UPPSALA DISSERTATIONS IN MATHEMATICS

140

Behind the training dynamics of neural networks

Analysis of Fokker-Planck equations and the path to metastability

Mingyi Hou

Department of Mathematics
Uppsala University
UPPSALA 2025

Dissertation presented at Uppsala University to be publicly examined in Polhemssalen, Ångströmlaboratoriet, Lägerhyddsvägen 1, Uppsala, Wednesday, 21 May 2025 at 13:15 for the degree of Doctor of Philosophy. The examination will be conducted in English. Faculty examiner: Professor Sergio Polidoro (Università di Modena e Reggio Emilia).

Abstract
Hou, M. 2025. Behind the training dynamics of neural networks. Analysis of Fokker-Planck equations and the path to metastability. *Uppsala Dissertations in Mathematics* 140. 53 pp. Uppsala: Uppsala University. ISBN 978-91-506-3102-9.

This thesis develops the theoretic mathematical foundations of Fokker-Planck-Kolmogorov operators within a potential theory framework to study metastability in stochastic processes. These operators generate Langevin dynamics, which model the motion of a particle influenced by thermal noise and an external potential field. A key focus of this study is the metastable transition time, the time required for a particle to move between local minima, which plays a fundamental role in applications such as chemical reactions, quantum tunneling, and neural network training.

The main contributions of this thesis include a comprehensive analysis of weak solutions to (possibly degenerate) Fokker-Planck-Kolmogorov equations. Specifically, we develop a Galerkin method for solving the Cauchy problem in a periodic setting, establish the existence and uniqueness of weak solutions for the stationary Kolmogorov operator in bounded product domains, and introduce Perron's solutions in general bounded domains. Additionally, we prove global boundedness results for time-dependent solutions in bounded time cylinders. Finally, we develop a variational framework for potential theory for stationary Kolmogorov operators. These results provide new mathematical tools for studying metastability and deepen our understanding of stochastic systems while also opening new research directions in Kolmogorov equations, including higher boundary regularity and obstacle problems.

Keywords: Fokker-Planck-Kolmogorov operators, Metastability, Weak solutions, Potential theory, Variational formulations, Eyring-Kramers formula, Dirichlet problems, Galerkin method

Mingyi Hou, Department of Mathematics, Box 480, Uppsala University, SE-75106 Uppsala, Sweden.

© Mingyi Hou 2025

ISSN 1401-2049
ISBN 978-91-506-3102-9
URN urn:nbn:se:uu:diva-553381 (http://urn.kb.se/resolve?urn=urn:nbn:se:uu:diva-553381)

List of papers

This thesis is based on the following papers, which are referred to in the text by their Roman numerals.

I B. Avelin, M. Hou, and K. Nyström. A Galerkin type method for kinetic Fokker-Planck equations based on Hermite expansions. *Kinet. Relat. Models*, 17 (4), 2024, pp. 634–658. DOI: 10.3934/krm.2023035.

II B. Avelin, and M. Hou. Weak and Perron Solutions for Stationary Kramers-Fokker-Planck Equations in Bounded Domains. *arXiv preprint*, 2024, arXiv:2405.04070.

III M. Hou. Boundedness of weak solutions to degenerate Kolmogorov equations of hypoelliptic type in bounded domains. *arXiv preprint*, 2024, arXiv:2407.00800.

IV M. Hou. Variational formulation and capacity estimates for non-self-adjoint Fokker-Planck operators in divergence form. *arXiv preprint*, 2025, arXiv:2502.12036.

Reprints were made with permission from the publishers.

Contents

1 Introduction 7
1.1 Behind the Training Process of Neural Networks 7
1.2 Chemical Reactions, Arrenhnuis Law, and Eyring-Kramers Law 10
1.3 Potential Theory in the Reversible Case 11
1.4 Aim of the Thesis: Towards the Underdamped Case 14

2 Fundaments of Fokker-Planck-Kolmogorov Operators 17
2.1 The Harnack Inequality and Interior Regularity of Solutions 18
2.1.1 Definition of Weak Solutions 18
2.1.2 The Harnack Inequality 19
2.2 Well-posedness and Boundary Regularity of Solutions 20
2.2.1 Existence of Perron's Solution and Boundary Regularity in Bounded Domains 21
2.2.2 Regular Points and Barriers 22
2.2.3 Maximum Principle 22
2.2.4 Dirichlet Problems in Unbounded Domains 23
2.2.5 Higher Regularity up to the Boundary 24
2.3 The Green Function 25
2.4 Definition of Capacity and Variational Formulation 26
2.4.1 Definition of Capacity 27
2.4.2 Symmetry Relations 27
2.4.3 Variational Formulation 28
2.5 Rough and Sharp Capacity Estimates 29
2.5.1 Rough Capacity Estimates 29
2.5.2 Sharp Capacity Estimates 30
2.6 What Has Been Done in This Thesis? 32

3 Summary of Articles 33
3.1 Paper I 33
3.2 Paper II 35
3.3 Paper III 40
3.4 Paper IV 43

4 Summary in Swedish 47

5 Acknowledgements 49

Bibliography 50

1. Introduction

Metastability refers to the behavior of a stochastic process confined in a potential landscape with multiple local minima. The process tends to remain in the vicinity of one minimum for a long time, experiencing small fluctuations due to noise, before making a rare, sudden transition to another minimum.

In this thesis, we build the mathematical foundations for Fokker-Planck-Kolmogorov equations. The goal is to develop a potential-theoretic toolbox to understand metastable transition times for general non-reversible Langevin dynamics.

Metastability has been a popular topic of research since its introduction earlier last century. It has been discovered in various fields, including physics (e.g., quantum mechanics) [53], chemistry (e.g., chemical kinetics) [26], and economics (e.g., asset pricing) [52], and more recently in machine learning (e.g., deep neural networks) [63]. The motivation of this thesis mainly comes from its recent application in machine learning.

1.1 Behind the Training Process of Neural Networks

Experiencing the boom of generative AIs, particularly large language models, the future has become full of surprises and uncertainty. While the application side of deep neural networks is promising, understanding their underlying mechanisms is still extremely challenging. One of the key aspects is understanding how the training process works.

A neural network is a non-linear regression model inspired by the human brain, where input signals pass through layers of interconnected nodes (neurons) to produce an output, as illustrated in Fig. 1.1. The model can be represented as a function $f_\theta(x) : \mathbb{R}^n \to \mathbb{R}^m$ for m, n integers, and where θ is a family of parameters defined in the model. Given an input x, the model generates an output $f_\theta(x)$, for example, the answer one gets after asking a question to an AI model. In particular, for large language models, the number of parameters θ is usually extremely large (on the order of millions or billions), and the model's complexity increases with the number of parameters, making it hard to understand and interpret.

To train a neural network, we usually need to set some objectives, which are known as the loss function. It measures how far away an output generated by the model is from the actual output. The aim of training is to minimize this

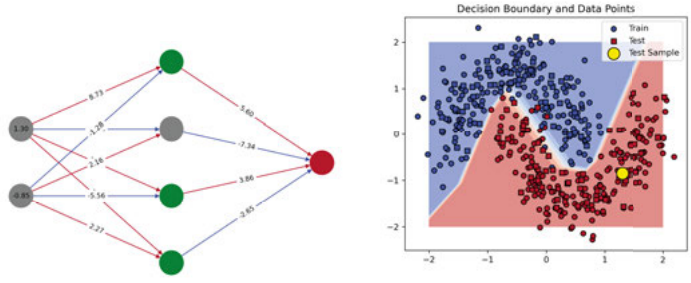

Figure 1.1. A simple neural network with one hidden layer (left) for a classification task, and its decision boundary (right).

loss function. Mathematically, let $\{x_i\}_{i\in I}$ be a set of inputs, $\{y_i\}_{i\in I}$ be the corresponding outputs, the simplest loss function is the mean square error:

$$L(\theta; x, y) := \frac{1}{\#I} \sum_{i\in I} l_i(\theta), \qquad \text{where } l_i(\theta) := (y_i - f_\theta(x_i))^2 ,$$

where $\#I$ is the size of the index set. Indeed, this is not much different from the mean square error for a linear regression model, except in the neural network case L is no longer convex and contains many local minima, see e.g., [45, 61, 49].

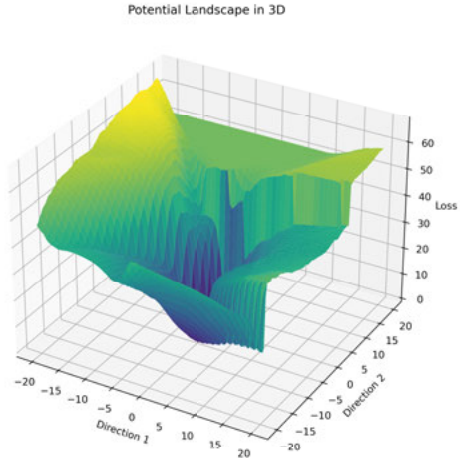

Figure 1.2. Illustration of the loss function for the above neural network.

When L is uniformly convex, the gradient descent method is guaranteed to find a global minimum. However, for neural networks, L is non-convex, mean-

ing that the standard gradient descent may get stuck in a local minimum. Moreover, due to the large number of parameters, performing gradient descent over the entire training dataset is impractical. Instead, we typically use stochastic gradient descent (SGD), which samples a subset of training examples at each iteration, known as a mini-batch $B_k \subsetneq I$, to update parameters. The standard SGD update rule is given by

$$\theta_{k+1} = \theta_k - \eta \frac{1}{\#B_k} \sum_{i \in B_k} \nabla_\theta l_i(\theta),$$

where η is the learning rate. According to [46], this discrete update scheme can be well approximated by the *overdamped Langevin process*:

$$\mathrm{d}\Theta_t = -\nabla L(\Theta_t)\mathrm{d}t + \sqrt{\eta}\mathrm{d}B_t,$$

where B_t is the d-dimensional Brownian motion, modelling the randomness introduced by mini-batch sampling. For simplicity, we assume that the mini-batch noise follows an isotropic Gaussian distribution, though in practice, it may be degenerate in some directions. Additionally, artificial noise can be added to the batch, a technique known as stochastic gradient Langevin dynamics.

The term "overdamped" comes from statistical physics, as the friction is large enough to dominate over the inertia of the particles. In contrast, there is another Langevin process known as the underdamped Langevin process, which will be introduced later in Section 1.4 for stochastic gradient descent with momentum, where the inertia is no longer negligible.

Fig. 1.3 is a demonstration of the metastability in a toy model where the loss function is a double well potential. It partially explains why SGD works

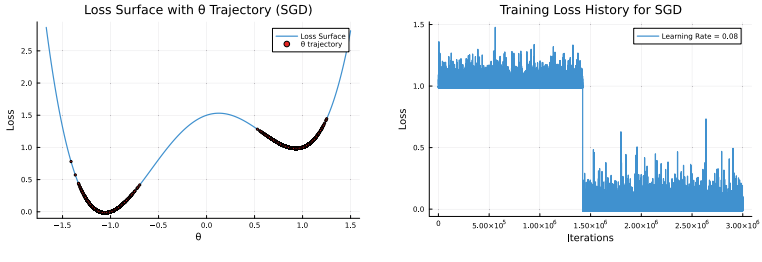

Figure 1.3. Metastability for overdamped Langevin dynamics with a double well potential. The transition happens at around 1.5×10^6 iterations ($t \approx 1.2 \times 10^5$).

when optimizing a loss function with multiple minima. Additionally, a better understanding of the metastability could potentially provide insights on when to stop the training.

1.2 Chemical Reactions, Arrenhnuis Law, and Eyring-Kramers Law

Prior to the advent of machine learning, metastable processes had been discovered in many physical systems. In particular, in the 1930s, transition state theory was proposed by H. Eyring [26] to describe how reactions occur at the molecular level. In this theory, reactants residing in one energy well must pass through a high energy barrier before forming products. The transition state, also called the activated complex, is located at the saddle point, and the energy required to reach this transition state is called the activation energy.

Let now $X \in \mathbb{R}^d$ denote the position, $U(X) : \mathbb{R}^d \to \mathbb{R}$ denote a double well potential that is a Morse function (a smooth function where all critical points have non-degenerate Hessians) with precisely two local minima, $\boldsymbol{m}_0$ and $\boldsymbol{m}_1$ (corresponding to the states of reactants and products), and, without loss of generality, a unique saddle point 0 (corresponding to the activation energy). Moreover, the potential blows up at infinity sufficiently fast. The transition state theory is described by the overdamped Langevin equation

$$\mathrm{d}X_t = -\nabla U(X_t)\mathrm{d}t + \sqrt{2\varepsilon}\mathrm{d}B_t, \tag{1.1}$$

where B_t is the d-dimensional Brownian motion, and ε relates to the thermal noise. The factor $\sqrt{2}$ is included for mathematical convenience. The average time for reactants to become products is mathematically known as the *mean transition time*:

$$\mathbb{E}^{\boldsymbol{m}_0}[\tau_{\boldsymbol{m}_1}], \quad \text{where } \tau_{\boldsymbol{m}_1} = \tau_{B_\delta(\boldsymbol{m}_1)} := \inf\{t > 0 : X_t \in B_\delta(\boldsymbol{m}_1)\}, \tag{1.2}$$

for δ a small positive constant. Here, $\tau_{B_\delta(\boldsymbol{m}_1)}$ is a stopping time, and the starting point of the process is $X_0 = \boldsymbol{m}_0$. In the context of chemical reactions, a relation for the mean transition time, which was physically observed and justified by Arrenhnuis [6] in 1889, states

$$\mathbb{E}^{\boldsymbol{m}_0}[\tau_{\boldsymbol{m}_1}] \approx \mathrm{e}^{(U(0)-U(\boldsymbol{m}_0))/\varepsilon},$$

where $a \approx b$ for $a, b > 0$ means that there exists a constant $c > 0$ such that $c^{-1}b \le a \le cb$. On the other hand, the Eyring-Kramers law [26, 37] improves the prefactor c, e.g., in the case $d = 1$,

$$\mathbb{E}^{\boldsymbol{m}_0}[\tau_{\boldsymbol{m}_1}] \approx \frac{2\pi}{\sqrt{U''(\boldsymbol{m}_0)|U''(0)|}}\mathrm{e}^{(U(0)-U(\boldsymbol{m}_0))/\varepsilon}.$$

This reveals the dependence on the geometry of the potential, namely, the curvatures at the starting minimum $\boldsymbol{m}_0$ and at the saddle point 0. However, the question remained as to what sense of "$\approx$" the Eyring-Kramers law is true. This remained an open problem until 2004, when Bovier et al. [15] gave a full characterization of the prefactor c and provided a solid proof based on potential

theory, which states

$$\mathbb{E}^{\boldsymbol{m}_0}[\tau_{\boldsymbol{m}_1}] = [1 + o_\varepsilon(1)]\,\frac{2\pi}{\lambda_1}\sqrt{\frac{|\det(\nabla^2 U(0))|}{\det(\nabla^2 U(\boldsymbol{m}_0))}}\mathrm{e}^{(U(0)-U(\boldsymbol{m}_0))/\varepsilon}, \tag{1.3}$$

where $-\lambda_1$ is the unique negative eigenvalue of $\nabla^2 U(0)$, and $o_\varepsilon(1)$ denotes a function $r(\varepsilon)$ such that $\lim_{\varepsilon\to 0} r(\varepsilon) = 0$. Indeed, the $o_\varepsilon(1)$ term is of the order $\varepsilon^{1/2}|\log\varepsilon|^{3/2}$.

1.3 Potential Theory in the Reversible Case

A beautiful story lies behind the potential theory proof of the Eyring-Kramers formula (1.3), where stochastic processes and partial differential equations interact in a beautiful way. We outline and highlight the proof strategy from [15] here.

Recall the overdamped Langevin process X_t given by (1.1). Its infinitesimal generator is known as the Fokker-Planck operator:

$$\mathscr{L} f = \varepsilon\Delta f - \nabla U\cdot\nabla f. \tag{1.4}$$

The process (1.1) is also *reversible*, meaning that the transition density kernels $p_t(x,y)$ with respect the stationary measure is a symmetric function, namely, $p_t(x,y) = p_t(y,x)$ (more precise definitions can be found in [10]). In other words, the operator $\mathscr{L}$ is self-adjoint with respect to the stationary measure $\exp(-U(x)/\varepsilon)\mathrm{d}x$ (this explicit exponential form is also known as the Gibbs measure).

Landscape Function and Mean Exit Time

By Dynkin's formula (cf. [57]), it follows that for any function f in the domain of $\mathscr{L}$, we have

$$\mathbb{E}^x[f(X_{\tau_{\mathcal{B}}})] = f(x) + \mathbb{E}^x\left[\int_0^{\tau_{\mathcal{B}}} \mathscr{L} f(X_s)\,\mathrm{d}s\right]. \tag{1.5}$$

Now, we consider a function $w(x)$ satisfying the following equation

$$\begin{cases} -\mathscr{L} w = 1, & \text{in } \left(\overline{\mathcal{B}}\right)^{\complement} \\ \quad w = 0, & \text{on } \partial\mathcal{B}, \end{cases} \tag{1.6}$$

where $\mathcal{B} := B_\delta(\boldsymbol{m}_1)$. The function $w(x)$ is also known as the landscape function in the context of localization phenomenon, see e.g., [27, 5]. Using $w = 0$ on $\partial\mathcal{B}$, and w satisfying $\mathscr{L} w = -1$, it yields that

$$w(x) = \mathbb{E}^x[\tau_{\mathcal{B}}]. \tag{1.7}$$

Indeed, the mean transition time is just evaluating $w(x)$ at the starting minimum, i.e. $w(\boldsymbol{m}_0) = \mathbb{E}^{\boldsymbol{m}_0}[\tau_{\boldsymbol{m}_1}]$.

Equilibrium Potential and Capacity
On the other hand, let $\mathcal{A} := B_\delta(\boldsymbol{m}_0)$, we consider a function $h(x) = h_{\mathcal{A},\mathcal{B}}(x)$, known as the equilibrium potential in classical potential theory. This function satisfies the equilibrium problem

$$\begin{cases} -\mathscr{L}h = 0, & \text{in } (\overline{\mathcal{A}\cup\mathcal{B}})^{\complement}, \\ h = 1, & \text{on } \partial\mathcal{A}, \\ h = 0, & \text{on } \partial\mathcal{B}. \end{cases} \tag{1.8}$$

If $\mathscr{L}$ were the Laplace operator, then $h_{\mathcal{A},\mathcal{B}}$ would represent the electric potential in an electrostatic setup, where the two disjoint sets $\mathcal{A}$ and $\mathcal{B}$ act as conductors held at potentials 1 and 0, respectively. Similarly, by Dynkin's formula, we obtain $h_{\mathcal{A},\mathcal{B}}(x) = \mathbb{P}^x[\tau_{\mathcal{A}} < \tau_{\mathcal{B}}]$, which states that the equilibrium potential at x represents the probability that the process X_t, starting from x, reaches $\mathcal{A}$ before $\mathcal{B}$.

In classical electrostatics, it is well known that free charges on a conductor in equilibrium concentrate on its surface, and the capacity of a conductor is the total charge required to raise its voltage to 1. Extending $h_{\mathcal{A},\mathcal{B}}$ to be 1 inside $\mathcal{A}$, we see that $-\mathscr{L}h_{\mathcal{A},\mathcal{B}}$ defines a non-negative measure supported on $\partial\mathcal{A}$, representing the charge distribution. The capacity of the capacitor $(\mathcal{A},\mathcal{B})$ is then given by

$$\operatorname{cap}(\mathcal{A},\mathcal{B}) := \int_{\overline{\mathcal{B}}^{\complement}} -\mathscr{L}h\, \mathrm{e}^{U/\varepsilon}\mathrm{d}x = \int_{\partial\mathcal{A}} \varepsilon\nabla h\cdot\boldsymbol{n}_\Omega\, \mathrm{e}^{U/\varepsilon}\sigma(\mathrm{d}x), \tag{1.9}$$

where $\boldsymbol{n}_\Omega$ is the unit outward normal vector to the domain $\Omega := (\overline{\mathcal{A}\cup\mathcal{B}})^{\complement}$, and $\sigma(\mathrm{d}x)$ denotes the standard $d-1$ dimensional surface measure.

Mean Transition Time in Terms of Equilibrium Potential and Capacity
The estimate of $w(\boldsymbol{m}_0)$ relies on the following simple observation, obtained via integration by parts:

$$\int_{\overline{\mathcal{B}}^{\complement}} h\, \mathrm{e}^{-U/\varepsilon}\mathrm{d}x = \int_{\overline{\mathcal{B}}^{\complement}} (-\mathscr{L}w)h\, \mathrm{e}^{-U/\varepsilon}\mathrm{d}x = \int_{\overline{\mathcal{B}}^{\complement}} w(-\mathscr{L}h)\, \mathrm{e}^{-U/\varepsilon}\mathrm{d}x. \tag{1.10}$$

The key idea is to extract $w(x)$ from the last integral, since $-\mathscr{L}h$ is a surface measure on $\partial\mathcal{A}$, and the total mass of this measure is the capacity.

It is well known from the regularity theory for second-order elliptic operators (see [28]) that a Harnack inequality holds for the equation $-\mathscr{L}f = g$. Using the Harnack inequality and the maximum principle for $-\mathscr{L}w = 1$, we can derive the following estimate: if x is a critical point of $U(x)$, and $U(x) \in \mathrm{C}^2(\mathbb{R}^d)$, then

$$\begin{aligned} \sup_{B_\varepsilon(x)} w(y) &\le \left(1 + C\varepsilon^{\alpha/2}\right) w(x), \\ \inf_{B_\varepsilon(x)} w(y) &\ge \left(1 - C\varepsilon^{\alpha/2}\right) w(x). \end{aligned} \tag{1.11}$$

Combining these two estimates with (1.10), we obtain the first key result:

$$w(\boldsymbol{m}_0) = [1 + o_1(\varepsilon)] \frac{\int_{\overline{\mathcal{B}}^{\complement}} h_{\mathcal{A},\mathcal{B}}(x)\, \mathrm{e}^{-U(x)/\varepsilon} \mathrm{d}x}{\mathrm{cap}(\mathcal{A}, \mathcal{B})}, \tag{1.12}$$

where we set $\delta = \varepsilon$, so that $\mathcal{A} = B_\varepsilon(\boldsymbol{m}_0)$ and $\mathcal{B} = B_\varepsilon(\boldsymbol{m}_1)$.

Renewal Estimate and Capacity Estimates

To estimate the integral in the numerator, we use the following observation, known as the renewal estimate. Its proof relies on the well-posedness of the Green function:

$$h_{\mathcal{A},\mathcal{B}}(x) \lesssim \frac{\mathrm{cap}\left(B_r(x), \mathcal{A}\right)}{\mathrm{cap}\left(B_r(x), \mathcal{B}\right)}. \tag{1.13}$$

Therefore, it suffices to estimate the capacity.

The estimation of capacity is based on its variational formulation, since $\mathscr{L}$ is self-adjoint with respect to the measure $\exp(-U(x)/\varepsilon)\mathrm{d}x$, namely

$$\mathrm{cap}(\mathcal{A}, \mathcal{B}) = \inf_{f \in \mathrm{H}_0^1(\overline{\mathcal{B}}^{\complement}),\, f \geq \mathcal{X}_{\mathcal{A}}} \varepsilon \int_{\overline{\mathcal{B}}^{\complement}} |\nabla f|^2\, \mathrm{e}^{U/\varepsilon} \mathrm{d}x, \tag{1.14}$$

where $\mathrm{H}_0^1(\overline{\mathcal{B}}^{\complement})$ is the standard Sobolev space of weakly differentiable functions whose trace vanishes on the boundary, and $\mathcal{X}_{\mathcal{A}}$ is the characteristic function of $\mathcal{A}$. Intuitively, the equilibrium potential is the unique function that minimizes the energy among all possible electrostatic potentials induced by charge distributions between the conductors $\mathcal{A}$ and $\mathcal{B}$.

To estimate the capacities appearing in (1.12) and (1.13), we carefully construct test functions and substitute them into the variational formulation (1.14), yielding

$$\mathrm{cap}(\mathcal{A}, \mathcal{B}) = [1 + o_\varepsilon(1)] \frac{(2\pi\varepsilon)^{d/2}}{2\pi} \frac{\lambda_1}{\sqrt{|\det(\nabla^2 U(0))|}} \mathrm{e}^{-U(0)/\varepsilon}. \tag{1.15}$$

The last step is to estimate the numerator in (1.12). From (1.13) and a rough capacity estimate, $h_{\mathcal{A},\mathcal{B}}$ is exponentially close to 1 in the basin of $\boldsymbol{m}_0$ and exponentially close to 0 in the basin of $\boldsymbol{m}_1$. Thus, the integral in the numerator of (1.12) can be evaluated using Laplace asymptotics:

$$\int_{\overline{\mathcal{B}}^{\complement}} h_{\mathcal{A},\mathcal{B}}(x)\, \mathrm{e}^{-U(x)/\varepsilon} \mathrm{d}x = [1 + o_\varepsilon(1)] \frac{(2\pi\varepsilon)^{d/2}}{\sqrt{\det(\nabla^2 U(\boldsymbol{m}_0))}} \mathrm{e}^{-U(\boldsymbol{m}_0)/\varepsilon}. \tag{1.16}$$

The key idea behind Laplace asymptotics is that, for small ε, the integral is dominated by the region near the maximum of the integrand $h_{\mathcal{A},\mathcal{B}} \exp(-U/\varepsilon)$, which is located at the critical point $\boldsymbol{m}_0$. Approximating $U(x)$ by its quadratic approximation around $\boldsymbol{m}_0$, and applying Gaussian integral estimates, we obtain the above result.

Finally, substituting (1.15) and (1.16) into (1.12) yields the final form of the Eyring-Kramers formula (1.3).

1.4 Aim of the Thesis: Towards the Underdamped Case

Metastability has been an active area of research for decades. A comprehensive introduction to different proofs of the Eyring-Kramers formula (1.3), including approaches based on large deviations and semiclassical analysis for the Witten Laplacian, can be found in N. Berglund's survey paper [11]. Many studies have explored generalizations of the process (1.1), including: the overdamped non-reversible case, see e.g., [39, 41]; the case where $U(x)$ has degenerate or singular critical points, see e.g., [13, 9, 8]; the case of a bounded domain with critical points on the boundary, see e.g., [56]; and the case where the potential is perturbed so that no explicit Gibbs measure of the form $\exp(-U/\varepsilon)$ exists, see e.g., [12]. This list is by no means exhaustive. There is also extensive research from a probabilistic perspective, which we do not detail here, but refer the reader to the references in the aforementioned works.

The Eyring-Kramers formula for the mean transition time of the underdamped Langevin process remained an open problem until the completion of this thesis. However, a full proof of the formula was very recently announced by Lee et al. [43], where they circumvent the potential theory approach and instead use probabilistic methods. Additionally, in their earlier work [42], they provided insights into the cutoff phenomenon for the underdamped Langevin process.

Consider the underdamped Langevin process, denoted by $Z_t = (X_t, V_t)$,

$$\begin{cases} \mathrm{d}X_t = V_t \mathrm{d}t, \\ \mathrm{d}V_t = -(\nabla_x U(x) + \gamma V_t)\mathrm{d}t + \sqrt{2\varepsilon}\mathrm{d}B_t, \end{cases} \tag{1.17}$$

where $(x, v) \in \mathbb{R}^n \times \mathbb{R}^n$ (so that $d = 2n$), B_t is the n-dimensional Brownian motion, and γ is the damping factor. The infinitesimal generator of this process is the Kramers-Fokker-Planck (KFP) operator, given by

$$\mathscr{K} f = \varepsilon \Delta_v f - \gamma v \cdot \nabla_v f - \nabla_x U \cdot \nabla_v f + v \cdot \nabla_x f. \tag{1.18}$$

Similar to (1.1), the underdamped Langevin process (1.17) can serve as a continuous model for stochastic gradient descent with momentum, following arguments similar to those in [46, 63]. It is also relevant as a statistical physics model for quantum tunneling, see e.g., [53].

Let $\boldsymbol{m}_1$ be the global minimum of $U(x)$, with $U(\boldsymbol{m}_0) > U(\boldsymbol{m}_1)$, and assume 0 is the unique saddle point. Under these conditions, the result from [43] states that the mean transition time can be estimated by

$$\mathbb{E}^{(\boldsymbol{m}_0,0)}[\tau_{(\boldsymbol{m}_1,0)}] \approx \frac{2\pi}{\mu_1} \sqrt{\frac{|\det \nabla^2 W(0,0)|}{\det \nabla^2 W(\boldsymbol{m}_0, 0)}} \mathrm{e}^{(W(0,0) - W(\boldsymbol{m}_0,0))/h}, \tag{1.19}$$

where τ is a stopping time defined analogously to (1.2) with the process replaced by (1.17), $h = \varepsilon/\gamma$, $W(x, v) := U(x) + V(v)$, and $V(v) := |v|^2/2$.

Moreover, $-\mu_1$ is the unique negative eigenvalue of the matrix

$$\begin{bmatrix} 0 & -\nabla_x^2 U(0) \\ \nabla_v^2 V(0) & \gamma \nabla_v^2 V(0) \end{bmatrix}. \tag{1.20}$$

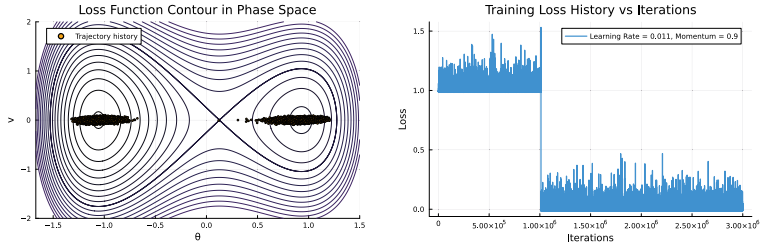

Figure 1.4. Metastability for SGDM with the same double well potential as in Fig. 1.3. The transition happens at around 1×10^6 iterations with a much smaller learning rate ($t \approx 1.1 \times 10^4$) which is much faster than SGD.

Both theoretic and empirical evidence suggest that, for a suitably chosen damping factor γ, the mean transition time should be shorter than in the reversible case. Fig. 1.4 presents a simulation of a toy model illustrating this behavior.

It is worth mentioning that an Eyring-Kramers formula for the low-lying spectrum of the KFP operator was also recently derived by Bony et al. [14] using semiclassical analysis. For the KFP operator, it is known from [33] that its nonzero spectrum lies within a sector of the complex plane, with its real part bounded below by a positive constant. See Fig. 1.5 for an illustration.

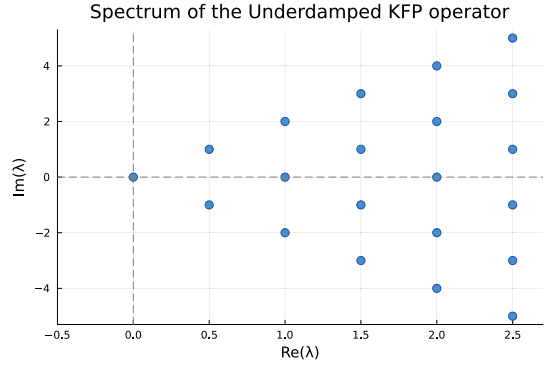

Figure 1.5. Example of eigenvalues of the KFP operator (1.18) with $U(x) = x^2$, $\varepsilon, \gamma = 1$, so that it's underdamped. If the friction is large enough eigenvalues will be real. For more details, see [60, 32].

We remark that, in the reversible case, (1.4) on a bounded set Ω defines a compact self-adjoint operator whose spectrum is discrete and consists only of real eigenvalues. The principle eigenvalue λ, which is the second eigenvalues

when ordered from 0 in increasing order, is closely related to the mean exit time via the Donsker-Varadhan inequality, cf. [24, 50]:

$$\lambda \geq \frac{1}{\sup_{x\in\Omega} \mathbb{E}^x[\tau_{\Omega^\complement}]}. \tag{1.21}$$

In fact, Bovier et al. [16] improved this inequality for the reversible case (1.4), showing that the low-lying eigenvalues can be estimated using capacity estimates and equilibrium potentials. Moreover, the corresponding eigenfunctions can be approximated by equilibrium potentials.

However, to the best of our knowledge, a connection between the low-lying spectrum and capacity estimates (or mean transition times) is not yet established for non-reversible cases. In the following, we explore a potential-theoretic approach to the metastability of (1.17), particularly through the study of Fokker-Planck-Kolmogorov equations of the form (1.18). Ultimately, our theory suggests a form of the Eyring-Kramers formula for the underdamped Langevin process, see Remark 3.4.5, that recovers (1.19) from [43]. This provides a promising direction for future research, offering a different approach to the Eyring-Kramers formula for the underdamped Langevin process based on potential theory and PDE tools. Furthermore, our findings open new research directions in Kolmogorov equations and potential theory, including investigations into higher-order boundary regularity and obstacle problems.

2. Fundaments of Fokker-Planck-Kolmogorov Operators

The general form of the Fokker-Planck-Kolmogorov operator studied in this thesis is given by, for $z \in \mathbb{R}^d$,

$$\mathscr{L} f(z) = \varepsilon \mathrm{e}^{W/\varepsilon} \nabla \cdot \left[\mathrm{e}^{-W/\varepsilon} \mathbf{A} \nabla f \right] + (\boldsymbol{b} + \mathbf{B} z) \cdot \nabla f, \tag{2.1}$$

where $W(z) : \mathbb{R}^d \to \mathbb{R}$ is a confining potential that grows sufficiently fast at infinity, and ε represents the thermal noise. Moreover, $\mathbf{A}(z)$ and $\boldsymbol{b}(z)$ satisfy certain structural conditions detailed in Paper IV. The matrix $\mathbf{B}$ is a constant matrix satisfying a homogeneous Lie group condition, which is essential in the regularity theory later on, see Paper IV, and also [2] for more details.

If $\mathbf{A}$ is divergence-free and has a symmetric part $\mathbf{S} = (\mathbf{A} + \mathbf{A}^\mathsf{T})/2$, then (2.1) is the generator of the following stochastic process:

$$\mathrm{d}Z_t = \left(-\mathbf{A}^\mathsf{T}(Z_t) \nabla W(Z_t) + \boldsymbol{b}(Z_t) + \mathbf{B} Z_t \right) \mathrm{d}t + \sqrt{2\varepsilon \mathbf{S}(Z_t)} \mathrm{d}B_t, \tag{2.2}$$

where B_t is the Brownian motion in $\mathbb{R}^d$.

The operator (2.1) includes the two special cases discussed in the previous chapter. In particular, when $\mathbf{A} = \mathbf{I}_d$, $\boldsymbol{b} = 0$, and $\mathbf{B} = 0$, it reduces to the overdamped reversible process (1.1) whose generator is given by (1.4).

Moreover, for $z = (x, v) \in \mathbb{R}^n \times \mathbb{R}^n$, $W(z) = U(x) + \gamma |v|^2/2$, when $U(x)$ is a confining potential, setting $\mathbf{A} = \operatorname{diag}(\mathbf{0}_n, \mathbf{I}_n)$, $\boldsymbol{b} = 0$ and $\mathbf{B} z = (v, 0)$ recovers the underdamped process (1.17), whose generator is given by (1.18).

Key Questions

To investigate the metastability of (2.2), we follow the proof strategy in Section 1.3 and consider the following questions:

- Does a Harnack inequality hold for (2.1)?
- Are the Dirichlet problems (1.6) and (1.8) well-posed for (2.1)? What is the regularity of their solutions up to the boundary?
- Does (2.1) admit a Green's function?
- What is a suitable definition of capacity induced by (2.1)? Is there a variational formulation applicable to non-self-adjoint operators?
- How can the capacity be effectively estimated?

We now address these questions one by one.

2.1 The Harnack Inequality and Interior Regularity of Solutions

The only question we can answer definitely is the interior Harnack inequality, which in turn establishes the interior regularity of solutions.

When **A** is uniformly elliptic, (2.1) reduces to a standard second order uniformly elliptic equation. In this case, the Harnack inequality and regularity results are well-established, see for instance, the classical textbook by Gilbarg and Trudinger [28].

For the degenerate case, where **A** is degenerate, significant progress has been made in the study of Kolmogorov equations. Our focus will primarily be on this case. Due to the local nature of the Harnack inequality, we can temporarily disregard the unboundedness of the coefficients in (2.1), as they are locally bounded. Throughout this section, we consider the following fundamental prototype operator, originally studied by Kolmogorov in the 1930s, see [36]. For $z = (x, v) \in \mathbb{R}^n \times \mathbb{R}^n$, we redefine the operator (2.1) as

$$\mathscr{L}u := \Delta_v u + v \cdot \nabla_x u \tag{2.3}$$

This is the generator of the stochastic process

$$\begin{cases} \mathrm{d}X_t = V_t \mathrm{d}t, \\ \mathrm{d}V_t = \sqrt{2}\mathrm{d}B_t, \end{cases} \tag{2.4}$$

where B_t is the Brownian motion in $\mathbb{R}^n$. A more commonly studied version of this operator appears in the kinetic Fokker-Planck equation, given by

$$(\partial_t - \mathscr{L}^*)u = f, \tag{2.5}$$

where $\mathscr{L}^* = \Delta_v - v \cdot \nabla_x$ is the formal adjoint of $\mathscr{L}$.

2.1.1 Definition of Weak Solutions

Before proceeding, we must specify the notion of solutions. Since (2.1) is in divergence form, the natural framework for solutions is that of weak solutions. Let $\mathcal{D} \subset \mathbb{R}^{2n}$ be an open set. We define the standard Sobolev space of functions as $\mathrm{H}^1(\mathcal{D}) = \mathrm{W}^{1,2}(\mathcal{D})$, and denote by $\mathrm{H}^1_0(\mathcal{D})$ the closure of $\mathrm{C}^\infty_c(\mathcal{D})$ in $\mathrm{H}^1(\mathcal{D})$. Additionally, we introduce the function space $\widehat{\mathrm{H}}^1(\mathcal{D}) := \{f(z) \in \mathrm{L}^2(\mathcal{D}) : \nabla_v f \in \mathrm{L}^2(\mathcal{D})\}$

Definition 2.1.1 *Let $\mathcal{D} \subset \mathbb{R}^d$ be an open set, and let $f \in \mathrm{L}^2(\mathcal{D})$. A function $u \in \widehat{\mathrm{H}}^1(\mathcal{D})$ is said to be a weak solution to the equation*

$$-\mathscr{L}u = f,$$

if, for all test functions $\varphi \in \mathrm{C}^1_c(\mathcal{D})$, the following holds:

$$\int_{\mathcal{D}} \left(\nabla_v u \cdot \nabla_v \varphi + u(v \cdot \nabla_x \varphi)\right) \mathrm{d}z = \int_{\mathcal{D}} f\varphi \, \mathrm{d}z.$$

2.1.2 The Harnack Inequality

The Harnack inequality plays a fundamental role in the regularity theory. It essentially says that a non-negative solution cannot grow too wildly within any small neighborhood. As a result, the oscillation of the solution can be controlled, leading to Hölder continuity.

For Kolmogorov equations, the Harnack inequality asserts that if u is a non-negative weak solution to $-\mathscr{L}u = f$, then there exists a constant C such that

$$\sup_{\mathcal{Q}_-} u \leq C \left(\inf_{\mathcal{Q}_+} u + \|f\|_{\mathrm{L}^\infty(\mathcal{Q})} \right),$$

where $\mathcal{Q}_-, \mathcal{Q}_+ \subset \mathcal{Q} \subset \mathbb{R}^{2n}$ are cubes shifted according to the Lie group structure associated with the equation.

Intuitively, equation (2.3) behaves like an elliptic equation in the v-direction, while exhibiting transport (or "parabolic") character in the x-direction. Due to this structure, the Kolmogorov equation is sometimes referred to as an ultraparabolic equation. Consequently, there is a "spatial lag" in the Harnack inequality, analogous to the time lag in the Harnack inequality for the heat equation. This spatial lag arises due to the transport of information by the vector field $v \cdot \nabla_x$, while the Laplacian Δ_v diffuses information in the v-direction.

A more intuitive understanding emerges from the stochastic representation of the solution, by means of the Dynkin's formula (1.5), where the value at a given point depends only on information carried along all possible trajectories. In the kinetic case, an *admissible curve* is one that moves in the direction of its current velocity, and the *attainable set* consists of points that can be reached by such curves within the domain $\mathcal{D}$.

Fig. 2.1 illustrates this concept for the case $n = 1$. It is worth noting that, for a cube centered at the origin, there is no spatial lag in the Harnack inequality, as any two points within the cube can be connected by an admissible curve.

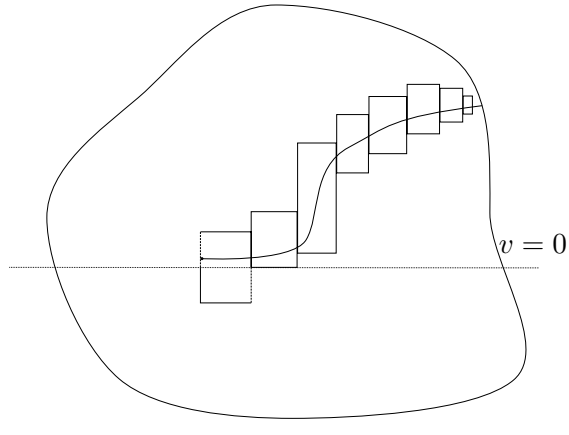

Figure 2.1. Illustration of admissible curves and Harnack chains.

Notably, the Harnack inequality for the non-divergence form of (2.5) was proved by Di Francesco and Polidoro in [22]. For weak solutions to the kinetic Fokker-Planck equation $(\partial_t - \mathscr{K}^*)u = f$, where $\mathscr{K}$ is defined in (1.18), the first proof was given in [29], followed by an alternative approach in [30] and a quantitative Harnack inequality in [31]. Furthermore, [3] extends the Harnack inequality to Kolmogorov equations of the form (2.5), where $\mathscr{L}$ satisfies (2.1).

Since a solution to the stationary equation $-\mathscr{L}u = f$ can be regarded as a time-independent solution to the corresponding evolutionary equation, the Harnack inequality for the latter naturally carries over to the stationary case.

2.2 Well-posedness and Boundary Regularity of Solutions

We can partially address the well-posedness and boundary regularity of solutions. However, the question of higher-order boundary regularity remains open.

It is well known in the field that the notion of weak solutions is more suitable for product domains, such as $\mathcal{D} = \mathcal{U} \times \mathcal{V}$ where $\mathcal{U}, \mathcal{V} \subset \mathbb{R}^n$. The canonical function space for solutions, known as the hypoelliptic space, was introduced in [1] and is defined as

$$\mathrm{H}^1_{\mathrm{hyp}}(\mathcal{U} \times \mathcal{V}) := \left\{ f(x,v) \in \mathrm{L}^2(\mathcal{U}; \mathrm{H}^1(\mathcal{V})) : v \cdot \nabla_x f \in \mathrm{L}^2(\mathcal{U}; \mathrm{H}^{-1}(\mathcal{V})) \right\},$$

where $\mathrm{H}^{-1}(\mathcal{V}) := (\mathrm{H}^1_0(\mathcal{V}))^*$ is the dual space.

The major challenge in this setting is that $v \cdot \nabla_x f$ is merely a distribution, making it difficult to define a suitable trace operator. Consequently, an integration-by-parts formula involving $v \cdot \nabla_x$ is not readily available. This issue becomes even more significant when the underlying domain is not a product domain.

We also remark that the well-posedness of weak solutions for the time-dependent equation $\partial_t u - \mathscr{L}u = f$ (2.5) or its variations in product domains is actively studied in the literature, see [19, 54, 55, 47, 65, 62, 7] and so on, as well as Paper I. The list here is not exhaustive, and we refer to the references therein for more details.

In Paper II, we establish the well-posedness of weak solutions to the Dirichlet problem for the stationary equation $-\mathscr{L}u = 0$ in bounded product domains, based on recent advances in the theory of weak solutions, particularly [64, 65].

To circumvent the difficulty of the lack of a notion of trace in arbitrary domains, we introduced the notion of Perron's solution, which builds upon the well-posedness theory in bounded domains. This approach is developed in Paper II, and we will outline its key ideas in this section.

Throughout this section, we continue to focus on the prototype operator (2.3).

2.2.1 Existence of Perron's Solution and Boundary Regularity in Bounded Domains

For the evolutionary equation $(\partial_t - \mathscr{L})u = f$, where $\mathscr{L}$ is in non-divergence form and with Hölder continuous coefficients in bounded domains, the Perron solution has been extensively studied within the framework of potential theory. Notably, Manfredini [51], and Di Francesco and Polidoro [22] established results by using the existence of a fundamental solution under the homogeneous Lie group condition on **B** and employing an abstract axiomatic β-harmonic space approach.

However, for the stationary equation $-\mathscr{L}u = 0$, the existence and properties of a fundamental solution remain unclear in the literature. This issue will be further discussed in Section 2.3. Additionally, since $\mathscr{L}$ is in divergence form and may have rough coefficients, classical solutions may not be well-defined when **A** lacks differentiability. This demands a more flexible notion of solution.

Definition 2.2.1 *Let $\Omega \subset \mathbb{R}^{2n}$ be an open set. A function $u \in \mathrm{C}(\Omega)$ is called a* K*-harmonic function if u is a weak solution to the equation $-\mathscr{L}u = 0$ in every product domain $\mathcal{D} = \mathcal{U} \times \mathcal{V}$ compactly contained in Ω.*

The weak solution theory in bounded product domains serves as the foundation for the construction of Perron's solutions, forming the first part of Paper II.

In potential theory, K-*subharmonic functions* and K-superharmonic functions are defined through the comparison principle. Intuitively, they can be understood as functions that are below or above possible solutions satisfying $-\mathscr{L}u \leq 0$ or $-\mathscr{L}u \geq 0$ respectively. The *Perron upper and lower solutions*, denoted by $\overline{H}_g$ and $\underline{H}_g$, are constructed via the Perron envelope. Specifically, $\overline{H}_g$ is the infimum of all K-superharmonic functions that are greater than or equal to g on the boundary, while $\underline{H}_g$ is the supremum of all K-subharmonic functions that are less than or equal to g on the boundary.

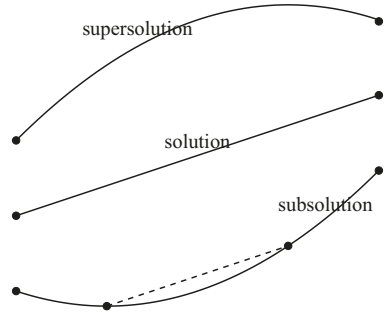

Figure 2.2. Illustration of the Perron envelope.

In Paper II, we obtain the existence of K-harmonic Perron's solutions for the Dirichlet problem

$$\begin{cases} -\mathscr{L}u = 0, & \text{in } \Omega, \\ u = g, & \text{on } \partial\Omega, \end{cases} \tag{2.6}$$

If **A** is differentiable, meaning that (2.1) is in non-divergence form with rough coefficients, then the upper and lower Perron solutions coincide. This property is known as the *resolutivity*.

Furthermore, by applying the Harnack inequality and regularity theory discussed earlier, we conclude that Perron's solutions are automatically Hölder continuous in the interior of the domain.

2.2.2 Regular Points and Barriers

To determine whether the Perron solution attains the prescribed boundary values, we must analyze the regularity of solutions up to the boundary. This is typically done by constructing a *barrier*, which is a subsolution that touches the boundary value from below.

A boundary point where the Perron solution attains the boundary value is called a *regular point*, meaning that for $z_0 \in \partial\Omega$,

$$\lim_{\Omega \ni z \to z_0} \overline{H}_g(z) = g(z_0).$$

The set of all regular points for a given domain Ω is denoted by $\Gamma(\Omega)$.

The characterization of regular points is a rich area of study, particularly in relation to the Wiener criterion; see [2] for a survey. However, for the purposes of this thesis, simple geometric conditions for regularity will suffice. Specifically, following [51], a boundary point is regular if the outward normal vector $\boldsymbol{n} = (\boldsymbol{n}_x, \boldsymbol{n}_v)$ satisfies $\boldsymbol{n}_v \neq 0$.

By leveraging the geometric criterion from [51], we establish in Paper II that certain velocity symmetric sets are regular, as illustrated in Fig. 2.3. In particular, a ball centered at the origin is entirely regular. Furthermore, for the domains considered in (1.6) and (1.8), their boundaries are almost everywhere regular. As a consequence, the Perron solution is automatically continuous up to the regular part of the boundary.

2.2.3 Maximum Principle

Another crucial property of solutions is the maximum principle, which states that the maximum of a continuous solution is always attained on the regular part of the boundary, namely

$$\sup_{\Omega} u = \sup_{\Gamma(\Omega)} u. \tag{2.7}$$

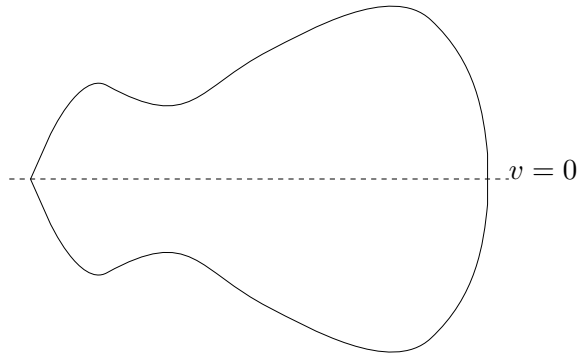

Figure 2.3. Illustration of velocity symmetric set when $n = 1$.

The primary role of the maximum principle is to establish a comparison principle, which in turn guarantees the uniqueness of solutions to the Dirichlet problem.

For the case where $\mathscr{L}$ is in non-divergence form, the maximum principle for the evolutionary equation $\partial_t u - \mathscr{L} u = 0$ was established in [51] using a classical argument by contradiction. The same reasoning applies to the stationary equation in non-divergence form as well.

For divergence form, applying classical arguments becomes more challenging: the standard contradiction argument, such as [51], relies on the non-divergence form which is not available for divergence form operators; energy estimates for weak solutions, as used in [28], require an integration by parts formula, which is problematic for non-product domains.

However, a strong maximum principle, which states that the supremum can only be attained on the attainable boundary (i.e., the intersection of the closure of the attainable set of all interior points and the boundary), can still be derived. This follows from the Harnack inequality along admissible curves (discussed in Section 2.1.2), as shown in [2]. For simple domains, such as balls, the attainable boundary coincides with the regular part of the boundary. Consequently, the above strong maximum principle is equivalent to the classical one and implies (2.7).

2.2.4 Dirichlet Problems in Unbounded Domains

For the equilibrium problem (1.8), where the domain has a smooth boundary but is unbounded, we construct a sequence of solutions by considering the problem in $B_R(0) \cap \Omega$ for sufficiently large R, imposing a boundary condition of 0 on $\partial B_R(0)$. Using the maximum principle, we can show that the sequence converges monotonically, and its limit is a Perron's solution to (1.8).

For the landscape problem (1.6), we need to handle the nonzero right-hand side. We reduce it to a problem with a homogeneous right-hand side as follows:

Recall the landscape equation (1.6) is given by

$$\begin{cases} -\mathscr{L}w = 1, & \text{in } \overline{\mathcal{B}}^{\complement}, \\ w = 0, & \text{on } \partial\mathcal{B}. \end{cases} \tag{2.8}$$

By setting $u = w - g$, we transform this into

$$\begin{cases} -\mathscr{L}u = 0, & \text{in } \overline{\mathcal{B}}^{\complement}, \\ u = -g, & \text{on } \partial\mathcal{B}, \end{cases} \tag{2.9}$$

where g is a solution to $-\mathscr{L}g = 1$ in the entire $\mathbb{R}^{2n}$. For existence of a solution to $-\mathscr{L}w = 1$ in $\mathbb{R}^{2n}$, we refer to [1], which applies the Lions-Lax-Milgram theorem.

2.2.5 Higher Regularity up to the Boundary

As mentioned earlier, the regularity of solutions up to the boundary is a challenging problem, and it is not yet clear how to obtain higher regularity up to the boundary for any notion of solutions.

For simplicity, let's consider the equation $-\mathscr{L}u = 0$ in a bounded product domain $\mathcal{D} = \mathcal{U} \times \mathcal{V}$, where the Perron solution coincides with the weak solution, as shown by the weak regularity theory presented in Paper II and also in [64, 65]. In particular, it is shown in the appendix of [65], without further restrictions on the boundary data, even for equations with smooth coefficients and smooth boundary data, the regularity in the x-direction is expected to be only $\mathrm{H}^{1/3}$.

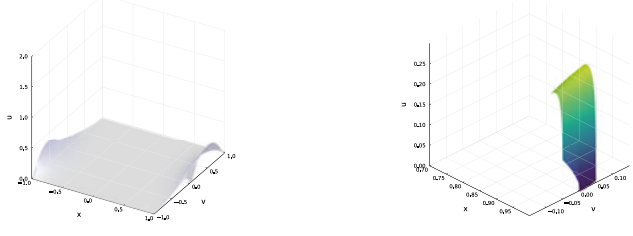

Figure 2.4. Monte Carlo simulation of a Dirichlet problem that the solution is $\mathrm{H}^1((-1,1)\times(-1,1))$. The left figure shows the solution in the entire domain, and the right figure is a zoomed-in view near the singular set.

Interestingly, (2.3) is also known as the forward-backward equation in fluid dynamics, and in a recent preprint [20], it is proven that the solution will be $\mathrm{H}^1(\mathcal{D})$, provided the boundary data vanishes fast enough near the singular set (or grazing set), where the velocity is tangent to the position boundary. However, the numerics suggest that, even in this case, the regularity of the solution

at the singular set is still roughly Hölder continuous of order $1/3$ in the x-direction.

The regularity in the v-direction is expected to be better, since the equation is elliptic in the v-direction. We expect the solution u to have $\nabla_v u$ continuous up to $\mathcal{U} \times \partial\mathcal{V}$ at least. However, no rigorous proof is available yet, and these are open questions for future research.

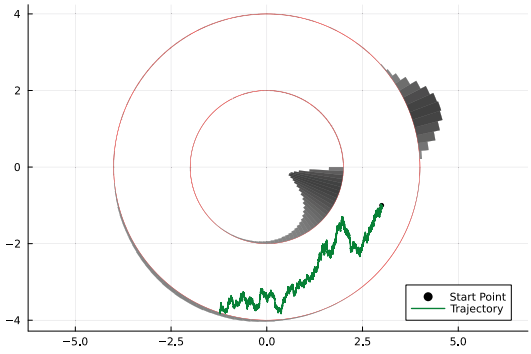

Figure 2.5. Monte Carlo simulation of the hitting probability starting from a point. This is also known as the harmonic measure and is relevant to the boundary regularity of solutions. Relevant studies can be found in Litsgård's thesis [48] and the references therein.

2.3 The Green Function

We now turn our attention to the existence of the Green function for the operator (2.1). In this section, we must take into account the unbounded nature of the operator (2.1), namely, the coefficients blow up as $|z| \to \infty$.

For the case when **A** is uniformly elliptic, the Green's function for (2.1) is known to exist, and it can be obtained by the semigroup method, as shown in [58].

For the case when **A** is degenerate and the coefficients are smooth, it is known that a fundamental solution $K(t, z, z')$ exists for the evolutionary equation $(\partial_t - \mathscr{L})K = 0$, as discussed in [59]. In fact, for the semigroup,

$$T_t g(z) = f(t, z)$$

where $f(t, z)$ is the solution to the Cauchy problem

$$\begin{cases} (\partial_t - \mathscr{L})f(t, z) = 0, & t > 0, \\ f(0, z) = g, & z \in \mathbb{R}^d, \end{cases} \tag{2.10}$$

there exists a transition probability

$$P_t(z, \mathrm{d}z') = p_t(z, z')\mathrm{d}z',$$

where $p_t(z, z')$ is a distributional solution to the evolutionary equation

$$(\partial_t - \mathscr{L})p_t(z, \cdot) = 0, \text{ and } (\partial_t - \mathscr{L}^*)p_t(\cdot, z') = 0.$$

By the Hömander property (see e.g., [34] and [18] for a modern version) of $\partial_t - \mathscr{L}$, the transition probability is smooth for $t > 0$ and $z \neq z'$. Regarding the well-posedness of the semigroup for smooth coefficients, Paper I provides a well-posedness result for (2.10) based on the Galerkin method. For rough coefficients, the well-posedness can be easily obtained via the Lions-Lax-Milgram theorem as shown in [1], where the Hömander gain of regularity for less smooth coefficients is also discussed.

Moreover, for the Kramers-Fokker-Planck operator $(\partial_t - \mathscr{K})$, where $\mathscr{K}$ is defined in (1.18), there is a Gaussian upper bound for the kernel, as shown in [44]. We expect that the fundamental solution of the stationary equation $-\mathscr{L}u = 0$ can be obtained by taking the time average of the fundamental solution of the evolutionary equation, i.e. $p(z, z') = \int_0^\infty K(t, z, z')\mathrm{d}t$. However, this has not been rigorously written down. On the other hand, from the hypocoercivity theory, we expect that the fundamental solution $K(t, z, z')$ decays exponentially in time, which supports the time-average method. These are open questions for future research.

We remark that for $\partial_t - \mathscr{L}$ with bounded coefficients, the fundamental solution is known to exist for a long time. In particular, for constant coefficients, the fundamental solution can be constructed explicitly, see e.g., [34]. For Hölder continuous coefficients, see [21], where further pointwise estimates are obtained. Recently, for equations with rough coefficients, this has been studied in [3]. For the stationary equation $-\mathscr{L}u = 0$ with bounded coefficients, nothing definitive has yet been written, but we expect the time average method to work in this case as well.

2.4 Definition of Capacity and Variational Formulation

The most important part of the theory is the definition of capacity and the variational formulation for non-self-adjoint operators.

In the literature, a definition of capacity for the evolutionary operator $\partial_t - \mathscr{L}$ based on kernels (or the fundamental solution) has been provided, see e.g., [35], for the study of Wiener criterion and regularity theory. However, for the stationary operator $-\mathscr{L}$, the definition of capacity remains unclear, as a fundamental solution for the operator is not yet known, as discussed in Section 2.3. On the other hand, the kernel definition is not convenient for calculations. Therefore, we define a capacity through the equilibrium potential $h_{\mathcal{A},\mathcal{B}}$, the solution to (1.8), following the classical potential theory, by finding the measure $-\mathscr{L}h_{\mathcal{A},\mathcal{B}}$ on $\partial\mathcal{A}$.

2.4.1 Definition of Capacity

In particular, let's consider the Kramers-Fokker-Planck operator $\mathscr{K}$ defined in (1.18), and, without loss of generality, set $\varepsilon, \gamma = 1$.

Similar to the classical potential theory, we extend the equilibrium potential $h = h_{\mathcal{A},\mathcal{B}}$ to the set $\mathcal{A}$ by setting $h \equiv 1$ on $\mathcal{A}$. By testing $-\mathscr{K}h$ against a test function $\varphi \in \mathrm{C}_c^\infty(\overline{\mathcal{B}}^{\complement})$, we obtain the capacity between $\mathcal{A}$ and $\mathcal{B}$ as

$$\mathrm{cap}(\mathcal{A},\mathcal{B}) = \int_{\partial\mathcal{A}} (\varepsilon \nabla_v h \cdot \boldsymbol{n}_{v,\Omega} + (h-1)v \cdot \boldsymbol{n}_{x,\Omega})\, \mathrm{e}^{-W/\varepsilon} \sigma(\mathrm{d}z),$$

where $\Omega = (\overline{\mathcal{A} \cup \mathcal{B}})^{\complement}$, $\boldsymbol{n}_\Omega = (\boldsymbol{n}_{x,\Omega}, \boldsymbol{n}_{v,\Omega})$ is the outward normal, $\sigma(\mathrm{d}z)$ is the surface measure on $\partial\mathcal{A}$, and $W(x,v) = U(x) + \gamma|v|^2/2$.

2.4.2 Symmetry Relations

A crucial property needed for a variational formulation is the self-adjointness of the operator, which is not the case for the operator (2.1). However, we can still identify certain symmetry relations that are crucial for understanding the variational formulation. Recall that

$$\mathscr{K} f = \Delta_v f - v \cdot \nabla_v f - \nabla_x U \cdot \nabla_v f + v \cdot \nabla_v f.$$

The adjoint operator $\mathscr{K}^*$ with respect to the Lebesgue measure is given by

$$\mathscr{K}^* f = \nabla_v \cdot (\nabla_v f + v f) + \nabla_x U \cdot \nabla_v f - v \cdot \nabla_x f.$$

It is easy to verify that $\mathscr{K}^* \rho(x,v) = 0$, where $\rho(x,v) = \exp(-W)$, hence there is a stationary measure

$$\rho(\mathrm{d}z) = \exp(-W(z))\mathrm{d}z.$$

The adjoint of $\mathscr{K}$ with respect to ρ is given by

$$\mathscr{K}^\dagger f = \Delta_v f - v \cdot \nabla_v f + \nabla_x U \cdot \nabla_v f - v \cdot \nabla_x f.$$

In fact, $\mathscr{K}^\dagger$ is the generator of the following process, where we denote $Z_t^\dagger = (X_t, V_t)$:

$$\begin{cases} \mathrm{d}X_t = -V_t \mathrm{d}t, \\ \mathrm{d}V_t = (\nabla_x U(X_t) - V_t)\mathrm{d}t + \sqrt{2}\mathrm{d}B_t. \end{cases}$$

The adjoint of $\mathscr{K}^\dagger$ with respect to the Lebesgue measure is given by

$$\mathscr{K}^{\dagger *} f = \nabla_v \cdot (\nabla_v f + v f) - \nabla_x U \cdot \nabla_v f + v \cdot \nabla_x f.$$

Moreover, $\mathscr{K}$ and $\mathscr{K}^{\dagger *}$ are conjugate equivalent, i.e. $\rho \mathscr{K}(\rho^{-1} f) = \mathscr{K}^{\dagger *} f$, and similarly for $\mathscr{K}^*$ and $\mathscr{K}^\dagger$. This duality relation among the operators can be illustrated in Fig. 2.6.

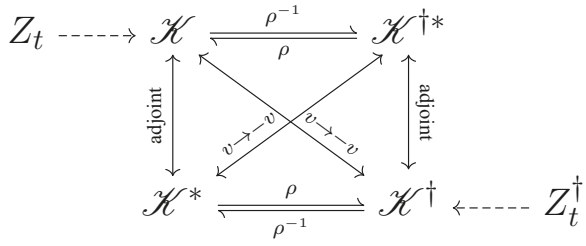

Figure 2.6. Duality relation among the operators.

Finally, we can define an adjoint capacity between two sets $\mathcal{A}$ and $\mathcal{B}$ as

$$\mathrm{cap}(\mathcal{A},\mathcal{B}) = \int_{\partial\mathcal{A}} \left(\varepsilon \nabla_v h^\dagger \cdot \boldsymbol{n}_{v,\Omega} + (1-h^\dagger) v \cdot \boldsymbol{n}_{x,\Omega}\right) \mathrm{e}^{-W/\varepsilon} \sigma(\mathrm{d}z),$$

where $h^\dagger$ is the equilibrium potential associated with $\mathscr{K}^\dagger$, namely

$$\begin{cases} -\mathscr{L}^\dagger h^\dagger = 0, & \text{in } (\overline{\mathcal{A}\cup\mathcal{B}})^{\complement}, \\ h^\dagger = 1, & \text{on } \partial\mathcal{A}, \\ h^\dagger = 0, & \text{on } \partial\mathcal{B}. \end{cases}$$

2.4.3 Variational Formulation

Inspired by [1], we introduce a variational formulation in Paper IV. For a nice smooth function f vanishing on $\mathcal{B}$, we define the functional

$$J[f] := \inf_{\boldsymbol{g}\in \mathrm{V}_f} \varepsilon \int_{\overline{\mathcal{B}}^{\complement}} |\nabla_v f - \boldsymbol{g}|^2 \, \mathrm{e}^{-W/\varepsilon} \mathrm{d}z, \tag{2.11}$$

where the admissible vector field $\boldsymbol{g} = (g_1, \ldots, g_n)$ satisfies

$$\varepsilon e^{W/\varepsilon} \nabla_v \cdot \left(e^{-W/\varepsilon} \boldsymbol{g}\right) = \nabla_x U \cdot \nabla_v f - v \cdot \nabla_x f.$$

It is shown in Paper IV that the capacity can be controlled by the functional $J[f]$ in (2.11), which is given by

$$\varepsilon \int_\Omega |\nabla_v h|^2 \mathrm{e}^{-W/\varepsilon} \mathrm{d}z \le \mathrm{cap}(\mathcal{A},\mathcal{B}) \le \inf_{f\ge \mathcal{X}_{\mathcal{A}}} J_{\overline{\mathcal{B}}^{\complement}}[f].$$

It is worth mentioning that if h is sufficiently regular, which is still unclear, as discussed in Section 2.2.5, then the above inequality becomes an equality, and the minimum is attained at $\bar{f} = (h + h^\dagger)/2$ and $\bar{g} = (\nabla h - \nabla h^\dagger)/2$. This observation also leads to the question of variational obstacle problems for Kolmogorov operators, as the minimization problem above naturally takes an obstacle form. This is an interesting direction for future research. While

some attention has been given to this topic in the literature, see recent works [4, 17], further investigation is needed to understand the variational obstacle problem for Kolmogorov operators.

We also remark that if **A** in (2.1) is uniformly elliptic, while (2.1) still being non-self-adjoint, a variational formulation for the capacity can be obtained using the same method. In this case, the inequality becomes an equality, with the minimum attained by the average of the equilibrium potentials. For further details, see Paper IV.

2.5 Rough and Sharp Capacity Estimates

In this section, we discuss rough and sharp capacity estimates for the operator (2.1), which are needed in (1.12) and (1.13). For simplicity, we consider the reversible operator (1.4), as the idea can be extended to the general operator (2.1). Recall from (1.14) that

$$\operatorname{cap}(\mathcal{A}, \mathcal{B}) = \inf_{f \in \mathrm{H}_0^1(\overline{\mathcal{B}}^{\complement}),\, f \geq \mathcal{X}_{\mathcal{A}}} J[f]$$

where

$$J[f] = \varepsilon \int_{\overline{\mathcal{B}}^{\complement}} |\nabla f|^2 \, \mathrm{e}^{-U/\varepsilon} \mathrm{d}x. \tag{2.12}$$

2.5.1 Rough Capacity Estimates

We first establish rough bounds for the capacity, which will be used in (1.13). There are two directions to consider.

Upper Bound

Let $\mathcal{A} \subset \mathbb{R}^d$ be a bounded smooth set around a local minimum, and let $z \in \overline{\mathcal{A}}^{\complement}$ be any point that is away from $\mathcal{A}$. To estimate $\operatorname{cap}(B_\varepsilon(z), \overline{\mathcal{A}})$, we proceed as follows: Define $z^\star$ as the point where the potential U attains its maximum along the lowest energy path connecting z and $\mathcal{A}$. In other words, $z^\star$ represents the highest energy barrier that must be overcome to reach $\mathcal{A}$ from z.

If $z^\star = z$, meaning there is no energy barrier to overcome, we can construct a test function f that is 1 on $B_\varepsilon(z)$ and decays to 0 on $B_{2\varepsilon}(z) \setminus B_\varepsilon(z)$, ensuring that $|\nabla f| \leq 1/\varepsilon$.

On the other hand, if $z^\star = 0$, meaning $\mathcal{A}$ and z lie in two different energy wells, we construct f such that it is 1 on the whole valley that contains z, denoted by $\mathcal{V}$, and transitions from 1 to 0 on $(\mathcal{V})_\varepsilon \setminus \mathcal{V}$, with $|\nabla f| \leq 1/\varepsilon$, where $(\mathcal{V})_\varepsilon := \{y \in \mathbb{R}^d : \operatorname{dist}(y, \mathcal{V}) < \varepsilon\}$. By substituting the constructed test functions into the variational formulation (2.12), the upper bound of $\operatorname{cap}(B_\varepsilon(z), \overline{\mathcal{A}})$ is roughly of order $\exp(-U(z^\star)/\varepsilon)$.

Lower Bound

For the lower bound, we restrict our analysis to a smaller set, constructed as follows: first, connect $\mathcal{A}$ and $B_\varepsilon(z)$ by a path $\omega(t)$ that passes through the minimum-height point $z^\star$. Then, consider a small neighborhood of the path, denoted by $(\omega)_r$ for $r > 0$. In this way, $J[f]$ is bounded below by the norm of the directional derivative of f along $(\omega)_r$:

$$J[f] \geq \int_{B_r} \mathrm{d}z_\perp \int_0^{|\omega|} |\partial_t f\left(\omega(t) + z_\perp\right)|^2 \mathrm{e}^{-U(\omega(t)+z_\perp)/\varepsilon} \mathrm{d}t.$$

The minimization problem is reduced to a one-dimensional problem, which can be solved by solving an ordinary differential equation. As a result, the lower bound is also roughly of order $\exp(-U(z^\star)/\varepsilon)$.

Combining the two bounds above and applying the renewal estimate (1.13), we conclude that the equilibrium potential $h_{\mathcal{A},\mathcal{B}}$ is exponentially close to one in the valley containing $\mathcal{A}$ and is exponentially close to zero in the valley containing $\mathcal{B}$.

We note that the rough capacity estimates above originate from [15] and were later extended to the elliptic but non-reversible case in Paper IV. A possible future direction is to extend these results to the degenerate case.

2.5.2 Sharp Capacity Estimates

For the sharp capacity, particularly in the reversible case, the idea is to construct a test function f that closely approximates the equilibrium potential $h_{\mathcal{A},\mathcal{B}}$, and then substitute it directly into the variational formulation (2.12).

However, in the non-reversible case, such as Kramers-Fokker-Planck operator and its variational formulation (2.11), we have to construct an admissible vector field $\boldsymbol{g}$ that is convenient to use in the variational formulation. While finding an admissible vector field $\boldsymbol{g}$ is straightforward since they are not unique, identifying a suitable one that allows for easy computation of the functional $J[f]$ is more challenging. To overcome this difficulty, we use an alternative approach motivated by the variational formulation, as discussed in [41].

Let $h = h_{\mathcal{A},\mathcal{B}}$ be the equilibrium potential associated with (1.4), and let $f \in \mathrm{H}_0^1(\overline{\mathcal{B}}^{\complement})$ satisfy $f = 1$ on $\partial\mathcal{A}$. Then we have

$$\mathrm{cap}(\mathcal{A},\mathcal{B}) = \int_{\overline{\mathcal{B}}^{\complement}} \nabla h \cdot \nabla f \, \mathrm{e}^{-U/\varepsilon} \mathrm{d}z = \int_{\overline{\mathcal{B}}^{\complement}} h(-\mathscr{L} f) \, \mathrm{e}^{-U/\varepsilon} \mathrm{d}z. \tag{2.13}$$

Since we know that the minimizer is h, we aim to construct a test function f that approximates h well.

Using the rough estimates from the previous section, we design f so that it takes the value 1 in the valley containing $\mathcal{A}$, 0 in the valley containing $\mathcal{B}$, and

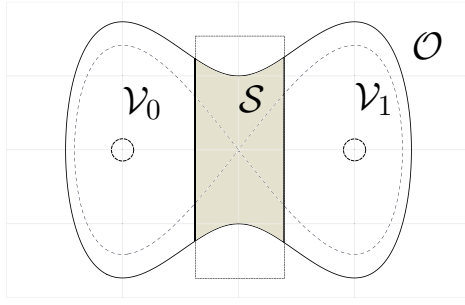

Figure 2.7. The separation of the domain into different regions.

transitions rapidly near the saddle point 0. Following this idea, we first split the domain into several regions, as illustrated in Fig. 2.7.

Next, we define the function

$$p(z) = \frac{1}{C}\int_{-\infty}^{z\cdot \boldsymbol{e}_1} \exp\left(-\frac{\beta}{2\varepsilon}t^2\right)\,\mathrm{d}t, \quad z \in \mathcal{S}, \tag{2.14}$$

where $\boldsymbol{e}_1 = (1, 0, \ldots, 0)$ is the eigenvector of $\nabla^2 U(0)$, and

$$C := \int_{-\infty}^{\infty} \exp\left(-\frac{\beta}{2\varepsilon}t^2\right)\,\mathrm{d}t = \sqrt{\frac{2\pi\varepsilon}{\beta}}.$$

Indeed, p solves the linearized equation around the saddle point 0, i.e.

$$\varepsilon\Delta p(z) - \nabla U(0)x \cdot \nabla p(z) = 0.$$

We then construct the test function f as follows:

$$f = \begin{cases} p(z), & z \in \mathcal{S}, \\ 1, & z \in \mathcal{V}_1, \\ 0, & z \in \mathcal{V}_0 \cup \mathcal{O}. \end{cases}$$

Finally, we mollify f to obtain a smooth function $f_\eta := f * \varphi_\eta$, where φ_η is a standard mollifier supported in $B_\eta(0)$. Through careful calculations on the right-hand side of (2.13), we ultimately obtain a sharp capacity estimate, leading to an Eyring-Kramers formula as given in (1.3).

We remark that the above sharp capacity estimate is originally from [39, 41], later extended to the general non-reversible overdamped case in Paper IV, and very recently extended to the underdamped case (1.17) in [43].

2.6 What Has Been Done in This Thesis?

In this thesis, we focus on the well-posedness and regularity of solutions, as well as the definition of capacity and the variational formulation for the operator (2.1).

In particular, in Paper I, we use the Galerkin method to construct solutions to the evolutionary equation $\partial_t u - \mathscr{L} u = f$ in periodic domains and in the whole space. We also analyze the dependence of regularity on the initial data and coefficients. The Galerkin approximation provides an alternative proof of well-posedness and offers a potential numerical approach for simulating the evolution of the underlying probability density function of the stochastic process (2.2).

In Paper II, we establish the well-posedness of weak solution in bounded product domains and prove the existence of Perron's solution to the stationary equation $-\mathscr{L} u = 0$ in bounded domains. Additionally, we study the regularity up to the boundary, as discussed in Section 2.2.

Paper III is not directly related to metastability, but this paper contributes to the regularity theory of the operator (2.1). We prove the boundedness of solutions to $\partial_t u - \mathscr{L} u = f$ in bounded product domains using De Giorgi iteration, where the operator $\mathscr{L}$ now includes lower-order divergence terms. As a consequence, the Harnack inequality can be extended to a broader class of Kolmogorov operators.

In Paper IV, we define the capacity through the equilibrium potential and provide a variational formulation for the operator (2.1). Furthermore, we prove the Eyring-Kramers formula in the case where $\mathbf{A}$ is uniformly elliptic, demonstrating the feasibility of our definition.

3. Summary of Articles

3.1 Paper I

In this paper, we study the well-posedness of the Kramers-Fokker-Planck equation in periodic bounded domains using the Galerkin method. We note that the Galerkin method is commonly used in the field of numerical analysis, where it is typically assumed that the problem is already well-posed, and the focus is on how accurately the method approximates the solution. However, the Galerkin method is also a powerful tool for proving well-posedness of parabolic PDEs, as demonstrated in [25, 38].

Our focus in this paper is on obtaining weak solutions to the Kramers-Fokker-Planck equation via Galerkin approximations. As a consequence, we also derive a quantitative error estimate.

Recall that the Kramers-Fokker-Planck operator, with $\varepsilon, \gamma = 1$, is given by

$$\mathscr{K} u = \Delta_v u - v \cdot \nabla_v u - \nabla_x U \cdot \nabla_v u + v \cdot \nabla_v u.$$

We consider the following Cauchy problem:

$$\begin{cases} (\partial_t - \mathscr{K})u = f & \text{in } (0,T) \times \mathcal{D} \times \mathbb{R}^n, \\ u = g & \text{on } \{t = 0\} \times \mathcal{D} \times \mathbb{R}^n, \end{cases} \tag{3.1}$$

where either $\mathcal{D} = \mathbb{T}^n$, the n-dimensional torus, or $\mathcal{D} = \mathbb{R}^n$. We also denote by $\Omega = \mathcal{D} \times \mathbb{R}^n$.

We introduce the Gibbs measures μ and η on $\mathbb{R}^n$,

$$\mu(\mathrm{d}v) = e^{-|v|^2/2}\mathrm{d}v, \quad \eta(\mathrm{d}x) = e^{-U(x)}\mathrm{d}x,$$

and assume that $\exp(-U(x)) \in \mathrm{L}^1(\mathbb{R}^n, \mathrm{d}x)$. For $k \geq 1$, $\mathcal{D} \subset \mathbb{R}^n$, and $T > 0$, we define the function space $\mathrm{H}^k_{t,x,v} := \mathrm{H}^k_{t,x,v}((0,T) \times \mathcal{D} \times \mathbb{R}^n)$ as

$$\mathrm{H}^k_{t,x,v}((0,T) \times \mathcal{D} \times \mathbb{R}^n) := \big\{ u : \partial_t^i \nabla_x^j \nabla_v^{2l} u \in \mathrm{L}^2(\mathrm{L}^2_\eta(\mathrm{L}^2_\mu)), \\ \forall i,j,k \in \mathbb{N} \text{ s.t. } i + j + l \leq k \big\},$$

where $\mathrm{L}^2(\mathrm{L}^2_\eta(\mathrm{L}^2_\mu)) = \mathrm{L}^2(0,T; \mathrm{L}^2_\eta(\mathcal{D}; \mathrm{L}^2_\mu(\mathbb{R}^n)))$. When there is no ambiguity, we omit the dependence on the domain. We equip this space with the norm

$$\|u\|^2_{\mathrm{H}^k_{t,x,v}} := \sum_{i+j+l=0}^{k} \|\partial_t^i \nabla_x^j \nabla_v^{2l} u\|^2_{\mathrm{L}^2(\mathrm{L}^2_\eta(\mathrm{L}^2_\mu))}.$$

The function spaces $\mathrm{H}^k_{t,x} := \mathrm{H}^k_{t,x}((0,T)\times\mathcal{D})$ and $\mathrm{H}^k_{x,v} := \mathrm{H}^k_{x,v}(\mathcal{D}\times\mathbb{R}^d)$ are defined analogously.

In the case $D = \mathbb{T}^d$, we consider the problem in (3.1) and the ansatz

$$u_m(t,x,v) = \sum_{|\alpha|=0}^{m} c_m^\alpha(t,x)\Psi_\alpha(v), \tag{3.2}$$

where $\{\Psi_\alpha(v)\}$ are normalized Hermite polynomials in $\mathbb{R}^n$. For instance, in the case $n = 1$, the standard (non-normalized) Hermite polynomials are given by

$$h_k(v) := (-1)^k \mathrm{e}^{v^2/2}\frac{\mathrm{d}^k}{\mathrm{d}v^k}(\mathrm{e}^{-v^2/2}).$$

A crucial observation is that

$$\|u_m\|^2_{\mathrm{L}^2(\mathrm{H}^k_\eta(\mathrm{L}^2_\mu))} = \sum_{|\alpha|=0}^{m}\|c_m^\alpha\|^2_{\mathrm{L}^2(\mathrm{H}^k_\eta)},$$

where α are multi-indices. The coefficients $\{c_m^\alpha\}$ satisfy a hyperbolic system derived from the equation. In the case $n = 1$, this system is given by

$$\begin{array}{llll}
\partial_t c_m^0 & & -\sqrt{1}\partial_x^* c_m^1 & = f^0,\\
\partial_t c_m^1 & +c_m^1 & -\sqrt{2}\partial_x^* c_m^2 \quad +\sqrt{1}\partial_x c_m^0 & = f^1,\\
\vdots & \vdots & \vdots \quad\quad\quad \vdots & \vdots\\
\partial_t c_m^{m-1} & +(m-1)c_m^{m-1} & -\sqrt{m}\partial_x^* c_m^m \quad +\sqrt{m-1}\partial_x c_m^{m-2} & = f^{m-1},\\
\partial_t c_m^m & +mc_m^m & \quad\quad +\sqrt{m}\partial_x c_m^{m-1} & = f^m,
\end{array}$$

where $\partial_x^* = -\partial_x + \partial_x U$.

The main result in this paper is the following Galerkin approximation theorem for the case $D = \mathbb{T}^d$.

Theorem 3.1.1 *Let $\mathcal{D} = \mathbb{T}^d$, $T > 0$, and let $k \in \mathbb{Z}$ be such that $k \geq 2$. If we assume that*

$$\begin{cases}
U(x) \in \mathrm{C}^\infty(\mathbb{T}^d),\ g \in \mathrm{H}^k_{x,v}(\mathbb{T}^d\times\mathbb{R}^d),\ f \in \mathrm{H}^k_{t,x,v}((0,T)\times\mathbb{T}^d\times\mathbb{R}^d).\\
\nabla_x^{k+1} g \in \mathrm{L}^2_\eta(\mathbb{T}^d;\mathrm{L}^2_\mu(\mathbb{R}^d)),\\
\partial_t^i\nabla_x^j f \in \mathrm{L}^2((0,T);\mathrm{L}^2_\eta(\mathbb{T}^d;\mathrm{L}^2_\mu(\mathbb{R}^d))),\ \textit{whenever } i+j = k+1,
\end{cases} \tag{3.3}$$

then there exists, for each $m \in \mathbb{Z}_+$, a unique function

$$u_m \in \mathrm{H}^k_{t,x,v}((0,T)\times\mathbb{T}^d\times\mathbb{R}^d)$$

of the form (3.2) *with $\{c_m^\alpha\}$ satisfying the hyperbolic system. Also, there exists a constant $C = C(d,T,\|\nabla_x^2 U\|_\infty,\ldots,\|\nabla_x^{k+1}U\|_\infty,k) \geq 1$, such that*

$$\|u_m\|^2_{\mathrm{H}^k_{t,x,v}} \leq C(\|g\|^2_{\mathrm{H}^k_{x,v}} + \|f\|^2_{\mathrm{H}^k_{t,x,v}}).$$

Furthermore, after passing to a subsequence,

$$u_m \rightharpoonup u \in \mathrm{H}^k_{t,x,v}((0,T)\times\mathbb{T}^d\times\mathbb{R}^d),$$

where u is the unique weak solution to (3.1)*, and for all* $0 \le l \le k-2$*,*

$$\|u-u_m\|^2_{\mathrm{H}^l_{t,x,v}} \le C\left(\frac{1}{(1+m)^{(k-l-2)+1/2}}\right)^2 \left(\|g\|^2_{\mathrm{H}^k_{x,v}} + \|f\|^2_{\mathrm{H}^k_{t,x,v}}\right), \quad (3.4)$$

where $C = C(d,T,\|\nabla_x^2 U\|_\infty,\dots,\|\nabla_x^{k+1}U\|_\infty,k) \ge 1$.

The key idea in the proof is that by substituting the Hermite expansion ansatz, we reduce the problem to a finite-dimensional hyperbolic system, as demonstrated above. This approach allows us to establish the existence of Galerkin approximations u_m and use energy estimates to control the approximation error. A crucial energy estimate is derived from the hyperbolic system via a bootstrapping argument, which constitutes a main contribution of the paper. As a result, we obtain an error estimate that depends on the regularity of the initial data.

Beyond the case of bounded torus, we also prove that under certain conditions on the potential $U(x)$, such as being quadratic at infinity, a sequence of bounded periodic problems can be used to approximate the solution to (3.1) when $\mathcal{D} = \mathbb{R}^n$. Another key idea in the energy estimates here is to use the logarithmic Sobolev inequality for the Gibbs measure $\eta(\mathrm{d}v)\mu(\mathrm{d}x)$ and an entropy-type inequality from [40].

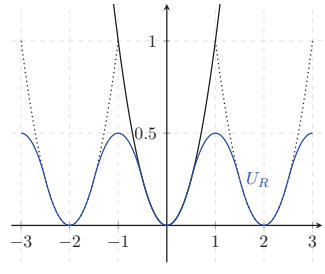

Figure 3.1. Illustration of the potential $U(x)$ and the truncated periodic potentials $U_R(x)$.

3.2 Paper II

In this paper, we study the well-posedness of the Kramers-Fokker-Planck equation in bounded domains using weak solutions and Perron's method. The operator under consideration is the Kramers-Fokker-Planck operator in divergence form with rough coefficients, given by

$$\mathscr{K}u := \nabla_v\cdot(\mathbf{A}\nabla_v u) + \boldsymbol{b}\cdot\nabla_v u + v\cdot\nabla_x u, \quad (3.5)$$

where the coefficients satisfy the condition $\nabla_v \cdot \boldsymbol{b} \geq 0$. For simplicity, we consider a prototype operator of the form

$$\mathscr{K} u = \Delta_v u + v \cdot \nabla_x u. \tag{3.6}$$

We begin by establishing the existence of weak solutions in bounded product domains $\mathcal{D} = \mathcal{U} \times \mathcal{V} \subset \mathbb{R}^n \times \mathbb{R}^n$. Building on this, we then construct Perron solutions in bounded domains of arbitrary shape and investigate the role of barriers in the analysis.

Weak solutions

The domain $\mathcal{D} = \mathcal{U} \times \mathcal{V}$ has two boundary components: the *v-boundary* $\partial^v \mathcal{D} = \mathcal{U} \times \partial \mathcal{V}$ and the *x-boundary* $\partial^x \mathcal{D} = \partial \mathcal{U} \times \mathcal{V}$.

The x-boundary is further divided into three parts: the *outgoing boundary* $\partial^x_+ \mathcal{D} = \{(x,v) \in \partial^x \mathcal{D} : v \cdot \boldsymbol{n}_x > 0\}$, the *tangential boundary* $\partial^x_0 \mathcal{D} = \{(x,v) \in \partial^x \mathcal{D} : v \cdot \boldsymbol{n}_x = 0\}$, and the *incoming boundary* $\partial^x_- \mathcal{D} = \{(x,v) \in \partial^x \mathcal{D} : v \cdot \boldsymbol{n}_x < 0\}$, where $\boldsymbol{n}_x$ is the unit outward normal of $\mathcal{U}$.

The *hypoelliptic boundary* of $\mathcal{D}$ is defined as $\partial_{\text{hyp}} \mathcal{D} := \overline{\partial^x_+ \mathcal{D} \cup \partial^x_0 \mathcal{D} \cup \partial^v \mathcal{D}}$.

Let $\Omega \subset \mathbb{R}^{2n}$ be a domain. The standard Sobolev space is denoted by $\mathrm{H}^1(\Omega)$, with the subspace

$$\mathrm{H}^1_0(\Omega) := \{f \in \mathrm{H}^1(\Omega) : \mathrm{Tr}(f) = 0\},$$

where $\mathrm{Tr} : \mathrm{H}^1(\Omega) \to \mathrm{L}^2(\partial\Omega)$ is the standard trace operator. The dual of $\mathrm{H}^1_0(\Omega)$ is denoted by $\mathrm{H}^{-1}(\Omega)$. We further define

$$\mathrm{H}^1_{\partial^v}(\mathcal{D}) := \mathrm{H}^1(\mathcal{D}) \cap \mathrm{L}^2(\mathcal{U}; \mathrm{H}^1_0(\mathcal{V})).$$

For $u \in \mathrm{L}^2(\mathcal{U}; \mathrm{H}^1(\mathcal{V}))$, the trace $\mathrm{Tr}_v(u) \in \mathrm{L}^2(\partial^v \mathcal{D})$ is given by

$$(\mathrm{Tr}_v(u))(x,v) := \mathrm{Tr}(u(x,\cdot)), \quad (x,v) \in \partial^v \mathcal{D}.$$

The *hypoelliptic space* associated with (3.6) is

$$\mathrm{H}^1_{\text{hyp}}(\mathcal{D}) := \left\{ u \in \mathrm{L}^2(\mathcal{U}; \mathrm{H}^1(\mathcal{V})) \text{ s.t. } v \cdot \nabla_x u \in \mathrm{L}^2(\mathcal{U}; \mathrm{H}^{-1}(\mathcal{V})) \right\},$$

where $\mathrm{H}^{-1}(\mathcal{V})$ is the dual of $\mathrm{H}^1_0(\mathcal{V})$, equipped with the norm

$$\|u\|_{\mathrm{H}^1_{\text{hyp}}} := \left(\|u\|^2_{\mathrm{L}^2(\mathcal{U};\mathrm{H}^1(\mathcal{V}))} + \|v \cdot \nabla_x u\|^2_{\mathrm{L}^2(\mathcal{U};\mathrm{H}^{-1}(\mathcal{V}))} \right)^{1/2}.$$

We also define the subspace

$$\mathrm{H}^1_{\text{hyp},\partial^v}(\mathcal{D}) := \left\{ u \in \mathrm{H}^1_{\text{hyp}}(\mathcal{D}) \text{ s.t. } \mathrm{Tr}_v(u) = 0 \text{ on } \partial^v \mathcal{D} \right\}.$$

It is observed in [64] that the classical H^1 trace can be extended to $\mathrm{H}^1_{\text{hyp}}$ in a weak sense:

Lemma 3.2.1 (Weak trace) *Let $\mathcal{D} \subset \mathbb{R}^{2n}$ with sufficiently regular boundaries, then there is a linear operator*

$$\mathrm{tr}_x : \mathrm{H}^1_{\mathrm{hyp}}(\mathcal{D}) \to \mathrm{L}^2_{\mathrm{loc}}(\partial^x\mathcal{D}, |v \cdot \boldsymbol{n}_x|^2),$$

such that, for $\varphi \in \mathrm{C}^1(\overline{\mathcal{D}})$, $\varphi|_{\partial^v\mathcal{D}} = 0$, we have the following integration by parts formula:

$$\langle v \cdot \nabla_x u,\, \varphi\rangle_{\mathrm{L}^2(\mathrm{H}^{-1}),\mathrm{L}^2(\mathrm{H}^1_0)} = \int_{\partial^x\mathcal{D}} v \cdot \boldsymbol{n}_x \mathrm{tr}_x(u)\varphi \,\mathrm{d}S - \iint_{\mathcal{D}} u(v \cdot \nabla_x\varphi)\,\mathrm{d}x\mathrm{d}v.$$

This integration by parts formula is crucial for the well-posedness theory. Next, the definition of weak solutions.

Definition 3.2.2 *Let Ω be an open set in $\mathbb{R}^{2n}$. We say that a function u is a weak solution to $-\mathscr{L}u = f$ in Ω if, for any product domain $\mathcal{D} = \mathcal{U} \times \mathcal{V} \subset \Omega$, we have $u \in \mathrm{L}^2(\mathcal{U}; \mathrm{H}^1(\mathcal{V}))$, $f \in \mathrm{L}^2(\mathcal{U}; \mathrm{H}^{-1}(\mathcal{V}))$ and for all $\varphi \in \mathrm{C}^1_c(\mathcal{D})$ the following holds*

$$\int_\Omega (\nabla_v u \cdot \nabla_v\varphi + u(v \cdot \nabla_x\varphi))\,\mathrm{d}z = \int_\Omega f\varphi\,\mathrm{d}z.$$

An important consequence of the weak trace is the following renormalization result.

Lemma 3.2.3 *Let $\mathcal{D} \subset \mathbb{R}^{2n}$ with sufficiently regular boundaries. Let $f \in \mathrm{L}^2(\mathcal{U}; \mathrm{H}^{-1}(\mathcal{V}))$, if $u \in \mathrm{H}^1_{\mathrm{hyp}}(\mathcal{D})$ is a weak solution of* (3.6)*, then for any $\Phi : \mathbb{R} \to \mathbb{R}$ such that $\Phi' \in \mathrm{W}^{1,\infty}(\mathbb{R})$ and either*

- *$\Phi'(\mathrm{Tr}_v(u)) = 0$ a.e. on $\partial^v\mathcal{D}$ and $\varphi \in \mathrm{C}^1(\overline{\mathcal{D}})$, or*
- *$\varphi \in \mathrm{C}^1(\overline{\mathcal{D}})$ with $\varphi|_{\partial^v\mathcal{D}} = 0$,*

we have

$$\int_{\mathcal{D}} \nabla_v\Phi(u) \cdot \nabla_v\varphi + \varphi\Phi''(u)|\nabla_v u|^2 + \Phi(u)v \cdot \nabla_x\varphi\,\mathrm{d}z$$
$$= \int_{\mathcal{D}} f\Phi'(u)\varphi\,\mathrm{d}z + \int_{\partial^x\mathcal{D}} v \cdot \mathbf{n}_x\Phi(\mathrm{tr}_x(u))\varphi\,\sigma(\mathrm{d}z),$$

where $\sigma(\mathrm{d}z)$ is the surface measure on $\partial^x\mathcal{D}$.

Consider the following Dirichlet problem,

$$\begin{cases} -\mathscr{K}u(x,v) = f(x,v), & \text{in } \mathcal{D}, \\ u(x,v) = g(x,v), & \text{on } \partial_{\mathrm{hyp}}\mathcal{D}, \end{cases} \tag{3.7}$$

where $f(x,v)$ and $g(x,v)$ are functions and

$$g(x,v) := \begin{cases} g_1(x,v), & \text{on } \partial^v\mathcal{D}, \\ g_2(x,v), & \text{on } \partial^x\mathcal{D}. \end{cases}$$

The first part of the paper is concerned with the following well-posedness result.

Theorem 3.2.4 *Let $\mathcal{D} \subset \mathbb{R}^{2n}$ with sufficiently regular boundaries. If $g_1 \in \mathrm{H}^1(\mathcal{D})$, $g_2 \in \mathrm{L}^2(\partial^x\mathcal{D}, |v \cdot \boldsymbol{n}_x|)$ and $f \in \mathrm{L}^2(\mathcal{U}; \mathrm{H}^{-1}(\mathcal{V}))$, then there exists a weak solution u to* (3.7) *such that $u - g_1 \in \mathrm{H}^1_{\mathrm{hyp},\partial^v}(\mathcal{D})$ and $\mathrm{tr}_x(u) = g_2$ on $\partial^x_+\mathcal{D}$.*

Moreover, let u, w be functions such that $-\mathscr{L}u \leq -\mathscr{L}w$ in $\mathcal{D}$, $tr_x(u) \leq \mathrm{tr}_x(\mathcal{D})$ on $\partial^x\mathcal{D}$ and $\mathrm{Tr}_v(u) \leq \mathrm{Tr}_v(w)$ on $\partial^v\mathcal{D}$. Then $u \leq w$ a.e. in $\mathcal{D}$.

Finally, if $g \in \mathrm{C}(\partial\mathcal{D})$, then we have that $u \in \mathrm{C}(\overline{\mathcal{D}})$.

We present two distinct approaches to establishing the existence of weak solutions: one based on the Lions-Lax-Milgram theorem and the other on the vanishing viscosity method. A key subtle point lies in that these methods handle $\mathcal{D}$ under slightly different regularity assumptions.

The comparison principle follows from careful energy estimates using Theorem 3.2.3. Notably, one cannot directly choose $\Phi(u) = (u-k)^2_+$ in Theorem 3.2.3; instead, it is necessary to use a sequence of smooth approximations of the function $(u-k)^2_+$, as illustrated in Fig. 3.2.

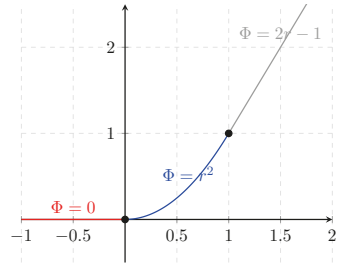

Figure 3.2. Illustration of the function $\Phi(r) = (r-k)^2_+$ and its smooth approximations.

Interior Hölder regularity result follows directly from the Harnack inequality, while boundary regularity is obtained by combining Theorem 3.2.3 with results already established in the literature.

Perron's method

We consider the problem, for a bounded domain $\Omega \subset \mathbb{R}^{2n}$,

$$\begin{cases} -\mathscr{K}u = 0, & \text{in } \Omega \subset \mathbb{R}^{2n}, \\ u = g, & \text{on } \partial\Omega, \end{cases} \tag{3.8}$$

where $\mathscr{K}$ is as in (3.6).

We begin with defining K-harmonic and K-superharmonic functions.

Definition 3.2.5 *Let $\Omega \subset \mathbb{R}^{2n}$ be an open set. We say that a function u is* K-harmonic *if u is a continuous weak solution to $-\mathscr{K}u = 0$ in Ω.*

From Theorem 3.2.4 we have the following: Let $\mathcal{D} = \mathcal{U} \times \mathcal{V} \subset \mathbb{R}^{2n}$ with sufficiently regular boundaries, for $h \in C(\partial\mathcal{D})$, there exists a unique K-harmonic function $u \in C(\overline{\mathcal{D}})$ such that $u = h$ on $\partial_{\mathrm{hyp}}\mathcal{D}$, satisfying $\sup_{\overline{\mathcal{D}}} u \leq \sup_{\partial_{\mathrm{hyp}}\mathcal{D}} h$.

The above is enough to guarantee that the notion of K-superharmonic functions is well-defined using the comparison principle.

Definition 3.2.6 *Let $\Omega \subset \mathbb{R}^{2n}$ be an open set. A function $u : \Omega \to (-\infty, \infty]$ is called K-*superharmonic *if*

1. *u is lower semi-continuous,*
2. *u is finite in a dense subset of Ω,*
3. *u satisfies the comparison principle on each product domain $\mathcal{D} \subset\subset \Omega$ as above: That is, for any* K*-harmonic function $h \in C(\overline{\mathcal{D}})$ in $\mathcal{D}$, such that $h \leq u$ on $\partial_{\text{hyp}}\mathcal{D}$, then $h \leq u$ in the whole $\mathcal{D}$.*

*A function w is K-*subharmonic *if $-w$ is* K*-superharmonic.*

With all these notions we can define the upper and lower Perron solutions respectively.

Definition 3.2.7 *Let $\Omega \subset \mathbb{R}^{2n}$ be an open set and let $g : \partial\Omega \to \mathbb{R}$ be a bounded function. A function u is said to belong to the* upper class $\mathcal{U}_g$ *with respect to* (3.8) *if u is* K*-superharmonic in Ω, bounded below and satisfies*

$$\liminf_{z \to z_0} u(z) \geq g(z_0)$$

at each point $z_0 \in \partial\Omega$. The upper Perron solution $\overline{H}_g$ *of* (3.8) *is defined by*

$$\overline{H}_g(z) = \inf\{u(z) : u \in \mathcal{U}_g\}.$$

The lower Perron solution $\underline{H}_g$ *is defined analogously. Finally, we say that g is* resolutive *with respect to Ω if $\overline{H}_g = \underline{H}_g$ in Ω.*

Theorem 3.2.8 *Let $\Omega \subset \mathbb{R}^{2n}$ be a bounded open set. If the boundary data $g : \partial\Omega \to \mathbb{R}$ is bounded, then the upper and lower Perron solutions $\underline{H}_g$ and $\overline{H}_g$ to the equation* (3.8) *are* K*-harmonic.*

Moreover, if the boundary data g is continuous on $\partial\Omega$ then g is resolutive for the problem (3.8)*.*

The proof of this theorem is based on the comparison principle and the definition of the Perron solutions, as discussed in Section 2.2.

We remark that, for the resolutivity of the general operator (3.5), we assume **A** is differentiable with bounded derivatives, leading to a Kramers-Fokker-Planck operator in non-divergence form with rough coefficients.

Finally, we discuss the regularity of the boundary of the domain Ω.

Definition 3.2.9 *We call a boundary point $z_0 \in \partial\Omega$ K-regular if*

$$\lim_{z \to z_0} \overline{H}_g(z) = g(z_0)$$

whenever $g : \partial\Omega \to \mathbb{R}$ is continuous.

We provide examples of regular boundaries of certain sets that are velocity symmetric (i.e., symmetric in v), as illustrated in Fig. 2.3. In particular, we show that a ball $B_r(0) \subset \mathbb{R}^{2n}$ is everywhere regular. For regions where $\boldsymbol{n}_{v,\Omega} \neq 0$, an explicit barrier can be constructed following [51]. The barrier on the singular part of the boundary is given by the solution to the Dirichlet problem in a cylindrical domain that locally touches the singular boundary. This result is new for the stationary Kramers-Fokker-Planck equation with general coefficients. In the case of constant coefficients, within the ultraparabolic setting, a Wiener-Landis criterion is available, see [35].

3.3 Paper III

In this paper, we establish a boundedness theory for the degenerate Fokker-Planck-Kolmogorov equation in divergence form within bounded product domains. The core of the theory relies on a renormalization formula up to the boundary, which is based on the weak trace results from Paper II. Additionally, we derive an embedding estimate for solutions using the explicit kernel and apply De Giorgi iterations to obtain boundedness results.

Let $d \geq 2$ and $0 < n_0 \leq d$ be positive integers, and let $z = (t, x) \in \mathbb{R} \times \mathbb{R}^d$ denote a point. We consider the following degenerate Fokker-Planck-Kolmogorov equation in divergence form:

$$\partial_t u - \langle \mathbf{B}x, Du \rangle = \mathscr{L}u + g + D_i f^i, \tag{3.9}$$

where, unless otherwise stated, the summation for indices i, j is from $d-n_0+1$ to d. Here, $\mathbf{B}$ is a constant real matrix, $D = (D_1, \ldots, D_d)$ is the gradient, and $\langle \cdot, \cdot \rangle$ is the Euclidean inner product. The functions $g(t, x)$ and $f^i(t, x)$ are measurable. The operator $\mathscr{L}$ is defined as:

$$\mathscr{L}u = D_i \left(a^{ij}(t, x) D_j u + b^i(t, x) u \right) + c^i(t, x) D_i u + d(t, x) u$$

where a^{ij}, b^i, c, d are measurable functions, and $a^{ij} = a^{ji}$. We may also denote the Kolmogorov operator by $\mathscr{K} := \mathscr{T} - \mathscr{L}$, where $\mathscr{T} = \partial_t - \langle \mathbf{B}x, D \rangle$ is the transport part.

Throughout this paper, we assume the following: for some positive constants λ and Λ:

$$|a^{ij}(t, x)| \leq \Lambda \text{ and } a^{ij}(t, x)\xi_i \xi_j \geq \lambda |\xi|^2, \quad \forall (t, x) \in \mathbb{R}^{d+1},\ \xi \in \mathbb{R}^{n_0}. \tag{3.10}$$

Additionally, we assume that the matrix $\mathbf{B}$ has the following form:

$$\mathbf{B} = \begin{bmatrix} \mathbf{0} & \mathbf{B}_\kappa & & & \\ & \mathbf{0} & \mathbf{B}_2 & & \\ & & \ddots & \ddots & \\ & & & \ddots & \mathbf{B}_1 \\ & & & & \mathbf{0} \end{bmatrix} \tag{3.11}$$

where $\mathbf{0}$ is the zero matrix and each $\mathbf{B}_j$ is an $n_j \times n_{j-1}$ constant matrix of rank n_j such that

$$n_0 + n_1 + \cdots + n_\kappa = d, \text{ and } n_0 \geq n_1 \geq \cdots \geq n_\kappa \geq 1.$$

The homogeneous dimension is defined as:

$$Q := n_0 + 3n_1 + \cdots + (2\kappa + 1)n_\kappa.$$

The necessary function space preliminaries are similar to those in Paper II. Throughout this paper, we consider $\Omega = \mathcal{U} \times \mathcal{V}$, where $\mathcal{U} \subset \mathbb{R}^{d-n_0}$ and $\mathcal{V} \subset \mathbb{R}^{n_0}$ are bounded domains with sufficiently regular boundaries. The time cylinder is defined as $\Omega_T := \Omega \times (0, T)$.

We introduce a function space $\mathrm{H}^1_{\text{kin}}(\Omega_T)$ analogous to the hypoelliptic space as follows:

$$\mathrm{H}^1_{\text{kin}}(\Omega_T) := \{u(t,x) \in \mathrm{L}^2(\mathcal{U}_T; \mathrm{H}^1(\mathcal{V})) \text{ such that } \partial_t u - \langle \mathbf{B}x, Du \rangle \in \mathrm{L}^2(\mathcal{U}_T; \mathrm{H}^{-1}(\mathcal{V}))\},$$

where $\mathcal{U}_T := (0, T) \times \mathcal{U}$. This space is equipped with the norm

$$\|u\|^2_{\mathrm{H}^1_{\text{kin}}} = \|u\|^2_{\mathrm{L}^2(\mathcal{U}_T; \mathrm{H}^1(\mathcal{V}))} + \|\partial_t u - \langle \mathbf{B}x, Du \rangle\|^2_{\mathrm{L}^2(\mathcal{U}_T; \mathrm{H}^{-1}(\mathcal{V}))},$$

where $\mathrm{H}^{-1}(\mathcal{V})$ is the dual space of $\mathrm{H}^1_0(\mathcal{V})$.

Next, we define a space $\mathrm{V}^0_{\text{kin}}(\Omega_T)$ suitable for time cylinders in the following way:

$$\mathrm{V}^0_{\text{kin}}(\Omega_T) := \mathrm{H}^1_{\text{kin}}(\Omega_T) \cap \mathrm{C}([0, T]; \mathrm{L}^2(\Omega))$$

equipped with the norm

$$\|u\|_{\mathrm{V}^0_{\text{kin}}(\Omega_T)} := \|u\|_{\mathrm{H}^1_{\text{kin}}(\Omega_T)} + \sup_{0<t<T} \|u(\cdot, t)\|_{\mathrm{L}^2(\Omega)}.$$

The weak trace operator in Paper II is extended to the time cylinder as follows:

$$\mathrm{tr}_{\Gamma_\mathrm{K}} : \mathrm{V}^0_{\text{kin}}(\Omega_T) \to \mathrm{L}^2_{\text{loc}}(\Gamma_\mathrm{K}, |\langle \mathbf{B}x, \boldsymbol{n}_x \rangle|^2),$$

where $\Gamma_\mathrm{K} := (0, T) \times (\partial\mathcal{U}) \times \mathcal{V}$, $\boldsymbol{n}_x$ denotes the outward unit normal for Ω, and $|\langle \mathbf{B}x, \boldsymbol{n}_x \rangle|^2$ is the weight and a similar renormalization formula as in Theorem 3.2.3 follows. Throughout this paper, we denote $u|_{\Gamma_\mathrm{K}} = \mathrm{tr}_{\Gamma_\mathrm{K}}(u)$, implying that boundary values are understood in the weak trace sense.

Definition 3.3.1 *A function $u \in \mathrm{V}^0_{\text{kin}}(\Omega_T)$ is called a subsolution to* (3.9) *if for all $v \in \mathrm{C}^1_c(\Omega_T)$ with $v \geq 0$, the inequality*

$$\iint_{\Omega_T} \left(-uv_t + u\langle \mathbf{B}x, Dv \rangle + \mathscr{L}[u, v] - gv + f^i D_i v\right) \mathrm{d}x\mathrm{d}t \leq 0$$

holds, where

$$\mathscr{L}[u, v] := a^{ij} D_j u D_i v + b^i u D_i v - c^i (D_i u) v - duv.$$

The separation of the boundary follows a similar approach to Paper II, where we define $\Gamma_{\mathrm{P}} := (\partial\mathcal{V} \times \mathcal{U} \times [0,T]) \cup (\Omega \times \{0\})$, $\Gamma_{\mathrm{K}}^{+} := \{(x,t) \in \Gamma_{\mathrm{K}} \mid \langle \mathbf{B}x, \boldsymbol{n}_x \rangle \geq 0\}$, and $\Gamma_{\mathrm{K}}^{-} := \{(x,t) \in \Gamma_{\mathrm{K}} \mid \langle \mathbf{B}x, \boldsymbol{n}_x \rangle < 0\}$.

Then we denote the maximum value of a function u on the boundary as

$$M := \sup_{\Gamma_{\mathrm{K}}^{+} \cup \Gamma_{\mathrm{P}}} u_{+},$$

where $u_{+} := \max\{u, 0\}$ denotes the positive part of a function. Define

$$p_0 := \frac{Q+2}{Q} \text{ and } q_0 := \frac{Q+2}{2}$$

the optimal embedding constants. Assume $\tilde{q} > q_0$ is fixed, and consider either:

$$c^i \in \mathrm{L}^{2\tilde{q}}(\Omega_T),\ d, g \in \mathrm{L}^{\tilde{q}}(\Omega_T),\ b^i, f^i \in \mathrm{L}^{2\tilde{q}}(\Omega_T); \tag{3.12}$$

or

$$c^i \in \mathrm{L}^{Q+2}(\Omega_T),\ d, g \in \mathrm{L}^{\tilde{q}}(\Omega_T),\ b^i, f^i = 0. \tag{3.13}$$

Moreover, $\|\cdot\|_{p,\Omega_T}$ denotes the $\mathrm{L}^p(\Omega_T)$ norm.

The main result of the paper is the following boundedness theorem.

Theorem 3.3.2 *Let $\Omega = \mathcal{U} \times \mathcal{V} \subset \mathbb{R}^{d-n_0} \times \mathbb{R}^{n_0}$ be a bounded product domain with sufficiently regular boundaries, and let the assumptions* (3.10) *and* (3.11) *hold. If $u \in \mathrm{V}^0_{\mathrm{kin}}(\Omega_T)$ is a subsolution to* (3.9)*, and either* (3.12) *or* (3.13) *holds, then there exists a constant $C = C(\lambda, \Lambda, \mathrm{Data})$ such that*

$$\sup_{\Gamma_{\mathrm{K}}^{-}} u,\ \sup_{\Omega_T} u \leq M + C \max\{1, M, \|(u-M)_{+}\|_{2,\Omega_T}\}, \tag{3.14}$$

and

$$\sup_{\Gamma_{\mathrm{K}}^{-}} u,\ \sup_{\Omega_T} u \leq (1+C)\max\{1, M\}. \tag{3.15}$$

The first inequality provides an L^2–L^∞ estimate which can be made local and is crucial for the Harnack inequality. The second inequality is a classical boundedness estimate up to the boundary. Therefore, the current theory is a canonical extension of the classical boundedness theory for uniformly parabolic equations. We believe that the result is optimal for the reason that for $n_0 = d \geq 2$, the Kolmogorov equation becomes the uniformly parabolic equation, and the optimal exponent $(d+2)/2$ is recovered under (3.13), see [38, 23].

An immediate corollary is the weak maximum principle.

Theorem 3.3.3 *Let Ω be defined as in Theorem 3.3.2, and let the assumptions* (3.10) *and* (3.11) *hold. Assume that $b^i, c^i \in \mathrm{L}^{Q+2}(\Omega_T)$, $d \in \mathrm{L}^2(\Omega_T)$, $g, f^i = 0$, and for all $v \in \mathrm{C}^1_c(\overline{\Omega_T})$ with $v \geq 0$ and $v = 0$ on $(0,T) \times \mathcal{U} \times (\partial\mathcal{V})$ that*

$$\int_{\Omega_T} dv - b^i D_i v \,\mathrm{d}z \leq 0.$$

If $u \in \mathrm{V}^0_{\mathrm{kin}}(\Omega_T)$ *is a subsolution to* (3.9)*, then*

$$\sup_{\Gamma_{\mathrm{K}}^-} u, \ \sup_{\Omega_T} u \leq M.$$

The key to the De Giorgi iteration is an embedding estimate. Since no a priori embedding inequality is known for the space $\mathrm{V}^2_{\mathrm{kin}}$, we instead use an estimate based on an explicit kernel. It is well-known that the constant coefficient equation admits a fundamental solution K such that

$$\partial_t K - \langle \mathbf{B}x, DK \rangle - \mathscr{L}_0 u = \delta,$$

where $\mathscr{L}_0 := D_i^2 u$ is the Laplace operator, and δ is the Dirac delta function. It can be shown by explicit computation that the kernel K is integrable in $\mathrm{L}^p((0,T) \times \mathbb{R}^d)$ for any $1 \leq p < (Q+2)/Q$, and the gradient $D_i K$ is integrable in $\mathrm{L}^p((0,T) \times \mathbb{R}^d)$ for any $1 \leq p < (Q+2)/(Q+1)$.

Using the linearity of the operators $\mathscr{L}$ and $\mathscr{L}_0$, we can add and subtract to find that a subsolution can be written as

$$\begin{aligned} \iint_{\Omega_T} (-uv_t + u\langle \mathbf{B}x, Dv \rangle + \mathscr{L}_0[u,v]) \, \mathrm{d}x\mathrm{d}t \\ \leq \iint_{\Omega_T} \left(\mathscr{L}_0[u,v] - \mathscr{L}[u,v] + gv - f^i D_i v \right) \mathrm{d}x\mathrm{d}t, \end{aligned}$$

which makes it possible to use the kernel K to derive an embedding estimate for subsolutions. It is worth mentioning that when performing energy estimates, one must carefully handle the renormalization term Φ as discussed in Fig. 3.2, which makes the energy estimates more involved.

3.4 Paper IV

In this paper, we present two main results. First, we establish a variational formula for the capacity in (2.1), which includes both degenerate and non-degenerate cases. Second, we provide quantitative estimates for the capacity in the uniformly elliptic case.

Recall the operator $\mathscr{L}$ in (2.1) is given by

$$\mathscr{L} f(z) = \varepsilon \mathrm{e}^{W/\varepsilon} \nabla \cdot \left[\mathrm{e}^{-W/\varepsilon} \mathbf{A} \nabla f \right] + (\boldsymbol{b} + \mathbf{B}z) \cdot \nabla f,$$

where $\mathbf{A}(z)$ is a matrix, $\boldsymbol{b}(z) : \mathbb{R}^d \to \mathbb{R}^d$ is a vector field of the form

$$\mathbf{A}(z) = \begin{bmatrix} \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \widehat{\mathbf{A}}(z) \end{bmatrix}, \quad \boldsymbol{b}(z) = \begin{bmatrix} \mathbf{0} \\ \widehat{\boldsymbol{b}}(z) \end{bmatrix}, \tag{3.16}$$

such that $\widehat{\mathbf{A}}(z)$ is a bounded $n_0 \times n_0$ matrix for fixed $n_0 \leq d$, Hölder continuous of order α for some $0 < \alpha < 1$, and uniformly elliptic, i.e., $\xi^\intercal \widehat{\mathbf{A}} \xi \geq \lambda |\xi|^2$ for

some $\lambda > 0$; $\widehat{\boldsymbol{b}}(z)$ is a continuous vector field of length n_0 and is locally bounded. Moreover, $\mathbf{B}$ is a constant matrix that satisfies the same assumptions as in (3.11).

Additionally, we denote $\widehat{\nabla} := (\partial_{d-n_0+1}, \ldots, \partial_d)$, the gradient of the last n_0 coordinates.

The following weighted divergence-free assumption is imposed on $\boldsymbol{b}(z)$ and $\mathbf{B}$:

$$\nabla \cdot \left((\boldsymbol{b}(z) + \mathbf{B}z)\, e^{-W(z)/\varepsilon} \right) = 0, \quad \text{for all } z \in \mathbb{R}^d. \tag{3.17}$$

It is easy to observe that if $\nabla \cdot (\boldsymbol{b} + \mathbf{B}z) = 0$ and $(\boldsymbol{b} + \mathbf{B}z) \cdot \nabla W(z) = 0$, then (3.17) holds. As a consequence, there exists a stationary measure ρ_ε for the operator $\mathscr{L}_\varepsilon$ given by

$$\rho_\varepsilon(z) = \frac{1}{Z_\varepsilon} e^{-W(z)/\varepsilon} \mathrm{d}z, \quad \text{where } Z_\varepsilon = \int_{\mathbb{R}^d} e^{-W(z)/\varepsilon} \mathrm{d}z,$$

which satisfies $\int_{\mathbb{R}^d} \mathscr{L}_\varepsilon f \, \rho_\varepsilon(\mathrm{d}z) = 0$ for all f such that the integral is well-defined. We denote by $\mathscr{L}_\varepsilon^\dagger$ the adjoint operator of $\mathscr{L}_\varepsilon$ with respect to the measure ρ_ε as in Section 2.4, given by

$$\mathscr{L}_\varepsilon^\dagger f(z) = \varepsilon e^{W/\varepsilon} \nabla \cdot \left[e^{-W/\varepsilon} \mathbf{A}^\intercal \nabla f \right] - (\boldsymbol{b} + \mathbf{B}z) \cdot \nabla f. \tag{3.18}$$

Let $\mathcal{D} \subset \mathbb{R}^d$ be an open set. We denote by $\mathrm{H}^1(\mathcal{D}) = \mathrm{W}^{1,2}(\mathcal{D})$ the Sobolev space of functions with respect to the measure ρ_ε, and $\mathrm{H}_0^1(\mathcal{D})$ be the closure of $\mathrm{C}_c^\infty(\mathcal{D})$ in $\mathrm{H}^1(\mathcal{D})$. We also introduce the function space $\widehat{\mathrm{H}}^1(\mathcal{D}) := \{f(z) \in \mathrm{L}^2(\mathcal{D}, \rho_\varepsilon) : \widehat{\nabla} f \in \mathrm{L}^2(\mathcal{D}, \rho_\varepsilon)\}$, equipped with the weighted norm, and $\widehat{\mathrm{H}}_0^1$ as the closure of $\mathrm{C}_c^\infty(\mathcal{D})$. In addition, we define $\widehat{\mathrm{H}}^1_{\mathrm{div}}(\mathcal{D}) := \{\boldsymbol{l} \in \left(\mathrm{L}^2(\mathcal{D}, \rho_\varepsilon)\right)^{n_0} : \widehat{\nabla} \cdot \boldsymbol{l} \in \mathrm{L}^2(\mathcal{D}, \rho_\varepsilon)\}$.

Next, we define the following inner product associated with the operator (2.1), for $\boldsymbol{u}, \boldsymbol{v} \in \left(\mathrm{L}^2(\mathbb{R}^d, \rho_\varepsilon)\right)^{n_0}$ vector-valued functions,

$$\langle \boldsymbol{u}, \boldsymbol{v} \rangle_{\mathcal{D}} := \int_{\mathcal{D}} \boldsymbol{u}^\intercal \widehat{\mathbf{S}}^{-1} \boldsymbol{v} \, \rho_\varepsilon(\mathrm{d}z),$$

where $\widehat{\mathbf{S}} = \left(\widehat{\mathbf{A}} + \widehat{\mathbf{A}}^\intercal\right)/2$ is the symmetric part of $\widehat{\mathbf{A}}$.

The capacity is defined via the equilibrium potential h_ε, as discussed in Section 2.4.

Definition 3.4.1 *Let the equilibrium potential $h_\varepsilon \in \mathrm{C}(\overline{\Omega}) \cap \mathrm{C}^1(\Omega)$ be such that $\widehat{\nabla} h_\varepsilon$ is sufficiently regular up the boundary, so that the following integral makes sense. The capacity between $\mathcal{A}$ and $\mathcal{B}$ is defined as*

$$\mathrm{cap}_\varepsilon(\mathcal{A}, \mathcal{B}) := \frac{1}{Z_\varepsilon} \int_{\partial \mathcal{A}} \left(\varepsilon (\widehat{\mathbf{A}} \widehat{\nabla} h_\varepsilon) \cdot \widehat{\boldsymbol{n}}_\Omega + (h_\varepsilon - 1)(\boldsymbol{b} + \mathbf{B}z) \cdot \boldsymbol{n}_\Omega \right) e^{-W/\varepsilon} \sigma(\mathrm{d}z),$$

where $\boldsymbol{n}_\Omega$ is the unit outward normal vector to Ω, $\widehat{\boldsymbol{n}}_\Omega$ corresponds to the part related to the last n_0 coordinates, and $\sigma(\mathrm{d}z)$ is the surface measure on $\partial\mathcal{A}$.

We define a functional for $f \in \mathrm{H}^1_0(\overline{\mathcal{B}}^{\complement})$ as

$$J_{\varepsilon,\overline{\mathcal{B}}^{\complement}}[f] := \inf_{\boldsymbol{g}\in \mathrm{V}_{\varepsilon,f}(\overline{\mathcal{B}}^{\complement})} \varepsilon\Big\langle \widehat{\mathbf{A}}\widehat{\nabla} f - \boldsymbol{g}, \widehat{\mathbf{A}}\widehat{\nabla} f - \boldsymbol{g}\Big\rangle_{\overline{\mathcal{B}}^{\complement}}, \tag{3.19}$$

where

$$\mathrm{V}_{\varepsilon,f}(\overline{\mathcal{B}}^{\complement}) := \Big\{ \boldsymbol{g} \in \widehat{\mathrm{H}}^1_{\mathrm{div}}(\overline{\mathcal{B}}^{\complement}) : \varepsilon e^{W/\varepsilon}\widehat{\nabla}\cdot\Big(e^{-W/\varepsilon}\boldsymbol{g}\Big) = -(\boldsymbol{b}+\mathbf{B}z)\cdot\nabla f\Big\}.$$

Our first result relates the variational formulation to the capacity.

Theorem 3.4.2 *Let the equilibrium potential $h_\varepsilon \in \mathrm{C}(\overline{\Omega})\cap\mathrm{C}^1(\Omega)$ be such that $\widehat{\nabla} h_\varepsilon$ is sufficiently regular up the boundary. Then the following relation holds:*

$$\varepsilon\Big\langle \widehat{\mathbf{S}}\widehat{\nabla} h_\varepsilon, \widehat{\mathbf{S}}\widehat{\nabla} h_\varepsilon\Big\rangle_{\Omega} \le \mathrm{cap}_\varepsilon(\mathcal{A},\mathcal{B}) \le \inf_{f\in\mathrm{H}^1_0(\overline{\mathcal{B}}^{\complement}),\, f\ge\mathcal{X}_{\mathcal{A}}} J_{\overline{\mathcal{B}}^{\complement}}[f]. \tag{3.20}$$

In particular, if $\mathbf{A}$ is non-degenerate and Hölder continuous, then the equality holds, and the minimizer for the variational formulation is attained at $\bar{f} = (h_\varepsilon + h_\varepsilon^\dagger)/2$, and $\bar{g} = (\mathbf{A}\nabla h_\varepsilon - \mathbf{A}^{\mathsf{T}}\nabla h_\varepsilon^\dagger)/2$.

Quantitative Capacity Estimates for the Elliptic Case

In the second part of the paper, we apply the aforementioned variational formula to the uniformly elliptic case to justify its feasibility. Let $\mathbf{A}$ be uniformly elliptic and Hölder continuous of order α. Set $\mathbf{B} = 0$ and $\boldsymbol{b}(z) = -\boldsymbol{l}(z)$ for some smooth vector field $\boldsymbol{l} : \mathbb{R}^d \to \mathbb{R}^d$ such that (3.17) holds. The operator becomes

$$\mathscr{L}_\varepsilon f = \varepsilon e^{W/\varepsilon}\nabla\cdot\Big[e^{-W/\varepsilon}\left(\mathbf{A}\nabla f - f\boldsymbol{l}/\varepsilon\right)\Big].$$

First, we derive a rough estimate for the capacity by following the outline presented in Section 2.4. Our approach, which fully utilizes the variational formula, is novel.

Theorem 3.4.3 *Let $\mathcal{A}$ be an open set, and let $x \in \overline{\mathcal{A}}^{\complement}$. Denote by $z^\star = z^\star(x,\overline{\mathcal{A}})$ a point such that*

$$W(z^\star) = \inf_{\gamma(t)}\sup_{t\in[0,1]} W(\gamma(t)),$$

where the infimum is taken over all continuous paths from x to $\overline{\mathcal{A}}$. Then, there exist constants c_1 and c_2 independent of ε such that

$$\mathrm{cap}_\varepsilon(B_\varepsilon(z),\overline{\mathcal{A}}) \ge c_1\varepsilon^d e^{-W(z^\star)/\varepsilon}, \text{ and } \mathrm{cap}_\varepsilon(B_\varepsilon(z),\overline{\mathcal{A}}) \le c_2\varepsilon^{d-4}e^{-W(z^\star)/\varepsilon}.$$

Next, for a sharp estimate, we assume that W is a smooth Morse function with two local minima $\boldsymbol{m}_0$ and $\boldsymbol{m}_1$, and without loss of generality, let 0 be the unique saddle point. We set $\mathcal{A} = B_\varepsilon(\boldsymbol{m}_1)$ and $\mathcal{B} = B_\varepsilon(\boldsymbol{m}_0)$. Moreover, we write $H := W(0)$, $h_0 := W(\boldsymbol{m}_0)$, and $h_1 := W(\boldsymbol{m}_1)$, and denote by $\mathbf{H}_z = \nabla^2 W(z)$ the Hessian of W and $\mathbf{L}_z = \nabla \boldsymbol{l}(z)$ the Jacobian of $\boldsymbol{l}$ at the point $z \in \mathbb{R}^d$. It can be shown that the matrix

$$\mathbf{H}_0\mathbf{A}_0 + \mathbf{L}_0^\intercal$$

has a unique negative eigenvalue, denoted by $-\mu$.

Theorem 3.4.4 *The following capacity estimate holds:*

$$\operatorname{cap}_\varepsilon(B_\varepsilon(\boldsymbol{m}_1), B_\varepsilon(\boldsymbol{m}_0)) = (1 + o_\varepsilon(1))\frac{1}{Z_\varepsilon}\frac{(2\pi\varepsilon)^{d/2}}{2\pi}\frac{\mu}{\sqrt{-\det(\mathbf{H}_0)}}e^{-H/\varepsilon}.$$

The proof of this result is outlined in Section 2.5.

Remark 3.4.5 *It can be readily checked that under the underdamped Langevin process, let $\boldsymbol{l}(z) = \boldsymbol{b}(z) + \mathbf{B}z$, and denote by $V(v) = |v|^2/2$, the structure from Theorem 3.4.4 suggests that we can set*

$$\mathbf{H}_0 = \begin{bmatrix} \nabla_x^2 U(0) & \mathbf{0} \\ \mathbf{0} & \nabla_v^2 V(0) \end{bmatrix}, \mathbf{A}_0 = \begin{bmatrix} \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \gamma \mathbf{I}_n \end{bmatrix},$$

and

$$\mathbf{L}_0^\intercal = \begin{bmatrix} \mathbf{0} & \nabla_x^2 U(0) \\ -\nabla_v^2 V(0) & \mathbf{0} \end{bmatrix}.$$

Hence,

$$\mathbf{H}_0\mathbf{A}_0 + \mathbf{L}_0^\intercal = \begin{bmatrix} \mathbf{0} & -\nabla_x^2 U(0) \\ -\mathbf{I}_n & \gamma \mathbf{I}_n \end{bmatrix}$$

The unique negative eigenvalue of $\mathbf{H}_0\mathbf{A}_0 + \mathbf{L}_0^\intercal$ is

$$-\mu = \frac{1}{2}(\gamma - \sqrt{\gamma^2 + 4\lambda_1})$$

where $-\lambda_1$ is the unique negative eigenvalue of $\nabla_x^2 U(0)$. This recovers the Eyring-Kramers formula in both [14, 43].

4. Summary in Swedish

Denna avhandling utvecklar de matematiska grunderna för Fokker-Planck-Kolmogorov-operatorer (FPK) inom en potentialteoretisk ram för att studera metastabilitet hos stokastiska processer. Metastabilitet är ett centralt fenomen inom flera områden, såsom statistisk fysik, kemi, sannolikhetsteori och maskininlärning, där system med flera jämviktstillstånd fastnar i ett av dem under en lägre tid följt av sällsynta men plötsliga övergångar till ett annat.

I många tillämpningar uppstår metastabilitet i stokastiska processer som styrs av ett potentialfält, vilket leder till övergångar mellan lokala minima i potentiallandskapet. Langevin-dynamik, en välstuderad stokastisk process, ger en naturlig modell för sådana fenomen, inklusive kemiska reaktioner, kvanttunnelering och träning av neurala nätverk.

Fokker-Planck-Kolmogorov-operatorn fungerar som generator för en stokastisk process och beskriver utvecklingen av dess sannolikhetsdensitet. Att förstå dess egenskaper särskilt svaga lösningar, variationsformuleringar och randbeteende är avgörande för att utveckla en rigorös potentialteori för metastabilitet.

Denna avhandling presenterar en omfattande analys av svaga lösningar till (möjligen degenererade) Fokker-Planck-Kolmogorov-ekvationer och utvecklar variationsmetoder för dessa problem.

I artikel I tillämpar vi Galerkinmetoden för att konstruera lösningar till Cauchyproblemet i begränsade periodiska områden och i hela rummet. Vi analyserar även hur lösningens regularitet beror på begynnelsevärden och koefficienter. Galerkinapproximationen ger ett alternativt bevis för välställdhet och kan även användas som en numerisk metod för att simulera utvecklingen av sannolikhetsdensiteten hos den underliggande stokastiska processen.

I artikel II visar vi att den stationära Kolmogorovoperatorn har en väldefinierad svag lösning i begränsade produktområden och visar existens av Perronlösningar till den stationära ekvationen i begränsade områden. Dessutom undersöker vi randregularitet och ger exempel som illustrerar lösningarnas beteende nära randen.

Artikel III är inte direkt kopplat till metastabilitet men bidrar till regularitetsteorin för Kolmogorovoperatorer. Vi visar att lösningar till den tidsberoende Kolmogorovekvationen i begränsade produktområden är globalt begränsade med hjälp av De Giorgi-iteration, där operatorn innehåller lägre ordningens divergenstermer. Som en konsekvens kan Harnacks olikhet generaliseras till en större klass av Kolmogorovoperatorer.

I artikel IV definierar vi kapacitet genom jämviktspotentialen och föreslår en variationsformulering av kapacitet för den stationära Kolmogorovoperatorn. Vidare visar vi Eyring-Kramers-formeln i fallet med en likformigt elliptisk men ickereversibel operator, vilket visar att vår definition är genomförbar.

Denna avhandling lägger en solid matematisk grund för potentialteori för stationära Kolmogorovoperatorer och introducerar nya verktyg för att studera variationsformulerade kapaciteter. Dessutom öppnar våra resultat nya forskningsriktningar inom Kolmogorovekvationer och potentialteori, inklusive studier av högre ordningens randregularitet och hinderproblem.

5. Acknowledgements

First and foremost, I would like to express my deepest gratitude to my supervisor, Prof. Benny Avelin, for his invaluable guidance, support, and encouragement throughout this research project. This journey has been a roller coaster, but we navigated it together. I would also like to thank my co-supervisor, Prof. Kaj Nyström, for his advice during the early stages of this project.

Furthermore, I am grateful to my colleagues and friends at the Department of Mathematics, Uppsala University. It has been a wonderful and peaceful place to work and study.

Most importantly, my wife, Hongyu, has been my greatest source of strength and support, and her presence has meant everything to me.

Finally, feel free to add your name here for your own convenience, as I have chosen not to include any.

Bibliography

[1] D. Albritton, S. Armstrong, J.-C. Mourrat, and M. Novack. Variational methods for the kinetic Fokker–Planck equation. *Anal. PDE*, 17(6):1953–2010, 2024.

[2] F. Anceschi and S. Polidoro. A survey on the classical theory for Kolmogorov equation. *Matematiche (Catania)*, 75(1):221–258, 2020.

[3] F. Anceschi and A. Rebucci. On the fundamental solution for degenerate Kolmogorov equations with rough coefficients. *Journal of Elliptic and Parabolic Equations*, 9:63–92, 2023.

[4] F. Anceschi and A. Rebucci. On the obstacle problem associated to the Kolmogorov-Fokker-Planck operator with rough coefficients. *Annali di Matematica*, 203:1979–1995, 2024.

[5] D. N. Arnold, G. David, M. Filoche, D. Jerison, and S. Mayboroda. Localization of eigenfunctions via an effective potential. *Comm. Partial Differential Equations*, 44(11):1186–1216, 2019.

[6] S. Arrhenius. Über die Reaktionsgeschwindigkeit bei der Inversion von Rohrzucker durch Säuren. *Zeitschrift für Physikalische Chemie*, 4U(1):226–248, 1889.

[7] P. Auscher, C. Imbert, and L. Niebel. Weak solutions to Kolmogorov-Fokker-Planck equations: regularity, existence and uniqueness. *arXiv preprint*, 2024.

[8] B. Avelin and V. Julin. A note on the capacity estimate in metastability for generic configurations. *Manuscripta Math.*, 175(1-2):245–266, 2024.

[9] B. Avelin, V. Julin, and L. Viitasaari. Geometric characterization of the Eyring-Kramers formula. *Comm. Math. Phys.*, 404(1):401–437, 2023.

[10] D. Bakry, I. Gentil, and M. Ledoux. *Analysis and geometry of Markov diffusion operators*, volume 348 of *Grundlehren der mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences]*. Springer, Cham, 2014.

[11] N. Berglund. Kramers' law: validity, derivations and generalisations. *Markov Process. Related Fields*, 19(3):459–490, 2013.

[12] N. Berglund. An Eyring-Kramers law for slowly oscillating bistable diffusions. *Probab. Math. Phys.*, 2(4):685–743, 2021.

[13] N. Berglund and B. Gentz. The Eyring-Kramers law for potentials with nonquadratic saddles. *Markov Process. Related Fields*, 16(3):549–598, 2010.

[14] J.-F. Bony, D. L. Peutrec, and L. Michel. Eyring–Kramers law for Fokker–Planck type differential operators. *Journal of the European Mathematical Society*, 2024. Published online first.

[15] A. Bovier, M. Eckhoff, V. Gayrard, and M. Klein. Metastability in reversible diffusion processes. I. Sharp asymptotics for capacities and exit times. *J. Eur. Math. Soc. (JEMS)*, 6(4):399–424, 2004.

[16] A. Bovier, V. Gayrard, and M. Klein. Metastability in reversible diffusion processes. II. Precise asymptotics for small eigenvalues. *J. Eur. Math. Soc. (JEMS)*, 7(1):69–99, 2005.

[17] D. Bowman. Hypoelliptic regularization in the obstacle problem for the Kolmogorov operator. *arXiv preprint*, 2025.
[18] M. Bramanti. On the proof of Hörmander's hypoellipticity theorem. *Matematiche (Catania)*, 75(1):3–26, 2020.
[19] J. A. Carrillo. Global weak solutions for the initial-boundary-value problems to the Vlasov-Poisson-Fokker-Planck system. *Math. Methods Appl. Sci.*, 21(10):907–938, 1998.
[20] A.-L. Dalibard, F. Marbach, and J. Rax. Linear and nonlinear parabolic forward-backward problems. *arXiv preprint*, 2024.
[21] M. Di Francesco and A. Pascucci. On a class of degenerate parabolic equations of Kolmogorov type. *AMRX Appl. Math. Res. Express*, (3):77–116, 2005.
[22] M. Di Francesco and S. Polidoro. Schauder estimates, Harnack inequality and Gaussian lower bound for Kolmogorov-type operators in non-divergence form. *Adv. Differential Equations*, 11(11):1261–1320, 2006.
[23] E. DiBenedetto. *Degenerate parabolic equations*. Universitext. Springer-Verlag, New York, 1993.
[24] M. D. Donsker and S. R. S. Varadhan. On the principal eigenvalue of second-order elliptic differential operators. *Comm. Pure Appl. Math.*, 29(6):595–621, 1976.
[25] L. C. Evans. *Partial differential equations*, volume 19 of *Graduate Studies in Mathematics*. American Mathematical Society, Providence, RI, second edition, 2010.
[26] H. Eyring. The activated complex in chemical reactions. *The Journal of Chemical Physics*, 3(2):107–115, 02 1935.
[27] M. Filoche and S. Mayboroda. Universal mechanism for Anderson and weak localization. *Proc. Natl. Acad. Sci. USA*, 109(37):14761–14766, 2012.
[28] D. Gilbarg and N. S. Trudinger. *Elliptic partial differential equations of second order*. Classics in Mathematics. Springer-Verlag, Berlin, 2001. Reprint of the 1998 edition.
[29] F. Golse, C. Imbert, C. Mouhot, and A. F. Vasseur. Harnack inequality for kinetic Fokker-Planck equations with rough coefficients and application to the Landau equation. *Ann. Sc. Norm. Super. Pisa Cl. Sci. (5)*, 19(1):253–295, 2019.
[30] J. Guerand and C. Imbert. Log-transform and the weak Harnack inequality for kinetic Fokker-Planck equations. *J. Inst. Math. Jussieu*, 22(6):2749–2774, 2023.
[31] J. Guerand and C. Mouhot. Quantitative De Giorgi methods in kinetic theory. *J. Éc. polytech. Math.*, 9:1159–1181, 2022.
[32] B. Helffer and F. Nier. *Hypoelliptic estimates and spectral theory for Fokker-Planck operators and Witten Laplacians*, volume 1862 of *Lecture Notes in Mathematics*. Springer-Verlag, Berlin, 2005.
[33] F. Hérau, M. Hitrik, and J. Sjöstrand. Tunnel effect for Kramers-Fokker-Planck type operators. *Ann. Henri Poincaré*, 9(2):209–274, 2008.
[34] L. Hörmander. Hypoelliptic second order differential equations. *Acta Math.*, 119:147–171, 1967.
[35] A. E. Kogoj, E. Lanconelli, and G. Tralli. Wiener-Landis criterion for Kolmogorov-type operators. *Discrete Contin. Dyn. Syst.*, 38(5):2467–2485, 2018.
[36] A. Kolmogorov. Zufällige Bewegungen (zur Theorie der Brownschen

Bewegung). *Ann. of Math. (2)*, 35(1):116–117, 1934.
[37] H. A. Kramers. Brownian motion in a field of force and the diffusion model of chemical reactions. *Physica*, 7:284–304, 1940.
[38] O. A. Ladyženskaja, V. A. Solonnikov, and N. N. Uralćeva. *Linear and quasilinear equations of parabolic type*, volume Vol. 23 of *Translations of Mathematical Monographs*. American Mathematical Society, Providence, RI, 1968. Translated from the Russian by S. Smith.
[39] C. Landim, M. Mariani, and I. Seo. Dirichlet's and Thomson's principles for non-selfadjoint elliptic operators with application to non-reversible metastable diffusion processes. *Arch. Ration. Mech. Anal.*, 231(2):887–938, 2019.
[40] M. Ledoux. Concentration of measure and logarithmic Sobolev inequalities. In *Séminaire de Probabilités, XXXIII*, volume 1709 of *Lecture Notes in Math.*, pages 120–216. Springer, Berlin, 1999.
[41] J. Lee and I. Seo. Non-reversible metastable diffusions with Gibbs invariant measure I: Eyring-Kramers formula. *Probab. Theory Related Fields*, 182(3-4):849–903, 2022.
[42] S. Lee, M. Ramil, and I. Seo. Asymptotic stability and cut-off phenomenon for the underdamped langevin dynamics. *arXiv preprint*, 2023.
[43] S. Lee, M. Ramil, and I. Seo. Eyring-kramers law for the underdamped langevin process. *arXiv preprint*, 2025.
[44] T. Lelièvre, M. Ramil, and J. Reygner. A probabilistic study of the kinetic Fokker-Planck equation in cylindrical domains. *J. Evol. Equ.*, 22(2):Paper No. 38, 74, 2022.
[45] H. Li, Z. Xu, G. Taylor, C. Studer, and T. Goldstein. Visualizing the loss landscape of neural nets. In S. Bengio, H. Wallach, H. Larochelle, K. Grauman, N. Cesa-Bianchi, and R. Garnett, editors, *Advances in Neural Information Processing Systems*, volume 31. Curran Associates, Inc., 2018.
[46] Q. Li, C. Tai, and W. E. Stochastic modified equations and dynamics of stochastic gradient algorithms i: Mathematical foundations. *Journal of Machine Learning Research*, 20(40):1–47, 2019.
[47] M. Litsgård and K. Nyström. The Dirichlet problem for Kolmogorov-Fokker-Planck type equations with rough coefficients. *J. Funct. Anal.*, 281(10):Paper No. 109226, 39, 2021.
[48] M. Litsgård. *On parabolic equations of Kolmogorov-Fokker-Planck type*. Phd dissertation, Department of Mathematics, Uppsala University, 2023.
[49] B. Liu. Spurious local minima are common for deep neural networks with piecewise linear activations. *IEEE Transactions on Neural Networks and Learning Systems*, 35(4):5382–5394, 2024.
[50] J. Lu and S. Steinerberger. A variation on the Donsker-Varadhan inequality for the principal eigenvalue. *Proc. A.*, 473(2204):20160877, 6, 2017.
[51] M. Manfredini. The Dirichlet problem for a class of ultraparabolic equations. *Adv. Differential Equations*, 2(5):831–866, 1997.
[52] D. Marcondes and A. Simonis. Metastable financial markets. *arXiv preprint*, 2023.
[53] V. I. Mel′nikov. The Kramers problem: fifty years of development. *Phys. Rep.*, 209(1-2):71, 1991.
[54] S. Mischler. On the trace problem for solutions of the Vlasov equation. *Comm.*

Partial Differential Equations, 25(7-8):1415–1443, 2000.
[55] S. Mischler. Kinetic equations with Maxwell boundary conditions. *Ann. Sci. Éc. Norm. Supér. (4)*, 43(5):719–760, 2010.
[56] B. Nectoux. Mean exit time for the overdamped Langevin process: the case with critical points on the boundary. *Comm. Partial Differential Equations*, 46(9):1789–1829, 2021.
[57] B. Øksendal. *Stochastic differential equations*. Universitext. Springer-Verlag, Berlin, sixth edition, 2003. An introduction with applications.
[58] R. G. Pinsky. *Positive harmonic functions and diffusion*, volume 45 of *Cambridge Studies in Advanced Mathematics*. Cambridge University Press, Cambridge, 1995.
[59] L. Rey-Bellet. Ergodic properties of Markov processes. In *Open quantum systems. II*, volume 1881 of *Lecture Notes in Math.*, pages 1–39. Springer, Berlin, 2006.
[60] H. Risken. *The Fokker-Planck equation*, volume 18 of *Springer Series in Synergetics*. Springer-Verlag, Berlin, 1984. Methods of solution and applications.
[61] I. Safran and O. Shamir. Spurious local minima are common in two-layer ReLU neural networks. In J. Dy and A. Krause, editors, *Proceedings of the 35th International Conference on Machine Learning*, volume 80 of *Proceedings of Machine Learning Research*, pages 4433–4441. PMLR, 10–15 Jul 2018.
[62] C. F. Sanchez, P. Gabriel, and S. Mischler. On the Krein-Rutman theorem and beyond. *arXiv preprint*, 2024.
[63] B. Shi, W. J. Su, and M. I. Jordan. On learning rates and Schrödinger operators. *J. Mach. Learn. Res.*, 24:Paper No. [379], 53, 2023.
[64] L. Silvestre. Hölder estimates for kinetic Fokker-Planck equations up to the boundary. *Ars Inven. Anal.*, pages Paper No. 6, 29, 2022.
[65] Y. Zhu. Regularity of kinetic Fokker–Planck equations in bounded domains. *Annales Henri Lebesgue*, 7:1323–1366, 2024.

UPPSALA DISSERTATIONS IN MATHEMATICS
Dissertations at the Department of Mathematics
Uppsala University

---

1. Torbjörn Lundh: *Kleinian groups and thin sets at the boundary*. 1995.
2. Jan Rudander: O*n the first occurrence of a given pattern in a semi-Markov process*. 1996.
3. Alexander Shumakovitch: *Strangeness and invariants of finite degree*. 1996
4. Stefan Halvarsson: *Duality in convexity theory applied to growth problems in complex analysis*. 1996.
5. Stefan Svanberg: *Random walk in random environment and mixing*. 1997.
6. Jerk Matero: *Nonlinear elliptic problems with boundary blow-up*. 1997.
7. Jens Blanck: *Computability on topological spaces by effective domain representations*. 1997.
8. Jonas Avelin: *Differential calculus for multifunctions and nonsmooth functions*. 1997.
9. Hans Garmo: *Random railways and cycles in random regular graphs*. 1998.
10. Vladimir Tchernov: *Arnold-type invariants of curves and wave fronts on surfaces*. 1998.
11. Warwick Tucker: *The Lorenz attractor exists*. 1998.
12. Tobias Ekholm: *Immersions and their self intersections*. 1998.
13. Håkan Ljung: *Semi Markov chain Monte Carlo*. 1999.
14. Pontus Andersson: *Random tournaments and random circuits*. 1999.
15. Anders Andersson: *Lindström quantifiers and higher-order notions on finite structures*. 1999.
16. Marko Djordjević: *Stability theory in finite variable logic*. 2000.
17. Andreas Strömbergsson: *Studies in the analytic and spectral theory of automorphic forms*. 2001.
18. Olof-Petter Östlund: *Invariants of knot diagrams and diagrammatic knot invariants*. 2001.
19. Stefan Israelsson: *Asymptotics of random matrices and matrix valued processes*. 2001.
20. Yacin Ameur: *Interpolation of Hilbert spaces*. 2001.
21. Björn Ivarsson: *Regularity and boundary behavior of solutions to complex Monge-Ampère equations*. 2002.
22. Lars Larsson-Cohn: *Gaussian structures and orthogonal polynomials*. 2002.
23. Sara Maad: *Critical point theory with applications to semilinear problems without compactness*. 2002.
24. Staffan Rodhe: *Matematikens utveckling i Sverige fram till 1731*. 2002.
25. Thomas Ernst: *A new method for q-calculus*. 2002.
26. Leo Larsson: *Carlson type inequalities and their applications*. 2003.

27. Tsehaye K. Araaya: *The symmetric Meixner-Pollaczek polynomials*. 2003.
28. Jörgen Olsén: *Stochastic modeling and simulation of the TCP protocol*. 2003.
29. Gustaf Strandell: *Linear and non-linear deformations of stochastic processes*. 2003.
30. Jonas Eliasson: *Ultrasheaves*. 2003.
31. Magnus Jacobsson: *Khovanov homology and link cobordisms*. 2003.
32. Peter Sunehag: *Interpolation of subcouples, new results and applications*. 2003.
33. Raimundas Gaigalas: *A non-Gaussian limit process with long-range dependence*. 2004.
34. Robert Parviainen: *Connectivity Properties of Archimedean and Laves Lattices*. 2004.
35. Qi Guo: *Minkowski Measure of Asymmetry and Minkowski Distance for Convex Bodies*. 2004.
36. Kibret Negussie Sigstam: *Optimization and Estimation of Solutions of Riccati Equations*. 2004.
37. Maciej Mroczkowski: *Projective Links and Their Invariants*. 2004.
38. Erik Ekström: *Selected Problems in Financial Mathematics*. 2004.
39. Fredrik Strömberg: *Computational Aspects of Maass Waveforms*. 2005.
40. Ingrid Lönnstedt: *Empirical Bayes Methods for DNA Microarray Data*. 2005.
41. Tomas Edlund: *Pluripolar sets and pluripolar hulls*. 2005.
42. Göran Hamrin: *Effective Domains and Admissible Domain Representations*. 2005.
43. Ola Weistrand: *Global Shape Description of Digital Objects*. 2005.
44. Kidane Asrat Ghebreamlak: *Analysis of Algorithms for Combinatorial Auctions and Related Problems*. 2005.
45. Jonatan Eriksson: *On the pricing equations of some path-dependent options*. 2006.
46. Björn Selander: *Arithmetic of three-point covers*. 2007.
47. Anders Pelander: *A Study of Smooth Functions and Differential Equations on Fractals*. 2007.
48. Anders Frisk: *On Stratified Algebras and Lie Superalgebras*. 2007.
49. Johan Prytz: *Speaking of Geometry*. 2007.
50. Fredrik Dahlgren: *Effective Distribution Theory*. 2007.
51. Helen Avelin: *Computations of automorphic functions on Fuchsian groups*. 2007.
52. Alice Lesser: *Optimal and Hereditarily Optimal Realizations of Metric Spaces*. 2007.
53. Johanna Pejlare: *On Axioms and Images in the History of Mathematics*. 2007.
54. Erik Melin: *Digital Geometry and Khalimsky Spaces*. 2008.
55. Bodil Svennblad: *On Estimating Topology and Divergence Times in Phylogenetics*. 2008.

56. Martin Herschend: *On the Clebsch-Gordan problem for quiver representations*. 2008.
57. Pierre Bäcklund: *Studies on boundary values of eigenfunctions on spaces of constant negative curvature*. 2008.
58. Kristi Kuljus: *Rank Estimation in Elliptical Models*. 2008.
59. Johan Kåhrström: *Tensor products on Category O and Kostant's problem*. 2008.
60. Johan G. *Granström: Reference and Computation in Intuitionistic Type Theory*. 2008.
61. Henrik Wanntorp: *Optimal Stopping and Model Robustness in Mathematical Finance*. 2008.
62. Erik Darpö: *Problems in the classification theory of non-associative simple algebras*. 2009.
63. Niclas Petersson: *The Maximum Displacement for Linear Probing Hashing*. 2009.
64. Kajsa Bråting: *Studies in the Conceptual Development of Mathematical Analysis*. 2009.
65. Hania Uscka-Wehlou: *Digital lines, Sturmian words, and continued fractions*. 2009.
66. Tomas Johnson: *Computer-aided computation of Abelian integrals and robust normal forms*. 2009.
67. Cecilia Holmgren: *Split Trees, Cuttings and Explosions*. 2010.
68. Anders Södergren: *Asymptotic Problems on Homogeneous Spaces*. 2010.
69. Henrik Renlund: *Recursive Methods in Urn Models and First-Passage Percolation*. 2011.
70. Mattias Enstedt: *Selected Topics in Partial Differential Equations*. 2011.
71. Anton Hedin: *Contributions to Pointfree Topology and Apartness Spaces*. 2011.
72. Johan Björklund: *Knots and Surfaces in Real Algebraic and Contact Geometry*. 2011.
73. Saeid Amiri: *On the Application of the Bootstrap*. 2011.
74. Olov Wilander: *On Constructive Sets and Partial Structures*. 2011.
75. Fredrik Jonsson: *Self-Normalized Sums and Directional Conclusions*. 2011.
76. Oswald Fogelklou: *Computer-Assisted Proofs and Other Methods for Problems Regarding Nonlinear Differential Equations*. 2012.
77. Georgios Dimitroglou Rizell: *Surgeries on Legendrian Submanifolds*. 2012.
78. Granovskiy, Boris: *Modeling Collective DecisionMaking in Animal Groups*. 2012
79. Bazarganzadeh, Mahmoudreza: *Free Boundary Problems of Obstacle Type, a Numerical and Theoretical Study*. 2012
80. Qi Ma: *Reinforcement in Biology*. 2012.
81. Isac Hedén: *Ga-actions on Complex Affine Threefolds*. 2013.
82. Daniel Strömbom: *Attraction Based Models of Collective Motion*. 2013.
83. Bing Lu: Calibration: *Optimality and Financial Mathematics*. 2013.

84. Jimmy Kungsman: *Resonance of Dirac Operators*. 2014.
85. Måns Thulin: *On Confidence Intervals and Two-Sided Hypothesis Testing*. 2014.
86. Maik Görgens: *Gaussian Bridges – Modeling and Inference*. 2014.
87. Marcus Olofsson: *Optimal Switching Problems and Related Equations*. 2015.
88. Seidon Alsaody: *A Categorical Study of Composition Algebras via Group Actions and Triality*. 2015.
89. Håkan Persson: *Studies of the Boundary Behaviour of Functions Related to Partial Differential Equations and Several Complex Variables*. 2015.
90. Djalal Mirmohades: *N-complexes and Categorification*. 2015.
91. Shyam Ranganathan: *Non-linear dynamic modelling for panel data in the social sciences*. 2015.
92. Cecilia Karlsson: *Orienting Moduli Spaces of Flow Trees for Symplectic Field Theory*. 2016.
93. Olow Sande: *Boundary Estimates for Solutions to Parabolic Equations*. 2016.
94. Jonathan Nilsson: *Simple Modules over Lie Algebras*. 2016.
95. Marta Leniec: *Information and Default Risk in Financial Valuation*. 2016.
96. Arianna Bottinelli: *Modelling collective movement and transport network formation in living systems*. 2016.
97. Katja Gabrysch: *On Directed Random Graphs and Greedy Walks on Point Processes*. 2016.
98. Martin Vannestål: *Optimal timing decisions in financial markets*. 2017.
99. Natalia Zabzina: *Mathematical modelling approach to collective decision-making*. 2017.
100. Hannah Dyrssen: *Valuation and Optimal Strategies in Markets Experiencing Shocks*. 2017.
101. Juozas Vaicenavicius: *Optimal Sequential Decisions in Hidden-State Models*. 2017.
102. Love Forsberg: *Semigroups, multisemigroups and representations*. 2017.
103. Anna Belova: *Computational dynamics – real and complex*. 2017.
104. Ove Ahlman: *Limit Laws, Homogenizable Structures and Their Connections*. 2018.
105. Erik Thörnblad: *Degrees in Random Graphs and Tournament Limits*. 2018.
106. Yu Liu: Modelling *Evolution. From non-life, to life, to a variety of life*. 2018.
107. Samuel Charles Edwards: *Some applications of representation theory in homogeneous dynamics and automorphic functions*. 2018.
108. Jakob Zimmermann: *Classification of simple transitive 2-representations*. 2018.
109. Azza Alghamdi: *Approximation of pluricomplex Green functions – A probabilistic approach*. 2018.
110. Brendan Frisk Dubsky: *Structure and representations of certain classes of infinite-dimensional algebras*. 2018.
111. Björn R. H. Blomqvist: *Gaussian process models of social change*. 2018.

112. Konstantinos Tsougkas: *Combinatorial and analytical problems for fractals and their graph approximations*. 2019.
113. Yevgen Ryeznik: *Optimal adaptive designs and adaptive randomization techniques for clinical trials*. 2019.
114. Andrea Pasquali: *Constructions in higher-dimensional Auslander-Reiten theory*. 2019.
115. Daniah Tahir: *Multi-trait Branching Models with Applications to Species Evolution*. 2019.
116. Andreas Rocén: *7D supersymmetric Yang-Mills theory on toric and hypertoric manifolds*. 2019.
117. Filipe Mussini: *Selected Topics in Continuum Percolation. Phase Transitions, Cover Times and Random Fractals*. 2019.
118. Laertis Vaso: *Constructions of n-cluster tilting subcategories using representation-directed algebras.* 2020.
119. Israelsson, Anders: *Regularity Properties of Oscillatory Integral Operators on Function Spaces.* 2020.
120. Asplund, Johan: *Chekanov-Eliashberg dg-algebras and partially wrapped Floer cohomology*. 2021.
121. Elin Persson Westin: *Stratified algebras and classification of tilting modules*. 2022.
122. Marcus Westerberg: *Prostate cancer incidence, treatment and mortality. Empirical longitudinal register-based studies and methods for handling missing data*. 2022.
123. Dan Lilja. *Properties of invariant sets in certain two-dimensional dynamical systems. Renormalization and beyond*. 2022.
124. Oleksandra Gasanova. *On properties of monomial ideals and algebras.* 2022.
125. Colin Desmarais. *Decorating trees grown in urns*. 2022
126. Malte Litsgård. *On parabolic equations of Kolmogorov-Fokker-Planck type*. 2023
127. Carmina Fjellström. *Selected Topics in Mathematical Modelling: Machine Learning and Tugs-of-War*. 2023.
128. Robert Moscrop. *Quivers, superconformal field theories and higher symmetries.* 2023.
129. Markus Thuresson. *Ext-algebras of standard modules over quasi-hereditary algebras*. 2023.
130. Aksel Bergfeldt. *Regularity of multilinear oscillatory integral operators and nonlinear dispersive partial differential equations*. 2023.
131. Fabian Burghart. *Building and Destroying Urns, Graphs, and Trees.* 2023
132. Yuqiong Wang. *Optimal stopping, incomplete information, and stochastic games.* 2023.
133. Helena Jonsson. *Structure and representations of bimodule categories*. 2023.

134. Christoffer Olsson. *Symmetry in Randomness. Additive Functionals and Symmetries of Random Trees and Tree-Like Graphs*. 2023.

135. Linnéa Gyllingberg. *The Art of Modelling Oscillations and Feedback across Biological Scales.* 2024.

136. Joel Dahne. *TComputer Assisted Studies in Fluid Mechanics and Spectral Geometry.* 2024.

137. Lukas Nakamura. *Skein-valued Gromov-Witten theory and Hofer geometry in contact manifolds*. 2024

138. Axel Husin. *Generalized local systems and the Maslov class of exact Lagrangians.* 2025

139. Rebekka Müller-Widmann. *Probabilistic Models of Genetic Variability in Sequence Evolution.* 2025

140. Mingyi Hou. *Behind the training dynamics of neural networks. Analysis of Fokker-Planck equations and the path to metastability*. 2025