Speeding up Convergence to Equilibrium for Diffusion Processes

G.A. Pavliotis
Department of Mathematics Imperial College London
Joint Work with T. Lelievre(CERMICS), F. Nier (CERMICS), M. Ottobre (Warwick), K. Pravda-Starov (Cergy)

Multiscale Inverse Problems
Warwick University
17/06/2013

- *Nonreversible optimization of the convergence to equilibrium for diffusions with linear drift*, T. Lelièvre, F. Nier, G.P. J Stat Phys. 2013.
- *Exponential Return to Equilibrium for Quadratic Hypoelliptic Systems* M. Ottobre, G.P., K. Pravda-Starov. J. Func. Analysis 262(9) pp. 4000-4039 (2012).
- *Asymptotic Analysis for the Generalized Langevin Equation*, (M. Ottobre and G. P.), Nonlinearity, 24 (2011) 1629-1653.

- Goal: sample from a distribution $\pi(x)$ that is known only up to a constant.
- Construct ergodic stochastic dynamics whose invariant distribution is $\pi(x)$.
- There are many different dynamics whose invariant distribution is given by $\pi(x)$.
- Different discretizations of the corresponding SDE can behave very differently, even fail to converge to $\pi(x)$.
- computational efficiency: choose the dynamics that converges to equilibrium as quickly as possible.

- Consider the long time asymptotics of finite dimensional Itô diffusions

$$dX_t = b(X_t)\,dt + \sigma(X_t)\,dW_t, \tag{1}$$

- where W_t is a standard Brownian motion in $\mathbb{R}^d$.
- This is a Markov process with generator

$$\mathcal{L} = b(x) \cdot \nabla + \frac{1}{2}\mathrm{Tr}\left(\Sigma(x) D^2\right), \tag{2}$$

- where $\Sigma = \sigma\sigma^T$.

- In order for $\pi(x)$ to be the invariant distribution of X_t, we require that it is the (unique) solution of the stationary Fokker-Planck equation

$$\nabla \cdot \left(-b(x)\pi(x) + \frac{1}{2}\nabla \cdot (\Sigma(x)\pi(x)) \right) = 0, \tag{3}$$

- together with appropriate boundary conditions (if we are in a bounded domain).
- If the **detailed balance** condition

$$\mathbf{J}_s := -b\pi + \frac{1}{2}\nabla \cdot (\Sigma\pi) = \mathbf{0}, \tag{4}$$

is satisfied, then the process X_t is reversible wrt $\pi(x)$.

- A stationary diffusion process X_t is **reversible** (wrt $\pi(x)$) if X_t and X_{T-t} have the same law for $t \in [0, T]$.
- Time reversibility is equivalent to:
  - The detailed balance condition (4);
  - the generator $\mathcal{L}$ being self-adjoint in $L^2(\mathbb{R}^d; \pi(x))$ (equivalently, the Fokker-Planck operator $\mathcal{L}^* = \nabla \cdot (-b(x) + \nabla \Sigma(x))$ being self-adjoint in $L^2(\mathbb{R}^d; \pi^{-1}(x)))$;
  - X_t having zero entropy production rate.

- There are (infinitely) many reversible diffusions that can be used in order to sample from $\pi(x)$:
  - Either fix $b(x)$ in the detailed balance equation (4) and choose the diffusion matrix $\Sigma(x)$, or,
  - fix $\Sigma(x)$ and choose the drift $b(x)$.
- The rate of convergence to equilibrium depends on the tails of the distribution $\pi(x)$ **and** on the choice of diffusion process X_t (Stramer and Tweedie 1999, Bakry, Cattiaux, Guillin, 2008).
- The overdampled Langevin dynamics, $\Sigma = 2I$, $b(x) = \nabla \log \pi(x)$

$$dX_t = \nabla \log \pi(X_t)\, dt + \sqrt{2} dW_t.$$

is not (always) the best choice.

- We can also use higher order Markovian models, for example the underdamped Langevin dynamics

$$dq_t = p_t\, dt, \quad dp_t = \nabla \log(\pi)(q_t)\, dt - \gamma p_t\, dt + \sqrt{2\gamma}\, dW_t. \qquad (5)$$

- The generator is

$$\mathcal{L} = p \cdot \nabla_q + \nabla_q \log \pi(q) \cdot \nabla_p + \gamma\big(- p \cdot \nabla_p + \Delta_p \big). \qquad (6)$$

- We have convergence to the equilibrium distribution

$$\psi_\infty(p, q) = \frac{1}{(2\pi)^{N/2} Z} \pi(q) e^{-p^2/2}.$$

- The parameter $\gamma > 0$ in (5) can be tuned in order to optimize the rate of convergence to equilibrium.

To speed up convergence to the target distribution $\pi(x)$:

- Choose optimally $(b(x), \Sigma(x))$ within the class of reversible diffusions.
- Consider higher order Markovian models (underdamped Langevin, generalized Langevin equations...) and optimize over the set of parameters in these models, e.g. the friction coefficient γ.
- Consider time dependent temperature (simulated annealing): Holley, Kusuoka, Stroock: **Asymptotics of the spectral gap with applications to the theory of simulated annealing** (1989).

To speed up convergence to the target distribution $\pi(x)$ (contd.):

- Choose an appropriate numerical scheme for the solution of the underlying SDE: Stramer, Tweedie, **Langevin-type models. I. Diffusions with given stationary distributions and their discretizations.** (1999).
- Use a Metropolis-Hastings alrgorithm (acceptance-rejection step): Stramer, Tweedie, **Langevin-type models. II. Self-targeting candidates for MCMC algorithms.** (1999).

Choose optimally $(b(x), \Sigma(x))$ within the class of reversible diffusions:

- Convergence to equilibrium is determined by the first nonzero eigenvalue of the generator of X_t, which is a self-adjoint operator in $L^2(\mathbb{R}^d; \pi(x))$:

$$\lambda_1 = \min_{\phi \in D(\mathcal{L}),\, \phi \neq const} \frac{\langle -\mathcal{L}\phi, \phi \rangle}{\|\phi\|^2}. \tag{7}$$

- We want to maximize λ_1 over the set of drift and diffusion coefficients that satisfy the detailed balance condition (4). This leads to a max-min problem for a self-adjoint operator:

$$\lambda_1 = \max_{b, \Sigma, J_s = 0} \min_{\phi \in D(\mathcal{L}),\, \phi \neq const} \frac{\langle -\mathcal{L}\phi, \phi \rangle}{\|\phi\|^2}. \tag{8}$$

- If we stay within the class of reversible diffusions then it is sufficient to consider spectral optimization problems for selfadjoint operators.
- When sampling from multimodal distributions (metastability...) it is reasonable to expect that a well chosen space-dependent diffusion coefficient can speed up convergence to equilibrium.

- Consider the overdamped Langevin dynamics

$$dX_t = -\nabla V(X_t)\, dt + \sqrt{2\beta^{-1}}\, dW_t. \tag{9}$$

- The law (PDF) ψ_t of X_t satisfies the Fokker-Planck equation

$$\partial_t \psi_t = \nabla \cdot (\nabla V \psi_t + \nabla \psi_t) \,. \tag{10}$$

- Under appropriate assumptions on the potential, ψ_∞ satisfies a Poincaré inequality: there exists $\lambda > 0$ such that for all probability density functions ϕ,

$$\int_{\mathbb{R}^N} \left(\frac{\phi}{\psi_\infty} - 1 \right)^2 \psi_\infty dx \leq \frac{2}{\lambda} \int_{\mathbb{R}^N} \left| \nabla \left(\frac{\phi}{\psi_\infty} \right) \right|^2 \psi_\infty dx \,. \tag{11}$$

- The optimal parameter λ in (11) is the spectral gap of the Fokker-Planck operator $\nabla \cdot (\nabla V \cdot + \nabla \cdot)$, which is self-adjoint in $L^2(\mathbb{R}^N, \psi_\infty^{-1}\, dx)$.
- (11) is equivalent to exponential convergence to equilibrium for (9): for all initial conditions $\psi_0 \in L^2(\mathbb{R}^N, \psi_\infty^{-1} dx)$, for all times $t \geq 0$,

$$\|\psi_t - \psi_\infty\|_{L^2(\psi_\infty^{-1} dx)} \leq e^{-\lambda t} \|\psi_0 - \psi_\infty\|_{L^2(\psi_\infty^{-1} dx)}, \tag{12}$$

- We will consider the nonreversible dynamics (Hwang et al 1993, 2005).
$$dX_t^b = \big(-\nabla V(X_t^b) + b(X_t^b) \big)\, dt + \sqrt{2}\, dW_t, \tag{13}$$
where b is taken to be divergence-free with respect to the invariant distribution $\psi_\infty\, dx$:
$$\nabla \cdot \big(b e^{-V} \big) = 0. \tag{14}$$
- This ensures that $\psi_\infty(x)\, dx$ is still the invariant measure of the dynamics (13).
- We can construct such vector fields by taking
$$b = J \nabla V, \quad J = -J^T. \tag{15}$$

- The dynamics (13) is non-reversible: $(X_t^b)_{0 \leq t \leq T}$ has the same law as $(X_{T-t}^{-b})_{0 \leq t \leq T}$ and thus not the same law as $(X_{T-t}^{b})_{0 \leq t \leq T}$.
- Equivalently, the system does not satisfy detailed balance–the stationary probability flux is not zero.
- From (15) it is clear that there are many (in fact, infinitely many) different ways for modifying the reversible dynamics without changing the invariant measure.

- We ask whether the addition of a nonreversible term can improve the rate of convergence to equilibrium and, if so, whether there exists an optimal choice for the perturbation.
- This question leads to a min-max problem:

$$\max_{b \in \mathcal{A}} \min \mathrm{Re}\big(\sigma(-\mathcal{L}_b)\big).$$

- where

$$\mathcal{L}_b = (-\nabla V + b) \cdot \nabla + \Delta.$$

- We are perturbing a self-adjoint operator $\mathcal{L}$ with discrete spectrum by adding an antisymmetric operator that preserves the null space of $\mathcal{L}$. We want to find the antisymmetric perturbation that maximizes the spectral gap.

- The reversible case is the worst (Hwang et al 2005):
  - Convergence to equilibrium (measured in $L^2(\mathbb{R}^d, \psi_\infty^{-1})$) is slowest for the reversible case.
  - In other words: the spectral gap is the smallest for the self-adoint problem.
  - Related work on Poincare and logarithmic Sobolev inequalities for nonreversible diffusions (Arnold, Carlen, Ju 2008).
- Intuition: the addition of a nonreversible (Hamiltonian) part introduces a drift that can help the system escape from metastable states.

- We consider the nonreversible dynamics

$$dX_t = (-I + \delta J)\nabla V(X_t)\, dt + \sqrt{2\beta^{-1}}\, dW_t, \tag{16}$$

- with $\delta \in \mathbb{R}$ and J the standard 2×2 antisymmetric matrix, i.e. $J_{12} = 1,\ J_{21} = -1$. For this class of nonreversible perturbations the parameter that we wish to choose in an optimal way is δ.
- However, the numerical experiments will illustrate that even a non-optimal choice of δ can significantly accelerate convergence to equilibrium.
- We will use the potential

$$V(x, y) = \frac{1}{4}(x^2 - 1)^2 + \frac{1}{2}y^2. \tag{17}$$

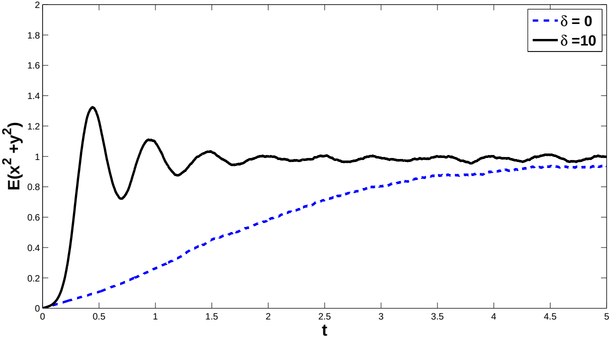

Figure: Second moment as a function of time for (16) with the potential (17). We take 0 initial conditions and $\beta^{-1} = 0.1$.

- The upper bound for the reversible dynamics (9) is still valid:

$$\|\psi_t^b - \psi_\infty\|_{L^2(\psi_\infty^{-1}dx)} \leq e^{-\lambda t}\|\psi_0^b - \psi_\infty\|_{L^2(\psi_\infty^{-1}dx)} . \tag{18}$$

- Adding a non-reversible part to the dynamics cannot be worse than the original dynamics (9) (where $b = 0$) in terms of exponential rate of convergence.
- Our main result is that for a linear drift it is possible to choose b in order to obtain convergence at exponential rate of the form:

$$\|\psi_t^b - \psi_\infty\|_{L^2(\psi_\infty^{-1}dx)} \leq C(V, b)e^{-\overline{\lambda}t}\|\psi_0^b - \psi_\infty\|_{L^2(\psi_\infty^{-1}dx)} , \tag{19}$$

- with $\overline{\lambda} > \lambda$ and $C(V, b) > 1$.

- Let $V(x) = \frac{1}{2}x^T Sx$. The nonreversible perturbations that satisfy (14) are of the form

$$b(x) = -Ax$$

$$A = JS, \quad \text{with } \; J = -J^T \,. \tag{20}$$

- The corresponding SDE is

$$dX_t^J = -(I+J)SX_t^J \, dt + \sqrt{2}\, dW_t. \tag{21}$$

- The invariant distribution is

$$\psi_\infty(x) = \frac{\det(S)^{1/2}}{(2\pi)^{N/2}} \exp\left(-\frac{x^T Sx}{2}\right) \,. \tag{22}$$

Theorem

Define $B_J = (I + J)S$. *Then*

$$\max_{J \in \mathcal{A}_N(\mathbb{R})} \min \mathrm{Re}\,(\sigma(B_J)) = \frac{\mathrm{Tr}(S)}{N}. \tag{23}$$

Furthermore, one can construct matrices $J_{opt} \in \mathcal{A}_N(\mathbb{R})$ *such that the maximum in (23) is attained. The matrix* J_{opt} *can be chosen so that the semigroup associated to* $B_{J_{opt}}$ *satisfies the bound*

$$\left\| e^{-(I+J_{opt})St} \right\| \leq C_N^{(1)} \kappa(S)^{1/2} \exp\left(-\frac{\mathrm{Tr}(S)}{N} t \right), \tag{24}$$

where $\kappa(S) = \|S\| \, \|S^{-1}\|$ *denotes the condition number.*

Theorem

For $B_J = (I + J)S$ with $J \in \mathcal{A}_N$, the drift-diffusion operator $\mathcal{L}_J = -(B_J x).\nabla + \Delta$ defined in $L^2(\mathbb{R}^N, \psi_\infty dx; \mathbb{C})$ with domain of definition

$$D(\mathcal{L}_J) = \left\{ u \in L^2(\mathbb{R}^N, \psi_\infty dx; \mathbb{C}),\ \mathcal{L}_J u \in L^2(\mathbb{R}^N, \psi_\infty dx; \mathbb{C}) \right\}$$

generates a contraction semigroup $(e^{t\mathcal{L}_J})_{t \geq 0}$ and it has a compact resolvent. Optimizing its spectrum with respect to J gives

$$\max_{J \in \mathcal{A}_N(\mathbb{R})} \min \mathrm{Re}\,(\sigma(-\mathcal{L}_J)) = \frac{\mathrm{Tr}(S)}{N}. \tag{25}$$

Theorem (contd.)

Furthermore, the maximum in (25) *is attained for the matrices* $J_{opt} \in \mathcal{A}_N(\mathbb{R})$ *constructed as before. The matrix* J_{opt} *can be chosen so that*

$$\begin{aligned}\Big\| e^{t\mathcal{L}_{J_{opt}}} u - \Big(\int_{\mathbb{R}^N} u \psi_\infty dx \Big) \Big\|_{L^2(\psi_\infty dx)} \\ \leq C_N^{(2)} \kappa(S)^{7/2} \exp\left(-\frac{\mathrm{Tr}(S)}{N} t \right) \Big\| u - \Big(\int_{\mathbb{R}^N} u \psi_\infty dx \Big) \Big\|_{L^2(\psi_\infty dx)}\end{aligned} \tag{26}$$

holds for all $u \in L^2(\mathbb{R}^N, \psi_\infty dx; \mathbb{C})$ *and all* $t \geq 0$.

Corollary

Let us consider the Fokker Planck equation associated to the dynamics (21) *on* X_t^J*:*

$$\partial_t \psi_t^J = \nabla \cdot \left(B_J x \, \psi_t^J + \nabla \psi_t^J \right), \tag{27}$$

where $B_J = (I + J)S$*. Let us assume that* $\psi_0^J \in L^2(\mathbb{R}^N, \psi_\infty^{-1}\, dx)$*. Then, by considering* $J = -J_{opt}$*, where* $J_{opt} \in \mathcal{A}_N(\mathbb{R})$ *refers to the matrix considered in Theorem 2 to get* (26)*. Then the inequality*

$$\left\| \psi_t^J - \psi_\infty \right\|_{L^2(\psi_\infty^{-1} dx)} \le C_N^{(2)} \kappa(S)^{7/2} \exp\left(-\frac{\mathrm{Tr}(S)}{N} t \right) \left\| \psi_0^J - \psi_\infty \right\|_{L^2(\psi_\infty^{-1} dx)},$$

holds for all $t \ge 0$*, when* ψ_∞ *is defined by* (22).

The proofs of these results consist of three steps:

- Solve the min-max problem for the matrix $B_J = (I + J)S$.
- Calculate the spectrum of $\mathcal{L}_J$ in terms of the eigenvalues of B_J.
- Control the constant and optimize wrt to N.

Proposition

Assume that $\tilde{J} = S^{1/2} J S^{1/2} \in \mathcal{A}_N(\mathbb{R})$ and that $S \in \mathcal{S}_N^{>0}(\mathbb{R})$. Then the following two conditions are equivalent:

(i) *The matrix $\tilde{B}_J = S + \tilde{J}$ is diagonalizable (in $\mathbb{C}$) and the spectrum of $\tilde{B}_J$ satisfies*

$$\sigma(\tilde{B}_J) \subset \frac{\mathrm{Tr}(S)}{N} + i\mathbb{R}\,. \tag{28}$$

(ii) *There exists a real symmetric positive definite matrix $Q = Q^T$ such that*

$$\tilde{J}Q - Q\tilde{J} = -QS - SQ + \frac{2\mathrm{Tr}(S)}{N} Q\,. \tag{29}$$

- Let $\{\lambda_k\}_{k=1}^N$ denote the (positive real) eigenvalues of Q (counted with multiplicity), and $\{\psi_k\}_{k=1}^N$ the associated eigenvectors, which form an orthonormal basis of $\mathbb{R}^N$.
- Equation (29) is equivalent to the following two conditions: for all k in $\{1, \dots, N\}$,

$$(\psi_k, S\psi_k)_{\mathbb{R}} = \frac{\mathrm{Tr}(S)}{N} \tag{30}$$

- and, for all $j \neq k$ in $\{1, \dots, N\}$,

$$(\lambda_j - \lambda_k)(\psi_j, \tilde{J}\psi_k)_{\mathbb{R}} = (\lambda_k + \lambda_j)(\psi_j, S\psi_k)_{\mathbb{R}} \,. \tag{31}$$

- This equivalence enables us to develop an algorithm for calculating the optimal nonreversible perturbation.

- The calculation of the eigenvalues of the matrix in the drift is sufficient in order to calculate the spectrum of the generator. Consider the linear SDE (Ornstein-Uhlenbeck process)

$$dX_t = BX_t\,dt + \sigma\,dW_t \tag{32}$$

- The generator is (with $\Sigma = \sigma\sigma^T$)

$$\mathcal{L} = \frac{1}{2}\mathrm{Tr}\left(\Sigma D^2\right) + \langle Bx, \nabla\rangle \tag{33}$$

- Σ can be degenerate, provided that $\mathcal{L}$ is hypoelliptic.
- We assume that (32) is ergodic with unique Gaussian invariant measure $\pi(x)\,dx$.

- The spectrum of $\mathcal{L}$ in $L^p(\mathbb{R}^N; \pi(x))$ has been calculated in Metafune, Pallara and Priola, J. Func. Analysis 196 (2002), pp. 40-60:
- For all $p > 1$ the spectrum consists of integer linear combinations of the drift matrix. It is independent of the diffusion matrix:

$$\sigma(\mathcal{L}) = \left\{ \sum_{j=1}^{r} n_j \lambda_j, \; n_j \in \mathbb{N} \right\}, \tag{34}$$

  where $\{\lambda_j\}_{j=1}^r$ denote the r (distinct) eigenvalues of A. In particular, the spectral gap of the generator $\mathcal{L}$ is determined by the eigenvalues of A.
- A new proof of this result (for $p = 2$) can be found in M Ottobre, G.P., K. Pravda-Starov, J. Func. analysis 262(9), 4000-4039 (2012).
- A similar result can also be proved in infinite dimensions: van Neerven, Infin. Dimens. Anal. Quantum Probab. Relat. Top. 8 (2005), no. 3, 473-495.

To prove the exponential convergence for the solution of the Fokker-Planck equation and to estimate the constant:

- Since $\mathcal{L}_J$ is not self-adjoint, in addition to spectral information we also need estimates on the resolvent.
- We use the second quantization formalism (Wick calculus):

$$-\tilde{\mathcal{L}}_J = a^{*,T}(S - \tilde{J})a \, . \tag{35}$$

- where a, a^* denote creation and annihilation operators.
- Use an appropriate expansion in Hermite polynomials.
- For the constant in front of the exponential we have that

$$C_N^{(2)} = \mathcal{O}(N^3).$$

Algorithm for constructing the optimal nonreversible perturbation

Start from an arbitrary orthonormal basis $(\psi_1, \dots, \psi_N)$.

for $n = 1 : N - 1$ **do**
begin

1. Make a permutation of $(\psi_n, \dots, \psi_N)$ so that
$$(\psi_n, S\psi_n)_{\mathbb{R}} = \max_{k=n,\dots,N} (\psi_k, S\psi_k)_{\mathbb{R}} > \mathrm{Tr}(S)/N$$
and
$$(\psi_{n+1}, S\psi_{n+1})_{\mathbb{R}} = \min_{k=n,\dots,N} (\psi_k, S\psi_k)_{\mathbb{R}} < \mathrm{Tr}(S)/N\,.$$
2. Compute t_* such that $\psi_{t_*} = \cos(t_*)\psi_n + \sin(t_*)\psi_{n+1}$ satisfies $(\psi_{t_*}, S\psi_{t_*})_{\mathbb{R}} = \mathrm{Tr}(S)/N$
3. Use a Gram-Schmidt procedure to change the set of vectors $(\psi_{t_*}, \psi_{n+1}, \dots, \psi_N)$ to an orthonormal basis $(\psi_{t_*}, \tilde{\psi}_{n+1}, \dots, \tilde{\psi}_N)$.
4. Increase n by one and go back to step 1.

end

We consider the three dimensional problem with the symmetric matrix

$$S = \text{diag}(1, 0.1, 0.01). \tag{36}$$

The spectral gap of the optimally perturbed nonreversible matrix (and of the generator of the semigroup) is given by

$$\frac{\mathrm{Tr} S}{3} = 0.37,$$

a substantial improvement over that of S, 0.01.

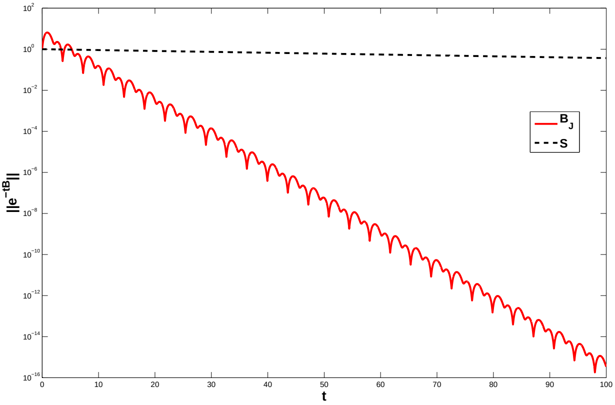

Figure: Norms of the matrix exponentials for the 3×3 diagonal matrix (36) and its optimal nonreversible perturbation.

We consider a 100×100 diagonal matrix with random entries, uniformly distributed on $[0, 1]$. For our example the minimum diagonal element (spectral gap) is 0.0012. On the contrary, the spectral gap of B_J with $J = J_{opt}$ is 0.4762.

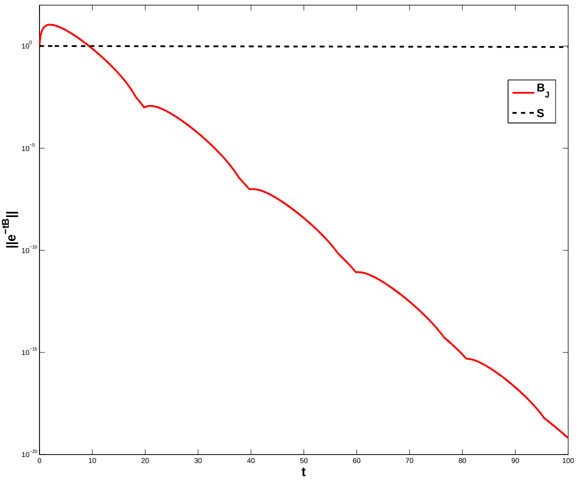

Figure: Norms of the matrix exponentials for a diagonal matrix with random uniformly distributed entries and its optimal nonreversible perturbation for $N = 100$.

- Consider a finite difference approximation of the SPDE

$$\partial_t u = \Delta u + \xi. \tag{37}$$

- We consider this SPDE in 1 dimension on $[0, 1]$ with Dirichlet boundary conditions.
- We can use the finite difference approximation to sample approximately from an infinite dimensional Gaussian measure.

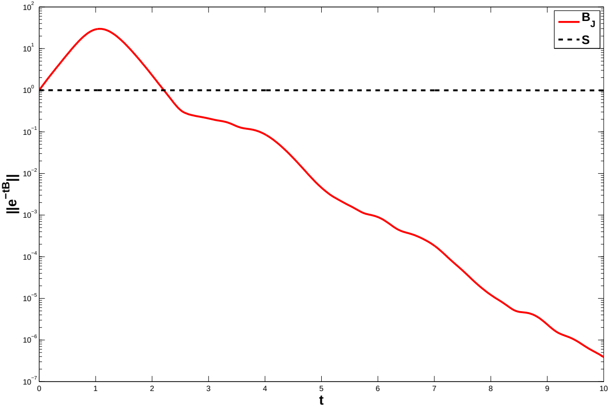

Figure: Norms of the matrix exponentials for the the discrete Laplacian and its optimal nonreversible perturbation for $N = 100$.

- There are many other stochastic dynamics that can be used in order to sample from a given distribution.
- Consider for example the second order Langevin dynamics

$$\ddot{q} = -\nabla V(q) - \gamma \dot{q} + \sqrt{2\gamma\beta^{-1}}\dot{W}.$$

- For $V(q) = \frac{1}{2}\omega_0^2 q^2$ the optimal spectral gap is (using (34))

$$\sigma_{opt} = \frac{\gamma}{2}, \quad for \gamma = 2\omega_0.$$

- We can perturb the Langevin dynamics without changing the invariant measure $\frac{1}{Z}e^{-\beta H(q,p)}$:

$$\dot{q} = p - J\nabla V(q), \quad \dot{p} = -\nabla_q V(q) + Jp - \Gamma p + \sqrt{2\Gamma\beta^{-1}}\,\dot{W}.$$

- We have also considered a general symmetric drift matrix Γ instead of a scalar.

- Consider the three models

$$\dot{q} = -V'(q) + \sqrt{2}\dot{W}, \tag{38a}$$

$$\ddot{q} = -V'(q) - \gamma\dot{q} + \sqrt{2\gamma}\dot{W}, \tag{38b}$$

$$\dot{q} = p, \; \dot{p} = -V'(q) + \lambda z, \; \dot{z} = -\alpha z - \lambda p + \sqrt{2\alpha}\dot{W}. \tag{38c}$$

- All these three models can be used in order to sample from the distribution $Z^{-1}e^{-V(q)}$.
- Notice that there are no control parameters in (38a), 1 (the friction coefficient) in (38b) and 2 (α and λ) in (38c).
- We would like to choose (α, λ) in (38c) in order to optimize the rate of convergence to equilibrium.

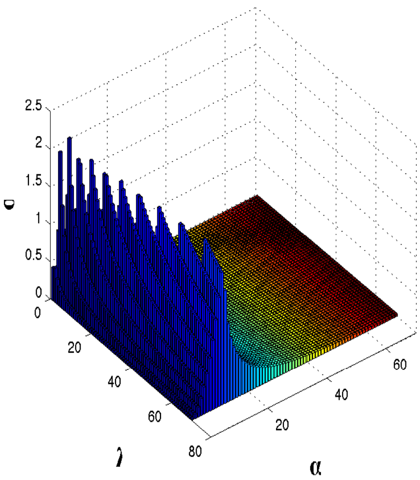

Figure: Spectral gap as a function of α and λ.

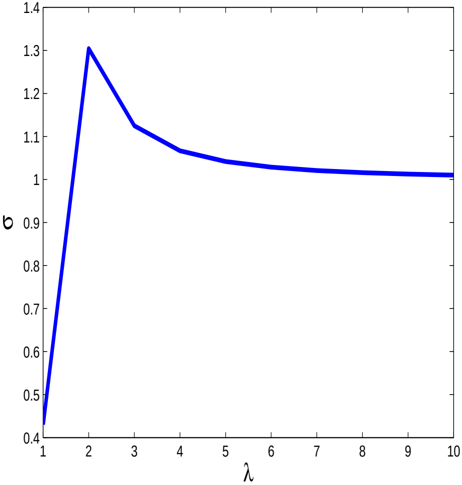

Figure: Spectral gap as a function of λ with $\gamma = \frac{\lambda^2}{\alpha}$ fixed.