

Inria



European Research Council
Established by the European Commission

Convergence of Adaptive Langevin (and other dynamics) using hypocoercivity

Gabriel STOLTZ

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

In collaboration with A. Iacobucci, B. Leimkuhler, S. Olla, S. Redon, J. Roussel, M. Sachs, Z. Trstanova, E. Vanden-Eijnden; and currently A. Levitt, I. Sekkat

- **Adaptive Langevin dynamics**
 - Motivation: Bayesian inference for large data sets
- **Convergence of Langevin type dynamics**
 - A review of possible approaches
 - A focus on a “direct” hypocoercive technique
- **Convergence of Adaptive Langevin dynamics**
 - Heuristic demonstration of convergence rates
 - Sharp bounds through hypocoercive techniques
 - Central Limit Theorem and “Langevin” limit

B. Leimkuhler, M. Sachs and G. Stoltz, Hypocoercivity properties of adaptive Langevin dynamics, *arXiv preprint* **1908.09363**

Adaptive Langevin dynamics

Bayesian inference in the large data context

- **Data** $\{x_i\}_{i=1,\dots,N}$ to be explained by a statistical model

- Parametrization by $q \in \mathbb{R}^n$: individual likelihoods $P(x_i|q)$
- Prior $\rho(q)$ on the parameters
- Sample q from $\nu(dq) = e^{-V(q)} dq = Z_\nu^{-1} \rho(q) \prod_{i=1}^N P(x_i|q) dq$
- For usual MCMC methods, **each step costs $O(N)$**

- **Mini-batching**: Stochastic gradient Langevin dynamics¹

- Fundamental assumption: for $1 \ll \mathcal{N} \ll N$ and $J_{\mathcal{N}} \in \{1, \dots, N\}^{\mathcal{N}}$,

$$\nabla(\ln \rho)(q) + \frac{N}{\mathcal{N}} \sum_{j \in J_{\mathcal{N}}} \nabla(\ln P(x_j|q)) = -\nabla V(q) + \mathcal{G}, \quad \mathcal{G} \sim \mathcal{N}(0, \Sigma(q))$$

- Amounts to introducing an **additional Brownian motion of unknown magnitude** \rightarrow **bias**
- Assume that $\Sigma(q)$ is constant [Work of Inass Sekkat...]

¹Welling/Teh, *ICML* (2011)

Removing the mini-batching bias

- Phase-space extension: momenta p and **variable friction** ζ

Adaptive Langevin dynamics¹: **unknown** σ (scalar, for simplicity)

$$\begin{aligned}dq_t &= M^{-1} p_t dt, \\dp_t &= (-\nabla V(q_t) - \zeta_t M^{-1} p_t) dt + \sigma dW_t, \\d\zeta_t &= \frac{1}{m} \left(p_t^T M^{-2} p_t - \beta^{-1} \text{Tr} (M^{-1}) \right) dt\end{aligned}$$

- Invariant measure π with density proportional to

$$\exp \left(-\beta \left[\frac{p^T M^{-1} p}{2} + V(q) + \frac{m}{2} \left(\zeta - \frac{\beta \sigma^2}{2} \right)^2 \right] \right) dq dp d\zeta$$

- **The marginal of π in q is indeed ν** whatever σ ... Prove convergence, in particular **Central Limit Theorem** for time averages?

¹A. Jones and B. Leimkuhler, *J. Chem. Phys.* (2011); Ding et al., *NIPS* (2014); B. Leimkuhler and X. Shang, *SIAM J. Sci. Comput.* (2015)

Standard Langevin dynamics

Langevin dynamics (1)

- Positions $q \in \mathcal{D} = (\mathbb{L}\mathbb{T})^d$ or \mathbb{R}^d and momenta $p \in \mathbb{R}^d$
→ phase-space $\mathcal{E} = \mathcal{D} \times \mathbb{R}^d$
- **Hamiltonian** $H(q, p) = V(q) + \frac{1}{2}p^T M^{-1}p$

Stochastic perturbation of the Hamiltonian dynamics

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

- Given (known) **friction** $\gamma > 0$ (could be a position-dependent matrix)

Langevin dynamics (2)

- Evolution semigroup $(e^{t\mathcal{L}}\varphi)(q, p) = \mathbb{E} \left[\varphi(q_t, p_t) \mid (q_0, p_0) = (q, p) \right]$
- Generator of the dynamics \mathcal{L}

$$\frac{d}{dt} \left(\mathbb{E} \left[\varphi(q_t, p_t) \mid (q_0, p_0) = (q, p) \right] \right) = \mathbb{E} \left[(\mathcal{L}\varphi)(q_t, p_t) \mid (q_0, p_0) = (q, p) \right]$$

Generator of the Langevin dynamics $\mathcal{L} = \mathcal{L}_{\text{ham}} + \gamma\mathcal{L}_{\text{FD}}$

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

- Existence and uniqueness of the invariant measure characterized by

$$\forall \varphi \in C_0^\infty(\mathcal{E}), \quad \int_{\mathcal{E}} \mathcal{L}\varphi d\mu = 0$$

- Here, **canonical measure**

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

Fokker–Planck equations

- Evolution of the law $\psi(t, q, p)$ of the process at time $t \geq 0$

$$\frac{d}{dt} \left(\int_{\mathcal{E}} \varphi \psi(t) \right) = \int_{\mathcal{E}} (\mathcal{L}\varphi) \psi(t)$$

- Fokker–Planck equation (with \mathcal{L}^\dagger adjoint of \mathcal{L} on $L^2(\mathcal{E})$)

$$\partial_t \psi = \mathcal{L}^\dagger \psi$$

- It is convenient to **work in** $L^2(\mu)$ with $f(t) = \psi(t)/\mu$
 - denote the adjoint of \mathcal{L} on $L^2(\mu)$ by \mathcal{L}^*

$$\mathcal{L}^* = -\mathcal{L}_{\text{ham}} + \gamma \mathcal{L}_{\text{FD}}$$

- Fokker–Planck equation $\partial_t f = \mathcal{L}^* f$
- Convergence results for $e^{t\mathcal{L}}$ on $L^2(\mu)$ are very similar to the ones for $e^{t\mathcal{L}^*}$

Hamiltonian and overdamped limits

- As $\gamma \rightarrow 0$, the **Hamiltonian** dynamics is recovered
- **Overdamped** limit $\gamma \rightarrow +\infty$ (or masses going to 0)

$$\begin{aligned}q_{\gamma t} - q_0 &= -\frac{1}{\gamma} \int_0^{\gamma t} \nabla V(q_s) ds + \sqrt{\frac{2}{\gamma\beta}} W_{\gamma t} - \frac{1}{\gamma} (p_{\gamma t} - p_0) \\ &= -\int_0^t \nabla V(q_{\gamma s}) ds + \sqrt{2\beta^{-1}} B_t - \frac{1}{\gamma} (p_{\gamma t} - p_0)\end{aligned}$$

which converges to the solution of $dQ_t = -\nabla V(Q_t) dt + \sqrt{2\beta^{-1}} dB_t$

- In both cases, **slow convergence to equilibrium**
 - it takes time to change energy levels in the Hamiltonian limit²
 - for fixed masses, time has to be rescaled by a factor γ

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

Ergodicity results (1)

- Almost-sure convergence³ of **ergodic averages** $\widehat{\varphi}_t = \frac{1}{t} \int_0^t \varphi(q_s, p_s) ds$
- **Asymptotic variance** of ergodic averages

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

where $\Pi_0 \varphi = \varphi - \mathbb{E}_\mu(\varphi)$

- A central limit theorem holds⁴ when the equation has a solution in $L^2(\mu)$

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

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

- Well-posedness of such equations?

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

⁴Bhattacharya, *Z. Wahrsch. Verw. Gebiete* **60**, 185-201 (1982)

Ergodicity results (2)

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

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

- Prove **exponential convergence** of the semigroup $e^{t\mathcal{L}}$
 - various Banach spaces $E \cap L_0^2(\mu)$
 - **Lyapunov** techniques⁵ $L_W^\infty(\mathcal{E}) = \left\{ \varphi \text{ measurable, } \left\| \frac{\varphi}{W} \right\|_{L^\infty} < +\infty \right\}$
 - standard **hypo-coercive**⁶ setup $H^1(\mu)$
 - $E = L^2(\mu)$ after hypoelliptic regularization⁷ from $H^1(\mu)$
 - Directly $E = L^2(\mu)$ (recently⁸ Poincaré using $\partial_t - \mathcal{L}_{\text{ham}}$)
 - **coupling** arguments⁹

⁵Wu ('01); Mattingly/Stuart/Higham ('02); Rey-Bellet ('06); Hairer/Mattingly ('11)

⁶Villani (2009) and before Talay (2002), Eckmann/Hairer (2003), Hérau/Nier (2004)

⁷F. Hérau, *J. Funct. Anal.* **244**(1), 95-118 (2007)

⁸Armstrong/Mourrat (2019)

⁹A. Eberle, A. Guillin and R. Zimmer, *Ann. Probab.* **47**(4), 1982-2010 (2019)

Direct $L^2(\mu)$ approach: lack of coercivity

- The generator, considered on $L^2(\mu)$, is the sum of...
 - a **degenerate** symmetric part $\mathcal{L}_{\text{FD}} = -p^T M^{-1} \nabla_p + \frac{1}{\beta} \Delta_p$
 - an **antisymmetric** part $\mathcal{L}_{\text{ham}} = p^T M^{-1} \nabla_q - \nabla V^T \nabla_p$
- Standard strategy for coercive generators: consider φ with average 0 with respect to μ and compute

$$\begin{aligned} \frac{d}{dt} \left(\|e^{t\mathcal{L}} \varphi\|_{L^2(\mu)}^2 \right) &= \langle e^{t\mathcal{L}} \varphi, \mathcal{L} e^{t\mathcal{L}} \varphi \rangle_{L^2(\mu)} = \langle e^{t\mathcal{L}} \varphi, \mathcal{L}_{\text{FD}} e^{t\mathcal{L}} \varphi \rangle_{L^2(\mu)} \\ &= -\frac{1}{\beta} \|\nabla_p e^{t\mathcal{L}} \varphi\|_{L^2(\mu)}^2 \leq 0, \end{aligned}$$

but no control of $\|\phi\|_{L^2(\mu)}$ by $\|\nabla_p \phi\|_{L^2(\mu)}$ for a Gronwall estimate...

- **Change of scalar product** in order to use the antisymmetric part

Almost direct $L^2(\mu)$ approach: convergence result

- Assume that the potential V is **smooth** and^{10,11}
 - the marginal measure ν satisfies a **Poincaré** inequality

$$\|\Pi_0 \varphi\|_{L^2(\nu)}^2 \leq \frac{1}{C_\nu} \|\nabla_q \varphi\|_{L^2(\nu)}^2$$

- there exist $c_1 > 0$, $c_2 \in [0, 1)$ and $c_3 > 0$ such that V satisfies

$$\Delta V \leq c_1 + \frac{c_2}{2} |\nabla V|^2, \quad |\nabla^2 V| \leq c_3 (1 + |\nabla V|)$$

There exist $C > 0$ and $\lambda_\gamma > 0$ such that, for any $\varphi \in L_0^2(\mu)$,

$$\forall t \geq 0, \quad \|e^{t\mathcal{L}} \varphi\|_{L^2(\mu)} \leq C e^{-\lambda_\gamma t} \|\varphi\|_{L^2(\mu)}.$$

with convergence rate of order $\min(\gamma, \gamma^{-1})$: there exists $\bar{\lambda} > 0$ such that

$$\lambda_\gamma \geq \bar{\lambda} \min(\gamma, \gamma^{-1}).$$

¹⁰Dolbeault, Mouhot and Schmeiser, *C. R. Math. Acad. Sci. Paris* (2009)

¹¹Dolbeault, Mouhot and Schmeiser, *Trans. AMS*, **367**, 3807–3828 (2015)

Sketch of proof

- Modified square norm $\mathcal{H}[\varphi] = \frac{1}{2}\|\varphi\|^2 - \varepsilon \langle A\varphi, \varphi \rangle$ for $\varepsilon \in (-1, 1)$ and
$$A = \left(1 + (\mathcal{L}_{\text{ham}} \Pi_p)^*(\mathcal{L}_{\text{ham}} \Pi_p)\right)^{-1} (\mathcal{L}_{\text{ham}} \Pi_p)^*, \quad \Pi_p \varphi = \int_{\mathbb{R}^D} \varphi d\kappa$$
- $A = \Pi_p A (1 - \Pi_p)$ and $\mathcal{L}_{\text{ham}} A$ are bounded so that $\mathcal{H} \sim \|\cdot\|_{L^2(\mu)}^2$

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

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

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

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

- Gronwall inequality $\frac{d}{dt} (\mathcal{H} [e^{t\mathcal{L}}\varphi]) = -\mathcal{D} [e^{t\mathcal{L}}\varphi] \leq -\frac{2\tilde{\lambda}_\gamma}{1+\varepsilon} \mathcal{H} [e^{t\mathcal{L}}\varphi]$

Direct L^2 hypocoercivity for modified Langevin (1)

- **General kinetic energy** function $U(p)$ in the Langevin dynamics
 - The generator \mathcal{L} **may not be hypoelliptic** (U constant on open sets)
 - Convergence using Lyapunov¹² or hypocoercive¹³ techniques
- **Nonequilibrium** Langevin dynamics
 - **Invariant measure not known...**
 - Perturbative results,¹⁴ with some uniformity on the range of perturbations¹⁵
 - Temperature accelerated molecular dynamics¹⁶

¹²S. Redon, G. Stoltz and Z. Trstanova, *J. Stat. Phys.* (2016)

¹³G. Stoltz and Z. Trstanova, *SIAM MMS* (2018)

¹⁴E. Bouin, F. Hoffmann, and C. Mouhot, *SIAM J. Math. Anal.* (2017)

¹⁵A. Iacobucci, S. Olla and G. Stoltz, *Ann. Math. Quebec* (2019)

¹⁶G. Stoltz and E. Vanden-Eijnden, *Nonlinearity* (2018)

Direct L^2 hypocoercivity for modified Langevin (2)

- **Spectral discretization** of generator of Langevin dynamics
 - Goal: approximate solutions of Poisson equation to build a control variate¹⁷
 - Bounds on convergence rates as a function of the basis size¹⁸
- **Current lines of work:**
 - An even more direct approach **avoiding the change of scalar product?** (with Antoine Levitt, Inria/CERMICS)
 - Quantitative bounds for **atom chains** (Very long term goal...)
 - **Non-perturbative** approach for nonequilibrium dynamics

¹⁷J. Roussel and G. Stoltz, *SIAM MMS* (2019)

¹⁸J. Roussel and G. Stoltz, *M2AN* (2018)

Convergence of Adaptive Langevin dynamics

Structure of Adaptive Langevin dynamics (1)

- **Normalization** of the dynamics, for the invariant measure to be independent of m (take $M = \text{Id}$ to simplify)

$$\begin{cases} dq_t = p_t dt, \\ dp_t = (-\nabla V(q_t) - \zeta_t p_t) dt + \sigma dW_t, \\ d\zeta_t = \frac{1}{m} \left(|p_t|^2 - \frac{n}{\beta} \right) dt \end{cases}$$

- Set $\varepsilon = \sqrt{m}$ and $\zeta = \gamma + \frac{\xi}{\varepsilon}$ with $\gamma = \beta\sigma^2/2$ (**unknown**)

Normalized Adaptive Langevin dynamics

$$\begin{cases} dq_t = p_t dt, \\ dp_t = \left(-\nabla V(q_t) - \frac{\xi_t}{\varepsilon} p_t - \gamma p_t \right) dt + \sqrt{\frac{2\gamma}{\beta}} dW_t, \\ d\xi_t = \frac{1}{\varepsilon} \left(|p_t|^2 - \frac{n}{\beta} \right) dt \end{cases}$$

Structure of Adaptive Langevin dynamics (2)

- Invariant measure π with density $Z^{-1} \exp\left(-\beta \left[\frac{|p|^2}{2} + V(q) + \frac{\xi^2}{2}\right]\right)$
- Generator $\mathcal{L}_{\text{AdL}} = \mathcal{L}_{\text{ham}} + \gamma \mathcal{L}_{\text{FD}} + \varepsilon^{-1} \mathcal{L}_{\text{NH}}$ with (adjoints on $L^2(\pi)$)

$$\mathcal{L}_{\text{ham}} = \frac{1}{\beta} (\nabla_p^* \nabla_q - \nabla_q^* \nabla_p) = \frac{1}{\beta} \sum_{i=1}^n \partial_{p_i}^* \partial_{q_i} - \partial_{q_i}^* \partial_{p_i},$$

$$\mathcal{L}_{\text{FD}} = -\frac{1}{\beta} \nabla_p^* \nabla_p = -\frac{1}{\beta} \sum_{i=1}^n \partial_{p_i}^* \partial_{p_i},$$

$$\mathcal{L}_{\text{NH}} = \left(|p|^2 - \frac{n}{\beta}\right) \partial_\xi - \xi p^T \nabla_p = \frac{1}{\beta^2} \left((\partial_\xi - \partial_\xi^*) \nabla_p^* \nabla_p + \Delta_p^* \partial_\xi - \Delta_p \partial_\xi^* \right)$$

- Antisymmetric parts \mathcal{L}_{ham} , \mathcal{L}_{NH} and symmetric one \mathcal{L}_{FD}
- Proof of exponential decay using Lyapunov techniques¹⁹

¹⁹D. Herzog, *Commun. Math. Sci.* (2018)

Expected scalings

- Generator \simeq **superposition** of $\mathcal{L}_{\text{ham}} + \gamma\mathcal{L}_{\text{FD}}$ and $\varepsilon^{-1}\mathcal{L}_{\text{NH}} + \gamma\mathcal{L}_{\text{FD}}$
 - Exponential rate of decay $\sim \min(\gamma, \gamma^{-1})$ for the **Langevin** part
 - **Nosé–Hoover**-like part rewritten as $\varepsilon^{-1}(\mathcal{L}_{\text{NH}} + \gamma\varepsilon\mathcal{L}_{\text{FD}})$
 - suggests rate of decay $\sim \varepsilon^{-1} \min(\gamma\varepsilon, (\gamma\varepsilon)^{-1})$

Exponential convergence of the semigroup

There exist $C, \bar{\lambda}$ such that, for any $\varepsilon, \gamma > 0$, there is $\lambda_{\varepsilon, \gamma} > 0$ for which

$$\forall t \geq 0, \forall \varphi \in L^2(\pi), \quad \left\| e^{t\mathcal{L}_{\text{AdL}}}\varphi - \int \varphi d\pi \right\|_{L^2(\pi)} \leq C e^{-\lambda_{\varepsilon, \gamma} t} \left\| \varphi - \int \varphi d\pi \right\|_{L^2(\pi)}$$

with the lower bound $\lambda_{\varepsilon, \gamma} \geq \bar{\lambda} \min\left(\gamma, \gamma\varepsilon^2, \frac{1}{\gamma}, \frac{1}{\gamma\varepsilon^2}\right)$. As a consequence,

$$\mathcal{L}_{\text{AdL}}^{-1} = - \int_0^\infty e^{t\mathcal{L}_{\text{AdL}}} dt, \quad \|\mathcal{L}_{\text{AdL}}^{-1}\|_{\mathcal{B}(L^2_0(\pi))} \leq \frac{C}{\bar{\lambda}} \max(\gamma, \gamma^{-1}, \gamma\varepsilon^2, \gamma^{-1}\varepsilon^{-2}).$$

Sharpness of the scaling and elements of proof

- Scaling of resolvent norm **sharp** for $\gamma, \gamma^{-1}, \gamma\varepsilon^2$: for instance,

$$\mathcal{L}_{\text{AdL}} \left(\gamma\varepsilon\xi + \frac{|p|^2}{2} - \frac{p^T \nabla V}{\gamma} \right) = -\frac{\xi|p|^2}{\varepsilon} + \frac{p^T \nabla V}{\gamma\varepsilon} - \frac{1}{\gamma} \left(p^T \nabla^2 V p - |\nabla V|^2 \right),$$

which shows that $\|\mathcal{L}_{\text{AdL}}^{-1}\|_{\mathcal{B}(L_0^2(\pi))} \geq c\gamma\varepsilon^2$ by choosing $\gamma \gg \varepsilon \gg 1$

- **Proof:** construction of **regularization operator with correct scaling**

- total antisymmetric part $\mathcal{A}_\varepsilon = \mathcal{L}_{\text{ham}} + \varepsilon^{-1}\mathcal{L}_{\text{NH}}$
- **distinguish** $\varepsilon \leq 1$ (Langevin limits convergence) or $\varepsilon \geq 1$

$$\begin{aligned} A_\varepsilon &:= -\min \left(1, \frac{1}{\varepsilon} \right) \left[\min \left(1, \frac{1}{\varepsilon^2} \right) - \Pi \mathcal{A}_\varepsilon^2 \Pi \right]^{-1} \Pi \mathcal{A}_\varepsilon \\ &= -\min \left(1, \frac{1}{\varepsilon} \right) \left[\min \left(1, \frac{1}{\varepsilon^2} \right) + \Pi \left(\frac{2n}{(\beta\varepsilon)^2} \partial_\xi^* \partial_\xi + \frac{1}{\beta} \nabla_q^* \nabla_q \right) \Pi \right]^{-1} \Pi \mathcal{A}_\varepsilon \end{aligned}$$

- technical estimates to control all terms uniformly in ε

Central Limit Theorem

- Consider $\varphi \in L^2(\pi)$ and $\bar{\varphi}_t := \frac{1}{t} \int_0^t \varphi(q_s, p_s, \xi_s) ds$

Central Limit Theorem

$$\sqrt{t}(\hat{\varphi}_t - \mathbb{E}_\pi \varphi) \xrightarrow[t \rightarrow +\infty]{\text{law}} \mathcal{N}(0, \sigma_{\varepsilon, \gamma}^2(\varphi)),$$

with the asymptotic variance (with $\Pi_0 \varphi = \varphi - \mathbb{E}_\pi(\varphi)$)

$$\sigma_{\varepsilon, \gamma}^2(\varphi) = 2 \int (-\mathcal{L}_{\text{AdL}}^{-1} \Pi_0 \varphi) \Pi_0 \varphi d\pi \leq \frac{2C \|\varphi\|_{L^2(\pi)}^2}{\bar{\lambda}} \max(\gamma, \gamma^{-1}, \gamma \varepsilon^2, \gamma^{-1} \varepsilon^{-2})$$

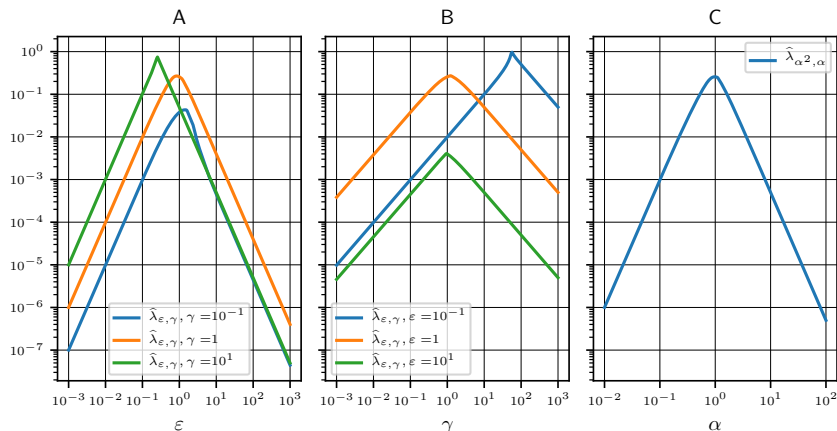
- Suggests taking** $\gamma = 1$ and $\varepsilon \sim 1$
- Langevin type limit** $\varepsilon \rightarrow +\infty$ for a function $\varphi(q, p)$ (independent of ξ)

$$|\sigma_{\varepsilon, \gamma}^2(\varphi) - \sigma_{\infty, \gamma}^2(\varphi)| \leq \frac{K}{\varepsilon}$$

Proof: [asymptotic analysis](#) and fine estimates²⁰ of $\mathcal{L}_{\text{Lang}} = \mathcal{L}_{\text{ham}} + \gamma \mathcal{L}_{\text{FD}}$

²⁰Talay, *Markov Proc. Rel. Fields* (2002); Kopec, *BIT* (2015)

Some numerical results



Spectral gap computed with a Galerkin method for V quadratic

A: Scaling $\min(\epsilon^2, \epsilon^{-2})$ for γ fixed.

B: Scaling $\min(\gamma, \gamma^{-1})$ for ϵ fixed.

C: Scaling $\min(\alpha^3, \alpha^{-3})$ for $\alpha = \gamma = \epsilon$.