École polytechnique fédérale de Lausanne

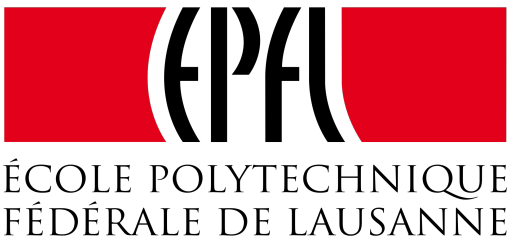
Semester Project

Master in Computational Science and Engineering

Optimization of the rate of convergence of diffusions

Author:
Andrea Scaglioni

Supervisors:
Dr. Sebastian Krumscheid
Prof. Fabio Nobile



Contents

1 Introduction

The rate of convergence to a stationary state of diffusion processes is a critical parameter in several contexts ranging from statistical physics (see [1]) to statistics. An example from statistics is given by the Markov Chain Monte Carlo algorithm, in which the rate of convergence to equilibrium of a Markov chain is the most relevant index of performance of the method. With this example in mind, it is natural to ask whether it is possible to accelerate (increase the rate of convergence) of the diffusion process while preserving the same invariant distribution.

In the present report we study this problem in the case of low-dimensional diffusion processes and propose a possible solution based on the maximization of the spectral of an appropriate differential operator. This approach is discussed and numerical experiments are presented as a verification.

The following is an outline of this document: in section 2 we present some aspects of the theory of diffusion processes, in particular we link them to the (backward) Kolmogorov operator and the Fokker-Planck operator, discuss some of their spectral gap properties and highlight result concerning the rate of convergence to a stationary state. In section 3 we discuss how the rate of convergence can be increased my introducing a suitable perturbation in the dynamics. An example in which the family of optimal perturbations can be determined analytically is presented as well. In section 4 we present the problem of approximating numerically the solution of the stationary Fokker-Planck equation with finite elements. In section 5 we give a general outline of the generalized eigenvalue problem that is required to be solved to determine the spectral gap of the above mentioned differential operators as well as their numericals approximation. In section 6 we present numerical experiments with the aim of verifying the facts discussed in the previous sections in the case of three test problems. Finally, in sections 7 and 8 we present the conclusions and some possible further development of the present project.

2 Setting

In the present section the general setting of the problem is introduced. The presentation is mainly based on [2], [3] and [6].

2.1 Reversible dynamics

Consider the "gradient dynamics":

$$\mathrm{d}X_t = -\nabla V(X_t)\mathrm{d}t + \sqrt{2}\mathrm{d}W_t \tag{1}$$

in which the drift term is given by the gradient of a smooth function $V : \Omega \subset \mathbb{R}^d \to \mathbb{R}$ such that:

$$\lim_{|\boldsymbol{x}|\to\infty} V(\boldsymbol{x}) = \infty \qquad \text{and} \qquad e^{-V(\boldsymbol{x})} \in L^1(\Omega). \tag{2}$$

Functions that satisfy (2) are colled *confining*. Some examples of confining potentials (the first from [3], the others from [7]) with $\Omega = \mathbb{R}^2$ and $\boldsymbol{x} = (x, y) \in \mathbb{R}^2$ are:

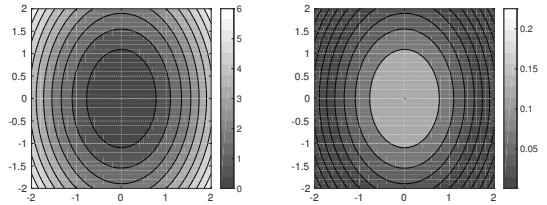
- Quadratic potential: $V(\boldsymbol{x}) := \frac{1}{2}\boldsymbol{x}^T S\boldsymbol{x}$ ($S \in \mathbb{R}^{2\times 2}$ symmetric positive definite);
- Two-wells potential: $V(\boldsymbol{x}) := \frac{1}{4}(x^2-1)^2 + \frac{1}{2}y^2$;
- Three-wells potential:

$$V(\boldsymbol{x}) := 3e^{-x^2-\left(y-\frac{1}{3}\right)^2} - 3e^{-x^2-\left(y-\frac{5}{3}\right)^2} - 5e^{-(x-1)^2-y^2} - 5e^{-(x+1)^2-y^2} + \frac{1}{5}x^4 + \frac{1}{5}\left(y-\frac{1}{3}\right)^4.$$

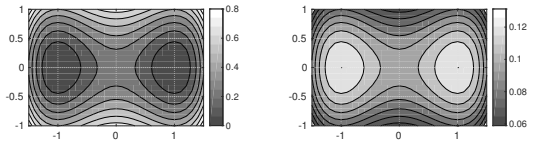
Plots of these potential can be seen in figure 1.

For a general, suitable potential V, the generator of the stochastic process X_t, i.e. the (Backward) Kolmogorov operator is

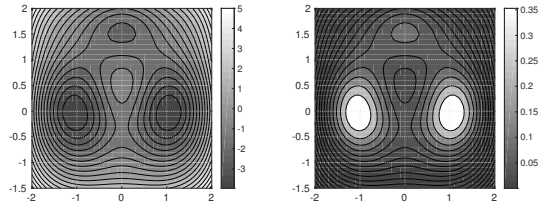
$$L(\psi) := -\nabla V \cdot \nabla\psi + \Delta\psi \tag{3}$$



(a) Quadratic potential



(b) Two-wells potential



(c) three-wells potential

Figure 1: Some example potentials V and the corresponding PDF $\frac{1}{Z}e^{-V}$ where Z is the normalization constant.

and its $L^2(\mathbb{R}^d)$-adjoint is the Fokker-Planck (Forward Kolmogorov) operator:

$$\begin{aligned} L^*(\psi) &:= \boldsymbol{\nabla} \cdot (\nabla V \psi + \nabla \psi) \\ &= \nabla V \cdot \nabla \psi + \Delta V \psi + \Delta \psi \end{aligned} \tag{4}$$

written in conservative and non-conservative form respectively ([2, chapter 4.1]).
Indeed, fixed $u, v \in H^1(\mathbb{R}^d)$

$$\begin{aligned} \langle Lu, v \rangle_{L^2(\mathbb{R}^d)} &= \int_{\mathbb{R}^d} \left(-\nabla V \cdot \nabla u + \Delta u\right) v \mathrm{d}x \\ \text{(integration by parts 2 times)} &= \int_{\mathbb{R}^d} \left(-\nabla V \cdot \nabla uv + \Delta vu\right) \mathrm{d}x \\ \text{(integration by parts)} &= \int_{\mathbb{R}^d} \left(\boldsymbol{\nabla} \cdot (\nabla V v) u + \Delta vu\right) \mathrm{d}x \\ &= \int_{\mathbb{R}^d} \boldsymbol{\nabla} \cdot (\nabla V v + \nabla v) u \mathrm{d}x \\ &= \langle u, L^* v \rangle_{L^2(\mathbb{R}^d)}. \end{aligned}$$

Let the solution of (1) have a PDF $\psi(\cdot, t)$ for all $t > 0$, (therefore $\psi(\cdot, t) \in L^1(\mathbb{R}^d)$, $\psi(\cdot, t) \geq 0$ and $\int_{\mathbb{R}^d} \psi(x, t) \mathrm{d}x = 1 \ \forall t > 0$). Then, ψ satisfies the Fokker-Planck (Forward Kolmogorov) equation:

$$\begin{cases} \frac{\partial \psi}{\partial t} = L^*(\psi) & \text{on } \mathbb{R}^d \\ \psi(0) = \psi_0 \end{cases} \tag{5}$$

where the initial condition ψ_0 is also a PDF.
In case the stochastic differential equation (1) is defined on Ω, an open bounded subset of $\mathbb{R}^d$, analogous definitions of Komogorov and Fokerr-Plank operators can be given, however the Fokker-Plank problem (5) needs to be additionally equipped with boundary conditions (see [2, chapter 4.1]).
The stationary solution of (5), ψ_∞, is (by direct substitution, and remembering it must be a PDF)

$$\psi_\infty = \frac{e^{-V}}{\int_\Omega e^{-V}} \mathrm{d}x.$$

Where the integral is defined since the potential V is assumed to be confining (2).
Observe that by re-defining $\tilde{V}(x) := V(x) + \log\left(\int_{\mathbb{R}^d} e^{-V} \mathrm{d}x\right)$ one simply has $\psi_\infty = e^{-\tilde{V}}$. Therefore, without loss of generality, from now on we will identify ψ_∞ and e^{-V}.

2.1.1 Spectral properties

It may be useful to study the spectral properties of L (as well as those of its L^2-adjoint L^*).
The following lemma will be used in the next proposition, which lists some spectral properties of the two operators.

Lemma 1. *Let $T_1 : H_1 \to H_1$, $T_2 : H_2 \to H_2$ be self-adjoint linear operators onH_1, H_2 respectively with a discrete spectrum. Let moreover T_1, T_2 be unitarily equivalent through the (unitary) operator $U : H_1 \to H_2$ (i.e. $T_1 = U^{-1}T_2U$). Then, T_1 and T_2 have the same eigenvalues.*

A proof can be found in [4]

Proposition 1. *The following properties on the spectrum of L and L^* hold:*

1. *L is self-adjoint in $L^2(\mathbb{R}^d, \psi_\infty)$;*
2. *L^* is self-adjoint in $L^2(\mathbb{R}^d, \psi_\infty^{-1})$;*
3. *$\lambda \in \mathbb{R}$ is an eigenvalue of L iff λ is an eigenvalue of L^**

Proof.

1. Fix $u, v \in L^2(\psi_\infty)$.

$$\begin{aligned}
\langle Lu, v\rangle_{L^2(\psi_\infty)} &= \int_{\mathbb{R}^d} \left(-\nabla V \cdot \nabla u + \Delta u\right) v e^{-V} \mathrm{d}x \\
&= \int_{\mathbb{R}^d} \left(-\nabla V \cdot \nabla u v e^{-V} + \Delta u v e^{-V}\right) \mathrm{d}x \\
\text{(integration by parts)} \quad &= -\int_{\mathbb{R}^d} \nabla V \cdot \nabla u v e^{-V} \mathrm{d}x - \int_{\mathbb{R}^d} \nabla u \nabla (v e^{-V}) \mathrm{d}x \\
&= -\int_{\mathbb{R}^d} \nabla V \cdot \nabla u v e^{-V} \mathrm{d}x - \int_{\mathbb{R}^d} \nabla u \nabla v e^{-V} \mathrm{d}x + \int_{\mathbb{R}^d} \nabla u \cdot \nabla V e^{-V} v \mathrm{d}x \\
&= -\int_{\mathbb{R}^d} \nabla u \nabla v e^{-V} \mathrm{d}x
\end{aligned}$$

   which is symmetric in u and v.

2. Fix $u, v \in L^2(\psi_\infty^{-1})$.

$$\begin{aligned}
\langle L^* u, v\rangle_{L^2(\psi_\infty^{-1})} &= \int_{\mathbb{R}^d} \boldsymbol{\nabla} \cdot \left(\nabla V u + \nabla u\right) v e^{V} \mathrm{d}x \\
\text{(integration by parts)} \quad &= -\int_{\mathbb{R}^d} \left(\nabla V u + \nabla u\right) \cdot \nabla \left(v e^{V}\right) \mathrm{d}x \\
&= -\int_{\mathbb{R}^d} \left(\left(\nabla V u + \nabla u\right) e^{V}\right) \cdot \nabla \left(v e^{V}\right) e^{-V} \mathrm{d}x.
\end{aligned}$$

   Finally, since $\nabla(u e^V) = \nabla u e^V + u e^V \nabla V = \left(\nabla V u + \nabla u\right) e^V$,

$$\langle L^* u, v\rangle_{L^2(\psi_\infty^{-1})} = -\int_{\mathbb{R}^d} \nabla \left(u e^{V}\right) \nabla \left(v e^{V}\right) e^{-V} \mathrm{d}x$$

   which is symmetric in u and v.

3. Due to lemma 1 it is sufficient to prove that L and L^* are unitarily equivalent. We choose $U : L^2(\mathbb{R}^d, e^{-V}) \to L^2(\mathbb{R}^d, e^{V})$ such that $U(f) := f e^{-V}$ (remember V is confining as in (2)). Let

then $v \in L^2(\mathbb{R}^d, e^{-V})$.

$$\begin{aligned}
U\left(L\left(U^{-1}(v)\right)\right) &= e^{-V} L(ve^V) \\
&= e^{-V}\left(-\nabla V \cdot \nabla(ve^V) + \Delta(ve^V)\right) \\
&= e^{-V}\left(-\nabla V \cdot \nabla(ve^V) + \boldsymbol{\nabla} \cdot \nabla(ve^V)\right) \\
&= e^{-V}\left(-\nabla V \cdot \nabla(ve^V) + \boldsymbol{\nabla} \cdot \left(\nabla ve^V + ve^V \nabla V\right)\right) \\
&= e^{-V}\left(-\nabla V \cdot \nabla(ve^V) + \boldsymbol{\nabla} \cdot \left(\nabla ve^V\right) + \nabla(ve^V) \cdot \nabla V + \Delta V ve^V\right) \\
&= e^{-V}\left(\boldsymbol{\nabla} \cdot \left(\nabla ve^V\right) + \Delta V ve^V\right) \\
&= e^{-V}\boldsymbol{\nabla} \cdot \left(\nabla ve^V\right) + e^{-V}\Delta V ve^V \\
\text{(integration by parts)} &= \boldsymbol{\nabla} \cdot (\nabla v) - \nabla\left(e^{-V}\right) \cdot \nabla ve^V + \Delta V v \\
&= \Delta v + e^{-V}\nabla V \cdot \nabla ve^V + \Delta V v \\
&= \Delta v + \nabla V \cdot \nabla v + \Delta V v.
\end{aligned}$$

□

A connection between the properties of the stochastic differential equation (1) and the spectral properties of Kolmogorov (resp. Fokker-Planck) operator is the concept of *reversibility*:

Definition 1. *A stochastic process X_t with invariant distribution ρ_s is called* reversible *if $\forall\ T > 0,\ \forall t < T$, X_t and X_{T-t} have the same distribution assigned initial condition $X_0 \sim \rho_s$.*

Proposition 2. *A stationary diffusion process X_t in $\mathbb{R}^d$ with generator L and invariant measure ψ_∞ is reversible if and only it L is self-adjoint in $L^2(\mathbb{R}^d, \psi_\infty)$*

A proof can be found in [2, theorem 4.5]. Therefore, the stochastic process defined by (1) is reversible due to (1) of proposition 1.

Proposition 1 allows to characterize the kernel of the Fokker-Planck and Kolmogorov operators in the L^2 spaces in which they are self-adjoint. Let us first study the kernel of the Fokker-Planck operator (as in [1]). It was proved in Proposition 1 that L^* is self-adjoint in $L^2(\mathbb{R}^d, \psi_\infty^{-1})$, where ψ_∞ is the invariant distribution of (1). More precisely, it was proved that:

$$\forall u, v \in L^2(\mathbb{R}^d, \psi_\infty^{-1}) \qquad \langle L^* u, v\rangle_{L^2(\psi_\infty^{-1})} = -\langle \nabla(u\psi_\infty^{-1}), \nabla(v\psi_\infty^{-1})\rangle_{L^2(\psi_\infty)}. \tag{6}$$

Therefore, choosing $u = v$

$$\langle L^* u, u\rangle_{L^2(\psi_\infty^{-1})} = -\|\nabla(u\psi_\infty^{-1})\|_{L^2(\psi_\infty)}. \tag{7}$$

Thus, if $u \in \ker(L^*)$ then $\|\nabla(u\psi_\infty^{-1})\|_{L^2(\psi_\infty^{-1})} = 0$ or equivalently $\nabla(u\psi_\infty^{-1}) = \mathbf{0}$ a. e., which implies that, assuming that ψ_∞ and u are continuous, it exists a constant $C \in \mathbb{R}$ such that

$$u = C\psi_\infty.$$

Finally, the kernel of L^* in $L^2(\mathbb{R}^d, \psi_\infty^{-1})$ is

$$\ker(L^*) = \{C\psi_\infty : C \in \mathbb{R}\}. \tag{8}$$

Proceeding analogously with the Kolmogorov operator (3) on $L^2(\mathbb{R}^d, \psi_\infty)$ one finds that its kernel is made of constants:

$$\ker(L) = \mathbb{R}.$$

2.1.2 Convergence to a stationary state

In this section we prove that the convergence to equilibrium of the dynamics (1) is exponentially fast given any initial condition $X_0 \sim \psi_0$. An important ingredient for this proof is the Poincaré inequality.

Definition 2 (Poincaré inequality)**.** *Let $V \in C^2(\mathbb{R}^d)$, $\rho = \frac{1}{Z}e^{-V}$ where $Z := \int_{\mathbb{R}^d} e^{-V}\,dx$. ρ (or, more simply, V) is said to satisfy the* Poincaré inequality *with constant $\lambda > 0$ if:*
$\forall \phi \in C^1(\mathbb{R}^d) \cap L^2(\mathbb{R}^d, \rho) : \int_{\mathbb{R}^d} \phi \rho dx = 0$

$$\lambda \|\phi\|_{L^2(\rho)} \leq \|\nabla \phi\|_{L^2(\rho)}.$$

Theorem 1 (Sufficient condition for Poincaré inequality)**.** *If*

$$\lim_{|x|\to+\infty} \frac{1}{2}|\nabla V(x)|^2 - \Delta V(x) = +\infty \tag{9}$$

then the Poincaré inequality *holds.*

A proof can be found in [5]. The following estimate on the rate of convergence to the stationary solution holds:

Proposition 3. *If the potential V satisfies the Poincaré inequality with constant λ, then the solution of the Fokker-Planck problem (5) ψ_t converges to its stationary distribution $\psi_\infty = e^{-V}$ at exponential rate:*

$$\|\psi(t) - \psi_\infty\|_{L^2(\psi_\infty^{-1})} \leq Ce^{-\lambda t}\|\psi_0 - \psi_\infty\|_{L^2(\psi_\infty^{-1})} \qquad \forall t > 0 \tag{10}$$

where $C > 0$ is a constant.

Proof.

$$\begin{aligned}
\frac{\mathrm{d}}{\mathrm{d}t}\|\psi_t - \psi_\infty\|_{L^2(\psi_\infty^{-1})} &= \int_{\mathbb{R}^d} \frac{\partial}{\partial t}(\psi_t - \psi_\infty)^2 \psi_\infty^{-1} \mathrm{d}x \\
&= \int_{\mathbb{R}^d} 2(\psi_t - \psi_\infty)\frac{\partial}{\partial t}(\psi_t - \psi_\infty)\psi_\infty^{-1}\mathrm{d}x \\
\left(\frac{\partial \psi_\infty}{\partial t} = 0\right) &= \int_{\mathbb{R}^d} 2(\psi_t - \psi_\infty)\frac{\partial \psi_t}{\partial t}\psi_\infty^{-1}\mathrm{d}x \\
(\psi_t \text{ solves } (5)) &= \int_{\mathbb{R}^d} 2(\psi_t - \psi_\infty)\boldsymbol{\nabla}\cdot(\nabla V \psi_t - \nabla \psi_t)\psi_\infty^{-1}\mathrm{d}x \\
(\text{use } (6)) &= -2\int_{\mathbb{R}^d} \nabla\left(\psi_t\psi_\infty^{-1}\right)\nabla\left((\psi_t - \psi_\infty)\psi_\infty^{-1}\right)\psi_\infty \mathrm{d}x \\
&= -2\int_{\mathbb{R}^d} \nabla\left(\psi_t\psi_\infty^{-1}\right)\cdot\nabla\left(\psi_t\psi_\infty^{-1}\right)\psi_\infty \mathrm{d}x \\
&= -2\int_{\mathbb{R}^d} |\nabla\left(\psi_t\psi_\infty^{-1}\right)|^2\psi_\infty \mathrm{d}x \\
(\text{Poincaré inequality}) &\leq -2\lambda\int_{\mathbb{R}^d}\left(\psi_t\psi_\infty^{-1} - 1\right)^2\psi_\infty \mathrm{d}x \\
&= -2\lambda\int_{\mathbb{R}^d}(\psi_t - \psi_\infty)^2\psi_\infty^{-1}\mathrm{d}x \\
&\leq -\lambda\|\psi_t - \psi_\infty\|^2_{L^2(\psi_\infty^{-1})}.
\end{aligned}$$

Finally, applying Gronwall's lemma, we obtain (10). □

Remark 1. *In Proposition 1 we proved that $\langle Lu, u\rangle_{L^2(\psi_\infty)} = -\|\nabla u\|^2_{L^2(\psi_\infty)}$. Due to this fact the Poincaré inequality can be re-written as:*

$$\exists \lambda > 0 : \forall u \in L^2(\mathbb{R}^d, \psi_\infty) \qquad \lambda \leq \frac{\langle -Lu, u\rangle_{L^2(\psi_\infty)}}{\|u\|^2_{L^2(\psi_\infty)}}$$

i.e. the Kolmogorov operator has a spectral gap*:*

$$0 \leq \lambda \leq \inf_{u \in L^2(\mathbb{R}^d, \psi_\infty),\ u \neq 0} \frac{\langle -Lu, u \rangle_{L^2(\psi_\infty)}}{\|u\|^2_{L^2(\psi_\infty)}}.$$

The spectral gap of L is therefore a bound from below of the optimal constant appearing in proposition 3.

2.2 Non-reversible dynamics

We can consider the family of SDEs:

$$\mathrm{d}X_t = \boldsymbol{b}(X_t)\mathrm{d}t + \sqrt{2}\mathrm{d}W_t \tag{11}$$

where $\boldsymbol{b} : \Omega \subset \mathbb{R}^d \to \mathbb{R}^d$. For example, the SDE (1) is obtained choosing $\boldsymbol{b} = -\nabla V$.
For a general drift $\boldsymbol{b}$ the generator (backward Kolmogorov operator) of (11) is

$$L(\psi) := \boldsymbol{b} \cdot \nabla \psi + \Delta \psi \tag{12}$$

and its Fokker-Planck operator (as before the $L^2(\mathbb{R}^d)$-adjoint of (12)) is

$$L^*(\psi) := \boldsymbol{\nabla} \cdot (-\boldsymbol{b}\psi + \nabla \psi). \tag{13}$$

We assume that the dynamics (11) has a unique stationary distribution, that is the solution of the Fokker-Planck equation:

$$L^*(\psi) = 0 \qquad \text{on } \mathbb{R}^d$$

where L^* is the Fokker-Plank operator:

$$L^*(\psi) = \boldsymbol{\nabla} \cdot (J(\psi)) \qquad \text{where} \qquad J(\psi) := -\boldsymbol{b}\psi + \nabla \psi \tag{14}$$

$J(\cdot)$ is called the *probability flux*.
Observe that in this general case the invariant distribution ρ_s cannot be expressed as a function of the drift $\boldsymbol{b}$ in a simple way.
A sufficient condition for ψ to solve $L^*(\psi) = 0$ is that $J(\psi) = 0$. Because of the importance of this condition, we give the following definition:

Definition 3. *We say that ψ satisfies the* detailed balance condition *if*

$$J(\psi) = 0. \tag{15}$$

One can ask under which conditions on $\boldsymbol{b}$ the solution of (11) is a reversible diffusion. Due to Proposition 2 it is sufficient to find for which drifts $\boldsymbol{b}$ the Kolmogorov operator (11) L is self-adjoint.

Proposition 4. *Consider the dynamics (11) and assume that it admits a unique invariant distribution ρ_s. Then the Kolmogorov operator (12) associated to (11) is self-adjoint in $L^2(\mathbb{R}, \rho_s)$ if and only if $J(\rho_s) = 0$, where J is the probability flux (14).*

Proof. Consider $f, h \in H^1(\mathbb{R}^d, \rho_s)$

$$\begin{aligned}
\langle Lf, h\rangle_{\rho_s} &= \int_{\mathbb{R}^d} L(f)h\rho_s \mathrm{d}x \\
&= \int_{\mathbb{R}^d} (\boldsymbol{b}\cdot\nabla f + \Delta f)h\rho_s \mathrm{d}x \\
&= \int_{\mathbb{R}^d} (\boldsymbol{b}\cdot\nabla f h\rho_s + \Delta f h\rho_s)\mathrm{d}x \\
\text{(integration by parts)} &= \int_{\mathbb{R}^d} (\boldsymbol{b}\cdot\nabla f h\rho_s - \nabla f\cdot\nabla(h\rho_s))\mathrm{d}x \\
&= \int_{\mathbb{R}^d} (\boldsymbol{b}\cdot\nabla f h\rho_s - \nabla f\cdot\nabla h\rho_s - \nabla f\cdot\nabla\rho_s h)\mathrm{d}x \\
&= \int_{\mathbb{R}^d} (-J(\rho_s)\cdot\nabla f h - \nabla f\cdot\nabla h\rho_s)\mathrm{d}x \\
\text{(integration by parts)} &= \int_{\mathbb{R}^d} (J(\rho_s)\cdot\nabla h f - \nabla f\cdot\nabla h\rho_s)\mathrm{d}x \\
&= -\langle\nabla h, \nabla f\rangle_{\rho_s} + \langle J(\rho_s)\cdot\nabla h\rho_s^{-1}, f\rangle_{\rho_s}.
\end{aligned}$$

Thus, a sufficient condition for (12) to be self-adjoint is $J(\rho_s) = 0$.
This condition is also necessary. To see this, first observe that the second term is antisymmetric in f and h:

$$\begin{aligned}
\langle J(\rho_s)\nabla h\rho_s^{-1}, f\rangle_{\rho_s} &= \int_{\mathbb{R}^d} J(\rho_s)\nabla h f \mathrm{d}x \\
\text{(integration by parts)} &= \int_{\mathbb{R}^d} \boldsymbol{\nabla\cdot} J(\rho_s) f h \mathrm{d}x - \int_{\mathbb{R}^d} J(\rho_s)\cdot\nabla f h \mathrm{d}x \\
(\ \rho_s \text{ is in the kernel of 13}) &= -\int_{\mathbb{R}^d} J(\rho_s)\cdot\nabla f h \mathrm{d}x.
\end{aligned}$$

Therefore, $\langle J(\rho_s)\cdot\nabla h\rho_s^{-1}, f\rangle_{\rho_s}$ must be zero, which is true only if $J(\rho_s) = 0$. □

Therefore, (11) is reversible if and only if the detailed balance condition (15) holds for ρ_s, i.e. $J(\rho_s) = 0$.
Assuming $\rho_s > 0$ on $\mathbb{R}^d$, we can write the invariant distribution as $\rho_s = e^{-\Phi}$, where $\Phi = -\log\rho_s$ is called a *generalized potential*. The detailed balance condition then simplifies to:

$$\boldsymbol{b} = -\nabla\Phi. \tag{16}$$

To conclude, the dynamics (11) is reversible if and only if its drift is the gradient of a potential.

2.2.1 Perturbed dynamics

Referring again to the dynamics (1) with invariant distribution ψ_∞, a natural non-reversible dynamics whose invariant distribution is still ψ_∞ is obtained with $\boldsymbol{b} = -\nabla V + \boldsymbol{p}$:

$$\mathrm{d}X_t^{\boldsymbol{p}} = (-\nabla V(X_t^{\boldsymbol{p}}) + \boldsymbol{p}(X_t^{\boldsymbol{p}}))\mathrm{d}t + \sqrt{2}\mathrm{d}W_t, \tag{17}$$

where the *perturbation* $\boldsymbol{p}$ is taken divergence free with respect to ψ_∞:

$$\mathrm{div}(\boldsymbol{p}e^{-V}) = 0. \tag{18}$$

Indeed, the Fokker-Planck operator associated to (17) is

$$L_{\boldsymbol{p}}^*(\psi) = \mathrm{div}((\nabla V - \boldsymbol{p})\psi + \nabla\psi)$$

and ψ_∞ is in its kernel.
The process $X_t^{\boldsymbol{p}}$ has density $\psi^{\boldsymbol{p}}(t)$ for all t, which solves the Fokker-Planck problem:

$$\begin{cases} \frac{\partial \psi^{\boldsymbol{p}}}{\partial t} = L^*_{\boldsymbol{p}}(\psi^{\boldsymbol{p}}) & \text{on } \Omega \\ \psi^{\boldsymbol{p}}(0) = \psi_0^{\boldsymbol{p}}. \end{cases} \tag{19}$$

An example of perturbation that satisfies (18) is $\boldsymbol{p} = J\nabla V$ where J is a constant, antisymmetric matrix.
It can be proved that introducing a perturbation cannot make the convergence slower:

Proposition 5. *Consider ψ_t, solution of the unperturbed Fokker-Planck problem(5) and $\psi_t^{\boldsymbol{p}}$ solution of (19) given the divergence free perturbation $\boldsymbol{p}$ as in (18). If the stationary distribution ψ_∞ satisfies the Poincaré inequality with constant λ, then*

$$\|\psi_t^{\boldsymbol{p}} - \psi_\infty\|_{L^2(\psi_\infty^{-1})} \leq e^{-\lambda t}\|\psi_0^{\boldsymbol{p}} - \psi_\infty\|_{L^2(\psi_\infty^{-1})} \qquad \forall t \tag{20}$$

Proof. The result in proposition 3 can be replicated. □

While in the reversible case the optimal constant λ could be identified with the spectral gap of the Fokker-Planck (or Kolmogorov) operator, in this case establishing an analogous result is not trivial (also observe that the Kolmogorov and Fokker-Planck operator are not self-adjoint for $\boldsymbol{p} \neq \boldsymbol{0}$, therefore they may not even have real eigenvalues). In view of proposition 5, an estimate similar to (10) can be postulated:

$$\|\psi_t^{\boldsymbol{p}} - \psi_\infty\|_{L^2(\psi_\infty^{-1})} \leq C(V,\boldsymbol{p})e^{-\bar{\lambda}t}\|\psi_0^{\boldsymbol{p}} - \psi_\infty\|_{L^2(\psi_\infty^{-1})} \qquad \forall t \geq 0 \tag{21}$$

where $C(V,\boldsymbol{p}) > 1$ and $\bar{\lambda} \geq \lambda$ (λ being the spectral gap of the unperturbed Kolmogorov operator L (3)).
In this case, an obvious but not trivial question is how to characterize the maximum $\bar{\lambda}$ for which (21) holds.
Finally, observe that in this case the rate of convergence also depends on the multiplicative constant C (which actually depends on V and $\boldsymbol{p}$). Also this quantity has no simple relation with the spectrum of L or L^*.

3 Maximizing the rate of convergence

The aim of the present work is to study the acceleration of the convergence of (17) induced by a perturbation $\boldsymbol{p}$ satisfying the divergence-free condition (18). Clearly, we would like to find a perturbation $\boldsymbol{p}$ that maximizes the rate of convergence. This task is not trivial for several reasons. Firstly, as remarked at the end of the previous section, there is no simple relation between rate of convergence and spectral gap of $L_{\boldsymbol{p}}$. We can however follow [8] and conjecture that a good indicator of the rate of convergence to equilibrium is the real part of the eigenvalue of $L^{\boldsymbol{p}}$ with smallest real part:

$$\lambda(\boldsymbol{p}) = \operatorname{Re}\left(\operatorname{argmin}_{\mu \in \sigma(L^{\boldsymbol{p}})\setminus\{0\}} |\operatorname{Re}(\mu)|\right) \tag{22}$$

where $\sigma(L)$ denotes the set of the eigenvalues of the differential operator L.
Moreover, consider that attempting to simply maximize $\lambda(\boldsymbol{p})$ by choosing a perturbation $\boldsymbol{p} \in P$ (P being the set of divergence free perturbations as in (18)):

$$\lambda(\boldsymbol{p}_{\text{opt}}) = \max_{\boldsymbol{p} \in P} \lambda(\boldsymbol{p})$$

may not be an effective strategy as at the same time the constant $C(V,\boldsymbol{p})$ may "blow up".
There are cases in which the set of optimal perturbations can be characterized. An example is the one of the quadratic potential. Moreover, in this case the quantity (22) determines the rate of convergence. We present some results about this case in the next section.

3.1 An example with the quadratic potential

In this section we present some results from [3]. Consider the case of the quadratic potential:

$$V(\boldsymbol{x}) := \frac{1}{2}\boldsymbol{x}^T S\boldsymbol{x} \ \forall \boldsymbol{x} \in \mathbb{R}^n \qquad (S \in \mathbb{R}^{d\times d} \text{ symmetric positive definite}),$$

the class of divergence-free perturbation $\boldsymbol{p} \in P := \{J\nabla V : J \in \mathcal{A}_d(\mathbb{R})\}$ ($\mathcal{A}_d(\mathbb{R})$ is the set of antisymmetric matrices in $\mathbb{R}^{d\times d}$) and defined $B_J := (I+J)S$. As $\nabla V(\boldsymbol{x}) = S\boldsymbol{x}$, the unperturbed Kolmogorov and Fokker-Planck operators reads

$$L = -(S\boldsymbol{x})\cdot\nabla + \Delta, \qquad L^* = \boldsymbol{\nabla}\cdot((S\boldsymbol{x})\cdot + \nabla),$$

where by $S\boldsymbol{x}$ we denote the linear operator on $\mathbb{R}^d$ induced by S, while their perturbed ($\boldsymbol{p} = J\nabla V = JS$) counterparts read:

$$L_J = -(B_J\boldsymbol{x})\cdot\nabla + \Delta, \qquad L_J^* = \boldsymbol{\nabla}\cdot((B_J\boldsymbol{x})\cdot + \nabla).$$

In this case there exists an algorithm to build an optimal perturbation $\boldsymbol{p}_{opt} = J_{opt}S$ (see [3, figure 1] for the pseudocode). Moreover, the convergence estimate (21) can be written as:

$$\|\psi_t^{\boldsymbol{p_{opt}}} - \psi_\infty\|_{L^2(\psi_\infty^{-1})} = C_d\kappa(S)^{\frac{7}{2}}e^{-\frac{Tr(S)}{N}t}\|\psi_0^{\boldsymbol{p_{opt}}} - \psi_\infty\|_{L^2(\psi_\infty^{-1})}$$

where C_d is a constant depending on the dimension of the domain and $\kappa(S)$ denotes the conditioning number of S.
Moreover, in this example the optimal rate of convergence $\bar{\lambda} = \frac{Tr(S)}{d}$ is actually the spectral gap of the matrix $B_{J\text{opt}}$:

$$\max_{J\in\mathcal{A}_d(\mathbb{R})} \min Re(\sigma(B_J)\setminus\{0\}) = \min Re(\sigma(B_{J\text{opt}})\setminus\{0\}) = \frac{Tr(S)}{d}.$$

Finally, thanks to the following lemma, this quantity is also equal to the spectral gap of L:

$$\max_{J\in\mathcal{A}_d(\mathbb{R})} \min Re(\sigma(L_J)\setminus\{0\}) = \min Re(\sigma(L_{J\text{opt}})\setminus\{0\}) = \frac{Tr(S)}{d}.$$

Lemma 2. *Given a positive definite matrix $B \in \mathbb{R}^{n\times n}$ with r distinct eigenvalues $\{\lambda_i\}_{i=1}^r$, the differential operator $L = -(Bx)\cdot\nabla + \Delta$ on $L^2(\mathbb{R}^n, \psi_\infty)$ (ψ_∞ the invariant measure of the dynamics generated by L) has a spectrum of the following form:*

$$\sigma(L) = \left\{ -\sum_{j=1}^r n_j\lambda_j, n_j \in \mathbb{N} \right\} \tag{23}$$

A proof can be found in [3]

4 Numerical solution of stationary Fokker-Planck problem

As we have seen in section 2.1.1, the kernel of the Fokker-Planck equations can be characterized by (8), which is a one-dimensional linear space. As the solution of the Fokker-Planck (5) problem is a PDF, we can we can find a unique solution by imposing $\int_{\mathbb{R}^d} C\psi_\infty \mathrm{d}x = 1$ or, equivalently $C = \left(\int_{\mathbb{R}^d} e^{-V}\mathrm{d}x\right)^{-1}$.

Despite the fact that the exact solution is known (indeed, V is given), one may want to verify numerically that the stationary Fokker-Planck equation is solved by $\frac{e^{-V}}{\int_{\mathbb{R}^d} e^{-V}\mathrm{d}x}$.
In order to formulate the numerical problem (for example, through finite elements) one has to face the problem that a direct application of the numerical method is not possible since the domain is unbounded. Therefore, the analytical problem on $\mathbb{R}^d$ has to be "modeled" with another problem on a bounded domain (and therefore introducing an "modelling" error). There are two properties we would like this model to satisfy:

- well posedness;
- its solution converges (in some sense) to the analytical solution of the Fokker-Planck problem in $\mathbb{R}^d$ as the bounded domain is enlarged.

Consider the stationary (unperturbed) Fokker-Planck equation on Ω bounded open set of $\mathbb{R}^d$:

$$\begin{cases} L^*(\psi) = \operatorname{div}(\nabla V \psi + \nabla \psi) = 0 & \text{on } \Omega \\ \psi = 0 & \text{on } \partial\Omega \end{cases} \tag{24}$$

The choice of homogeneous Dirichlet boundary conditions intuitively comes from the fact that, since the solution of the stationary Fokker-Planck equation on $\mathbb{R}^d$ $\psi_\infty \in L^1(\mathbb{R}^d)$, $\lim_{|\boldsymbol{x}|\to+\infty} \psi_\infty(\boldsymbol{x}) = 0$.

Let us first attempt to prove well posedness.Consider the following theorem (from [10], where a proof can also be found), which gives sufficient condition for well posedness for elliptic problems with Dirichlet boundary conditions:

Theorem 2. *Consider Ω open bounded set of $\mathbb{R}^d$ with piecewise smooth boundary $\partial\Omega$. and the following elliptic PDE equipped with Dirichlet boundary conditions*

$$-\boldsymbol{\nabla} \cdot (\sigma \nabla u) + \beta \cdot \nabla u + \mu u = f$$

where $f \in L^2(\Omega)$, $\sigma \in [L^\infty(\Omega)]^{d\times d}$ satisfies

$$\forall \xi \in \mathbb{R}^d \ \sum_{i,j=1}^{d} \sigma_{ij}\xi_i\xi_j \geq \sigma_0 \|\xi\|_d^2 \ \text{a.e. in } \Omega,$$

$\beta \in [L^\infty(\Omega)]^d$ with $\boldsymbol{\nabla} \cdot \beta \in L^\infty(\Omega)$ and $\mu \in L^\infty(\Omega)$. Set

$$p = essinf_{x\in\Omega} \left(\mu - \frac{1}{2} \boldsymbol{\nabla} \cdot \beta \right)$$

and let C_Ω be the Poincaré constant of Ω:

$$\forall u \in H_0^1(\Omega) \qquad C_\Omega \|u\|_{L^2(\Omega)} \leq \|\nabla u\|_{L^2(\Omega)}. \tag{25}$$

The problem is well posed if

$$\sigma_0 + min\left(0, \frac{p}{C_\Omega}\right) > 0.$$

Therefore, in the case considered in this report($f = 0$, $\sigma = -I$, $\sigma_0 = 1$, $\beta = -\nabla V$, $\mu = -\Delta V$), the theorem leads to:

$$\begin{aligned} p &= \operatorname{essinf}_{x\in\Omega} \left(-\Delta V - \frac{1}{2} \boldsymbol{\nabla} \cdot (-\nabla V) \right) \\ &= \operatorname{essinf}_{x\in\Omega} \left(-\frac{1}{2} \Delta V \right) \\ &= -\frac{1}{2} \operatorname{esssup}_{x\in\Omega} (\Delta V) \end{aligned}$$

and a sufficient condition for well posedness is:

$$1 + \min\left(0, -\frac{\operatorname{esssup}_{x\in\Omega} \Delta V}{2C_\Omega}\right) > 0$$

which simplifies to:

$$\operatorname{esssup}_{x\in\Omega} \Delta V < 2C_\Omega. \tag{26}$$

Remark 2. *Observe that in the case (26) is satisfied, the unique solution of (24) is the identically zero function. Contrarily, if the condition is not satisfied the problem may be ill posed, and in particular admit more than one solution.*

To assess whether the theorem can be applied or not, the value of the Poincaré constant is needed. The following theorem from [9] allows to quantify it (given some hypothesis on the domain Ω):

Theorem 3. *If the domain $\Omega \subset \mathbb{R}^d$ is a bounded, convex, Lipschitz domain with diameter d then the Poincaré constant defined in theorem 2 satisfies: $C_\Omega < \frac{\pi}{d}$.*

Therefore, as the diameter of the domain increases (the limit in which we are interested in this context), the Poincaré constant decreases. Consequently, if $\Delta V > 0$ theorem 2 cannot be used to grant well posedness.

Loosely speaking, since $\psi \equiv 0$ is always a solution, we actually want the problem not to have a unique solution. This way, among the possible solutions we hope to be able to identify one that is also a PDF. If the potential is *confining* (2 a solution that is also a PDF can be retrieved by imposing an integral constraint:

$$\int_\Omega \psi \mathrm{d}x = 1 \tag{27}$$

at the cost of introducing an error in (24) (but fulfilling the boundary condition exactly).
This can be expressed as a constrained minimization problem

$$\text{Find } u \in H_0^1(\Omega) : \frac{1}{2}|a(u,u)|^2 = \min_v \left\{ \frac{1}{2}|a(v,v)|^2 : v \in H_0^1(\Omega), v_{|\partial\Omega} = 0, \int_\Omega v \mathrm{d}x = 1 \right\} \tag{28}$$

where $a(\psi, v) := \int_\Omega (-\nabla\psi \cdot \nabla v + \nabla V \cdot \nabla \psi v + \Delta V \psi v)\, \mathrm{d}x$ is the bilinear form appearing in the weak formulation of (24).
The problem can be "unconstrained" as follows:
Find $\psi \in H_0^1(\Omega), \lambda \in \mathbb{R}$ such that

$$\begin{cases} a(\psi, v) & = \lambda \int_\Omega v \mathrm{d}x \qquad \forall v \in H_0^1(\Omega) \\ \int_\Omega \psi \mathrm{d}x & = 1 \end{cases} \tag{29}$$

Problem (29) can be tackled numerically using the finite elements method. Let V_h be a finite dimensional linear subspace of $H_0^1(\Omega)$ of dimension N with basis $\{\phi_i\}_{i=1}^N$. Then, calling ψ_h the unknown of the numerical problem, it exists $\boldsymbol{\psi} \in \mathbb{R}^N$ such that the solution $\psi_h = \sum_{i=1}^N \psi_i \phi_i$ and choosing $v = \phi_i$ problem (29) can be approximated as:

$$\begin{aligned} &\sum_{j=1}^N \psi_j a(\phi_j, \phi_i) = \lambda \int_\Omega \phi_i \mathrm{d}x \qquad \forall i = 1, ..., N \\ &\sum_{j=1}^N \psi_i \int_\Omega \phi_i \mathrm{d}x = 1 \end{aligned} \tag{30}$$

or, in matrix form:

$$\left(\begin{array}{c|c} A & \boldsymbol{b} \\ \hline \boldsymbol{b}^T & 0 \end{array}\right) \begin{pmatrix} \boldsymbol{\psi} \\ \lambda \end{pmatrix} = \begin{pmatrix} \mathbf{0} \\ 1 \end{pmatrix} \tag{31}$$

where $A \in \mathbb{R}^{N \times N}$, $(A)_{i,j} = a(\phi_j, \phi_i)$, $\boldsymbol{b} \in \mathbb{R}^N$, $b_i = \int_\Omega \phi_i \mathrm{d}x$ and $\mathbf{0}$ is the zero vector of $\mathbb{R}^N$.
To summarize, by truncating the domain (and imposing boundary conditions) and introducing an error due to the Lagrange multiplier strategy adopted above, a "modelling" error is made. In addition, a numerical error is introduced by the finite elements approximation. Calling ψ the analytical solution on $\mathbb{R}^d$, ψ_m the analytical model of the solution (i.e. solution of (29) on Ω) and ψ^h the numerical solution we have the error splitting in "modelling error" and "numerical error":

$$\|\psi - \psi^h\|_{L^2(\mathbb{R}^d)} \leq \|\psi - \psi_m\|_{L^2(\mathbb{R}^d)} + \|\psi_m - \psi^h\|_{L^2(\Omega)} \tag{32}$$

where in the first term of the right hand side ψ_m is extended with the identically zero function outside Ω. Finally, observe that the second term can be identified through basic finite elements analysis as a $\mathcal{O}(h^{m+1})$ where m is the order of finite elements used.

5 Numerical eigenvalue problems

To tackle the research of eigenvalues of the operators L and L^* numerically we first formulate the analytical eigenvalue problems as:

$$\text{Find } \lambda \in \mathbb{R}, \boldsymbol{v} \in L^2(\Omega, \psi_\infty) : \begin{cases} Lv = \lambda v & \text{in } \Omega \\ \partial_{\boldsymbol{n}} v = 0 & \text{on } \partial\Omega \end{cases} \tag{33}$$

and

$$\text{Find } \lambda \in \mathbb{R}, \boldsymbol{v} \in L^2(\Omega, \psi_\infty^{-1}) : \begin{cases} L^*v = \lambda v & \text{in } \Omega \\ v = 0 & \text{on } \partial\Omega \end{cases} \tag{34}$$

where $\boldsymbol{n}$ denotes the outward unit normal vector to $\partial\Omega$ and ψ_∞ is now the restriction of ψ_∞ to Ω. The choice of boundary conditions can be motivated as follows: as the functions in the kernel of each operator are the eigenvectors corresponding to the eigenvalue 0, we impose the boundary condition that apply to those: being $\ker(L) = \mathbb{R}$ (see (2.1.1)) we choose homogeneous Neumann boundary conditions in (33) and begin $\ker(L^*) = \mathbb{R}\psi_\infty$ (see (8)) we choose homogeneous Dirichlet boundary conditions in (34).
The operators L and L^* defined on $L^2(\Omega, \psi_\infty)$ and $L^2(\Omega, \psi_\infty^{-1})$ are (still) such that:

- L, L^* are adjoint in $L^2(\Omega)$;
- L is self-adjoint in $L^2(\Omega, \psi_\infty)$;
- L^* is self-adjoint in $L^2(\Omega, \psi_\infty^{-1})$;
- λ is an eigenvalue of L in $L^2(\Omega, \psi_\infty)$ iff λ is an eigenvalue of L^* in $L^2(\Omega, \psi_\infty^{-1})$

As a matter of facts, the derivations obtained in the previous sections (in particular, integrations by parts) with $\mathbb{R}^d$ as domain of definition of the operators can be replicated in this setting thanks to boundary conditions.

Let us formulate the finite elements approximation of the eigenvalue problem for the Kolmogorov operator (the case of the Fokker-Planck operator is analogous) following [11] (chapter 7). The weak for of problem (33) reads:

$$\text{Find } \lambda \in \mathbb{R}, u \in H^1(\Omega, \psi_\infty) : a(u, v) = \lambda \langle u, v\psi_\infty \rangle_{L^2(\Omega)} \tag{35}$$

where

$$a(u, v) := -\int_\Omega \nabla V \cdot \nabla u v \psi_\infty \mathrm{d}x + \int_\Omega \nabla u \cdot \nabla v \psi_\infty \mathrm{d}x + \int_\Omega \nabla u \nabla \psi_\infty v \mathrm{d}x$$

is the bilinear form associated to the weak formulation of the Kolmogorov operator. Proceeding as in section 4, we obtain the generalized eigenvalue problem:

$$\begin{aligned} &\text{Find } \lambda \in \mathbb{R}, \boldsymbol{u} \in \mathbb{R}^N : \\ &\qquad Au = \lambda B u \end{aligned}$$

where $A, B \in \mathbb{R}^{n \times n}$, $A_{i,j} = a(\phi_j, \phi_i)$, $B_{i,j} = \int_\Omega \phi_j \phi_i \psi_\infty \mathrm{d}x$.

We wish as well to solve numerically the eigenvalue problems associated to the perturbed Kolmogorov and Fokker-Planck operators $L_{\boldsymbol{p}}$ and $L_{\boldsymbol{p}}^*$. In this case the situation is more involved. As mentioned in the previous sections, the operators are not self-adjoint, therefore eigenvalues (and eigenvectors) may be complex valued quantities. It must however be remarked that the two operators still share the same eigenvalues, in the sense that $\lambda \in \mathbb{C}$ is an eigenvalue of L_p if and only if it is an eigenvalue of L_p^*.

6 Numerical experiments

In the present section we present implementation and result of the numerical experiments mentioned in the previous sections. In particular numerical solvers for the following problems are implemented:

- Stationary Fokker-Planck problem;
- eigenvalue problem for Kolmogorov and Fokker-Planck operators;
- unsteady Fokker-Planck problem.

For each of them, we considered three test cases corresponding to the three potentials listed in the first section:

- Quadratic potential:
$$V(\boldsymbol{x}) := \frac{1}{2}\boldsymbol{x}^T S\boldsymbol{x} \text{ with } S = \begin{pmatrix} 1 & 0 \\ 0 & \frac{1}{2} \end{pmatrix}; \tag{36}$$
- Two-wells potential:
$$V(\boldsymbol{x}) := \frac{1}{4}(x^2-1)^2 + \frac{1}{2}y^2; \tag{37}$$
- Three-wells potential:
$$V(\boldsymbol{x}) := 3e^{-x^2-\left(y-\frac{1}{3}\right)^2} - 3e^{-x^2-\left(y-\frac{5}{3}\right)^2} - 5e^{-(x-1)^2-y^2} - 5e^{-(x+1)^2-y^2} + \frac{1}{5}x^4 + \frac{1}{5}\left(y-\frac{1}{3}\right)^4. \tag{38}$$

All numerical solvers are implemented in the programming language `FreeFem++` (see [12]).

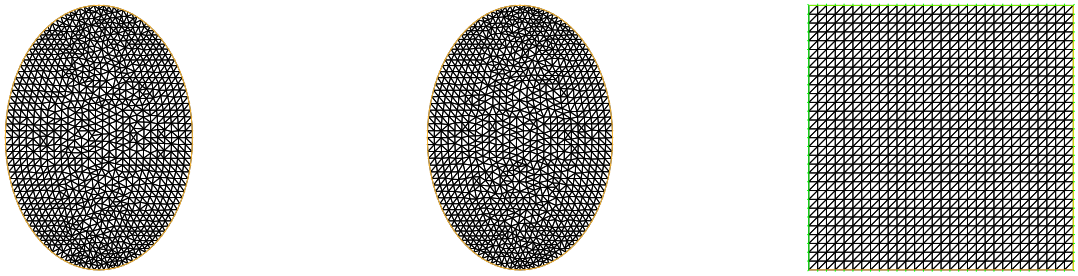
6.1 Implementation

6.1.1 Stationary Fokker-Planck problem

The finite elements problem derived in section (4) is implemented using $P1$ finite elements on triangular meshes on the domains described above and assigned to each potential. We recall that the aim of this numerical experiment is to asses if an approximation of the invariant distribution can be obtained with a finite elements solver. Moreover, we wish to assess what is the effect on the accuracy of the solution of the multiple approximations we have introduced (in particular, the truncation of the domain, the use of a Lagrange multiplier to impose the integral of the solution and the numerical error). The linear system arising from the finite elements approximation is solved with the library `UMFPACK` ([13]), which is suited for the solution of non-symmetric sparse linear systems. The computational domain is chosen in such a way that it its boundary corresponds (or at least is close to) an isoline of the analytical solution on ψ_∞. In particular we choose

- for the quadratic potential: an ellipse with semiaxes of length a an b (being a the semiaxis parallel to x, b the one parallel to y) such that their ration is $\frac{a}{b} = \sqrt{\frac{S_{2,2}}{S_{1,1}}}$ where $S_{i,j}$ are the entries of the matrix S appearing in (36);
- for the two-wells potential: an ellipse with semiaxes of length a an b (as above) such that their ration is $\frac{a}{b} = 0.7$ (this values is determined empirically);
- for the three-wells potential: a square with antipodal vertices $(\frac{1}{2}+a, a)$, $(\frac{1}{2}-a, -a)$.

See figure **??** for plots of meshes generated starting from these domains. Observe that meshes for the quadratic and tow-wells potential lead to a non-conformal finite elements problem.
Each of the following numerical experiments as well are carried out for each test potential on the assigned domain.



(a) Mesh for quadratic potential (b) Mesh for two-wells potential (c) Mesh for three-wells potential

Figure 2: Meshes used to carry out numerical experiments using quadratic (a), 2-wells (b) and 3-wells (c) potential.

6.1.2 Eigenvalue problems

The finite elements approximation of the generalized eigenvalue problems derived in section 5 are implemented. We remind that the aim of this numerical experiment is primarily to determine what is the dependence of the spectral gap on the perturbation and in particular to find a perturbation that leads to the maximum spectral gap (where by spectral gap we always refer to the minimum real part of the non-zero eigenvalue of $L^{\boldsymbol{p}}$). Moreover, these numerical experiments will allow us to verify numerically other properties of $L_{\boldsymbol{p}}$ and $L^*_{\boldsymbol{p}}$ such as the fact that they have the same eigenvalues. The generalized eigenvalue problem that is derived through the finite elements approximation is solved using the library `ARPACK` ([14]). The algorithm implemented in this library is the Arnoldi method for eigenvalue research and is suited to determine few eigenvalues from sparse generalized eigenvalue problems. The algorithm allows for the choice of a "shift", i.e. a complex number in a neighbourhood of which a given number of eigenvalue should be researched. In the present case, we always compute the 10 eigenvalues and choose a shift equal to zero, since our aim is to determine the eigenvalue with smallest non-zero real part.

6.1.3 Unsteady Fokker-Planck problem

The implementation is largely inspired by [15]. The unsteady Fokker-Planck problem is approximated by P1 finite elements in space and with an explicit Euler scheme in time. Being the time advancing scheme conditionally stable, care must be taken to use a sufficiently small time-step. As the solution of the time-dependent Fokker-Planck problem is supposed to be a PDF at every time, at each timestep we normalize it before advancing. The initial condition as well is a PDF (we choose a second order polynomial that is zero on the boundary to match the boundary conditions). As for the stationary problem, the linear system arising at every timestep is solved with `UMFPACK`. We remind that the aim of this numerical experiment is to determine the rate of convergence of the solution (as a function of time) to the stationary state. To do so, we measure the distance with the L^2 norm: $\|\psi_t - \bar{\psi}_\infty\|_{L^2(\Omega,\psi_\infty^{-1})}$, where $\bar{\psi}_\infty$ is the numerical solution computed at a final time.

6.2 Results

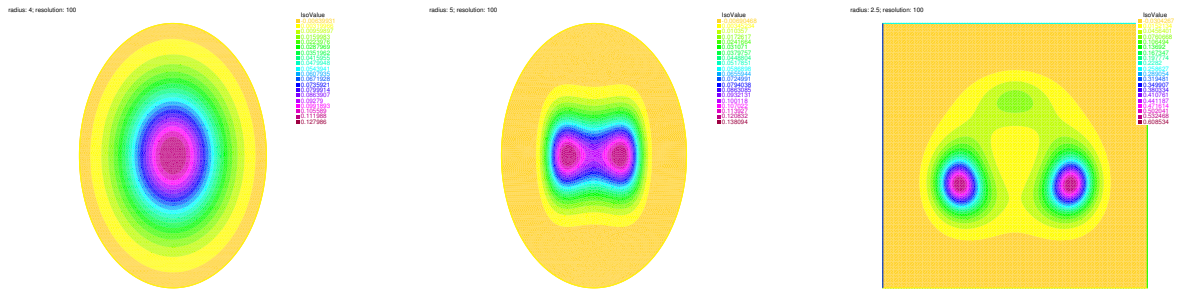
6.2.1 Stationary Fokker-Planck problem

Results for the three potentials can be seen in figure 3. The numerical solutions are visually similar to the exact solutions of the problem on $\mathbb{R}^2$.

To evaluate the error due to the domain truncation we compute the "modelling" error $\|\psi - \psi_m\|_{L^2(\mathbb{R}^d)}$ (as in (32)) by splitting it in "internal" and "external" errors:

$$\|\psi - \psi_m\|_{L^2(\mathbb{R})} = \|\psi - \psi_m\|_{L^2(\Omega)} + \|\psi\|_{L^2(\mathbb{R}^d\setminus\Omega)} \tag{39}$$

(we supposed the model solution is extended with the identically zero function on $\mathbb{R}^d\setminus\Omega$) and computing the two terms separately. We use the square potential with a square domain with antipodal vertices



(a) Solution with quadratic potential (b) Solution with two-wells potential (c) Solution with three-wells potential

Figure 3: Numerical solutions of the stationary Fokker-Planck equation for the quadratic (a), 2-wells (b) and 3-wells (c) potentials.

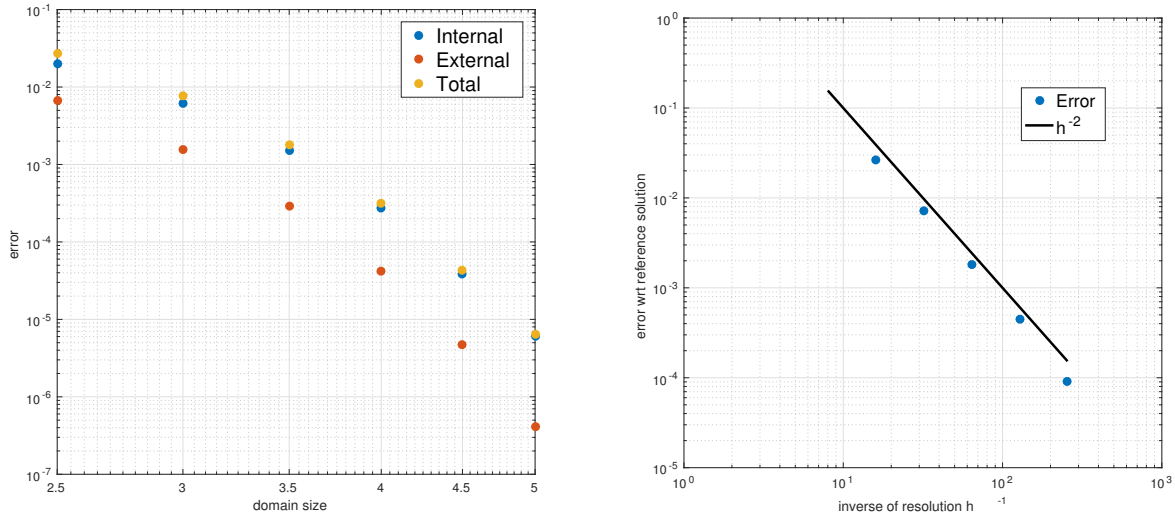


Figure 4: Modelling error for various domain dimensions as a sum of internal and external error (a) and numerical error for various mesh spacings (b). Both are computed using the quadratic potential on a square structured mesh with antipodal vertices $(-a, a)$ and (a, a).

$(-a, a)$ and $a, a)$ and let $a > 0$ vary. The first term of (39) is computed by first computing an approximation of ψ_m on a fine mesh, then the exact solution ψ is restricted to Ω and projected in the $P1$ finite elements space of the approximation of ψ_m, finally the L^2 error is computed numerically if `FreeFem++`. The "external error" is computed with a high order quadrature formula of `Matlab`. Results can be seen in figure 4(a). It can be observed that the external error is always around one order of magnitude smaller than the internal one. Moreover, both (and therefore the total error) decrease exponentially as the size of the mesh is increased.

As for the numerical error, $\|\psi_m - \psi^h\|_{L^2(\Omega)}$, we find an $\mathcal{O}(h^2)$ dependence with the mesh spacing h as expected for P1 finite elements [10] (see figure 4(b)).

Finally, we evaluate the effect of the imposition of the integral constraint by checking the value of the numerically computed lagrange multiplier λ for different mesh sizes:

radius	3	4	5	6	7	8
λ	$2.70 \cdot 10^{-2}$	$3.39 \cdot 10^{-3}$	$3.28 \cdot 10^{-4}$	$2.04 \cdot 10^{-5}$	$7.60 \cdot 10^{-7}$	$1.65 \cdot 10^{-8}$

These results were obtained using the quadratic potential on the elliptic domain we assigned to it in the previous section. The mesh spacing was maintained constant in all experiments (despite the fact that the domain grows) by imposing on the domain's boundary a number of vertices proportional to the perimeter of the domain.
It can be seen that the Lagrange multiplier approaches zero exponentially fast as the domain grows.

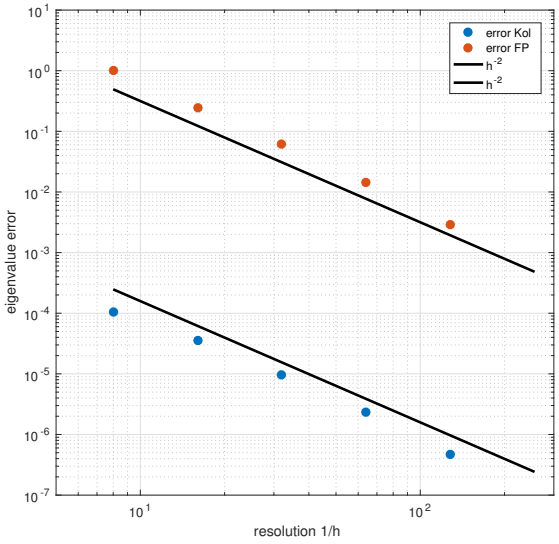


Figure 5: Convergence study for finite elements approximations of the generalized eigenvalue problem for Kolmogorov and Fokker-Planck operators. the quadratic potential is used on the elliptic domain

This result suggests that the the integral constraint becomes less and less restrictive as the size of the domain grows.

6.2.2 Eigenvalue problems

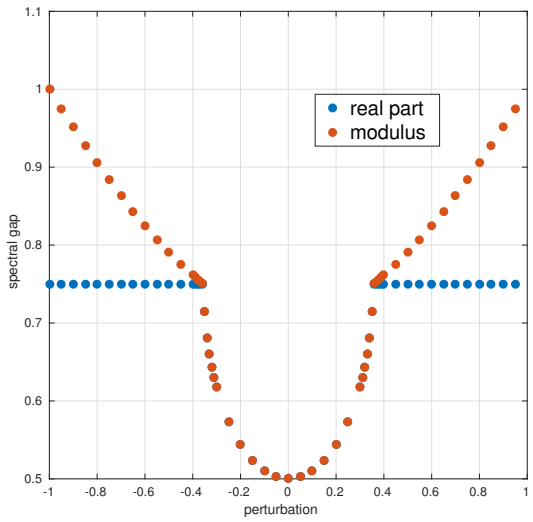
Firstly, we verify some properties regarding the spectrum of the unperturbed Kolmogorov and Fokker-Planck operators $L_{\mathbf{0}}$, $L_{\mathbf{0}}^*$, in particular that their eigenvalues are real, negative and identical. This is done using the quadratic potential (36) on the elliptic domain. The Kolmogorov operator seems to approximate the eigenvalues more accurately.
To confirm this observation we carry out a convergence study of the spectral gap with respect to the mesh spacing h for the two problems. The error can be computed exactly since the exact spectral gap is known to be $\lambda = 0.5$ (see section 3.1). Results are presented in figure 5 where it can be seen that both numerical approximations are $\mathcal{O}(h^2)$ (as expected for P1 finite elements [11]) but the error using the Kolmogorov operator is smaller. The same behaviour can be observed for the other potentials (we do not present the results here).

From this observation, since Kolmogorov and Fokker-Planck operators have the same eigenvalues, from now on we will compute the spectral gap using the Kolmogorov operator.

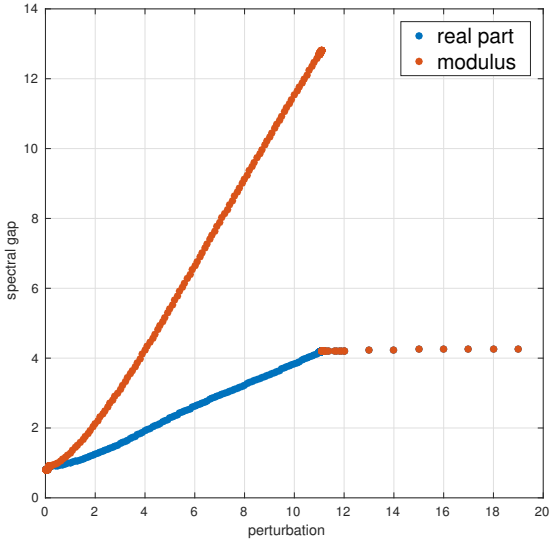
We present a study on the dependence of the spectral gap on the perturbation. We recall that the perturbation is in the form $J\nabla V$ where $J = \begin{pmatrix} a & 0 \\ 0 & -a \end{pmatrix}$ with $a \in \mathbb{R}$. Therefore the set of perturbations is a one-dimensional linear space. We can therefore produce some graphs of the spectral gap as a function of the parameter a. Examples of these graphs can be seen in figure 6 (where we plot the *opposite* of the smallest non-zero real part of the eigenvalue as well as its modulus for the three test potentials on the same domain listed above.

Some general observations can be made for all three potentials: both spectral gap and modulus are convex and symmetric (this is the reason why in some cases we plot only for $a > 0$ or only with few negative values) with respect to the independent variable a. Moreover, as expected, the minimum is always in $a = 0$.

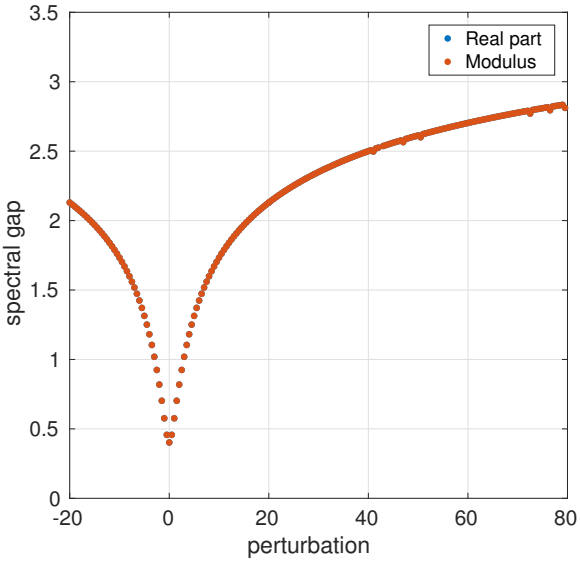
As for the quadratic potential, the theoretical results presented in section 3.1 seems to be realized. In particular, $\lambda(0) = 0.5$ and the maximum spectral gap is $\lambda_{opt} = 0.75 = \frac{Tr(S)}{2}$. The maximum non-zero real part eigenvalue is real for $|a| \leq 0.35$, while for the other values it has non-zero imaginary part, the real part is constant and the modulus increases indefinitely as a linear function of slope



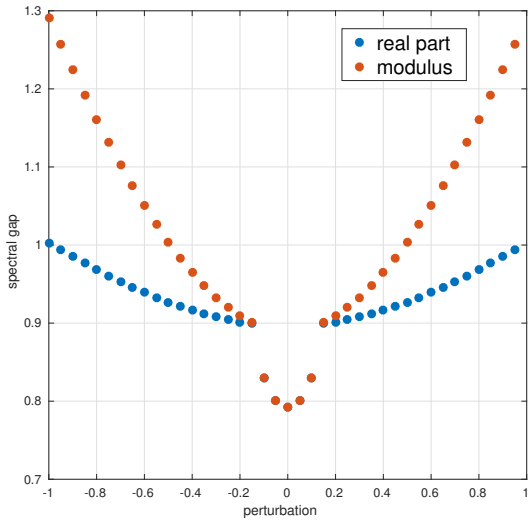
(a) Spectral gap for quadratic potential



(b) Spectral gap for two-wells potential



(c) Spectral gap for three-wells potential



(d) Spectral gap for quadratic potential (close-up)

Figure 6: Spectral gap as a function of the perturbation parameter a for the quadratic (a), two-wells (b) and three-wells (c) potentials. In (d) we can see a close-up of the values for the two-wells potential around $a = 0$.

$\cong 0.67$ as $a \to +\infty$ (and symmetrically for $a \to -\infty$). Both real part and modulus, as a function of a, look differentiable for all a except $|a| \cong 0.355$, where they seem only Lipschitz.

As for the two-wells potential, the (minimum) spectral gap is $\lambda(0) \cong 0.79$, it seems not to have a maximum, while the supremum is about 4.2 (attained in the limit $|a| \to +\infty$). The smallest non-zero real part eigenvalue is real for $|a| < 0.13$ and $|a| > 11.1$, while it has non-zero imaginary part for other values. The function $\mathrm{Re}(\lambda)(a)$ appears differentiable everywhere except for the values $|a| = 0.13$, where it seems only Lipschitz, and $|a| = 11.1$, where, surprisingly, it seems discontinuous.

As for the three-wells potential, all eigenvalues are real. The minimum is $\lambda(0) \cong 0.4$, the maximum seems not to be attained and the supremum is about 3. The function seems differentiable everywhere.

6.2.3 Unsteady Fokker-Planck problem

The time-dependent Fokker-Planck equation seems to converge to the numerical the solution of the stationary problem, which in turn has been proven (at least numerically) to converge to the analytical solution of the problem as the mesh is refined and the domain enlarged.

As for the rate of convergence to the stationary state, results can be seen in figure 7. The rate of convergence to the stationary solution is, in all three cases, described fairly accurately by the spectral gap of the corresponding Kolmogorov operator. Moreover, it can be clearly seen that the solution reaches the stationary state faster if a perturbation is applied. The reduction of this time is quite dramatic when the optimal perturbation is applied. we recall the distance from the stationary state is measured as $\|\psi_t - \psi_\infty\|_{L^2(\psi_\infty^{-1})}$. As for the case of the quadratic potential, the following table lists times to convergence (within a tolerance) for various perturbations:

perturbation a	0	0.1	0.35	1
time to convergence	42.65	42.18	32.69	30.68

It can be seen that for the perturbation given by $a = 1$ (the only one for which the minimum non-zero real part eigenvalue has a non-zero imaginary part) the distance to the stationary solution has an oscillatory behaviour. This does not happen for the similar cases for the two-wells potential.

As for the case of the two-wells potential, the following table lists time to convergence (within a tolerance) for various perturbations:

perturbation a	0	0.14	8	20
time to convergence	30.08	29.10	8.08	7.66

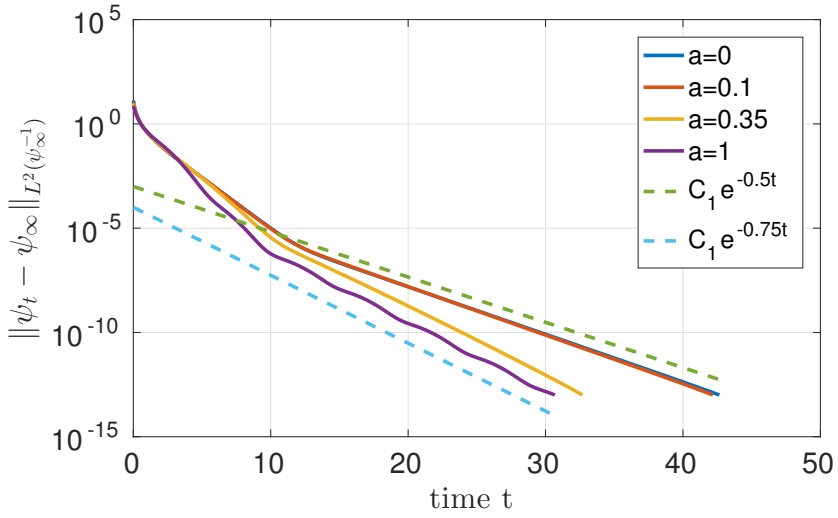
As for the case of the three-wells potential, the following table lists time to convergence (within a tolerance) for various perturbations:

perturbation a	0	10	50	80
time to convergence	55.8	16.44	10.6	9.72

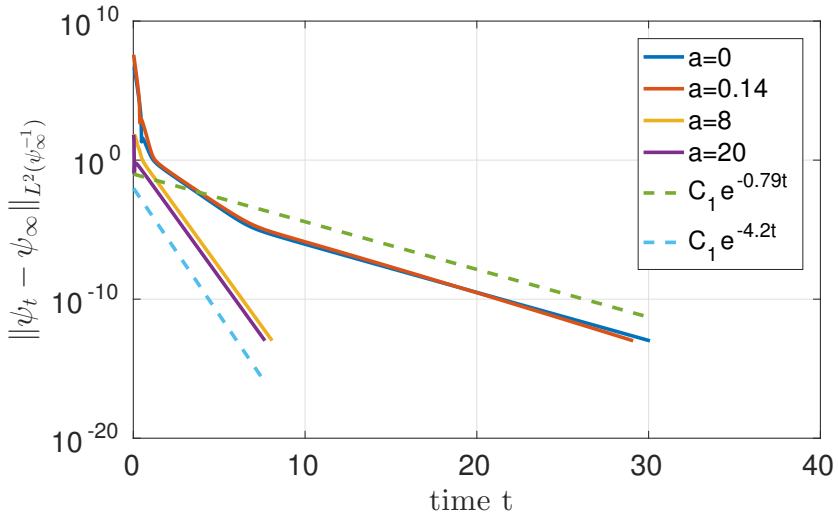
7 Conclusions

As for the stationary Fokker-Plank problem, we have observed that the approximations we have introduced (in particular, the domain truncation, the introduction of a Lagrange multiplier and the numerical approximation) cause an error on the solution. However, we have also verified (at least numerically for the quadratic potential) that this error vanishes as the domain is enlarged and the mesh spacing is reduced. This result leads us to think that the same holds for the other problems presented here (eigenvalue problems and unsteady Fokker-Plank problem). Therefore, the numerical solutions presented above can also be considered reliable.

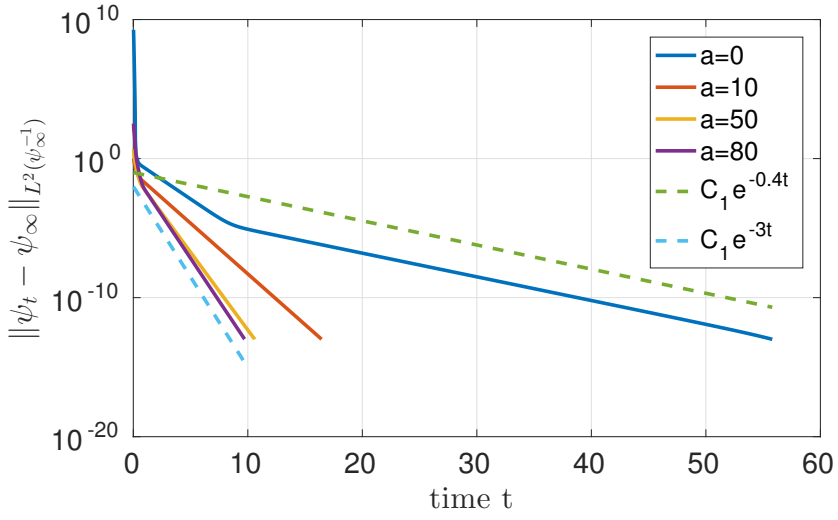
We have verified that there is actually a tight link between the spectral gap of the perturbed Kolmogorov operator and the rate of convergence to a stationary state of the associate Fokker-Plank equation. Indeed, it is proven (see [2] or section 3.1 where we report the results) that in the case of the (perturbed or not) quadratic potential the spectral gap (that we defined as the real part of the minimum non-zero real part eigenvalue) describes the rate of convergence to equilibrium of the dynamics. In addition, we verified this fact numerically with high accuracy. As for the other test cases



(a)



(b)



(c)

Figure 7: Distance from the stationary state of the solution of the time-dependent Fokker-Planck equation as a function of time. Results for the quadratic (a), two-wells (b) and three-wells (c) potentials are presented.

(two and three-wells potentials), in which no analytical result is known, we also achieved good results, as also in these cases the spectral gap predicts accurately the rate of convergence to equilibrium of the related dynamics. In particular, the cases of the perturbed dynamics for these two potentials deserve special attention. As a matter of facts, in section 2.2.1 (and 3) we assume that the spectral gap is a good estimator of the rate of convergence to equilibrium of the dynamics (following [8]), despite the fact that no analytical result is available in these cases. What we find from numerical experiments is that the link between the two quantities is as tight as in the case of the quadratic potential. Moreover, the fact that in the case of the quadratic potential theoretical and numerical results coincide with a high accuracy leads us once more to think that these numerical results are not biased by systematic "modelling" errors (e.g. domain truncation).

The three test cases we have considered suggest that the optimal spectral gap is achieved for "large" perturbations (e.g. for $|a| > 0.35$ for the quadratic potential) or even only asymptotically (for two and three-wells potentials). As a matter of facts, in all three test cases the spectral gap as a function of the "magnitude"of the perturbation is a convex function. Since maximizing the spectral gap (and therefore the rate of convergence of the dynamics) is what we aim at, it would be important to find out how general these properties are.

Finally, we have observed that the choice of an optimal perturbation can lead to a dramatic acceleration of the diffusion. For instance, choosing a (close to) optimal perturbation for the thee-wells potential leads the dynamics to convergence in less than one fifth of the time taken by the unperturbed diffusion. These results may lead to interesting applications, for instance acceleration of the Markov Chain Monte Carlo algorithm.

8 Outlook

The work presented in this report can be extended and improved in many ways.

As for the theory, it should be possible to prove analytically that enlarging the domain of the problem (for instance stationary Fokker-Plank) the solution converges to the solution on $\mathbb{R}^d$. Other important developments may come from the study of the spectral properties of the Kolmogorov operator perturbed with a divergence-free function (e.g. the asymptotic behaviour as done in [17]). As mentioned above, the test cases considered in this context suggest that relatively "large" perturbations are needed to achieve an optimal spectral gap (or, as for instance for the three and two-wells potentials, the optimum can be reached only asymptotically). A final theoretical possible development (possibly the most relevant of this work) is to prove that the rate of convergence to equilibrium of the dynamics is given by the spectral gap also for the perturbed (therefore non-self adjoint) dynamics. This is in facts the result we obtain in numerical experiments. However, it may be wise to evaluate other test cases (i.e. potentials) first.

As for the numerical methods, other, more effective boundary conditions may be employed in order to better approximate the unbounded problem on a bounded domain and therefore avoid the necessity to enlarge the domain (and therefore the complexity) of the numerical problem in order to reduce "modelling" errors. Regarding the methods we have used here to approximate the stationary Fokker-Plank equation, it may be desirable to also enforce a positivity constraint on the solution (despite the fact that the imposition of the integral of the solution alone leads to strictly positive solutions). Another possible development may come from the use of spectral methods rather than the finite elements methods as we have done here. Indeed, employing spectral methods removes the problem of approximating the domain with a bounded sets. Finally, multidimensional perturbations can be considered to get closer to applications in chemistry, statistical physics and statistics.

Finally several applications of the concepts presented here can be mentioned. The acceleration of the Markov Chain Monte Carlo algorithm is an example. Another aspect of the problem that can be tackled is how to *efficiently* find an optimal perturbation. Clearly, this can be viewed as an optimization problem. Due to the regularity of the spectral gap as a function of the perturbation (even for the simple examples considered here it presents discontinuities and point of non-differentiability) gradient based methods are unlikely to be an effective way to find a solution. Therefore, one is likely to have to rely on global gradient-free optimization methods such as the golden-search method (for

one-dimensional perturbations only) or simulated annealing (see [16]).

References

[1] Markowich, Peter A., and Cédric Villani. "On the trend to equilibrium for the Fokker-Planck equation: an interplay between physics and functional analysis." Mat. Contemp 19 (2000): 1-29.

[2] Pavliotis, Grigorios A. Stochastic processes and applications. Springer-Verlag New York, 2016.

[3] Lelièvre, Tony, Francis Nier, and Grigorios A. Pavliotis. "Optimal non-reversible linear drift for the convergence to equilibrium of a diffusion." Journal of Statistical Physics 152, no. 2 (2013): 237-274.

[4] Yosida, Kosaku. "Functional analysis." Grundlehren Math. Wiss. 123 (1980).

[5] Cédric Villani, C. Hypocoercivity. No. 949-951. American Mathematical Soc., 2009.

[6] Brezis, Haim. Functional analysis, Sobolev spaces and partial differential equations. Springer Science & Business Media, 2010.

[7] Metzner, Philipp, Christof Schütte, and Eric Vanden-Eijnden. "Illustration of transition path theory on a collection of simple examples." The Journal of chemical physics 125, no. 8 (2006): 084110.

[8] Hwang, Chii-Ruey, Shu-Yin Hwang-Ma, and Shuenn-Jyi Sheu. "Accelerating diffusions." The Annals of Applied Probability 15, no. 2 (2005): 1433-1444.

[9] Payne, Lawrence E., and Hans F. Weinberger. "An optimal Poincaré inequality for convex domains." Archive for Rational Mechanics and Analysis 5, no. 1 (1960): 286-292.

[10] Ern, Alexandre, and Jean-Luc Guermond. "Theory and practice of finite elements. Number 159 in Applied Mathematical Sciences." (2004).

[11] Larsson, Stig, and Vidar Thomée. Partial differential equations with numerical methods. Vol. 45. Springer Science & Business Media, 2008.

[12] Hecht, Frédéric, Olivier Pironneau, A. Le Hyaric, and K. Ohtsuka. "FreeFem++." Numerical Mathematics and Scientific Computation. Laboratoire JL Lions, Université Pierre et Marie Curie, http://www. freefem. org/ff+ 3 (2007).

[13] Davis, Timothy A. "Umfpack version 4.1 user guide." Department of Computer and Information Science and Engineering, University of Florida (2003).

[14] Lehoucq, Richard B., Danny C. Sorensen, and Chao Yang. ARPACK users' guide: solution of large-scale eigenvalue problems with implicitly restarted Arnoldi methods. Society for Industrial and Applied Mathematics, 1998.

[15] De Vuyst, Florian. "Numerical modeling of transport problems using freefem++ software–with examples in biology, CFD, traffic flow and energy transfer." (2013): 162.

[16] Press, William H. Numerical recipes 3rd edition: The art of scientific computing. Cambridge university press, 2007.

[17] Franke, B., C-R. Hwang, H-M. Pai, and S-J. Sheu. "The behavior of the spectral gap under growing drift." Transactions of the American Mathematical Society 362, no. 3 (2010): 1325-1350.