Hypocoercivity with Schur complements. *arXiv preprint*, 2003.00726, 2020.

[9] Y. Cao, J. Lu, and L. Wang. On explicit L^2-convergence rate estimate for underdamped Langevin dynamics. *arXiv preprint*, 1908.04746, 2019.

[10] P. Cattiaux, A. Guillin, P. Monmarché, and C. Zhang. Entropic multipliers method for Langevin diffusion and weighted log Sobolev inequalities. *J. Funct. Anal.*, 277(11), 2019.

[11] M. Ceriotti. *A Novel Framework for Enhanced Molecular Dynamics Based on the Generalized Langevin Equation*. PhD thesis, ETH Zurich, 2010.

[12] M. Ceriotti, G. Bussi, and M. Parrinello. Langevin equation with colored noise for constant-temperature molecular dynamics simulations. *Phys. Rev. Lett.*, 102:020601, 2009.

[13] M. Ceriotti, G. Bussi, and M. Parrinello. Colored-noise thermostats à la carte. *Journal of Chemical Theory and Computation*, 6(4):1170–1180, 2010.

[14] M. Chak, N. Kantas, and G. A. Pavliotis. On the generalised Langevin equation for simulated annealing. *arXiv preprint*, 2003.06448, 2020.

[15] L. Desvillettes and C. Villani. On the trend to global equilibrium in spatially inhomogeneous entropy-dissipating systems: The linear Fokker-Planck equation. *Comm. Pure Appl. Math.*, 54(1):1–42, 2001.

[16] J. Dolbeault, A. Klar, C. Mouhot, and C. Schmeiser. Exponential rate of convergence to equilibrium for a model describing fiber lay-down processes. *Appl. Math. Res. eXpress*, 2013(2):165–175, 2013.

[17] J. Dolbeault, C. Mouhot, and C. Schmeiser. Hypocoercivity for kinetic equations with linear relaxation terms. *C. R. Math. Acad. Sci. Paris*, 347(9-10):511–516, 2009.

[18] J. Dolbeault, C. Mouhot, and C. Schmeiser. Hypocoercivity for linear kinetic equations conserving mass. *Trans. Amer. Math. Soc.*, 367(6):3807–3828, 2015.

[19] J. D. Doll, L. E. Myers, and S. A. Adelman. Generalized Langevin equation approach for atom/solid-surface scattering: Inelastic studies. *J. Chem. Phys.*, 63(11):4908–4914, 1975.

[20] A. Eberle, A. Guillin, and R. Zimmer. Coupling and quantitative contraction rates for Langevin dynamics. *Ann. Probab.*, 47(4):1982–2010, 2019.

[21] J.-P. Eckmann and M. Hairer. Non-equilibrium statistical mechanics of strongly anharmonic chains of oscillators. *Comm. Math. Phys.*, 212(1):105–164, 2000.

[22] J.-P. Eckmann, C.-A. Pillet, and L. Rey-Bellet. Entropy production in nonlinear, thermally driven Hamiltonian systems. *J. Statist. Phys.*, 95(1-2):305–331, 1999.

[23] J.-P. Eckmann, C.-A. Pillet, and L. Rey-Bellet. Non-equilibrium statistical mechanics of anharmonic chains coupled to two heat baths at different temperatures. *Comm. Math. Phys.*, 201(3):657–697, 1999.

[24] M. Fathi, A.-A. Homman, and G. Stoltz. Error analysis of the transport properties of Metropolized schemes. *ESAIM Proc.*, 48:341–363, 2015.

[25] M. Freidlin and M. Weber. A remark on random perturbations of the nonlinear pendulum. *Ann. Appl. Probab.*, 9(3):611–628, 1999.

[26] B. Gidas. Global optimization via the Langevin equation. In *1985 24th IEEE Conference on Decision and Control*, pages 774–778. IEEE, 1985.

[27] R. Gomer. Diffusion of adsorbates on metal surfaces. *Rep. Prog. Phys.*, 53(7):917, 1990.

[28] M. Grothaus and P. Stilgenbauer. Hilbert space hypocoercivity for the Langevin dynamics revisited. *Methods Funct. Anal. Topology*, 22(2):152–168, 2016.

[29] M. Hairer and J. C. Mattingly. Yet another look at Harris' ergodic theorem for Markov chains. In *Seminar on Stochastic Analysis, Random Fields and Applications VI*, volume 63 of *Progr. Probab.*, pages 109–117. Birkhäuser/Springer Basel AG, Basel, 2011.

[30] M. Hairer and G. A. Pavliotis. From ballistic to diffusive behavior in periodic potentials. *J. Stat. Phys.*, 131(1):175–202, 2008.

[31] F. Hérau. Hypocoercivity and exponential time decay for the linear inhomogeneous relaxation Boltzmann equation. *Asymptot. Anal.*, 46(3-4):349–359, 2006.

[32] F. Hérau. Short and long time behavior of the Fokker–Planck equation in a confining potential and applications. *J. Funct. Anal.*, 244(1):95–118, 2007.

[33] F. Hérau. Introduction to hypocoercive methods and applications for simple linear inhomogeneous kinetic models. In *Lectures on the Analysis of Nonlinear Partial Differential Equations. Part 5*, volume 5 of *Morningside Lect. Math.*, pages 119–147. Int. Press, Somerville, MA, 2018.

[34] A. Iacobucci, S. Olla, and G. Stoltz. Convergence rates for nonequilibrium Langevin dynamics. *Ann. Math. Qué.*, 43(1):73–98, 2019.

[35] A. Igarashi, P. V. E. McClintock, and N. G. Stocks. Velocity spectrum for non-Markovian Brownian motion in a periodic potential. *J. Statist. Phys.*, 66(3-4):1059–1070, 1992.

[36] A. Igarashi and T. Munakata. Non-Markovian Brownian motion in a periodic potential. *J. Phys. Soc. Jpn.*, 57(7):2439–2447, 1988.

[37] E. Jones, T. Oliphant, P. Peterson, et al. SciPy: Open source scientific tools for Python, 2001–2018.

[38] T. Kato. *Perturbation Theory for Linear Operators.* Classics in Mathematics. Springer-Verlag, Berlin, 1995.

[39] J. Kierzenka and L. F. Shampine. A BVP solver based on residual control and the MATLAB PSE. *ACM Trans. Math. Software*, 27(3):299–316, 2001.

[40] M. Kopec. Weak backward error analysis for Langevin process. *BIT*, 55(4):1057–1103, 2015.

[41] R. Kupferman, A. M. Stuart, J. R. Terry, and P. F. Tupper. Long-term behaviour of large mechanical systems with random initial data. *Stoch. Dynam.*, 2(4):1–30, 2002.

[42] A. M. Lacasta, J. M. Sancho, A. H. Romero, I. M. Sokolov, and K. Lindenberg. From subdiffusion to superdiffusion of particles on solid surfaces. *Phys. Rev. E*, 70:051104, 2004.

[43] B. Leimkuhler, C. Matthews, and G. Stoltz. The computation of averages from equilibrium and nonequilibrium Langevin molecular dynamics. *IMA J. Numer. Anal.*, 36(1):13–79, 2016.

[44] B. Leimkuhler and M. Sachs. Ergodic properties of quasi-Markovian generalized Langevin equations with configuration dependent noise and non-conservative force. In G. Giacomin, S. Olla, E. Saada, H. Spohn, and G. Stoltz, editors, *Stochastic dynamics out of equilibrium*, volume 282 of *Springer Proc. Math. Stat.*, pages 282–330. Springer, Cham, 2019.

[45] T. Lelièvre and G. Stoltz. Partial differential equations and stochastic methods in molecular dynamics. *Acta Numer.*, 25:681–880, 2016.

[46] V. Letizia and S. Olla. Non-equilibrium isothermal transformations in a temperature gradient from a microscopic dynamics. *Ann. Probab.*, 45(6A):3987–4018, 2017.

[47] A. Lindquist and G. Picci. Realization theory for multivariate stationary Gaussian processes. *SIAM J. Control Optim.*, 23(6):809–857, 1985.

[48] L. Lorenzi and M. Bertoldi. *Analytical Methods for Markov Semigroups*, volume 283 of *Pure and Applied Mathematics (Boca Raton)*. Chapman & Hall/CRC, Boca Raton, FL, 2007.

[49] J. C. Mattingly and A. M. Stuart. Geometric ergodicity of some hypo-elliptic diffusions for particle motions. *Markov Proc. Rel. Fields*, 8(2):199–214, 2002.

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

[51] G. Metafune, D. Pallara, and E. Priola. Spectrum of Ornstein-Uhlenbeck operators in L^p spaces with respect to invariant measures. *J. Funct. Anal.*, 196(1):40–60, 2002.

[52] H. Mori. A continued-fraction representation of the time-correlation functions. *Progr. Theoret. Phys.*, 34:399–416, 1965.

[53] H. Mori. Transport, collective motion, and Brownian motion. *Progr. Theoret. Phys.*, 33(3):423–455, 1965.

[54] C. Mouhot and L. Neumann. Quantitative perturbative study of convergence to equilibrium for collisional kinetic models in the torus. *Nonlinearity*, 19(4):969–998, 2006.

[55] M. Ottobre and G. A. Pavliotis. Asymptotic analysis for the generalized Langevin equation. *Nonlinearity*, 24(5):1629–1653, 2011.

[56] M. Ottobre, G. A. Pavliotis, and K. Pravda-Starov. Exponential return to equilibrium for hypoelliptic quadratic systems. *J. Funct. Anal.*, 262(9):4000–4039, 2012.

[57] G. A. Pavliotis. *Stochastic Processes and Applications*, volume 60 of *Texts in Applied Mathematics*. Springer, New York, 2014.

[58] G. A. Pavliotis and A. M. Stuart. *Multiscale Methods: Averaging and Homogenization*, volume 53 of *Texts in Applied Mathematics*. Springer, New York, 2008.

[59] G. A. Pavliotis and A. Vogiannou. Diffusive transport in periodic potentials: underdamped dynamics. *Fluct. Noise Lett.*, 8(2):L155–L173, 2008.

[60] M. Reed and B. Simon. *Methods of Modern Mathematical Physics. Analysis of Operators*, volume IV. Academic Press, 1978.

[61] P. Reimann. Brownian motors: noisy transport far from equilibrium. *Phys. Rep.*, 361(2-4):57–265, 2002.

[62] L. Rey-Bellet. Ergodic properties of Markov processes. In S. Attal, A. Joye, and C.-A. Pillet, editors, *Open Quantum Systems II*, volume 1881 of *Lecture Notes in Mathematics*, pages 1–39. Springer, 2006.

[63] L. Rey-Bellet. Open classical systems. In S. Attal, A. Joye, and C.-A. Pillet, editors, *Open Quantum Systems II*, volume 1881 of *Lecture Notes in Mathematics*, pages 41–78. Springer, 2006.

[64] L. Rey-Bellet and L. E. Thomas. Exponential convergence to non-equilibrium stationary states in classical statistical mechanics. *Comm. Math. Phys.*, 225(2):305–329, 2002.

[65] H. Risken. *The Fokker-Planck Equation*, volume 18 of *Springer Series in Synergetics*. Springer-Verlag, Berlin, second edition, 1989.

[66] J. Roussel and G. Stoltz. Spectral methods for Langevin dynamics and associated error estimates. *ESAIM: Math. Model. Numer. Anal.*, 52(3):1051–1083, 2018.

[67] J. Roussel and G. Stoltz. A perturbative approach to control variates in molecular dynamics. *Multiscale Model. Simul.*, 17(1):552–591, 2019.

[68] Y. Saad and M. H. Schultz. GMRES: a generalized minimal residual algorithm for solving nonsymmetric linear systems. *SIAM J. Sci. Statist. Comput.*, 7(3):856–869, 1986.

[69] J. M. Sancho, A. M. Lacasta, K. Lindenberg, I. M. Sokolov, and A. H. Romero. Diffusion on a solid surface: anomalous is normal. *Phys. Rev. Lett.*, 92(25):250601, 2004.

[70] I. Snook. *The Langevin and Generalised Langevin Approach to the Dynamics of Atomic, Polymeric and Colloidal Systems*. Elsevier, 2006.

[71] R. L. Stratonovich. *Topics in the Theory of Random Noise. Vol. I.* Gordon and Breach Science Publishers, New York, 1963.

[72] R. L. Stratonovich. *Topics in the theory of random noise. Vol. II.* Gordon and Breach Science Publishers, New York, 1967.

[73] D. Talay. Stochastic Hamiltonian systems: exponential convergence to the invariant measure, and discretization by the implicit Euler scheme. *Markov Proc. Rel. Fields*, 8(2):163–198, 2002.

[74] T. Tang. The Hermite spectral method for Gaussian-type functions. *SIAM J. Sci. Comput.*, 14(3):594–606, 1993.

[75] U. Vaes. *Topics in Multiscale Modelling: Numerical Analysis and Applications*. PhD thesis, Imperial College London, 2019.

[76] C. Villani. Hypocoercivity. *Mem. Amer. Math. Soc.*, 202(950), 2009.

[77] R. Zwanzig. Nonlinear generalized Langevin equations. *J. Stat. Phys.*, 9(3):215–220, 1973.