(Non)equilibrium Langevin dynamics: convergence and numerical approximation

Gabriel STOLTZ

gabriel.stoltz@enpc.fr

(CERMICS, Ecole des Ponts & MATHERIALS team, INRIA Paris)

Work also supported by ANR Funding ANR-14-CE23-0012 ("COSMOS")

UMass, February 2017

Outline

- **Equilibrium Langevin dynamics**
  - Various convergence results
  - A focus on the approach by Dolbeault, Mouhot and Schmeiser

- **Discretization by a spectral Galerkin method**
  - A priori error estimates for Poisson equations
  - Explicit convergence rates for a representative system
  - Numerical results

- **Nonequilibrium Langevin dynamics**
  - Non-gradient perturbation (constant force)
  - Convergence results
  - Numerical estimation of the spectral gap

Equilibrium Langevin dynamics

Langevin dynamics

- Positions $q \in \mathcal{D} = (L\mathbb{T})^d$ or $\mathbb{R}^d$ and momenta $p \in \mathbb{R}^d$

$\rightarrow$ phase-space $\mathcal{E} = \mathcal{D} \times \mathbb{R}^d$

- Hamiltonian $H(q,p) = V(q) + \frac{1}{2} p^T M^{-1} p$ (more general kinetic energies $U(p)$ can be considered[1])

Stochastic perturbation of the Hamiltonian dynamics

$$\begin{cases} dq_t = M^{-1} p_t \, dt \\ dp_t = -\nabla V(q_t)\, dt - \gamma M^{-1} p_t \, dt + \sqrt{\dfrac{2\gamma}{\beta}}\, dW_t \end{cases}$$

- Friction $\gamma > 0$ (could be a position-dependent matrix)
- Existence and uniqueness of the invariant measure (canonical measure)

$$\mu(dq\,dp) = Z^{-1} \mathrm{e}^{-\beta H(q,p)}\, dq\, dp = \nu(dq)\,\kappa(dp)$$

[1] Redon, Stoltz, Trstanova, *J. Stat. Phys.* (2016)

Hamiltonian and overdamped limits

- As $\gamma \to 0$, the Hamiltonian dynamics is recovered
- Overdamped limit $\gamma \to +\infty$ or $m \to 0$

$$q_t - q_0 = -\frac{1}{\gamma} \int_0^t \nabla V(q_s)\, ds + \sqrt{\frac{2}{\gamma\beta}} W_t - \frac{1}{\gamma}(p_t - p_0)$$

which converges to the solution of $dQ_t = -\nabla V(Q_t)\, dt + \sqrt{\frac{2}{\beta}}\, dW_t$

- In both cases, slow convergence to equilibrium
  - it takes time to change energy levels in the Hamiltonian limit[2]
  - for m fixed, time has to be rescaled by a factor γ

[2] Hairer and Pavliotis, *J. Stat. Phys.*, **131**(1), 175-202 (2008)

Ergodicity results (1)

- Almost-sure convergence[3] of ergodic averages $\widehat{\varphi}_t = \dfrac{1}{t}\int_0^t \varphi(q_s,p_s)\,ds$
- Asymptotic variance of ergodic averages

$$\sigma_\varphi^2 = \lim_{t\to+\infty} t\mathbb{E}\left[\widehat{\varphi}_t^2\right] = 2\int_{\mathcal{E}} \left(-\mathcal{L}^{-1}\Pi_0\varphi\right)\Pi_0\varphi\,d\mu$$

where $\Pi_0\varphi = \varphi - \mathbb{E}_\mu(\varphi)$ and $\mathcal{L} = \mathcal{L}_{\rm ham} + \gamma\mathcal{L}_{\rm FD}$ with

$$\mathcal{L}_{\rm ham} = p^T M^{-1}\nabla_q - \nabla V^T\nabla_p, \qquad \mathcal{L}_{\rm FD} = -p^T M^{-1}\nabla_p + \frac{1}{\beta}\Delta_p$$

- A central limit theorem holds[4] when the equation has a solution in $L^2(\mu)$

Poisson equation in $L^2(\mu)$

$$-\mathcal{L}\Phi = \Pi_0\varphi$$

[3] Kliemann, *Ann. Probab.* **15**(2), 690-707 (1987)

[4] Bhattacharya, *Z. Wahrsch. Verw. Gebiete* **60**, 185–201 (1982)

Ergodicity results (2)

- Invertibility of $\mathcal{L}$ on subsets of $L_0^2(\mu) = \left\{ \varphi \in L^2(\mu) \,\middle|\, \int_{\mathcal{E}} \varphi \, d\mu = 0 \right\}$?

$$-\mathcal{L}^{-1} = \int_0^{+\infty} \mathrm{e}^{t\mathcal{L}} \, dt$$

- Prove exponential convergence of the semigroup $\mathrm{e}^{t\mathcal{L}}$
  - various Banach spaces $E \cap L_0^2(\mu)$
  - Lyapunov techniques[5,6] $L_W^\infty(\mathcal{E}) = \left\{ \varphi \, \text{measurable}, \, \left\| \frac{\varphi}{W} \right\|_{L^\infty} < +\infty \right\}$
  - standard hypocoercive[7] setup $H^1(\mu)$
  - $E = L^2(\mu)$ after hypoelliptic regularization[8] from $H^1(\mu)$

[5] L. Rey-Bellet, *Lecture Notes in Mathematics* (2006)
[6] Hairer and Mattingly, *Progr. Probab.* **63** (2011)
[7] Villani (2009) and before Talay (2002), Eckmann/Hairer (2003), Hérau/Nier (2004)
[8] F. Hérau, *J. Funct. Anal.* **244**(1), 95-118 (2007)

Direct $L^2(\mu)$ approach

- Assume that the potential V is smooth and[9,10]
  - the marginal measure ν satisfies a Poincaré inequality
$$\|\Pi_0\varphi\|^2_{L^2(\nu)} \leqslant \frac{1}{C_\nu}\|\nabla_q\varphi\|^2_{L^2(\nu)}.$$
  - there exist $c_1 > 0$, $c_2 \in [0,1)$ and $c_3 > 0$ such that V satisfies
$$\Delta V \leqslant c_1 + \frac{c_2}{2}|\nabla V|^2, \quad |\nabla^2 V| \leqslant c_3\left(1 + |\nabla V|\right).$$

There exist $C > 0$ and $\lambda_\gamma > 0$ such that, for any $\varphi \in L^2_0(\mu)$,
$$\forall t \geqslant 0, \qquad \left\|\mathrm{e}^{t\mathcal{L}}\varphi\right\|_{L^2(\mu)} \leqslant C\mathrm{e}^{-\lambda_\gamma t}\|\varphi\|_{L^2(\mu)}.$$
with convergence rate of order $\min(\gamma,\gamma^{-1})$: there exists $\overline{\lambda} > 0$ such that
$$\lambda_\gamma \geqslant \overline{\lambda}\min(\gamma,\gamma^{-1}).$$

[9] Dolbeault, Mouhot and Schmeiser, *C. R. Math. Acad. Sci. Paris* (2009)

[10] Dolbeault, Mouhot and Schmeiser, *Trans. AMS*, **367**, 3807–3828 (2015)

Sketch of proof

- Entropy functional $\mathcal{H}[\varphi] = \dfrac{1}{2}\|\varphi\|^2 - \varepsilon \langle A\varphi, \varphi \rangle$ for $\varepsilon \in (-1, 1)$ and

$$A = \Big(1 + (\mathcal{L}_{\rm ham}\Pi_p)^*(\mathcal{L}_{\rm ham}\Pi_p)\Big)^{-1}(\mathcal{L}_{\rm ham}\Pi_p)^*, \qquad \Pi_p\varphi = \int_{\mathbb{R}^D} \varphi \, d\kappa$$

- $A = \Pi_p A(1 - \Pi_p)$ and $\mathcal{L}_{\rm ham}A$ are bounded so that $\mathcal{H} \sim \| \cdot \|^2_{L^2(\mu)}$

Coercivity in the scalar product $\langle\langle \cdot, \cdot \rangle\rangle$ induced by $\mathcal{H}$

$$\mathscr{D}[\varphi] := \langle\langle -\mathcal{L}\varphi, \varphi \rangle\rangle \geqslant \widetilde{\lambda}_\gamma \|\varphi\|^2,$$

- Idea: control of $\|(1 - \Pi_p)\varphi\|^2$ by $\langle -\mathcal{L}_{\rm FD}\varphi, \varphi \rangle$ (Poincaré); for $\|\Pi_p\varphi\|^2$,

$$\|\mathcal{L}_{\rm ham}\Pi_p\varphi\|^2 \geqslant \frac{DC_\nu}{\beta m}\|\Pi_p\varphi\|^2, \qquad \text{hence } A\mathcal{L}_{\rm ham}\Pi_p \geqslant \lambda_{\rm ham}\Pi_p$$

- Gronwall inequality $\dfrac{d}{dt}\left(\mathcal{H}\left[\mathrm{e}^{t\mathcal{L}}\varphi\right]\right) = -\mathscr{D}\left[\mathrm{e}^{t\mathcal{L}}\varphi\right] \leqslant -\dfrac{2\widetilde{\lambda}_\gamma}{1+\varepsilon}\mathcal{H}\left[\mathrm{e}^{t\mathcal{L}}\varphi\right]$

Discretization by a spectral Galerkin method

Principle of Galerkin approximation

- Reference Poisson equation $-\mathcal{L}\Phi = \Pi_0 R$
- Galerkin space V_M: conformal ($V_M \subset L^2_0(\mu)$) or non-conformal

Variational formulation

$$\left\{ \begin{array}{l} \text{Find } \Phi_M \in V_M \cap L^2_0(\mu) \text{ such that} \\ \forall \psi \in V_M, \ -\langle \psi, \mathcal{L}\Phi_M \rangle = \langle \psi, \Pi_0 R \rangle \, . \end{array} \right.$$

- In fact, $-\Pi_M \mathcal{L} \Pi_M \Phi_M = \Pi_M R$
- Error = (related to) consistency error + approximation error

$$\Phi_M - \Phi = (\Phi_M - \Pi_M \Phi) - (1 - \Pi_M)\Phi$$

- Well-posedness? Cannot apply Lax–Milgram...
- Approximation error: for specific models

Well posedness of the Galerkin procedure (conformal case)

- Conformal space and $\|(A + A^*)(1 - \Pi_M)\mathcal{L}\Pi_M\| \xrightarrow[M\to\infty]{} 0$

Invertibility of $-\Pi_M\mathcal{L}\Pi_M$

There exist $C \geqslant 1$ (independent of M, γ) and $M_0 \in \mathbb{N}$ such that, for any $M \geqslant M_0$, there is $\lambda_{\gamma,M} > 0$ for which

$$\forall \varphi \in V_M, \quad \forall t \geqslant 0, \qquad \left\| \mathrm{e}^{t\Pi_M\mathcal{L}\Pi_M}\varphi \right\| \leqslant C\mathrm{e}^{-\lambda_{\gamma,M}t}\|\varphi\|.$$

Moreover, $\lambda_{\gamma,M} \xrightarrow[M\to\infty]{} \lambda_\gamma$ where $\lambda_\gamma > 0$.

- If $\mathcal{L}_{\mathrm{FD}}$ stabilizes V_M (i.e. $\Pi_M\mathcal{L}_{\mathrm{FD}} = \mathcal{L}_{\mathrm{FD}}\Pi_M$), uniform lower bound

$$\forall \gamma > 0, \qquad \lambda_{\gamma,M} \geqslant \overline{\lambda}_M \min(\gamma, \gamma^{-1}),$$

- Key inequality for the proof: $\mathscr{D}_M[\varphi] = -\langle\langle \varphi, \Pi_M\mathcal{L}\Pi_M\varphi \rangle\rangle$ such that

$$\mathscr{D}_M[\varphi] = \mathscr{D}[\varphi] + \varepsilon \langle A(1 - \Pi_M)\mathcal{L}\varphi, \varphi \rangle + \varepsilon \langle \varphi, A^*(1 - \Pi_M)\mathcal{L}\varphi \rangle$$

Well posedness in the non-conformal case

- Work on $V_{M,0} = V_M \cap L_0^2(\mu)$
- Projector $\Pi_M = \Pi_{M,0} + \Pi_{u_M}$ with $u_M = \dfrac{\Pi_M \mathbf{1}}{\|\Pi_M \mathbf{1}\|} \in V_M$
- Invertibility of $-\Pi_{M,0}\mathcal{L}\Pi_{M,0}$ under the additional condition

$$\|\mathcal{L}^* u_M\| \xrightarrow[M\to\infty]{} 0$$

Saddle-point formulation

For any $R \in L^2(\mu)$, there exist a unique $\Phi_M \in V_M$ and a unique $\alpha_M \in \mathbb{R}$ such that

$$\left\{\begin{aligned} -\Pi_M\mathcal{L}\Pi_M\Phi_M + \alpha_M u_M &= \Pi_M R, \\ \langle \Phi_M, u_M \rangle &= 0. \end{aligned}\right.$$

Consistency error

- Poisson equation implies $-\Pi_{M,0}\mathcal{L}\Pi_{M,0}\Phi = \Pi_{M,0}R + \Pi_{M,0}\mathcal{L}(1-\Pi_{M,0})\Phi$

- Subtracting the equation for Φ_M,

$$\Pi_{M,0}\mathcal{L}\Pi_{M,0}(\Phi_M - \Pi_{M,0}\Phi) = \Pi_{M,0}\mathcal{L}(1-\Pi_{M,0})\Phi$$

Consistency error

$$\|\Phi_M - \Pi_{M,0}\Phi\|_{L^2(\mu)} \leqslant \frac{C}{\widehat{\lambda}_{\gamma,M}}\left(\|\Pi_M\mathcal{L}(1-\Pi_M)\Phi\|_{L^2(\mu)} + \|\mathcal{L}u_M\|\|\Phi\|\right)$$

- Explicit estimates/rates for specific models... Typically,

$$\|\Pi_M\mathcal{L}(1-\Pi_M)\Phi\|_{L^2(\mu)} \leqslant \|\Pi_M\mathcal{L}(1-\Pi_M)\|_{\mathcal{B}(L^2(\mu))}\|(1-\Pi_M)\Phi\|_{L^2(\mu)}$$

provided the second term is not too large and the last one is small

Spectral basis for Langevin operators

- Weighted Fourier modes in position

$$G_{2k}(q) = \sqrt{\frac{Z_{\beta,\nu}}{\pi}} \cos(kq)\, \mathrm{e}^{\beta V(q)/2}, \qquad G_{2k-1}(q) = \sqrt{\frac{Z_{\beta,\nu}}{\pi}} \sin(kq)\, \mathrm{e}^{\beta V(q)/2}$$

- Hermite functions H_ℓ for momenta (eigenfunctions of $\mathcal{L}_{\mathrm{FD}}$)

- Tensor basis of $2K-1$ Fourier and L Hermite modes: projector Π_{KL}

Approximation error

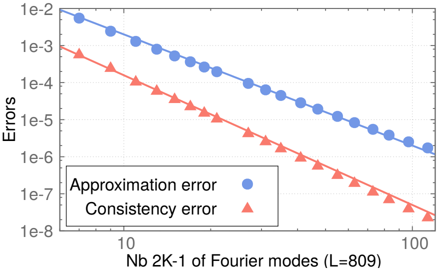
For any $s \in \mathbb{N}$, there exists $A_s \in \mathbb{R}_+$ such that, for all $\varphi \in H^s(\mu)$,

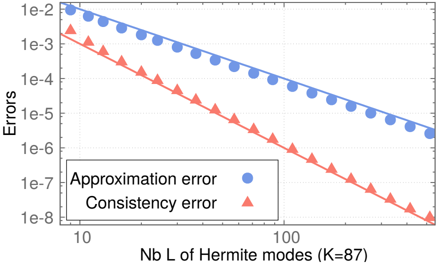
$$\forall K \geqslant 1,\ L \geqslant s, \qquad \|\varphi - \Pi_{KL}\varphi\|_{L^2(\mu)} \leqslant A_s \left(\frac{1}{K^s} + \frac{1}{L^{s/2}} \right) \|\varphi\|_{H^s(\mu)}$$

- Consistency error: bounds on $\|\Pi_{KL}\mathcal{L}(1-\Pi_{KL})\|_{\mathcal{B}(L^2(\mu))} = \mathrm{O}(K\sqrt{L})$

Numerical results

Observable $R = \sum_{k\in\mathbb{N},\ell\in\mathbb{N}} \max(1,k)^{-5/2}\max(1,\ell)^{-3/2} G_k H_\ell$ (almost $H^2(\mu)$)

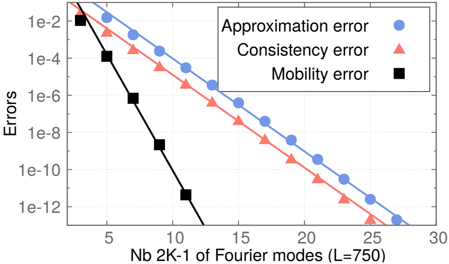


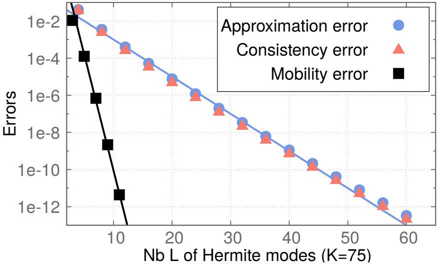


Left: approximation error $\sim K^{-3}$, consistency error $\sim K^{-7/2}$
Right: approximation error $\sim L^{-2}$, consistency error $\sim L^{-3}$
$\rightarrow$ "full" regularization by $\mathcal{L}^{-1}$!

Estimation of the self-diffusion (1)

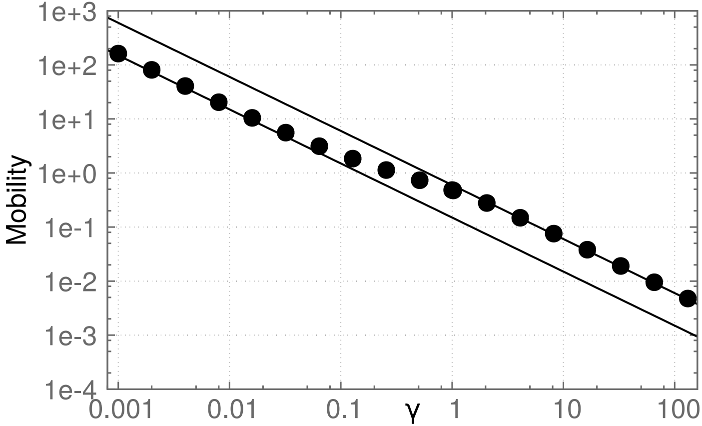
Definition $D(\gamma) = \int_0^{+\infty} \mathbb{E}(p_t p_0)\, dt = \left\langle -\mathcal{L}^{-1} p, p \right\rangle_{L^2(\mu)}$





Exponential rate of convergence

Estimation of the self-diffusion (2)

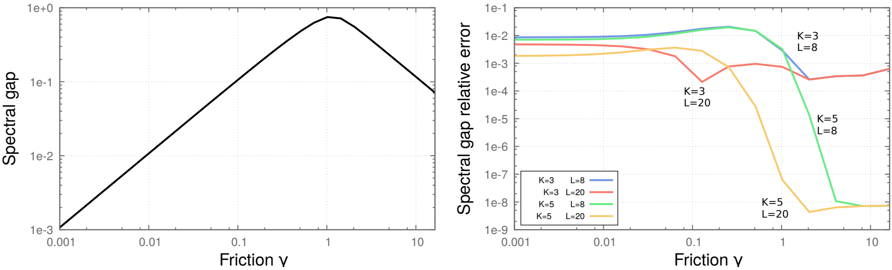


Scaling $D(\gamma) \sim \gamma^{-1}$ both for small and large frictions[11]

[11] M. Hairer and G. Pavliotis, *J. Stat. Phys.* (2008)

Errors on the spectral gap

Also exponential decay of the error on the spectral gap



Left: Spectral gap as a function of the friction γ

Right: Relative error on the spectral gap for several couples K, L. Note that it depends only on K in the overdamped regime

Nonequilibrium Langevin dynamics

Nonequilibrium perturbation of equilibrium dynamics

- Compact position space $\mathcal{D} = (2\pi\mathbb{T})^d$, constant force $|F| = 1$

Langevin dynamics perturbed by a constant force term

$$\begin{cases} dq_t = \dfrac{p_t}{m} dt, \\ dp_t = (-\nabla V(q_t) + \tau F) dt - \gamma \dfrac{p_t}{m} dt + \sqrt{\dfrac{2\gamma}{\beta}} dW_t, \end{cases}$$

- Non-zero velocity in the direction F is expected in the steady-state

- F **does not derive from the gradient of a periodic function**
  - of course, $F = -\nabla W_F(q)$ with $W_F(q) = -F^T q$
  - ...but W_F is not periodic!

Existence and uniqueness of the steady state

- Generator $\mathcal{L}_{\gamma,\tau} = \mathcal{L}_{\mathrm{ham}} + \gamma \mathcal{L}_{\mathrm{FD}} + \tau \mathcal{L}_{\mathrm{pert}}$ with $\mathcal{L}_{\mathrm{pert}} = F \cdot \nabla_p$

- Lyapunov functions[12] $\mathcal{K}_n(q,p) = 1 + |p|^n$ for $n \geqslant 2$

Exponential convergence to equilibrium

Consider $\tau_* > 0$ and fix $\gamma > 0$. For any $\tau \in [-\tau_*, \tau_*]$, there is a unique invariant probability measure which admits a C^∞ density $\psi_\tau(q,p)$. Moreover, for any $n \geqslant 2$, there exist $C_n, \lambda_n > 0$ (depending on τ_*) such that, for any $\tau \in [-\tau_*, \tau_*]$ and for any $\varphi \in L^\infty_{\mathcal{K}_n}(\mathcal{E})$,

$$\forall t \geqslant 0, \qquad \left\| \mathrm{e}^{t\mathcal{L}_\tau} \varphi - \int_{\mathcal{E}} \varphi \, \psi_\tau \right\|_{L^\infty_{\mathcal{K}_n}} \leqslant C_n \mathrm{e}^{-\lambda_n t} \|\varphi\|_{L^\infty_{\mathcal{K}_n}} .$$

- Non-perturbative result (also non-quantitative unfortunately...)

[12] M. Hairer and J. Mattingly, *Progr. Probab.* (2011); Meyn and Tweedie (2009); Rey–Bellet (2006)

Perturbative expansion of the steady state

- Perturbative framework: operators considered on $L^2(\mu)$

$$\psi_\tau = h_\tau \mu, \qquad h_\tau \in L^2(\mu)$$

- Fokker–Planck equation $\mathcal{L}_{\gamma,\tau}^* h_\tau = 0$ with $\displaystyle\int_{\mathcal{E}} h_\tau \, d\mu = 1$

Power expansion of the invariant measure

For $|\tau| < r^{-1}$, it holds $h_\tau \in L^2(\mu)$ and

$$h_\tau = \left(1 + \tau \left(\mathcal{L}_{\mathrm{pert}} \mathcal{L}_{\gamma,0}^{-1}\right)^*\right)^{-1} \mathbf{1} = \left(1 + \sum_{n=1}^{+\infty} (-\tau)^n \left[\left(\mathcal{L}_{\mathrm{pert}} \mathcal{L}_{\gamma,0}^{-1}\right)^*\right]^n\right) \mathbf{1}.$$

- Spectral radius $r = \lim_{n \to +\infty} \left\| \left[\left(\mathcal{L}_{\mathrm{pert}} \mathcal{L}_{\gamma,0}^{-1}\right)^*\right]^n \right\|_{\mathcal{B}(L_0^2(\mu))}^{1/n}$. In fact,

$$\frac{1}{r} \geqslant \frac{\min(1,\gamma)}{\sqrt{\beta K}}$$

Provides magnitude of admissible perturbations for $L^2(\mu)$ convergence

Convergence rates

- Suboptimal results by the standard hypocoercive approach in $H^1(\mu)$
$\rightarrow$ nonequilibrium perturbation[13] of direct $L^2(\mu)$ strategy

Uniform rates for nonequilibrium perturbations

There exist $C, \delta_* > 0$ such that, for any $\delta \in [0, \delta^*]$, there is $\overline{\lambda}_\delta > 0$ for which, for all $\gamma \in (0, +\infty)$ and all $\tau \in [-\delta \min(\gamma, 1), \delta \min(\gamma, 1)]$,

$$\left\| \mathrm{e}^{t\mathcal{L}^*_{\gamma,\tau}} f - h_\tau \right\|_{L^2(\mu)} \leqslant C \mathrm{e}^{-\overline{\lambda}_\delta \min(\gamma,\gamma^{-1})t} \| f - h_\tau \|_{L^2(\mu)}.$$

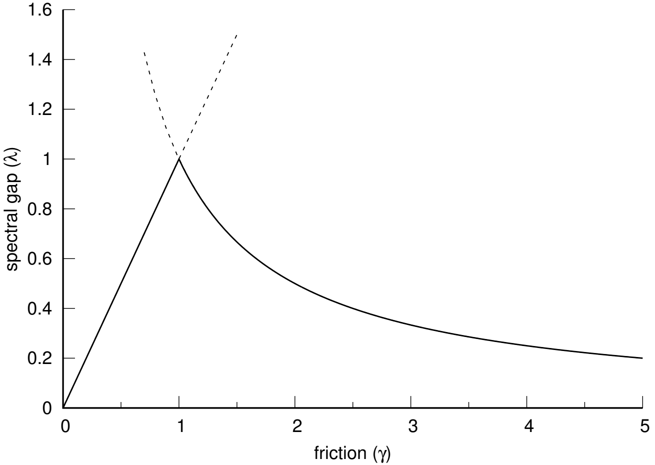
Moreover, $\overline{\lambda}_\delta = \overline{\lambda}_0 + \mathrm{O}(\delta)$.

- As a corollary: lower bounds on the spectral gap of order $\min(\gamma, \gamma^{-1})$

- Some elements on hypocoercive entropy estimates

[13] E. Bouin, F. Hoffmann, and C. Mouhot, *arXiv preprint* **1605.04121**

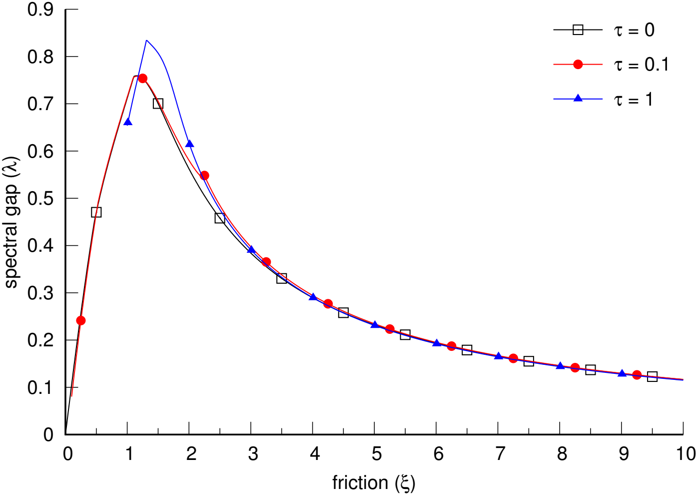
Numerical results (1)



Predicted spectral gap as a function of the friction γ when $V = 0$, $\beta = 1$ and $m = 1$ (solid line) vs. theoretical prediction[14]

[14] S. M. Kozlov, *Math. Notes* **45**, 360-368 (1989)

Numerical results (2)



Spectral gap as a function of γ for $\tau = 0, 0.1, 1$ when $V(q) = 1 - \cos(q)$

References

References

- J. Roussel and G. Stoltz, Spectral methods for the Langevin dynamics and associated error estimates, *arXiv preprint* **1702.????**

- A. Iacobucci, S. Olla and G. Stoltz, Convergence rates for nonequilibrium Langevin dynamics, *arXiv preprint* **1702.03685**

- T. Lelièvre and G. Stoltz, Partial differential equations and stochastic methods in molecular dynamics, *Acta Numerica* **25**, 681–880 (2016)