Overdamped limit at stationarity for non-equilibrium Langevin diffusions[*]

Pierre Monmarché[†]   Mouad Ramil[‡]

Abstract

In this note, we establish that the stationary distribution of a possibly non-equilibrium Langevin diffusion converges, as the damping parameter goes to infinity (or equivalently in the Smoluchowski-Kramers vanishing mass limit), toward a tensor product of the stationary distribution of the corresponding overdamped process and of a Gaussian distribution.

Keywords: Langevin diffusion; overdamped limit; Smoluchowski-Kramers limit; Wasserstein distance.
MSC2020 subject classifications: 60J60; 82C31.
Submitted to ECP on October 4, 2021, final version accepted on January 8, 2022.

1 Introduction

Consider the Langevin diffusion on $\mathbb{T}^d \times \mathbb{R}^d$ ($\mathbb{T} = \mathbb{R}/\mathbb{Z}$) solution of

$$\begin{cases} \mathrm{d}X_t &= Y_t \mathrm{d}t \\ \mathrm{d}Y_t &= F(X_t)\mathrm{d}t - \gamma Y_t \mathrm{d}t + \sqrt{2\gamma}\Sigma \mathrm{d}B_t, \end{cases} \tag{1.1}$$

with a standard d-dimensional Brownian motion $(B_t)_{t\geqslant 0}$, a force field $F \in \mathcal{C}^1(\mathbb{T}^d, \mathbb{R}^d)$, a damping parameter $\gamma > 0$ and a constant symmetric positive definite diffusion matrix Σ. For all $\gamma > 0$, this Markov process admits a unique invariant measure μ_γ. Besides, as $\gamma \to +\infty$, $(X_{\gamma t})_{t\geqslant 0}$ converges to the overdamped Langevin diffusion on $\mathbb{T}^d$, solution of

$$\mathrm{d}Z_t = F(Z_t)\mathrm{d}t + \sqrt{2}\Sigma \mathrm{d}B_t\,, \tag{1.2}$$

which admits a unique invariant measure μ_O. These facts are well-known, see e.g. [12]. Up to a rescaling, this overdamped limit is equivalent to the so-called Smoluchowski-Kramers limit, where a mass parameter is sent to zero. In the so-called equilibrium case, in which $F = -\nabla U$ is a conservative force ($U \in \mathcal{C}^2(\mathbb{T}^d)$) and $\Sigma = I$, μ_γ and μ_O are explicit, and more precisely, for all $\gamma > 0$, $\mu_\gamma = \mu_O \otimes g$ with

$$\mu_O(\mathrm{d}x) = \frac{e^{-U(x)}}{\int_{\mathbb{T}^d} e^{-U(w)}\mathrm{d}w}\mathrm{d}x\,, \qquad g(\mathrm{d}y) = \frac{e^{-|y|^2/2}}{(2\pi)^{d/2}}\mathrm{d}y\,.$$

[†]LJLL and LCT, Sorbonne Université, Paris, France. E-mail: pierre.monmarche@sorbonne-universite.fr
[‡]CERMICS, Ecole des Ponts, Marne-la-Vallée, France. E-mail: ramil.mouad@gmail.com

Here and in the rest of this work, $\mathbb{T}^d$ is endowed with the Lebesgue measure $\mathrm{d}x$ normalized so that its total mass is 1. In the general non-equilibrium case, however, there is no reason for any of this to be true, namely μ_O and μ_γ are not explicit, μ_γ does depend on γ, it is not a tensor product and its position marginal is not μ_O (see Section 2). The main result of this note is that, nevertheless, these properties are recovered in the overdamped limit, i.e. that μ_γ converges as $\gamma \to +\infty$ toward $\mu_O \otimes g_\Sigma$ where g_Σ is the centered Gaussian distribution on $\mathbb{R}^d$ with covariance matrix Σ^2. The convergence is quantitative, and stated in term of Wasserstein distance. The Wasserstein distance between two probability laws μ, ν on a metric space (E, dist) is defined as

$$\mathcal{W}(\mu, \nu) = \inf_{\pi \in \Gamma(\mu, \nu)} \int_{E^2} \mathrm{dist}(y, z) \pi(\mathrm{d}z, \mathrm{d}y) \,,$$

where $\Gamma(\mu, \nu)$ is the set of probability measures on E^2 with marginals μ and ν. Our result is the following.

Theorem 1.1. *There exists $C > 0$, which depends on F, Σ and d, such that, for all $\gamma \geqslant 2$,*

$$\mathcal{W}(\mu_\gamma, \mu_O \otimes g_\Sigma) \leqslant C \frac{\sqrt{\log \gamma}}{\gamma} \,.$$

Remark 1.2. We consider a position in the compact torus since this is often the case considered in previous works on non-equilibrium Langevin processes, see e.g. [9]. Moreover, standard molecular dynamics simulations use periodic boundary conditions [11, 12, 15]. However, it is straightforward to adapt the proof to a position on $\mathbb{R}^d$ under the additional assumptions that there exist $R, \lambda > 0$ such that $F(x) \cdot x \leqslant -\lambda |x|^2$ for all $|x| \geqslant R$ and that the Jacobian matrix of F is bounded.

This note is organized as follows. We finish this introduction by discussing applications and previous related works. In Section 2 we characterize the cases where, as in the equilibrium case, position and velocity are independent under μ_γ, or μ_γ is independent from γ. Section 3 is devoted to the proof of Theorem 1.1.

Non-equilibrium Langevin processes appear in various situations where a physical system undergoes an external force or a gradient of temperature, see e.g. [6, 1, 13, 15, 5] or [12, Chapter 5] and references within. In these cases, macroscopic quantities of interest such as transport coefficients (e.g. mobility, thermal conductivity or shear viscosity) can be expressed in term of expectations with respect to the stationary state of the non-equilibrium process, see e.g. [5] for details.

More precisely, with $\Sigma = I$, the question of the linear response of a system at equilibrium to a non-conservative perturbation leads to the study of forces of the form $F(x) = -\nabla U(x) + \tau \tilde{F}(x)$ where $\tilde{F}$ is non-conservative and τ is small. For instance, [12, Theorem 5.2] gives an asymptotic expansion of μ_γ as τ vanishes (at a fixed γ). Similarly, the long-time convergence of non-equilibrium Langevin processes is studied in [9] in this case, for τ small enough. Our result is not restricted to this perturbative regime.

Chains of oscillators such as those studied in [1, 13] and references within, which are models for thermal conductivity, correspond to a case where $F(x) = -\nabla U(x)$ for some potential U but Σ is not an homothety. Hence, the Jacobian matrix of $\Sigma^2 F$ is not symmetric, which shows that $\Sigma^2 F$ cannot be a gradient (which means this is a non-equilibrium process, cf. Section 2).

For more details and a general discussion on non-equilibrium Langevin processes from a mathematical perspective, we refer to [12, Chapter 5].

To the best of our knowledge, Theorem 1.1, or even the weak convergence of the first marginal of μ_γ toward μ_O, is new. A similar result as been proven by one of the

authors in [14] but, rather than the stationary measure of the process, it concerns the quasi-stationary measure of the process killed when it exists a given domain. This is a different and more difficult problem, besides the convergence in [14] is not quantitative. In infinite dimension, a result similar to Theorem 1.1 (although not quantitative and concerning only the convergence of the position marginal), with some similarity in the proof, is established in [2] for stochastic damped wave equations.

Distinct but somewhat related results are the overdamped limit of the diffusion coefficient proven in [8, 11] in dimension 1 and of the large deviation quasi-potential in [3]. Indeed, these two works study the overdamped limit of a stationary quantity in some asymptotic regime, either $\tau \to 0$ (with $F = -\nabla U + \tau \tilde{F}$) or $\Sigma \to 0$.

2 Particular cases

In this section, we investigate under which conditions on F and Σ each of the following properties holds true: a) μ_γ is independent from γ, b) μ_γ is the tensor product of two probability laws respectively on $\mathbb{T}^d$ and $\mathbb{R}^d$, and c) the first marginal of μ_γ is μ_O. Before addressing the general case, let us highlight the following two simple cases:

- *The equilibrium case.* If $F(x) = -\Sigma^2 \nabla U(x)$ for some $U \in \mathcal{C}^2(\mathbb{T}^d)$ then the properties a), b) and c) hold since it is easily checked that

$$\mu_\gamma(\mathrm{d}x, \mathrm{d}y) = \frac{e^{-U(x) - |\Sigma^{-1}y|^2/2}}{(2\pi)^{d/2} \det(\Sigma) \int_{\mathbb{T}^d} e^{-U(w)} \mathrm{d}w} \mathrm{d}x \mathrm{d}y$$

  for all $\gamma > 0$. Considering on $\Sigma^{-1}\mathbb{T}^d$ the potential $\tilde{U}(z) = U(\Sigma z)$, the process $(Z, W) = (\Sigma^{-1}X, \Sigma^{-1}Y)$ on $(\Sigma^{-1}\mathbb{T}^d, \mathbb{R}^d)$ solves

$$\begin{cases} \mathrm{d}Z_t &= W_t \mathrm{d}t \\ \mathrm{d}W_t &= -\nabla \tilde{U}(Z_t)\mathrm{d}t - \gamma W_t \mathrm{d}t + \sqrt{2}\mathrm{d}B_t\,. \end{cases}$$

- *The space-homogeneous case.* If $F(x) = \eta$ for all $x \in \mathbb{T}^d$ for some constant $\eta \in \mathbb{R}^d$, then

$$\mu_\gamma(\mathrm{d}x, \mathrm{d}y) = \frac{e^{-|\Sigma^{-1}(y - \eta/\gamma)|^2/2}}{(2\pi)^{d/2} \det(\Sigma)} \mathrm{d}x \mathrm{d}y$$

  for all $\gamma > 0$. In particular, the properties b) and c) hold but a) holds iff $\eta = 0$ (which corresponds to the equilibrium case with a constant U). Moreover, since the Wasserstein distance between two Gaussian distributions with the same variance is given by the distance between their means (see e.g. [7, Proposition 7]),

$$\mathcal{W}(\mu_\gamma, \mu_O \otimes g_\Sigma) = \frac{|\Sigma^{-1}\eta|}{\gamma}\,.$$

Let us state a characterization of the cases where a) or b) hold.

Proposition 2.1. *First, the two following properties are equivalent.*

(i) *There exist $\gamma_1 > \gamma_2 > 0$ such that $\mu_{\gamma_1} = \mu_{\gamma_2}$.*

(ii) *There exists $U \in \mathcal{C}^2(\mathbb{T}^d)$ such that $F = -\Sigma^2 \nabla U$.*

Second, the two following properties are equivalent.

(iii) *There exists $\gamma > 0$ such that μ_γ is the tensor product of two probability laws respectively on $\mathbb{T}^d$ and $\mathbb{R}^d$.*

(iv) There exists $\eta \in \mathbb{R}^d$ and $U \in \mathcal{C}^2(\mathbb{T}^d)$ such that, for all $x \in \mathbb{T}^d$, $F(x) = -\Sigma^2 \nabla U(x) + \eta$ and $\nabla U(x) \cdot \eta = 0$.

Moreover, if (iv) holds, then

$$\mu_O(\mathrm{d}x) = \frac{e^{-U(x)}}{\int_{\mathbb{T}^d} e^{-U(w)}\mathrm{d}w}\mathrm{d}x \qquad \text{and for all } \gamma > 0, \qquad \mu_\gamma = \mu_O \otimes \frac{e^{-|\Sigma^{-1}(y-\eta/\gamma)|^2/2}}{(2\pi)^{d/2}\det(\Sigma)}\mathrm{d}y.$$

Before proving this result, let us discuss the condition $\nabla U \cdot \eta = 0$ in (iv), which is equivalent to say that $U(x + t\eta) = U(x)$ for all $x \in \mathbb{T}^d, t \in \mathbb{R}$. Since U is $\mathbb{Z}$-periodic, depending on η, this may be more or less restrictive. For instance:

- If $\mathbb{R}\eta$ is dense in $\mathbb{T}^d$ (which is in particular the case if $d = 1$ and $\eta \neq 0$), then necessarily U is constant, hence so is F.

- If $\eta = \tau\mathbf{e}$ with $\tau \in \mathbb{R}$ and $\mathbf{e}$ a vector of the canonical basis (say, the first one), then the condition is satisfied as soon as U is periodic function of $x^{\neq 1} = (x_j)_{j \in [\![2,d]\!]}$. If moreover, for instance, $\Sigma = I$, the process can be decomposed as two independent parts, a position homogenous coordinate and the other coordinates at equilibrium, namely

$$\begin{cases} \mathrm{d}X_t^1 &=& Y_t^1\mathrm{d}t \\ \mathrm{d}Y_t^1 &=& \tau\mathrm{d}t - \gamma Y_t^1\mathrm{d}t + \sqrt{2\gamma}\mathrm{d}B_t^1 \\ \mathrm{d}X_t^{\neq 1} &=& Y_t^{\neq 1}\mathrm{d}t \\ \mathrm{d}Y_t^{\neq 1} &=& -\nabla_{\neq 1}U(X^{\neq 1})\mathrm{d}t - \gamma Y_t^{\neq 1}\mathrm{d}t + \sqrt{2\gamma}\mathrm{d}B_t^{\neq 1}\,. \end{cases} \tag{2.1}$$

In any cases, assuming (iv) with $\eta \neq 0$, we see that $(Z,W) = (Q\Sigma^{-1}X, Q\Sigma^{-1}W)$, the process on $Q\Sigma^{-1}\mathbb{T}^d \times \mathbb{R}^d$ where Q is an orthonomal matrix that sends $\eta/|\eta|$ to $\mathbf{e}$, solves an equation of the form (2.1). In other words, apart from the equilibrium case, the only cases where μ_γ is of a tensor form for some $\gamma > 0$ is a combination of a one-dimensional space homogeneous system with an independent $d - 1$ dimensional equilibrium system.

Proof. The Langevin process being hypoelliptic and irreducible, its invariant measure admits a smooth positive density with respect to the Lebesgue measure, which we write $e^{-H(x,y)}$ (H possibly depends on γ but we do not explicit this dependency to simplify notation). The invariance of μ_γ is equivalent to

$$\forall f \in \mathcal{C}^\infty(\mathbb{T}^d \times \mathbb{R}^d), \qquad \int_{\mathbb{T}^d \times \mathbb{R}^d} Lf(x,y)e^{-H(x,y)}\mathrm{d}x\mathrm{d}y = 0\,,$$

where, using the notation $\Sigma^2 : \nabla_y^2 = \sum_{i,j=1}^d (\Sigma^2)_{i,j}\partial_{y_i}\partial_{y_j}$,

$$L = y \cdot \nabla_x + (F(x) - \gamma y) \cdot \nabla_y + \gamma\Sigma^2 : \nabla_y^2 \tag{2.2}$$

is the generator of the process (1.1). Integrating by parts, this is equivalent to

$$\forall(x,y) \in \mathbb{T}^d \times \mathbb{R}^d, \qquad 0 = y \cdot \nabla_x H(x,y) + (F(x) - \gamma y) \cdot \nabla_y H(x,y)$$
$$+ \gamma d + \gamma|\Sigma\nabla_y H(x,y)|^2 - \gamma\Sigma^2 : \nabla_y^2 H(x,y). \tag{2.3}$$

The implications $(ii) \to (i)$ and $(iv) \to (iii)$ immediately follow from the fact that (2.3) holds with $H(x,y) = U(x) + |\Sigma^{-1}(y - \eta/\gamma)|^2/2$ if $F = -\Sigma^2\nabla U + \eta$ and $\nabla U \cdot \eta = 0$.

Let us prove that $(i) \to (ii)$. Assuming (i) and applying (2.3) for $\gamma \in \{\gamma_1, \gamma_2\}$ yields

$$\forall(x,y) \in \mathbb{T}^d \times \mathbb{R}^d \quad \begin{cases} y \cdot \nabla_x H(x,y) + F(x) \cdot \nabla_y H(x,y) &=& 0 \\ -y \cdot \nabla_y H(x,y) + d + |\Sigma\nabla_y H(x,y)|^2 - \Sigma^2 : \nabla_y^2 H(x,y) &=& 0\,. \end{cases}$$

The second equation is equivalent to say that for all $x \in \mathbb{T}^d$, $y \mapsto e^{-H(x,y)} / \int_{\mathbb{R}^d} e^{-H(x,z)} \mathrm{d}z$ is the invariant probability density of the Ornstein-Uhlenbeck process $\mathrm{d}Y = -Y \mathrm{d}t + \sqrt{2}\Sigma \mathrm{d}B_t$, in other words is the d-dimensional Gaussian distribution with mean 0 and covariance matrix Σ^2. This means that $H(x,y) = U(x) + |\Sigma^{-1}y|^2/2$ for some $U \in \mathcal{C}^1(\mathbb{T}^d)$. Plugging this in the first equation yields $F(x) = -\Sigma^2 \nabla U(x)$ for all $x \in \mathbb{T}^d$.

We now turn to the proof of the implication $(iii) \rightarrow (iv)$. Let γ be such that μ_γ is a tensor product, namely $H(x,y) = U(x) + V(y)$ for some $U \in \mathcal{C}^1(\mathbb{T}^d)$, $V \in \mathcal{C}^1(\mathbb{R}^d)$. Let $x_* \in \mathbb{T}^d$ be a critical point of U. Applying (2.3) for $x = x_*$ reads

$$\forall y \in \mathbb{R}^d \qquad (F(x_*) - \gamma y) \cdot \nabla V(y) + \gamma d + \gamma |\Sigma \nabla V(y)|^2 - \gamma \Sigma : \nabla^2 V(y) = 0\,,$$

in other words e^{-V} is an invariant measure for the Ornstein-Uhlenbeck process

$$\mathrm{d}Y_t = (F(x_*) - \gamma Y_t)\mathrm{d}t + \sqrt{2\gamma}\Sigma \mathrm{d}B_t\,.$$

Hence, $V(y) = |\Sigma^{-1}(y - F(x_*)/\gamma)|^2/2$. Plugging this back in (2.3) yields

$$\forall (x,y) \in \mathbb{T}^d \times \mathbb{R}^d, \qquad y \cdot \nabla_x U(x) + (F(x) - F(x_*)) \cdot \Sigma^{-2}(y - F(x_*)/\gamma) = 0\,,$$

which is equivalent to

$$\forall x \in \mathbb{T}^d, \qquad F(x) - F(x_*) = -\Sigma^2 \nabla U(x)\,, \qquad (F(x) - F(x_*)) \cdot \Sigma^{-2} F(x_*) = 0\,,$$

and thus $\nabla U(x) \cdot F(x_*) = 0$ for all $x \in \mathbb{T}^d$.

Finally, we prove the last statement of the proposition as follows. The expression of μ_γ under (iv) is proven by checking (2.3). Since the position marginal doesn't depend on γ, the fact that it is necessarily equal to μ_O follows from letting $\gamma \rightarrow +\infty$ in Theorem 1.1 (it can also be checked directly on the overdamped stationary equation). $\qquad \square$

3 Proof of Theorem 1.1

Let $(P_t^\gamma)_{t \geqslant 0}$ be the Markov semi-group associated with the Langevin process (1.1) with friction γ and $(\tilde{P}_t)_{t \geqslant 0}$ be the semi-group associated with the overdamped process (1.2), namely

$$P_t^\gamma f(x,y) = \mathbb{E}_{(x,y)}\left(f(X_t, Y_t)\right)\,, \qquad \tilde{P}_t f(z) = \mathbb{E}_z\left(f(Z_t)\right)$$

for all bounded measurable functions f respectively on $\mathbb{T}^d \times \mathbb{R}^d$ and $\mathbb{T}^d$. If ν is a probability on $\mathbb{T}^d$ then $\nu \tilde{P}_t$ is the law of Z_t if $Z_0 \sim \nu$, and similarly for P_t^γ.

Lemma 3.1. *There exists $C_1, \rho > 0$, which depends only on F, d and Σ, such that, for all probability measures μ, ν on $\mathbb{T}^d$ and all $t \geqslant 0$,*

$$\mathcal{W}(\mu \tilde{P}_t, \nu \tilde{P}_t) \leqslant C_1 e^{-\rho t} \mathcal{W}(\mu, \nu)\,.$$

Proof. This is proven in [4], although the latter is written in $\mathbb{R}^d$ and not $\mathbb{T}^d$. The adaptation is straightforward. $\qquad \square$

The following is a simple fact on Wasserstein distances.

Lemma 3.2. *Let (E_i, r_i) for $i = 1, 2$ be two metric spaces and endow $E_1 \times E_2$ with a distance r such that $r((x,y),(z,w)) \geqslant r_1(x,z)$ for all $(x,y),(z,w) \in E_1 \times E_2$. Let μ, ν be two probability measures on $E_1 \times E_2$ and μ^1, ν^1 be their marginal laws on E_1. Then*

$$\mathcal{W}(\mu^1, \nu^1) \leqslant \mathcal{W}(\mu, \nu).$$

Proof. Let $((X,Y),(Z,W)) \sim \pi \in \Gamma(\mu, \nu)$ be a coupling of μ and ν. Then the law of (X,Z) is in $\Gamma(\mu^1, \nu^1)$, and thus

$$\mathcal{W}(\mu^1, \nu^1) \leqslant \mathbb{E}(r_1(X,Z)) \leqslant \mathbb{E}(r((X,Y),(Z,W))).$$

Taking the infimum over $\Gamma(\mu, \nu)$ concludes the proof. $\qquad \square$

Next, we state a moment bound, uniform in $\gamma \geqslant 1$.

Lemma 3.3. *For all $\gamma > 0$,*

$$\mathbb{E}_{\mu_\gamma}\left(|Y|^2\right) \leqslant 2\mathrm{Tr}(\Sigma^2) + \frac{\|F\|_\infty^2}{\gamma^2}\,.$$

Proof. Since μ_γ is a stationary measure for the Langevin process, considering $f(x,v) = |v|^2/2$ and the generator L given by (2.2),

$$0 = \int_{\mathbb{T}^d \times \mathbb{R}^d} Lf(x,v)\mu_\gamma(\mathrm{d}x,\mathrm{d}v) \leqslant \int_{\mathbb{T}^d \times \mathbb{R}^d} \left(\|F\|_\infty|y| - \gamma|y|^2 + \gamma\mathrm{Tr}(\Sigma^2)\right)\mu_\gamma(\mathrm{d}x,\mathrm{d}v)$$

$$\leqslant -\frac{\gamma}{2}\int_{\mathbb{T}^d \times \mathbb{R}^d} |y|^2\mu_\gamma(\mathrm{d}x,\mathrm{d}v) + \gamma\mathrm{Tr}(\Sigma^2) + \frac{\|F\|_\infty^2}{2\gamma}\,,$$

which concludes the proof. $\qquad\square$

The next result is based on some arguments from [14].

Proposition 3.4. *For all $t > 0$, there exists $C_2 > 0$, which depends only on t, F, d and Σ, such that, for all $\gamma \geqslant 2$ and all probability measure ν on $\mathbb{T}^d \times \mathbb{R}^d$, denoting by ν^1 the marginal of ν on $\mathbb{T}^d$,*

$$\mathcal{W}\left(\nu P_{\gamma t}^\gamma, \nu^1 \tilde{P}_t \otimes g_\Sigma\right) \leq C_2\frac{\sqrt{\log(\gamma)}}{\gamma}\left(1 + \mathbb{E}_\nu\left(|Y|\right)\right).$$

Before proving this proposition, let us first conclude the proof of our main result.

Proof of Theorem 1.1. Notice that in the statement of the theorem, the distance used on $\mathbb{T}^d \times \mathbb{R}^d$ is implicit, namely it is any distance equivalent to the standard $r((x,y),(z,w)) = \mathrm{dist}(x,z) + |y-w|$, where dist is the image of the Euclidean norm $|\cdot|$ on the torus. We prove the theorem with this distance.

Considering C_1 and ρ as in Lemma 3.1, let $t = \log(2C_1)/\rho$. Using that μ_γ and μ_O are invariant respectively for $P_{\gamma t}^\gamma$ and $\tilde{P}_t$, and denoting by μ_γ^1 the marginal law of μ_γ on $\mathbb{T}^d$, we get

$$\begin{aligned}
\mathcal{W}\left(\mu_\gamma, \mu_O \otimes g_\Sigma\right) &= \mathcal{W}\left(\mu_\gamma P_{\gamma t}^\gamma, \left(\mu_O \tilde{P}_t\right) \otimes g_\Sigma\right)\\
&\leqslant \mathcal{W}\left(\mu_\gamma P_{\gamma t}^\gamma, \left(\mu_\gamma^1 \tilde{P}_t\right) \otimes g_\Sigma\right) + \mathcal{W}\left(\left(\mu_\gamma^1 \tilde{P}_t\right) \otimes g_\Sigma, \left(\mu_O \tilde{P}_t\right) \otimes g_\Sigma\right)\\
&\leqslant \mathcal{W}\left(\mu_\gamma P_{\gamma t}^\gamma, \left(\mu_\gamma^1 \tilde{P}_t\right) \otimes g_\Sigma\right) + C_1 e^{-\rho t}\mathcal{W}\left(\mu_\gamma^1, \mu_O\right)\\
&\leqslant \mathcal{W}\left(\mu_\gamma P_{\gamma t}^\gamma, \left(\mu_\gamma^1 \tilde{P}_t\right) \otimes g_\Sigma\right) + \frac{1}{2}\mathcal{W}\left(\mu_\gamma, \mu_O \otimes g_\Sigma\right),
\end{aligned}$$

where we used Lemma 3.2 for the last inequality. Proposition 3.4 then yields

$$\mathcal{W}\left(\mu_\gamma, \mu_O \otimes g_\Sigma\right) \leqslant 2C_2\frac{\sqrt{\log(\gamma)}}{\gamma}\left(1 + \mathbb{E}_{\mu_\gamma}\left(|Y|\right)\right)$$

and Lemma 3.3 concludes the proof. $\qquad\square$

Proof of Proposition 3.4. We are going to construct an explicit coupling introduced in [14]. The time $t > 0$ is fixed. Let (X_0, Y_0) be a random variable with law ν and $(B_s)_{s\geqslant 0}$ be a standard Brownian motion independent from (X_0, Y_0). For $\gamma > 0$, let $(X_s^{(\gamma)}, Y_s^{(\gamma)})_{s\geqslant 0}$ be the solution of (1.1) driven by $(B_s)_{s\geqslant 0}$ with initial condition (X_0, Y_0), define the

Brownian motion $(B_s^{(\gamma)})_{s\geq 0} = (B_{\gamma s}/\sqrt{\gamma})_{s\geq 0}$ and let $(\overline{X}_s^{(\gamma)})_{s\geq 0}$ be the overdamped process starting from X_0 and solving (1.2) where the Brownian motion is replaced by the γ-dependent Brownian motion $(B_s^{(\gamma)})_{s\geq 0}$. Let also

$$\forall s \geq 0, \qquad A_s^{(\gamma)} := \sqrt{2}\Sigma e^{-\gamma^2 s}\int_0^{\gamma s} e^{\gamma r}\mathrm{d}B_r. \tag{3.1}$$

Then, for all $s \geq 0$,

$$Y_{\gamma s}^{(\gamma)} = e^{-\gamma^2 s}Y_0 + \gamma e^{-\gamma^2 s}\int_0^s e^{\gamma^2 r}F(X_{\gamma r}^{(\gamma)})\mathrm{d}r + A_s^{(\gamma)}. \tag{3.2}$$

Moreover, let $h_t^{(\gamma)} : [0,t] \mapsto \mathbb{R}$ and the process $(Z_{s,t}^{(\gamma)})_{s\in[0,t]}$ be defined as follow:

$$\forall s \in [0,t], \qquad h_t^{(\gamma)}(s) := \frac{2}{\gamma}\frac{e^{-\gamma^2(t-s)} - e^{-\gamma^2 t}}{1 - e^{-2\gamma^2 t}}, \qquad Z_{s,t}^{(\gamma)} := \Sigma B_s^{(\gamma)} - h_t^{(\gamma)}(s)A_t^{(\gamma)}. \tag{3.3}$$

Let $(\mathcal{F}_s^{(\gamma),Z})_{s\in[0,t]}$ be the natural filtration of $(Z_{s,t}^{(\gamma)})_{s\in[0,t]}$. Using a Itô's fixed point argument, see for instance [10, Thm 2.9 p. 289], the following equation

$$\mathrm{d}W_s^{(\gamma)} = F(W_s^{(\gamma)})\mathrm{d}s + \sqrt{2}\mathrm{d}Z_{s,t}^{(\gamma)}, \qquad W_0^{(\gamma)} = X_0$$

possesses a unique strong solution on $[0,t]$ which is adapted to $(\mathcal{F}_s^{(\gamma),Z})_{s\in[0,t]}$. Moreover, it is shown in [14, Lemma 3.1] that the process $(W_s^{(\gamma)})_{s\in[0,t]}$ is independent of the random variable $A_t^{(\gamma)}$. Denoting by $\mathcal{L}(Q)$ the law of a random variable Q, we bound

$$\mathcal{W}\left(\nu P_{\gamma t}^\gamma, \nu^1 \tilde{P}_t \otimes g_\Sigma\right) \leq \mathcal{W}\left(\nu P_{\gamma t}^\gamma, \mathcal{L}(W_t^{(\gamma)}, A_t^{(\gamma)})\right) + \mathcal{W}\left(\mathcal{L}(W_t^{(\gamma)}) \otimes \mathcal{L}(A_t^{(\gamma)}), \nu^1 \tilde{P}_t \otimes g_\Sigma\right)$$

$$\leq \mathbb{E}\left(\mathrm{dist}\left(X_{\gamma t}^{(\gamma)}, W_t^{(\gamma)}\right) + |Y_{\gamma t}^{(\gamma)} - A_t^{(\gamma)}|\right) + \mathbb{E}\left(\mathrm{dist}\left(\overline{X}_t^{(\gamma)}, W_t^{(\gamma)}\right)\right) + \mathcal{W}\left(\mathcal{L}(A_t^{(\gamma)}), g_\Sigma\right).$$

It remains to bound each of these terms. First, it follows from [14, (i) Lemma 3.2] that there exists a constant $C > 0$ such that for all $\gamma \geqslant 1$,

$$\mathbb{E}\left[\mathrm{dist}\left(X_{\gamma t}^{(\gamma)}, W_t^{(\gamma)}\right)\right] \leq \frac{C}{\gamma}\left(1 + \mathbb{E}(|Y_0|) + \sqrt{\log(1 + \gamma^2 t)}\right)e^{Ct}.$$

Second, from (3.2),

$$\mathbb{E}\left[\left|Y_{\gamma t}^{(\gamma)} - A_t^{(\gamma)}\right|\right] \leq \mathbb{E}(|Y_0|)e^{-\gamma^2 t} + \frac{\|F\|_\infty}{\gamma}.$$

Third, it follows from [14, (ii) Lemma 3.2] that there exists a constant $C > 0$ such that for all $\gamma \geqslant 1$,

$$\mathbb{E}\left(\mathrm{dist}\left(W_t^{(\gamma)}, \overline{X}_t^{(\gamma)}\right)\right) \leq \frac{C}{\gamma}e^{Ct}.$$

Finally, since $\mathcal{L}(A_t^{(\gamma)})$ is a centered Gaussian distribution with variance $\Sigma^2(1 - e^{-2\gamma^2 t})$, the formula for the L^2 Wasserstein distance (which is larger than the L^1 distance $\mathcal{W}$) between Gaussian distributions (see [7]) yields

$$\mathcal{W}\left(\mathcal{L}(A_t^{(\gamma)}), \mu_\Sigma\right) \leq \sqrt{\mathrm{Tr}(\Sigma^2)}\left(1 - \sqrt{1 - e^{-2\gamma^2 t}}\right).$$

Summing the four last inequalities concludes the proof. $\qquad\qquad\square$

References

[1] Federico Bonetto, Joel L. Lebowitz, and Luc Rey-Bellet, *Fourier's law: a challenge to theorists*, pp. 128–150. MR1773043

[2] Sandra Cerrai and Nathan Glatt-Holtz, *On the convergence of stationary solutions in the Smoluchowski-Kramers approximation of infinite dimensional systems*, Journal of Functional Analysis **278** (2020), no. 8, 108421. MR4056993

[3] Zhiwei Chen and Mark Freidlin, *Smoluchowski-Kramers approximation and exit problems*, Stochastics and Dynamics **05** (2005), no. 04, 569–585. MR2185506

[4] Andreas Eberle, *Reflection couplings and contraction rates for diffusions*, Probab. Theory Relat. Fields **166** (2016), 851–886. MR3568041

[5] Denis J. Evans and Gary P. Morriss, *Statistical mechanics of nonequilibrium liquids*, Academic Press, 1990.

[6] Giambattista Giacomin, Stefano Olla, Ellen Saada, Herbert Spohn, and Gabriel Stoltz, *Stochastic dynamics out of equilibrium*, Springer Proceedings in Mathematics & Statistics, vol. 282, Springer Nature Switzerland AG 2019. MR3986060

[7] Clark R. Givens and Rae Michael Shortt, *A class of Wasserstein metrics for probability distributions.*, Michigan Mathematical Journal **31** (1984), no. 2, 231 – 240. MR0752258

[8] Martin Hairer and Grigorios Pavliotis, *From ballistic to diffusive behavior in periodic potentials*, J Stat Phys **131** (2008), 175–202. MR2394704

[9] Alessandra Iacobucci, Stefano Olla, and Gabriel Stoltz, *Convergence rates for nonequilibrium Langevin dynamics*, Ann. Math. Québec **43** (2019), 73–98. MR3925138

[10] Ioannis Karatzas and Steven E. Shreve, Brownian Motion and Stochastic Calculus, Springer, 1998, pp. 47–127. MR1121940

[11] Benedict Leimkuhler, Charles Matthews, and Gabriel Stoltz, *The computation of averages from equilibrium and nonequilibrium Langevin molecular dynamics*, IMA Journal of Numerical Analysis **36** (2015), no. 1, 13–79. MR3463433

[12] Tony Lelièvre and Gabriel Stoltz, *Partial differential equations and stochastic methods in molecular dynamics*, Acta Numerica **25** (2016), 681–880. MR3509213

[13] Viviana Letizia and Stefano Olla, *Nonequilibrium isothermal transformations in a temperature gradient from a microscopic dynamics*, The Annals of Probability **45** (2017), no. 6A, 3987 – 4018. MR3729621

[14] Mouad Ramil, *Quasi-stationary distribution for the Langevin process in cylindrical domains, part II: overdamped limit*, arXiv e-prints (2021), arXiv:2103.00338. MR4347490

[15] Julien Roussel and Gabriel Stoltz, *A perturbative approach to control variates in molecular dynamics*, Multiscale Modeling & Simulation **17** (2019), no. 1, 552–591. MR3932612