

TOPICS IN SAMPLING SCHEMES BASED ON MARKOV PROCESSES

A THESIS PRESENTED FOR THE DEGREE OF
DOCTOR OF PHILOSOPHY OF IMPERIAL COLLEGE LONDON
AND THE
DIPLOMA OF IMPERIAL COLLEGE
BY
NIKOLAS NÜSKEN

DEPARTMENT OF MATHEMATICS
IMPERIAL COLLEGE
180 QUEEN'S GATE, LONDON SW7 2BZ

MAY 2018

I certify that this thesis is the product of my own work and that any ideas or quotations from the work of other people, published or otherwise, are properly acknowledged.

Nikolas Nüsken

COPYRIGHT

The copyright of this thesis rests with the author and is made available under a Creative Commons Attribution Non-Commercial No Derivatives licence. Researchers are free to copy, distribute or transmit the thesis on the condition that they attribute it, that they do not use it for commercial purposes and that they do not alter, transform or build upon it. For any reuse or redistribution, researchers must make clear to others the licence terms of this work.

Topics in sampling schemes based on Markov processes

ABSTRACT

In this thesis we consider several topics related to the construction of optimal Markovian dynamics in the context of sampling from high-dimensional probability distributions.

Firstly, we introduce and analyse Langevin samplers that consist of perturbations of the standard overdamped and underdamped Langevin dynamics. The perturbed dynamics is such that its invariant measure is the same as that of the unperturbed dynamics. We show that appropriate choices of the perturbations can lead to samplers that have improved properties, at least in terms of reducing the asymptotic variance. We present a detailed analysis of the new Langevin samplers for Gaussian target distributions. Our theoretical results are supported by numerical experiments with non-Gaussian target measures.

Secondly, we present a general framework for the analysis and development of ensemble based methods, encompassing both diffusion and piecewise deterministic Markov processes. For many performance criteria of interest, including the asymptotic variance, the task of finding efficient couplings can be phrased in terms of problems related to the theory of optimal transportation. We investigate general structural properties, proving a singularity theorem that has both geometric and probabilistic interpretations. Moreover, we show that those problems can often be solved approximately and support our findings with numerical experiments. Addressing the convergence to equilibrium of coupled processes we furthermore derive a modified Poincaré inequality.

Finally, under some conditions, we prove exponential ergodicity for the zigzag process using hypocoercivity techniques.

PUBLICATIONS AND COLLABORATIONS

The work presented in this thesis is based on the author's work in collaboration with his supervisor Professor Grigorios Pavliotis. Parts of the thesis have resulted from further collaborations with Claudio Bellani, Dr. Andrew Duncan and Julien Roussel. In particular:

- The results in Chapter 3 have been obtained together with Claudio Bellani on a visit to Professor Andreas Eberle at the University of Bonn, the support and hospitality of which are greatly acknowledged. A publication containing these results is currently in preparation.
- The contents of Chapters 4-8 have been developed together with Dr. Andrew Duncan and Professor Grigorios Pavliotis and were published in [42].
- Chapters 9-16 have resulted from the collaboration with Professor Grigorios Pavliotis. A publication containing these results is currently in preparation.
- The results in Chapter 17 have been obtained in collaboration with Julien Roussel on his visit to Imperial College London. A publication containing these results is currently in preparation (in joint work with Professor Christophe Andrieu and Dr. Alain Durmus).

ACKNOWLEDGEMENTS

First of all, I would like to express my deep gratitude to my supervisor Greg Pavliotis for his enthusiasm and optimism in every meeting, for his guidance and support, for enabling many collaborations, and for sharing his view on Markov processes and spectral theory with me. It has been a complete privilege working in his research group for the last three and a half years.

I am also very grateful for the opportunity to work with Andrew Duncan, the many stimulating discussions with him, and especially his extremely careful and critical reading of some of my proofs. My warmest thanks go to Claudio Bellani and Julien Roussel for their great ideas contributing to our joint research projects and to Maximilian Engel and Onur Teymur for their great friendship and extensive debates about mathematics and statistics in general. Additionally, I would like to thank Pedro Aceves, Andreas Eberle, Marco Francischello, Franca Hoffmann, Pierre Monmarché, Martin Rasmussen, Sebastian Reich, Gabriel Stoltz, Urbain Vaes, Martin Weidner and Raphael Zimmer for useful discussions.

This research was funded by the Department of Mathematics, Imperial College London, through an EPSRC Roth scholarship. I would like to furthermore acknowledge support from the Institut Henri Poincaré and the CERMICS research centre at the Ecole des Ponts in Paris, as well as from the University of Bonn. Some of the results in this thesis were obtained during academic visits to these institutions.

Finally, thanks are due to my family, in particular to my sister Johanna, my grandmothers, Doris, Dieter and Kirstin, and, most of all, to my parents, to whom I owe everything.

CONTENTS

LIST OF FIGURES	17
NOTATION AND BASIC CONVENTIONS	19
I Introduction	21
1 THE SAMPLING PROBLEM	23
1.1 Preliminaries on Markov processes	24
1.2 Comparison criteria	32
1.2.1 Convergence to equilibrium and spectral gap	32
1.2.2 Fluctuations and asymptotic variance	35
1.3 A selection of ergodic Markov processes	36
1.3.1 Diffusion processes	37
1.3.2 The zigzag process	43
1.3.3 MCMC with many particles: coupling	45
1.4 Outline of the thesis	46
II Nonreversibility	49
2 NONREVERSIBLE PERTURBATIONS AND SAMPLING – AN OVERVIEW	51
3 PERTURBED OVERDAMPED LANGEVIN DYNAMICS REVISITED	55
3.1 Exponential decay to equilibrium in Wasserstein distances	56
3.2 Discussion	60
4 PERTURBED UNDERDAMPED LANGEVIN DYNAMICS	61
4.1 Ergodicity and exponential convergence to equilibrium	63

4.2	The overdamped limit	64
5	SAMPLING FROM A GAUSSIAN MEASURE	67
5.1	Unit covariance – small perturbations	67
5.2	Exponential decay rate	73
5.3	Unit covariance – large perturbations	76
5.4	Optimal choices of J for linear and quadratic observables	85
5.5	Gaussians with arbitrary covariance and preconditioning	87
6	NUMERICAL EXPERIMENTS	93
6.1	Numerical scheme	93
6.2	Diffusion bridge sampling	95
7	DISCUSSION	99
8	PROOFS AND ADDITIONAL MATERIAL	101
8.1	Hypoellipticity	101
8.2	The overdamped limit	102
8.3	Hypocoercivity	103
8.4	The asymptotic variance for small perturbations in the linear case	108
8.5	Orthogonal transformation of tracefree symmetric matrices into a matrix with zeros on the diagonal	113
III	Couplings	115
9	INTRODUCTION TO ENSEMBLE BASED APPROACHES	117
10	COUPLED PROCESSES AND COUPLING OPERATORS	123
10.1	Preliminaries, notation and setting	123
10.2	Coupled processes and coupling operators	128
10.3	Ergodicity of couplings	131
10.4	A general way of constructing coupling operators	133
11	EXAMPLES OF COUPLED PROCESSES	137
11.1	Overdamped Langevin dynamics	137

11.2 Underdamped Langevin dynamics	143
11.3 The zigzag process	145
12 ASYMPTOTIC VARIANCE AND OPTIMAL TRANSPORT	149
12.1 A central limit theorem for coupled processes	149
12.2 Connections to the theory of optimal transportation	151
13 A PERTURBATIVE APPROACH FOR THE STUDY OF THE ASYMPTOTIC VARIANCE	161
13.1 Overdamped Langevin dynamics in one dimension with two particles	162
13.2 Overdamped Langevin dynamics with multiple particles in arbitrary dimen- sions	167
13.3 The zigzag process	171
14 A REMARK ON THE RATE OF CONVERGENCE TO EQUILIBRIUM	175
15 DISCUSSION	181
16 PROOFS AND ADDITIONAL MATERIAL	183
16.1 Tightness of coupled measures	183
16.2 Random orthogonal transformations of Brownian motions	184
16.3 The derivative formula for invariant measures	184
16.4 Properties of the solutions to one-dimensional Poisson equations	186
IV Ergodicity of piecewise deterministic Markov processes	189
17 HYPOCOERCIVITY FOR THE ZIGZAG PROCESS	191
17.1 Functional-analytic setting and outline of the proof	192
17.2 Microscopic and macroscopic coercivity	196
17.3 Auxiliary lemmas and bounds	199
17.4 Coercivity of the entropy dissipation functional	202
17.5 Discussion	203
REFERENCES	215

LIST OF FIGURES

5.1	Asymptotic variance for linear and quadratic observables, depending on relative perturbation and friction strengths	72
5.2	Effects of the perturbation on the spectra of $-\mathcal{L}$ and B	75
5.3	Autocorrelation plots for the perturbed and unperturbed dynamics	76
6.1	Standard deviation of $\hat{\pi}(f)$ for a linear observable as a function of the friction γ and the perturbation strength μ	98
6.2	Standard deviation of $\hat{\pi}(f)$ for a quadratic observable as a function of the friction γ and the perturbation strength μ	98
13.1	Comparison of invariant measures associated to solutions of Problem 13.1 and optimal transport maps as solutions to the Kantorovich problem for the example of overdamped Langevin dynamics in dimension one with quadratic potential.	164
13.2	Dependence of the asymptotic variance on different coupling schemes and coupling strengths.	168
13.3	Comparison between pairwise couplings with and without sorting according to $ \nabla\phi $ for a Gaussian target measure in $d = 10$ dimensions and with $n = 10$ particles.	171
13.4	Coupling for the zigzag process according to Corollary 13.3.	174

NOTATION AND BASIC CONVENTIONS

For a vector $x \in \mathbb{R}^d$, its Euclidean norm is denoted by $|x| = \sqrt{\sum_{i=1}^d x_i^2}$. The identity operator or identity matrix is denoted by I . In some occasions we use the notation $I_{d \times d}$ for the $d \times d$ -identity matrix to avoid confusion. For a matrix A , its transpose is denoted by A^T . The set of orthogonal matrices (i.e. those satisfying $AA^T = A^T A = I$) of size $d \times d$ is referred to as $O(\mathbb{R}^d)$. The sets of symmetric ($A^T = A$) and skew-symmetric ($A^T = -A$) matrices are denoted by $\mathbb{R}_{\text{sym}}^{d \times d}$ and $\mathbb{R}_{\text{skew}}^{d \times d}$ respectively. The trace of a matrix A is denoted by $\text{Tr } A$. For two quadratic matrices A and B (of the same size) their Frobenius inner product is given by $A : B = \text{Tr}(AB)$ and their commutator by $[A, B] = AB - BA$. The range and null spaces of a linear operator A are denoted by $\text{Ran } A = \{Ax\}$ and $\ker A = \{x : Ax = 0\}$ respectively.

Given a measure π and a measurable function f , we employ the short-hand notation $\pi(f) = \int f d\pi$. The Dirac delta measure centred at x will be denoted by δ_x , i.e. $\delta_x(f) = f(x)$. The total variation norm between two probability measures μ and ν is given by

$$\|\mu - \nu\|_{TV} = \sup_A |\mu(A) - \nu(A)|.$$

If two random variables X and Y have the same law, this fact is indicated by $X \sim Y$. Furthermore, $\xrightarrow{d}$ denotes convergence of random variables in distribution. The law of a d -dimensional Gaussian random variable with mean x and covariance matrix Σ is denoted by $\mathcal{N}(x, \Sigma)$.

The space of real-valued k -times continuously differentiable functions is denoted by $C^k(\mathbb{R}^d)$ and the corresponding subspace of functions with compact support by $C_c^k(\mathbb{R}^d)$. In particular, the spaces of smooth and compactly supported smooth functions are denoted by $C^\infty(\mathbb{R}^d)$ and $C_c^\infty(\mathbb{R}^d)$ respectively. The Hessian $(\partial_i \partial_j f)_{i,j=1}^d$ of a function $f \in C^2(\mathbb{R}^d)$ is denoted by either $\nabla \nabla f$ or $\text{Hess } f$, depending on the context. Given a measure π on $\mathbb{R}^d$, we make use of the space $L^2(\pi)$ of π -square integrable functions on $\mathbb{R}^d$, equipped with the scalar product

$$\langle f, g \rangle_{L^2(\pi)} = \int_{\mathbb{R}^d} f g d\pi, \quad f, g \in L^2(\pi)$$

and associated norm $\|f\|_{L^2(\pi)}^2 = \langle f, f \rangle_{L^2(\pi)}$. Moreover, the corresponding Sobolev spaces $H^k(\pi)$ are needed occasionally, consisting of those functions whose weak derivatives up to order k belong to $L^2(\pi)$.

Part I

Introduction

1. THE SAMPLING PROBLEM

Sampling from probability measures in high-dimensional spaces is a ubiquitous problem in the applied sciences, e.g. in machine learning, computational molecular dynamics and in Bayesian statistics. In particular, we are faced with the problem of computing expectations with respect to a probability measure π on $\mathbb{R}^d$, i.e. we wish to evaluate integrals of the form

$$\pi(f) := \int_{\mathbb{R}^d} f(x) \pi(\mathrm{d}x). \quad (1.1)$$

As is typical in many applications, particularly in molecular dynamics and Bayesian inference, the density (for convenience denoted by the same symbol π) is known only up to a normalisation constant; furthermore, the dimension of the underlying space is quite often large enough to render deterministic quadrature schemes computationally infeasible.

A standard approach to approximating such integrals is Markov chain Monte Carlo (MCMC) techniques [51, 94, 128], where a Markov process $(X_t)_{t \geq 0}$ is constructed which is ergodic with respect to the probability measure π . Then, defining the long-time average

$$\tilde{\pi}_T(f) := \frac{1}{T} \int_0^T f(X_s) \mathrm{d}s \quad (1.2)$$

for $f \in L^1(\pi)$, the celebrated ergodic theorem guarantees almost sure convergence of $\tilde{\pi}_T(f)$ to $\pi(f)$ as $T \rightarrow \infty$.

There are infinitely many Markov processes that are ergodic with respect to a given target measure, leading to challenging and mathematically rich inverse problems. Naturally, the choice should be dictated by the requirement that $\tilde{\pi}_T(f)$ can provide a good estimator of $\pi(f)$, for a fixed computational cost. More precisely, the process $(X_t)_{t \geq 0}$ shall converge to equilibrium as fast as possible, exploring the state space thoroughly and effectively, and, moreover, lend itself to efficient numerical implementations. The goal of providing principled guidelines for the construction of suitable Markov processes has spurred mathematical research across a wide range of disciplines, including the analysis of parabolic partial differential equations, spectral theory, functional inequalities and the theory of optimal transportation. This thesis shall make a contribution towards these lines of investigation.

In this introduction, we shall clarify some of the topics alluded to in the preceding paragraph. Towards this end, Section 1.1 provides a brief review of the mathematical background needed in the analysis of Markov processes, leading to natural comparison criteria for MCMC samplers outlined in Section 1.2. In Section 1.3 we summarise the classes of Markov processes considered in this thesis. An overview of the main results and the structure of this work can be found in Section 1.4.

1.1 PRELIMINARIES ON MARKOV PROCESSES

We start our review by briefly recalling some background material on continuous-time Markov processes and by fixing relevant conventions and notations. Further details concerning the topics in this section can be found in [10, 24, 45, 48, 113].

Let E be a Polish space*, equipped with its Borel σ -algebra $\mathcal{B}(E)$. We will denote the space of probability measures on $(E, \mathcal{B})$ by $\mathcal{P}(E)$ and the space of signed measures with bounded variation by $\mathcal{M}(E)$. The space of real-valued bounded measurable functions on E will be denoted by $\mathcal{B}_b(E)$ and the space of bounded continuous functions on E by $C_b(E)$. Both $\mathcal{B}_b(E)$ and $C_b(E)$ turn into Banach spaces when endowed with the supremum norm

$$\|f\|_\infty = \sup_{x \in E} |f(x)|. \quad (1.3)$$

On some probability space $(\Omega, \mathcal{F}, \mathbb{P})$, let us consider a continuous-time stochastic process $(X_t)_{t \geq 0}$ taking values in E , and denote its associated natural filtration by

$$\mathcal{F}_t^X := \sigma(X_s : s \leq t). \quad (1.4)$$

The process $(X_t)_{t \geq 0}$ satisfies the *Markov property* if

$$\mathbb{P}[X_t \in B | \mathcal{F}_s^X] = \mathbb{P}[X_t \in B | X_s], \quad \text{for all } B \in \mathcal{B}(E), 0 \leq s \leq t. \quad (1.5)$$

A convenient way of describing (time-homogeneous) Markov processes is via their transition functions:

Definition 1.1. *A time-homogeneous transition function is a mapping $\rho : [0, \infty) \times E \times \mathcal{B}(E) \rightarrow \mathbb{R}$ such that the following properties hold:*

$$1. \quad \rho_t(x, \cdot) \in \mathcal{P}(E), \quad \text{for all } (t, x) \in [0, \infty) \times E, \quad (1.6a)$$

$$2. \quad \rho_0(x, \cdot) = \delta_x, \quad \text{for all } x \in E, \quad (1.6b)$$

$$3. \quad \rho(\cdot, B) \in \mathcal{B}([0, \infty) \times E), \quad \text{for all } B \in \mathcal{B}(E), \quad (1.6c)$$

*A Polish space is a separable and completely metrisable topological space.

as well as the Chapman-Kolmogorov equation:

$$\rho_{t+s}(x, B) = \int_E \rho_s(y, B) \rho_t(x, dy), \quad s, t \geq 0, x \in E, B \in \mathcal{B}(E). \quad (1.7)$$

The process $(X_t)_{t \geq 0}$ is said to be a *time-homogeneous Markov process with transition function* ρ if $(X_t)_{t \geq 0}$ and ρ are related via

$$\mathbb{P}[X_{t+s} \in B | \mathcal{F}_t^X] = \rho_s(X_t, B), \quad \mathbb{P}\text{-a.s. for all } s, t \geq 0, B \in \mathcal{B}(E). \quad (1.8)$$

Note that $\rho_t(x, \cdot)$ is the law of X_t if the process is started at x (i.e. $X_0 = x$, $\mathbb{P}$ -almost surely). In the following, we will always restrict our attention to Markov processes constructed in the way just described.

Associated to $(X_t)_{t \geq 0}$ we define a semigroup of bounded linear operators $(S_t)_{t \geq 0}$ on $\mathcal{B}_b(E)$ via

$$(S_t f)(x) = \mathbb{E}[f(X_t) | X_0 = x], \quad t \geq 0, x \in E, f \in \mathcal{B}_b(E). \quad (1.9)$$

Observe that by virtue of (1.8), the above can equivalently be stated as

$$(S_t f)(x) = \int_E \rho_t(x, dy) f(y), \quad t \geq 0, x \in E, f \in \mathcal{B}_b(E). \quad (1.10)$$

According to the Chapman-Kolmogorov relation (1.7), $(S_t)_{t \geq 0}$ indeed satisfies the semigroup properties

$$S_0 = I, \quad S_t S_s = S_{t+s}, \quad t, s \geq 0. \quad (1.11)$$

Furthermore, $(S_t)_{t \geq 0}$ is *contractive*,

$$\|S_t f\|_\infty \leq \|f\|_\infty, \quad f \in \mathcal{B}_b(E), t \geq 0, \quad (1.12)$$

positive

$$f \geq 0 \implies S_t f \geq 0, \quad (1.13)$$

and *conservative*,

$$S_t \mathbf{1} = \mathbf{1}, \quad t \geq 0, \quad (1.14)$$

where $\mathbf{1}$ denotes the constant function on E , taking the value 1. Let us also introduce the adjoint semigroup $(S_t^*)_{t \geq 0}$ acting on $\mathcal{M}(E)$ by

$$(S_t^* \mu)(A) = \int_E \rho_t(x, A) \mu(dx), \quad A \in \mathcal{B}(E), \mu \in \mathcal{M}(E). \quad (1.15)$$

The semigroups $(S_t)_{t \geq 0}$ and $(S_t^*)_{t \geq 0}$ are dual to each other, in the sense that

$$\int_E S_t f d\mu = \int_E f d(S_t^* \mu), \quad (1.16)$$

for $f \in \mathcal{B}_b(E)$ and $\mu \in \mathcal{M}(E)$.

Let us now introduce two of the central themes in this thesis: invariant measures and ergodic Markov processes.

Definition 1.2 (Invariant measures). *A probability measure $\pi \in \mathcal{P}(E)$ is called invariant for $(X_t)_{t \geq 0}$ if*

$$\int_E S_t f \, d\pi = \int_E f \, d\pi, \quad (1.17)$$

for all $f \in \mathcal{B}_b(E)$. Equivalently, π is invariant if $S_t^ \pi = \pi$, for all $t \geq 0$.*

Invariant measures should be thought of as equilibria of the dynamics; in probabilistic terms, a measure $\pi \in \mathcal{P}(E)$ is invariant for $(X_t)_{t \geq 0}$ if $X_0 \sim \pi$ implies $X_t \sim \pi$ for all $t \geq 0$. A Markov process is ergodic if there is a unique equilibrium that is approached by the law of the process for large times. This can be formalised as follows:

Definition 1.3 (Ergodicity). *The Markov process $(X_t)_{t \geq 0}$ is said to be ergodic if the following conditions hold:*

1. $(X_t)_{t \geq 0}$ admits a unique invariant measure $\pi \in \mathcal{P}(E)$,
2. for every initial measure $\pi_0 \in \mathcal{P}(E)$ the time averages $\frac{1}{T} \int_0^T (S_t^* \pi_0) \, dt$ converge weakly to π , i.e.

$$\frac{1}{T} \int_0^T \int_E f \, d(S_t^* \pi_0) \, dt \xrightarrow{T \rightarrow \infty} \int_E f \, d\pi, \quad (1.18)$$

for all $f \in C_b(E)$.

In this case, $(X_t)_{t \geq 0}$ is also said to be ergodic with respect to π , and π is called an ergodic invariant measure.

Remark 1.1. *The precise notion of ergodicity varies considerably across the literature. In the theory of dynamical systems, a probability measure $\pi \in \mathcal{P}(E)$ is usually said to be ergodic if (1.18) holds, provided that $X_0 \sim \pi$. This concept is closely connected to certain irreducibility properties of the process $(X_t)_{t \geq 0}$ on the support of π . Note that in this framework there can be multiple ergodic measures. We refer the reader to [143] for more details. In our context, it is unreasonable to assume that the process is started in equilibrium. Rather, the essence of Definition 1.3 is the convergence of the process towards a unique equilibrium for large times. The precise notion of this convergence also varies according to the setting. In (1.18), the limit is taken in the Césaro sense (i.e. considering time averages) which translates naturally into the statement*

$$\frac{1}{T} \int_0^T f(X_t) \, dt \xrightarrow{T \rightarrow \infty} \int_E f \, d\pi, \quad (1.19)$$

almost surely and in $L^1(\pi)$, and this formulation is widely known as Birkhoff's ergodic theorem. As explained in the introduction, (1.19) is the linchpin for Markov chain Monte Carlo algorithms as considered in this thesis. Let us remark that Césaro convergence is a particularly weak requirement; for instance it is implied by ordinary weak convergence of probability measures ($S_t^* \pi_0 \rightarrow \pi$) and of course also by strong convergence in total variation norm ($\|S_t^* \pi_0 - \pi\|_{\text{TV}} \rightarrow 0$). Further information concerning different notions of convergence in relation to ergodic properties can be found in [99] and [133].

STRONGLY CONTINUOUS SEMIGROUPS

The study of continuous-time Markov processes is greatly enriched by the theory of strongly continuous semigroups on arbitrary Banach spaces. Here, we content ourselves with highlighting the crucial role played by the concept of strong continuity in the definition of generators and refer the reader to [47] and [116] for much more thorough treatments of the subject.

Definition 1.4 (Strong continuity). *Let $(B, \|\cdot\|_B)$ be a Banach space, and $(S_t)_{t \geq 0}$ be a one-parameter family of bounded linear operators on B . Then $(S_t)_{t \geq 0}$ is said to be strongly continuous if*

$$\|S_t f - f\|_B \xrightarrow[t \rightarrow 0]{} 0, \quad (1.20)$$

for all $f \in B$.

For a strongly continuous semigroup on a Banach space $(B, \|\cdot\|_B)$, we can define its generator $(\mathcal{L}, \mathcal{D}(\mathcal{L}))$ by

$$\mathcal{L}f = \lim_{t \rightarrow 0} \frac{S_t f - f}{t}, \quad f \in \mathcal{D}(\mathcal{L}), \quad (1.21a)$$

$$\mathcal{D}(\mathcal{L}) = \left\{ f \in B : \lim_{t \rightarrow 0} \frac{S_t f - f}{t} \text{ exists in } (B, \|\cdot\|_B) \right\}. \quad (1.21b)$$

The theory of strongly continuous semigroups then shows that $(\mathcal{L}, \mathcal{D}(\mathcal{L}))$ is closed and densely defined. Unfortunately, the semigroup $(S_t)_{t \geq 0}$ defined in (1.9) usually fails to be strongly continuous on $(\mathcal{B}_b(E), \|\cdot\|_\infty)$. In particular, the domain $\mathcal{D}(\mathcal{L})$ as defined above might not be dense in B . There are (at least) two standard ways of dealing with this problem:

1. If an invariant measure π is available, then for any $p \in [1, \infty]$, formula (1.10) extends straightforwardly to a contraction semigroup on $L^p(\pi)$ by Jensen's inequality. For $1 \leq p < \infty$, the extended semigroup is then strongly continuous in $L^p(\pi)$ under minimal regularity assumptions on the paths of $(X_t)_{t \geq 0}$, for instance, if the paths are right-continuous almost surely. Henceforth, we will assume that all Markov processes under consideration satisfy this property. For the bulk of Part II an invariant

measure π will be fixed and therefore it will be convenient to work with the extended semigroups on $L^p(\pi)$ and, in particular, on $L^2(\pi)$, since the latter is a Hilbert space.

2. In Part III, we will refrain from fixing an invariant measure and instead make the further regularity assumption that $(S_t)_{t \geq 0}$ satisfies the *Feller property*, i.e. that $(S_t)_{t \geq 0}$ can be realised as a strongly continuous semigroup on $C_0(E)$, the space of functions that vanish at infinity. For a more detailed introduction to the theory of Feller semigroups and precise statements, we refer to Section 10.1.

In applications, the domain $\mathcal{D}(\mathcal{L})$ of a generator is usually very hard to determine explicitly. Fortunately, it is almost always sufficient to perform manipulations on suitable dense subspaces:

Definition 1.5 (Core). *Let $(\mathcal{L}, \mathcal{D}(\mathcal{L}))$ be the generator of a strongly continuous semigroup on a Banach space $(B, \|\cdot\|_B)$. Then the graph norm on $\mathcal{D}(\mathcal{L})$ is given by*

$$\|f\|_{\mathcal{D}(\mathcal{L})} = \|f\|_B + \|\mathcal{L}f\|_B, \quad f \in \mathcal{D}(\mathcal{L}). \quad (1.22)$$

A subset $\mathcal{D} \subset \mathcal{D}(\mathcal{L})$ is called a core if it is dense in $\mathcal{D}(\mathcal{L})$ with respect to $\|\cdot\|_{\mathcal{D}(\mathcal{L})}$.

The significance of the foregoing definition is the following. If $\mathcal{D}$ is a core for $(\mathcal{L}, \mathcal{D}(\mathcal{L}))$, then $\mathcal{L}|_{\mathcal{D}}$ contains all the information on the corresponding semigroup in the sense that the smallest closed extension of $\mathcal{L}|_{\mathcal{D}}$ is precisely $(\mathcal{L}, \mathcal{D}(\mathcal{L}))$, see [4, Section A-I-1].

FOKKER-PLANCK AND KOLMOGOROV EQUATIONS

Let us return to the Markov semigroup $(S_t)_{t \geq 0}$ defined in (1.9) and suppose that strong continuity holds, either for the extended semigroup on $L^p(\pi)$, or in terms of the Feller property on $C_0(E)$. In both cases, the generator $(\mathcal{L}, \mathcal{D}(\mathcal{L}))$ is well defined and the *backward Kolmogorov equation*

$$\partial_t S_t f = \mathcal{L} S_t f, \quad f \in \mathcal{D}(\mathcal{L}), \quad (1.23)$$

holds, governing the evolution of observables of the dynamics. We remark that usually $(S_t)_{t \geq 0}$ has sufficient regularising properties to ensure that $S_t(B) \subset \mathcal{D}(\mathcal{L})$ for $t > 0$, such that (1.23) in fact holds for all $f \in B$. In the preceding statement, it is understood that either $B = L^p(\pi)$ or $B = C_0(E)$, depending on the context.

A similar relation holds between the law of $(X_t)_{t \geq 0}$ and the adjoint of $\mathcal{L}$, defined in a suitable sense. To make this precise, let us now assume that the transition functions are absolutely continuous with respect to a reference measure m on E , i.e.

$$\rho_t(x, dy) = \rho_t(x, y) m(dy), \quad (1.24)$$

and that the law of X_0 is given by $\pi_0(x)m(dx)$. Then, at least formally, the law of X_t is given by $\pi_t(x)m(dx)$, where π_t satisfies the *Fokker-Planck equation*

$$\partial_t \pi_t = \mathcal{L}^\dagger \pi_t. \quad (1.25)$$

Here, $\mathcal{L}^\dagger$ denotes the (formal) adjoint of $\mathcal{L}$ in $L^2(E, m)$, i.e.

$$\int_E f(\mathcal{L}g) dm = \int_E (\mathcal{L}^\dagger f)g dm \quad (1.26)$$

for functions f and g in a suitably rich class of test functions. If π is an invariant measure for $(X_t)_{t \geq 0}$, absolutely continuous with respect to m , then clearly its density (for convenience denoted by the same letter) has to satisfy the stationary Fokker-Planck equation

$$\mathcal{L}^\dagger \pi = 0. \quad (1.27)$$

Let us already mention that in the case of Feller semigroups (rigorously introduced in Definition 10.2), the preceding statement has the following rigorous version (see for instance [97, Proposition 8.1.2]):

Theorem 1.1. *Let $\mathcal{L}$ be the generator of a Feller semigroup. Then $\pi \in \mathcal{P}(E)$ is an invariant measure for the corresponding process if and only if*

$$\int_E \mathcal{L}f d\pi = 0, \quad \text{for all } f \in \mathcal{D}(\mathcal{L}). \quad (1.28)$$

Let us seize the opportunity to stress the crucial conceptual role that (1.27) plays in the context of this thesis. Given a measure $\pi \in \mathcal{P}(E)$, our aim is to find Markov processes that are ergodic with respect to it and enjoy favourable (statistical) properties. In terms of the stationary Fokker-Planck equation, we are led to study subclasses of operators that satisfy (1.27).

PROPERTIES OF THE GENERATOR IN $L^2(\pi)$

In the two final sections of this review, let us assume that $(X_t)_{t \geq 0}$ is ergodic with respect to π , and that $(X_t)_{t \geq 0}$ has right-continuous sample paths almost surely. As explained previously, $(S_t)_{t \geq 0}$ then extends to a strongly continuous semigroup on the Hilbert space $L^2(\pi)$. In this situation, the generator $\mathcal{L}$ enjoys a number of useful properties that we will summarise here. First of all, let us introduce the following subspace of $L^2(\pi)$, consisting of π -centred functions:

$$L_0^2(\pi) = \{f \in L^2(\pi) : \pi(f) = 0\}. \quad (1.29)$$

Note that we naturally have the decomposition

$$L^2(\pi) = L_0^2(\pi) \oplus \text{span}\{\mathbf{1}\}. \quad (1.30)$$

According to (1.14) and (1.17), the semigroup $(S_t)_{t \geq 0}$ leaves both $\text{span}\{\mathbf{1}\}$ and $L_0^2(\pi)$ invariant. In particular, it is often convenient to think of $(S_t)_{t \geq 0}$ as a strongly continuous semigroup on $L_0^2(\pi)$. The following proposition shows that the generator $\mathcal{L}$ interacts nicely with the decomposition (1.30). The proofs are contained in [19].

Proposition 1.1 (Properties of the generator in $L^2(\pi)$). *Let $(X_t)_{t \geq 0}$ be a Markov process with right-continuous paths, ergodic with respect to π , and $(S_t)_{t \geq 0}$ the corresponding strongly continuous semigroup on $L^2(\pi)$ with generator $(\mathcal{L}, \mathcal{D}(\mathcal{L}))$. Then the following hold:*

1. *The spectrum of $\mathcal{L}$ is contained in the left half-plane:*

$$\sigma(\mathcal{L}) \subset \mathbb{R}_{\leq 0} \times i\mathbb{R}. \quad (1.31)$$

2. *$\mathcal{L}\mathbf{1} = 0$, and 0 is a simple eigenvalue of $\mathcal{L}$, i.e.*

$$\ker \mathcal{L} = \text{span}\{\mathbf{1}\}. \quad (1.32)$$

3. *The range $\text{Ran } \mathcal{L}$ is dense in $L_0^2(\pi)$, so in particular*

$$\int_E \mathcal{L}f \, d\pi = 0, \quad f \in \mathcal{D}(\mathcal{L}). \quad (1.33)$$

Moreover, $\text{Ran } \mathcal{L} = L_0^2(\pi)$ if and only if 0 is an isolated point in the spectrum $\sigma(\mathcal{L})$. In the latter case, $\mathcal{L}$ is invertible on $L_0^2(\pi)$ and the inverse is given by

$$\mathcal{L}^{-1}f = - \int_0^\infty S_t f \, dt, \quad f \in L_0^2(\pi). \quad (1.34)$$

Conversely, if the integral (1.34) converges in $L^2(\pi)$ for all $f \in L_0^2(\pi)$, then 0 is an isolated eigenvalue in $\sigma(\mathcal{L})$.

REVERSIBILITY AND SELF-ADJOINTNESS

As in the previous section, let us assume that $(X_t)_{t \geq 0}$ is ergodic with respect to $\pi \in \mathcal{P}(E)$, inducing a strongly continuous semigroup $(S_t)_{t \geq 0}$ with generator $\mathcal{L}$ on $L^2(\pi)$. Evidently, if $X_0 \sim \pi$, then $X_t \sim \pi$ for all $t \geq 0$. A stochastic process with this property is called *stationary*. As we will see in Definition 1.6 and Theorem 1.2, the dynamical behaviour of $(X_t)_{t \geq 0}$ with respect to time reversal in stationarity is intimately related to self-adjointness properties of $\mathcal{L}$.

Definition 1.6 (Reversible process). *A stationary stochastic process $(X_t)_{t \geq 0}$ is called reversible if for all $T > 0$, the processes $(X_t)_{0 \leq t \leq T}$ and $(X_{T-t})_{0 \leq t \leq T}$ have the same law.*

In terms of the transition functions, reversibility can be expressed in the form

$$\rho_t(x, dy)\pi(dx) = \rho_t(y, dx)\pi(dy), \quad t \geq 0, \quad x, y \in E. \quad (1.35)$$

Roughly speaking, at equilibrium the probability of the process making a transition from x to y equals the probability of the backward transition from y to x . In the physics literature, reversibility is often called ‘detailed balance’, hinting at the fact that π is not only (globally) invariant, but also locally, i.e. when only pairwise transition probabilities are considered. The proof of the following result can be found in [77, Section 3.3.2].

Theorem 1.2. *The following are equivalent:*

1. *If $X_0 \sim \pi$, then the process $(X_t)_{t \geq 0}$ is reversible.*
2. *The semigroup $(S_t)_{t \geq 0}$ is self-adjoint in $L^2(\pi)$, i.e.*

$$\langle f, S_t g \rangle_{L^2(\pi)} = \langle S_t f, g \rangle_{L^2(\pi)}, \quad t \geq 0, \quad f, g \in L^2(\pi). \quad (1.36)$$

3. *The generator $\mathcal{L}$ is self-adjoint in $L^2(\pi)$.*

Remark 1.2 (Symmetry and self-adjointness). *Since the generator of a self-adjoint semigroup is necessarily self-adjoint, in order to verify the third condition in Theorem 1.2 it is sufficient to show that $\mathcal{L}$ is symmetric on a core $\mathcal{D}$, i.e.*

$$\langle f, \mathcal{L}g \rangle_{L^2(\pi)} = \langle \mathcal{L}f, g \rangle_{L^2(\pi)}, \quad f, g \in \mathcal{D}. \quad (1.37)$$

Remark 1.3. *Although the notion of reversibility has only been defined for stationary processes, the second and third conditions of Theorem 1.2 make perfect sense for processes out of equilibrium. It is therefore natural and meaningful to discuss reversibility properties of $(X_t)_{t \geq 0}$ even if $X_0 \approx \pi$.*

We close this review by observing that any generator of a strongly continuous semigroup on $L^2(\pi)$ can be split into its symmetric and antisymmetric parts,

$$\mathcal{L} = \mathcal{S} + \mathcal{T}, \quad (1.38)$$

where

$$\mathcal{S} = \frac{1}{2}(\mathcal{L} + \mathcal{L}^*), \quad \mathcal{T} = \frac{1}{2}(\mathcal{L} - \mathcal{L}^*), \quad (1.39)$$

denoting by $\mathcal{L}^*$ the adjoint of $\mathcal{L}$ in $L^2(\pi)$. As Theorem 1.2 shows, this decomposition carries meaning with respect to the dynamical properties of $(X_t)_{t \geq 0}$. In particular, the

antisymmetric part $\mathcal{T}$ is related to nonreversible behaviour. Informally speaking, the latter manifests itself in directed motion as commonly observed in transport phenomena.

1.2 COMPARISON CRITERIA

In this section, we will connect the theoretical background reviewed in the previous section to the task of selecting appropriate Markovian dynamics for MCMC samplers. More specifically, we will see how key algorithmic performance criteria are naturally related to certain properties of the corresponding generators.

Let $(X_t)_{t \geq 0}$ be a continuous-time Markov process with right-continuous paths, ergodic with respect to $\pi \in \mathcal{P}(E)$. For a fixed observable $f \in L^1(\pi)$, our goal is to quantify the accuracy of the estimator

$$\tilde{\pi}_T(f) = \frac{1}{T} \int_0^T f(X_t) dt. \quad (1.40)$$

Recall that according to Remark 1.1, $\tilde{\pi}_T(f) \rightarrow \pi(f)$ as $T \rightarrow \infty$. A natural measure capturing the quality of the approximation for finite T is the *mean square error* (MSE) defined by

$$MSE(f, T) := \mathbb{E}_x |\tilde{\pi}_T(f) - \pi(f)|^2, \quad (1.41)$$

where $\mathbb{E}_x$ denotes the expectation conditioned on the process $(X_t)_{t \geq 0}$ starting at $x \in E$. It is instructive to study the decomposition

$$MSE(f, T) = \mu^2(f, T) + \sigma^2(f, T), \quad (1.42)$$

introducing

$$\begin{aligned} \mu(f, T) &= |\mathbb{E}_x[\tilde{\pi}_T(f)] - \pi(f)| \quad \text{and} \quad \sigma^2(f, T) = \mathbb{E}_x |\tilde{\pi}_T(f) - \mathbb{E}_x \tilde{\pi}_T(f)|^2 \\ &= \text{Var}[\tilde{\pi}_T(f)]. \end{aligned} \quad (1.43)$$

Here $\mu(f, T)$ measures the bias of the estimator $\tilde{\pi}_T(f)$ and $\sigma^2(f, T)$ measures the variance of fluctuations of $\tilde{\pi}_T(f)$ around the mean. In the sequel, we will outline the connections between these two quantities and properties of the generator $\mathcal{L}$ corresponding to $(X_t)_{t \geq 0}$.

1.2.1 CONVERGENCE TO EQUILIBRIUM AND SPECTRAL GAP

Let us denote the law of X_t by π_t and the initial distribution by π_0 . Often, we can then establish a decay estimate of the form

$$d(\pi_t, \pi) \leq C e^{-\lambda t} d(\pi_0, \pi), \quad t \geq 0, \quad (1.44)$$

where $C \geq 1$, $\lambda > 0$ and d is an appropriate distance on the space of probability measures $\mathcal{P}(E)$. The constants λ and C measure the speed of convergence to equilibrium, and in particular allow to control $\mu(f, T)$ in many situations (see for instance Lemma 1.1 below). It is clearly desirable to find a process $(X_t)_{t \geq 0}$ that allows C to be chosen as small and λ as large as possible.

The metric d can be selected among a wide range of options (see for instance [52] for an overview). Typically, this choice will greatly impact both the mathematical tools available as well as the scope of the results, particularly in high-dimensional settings. Let us already refer the reader to Chapter 3 where a comparison between Wasserstein distances, relative entropies and $L^2(\pi)$ -based distances is made in the context of overdamped Langevin diffusions targeting a Gaussian measure (see in particular the discussion in Section 3.2).

In the case when the measures π_0 , $(\pi_t)_{t \geq 0}$ and π have densities with respect to a reference measure m on E (for convenience denoted by the same letter) and the density associated to π is strictly positive, we can choose

$$d(\pi^1, \pi^2) = \|\pi^1 - \pi^2\|_{L^2(\frac{m}{\pi})} = \left(\int_E (\pi^1(x) - \pi^2(x))^2 \frac{dm}{\pi(x)} \right)^{1/2}. \quad (1.45)$$

With this distance, (1.44) turns out to be equivalent to the following decay estimate of the semigroup $(S_t)_{t \geq 0}$:

$$\|S_t f - \pi(f)\|_{L^2(\pi)} \leq C e^{-\lambda t} \|f - \pi(f)\|_{L^2(\pi)}, \quad f \in L^2(\pi). \quad (1.46)$$

Note that the above is equivalent to

$$\|S_t\|_{L_0^2(\pi) \rightarrow L_0^2(\pi)} \leq C e^{-\lambda t}, \quad (1.47)$$

where $(S_t)_{t \geq 0}$ is considered as a semigroup on $L_0^2(\pi)$. This setting is quite restrictive, since the requirement $\|\pi_0 - \pi\|_{L^2(\frac{m}{\pi})} < \infty$ (equivalently, $\pi_0 \in L^2(\frac{m}{\pi})$) enforces very stringent decay conditions on the tails of π_0 ; in a sense the process has to be started already close to equilibrium. On the positive side, this setting is very amenable to analysis since (1.47) is closely related to the spectrum of the generator $\mathcal{L}$. In order to make this precise, let us denote the *optimal decay rate* in (1.47) by $\lambda^*(\mathcal{L})$, i.e.

$$\lambda^*(\mathcal{L}) = \inf_{t > 0} \frac{\log \|S_t\|_{L_0^2(\pi) \rightarrow L_0^2(\pi)}}{t} = \lim_{t \rightarrow \infty} \frac{\log \|S_t\|_{L_0^2(\pi) \rightarrow L_0^2(\pi)}}{t}, \quad (1.48)$$

and the *spectral gap* of the generator by $\lambda_{\text{gap}}(\mathcal{L})$, i.e.

$$\lambda_{\text{gap}}(\mathcal{L}) = \inf (\operatorname{Re} \sigma(-\mathcal{L}) \setminus \{0\}). \quad (1.49)$$

We then have that

$$0 \leq \lambda^*(\mathcal{L}) \leq \lambda_{\text{gap}}(\mathcal{L}), \quad (1.50)$$

by the general theory of strongly continuous semigroups, see [47, Corollary 1.13]. For Markov processes, the semigroup $(S_t)_{t \geq 0}$ usually has additional smoothing properties, for instance in the sense that $S_t f \in \mathcal{D}(\mathcal{L})$ for all $f \in L^2(\pi)$ and $t > 0$. In this case, an appropriate spectral mapping theorem holds, and $\lambda^*(\mathcal{L}) = \lambda_{\text{gap}}(\mathcal{L})$ follows (for an in-depth discussion, see [4, Chapter 6]).

If we restrict our attention to decay estimates of the form (1.46) with $C = 1$, then the decay rate can be characterised in terms of a functional inequality [10, Theorem 4.2.5]:

Theorem 1.3. *For $\lambda > 0$, the following are equivalent:*

1. Poincaré inequality:

$$\langle f, f \rangle_{L^2(\pi)} \leq \frac{1}{\lambda} \langle f, (-\mathcal{L})f \rangle_{L^2(\pi)}, \quad f \in \mathcal{D}(\mathcal{L}) \cap L_0^2(\pi). \quad (1.51)$$

2. The decay estimate (1.46) holds with $C = 1$.

Strikingly, the validity of the Poincaré inequality only depends on the symmetric part of the generator. This phenomenon will be discussed in more detail in Part II, see in particular the discussion following Theorem 2.1.

Finally, let us show that the decay of the bias $\mu(f, T)$ as $T \rightarrow \infty$ can be controlled in terms of the constants λ and C in (1.46).

Lemma 1.1. *Let $(X_t)_{t \geq 0}$ be ergodic with respect to π , such that for $t \geq 0$, $X_t \sim \pi_t$ is absolutely continuous with respect to π , with $\frac{d\pi_0}{d\pi} \in L^2(\pi)$. Suppose that the semigroup $(S_t)_{t \geq 0}$ satisfies (1.46). Then for $f \in L^\infty(\pi)$,*

$$\mu(f, T) \leq \frac{C}{2\lambda T} \left(1 - e^{-\lambda T}\right) \|f\|_{L^\infty} \text{Var}_\pi \left[\frac{d\pi_0}{d\pi} \right]^{\frac{1}{2}}.$$

Proof. We denote by π_t the law of X_t given $X_0 \sim \pi_0$. Furthermore, let $h := \frac{d\pi_0}{d\pi} \in L^2(\pi)$. Then

$$\|\pi_t - \pi\|_{TV} = \frac{1}{2} \|S_t^* h - \mathbf{1}\|_{L^1(\pi)} \leq \frac{1}{2} \|S_t^*\|_{L^2(\pi) \rightarrow L^2(\pi)} \|h - \mathbf{1}\|_{L^2(\pi)} \leq \frac{C}{2} e^{-\lambda t} \|h - \mathbf{1}\|_{L^2(\pi)},$$

where S_t^* denotes the $L^2(\pi)$ -adjoint of S_t . Since f is assumed to be bounded, we immediately obtain

$$|\mathbb{E}[f(X_t)|X_0 \sim \pi_0] - \pi(f)| \leq \frac{C}{2} \|f\|_{L^\infty} e^{-\lambda t} \left(\text{Var}_\pi \left[\frac{d\pi_0}{d\pi} \right] \right)^{1/2},$$

and so, for $X_0 \sim \pi_0$,

$$|\tilde{\pi}_T(f) - \pi(f)| \leq \frac{C}{2\lambda T} (1 - e^{-\lambda t}) \|f\|_{L^\infty} \left(\text{Var}_\pi \left[\frac{d\pi_0}{d\pi} \right] \right)^{1/2},$$

as required. $\square$

1.2.2 FLUCTUATIONS AND ASYMPTOTIC VARIANCE

In order to study the variance term $\sigma^2(f, T)$ in the decomposition (1.42) let us present the following result [19]:

Theorem 1.4 (Central limit theorem). *Let $(X_t)_{t \geq 0}$ be ergodic with respect to π , started at equilibrium, i.e. $X_0 \sim \pi$, and consider $f \in L^2(\pi)$. If $f \in \text{Ran } \mathcal{L}$, then*

$$\sqrt{T} \left(\frac{1}{T} \int_0^T f(X_t) dt - \pi(f) \right) \xrightarrow[T \rightarrow \infty]{d} \mathcal{N}(0, 2\sigma_f^2), \quad (1.52)$$

where the asymptotic variance is given by

$$\sigma_f^2 = \langle f, \phi \rangle_{L^2(\pi)}, \quad (1.53)$$

and $\phi \in \mathcal{D}(\mathcal{L})$ is a solution to the Poisson equation

$$-\mathcal{L}\phi = f, \quad \pi(\phi) = 0. \quad (1.54)$$

Remark 1.4. *The condition $X_0 \sim \pi$ can be weakened in many situations. For instance, it has been proved in [19, Theorem 2.6] that the former can be dropped under a stronger ergodicity assumption, namely if the limit in (1.18) is required to hold with respect to the total variation distance of probability measures (instead of merely in terms of weak convergence of time averages).*

In the above theorem, $\mathcal{N}(0, 2\sigma_f^2)$ denotes a Gaussian random variable with mean zero and variance $2\sigma_f^2$. The result hence shows that $\sigma^2(f, T) \approx 2T^{-1/2}\sigma_f^2$ for large T . Let us comment on the condition $f \in \text{Ran } \mathcal{L}$. From Proposition 1.1 we know that this requirement necessarily implies $\pi(f) = 0$. However, observables that only differ by a constant clearly have associated central limit theorems with the same asymptotic variances. Replacing f by $f - \pi(f)$, we are therefore entitled to assume $f \in L_0^2(\pi)$ without loss of generality. Equivalently, we can alter the Poisson equation (1.54) in the following way:

$$-\mathcal{L}\phi = f - \pi(f), \quad \pi(\phi) = 0. \quad (1.55)$$

From (1.50) and Proposition 1.1, we see that establishing the decay estimate (1.46) is sufficient to ensure that $\text{Ran } \mathcal{L} = L_0^2(\pi)$, guaranteeing that the central limit theorem

(1.52) holds for all $f \in L^2(\pi)^\dagger$. Moreover, in this case (1.34) immediately implies the estimate

$$\sigma_f^2 = \langle f, (-\mathcal{L})^{-1} f \rangle_{L^2(\pi)} \leq \frac{C}{\lambda} \|f\|_{L^2(\pi)}^2, \quad f \in L_0^2(\pi). \quad (1.56)$$

Consequently, the constants λ and C in (1.46) not only control the bias $\mu(f, T)$, but also the asymptotic variance σ_f^2 . Combining (1.53) with (1.34) we see that the asymptotic variance can be expressed as an integrated autocorrelation function (see also [7, Proposition IV 1.3]):

$$\sigma_f^2 = \int_0^\infty \mathbb{E}_\pi [(f(X_0) - \pi(f))(f(X_t) - \pi(f))] dt. \quad (1.57)$$

Intuitively speaking, reducing the asymptotic variance therefore amounts to reducing the correlation (or increasing negative correlations) of the values $(f(X_t))_{t \geq 0}$.

Remark 1.5. *For Markov chains, criteria involving the asymptotic variance are strongly related to the Peskun ordering (see [105] and references therein). Let us mention that the comparison of Markovian dynamics for MCMC sampling is not limited to the study of the asymptotic variance or the speed of convergence to equilibrium. Other criteria include large deviation rate functionals for the ensemble measure $\frac{1}{t} \int_0^t \delta_{X(t)}(dx)$ [44, 125] and scaling limits for Markov chains [14, 129]. Also of interest are nonasymptotic estimates as provided in [78].*

1.3 A SELECTION OF ERGODIC MARKOV PROCESSES

After having outlined the general mathematical framework in Section 1.1 as well as natural comparison criteria between Markov processes in the context of MCMC sampling in Section 1.2, we will now present the concrete examples that will be considered in this thesis. Throughout, we will assume that the target measure is given by

$$\pi(dx) = \frac{1}{Z} e^{-V(x)} dx, \quad (1.58)$$

where the potential $V \in C^\infty(\mathbb{R}^d)$ is such that

$$Z := \int_{\mathbb{R}^d} e^{-V(x)} dx < \infty. \quad (1.59)$$

In applications, the normalisation constant Z is typically unknown and the potential V (and some of its derivatives) can be evaluated pointwise. For convenience, the density $\pi(x) = \frac{1}{Z} e^{-V(x)}$ will be denoted by the same symbol. The Markov processes in this section will be constructed in such a way that π (or an extension of it) is invariant, and usually ergodic. These should therefore be considered as candidates for the use in Markov chain Monte Carlo sampling schemes.

[†]More generally, any decay estimate showing convergence of the integral in (1.34) is sufficient.

1.3.1 DIFFUSION PROCESSES

Let us consider the case when $(X_t)_{t \geq 0}$ is a diffusion process defined by the solution of the following Itô SDE,

$$dX_t = a(X_t) dt + \sqrt{2}b(X_t) dW_t, \quad (1.60)$$

with drift coefficient $a : \mathbb{R}^d \rightarrow \mathbb{R}^d$ and diffusion coefficient $b : \mathbb{R}^d \rightarrow \mathbb{R}^{d \times m}$, both having smooth components, and where $(W_t)_{t \geq 0}$ is a standard $\mathbb{R}^m$ -valued Brownian motion. We assume that the (strong) solution to (1.60) exists for all $t \geq 0$. Associated with (1.60) is the infinitesimal generator $\mathcal{L}$ formally given by

$$\mathcal{L}f = a \cdot \nabla f + Q : \nabla \nabla f, \quad f \in C_c^2(\mathbb{R}^d), \quad (1.61)$$

where $Q = bb^\top$, $\nabla \nabla f$ denotes the Hessian of the function f and $:$ denotes the Frobenius inner product. In general, Q is nonnegative definite, and could possibly be degenerate. In particular, the infinitesimal generator (1.61) need not be uniformly elliptic. To ensure that the corresponding semigroup exhibits sufficient smoothing behaviour, we shall require that the generator (1.61) is hypoelliptic in the sense of Hörmander [70]:

Definition 1.7 (Hypoellipticity). *Let $\mathcal{L}$ be a linear differential operator on $\mathbb{R}^d$, acting on the space of distributions $(C_c^\infty(\mathbb{R}^d))'$. Then $\mathcal{L}$ is called hypoelliptic if $\mathcal{L}f \in C^\infty(\mathbb{R}^d)$ implies $f \in C^\infty(\mathbb{R}^d)$.*

If Q is positive definite at every point, then $\mathcal{L}$ is elliptic, so in particular hypoelliptic (see [53, Chapter 7]). In the case when Q is degenerate, a priori $\mathcal{L}$ possesses smoothing properties only in certain directions. However, the interplay between a and Q (drift and diffusion) is often sufficient to ensure hypoellipticity also in this situation. The following classical theorem makes this precise. Note that a set $\{\mathbf{X}_k\}_{k=0}^K$ of homogeneous first-order linear differential operators with smooth coefficients on $\mathbb{R}^d$ (i.e. operators of the form $\mathbf{X} = \sum_i a_i \partial_i$) naturally gives rise to a Lie algebra by considering linear combinations and (iterated) commutators $[A, B] = AB - BA$. More precisely, we define $\text{Lie}(\mathbf{X}_0, \dots, \mathbf{X}_K)$ to be the linear span of the family

$$\{\mathbf{X}_i\}_{i=0}^K, \quad \{[\mathbf{X}_i \mathbf{X}_j]\}_{i,j=0}^K, \quad \{[[\mathbf{X}_i, \mathbf{X}_j], \mathbf{X}_k]\}_{i,j,k=0}^K, \quad \dots \quad (1.62)$$

Clearly, $\text{Lie}(\mathbf{X}_0, \dots, \mathbf{X}_K)$ is closed under the operation of taking commutators, and hence indeed a Lie algebra.

Theorem 1.5 (Hörmander). *Let $\mathcal{L}$ be a linear differential operator on $\mathbb{R}^d$ of the form*

$$\mathcal{L} = \mathbf{X}_0 + \frac{1}{2} \sum_{k=1}^K \mathbf{X}_k^2, \quad (1.63)$$

where $\{\mathbf{X}_k\}_{k=0}^K$ is a set of homogeneous first-order linear differential operators with smooth

coefficients. Assume that the Hörmander condition

$$\text{Lie}(\mathbf{X}_0, \dots, \mathbf{X}_K) = \text{span}(\partial_{x_1}, \dots, \partial_{x_d}) \quad (1.64)$$

holds. Then $\mathcal{L}$ is hypoelliptic.

Let us denote by $\mathcal{L}^\dagger$ the (formal) adjoint of $\mathcal{L}$ in $L^2(\mathbb{R}^d)$. On account of the Fokker-Planck equation (1.25), the process $(X_t)_{t \geq 0}$ has convenient regularity properties if $\partial_t - \mathcal{L}^\dagger$ is hypoelliptic as an operator on $(C_c^\infty(\mathbb{R}^{d+1}))'$. In this case, ergodicity properties can be deduced under natural conditions, leading to the following result (see [82], [90, Theorem 2.7], [124, Corollary 7.2]).

Theorem 1.6. *Let the generator $\mathcal{L}$ be of the form (1.63) and assume that the span of the family*

$$\{\mathbf{X}_i\}_{i=1}^K, \quad \{[\mathbf{X}_i \mathbf{X}_j]\}_{i,j=0}^K, \quad \{[[\mathbf{X}_i, \mathbf{X}_j], \mathbf{X}_k]\}_{i,j,k=0}^K, \quad \dots \quad (1.65)$$

equals $\text{span}(\partial_{x_1}, \dots, \partial_{x_d})$. Then the transition functions $\rho_t(x, dy)$ associated to (1.60) as well as any invariant measures have smooth densities with respect to the Lebesgue measure. Furthermore, if there exists an invariant measure with strictly positive density, then this measure is unique, and the process $(X_t)_{t \geq 0}$ is ergodic.

Remark 1.6. *Observe that the family (1.65) differs from (1.62) by the omission of $\mathbf{X}_0$ in the set $\{\mathbf{X}_i\}_{i=0}^K$. The generation property corresponding to (1.65) in Theorem 1.6 is often termed the parabolic Hörmander condition. If the latter is satisfied, then the associated diffusion process as well as (slightly abusively) the generator $\mathcal{L}$ are called hypoelliptic.*

The following result provides a necessary and sufficient condition on the coefficients of (1.60) for the diffusion process to be invariant with respect to a target distribution π as given in (1.58).

Theorem 1.7. *Consider a diffusion process $(X_t)_{t \geq 0}$ on $\mathbb{R}^d$ defined by the unique, non-explosive solution to the Itô SDE (1.60) with drift coefficient $a \in C^1(\mathbb{R}^d, \mathbb{R}^d)$ and diffusion coefficient $b \in C^1(\mathbb{R}^d, \mathbb{R}^{d \times m})$. Then $(X_t)_{t \geq 0}$ is invariant with respect to π if and only if*

$$a = Q \nabla \log \pi + \nabla \cdot Q + \gamma, \quad (1.66)$$

where $Q = bb^\top$ and $\gamma \in C^1(\mathbb{R}^d, \mathbb{R}^d)$ is a vector field satisfying

$$\nabla \cdot (\pi \gamma) = 0. \quad (1.67)$$

If additionally $\gamma \in L^1(\pi)$, then there exists a skew-symmetric matrix function $C : \mathbb{R}^d \rightarrow \mathbb{R}^{d \times d}$ such that $\gamma = \frac{1}{\pi} \nabla \cdot (\pi C)$. In this case the infinitesimal generator can be written as an $L^2(\pi)$ -extension of

$$\mathcal{L}f = \frac{1}{\pi} \nabla \cdot ((Q + C)\pi \nabla f), \quad f \in C_c^2(\mathbb{R}^d). \quad (1.68)$$

The proof of the first part of this result can be found in [113, Ch. 4]; similar versions of this characterisation can be found in [141] and [72]. For the existence of the skew-symmetric matrix C see, e.g., [49, Sec.4, Prop. 1]. See also [98].

Remark 1.7. *The matrices Q and C in (1.68) are related to the decomposition (1.38) of $\mathcal{L}$ into $L^2(\pi)$ -symmetric and skew-symmetric parts, i.e.*

$$\mathcal{S}f = \frac{1}{\pi} \nabla \cdot (Q\pi \nabla f), \quad \mathcal{T}f = \frac{1}{\pi} \nabla \cdot (C\pi \nabla f) \quad (1.69)$$

satisfy $\mathcal{S}^* = \mathcal{S}$ and $\mathcal{T}^* = -\mathcal{T}$, where the adjoint is taken in $L^2(\pi)$. In particular, the process generated by (1.68) is reversible if and only if $C \equiv 0$. Denoting by $Q^{1/2}$ the unique positive semidefinite square root of Q and writing $A_i = \sum_{j=1}^d Q_{ij}^{1/2} \partial_j$, we have the representation

$$\mathcal{S} = \sum_{i=1}^d A_i^* A_i, \quad (1.70)$$

which will turn out to be useful in Part III, Chapter 11 in the construction of coupled processes. Let us point out that the process associated to (1.68) can be interpreted as a diffusion on a Riemannian manifold with metric tensor Q (see [10, Section 3.2] and [55, 96]).

Remark 1.8. *If (1.66) and (1.67) hold and $\mathcal{L}$ is hypoelliptic then it follows immediately from Theorem 1.6 that $(X_t)_{t \geq 0}$ is ergodic with unique invariant measure π .*

Generalising the above framework, we can consider Itô diffusions in an extended phase space of dimension $D \geq d$, ergodic with respect to a measure $\hat{\pi}$. We then require that $\hat{\pi}$ has marginal π , i.e.

$$\int_{\mathbb{R}^{d'}} \hat{\pi}(x, y) dy = \pi(x), \quad (1.71)$$

where $x \in \mathbb{R}^d$, $y \in \mathbb{R}^{d'}$ and $d + d' = D$.

There are various well-known choices of dynamics which are invariant (and indeed ergodic) with respect to the target measure π .

1. *Overdamped Langevin dynamics.* Choosing $b = I$ and $\gamma = 0$ we obtain

$$dX_t = -\nabla V(X_t)dt + \sqrt{2}dW_t, \quad (1.72)$$

see for instance [113, Section 4.5]. A suitable discretisation of (1.72) accompanied by a Metropolis-Hastings step leads to the popular *Metropolis adjusted Langevin algorithm* (MALA) [128, Section 6.5.2]. The analysis of (1.72) also plays a crucial role in modern approaches to optimisation such as *stochastic gradient descent* (SGD) [58, Chapter 2.8].

2. *Nonreversible overdamped Langevin dynamics.* Choosing $b = I$, and $\gamma \neq 0$ such that (1.67) holds gives rise to

$$dX_t = -(\nabla V(X_t) - \gamma(X_t)) dt + \sqrt{2} dW_t. \quad (1.73)$$

Following Remark 1.8, the dynamics (1.73) is ergodic with respect to π . In particular choosing $\gamma(x) = J\nabla V(x)$ for a constant skew-symmetric matrix J we obtain

$$dX_t = -(I + J)\nabla V(X_t) dt + \sqrt{2} dW_t, \quad (1.74)$$

which will be studied in more detail in Part II, Chapter 3.

3. *Preconditioned and geometric Langevin dynamics.* Choosing $\gamma = 0$ and taking an arbitrary positive definite matrix Q leads to the dynamics

$$dX_t = -Q\nabla V(X_t)dt + \sqrt{2Q}dW_t. \quad (1.75)$$

The matrix Q can be thought of as a preconditioner, taking account of different scales inherent in the potential V . More generally, we can consider $Q : \mathbb{R}^d \rightarrow \mathbb{R}_{\text{sym}}^{d \times d}$, taking values in the symmetric positive definite matrices, allowing for locally adaptive preconditioning. The dynamics associated to this choice read

$$dX_t = -Q\nabla V(X_t)dt + (\nabla \cdot Q(X_t)) dt + \sqrt{2Q}dW_t, \quad (1.76)$$

which can be interpreted as Langevin diffusion on a Riemannian manifold, the metric tensor of which is given by Q , see Remark 1.7.

4. *Underdamped Langevin dynamics.* For an arbitrary symmetric positive definite ‘mass’ matrix $M \in \mathbb{R}_{\text{sym}}^{d \times d}$, let us consider the augmented target density

$$\hat{\pi}(q, p) = \frac{1}{\hat{Z}} \exp \left(-V(q) - \frac{1}{2} p \cdot M^{-1} p \right), \quad (1.77)$$

where $(q, p) \in \mathbb{R}^{2d}$, and $\hat{Z}$ is an appropriate normalisation constant. A short calculation shows that (1.71) is satisfied. Choosing

$$\gamma((q, p)) = \begin{pmatrix} M^{-1}p \\ -\nabla V(q) \end{pmatrix} \quad \text{and} \quad b = \begin{pmatrix} \mathbf{0} \\ \sqrt{\Gamma} \end{pmatrix} \in \mathbb{R}^{2d \times d}, \quad (1.78)$$

where $\Gamma \in \mathbb{R}_{\text{sym}}^{d \times d}$ is a symmetric positive definite ‘friction’ matrix, the conditions of Theorem 1.7 are satisfied for the target density $\hat{\pi}$. The resulting dynamics $(q_t, p_t)_{t \geq 0}$

is determined by the underdamped Langevin equation

$$dq_t = M^{-1}p_t dt, \quad (1.79a)$$

$$dp_t = -\nabla V(q_t) dt - \Gamma M^{-1}p_t dt + \sqrt{2\Gamma} dW_t, \quad (1.79b)$$

It is straightforward to verify that the generator is hypoelliptic, [90, Sec 2.2.3.1], and thus $(q_t, p_t)_{t \geq 0}$ is ergodic. The popular Hamiltonian Monte Carlo algorithm [31, Chapter 5] can be roughly understood as a particular implementation of the dynamics (1.79), involving a splitting scheme between the Hamiltonian part and the noise (Ornstein-Uhlenbeck) part.

5. More generally, consider the augmented target density $\hat{\pi}$ on $\mathbb{R}^{2d}$ as above, and choose

$$\gamma((q, p)) = \begin{pmatrix} M^{-1}p - \mu J_1 \nabla V(q) \\ -\nabla V(q) - \nu J_2 M^{-1}p \end{pmatrix} \quad \text{and} \quad b = \begin{pmatrix} \mathbf{0} \\ \sqrt{\Gamma} \end{pmatrix} \in \mathbb{R}^{2d \times d}, \quad (1.80)$$

where μ and ν are scalar constants and $J_1, J_2 \in \mathbb{R}_{\text{skew}}^{d \times d}$ are constant skew-symmetric matrices. With this choice we arrive at the perturbed underdamped Langevin dynamics

$$dq_t = M^{-1}p_t dt - \mu J_1 \nabla V(q_t) dt, \quad (1.81a)$$

$$dp_t = -\nabla V(q_t) dt - \nu J_2 M^{-1}p_t dt - \Gamma M^{-1}p_t dt + \sqrt{2\Gamma} dW_t, \quad (1.81b)$$

which is the main object of study in Part II. It is straightforward to check that (1.80) satisfies the invariance condition (1.67), and thus Theorem 1.7 guarantees that (1.81) is invariant with respect to $\hat{\pi}$.

6. In a similar fashion, one can introduce an augmented target density on $\mathbb{R}^{(m+2)d}$, with

$$\hat{\hat{\pi}}(q, p, u_1, \dots, u_m) \propto e^{-\left(\frac{|p|^2}{2} + \frac{|u_1|^2 + \dots + |u_m|^2}{2} + V(q)\right)},$$

where $p, q, u_i \in \mathbb{R}^d$, for $i = 1, \dots, m$. Clearly

$$\int_{\mathbb{R}^d \times \mathbb{R}^{md}} \hat{\hat{\pi}}(q, p, u_1, \dots, u_m) dp du_1 \dots du_m = \pi(q).$$

We now define $\gamma : \mathbb{R}^{(m+2)d} \rightarrow \mathbb{R}^{(m+2)d}$ by

$$\gamma(q, p, u_1, \dots, u_m) = \begin{pmatrix} p & -\nabla_q V(q) + \sum_{j=1}^m \lambda_j u_j & -\lambda_1 p & \dots & -\lambda_m p \end{pmatrix}^T$$

and $b : \mathbb{R}^{(m+2)d} \rightarrow \mathbb{R}^{(m+2)d \times (m+2)d}$ by

$$b(q, p, u_1, \dots, u_m) = \begin{pmatrix} \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \dots & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \dots & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \sqrt{\alpha_1} I_{d \times d} & \mathbf{0} & \dots & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \sqrt{\alpha_2} I_{d \times d} & \dots & \mathbf{0} \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \dots & \sqrt{\alpha_m} I_{d \times d} \end{pmatrix},$$

where $\lambda_i \in \mathbb{R}$ and $\alpha_i > 0$, for $i = 1, \dots, m$. The resulting process (1.60) is given by

$$\begin{aligned} dq_t &= p_t dt, & dp_t &= -\nabla_q V(q_t) dt + \sum_{j=1}^d \lambda_j u^j(t) dt, \\ du_t^1 &= -\lambda_1 p_t dt - \alpha_1 u_t^1 dt + \sqrt{2\alpha_1} dW_t^1, \\ &\vdots \\ du_t^m &= -\lambda_m p_t dt - \alpha_m u_t^m dt + \sqrt{2\alpha_m} dW_t^m, \end{aligned} \tag{1.82}$$

where $(W_t^j)_{t \geq 0, j=1, \dots, m}$ are independent $\mathbb{R}^d$ -valued Brownian motions. This process is ergodic with unique invariant distribution $\hat{\pi}$, and under appropriate conditions on V , converges exponentially fast to equilibrium in relative entropy [109]. Equation (1.82) is a Markovian representation of a generalised Langevin equation of the form

$$dq_t = p_t dt, \quad dp_t = -\nabla_q V(q_t) dt - \int_0^t F(t-s) p_s ds + N(t),$$

where $N(t)$ is a mean-zero stationary Gaussian process with autocorrelation function $F(t)$, i.e. $\mathbb{E}[N(t) \otimes N(s)] = F(t-s) I_{d \times d}$ and $F(t) = \sum_{i=1}^m \lambda_i^2 e^{-\alpha_i |t|}$.

7. Let $\tilde{\pi}(z) \propto \exp(-\Phi(z))$ be a positive density on $\mathbb{R}^D$ where $D > d$ such that

$$\pi(x) = \int_{\mathbb{R}^{D-d}} \tilde{\pi}(x, y) dy,$$

with $(x, y) \in \mathbb{R}^d \times \mathbb{R}^{D-d}$. Then choosing $b = I_{D \times D}$ and $\gamma = 0$ we obtain the dynamics

$$dX_t = -\nabla_x \Phi(X_t, Y_t) dt + \sqrt{2} dW_t^x, \quad dY_t = -\nabla_y \Phi(X_t, Y_t) dt + \sqrt{2} dW_t^y,$$

for two independent Brownian motions $(W_t^x)_{t \geq 0}$ and $(W_t^y)_{t \geq 0}$. The joint process $(X_t, Y_t)_{t \geq 0}$ is immediately seen to be ergodic with respect to $\tilde{\pi}$.

1.3.2 THE ZIGZAG PROCESS

In recent years, there has been a growing interest in using piecewise deterministic Markov processes (PDMPs) [36] in the context of sampling problems. These are processes that move deterministically between random events, usually along the trajectories of an ODE. At those events, a random transition (e.g. a ‘jump’) occurs. Both the deterministic dynamics as well as the random transitions can be chosen with a great deal of flexibility, resulting in a range of possible PDMP algorithms. Let us mention here the bouncy particle sampler (BPS) [28], the zigzag sampler [22], randomised Hamiltonian Monte Carlo [27], and event-chain Monte Carlo techniques [102, 103]. The recent papers [50] and [139] provide good overviews in a general framework.

In this thesis we will concentrate on the zigzag process in order to illustrate some of our results. Hence, let us give a brief outline of its dynamics. The exposition here follows [22] and [23] in style and notation. Similarly to underdamped Langevin dynamics (see (1.79)), the zigzag process targets a probability measure in an enlarged state space. Specifically, consider $E = \mathbb{R}^d \times \{-1, 1\}^d$, and denote by κ the uniform probability measure on $\{-1, 1\}^d$. The extended target measure is then given by

$$\widehat{\pi} = \pi \otimes \kappa, \quad (1.83)$$

where we recall that π has been defined in (1.58). A crucial ingredient are the ‘flip operators’ $F_i : \{-1, 1\}^d \rightarrow \{-1, 1\}^d$ given by

$$(F_i \theta)_j = \begin{cases} \theta_j & j \neq i \\ -\theta_j & j = i. \end{cases} \quad (1.84)$$

Furthermore, let us introduce the *switching rates* $\lambda_i : E \rightarrow [0, \infty)$, satisfying the following relation with respect to the potential V :

$$\lambda_i(x, \theta) - \lambda_i(x, F_i \theta) = \theta_i \partial_i V(x), \quad (x, \theta) \in E, i = 1, \dots, d. \quad (1.85)$$

Equivalently (see [22]), there exists a continuous function $\gamma : E \rightarrow [0, \infty)^d$, satisfying

$$\gamma_i(x, F_i \theta) = \gamma_i(x, \theta), \quad i = 1, \dots, d, \quad (1.86)$$

and

$$\lambda_i(x, \theta) = (\theta \partial_i V(x))_+ + \gamma_i(x, \theta), \quad (1.87)$$

where $a_+ = \max(a, 0)$. Following [23], we call γ the *excess switching rate* and $\gamma \equiv 0$ the

canonical switching rate. The zigzag process is defined by its generator

$$\mathcal{L}f(x, \theta) = \sum_{i=1}^d (\theta_i \partial_i f(x, \theta) + \lambda_i(x, \theta) \cdot (f(x, F_i \theta) - f(x, \theta))), \quad (x, \theta) \in E, \quad (1.88)$$

acting on functions $f \in \mathcal{D}$, where

$$\mathcal{D} = \left\{ f : E \rightarrow \mathbb{R} : f(\cdot, \theta) \in C_c^\infty(\mathbb{R}^d) \text{ for all } \theta \in \{-1, 1\}^d \right\} \quad (1.89)$$

is a core for $\mathcal{L}$.

Roughly speaking, the zigzag process moves along straight lines in directions determined by $\theta \in \{-1, 1\}^d$. At random times a switch occurs, i.e. θ_i is replaced by $-\theta_i$ for some $i \in \{1, \dots, d\}$. Those events are sampled according to the switching rates λ_i , i.e. at a point $(x, \theta) \in E$, the probability for the switch $\theta_i \mapsto -\theta_i$ in the time span $[t, t + \varepsilon]$ is given by $\lambda_i(x, t)\varepsilon + o(\varepsilon)$. For more details on the construction and simulation of zigzag processes we refer the reader to [21] and [22], as well as to [36] for more general piecewise deterministic Markov processes.

It is straightforward (see [22, Theorem 2.2]) to show that

$$\int_E \mathcal{L}f \, d\hat{\pi} = 0, \quad f \in \mathcal{D}, \quad (1.90)$$

i.e. $\hat{\pi}$ is indeed an invariant measure for the zigzag process. Ergodicity is much harder to show and will be addressed in Part IV, Chapter 17. In Part III, we will show how to use coupling ideas in the context of PDMP samplers.

For future reference, it will be convenient to have the decomposition of $\mathcal{L}$ into $L^2(\hat{\pi})$ -symmetric and skew-symmetric parts available (see (1.38)). To this end, let us introduce the operators

$$R_i : L^2(\hat{\pi}) \rightarrow L^2(\hat{\pi}), \quad (R_i f)(x, \theta) = f(x, F_i \theta), \quad i \in \{1, \dots, d\}. \quad (1.91)$$

Then, the symmetric and skew-symmetric parts can be expressed in the form

$$\mathcal{S} = \frac{1}{2}(\mathcal{L} + \mathcal{L}^*) = \frac{1}{2} \sum_{i=1}^d (|\partial_i V| + 2\gamma_i) (R_i - I), \quad (1.92a)$$

$$\mathcal{T} = \frac{1}{2}(\mathcal{L} - \mathcal{L}^*) = \frac{1}{2} \sum_{i=1}^d (2\theta_i \partial_i + \theta_i (\partial_i V) (R_i - I)), \quad (1.92b)$$

see Chapter 17, Section 17.1 for more details. Furthermore, $\mathcal{S}$ admits the representation $\mathcal{S} = \sum_{i=1}^d A_i^* A_i$, with

$$A_i = \frac{1}{2} (|\partial_i V| + 2\gamma_i)^{1/2} (R_i - I), \quad (1.93)$$

in direct analogy with (1.70) in the context of diffusion processes.

1.3.3 MCMC WITH MANY PARTICLES: COUPLING

Taking into account recent advances in computing architecture, it is tempting to run the dynamics put forward in Sections 1.3.1 and 1.3.2 in parallel. For instance, in [88] it was proposed to propagate an ensemble of paths following the underdamped Langevin dynamics (1.79) and adjust the parameters $M \in \mathbb{R}_{\text{sym}}^{d \times d}$ and $\Gamma \in \mathbb{R}_{\text{sym}}^{d \times d}$ according to the empirical covariance of neighbouring replicas. Another approach is to use coupled noise for the different copies of the dynamics in order to produce anti-correlated samples. Here, we will give an illustrative example of this idea; the topic will be explored in much greater detail in Part III.

Let $n \in \mathbb{N}$ be the number of particles (or ‘walkers’) and consider a matrix-valued function of n arguments $G : (\mathbb{R}^d)^n \rightarrow \mathbb{R}^{nd \times nd}$. We assume that G is symmetric at every point, i.e.

$$G(x_1, \dots, x_n)^T = G(x_1, \dots, x_n), \quad (x_1, \dots, x_n) \in (\mathbb{R}^d)^n. \quad (1.94)$$

Furthermore, let us view G as an $n \times n$ -matrix with $\mathbb{R}^{d \times d}$ -entries, i.e. G_{ij} is a $d \times d$ -matrix for $i, j \in \{1, \dots, n\}$. With this notation, we require that

$$\sum_{j=1}^n G_{ij}(x_1, \dots, x_n) G_{ij}(x_1, \dots, x_n)^T = I_{d \times d} \quad (1.95)$$

for all $i \in \{1, \dots, n\}$. Consider the ‘interacting’ overdamped Langevin dynamics for n particles:

$$dX_t^i = -\nabla V(X_t^i) dt + \sqrt{2} \sum_{j=1}^n G_{ij}(X_t^1, \dots, X_t^n) dB_t^j, \quad i = 1, \dots, n, \quad (1.96)$$

together with the extended observable

$$F(x_1, \dots, x_n) = \frac{1}{n} \sum_{i=1}^n f(x_i). \quad (1.97)$$

Using (1.94) and (1.95), it can be shown by Levy’s characterisation of Brownian motion that (1.96) can be written as

$$dX_t^i = -\nabla V(X_t^i) dt + \sqrt{2} d\tilde{B}_t^i, \quad i = 1, \dots, n, \quad (1.98)$$

where $(\tilde{B}_t^i)_{t \geq 0}$ are standard Brownian motions that are however not necessarily independent. Since for each $i \in \{1, \dots, n\}$, the process $(X_t^i)_{t \geq 0}$ is ergodic with respect to π , the coupled dynamics $(X_t^1, \dots, X_t^n)_{t \geq 0}$ is (under certain conditions) expected to be ergodic

with respect to a measure $\bar{\pi}$ on $(\mathbb{R}^d)^n$ satisfying

$$\int_{(\mathbb{R}^d)^n} \bar{\pi}(x_1, \dots, x_n) dx_1 \dots \widehat{dx_i} \dots dx_n = \pi(x_i), \quad i \in \{1, \dots, n\}. \quad (1.99)$$

In other words, $\bar{\pi}$ is a coupling of n copies of π . Observe furthermore that

$$\lim_{T \rightarrow \infty} \frac{1}{T} \int_0^T F(X_t^1, \dots, X_t^n) dt = \int_{\mathbb{R}^d} f d\pi, \quad (1.100)$$

i.e. the dynamics (1.96) in conjunction with the observable (1.97) leads to an asymptotically correct estimator. It should be noted that this approach relies on a Markov process targeting an augmented invariant measure and is hence somewhat related to the underdamped Langevin dynamics (1.79) (leaving (1.77) invariant) as well as to the zigzag process (leaving (1.83) invariant). An obvious question is which choices of the matrix-valued function $G : (\mathbb{R}^d)^n \rightarrow \mathbb{R}^{nd \times nd}$ lead to efficient sampling algorithms, both in terms of asymptotic variance and speed of convergence to equilibrium. This and similar problems will be studied extensively in Part III.

1.4 OUTLINE OF THE THESIS

Let us come back to the sampling problem introduced at the very beginning of this introduction, i.e. to the task of computing the integral $\int_{\mathbb{R}^d} f d\pi$ by means of the long-time ergodic average

$$\lim_{T \rightarrow \infty} \frac{1}{T} \int_0^T f(X_t) dt = \int_{\mathbb{R}^d} f d\pi. \quad (1.101)$$

In Section 1.1 we have seen that suitable time-continuous Markov processes $(X_t)_{t \geq 0}$ can be described effectively in terms of the corresponding semigroups $(S_t)_{t \geq 0}$ and their generators $(\mathcal{L}, \mathcal{D}(\mathcal{L}))$. Note in particular that $\mathcal{L}$ and π have to be related through the stationary Fokker-Planck equation (1.27) in order for (1.101) to be valid. As Section 1.2 showed, the speed of convergence to equilibrium (quantified in terms of the constants C and λ in (1.44)), as well as the asymptotic variance (quantified in terms of central limit theorems of the form (1.52)) are appropriate (and complementary) measures for the accuracy of MCMC estimators. Moreover, they are directly related to the spectral gap of $\mathcal{L}$ (see (1.49)) and to the study of the solutions to the associated Poisson equation (1.54). In Section 1.3 we presented a wide variety of Markov processes for which (1.101) holds. This thesis shall make a contribution to developing the mathematical understanding according to which a Markov process among those shall be chosen. As a guideline, we can informally state our objective as follows, highlighting its essence as an inverse problem:

Problem 1.1. *Let $\pi \in \mathcal{P}(\mathbb{R}^d)$ be a target probability measure of interest. Find a generator*

$\mathcal{L}$ of a Markov process satisfying

$$\mathcal{L}^\dagger \pi = 0, \quad (1.102)$$

such that the spectral gap $\lambda_{\text{gap}}(\mathcal{L})$ is maximised and the asymptotic variance $\sigma_f^2(\mathcal{L})$ is minimised, either for a particular observable f or across a class of observables. More generally, we can consider an augmented target measure $\hat{\pi} \in \mathcal{P}(\mathbb{R}^D)$, with $D > d$, such that its marginal on $\mathbb{R}^d$ is given by π .

We wish to emphasize that Problem 1.1 is not well posed as stated here, mainly because it does not take the computational cost of simulating a given Markov process into account[‡]. Indeed, observe that under the time rescaling $\mathcal{L} \mapsto \tau\mathcal{L}$ the stationary Fokker-Planck equation (1.102) remains invariant, while the spectral gap and the asymptotic variance behave as

$$\lambda_{\text{gap}}(\tau\mathcal{L}) = \tau\lambda_{\text{gap}}(\mathcal{L}), \quad \sigma_f^2(\tau\mathcal{L}) = \frac{1}{\tau}\sigma_f^2(\mathcal{L}),$$

i.e. they improve if $\tau > 1$. Clearly, this rescaling does not lead to an improved sampler however since it requires a smaller discretisation time step when applied in practice.

Let us now describe the structure and main results of the thesis.

- Part II is dedicated to the study of nonreversible perturbations of well-established sampling schemes, in particular to those based on overdamped and underdamped Langevin dynamics (see (1.72) and (1.79)). In Chapter 3 we revisit the nonreversible overdamped Langevin dynamics (1.73) that has already been the subject of various works [41, 71, 72, 89, 144]. In particular, the authors of [89] have performed a detailed study of the spectral gap corresponding to the generator in the Gaussian case. Our contribution (Theorems 3.1 and 3.2) is a similar decay estimate in Wasserstein distances, recovering the same rate λ , but achieving sharper control of the constant C , especially in high-dimensional settings. The bulk of Part II is concerned with the perturbed underdamped Langevin dynamics (1.81), based on the published work of the author [42]. In Chapter 4 we investigate the ergodic properties of this dynamics and prove exponential decay to equilibrium using the theoretical framework developed in [141]. Moreover, we obtain an overdamped limit, connecting the perturbed overdamped and underdamped dynamics. The perturbed underdamped Langevin process introduces a number of parameters in addition to the mass and friction tensors which must be tuned to ensure that the process is an efficient sampler. In Chapter 5 we hence consider the task of sampling from a Gaussian measure, analysing the limiting cases of both small and large perturbations. In this way, we obtain guidelines and algorithmic procedures for adjusting the parameters of the dynamics according to the covariance of the target distribution, leading to a decrease in

[‡]In general, efficient numerical implementations are not discussed at great length in this work (see however Chapter 6). We take the opportunity to refer the reader to [43] and [120] where suitable algorithms aimed towards the simulation of dynamics similar to those featuring in Part II are developed.

asymptotic variance. While these results are valid for Gaussian target densities, we advocate these particular parameter choices also for more complex target measures. To demonstrate their efficacy, in Chapter 6 we present a discretisation scheme for the simulation of perturbed underdamped Langevin dynamics and perform numerical experiments in the context of diffusion bridge sampling. The corresponding target measure is a perturbation of a high-dimensional Gaussian distribution, allowing us to connect this setting to our results from Chapter 5. Chapter 8 contains additional proofs and supplementary material. For a more detailed overview of the results concerning the perturbed underdamped Langevin dynamics we refer the reader to the introductory section of Chapter 4.

- Part III will pick up the thread from the example presented in Section 1.3.3 and develop a general framework for the construction of coupled samplers based on the ergodic Markov processes considered in Sections 1.3.1 and 1.3.2. Chapter 9 gives a survey of this approach and provides an illustration of the main results in terms of a simple example. We refer the reader to this chapter for a more thorough exposition of the obtained results. Chapter 10 lays the foundation of the developed theory, giving a characterisation of coupled processes in terms of a particular set of operators (see Proposition 10.1) and providing a practical approach to the construction of couplings (see Section 10.4). In Chapter 11 we show how various examples fit into the general framework including coupled overdamped and underdamped Langevin dynamics as well as zigzag processes. Chapter 12 contains a central limit theorem for coupled processes, yielding an explicit expression for the associated asymptotic variance. It turns out that the latter can be given a natural interpretation in terms of the theory of optimal transportation. This connection is explored in detail in Section 12.2. In Chapter 13 we introduce a perturbative approach to finding optimal couplings in terms of reducing the asymptotic variance. The results are illustrated in terms of numerical examples. The topic of Chapter 14 is the study of the convergence to equilibrium for coupled processes. In particular, we derive a suitable Poincaré inequality adapted to our setting.
- Part IV is devoted to the study of exponential convergence to equilibrium for the zigzag process introduced in Section 1.3.2, using the hypocoercivity techniques developed in [40] in the context of underdamped Langevin dynamics. Under the assumption that the state space is compact and given a certain nonvanishing condition on the excess switching rate, we are able to prove a decay estimate of the form (1.46).

Part II

Nonreversibility

2. NONREVERSIBLE PERTURBATIONS AND SAMPLING — AN OVERVIEW

Reversibility plays a central role in the study of Markovian dynamics, connecting the stochastic behaviour of a process under consideration (namely the invariance of its law under time reversal) with functional-analytic properties of the corresponding generator (namely self-adjointness). This fundamental relationship has already been exposed in Theorem 1.2; furthermore, the decomposition of the generator into symmetric and skew-symmetric parts (and therefore into reversible and nonreversible contributions) was a recurring theme in Section 1.3.

In order to subsume the work in this part of the thesis under a common framework, it is helpful to generalise the notion of reversibility. To this end, let us assume that the Markov process $(X_t)_{t \geq 0}$ is ergodic with respect to $\pi \in \mathcal{P}(E)$, with associated strongly continuous semigroup $(S_t)_{t \geq 0}$ and generator $\mathcal{L}$ acting on $L^2(\pi)$. Moreover, suppose that $\Psi : E \rightarrow E$ is a bijective measurable involution (meaning that $\Psi \circ \Psi = I$) that leaves π invariant, i.e.

$$\int_E f \, d\pi = \int_E f \circ \Psi \, d\pi, \quad f \in \mathcal{B}_b(E). \quad (2.1)$$

Clearly, Ψ induces a bounded linear operator $\bar{\Psi} : L^2(\pi) \rightarrow L^2(\pi)$ acting via $f \mapsto f \circ \Psi$. Following [90, Section 2.2.1.2], we then say that $(X_t)_{t \geq 0}$ is Ψ -reversible (or reversible up to Ψ) if the semigroup $(S_t)_{t \geq 0}$ is Ψ -symmetric in $L^2(\pi)$, i.e. if

$$\langle f, \bar{\Psi} S_t g \rangle_{L^2(\pi)} = \langle S_t f, \bar{\Psi} g \rangle_{L^2(\pi)}, \quad t \geq 0, \quad f, g \in L^2(\pi), \quad (2.2)$$

or, equivalently, if the generator $\mathcal{L}$ is $\bar{\Psi}$ -symmetric,

$$\langle f, \bar{\Psi} \mathcal{L} g \rangle_{L^2(\pi)} = \langle \mathcal{L} f, \bar{\Psi} g \rangle_{L^2(\pi)}, \quad t \geq 0, \quad f, g \in \mathcal{D}(\mathcal{L}). \quad (2.3)$$

We remark that relative symmetry (such as Ψ -symmetry) of operators has been investigated extensively in [69]. On the level of the stochastic process, Ψ -reversibility means that if the process is started in stationarity (i.e. $X_0 \sim \pi$), then $(X_t)_{0 \leq t \leq T}$ and $(\Psi(X_{T-t}))_{0 \leq t \leq T}$ have the same law, for all $T > 0$.

The concept of Ψ -reversibility is a direct generalisation of that of ordinary reversibility; indeed a I -reversible process is just reversible in the sense of Definition 1.6. One of the most

topical representatives of reversible Markov processes in our context is the overdamped Langevin dynamics (1.72). The following examples illustrate the notion of Ψ -reversibility (with nontrivial involutions).

Example 2.1 (Underdamped Langevin dynamics). Introducing the momentum flip operator

$$P(q, p) = (q, -p), \quad (q, p) \in \mathbb{R}^{2d}, \quad (2.4)$$

it is well-known (see [90, Section 2.2.3] or Lemma 5.2(a) below) that the underdamped Langevin dynamics (1.79) is P -reversible, i.e. its generator $\mathcal{L}_0$ has the symmetry property

$$\langle f, \bar{P}\mathcal{L}_0 g \rangle_{L^2(\hat{\pi})} = \langle \mathcal{L}_0 f, \bar{P}g \rangle_{L^2(\hat{\pi})}, \quad f, g \in \mathcal{D}(\mathcal{L}_0). \quad (2.5)$$

The notation $\mathcal{L}_0$ has been chosen in anticipation of the developments below, suggesting that the corresponding dynamics should be considered as ‘unperturbed’. For any $T > 0$, and under the assumption that $(q_0, p_0) \sim \hat{\pi}$ (where the invariant measure $\hat{\pi}$ has been defined in (1.77)), the processes $(q_t, p_t)_{0 \leq t \leq T}$ and $(q_{T-t}, -p_{T-t})_{0 \leq t \leq T}$ have the same law.

Example 2.2 (Zigzag process). Assuming that the excess switching rate $\gamma : E \rightarrow \mathbb{R}$ does not depend on θ , the zigzag process (see Section 1.3.2) is reversible with respect to

$$P(x, \theta) = (x, -\theta), \quad (x, \theta) \in E. \quad (2.6)$$

Let us now comment on the role of reversibility (and nonreversibility) in the context of Markov chain Monte Carlo sampling, focussing for the moment on ordinary reversibility (i.e. I -reversibility). It has been observed in various works [20, 41, 71, 72, 73, 89, 108, 126, 127, 138] that nonreversible processes have improved convergence properties in terms of asymptotic variance and spectral gap. To give an impression, we recall here the setting and results laid out in [127]. The framework involves a fixed reversible process, the generator of which is perturbed by a skew-adjoint operator in such a way that the target measure of interest is still invariant. Clearly, the process obtained in this way is nonreversible. To be precise, the authors of [127] work under the following assumption:

Assumption 2.1. *Suppose that $\mathcal{L}_0$ is the generator of a reversible stochastic process, ergodic with respect to $\pi \in \mathcal{P}(E)$. Assume furthermore that 0 is an isolated point of $\sigma(\mathcal{L})$, i.e. $\lambda_{\text{gap}}(\mathcal{L}) > 0$. Suppose that $\mathcal{A}$ is skew-symmetric in $L^2(\pi)$, i.e.*

$$\langle f, \mathcal{A}g \rangle_{L^2(\pi)} = -\langle \mathcal{A}f, g \rangle_{L^2(\pi)}, \quad f, g \in \mathcal{D}(\mathcal{A}). \quad (2.7)$$

Assume that $\mathcal{L} = \mathcal{L}_0 + \mathcal{A}$ is the generator of a Markov process, ergodic with respect to π , and that $\mathcal{D}(\mathcal{L}) \subset \mathcal{D}(\mathcal{A})$.

Remark 2.1. *In order for the process generated by $\mathcal{L}$ to be ergodic with respect to π , it is*

clearly necessary that

$$\int_E \mathcal{A}f \, d\pi = 0, \quad f \in \mathcal{D}(\mathcal{A}), \quad (2.8)$$

or, in the case when π is absolutely continuous with respect to a reference measure,

$$\mathcal{A}^\dagger \pi = 0. \quad (2.9)$$

In this sense, the skew-symmetric perturbation has to leave the target measure invariant.

The following two results show that the performance of the sampler, measured in terms of spectral gap and asymptotic variance, is improved by the perturbation:

Theorem 2.1 (Spectral gap for nonreversible perturbations). *Let Assumption 2.1 be satisfied. Then*

$$\lambda_{\text{gap}}(\mathcal{L}_0 + \mathcal{A}) \geq \lambda_{\text{gap}}(\mathcal{L}_0). \quad (2.10)$$

This result is a direct consequence of the fact that the validity of the Poincaré inequality (1.51) in Theorem 1.3 does not depend at all on the skew-symmetric part of the generator. We wish to emphasize that the equivalence between (1.51) and (1.46) only holds in the case when $C = 1$. In other words, by adding a nonreversible perturbation $\mathcal{A}$, one may hope to achieve an improved exponential convergence rate λ at the cost of allowing a constant $C > 1$. Quantifying this trade-off is challenging and will be addressed in Section 3.1 (in the context of overdamped Langevin dynamics targeting a Gaussian measure).

Theorem 2.2 (Asymptotic variance for nonreversible perturbations). *Let Assumption 2.1 be satisfied and assume that the operator $(-\mathcal{L})^{-1/2} \mathcal{A} (-\mathcal{L})^{-1/2}$ is bounded. Then*

$$\sigma_f^2(\mathcal{L}_0 + \mathcal{A}) \leq \sigma_f^2(\mathcal{L}_0), \quad f \in L^2(\pi). \quad (2.11)$$

Versions of the foregoing result can be found in [41, 73, 127].

The primary example illustrating the framework established in Assumption 2.1 is the nonreversible overdamped Langevin dynamics as written in (1.74), the generator of which can be decomposed as $\mathcal{L} = \mathcal{L}_0 + \mathcal{A}$, where

$$\mathcal{L}_0 = -\nabla V \cdot \nabla + \Delta \quad \text{and} \quad \mathcal{A} = J \nabla V \cdot \nabla, \quad (2.12)$$

recalling that J is a skew-symmetric matrix, i.e. $J^T = -J$. Under appropriate growth on conditions on V (equivalently, decay conditions on the tails of π), it turns out that Assumption 2.1 is satisfied (especially, $\lambda_{\text{gap}} > 0$). For precise statements, see for instance [41]. We remark here that those conditions are satisfied in particular if π is a Gaussian measure. In those cases therefore, Theorems 2.1 and 2.2 immediately show that adding a nonreversible term to the dynamics never decreases the performance of the sampler when

measured in terms of spectral gap and asymptotic variance, independently of the chosen matrix J . Another (perhaps more striking) way of saying this is that the choice $J = 0$ is the worst* possible (at least according to the criteria based on the spectral gap and asymptotic variance).

Naturally, the question arises of whether the preceding ‘success story’ can be continued in the more general setting of Ψ -reversible processes. To be specific, consider the generator of the perturbed underdamped Langevin dynamics (1.81),

$$\mathcal{L} = \underbrace{M^{-1}p \cdot \nabla_q - \nabla_q V \cdot \nabla_p}_{\mathcal{L}_{\text{ham}}} \underbrace{- \Gamma M^{-1}p \cdot \nabla_p + \Gamma : \nabla_p \nabla_p}_{\mathcal{L}_{\text{therm}}} \underbrace{- \mu J_1 \nabla V \cdot \nabla_q - \nu J M^{-1}p \cdot \nabla_p}_{\mathcal{L}_{\text{pert}}}, \quad (2.13)$$

decomposed into the perturbation $\mathcal{L}_{\text{pert}}$ and the unperturbed operator $\mathcal{L}_0$, which can be further split into the Hamiltonian part $\mathcal{L}_{\text{ham}}$ and the thermostat (Ornstein-Uhlenbeck) part $\mathcal{L}_{\text{therm}}$, see [113, 90, 91]. As already observed in Example 2.1, $\mathcal{L}_0$ enjoys the P -reversibility property (2.5) with respect to the momentum flip (2.4). A short calculation (see Lemma 5.2(c) below) shows that

$$\langle f, \bar{P} \mathcal{L}_{\text{pert}} g \rangle_{L^2(\hat{\pi})} = -\langle \mathcal{L}_{\text{pert}} f, \bar{P} g \rangle_{L^2(\hat{\pi})}, \quad f, g \in \mathcal{D}(\mathcal{L}_{\text{pert}}), \quad (2.14)$$

i.e. $\mathcal{L}_{\text{pert}}$ is ‘skew-adjoint with respect to $\bar{P}$ ’. The setting is therefore very similar to the one described in Assumption 2.1, the main difference being that the scalar product $\langle \cdot, \cdot \rangle$ has been replaced by the indefinite symmetric bilinear form $\langle \cdot, \bar{P} \cdot \rangle_{L^2(\hat{\pi})}$. The detailed study of perturbed underdamped Langevin dynamics is the subject of the Chapters 4–6; for a more detailed summary of our results we refer the reader to the introductory section of Chapter 4. Let us already remark here that the analogues of Theorems 2.1 and 2.2 unfortunately fail to hold. However, we will be able to find suitable choices for the matrices J_1 , J_2 , M and Γ that improve the performance of the sampler in certain situations. Moreover, our analysis sheds light on the general principles underlying the design of optimal perturbations for Ψ -reversible dynamics.

*We remark that adding a nonreversible perturbation makes the accurate simulation of the ensuing dynamics computationally more intricate, hence in this sense, $J = 0$ is the ‘best choice’. The demands of finding a suitable dynamics that is both optimal in terms of (continuous-time) convergence properties as well as efficient implementations are therefore opposing. In [43], the authors discuss how to strike this balance in practice.

3. PERTURBED OVERDAMPED LANGEVIN DYNAMICS REVISITED

The topic of this chapter is a careful analysis of the decay properties of the nonreversible overdamped Langevin dynamics (1.74),

$$dX_t = -\nabla V(X_t) dt + J\nabla V(X_t) dt + \sqrt{2}dW_t, \quad (3.1)$$

where we recall that $J \in \mathbb{R}_{\text{skew}}^{d \times d}$ is a skew-symmetric matrix. This process has been the subject of various works [41, 71, 72, 73, 89, 125, 126]. On account of Theorems 2.1 and 2.2 in the preceding chapter, the perturbation $J\nabla V(X_t) dt$ generically leads to an improvement, both in terms of the spectral gap and the asymptotic variance. A pertinent question, however, is how to optimally choose the skew-symmetric matrix J . In the case when the target measure is Gaussian, i.e.

$$V(x) = \frac{1}{2}x \cdot Sx, \quad x \in \mathbb{R}^d, \quad (3.2)$$

where $S \in \mathbb{R}_{\text{sym}}^{d \times d}$ is a fixed symmetric positive definite matrix, this question has been studied in [89]. The authors show that it is possible to choose J in such a way that the decay estimate (1.46) holds with optimal decay rate $\lambda = \frac{\text{Tr } S}{d}$, and that a matrix J with this property can be obtained in an algorithmic way. However, their analysis leads to a complicated expression for the constant C that moreover tends to infinity when the dimension d grows large. A similar analysis (considering more general dynamics) for the decay in relative entropy was performed in [62]. In this chapter we obtain estimates for the convergence to equilibrium in Wasserstein distances, yielding a much more accessible expression for C that moreover is dimension-independent*. The method of the proof relies on coupling techniques that have been used in similar contexts (see [46] and references therein). A key additional ingredient is an appropriate change of the Euclidean metric (see Lemma 3.1 below), a technique that is reminiscent of those employed in Section 4.1 and Chapter 17 for establishing exponential convergence to equilibrium in $L^2(\pi)$.

The results in this chapter were obtained in collaboration with Claudio Bellani (now a PhD student at Imperial College London) on a visit to Prof. Andreas Eberle at the University of Bonn, the hospitality of which is greatly acknowledged.

*More precisely, the optimal value for C is determined by the condition number $\kappa(S)$, which, depending on the application, might introduce a dimension dependence.

3.1 EXPONENTIAL DECAY TO EQUILIBRIUM IN WASSERSTEIN DISTANCES

We begin by recalling the definitions of the Wasserstein distances under consideration:

Definition 3.1 (Couplings and Wasserstein distances). *Let $p \in [1, \infty)$. The space of probability measures on $\mathbb{R}^d$ with finite p -th moment will be denoted by $\mathcal{P}_p(\mathbb{R}^d)$, i.e.*

$$\mathcal{P}_p(\mathbb{R}^d) = \left\{ \pi \in \mathcal{P}(\mathbb{R}^d) : \int_{\mathbb{R}^d} |x|^p \pi(\mathrm{d}x) < \infty \right\}. \quad (3.3)$$

The set of couplings $\mathcal{C}(\mu, \nu) \subset \mathcal{P}(\mathbb{R}^d \times \mathbb{R}^d)$ between two probability measures $\mu, \nu \in \mathcal{P}(\mathbb{R}^d)$ consists of those probability measures whose first and second marginals are given by μ and ν respectively. The L^p -Wasserstein distances are then defined by

$$\mathcal{W}_p(\mu, \nu) = \left(\inf_{\gamma \in \mathcal{C}(\mu, \nu)} \int_{\mathbb{R}^d} |x - y|^p \gamma(\mathrm{d}x \mathrm{d}y) \right)^{\frac{1}{p}}. \quad (3.4)$$

Throughout, we will denote by $(X_t)_{t \geq 0}$ the solution to (3.1) and assume that the potential V is given by (3.2). Our main result shows that the optimal decay of the nonreversible dynamics as measured by the Wasserstein distances can be accurately quantified in terms of the condition number of the precision matrix S (in the following denoted by $\kappa(S)$):

Theorem 3.1 (Optimal convergence to equilibrium in Wasserstein distances). *Let $p \in [1, \infty)$ and denote $X_t \sim \pi_t$. For every $S \in \mathbb{R}_{\text{sym}}^{d \times d}$ and $\varepsilon > 0$, there exists a skew-symmetric matrix $J \in \mathbb{R}_{\text{skew}}^{d \times d}$ such that*

$$\mathcal{W}_p(\pi_t, \pi_0) \leq C e^{-\lambda t} \mathcal{W}_p(\pi_0, \pi), \quad (3.5)$$

where

$$\lambda = \frac{\text{Tr } S}{d}, \quad C \leq \sqrt{\kappa(S)} + \varepsilon. \quad (3.6)$$

Remark 3.1. *The skew-symmetric matrix J can be obtained by the same algorithm as introduced in [89]. Using this method, $J \rightarrow \infty$ as $\varepsilon \rightarrow 0$. In the limit, this would require a vanishing time-step for the discretisation. Hence, in practice there is a trade-off between optimising the constant C while maintaining a reasonable order of magnitude for the matrix J . The exponential rate λ is the same as the one obtained in [89] and [62] for the decay in $L^2(\pi)$ and relative entropy, respectively. Using [56, Proposition 3], it is possible to extend this result to the limiting case $p = \infty$.*

Remark 3.2 (Reversible versus nonreversible decay estimates). *In the reversible case (i.e. when $J = 0$), the decay estimate (3.5) holds with $\lambda = \min \text{Re } \sigma(S) \leq \frac{\text{Tr } S}{d}$ and $C = 1$. This is an instance of the trade-off between C and λ that was already alluded to in the discussion following Theorem 2.1.*

Before proving Theorem 3.1, we will present another result that might be of independent interest. Consider the linear SDE

$$dX_t = -BX_t dt + \sqrt{2} Q dW_t, \quad X_0 \sim \pi_0, \quad (3.7)$$

where $Q \in \mathbb{R}^{d \times d}$, $X_t \in \mathbb{R}^d$, and $B \in \mathbb{R}^{d \times d}$ is assumed to satisfy the following assumption:

Assumption 3.1. *The matrix B is diagonalisable and positively stable, i.e.*

$$\sigma(B) \subset (0, \infty) \times i\mathbb{R}.$$

Under Assumption 3.1, the process defined by (3.7) is ergodic with respect to a (possibly degenerate) Gaussian measure $\pi \in \mathcal{P}(\mathbb{R}^d)$.

Remark 3.3. *The condition that B is diagonalisable is not necessary, but it allows us to state the results in a simplified form. Our methods can be extended to the nondiagonalisable case by using for instance Lemma 4.7 in [5].*

We have the following contraction result for solutions of (3.7).

Theorem 3.2 (Wasserstein contraction). *Suppose that Assumption 3.1 is satisfied and let $p \in [1, \infty)$. Define*

$$\lambda = \min \operatorname{Re} \sigma(B). \quad (3.8)$$

Then there exists a constant $C > 0$, explicitly computable and only depending on B , such that

$$\mathcal{W}_p(\mu_t, \nu_t) \leq C e^{-\lambda t} \mathcal{W}_p(\mu_0, \nu_0), \quad (3.9)$$

where $\mu_0, \nu_0 \in \mathcal{P}_p(\mathbb{R}^d)$ are arbitrary initial laws for the dynamics (3.7), and μ_t, ν_t denote the laws of the corresponding solutions for $t \geq 0$.

Remark 3.4. *By choosing $\nu_0 = \pi$, we immediately obtain a decay estimate quantifying the rate of convergence to equilibrium of the form (1.44). The proof of the above theorem will provide an explicit expression for the constant C . In fact, it is given by $\sqrt{\kappa(G)}$, where $\kappa(G)$ is the condition number of the matrix G obtained in Lemma 3.1 below. Theorem 3.2 is analogous to [5, Theorem 1.3] (see also [6, Theorem 4.9]), where a similar decay estimate has been obtained for a class of relative entropies.*

For the proof of the foregoing theorem, we will need the following lemma:

Lemma 3.1. *Suppose Assumption 3.1 is satisfied and let λ be defined as in (3.8). Then there exists a symmetric positive definite matrix G such that*

$$GB + B^T G \geq 2\lambda G \quad (3.10)$$

in the sense of symmetric matrices.

Proof. Let $\{w_i\}_{i=1}^d$ be a (normalised) set of linearly independent eigenvectors of B . A short calculation shows that

$$G = \sum_{i=1}^d w_i \overline{w_i}^T \quad (3.11)$$

fulfills condition (3.10). For details, see the proof of Lemma 4.3 in [6]. $\square$

We now proceed to the proof of Theorem 3.2:

Proof of Theorem 3.2. Let G be a positive definite matrix satisfying (3.10), the existence of which is guaranteed by Lemma 3.1. Associated to G , we define the distorted norm

$$|x|_G = |G^{1/2}x| = \sqrt{x \cdot Gx}, \quad x \in \mathbb{R}^d, \quad (3.12)$$

and the corresponding Wasserstein distances

$$\mathcal{W}_{p,G}(\mu, \nu) = \left(\inf_{\gamma \in \mathcal{C}(\mu, \nu)} \int_{\mathbb{R}^d} |x - y|_G^p \gamma(\mathrm{d}x \mathrm{d}y) \right)^{\frac{1}{p}}. \quad (3.13)$$

Observe that

$$\sqrt{g_{\min}}|x| \leq |x|_G \leq \sqrt{g_{\max}}|x|, \quad x \in \mathbb{R}^d, \quad (3.14)$$

and correspondingly

$$\sqrt{g_{\min}}\mathcal{W}_p(\mu, \nu) \leq \mathcal{W}_{p,G}(\mu, \nu) \leq \sqrt{g_{\max}}\mathcal{W}_p(\mu, \nu), \quad \mu, \nu \in \mathcal{P}_p(\mathbb{R}^d), \quad (3.15)$$

where $g_{\min}$ and $g_{\max}$ denote the smallest and the largest eigenvalues of G respectively. Now consider two solutions $(X_t)_{t \geq 0}$ and $(Y_t)_{t \geq 0}$ of (3.7), with the initial distributions $X_0 \sim \mu_0$ and $Y_0 \sim \nu_0$, driven by the same Brownian motion. We realise those solutions on a joint probability space, i.e. the initial law of the joint process is given by a coupling of μ_0 and ν_0 . The difference $Z_t := X_t - Y_t$ satisfies the ODE

$$Z'_t = -BZ_t. \quad (3.16)$$

For $Z_t \neq 0$, we immediately deduce that

$$\frac{\mathrm{d}}{\mathrm{d}t}|Z_t|_G = -\frac{1}{2} \frac{Z_t \cdot (GB + B^T G)Z_t}{|Z_t|_G} \leq -\lambda|Z_t|_G, \quad (3.17)$$

using (3.10). By Gronwall's inequality, this implies

$$\mathbb{E}|Z_t|_G^p \leq e^{-\lambda p t} \mathbb{E}|Z_0|_G^p. \quad (3.18)$$

Clearly, $\mathbb{E}|Z_t|_G^p$ is an upper bound for $\mathcal{W}_{p,G}^p(\mu_t, \nu_t)$. Taking the infimum over couplings of

the initial laws μ_0 and ν_0 we hence arrive at

$$\mathcal{W}_{p,G}(\mu_t, \nu_t) \leq e^{-\lambda t} \mathcal{W}_{p,G}(\mu_0, \nu_0). \quad (3.19)$$

From the estimates (3.15) it follows that

$$\mathcal{W}_p(\mu_t, \nu_t) \leq C e^{-\lambda t} \mathcal{W}_p(\mu_0, \nu_0), \quad (3.20)$$

with $C = \sqrt{\frac{g_{\max}}{g_{\min}}} = \sqrt{\kappa(G)}$, where $\kappa(G)$ denotes the condition number of G . $\square$

Remark 3.5. *The proof technique of changing the norm (or the scalar product) in the context of nonreversible stochastic processes is now routinely applied in many contexts. To illustrate the idea, let us mention that the above proof without introducing the distorted Wasserstein distance $\mathcal{W}_{p,G}$ would allow us to deduce a contraction with rate*

$$\tilde{\lambda} = \min \operatorname{Re} \sigma \left(\frac{1}{2} (B + B^T) \right) \leq \lambda, \quad (3.21)$$

the spectral gap of the symmetric part of B , while the skew-symmetric part would drop out in the calculation. The latter is not skew-symmetric with respect to $|\cdot|_G$, allowing us to recover the improved contraction rate. The proofs in [5] and [6] proceed along similar lines, introducing distorted relative entropy functionals. We also refer to [11] and [62] for similar approaches. Although in a different context, the calculations in Sections 8.3 and Chapter 17 use very similar ideas.

We finally return to the proof of Theorem 3.1.

Proof of Theorem 3.1. The claim follows by carefully tracing the arguments in [89]. Introducing $Y_t = S^{1/2} X_t$ and $\tilde{J} = S^{1/2} J S^{1/2}$, the original dynamics

$$dX_t = -(I + J) S X_t dt + \sqrt{2} dW_t \quad (3.22)$$

is equivalent to

$$dY_t = -(S + \tilde{J}) Y_t dt + \sqrt{2} S^{1/2} dW_t. \quad (3.23)$$

The authors of [89] then proceed to show that it is possible to construct (algorithmically) an orthonormal basis $\{\psi_i\}_{i=1}^d$ of $\mathbb{R}^d$ such that

$$\psi_i \cdot S \psi_i = \frac{\operatorname{Tr} S}{d}, \quad i = 1, \dots, d. \quad (3.24)$$

Let G be the symmetric positive definite matrix with eigenvectors $\{\psi_i\}_{i=1}^d$ and associated eigenvalues $\{\lambda_i\}_{i=1}^d$, chosen such that $\lambda_i > 0$ and $\lambda_j \neq \lambda_k$ for $j \neq k$. The requirement

$$\psi_j \cdot \tilde{J} \psi_k = -\frac{\lambda_k + \lambda_j}{\lambda_k - \lambda_j} \psi_j \cdot S \psi_k, \quad j \neq k, \quad (3.25)$$

together with $\tilde{J}^T = -\tilde{J}$, determines $\tilde{J}$ uniquely. Furthermore,

$$\min \operatorname{Re} \sigma \left(S + \tilde{J} \right) = \frac{\operatorname{Tr} S}{d} \equiv \lambda, \quad (3.26)$$

and 3.10 holds (see [89, Section 3.2]). Clearly, by choosing the eigenvalues of G closely together, we can enforce that $\kappa(G) < 1 + \varepsilon$, for any $\varepsilon > 0$. By Theorem 3.2, it follows that the solutions to (3.23) satisfy

$$\mathcal{W}_p(\tilde{\pi}_t, \tilde{\pi}) \leq C e^{-\lambda t} \mathcal{W}_p(\tilde{\pi}_0, \tilde{\pi}), \quad (3.27)$$

where $Y_t \sim \tilde{\pi}_t$, $\tilde{\pi}$ denotes the corresponding invariant measure, and $1 < C < 1 + \varepsilon$. The result now follows from noticing that $\mathcal{W}_p(\tilde{\pi}_t, \tilde{\pi}) = \mathcal{W}_{p,S}(\pi_t, \pi)$ and the equivalence estimate (3.15). $\square$

3.2 DISCUSSION

As the results in this chapter show, the exponential decay rate λ for linear diffusions is remarkably stable in the sense that it does not depend on whether it is measured in terms of $L^2(\pi)$, relative entropies or Wasserstein distances. The same cannot be said about the constant C . In fact, its exact numerical value as well as its dependence on the dimensionality of the problem are highly sensitive to the chosen distance between probability distributions. According to Theorem 3.1 the Wasserstein distances are well suited to quantify the exact dependence of C on the nonreversible perturbation. This result opens up various avenues for further research. Firstly, the class of skew-symmetric matrices J that lead to an optimal exponential rate λ is relatively large (for instance, the algorithm presented in [89] contains a significant number of arbitrary parameters). It is therefore a natural problem to optimise the choice of J taking into account both λ and the estimate for C obtained in this chapter. Secondly, the proof of Theorem 3.2 appears to be valid also in infinite dimensions (under suitable conditions), i.e. in the context of SPDEs. This observation suggests us to employ and analyse the proposed methodology in the settings of inverse problems [135] as well as sampling techniques for Gaussian random fields [132]. Those topics are currently under investigation.

4. PERTURBED UNDERDAMPED LANGEVIN DYNAMICS

The objective of the remainder of Part II is to compare the performances of the perturbed underdamped Langevin dynamics (1.81) and the unperturbed dynamics (1.79) according to the criteria outlined in Section 1.2 (namely spectral gap and asymptotic variance). For the reader's convenience, let us recall that the former is defined by

$$dq_t = M^{-1}p_t dt - \mu J_1 \nabla V(q_t) dt, \quad (4.1a)$$

$$dp_t = -\nabla V(q_t) dt - \nu J_2 M^{-1}p_t dt - \Gamma M^{-1}p_t dt + \sqrt{2\Gamma}dW_t, \quad (4.1b)$$

where $M \in \mathbb{R}_{\text{sym}}^{d \times d}$ and $\Gamma \in \mathbb{R}_{\text{sym}}^{d \times d}$ denote the symmetric positive definite mass and friction matrices, $J_1, J_2 \in \mathbb{R}_{\text{skew}}^{d \times d}$ the skew-symmetric perturbations, and $\mu, \nu \in \mathbb{R}$ the corresponding perturbation strengths. We begin our investigations of (4.1) by establishing ergodicity and exponentially fast return to equilibrium in Section 4.1 and by studying the overdamped limit of (4.1) in Section 4.2. As the latter turns out to be nonreversible and therefore according to Theorems 2.1 and 2.2 in principle superior to the usual overdamped limit (1.72), this calculation provides us with further motivation to study the proposed dynamics.

For the bulk of this line of work, we consider the same setting as in Chapter 3 and focus on the particular case when the target measure is Gaussian, i.e. when the potential is given by $V(q) = \frac{1}{2}q \cdot Sq$ with a symmetric and positive definite precision matrix S (i.e. the covariance matrix is given by S^{-1}). Our analysis partly relies on the fact that under this assumption the dynamics (4.1) becomes a generalised Ornstein-Uhlenbeck process for which the entire spectrum of the generator – or, equivalently, the Fokker-Planck operator – can be computed analytically and, in particular, an explicit formula for the $L^2(\pi)$ -spectral gap can be obtained [101, 110, 111]. In the Gaussian setting, we advocate the following conditions for the choice of parameters:

$$M = S, \quad \Gamma = \gamma S, \quad SJ_1S = J_2, \quad \mu = \nu. \quad (4.2)$$

Under the above choices (4.2), we show that the large perturbation limit $\lim_{\mu \rightarrow \infty} \sigma_f^2$ exists and is finite and we provide an explicit expression for it (see Corollary 5.5). From this result, we derive an algorithm for finding optimal choices for J_1 in the case of quadratic observables (see Algorithm 2).

If the friction coefficient is not too small ($\gamma > \sqrt{2}$), and under certain mild nondegeneracy conditions, we prove that adding a small perturbation will always decrease the asymptotic variance for observables of the form $f(q) = q \cdot Kq + l \cdot q + C$:

$$\left. \frac{d}{d\mu} \sigma_f^2 \right|_{\mu=0} = 0 \quad \text{and} \quad \left. \frac{d^2}{d\mu^2} \sigma_f^2 \right|_{\mu=0} < 0,$$

see Theorem 5.1. In fact, we conjecture that this statement is true for arbitrary observables $f \in L^2(\pi)$, but we have not been able to prove this. The dynamics (4.1) (used in conjunction with the conditions (4.2)) proves to be especially effective when the observable is antisymmetric (i.e. when it is invariant under the substitution $q \mapsto -q$) or when it has a significant antisymmetric part. In particular, in Proposition 5.4 we show that under certain conditions on the spectrum of J_1 , for any antisymmetric observable $f \in L^2(\pi)$ it holds that $\lim_{\mu \rightarrow \infty} \sigma_f^2 = 0$.

Numerical experiments and analysis show that departing significantly from (4.2) in fact possibly decreases the performance of the sampler. This is in stark contrast to the perturbed overdamped Langevin dynamics (1.73), where it is not possible to decrease the spectral gap or increase the asymptotic variance by *any* perturbation (see Theorems 2.1 and 2.2). For that reason, until now it seems practical to use (4.1) as a sampler only when a reasonable estimate of the global covariance of the target distribution is available. In the case of Bayesian inverse problems and diffusion bridge sampling, the target measure π is given with respect to a Gaussian prior. In Chapter 6 we demonstrate the effectiveness of our approach in these applications, taking the prior Gaussian covariance as S in (4.2).

Remark 4.1. In [89, Remark 3] another modification of (1.79) was suggested (albeit with the simplifications $\Gamma = \gamma \cdot I$ and $M = I$):

$$dq_t = (I - J)M^{-1}p_t dt, \tag{4.3a}$$

$$dp_t = -(I + J)\nabla V(q_t) dt - \Gamma M^{-1}p_t dt + \sqrt{2\Gamma}dW_t, \tag{4.3b}$$

J again denoting a skew-symmetric matrix. However, under the change of variables $p \mapsto (I + J)\tilde{p}$ the above equations transform into

$$dq_t = \tilde{M}^{-1}p_t dt, \tag{4.4a}$$

$$d\tilde{p}_t = -\nabla V(q_t) dt - \tilde{\Gamma}\tilde{M}^{-1}\tilde{p}_t dt + \sqrt{2\tilde{\Gamma}}d\tilde{W}_t, \tag{4.4b}$$

where $\tilde{M} = (I + J)^{-1}M(I - J)^{-1}$ and $\tilde{\Gamma} = (I + J)^{-1}\Gamma(I - J)^{-1}$. Since any observable f depends only on q (the p -variables are merely auxiliary), the estimator $\tilde{\pi}_T(f)$ as well as its associated convergence characteristics (i.e. asymptotic variance and speed of convergence to equilibrium) are invariant under this transformation. Therefore, (4.3) reduces to the underdamped Langevin dynamics (1.79) and does not represent an independent approach to sampling. Suitable choices of M and Γ will be discussed in Section 5.5.

4.1 ERGODICITY AND EXPONENTIAL CONVERGENCE TO EQUILIBRIUM

In this section we study some of the ergodicity properties of the perturbed underdamped dynamics (4.1). We recall from (2.13) that its generator is of the form

$$\mathcal{L} = \underbrace{M^{-1}p \cdot \nabla_q - \nabla_q V \cdot \nabla_p}_{\mathcal{L}_{\text{ham}}} - \underbrace{\Gamma M^{-1}p \cdot \nabla_p + \Gamma : \nabla_p \nabla_p}_{\mathcal{L}_{\text{therm}}} - \underbrace{\mu J_1 \nabla V \cdot \nabla_q - \nu J M^{-1}p \cdot \nabla_p}_{\mathcal{L}_{\text{pert}}}. \quad (4.5)$$

We have the following regularity result:

Lemma 4.1. *The infinitesimal generator $\mathcal{L}$ is hypoelliptic.*

Proof. The proof consists of verifying the conditions of Hörmander's theorem (Theorem 1.5) for the generator (4.5) and can be found in Section 8.1. $\square$

An immediate corollary of this result and of Theorem 1.7 is that the perturbed underdamped Langevin process (4.1) is ergodic with unique invariant distribution $\hat{\pi}$ given by (1.77).

As explained in Section 1.2.2, the exponential decay estimate (1.46) is crucial for our approach, as in particular it guarantees the well-posedness of the Poisson equation (1.54). From now on, we will therefore make the following assumption on the potential V , required to prove exponential decay in $L^2(\pi)$:

Assumption 4.1. *Assume that the Hessian of V is bounded and that the target measure $\pi(dq) = \frac{1}{Z}e^{-V}dq$ satisfies a Poincaré inequality, i.e. there exists a constant $C_P > 0$ such that*

$$\left(\int_{\mathbb{R}^d} f^2 d\pi - \int_{\mathbb{R}^d} f d\pi \right)^2 \leq C_P \int_{\mathbb{R}^d} |\nabla f|^2 d\pi, \quad f \in H^1(\pi). \quad (4.6)$$

The validity of (4.6) is related to sufficiently fast growth of V at infinity (see for example [141, Theorem A.1]). Further conditions (including e.g. the Bakry-Emery criterion) are presented in [10, Chapter 4]. The fact that the inequalities (1.51) and (4.6) bear the same name is no coincidence. In fact, the former reduces to the latter when $\mathcal{L}$ is the generator of overdamped Langevin dynamics (1.72) and π is the corresponding invariant measure.

Theorem 4.1. *Under Assumption 4.1 there exist constants $C \geq 1$ and $\lambda > 0$ such that the semigroup $(S_t)_{t \geq 0}$ generated by $\mathcal{L}$ exhibits exponential decay in $L^2(\pi)$ as in (1.46).*

Proof. The proof uses the machinery of hypocoercivity developed in [141] and can be found in Section 8.3. Using the framework of [40] employed in Chapter 17, we conjecture that the assumption on the boundedness of the Hessian of V can be substantially weakened

and more quantitative decay estimates (in particular with respect to μ , ν and γ) can be obtained. $\square$

4.2 THE OVERDAMPED LIMIT

In this section we develop a connection between the perturbed underdamped Langevin dynamics (4.1) and the nonreversible overdamped Langevin dynamics (1.73). The analysis is very similar to the one presented in [90, Section 2.2.2] and we will be brief. For convenience in this section we will perform the analysis on the d -dimensional torus $\mathbb{T}^d \cong (\mathbb{R}/\mathbb{Z})^d$, i.e. we will assume $q \in \mathbb{T}^d$. Consider the following scaling of (4.1):

$$dq_t^\varepsilon = \frac{1}{\varepsilon} M^{-1} p_t^\varepsilon dt - \mu J_1 \nabla_q V(q_t) dt, \quad (4.7a)$$

$$dp_t^\varepsilon = -\frac{1}{\varepsilon} \nabla_q V(q_t^\varepsilon) dt - \frac{1}{\varepsilon^2} \nu J_2 M^{-1} p_t^\varepsilon dt - \frac{1}{\varepsilon^2} \Gamma M^{-1} p_t^\varepsilon dt + \frac{1}{\varepsilon} \sqrt{2\Gamma} dW_t, \quad (4.7b)$$

valid for the small mass/small momentum regime $M \rightarrow \varepsilon^2 M$, $p_t \rightarrow \varepsilon p_t$. Equivalently, those modifications can be obtained from substituting $\Gamma \rightarrow \varepsilon^{-1} \Gamma$ and $t \mapsto \varepsilon^{-1} t$, and so in the limit as $\varepsilon \rightarrow 0$ the dynamics (4.7) describes the limit of large friction with rescaled time. It turns out that as $\varepsilon \rightarrow 0$, the dynamics (4.7) converges to the limiting SDE

$$dq_t = -(\nu J_2 + \Gamma)^{-1} \nabla_q V(q_t) dt - \mu J_1 \nabla_q V(q_t) dt + (\nu J_2 + \Gamma)^{-1} \sqrt{2\Gamma} dW_t. \quad (4.8)$$

The following proposition makes this statement precise.

Proposition 4.1. *Denote by $(q_t^\varepsilon, p_t^\varepsilon)_{t \geq 0}$ the solution to (4.7) with (deterministic) initial conditions $(q_0^\varepsilon, p_0^\varepsilon) = (q_{\text{init}}, p_{\text{init}})$ and by $(q_t^0)_{t \geq 0}$ the solution to (4.8) with initial condition $q_0^0 = q_{\text{init}}$. For any $T > 0$, $(q_t^\varepsilon)_{0 \leq t \leq T}$ converges to $(q_t^0)_{0 \leq t \leq T}$ in $L^2(\Omega, C([0, T], \mathbb{T}^d))$ as $\varepsilon \rightarrow 0$, i.e.*

$$\lim_{\varepsilon \rightarrow 0} \mathbb{E} \left(\sup_{0 \leq t \leq T} |q_t^\varepsilon - q_t^0|^2 \right) = 0.$$

Proof. The proof can be found in Section 8.2. By a more refined analysis, it is also possible to get information on the rate of convergence; see, e.g. [114, 115]. $\square$

Remark 4.2. *The overdamped limit (4.8) respects the invariant distribution, in the sense that $\pi(dq) = \frac{1}{Z} e^{-V} dq$ is invariant. To see this, we have to check that*

$$\begin{aligned} \mathcal{L}^\dagger (e^{-V}) &= -\nabla_q \cdot ((\nu J_2 + \Gamma)^{-1} \nabla_q e^{-V}) + \nabla_q \cdot (\mu J_1 \nabla_q e^{-V}) \\ &\quad + \nabla_q \cdot ((\nu J_2 + \Gamma)^{-1} \Gamma (-\nu J_2 + \Gamma)^{-1} \nabla_q e^{-V}) = 0, \end{aligned}$$

where $\mathcal{L}^\dagger$ refers to the $L^2(\mathbb{R}^d)$ -adjoint of the generator of (4.8), i.e. to the associated Fokker-Planck operator. Indeed, the term $\nabla_q \cdot (\mu e^{-V} J_1 \nabla_q V)$ vanishes because of the skew-

symmetry of J_1 . Therefore, it remains to show that

$$\nabla_q \cdot ((\nu J_2 + \Gamma)^{-1} \Gamma (-\nu J_2 + \Gamma)^{-1} - (\nu J_2 + \Gamma)^{-1} \nabla_q e^{-V}) = 0,$$

i.e. that the matrix $(\nu J_2 + \Gamma)^{-1} \Gamma (-\nu J_2 + \Gamma)^{-1} - (\nu J_2 + \Gamma)^{-1}$ is antisymmetric. Clearly, the first term is symmetric and furthermore it turns out to be equal to the symmetric part of the second term:

$$\begin{aligned} \frac{1}{2}((\nu J_2 + \Gamma)^{-1} + (-\nu J_2 + \Gamma)^{-1}) &= \frac{1}{2}((\nu J_2 + \Gamma)^{-1}(-\nu J_2 + \Gamma)(-\nu J_2 + \Gamma)^{-1} \\ &\quad + (\nu J_2 + \Gamma)^{-1}(\nu J_2 + \Gamma)(-\nu J_2 + \Gamma)^{-1}) \\ &= (\nu J_2 + \Gamma)^{-1} \Gamma (-\nu J_2 + \Gamma)^{-1}, \end{aligned}$$

so π is indeed invariant under the limiting dynamics (4.8).

The limiting SDE (4.8) is nonreversible due to the term $-\mu J_1 \nabla_q V(q_t) dt$ and also because the matrix $(\nu J_2 + \Gamma)^{-1}$ is in general neither symmetric nor skew-symmetric. This result, together with the fact that nonreversible perturbations of overdamped Langevin dynamics of the form (1.73) have improved performance properties (see Theorems 2.1 and 2.2), motivates further investigation of the dynamics (4.1).

5. SAMPLING FROM A GAUSSIAN MEASURE

In this chapter we study in detail the performance of the Langevin sampler (4.1) for Gaussian target densities, first considering the case of unit covariance. In particular, we study the optimal choice for the parameters in the sampler, the exponential decay rate and the asymptotic variance. We then extend our results to Gaussian target densities with arbitrary covariance matrices.

5.1 UNIT COVARIANCE – SMALL PERTURBATIONS

In our study of the dynamics given by (4.1) we first consider the simple case when $V(q) = \frac{1}{2}|q|^2$, i.e. the task of sampling from a Gaussian measure with unit covariance. We will assume that $M = I$, $\Gamma = \gamma I$ and $J_1 = J_2 =: J$ (so that the q - and p -dynamics are perturbed in the same way, albeit possibly with different strengths μ and ν). Using these simplifications, (4.1) reduces to the linear system

$$\begin{aligned} dq_t &= p_t dt - \mu J q_t dt, \\ dp_t &= -q_t dt - \nu J p_t dt - \gamma p_t dt + \sqrt{2\gamma} dW_t. \end{aligned} \tag{5.1}$$

We will consider quadratic observables of the form

$$f(q) = q \cdot K q + l \cdot q + C, \tag{5.2}$$

with $K \in \mathbb{R}_{\text{sym}}^{d \times d}$, $l \in \mathbb{R}^d$ and $C \in \mathbb{R}$ (however it is worth recalling that the asymptotic variance σ_f^2 does not depend on C , see Section 1.2.2). We also stress that f is naturally assumed to be independent of the auxiliary variable p . Let us remark that an extension of our results to more general observables would be interesting, but not straightforward. Our first aim will be to study the dependence of the asymptotic variance σ_f^2 (see (1.52)) on the parameters μ and ν , i.e. the function

$$\Theta : \mathbb{R}^2 \rightarrow \mathbb{R}, \quad (\mu, \nu) \mapsto \sigma_f^2,$$

assuming a fixed observable f and perturbation matrix J . We will furthermore focus on small perturbations, i.e. on the behaviour of the function Θ in the neighbourhood

of the origin. Our main theoretical tool will be the Poisson equation (1.54), as can be seen from the proofs in Section 8.4. Anticipating the forthcoming analysis, let us already state our main result, showing that close to the origin, the function Θ has favourable properties along the diagonal $\mu = \nu$. In other words, perturbations of equal strength ($\mu = \nu$) always lead to an improvement in asymptotic variance under the nondegeneracy conditions $[J, K] \neq 0$ and $l \notin \ker J$, and when the damping is strong enough ($\gamma > \sqrt{2}$):

Theorem 5.1. *Consider the dynamics*

$$dq_t = p_t dt - \mu J q_t dt, \quad (5.3a)$$

$$dp_t = -q_t dt - \mu J p_t dt - \gamma p_t dt + \sqrt{2\gamma} dW_t, \quad (5.3b)$$

with $\gamma > \sqrt{2}$ and an observable of the form $f(q) = q \cdot K q + l \cdot q + C$, where $K \in \mathbb{R}_{\text{sym}}^{d \times d}$, $l \in \mathbb{R}^d$ and $C \in \mathbb{R}$. If at least one of the conditions $[J, K] \neq 0$ and $l \notin \ker J$ is satisfied, then the asymptotic variance of the unperturbed sampler is at a local maximum independently of K and J (and γ , as long as $\gamma > \sqrt{2}$), i.e.

$$\partial_\mu \sigma_f^2|_{\mu=0} = 0 \quad \text{and} \quad \partial_\mu^2 \sigma_f^2|_{\mu=0} < 0. \quad (5.4)$$

Remark 5.1. We wish to stress that in (5.4), σ_f^2 is a function of μ only, since $\mu = \nu$ has been assumed. Addressing the nondegeneracy conditions, we will see in Section 5.3, Example 5.3, that if $[J, K] = 0$ and $l \in \ker J$, then the asymptotic variance is constant as a function of μ , i.e. the perturbation has no effect.

Theorem 5.1 will follow from Corollary 5.1 and the discussion following Proposition 5.2 below. In the sequel, we will divide our study and investigate purely quadratic and purely linear observables separately.

PURELY QUADRATIC OBSERVABLES

We have the following result:

Proposition 5.1. *Let $l = 0$, i.e. $f(q) = q \cdot K q + C$. Then Θ satisfies*

$$\nabla \Theta|_{(\mu, \nu)=(0,0)} = 0 \quad (5.5)$$

and

$$\text{Hess } \Theta|_{(\mu, \nu)=(0,0)} = \begin{pmatrix} -(\gamma + \frac{1}{\gamma^3} + \gamma^3) (\text{Tr}(JKJK) - \text{Tr}(J^2K^2)) & (\frac{1}{\gamma^3} + \frac{1}{\gamma} - \gamma) \text{Tr}(J^2K^2) \\ -\frac{2}{\gamma} \text{Tr}(JKJK) & +(-\frac{1}{\gamma^3} + \frac{1}{\gamma} + \gamma) \text{Tr}(JKJK) \\ (\frac{1}{\gamma^3} + \frac{1}{\gamma} - \gamma) \text{Tr}(J^2K^2) & (\frac{1}{\gamma^3} - \frac{1}{\gamma}) \text{Tr}(J^2K^2) \\ +(-\frac{1}{\gamma^3} + \frac{1}{\gamma} + \gamma) \text{Tr}(JKJK) & -(\frac{1}{\gamma^3} + \frac{1}{\gamma}) \text{Tr}(JKJK) \end{pmatrix}. \quad (5.6)$$

Proof. See Section 8.4. □

The above proposition shows that the unperturbed dynamics represents a critical point of Θ , independently of the choice of K , J and γ . In general though, $\text{Hess } \Theta|_{(\mu, \nu)=(0,0)}$ can have both positive and negative eigenvalues. In particular, this implies that an unfortunate choice of the perturbations will actually increase the asymptotic variance (in contrast to the situation of perturbed *overdamped* Langevin dynamics (1.74), where any nonreversible perturbation leads to an improvement in asymptotic variance as detailed in Theorem 2.2). Furthermore, the nondiagonality of $\text{Hess } \Theta|_{(\mu, \nu)=(0,0)}$ hints at the fact that the interplay of the perturbations J_1 and J_2 (or rather their relative strengths μ and ν) is crucial for the performance of the sampler and, consequently, the effect of these perturbations cannot be satisfactorily studied independently.

Example 5.1. Assuming $J^2 = -I$ and $[J, K] = 0$, it follows that

$$\partial_\mu^2 \Theta|_{(\mu, \nu)=(0,0)} = \partial_\nu^2 \Theta|_{(\mu, \nu)=(0,0)} = \frac{1}{\gamma} \text{Tr}(K^2) > 0,$$

for all nonzero K . Therefore in this case, a small perturbation using either only J_1 or only J_2 will increase the asymptotic variance, uniformly over all choices of K and γ .

As already hinted at in the introduction to this chapter, it turns out that it is possible to construct an improved sampler by combining both perturbations in a suitable way. Indeed, the function Θ can be seen to have good properties along $\mu = \nu$. We set $\mu(s) = s$, $\nu(s) := s$ and compute

$$\begin{aligned}
\left. \frac{d^2}{ds^2} \Theta \right|_{s=0} &= (1, 1) \cdot \text{Hess } \Theta|_{(\mu, \nu)=(0,0)}(1, 1) \\
&= - \left(\gamma + \frac{1}{\gamma^3} + \gamma^3 \right) (\text{Tr}(JKJK) - \text{Tr}(J^2K^2)) - \frac{2}{\gamma} \text{Tr}(JKJK) \\
&\quad + 2 \cdot \left(\left(\frac{1}{\gamma^3} + \frac{1}{\gamma} - \gamma \right) \text{Tr}(J^2K^2) + \left(-\frac{1}{\gamma^3} + \frac{1}{\gamma} + \gamma \right) \text{Tr}(JKJK) \right) \\
&\quad + \left(\frac{1}{\gamma^3} - \frac{1}{\gamma} \right) \text{Tr}(J^2K^2) - \left(\frac{1}{\gamma^3} + \frac{1}{\gamma} \right) \text{Tr}(JKJK) \\
&= \left(\gamma - \frac{4}{\gamma^3} - \gamma^3 - \frac{1}{\gamma} \right) \cdot (\text{Tr}(JKJK) - \text{Tr}(J^2K^2)) \leq 0.
\end{aligned}$$

The last inequality follows from

$$\gamma - \frac{4}{\gamma^3} - \gamma^3 - \frac{1}{\gamma} < 0 \quad \text{and} \quad \text{Tr}(JKJK) - \text{Tr}(J^2K^2) \geq 0.$$

Both inequalities are proven in Section 8.4, Lemma 8.3, where the last inequality is strict if $[J, K] \neq 0$. Consequently, choosing both perturbations to be of the same magnitude ($\mu = \nu$) and assuring that J and K do not commute always leads to a smaller asymptotic variance, independently of the choice of K , J and γ . We state this result in the following corollary:

Corollary 5.1. *Consider the dynamics (5.3) and a quadratic observable*

$$f(q) = q \cdot Kq + C.$$

If $[J, K] \neq 0$, then the asymptotic variance of the unperturbed sampler is at a local maximum independently of K , J and γ , i.e.

$$\partial_\mu \sigma_f^2|_{\mu=0} = 0, \quad \text{and} \quad \partial_\mu^2 \sigma_f^2|_{\mu=0} < 0.$$

Example 5.2. Let us set $\mu(s) := s$ and $\nu(s) := -s$ (this corresponds to a small perturbation with $J\nabla V(q_t) dt$ in q and $-Jp_t dt$ in p). In this case we get

$$\left. \frac{d^2 \Theta}{ds^2} \right|_{s=0} = - \underbrace{\frac{1}{2} \cdot \frac{\gamma^4 + 3\gamma^2 + 5}{\gamma} (\text{Tr}(JKJK) - \text{Tr}(J^2K^2))}_{\leq 0} \underbrace{- 4 \frac{\text{Tr}(J^2K^2)}{\gamma}}_{\geq 0},$$

which changes its sign depending on J and K as the first term is negative and the second one is positive. Whether the perturbation improves the performance of the sampler in terms of asymptotic variance therefore depends on the specifics of the observable and the perturbation in this case.

We have the following result:

Proposition 5.2. *Let $K = 0$, i.e. $f(q) = l \cdot q + C$. Then Θ satisfies*

$$\nabla \Theta|_{(\mu, \nu)=(0,0)} = 0$$

and

$$\text{Hess } \Theta|_{(\mu, \nu)=(0,0)} = \begin{pmatrix} -2\gamma^3 |Jl|^2 & 2\gamma |Jl|^2 \\ 2\gamma |Jl|^2 & 0 \end{pmatrix}.$$

Proof. See Section 8.4. □

Let us assume that $l \notin \ker J$. Then $\partial_\mu^2 \Theta|_{\mu, \nu=0} < 0$, and hence Proposition 5.2 shows that a small perturbation by $\mu J \nabla V(q_t) dt$ alone always results in an improvement of the asymptotic variance. However, if we combine both perturbations $\mu J \nabla V(q_t) dt$ and $\nu J p_t dt$, then the effect depends on the sign of

$$\begin{pmatrix} \mu & \nu \end{pmatrix} \begin{pmatrix} -2\gamma^3 |Jl|^2 & 2\gamma |Jl|^2 \\ 2\gamma |Jl|^2 & 0 \end{pmatrix} \begin{pmatrix} \mu \\ \nu \end{pmatrix} = -(2\mu^2 \gamma^3 - 4\mu\nu\gamma) |Jl|^2.$$

This quantity will be negative if μ and ν have different signs, and also if they have the same sign and γ is large enough. Following the discussion about purely quadratic observables (see Proposition 5.1 and Corollary 5.1), we impose the condition $\mu = \nu$. We then end up with the requirement

$$2\mu^2 \gamma^3 - 4\mu\nu\gamma > 0,$$

which is satisfied if and only if $\gamma > \sqrt{2}$.

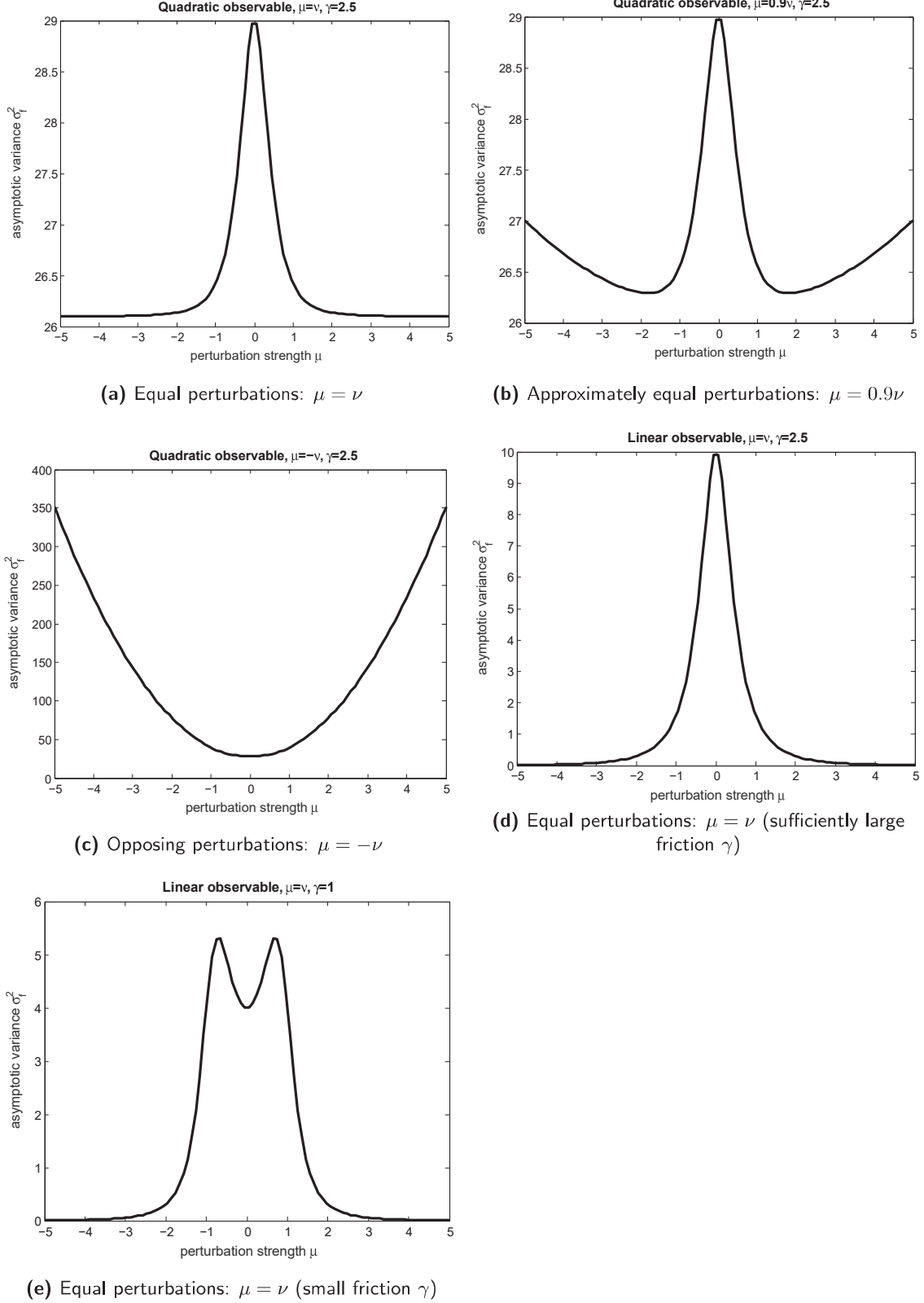
Summarising the results of this section, for observables of the form $f(q) = q \cdot K q + l \cdot q + C$, choosing equal perturbations ($\mu = \nu$) with a sufficiently strong damping ($\gamma > \sqrt{2}$) always leads to an improvement in asymptotic variance under the conditions $[J, K] \neq 0$ and $l \notin \ker J$. This is finally the content of Theorem 5.1.

Numerical examples show that the conditions $\gamma > \sqrt{2}$ and $\mu = \nu$ are indeed necessary for the conclusions of Theorem 5.1 to hold (an explanation in terms of the spectrum of the generator will be provided in Section 5.2 below). In particular, an unfortunate choice of the perturbations will actually increase the asymptotic variance.

Let us illustrate this by plotting the asymptotic variance as a function of the perturbation strength μ (see Figure 5.1), making the choices $d = 2$, $l = (1, 1)^T$,

$$K = \begin{pmatrix} 2 & 0 \\ 0 & 1 \end{pmatrix} \quad \text{and} \quad J = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}. \tag{5.7}$$

Figure 5.1: Asymptotic variance for linear and quadratic observables, depending on relative perturbation and friction strengths



The asymptotic variance has been computed according to (8.18a) and (8.19). Going beyond the results of this section, the graphs give an impression of the behaviour of the asymptotic variance for large values of μ , discussed further in Section 5.3.

Figures 5.1a, 5.1b and 5.1c show the asymptotic variance associated with the quadratic observable $f(q) = q \cdot Kq$. In accordance with Theorem 5.1, the asymptotic variance is at a local maximum at zero perturbation in the case $\mu = \nu$ (see Figure 5.1a). For increasing perturbation strength, the graph shows monotone decay for $\mu \rightarrow \infty$ (this limiting behaviour will be explored analytically in Section 5.3). If the condition $\mu = \nu$ is only approximately satisfied (Figure 5.1b), our numerical example still exhibits decaying asymptotic variance in the neighbourhood of the critical point. In this case, however, the asymptotic variance diverges for growing values of the perturbation μ . If the perturbations are opposed ($\mu = -\nu$), it is possible for certain observables that the unperturbed dynamics represents a global minimum. Such a case is observed in Figure 5.1c. In Figures 5.1d and 5.1e the observable $f(q) = l \cdot q$ is considered. If the damping is sufficiently strong ($\gamma > \sqrt{2}$), the unperturbed dynamics is at a local maximum of the asymptotic variance (Figure 5.1d). Furthermore, the asymptotic variance approaches zero as $\mu \rightarrow \infty$ (for a theoretical explanation see again Section 5.3). The graph in Figure 5.1e shows that the assumption of γ not being too small cannot be dropped from Theorem 5.1. Even in this case though the example shows decay of the asymptotic variance for large values of μ .

5.2 EXPONENTIAL DECAY RATE

Recall the definitions of the optimal decay rate $\lambda^*(\mathcal{L})$ and the spectral gap $\lambda_{\text{gap}}(\mathcal{L})$ from (1.48) and (1.49). In [101, Proposition 2.1] it was proven that the Ornstein-Uhlenbeck semigroup $(S_t)_{t \geq 0}$ considered in this section is differentiable. In this case (see [47, Corollary 3.12]), it is known that the optimal decay rate and the spectral gap coincide, i.e. $\lambda^*(\mathcal{L}) = \lambda_{\text{gap}}(\mathcal{L})$. In this section we will therefore analyse the spectral properties of the generator $\mathcal{L}$ in the Gaussian case. In particular, this leads to some intuition of why choosing equal perturbations ($\mu = \nu$) is crucial for the performance of the sampler.

First, notice that the dynamics (5.1) is of Ornstein-Uhlenbeck type, i.e. we can write

$$dX_t = -BX_t dt + \sqrt{2Q} d\bar{W}_t, \quad (5.8)$$

with $X = (q, p)^T$,

$$B = \begin{pmatrix} \mu J & -I \\ I & \gamma I + \nu J \end{pmatrix}, \quad Q = \begin{pmatrix} \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \gamma I \end{pmatrix}, \quad (5.9)$$

and $(\bar{W}_t)_{t \geq 0}$ denoting a standard Wiener process on $\mathbb{R}^{2d}$. The generator of (5.8) is then given by

$$\mathcal{L} = -Bx \cdot \nabla + Q : \nabla \nabla. \quad (5.10)$$

In [101] (see also [110]), it was proven that the spectrum of $\mathcal{L}$ in $L^2(\widehat{\pi})$ is given by

$$\sigma(\mathcal{L}) = \left\{ -\sum_{j=1}^r n_j \lambda_j : n_j \in \mathbb{N}, \lambda_j \in \sigma(B) \right\}. \quad (5.11)$$

Note that $\sigma(\mathcal{L})$ only depends on the drift matrix B . In the case where $\mu = \nu$, the spectrum of B can be computed explicitly.

Lemma 5.1. *Assume that $\mu = \nu$. Then the spectrum of B is given by*

$$\sigma(B) = \left\{ \mu\lambda + \sqrt{\left(\frac{\gamma}{2}\right)^2 - 1} + \frac{\gamma}{2}|\lambda \in \sigma(J) \right\} \cup \left\{ \mu\lambda - \sqrt{\left(\frac{\gamma}{2}\right)^2 - 1} + \frac{\gamma}{2}|\lambda \in \sigma(J) \right\}^*. \quad (5.12)$$

Proof. We will compute $\sigma(B - \frac{\gamma}{2}I)$ and then use the identity

$$\sigma(B) = \left\{ \lambda + \frac{\gamma}{2}|\lambda \in \sigma\left(B - \frac{\gamma}{2}I\right) \right\}. \quad (5.13)$$

We have

$$\begin{aligned} \det\left(B - \frac{\gamma}{2}I - \lambda I\right) &= \det\left(\left(\mu J - \frac{\gamma}{2}I - \lambda I\right)\left(\mu J + \frac{\gamma}{2}I - \lambda I\right) + I\right) \\ &= \det\left(\mu J - \lambda I + \sqrt{\left(\frac{\gamma}{2}\right)^2 - 1} \cdot I\right) \cdot \det\left(\mu J - \lambda I - \sqrt{\left(\frac{\gamma}{2}\right)^2 - 1} \cdot I\right), \end{aligned}$$

where I is understood to denote the identity matrix of appropriate dimension. The above quantity is zero if and only if

$$\lambda - \sqrt{\left(\frac{\gamma}{2}\right)^2 - 1} \in \sigma(\mu J) \quad \text{or} \quad \lambda + \sqrt{\left(\frac{\gamma}{2}\right)^2 - 1} \in \sigma(\mu J).$$

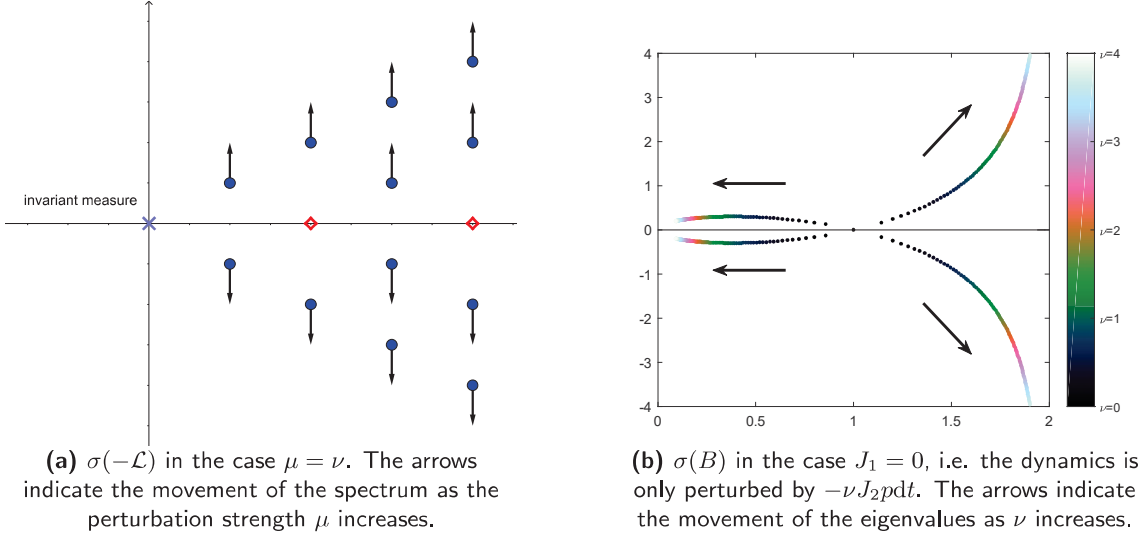
Together with (5.13), the claim follows. $\square$

Using formula (5.12), in Figure 5.2a we show a sketch of the spectrum $\sigma(-\mathcal{L})$ for the case of equal perturbations ($\mu = \nu$) with the convenient choices $n = 1$ and $\gamma = 2$. Of course, the eigenvalue at 0 is associated to the invariant measure since $\mathcal{L}^\dagger \widehat{\pi} = 0$. The arrows indicate the movement of the eigenvalues as the perturbation μ increases in accordance with Lemma 5.1. Clearly, the spectral gap of $\mathcal{L}$ is not affected by the perturbation. Note that the eigenvalues on the real axis stay invariant under the perturbation. The subspace of $L_0^2(\widehat{\pi})$ associated to those will turn out to be crucial for the characterisation of the limiting asymptotic variance as $\mu \rightarrow \infty$ (see Remark 5.6).

To illustrate the suboptimal properties of the perturbed dynamics when the perturbations are not equal, we plot the spectrum of the drift matrix $\sigma(B)$ in the case when the dynamics

*Notice that $\sqrt{\left(\frac{\gamma}{2}\right)^2 - 1}$ is understood to be a complex number for $\gamma < 2$.

Figure 5.2: Effects of the perturbation on the spectra of $-\mathcal{L}$ and B .



is only perturbed by the term $\nu J_2 p dt$ (i.e. $\mu = 0$) for $n = 2$, $\gamma = 2$ and

$$J_2 = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}, \quad (5.14)$$

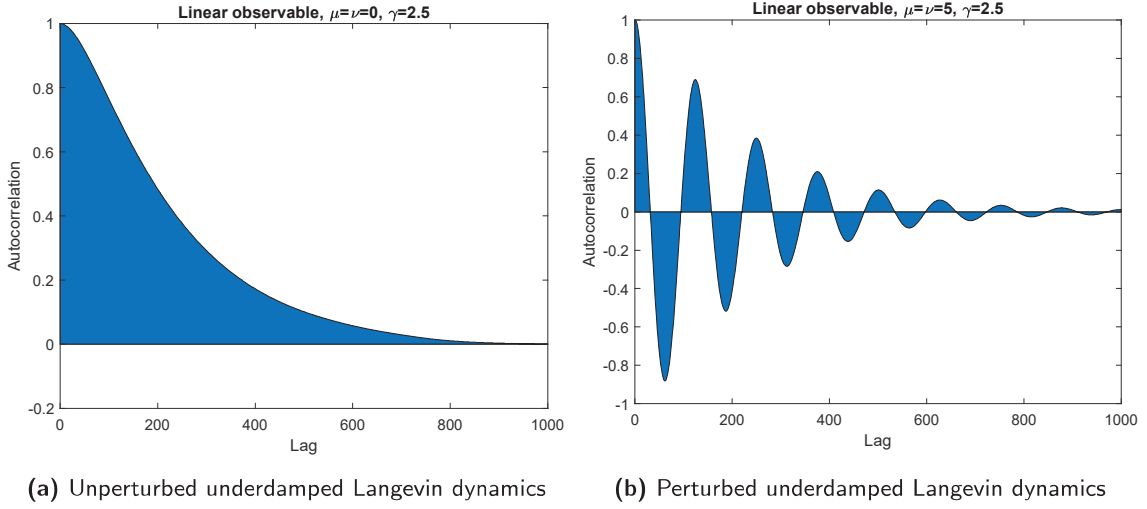
(see Figure 5.2b). Note that the full spectrum $\sigma(-\mathcal{L})$ can be inferred from (5.11). For $\nu = 0$ we have that the spectrum $\sigma(B)$ consists only of the (degenerate) eigenvalue 1. For increasing ν , the figure shows that the degenerate eigenvalue splits up into four eigenvalues, two of which get closer to the imaginary axis as ν increases, leading to a smaller spectral gap and therefore to a decrease in the speed of convergence to equilibrium. Figures (5.2a) and (5.2b) give an intuitive explanation of why the fine-tuning of the perturbation strengths is crucial.

We close this subsection by providing autocorrelation plots (see Figure 5.3) for the linear observable considered in Figure 5.1d (with a friction coefficient of $\gamma = 2.5$). It is well-known that the asymptotic variance is given by the integrated autocorrelation function (see Section 1.2.2),

$$\sigma_f^2 = \int_0^\infty \mathbb{E}_\pi [(f(q_0) - \pi(f)) (f(q_t) - \pi(f))] dt. \quad (5.15)$$

Comparing Figures 5.3a and 5.3b yields additional insight into the mechanics of the variance reduction: the increase of the imaginary part of the eigenvalues of $\mathcal{L}$ (as indicated in Figure 5.2a) leads to oscillations in the autocorrelation function and therefore to cancellations in (5.15). A similar effect has already been observed in [125] for the nonreversible overdamped Langevin dynamics (1.74).

Figure 5.3: Autocorrelation plots for the perturbed and unperturbed dynamics



5.3 UNIT COVARIANCE – LARGE PERTURBATIONS

In the previous section we observed that for the particular choice of perturbation $J_1 = J_2$ and $\mu = \nu$ the perturbed Langevin dynamics demonstrated an improvement in performance for μ in a neighbourhood of 0, when the observable is linear or quadratic (see Theorem 5.1). Recall that this dynamics is ergodic with respect to a standard Gaussian measure $\hat{\pi}$ on $\mathbb{R}^{2d}$ with marginal π with respect to the q -variable. As before, we shall consider only observables that do not depend on p . Moreover, we assume without loss of generality that $\pi(f) = 0$. For such observables we will write $f \in L_0^2(\pi)$ and consider the canonical embedding $L_0^2(\pi) \subset L^2(\hat{\pi})$. We emphasize that $L_0^2(\pi)$ consists of functions that only depend on q , whereas functions in $L^2(\hat{\pi})$ may depend on both q and p .

In this section we will analyse the asymptotic variance for large values of μ . The generator of (5.3) can be written as

$$\mathcal{L} = \underbrace{p \cdot \nabla_q - q \cdot \nabla_p + \gamma(-p \cdot \nabla_p + \Delta_p)}_{\mathcal{L}_0} + \underbrace{\mu(-Jq \cdot \nabla_q - Jp \cdot \nabla_p)}_{\mathcal{A}} =: \mathcal{L}_0 + \mu\mathcal{A}, \quad (5.16)$$

where we have introduced the notation $\mathcal{L}_{\text{pert}} = \mu\mathcal{A}$. In the sequel, the adjoint of an operator $\mathcal{B}$ in $L^2(\hat{\pi})$ will be denoted by $\mathcal{B}^*$. In the rest of this section we will make repeated use of the Hermite polynomials

$$h_\alpha(x) = (-1)^{|\alpha|} e^{\frac{|x|^2}{2}} \nabla^\alpha e^{-\frac{|x|^2}{2}}, \quad \alpha \in \mathbb{N}^{2d}, \quad (5.17)$$

invoking the notation $x = (q, p) \in \mathbb{R}^{2d}$. For $m \in \mathbb{N}_0$ we define the Hilbert spaces

$$H_m = \text{span}\{h_\alpha : |\alpha| = m\}, \quad \langle f, g \rangle_m := \langle f, g \rangle_{L^2(\hat{\pi})}, \quad f, g \in H_m.$$

The following result (Theorem 5.2) holds for operators of the form (5.10), providing an orthogonal decomposition of $L^2(\widehat{\pi})$ into invariant subspaces. The drift and diffusion matrices B and Q are assumed to be such that $\mathcal{L}$ is the generator of an ergodic stochastic process with smooth transition functions (see [6, Definition 2.1] for precise conditions).

Theorem 5.2. *[6, Section 5]. The following hold:*

(a) *The space $L^2(\widehat{\pi})$ has a decomposition into mutually orthogonal subspaces:*

$$L^2(\widehat{\pi}) = \bigoplus_{m \in \mathbb{N}_0} H_m.$$

(b) *For all $m \in \mathbb{N}_0$, H_m is invariant under $\mathcal{L}$ as well as under the corresponding semigroup $(S_t)_{t \geq 0}$.*

(c) *The spectrum of $\mathcal{L}$ has the following decomposition:*

$$\sigma(\mathcal{L}) = \bigcup_{m \in \mathbb{N}_0} \sigma(\mathcal{L}|_{H_m}), \quad \sigma(\mathcal{L}|_{H_m}) = \left\{ -\sum_{j=1}^{2d} \alpha_j \lambda_j : |\alpha| = m, \lambda_j \in \sigma(B) \right\}.$$

Remark 5.2. *Note that by the ergodicity of the dynamics, $\ker \mathcal{L}$ consists of constant functions (see Proposition 1.1) and so $\ker \mathcal{L} = H_0$. Therefore, $L_0^2(\widehat{\pi})$ has the decomposition*

$$L_0^2(\widehat{\pi}) = L^2(\widehat{\pi}) / \ker \mathcal{L} = \bigoplus_{m \geq 1} H_m.$$

Our first main result of this section is an expression for the asymptotic variance in terms of the unperturbed operator $\mathcal{L}_0$ and the perturbation $\mathcal{A}$:

Proposition 5.3. *Let $f \in L_0^2(\pi)$ (so in particular $f = f(q)$). Then the associated asymptotic variance is given by*

$$\sigma_f^2 = \langle f, -\mathcal{L}_0(\mathcal{L}_0^2 + \mu^2 \mathcal{A}^* \mathcal{A})^{-1} f \rangle_{L^2(\widehat{\pi})}. \quad (5.18)$$

Remark 5.3. *The proof of the preceding Proposition will show that $\mathcal{L}_0^2 + \mu^2 \mathcal{A}^* \mathcal{A}$ is invertible on $L_0^2(\widehat{\pi})$ and that $(\mathcal{L}_0^2 + \mu^2 \mathcal{A}^* \mathcal{A})^{-1} f \in \mathcal{D}(\mathcal{L}_0)$ for all $f \in L_0^2(\widehat{\pi})$.*

To prove Proposition 5.3 we will make use of the *generator with reversed perturbation*

$$\mathcal{L}_- = \mathcal{L}_0 - \mu \mathcal{A}$$

and the momentum flip operator

$$P : L_0^2(\widehat{\pi}) \rightarrow L_0^2(\widehat{\pi}), \quad f(q, p) \mapsto f(q, -p), \quad (5.19)$$

already presented in Example 2.1. Here, we drop the bar for notational convenience. Clearly, $P^2 = I$ and $P^* = P$. Further properties of $\mathcal{L}_0$, $\mathcal{A}$ and the auxiliary operators $\mathcal{L}_-$ and P are gathered in the following lemma:

Lemma 5.2. *For all $f, g \in C^\infty(\mathbb{R}^{2d}) \cap L^2(\widehat{\pi})$ the following hold:*

(a) *The generator $\mathcal{L}_0$ is symmetric in $L^2(\widehat{\pi})$ with respect to P :*

$$\langle f, P\mathcal{L}_0 g \rangle_{L^2(\widehat{\pi})} = \langle \mathcal{L}_0 f, P g \rangle_{L^2(\widehat{\pi})}.$$

(b) *The perturbation $\mathcal{A}$ is skew-adjoint in $L^2(\widehat{\pi})$:*

$$\mathcal{A}^* = -\mathcal{A}.$$

(c) *The perturbation $\mathcal{A}$ is skew-symmetric in $L^2(\widehat{\pi})$ with respect to P :*

$$\langle f, P\mathcal{A}g \rangle_{L^2(\widehat{\pi})} = -\langle \mathcal{A}f, P g \rangle_{L^2(\widehat{\pi})}.$$

(d) *The operators $\mathcal{L}_0$ and $\mathcal{A}$ commute:*

$$[\mathcal{L}_0, \mathcal{A}]f = 0.$$

(e) *The perturbation $\mathcal{A}$ satisfies*

$$P\mathcal{A}P f = \mathcal{A}f.$$

(f) *$\mathcal{L}$ and $\mathcal{L}_-$ commute,*

$$[\mathcal{L}, \mathcal{L}_-]f = 0,$$

and the following relation holds:

$$\langle f, P\mathcal{L}P g \rangle_{L^2(\widehat{\pi})} = \langle \mathcal{L}_- f, g \rangle_{L^2(\widehat{\pi})}. \quad (5.20)$$

(g) *The operators $\mathcal{L}$, $\mathcal{L}_0$, $\mathcal{L}_-$, $\mathcal{A}$ and P leave the Hermite spaces H_m invariant.*

Remark 5.4. *The claim (d) in the above lemma is crucial for our approach, which itself rests heavily on the fact that the q - and p -perturbations match ($J_1 = J_2$).*

Proof of Lemma 5.2. To prove (a), consider the following decomposition of $\mathcal{L}_0$ as in (4.5):

$$\mathcal{L}_0 = \underbrace{p \cdot \nabla_q - q \cdot \nabla_p}_{\mathcal{L}_{\text{ham}}} + \underbrace{\gamma(-p \cdot \nabla_p + \Delta_p)}_{\mathcal{L}_{\text{therm}}}.$$

By partial integration it is straightforward to see that

$$\langle f, \mathcal{L}_{\text{ham}} g \rangle_{L^2(\widehat{\pi})} = -\langle \mathcal{L}_{\text{ham}} f, g \rangle_{L^2(\widehat{\pi})}$$

and

$$\langle f, \mathcal{L}_{\text{therm}} g \rangle_{L^2(\widehat{\pi})} = \langle \mathcal{L}_{\text{therm}} f, g \rangle_{L^2(\widehat{\pi})},$$

for all $f, g \in C^\infty(\mathbb{R}^{2d}) \cap L^2(\widehat{\pi})$, i.e. $\mathcal{L}_{\text{ham}}$ and $\mathcal{L}_{\text{therm}}$ are antisymmetric and symmetric in $L^2(\widehat{\pi})$ respectively. Furthermore, we immediately see that $\mathcal{L}_{\text{ham}} P f = -P \mathcal{L}_{\text{ham}} f$ and $\mathcal{L}_{\text{therm}} P f = P \mathcal{L}_{\text{therm}} f$, so that

$$\langle f, \mathcal{L}_0 P g \rangle_{L^2(\widehat{\pi})} = \langle f, -\mathcal{L}_{\text{ham}} P g + \mathcal{L}_{\text{therm}} P g \rangle_{L^2(\widehat{\pi})} = \langle \mathcal{L}_0 f, P g \rangle_{L^2(\widehat{\pi})}.$$

We note that this result is well known (see [90, Section 2.2.3.1] for instance) and holds in the more general setting of Chapter 4 for the infinitesimal generator (4.5). The claim (b) follows by noting that the flow vector field $b(q, p) = (-Jq, -Jp)$ associated to $\mathcal{A}$ is divergence-free with respect to $\widehat{\pi}$, i.e. $\nabla \cdot (\widehat{\pi} b) = 0$. Therefore, $\mathcal{A}$ is the generator of a strongly continuous unitary semigroup on $L^2(\widehat{\pi})$ and hence skew-adjoint by Stone's Theorem [47, Theorem 3.24]. The claim (c) follows from combining (b) and (e) (proved below). To prove (d) we use the decomposition $\mathcal{L}_0 = \mathcal{L}_{\text{ham}} + \mathcal{L}_{\text{therm}}$ to obtain

$$[\mathcal{L}_0, \mathcal{A}]f = [\mathcal{L}_{\text{ham}}, \mathcal{A}]f + [\mathcal{L}_{\text{therm}}, \mathcal{A}]f, \quad f \in C^\infty(\mathbb{R}^{2d}) \cap L^2(\widehat{\pi}). \quad (5.21)$$

The first term of (5.21) gives

$$\begin{aligned} & [p \cdot \nabla_q - q \cdot \nabla_p, -Jq \cdot \nabla_q - Jp \cdot \nabla_p]f \\ &= ([p \cdot \nabla_q, -Jq \cdot \nabla_q] + [p \cdot \nabla_q, -Jp \cdot \nabla_p] + [-q \cdot \nabla_p, -Jq \cdot \nabla_q] \\ &\quad + [-q \cdot \nabla_p, -Jp \cdot \nabla_p])f \\ &= Jp \cdot \nabla_q f - Jp \cdot \nabla_q f + Jq \cdot \nabla_p f - Jq \cdot \nabla_p f = 0. \end{aligned}$$

The second term of (5.21) gives

$$[-p \cdot \nabla_p + \Delta_p, \mathcal{A}]f = [-p \cdot \nabla_p, -Jp \cdot \nabla_p]f + [\Delta_p, -Jp \cdot \nabla_p]f, \quad (5.22)$$

since $Jq \cdot \nabla_q$ commutes with $p \cdot \nabla_p + \Delta_p$. Both terms in (5.22) are clearly zero due the skew-symmetry of J and the symmetry of the Hessian $\nabla_p \nabla_p f$.

The claim (e) follows from a short calculation similar to the proof of (a). To prove (f), note that the fact that $\mathcal{L}$ and $\mathcal{L}_-$ commute follows from (d), as

$$[\mathcal{L}, \mathcal{L}_-]f = [\mathcal{L}_0 + \mu \mathcal{A}, \mathcal{L}_0 - \mu \mathcal{A}]f = -2\mu [\mathcal{L}_0, \mathcal{A}]f = 0, \quad f \in C^\infty(\mathbb{R}^{2d}) \cap L^2(\widehat{\pi}),$$

while the property $\langle f, P\mathcal{L}_0Pg \rangle_{L^2(\widehat{\pi})} = \langle \mathcal{L}_-f, g \rangle_{L^2(\widehat{\pi})}$ follows from properties (a), (b) and (e). Indeed,

$$\begin{aligned} \langle f, P\mathcal{L}Pg \rangle_{L^2(\widehat{\pi})} &= \langle f, P(\mathcal{L}_0 + \mu\mathcal{A})Pg \rangle_{L^2(\widehat{\pi})} = \langle f, (P\mathcal{L}_0P + \mu\mathcal{A})g \rangle_{L^2(\widehat{\pi})} \\ &= \langle (\mathcal{L}_0 - \mu\mathcal{A})f, g \rangle_{L^2(\widehat{\pi})} = \langle \mathcal{L}_-f, g \rangle_{L^2(\widehat{\pi})}, \end{aligned}$$

as required. To prove (g) first notice that $\mathcal{L}$, $\mathcal{L}_0$ and $\mathcal{L}_-$ are of the form (5.10) and therefore leave the spaces H_m invariant by Theorem 5.2. It follows immediately that also $\mathcal{A}$ leaves those spaces invariant. The fact that P leaves the spaces H_m invariant follows directly by inspection of (5.17). $\square$

Now we proceed with the proof of Proposition 5.3:

Proof of Proposition 5.3. Since the potential V is quadratic, Assumption 4.1 clearly holds and thus Theorem 4.1 together with Proposition 1.1 ensures that $\mathcal{L}$ and $\mathcal{L}_-$ are invertible on $L_0^2(\widehat{\pi})$ with

$$(-\mathcal{L})^{-1} = \int_0^\infty S_t dt, \quad (5.24)$$

and analogously for $\mathcal{L}_-^{-1}$. In particular, the asymptotic variance can be written as

$$\sigma_f^2 = \langle f, (-\mathcal{L})^{-1}f \rangle_{L^2(\widehat{\pi})}.$$

Due to the representation (5.24) and Theorem 5.2, the inverses of $\mathcal{L}$ and $\mathcal{L}_-$ leave the Hermite spaces H_m invariant. We will prove the claim from Proposition 5.3 under the assumption that $Pf = f$ which includes the case $f = f(q)$. For the following calculations we will assume $f \in H_m$ for fixed $m \geq 1$. Combining statement (g) with (a) and (f) of Lemma 5.2 (and noting that $H_m \subset C^\infty(\mathbb{R}^{2d}) \cap L^2(\widehat{\pi})$) we see that

$$P\mathcal{L}P = \mathcal{L}_-^* \quad \text{and} \quad P\mathcal{L}_0P = \mathcal{L}_0^* \quad (5.25)$$

when restricted to H_m . Therefore, the following calculations are justified:

$$\begin{aligned} \langle f, (-\mathcal{L})^{-1}f \rangle_{L^2(\widehat{\pi})} &= \frac{1}{2} (\langle f, (-\mathcal{L})^{-1}f \rangle_{L^2(\widehat{\pi})} + \langle f, (-\mathcal{L}^*)^{-1}f \rangle_{L^2(\widehat{\pi})}) \\ &= \frac{1}{2} (\langle f, (-\mathcal{L})^{-1}f \rangle_{L^2(\widehat{\pi})} + \langle Pf, (-\mathcal{L}^*)^{-1}Pf \rangle_{L^2(\widehat{\pi})}) \\ &= \frac{1}{2} (\langle f, (-\mathcal{L})^{-1}f \rangle_{L^2(\widehat{\pi})} + \langle f, (-\mathcal{L}_-)^{-1}f \rangle_{L^2(\widehat{\pi})}) \\ &= \frac{1}{2} \langle f, ((-\mathcal{L})^{-1} + (-\mathcal{L}_-)^{-1})f \rangle_{L^2(\widehat{\pi})}, \end{aligned}$$

where in the second line we have used the assumption $Pf = f$ and in the third line the properties $P^2 = I$, $P^* = P$ and equation (5.25). Since $\mathcal{L}$ and $\mathcal{L}_-$ commute on H_m

according to the statements (f) and (g) of Lemma 5.2 we can write

$$(-\mathcal{L})^{-1} + (-\mathcal{L}_-)^{-1} = \mathcal{L}_-(-\mathcal{L}\mathcal{L}_-)^{-1} + \mathcal{L}(-\mathcal{L}\mathcal{L}_-)^{-1} = -2\mathcal{L}_0(\mathcal{L}\mathcal{L}_-)^{-1}$$

for the restrictions on H_m , using $\mathcal{L} + \mathcal{L}_- = 2\mathcal{L}_0$. We also have

$$\mathcal{L}\mathcal{L}_- = (\mathcal{L}_0 + \mu\mathcal{A})(\mathcal{L}_0 - \mu\mathcal{A}) = \mathcal{L}_0^2 + \mu^2\mathcal{A}^*\mathcal{A},$$

since $\mathcal{L}_0$ and $\mathcal{A}$ commute. We thus arrive at the formula

$$\sigma_f^2 = \langle f, -\mathcal{L}_0(\mathcal{L}_0^2 + \mu^2\mathcal{A}^*\mathcal{A})^{-1}f \rangle_{L^2(\hat{\pi})}, \quad f \in H_m. \quad (5.26)$$

Now since $(\mathcal{L}_0^2 + \mu^2\mathcal{A}^*\mathcal{A})^{-1}f = (\mathcal{L}\mathcal{L}_-)^{-1}f \in \mathcal{D}(\mathcal{L}_0)$ for all $f \in L^2(\hat{\pi})$, it follows that the operator $-\mathcal{L}_0(\mathcal{L}_0^2 + \mu^2\mathcal{A}^*\mathcal{A})^{-1}$ is bounded. We can therefore extend formula (5.26) to the whole of $L^2(\hat{\pi})$ by continuity, using the fact that $L_0^2(\hat{\pi}) = \bigoplus_{m \geq 1} H_m$. $\square$

Applying Proposition 5.3, we can analyse the behaviour of σ_f^2 in the limit of large perturbation strength $\mu \rightarrow \infty$. To this end, we introduce the orthogonal decomposition

$$L_0^2(\pi) = \ker(Jq \cdot \nabla_q) \oplus \ker(Jq \cdot \nabla_q)^\perp, \quad (5.27)$$

where $Jq \cdot \nabla_q$ is understood as an unbounded operator acting on $L_0^2(\pi)$, obtained as the smallest closed extension of $Jq \cdot \nabla_q$ acting on $C_c^\infty(\mathbb{R}^d)$. In particular, $\ker(Jq \cdot \nabla_q)$ is a closed linear subspace of $L_0^2(\pi)$. Let Π denote the $L_0^2(\pi)$ -orthogonal projection onto $\ker(Jq \cdot \nabla_q)$. We will write $\sigma_f^2(\mu)$ to stress the dependence of the asymptotic variance on the perturbation strength. The following result shows that for large perturbations, the limiting asymptotic variance is always smaller than the asymptotic variance in the unperturbed case. Furthermore, the limit is given as the asymptotic variance of the projected observable Πf for the unperturbed dynamics.

Theorem 5.3. *Let $f \in L_0^2(\pi)$ (so in particular $f = f(q)$). Then*

$$\lim_{\mu \rightarrow \infty} \sigma_f^2(\mu) = \sigma_{\Pi f}^2(0) \leq \sigma_f^2(0).$$

Remark 5.5. *Note that the fact that the limit exists and is finite is nontrivial. In particular, as Figures 5.1b and 5.1c demonstrate, it is often the case that $\lim_{\mu \rightarrow \infty} \sigma_f^2(\mu) = \infty$ if the condition $\mu = \nu$ is not satisfied.*

Remark 5.6. *The decomposition (5.27) can be interpreted in terms of the spectrum $\sigma(\mathcal{L})$ as follows: First observe that for functions f that only depend on q , $f \in \ker(Jq \cdot \nabla_q)$ is equivalent to $f \in \ker \mathcal{A}$. Let us denote by $\bar{\sigma} = \bigcap_{\mu \in \mathbb{R}} \sigma(\mathcal{L}_0 + \mu\mathcal{A})$ the part of $\sigma(\mathcal{L}_0)$ that is not affected by the perturbation and by*

$$\bar{E} = \overline{\text{span}\{f \in L^2(\hat{\pi}) : \text{there exists } \lambda \in \bar{\sigma} \text{ such that } \mathcal{L}_0 f = \lambda f\}} \subset L^2(\hat{\pi})$$

the corresponding subspace. In the case when the eigenvectors of $\mathcal{L}$ span $L^2(\widehat{\pi})^\dagger$ it is straightforward to see that $\ker(\mathcal{A}) = \bar{E}^\dagger$. In Figure 5.2a, $\bar{\sigma}$ has been highlighted by diamonds.

Proof of Theorem 5.3. Note that $\mathcal{L}_0$ and $\mathcal{A}^*\mathcal{A}$ leave the Hermite spaces H_m invariant and their restrictions to those spaces commute (see Lemma 5.2, (b), (d) and (g)). Furthermore, as the Hermite spaces H_m are finite-dimensional, those operators have discrete spectrum. As $\mathcal{A}^*\mathcal{A}$ is nonnegative self-adjoint, there exists an orthogonal decomposition $L_0^2(\pi) = \bigoplus_i W_i$ into eigenspaces of the operator $-\mathcal{L}_0(\mathcal{L}_0^2 + \mu^2 \mathcal{A}^*\mathcal{A})^{-1}$, the decomposition $\bigoplus W_i$ being finer than $\bigoplus H_m$ in the sense that every W_i is a subspace of some H_m . Moreover,

$$-\mathcal{L}_0(\mathcal{L}_0^2 + \mu^2 \mathcal{A}^*\mathcal{A})^{-1}|_{W_i} = -\mathcal{L}_0(\mathcal{L}_0^2 + \mu^2 \lambda_i)^{-1}|_{W_i},$$

where $\lambda_i \geq 0$ is the eigenvalue of $\mathcal{A}^*\mathcal{A}$ associated to the subspace W_i . Consequently, formula (5.18) can be written as

$$\sigma_f^2 = \sum_i \langle f_i, -\mathcal{L}_0(\mathcal{L}_0^2 + \mu^2 \lambda_i)^{-1} f_i \rangle_{L^2(\widehat{\pi})}, \quad (5.28)$$

where $f = \sum_i f_i$ and $f_i \in W_i$. Let us assume now without loss of generality that $W_0 = \ker \mathcal{A}^*\mathcal{A}$, so in particular $\lambda_0 = 0$. Then clearly

$$\lim_{\mu \rightarrow \infty} \sigma_f^2 = \langle f_0, -\mathcal{L}_0(\mathcal{L}_0^2)^{-1} f_0 \rangle_{L^2(\widehat{\pi})} = \langle f_0, (-\mathcal{L}_0)^{-1} f_0 \rangle_{L^2(\widehat{\pi})} = \sigma_{f_0}^2(0).$$

Now notice that $W_0 = \ker \mathcal{A}^*\mathcal{A} = \ker \mathcal{A}$, proving the equality in the claim. It remains to show that $\sigma_{\Pi f}^2(0) \leq \sigma_f^2(0)$. To see this, we write

$$\begin{aligned} \sigma_f^2(0) &= \langle f, (-\mathcal{L}_0)^{-1} f \rangle_{L^2(\widehat{\pi})} = \langle \Pi f + (1 - \Pi)f, (-\mathcal{L}_0)^{-1} (\Pi f + (1 - \Pi)f) \rangle_{L^2(\widehat{\pi})} \\ &= \sigma_{\Pi f}^2(0) + \sigma_{(1 - \Pi)f}^2(0) + R, \end{aligned}$$

where

$$R = \langle \Pi f, (-\mathcal{L}_0)^{-1} (1 - \Pi)f \rangle_{L^2(\widehat{\pi})} + \langle (1 - \Pi)f, (-\mathcal{L}_0)^{-1} \Pi f \rangle_{L^2(\widehat{\pi})}.$$

Note that since we only consider observables that do not depend on p , $\Pi f \in \ker(Jq \cdot \nabla_q)$ and $(1 - \Pi)f \in \bigoplus_{i \geq 1} W_i$. Since $\mathcal{L}_0$ commutes with $\mathcal{A}$, it follows that $(-\mathcal{L}_0)^{-1}$ leaves both W_0 and $\bigoplus_{i \geq 1} W_i$ invariant. Therefore, as the latter spaces are orthogonal to each other, we deduce that $R = 0$, from which the result follows. $\square$

From Theorem 5.3, it follows that in the limit as $\mu \rightarrow \infty$, the asymptotic variance $\sigma_f^2(\mu)$

[†]Using the decomposition of $L^2(\widehat{\pi})$ into Hermite polynomials provided by Theorem 5.2, this is the case if and only if the corresponding drift matrix is diagonalisable, i.e. if and only if $\gamma \neq 2$.

[‡]Indeed, the fact that $f \in \ker \mathcal{A}$ is equivalent to $f \in \bar{E}$ is easy to check if f is an eigenvector of $\mathcal{L}_0$ (recall that f is then an eigenvector of $\mathcal{L}_0 + \mu \mathcal{A}$ as well, using Lemma 5.2(d)). The claim then follows by extending linearly.

is not decreased by the perturbation if $f \in \ker(Jq \cdot \nabla_q)$. In fact, this result also holds true non-asymptotically, i.e. observables in $\ker(Jq \cdot \nabla_q)$ are not affected at all by the perturbation:

Lemma 5.3. *Let $f \in \ker(Jq \cdot \nabla_q)$. Then $\sigma_f^2(\mu) = \sigma_f^2(0)$ for all $\mu \in \mathbb{R}$.*

Proof. From $f \in \ker(Jq \cdot \nabla_q)$ it follows immediately that $f \in \ker \mathcal{A}^* \mathcal{A}$. Then the claim follows from the expression (5.28). $\square$

Example 5.3. Recall the case of observables of the form $f(q) = q \cdot Kq + l \cdot q + C$ with $K \in \mathbb{R}_{\text{sym}}^{d \times d}$, $l \in \mathbb{R}^d$ and $C \in \mathbb{R}$ from Section 5.1. If $[J, K] = 0$ and $l \in \ker J$, then $f \in \ker(Jq \cdot \nabla_q)$ as

$$Jq \cdot \nabla_q(q \cdot Kq + l \cdot q + C) = 2Jq \cdot Kq + Jq \cdot l = q \cdot (KJ - JK)q - q \cdot Jl = 0.$$

From the preceding lemma it follows that $\sigma_f^2(\mu) = \sigma_f^2(0)$ for all $\mu \in \mathbb{R}$, showing that the assumption in Theorem 5.1 does not exclude nontrivial cases.

The following result shows that the dynamics (5.3) is particularly effective for antisymmetric observables (at least in the limit of large perturbations):

Proposition 5.4. *Let $f \in L_0^2(\pi)$ satisfy $f(-q) = -f(q)$ and assume that $\ker J = \{0\}$. Furthermore, assume that the eigenvalues of J are rationally independent, i.e.*

$$\sigma(J) = \{\pm i\lambda_1, \pm i\lambda_2, \dots, \pm i\lambda_d\} \quad (5.29)$$

with $\lambda_i \in \mathbb{R}_{>0}$ and $\sum_i k_i \lambda_i \neq 0$ for all $(k_1, \dots, k_d) \in \mathbb{Z}^d \setminus (0, \dots, 0)$. Then

$$\lim_{\mu \rightarrow \infty} \sigma_f^2(\mu) = 0.$$

Proof of Proposition 5.4. The claim would immediately follow from $f \in \ker(Jq \cdot \nabla)^\perp$ according to Theorem 5.3, but that does not seem to be so easy to prove directly. Instead, we again make use of the Hermite polynomials.

Recall from the proof of Proposition 5.3 that $\mathcal{L}$ is invertible on $L_0^2(\widehat{\pi})$ and its inverse leaves the Hermite spaces H_m invariant. Consequently, the asymptotic variance of an observable $f \in L_0^2(\widehat{\pi})$ can be written as

$$\sigma_f^2 = \langle f, (-\mathcal{L})^{-1} f \rangle_{L^2(\widehat{\pi})} = \sum_{m=1}^{\infty} \langle \Pi_m f, (-\mathcal{L}|_{H_m})^{-1} \Pi_m f \rangle_{L^2(\widehat{\pi})}, \quad (5.30)$$

where $\Pi_m : L_0^2(\widehat{\pi}) \rightarrow H_m$ denotes the orthogonal projection onto H_m . From (5.17) it is clear that h_a is symmetric for $|\alpha|$ even and antisymmetric for $|\alpha|$ odd. Therefore, from f

being antisymmetric it follows that $f \in \bigoplus_{m \geq 1, m \text{ odd}} H_m$. In view of (5.12), statement (c) in Theorem 5.2 and (5.29) the spectrum of $\mathcal{L}|_{H_m}$ can be written as

$$\sigma(\mathcal{L}|_{H_m}) = \left\{ \mu \sum_{j=1}^{2d} \alpha_j \beta_j + C_{\alpha, \gamma} : |\alpha| = m, \beta_j \in \sigma(J) \right\} \quad (5.31a)$$

$$= \left\{ i\mu \sum_{j=1}^d (\alpha_j - \alpha_{j+d}) \lambda_j + C_{\alpha, \gamma} : |\alpha| = m \right\}, \quad (5.31b)$$

with appropriate real constants $C_{\alpha, \gamma} \in \mathbb{R}$ that depend on α and γ , but not on μ . For $|\alpha| = \sum_{j=1}^{2d} \alpha_j = m$ odd, we have that

$$\sum_{j=1}^d (\alpha_j - \alpha_{j+d}) \lambda_j \neq 0. \quad (5.32)$$

Indeed, assume to the contrary that the above expression is zero. Then it would follow that $\alpha_j = \alpha_{j+d}$ for all $j = 1, \dots, d$ by rational independence of $\lambda_1, \dots, \lambda_d$ and $|m|$ would have to be even. From (5.31b) and (5.32) it is clear that

$$\sup \{r > 0 : B(0, r) \cap \sigma(\mathcal{L}|_{H_m}) = \emptyset\} \xrightarrow{\mu \rightarrow \infty} \infty,$$

where $B(0, r)$ denotes the ball of radius r centered at the origin in $\mathbb{C}$. Consequently, the spectral radius of $(-\mathcal{L}|_{H_m})^{-1}$ and hence $(-\mathcal{L}|_{H_m})^{-1}$ itself converges to zero as $\mu \rightarrow \infty$. The result then follows from (5.30). $\square$

Remark 5.7. *The idea of the preceding proof can be explained using Figure 5.2a and Remark 5.6. The eigenvalues in the fixed spectrum $\bar{\sigma}$ (on the real axis, highlighted by diamonds) correspond to Hermite polynomials of even order. The independence condition on the eigenvalues of J prevents cancellations that would lead to fixed eigenvalues associated to Hermite polynomials of odd order. Therefore, antisymmetric observables are orthogonal to $\bar{E} = \ker \mathcal{A}$.*

The following corollary gives a version of the converse of Proposition 5.4 and provides further intuition into the mechanics of the variance reduction achieved by the perturbation.

Corollary 5.2. *Let $f \in L_0^2(\pi)$ and assume that $\lim_{\mu \rightarrow \infty} \sigma_f^2(\mu) = 0$. Then*

$$\int_{B(0, r)} f \, dq = 0$$

for all $r \in (0, \infty)$, where $B(0, r)$ denotes the ball centered at 0 with radius r .

Proof. According to Theorem 5.3, $\lim_{\mu \rightarrow \infty} \sigma_f^2(\mu) = 0$ implies $\sigma_{\Pi f}^2(0) = 0$. We can write

$$\sigma_{\Pi f}^2(0) = \langle \Pi f, (-\mathcal{L}_0)^{-1} \Pi f \rangle_{L^2(\widehat{\pi})} = \frac{1}{2} \langle \Pi f, ((-\mathcal{L}_0)^{-1} + (-\mathcal{L}_0^*)^{-1}) \Pi f \rangle_{L^2(\widehat{\pi})} \quad (5.33)$$

and recall from the proof of Proposition 5.3 that $(-\mathcal{L}_0)^{-1}$ and $(-\mathcal{L}_0^*)^{-1}$ leave the Hermite spaces H_m invariant. Therefore $\ker((-\mathcal{L}_0)^{-1} + (-\mathcal{L}_0^*)^{-1}) = \{0\}$ in $L_0^2(\widehat{\pi})$, and in particular $\sigma_{\Pi f}^2(0) = 0$ is equivalent to $\Pi f = 0$. This in turn shows that $f \in \ker(Jq \cdot \nabla)^\perp$. From $(Jq \cdot \nabla)^* = -Jq \cdot \nabla$, it follows that

$$\ker(Jq \cdot \nabla)^\perp = \overline{\operatorname{im}(Jq \cdot \nabla)^*} = \overline{\operatorname{im}(Jq \cdot \nabla)}. \quad (5.34)$$

Hence, there exists a sequence $(\phi_j)_j \in C_c^\infty(\mathbb{R}^d)$ such that $Jq \cdot \nabla \phi_j \rightarrow f$ in $L^2(\pi)$. Taking a subsequence if necessary, we can assume that the convergence is pointwise π -almost everywhere and that the sequence is pointwise bounded by a function in $L^1(\pi)$. Since J is antisymmetric, we have that $Jq \cdot \nabla \phi_j = \nabla \cdot (\phi_j Jq)$. Now Gauss's theorem yields

$$\int_{B(0,r)} \nabla \cdot (\phi_j Jq) \, dq = \int_{\partial B(0,r)} \phi_j Jq \cdot n \, dn,$$

where n denotes the outward normal to the sphere $\partial B(0,r)$. This quantity is zero due to the orthogonality of Jq and n , and so the result follows from Lebesgue's dominated convergence theorem. $\square$

5.4 OPTIMAL CHOICES OF J FOR LINEAR AND QUADRATIC OBSERVABLES

Assume that $f \in L_0^2(\pi)$ is given by

$$f(q) = q \cdot Kq + l \cdot q - \operatorname{Tr} K,$$

with $K \in \mathbb{R}_{\text{sym}}^{d \times d}$ and $l \in \mathbb{R}^d$ (note that the constant term is adjusted to guarantee that $\pi(f) = 0$). Our objective is to choose J in such a way that $\lim_{\mu \rightarrow \infty} \sigma_f^2(\mu)$ becomes as small as possible. To stress the dependence on the choice of J , we introduce the notation $\sigma_f^2(\mu, J)$. Also, we denote the orthogonal projection onto $(\ker J)^\perp$ by $\Pi_{\ker J}^\perp$.

Lemma 5.4 (Zero variance limit for linear observables). *Assume that $K = 0$ and $\Pi_{\ker J}^\perp l = 0$. Then*

$$\lim_{\mu \rightarrow \infty} \sigma_f^2(\mu, J) = 0.$$

Proof. According to Theorem 5.3, we have to show that $\Pi f = 0$, where Π is the $L^2(\pi)$ -orthogonal projection onto $\ker(Jq \cdot \nabla)$. Let us thus use (5.34) and prove that $f \in \overline{\operatorname{im}(Jq \cdot \nabla)}$. Indeed, since $\Pi_{\ker J}^\perp l = 0$, by Fredholm's alternative there exists $u \in \mathbb{R}^d$ such that $Ju = l$. Now define $\phi \in L_0^2(\pi)$ by $\phi(q) = -u \cdot q$, leading to $f = Jq \cdot \nabla \phi$, so the result

follows. □

Lemma 5.5 (Zero variance limit for purely quadratic observables). *Let $l = 0$ and consider the decomposition $K = K_0 + K_1$ into the traceless part $K_0 = K - \frac{\text{Tr } K}{d} \cdot I$ and the trace-part $K_1 = \frac{\text{Tr } K}{d} \cdot I$. For the corresponding decomposition of the observable*

$$f(q) = f_0(q) + f_1(q) = q \cdot K_0 q + q \cdot K_1 q - \text{Tr } K$$

the following holds:

(a) *There exists a skew-symmetric matrix J such that*

$$\lim_{\mu \rightarrow \infty} \sigma_{f_0}^2(\mu, J) = 0, \quad (5.35)$$

and there is an algorithmic way (see Algorithm 1) to compute an appropriate matrix J in terms of K .

(b) *The trace-part is not effected by the perturbation, i.e. $\sigma_{f_1}^2(\mu, J) = \sigma_{f_1}^2(0)$ for all $\mu \in \mathbb{R}$.*

Proof. To prove the first claim, according to Theorem 5.3 it is sufficient to show that $f_0 \in \ker(Jq \cdot \nabla)^\perp = \overline{\text{im}(Jq \cdot \nabla)}$. Let us consider the function $\phi(q) = q \cdot Aq$, with $A \in \mathbb{R}_{\text{sym}}^{d \times d}$. It holds that $Jq \cdot \nabla \phi = 2q \cdot (J^T A q) = q \cdot [A, J]q$. The task of finding a skew-symmetric matrix J such that $\lim_{\mu \rightarrow \infty} \sigma_{f_0}^2(\mu, J) = 0$ can therefore be accomplished by constructing an antisymmetric matrix J such that there exists a symmetric matrix A with the property $K_0 = [A, J]$. Given any traceless matrix K_0 there exists an orthogonal matrix $U \in O(\mathbb{R}^d)$ such that UK_0U^T has zero entries on the diagonal, and U can be obtained in an algorithmic manner (see Section 8.5). Assume thus that such a matrix $U \in O(\mathbb{R}^d)$ has been found and choose real numbers $a_1, \dots, a_d \in \mathbb{R}$ such that $a_i \neq a_j$ if $i \neq j$. We now set $\bar{A} = \text{diag}(a_1, \dots, a_d)$, and

$$\bar{J}_{ij} = \begin{cases} \frac{(UK_0U^T)_{ij}}{a_i - a_j} & \text{if } i \neq j, \\ 0 & \text{if } i = j. \end{cases} \quad (5.36)$$

Observe that since UK_0U^T is symmetric, $\bar{J}$ is skew-symmetric. A short calculation shows that $[\bar{A}, \bar{J}] = UK_0U^T$. We can thus define $A = U^T \bar{A} U$ and $J = U^T \bar{J} U$ to obtain $[A, J] = K_0$. Therefore, the matrix J constructed in this way indeed satisfies (5.35). For the second claim, note that $f_1 \in \ker(Jq \cdot \nabla)$, since $Jq \cdot \nabla (q \cdot \frac{\text{Tr } K}{d} q) = 2 \frac{\text{Tr } K}{d} q \cdot Jq = 0$ due to the skew-symmetry of J . The result then follows from Lemma 5.3. □

We would like to stress that the perturbation J constructed in the previous lemma is far from unique due to the freedom of choice of U and $a_1, \dots, a_d \in \mathbb{R}$ in the proof. However, it is asymptotically optimal:

Corollary 5.3. *In the setting of Lemma 5.5 the following holds:*

$$\min_{J^T = -J} \left(\lim_{\mu \rightarrow \infty} \sigma_f^2(\mu, J) \right) = \sigma_{f_1}^2(0).$$

Proof. The claim follows immediately since $f_1 \in \ker(Jq \cdot \nabla)$ for arbitrary skew-symmetric J as shown in (5.4), and therefore the contribution of the trace part f_1 to the asymptotic variance cannot be reduced by any choice of J according to Lemma 5.3. $\square$

As the proof of Lemma 5.5 is constructive, we obtain the following algorithm to determine optimal perturbations for quadratic observables:

Algorithm 1. Given $K \in \mathbb{R}_{\text{sym}}^{d \times d}$, determine an optimal skew-symmetric perturbation J as follows:

1. Set $K_0 = K - \frac{\text{Tr} K}{d} \cdot I$.
2. Find $U \in O(\mathbb{R}^d)$ such that UK_0U^T has zero entries on the diagonal (see Section 8.5).
3. Choose $a_i \in \mathbb{R}$, $i = 1, \dots, d$ such that $a_i \neq a_j$ for $i \neq j$ and set

$$\bar{J}_{ij} = \frac{(UK_0U^T)_{ij}}{a_i - a_j}$$

for $i \neq j$ and $\bar{J}_{ii} = 0$ otherwise.

4. Set $J = U^T \bar{J} U$.

Remark 5.8. *In [41], the authors consider the task of finding optimal perturbations J for the nonreversible overdamped Langevin dynamics given in (1.74). In the Gaussian case this optimisation problem turns out to be equivalent to the one considered in this section. Indeed, equation (39) of [41] can be rephrased as $f \in \ker(Jq \cdot \nabla)^\perp$. Therefore, Algorithm 1 and its generalisation Algorithm 2 (described in Section 5.5) can be used without modifications to find optimal perturbations of overdamped Langevin dynamics.*

5.5 GAUSSIANS WITH ARBITRARY COVARIANCE AND PRECONDITIONING

In this section we extend the results of the preceding sections to the case when the target measure π is given by a Gaussian with arbitrary covariance, i.e. $V(q) = \frac{1}{2}q \cdot Sq$ with $S \in \mathbb{R}_{\text{sym}}^{d \times d}$ symmetric and positive definite. The dynamics (4.1) then takes the form

$$dq_t = M^{-1}p_t dt - \mu J_1 S q_t dt, \tag{5.37a}$$

$$dp_t = -S q_t dt - \nu J_2 M^{-1}p_t dt - \Gamma M^{-1}p_t dt + \sqrt{2\Gamma} dW_t. \tag{5.37b}$$

The key observation is now that the choices $M = S$ and $\Gamma = \gamma S$ together with the transformation $\tilde{q} = S^{1/2}q$ and $\tilde{p} = S^{-1/2}p$ lead to the dynamics

$$d\tilde{q}_t = \tilde{p}_t dt - \mu S^{1/2} J_1 S^{1/2} \tilde{q}_t dt, \quad (5.38a)$$

$$d\tilde{p}_t = -\tilde{q}_t dt - \mu S^{-1/2} J_2 S^{-1/2} \tilde{p}_t dt - \gamma \tilde{p}_t dt + \sqrt{2\gamma} dW_t, \quad (5.38b)$$

which is of the form (5.3) if J_1 and J_2 obey the condition $SJ_1S = J_2$. Clearly the dynamics (5.38) is ergodic with respect to a Gaussian measure with unit covariance, in the following denoted by $\tilde{\pi}$. The connection between the asymptotic variances associated to (5.37) and (5.38) is as follows:

For an observable $f \in L_0^2(\pi)$ we can write

$$\sqrt{T} \left(\frac{1}{T} \int_0^T f(q_s) ds - \pi(f) \right) = \sqrt{T} \left(\frac{1}{T} \int_0^T \tilde{f}(\tilde{q}_s) ds - \tilde{\pi}(\tilde{f}) \right),$$

where $\tilde{f}(q) = f(S^{-1/2}q)$. Therefore, the asymptotic variances satisfy $\sigma_f^2 = \tilde{\sigma}_{\tilde{f}}^2$, where $\tilde{\sigma}_{\tilde{f}}^2$ denotes the asymptotic variance associated to the process $(\tilde{q}_t)_{t \geq 0}$. Because of this, the results from the previous sections generalise to (5.37), subject to the condition that the choices $M = S$, $\Gamma = \gamma S$ and $SJ_1S = J_2$ are made. We formulate our results in this general setting as corollaries:

Corollary 5.4. *Consider the dynamics*

$$dq_t = M^{-1}p_t dt - \mu J_1 \nabla V(q_t) dt, \quad (5.39a)$$

$$dp_t = -\nabla V(q_t) dt - \mu J_2 M^{-1}p_t dt - \Gamma M^{-1}p_t dt + \sqrt{2\Gamma} dW_t, \quad (5.39b)$$

with $V(q) = \frac{1}{2}q \cdot Sq$. Assume that $M = S$, $\Gamma = \gamma S$ with $\gamma > \sqrt{2}$ and $SJ_1S = J_2$. Let $f \in L^2(\pi)$ be an observable of the form

$$f(q) = q \cdot Kq + l \cdot q + C \quad (5.40)$$

with $K \in \mathbb{R}_{\text{sym}}^{d \times d}$, $l \in \mathbb{R}^d$ and $C \in \mathbb{R}$. If at least one of the conditions $KJ_1S \neq SJ_1K$ and $l \notin \ker J$ is satisfied, then the asymptotic variance is at a local maximum for the unperturbed sampler, i.e.

$$\partial_\mu \sigma_f^2|_{\mu=0} = 0 \quad \text{and} \quad \partial_\mu^2 \sigma_f^2|_{\mu=0} < 0.$$

Proof. Note that

$$\tilde{f}(q) = f(S^{-1/2}q) = q \cdot S^{-1/2}KS^{-1/2}q + S^{-1/2}l \cdot q + C = q \cdot \tilde{K}q + \tilde{l} \cdot q + C$$

is again of the form (5.40) (where in the last equality, $\tilde{K} = S^{-1/2}KS^{-1/2}$ and $\tilde{l} = S^{-1/2}l$ have been defined). From (5.38), $\sigma_f^2 = \tilde{\sigma}_{\tilde{f}}^2$, and Theorem 5.1 the claim follows if at

least one of the conditions $[\tilde{K}, S^{1/2}J_1S^{1/2}] \neq 0$ and $\tilde{l} \notin \ker(S^{1/2}J_1S^{1/2})$ is satisfied. The first of those can easily be seen to be equivalent to $S^{-1/2}(KJS - SJK)S^{-1/2} \neq 0$, which is equivalent to $KJ_1S \neq SJ_1K$ since S is nondegenerate. The second condition is equivalent to $S^{1/2}J_1l \neq 0$, which is equivalent to $J_1l \neq 0$, again by nondegeneracy of S . $\square$

Corollary 5.5. *Assume the setting from the previous corollary and denote by Π the orthogonal projection onto $\ker(J_1Sq \cdot \nabla)$. For $f \in L^2(\pi)$ it holds that*

$$\lim_{\mu \rightarrow \infty} \sigma_f^2(\mu) = \sigma_{\Pi f}^2(0) \leq \sigma_f^2(0).$$

Proof. Theorem 5.3 implies $\lim_{\mu \rightarrow \infty} \tilde{\sigma}_{\tilde{f}}^2(\mu) = \tilde{\sigma}_{\tilde{\Pi} \tilde{f}}^2(0) \leq \tilde{\sigma}_{\tilde{f}}^2(0)$ for the transformed system (5.38). Here $\tilde{f}(q) = f(S^{-1/2}q)$ is the transformed observable and $\tilde{\Pi}$ denotes the $L^2(\pi)$ -orthogonal projection onto $\ker(S^{1/2}J_1S^{1/2}q \cdot \nabla)$. According to the relation $\sigma_f^2 = \tilde{\sigma}_{\tilde{f}}^2$, it is sufficient to show that $(\Pi f) \circ S^{-1/2} = \tilde{\Pi} \tilde{f}$. This however follows directly from the fact that the linear transformation $\phi \mapsto \phi \circ S^{1/2}$ maps $\ker(S^{1/2}J_1S^{1/2}q \cdot \nabla)$ bijectively onto $\ker(J_1Sq \cdot \nabla)$. $\square$

Let us also reformulate Algorithm 1 for the case of a Gaussian target with arbitrary covariance.

Algorithm 2. Given $K, S \in \mathbb{R}_{\text{sym}}^{d \times d}$ with $f(q) = q \cdot Kq$ and $V(q) = \frac{1}{2}q \cdot Sq$ (assuming S is nondegenerate), determine optimal perturbations J_1 and J_2 as follows:

1. Set $\tilde{K} = S^{-1/2}KS^{-1/2}$ and $\tilde{K}_0 = \tilde{K} - \frac{\text{Tr} \tilde{K}}{d} \cdot I$.
2. Find $U \in O(\mathbb{R}^d)$ such that $U\tilde{K}_0U^T$ has zero entries on the diagonal (see Section 8.5).
3. Choose $a_i \in \mathbb{R}$, $i = 1, \dots, d$ such that $a_i \neq a_j$ for $i \neq j$ and set

$$\bar{J}_{ij} = \frac{(U\tilde{K}_0U^T)_{ij}}{a_i - a_j}.$$

4. Set $\tilde{J} = U^T \bar{J} U$.
5. Put $J_1 = S^{-1/2}\tilde{J}S^{-1/2}$ and $J_2 = S^{1/2}\tilde{J}S^{1/2}$.

Finally, we obtain the following optimality result from Lemma 5.4 and Corollary 5.3.

Corollary 5.6. *Let $f(q) = q \cdot Kq + l \cdot q - \text{Tr} K$ and assume that $\Pi_{\ker J}^\perp l = 0$. Then*

$$\min_{J_1^T = -J_1, J_2 = SJ_1S} \left(\lim_{\mu \rightarrow \infty} \sigma_f^2(\mu, J_1, J_2) \right) = \sigma_{f_1}^2(0),$$

where $f_1(q) = q \cdot K_1 q$, $K_1 = \frac{\text{Tr}(S^{-1}K)}{d}S$. Optimal choices for J_1 and J_2 can be obtained using Algorithm 2.

Remark 5.9. Since in Section 5.1 we analysed the case where J_1 and J_2 are proportional, we are not able to drop the restriction $J_2 = SJ_1S$ from the above optimality result. Analysis of completely arbitrary perturbations might be the subject of future work.

Remark 5.10. The choices $M = S$ and $\Gamma = \gamma S$ have been introduced to make the considered perturbations lead to samplers that perform well in terms of reducing the asymptotic variance. However, adjusting the mass and friction matrices according to the target covariance in this way (i.e. $M = S$ and $\Gamma = \gamma S$) is a popular way of preconditioning the dynamics, see for instance the geometric approach in [55] and mass-tensor molecular dynamics [13]. Here we will present an argument as to why such a preconditioning is indeed beneficial in terms of the convergence rate of the dynamics. Let us first assume that S is diagonal, i.e. $S = \text{diag}(s^{(1)}, \dots, s^{(d)})$ and that $M = \text{diag}(m^{(1)}, \dots, m^{(d)})$ and $\Gamma = \text{diag}(\gamma^{(1)}, \dots, \gamma^{(d)})$ are chosen diagonally as well. Then (5.37) decouples into one-dimensional SDEs of the following form:

$$dq_t^{(i)} = \frac{1}{m^{(i)}} p_t^{(i)} dt, \quad (5.41a)$$

$$dp_t^{(i)} = -s^{(i)} q_t^{(i)} dt - \frac{\gamma^{(i)}}{m^{(i)}} p_t^{(i)} dt + \sqrt{2\gamma^{(i)}} dW_t, \quad i = 1, \dots, d. \quad (5.41b)$$

Let us write those Ornstein-Uhlenbeck processes as

$$dX_t^{(i)} = -B^{(i)} X_t^{(i)} dt + \sqrt{2Q^{(i)}} dW_t^{(i)}, \quad (5.42a)$$

$$B^{(i)} = \begin{pmatrix} 0 & -\frac{1}{m^{(i)}} \\ s^{(i)} & \frac{\gamma^{(i)}}{m^{(i)}} \end{pmatrix}, \quad Q^{(i)} = \begin{pmatrix} 0 & 0 \\ 0 & \gamma^{(i)} \end{pmatrix}. \quad (5.42b)$$

As in Section 5.2, the rate of the exponential decay of (5.42a) is given by $\min \text{Re } \sigma(B^{(i)})$. A short calculation shows that the eigenvalues of $B^{(i)}$ equal

$$\lambda_{1,2}^{(i)} = \frac{\gamma^{(i)}}{2m^{(i)}} \pm \sqrt{\left(\frac{\gamma^{(i)}}{2m^{(i)}}\right)^2 - \frac{s^{(i)}}{m^{(i)}}}.$$

Therefore, the rate of the exponential decay is maximal when

$$\left(\frac{\gamma^{(i)}}{2m^{(i)}}\right)^2 - \frac{s^{(i)}}{m^{(i)}} = 0, \quad (5.43)$$

in which case it is given by

$$(\lambda^{(i)})^* = \sqrt{\frac{s^{(i)}}{m^{(i)}}}.$$

Naturally, it is reasonable to choose $m^{(i)}$ in such a way that the exponential rate $(\lambda^{(i)})^*$

is the same for all $i = 1, \dots, d$, leading to the restriction $M = cS$ with $c > 0$. Choosing c small will result in fast convergence to equilibrium, but also make the dynamics (5.41) quite stiff, requiring a very small timestep Δt in a discretisation scheme. The choice of c will therefore need to strike a balance between those two competing effects. The constraint (5.43) then implies $\Gamma = 2cS$. By a coordinate transformation, the preceding argument also applies if S , M and Γ are diagonal in the same basis, and of course M and Γ can always be chosen that way. Numerical experiments show that it is possible to increase the rate of convergence to equilibrium even further by choosing M and Γ nondiagonally with respect to S (although only by a small margin). A clearer understanding of this is a topic for further investigation.

6. NUMERICAL EXPERIMENTS

In this chapter we will illustrate our analytical results for the perturbed underdamped Langevin dynamics (4.1) by some numerical experiments in a high-dimensional setting. To this end, we introduce an appropriate integration scheme in Section 6.1. The actual experiment will be detailed in Section 6.2.

6.1 NUMERICAL SCHEME

In this section we introduce a splitting scheme for simulating the perturbed underdamped Langevin dynamics given by equation (4.1). In the unperturbed case, i.e. when $J_1 = J_2 = 0$, the right-hand side can be decomposed into parts A , B and C according to

$$d \begin{pmatrix} q_t \\ p_t \end{pmatrix} = \underbrace{\begin{pmatrix} M^{-1}p_t \\ \mathbf{0} \end{pmatrix} dt}_A + \underbrace{\begin{pmatrix} \mathbf{0} \\ -\nabla V(q_t) \end{pmatrix} dt}_B + \underbrace{\begin{pmatrix} \mathbf{0} \\ -\Gamma M^{-1} + \sqrt{2\Gamma} dW_t \end{pmatrix}}_O,$$

i.e. O refers to the Ornstein-Uhlenbeck part of the dynamics, whereas A and B stand for the momentum and position updates, respectively.

One particular splitting scheme which has proven to be efficient is the *BAOAB* scheme, (see [86] and references therein). The string of letters refers to the order in which the different parts are integrated, namely

$$\begin{aligned} p_{n+1/2} &= p_n - \frac{1}{2}\Delta t \nabla V(q_n), \\ q_{n+1/2} &= q_n + \frac{1}{2}\Delta t \cdot M^{-1} p_{n+1/2}, \\ \hat{p} &= \exp(-\Delta t \Gamma M^{-1}) p_{n+1/2} + \sqrt{I - e^{-2\Gamma \Delta t}} \mathcal{N}(0, I), \\ q_{n+1} &= q_{n+1/2} + \frac{1}{2}\Delta t \cdot M^{-1} \hat{p}, \\ p_{n+1} &= \hat{p} - \frac{1}{2}\Delta t \cdot \nabla V(q_{n+1}). \end{aligned}$$

We note that many different discretisation schemes such as *ABOBA*, *OABAO*, etc. are viable, but that analytical and numerical evidence has shown that the *BAOAB*-ordering

has particularly good properties to compute long-time ergodic averages with respect to q -dependent observables. Motivated by this, we introduce the following perturbed scheme, introducing additional Runge-Kutta integration steps between the A , B and O parts:

$$p_{n+1/2} = p_n - \frac{1}{2}\Delta t \nabla V(q_n), \quad (6.2a)$$

$$q_{n+1/2} = q_n + \frac{1}{2}\Delta t \cdot M^{-1}p_{n+1/2}, \quad (6.2b)$$

$$q'_{n+1/2} = RK_4\left(\frac{1}{2}\Delta t, q_{n+1/2}\right), \quad (6.2c)$$

$$\hat{p} = \exp(-\Delta t(\Gamma M^{-1} + \nu J_2 M^{-1}))p_{n+1/2} + \sqrt{I - e^{-2\Gamma\Delta t}}\mathcal{N}(0, 1), \quad (6.2d)$$

$$q''_{n+1/2} = RK_4\left(\frac{1}{2}\Delta t, q'_{n+1/2}\right), \quad (6.2e)$$

$$q_{n+1} = q''_{n+1/2} + \frac{1}{2}\Delta t \cdot M^{-1}\hat{p}, \quad (6.2f)$$

$$p_{n+1} = \hat{p} - \frac{1}{2}\Delta t \cdot \nabla V(q_{n+1}), \quad (6.2g)$$

where $RK_4(\Delta t, q_0)$ refers to fourth order Runge-Kutta integration of the ODE

$$\dot{q} = -J_1 \nabla V(q), \quad q(0) = q_0, \quad (6.3)$$

up until time Δt . We remark that the J_2 -perturbation is linear and can therefore be included in the O -part without much computational overhead. We emphasize the fact that many different splitting schemes could be investigated: although the BAOAB-scheme works well for unperturbed underdamped Langevin dynamics, it is not clear whether this remains true for the perturbed version. Moreover, the perturbations introduced by J_1 and J_2 can be added in various places. Other discretisation schemes for the ODE (6.3) could be useful as well, for instance one could use a symplectic integrator, taking advantage of the Hamiltonian structure of (6.3). However, since V as the Hamiltonian for (6.3) is not separable in general, such a symplectic integrator would have to be implicit. Note that (6.2c) and (6.2e) could be merged since (6.2e) commutes with (6.2d). Here, we content ourselves with the above scheme for the numerical experiments. Investigation of optimal numerical schemes for perturbed Langevin dynamics is an interesting problem for further research.

Remark 6.1. *The aforementioned schemes lead to an error in the approximation for $\pi(f)$ since the invariant measure π is not preserved exactly by the numerical scheme. In practice, the BAOAB-scheme can therefore be accompanied by an accept-reject Metropolis step as in [88], leading to an unbiased estimate of $\pi(f)$, albeit with an inflated variance. In this case, after every rejection the momentum variable has to be flipped ($p \mapsto -p$) in order to keep the correct invariant measure. We note here that our perturbed scheme can be ‘Metropolised’ in a similar way by ‘flipping’ the matrices J_1 and J_2 after every rejection ($J_1 \mapsto -J_1$ and $J_2 \mapsto -J_2$) and using an appropriate (volume-preserving and*

time-reversible) integrator for the dynamics given by (6.3). A similar approach in the context of nonreversible overdamped Langevin dynamics has been investigated in [120].

6.2 DIFFUSION BRIDGE SAMPLING

To numerically test our analytical results, we will apply the dynamics (4.1) to sample a measure on path space associated to a diffusion bridge. Specifically, consider the SDE

$$dX_s = -\nabla U(X_s) ds + \sqrt{2\beta^{-1}} dW_s,$$

with $X_s \in \mathbb{R}^n$, $\beta > 0$ and the potential $U : \mathbb{R}^n \rightarrow \mathbb{R}$ obeying adequate growth and smoothness conditions (see [65, Section 5] for precise statements). The law of the solution to this SDE conditioned on the events $X(0) = x_-$ and $X(s_+) = x_+$ is a probability measure π on $L^2([0, s_+], \mathbb{R}^n)$ which poses a challenging and important sampling problem, especially if U is multimodal. This setting has been used as a test case for sampling probability measures in high dimensions (see for example [15] and [112]). For a more detailed introduction (including applications) see [17] and for a rigorous theoretical treatment the works [17, 64, 65, 66].

In the case $U \equiv 0$, it can be shown that the law of the conditioned process is given by a Gaussian measure π_0 with mean zero and precision operator $S = -\frac{\beta}{2}\Delta$ on the Sobolev space $H^1([0, s_+], \mathbb{R}^n)$ equipped with appropriate boundary conditions. The general case can then be understood as a perturbation thereof: The measure π is absolutely continuous with respect to π_0 with Radon-Nikodym derivative

$$\frac{d\pi}{d\pi_0} \propto \exp(-\Psi), \quad (6.4)$$

where

$$\Psi(x) = \frac{\beta}{2} \int_0^{s_+} G(x(s), \beta) ds \quad \text{and} \quad G(x, \beta) = \frac{1}{2} |\nabla U(x)|^2 - \frac{1}{\beta} \Delta U(x).$$

We will make the choice $x_- = x_+ = 0$, which is possible without loss of generality as explained in [16, Remark 3.1], leading to Dirichlet boundary conditions on $[0, s_+]$ for the precision operator S . Furthermore, we choose $s_+ = 1$ and discretise the ensuing s -interval $[0, 1]$ according to

$$[0, 1] = [0, s_1) \cup [s_1, s_2) \cup \dots \cup [s_{d-1}, s_d) \cup [s_d, 1],$$

in an equidistant way with stepsize $s_{j+1} - s_j \equiv \delta = \frac{1}{d+1}$. Functions on this grid are determined by the values $x(s_1) = x_1, \dots, x(s_d) = x_d$, recalling that $x(0) = x(1) = 0$ by

the Dirichlet boundary conditions. We discretise the functional Ψ as

$$\tilde{\Psi}(x_1, \dots, x_d) = \frac{\beta}{2} \delta \sum_{i=1}^d G(x_i, \beta) = \frac{\beta}{2} \delta \sum_{i=1}^d \left(\frac{1}{2} U'(x_i)^2 - \frac{1}{\beta} U''(x_i) \right),$$

such that its gradient is given by

$$(\nabla \tilde{\Psi})_i = \frac{\beta}{2} \delta (2U'(x_i)U''(x_i) - \frac{1}{\beta} U'''(x_i)), \quad i = 1, \dots, d.$$

We denote by A_δ the discretised Dirichlet Laplacian on $[0, 1]$ with stepsize δ . Following (6.4), the discretised target measure $\hat{\pi}$ has the form

$$\hat{\pi} = \frac{1}{Z} e^{-V} dx \quad \text{with} \quad V(x) = \tilde{\Psi}(x) - \frac{\beta\delta}{4} x \cdot A_\delta x, \quad x \in \mathbb{R}^d.$$

In the following we will consider the case $n = 1$ with potential $U : \mathbb{R} \rightarrow \mathbb{R}$ given by $U(x) = \frac{1}{2}(x^2 - 1)^2$ and set $\beta = 1$. To test our algorithm we adjust the parameters M , Γ , J_1 and J_2 according to the recommended choice in the Gaussian case, (4.2), where we take $S = \frac{\beta}{2} \delta \cdot A_\delta$ as the precision operator of the Gaussian target. We will consider the linear observable $f_1(x) = l \cdot x$ with $l = (1, \dots, 1)$ and the quadratic observable $f_2(x) = |x|^2$. In a first experiment we adjust the perturbations J_1 and J_2 to the observable f_2 according to Algorithm 2. The dynamics (4.1) is integrated using the splitting scheme introduced in Section 6.1 with a stepsize of $\Delta t = 10^{-4}$ over the time interval $[0, T]$ with $T = 10^2$. Furthermore, we choose initial conditions $q_0 = (1, \dots, 1)$, $p_0 = (0, \dots, 0)$ and introduce a burn-in time $T_0 = 1$, i.e. we take the estimator to be $\hat{\pi}(f) \approx \frac{1}{T-T_0} \int_{T_0}^T f(q_t) dt$. We compute the variance of the above estimator from $N = 500$ realisations and compare the results for different choices of the friction coefficient γ and of the perturbation strength μ .

The numerical experiments show that the perturbed dynamics generally outperform the unperturbed dynamics independently of the choices of μ and γ , both for linear and quadratic observables. One notable exception is the behaviour of the linear observable for small friction $\gamma = 10^{-3}$ (see Figure 6.1a), where the asymptotic variance initially increases for small perturbation strengths μ . This does not contradict our analytical results, since the small perturbation results Theorem 5.1 and Corollary 5.4 are only valid if $\gamma > \sqrt{2}$. We remark here that the condition $\gamma > \sqrt{2}$, while necessary for the theoretical results from Section 5.1, is not a very advisable choice in practice (at least in this experiment), since Figures 6.1b and 6.2b clearly indicate that the optimal friction is located around $\gamma \approx 10^{-1}$. Interestingly, the problem of choosing a suitable value for the friction coefficient γ becomes mitigated by the introduction of the perturbation: While the performance of the unperturbed sampler depends quite sensitively on γ , the asymptotic variance of the perturbed dynamics is a lot more stable with respect to variations of γ . A somewhat surprising phenomenon is that the standard deviation σ associated to the linear observable decays in the range $\gamma \in [10, 100]$ for the unperturbed sampler (see Figure 6.1b).

We confirmed this behaviour by further numerical experiments and remark that as the target measure $\hat{\pi}$ is fairly complicated, convexity of the function $\gamma \mapsto \sigma$ should not be expected.

In the regime of growing values of μ , the experiments confirm the results from Section 5.3, i.e. the asymptotic variance approaches a limit that is smaller than the asymptotic variance of the unperturbed dynamics.

As a final remark we report our finding that the performance of the sampler for the linear observable is qualitatively independent of the choice of J_1 (as long as J_2 is adjusted according to (4.2)). This result is in alignment with Proposition 5.4 which predicts good properties of the sampler for antisymmetric observables. In contrast to this, a judicious choice of J_1 is critical for quadratic observables. In particular, applying Algorithm 2 significantly improves the performance of the perturbed sampler in comparison to choosing J_1 arbitrarily.

Figure 6.1: Standard deviation of $\hat{\pi}(f)$ for a linear observable as a function of the friction γ and the perturbation strength μ

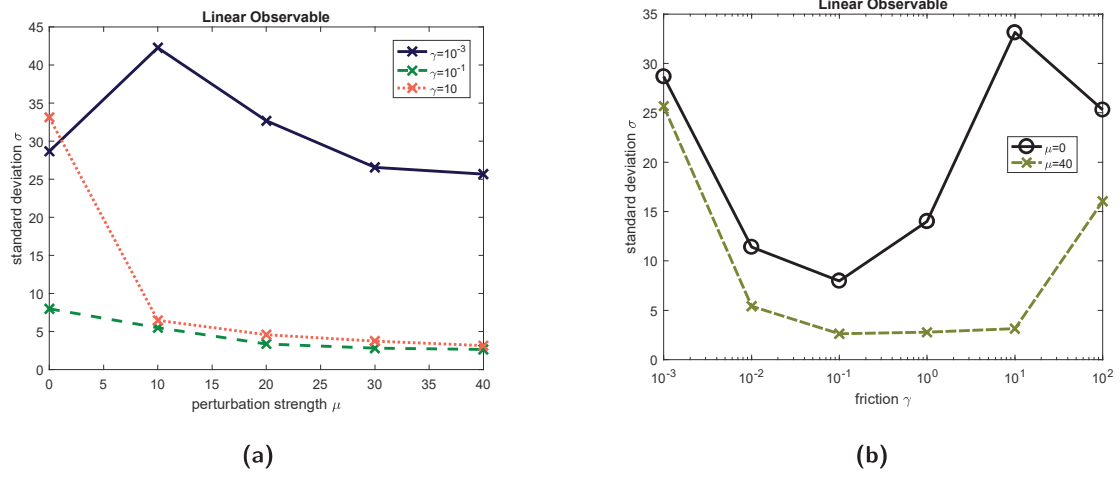
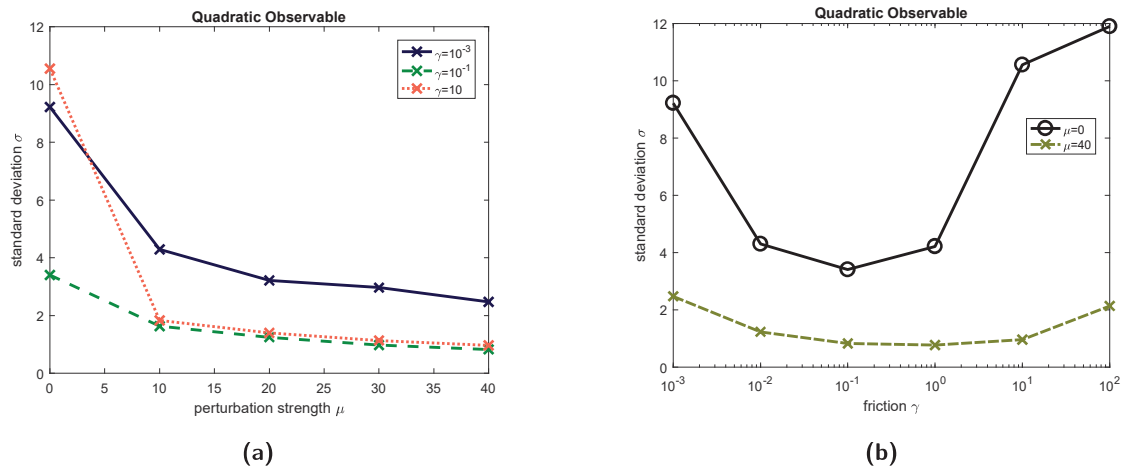


Figure 6.2: Standard deviation of $\hat{\pi}(f)$ for a quadratic observable as a function of the friction γ and the perturbation strength μ



7. DISCUSSION

Starting in Chapter 4 we introduced and analysed a new family of Langevin samplers consisting of perturbations of the underdamped Langevin dynamics, adding auxiliary drift terms in the equations for both the position and the momentum variables in such a way that ergodicity with respect to the fixed invariant distribution is maintained. These new Langevin samplers were studied in detail for Gaussian target distributions where it was shown, using tools from spectral theory for differential operators, that an appropriate choice of the parameters can improve the performance of the sampler, at least in terms of reducing the asymptotic variance. The performance of the perturbed Langevin sampler when applied to non-Gaussian target densities was tested numerically on the problem of diffusion bridge sampling.

The work presented here can be improved and extended in several directions. First, a rigorous analysis of the perturbed underdamped Langevin dynamics for non-Gaussian target densities is needed. The analytical tools developed in [41] can be used as a starting point. Furthermore, the study of the actual computational cost and its minimisation by an appropriate choice of the numerical scheme and of the perturbations in position and momentum would be of interest to practitioners. In addition, the analysis of our proposed samplers can be facilitated by using tools from symplectic and differential geometry. Finally, the mathematical formalism underlying the perturbation of Ψ -reversible processes introduced in Chapter 2 and leveraged in the proofs of Chapter 5 is not exclusive to underdamped Langevin dynamics. As pointed out in Example 2.2, the zigzag process enjoys very similar structural properties. We hence envisage that an analogous approach might be fruitful in the investigation of perturbations of piecewise deterministic Markov processes. All of these topics are currently under investigation.

8. PROOFS AND ADDITIONAL MATERIAL

This chapter contains various proofs referred to in the main body of Part II. In Section 8.1 we show that the generator (4.5) is hypoelliptic (Lemma 4.1) and hence the associated process possesses the smoothing properties guaranteed by Theorem 1.6. Section 8.2 is devoted to the proof of the overdamped limit for perturbed underdamped Langevin dynamics (Proposition 4.1), connecting it to the nonreversible overdamped Langevin dynamics. In Section 4.1 we give an overview of the theory of hypocoercivity [141], used in order to prove the crucial exponential decay estimate (1.46) for the perturbed underdamped Langevin dynamics. Section 8.4 is comprised of the computations for the derivatives of the asymptotic variance in the Gaussian case as presented in Section 5.1 of Chapter 5. Finally, we present in Section 8.5 an algorithmic way of orthogonally transforming a symmetric tracefree matrix into one with zeros on the diagonal. The former is a crucial ingredient in the Algorithms 1 and 2 presented in Section 5.4 of Chapter 5.

8.1 HYPOELLIPTICITY

The proof of Lemma 4.1 proceeds by checking the conditions of Hörmander's theorem (Theorem 1.5). A very similar argument showing the hypocoercivity of unperturbed underdamped Langevin dynamics can be found in [90, Section 2.2.3.1].

Proof of Lemma 4.1. We first note that $\mathcal{L}$ in (4.5) can be written in the ‘sum of squares’ form,

$$\mathcal{L} = \mathbf{X}_0 + \frac{1}{2} \sum_{k=1}^d \mathbf{X}_k^2,$$

where

$$\mathbf{X}_0 = M^{-1}p \cdot \nabla_q - \nabla_q V \cdot \nabla_p - \mu J_1 \nabla_q V \cdot \nabla_q - \nu J_2 M^{-1}p \cdot \nabla_p - \Gamma M^{-1}p \cdot \nabla_p$$

and

$$\mathbf{X}_k = e_k \cdot \Gamma^{1/2} \nabla_p, \quad k = 1, \dots, d.$$

Here $\{e_k\}_{k=1}^d$ denotes the standard Euclidean basis and $\Gamma^{1/2}$ is the unique positive definite

square root of the matrix Γ . The relevant commutators turn out to be

$$[\mathbf{X}_0, \mathbf{X}_k] = e_k \cdot \Gamma^{1/2} M^{-1} (\Gamma \nabla_p - \nabla_q - \nu J_2 \nabla_p), \quad k = 1, \dots, d.$$

Because Γ has full rank on $\mathbb{R}^d$, it follows that

$$\text{span}\{\mathbf{X}_k : k = 1, \dots, d\} = \text{span}\{\partial_{p_k} : k = 1, \dots, d\}.$$

Since

$$e_k \cdot \Gamma^{1/2} M^{-1} (\Gamma \nabla_p - \nu J_2 \nabla_p) \in \text{span}\{\mathbf{X}_j : j = 1, \dots, d\}, \quad k = 1, \dots, d,$$

and $\text{span}(\{(\Gamma^{1/2} M^{-1} \nabla_q)_k : k = 1, \dots, d\}) = \text{span}\{\partial_{q_k} : k = 1, \dots, d\}$, it follows that

$$\begin{aligned} & \text{span}(\{\mathbf{X}_k : k = 1, \dots, d\} \cup \{[\mathbf{X}_0, \mathbf{X}_k] : k = 1, \dots, d\}) \\ &= \text{span}(\{\partial_{p_k} : k = 1, \dots, d\} \cup \{\partial_{q_k} : k = 1, \dots, d\}), \end{aligned}$$

so the assumptions of Hörmander's theorem hold. $\square$

8.2 THE OVERDAMPED LIMIT

We begin with a technical lemma required for the proof of Proposition 4.1:

Lemma 8.1. *Assume that the conditions from Proposition 4.1 are satisfied. Then for every $T > 0$ there exists $C > 0$ such that*

$$\mathbb{E} \left(\sup_{0 \leq t \leq T} |p_t^\varepsilon|^2 \right) \leq C.$$

Proof. Using variation of constants, we can write the second line of (4.7) as

$$\begin{aligned} p_t^\varepsilon &= e^{-\frac{t}{\varepsilon^2}(\nu J_2 + \Gamma)M^{-1}} p_0 - \frac{1}{\varepsilon} \int_0^t e^{-\frac{(t-s)}{\varepsilon^2}(\nu J_2 + \Gamma)M^{-1}} \nabla_q V(q_s^\varepsilon) ds \\ &\quad + \frac{1}{\varepsilon} \sqrt{2\Gamma} \int_0^t e^{-\frac{(t-s)}{\varepsilon^2}(\nu J_2 + \Gamma)M^{-1}} dW_s. \end{aligned}$$

By Youngs's inequality, we see that

$$\mathbb{E} \sup_{0 \leq t \leq T} |p_t^\varepsilon|^2 \leq 3 \cdot \left(\sup_{0 \leq t \leq T} \left| e^{-\frac{t}{\varepsilon^2}(\nu J_2 + \Gamma)M^{-1}} p_0 \right|^2 \right. \tag{8.3a}$$

$$\left. + \frac{1}{\varepsilon^2} \mathbb{E} \sup_{0 \leq t \leq T} \left| \int_0^t e^{-\frac{(t-s)}{\varepsilon^2}(\nu J_2 + \Gamma)M^{-1}} \nabla_q V(q_s^\varepsilon) ds \right|^2 \right. \tag{8.3b}$$

$$\left. + \frac{1}{\varepsilon^2} \mathbb{E} \sup_{0 \leq t \leq T} \left| \sqrt{2\Gamma} \int_0^t e^{-\frac{(t-s)}{\varepsilon^2}(\nu J_2 + \Gamma)M^{-1}} dW_s \right|^2 \right). \tag{8.3c}$$

Since $\sigma((\nu J_2 + \Gamma)M^{-1}) = \sigma(M^{-1/2}(\nu J_2 + \Gamma)M^{-1/2}) \subset \mathbb{R}_{>0} \times i\mathbb{R}$, the term on the right-hand side of (8.3a) is bounded. For the term in (8.3b), observe that

$$\frac{1}{\varepsilon^2} \mathbb{E} \sup_{0 \leq t \leq T} \left| \int_0^t e^{-\frac{(t-s)}{\varepsilon^2}(\nu J_2 + \Gamma)M^{-1}} \nabla_q V(q_s^\varepsilon) ds \right|^2 \leq \frac{1}{\varepsilon^2} \sup_{0 \leq t \leq T} \int_0^t \left\| e^{-\frac{(t-s)}{\varepsilon^2}(\nu J_2 + \Gamma)M^{-1}} \right\|^2 ds, \quad (8.4)$$

since $V \in C^1(\mathbb{T}^d)$ and therefore $\nabla_q V$ is bounded. By the basic matrix exponential estimate $\|e^{-t(\nu J_2 + \Gamma)M^{-1}}\| \leq Ce^{-\omega t}$ for suitable C and ω , we see that (8.4) can further be bounded by

$$\frac{1}{\varepsilon^2} C \sup_{0 \leq t \leq T} \int_0^t e^{-2\omega \frac{(t-s)}{\varepsilon^2}} ds = \frac{C}{2\omega} \left(1 - e^{-2\omega \frac{T}{\varepsilon^2}}\right),$$

so this term is bounded as well. The term in (8.3c) is bounded by the Burkholder–Davis–Gundy inequality and a similar argument to the one used for the term in (8.3b) applies. $\square$

We now proceed to the proof of Proposition 4.1:

Proof of Proposition 4.1. Equations (4.7) can be written in integral form as

$$(\nu J_2 + \Gamma)q_t^\varepsilon = (\nu J_2 + \Gamma)q_0^\varepsilon + \frac{1}{\varepsilon} \int_0^t (\nu J_2 + \Gamma)M^{-1}p_s^\varepsilon ds - \mu \int_0^t (\nu J_2 + \Gamma)J_1 \nabla_q V(q_s^\varepsilon) ds$$

and

$$- \int_0^t \nabla V(q_s^\varepsilon) ds - \frac{1}{\varepsilon} \int_0^t (\nu J_2 + \Gamma)M^{-1}p_s^\varepsilon ds + \sqrt{2\Gamma}W(t) = \varepsilon(p_t^\varepsilon - p_0), \quad (8.5)$$

where the first line has been multiplied by the matrix $\nu J_2 + \Gamma$. Combining both equations yields

$$\begin{aligned} q_t^\varepsilon &= q_0^\varepsilon - \int_0^t (\nu J_2 + \Gamma) \nabla_q V(q_s^\varepsilon) ds - \varepsilon(\nu J_2 + \Gamma)^{-1}(p_t^\varepsilon - p_0) \\ &\quad - \mu \int_0^t J_1 \nabla_q V(q_s) ds + (\nu J_2 + \Gamma)^{-1} \sqrt{2\Gamma}W_t. \end{aligned}$$

Now applying Lemma 8.1 gives the desired result, since the above equation differs from the integral version of (4.8) only by the term $\varepsilon(\nu J_2 + \Gamma)^{-1}(p_t^\varepsilon - p_0)$ which vanishes in the limit as $\varepsilon \rightarrow 0$. $\square$

8.3 HYPOCOERCIVITY

The objective of this section is to prove that the perturbed dynamics (4.1) converges to equilibrium exponentially fast, i.e. that the associated semigroup $(S_t)_{t \geq 0}$ satisfies the estimate (1.46). We will use the theory of hypocoercivity introduced in [141] (see also the exposition in [113, Section 6.2])*. We start by providing a brief review of this theory.

*An alternative would be to employ the framework developed in [40] as we do in Chapter 17.

Let $(\mathcal{H}, \langle \cdot, \cdot \rangle)$ be a real separable Hilbert space and consider two unbounded operators $\mathcal{A} : \mathcal{H} \rightarrow \mathcal{H}^m$ and $\mathcal{T} : \mathcal{H} \rightarrow \mathcal{H}$ with domains $\mathcal{D}(\mathcal{A})$ and $\mathcal{D}(\mathcal{T})$ respectively, $\mathcal{T}$ skew-symmetric. Let $\mathcal{D} \subset \mathcal{H}$ be a dense vectorspace such that $\mathcal{D} \subset \mathcal{D}(\mathcal{A}) \cap \mathcal{D}(\mathcal{T})$, i.e. the operations of $\mathcal{A}$ and $\mathcal{T}$ are authorised on $\mathcal{D}$. The theory of hypocoercivity under consideration is concerned with equations of the form

$$\partial_t h + Lh = 0, \quad (8.7)$$

as well as with the associated semigroup $(S_t)_{t \geq 0}$ generated by $L = \mathcal{A}^* \mathcal{A} - \mathcal{T}$, where it is understood that $\mathcal{A}^* \mathcal{A} = \sum_{i=1}^m \mathcal{A}_i^* \mathcal{A}_i$. Let us also introduce the notation $K = \ker L$. With the choices $\mathcal{H} = L^2(\widehat{\pi})$, $\mathcal{A} = \Gamma^{1/2} \nabla_p$ and $\mathcal{T} = M^{-1} p \cdot \nabla_q - \nabla_q V \cdot \nabla_p - \mu J_1 \nabla_q V \cdot \nabla_q - \nu J_2 M^{-1} p \cdot \nabla_p$, it turns out that L is the (flat) $L^2(\mathbb{R}^{2d})$ -adjoint of the generator $\mathcal{L}$ given in (4.5) and therefore equation (8.7) is the Fokker-Planck equation (see (1.25)) associated to the dynamics (4.1). For convenience, we will use the notation $\sigma = \Gamma^{1/2}$ for the unique symmetric positive definite square root of Γ . In many situations of practical interest, the operator $\mathcal{A}^* \mathcal{A}$ is coercive only in certain directions of the state space, and therefore exponential return to equilibrium does not follow in general. In our case for instance, the noise acts only in the p -variables and therefore relaxation in the q -variables cannot be concluded a priori. However, intuitively speaking, the noise gets transported through the equations by the transport part of the dynamics. This is what the theory of hypocoercivity makes precise. Under some conditions on the interactions between $\mathcal{A}$ and $\mathcal{T}$ (encoded in their iterated commutators), exponential return to equilibrium can be proved. This idea is strongly related to Hörmander's theorem (Theorem 1.5), in the context of regularisation of solutions to hypoelliptic equations (hence the name 'hypocoercivity'). Let us remark that in Chapter 17 the problem of missing coercivity in certain directions is circumvented in a somewhat different manner. To state the main abstract theorem, we need the following definitions:

Definition 8.1 (Coercivity). *Let T be an unbounded operator on $\mathcal{H}$ with domain $\mathcal{D}(T)$ and kernel K . Assume that there exists another Hilbert space $(\tilde{\mathcal{H}}, \langle \cdot, \cdot \rangle_{\tilde{\mathcal{H}}})$, continuously and densely embedded in $K^\perp$. For a constant $\lambda > 0$, the operator T is said to be λ -coercive if*

$$\langle Th, h \rangle_{\tilde{\mathcal{H}}} \geq \lambda \|h\|_{\tilde{\mathcal{H}}}^2,$$

for all $h \in K^\perp \cap \mathcal{D}(T)$.

Definition 8.2 (Relative boundedness). *An operator T on $\mathcal{H}$ is said to be relatively bounded with respect to the operators $T_1, \dots, T_n$ if the intersection of the domains $\cap \mathcal{D}(T_j)$ is contained in $\mathcal{D}(T)$ and there exists a constant $\alpha > 0$ such that*

$$\|Th\| \leq \alpha (\|T_1 h\| + \dots + \|T_n h\|)$$

holds for all $h \in \mathcal{D}(T)$.

We can now proceed to the main result of the theory.

Theorem 8.1. [141, Theorem 24] *Assume there exists $N \in \mathbb{N}$ and possibly unbounded operators*

$$C_0, C_1, \dots, C_{N+1}, R_1, \dots, R_{N+1}, Z_1, \dots, Z_{N+1},$$

such that $C_0 = \mathcal{A}$,

$$[C_j, \mathcal{T}] = Z_{j+1}C_{j+1} + R_{j+1}, \quad (0 \leq j \leq N), \quad C_{N+1} = 0, \quad (8.8)$$

and for all $k = 0, 1, \dots, N$

- (a) $[\mathcal{A}, C_k]$ *is relatively bounded with respect to $\{C_j\}_{0 \leq j \leq k}$ and $\{C_j \mathcal{A}\}_{0 \leq j \leq k-1}$,*
- (b) $[C_k, \mathcal{A}^*]$ *is relatively bounded with respect to I and $\{C_j\}_{0 \leq j \leq k}$,*
- (c) R_k *is relatively bounded with respect to $\{C_j\}_{0 \leq j \leq k-1}$ and $\{C_j \mathcal{A}\}_{0 \leq j \leq k-1}$, and*
- (d) *there are positive constants $(\lambda_j, \Lambda_j)_{1 \leq j \leq N+1}$ such that $\lambda_j I \leq Z_j \leq \Lambda_j I$.*

Furthermore, assume that $\sum_{j=0}^N C_j^ C_j$ is κ -coercive for some $\kappa > 0$. Then, there exists $C \geq 0$ and $\lambda > 0$ such that*

$$\|S_t\|_{\mathcal{H}^1/K \rightarrow \mathcal{H}^1/K} \leq C e^{-\lambda t}, \quad (8.9)$$

where $\mathcal{H}^1 \subset \mathcal{H}$ is the subspace associated to the norm

$$\|h\|_{\mathcal{H}^1} = \sqrt{\|h\|^2 + \sum_{k=0}^N \|C_k h\|^2}. \quad (8.10)$$

Remark 8.1. *The property (8.9) is called hypocoercivity of L on $\mathcal{H}^1 := (K^\perp, \|\cdot\|_{\mathcal{H}^1})$.*

If the conditions of the above theorem hold, we also get a regularisation result for the semigroup $(S_t)_{t \geq 0}$ (see [141, Theorem A.12]):

Theorem 8.2 (Hypocoercive regularisation). *Assume the setting and notation of Theorem 8.1. Then there exists a constant $C > 0$ such that for all $k = 0, 1, \dots, N$ and $t \in (0, 1]$ the following holds:*

$$\|C_k S_t h\| \leq C \frac{\|h\|}{t^{k+\frac{1}{2}}}, \quad h \in \mathcal{H}.$$

We are now well-prepared to give the proof of Theorem 4.1:

Proof of Theorem 4.1. We prove the claim by verifying the conditions of Theorem 8.1. Recall that $C_0 = \mathcal{A} = \sigma \nabla_p$ and

$$\mathcal{T} = M^{-1} p \cdot \nabla_q - \nabla_q V \cdot \nabla_p - \mu J_1 \nabla_q V \cdot \nabla_q - \nu J_2 M^{-1} p \cdot \nabla_p.$$

A quick calculation shows that

$$\mathcal{A}^* = \sigma M^{-1} p - \sigma \nabla_p,$$

so that indeed

$$\mathcal{A}^* \mathcal{A} = \Gamma M^{-1} p \cdot \nabla_p - \Gamma : \nabla_p \nabla_p = \mathcal{L}_{\text{therm}}$$

and

$$\mathcal{A}^* \mathcal{A} - \mathcal{T} = -\mathcal{L}^*.$$

We make the choice $N = 1$ and calculate the commutator

$$[\mathcal{A}, \mathcal{T}] = \sigma M^{-1} (\nabla_q + \nu J_2 \nabla_p).$$

Let us now set $C_1 = \sigma M^{-1} \nabla_q$, $Z_1 = I$ and $R_1 = \nu \sigma M^{-1} J_2 \nabla_p$, such that (8.8) holds for $j = 0$. Note that $[\mathcal{A}, \mathcal{A}] = 0^\dagger$, $[\mathcal{A}, C_1] = 0$ and $[\mathcal{A}^*, C_1] = 0$. Furthermore, we have that

$$[\mathcal{A}, \mathcal{A}^*] = \sigma M^{-1} \sigma.$$

We now compute

$$[C_1, \mathcal{T}] = -\sigma M^{-1} (\text{Hess } V) \nabla_p + \mu \sigma M^{-1} (\text{Hess } V) J_1 \nabla_q,$$

choose $R_2 = [C_1, \mathcal{T}]$, $Z_2 = I$ and recall that $C_2 = 0$ by assumption (of Theorem 8.1). With those choices, Assumptions (a)-(d) of Theorem 8.1 are fulfilled. Indeed, Assumption (a) holds trivially since all relevant commutators are zero. Assumption (b) follows from the fact that $[\mathcal{A}, \mathcal{A}^*] = \sigma M^{-1} \sigma$ is clearly bounded relative to I . To verify Assumption (c), let us start with the case $k = 1$. It is necessary to show that $R_1 = \nu \sigma M^{-1} J_2 \nabla_p$ is bounded relatively to $\mathcal{A} = \sigma \nabla_p$ and $\mathcal{A}^2$. This is obvious since the p -derivatives appearing in R_1 can be controlled by the p -derivatives appearing in $\mathcal{A}$. For $k = 2$, a similar argument shows that $R_2 = -\sigma M^{-1} (\text{Hess } V) \nabla_p + \mu \sigma M^{-1} (\text{Hess } V) J_1 \nabla_q$ is bounded relatively to $\mathcal{A} = \sigma \nabla_p$ and $C_1 = \sigma M^{-1} \nabla_q$ because of the assumption that $\text{Hess } V$ is bounded. Note that it is crucial for the preceding arguments to assume that the matrices σ and M have full rank. Assumption (d) is trivially satisfied, since Z_1 and Z_2 are equal to the identity. It remains to show that

$$\bar{C} := \sum_{j=0}^N C_j^* C_j$$

is κ -coercive for some $\kappa > 0$. It is straightforward to see that the kernel of $\bar{C}$ consists of constant functions and therefore

$$(\ker \bar{C})^\perp = L_0^2(\widehat{\pi}).$$

[†]This is not true automatically, since $[\mathcal{A}, \mathcal{A}]$ stands for the array $([\mathcal{A}_j, \mathcal{A}_k])_{jk}$.

Hence, κ -coercivity of $\bar{C}$ amounts to the functional inequality

$$\int_{\mathbb{R}^{2d}} (|\sigma M^{-1} \nabla_q f|^2 + |\sigma \nabla_p f|^2) d\hat{\pi} \geq \kappa \left(\int_{\mathbb{R}^{2d}} f^2 d\hat{\pi} - \left(\int_{\mathbb{R}^{2d}} f d\hat{\pi} \right)^2 \right), \quad f \in H^1(\hat{\pi}).$$

By positive definiteness of M and σ , there exist constants $c_1, c_2 > 0$ such that

$$c_1 I < \sigma \sigma^T < c_2 I, \quad c_1 I < M^{-1} \sigma \sigma^T M^{-1} < c_2 I,$$

in the sense of positive definite matrices. Therefore, the above is equivalent to

$$\int_{\mathbb{R}^{2d}} (|\nabla_q f|^2 + |\nabla_p f|^2) d\hat{\pi} \geq \tilde{\kappa} \left(\int_{\mathbb{R}^{2d}} f^2 d\hat{\pi} - \left(\int_{\mathbb{R}^{2d}} f d\hat{\pi} \right)^2 \right), \quad f \in H^1(\hat{\pi}),$$

with a possibly different constant $\tilde{\kappa}$, i.e. a Poincaré inequality for $\hat{\pi}$. Since $\hat{\pi} = \pi \otimes \mathcal{N}(\mathbf{0}, M)$, coercivity of $\bar{C}$ boils down to a Poincaré inequality for π as in Assumption 4.1. This concludes the proof of the hypocoercive decay estimate (8.9). Clearly, the abstract $\mathcal{H}^1$ -norm from (8.10) is equivalent to the Sobolev norm in $H^1(\hat{\pi})$, and therefore it follows that there exist constants $C \geq 0$ and $\lambda \geq 0$ such that

$$\|S_t f\|_{H^1(\hat{\pi})} \leq C e^{-\lambda t} \|f\|_{H^1(\hat{\pi})}, \quad (8.11)$$

for all $f \in H^1(\hat{\pi})/K$, where $K = \ker \bar{C}$ consists of constant functions. Let us now lift this estimate to $L^2(\hat{\pi})$. There exists a constant $\tilde{C} \geq 0$ such that

$$\|h\|_{H^1(\hat{\pi})} \leq \tilde{C} \sum_{k=0}^2 \|C_k h\|_{L^2(\hat{\pi})}, \quad f \in H^1(\hat{\pi}). \quad (8.12)$$

Therefore, Theorem 8.2 implies

$$\|S_1 f\|_{H^1(\hat{\pi})} \leq \tilde{C} \|f\|_{L^2(\hat{\pi})}, \quad f \in L^2(\hat{\pi}), \quad (8.13)$$

for $t = 1$ and a possibly different constant $\tilde{C}$. Let us now assume that $t \geq 1$ and $f \in L^2(\hat{\pi})/K$. It holds that

$$\|S_t f\|_{L^2(\hat{\pi})} \leq \|S_t f\|_{H^1(\hat{\pi})} = \|S_{t-1} S_1 f\|_{H^1(\hat{\pi})} \leq C e^{-\lambda(t-1)} \|S_1 f\|_{H^1(\hat{\pi})}, \quad (8.14)$$

where the last inequality follows from (8.11). Now applying (8.13) and gathering constants results in

$$\|S_t f\|_{L^2(\hat{\pi})} \leq C e^{-\lambda t} \|f\|_{L^2(\hat{\pi})}, \quad f \in L^2(\hat{\pi})/K. \quad (8.15)$$

Note that although we assumed $t \geq 1$, the above estimate also holds for $t \geq 0$ (although possibly with a different constant C) since $\|S_t\|_{L^2(\hat{\pi}) \rightarrow L^2(\hat{\pi})}$ is bounded on $[0, 1]$. $\square$

8.4 THE ASYMPTOTIC VARIANCE FOR SMALL PERTURBATIONS IN THE LINEAR CASE

Our objective in this section is to give the proofs for the results discussed in Section 5.1. Recall the notations from (5.8), (5.9) and (5.10), in particular the definitions of the matrices B and Q . Moreover, let us introduce the notation

$$\bar{K} = \begin{pmatrix} K & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{pmatrix} \in \mathbb{R}^{2d \times 2d}, \quad \bar{l} = \begin{pmatrix} l \\ \mathbf{0} \end{pmatrix} \in \mathbb{R}^{2d},$$

such that the observable (5.2) reads

$$f(x) = x \cdot \bar{K}x + \bar{l} \cdot x - \text{Tr } \bar{K},$$

where, without loss of generality, we have chosen $C = -\text{Tr } \bar{K}$ in order to have $\hat{\pi}(f) = 0$ (recall from Section 1.2.2 that the asymptotic variance does not depend on additive constants). The following calculations proceed very much along the lines of [41, Section 4]. Since the target measure π is Gaussian (and so in particular, the Hessian of V is bounded), Assumption 4.1 is satisfied and exponential decay of the semigroup $(S_t)_{t \geq 0}$ as in (1.46) follows by Theorem 4.1. According to Theorem 2.2, the central limit theorem (1.52) holds and the asymptotic variance is given by

$$\sigma_f^2 = \langle \phi, f \rangle_{L^2(\hat{\pi})}, \quad (8.16)$$

where ϕ is the solution to the Poisson equation

$$-\mathcal{L}\phi = f, \quad \hat{\pi}(\phi) = 0. \quad (8.17)$$

In the sequel we will solve (8.17) analytically. By the uniqueness of the solution to the Poisson equation (8.17) and linearity of $\mathcal{L}$, ϕ has to be a quadratic polynomial, so we can write

$$\phi(x) = x \cdot Cx + D \cdot x - \text{Tr } C,$$

where $C \in \mathbb{R}_{\text{sym}}^{2d \times 2d}$ and $D \in \mathbb{R}^{2d}$ (notice that C can be chosen to be symmetrical since $x \cdot Cx$ does not depend on the skew-symmetric part of C). Plugging this ansatz into (8.17) yields

$$-\mathcal{L}\phi(x) = x \cdot A(2Cx + D) - \gamma \text{Tr}_p C = x \cdot \bar{K}x + \bar{l} \cdot x - \text{Tr } \bar{K},$$

where

$$\text{Tr}_p C = \sum_{i=n+1}^{2n} C_{ii}$$

denotes the trace of the momentum component of C , and we have introduced the matrix $A = B^T$. Comparing different powers of x , this leads to the conditions

$$AC + CA^T = \bar{K}, \quad (8.18a)$$

$$AD = \bar{l}, \quad (8.18b)$$

$$\gamma \text{Tr}_p C = \text{Tr } \bar{K}. \quad (8.18c)$$

Note that (8.18c) will eventually be satisfied by existence and uniqueness of the solution to (8.17). Then, by the calculations in [41, Section 4], the asymptotic variance is given by

$$\sigma_f^2 = 2 \text{Tr}(C\bar{K}) + D \cdot \bar{l}. \quad (8.19)$$

For the calculations in the proof of Proposition 5.1, the following lemma will come in handy:

Lemma 8.2. *Consider the matrix equation*

$$UY + YU^T = W, \quad (8.20)$$

for symmetric matrices $Y \in \mathbb{R}_{\text{sym}}^{2d \times 2d}$ and $W \in \mathbb{R}_{\text{sym}}^{2d \times 2d}$ in block form

$$Y = \begin{pmatrix} Y_1 & Y_2 \\ Y_2^T & Y_3 \end{pmatrix}, \quad W = \begin{pmatrix} W_1 & W_2 \\ W_2^T & W_3 \end{pmatrix}, \quad (8.21)$$

i.e. $Y_i, W_i \in \mathbb{R}^{d \times d}$, $Y_1^T = Y_1$, $Y_3^T = Y_3$, $W_1^T = W_1$ and $W_3^T = W_3$. Furthermore, let

$$U = \begin{pmatrix} \mathbf{0} & I \\ -I & \gamma I \end{pmatrix} \in \mathbb{R}^{2d \times 2d}, \quad \gamma > 0. \quad (8.22)$$

Then Y is given by

$$Y_1 = \frac{1}{2} \left(\gamma W_1 + \frac{1}{\gamma} (W_1 + W_3) - W_2 - W_2^T \right), \quad (8.23a)$$

$$Y_2 = \frac{1}{2} \left(W_1 + \frac{1}{\gamma} (W_2 - W_2^T) \right), \quad (8.23b)$$

$$Y_3 = \frac{1}{2\gamma} (W_1 + W_3). \quad (8.23c)$$

Proof. Inserting (8.21) and (8.22) into (8.20), we obtain the system of matrix equations

$$Y_2 + Y_2^T = W_1, \quad (8.24a)$$

$$-Y_1 + \gamma Y_2 + Y_3 = W_2, \quad (8.24b)$$

$$-Y_1 + \gamma Y_2^T + Y_3 = W_2^T, \quad (8.24c)$$

$$-Y_2 - Y_2^T + 2\gamma Y_3 = W_3. \quad (8.24d)$$

Note that (8.24b) and (8.24c) are equivalent by taking the transpose. Combining (8.24a) and (8.24d) yields (8.23c). By adding (8.24b) and (8.24c), and using (8.23c), we obtain (8.23a). Finally, we compute (8.23b) by inserting (8.23a) and (8.23c) into (8.24b). $\square$

Proof of Proposition 5.1. Recall that the asymptotic variance is given by (8.19), where C solves (8.18a), with $A = B^T$, and B is the drift matrix in (5.9). We will use the notation

$$C(\mu, \nu) = \begin{pmatrix} C_1(\mu, \nu) & C_2(\mu, \nu) \\ C_2^T(\mu, \nu) & C_3(\mu, \nu) \end{pmatrix},$$

and the abbreviations $C(0) := C(0, 0)$, $C^\mu(0) := \partial_\mu C|_{\mu, \nu=0}$ and $C^\nu(0) := \partial_\nu C|_{\mu, \nu=0}$, and analogously for $A(0)$, $A^\mu(0)$ and $A^\nu(0)$. Let us first determine $C(0)$, i.e. the solution to the equation

$$\begin{pmatrix} \mathbf{0} & I \\ -I & \gamma I \end{pmatrix} C(0) + C(0) \begin{pmatrix} \mathbf{0} & I \\ -I & \gamma I \end{pmatrix}^T = \begin{pmatrix} K & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{pmatrix}.$$

Employing Lemma 8.2, we see that

$$C(0) = \begin{pmatrix} \frac{1}{2\gamma}K + \frac{\gamma}{2}K & \frac{1}{2}K \\ \frac{1}{2}K & \frac{1}{2\gamma}K \end{pmatrix}. \quad (8.25)$$

Taking the μ -derivative of (8.18a) and setting $\mu = \nu = 0$ yields

$$A^\mu(0)C(0) + A(0)C^\mu(0) + C^\mu(0)A(0)^T + C(0)A^\mu(0)^T = \mathbf{0}. \quad (8.26)$$

Notice that

$$\begin{aligned} A^\mu(0)C(0) + C(0)A^\mu(0)^T &= \begin{pmatrix} -J & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{pmatrix} C(0) + C(0) \begin{pmatrix} J & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{pmatrix} \\ &= \begin{pmatrix} (\frac{1}{2\gamma} + \frac{\gamma}{2})[K, J] & -\frac{1}{2}JK \\ \frac{1}{2}KJ & \mathbf{0} \end{pmatrix}. \end{aligned}$$

Again using Lemma 8.2, the solution to equation (8.26) is seen to be

$$C^\mu(0) = \begin{pmatrix} -(\frac{\gamma^2}{4} + \frac{1}{4\gamma^2} + \frac{1}{4})[K, J] & \frac{1}{2\gamma}JK - \frac{\gamma}{4}[K, J] \\ -\frac{1}{2\gamma}KJ - \frac{\gamma}{4}[K, J] & -(\frac{1}{4\gamma^2} + \frac{1}{4})[K, J] \end{pmatrix}. \quad (8.28)$$

We employ a similar strategy to determine $C^\nu(0)$: Taking the ν -derivative in equation (8.18a), setting $\mu = \nu = 0$ and inserting $C(0)$ and $A(0)$ leads to the equation

$$\begin{pmatrix} \mathbf{0} & I \\ -I & \gamma I \end{pmatrix} C^\nu(0) + C^\nu(0) \begin{pmatrix} \mathbf{0} & I \\ -I & \gamma I \end{pmatrix} = \begin{pmatrix} \mathbf{0} & -\frac{1}{2}KJ \\ \frac{1}{2}JK & -\frac{1}{2\gamma}[K, J] \end{pmatrix},$$

which can be solved by

$$C^\nu(0) = \begin{pmatrix} (-\frac{1}{4\gamma^2} + \frac{1}{4})[K, J] & \frac{1}{\gamma}(-\frac{1}{2}KJ + \frac{1}{4}[K, J]) \\ \frac{1}{\gamma}(\frac{1}{2}KJ - \frac{1}{4}[K, J]) & -\frac{1}{4\gamma^2}[K, J] \end{pmatrix}, \quad (8.29)$$

using again Lemma 8.2. Note that $\text{Tr}(C\bar{K}) = \text{Tr}(C_1K)$, and so

$$\partial_\mu \Theta|_{\mu, \nu=0} = 2 \text{Tr}(C_1^\mu(0)K) = -(\frac{\gamma^2}{4} + \frac{1}{4\gamma^2} + \frac{1}{4}) \cdot \text{Tr}([K, J]K) = 0,$$

since clearly $\text{Tr}([K, J], K) = \text{Tr}(KJK) - \text{Tr}(JK^2) = 0$. In the same way it follows that $\partial_\nu \Theta|_{\mu, \nu=0} = 0$, proving (5.5).

Taking the second μ -derivative of (8.18a) and setting $\mu = \nu = 0$ yields

$$2A^\mu(0)C^\mu(0) + A(0)C^{\mu\mu}(0) + C^{\mu\mu}(0)A(0)^T + 2C^\mu(0)A^\mu(0)^T = \mathbf{0}, \quad (8.30)$$

employing the notation $C^{\mu\mu}(0) = \partial_\mu^2 C|_{\mu, \nu=0}$ and noticing that $\partial_\mu^2 A = 0$. Using (8.28) we calculate

$$A^\mu(0)C^\mu(0) + C^\mu(0)A^\mu(0)^T = \begin{pmatrix} (\frac{\gamma^2}{4} + \frac{1}{4\gamma^2} + \frac{1}{4})[J, [K, J]] & -\frac{1}{2\gamma}J^2K + \frac{\gamma}{4}J[K, J] \\ -\frac{1}{2\gamma}KJ^2 - \frac{\gamma}{4}[K, J]J & \mathbf{0} \end{pmatrix}.$$

As before, we make the ansatz

$$C^{\mu\mu}(0) = \begin{pmatrix} C_1^{\mu\mu}(0) & C_2^{\mu\mu}(0) \\ (C_2^{\mu\mu}(0))^T & C_3^{\mu\mu}(0) \end{pmatrix},$$

and apply Lemma 8.2 to obtain

$$C_1^{\mu\mu}(0) = \left(\frac{\gamma}{4} + \frac{1}{4\gamma^3} + \frac{\gamma^3}{4} \right) (2JKJ - J^2K - KJ^2) - \frac{1}{\gamma}JKJ.$$

Now

$$\partial_\mu^2 \Theta|_{\mu, \nu=0} = 2 \operatorname{Tr}(C_1^{\mu\mu}(0)K) = - \left(\gamma + \frac{1}{\gamma^3} + \gamma^3 \right) \left(\operatorname{Tr}(JKJK) - \operatorname{Tr}(J^2K^2) \right) - \frac{2}{\gamma} \operatorname{Tr}(JKJK)$$

gives the first entry of the Hessian appearing in (5.6). We proceed in the same way to determine $C_1^{\nu\nu}(0)$. Analogously, we get

$$A^\nu(0)C^\nu(0) + C^\nu(0)A^\nu(0)^T = \begin{pmatrix} \mathbf{0} & \frac{1}{\gamma}(KJ^2 - \frac{1}{2}[K, J]J) \\ \frac{1}{\gamma}(JKJ - \frac{1}{2}J[K, J]) & \frac{1}{2\gamma^2}([K, J]J - J[K, J]) \end{pmatrix}.$$

Solving the resulting matrix equation (similar to (8.30)) results in

$$C_1^{\nu\nu}(0) = \left(\frac{1}{4\gamma^3} - \frac{1}{4\gamma} \right) (KJ^2 + J^2K) - \left(\frac{1}{2\gamma^3} + \frac{1}{2\gamma} \right) JKJ,$$

leading to

$$\partial_\nu^2 \Theta|_{\mu, \nu=0} = 2 \operatorname{Tr}(C_1^{\nu\nu}(0)K) = \left(\frac{1}{\gamma^3} - \frac{1}{2\gamma} \right) \operatorname{Tr}(J^2K^2) - \left(\frac{1}{2\gamma^3} + \frac{1}{2\gamma} \right) \operatorname{Tr}(JKJK),$$

in accordance with the corresponding term in (5.6). To compute the cross term $C_1^{\mu\nu}(0)$, we take the mixed derivative $\partial_{\mu\nu}^2$ of (8.18a) and set $\mu = \nu = 0$ to arrive at

$$A^\mu(0)C^\nu(0) + A^\nu(0)C^\mu(0) + A(0)C^{\mu\nu}(0) + C^{\mu\nu}(0)A(0)^T + C^\mu(0)A^\nu(0)^T + C^\nu(0)A^\mu(0)^T = \mathbf{0}.$$

Using (8.28) and (8.29) we see that

$$\begin{aligned} & A^\mu(0)C^\nu(0) + A^\nu(0)C^\mu(0) + C^\mu(0)A^\nu(0)^T + C^\nu(0)A^\mu(0)^T \\ &= \begin{pmatrix} \left(\frac{1}{4\gamma^2} - \frac{1}{4} \right) [J, [K, J]] & \frac{1}{\gamma} JKJ - \frac{1}{4\gamma} J[K, J] - \frac{\gamma}{4} [K, J]J \\ \frac{1}{2\gamma} JKJ + \frac{1}{2\gamma} KJ^2 + \frac{\gamma}{4} J[K, J] - \frac{1}{4\gamma} [K, J]J & \left(\frac{1}{4\gamma^2} + \frac{1}{4} \right) [J, [K, J]] \end{pmatrix}. \end{aligned}$$

Lemma 8.2 yields the solution

$$C_1^{\mu\nu}(0) = \left(-\frac{1}{4\gamma^3} + \frac{\gamma}{4} - \frac{1}{4\gamma} \right) [J, [K, J]] + \frac{1}{\gamma} JKJ,$$

leading to

$$\partial_{\mu\nu}^2 \Theta|_{\mu, \nu=0} = 2 \operatorname{Tr}(C_1^{\mu\nu}(0)K) = \left(\frac{1}{\gamma^3} + \frac{1}{\gamma} - \gamma \right) \operatorname{Tr}(J^2K^2) + \left(-\frac{1}{\gamma^3} + \frac{1}{\gamma} + \gamma \right) \operatorname{Tr}(JKJK).$$

This completes the proof. $\square$

Proof of Proposition 5.2. By (8.18b) and (8.19) the function Θ satisfies

$$\Theta(\mu, \nu) = \bar{l} \cdot A^{-1} \bar{l}.$$

Recall the following formula for the blockwise inversion of matrices using the Schur complement:

$$\begin{pmatrix} U & V \\ W & X \end{pmatrix}^{-1} = \begin{pmatrix} (U - VX^{-1}W)^{-1} & \dots \\ \dots & \dots \end{pmatrix}, \quad (8.31)$$

provided that X and $U - VX^{-1}W$ are invertible. Using this, we obtain

$$\Theta(\mu, \nu) = l \cdot (-\mu J + (\gamma I - \nu J)^{-1})^{-1} l. \quad (8.32)$$

We remark that $\ker(\gamma I - \nu J) = \{0\}$, and so the matrix $\gamma I - \nu J$ is invertible: Indeed, taking the scalar product of the equation $(\gamma I - \nu J)x = 0$ with x shows that necessarily $x = 0$. Taking derivatives in (8.32), setting $\mu = \nu = 0$, and using the fact that $J^T = -J$ leads to the desired result. $\square$

The following lemma is used in the calculation leading up to Corollary 5.1.

Lemma 8.3. *The following holds:*

- (a) $\gamma - \frac{4}{\gamma^3} - \gamma^3 - \frac{1}{\gamma} < 0$ for $\gamma \in (0, \infty)$.
- (b) Let $J, K \in \mathbb{R}^{d \times d}$ with $J = -J^T$ and $K = K^T$. Then $\text{Tr}(JKJK) - \text{Tr}(J^2K^2) \geq 0$. Furthermore, equality holds if and only if $[J, K] = 0$.

Proof. To show (a), we note that $\gamma - \frac{4}{\gamma^3} - \gamma^3 - \frac{1}{\gamma} < \gamma - \frac{4}{\gamma^3} - \gamma^3 = \gamma(1 - \frac{4}{\gamma^4} - \gamma^2)$. The function $f(\gamma) := 1 - \frac{4}{\gamma^4} - \gamma^2$ has a unique global maximum on $(0, \infty)$ at $\gamma_{\min} = 8^{1/6}$ with $f(\gamma_{\min}) = -2$, so the result follows.

For (b), we note that $[J, K]^T = [J, K]$, and that therefore $[J, K]^2$ is symmetric and non-negative definite. It immediately follows that

$$\text{Tr}([J, K]^2) \geq 0,$$

with equality if and only if $[J, K] = 0$. Now expand

$$\text{Tr}([J, K]^2) = 2(\text{Tr}(JKJK) - \text{Tr}(J^2K^2)),$$

which implies the advertised claim. $\square$

8.5 ORTHOGONAL TRANSFORMATION OF TRACEFREE SYMMETRIC MATRICES INTO A MATRIX WITH ZEROS ON THE DIAGONAL

Given a symmetric matrix $K \in \mathbb{R}_{\text{sym}}^{d \times d}$ with $\text{Tr } K = 0$, we seek to find an orthogonal matrix $U \in O(\mathbb{R}^d)$ such that UKU^T has zeros on the diagonal. This is a crucial step in Algorithms

1 and 2 and has been addressed in various places in the literature (see for instance [81] or [18, Section 2.2]). For the convenience of the reader, in the following we summarise an algorithm very similar to the one in [81].

Since K is symmetric, there exists an orthogonal matrix $U_0 \in O(\mathbb{R}^d)$ such that $U_0 K U_0^T = \text{diag}(\lambda_1, \dots, \lambda_d)$. Now the algorithm proceeds iteratively, orthogonally transforming this matrix into one with the first diagonal entry vanishing, then the first two diagonal entries vanishing, etc, until after d steps we are left with a matrix with zeros on the diagonal. Starting with λ_1 , assume that $\lambda_1 \neq 0$ (otherwise proceed with λ_2). Since $\text{Tr}(K) = \text{Tr}(U_0 K U_0^T) = \sum_i \lambda_i = 0$, there exists λ_j , $j \in \{2, \dots, d\}$ such that $\lambda_1 \lambda_j < 0$ (i.e. λ_1 and λ_j have opposing signs). We now apply a rotation in the $1j$ -plane to transform the first diagonal entry into zero. More specifically, let

$$U_1 = \begin{pmatrix} \cos \alpha & 0 & \dots & 0 & -\sin \alpha & 0 & \dots & 0 \\ 0 & 1 & & & 0 & & & \vdots \\ \vdots & & \ddots & & & & & \\ 0 & & & 1 & 0 & & & \\ \sin \alpha & 0 & & 0 & \cos \alpha & 0 & & \\ 0 & & & & 0 & 1 & & \vdots \\ \vdots & & & & & & \ddots & 0 \\ 0 & 0 & \dots & & \dots & 0 & & 1 \end{pmatrix} \in O(\mathbb{R}^d)$$

with $\alpha = \arctan \sqrt{-\frac{\lambda_1}{\lambda_j}}$. We then have $(U_1 U_0 K U_0^T U_1^T)_{11} = 0$. Now the same procedure can be applied to the second diagonal entry λ_2 , leading to the matrix $U_2 U_1 U_0 K U_0^T U_1^T U_2^T$ with

$$(U_2 U_1 U_0 K U_0^T U_1^T U_2^T)_{11} = (U_2 U_1 U_0 K U_0^T U_1^T U_2^T)_{22} = 0.$$

Iterating this process, we obtain that $U_d \dots U_1 U_0 K U_0^T U_1^T \dots U_d^T$ has zeros on the diagonal, so $U_d \dots U_1 U_0 \in O(\mathbb{R}^d)$ is the required orthogonal transformation.

Part III

Couplings

9. INTRODUCTION TO ENSEMBLE BASED APPROACHES

In this part of the thesis, our aim is to construct and analyse MCMC samplers based on a collection of interacting particles. Before describing our approach in greater detail, let us take the opportunity to place our work in a broader context.

The task of approximating (high-dimensional) probability distributions is often approached by considering empirical measures associated to an ensemble of n particles, i.e.

$$\pi \approx \frac{1}{n} \sum_{i=1}^n \delta_{X^{(i)}} =: \tilde{\pi}, \quad (9.1)$$

where $X^{(i)}$ stands for the location of the i th particle and δ_x denotes the Dirac measure centred at x . Usually, the particles are moved according to some (more often than not stochastic) dynamics, judiciously crafted in order for the empirical measure $\tilde{\pi}_t := \frac{1}{n} \sum_{i=1}^n \delta_{X_t^{(i)}}$ to approach π when t reaches a terminal value (finite or infinite) or the number of particles tends to infinity (or both). This general idea and various modifications thereof have been particularly influential in the statistical inference of hidden-state Markov models (stochastic filtering or sequential Monte Carlo techniques, see for instance [38, 122] and references therein). Ensemble based methods have also been employed in the contexts of optimisation [119, 123], molecular dynamics [131], Markov chain Monte Carlo [88, 106], or variational Bayesian inference [95]. Let us also mention the works [2] and [68], combining different aspects of various sampling strategies. The increasing availability of parallel-processing computational architectures has further encouraged the development and analysis of similar methodologies.

From a mathematical standpoint, many of the aforementioned algorithms targeting a probability measure π on some state space E naturally produce probability measures $\bar{\pi}$ on the product space $\bar{E} = \prod_{i=1}^n E_i$, where E_i is an identical copy of E , standing for the state space of the i th particle. Denoting by $P_i : \mathcal{P}(\bar{E}) \rightarrow \mathcal{P}(E_i)$ the mappings that send probability measures on $\bar{E}$ to their marginals on E_i , one then obtains the measure

$$\frac{1}{n} \sum_{i=1}^n P_i(\bar{\pi}) \quad (9.2)$$

as an approximation for π . Clearly, the map

$$\Pi : \mathcal{P}(\bar{E}) \rightarrow \mathcal{P}(E), \quad \bar{\pi} \mapsto \frac{1}{n} \sum_{i=1}^n P_i(\bar{\pi}) \quad (9.3)$$

is far from injective since $\Pi(\bar{\pi})$ only depends on the marginals of $\bar{\pi}$. This viewpoint shows that there is a considerable flexibility when generating the joint measure $\bar{\pi}$, immediately suggesting fruitful connections to the theory of couplings of probability measures [93, 137] prominently encountered for instance in relation to optimal transport problems [140, 142] or decay estimates in Wasserstein distances (see Chapter 3). Since in applications $\Pi(\bar{\pi})$ is only an approximation of the target measure of interest, the freedom to design appropriate couplings can be used to suppress bias, variance and discretisation errors. This general idea has proved to be very versatile, leading to powerful simulation techniques such as multilevel Monte Carlo [54], coupling from the past [121] and antithetic variates [83, Section 9.2].

Let us now more specifically describe our approach in this part of the thesis, applying the general ideas outlined above for the construction of coupled MCMC samplers (recall that a particular example was already presented in Section 1.3.3). To this end, suppose that we are presented with a Markov process $(X_t)_{t \geq 0}$ on an appropriate state space E , ergodic with respect to a measure $\pi \in \mathcal{P}(E)$, i.e.

$$\lim_{T \rightarrow \infty} \frac{1}{T} \int_0^T f(X_t) dt = \int_E f d\pi, \quad f \in L^1(\pi). \quad (9.4)$$

Next, let us consider n identical copies of E , $(X_t)_{t \geq 0}$ and π , denoted by E_i , $(X_t^i)_{t \geq 0}$ and π_i , for $i \in \{1, \dots, n\}$. Our main object of study is the class of Markovian couplings $(\bar{X}_t)_{t \geq 0}$ of $\{(X_t^i)_{t \geq 0} : i \in \{1, \dots, n\}\}$ on the product space $\bar{E} = \prod_{i=1}^n E_i$ that obey certain mild regularity assumptions. In particular, we characterise those couplings in terms of their infinitesimal generators in Chapter 10. For an observable $f \in L^1(\pi)$, we can then define the extended observable $F : \bar{E} \rightarrow \mathbb{R}$ by

$$F(x_1, \dots, x_n) = \frac{1}{n} \sum_{i=1}^n f(x_i). \quad (9.5)$$

From (9.4) it is then immediate that

$$\lim_{T \rightarrow \infty} \frac{1}{T} \int_0^T F(\bar{X}_t) dt = \int_E f d\pi, \quad (9.6)$$

i.e. the coupled process $(\bar{X}_t)_{t \geq 0}$ in conjunction with the observable (9.5) provides a valid sampling scheme. Let us remark that the framework we develop starting in Chapter 10 accommodates the case when the spaces E_i , the processes $(X_t^i)_{t \geq 0}$ and the measures π_i are not identical, allowing for almost arbitrary combinations of the examples encountered

in Section 1.3.

The study of observables of the form (9.5) provides a compelling dual perspective on the ‘sum of marginals’ operator (9.3). Recall that we denote by $\mathcal{B}_b(E)$ the space of bounded measurable functions and consider the ‘extension operator’

$$\Pi^* : \mathcal{B}_b(E) \rightarrow \mathcal{B}_b(\bar{E}), \quad f \mapsto \frac{1}{n} \sum_{i=1}^n f_i, \quad (9.7)$$

provided by (9.5). For $\bar{\pi} \in \mathcal{P}(\bar{E})$ and $f \in \mathcal{B}_b(E)$ we clearly have $(\Pi\bar{\pi})(f) = \bar{\pi}(\Pi^*f)$, showing that understanding the class of observables given by (9.5) is sufficient for analysing the properties of measures of the form (9.2). This idea features in particular in Chapter 14 in the analysis of the exponential convergence to equilibrium for coupled processes.

OVERVIEW OF THE MAIN RESULTS BY MEANS OF A SIMPLE EXAMPLE

In this section we present our main findings informally by means of a very simple example, pointing to the exact statements in the forthcoming chapters. Let us stress that our results hold in much greater generality, in particular also including the recently fashionable piecewise deterministic Markov processes (PDMPs) introduced in Section 1.3.2.

Let us consider $n = 2$ particles (the locations of which are denoted by X and Y) moving in one dimension respectively according to the overdamped Langevin dynamics (1.72). In other words, $(X_t)_{t \geq 0}$ and $(Y_t)_{t \geq 0}$ satisfy the SDEs

$$dX_t = -V'(X_t) dt + \sqrt{2} dB_t^x, \quad (9.8a)$$

$$dY_t = -V'(Y_t) dt + \sqrt{2} dB_t^y, \quad (9.8b)$$

where we recall that the potential $V \in C^\infty(\mathbb{R})$ is such that

$$Z := \int_{\mathbb{R}} e^{-V} dx < \infty, \quad (9.9)$$

and $(B_t^x)_{t \geq 0}, (B_t^y)_{t \geq 0}$ denote standard one-dimensional Brownian motions. Still following Chapter 1, each of these processes considered separately is ergodic with respect to $\pi = \frac{1}{Z} e^{-V} dx$, i.e. (9.4) holds. Note that we have deliberately refrained from stating that the Brownian motions $(B_t^x)_{t \geq 0}$ and $(B_t^y)_{t \geq 0}$ are independent. Indeed, notwithstanding any dependence between these, it is immediate that (9.6) holds for the extended observable $F(x, y) = \frac{1}{2}(f(x) + f(y))$, as defined in (9.5). One of the main objectives of our analysis is to find couplings between $(B_t^x)_{t \geq 0}$ and $(B_t^y)_{t \geq 0}$ such that the induced joint process $(X_t, Y_t)_{t \geq 0}$ has favourable properties, in terms of the asymptotic variance associated to (9.6) as well as in terms of convergence to equilibrium of $\Pi(\bar{\pi}_t)$, where $\bar{\pi}_t$ denotes the joint law of (X_t, Y_t) and Π has been defined in (9.3). The dependence between $(B_t^x)_{t \geq 0}$ and

$(B_t^y)_{t \geq 0}$ can be conveniently encoded in a suitable matrix-valued function $G : \mathbb{R}^2 \rightarrow \mathbb{R}^{2 \times 2}$, writing

$$d \begin{pmatrix} X_t \\ Y_t \end{pmatrix} = \begin{pmatrix} -V'(X_t) \\ -V'(Y_t) \end{pmatrix} dt + \sqrt{2} G(X_t, Y_t) \begin{pmatrix} dW_t^x \\ dW_t^y \end{pmatrix}, \quad (9.10)$$

for two independent Brownian motions $(W_t^x)_{t \geq 0}$ and $(W_t^y)_{t \geq 0}$. In this sense, the optimisation problem alluded to above is naturally posed over an appropriate set of matrix-valued functions.

In Chapter 10 we introduce the general framework, leading to a characterisation of possible couplings in terms of infinitesimal generators of the dynamics. In the present example, recall that the generators of the one-particle dynamics are given by

$$\mathcal{L}_x = -V'(x)\partial_x + \partial_x^2, \quad \mathcal{L}_y = -V'(y)\partial_y + \partial_y^2. \quad (9.11)$$

The generators of possible couplings $(X_t, Y_t)_{t \geq 0}$ turn out to be of the form

$$\bar{\mathcal{L}}_\Gamma := \mathcal{L}_x + \mathcal{L}_y + \Gamma, \quad \Gamma = 2\alpha \partial_x \partial_y, \quad (9.12)$$

where $\alpha : \mathbb{R}^2 \rightarrow [-1, 1]$ is a function with suitable regularity properties. The connection between α and G will be made precise in Chapter 11. We use the term ‘coupling operator’ when referring to Γ and denote the set of such operators by $\mathcal{G}$. Note that Γ as defined in (9.12) vanishes on functions that depend either only on x or only on y . In Proposition 10.1 we will see that this property essentially characterises coupling operators in general. As it turns out (see the discussion in Section 10.3), not every coupling of ergodic Markov processes is such that the joint process is ergodic. Hence, we introduce the subset $\mathcal{G}^0 \subset \mathcal{G}$ of ergodic coupling operators that do preserve ergodicity. In the present example, $\bar{\mathcal{L}}_\Gamma$ is elliptic whenever $-1 < \alpha < 1$ pointwise and therefore the corresponding coupling operators are ergodic according to Theorem 1.6. Intuitively, the nonergodic coupling operators in $\mathcal{G} \setminus \mathcal{G}^0$ can hence be thought of as lying ‘at the boundary’ of $\mathcal{G}^0$.

On $\mathcal{G}^0$ we can consider the map $\Gamma \mapsto \bar{\pi}_\Gamma$, where $\bar{\pi}_\Gamma$ stands for the unique invariant measure associated with $\bar{\mathcal{L}}_\Gamma$. It is immediately clear from the construction that any $\bar{\pi}_\Gamma$ arising in this way is a coupling of π to itself (i.e. $\bar{\pi}_\Gamma$ has marginal π in both directions). We argue in Chapter 12 that a wide range of optimisation problems in our context can be cast in the following form, very closely linked to the theory of optimal transportation,

$$\min_{\Gamma \in \mathcal{G}^0} \int_{\bar{E}} c d\bar{\pi}_\Gamma, \quad (9.13)$$

where c is an appropriate cost function. Indeed, we show in Section 12.1 that the task of optimising the asymptotic variance of a coupled process with respect to a given observable is equivalent to (9.13), for a cost function that is constructed from the solution of a related Poisson equation. Addressing the problem (9.13), we first note that the dependence

$\Gamma \mapsto \int_{\bar{E}} c d\bar{\pi}_{\Gamma}$ is highly nonlinear, in particular, for $\Gamma \in \mathcal{G}^0$ and $\lambda \in [0, 1]$, the mapping $\lambda \mapsto \int_{\bar{E}} c d\bar{\pi}_{\lambda\Gamma}$ generally exhibits many local minima and maxima*. Nevertheless, we find that under suitable conditions the function $\Gamma \mapsto \int_{\bar{E}} c d\bar{\pi}_{\Gamma}$ does not attain its extrema on interior points. This is the main result of Section 12.2 and is stated rigorously in Theorem 12.3. In the example under consideration, this implies that optimal couplings necessarily satisfy $\|\alpha\|_{\infty} = 1$, leading to singular (i.e. degenerately elliptic) generators $\tilde{\mathcal{L}}_{\Gamma}$. This conclusion is interesting in two respects: Firstly, it complements standard results from optimal transport theory showing that optimal couplings are typically singular in a certain sense. We stress, however, that the problem (9.13) is genuinely different from problems occurring in optimal transport theory, and that our proof uses fundamentally different techniques. Secondly, this result supports the folklore that optimal MCMC samplers use the least amount of noise necessary to guarantee their ergodicity.

While the results from Chapter 12 indicate the possible locations of optimal coupling operators Γ in the set $\mathcal{G}$, they do not help to actually find them in practice. In Chapter 13 we address this problem by considering small perturbations around the trivial coupling $\mathbf{0} \in \mathcal{G}$ corresponding to independent Brownian motions. This leads to a much more tractable optimisation problem that can be solved explicitly in concrete examples and gives promising results in our numerical experiments. In the present example, ‘mirror coupling’ ($B_t^x = -B_t^y$) turns out to be optimal in terms of reducing the asymptotic variance of monotone observables, in the sense of the optimisation problem just referred to. However, for different observables (perhaps exhibiting other types of symmetries) more intricate coupling strategies turn out to be advisable.

In Chapter 14 we analyse the rate of convergence to equilibrium for coupled processes. As we will see, the former can be characterised in terms of an inequality of Poincaré type that is in turn related to an appropriate Hilbert space constructed in terms of the coupling. Applied to the present example, this result shows that the rate of convergence can be improved if the potential V is symmetric, i.e. $V(x) = V(-x)$. In general, the speed of convergence to equilibrium can also be slower, in the sense that the constant C in (1.46) might be greater than one. We leave an exploration of this phenomenon for future study.

* This claim is made assuming that $\lambda\Gamma \in \mathcal{G}^0$ for all $\lambda \in [0, 1]$.

10. COUPLED PROCESSES AND COUPLING OPERATORS

This chapter is devoted to the interplay between couplings of Markov processes and their infinitesimal generators. We start by specifying the setting and notation.

10.1 PRELIMINARIES, NOTATION AND SETTING

In this section we extend on some of the background material from Section 1.1, and in particular give a more thorough treatment of Feller processes. The Feller framework is appropriate since it allows us to proceed in the analysis without fixing an invariant measure (or other reference measure) as we did in Part II. For more details concerning the mathematical setting we refer to [26, Chapter 1], [75, Chapter 3 and 4], [76, Chapter 4], and [79, Chapter 17].

FELLER SEMIGROUPS

Recall that for a given locally compact Polish space E we denote the space of bounded, Borel measurable functions by $\mathcal{B}_b(E)$ and the space of bounded continuous functions by $C_b(E)$. We will furthermore need the space of functions vanishing at infinity:

Definition 10.1. *A function $f \in C_b(E)$ vanishes at infinity if for all $\varepsilon > 0$ there exists a compact set $K \subset E$ such that $|f(x)| \leq \varepsilon$ for all $x \in E \setminus K$. The space of functions vanishing at infinity is denoted by $C_0(E)$.*

The spaces $\mathcal{B}_b(E)$, $C_b(E)$ and $C_0(E)$ become Banach spaces when equipped with the supremum norm, denoted by $\|\cdot\|_\infty$. Let $(X_t)_{t \geq 0}$ be a Markov process on E with semigroup $(S_t)_{t \geq 0}$ as defined in (1.9). As already remarked in Section 1.1, $(S_t)_{t \geq 0}$ is usually not strongly continuous on $\mathcal{B}_b(E)$ (recall Definition 1.4). If the Feller property is satisfied it allows us to recover strong continuity on the smaller space $C_0(E)$:

Definition 10.2 (Feller process). *The Markov process $(X_t)_{t \geq 0}$ is a Feller process, if the corresponding semigroup $(S_t)_{t \geq 0}$ satisfies the following properties:*

1. $(S_t)_{t \geq 0}$ leaves $C_0(E)$ invariant, i.e. $S_t f \in C_0(E)$ for all $f \in C_0(E)$ and $t \geq 0$.

2. $(S_t)_{t \geq 0}$ is strongly continuous on $C_0(E)$, i.e.

$$\|S_t f - f\|_\infty \xrightarrow{t \rightarrow 0} 0 \quad (10.1)$$

for all $f \in C_0(E)$.

In the following, we will denote the semigroup $(S_t)_{t \geq 0}$ restricted to $C_0(E)$ by the same letter and call it a *Feller semigroup*. To put the foregoing definition into context let us cite the following result:

Lemma 10.1. [26, Lemma 1.4] *If $(S_t)_{t \geq 0}$ satisfies the first condition of Definition 10.2, then each of the following are equivalent to strong continuity on $C_0(E)$:*

1. For every $f \in C_0(E)$ and $x \in E$:

$$\lim_{t \rightarrow 0} S_t f(x) = f(x).$$

2. For every $f \in C_0(E)$ the mapping

$$\begin{aligned} [0, \infty) \times E &\rightarrow \mathbb{R} \\ (t, x) &\mapsto S_t f(x) \end{aligned}$$

is continuous.

3. The mapping

$$\begin{aligned} [0, \infty) \times E \times C_0(E) &\rightarrow \mathbb{R} \\ (t, x, f) &\mapsto S_t f(x) \end{aligned}$$

is continuous.

Given a Feller semigroup, we can straightforwardly define its generator $(\mathcal{L}, \mathcal{D}(\mathcal{L}))$ as in (1.21), noting that due to strong continuity, $\mathcal{L}$ is closed and densely defined in $C_0(E)$. Throughout this thesis, we will always assume that E is a manifold, not necessarily connected, such that the space of smooth compactly supported functions $C_c^\infty(E)$ is meaningfully defined. The state spaces of all Markov processes considered in Section 1.3 naturally fulfil this requirement. Let us remark, however, that our framework can be extended to more general spaces (in particular of infinite dimensionality), replacing $C_c^\infty(E)$ by suitable function spaces adapted to the particular setting.

We will restrict our attention to the class of rich Feller processes:

Assumption 10.1. *All considered Feller processes are rich*, i.e. $C_c^\infty(E) \subset \mathcal{D}(\mathcal{L})$.*

*We adopt this terminology following for instance [84, Section 1.5] and references therein.

Clearly, $C_0(E)$ does not contain constant functions (apart from the zero function). In preparation for Section 10.2, we therefore need to extend the domain $\mathcal{D}(\mathcal{L})$ to a suitable (dense) subset of $C_b(E)$. Following [134] (see also [76, Section 4.8]), we start with the following definition.

Definition 10.3 (Locally uniform convergence). *Let $(f_n)_{n \in \mathbb{N}}$ be a sequence in $C_b(E)$. Then f_n is said to converge locally uniformly to $f \in C_b(E)$ if for every compact set $K \subset E$, the sequence $f_n|_K$ converges uniformly to $f|_K$, i.e.*

$$\sup_{x \in K} |f_n(x) - f(x)| \rightarrow 0. \quad (10.4)$$

We can now equip $C_b(E)$ with the topology of locally uniform convergence (note however that this does not lead to a Banach space) and define the extended generator $(\tilde{\mathcal{L}}, \mathcal{D}(\tilde{\mathcal{L}}))$ as in (1.21), i.e.

$$\tilde{\mathcal{L}}f = \lim_{t \rightarrow 0} \frac{S_t f - f}{t} \quad (10.5a)$$

$$\mathcal{D}(\tilde{\mathcal{L}}) = \left\{ f \in C_b(E) : \lim_{t \rightarrow 0} \frac{S_t f - f}{t} \text{ exists locally uniformly} \right\}. \quad (10.5b)$$

By the conservativeness of $(S_t)_{t \geq 0}$ (see (1.14)), we immediately see that $\mathbf{1} \in \mathcal{D}(\tilde{\mathcal{L}})$ and $\tilde{\mathcal{L}}\mathbf{1} = 0$, i.e. $\tilde{\mathcal{L}}$ vanishes on constant functions. Moreover, $(\tilde{\mathcal{L}}, \mathcal{D}(\tilde{\mathcal{L}}))$ is an extension of $(\mathcal{L}, \mathcal{D}(\mathcal{L}))$, i.e. $\mathcal{D}(\mathcal{L}) \subset \mathcal{D}(\tilde{\mathcal{L}})$ and $\tilde{\mathcal{L}}|_{\mathcal{D}(\mathcal{L})} = \mathcal{L}$. Hence, we will from now on drop the tilde when no confusion is possible.

Through Itô's formula, the generators of Feller processes are naturally connected to martingale problems, as can be seen from the following definition and theorem[†]:

Definition 10.4 (Martingale problem). *Let $\mathcal{L} : \mathcal{D}(\mathcal{L}) \subset \mathcal{B}_b(E) \rightarrow \mathcal{B}_b(E)$ be a linear operator. A pair $((\Omega, \mathcal{F}, \mathbb{P}), (X_t)_{t \geq 0})$, where $(\Omega, \mathcal{F}, \mathbb{P})$ is a probability space and $(X_t)_{t \geq 0}$ is a stochastic process, is a solution to the martingale problem for $(\mathcal{L}, \mathcal{D}(\mathcal{L}))$, if for all $f \in \mathcal{D}(\mathcal{L})$ the process*

$$M_t^f = f(X_t) - f(X_0) - \int_0^t (\mathcal{L}f)(X_s) ds, \quad t \geq 0, \quad (10.6)$$

is a martingale with respect to the natural filtration $(\mathcal{F}_t^X)_{t \geq 0}$ generated by $(X_t)_{t \geq 0}$ (see (1.4)).

The martingale problem is said to exhibit uniqueness, if whenever $((\Omega, \mathcal{F}, \mathbb{P}), (X_t)_{t \geq 0})$ and $((\tilde{\Omega}, \tilde{\mathcal{F}}, \tilde{\mathbb{P}}), (\tilde{X}_t)_{t \geq 0})$ are two solutions, it follows that $\mathbb{P} \circ X^{-1} = \tilde{\mathbb{P}} \circ \tilde{X}^{-1}$.

[†]In fact, martingale problems offer a powerful approach to continuous-time Markov processes, reaching far beyond those satisfying the Feller property [48]. However, without the latter, uniqueness of the solutions to martingale problems is usually quite challenging to obtain.

The martingale problem is said to be well-posed, if for all $\pi_0 \in \mathcal{P}(E)$ there exists a solution with $X_0 \sim \pi_0$, and uniqueness holds.

Theorem 10.1 (Well-posedness for Feller generators). *Let $(\mathcal{L}, \mathcal{D}(\mathcal{L}))$ be the generator of a Feller process. Then (10.6) is a martingale for all $f \in \mathcal{D}(\mathcal{L})$. Conversely, let $\mathcal{D} \subset \mathcal{D}(\mathcal{L})$ be a core (see Definition 1.5). Then the martingale problem for $(\mathcal{L}, \mathcal{D})$ is well-posed.*

The generators of Feller processes can essentially be characterised by the so-called ‘maximum principle’. We conclude our overview by stating this fact, starting with the following definition:

Definition 10.5 (Positive maximum principle). *[26, Lemma 1.28] A linear operator $(\mathcal{L}, \mathcal{D}(\mathcal{L}))$, $\mathcal{D}(\mathcal{L}) \subset \mathcal{B}_b(E)$, satisfies the positive maximum principle if*

$$f(x_0) = \sup_{x \in E} f(x) \geq 0 \quad \text{implies} \quad \mathcal{L}f(x_0) \leq 0, \quad (10.7)$$

for all $f \in \mathcal{D}(\mathcal{L})$ and $x_0 \in E$.

We now have the following characterisation:

Theorem 10.2 (Hille-Yosida-Ray). *[26, Theorem 1.30] A linear operator $(\mathcal{L}, \mathcal{D}(\mathcal{L}))$ on $C_0(E)$ is closable and its closure is the generator of a Feller semigroup if and only if the following three conditions hold:*

1. $\mathcal{D}(\mathcal{L}) \subset C_0(E)$ is dense.
2. $(\mathcal{L}, \mathcal{D}(\mathcal{L}))$ satisfies the positive maximum principle.
3. $\text{Ran}(\lambda - \mathcal{L})$ is dense in $C_0(E)$ for some $\lambda > 0$.

Remark 10.1. *Once $(\mathcal{L}, \mathcal{D}(\mathcal{L}))$ is the generator of a strongly continuous semigroup $(S_t)_{t \geq 0}$ on $C_0(E)$, there is a canonical extension to $\mathcal{B}_b(E)$. Indeed, by the Riesz representation theorem there exist unique transition functions $(\rho_t(x, \cdot))_{t \geq 0, x \in E}$ such that $(S_t)_{t \geq 0}$ has the representation (1.10) which can then be trivially extended to $\mathcal{B}_b(E)$.*

Local[‡] linear operators acting on $C_c^\infty(\mathbb{R}^d)$ and satisfying the positive maximum principle are necessarily (possibly degenerately) elliptic differential operators of second order. This can be summarised in the following form ([34, Theorem 0.1], [76, Section 4.5]):

Theorem 10.3 (Courrège). *Let $(\mathcal{L}, \mathcal{D}(\mathcal{L}))$ be the generator of a rich[§] Feller process on $E = \mathbb{R}^d$ with almost surely continuous sample paths. Then there exist functions $Q : \mathbb{R}^d \rightarrow$*

[‡]Locality of a linear operator $\mathcal{L}$ acting on $C_c^\infty(\mathbb{R}^d)$ is defined by the condition that $\text{supp } \mathcal{L}\phi \subset \text{supp } \phi$, for all $\phi \in C_c^\infty(\mathbb{R}^d)$. By Peetre’s theorem [117], $\mathcal{L}$ is then necessarily a differential operator. Furthermore, the locality property is directly related to the continuity of the paths of the associated stochastic process if $\mathcal{L}$ is a generator.

[§]See Assumption 10.1.

$\mathbb{R}_{\text{sym}}^{d \times d}$ and $a : \mathbb{R}^d \rightarrow \mathbb{R}^d$ such that

$$(\mathcal{L}f)(x) = \sum_{i,j=1}^d Q_{ij}(x) \frac{\partial^2 f(x)}{\partial x_i \partial x_j} + \sum_{i=1}^d a_i(x) \frac{\partial f(x)}{\partial x_i}, \quad f \in C_c^\infty(\mathbb{R}^d). \quad (10.8)$$

The matrix-valued function Q is pointwise positive semidefinite.

Remark 10.2. The structure of Feller processes with discontinuous paths is determined by Courrègue's theorem as well, but since we will not make use of this fact we prefer to present a simplified version. Note also that Courrègue's theorem does not guarantee that an operator of the form (10.8) can be extended to the generator of a Feller process. In particular, further regularity assumptions on the coefficients a and Q need to be made in order for this conclusion to hold.

PRODUCT SPACES

We will be dealing with a collection of locally compact Polish spaces E_i , indexed by $i \in \{1, \dots, n\}$, and denote their cartesian product by $\bar{E} := E_1 \times \dots \times E_n$. For $f \in \mathcal{B}_b(E_i)$, it is of course understood that also $f \in \mathcal{B}_b(\bar{E})$, then depending only on the coordinate x_i in $\bar{x} \equiv (x_1, \dots, x_n)$. To a given function $f \in C_c^\infty(E_i)$ or $f \in C_0(E_i)$, we will also associate the canonical element in $\mathcal{B}_b(\bar{E})$. We wish to emphasize though that clearly f does not in general have compact support or does not vanish at infinity when considered as a function on $\bar{E}$. Frequently, the spaces E_i will be identical copies of each other, i.e. $\bar{E} = E^n$. Given $f \in \mathcal{B}_b(E)$, we will then write $f_i \in \mathcal{B}_b(\bar{E})$ for the function given by

$$f_i(x_1, \dots, x_n) = f(x_i), \quad (x_1, \dots, x_n) \in \bar{E}. \quad (10.9)$$

Sums of unbounded operators are defined in the usual way: For two operators $(A, \mathcal{D}(A))$ and $(B, \mathcal{D}(B))$ defined on the same Banach space X (i.e. $\mathcal{D}(A) \subset X$ and $\mathcal{D}(B) \subset X$), their sum is defined via

$$(A + B)f := Af + Bf, \quad f \in \mathcal{D}(A + B) := \mathcal{D}(A) \cap \mathcal{D}(B), \quad (10.10)$$

see for instance [47, Chapter III]. In the case when $(A, \mathcal{D}(A))$ and $(B, \mathcal{D}(B))$ are defined on two distinct spaces $\mathcal{B}_b(E_i)$ and $\mathcal{B}_b(E_j)$, $i \neq j$, (i.e. $\mathcal{D}(A) \subset \mathcal{B}_b(E_i)$ and $\mathcal{D}(B) \subset \mathcal{B}_b(E_j)$), their sum is defined as

$$A + B := A \otimes I + I \otimes B, \quad \mathcal{D}(A + B) := \mathcal{D}(A) \widehat{\otimes} \mathcal{D}(B), \quad (10.11)$$

where $\widehat{\otimes}$ denotes the canonical topological tensor product on $\mathcal{B}_b(E_i \times E_j)$ (see [1] and [4, A-I 3.7]).

10.2 COUPLED PROCESSES AND COUPLING OPERATORS

Assume that for $i \in \{1, \dots, n\}$ we are given locally compact Polish spaces E_i , representing the state spaces of n distinct particles. Furthermore, for $i \in \{1, \dots, n\}$, let us fix Feller semigroups $(S_t^i)_{t \geq 0}$ on E_i with generators $(\mathcal{L}_i, \mathcal{D}(\mathcal{L}_i))$ and associated Feller processes $(X_t^i)_{t \geq 0}$ on appropriate stochastic bases $(\Omega_i, \mathbb{P}_i, \mathcal{F}_i)$, representing the dynamics of those particles (in the following these processes will be referred to as the ‘one-particle dynamics’). Let us also assume that the spaces $C_c^\infty(E_i)$ are cores for the semigroups $(S_t^i)_{t \geq 0}$.

Remark 10.3. *By Watanabe’s Theorem (see for instance [79, Proposition 17.9]), $C_c^\infty(E_i)$ is a core for $\mathcal{L}_i$ if it is dense in $\mathcal{D}(\mathcal{L}_i)$ and invariant under $(S_t^i)_{t \geq 0}$. It is possible to extend our framework by exchanging $C_c^\infty(E_i)$ for other cores, say $\mathcal{D}_i$, as long as the first condition in Definition 10.7 is altered accordingly.*

Consider now a Feller process $(\bar{X}_t)_{t \geq 0}$ on the product space $\bar{E} := E_1 \times \dots \times E_n$, together with its associated semigroup $(\bar{S}_t)_{t \geq 0}$ on $\mathcal{B}_b(\bar{E})$ and generator $(\bar{\mathcal{L}}, \mathcal{D}(\bar{\mathcal{L}}))$ in $C_0(\bar{E})$. We will denote the E_i -valued coordinate processes of $(\bar{X}_t)_{t \geq 0}$ by $(\bar{X}_t^i)_{t \geq 0}$.

Definition 10.6 (Feller couplings). *The process $(\bar{X}_t)_{t \geq 0}$ is called a Feller coupling of the processes $(X_t^i)_{t \geq 0}$, if its marginals are given by these processes, i.e. if for all $i \in \{1, \dots, n\}$, the processes $(\bar{X}_t^i)_{t \geq 0}$ and $(X_t^i)_{t \geq 0}$ induce the same law on the space of càdlàg functions $D([0, \infty), E_i)$ [¶].*

Our aim in this section is to characterise the infinitesimal generators of such coupled processes.

Remark 10.4. *We are making the following two assumptions when considering the class of processes described above. Firstly, we assume certain continuity properties of the process $(\bar{X}_t)_{t \geq 0}$, see Lemma 10.1. Restricting our attention to the class of Feller processes allows us to use the theory of strongly continuous semigroups on Banach spaces as briefly hinted at in Section 1.1 for the development of the theoretical framework. In examples and applications however (see Chapters 11 and 13), we will relax this assumption a bit, allowing for more general processes.*

Secondly, we consider processes $(\bar{X}_t)_{t \geq 0}$ that are Markovian. Obviously there are many non-Markovian couplings of the underlying processes $(X_t^i)_{t \geq 0}$ and indeed those might be of particular interest for applications. Hence we plan to investigate the possibility of extending our framework in this direction in a forthcoming project.

We now proceed to introduce a class of linear (unbounded) operators $(\Gamma, \mathcal{D}(\Gamma))$ on $\mathcal{B}_b(\bar{E})$:

[¶]Every Feller process has a càdlàg modification, see [26, Theorem 1.19].

Definition 10.7 (Coupling operators). *Let $(\Gamma, \mathcal{D}(\Gamma))$ be a (possibly unbounded) operator on $\mathcal{B}_b(\bar{E})$. Then $(\Gamma, \mathcal{D}(\Gamma))$ is called a coupling operator if the following conditions are satisfied:*

1. *Test functions that only depend on one component of $\bar{x} = (x_1, \dots, x_n)$ are in the kernel of Γ : For all $i \in \{1, \dots, n\}$ and $f \in C_c^\infty(E_i)$ it holds that $f \in \mathcal{D}(\Gamma)$ and*

$$\Gamma f = 0. \quad (10.12)$$

2. *The operator*

$$\bar{\mathcal{L}}_\Gamma := \sum_{i=1}^n \mathcal{L}_i + \Gamma \quad (10.13)$$

with domain $\mathcal{D}(\bar{\mathcal{L}}_\Gamma) = \bigotimes_{i=1}^n \mathcal{D}(\mathcal{L}_i) \cap \mathcal{D}(\Gamma)$ is closable and its closure is the generator of a Feller process on $\bar{E}$.

The Feller semigroup corresponding to a coupling operator will be referred to by $(\bar{S}_t^\Gamma)_{t \geq 0}$. Furthermore, the set of coupling operators will be denoted by $\mathcal{G}$, i.e.

$$\mathcal{G} = \{\Gamma : \mathcal{D}(\Gamma) \subset \mathcal{B}_b(\bar{E}) \rightarrow \mathcal{B}_b(\bar{E}) : \text{Conditions 1 and 2 are satisfied.}\}.$$

Remark 10.5. *We will not distinguish (notationally) between $\bar{\mathcal{L}}_\Gamma$ and its closure. Notice also that the first condition in Definition 10.7 necessitates thinking of Γ as an operator defined on (a subspace of) $\mathcal{B}_b(\bar{E})$ (rather than $C_0(\bar{E})$), because $C_c^\infty(E_i) \not\subset C_0(\bar{E})$. The second condition is naturally concerned with $\bar{\mathcal{L}}_\Gamma$ being the generator of a semigroup on $C_0(\bar{E})$ (and hence with the natural restriction of Γ).*

We have the following result, characterising completely the set of rich Feller couplings in terms of the coupling operators $\mathcal{G}$:

Proposition 10.1. *For any $\Gamma \in \mathcal{G}$, the Feller process generated by $\bar{\mathcal{L}}_\Gamma$ as defined in (10.13) is a coupling of the processes $((X_t^i)_{t \geq 0}, i \in \{1, \dots, n\})$. Conversely, if $(\bar{X}_t)_{t \geq 0}$ is a Feller coupling of the processes $((X_t^i)_{t \geq 0}, i \in \{1, \dots, n\})$, then its generator is of the form (10.13), with $\Gamma \in \mathcal{G}$.*

Proof. Let $\Gamma \in \mathcal{G}$, and consider the process $(\bar{X}_t)_{t \geq 0}$ generated by the corresponding operator $\bar{\mathcal{L}}_\Gamma$ as defined in (10.13). Let us fix $i \in \{1, \dots, n\}$. Clearly, $C_c^\infty(E_i) \subset \mathcal{D}(\bar{\mathcal{L}}_\Gamma)$ and $\bar{\mathcal{L}}_\Gamma f = \mathcal{L}_i f$ for $f \in C_c^\infty(E_i)$. Hence, for all $f \in C_c^\infty(E_i)$, the process

$$f(\bar{X}_t) - f(\bar{X}_0) - \int_0^t (\mathcal{L}_i f)(\bar{X}_s) ds, \quad t \geq 0, \quad (10.14)$$

is a martingale with respect to the natural filtration $(\mathcal{F}_t^{\bar{X}})_{t \geq 0}$ generated by $(\bar{X}_t)_{t \geq 0}$, according to Theorem 10.1. From the uniqueness of the martingale problem for the generator $\mathcal{L}_i$

(see again Theorem 10.1) and the fact that $C_c^\infty(E_i)$ is a core for $\mathcal{L}_i$, it follows that $(\bar{X}_t^i)_{t \geq 0}$ has indeed the same law as $(X_t^i)_{t \geq 0}$.

Conversely, assume that $(\bar{X}_t)_{t \geq 0}$ is a Feller coupling of the processes $((X_t^i)_{t \geq 0}, i \in \{1, \dots, n\})$ and denote its generator by $(\tilde{\mathcal{L}}, \mathcal{D}(\tilde{\mathcal{L}}))$. For fixed $i \in \{1, \dots, n\}$, we first argue that $C_c^\infty(E_i) \subset \mathcal{D}(\tilde{\mathcal{L}})$, referring to the domain of the extended generator defined in (10.5b). Indeed, this amounts to showing that for all $f \in C_c^\infty(E_i)$ the limit

$$\lim_{t \rightarrow 0} \frac{1}{t} (\tilde{S}_t f - f) \quad (10.15)$$

exists uniformly on compact sets. By the coupling property (Definition 10.6), we have that $\tilde{S}_t f = S_t^i f$ for $f \in C_c^\infty(E_i)$. Therefore (and since $C_c^\infty(E_i) \subset \mathcal{D}(\mathcal{L}_i)$ by assumption), it follows that the limit (10.15) even exists uniformly on the whole of $\bar{E}$.

We can now define $\tilde{\Gamma} := \tilde{\mathcal{L}} - \sum_{i=1}^n \mathcal{L}_i$ on $\mathcal{D}(\tilde{\Gamma}) := \mathcal{D}(\tilde{\mathcal{L}}) \cap \bigotimes_{i=1}^n \mathcal{D}(\mathcal{L}_i)$. It is then sufficient to show that $\tilde{\Gamma}$ satisfies the first condition of Definition 10.7. To this end, take $f \in C_c^\infty(E_i)$ in the martingale problem for $\tilde{\mathcal{L}}$ to see that

$$f(\bar{X}_t^i) - f(\bar{X}_0^i) - \int_0^t (\mathcal{L}_i f)(\bar{X}_s^i) ds - \int_0^t (\tilde{\Gamma} f)(\bar{X}_s) ds, \quad t \geq 0, \quad (10.16)$$

is a martingale, again with respect to the natural filtration $(\mathcal{F}_t^{\bar{X}})_{t \geq 0}$ generated by $(\bar{X}_t)_{t \geq 0}$. Since $(\bar{X}_t^i)_{t \geq 0}$ and $(X_t^i)_{t \geq 0}$ are equal in law by assumption, it follows that $(\bar{X}_t^i)_{t \geq 0}$ is a solution to the martingale problem for $\mathcal{L}_i$. Hence, $\int_0^t (\tilde{\Gamma} f)(\bar{X}_s) ds$ has to be a martingale as well. Since this process is of finite variation (and the initial condition for the process $(\bar{X}_t)_{t \geq 0}$ can be chosen arbitrarily), this implies $\tilde{\Gamma} f = 0$. $\square$

Remark 10.6. *Similar approaches, describing couplings in terms of coupling operators, are known in the literature. See for instance [32, Chapter 2, Definition 2.7] and references therein. The exact result of Proposition 10.1 and its proof using martingale problems seems to be new and in particular relevant for Conjecture 2.18 and Open Problem 2.19 in [32].*

Example 10.1. *Independent (or trivial) coupling:* The zero operator $\Gamma = \mathbf{0}$ is always in $\mathcal{G}$, as the conditions of Definition 10.7 clearly hold. Indeed, consider the operator

$$\bar{\mathcal{L}}_0 := \sum_{i=1}^n \mathcal{L}_i \quad (10.17)$$

on the domain $\mathcal{D}(\bar{\mathcal{L}}_0) := \bigoplus_i \mathcal{D}(\mathcal{L}_i)$. It is straightforward (see for instance [4, A-I 3.7]) to show that $\bar{\mathcal{L}}_0$ is the generator of a Feller semigroup $(\bar{S}_t^0)_{t \geq 0}$ given by

$$\bar{S}_t^0 f = \left(\prod_{i=1}^n S_t^i \right) f, \quad f \in \mathcal{B}_b(\bar{E}), \quad t \geq 0, \quad (10.18)$$

and that the associated Feller process is just $(\bar{X}_t^0)_{t \geq 0} = (X_t^1, \dots, X_t^n)_{t \geq 0}$, i.e. it is obtained

from independent copies of the underlying processes.

Let us briefly discuss some of the implications of the conditions in Definition 10.7. As can be seen from the proof of Proposition 10.1, the first condition is instrumental in guaranteeing that the coupled process $(\bar{X}_t)_{t \geq 0}$ has the correct marginals. To put the second condition into context, we recall that generators of Feller semigroups can be characterised by means of the Hille-Yosida-Ray theorem in terms of the positive maximum principle (see Definition 10.5 and Theorem 10.2). As we will see in examples (Chapter 11), the latter often restricts the ‘size’ of coupling operators, so that the set $\mathcal{G}$ often turns out to be ‘bounded’ in a certain sense. Let us close this section by mentioning the following conjecture:

Conjecture 10.1. *The set $\mathcal{G}$ is convex.*

It is obvious that the first condition of Definition 10.7 is closed under convex combinations. The second condition could be addressed by appealing to the Hille-Yosida-Ray theorem (Theorem 10.2). Indeed, the first condition of this result is trivially satisfied by our assumption that $C_c^\infty(\bar{E}) \subset \mathcal{D}(\bar{\mathcal{L}}_\Gamma)$. The second condition (positive maximum principle) is clearly closed under convex combinations. Proving that the third condition is closed under convex combinations poses some technical challenges that we are not presently able to overcome.

10.3 ERGODICITY OF COUPLINGS

From here on, let us make the following assumption, natural in the context of MCMC samplers:

Assumption 10.2. *The one-particle processes $(X_t^i)_{t \geq 0}$ are ergodic.*

Recall that according to Definition 1.3, ergodicity entails that each process $(X_t^i)_{t \geq 0}$ admits a unique invariant measure; those will be denoted by π_i in the sequel. Following up on Example 10.1, we see that the semigroup $(\bar{S}_t^0)_{t \geq 0}$ as given in (10.18) with generator $\bar{\mathcal{L}}_0$ as defined in (10.17) is ergodic with respect to the product measure

$$\bar{\pi}_0 := \bigotimes_{i=1}^n \pi_i \tag{10.19}$$

on $\bar{E}$. Unfortunately, it turns out that not all coupling operators $\Gamma \in \mathcal{G}$ induce ergodic coupled processes, even under the Assumption 10.2 (for an example, see [92, Section 3.1]). We therefore make the following definition:

Definition 10.8 (Ergodic couplings). *A coupling operator $\Gamma \in \mathcal{G}$ is called ergodic, if the Feller process generated by $\bar{\mathcal{L}}_\Gamma$ is ergodic. The corresponding subset of ergodic coupling*

operators will be denoted by $\mathcal{G}^0$. The unique invariant measure associated to $\Gamma \in \mathcal{G}^0$ will be denoted by $\bar{\pi}_\Gamma$.

Remark 10.7. By construction, the measures $\bar{\pi}_\Gamma$ are couplings of the one-particle invariant measures $(\pi_i)_{i=1}^n$.

Remark 10.8. For the analysis, ergodicity of the coupling is a crucial requirement (although with more work it might be possible to extend some of the results to the case when ergodicity fails to hold). Let us emphasize however that the validity of (9.6) does not depend on this, as only the marginal property of the coupling is used in its derivation. Hence in practice it is harmless to use nonergodic couplings, and in fact our results obtained in Chapter 12 (in particular, Theorem 12.3) suggest using couplings that are at least not straightforwardly seen to be ergodic. In this case, quantities measuring the performance of the sampler (such as the asymptotic variance corresponding to certain observables) might be undefined or depend on the initial condition.

In general, ergodicity might fail in various ways. For instance, the process might not admit any invariant measure at all, or convergence towards a unique invariant measure (in the Césaro sense) might not hold. The following result shows that the situation is simpler in our context:

Lemma 10.2. *Let Assumption 10.2 be satisfied. Then the following hold:*

1. *Every Feller coupling admits at least one invariant measure.*
2. *If a Feller coupling admits a unique invariant measure, then the process is ergodic, i.e. (1.18) holds.*

Proof. We proceed along the lines of the proof of the Krylov-Bogolyubov theorem [35, Section 3.1]. Let $\nu_i \in \mathcal{P}(E_i)$ be arbitrary initial conditions for the processes $(X_t^i)_{t \geq 0}$. By ergodicity, the families $(\tilde{\pi}_t^i)_{t \geq 0}$ of Césaro averages

$$\tilde{\pi}_t^i(A) = \frac{1}{t} \int_0^t ((S_s^i)^* \nu_i)(A) ds, \quad A \in \mathcal{B}(E_i),$$

are convergent, and therefore tight. Let $(\bar{X}_t)_{t \geq 0}$ be a Feller coupling and denote the corresponding Césaro averages by $(\tilde{\bar{\pi}}_t)_{t \geq 0}$. For any $t \geq 0$, $\tilde{\bar{\pi}}_t$ is a coupling of $(\tilde{\pi}_t^i)_{i=1}^n$. Using Lemma 16.1 (proved in Section 16.1), we see that $(\tilde{\bar{\pi}}_t)_{t \geq 0}$ is tight. By Prokhorov's theorem, there exists a weakly converging subsequence, the limit of which is an invariant measure (as in the proof of the Krylov-Bogolyubov theorem). This proves the first claim. Now let us assume that there exists a unique invariant measure. Since any convergent subsequence of $(\tilde{\bar{\pi}}_t)_{t \geq 0}$ has to converge to the same limit, and the sequence is tight, the second claim follows. $\square$

Corollary 10.1 (Regular couplings). *Let Assumption 10.2 be satisfied and consider a Feller coupling $(\bar{X}_t)_{t \geq 0}$. If the corresponding transition functions $(\rho_t(x, \cdot))_{t \geq 0, x \in \bar{E}}$ are mutually absolutely continuous (i.e. if the process is regular), then $(\bar{X}_t)_{t \geq 0}$ is ergodic.*

The measures π_i , as well as $\bar{\pi}_\Gamma$ (for $\Gamma \in \mathcal{G}^0$) induce the usual Hilbert spaces $L^2(\pi_i)$ and $L^2(\bar{\pi}_\Gamma)$ of square-integrable functions. Let us also recall the subspaces of centred functions, defined by

$$L_0^2(\bar{\pi}_\Gamma) = \{f \in L^2(\bar{\pi}_\Gamma) \mid \bar{\pi}_\Gamma(f) = 0\}, \quad (10.20)$$

and analogously for $L_0^2(\pi_i)$. Since any Feller process has a right-continuous version, the semigroups $(\bar{S}_t^\Gamma)_{t \geq 0}$ have unique extensions to strongly continuous semigroups on $L^2(\bar{\pi}_\Gamma)$, see Section 1.1. Before moving on to a somewhat more explicit description of coupling operators, let us mention the following open question, related to Conjecture 10.1:

Conjecture 10.2. *The set $\mathcal{G}^0$ is convex.*

10.4 A GENERAL WAY OF CONSTRUCTING COUPLING OPERATORS

In this section we describe an approach to construct coupling operators explicitly in applications. The particular form presented here is also theoretically important since some of the calculations in later sections depend on it (especially the proof of Theorem 12.3).

As in the previous section, we assume that the marginal processes are ergodic with respect to invariant measures π_i (see Assumption 10.2). Furthermore, let us assume that the generators $\mathcal{L}_i$ can be written as

$$\mathcal{L}_i = - \sum_{k=1}^{K_i} (A_k^i)^* A_k^i + \mathcal{T}^i, \quad (10.21)$$

where $(A_k^i)^*$ denotes the adjoint of A_k^i in $L^2(\pi_i)$, and $\mathcal{T}^i$ is skew-symmetric in $L^2(\pi_i)$. Clearly, this decomposition into symmetric and skew-symmetric parts is always possible, and in many cases the operators A_k^i and $\mathcal{T}^i$ can be chosen to have convenient forms (see Section 1.3). Note however that the decomposition (10.21) is not unique since there are (infinitely) many ways of choosing the operators A_k^i . A particular choice of decomposing the generators $\mathcal{L}_i$ as in (10.21) hence essentially amounts to the choice of square roots for the symmetric parts. We remark here that naturally $C_c^\infty(\bar{E}) \subset \mathcal{D}(A_k^i)$ and $C_c^\infty(\bar{E}) \subset \mathcal{D}(\mathcal{T}^i)$ are implicitly assumed, authorising the computations in later sections. The following lemma is essential for the construction in this section:

Lemma 10.3. *Let A_k^i and $\mathcal{T}^i$ be given as in equation (10.21). Then*

$$\text{span } \mathbf{1} \subset \ker A_k^i, \quad \text{span } \mathbf{1} \subset \ker \mathcal{T}^i \quad (10.22)$$

for all $i \in \{1, \dots, n\}$, $k \in \{1, \dots, K_i\}$.

Proof. See [141, Proposition 2]. By the conservativeness of $(S_t^i)_{t \geq 0}$, $\mathbf{1} \in \ker \mathcal{L}_i$ for all $i \in \{1, \dots, n\}$. It follows that

$$\langle \mathbf{1}, \mathcal{L}_i \mathbf{1} \rangle_{L^2(\pi_i)} = - \sum_{k=1}^{K_i} \|A_k^i \mathbf{1}\|_{L^2(\pi_i)}^2 = 0,$$

and hence $A_k^i \mathbf{1} = 0$. For the second claim, notice that

$$\mathcal{T}^i \mathbf{1} = \left(\mathcal{L}_i + \sum_{k=1}^{K_i} (A_k^i)^* A_k^i \right) \mathbf{1} = 0.$$

□

We may now set

$$\Gamma = \sum_{(i,j,k,l) \in \mathcal{J}} \alpha_{ijkl}(x_1, \dots, x_n) A_k^i A_l^j \quad (10.23)$$

for appropriate measurable functions $\alpha_{ijkl} : \bar{E} \rightarrow \mathbb{R}$ and where we have introduced the set of *admissible indices*

$$\mathcal{J} = \{(i, j, k, l) \in \mathbb{N}^4 : 1 \leq i, j \leq n, i \neq j, 1 \leq k \leq K_i, 1 \leq l \leq K_j\} \quad (10.24)$$

associated to the decomposition (10.21). Applying Lemma 10.3, we see immediately that the first condition of Definition 10.7 is satisfied. The second condition will typically enforce certain bounds on the functions α_{ijkl} via the positive maximum principle (Definition 10.5) as well as regularity constraints if we are interested in Feller couplings. Those properties will have to be determined according to the particular form of the generators $\mathcal{L}_i$. Furthermore, whether Γ as defined in (10.23) belongs to $\mathcal{G}^0$ will also depend on the choice of the functions α_{ijkl} .

It is not clear whether the construction presented in this section exhausts the class of coupling operators $\mathcal{G}$. We present this problem as a conjecture:

Conjecture 10.3. *For $i \in \{1, \dots, n\}$, assume that we are given generators $(\mathcal{L}_i, \mathcal{D}(\mathcal{L}_i))$ of ergodic Feller semigroups. Then there exist decompositions of the form (10.21) and a set of functions*

$$\mathcal{U} = \{(\alpha_{ijkl})_{(i,j,k,l) \in \mathcal{J}} : \bar{E} \rightarrow \mathbb{R}\} \quad (10.25)$$

such that

$$\mathcal{G} = \left\{ \Gamma = \sum_{(i,j,k,l) \in \mathcal{J}} \alpha_{ijkl} A_k^i A_l^j : \alpha_{ijkl} \in \mathcal{U} \right\}. \quad (10.26)$$

In the case when the underlying processes $(X_t^i)_{t \geq 0}$ are $\mathbb{R}^{d_i}$ -valued (i.e. $E_i = \mathbb{R}^{d_i}$) and have continuous paths almost surely, Courrège's theorem (Theorem 10.3) provides an explicit characterisation of Feller generators. If furthermore these processes are ergodic with respect to given invariant measures, the decomposition of their generators into symmetric and skew-symmetric parts can be made explicit (see Theorem 1.7 and Remark 1.7). Combining these theorems, we obtain the following partial result:

Proposition 10.2. *Let $E_i = \mathbb{R}^{d_i}$ for positive integers $d_i \in \mathbb{N}$ and assume that the processes $(X_t^i)_{t \geq 0}$ are ergodic and solve the Itô SDEs*

$$dX_t^i = a^i(X_t^i) dt + \sqrt{2}b^i(X_t^i) dW_t^i, \quad (10.27)$$

where $a^i \in C^1(\mathbb{R}^{d_i}, \mathbb{R}^{d_i})$, $b^i \in C^1(\mathbb{R}^{d_i}, \mathbb{R}^{d_i \times m_i})$, and $(W_t^i)_{t \geq 0}$ are standard m_i -dimensional Brownian motions. Then the conclusion of Conjecture 10.3 holds.

11. EXAMPLES OF COUPLED PROCESSES

Here we will illustrate the framework developed in the last chapter with concrete examples, revisiting some of the dynamics already introduced in Section 1.3 to address the task of sampling from the measure

$$\pi(\mathrm{d}x) = \frac{1}{Z} e^{-V(x)} \mathrm{d}x, \quad x \in \mathbb{R}^d, \quad (11.1)$$

where $V \in C^\infty(\mathbb{R}^d)$ is a potential satisfying

$$Z := \int_{\mathbb{R}^d} e^{-V(x)} \mathrm{d}x < \infty. \quad (11.2)$$

11.1 OVERDAMPED LANGEVIN DYNAMICS

Our first group of examples is concerned with the overdamped Langevin dynamics, introduced in (1.72). Let us start with the one-dimensional case, already encountered in the introduction to this part (Chapter 9).

TWO PARTICLES IN ONE DIMENSION

We consider $n = 2$ particles moving in dimension $d = 1$, each of them according to the dynamics

$$\mathrm{d}X_t = -V'(X_t) \mathrm{d}t + \sqrt{2} \mathrm{d}W_t. \quad (11.3)$$

Note that in order to precisely fit into our framework developed in Chapter 10, the process $(X_t)_{t \geq 0}$ is required to be a Feller process according to Definition 10.2. This property can be guaranteed by imposing certain growth conditions on the potential V , see [87, Proposition 5.9] and [97, Theorem 5.3.2, Example 5.3.3]. However, we wish to remark that the Feller property is not crucial in practice and dispensing with this regularity requirement still leads to perfectly well-defined couplings as will become clear in Lemma 11.1 below. According to Remark 1.7, the generator of (11.3) is given by

$$\mathcal{L} = -V'(x) \partial_x + \partial_x^2 = -\partial_x^* \partial_x, \quad (11.4)$$

where the adjoint is taken in $L^2(\pi)$. In particular, $\mathcal{L}$ can naturally be written in the form (10.21), with $A = \partial_x$ and $\mathcal{T} = 0$. To illustrate trivial couplings (see Example 10.1), consider first two independent identical copies of (11.3), denoted by $(X_t, Y_t)_{t \geq 0}$, hence evolving according to the dynamics

$$dX_t = -V'(X_t) dt + \sqrt{2} dW_t^x, \quad (11.5a)$$

$$dY_t = -V'(Y_t) dt + \sqrt{2} dW_t^y, \quad (11.5b)$$

on the product space $\bar{E} = \mathbb{R}^2$. Since for now the processes $(W_t^x)_{t \geq 0}$ and $(W_t^y)_{t \geq 0}$ are supposed to be two independent standard Brownian motions, the generator of (11.5) is given by

$$\bar{\mathcal{L}}_0 = \mathcal{L}_x + \mathcal{L}_y, \quad (11.6)$$

with $\mathcal{L}_x = -V'(x)\partial_x + \partial_x^2 = -\partial_x^* \partial_x$ and $\mathcal{L}_y = -V'(y)\partial_y + \partial_y^2 = -\partial_y^* \partial_y$, in agreement with Example 10.1, equation (10.17). The invariant measure of (11.5) is given by the product

$$\bar{\pi}_0 = \pi_x \otimes \pi_y = \frac{1}{Z^2} e^{-(V(x)+V(y))} dx dy. \quad (11.7)$$

Now let us consider nontrivial couplings. Following Section 10.4, we may set

$$\Gamma = 2\alpha(x, y)\partial_x \partial_y, \quad (11.8)$$

for an appropriate measurable function $\alpha : \mathbb{R}^2 \rightarrow \mathbb{R}$ (we have inserted a factor of 2 for convenience). According to Proposition 10.2, the set of operators of the form (11.8) exhausts the set of coupling operators $\mathcal{G}$. Clearly, the first condition of Definition 10.7 is satisfied for this set of operators (this is already guaranteed by using the construction from Section 10.4). The second condition enforces

$$-1 \leq \alpha(x, y) \leq 1, \quad \text{for all } x, y \in \mathbb{R}. \quad (11.9)$$

Indeed, observe that

$$\bar{\mathcal{L}}_\Gamma = \bar{\mathcal{L}}_0 + \Gamma = \nabla_z U(z) \cdot \nabla_z + Q(z) : \nabla_z \nabla_z, \quad (11.10)$$

where $z = (x, y)$, $U(z) = U(x, y) = V(x) + V(y)$,

$$Q(z) = Q(x, y) = \begin{pmatrix} 1 & \alpha(x, y) \\ \alpha(x, y) & 1 \end{pmatrix}, \quad (11.11)$$

and where $:$ denotes the Frobenius inner product of matrices. According to Courrège's theorem (see Theorem 10.3), $\bar{\mathcal{L}}_\Gamma$ satisfies the positive maximum principle (required by the Hille-Yosida-Ray Theorem 10.2) only if $Q(z)$ is nonnegative definite for every $z \in \mathbb{R}^2$. From this, we immediately deduce the constraint (11.9).

Remark 11.1. *We are deliberately vague about the regularity properties of α . If we restrict our attention to Feller processes, α certainly has to be at least continuous, and there are multiple results in the literature guaranteeing the Feller property under mild further assumptions on α , in particular, Hölder regularity [3, 97]. For a discussion of the martingale problem for generators with discontinuous coefficients see the recent preprint [85] and references therein. Note that even in this simple case, it is very challenging to characterise exactly the set $\mathcal{G}$ as introduced in Definition 10.7. In applications, however, the Feller property is not crucial. Lemma 11.1 below shows that measurability of α is sufficient to ensure that a reasonable coupled process can be constructed.*

Assuming that (11.9) is satisfied, the dynamics induced by the generator (11.10) are (at least formally) given by

$$d \begin{pmatrix} X_t \\ Y_t \end{pmatrix} = \begin{pmatrix} -V'(X_t) \\ -V'(Y_t) \end{pmatrix} dt + \sqrt{2}G(X_t, Y_t) \begin{pmatrix} dW_t^x \\ dW_t^y \end{pmatrix}, \quad (11.12)$$

where $G(x, y)G(x, y)^T = Q(x, y)$, for instance

$$G(x, y) = \begin{pmatrix} \cos \beta(x, y) & g(x, y) \sin \beta(x, y) \\ g(x, y) \sin \beta(x, y) & \cos \beta(x, y) \end{pmatrix}, \quad (11.13)$$

with $\beta(x, y) = \frac{1}{2} \arcsin |\alpha(x, y)|$ and $g(x, y) = \operatorname{sgn} \alpha(x, y)$. We have chosen this parametrisation since it generalises readily to higher dimensions (see below). Let us stress that writing the dynamics in the form (11.12) is vital for applications, since it enables its simulation in a straightforward manner. The following lemma shows that the process constructed in this way is indeed a coupling in the sense of Definition 10.6. Furthermore, it turns out that only minimal regularity of α is required.

Lemma 11.1. *Let α be measurable. Then the dynamics (11.12) can be written as*

$$d \begin{pmatrix} X_t \\ Y_t \end{pmatrix} = \begin{pmatrix} -V'(X_t) \\ -V'(Y_t) \end{pmatrix} dt + \sqrt{2} \begin{pmatrix} dB_t^x \\ dB_t^y \end{pmatrix}, \quad (11.14)$$

with two Brownian motions $(B_t^x)_{t \geq 0}$ and $(B_t^y)_{t \geq 0}$ that are in general not independent.

Proof. The claim follows from applying Lemma 16.2 from Chapter 16 to the components of the SDE

$$\begin{pmatrix} dB_t^x \\ dB_t^y \end{pmatrix} = G \begin{pmatrix} dW_t^x \\ dW_t^y \end{pmatrix}, \quad (11.15)$$

noting that $G^x := (g \cos \beta \sin \beta)$ and $G^y := (g \sin \beta \cos \beta)$ indeed satisfy condition (16.3) with $N = 1$ and $M = 2$. $\square$

Note that the function α (equivalently the pair β and g) encodes the coupling between the

Brownian motions $(B_t^x)_{t \geq 0}$ and $(B_t^y)_{t \geq 0}$. The parameter $\beta \in [0, \frac{\pi}{4}]$ is related to the *strength* of the coupling, whereas $g \in \{-1, 1\}$ is related to its *direction*. Indeed, if $\beta \equiv 0$, then $(B_t^x)_{t \geq 0}$ and $(B_t^y)_{t \geq 0}$ are independent ('trivial coupling', see (11.5) and Example 10.1). If $\beta \equiv \frac{\pi}{4}$ and $g \equiv 1$, then $B_t^x = B_t^y$ for $t \geq 0$, almost surely ('synchronous coupling'). Likewise, if $\beta \equiv \frac{\pi}{4}$ and $g \equiv -1$, then $B_t^x = -B_t^y$ ('mirror coupling').

Remark 11.2 (Ergodic couplings). *If the bound (11.9) is satisfied with strict inequalities, then the generator (11.10) is elliptic, and hence, by Theorem 1.6 and Corollary 10.1, the coupled process is ergodic. Let us mention that this condition is not necessary for ergodicity. Indeed, consider the coupling operator $\Gamma = -2\partial_x\partial_y$, inducing the so-called 'two-point motion' [12], i.e. $(X_t)_{t \geq 0}$ and $(Y_t)_{t \geq 0}$ are driven by the same Brownian motion, only differing by their initial laws. Under mild regularity conditions (i.e. Lipschitz continuity of the coefficients), it can be shown that $(X_t, Y_t)_{t \geq 0}$ is ergodic with respect to $\bar{\pi}_\Gamma = \frac{1}{2}e^{-V(x)}\delta_{x-y}(dx dy)$, see for instance [92, Theorem 2.1].*

THE GENERAL CASE

Here, we will extend the discussion from the previous section to the general case of n particles moving in d dimensions, i.e. we are concerned with couplings of the dynamics

$$dX_t^i = -\nabla V(X_t^i) dt + \sqrt{2} dW_t^i, \quad i = 1, \dots, n, \quad (11.16)$$

the processes $(X_t^i)_{t \geq 0}$ being $\mathbb{R}^d$ -valued. If the Brownian motions $(W_t^i)_{t \geq 0}$ are independent, then the joint process $(\bar{X}_t)_{t \geq 0} = (X_t^1, \dots, X_t^n)_{t \geq 0}$ is ergodic with respect to the product measure $\bar{\pi}_0 = \bigotimes_{i=1}^n \pi_i$ on $\mathbb{R}^{nd}$ (see Example 10.1) and the corresponding generator is given by

$$\bar{\mathcal{L}}_0 = \sum_{i=1}^n \mathcal{L}_i, \quad \mathcal{L}_i = -\nabla V(x_i) \cdot \nabla_{x_i} + \Delta_{x_i} = -\sum_{k=1}^d (\partial_k^i)^* \partial_k^i. \quad (11.17)$$

Here, ∂_k^i denotes the derivative with respect to the k -th component of x_i , and the adjoints are taken in the spaces $L^2(\pi_i)$. Clearly, the generators $\mathcal{L}_i$ are decomposed as in (10.21), with $A_k^i = \partial_k^i$ and $\mathcal{T}_i = 0$.

Remark 11.3. *Instead of (11.16), we can also consider the more general dynamics*

$$dX_t^i = -Q_i(X_t^i) \nabla V(X_t^i) dt + (\nabla \cdot Q_i)(X_t^i) dt + J_i \nabla V(X_t^i) dt + \sqrt{2Q_i(X_t^i)} dW_t^i, \quad (11.18)$$

with $i = 1, \dots, n$, $J_i \in \mathbb{R}_{\text{skew}}^{d \times d}$ being skew-symmetric matrices and $Q_i : \mathbb{R}^d \rightarrow \mathbb{R}_{\text{sym}}^{d \times d}$ being positive definite matrix-valued functions, as discussed in Section 1.3. Note that in this case the processes $(X_t^i)_{t \geq 0}$ are not copies of each other, since Q_i and J_i may not be the same for different particles.

To construct nontrivial couplings, we may set

$$\Gamma = \sum_{i,j=1, i \neq j}^n \sum_{k,l=1}^d \alpha_{ijkl} \partial_k^i \partial_l^j \quad (11.19)$$

for appropriate* functions $\alpha_{ijkl} : \mathbb{R}^{nd} \rightarrow \mathbb{R}$, following Section 10.4. Note that $\partial_k^i \partial_l^j$ is symmetric with respect to the interchange of indices $(i, k) \leftrightarrow (j, l)$, and so we may assume that $\alpha_{ijkl} = \alpha_{jilk}$. As in the one-dimensional case, the generator of the coupled system $\tilde{\mathcal{L}}_\Gamma = \tilde{\mathcal{L}}_0 + \Gamma$ is a second order differential operator which we require to be (possibly degenerately) elliptic in order for the second condition of Definition 10.7 to be satisfied (again with reference to Courrège's theorem (Theorem 10.3)).

To derive more easily verifiable conditions on the functions α_{ijkl} , let us introduce a matrix-valued function (or *matrix field*) $Q : (\mathbb{R}^d)^n \rightarrow \mathbb{R}^{nd \times nd}$ as follows. Firstly, it is helpful to view the target space of Q as $\mathbb{R}^{nd \times nd} \equiv (\mathbb{R}^{d \times d})^{n \times n}$, i.e. we think of $Q(x_1, \dots, x_n)$ as an $n \times n$ -matrix the entries of which are themselves $d \times d$ matrices. In other words, $Q_{ij}(x_1, \dots, x_n)$ is a $d \times d$ matrix for every pair $(i, j) \in \{1, \dots, n\}^2$. The matrix field Q can then be defined by

$$Q_{ij}(x_1, \dots, x_n) = \begin{cases} I_{d \times d}, & i = j, \\ \alpha_{ij}(x_1, \dots, x_n), & i \neq j, \end{cases} \quad (11.20)$$

where $\alpha_{ij}(x_1, \dots, x_n) \in \mathbb{R}^{d \times d}$ denotes the matrix with entries $(\alpha_{ijkl}(x_1, \dots, x_n))_{k,l=1, \dots, d}$. Note that $Q(x_1, \dots, x_n)$ as defined in (11.20) is symmetric as a matrix in $\mathbb{R}^{nd \times nd}$ by our assumption that $\alpha_{ijkl} = \alpha_{jilk}$.

Remark 11.4. *The matrices Q_{ij} can now be thought of as describing the coupling between the particles i and j .*

As in the one-dimensional case, the generator of the fully coupled system can be written as

$$\tilde{\mathcal{L}}_\Gamma = \tilde{\mathcal{L}}_0 + \Gamma = \nabla_z U(z) \cdot \nabla_z + Q(z) : \nabla_z \nabla_z, \quad (11.21)$$

introducing the notation $z \equiv (x_1, \dots, x_n) \in \mathbb{R}^{nd \times nd}$ and $U(z) = \sum_{i=1}^n V(x_i)$. To make the connection to SDEs and arrive at a description analogous to (11.12), let us consider matrix fields $G : (\mathbb{R}^d)^n \rightarrow (\mathbb{R}^{d \times d})^{n \times n}$ satisfying

$$\sum_{j=1}^n G_{ij}(x_1, \dots, x_n) G_{ij}(x_1, \dots, x_n)^T = I_{d \times d}, \quad (11.22)$$

for all $i = 1, \dots, n$. Here the transposition T is taken in $\mathbb{R}^{d \times d}$. Note that matrix fields of

*Concerning the regularity of these functions, the discussion from the previous section applies, see in particular Remark 11.1.

this form give rise to the matrix fields $Q : (\mathbb{R}^d)^n \rightarrow \mathbb{R}^{nd \times nd}$ defined in (11.20) via $Q = GG^T$ (where the transposition is taken in $\mathbb{R}^{nd \times nd}$).

Remark 11.5. *The advantage of constructing the coupling in terms of the matrix field G is that pointwise positive semi-definiteness of Q is automatically satisfied. A practical way to fulfil the constraint (11.22) is to choose matrix fields $g_{ij} : (\mathbb{R}^d)^n \rightarrow \mathbb{R}^{d \times d}$ that are orthogonal pointwise, i.e.*

$$g_{ij}(x_1, \dots, x_n)g_{ij}(x_1, \dots, x_n)^T = g_{ij}(x_1, \dots, x_n)^T g_{ij}(x_1, \dots, x_n) = I_{d \times d}, \quad (11.23)$$

for all $i, j = 1, \dots, n$ and $(x_1, \dots, x_n) \in (\mathbb{R}^d)^n$, as well as weights $w_{ij} : (\mathbb{R}^d)^n \rightarrow \mathbb{R}$ satisfying

$$\sum_{j=1}^n w_{ij}^2(x_1, \dots, x_n) = 1, \quad (11.24)$$

for all $i = 1, \dots, n$. Then, setting $G_{ij} = w_{ij}g_{ij}$, condition (11.22) holds. Intuitively, the orthogonal matrices g_{ij} encode the coupling between particle i and j through a rotation of the noise. The weights w_{ij} can be interpreted as the relative coupling strengths between the particles. Observe that both g_{ij} and w_{ij} may depend on $(x_1, \dots, x_n)$, i.e. on the locations of all the particles. This construction is a direct generalisation of (11.13).

Assuming $GG^T = Q$, the dynamics associated to the generator (11.21) is (again, at least formally) given by

$$dX_t^i = -\nabla V(X_t^i) dt + \sqrt{2} \sum_{j=1}^n G_{ij}(X_t^1, \dots, X_t^n) dW_t^j, \quad i = 1, \dots, n, \quad (11.25)$$

where $\left((W_t^j)_{t \geq 0}\right)_{j=1}^n$ are assumed to be independent standard Brownian motions. As in the one-dimensional case, we have the following lemma:

Lemma 11.2. *There exist $\mathbb{R}^d$ -valued standard Brownian motions $\left((B_t^j)_{t \geq 0}\right)_{j=1}^n$, not necessarily independent, such that the dynamics (11.25) can be written as*

$$dX_t^i = -\nabla V(X_t^i) dt + \sqrt{2} dB_t^i, \quad i = 1, \dots, n. \quad (11.26)$$

Proof. The argument is identical to the one used in the proof of Lemma 11.1. $\square$

Concerning the ergodicity of couplings, we have similar findings to those of Remark 11.2. If the matrix field (11.20) is positive definite at every point (i.e. the generator (11.21) is elliptic), then Theorem 1.6 and Corollary 10.1 imply that the coupled process is ergodic. An example of nonergodic couplings can be found in [92, Section 3.1].

Example 11.1 (Two particles). The foregoing constructions become more explicit when considering only $n = 2$ particles. To simplify the notation, we denote their positions by

$x \equiv x_1$ and $y \equiv x_2$. The diffusion matrices (11.20) reduce to

$$Q(x, y) = \begin{pmatrix} I_{d \times d} & \alpha(x, y) \\ \alpha^T(x, y) & I_{d \times d} \end{pmatrix}, \quad (11.27)$$

with $\alpha : (\mathbb{R}^d)^2 \rightarrow \mathbb{R}^{d \times d}$. Using Schur complements [30, Appendix 5.5], we see that $Q(x, y)$ is positive semidefinite if and only if

$$\alpha(x, y)^T \alpha(x, y) \leq I_{d \times d}, \quad (11.28)$$

in the sense of symmetric matrices (Loewner ordering). The corresponding coupling operator is given by

$$(\Gamma f)(x, y) = 2 \operatorname{Tr} (\alpha^T(x, y) \nabla_{xy}^2 f(x, y)), \quad f \in C_c^\infty(\mathbb{R}^d \times \mathbb{R}^d), \quad (11.29)$$

where the matrix of mixed derivatives $\nabla_{xy}^2 \phi$ is given by $(\nabla_{xy}^2 \phi)_{ij} = (\partial_{x_i} \partial_{y_j} \phi)_{ij}$. In order to illustrate the construction from Remark 11.5, let $g : (\mathbb{R}^d \times \mathbb{R}^d) \rightarrow \mathbb{R}^{d \times d}$ be a field of orthogonal matrices (see (11.23)) and put

$$G(x, y) = \begin{pmatrix} \cos \beta(x, y) I_{d \times d} & \sin \beta(x, y) g(x, y) \\ \sin \beta(x, y) g^T(x, y) & \cos \beta(x, y) I_{d \times d} \end{pmatrix}, \quad (11.30)$$

the function $\beta : \mathbb{R}^d \times \mathbb{R}^d \rightarrow [0, \frac{\pi}{4}]$ again regulating the strength of the coupling, and being associated with the weights w_{ij} in (11.24). From $GG^T = Q$ it follows that α and g are connected via

$$\alpha = 2 \cos \beta \sin \beta \cdot g. \quad (11.31)$$

11.2 UNDERDAMPED LANGEVIN DYNAMICS

For fixed $\gamma > 0$ ('friction') and positive definite $M \in \mathbb{R}_{\text{sym}}^{d \times d}$ ('mass'), the dynamics

$$dq_t = M^{-1} p_t dt, \quad (11.32a)$$

$$dp_t = -\nabla V(q_t) dt - \gamma p_t dt + \sqrt{2\gamma} dW_t, \quad (11.32b)$$

is ergodic with respect to the measure

$$\pi = \frac{1}{Z} e^{-(V(q) + \frac{1}{2} p^T M^{-1} p)} dq dp, \quad (q, p) \in \mathbb{R}^{2d}, \quad (11.33)$$

$\bar{Z}$ being an appropriate normalisation constant (see Section 1.3 for details). The generator is given by

$$\mathcal{L} = M^{-1}p \cdot \nabla_q - \nabla_q V(q) \cdot \nabla_p + \gamma(-p \cdot \nabla_p + \Delta_p) \quad (11.34a)$$

$$= \mathcal{T} - \sum_{k=1}^d A_k^* A_k, \quad (11.34b)$$

where $\mathcal{T} = M^{-1}p \cdot \nabla_q - \nabla_q V(q) \cdot \nabla_p$ is skew-symmetric in $L^2(\pi)$, $A_k = \sqrt{\gamma} \partial_{p_k}$, and the adjoint is taken in $L^2(\pi)$. To construct a coupled sampler of n processes, we may proceed as in the overdamped case and set

$$\Gamma = \sum_{i,j=1, i \neq j}^n \sum_{k,l=1}^d \alpha_{ijkl} \partial_{p_k}^i \partial_{p_l}^j, \quad (11.35)$$

for appropriate functions $\alpha_{ijkl} : (\mathbb{R}^{2d})^n \rightarrow \mathbb{R}$, denoting by $\partial_{p_k}^i$ the derivative with respect to the k -th component of p of the i -th particle. Following very closely the discussion in Section 11.1, we can introduce matrix fields $Q : (\mathbb{R}^{2d})^n \rightarrow (\mathbb{R}^{d \times d})^{n \times n}$ and $G : (\mathbb{R}^{2d})^n \rightarrow (\mathbb{R}^{d \times d})^{n \times n}$ satisfying (11.20) and (11.22) (with $(x_1, \dots, x_n)$ replaced by $(q_1, p_1, \dots, q_n, p_n)$) such that the generator of the coupled system is given by

$$\bar{\mathcal{L}}_\Gamma = \sum_{i=1}^n (M^{-1}p_i \cdot \nabla_{q_i} - \nabla_{q_i} V(q_i) \cdot \nabla_{p_i} - \gamma p_i \cdot \nabla_{p_i}) + Q(z) : \nabla_{z_p} \nabla_{z_p}, \quad (11.36)$$

making again use of the notations $z \equiv (q_1, p_1, \dots, q_n, p_n)$, $z_p \equiv (p_1, \dots, p_n)$ and such that the associated dynamics are given by

$$dq_t^i = M^{-1}p_t^i dt, \quad (11.37a)$$

$$dp_t^i = -\nabla_{q_i} V(q_t^i) dt - \gamma p_t^i dt + \sqrt{2\gamma} \sum_{j=1}^n G_{ij}(q_1, p_1, \dots, q_n, p_n) dW_t^j, \quad (11.37b)$$

for $i = 1, \dots, n$. Also in this case, it is straightforward to see that an appropriate version of the Lemmas 11.1 and 11.2 holds.

Remark 11.6. *Generalising the above to the case where the particles have different frictions γ_i and masses M_i or some or all of them move according to perturbed versions of underdamped Langevin dynamics as considered in Part II is straightforward (see also Remark 11.3).*

Example 11.2 (Overdamped and underdamped Langevin dynamics). It is also possible to couple different types of dynamics (such as overdamped and underdamped Langevin

dynamics). For instance, consider the generators

$$\mathcal{L}_1 = -\nabla V(x) \cdot \nabla_x + \Delta_x, \quad (11.38a)$$

$$\mathcal{L}_2 = M^{-1}p \cdot \nabla_q - \nabla_q V(q) \cdot \nabla_p + \gamma(-p \cdot \nabla_p + \Delta_p), \quad (11.38b)$$

as in (11.17) and (11.34a). Setting

$$\Gamma = \sum_{k,l=1}^d \alpha_{kl}(x, q, p) \partial_{x_k} \partial_{p_l} \quad (11.39)$$

and following along the lines of Sections 11.1 and 11.2 will result in the coupled dynamics

$$dX_t = -\nabla V(X_t) dt + \sqrt{2} dB_t^{(1)}, \quad (11.40a)$$

$$dq_t = M^{-1} p_t dt, \quad (11.40b)$$

$$dp_t = -\nabla_q V(q) dt - \gamma p_t dt + \sqrt{2\gamma} dB_t^{(2)}, \quad (11.40c)$$

where the Brownian motions $(B_t^{(1)})_{t \geq 0}$ and $(B_t^{(2)})_{t \geq 0}$ are in general not independent (and the exact dependence results from the choice of the functions α_{kl}).

11.3 THE ZIGZAG PROCESS

The objective of this section is to show how the framework from Chapter 10 can be employed in the construction of coupled samplers from piecewise deterministic Markov processes, using the example of the zigzag process introduced in Section 1.3.2. For ease of exposition, we restrict our attention to the one-dimensional case. Let us briefly recall the setting from Section 1.3.2, specialising to the case $d = 1$.

The state space under consideration is $E = \mathbb{R} \times \{-1, +1\}$, and the generator of the zigzag process reads

$$\mathcal{L}f(x, \theta) = \theta \partial_x f(x, \theta) + \lambda(x, \theta) (f(x, -\theta) - f(x, \theta)), \quad f \in C_c^\infty(E), \quad (11.41)$$

where the switching rate λ is given by

$$\lambda(x, \theta) = \max(0, \theta V'(x)) + \gamma(x). \quad (11.42)$$

We refer to Section 1.3.2 for more details concerning the dynamics of the process generated by (11.41). According to [21, Proposition 1] and [139, Appendix B.2], the zigzag process satisfies the Feller property; for more general piecewise deterministic Markov processes

this topic has been studied in [37, see Theorem 27.6]. The measure

$$\pi = \frac{1}{2Z} e^{-V(x)} dx \otimes (\delta_{-1} + \delta_{+1}) \quad (11.43)$$

is invariant, and, under some additional assumptions[†], ergodic. Recall from Section 1.3.2 that $\mathcal{L}$ can be decomposed in the form

$$\mathcal{L} = -A^*A + \mathcal{T}, \quad (11.44)$$

where

$$A = \left(\frac{1}{4}|V'| + \frac{1}{2}\gamma \right)^{1/2} (R-1), \quad \mathcal{T} = \theta \partial_x + \frac{1}{2} \theta V' (R-1), \quad (11.45)$$

and the ‘flip operator’ R is given by

$$(Rf)(x, \theta) = f(x, -\theta), \quad f \in C_c^\infty(\bar{E}). \quad (11.46)$$

To construct a coupled sampler from two zigzag processes, let us introduce the following notation: We consider the state space $\bar{E} = \mathbb{R}^2 \times \{-1, +1\}^2$, denoting its elements by $(x, y, \theta_x, \theta_y)$. Furthermore, we will make use of the flip operators

$$(R_x f)(x, y, \theta_x, \theta_y) = f(x, y, -\theta_x, \theta_y), \quad (R_y f)(x, y, \theta_x, \theta_y) = f(x, y, \theta_x, -\theta_y). \quad (11.47)$$

Following Section 10.4, let us set

$$\Gamma = \alpha(x, y, \theta_x, \theta_y)(R_x - 1)(R_y - 1), \quad (11.48)$$

for an appropriate function $\alpha : \bar{E} \rightarrow \mathbb{R}$, i.e. Γ acts as

$$\begin{aligned} (\Gamma f)(x, y, \theta_x, \theta_y) &= \alpha(x, y, \theta_x, \theta_y) \cdot \\ &\cdot (f(x, y, \theta_x, \theta_y) - f(x, y, -\theta_x, \theta_y) - f(x, y, \theta_x, -\theta_y) + f(x, y, -\theta_x, -\theta_y)) \end{aligned} \quad (11.49)$$

on test functions $f \in C_c^\infty(\bar{E})$. Note that Γ vanishes on functions that either depend on only x and θ_x or only on y and θ_y . The next task is to obtain bounds on α that ensure

[†]See Chapter 17, as well as [21] and [23]. Let us mention in particular that ergodicity is guaranteed whenever the excess switching rate γ is strictly positive.

that the second condition in Definition 10.7 is satisfied. To this end, let us expand

$$(\bar{\mathcal{L}}_\Gamma f)(x, y, \theta_x, \theta_y) = (\mathcal{L}_x + \mathcal{L}_y + \Gamma)f(x, y, \theta_x, \theta_y) \quad (11.50a)$$

$$= \theta_x \partial_x f(x, y, \theta_x, \theta_y) + \theta_y \partial_y f(x, y, \theta_x, \theta_y) \quad (11.50b)$$

$$- (\lambda(x, \theta_x) + \lambda(y, \theta_y) - \alpha(x, y, \theta_x, \theta_y)) f(x, y, \theta_x, \theta_y) \quad (11.50c)$$

$$+ (\lambda(x, \theta_x) - \alpha(x, y, \theta_x, \theta_y)) f(x, y, -\theta_x, \theta_y) \quad (11.50d)$$

$$+ (\lambda(y, \theta_y) - \alpha(x, y, \theta_x, \theta_y)) f(x, y, \theta_x, -\theta_y) \quad (11.50e)$$

$$+ \alpha(x, y, \theta_x, \theta_y) f(x, y, -\theta_x, -\theta_y). \quad (11.50f)$$

For (11.50) to be the generator of a Markov process (in particular, for it to satisfy the positive maximum principle), the following inequalities have to be satisfied:

$$\lambda(x, \theta_x) + \lambda(y, \theta_y) - \alpha(x, y, \theta_x, \theta_y) \geq 0, \quad (11.51a)$$

$$\lambda(x, \theta_x) - \alpha(x, y, \theta_x, \theta_y) \geq 0, \quad (11.51b)$$

$$\lambda(y, \theta_y) - \alpha(x, y, \theta_x, \theta_y) \geq 0, \quad (11.51c)$$

$$\alpha(x, y, \theta_x, \theta_y) \geq 0, \quad (x, y, \theta_x, \theta_y) \in \bar{E}. \quad (11.51d)$$

These conditions can be interpreted as saying that the transition probabilities for the coupled piecewise deterministic Markov process cannot be negative. Clearly, the conditions (11.51) are equivalent to

$$0 \leq \alpha(x, y, \theta_x, \theta_y) \leq \min(\lambda(x, \theta_x), \lambda(y, \theta_y)), \quad (x, y, \theta_x, \theta_y) \in \bar{E}. \quad (11.52)$$

Let us briefly comment on the dynamical behaviour that the coupling operator (11.49) introduces. As can be seen from (11.50f), α is connected to ‘double flips’, i.e. the event that both particles change their directions at the same time. Setting α to either the lower or the upper bound in (11.52) will either discourage or encourage those double flips. As in the case of the overdamped and underdamped Langevin dynamics, the coupling behaviour (encoded in α) is allowed to depend on the point $(x, y, \theta_x, \theta_y) \in \bar{E}$. We also remark that the process generated by $\bar{\mathcal{L}}_\Gamma$ as in (11.50) can be simulated conveniently by using the methods summarised in [21, Appendix B].

Remark 11.7. *The construction in this section can be generalised to couplings of multiple zigzag processes in arbitrary dimensions by following a similar approach to the one taken in Section 11.1.*

12. ASYMPTOTIC VARIANCE AND OPTIMAL TRANSPORT

In this chapter we analyse the asymptotic variance associated to estimators based on coupled processes (Section 12.1) and connect the result to the theory of optimal transportation (Section 12.2).

12.1 A CENTRAL LIMIT THEOREM FOR COUPLED PROCESSES

The objective of this section is to establish a central limit theorem characterising the convergence in (9.6) and to find an expression for the associated asymptotic variance in terms of ergodic coupling operators $\Gamma \in \mathcal{G}^0$ and invariant measures $\bar{\pi}_\Gamma$. In particular, our aim is to compare between estimators based on couplings (as in (9.6)) and the one-particle estimators (9.4). Naturally, Assumption 10.2 is still in force. Moreover, let us assume the following:

Assumption 12.1 (Invertibility of the one-particle generators). *The generators $\mathcal{L}_i$ are invertible on $L_0^2(\pi_i)$, i.e. for all $f \in L_0^2(\pi_i)$ there exists $\phi \in \mathcal{D}(\mathcal{L}_i) \cap L_0^2(\pi_i)$ such that*

$$-\mathcal{L}_i \phi = f. \quad (12.1)$$

We refer to Section 1.2.2 for sufficient conditions ensuring the validity of the foregoing assumption, mentioning in particular exponential decay of the semigroups $(S_t^i)_{t \geq 0}$ in $L^2(\pi_i)$ as in (1.46). In the following, let us fix observables of interest $f_i \in L_0^2(\pi_i)$, recalling from the discussion in Section 1.2.2 that the assumption $\pi_i(f_i) = 0$ can be made without loss of generality. Denoting the corresponding solutions to the Poisson equations (12.1) by ϕ_i and supposing $X_0^i \sim \pi_i$, Theorem 1.4 implies the central limit theorems

$$\sqrt{T} \left(\frac{1}{T} \int_0^T f_i(X_t^i) dt - \pi_i(f_i) \right) \xrightarrow[T \rightarrow \infty]{d} \mathcal{N}(0, 2\sigma_{f_i}^2), \quad (12.2)$$

where the asymptotic variances are given by

$$\sigma_{f_i}^2 = \langle f_i, \phi_i \rangle_{L^2(\pi_i)}. \quad (12.3)$$

We will now establish a central limit theorem for the coupled process $(\bar{X}_t)_{t \geq 0}$ induced by

ergodic coupling operators $\Gamma \in \mathcal{G}^0$ and associated to extended observables of the form

$$F = \frac{1}{n} \sum_{i=1}^n f_i. \quad (12.4)$$

Theorem 12.1. Central limit theorem for coupled processes. *Let Assumption 12.1 be satisfied and assume that $\Gamma \in \mathcal{G}^0$. Furthermore, let $\bar{X}_0 \sim \bar{\pi}_\Gamma$. Then*

$$\sqrt{T} \left(\frac{1}{T} \int_0^T F(\bar{X}_t) dt - \bar{\pi}_\Gamma(F) \right) \xrightarrow[T \rightarrow \infty]{d} \mathcal{N}(0, 2\sigma_F^2). \quad (12.5)$$

The asymptotic variance σ_F^2 is given by

$$\sigma_F^2 = \frac{1}{n^2} \sum_{i=1}^n \sigma_{f_i}^2 - \int_{\bar{E}} \bar{\mathcal{L}}_0 \xi d\bar{\pi}_\Gamma, \quad (12.6)$$

where

$$\xi = \frac{1}{n^2} \sum_{i < j} \phi_i \phi_j, \quad (12.7)$$

the functions $(\phi_i)_{i=1}^n$ being the solutions to the Poisson equations (12.1).

Remark 12.1. Note that in the case of the trivial coupling $\Gamma = \mathbf{0}$ (see Example 10.1), we have that

$$\int_{\bar{E}} \bar{\mathcal{L}}_0 \xi d\bar{\pi}_0 = 0, \quad (12.8)$$

since $\bar{\pi}_0$ is the invariant measure associated to the process generated by $\bar{\mathcal{L}}_0$. Therefore in this case, we obtain the result that the asymptotic variance σ_F^2 is given by the arithmetic mean of the asymptotic variances $\sigma_{f_i}^2$ of the one-particle processes, divided by n . This result is expected, since the computational cost of computing the evolution of the processes $(X_t^i)_{t \geq 0}$ is likewise increased by a factor of n .

Remark 12.2. Observe furthermore that

$$\int_{\bar{E}} \bar{\mathcal{L}}_\Gamma \xi d\bar{\pi}_\Gamma = 0, \quad (12.9)$$

hence using $\bar{\mathcal{L}}_\Gamma = \bar{\mathcal{L}}_0 + \Gamma$ we can equivalently express the asymptotic variance as

$$\sigma_F^2 = \frac{1}{n^2} \sum_{i=1}^n \sigma_{f_i}^2 + \int_{\bar{E}} \Gamma \xi d\bar{\pi}_\Gamma. \quad (12.10)$$

Proof of Theorem 12.1. First observe that by the fact that $\bar{\pi}_\Gamma$ is a coupling of $(\pi_i)_{i=1}^n$, we have that $\bar{\pi}_\Gamma(F) = 0$. The Poisson equation

$$-\bar{\mathcal{L}}_\Gamma \Phi = F, \quad \bar{\pi}_\Gamma(\Phi) = 0, \quad (12.11)$$

has a solution given by

$$\Phi = \frac{1}{n} \sum_{i=1}^n \phi_i, \quad (12.12)$$

resting on the fact that $\Gamma\Phi = 0$ by the first condition of Definition 10.7. Again, the condition $\bar{\pi}_\Gamma(\Phi) = 0$ is satisfied by the coupling property of $\bar{\pi}_\Gamma$. Using Theorem 1.4, we see that the central limit theorem (12.5) holds with asymptotic variance

$$\sigma_F^2 = \langle F, \Phi \rangle_{L^2(\bar{\pi}_\Gamma)}. \quad (12.13)$$

Expanding the above yields

$$\begin{aligned} \sigma_F^2 &= \int_{\bar{E}} \left(\frac{1}{n} \sum_{i=1}^n f_i \right) \left(\frac{1}{n} \sum_{j=1}^n \phi_j \right) d\bar{\pi}_\Gamma = \frac{1}{n^2} \sum_{i=1}^n \int_{\bar{E}} f_i \phi_i d\bar{\pi}_\Gamma + \frac{1}{n^2} \sum_{\substack{i,j=1 \\ i \neq j}}^n \int_{\bar{E}} f_i \phi_j d\bar{\pi}_\Gamma \\ &= \frac{1}{n^2} \sum_{i=1}^n \sigma_{f_i}^2 - \int_{\bar{E}} \bar{\mathcal{L}}_0 \xi d\bar{\pi}_\Gamma, \end{aligned}$$

where in the last equation we used the fact that $\bar{\pi}_\Gamma$ has marginal π_i in the i -th coordinate, expression (12.3), as well as the definition of ξ in (12.7). $\square$

12.2 CONNECTIONS TO THE THEORY OF OPTIMAL TRANSPORTATION

In this section we will always assume that Assumption 12.1 is satisfied, so that the central limit theorems from the previous section hold. Theorem 12.1 then shows that, in order to reduce the asymptotic variance, we are led to the problem of minimizing the expression

$$\int_{\bar{E}} (-\bar{\mathcal{L}}_0 \xi) d\bar{\pi}_\Gamma.$$

Remarkably, this expression depends on Γ through the measure $\bar{\pi}_\Gamma$ only*. Since $\bar{\pi}_\Gamma$ has fixed marginals (i.e. they do not depend on Γ), this task is very reminiscent of the Kantorovich problem [142, Chapter 1] appearing in the theory of optimal transportation [140, 142]. To make this connection more precise, let us introduce the following terminology:

Definition 12.1 (Admissible couplings). *The set of couplings of the marginal invariant measures $(\pi_i)_{i=1}^n$ will be denoted by $\mathcal{C}$. A coupling $\bar{\pi} \in \mathcal{C}$ is called admissible, if it arises as the invariant measure of an ergodically coupled process, i.e. if there exists $\Gamma \in \mathcal{G}^0$ such that*

$$\int_{\bar{E}} \bar{\mathcal{L}}_\Gamma f d\bar{\pi} = 0, \quad (12.15)$$

*Another way of saying this is that the map $\Gamma \mapsto \sigma_F^2$ factors through the map $\Gamma \mapsto \bar{\pi}_\Gamma$, i.e. with respect to the asymptotic variance, no information is lost by considering only the invariant measure of the joint process.

for all $f \in \mathcal{D}(\bar{\mathcal{L}}_\Gamma)$. The set of admissible measures will be denoted by $\mathcal{C}^0$.

Let us first recall the formulation of the Kantorovich problem [142, Chapter 1]:

Problem 12.1 (Kantorovich). *For a fixed cost function $c \in C_b(\bar{E})$, find $\bar{\pi} \in \mathcal{C}$ such that*

$$\bar{\pi} \in \arg \min_{\bar{\pi} \in \mathcal{C}} \int_{\bar{E}} c \, d\bar{\pi}. \quad (12.16)$$

Our aim in this section can be summarised in the following form, only replacing $\mathcal{C}$ by the subset $\mathcal{C}^0$:

Problem 12.2. *For a fixed cost function $c \in C_b(\bar{E})$, find $\bar{\pi} \in \mathcal{C}^0$ such that*

$$\bar{\pi} \in \arg \min_{\bar{\pi} \in \mathcal{C}^0} \int_{\bar{E}} c \, d\bar{\pi}. \quad (12.17)$$

Equivalently, find minimisers of the function

$$\mathcal{G}^0 \rightarrow \mathbb{R}, \quad \Gamma \mapsto \int_{\bar{E}} c \, d\bar{\pi}_\Gamma. \quad (12.18)$$

Remark 12.3. *As already pointed out, setting $c = -\bar{\mathcal{L}}_0 \xi$, with ξ as defined in (12.7), is equivalent to the problem of optimising the asymptotic variance for a particular observable. Other choices for c might be of interest. For instance, one might aim to optimise the asymptotic variance across a set of observables simultaneously. In this case, it seems reasonable to consider cost functions of the form $c = \sum_j c_j$, where c_j is the cost function associated with the j th observable. Assuming that all the particles evolve in the same state space E , another natural objective would be to maximise the average distance of the particles at equilibrium, leading to a cost function of the type*

$$c(x_1, \dots, x_n) = - \sum_{\substack{i,j=1 \\ i \neq j}}^n d(x_i, x_j), \quad (12.19)$$

for some metric d on E . More generally, for some function $g : \mathbb{R} \rightarrow \mathbb{R}$ it might be worthwhile to consider

$$c(x_1, \dots, x_n) = - \sum_{\substack{i,j=1 \\ i \neq j}}^n g(d(x_i, x_j)). \quad (12.20)$$

A cost function of this type would be reasonable if one aims to use the empirical measure of an ensemble of particles in order to precondition the dynamics (see [88]), in which case the particles should neither be too close nor too far away from each other. We emphasize that since in our framework the marginal processes are held fixed, our results are not directly applicable to the algorithm presented in [88]. However, we expect that the results might be generalised to this context.

Let us also remark that the assumption $c \in C_b(\bar{E})$ is mainly for technical convenience and both the continuity and the boundedness assumption can be weakened. Since we are interested in situations where the process $(\bar{X})_{t \geq 0}$ takes values in a compact set with high probability (i.e. the target measures π_i are concentrated in a compact set), boundedness of c is not a severe restriction.

Remark 12.4. Clearly, it holds that

$$\inf_{\bar{\pi} \in \mathcal{C}^0} \int_{\bar{E}} c \, d\bar{\pi} \geq \inf_{\bar{\pi} \in \mathcal{C}} \int_{\bar{E}} c \, d\bar{\pi}, \quad (12.21)$$

so the solutions to the usual optimal transport problems provide lower bounds for Problem 12.2.

In the Kantorovich formulation, the cost function c is often induced by a distance (for instance $c(x, y) = (d(x, y))^p$ for $n = 2$, $1 \leq p < \infty$), penalising couplings that put probability mass on pairs of points (x, y) where x and y are far apart from each other (hence the name optimal transport). In the setting of MCMC (in particular in the context of variance reduction), it is plausible to encourage the particles to stay away from each other, leading to sample diversity and improved exploration of the state space. In this respect, our setting bears certain similarities with the use of optimal transport problems in functional density theory, see [33].

The set $\mathcal{C}^0$ depends on the generators $\mathcal{L}_i$. Furthermore, $\mathcal{C}^0$ is a strict subset of $\mathcal{C}$. The support of an ergodic invariant measure for a Markov process with continuous paths is necessarily connected, for instance, while in general the support of a coupling is not. The following example illustrates that $\mathcal{C}^0$ is indeed usually significantly smaller than $\mathcal{C}$:

Example 12.1 (The set $\mathcal{C}^0$ contains only few singular measures). Consider the setting from Section 11.1, i.e. two particles moving in one dimension according to overdamped Langevin dynamics. Let us fix a coupling $\Gamma \in \mathcal{G}^0$ and assume that the invariant measure $\bar{\pi}_\Gamma$ is supported on the zero set of a smooth function $H : \bar{E} \rightarrow \mathbb{R}$ with nowhere vanishing gradient, i.e.

$$\text{supp } \bar{\pi}_\Gamma \subseteq \{(x, y) \in \bar{E} : H(x, y) = 0\}. \quad (12.22)$$

This implies that $\bar{\pi}_\Gamma$ is supported on a submanifold of $\bar{E}$ and is hence necessarily singular with respect to the Lebesgue measure. Frequently, optimisers of standard optimal transport problems are of this type (see Theorem 12.2 below). Itô's formula implies that

$$H(\bar{X}_t) = H(\bar{X}_0) + \int_0^t (G^T \nabla H)(\bar{X}_s) \cdot dW_s + \int_0^t (\bar{\mathcal{L}}_\Gamma H)(\bar{X}_s) \, ds, \quad t \geq 0, \quad (12.23)$$

where G is given in (11.13). Choosing the initial condition $\bar{X}_0 \sim \bar{\pi}_\Gamma$ results in $H(\bar{X}_t) = H(\bar{X}_0) = 0$ almost surely, for all $t \geq 0$. It then follows that both of the remaining integral terms individually have to be zero (owing to the decomposition into martingale

and bounded variation part). The quadratic variation of the martingale part is given by

$$\int_0^t w(\bar{X}_s) ds, \quad w = (\partial_x H)^2 + 4 \cos \beta \sin \beta \cdot (\partial_x H)(\partial_y H) + (\partial_y H)^2. \quad (12.24)$$

Since the quadratic variation has to be zero for all $t \geq 0$, it follows that $\partial_x H = \pm \partial_y H$ on $\text{supp } \bar{\pi}_\Gamma$. Since H is smooth with nonvanishing gradient, it turns out that $\text{supp } \bar{\pi}_\Gamma$ is contained in either one of the diagonals $x = y$ or $x = -y$, in fact either $\bar{\pi}_\Gamma = \frac{1}{Z} e^{-V(x)} \delta_{x=y}(\text{d}x \text{d}y)$ or $\bar{\pi}_\Gamma = \frac{1}{Z} e^{-V(x)} \delta_{x=-y}(\text{d}x \text{d}y)$, noting that the latter is only possible if V has the symmetry property $V(x) = V(-x)$. We conclude that, at least in the example considered here, $\mathcal{C}^0$ contains only very few singular measures.

From the theory of optimal transportation it is known that solutions of the Kantorovich problem (Problem 12.1) are typically quite singular, in the sense that they are supported on small sets. Consider for instance the following version of [100, Theorem 1.2]:

Theorem 12.2. *Let $n = 2$, $E_1 = E_2 = \mathbb{R}^d$ and consider a cost function $c \in C^2(\mathbb{R}^{2d})$. Suppose that $\pi_1, \pi_2 \in \mathcal{P}(\mathbb{R}^d)$ are two compactly supported probability measures and that $\bar{\pi}$ is a solution of Problem 12.1. Let $(x_0, y_0) \in \mathbb{R}^{2d}$ be such that $\det(\nabla_{xy}^2 c(x_0, y_0)) \neq 0$, where $\nabla_{xy}^2 c$ denotes the matrix of mixed partial derivatives. Then there exists a neighbourhood $\mathcal{N}$ of (x_0, y_0) such that $\mathcal{N} \cap \text{supp } \bar{\pi}$ is contained in an n -dimensional Lipschitz submanifold of $\mathbb{R}^{2d}$. In particular, the restriction of $\bar{\pi}$ to $\mathcal{N}$ is singular with respect to the Lebesgue measure on $\mathbb{R}^{2d}$.*

As Example 12.1 shows, singular measures often do not belong to $\mathcal{C}^0$. The aim of this section is to show a similar singularity property for Problem 12.2. Informally speaking, we will see that under reasonable conditions, the optimisers of (12.18) are not attained for coupling operators in the interior of $\mathcal{G}^0$. To make this statement precise, let us fix the decompositions

$$\mathcal{L}_i = - \sum_{k=1}^{K_i} (A_k^i)^* A_k^i + \mathcal{T}^i \quad (12.25)$$

of the underlying generators (see Section 10.4) and lay the focus on coupling operators of the form (10.23), denoting this set by $\mathcal{G}(A)$:

$$\mathcal{G}(A) = \left\{ \Gamma = \sum_{(i,j,k,l) \in \mathcal{J}} \alpha_{ijkl} A_k^i A_l^j : \alpha_{ijkl} : \bar{E} \rightarrow \mathbb{R} \right\} \cap \mathcal{G}, \quad (12.26)$$

where we recall the set $\mathcal{J}$ of admissible indices, defined in (10.24). We wish to stress however that the distinction between $\mathcal{G}(A)$ and $\mathcal{G}$ is often obsolete (see Proposition 10.2). The subset of ergodic coupling operators will similarly be denoted by $\mathcal{G}^0(A)$. Let us now

introduce the ‘tangent space’ to $\mathcal{G}(A)$:

$$T\mathcal{G}(A) = \left\{ \sum_{(i,j,k,l) \in \mathcal{J}} \alpha_{ijkl} A_k^i A_l^j \mid \alpha_{ijkl} \in C_c^\infty(\bar{E}) \right\}. \quad (12.27)$$

Remark 12.5. *The definition of $T\mathcal{G}(A)$ encapsulates the first condition of Definition 10.7 in the sense that elements of $T\mathcal{G}(A)$ vanish on functions that only depend on one variable, whereas the second condition is not accounted for.*

In order to state our main result, we need the following definition:

Definition 12.2 (Interior points). *An operator $\Gamma \in \mathcal{G}^0(A)$ is called an interior point, if*

1. *for all $d\Gamma \in T\mathcal{G}(A)$ there exists $C > 0$ such that $\bar{\mathcal{L}}_\Gamma + \varepsilon d\Gamma \in \mathcal{G}^0(A)$ for all $\varepsilon \in (-C, C)$,*
2. *the operator $\bar{\mathcal{L}}_\Gamma$ is invertible on $L_0^2(\bar{\pi}_\Gamma)$,*
3. *$d\Gamma \mathcal{L}_\Gamma^{-1} f \in C(\bar{E})$ for all $d\Gamma \in T\mathcal{G}(A)$ and all $f \in C(\bar{E}) \cap L_0^2(\bar{\pi}_\Gamma)$.*

Remark 12.6. *The first condition is the essence of the foregoing definition, describing the geometric intuition of interior points. The third condition is mostly technical, since in applications $\mathcal{L}_\Gamma^{-1}$ usually possesses sufficient smoothing properties in order for the composition $d\Gamma \mathcal{L}_\Gamma^{-1}$ to preserve continuity.*

Remark 12.7 (Lyapunov functions). *It is possible and often convenient to replace the second condition by the weaker requirement that invertibility holds on a suitable subspace $\mathcal{V}_\Gamma$ of $L_0^2(\bar{\pi}_\Gamma)$. Our results in this section will then continue to hold, provided that the cost function c satisfies $c - \bar{\pi}_\Gamma(c) \in \mathcal{V}_\Gamma$ for all interior points Γ . As an example, assume that there exist Lyapunov functions $\mathcal{K}_i : E_i \rightarrow [1, \infty)$ for the one-particle dynamics, i.e.*

$$\mathcal{L}_i \mathcal{K}_i \leq -a_i \mathcal{K}_i + b_i, \quad (12.28)$$

for suitable constants $a_i > 0$, $b_i \geq 0$. Defining $\bar{\mathcal{K}} = \sum_{i=1}^n \mathcal{K}_i$, it follows immediately from $\Gamma \bar{\mathcal{K}} = 0$ that $\bar{\mathcal{K}}$ is a Lyapunov function for $\bar{\mathcal{L}}_\Gamma$, independently of the coupling operator Γ . Under certain minorisation (irreducibility) conditions (see [63], [91, Chapter 2.4]), one can show that $\mathcal{L}_\Gamma^{-1}$ is invertible on

$$\mathcal{V}_\Gamma = \left\{ f \in L_0^2(\bar{\pi}_\Gamma) : \left\| \frac{f}{\bar{\mathcal{K}}} \right\|_\infty < \infty \right\}. \quad (12.29)$$

See also [57, Theorem 3.2].

Example 12.2. In the setting of Section 11.1 (overdamped Langevin dynamics), it is straightforward to see that Γ as defined in (11.8) satisfies condition 1 of Definition 12.2

if and only if $-1 < \alpha(x, y) < 1$ for all $(x, y) \in \mathbb{R}^2$, i.e. if and only if the bound (11.9) is strict. More generally, Γ as defined in (11.19) satisfies condition 1 if and only if the matrix Q as defined in (11.20) is (strictly) positive definite pointwise. Those conditions are clearly equivalent to the (pointwise) ellipticity of the corresponding generators $\tilde{\mathcal{L}}_\Gamma$. In the case of underdamped Langevin dynamics (Section 11.2), analogous statements are valid. Similarly, couplings of zigzag processes (Section 11.3) satisfy condition 1 if and only if the bound (11.52) is strictly satisfied.

For our further discussion, we will need the following derivative formula:

Proposition 12.1. *Let $\Gamma \in \mathcal{G}^0(A)$ be an interior point, $d\Gamma \in T\mathcal{G}(A)$, and consider the family of operators $\tilde{\mathcal{L}}_\Gamma + \varepsilon d\Gamma \in \mathcal{G}^0$, for ε small enough. Let the associated family of invariant measures be denoted by $\bar{\pi}_\Gamma^\varepsilon$ and fix $c \in C_b(\bar{E})$. Then the function $\varepsilon \mapsto \int_{\bar{E}} c d\bar{\pi}_\Gamma^\varepsilon$ is differentiable in $\varepsilon = 0$, and the derivative is given by*

$$\left. \frac{d}{d\varepsilon} \right|_{\varepsilon=0} \left(\int_{\bar{E}} c d\bar{\pi}_\Gamma^\varepsilon \right) = - \int_{\bar{E}} c [\tilde{\mathcal{L}}_\Gamma^*]^{-1} (d\Gamma^* \mathbf{1}) d\bar{\pi}_\Gamma^0, \quad (12.30)$$

where the adjoints are taken in $L^2(\bar{\pi}_\Gamma^0)$.

Proof. The proof can be found in Section 16.3. □

Remark 12.8. *Notice that our notation entails that $\bar{\pi}_\Gamma \equiv \bar{\pi}_\Gamma^0$. Moreover, the right-hand side of (12.30) is well defined. Indeed, by the second condition in Definition 12.2, $[\tilde{\mathcal{L}}_\Gamma^*]^{-1}$ is well defined on $L_0^2(\bar{\pi}_\Gamma^0)$ and furthermore $\text{Ran } d\Gamma^* \subseteq L_0^2(\bar{\pi}_\Gamma^0)$ due to*

$$\int_{\bar{E}} d\Gamma^* f d\bar{\pi}_\Gamma^0 = \int_{\bar{E}} (d\Gamma \mathbf{1}) f d\bar{\pi}_\Gamma^0 = 0, \quad f \in C_c^\infty(\bar{E}), \quad (12.31)$$

using that $d\Gamma \mathbf{1} = 0$ according to Lemma 10.3.

For an interior point $\Gamma \in \mathcal{G}^0(A)$ and $d\Gamma \in T\mathcal{G}(A)$ let us introduce the suggestive notation

$$\frac{d}{d\Gamma} \int_{\bar{E}} c d\bar{\pi}_\Gamma := \left. \frac{d}{d\varepsilon} \right|_{\varepsilon=0} \left(\int_{\bar{E}} c d\bar{\pi}_\Gamma^\varepsilon \right), \quad (12.32)$$

as well as the following terminology:

Definition 12.3 (Critical points). *Let $\Gamma \in \mathcal{G}(A)$ be an interior point. Then Γ is called critical if*

$$\frac{d}{d\Gamma} \int_{\bar{E}} c d\bar{\pi}_\Gamma = 0 \quad (12.33)$$

for all $d\Gamma \in T\mathcal{G}(A)$.

In our aim to find minimisers of the function $\Gamma \mapsto \int_{\bar{E}} c d\bar{\pi}_\Gamma$, it is natural to seek critical points. The following is our main result in this section:

Theorem 12.3. *Let $c \in C_b(\bar{E})$. Then either all interior points are critical, or no interior point is critical.*

Example 12.3. Let c be of the form

$$c(x_1, \dots, x_n) = g_1(x_1) + \dots + g_n(x_n), \quad (12.34)$$

for appropriate functions $g_i : E_i \rightarrow \mathbb{R}$. Then, since $\bar{\pi}_\Gamma$ is a coupling of the fixed marginals $(\pi_i)_{i=1}^n$, the function $\Gamma \mapsto \int_{\bar{E}} c d\bar{\pi}_\Gamma = \sum_{i=1}^n \int_{E_i} g_i d\pi_i$ is constant, and hence all interior points are critical.

Before proceeding to the proof of the theorem, let us give a few remarks:

Remark 12.9. *Informally, Theorem 12.3 states that the mapping $\mathcal{G}^0(A) \ni \Gamma \mapsto \int_{\bar{E}} c d\bar{\pi}_\Gamma$ is either locally constant or does not attain its extrema on interior points. In other words, if $\Gamma \mapsto \int_{\bar{E}} c d\bar{\pi}_\Gamma$ is not constant, then its extrema lie ‘at the boundary’ of $\mathcal{G}^0(A)$, although we have not rigorously defined this term, and moreover, $\int_{\bar{E}} c d\bar{\pi}_\Gamma$ is not even well defined for nonergodic couplings.*

Remark 12.10. *A striking consequence of Theorem 12.3 is that under mild conditions, independent coupling (associated to $\mathbf{0} \in \mathcal{G}^0$) of overdamped or underdamped Langevin dynamics is not optimal for any criterion of the form $\int_{\bar{E}} c d\bar{\pi}_\Gamma$. Theorem 12.3 complements results from the theory of optimal transportation that state that optimal couplings are generically singular in terms of their support (see Theorem 12.2). Indeed, considering the example of overdamped or underdamped Langevin dynamics, the ‘boundary of $\mathcal{G}^0(A)$ ’ consists of couplings that lead to degenerately elliptic generators that are in general not hypoelliptic (see Definition 1.7 and Theorem 1.5). In particular, the corresponding invariant measures are not in general absolutely continuous with respect to the Lebesgue measure (compare this with Theorem 1.6).*

Remark 12.11. *Theorem 12.3 also supports the folklore that optimal Markov chain Monte Carlo samplers use as little noise as possible to guarantee ergodicity, as degenerately elliptic operators correspond to dynamics where noise only acts in certain directions (at a given point in space and time). This should also be contrasted with Theorems 2.1 and 2.2, as the process of making a reversible sampler nonreversible can be thought of informally as decreasing the ratio between random and deterministic behaviour. For a similar line of thought, see the discussion of the excess switching rate γ in Chapter 17.*

Remark 12.12. *Let us examine the function $\mathcal{G}^0(A) \ni \Gamma \mapsto \int_{\bar{E}} c d\bar{\pi}_\Gamma$ along a ray. More precisely, fix $d\Gamma \in T\mathcal{G}(A)$, set $\bar{\mathcal{L}}_\varepsilon := \bar{\mathcal{L}}_0 + \varepsilon d\Gamma$ for ε small enough, and consider the function $\varepsilon \mapsto \int_{\bar{E}} c d\bar{\pi}_\varepsilon$, where $(\bar{\pi}_\varepsilon)_\varepsilon$ denotes the corresponding family of invariant measures. Since $d\Gamma$ is relatively bounded with respect to $\bar{\mathcal{L}}_0$ in $L^2(\bar{\pi}_0)$, we have the following Neumann*

power expansion for ε small enough:

$$\int_{\bar{E}} c \, d\bar{\pi}_\varepsilon = \int_{\bar{E}} c \left(1 + \sum_{j=1}^{\infty} (-\varepsilon)^j [(\mathrm{d}\Gamma \bar{\mathcal{L}}_0^{-1})^*]^j \right) \mathbf{1} \, d\bar{\pi}_0. \quad (12.35)$$

For details, see [91, Theorem 5.2]. The factor of $(-\varepsilon)^j$ in expression (12.35) signals oscillatory behaviour, and indeed it is straightforward to construct examples (for instance in the Gaussian case), where (12.35) exhibits multiple local minima and maxima as a function of ε (see for instance the graph related to linear coupling in Figure 13.2c below). This finding is not in contradiction with Theorem 12.3. Indeed, as Theorem 12.3 shows, at those extrema there are directions of ascent (or descent) in $T\mathcal{G}(A)$ not aligned with the considered ray and thus, those extrema turn out not to be critical when considered in the whole of $\mathcal{G}^0(A)$.

Let us now prove Theorem 12.3 and start with the following key lemma. Its significance derives from the fact that the second statement manifestly does not depend on Γ .

Lemma 12.1. *Let $c \in C_b(\bar{E}) \cap L_0^2(\bar{\pi}_0)$ and $\Gamma \in \mathcal{G}^0(A)$ be an interior point. Then the following conditions are equivalent:*

1. *The following holds for all admissible indices $(i, j, k, l) \in \mathcal{J}$:*

$$A_k^i A_l^j \bar{\mathcal{L}}_\Gamma^{-1} (c - \bar{\pi}_\Gamma(c)) = 0. \quad (12.36)$$

2. *There exists $f \in \mathcal{D}(\bar{\mathcal{L}}_0)$ such that both of the following hold:*

- (a) *for all admissible indices $(i, j, k, l) \in \mathcal{J}$ it holds that*

$$A_k^i A_l^j f = 0, \quad (12.37)$$

- (b)

$$\bar{\mathcal{L}}_0 f = c. \quad (12.38)$$

Proof. First assume that (12.36) holds for all $(i, j, k, l) \in \mathcal{J}$. Then setting

$$f = \bar{\mathcal{L}}_\Gamma^{-1} (c - \bar{\pi}_\Gamma(c))$$

immediately implies (12.37). Furthermore, from (12.37) and (12.26) it follows that $\bar{\mathcal{L}}_\Gamma f = \bar{\mathcal{L}}_0 f$, implying

$$\bar{\mathcal{L}}_0 f = (c - \bar{\pi}_\Gamma(c)), \quad (12.39)$$

as well as $f \in \mathcal{D}(\bar{\mathcal{L}}_0)$. Equation (12.39) clearly implies that $\bar{\pi}_0(c) = \bar{\pi}_\Gamma(c)$, and hence $\bar{\pi}_\Gamma(c) = 0$, leading to (12.38).

The reverse implication follows similarly by first observing that (12.37) and (12.38) imply that $\tilde{\mathcal{L}}_\Gamma f = c$, and hence $\bar{\pi}_\Gamma(c) = 0$. Combining this with (12.37) shows that (12.36) holds. $\square$

Proof of Theorem 12.3. Clearly, we can without loss of generality assume that $\bar{\pi}_0(c) = 0$. According to Definition 12.3 and Proposition 12.1, an interior point $\Gamma \in \mathcal{G}^0(A)$ is critical if and only if

$$\int_{\bar{E}} c [\tilde{\mathcal{L}}_\Gamma^*]^{-1} (d\Gamma^* 1) d\bar{\pi}_\Gamma = 0, \quad (12.40)$$

for all $d\Gamma \in T\mathcal{G}(A)$, which is equivalent to

$$\sum_{(i,j,k,l) \in \mathcal{J}} \int_{\bar{E}} \left(A_k^i A_l^j \tilde{\mathcal{L}}_\Gamma^{-1} (c - \bar{\pi}_\Gamma(c)) \right) \cdot \alpha_{ijkl} d\bar{\pi}_\Gamma = 0, \quad (12.41)$$

for all $(\alpha_{ijkl})_{(i,j,k,l) \in \mathcal{J}} \subset C_c^\infty(\bar{E})$. The latter statement is clearly equivalent to the first statement in Lemma 12.1. The result now follows by noting that the second statement in Lemma 12.1 does not depend on Γ . $\square$

13. A PERTURBATIVE APPROACH FOR THE STUDY OF THE ASYMPTOTIC VARIANCE

Informally speaking, Theorem 12.3 shows that the objective of optimising the asymptotic variance σ_F^2 leads to the requirement that $\Gamma \in \mathcal{G}^0(A)$ should be chosen to be a ‘boundary point’ (see the discussion in Remark 12.9). While being an interesting theoretical result, it does not give much guidance about how to choose a suitable coupling in practice (after all, both minima and maxima are obtained ‘at the boundary’ of $\mathcal{G}^0(A)$). In this section we therefore develop a perturbative approach, based on operators of the form

$$\bar{\mathcal{L}}_\Gamma^\varepsilon = \bar{\mathcal{L}}_0 + \varepsilon d\Gamma, \quad d\Gamma \in T\mathcal{G}(A), \quad (13.1)$$

for ε small enough. In the following, we will assume that $\varepsilon \in \mathcal{I}$, where $\mathcal{I} \subset \mathbb{R}$ is an appropriate interval such that $\bar{\mathcal{L}}_\Gamma^\varepsilon \in \mathcal{G}^0$ for all $\varepsilon \in \mathcal{I}$. As usual, we consider observables of the form $F = \frac{1}{n} \sum_{i=1}^n f_i$, for some $f_i \in L_0^2(\pi_i)$, and suppose that Assumption 12.1 is satisfied. To stress the dependence of the asymptotic variance on the parameter ε we will write $\sigma_F^2(\varepsilon)$. Note that a similar setting has already been considered in Remark 12.12. There, we investigated the dependence of the asymptotic variance (or more generally, of the quantity $\int_{\bar{E}} c d\bar{\pi}_\varepsilon$) on the parameter ε . Here, we are rather interested in the choice of the ‘direction’ $d\Gamma \in T\mathcal{G}(A)$, starting from the trivial (independent) coupling $\bar{\mathcal{L}}_0$.

Combining the expression (12.6) with either (12.35) or (12.30) we see that

$$\frac{d}{d\varepsilon} \sigma_F^2 \Big|_{\varepsilon=0} = \int_{\bar{E}} d\Gamma \xi d\bar{\pi}_0, \quad (13.2)$$

where ξ is given by

$$\xi = \frac{1}{n^2} \sum_{i < j} \phi_i \phi_j, \quad (13.3)$$

in terms of the solutions to the Poisson equations (12.1). The benefit of (13.2) is that its right-hand side consists of expressions that are known in principle, as the measure $\bar{\pi}_0$ is given by the product $\bar{\pi}_0 = \bigotimes_{i=1}^n \pi_i$. It therefore serves as a starting point for finding a suitable coupling operator $\Gamma \in \mathcal{G}$. Let us summarise our approach in this section in the following form:

Problem 13.1. *Given invariant measures $\pi_i \in \mathcal{P}(E_i)$ and observables $f_i \in L_0^2(\pi_i)$, find a*

coupling operator $\Gamma \in \mathcal{G}$ such that

$$\int_{\bar{E}} \Gamma \xi \, d\bar{\pi}_0 \quad (13.4)$$

is minimised.

Problem 13.1 can be thought of in two different ways: Firstly, it can be interpreted as a linearisation of Problem 12.2. Indeed, (13.4) depends linearly on Γ , whereas (12.18) is highly nonlinear (for an illustration of this fact, see the power expansion (12.35)). Another way of seeing this is by noting the similarity between (13.4) and the second term appearing on the right-hand side of (12.10). Not surprisingly, Problem 13.1 turns out to be much easier to (approximately) solve in practice. Note that by linearity, properties similar to the one expressed in Theorem 12.3 hold for Problem 13.1 (at least if $\mathcal{G}$ is convex). Choosing a coupling $\Gamma \in \mathcal{G}$ according to the formulation of Problem 13.1 is clearly heuristic. However, we have had good results with it in numerical experiments (see below).

Secondly, when a solution of Problem 13.1 is available, it is reasonable in practice to only implement a small perturbation of the independent sampler (i.e. choose ε to be small). Such a choice will not be optimal over all couplings in $\mathcal{G}$ according to Theorem 12.3. However, it is then guaranteed that the performance of the sampler is at least slightly improved. Let us also note that the formulation of Problem 13.1 does not require the coupling to be ergodic, as opposed to the formulation of Problem 12.2.

The aim of this section is to analyse Problem 13.1 for some of the examples presented in Chapter 11 and to present some numerical experiments. To this end, let us introduce the shorthand notation

$$\delta\sigma_F^2(\Gamma) := \int_{\bar{E}} \Gamma \xi \, d\bar{\pi}_0, \quad (13.5)$$

stressing the infinitesimal (approximate) nature of the objective in Problem 13.1. In the sequel, Γ will be given in terms of a function α , belonging to a set $\mathcal{A}$. To emphasize this dependence we will write Γ_α . We will not impose regularity constraints on the function α (beyond measurability), so that the operators Γ_α will in general not induce couplings that satisfy the Feller property (see Remark 11.1) and hence strictly speaking do not belong to $\mathcal{G}$.

13.1 OVERDAMPED LANGEVIN DYNAMICS IN ONE DIMENSION WITH TWO PARTICLES

Consider the setting from the first example presented in Section 11.1. Then, (13.5) takes the form

$$\delta\sigma_F^2(\Gamma_\alpha) = \int_{\mathbb{R}^2} \alpha(x, y) \phi'(x) \phi'(y) e^{-(V(x)+V(y))} \, dx dy, \quad (13.6)$$

where ϕ is the solution to the Poisson equation

$$-(-V'\phi' + \phi'') = f, \quad \pi(\phi) = 0, \quad (13.7)$$

and $f \in L_0^2(\pi)$ is an observable of interest. Furthermore, Γ is given as in (11.8), with

$$\alpha \in \mathcal{A} := \{\alpha : \mathbb{R}^2 \rightarrow \mathbb{R} \mid \alpha \text{ measurable, } -1 \leq \alpha \leq 1\}. \quad (13.8)$$

Recall from Section 11.1 that $\alpha \in \mathcal{A}$ induces a well-defined coupled process (Lemma 11.1) that however does not satisfy the Feller property in general (further regularity assumptions would be required). The following optimality result is immediate from an inspection of (13.6):

Proposition 13.1. *Let $\alpha^* \in \mathcal{A}$ be given by*

$$\alpha^*(x, y) = \begin{cases} 1 & \text{if } \phi'(x)\phi'(y) \leq 0, \\ -1 & \text{if } \phi'(x)\phi'(y) > 0, \end{cases} \quad (13.9)$$

Then Γ_{α^} solves Problem 13.1 in the sense that $\delta\sigma_F^2(\Gamma_{\alpha^*}) \leq \delta\sigma_F^2(\Gamma_\alpha)$ for all $\alpha \in \mathcal{A}$.*

Remark 13.1. *Clearly, we have*

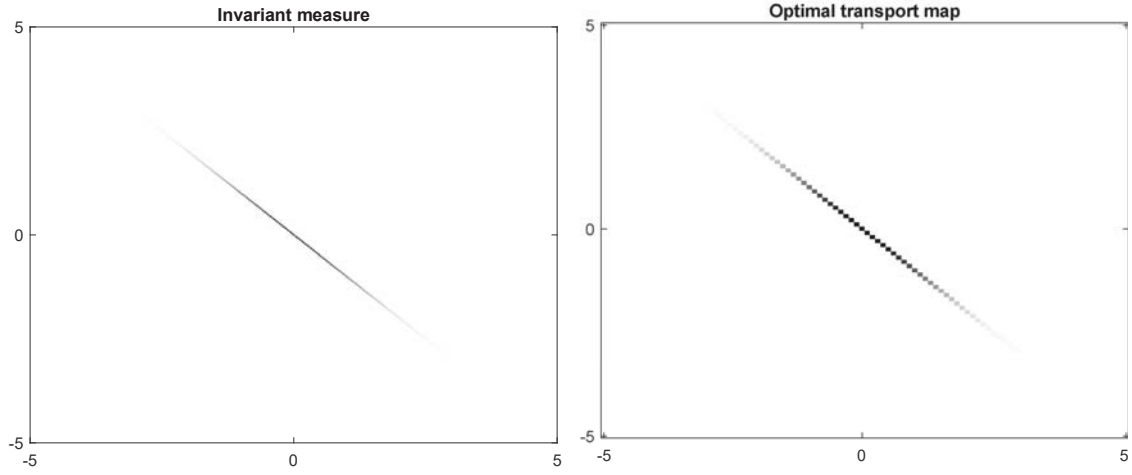
$$\int_{\bar{E}} \Gamma_{\alpha^*} \xi \, d\bar{\pi}_\Gamma \leq 0, \quad (13.10)$$

for α^ as defined in (13.9). We stress the difference between this expression and (13.6), where we compute the same integral, but with respect to $\bar{\pi}_0$. By comparison with (12.10), we see that $\sigma_F^2(\Gamma_{\alpha^*}) \leq \sigma_F^2(\mathbf{0})$, i.e. Γ_{α^*} always improves on independent coupling. However, we do not know whether Γ_{α^*} is optimal in the sense of Problem 12.2.*

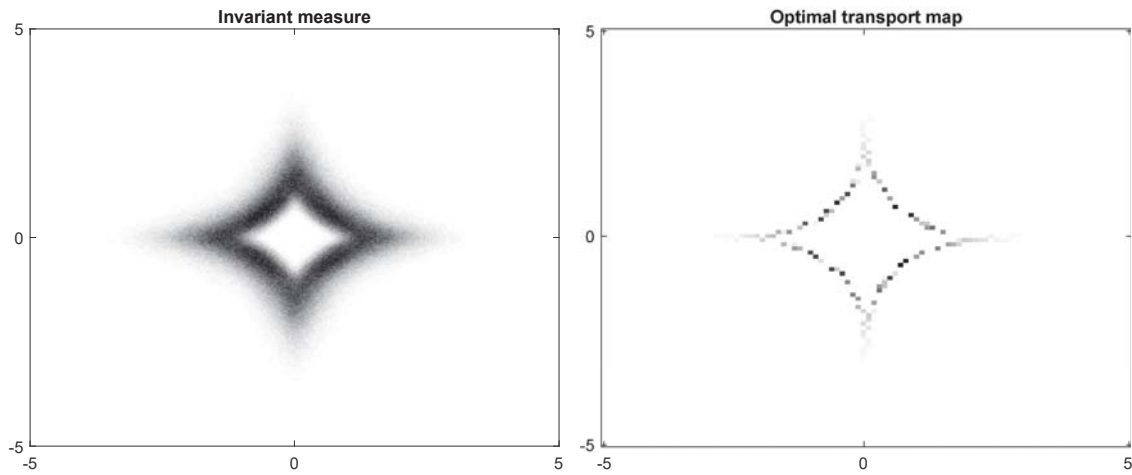
It is instructive to compare the solution of Problem 13.1 found in Proposition 13.1 to the solution of the Kantorovich problem (see Problem 12.1). Recall that Problem 13.1 can be considered to be a linearisation of Problem 12.2, which in turn is related to the Kantorovich problem in the sense that the minimisation is carried out over a smaller set of couplings (namely those couplings that are invariant measures of coupled processes, see Definition 12.1). For our experiments, we choose the quadratic potential $V(x) = \frac{1}{2}x^2$, i.e. the task of sampling from a Gaussian measure. Furthermore, we consider the linear observable $f_1(x) = x$, the quadratic observable $f_2(x) = x^2$ and the ‘mixed’ observable $f_3(x) = x^2 - x$. In Figure 13.1, we plot the invariant measure of the coupled processes induced by (13.9) (left-hand side) and compare them to the solutions of the Kantorovich problem* (right-hand side), with the appropriate cost function $c = -\bar{\mathcal{L}}_0\xi$ as given in Section 12.2. As it turns out, the solutions to Problem 13.1 and the standard Kantorovich problem look

*The optimal transport map was computed using the Python library POT 0.4.0 (accessible from <https://pypi.python.org/pypi/POT/0.4.0>) which is based on the algorithm proposed in [25].

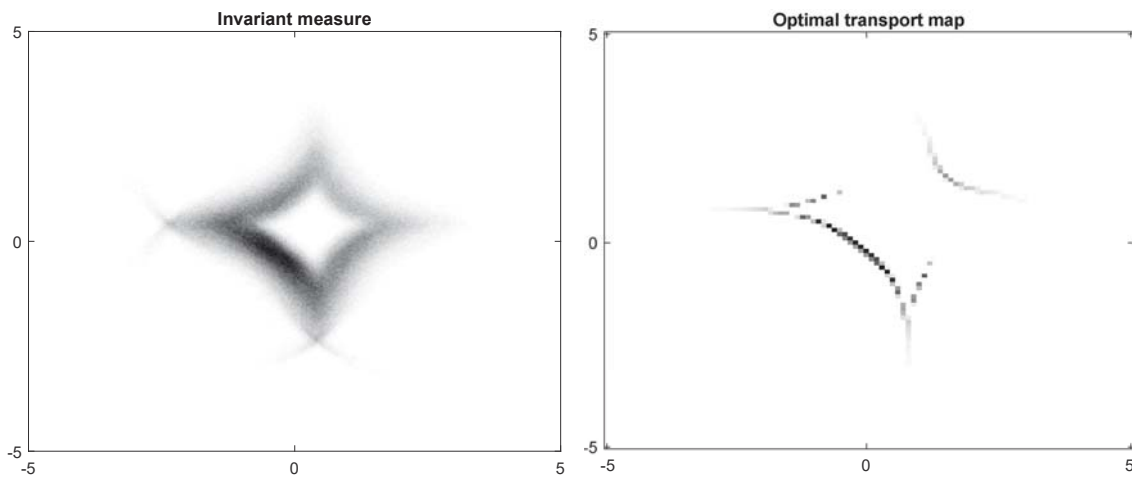
Figure 13.1: Comparison of invariant measures associated to solutions of Problem 13.1 and optimal transport maps as solutions to the Kantorovich problem for the example of overdamped Langevin dynamics in dimension one with quadratic potential.



(a) Linear observable: $f_1(x) = x$.



(b) Quadratic observable: $f_2(x) = x^2$.



(c) Mixed observable: $f_3(x) = x^2 - x$.

remarkably similar (at least in shape). We hence conclude that in the example considered here, Problem 13.1 is a good approximation of Problem 12.2, keeping in mind that the solution of the Kantorovich problem provides a lower bound for the objective function of Problem 12.2 (see Remark 12.4).

The following lemma serves to examine a few test cases and gain further intuition. For convenience, let us assume that f (and therefore, by elliptic regularity ϕ) are smooth.

Lemma 13.1. *Let $\phi \in L_0^2(\pi)$ solve the Poisson equation (13.7).*

1. *Assume that $f \in L_0^2(\pi)$ is monotonically increasing (decreasing). Then ϕ' is non-negative (nonpositive).*
2. *Assume that V and f are symmetric, i.e. $V(-x) = V(x)$ and $f(-x) = f(x)$ for all $x \in \mathbb{R}$. Furthermore, let f be monotonically decreasing (increasing) on $(-\infty, 0]$. Then $\phi'(x) \cdot x \leq 0$ ($\phi'(x) \cdot x \geq 0$) for all $x \in \mathbb{R}$.*

The proof can be found in Section 16.4. The following two corollaries are direct consequences of Lemma 13.1 and Proposition 13.1:

Corollary 13.1 (‘Mirror coupling’). *In the setting from the first part of Lemma 13.1,*

$$\alpha^* \equiv -1$$

solves Problem 13.1, in the sense that $\delta\sigma_F^2(\Gamma_{\alpha^}) \leq \delta\sigma_F^2(\Gamma_\alpha)$ for all $\alpha \in \mathcal{A}$.*

Corollary 13.2 (‘Symmetric coupling’). *In the setting from the second part of Lemma 13.1,*

$$\alpha^*(x, y) = \begin{cases} 1 & \text{if } x \cdot y \leq 0, \\ -1 & \text{if } x \cdot y > 0, \end{cases} \quad (13.11)$$

solves Problem 13.1, in the sense that $\delta\sigma_F^2(\Gamma_{\alpha^}) \leq \delta\sigma_F^2(\Gamma_\alpha)$ for all $\alpha \in \mathcal{A}$.*

A few comments on the findings from Corollaries 13.1 and 13.2 are in order. If the observable is monotone (Corollary 13.1), then it turns out that choosing the ‘mirror coupling’ $dB_t^x = -dB_t^y$ in (11.14) is optimal in the sense of Problem 13.1. This result has a clear connection to popular variance reduction techniques such as ‘antithetic variates’ [83, Chapter 9.2], where correlations between random variables are used to produce cancellations. In the case of symmetric observables (Corollary 13.2), optimal coupling in the sense of Problem 13.1 leads to a more sophisticated strategy: When the two particles (the locations of which are again denoted by x and y) ‘are on the same side of the potential’ (meaning that $x \geq 0$ and $y \geq 0$ or $x \leq 0$ and $y \leq 0$), then the Brownian motions should be coupled according to $dB_t^x = -dB_t^y$, as in the case of monotone observables. When the particles are on opposite sides ($x \geq 0$ and $y \leq 0$ or $x \leq 0$ and $y \geq 0$), according to Corollary 13.2 it is

best to switch to ‘synchronous coupling’, $dB_t^x = dB_t^y$. Intuitively this can be understood as follows: By symmetry, the situation where $x \geq 0$ and $y \leq 0$ with synchronous coupling ($dB_t^x = dB_t^y$) is equivalent to $x \geq 0$ and $y \geq 0$, with mirror coupling ($dB_t^x = -dB_t^y$). Since f is monotone on $[0, \infty)$, this argument provides a plausible explanation for optimality by appealing to Corollary 13.1. Finally, let us mention that numerical experiments show that using mirror coupling in the case of observables of the type encountered in Corollary 13.2 (‘naive antithetic variates’) actually leads to a less effective sampler in terms of the asymptotic variance (see Figure 13.2b).

Let us consider now the same set-up as in the numerical experiments presented in Figure 13.1, i.e. we consider a Gaussian target measure ($V(x) = \frac{1}{2}x^2$), and the observables $f_1(x) = x$ (‘linear’), $f_2(x) = x^2$ (‘quadratic’) and $f_3(x) = x^2 - x$ (‘mixed’). In Figure 13.2, we plot the asymptotic variances for f_1 , f_2 and f_3 , associated to different coupling schemes as a function of the coupling strength β . To be precise, the ‘Poisson’ coupling is defined by

$$\alpha(x, y) = \begin{cases} 2 \sin \beta \cos \beta & \text{if } \phi'(x)\phi'(y) \leq 0, \\ -2 \sin \beta \cos \beta & \text{if } \phi'(x)\phi'(y) > 0, \end{cases} \quad (13.12)$$

ϕ being the solution to the Poisson equation (13.7) for the corresponding observable, and $\beta \in [0, \frac{\pi}{4}]$ denoting the coupling strength[†]. For $\beta = 0$, we recover independent coupling, whereas $\beta = \frac{\pi}{4}$ leads to the optimal coupling from Proposition 13.1. According to Corollaries 13.1 and 13.2, Poisson coupling coincides with mirror coupling for f_1 and with symmetric coupling for f_2 . To illustrate the effect of couplings that are not tailored to the observable of interest, we also plot the asymptotic variances associated to symmetric coupling for f_1 , mirror coupling for f_2 , and both mirror and symmetric coupling for f_3 . For f_3 , we furthermore consider a coupling strategy that uses the derivative of the observable instead of the derivative of the solution to the Poisson equation, specifically, the coupling induced by

$$\alpha(x, y) = \begin{cases} 2 \sin \beta \cos \beta & \text{if } f'(x)f'(y) \leq 0, \\ -2 \sin \beta \cos \beta & \text{if } f'(x)f'(y) > 0. \end{cases} \quad (13.13)$$

The motivation for this is that in applications, the solution to the Poisson equation is often hard to obtain[‡], whereas the gradient of the observable is readily available. By integration by parts we have

$$\int_{\mathbb{R}^d} \nabla f \cdot \nabla \phi \, d\pi = \int_{\mathbb{R}^d} f^2 \, d\pi \geq 0, \quad (13.14)$$

suggesting us to use ∇f as a surrogate for $\nabla \phi$ (at equilibrium, the scalar product of ∇f and $\nabla \phi$ is positive on average).

[†]For $\beta \in [0, \frac{\pi}{4}]$, the function $2 \sin \beta \cos \beta$ is monotone, taking values in $[0, 1]$. We chose this parametrisation in order for it to be consistent with (11.13).

[‡]However, often one aims to approximate the solution to the Poisson equation in order to use it as a control variate, see for instance [39, 104, 130]. It suggests itself to use those approaches in conjunction with the coupling strategy developed here.

In all the cases considered, the Poisson coupling turns out to be the most efficient, uniformly in the coupling strength β . The fact that the absolute value of the derivative $\frac{d\sigma_F^2}{d\beta}|_{\beta=0}$ is maximal for Poisson coupling is precisely the content of Proposition 13.1, whereas the fact that the asymptotic variance for Poisson coupling is maximal at $\beta = 0$ follows from Remark 13.1. It is interesting to note the monotonicity of the asymptotic variance associated to Poisson coupling with respect to the coupling strength β ; this phenomenon is not covered by our theory. Importantly, the efficiency of a certain coupling strongly depends on the considered observable. Indeed, the mirror coupling (which is excellent for the linear observable, see Figure 13.2a) leads to an increase of the asymptotic variance for the quadratic observable (see Figure 13.2b). Similarly, the symmetric coupling (suited for the quadratic observable), does not improve the performance for the linear observable (but the performance is also not worsened). In Figure 13.2c, we observe that the coupling based on the derivative of the observable (see (13.13)) works almost as well as the Poisson coupling, so this might be a reasonable choice in applications, although further studies are needed. For a comment about the minimum of the gap associated to mirror coupling for the mixed observable (see Figure 13.2c) we refer to Remark 12.12.

13.2 OVERDAMPED LANGEVIN DYNAMICS WITH MULTIPLE PARTICLES IN ARBITRARY DIMENSIONS

Let us extend the discussion from the previous section to arbitrary dimensions, first considering the case of two particles (as done in Example 11.1). Using the expression (11.29), we see that

$$\delta\sigma_F^2(\Gamma_\alpha) = \int_{\mathbb{R}^d \times \mathbb{R}^d} (\nabla\phi(y) \cdot \alpha(x, y) \nabla\phi(x)) e^{-(V(x)+V(y))} dx dy, \quad (13.15)$$

where $\alpha : \mathbb{R}^d \times \mathbb{R}^d \rightarrow \mathbb{R}^{d \times d}$ is a matrix-valued function satisfying (11.28), i.e.

$$\alpha \in \mathcal{A} := \left\{ \alpha : \mathbb{R}^d \times \mathbb{R}^d \rightarrow \mathbb{R}^{d \times d} \text{ measurable, } \alpha(x, y)^T \alpha(x, y) \leq I_{d \times d} \text{ for all } x, y \in \mathbb{R}^d \right\}, \quad (13.16)$$

and ϕ is the solution to the Poisson equation

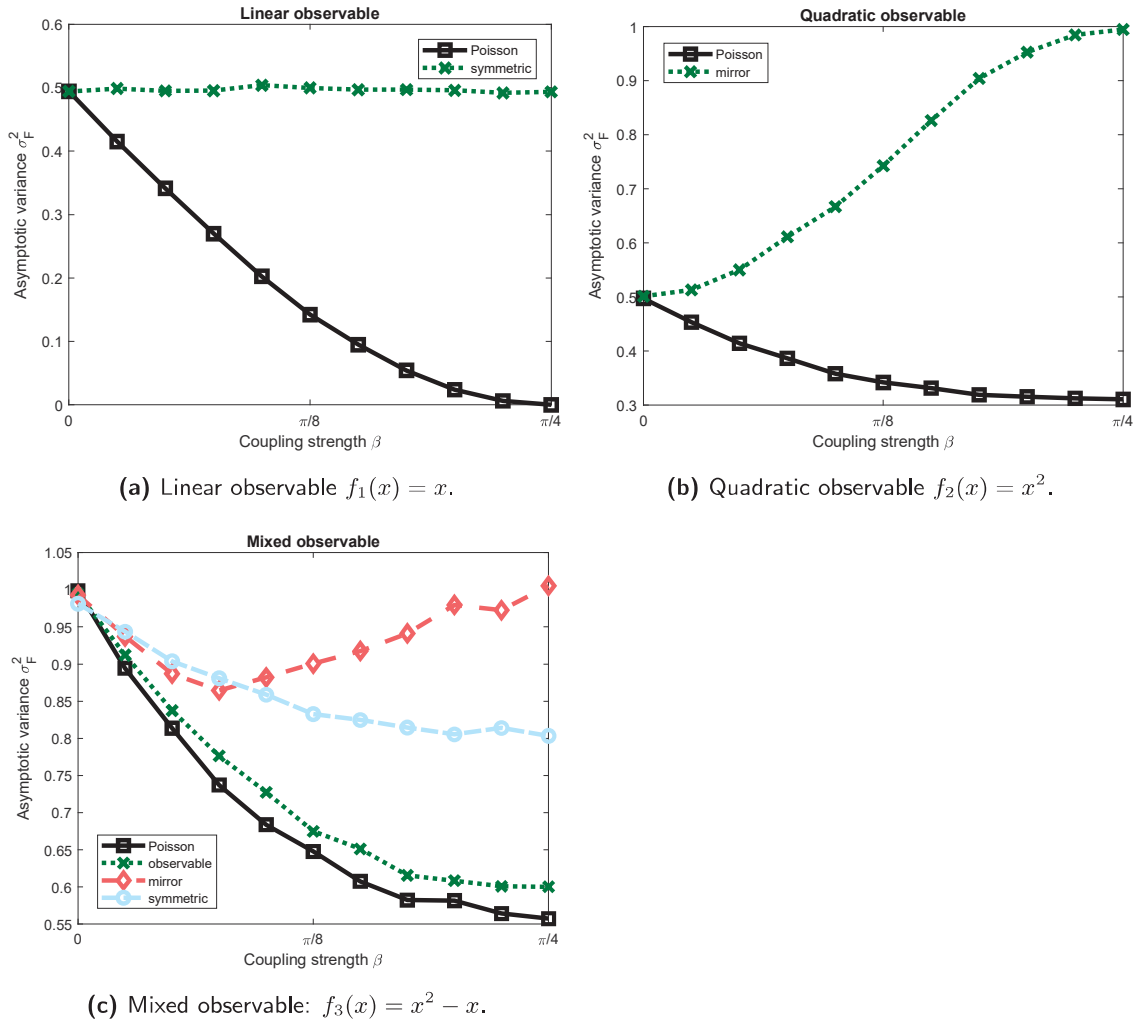
$$-(-\nabla V \cdot \nabla + \Delta)\phi = f, \quad \pi(\phi) = 0. \quad (13.17)$$

Since (11.28) implies

$$|\nabla\phi(y) \cdot \alpha(x, y) \nabla\phi(x)| \leq |\nabla\phi(x)| |\nabla\phi(y)|, \quad x, y \in \mathbb{R}, \quad (13.18)$$

we get the following optimality result:

Figure 13.2: Dependence of the asymptotic variance on different coupling schemes and coupling strengths.



Proposition 13.2. Assume that $\alpha^* \in \mathcal{A}$ is chosen such that

$$\nabla\phi(y) \cdot \alpha^*(x, y) \nabla\phi(x) = -|\nabla\phi(x)| |\nabla\phi(y)|, \quad x, y \in \mathbb{R}^d. \quad (13.19)$$

Then α^* solves Problem 13.1, i.e. $\delta\sigma_F^2(\Gamma_{\alpha^*}) \leq \delta\sigma_F^2(\Gamma_{\alpha})$ for all $\alpha \in \mathcal{A}$.

In the case when the solution ϕ to the Poisson equation is known it is straightforward to construct a matrix-valued function α^* such that both (11.28) and (13.19) are satisfied. For instance, any orthogonal matrix trivially satisfies (11.28), and (13.19) can be dealt with by choosing an appropriate rotation or reflection. As an example let us mention the following reflection in the plane spanned by $\nabla\phi(x)$ and $\nabla\phi(y)$:

$$\alpha^*(x, y) = \begin{cases} I_{d \times d} - 2 \frac{(\widehat{\nabla\phi(x)} + \widehat{\nabla\phi(y)}) (\widehat{\nabla\phi(x)} + \widehat{\nabla\phi(y)})^T}{(\widehat{\nabla\phi(x)} + \widehat{\nabla\phi(y)})^2} & \text{if } \nabla\phi(x) \neq 0, \nabla\phi(y) \neq 0, \\ & \widehat{\nabla\phi(x)} + \widehat{\nabla\phi(y)} \neq 0, \\ I_{d \times d} & \text{otherwise.} \end{cases} \quad (13.20)$$

Here, $\widehat{\nabla\phi} = \frac{\nabla\phi}{|\nabla\phi|}$ is used to denote the normalised gradient of ϕ . As mentioned in Section 13.1, the solution to the Poisson equation is usually hard to obtain in applications (but the popular methodology using control variates relies on approximations thereof). Inspired by the integration by parts formula (13.14), it seems reasonable to use the normalised gradient of the observable $\widehat{\nabla}f$ as a surrogate for $\widehat{\nabla\phi}$, i.e.

$$\alpha_f(x, y) = \begin{cases} I_{d \times d} - 2 \frac{(\widehat{\nabla f(x)} + \widehat{\nabla f(y)}) (\widehat{\nabla f(x)} + \widehat{\nabla f(y)})^T}{(\widehat{\nabla f(x)} + \widehat{\nabla f(y)})^2} & \text{if } \nabla f(x) \neq 0, \nabla f(y) \neq 0, \\ & \widehat{\nabla f(x)} + \widehat{\nabla f(y)} \neq 0, \\ I_{d \times d} & \text{otherwise.} \end{cases} \quad (13.21)$$

We recall that a comparison of the couplings associated to (13.20) and (13.21) was performed in the one-dimensional case (see Figure 13.2c) where α_f almost achieved the same reduction of the asymptotic variance as α^* . Based on (13.14) and numerical experiments we conjecture that choosing α_f guarantees an improvement in terms of the asymptotic variance for small perturbations:

Conjecture 13.1. Let $f \in L_0^2(\pi)$ and $\phi \in L_0^2(\pi)$ be the corresponding solution to the Poisson equation (13.17). Then

$$\delta\sigma_F^2(\Gamma_{\alpha}) = \int_{\mathbb{R}^d \times \mathbb{R}^d} (\nabla\phi(y) \cdot \alpha_f(x, y) \nabla\phi(x)) e^{-(V(x)+V(y))} dx dy \leq 0. \quad (13.22)$$

The complexity of the foregoing optimisation problems is increased substantially when considering more than two particles. From a practical perspective, it is desirable to specify

the coupling in terms of the matrix-valued function G appearing in (11.25) since this formulation is needed for the implementation of the numerical scheme. The linearised optimisation objective (Problem 13.1) however is formulated in terms of the coupling operator Γ . Passing from the latter to G involves the computationally expensive task of computing the square root of the matrix Q defined in (11.20). The construction and effective implementation of optimally coupled samplers with multiple particles therefore remains a subject for future work, but could be based on the results for the case of two particles. To give an impression, let us outline an idea based on the notation introduced in Remark 11.5. It is natural to choose the orthogonal matrices g_{ij} describing the coupling between the i th and the j th particle according to (13.20), i.e.

$$g_{ij}(x_1, \dots, x_n) = \begin{cases} I_{d \times d} - 2 \frac{(\widehat{\nabla \phi(x_i)} + \widehat{\nabla \phi(x_j)}) (\widehat{\nabla \phi(x_i)} + \widehat{\nabla \phi(x_j)})^T}{(\widehat{\nabla \phi(x_i)} + \widehat{\nabla \phi(x_j)})^2} & \text{if } \nabla \phi(x_i) \neq 0, \nabla \phi(x_j) \neq 0, \\ & \widehat{\nabla \phi(x_i)} + \widehat{\nabla \phi(x_j)} \neq 0, \\ I_{d \times d} & \text{otherwise,} \end{cases} \quad (13.23)$$

or, when the solution ϕ to the Poisson equation (or an approximation thereof) is not available, according to (13.21). Since the benefit of the coupling in terms of reducing the asymptotic variance is directly related to the value of the expression in (13.19), it is plausible to choose the weights w_{ij} (see (11.24)) in such a way that particle x_i is preferentially coupled to particle x_j if $|\nabla \phi(x_i)|$ and $|\nabla \phi(x_j)|$ are similar in magnitude. To make this precise, denote by $\sigma : \{1, \dots, n\} \rightarrow \{1, \dots, n\}$ the permutation that orders the particles according to $|\nabla \phi|$, i.e.

$$|\nabla \phi(x_{\sigma(1)})| \geq |\nabla \phi(x_{\sigma(2)})| \geq \dots \geq |\nabla \phi(x_{\sigma(n-1)})| \geq |\nabla \phi(x_{\sigma(n)})|. \quad (13.24)$$

Then, denoting the coupling strength by $\beta \in [0, \frac{\pi}{4}]$, we can set the weights as follows:

$$\begin{aligned} w_{ii} &= \cos \beta, & i &= 1, \dots, n, \\ w_{\sigma(1)\sigma(2)} &= \sin \beta, & w_{\sigma(3)\sigma(4)} &= \sin \beta, \quad \dots \quad w_{\sigma(n-1)\sigma(n)} = \sin \beta, \\ w_{ij} &= 0 & \text{otherwise.} \end{aligned} \quad (13.25)$$

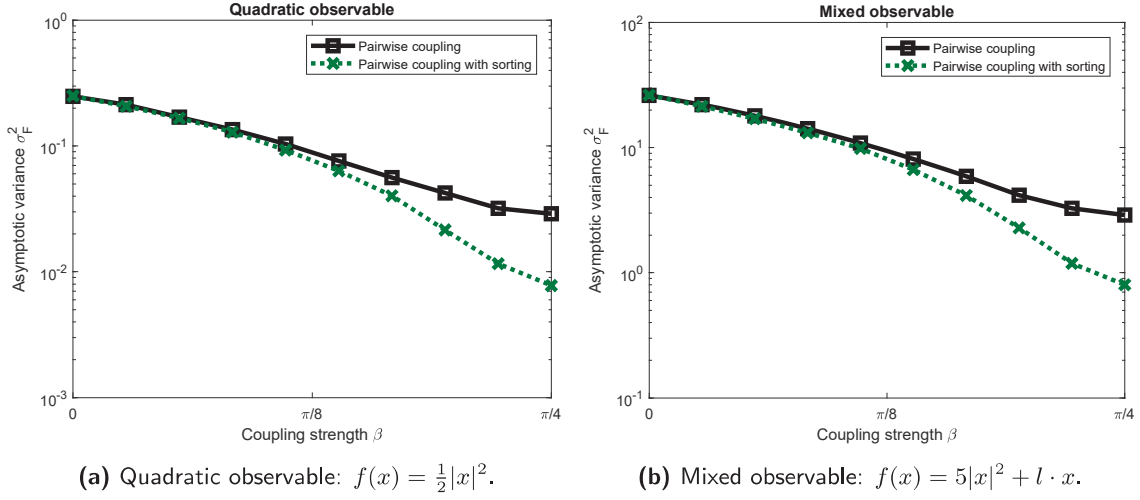
Let us emphasize that the sorting of the particles according to $|\nabla \phi|$ is supposed to be performed at every time step. We have compared this coupling strategy to simple pairwise coupling without sorting[§], i.e. replacing the second line of (13.25) by

$$w_{12} = \sin \beta, \quad w_{34} = \sin \beta, \quad \dots \quad w_{n-1,n} = \sin \beta, \quad (13.26)$$

for the example of sampling a standard Gaussian measure ($V = \frac{1}{2}|x|^2$) in $d = 10$ dimensions

[§]This is equivalent to running $n/2$ two-particle samplers independently in parallel.

Figure 13.3: Comparison between pairwise couplings with and without sorting according to $|\nabla\phi|$ for a Gaussian target measure in $d = 10$ dimensions and with $n = 10$ particles.



with $n = 10$ particles for the quadratic observable $f_1(x) = \frac{1}{2}|x|^2$ and the mixed observable $f_2(x) = 5|x|^2 + l \cdot x$, where $l = (1, \dots, 1)$. As Figure 13.3 shows, the sorting strategy as detailed in (13.25) leads to a smaller asymptotic variance in comparison to simple pairwise couplings.

13.3 THE ZIGZAG PROCESS

Recall the setting from Section 11.3 and fix an observable of interest $f \in L_0^2(\tilde{\pi})$ [¶]. For notational convenience, let us introduce the shorthands

$$\begin{aligned} \alpha_{++}(x, y) &:= \alpha(x, y, +1, +1), & \alpha_{+-}(x, y) &:= \alpha(x, y, +1, -1), \\ \alpha_{-+}(x, y) &:= \alpha(x, y, -1, +1), & \alpha_{--}(x, y) &:= \alpha(x, y, -1, -1). \end{aligned}$$

Taking the constraint (11.52) into account, we will optimise over the set

$$\mathcal{A} = \{\alpha : \mathbb{R}^2 \times \{-1, 1\}^2 \rightarrow \mathbb{R} \mid \alpha \text{ measurable, } 0 \leq \alpha \leq \min(\lambda(x, \theta_x), \lambda(y, \theta_y))\}.$$

The corresponding coupling operators (see (11.49)) will be denoted by Γ_α . We have the following lemma the proof of which can be found in Section 16.4.

[¶]We use the notation $\tilde{\pi}(dx) = \frac{1}{Z}e^{-V(x)}dx$ to distinguish it from the invariant measure π on the full space $\mathbb{R} \times \{-1, 1\}$, given in (11.43).

Lemma 13.2. *The zigzag process satisfies*

$$\begin{aligned}\delta\sigma_F^2(\Gamma_\alpha) &= \frac{1}{4} \int_{\mathbb{R}^2} \tilde{\alpha}(x, y) \cdot \tilde{\phi}'(x) \tilde{\phi}'(y) e^{-(V(x)+V(y))} dx dy, \\ \tilde{\alpha}(x, y) &= \alpha_{++}(x, y) + \alpha_{--}(x, y) - \alpha_{+-}(x, y) - \alpha_{-+}(x, y),\end{aligned}\tag{13.28}$$

where $\tilde{\phi} \in L^2(\tilde{\pi})$ is a solution to

$$-(-V'\tilde{\phi}' + \tilde{\phi}'') = f.\tag{13.29}$$

Remark 13.2. *Observe the remarkable coincidence that (13.29) coincides with the Poisson equation (13.7) for the overdamped Langevin dynamics. We employ the notation $\tilde{\phi}$ to distinguish (13.29) from the Poisson equation (16.25) in the whole space $\bar{E} = \mathbb{R}^2 \times \{-1, 1\}^2$.*

The following result is immediate from the expression (13.28):

Proposition 13.3. *Let $\alpha^* \in \mathcal{A}$ be given by*

$$\alpha^*(x, y, \theta_x, \theta_y) = \begin{cases} \min(\lambda(x, \theta_x), \lambda(y, \theta_y)), & \text{if } \tilde{\phi}'(x)\tilde{\phi}'(y)\theta_x\theta_y \leq 0 \\ 0, & \text{otherwise.} \end{cases}\tag{13.30}$$

Then Γ_{α^} solves Problem 13.1 in the sense that $\delta\sigma_F^2(\Gamma_{\alpha^*}) \leq \delta\sigma_F^2(\Gamma_\alpha)$ for all $\alpha \in \mathcal{A}$.*

Remark 13.3. *The comment from Remark 13.1 applies here as well.*

Combining Lemma 13.1 with Proposition 13.3 immediately yields the following corollaries:

Corollary 13.3. *In the setting from the first part of Lemma 13.1,*

$$\alpha^*(x, y, \theta_x, \theta_y) = \begin{cases} \min(\lambda(x, \theta_x), \lambda(y, \theta_y)), & \text{if } \theta_x\theta_y \leq 0 \\ 0, & \text{otherwise,} \end{cases}\tag{13.31}$$

solves Problem 13.1, in the sense that $\delta\sigma_F^2(\Gamma_{\alpha^}) \leq \delta\sigma_F^2(\Gamma_\alpha)$ for all $\alpha \in \mathcal{A}$.*

Corollary 13.4. *In the setting from the second part of Lemma 13.1,*

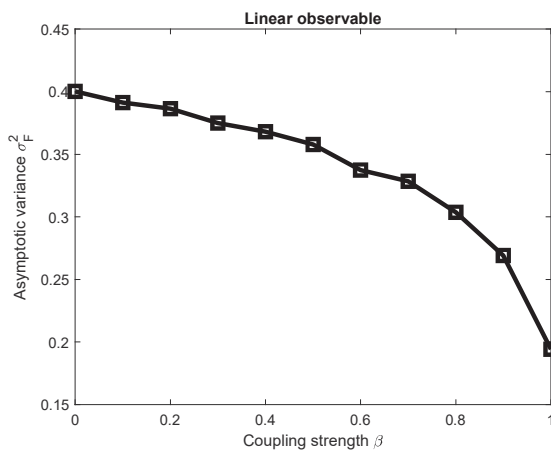
$$\alpha^*(x, y, \theta_x, \theta_y) = \begin{cases} \min(\lambda(x, \theta_x), \lambda(y, \theta_y)), & \text{if } xy\theta_x\theta_y \leq 0 \\ 0, & \text{otherwise,} \end{cases}\tag{13.32}$$

solves Problem 13.1, in the sense that $\delta\sigma_F^2(\Gamma_{\alpha^}) \leq \delta\sigma_F^2(\Gamma_\alpha)$ for all $\alpha \in \mathcal{A}$.*

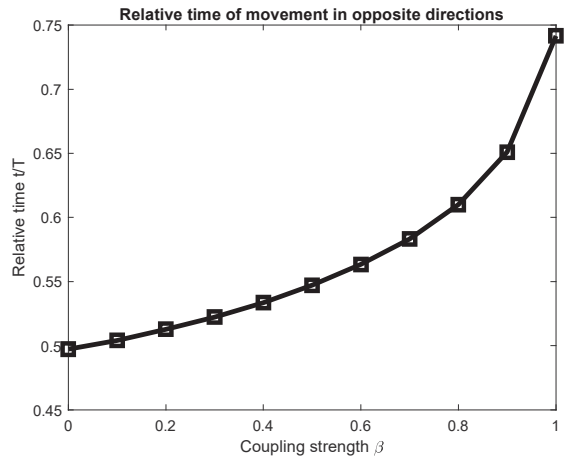
The results from Corollaries 13.3 and 13.4 can be interpreted intuitively in the following way. As already pointed out in Section 11.3, setting $\alpha(x, y, \theta_x, \theta_y) = \min(\lambda(x, \theta_x), \lambda(y, \theta_y))$ encourages simultaneous flips of the velocities θ_x and θ_y (when the particles are at locations x and y , with velocities θ_x and θ_y), whereas the flips occur independently if

$\alpha(x, y, \theta_x, \theta_y) = 0$. The coupling associated to (13.31) therefore leads to an increased probability of simultaneous flips precisely when the two particles move in opposite directions. Observe that simultaneous flips preserve the value of $\theta_x \theta_y$, while single flips change its sign. As a consequence, the relative amount of time during which the two particles move in opposite directions is increased by the coupling associated to (13.31). Similarly to the case of mirror coupling for overdamped Langevin diffusions (see the discussion following Corollary 13.2), it is plausible that this dynamics leads to cancellations for monotone observables in the spirit of antithetic variates. The interpretation of Corollary 13.4 is analogous to the one of Corollary 13.2. For illustration, we consider again the case of a quadratic potential $V(x) = \frac{1}{2}x^2$ (i.e. a Gaussian target measure) and a linear observable $f(x) = x$. The coupling is chosen according to Corollary 13.3, i.e. in a suitable manner for the linear observable, modulated by a parameter $\beta \in [0, 1]$, analogously to (13.12) and (13.13). In Figures 13.4a, 13.4b, and 13.4c we plot the associated asymptotic variance, the relative time the particles move in opposite directions, as well as the average distance between the particles. Those graphs support the foregoing intuitive arguments. The fact that the average distance between the particles increases with the strength of the coupling is interesting, since it suggests that the state space can be explored more efficiently by using appropriate couplings. In Figure 13.4d we plot a typical trajectory of the joint system. Comparing this graph with the optimal transport map depicted in 13.1a, we conclude that the solution to Problem 13.1 found in Proposition 13.3 is somewhat close to the solution of the Kantorovich problem, but not nearly as much as the corresponding solution in the case of overdamped Langevin dynamics. Interestingly, the aforementioned similarity is much more pronounced in the case when the target distribution is heavy-tailed. As an example, we plotted a typical trajectory of a mirror-coupled zigzag process targeting a Cauchy distribution in Figure 13.4e. We did perform numerical experiments for quadratic observables. For them, an improvement in the asymptotic variance is hardly noticeable. Furthermore, a typical trajectory for the coupling induced by (13.32) very much resembles the typical trajectories for the independent coupling. As it seems, couplings of zigzag processes are not very efficient in the setting of Lemma 13.1.2. A possible explanation is that piecewise deterministic Markov processes are more ‘rigid’ than diffusions (in fact, by definition, they move deterministically during a considerable time span), allowing less flexibility in terms of couplings.

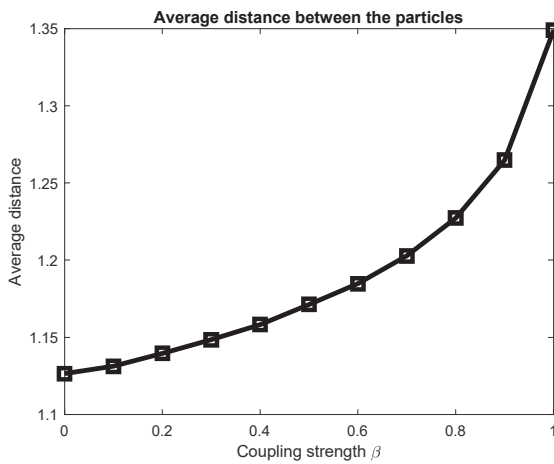
Figure 13.4: Coupling for the zigzag process according to Corollary 13.3.



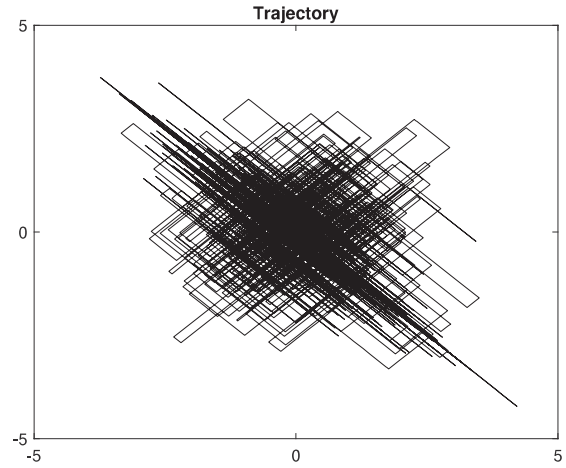
(a) Asymptotic variance associated to the linear observable $f(x) = x$, depending on the coupling strength β .



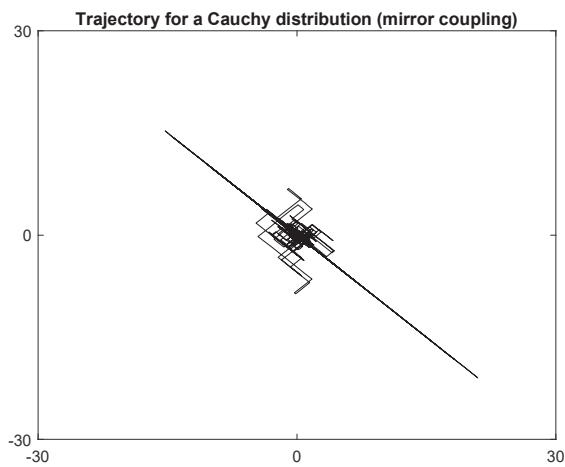
(b) Relative amount of time that the particles move in opposite directions, depending on the coupling strength β .



(c) Average distance between the two particles, depending on the coupling strength β .



(d) Part of the trajectory for a Gaussian target with mirror coupling.



(e) Part of the trajectory for a Cauchy target with mirror coupling.

14. A REMARK ON THE RATE OF CONVERGENCE TO EQUILIBRIUM

In this chapter, we study the rate of convergence to equilibrium for coupled processes. For convenience, let us assume that the spaces E_i and the operators $\mathcal{L}_i$ are identical copies of each other. When addressing the marginal process(es), we will usually suppress the indices and write $\mathcal{L}$ and E . Furthermore, let us fix an ergodic coupling operator $\Gamma \in \mathcal{G}^0$ and denote as usual the corresponding generator and semigroup by $\tilde{\mathcal{L}}_\Gamma = \tilde{\mathcal{L}}_0 + \Gamma$ and $(\tilde{S}_t^\Gamma)_{t \geq 0}$ respectively. In the sequel, we will make use of the following subspace of centred observables in $L_0^2(\bar{\pi}_\Gamma)$:

$$\tilde{L}_0^2(\bar{\pi}_\Gamma) := \left\{ F \in L_0^2(\bar{\pi}_\Gamma) \mid \text{there exists } f \in L_0^2(\pi) \text{ such that } F = \frac{1}{n} \sum_{i=1}^n f_i \right\} \subseteq L_0^2(\bar{\pi}_\Gamma). \quad (14.1)$$

Clearly, the space $\tilde{L}_0^2(\bar{\pi}_\Gamma)$ comprises the observables of interest in our framework. By using the extension operator Π^* from (9.7), $\tilde{L}_0^2(\bar{\pi}_\Gamma)$ can equivalently be defined via $\tilde{L}_0^2(\bar{\pi}_\Gamma) = \Pi^* L_0^2(\bar{\pi})$. The main result of this chapter is the following characterisation of exponential convergence to equilibrium.

Theorem 14.1. *For $\Lambda > 0$, the following are equivalent:*

1. Poincaré inequality:

$$\langle F, F \rangle_{L^2(\bar{\pi}_\Gamma)} \leq \frac{1}{\Lambda} \langle F, (-\tilde{\mathcal{L}}_\Gamma) F \rangle_{L^2(\bar{\pi}_\Gamma)} \quad (14.2)$$

for all $F \in \mathcal{D}(\tilde{\mathcal{L}}_\Gamma) \cap \tilde{L}_0^2(\bar{\pi}_\Gamma)$.

2. Exponential decay:

$$\|\tilde{S}_t^\Gamma F\|_{L^2(\bar{\pi}_\Gamma)}^2 \leq e^{-2\Lambda t} \|F\|_{L^2(\bar{\pi}_\Gamma)}^2, \quad t \geq 0, \quad (14.3)$$

for all $F \in \tilde{L}_0^2(\bar{\pi}_\Gamma)$.

Remark 14.1. *Theorem 14.1 is well known if $\tilde{L}_0^2(\bar{\pi}_\Gamma)$ is replaced by the whole space $L_0^2(\bar{\pi}_\Gamma)$, see Theorem 1.3. For our purposes however, it is natural to restrict attention to the smaller space $\tilde{L}_0^2(\bar{\pi}_\Gamma)$. In particular, by the duality $(\Pi\bar{\pi})(f) = \bar{\pi}(\Pi^*f)$ explained in Chapter 9, the decay estimate (14.3) implies exponential convergence of the laws $(\Pi\bar{\pi}_t)_{t \geq 0}$, with the same rate.*

The proof of Theorem 14.1 relies on the following lemma:

Lemma 14.1. *Let $F \in \tilde{L}_0^2(\bar{\pi}_\Gamma)$ with $F = \frac{1}{n} \sum_{i=1}^n f_i$, $f_i \in L_0^2(\pi)$. Then*

$$\bar{S}_t^\Gamma F = \frac{1}{n} \sum_{i=1}^n (S_t f)_i, \quad t \geq 0. \quad (14.4)$$

In particular, $\tilde{L}_0^2(\bar{\pi}_\Gamma)$ is invariant under the flow of $(\bar{S}_t^\Gamma)_{t \geq 0}$.

Proof. For $F \in C_b(\bar{E}) \cap \tilde{L}_0^2(\bar{\pi}_\Gamma)$ we have

$$\begin{aligned} (\bar{S}_t^\Gamma F)(x_1, \dots, x_n) &= \mathbb{E}[F(\bar{X}_t) | \bar{X}_0 = (x_1, \dots, x_n)] \\ &= \frac{1}{n} \sum_{i=1}^n \mathbb{E}[f(X_t^i) | \bar{X}_0 = (x_1, \dots, x_n)] \\ &= \frac{1}{n} \sum_{i=1}^n \mathbb{E}[f(X_t^i) | X_0^i = x_i] = \frac{1}{n} \sum_{i=1}^n (S_t f)(x_i). \end{aligned}$$

Between the second and the third line, we used the fact that the process $(\bar{X}_t)_{t \geq 0}$ has $(X_t^i)_{t \geq 0}$ as its i th marginal, so in particular, the law of $f(X_t^i)$ depends on the initial condition $\bar{X}_0$ only through $X_0^i = x_i$. For arbitrary $F \in \tilde{L}_0^2(\bar{\pi}_\Gamma)$ the result follows by a standard density argument. $\square$

Proof of Theorem 14.1. The proof is verbatim the same as for the usual result [9, Theorem 1.3], [10, Theorem 4.2.5]. However, the fact that $\tilde{L}_0^2(\bar{\pi}_\Gamma)$ is invariant under the flow $(\bar{S}_t^\Gamma)_{t \geq 0}$ is crucial. For completeness let us sketch the proof: Let $F \in \tilde{L}_0^2(\bar{\pi}_\Gamma)$ and assume that the Poincaré inequality (14.2) holds for some constant $\Lambda > 0$. Then

$$\frac{d}{dt} \left(\frac{1}{2} \|\bar{S}_t^\Gamma F\|_{L^2(\bar{\pi}_\Gamma)}^2 \right) = \langle \bar{S}_t^\Gamma F, \bar{\mathcal{L}}_\Gamma \bar{S}_t^\Gamma F \rangle_{L^2(\bar{\pi}_\Gamma)} \leq -\Lambda \langle \bar{S}_t^\Gamma F, \bar{S}_t^\Gamma F \rangle_{L^2(\bar{\pi}_\Gamma)}, \quad (14.6)$$

where the last inequality uses the fact that $\bar{S}_t^\Gamma f \in \tilde{L}_0^2(\bar{\pi}_\Gamma)$ according to Lemma 14.1. Exponential decay as in (14.3) follows by Gronwall's Lemma. The converse direction follows by performing a Taylor expansion of the decay estimate (14.3) around $t = 0$. $\square$

To explain the significance of Theorem 14.1, let us start by writing (14.2) in the form

$$\langle f, f \rangle_{L^2(\pi)} + \frac{1}{n} \sum_{\substack{i,j=1 \\ i \neq j}}^n \langle f_i, f_j \rangle_{L^2(\bar{\pi}_\Gamma)} \leq \frac{1}{\Lambda} \left(\langle f, (-\mathcal{L})f \rangle_{L^2(\pi)} + \frac{1}{n} \sum_{\substack{i,j=1 \\ i \neq j}}^n \langle f_i, (-\mathcal{L}_j)f_j \rangle_{L^2(\bar{\pi}_\Gamma)} \right), \quad (14.7)$$

using the marginal property of $\bar{\pi}_\Gamma$. Clearly, (14.7) deviates from the usual one-particle Poincaré inequality (1.51) by the additional terms involving summation over pairs of particles. To make this more precise and analyse the impact of these terms, let us define the

following bilinear form on $L_0^2(\pi)$:

$$\langle\langle f, g \rangle\rangle := \langle f, g \rangle_{L^2(\pi)} + \frac{1}{n} \sum_{\substack{i,j=1 \\ i \neq j}}^n \langle f_i, g_j \rangle_{L^2(\bar{\pi}_\Gamma)}, \quad f, g \in L_0^2(\pi). \quad (14.8)$$

For $F = \sum_{i=1}^n f_i$ and $G = \sum_{i=1}^n g_i$ we have that $\langle F, G \rangle_{L^2(\bar{\pi}_\Gamma)} = n \langle\langle f, g \rangle\rangle$. Hence, $\langle\langle \cdot, \cdot \rangle\rangle$ is both symmetric and nonnegative definite, but $\langle\langle f, f \rangle\rangle = 0$ is possible for $f \neq 0$. It is therefore natural to define the equivalence relation

$$f \sim g : \Longleftrightarrow \langle\langle f - g, f - g \rangle\rangle = 0, \quad f, g \in L_0^2(\pi), \quad (14.9)$$

and the corresponding Hilbert space

$$\mathcal{H} := L_0^2(\pi) / \sim. \quad (14.10)$$

Using again the correspondence $F = \sum_{i=1}^n f_i$, we see that $\langle\langle f, f \rangle\rangle = 0$ if and only if $F = 0$ $\bar{\pi}_\Gamma$ -almost surely. By ergodicity, this is also equivalent to $\bar{\mathcal{L}}_\Gamma F = 0$, $\bar{\pi}_\Gamma$ -almost surely. We hence see that $\mathcal{L}$ respects $\sim$ -equivalence classes, i.e. $f \sim g$ if and only if $\mathcal{L}f \sim \mathcal{L}g$. Denoting the induced operator on $\mathcal{H}$ by $\mathcal{L}_\mathcal{H}$, it is then immediate each of (14.2) and (14.7) is equivalent to

$$\langle\langle f, f \rangle\rangle \leq \frac{1}{\Lambda} \langle\langle f, (-\mathcal{L}_\mathcal{H})f \rangle\rangle, \quad f \in \mathcal{H} \cap \mathcal{D}(\mathcal{L}). \quad (14.11)$$

By its similarity to the one-particle Poincaré inequality (1.51), the formulation (14.11) is convenient for the comparison between the spectral gaps of the underlying and the coupled dynamics.

Let us assume from now on that $\mathcal{L}$ is self-adjoint in $L_0^2(\pi)$ with discrete spectrum, with

$$-\mathcal{L}e_i = \mu_i e_i, \quad 0 < \mu_1 \leq \mu_2 \leq \dots \quad (14.12)$$

where the eigenvectors $(e_i)_{i \in \mathbb{N}}$ form an orthonormal basis in $L_0^2(\pi)$. The optimal constant in the one-particle Poincaré inequality (1.51) is then clearly given by $\lambda = \mu_1$. In the study of the coupled Poincaré inequality (14.11), two interesting effects might occur. Firstly, the spectrum of $\mathcal{L}_\mathcal{H}$ might be different from the spectrum of $\mathcal{L}$. Secondly, $\mathcal{L}_\mathcal{H}$ might not be symmetric with respect to $\langle\langle \cdot, \cdot \rangle\rangle$. Let us start with the first point. Clearly, $\sigma(\mathcal{L}_\mathcal{H}) \subseteq \sigma(\mathcal{L})$, more precisely

$$\sigma(\mathcal{L}_\mathcal{H}) = \{\lambda_i : \langle\langle e_i, e_i \rangle\rangle \neq 0\}. \quad (14.13)$$

Example 14.1. Consider the dynamics

$$dX_t = -\nabla V(X_t) dt + \sqrt{2} dB_t, \quad (14.14a)$$

$$dY_t = -\nabla V(Y_t) dt - \sqrt{2} dB_t, \quad (14.14b)$$

with a standard $\mathbb{R}^d$ -valued Brownian motion $(B_t)_{t \geq 0}$. Let us assume that the potential V grows sufficiently fast at infinity such that the one-particle generator $\mathcal{L}$ has compact resolvent and hence discrete spectrum in $L_0^2(\pi)$. Furthermore, suppose that the eigenvalues and eigenfunctions are labelled and ordered as in (14.12). Let us now assume that V is even, i.e. $V(x) = V(-x)$, and that the process is ergodic. The invariant measure is then given by

$$\bar{\pi}_\Gamma(\mathrm{d}x\mathrm{d}y) = \frac{1}{Z} e^{-V(x)} \delta_{x+y} \mathrm{d}x\mathrm{d}y, \quad (14.15)$$

and the corresponding new (degenerate) scalar product in $L_0^2(\pi)$ turns out to be

$$\langle\langle f, g \rangle\rangle = \frac{1}{Z} \int_{\mathbb{R}^d} f(x)g(x)e^{-V(x)} \mathrm{d}x \quad (14.16a)$$

$$+ \frac{1}{2Z} \left(\int_{\mathbb{R}^d} f(x)g(-x)e^{-V(x)} \mathrm{d}x + \int_{\mathbb{R}^d} f(-x)g(x)e^{-V(x)} \mathrm{d}x \right). \quad (14.16b)$$

Notice that by the symmetry of V , all the eigenfunctions of $\mathcal{L}$ are either even or odd. Moreover, a short calculation shows that $\langle\langle f, f \rangle\rangle = 0$ if and only if f is odd (meaning that $-f(x) = f(-x)$). Using (14.13), we see that

$$\sigma(\mathcal{L}_\mathcal{H}) = \{\lambda_i : e_i \text{ is odd}\}. \quad (14.17)$$

Another short calculation shows that $\mathcal{L}_\mathcal{H}$ is symmetric with respect to $\langle\langle \cdot, \cdot \rangle\rangle$, i.e.

$$\langle\langle f, \mathcal{L}_\mathcal{H}g \rangle\rangle = \langle\langle \mathcal{L}_\mathcal{H}f, g \rangle\rangle, \quad f \in \mathcal{H} \cap \mathcal{D}(\mathcal{L}). \quad (14.18)$$

If the first eigenfunction e_1 is odd*, it therefore follows that the coupled Poincaré inequality (14.7) holds with the constant $\Lambda = \mu_2$, showing an improved rate of convergence for the coupled dynamics.

Example 14.2. Let us examine the second point, i.e. the possibility of $\mathcal{L}_\mathcal{H}$ not being symmetric with respect to $\langle\langle \cdot, \cdot \rangle\rangle$. For simplicity, assume that $\sigma(\mathcal{L}_\mathcal{H}) = \sigma(\mathcal{L})$, i.e. $\langle\langle e_i, e_i \rangle\rangle \neq 0$ for all $i \in \mathbb{N}$. Consider the case when the measures $\bar{\pi}_0$ and $\bar{\pi}_\Gamma$ have densities with respect to a common dominating measure m (for convenience denoted by the same symbols), and suppose there exist constants $c_1, c_2 > 0$ such that

$$c_1 \bar{\pi}_0(x) \leq \bar{\pi}_\Gamma(x) \leq c_2 \bar{\pi}_0(x), \quad x \in E. \quad (14.19)$$

This is the case precisely when the norms in $L^2(\bar{\pi}_0)$ and $L^2(\bar{\pi}_\Gamma)$ are equivalent. For $F = \sum_{i=1}^n f_i$, we have that $n\langle\langle f, f \rangle\rangle = \langle F, F \rangle_{L^2(\bar{\pi}_\Gamma)}$ as well as $n\langle f, f \rangle_{L^2(\pi)} = \langle F, F \rangle_{L^2(\bar{\pi}_0)}$.

*In one dimension, it can be proved that the first eigenfunction is always odd by appealing to the node theorem for Schrödinger operators in Sturm-Liouville theory [136, Chapter 9]. We conjecture that this fact might also be true in higher dimensions, but are not able to give a proof or a reference. Our special thanks go to Sabine Bögli and Ari Laptev for discussing this question with us.

Using (14.19), we hence conclude that

$$c_1 \langle f, f \rangle_{L^2(\pi)} \leq \langle\langle f, f \rangle\rangle \leq c_2 \langle f, f \rangle_{L^2(\pi)}, \quad f \in L^2(\pi). \quad (14.20)$$

By assumption, the marginal process satisfies the Poincaré inequality (1.51) as well as the equivalent decay estimate

$$\|S_t f\|_{L^2(\pi)}^2 \leq e^{-2\lambda t} \|f\|_{L^2(\pi)}^2, \quad f \in L_0^2(\pi), \quad (14.21)$$

with $\lambda = \mu_1$. By the equivalence (14.20), we conclude that

$$\|\bar{S}_t^\Gamma F\|_{L^2(\bar{\pi}_\Gamma)}^2 \leq C e^{-2\lambda t} \|F\|_{L^2(\bar{\pi}_\Gamma)}^2, \quad t \geq 0, \quad F \in \tilde{L}_0^2(\bar{\pi}_\Gamma), \quad (14.22)$$

with $C = \frac{c_2}{c_1}$. Comparing (14.21) and (14.22), we see that the coupled process achieves the same exponential rate of convergence as the one-particles processes, but possibly with a worse constant C in front of the exponential. The latter can be characterised in terms of the equivalence estimate (14.19).

Conclusively, the speed of convergence to equilibrium can be both faster (as in Example 14.1) and slower (as in Example 14.2) for coupled processes, in comparison with the underlying one-particle processes. We leave a more thorough investigation of the Poincaré inequality (14.7) for future work.

15. DISCUSSION

In this part of the thesis we introduced a general framework for the construction and analysis of coupled MCMC samplers, most notably observing an interesting connection with the theory of optimal transportation. Our work can be extended in several directions. On the theoretical side, proving or disproving the Conjectures 10.1, 10.2 and 10.3 would further illuminate the structural properties of the developed theory. Moreover, establishing a more rigorous connection between the optimal transport problems 12.1, 12.2 and 13.1 might lead to further developments bridging the theories of Markov processes and optimal transportation.

In terms of applications in sampling, a more detailed study of the couplings between many particles is needed, a starting point being the results in Section 13.2. Furthermore, it would be desirable to relax our assumption that the laws of the marginal processes remain unchanged, as this would allow for more pronounced interactions between the particles. In this regard, the inclusion of the methodology put forward in [88] in our framework would be of particular interest for practitioners.

In the broader context of statistical computation, it seems that coupling approaches along the lines developed here could be fruitfully applied in the context of the calculation of transport coefficients and sensitivities [8, 67]. More speculatively, it would be interesting to investigate the use of our ideas in the context of multilevel Monte Carlo [54] or computational optimal transport [118]. We leave these directions for future investigations.

16. PROOFS AND ADDITIONAL MATERIAL

In this chapter we provide proofs and additional material related to the content of Part III. In Section 16.1 we provide an additional lemma required in the proof of Lemma 10.2, addressing the ergodicity of coupled processes. Section 16.2 is concerned with orthogonal transformations of Brownian motions, playing a crucial role in the proofs of Lemmas 11.1 and 11.2. The purpose of Section 16.3 is to show the proof of Proposition 12.1 which is concerned with the derivative of the invariant measure with respect to an infinitesimal change of the coupling. Finally, Section 16.4 gathers the proofs of Lemmas 13.1 and 13.2 addressing the properties of solutions to one-dimensional Poisson equations related to the overdamped Langevin dynamics and the zigzag process.

16.1 TIGHTNESS OF COUPLED MEASURES

Let us first recall the definition of tightness of a set of probability measures:

Definition 16.1 (Tightness). *Let E be a Polish space and $\mathcal{P} \subset \mathcal{P}(E)$ be a set of probability measures. Then $\mathcal{P}$ is said to be tight, if for all $\varepsilon > 0$ there exists a compact set $K_\varepsilon \subset E$ such that $\mu(E \setminus K_\varepsilon) \leq \varepsilon$ for all $\mu \in \mathcal{P}$.*

The following lemma shows that tightness is inherited by couplings of sets of probability measures. This property is convenient when addressing ergodicity properties of coupled processes (see Lemma 10.2).

Lemma 16.1 (Tightness of couplings). *Let $(E_i)_{i=1}^n$ be a collection of Polish spaces and $\mathcal{P}_i \subset \mathcal{P}(E_i)$ be tight sets of probability measures. Denote by $\mathcal{C} \subset \mathcal{P}(E_1 \times \dots \times E_n)$ the set of couplings of $(\mathcal{P}_i)_{i=1}^n$, i.e. $\bar{\mu} \in \mathcal{C}$ if and only if for $i = 1, \dots, n$ there exist $\mu_i \in \mathcal{P}_i$ such that $\bar{\mu}$ has the marginals $(\mu_i)_{i=1}^n$. Then $\mathcal{C}$ is tight.*

Proof. The proof is an adaptation of [142, Lemma 4.4]. Let $\varepsilon > 0$. By assumption, there exist compact sets $K_\varepsilon^i \subset E_i$ such that $\mu_i(E_i \setminus K_\varepsilon^i) \leq \varepsilon$, for all $i \in \{1, \dots, n\}$ and $\mu_i \in \mathcal{P}_i$. Let $\bar{E} := E_1 \times \dots \times E_n$ and $\bar{K} = K_1 \times \dots \times K_n$. For fixed marginals $\mu_i \in \mathcal{P}_i$ and any

coupling $\bar{\mu}$ of $(\mu_i)_{i=1}^n$, we have that

$$\bar{\mu}(\bar{E} \setminus \bar{K}) \leq \sum_{i=1}^n \mu_i(E_i \setminus K_i) \leq n \cdot \varepsilon. \quad (16.1)$$

The claim now follows since this bound is independent of the choice of $(\mu_i)_{i=1}^n$. $\square$

16.2 RANDOM ORTHOGONAL TRANSFORMATIONS OF BROWNIAN MOTIONS

The following lemma has been extracted from [45, page 56], see also [107, Theorem 8.4.2]. This result states that the set of Brownian motions is preserved under possibly time-dependent linear transformations possessing certain orthogonality properties. Importantly, no regularity constraints with regard to the time-dependence are required beyond measurability. This fact is crucial in the proofs of Lemmas 11.1 and 11.2.

Lemma 16.2 (Random orthogonal transformations). *Suppose that the $\mathbb{R}^N$ -valued stochastic process $(X_t)_{t \geq 0}$ is a solution to the SDE*

$$dX_t = O_t dW_t, \quad X_0 = x_0, \quad (16.2)$$

where $(W_t)_{t \geq 0}$ is an M -dimensional standard Brownian motion generating the filtration $(\mathcal{F}_t)_{t \geq 0}$, and $(O_t)_{t \geq 0}$ is a product-measurable $(\mathcal{F}_t)_{t \geq 0}$ -adapted process taking values in $\mathbb{R}^{N \times M}$. Assume furthermore that

$$O_t O_t^T = I_{N \times N} \quad (16.3)$$

for all $t \geq 0$, almost surely. Then $(X_t)_{t \geq 0}$ is an N -dimensional standard Brownian motion.

Proof. The components

$$X_t^i = x_0^i + \sum_{k=1}^M \int_0^t O_s^{ik} dW_s^k \quad (16.4)$$

are continuous local martingales. Their covariations are given by

$$[X^i, X^j]_t = \sum_{kl} \int_0^t O_s^{ik} O_s^{jl} d[B^k, B^l]_s = \int_0^t \sum_k O_s^{ik} O_s^{jk} ds = \delta_{ij} t. \quad (16.5)$$

The claim follows therefore by Lévy's characterisation of Brownian motion. $\square$

16.3 THE DERIVATIVE FORMULA FOR INVARIANT MEASURES

Here, we provide the proof of the derivative formula (12.30) that allows us to compute the change of the invariant measure under an infinitesimal change of the coupling.

Proof of Proposition 12.1. The idea of the proof stems from [87] in the context of invariant measures for discretised SDEs and was also advertised in [91, Remark 5.5].

For convenience, let us first introduce the notation $g = -(\bar{\mathcal{L}}_\Gamma^*)^{-1}d\Gamma^*1$. Furthermore, we will make use of the projection operators

$$\Pi\phi = \int_{\bar{E}} \phi d\bar{\pi}_\Gamma^0, \quad \Pi^\perp\phi = \phi - \Pi\phi, \quad (16.6)$$

acting on $L^2(\bar{\pi}_\Gamma)$. Using $c = (\Pi + \Pi^\perp)c$ in (12.30), we see that (12.30) is equivalent to

$$\left. \frac{d}{d\varepsilon} \right|_{\varepsilon=0} \int_{\bar{E}} \Pi^\perp c (d\bar{\pi}_\Gamma^\varepsilon) = - \int_{\bar{E}} (\Pi^\perp c) [\bar{\mathcal{L}}_\Gamma^*]^{-1} (d\Gamma^*1) d\bar{\pi}_\Gamma^0. \quad (16.7)$$

We may thus without loss of generality assume that $\Pi^\perp c = c$ (i.e. $\Pi c = 0$), and will do so in the following. Furthermore, let us also assume that $c \in C^\infty(\bar{E})$ such that the calculations in the sequel are justified. The general case then follows by a standard approximation argument. A short calculation (using the fact that $\bar{\pi}_\Gamma^0(\bar{\mathcal{L}}_\Gamma c) = 0$) shows that

$$\int_{\bar{E}} \bar{\mathcal{L}}_\Gamma^\varepsilon c \cdot (1 + \varepsilon g) d\bar{\pi}_\Gamma^0 = \varepsilon^2 \int_{\bar{E}} (d\Gamma c) g d\bar{\pi}_\Gamma^0 = \mathcal{O}(\varepsilon^2). \quad (16.8)$$

Inserting $\Pi + \Pi^\perp = I$, we see that the above is equivalent to

$$\int_{\bar{E}} \Pi^\perp \bar{\mathcal{L}}_\Gamma^\varepsilon \Pi^\perp c \cdot (1 + \varepsilon g) d\bar{\pi}_\Gamma^0 + \varepsilon \int_{\bar{E}} \Pi d\Gamma c \cdot (1 + \varepsilon g) d\bar{\pi}_\Gamma^0 = \mathcal{O}(\varepsilon^2). \quad (16.9)$$

At the same time, we have that

$$\int_{\bar{E}} \bar{\mathcal{L}}_\Gamma^\varepsilon c d\bar{\pi}_\Gamma^\varepsilon = 0. \quad (16.10)$$

Using again $\Pi + \Pi^\perp = I$ and $\Pi \bar{\mathcal{L}}_\Gamma = 0$, (16.10) can be expressed as

$$\int_{\bar{E}} \Pi^\perp \bar{\mathcal{L}}_\Gamma^\varepsilon \Pi^\perp c d\bar{\pi}_\Gamma^\varepsilon = -\varepsilon \int_{\bar{E}} \Pi d\Gamma c d\bar{\pi}_\Gamma^\varepsilon. \quad (16.11)$$

We can now combine (16.9) and (16.11) to arrive at

$$\int_{\bar{E}} \Pi^\perp \bar{\mathcal{L}}_\Gamma^\varepsilon \Pi^\perp c \cdot (1 + \varepsilon g) d\bar{\pi}_\Gamma^0 - \int_{\bar{E}} \Pi^\perp \bar{\mathcal{L}}_\Gamma^\varepsilon \Pi^\perp c d\bar{\pi}_\Gamma^\varepsilon = \mathcal{O}(\varepsilon^2). \quad (16.12)$$

Let us introduce the ‘pseudo-inverse’

$$Q_\varepsilon = \Pi^\perp \bar{\mathcal{L}}_\Gamma^{-1} \Pi^\perp - \varepsilon \Pi^\perp \bar{\mathcal{L}}_\Gamma^{-1} \Pi^\perp d\Gamma \Pi^\perp \bar{\mathcal{L}}_\Gamma^{-1} \Pi^\perp, \quad (16.13)$$

acting on $L_0^2(\bar{\pi}_\Gamma^0)$. We have that

$$\Pi^\perp \bar{\mathcal{L}}_\Gamma^\varepsilon \Pi^\perp Q_\varepsilon = \Pi^\perp - \varepsilon^2 \Pi^\perp d\Gamma \Pi^\perp \bar{\mathcal{L}}_\Gamma^{-1} \Pi^\perp d\Gamma \Pi^\perp \bar{\mathcal{L}}_\Gamma^{-1} \Pi^\perp, \quad (16.14)$$

i.e. in $L_0^2(\bar{\pi}_\Gamma^0)$, Q_ε inverts $\bar{\mathcal{L}}_\Gamma^\varepsilon$ up to an error of order ε^2 . Upon replacing c by $Q_\varepsilon c$ in (16.12), it follows that

$$\int_{\bar{E}} c(1 + \varepsilon g) d\bar{\pi}_\Gamma^0 - \int_{\bar{E}} c d\bar{\pi}_\Gamma^\varepsilon = \mathcal{O}(\varepsilon^2), \quad (16.15)$$

recalling that $\Pi^\perp c = c$ by assumption. In the last step, we have used the fact that there exists a constant $C > 0$ such that

$$\left| \int_{\bar{E}} \Pi^\perp d\Gamma \Pi^\perp \bar{\mathcal{L}}_\Gamma^{-1} \Pi^\perp d\Gamma \Pi^\perp \bar{\mathcal{L}}_\Gamma^{-1} \Pi^\perp c d\bar{\pi}_\Gamma^\varepsilon \right| \leq C, \quad (16.16)$$

uniformly in ε . Indeed, the integrand is bounded by the third condition of Definition 12.2 and the fact that the coefficients of $d\Gamma$ have compact support. The bound (16.16) is required to ensure that the corresponding integral expression hidden on the right-hand side of (16.15) is indeed of order ε^2 . Finally, deviding by ε in (16.15) and letting $\varepsilon \rightarrow 0$ yields the desired result. $\square$

16.4 PROPERTIES OF THE SOLUTIONS TO ONE-DIMENSIONAL POISSON EQUATIONS

The proofs in this section essentially leverage the fact that the Poisson equations under consideration can be solved up to quadratures in one dimension.

Proof of Lemma 13.1. Variation of constants shows that ϕ' is given by

$$\phi'(x) = \left(- \int_{-\infty}^x f(s) e^{-V(s)} ds + C \right) e^{V(x)}, \quad (16.17)$$

for some constant $C \in \mathbb{R}$. The requirement that $\pi(\phi) = 0$ necessitates $C = 0$. Indeed, from $\pi(f) = 0$ it follows that the integral term in (16.17) goes to zero as $x \rightarrow \pm\infty$, and therefore

$$\lim_{x \rightarrow \pm\infty} \frac{\phi'(x)}{e^{V(x)}} = C. \quad (16.18)$$

By L'Hôpital's rule, we have that

$$\lim_{x \rightarrow \pm\infty} \frac{\phi(x) e^{-V(x)}}{\int_0^x e^{V(s)} ds \cdot e^{-V(x)}} = C. \quad (16.19)$$

The requirement that ϕ is integrable with respect to $\pi(dx) \propto e^{-V(x)} dx$ implies that

$$\lim_{x \rightarrow \pm\infty} \phi(x) e^{-V(x)} = 0.$$

Furthermore (again by L'Hôpital's rule),

$$\lim_{x \rightarrow \pm\infty} \int_0^x e^{V(s)} ds \cdot e^{-V(x)} = - \lim_{x \rightarrow \pm\infty} V'(x), \quad (16.20)$$

which cannot be zero since $\int_{-\infty}^{\infty} e^{-V(x)} dx < \infty$. Hence, $C = 0$.

To prove 1.), notice that from $\pi(f) = 0$ and monotonicity, it follows that there exists $x^* \in \mathbb{R}$ such that $f(x^*) = 0$. Let us assume that f is monotonically increasing (for monotonically decreasing f the reasoning is analogous). We then have that $f \leq 0$ on $(-\infty, x^*]$ and $f \geq 0$ on $[x^*, \infty)$. Consider now the function

$$\Phi(x) = - \int_{-\infty}^x f(s) e^{-V(s)} ds, \quad x \in \mathbb{R}. \quad (16.21)$$

Clearly Φ is increasing on $(-\infty, x^*]$ and decreasing on $[x^*, \infty)$. From $\pi(f) = 0$ it follows that $\lim_{x \rightarrow \pm\infty} \Phi(x) = 0$ and hence $\Phi(x) \geq 0$ for all $x \in \mathbb{R}$. This proves the claim since $\phi'(x) = \Phi(x) e^{V(x)}$.

To prove 2.), first observe that $\pi(f) = 0$ implies

$$\int_{-\infty}^0 f(s) e^{-V(s)} ds + \int_0^{\infty} f(s) e^{-V(s)} ds = 0. \quad (16.22)$$

Furthermore, the symmetry properties of f and V show that

$$\int_{-\infty}^0 f(s) e^{-V(s)} ds - \int_0^{\infty} f(s) e^{-V(s)} ds = 0, \quad (16.23)$$

using the substitution $s \mapsto -s$. Equations (16.22) and (16.23) together imply that $\Phi(0) = 0$, for Φ as defined in (16.21). The claim now follows using an analogous argument to the one used in the proof of 1.). $\square$

Proof of Lemma 13.2. Recall from (13.4) that

$$\delta\sigma_F^2(\Gamma_\alpha) = \int_{\bar{E}} \xi d\bar{\pi}_0, \quad (16.24)$$

where $\xi(x, y) = \phi(x)\phi(y)$ and $\phi : \mathbb{R} \times \{-1, +1\} \rightarrow \mathbb{R}$ is the solution to the (one-particle) Poisson equation

$$\theta \partial_x \phi(x, \theta) + \lambda(x, \theta) (\phi(x, -\theta) - \phi(x, \theta)) = f(x), \quad \pi(\phi) = 0. \quad (16.25)$$

Note that for convenience, we have assumed without loss of generality that $\tilde{\pi}(f) = 0$. Let

us now calculate

$$\begin{aligned}
\int_{\bar{E}} \Gamma \xi \, d\pi_0 &= \frac{1}{4} \sum_{\theta_x=\pm 1, \theta_y=\pm 1} \int_{\mathbb{R}^2} \left(\alpha(x, y, \theta_x, \theta_y) \cdot \left[\phi(x, \theta_x) \phi(y, \theta_y) - \phi(x, \theta_x) \phi(y, -\theta_y) - \right. \right. \\
&\quad \left. \left. - \phi(x, -\theta_x) \phi(y, \theta_y) + \phi(x, -\theta_x) \phi(y, -\theta_y) \right] \right) e^{-(V(x)+V(y))} \, dx dy \\
&= \frac{1}{4} \int_{\mathbb{R}^2} \left(\left[\alpha_{++}(x, y) + \alpha_{--}(x, y) - \alpha_{+-}(x, y) - \alpha_{-+}(x, y) \right] \cdot \right. \\
&\quad \left. \cdot \left[\phi_+(x) \phi_+(y) + \phi_-(x) \phi_-(y) - \phi_+(x) \phi_-(y) - \phi_-(x) \phi_+(y) \right] \right) e^{-(V(x)+V(y))} \, dx dy \\
&= \frac{1}{4} \int_{\mathbb{R}^2} \left(\left[\alpha_{++}(x, y) + \alpha_{--}(x, y) - \alpha_{+-}(x, y) - \alpha_{-+}(x, y) \right] \cdot \right. \\
&\quad \left. \cdot \left[\phi_+(x) - \phi_-(x) \right] \cdot \left[\phi_+(y) - \phi_-(y) \right] \right) e^{-(V(x)+V(y))} \, dx dy, \tag{16.26}
\end{aligned}$$

where again we employed the notation $\phi_{\pm}(x) = \phi(x, \pm 1)$. Observe now that equation (16.25) can be recast as

$$\partial_x \phi_+(x) + \lambda_+(x) (\phi_-(x) - \phi_+(x)) = f(x), \tag{16.27a}$$

$$-\partial_x \phi_-(x) + \lambda_-(x) (\phi_+(x) - \phi_-(x)) = f(x), \tag{16.27b}$$

where both ϕ_+ and ϕ_- have to be integrable with respect to the measure $\frac{1}{Z} e^{-V(x)} dx$ and satisfy

$$\int_{\mathbb{R}} (\phi_+(x) + \phi_-(x)) e^{-V(x)} dx = 0. \tag{16.28}$$

Adding (16.27a) and (16.27b) leads to

$$(\partial_x - V') (\phi_+ - \phi_-) = 2f, \tag{16.29}$$

using $\lambda_+(x) - \lambda_-(x) = V'(x)$. Finally setting $\partial_x \tilde{\phi} = \frac{1}{2}(\phi_+ - \phi_-)$ and comparing with (16.26) leads to the desired result. Note that as in the proof of Lemma 13.1, (16.29) determines $\phi_+ - \phi_-$ uniquely under the condition that $\phi_+ - \phi_-$ is integrable with respect to $e^{-V(x)} dx$. $\square$

Part IV

Ergodicity of piecewise deterministic Markov processes

17. HYPOCOERCIVITY FOR THE ZIGZAG PROCESS

The aim of this chapter is to investigate the ergodic properties of the zigzag process introduced in Section 1.3.2. In particular, under certain conditions we are able to prove exponentially fast convergence to equilibrium of the corresponding semigroup. We will consider the case of a compact state space, replacing $\mathbb{R}^d$ by the torus $\mathbb{T}^d$. In situations where the target measure is sufficiently concentrated this setting is a reasonable approximation. Extending the analysis to unbounded state spaces poses significant technical difficulties and will be addressed in future work. The results in this chapter have been obtained in collaboration with Julien Roussel (PhD student at CERMICS Paris).

Let us briefly recall the setting already outlined in Section 1.3.2, but considering the compact state space $E = \mathbb{T}^d \times \{-1, 1\}^d$. The target measure is $\hat{\pi} = \pi \otimes \kappa$, where $\pi \propto e^{-V(x)} dx$ is the measure of interest, associated with a smooth potential $V \in C^\infty(\mathbb{T}^d)$, and κ is the uniform probability measure on $\{-1, 1\}^d$. Recall that the generator of the zigzag process is given by

$$\mathcal{L}f(x, \theta) = \sum_{i=1}^d (\theta_i \partial_i f(x, \theta) + \lambda_i(x, \theta) \cdot (f(x, F_i \theta) - f(x, \theta))), \quad (x, \theta) \in E, \quad (17.1a)$$

$$f \in \mathcal{D} = \left\{ f : E \rightarrow \mathbb{R} : \quad f(\cdot, \theta) \in C^\infty(\mathbb{T}^d) \text{ for all } \theta \in \{-1, 1\}^d \right\}, \quad (17.1b)$$

where the flip operators $F_i : \{-1, 1\}^d \rightarrow \{-1, 1\}^d$ are defined as follows:

$$(F_i \theta)_j = \begin{cases} \theta_j & j \neq i \\ -\theta_j & j = i. \end{cases} \quad (17.2)$$

The switching rates $\lambda_i : E \rightarrow [0, \infty)$ are given by

$$\lambda_i(x, \theta) = (\theta \partial_i V(x))_+ + \gamma_i(x, \theta), \quad (17.3)$$

where we recall that $a_+ = \max(a, 0)$ denotes the positive part. The excess switching rate is a continuous function $\gamma : E \rightarrow [0, \infty)^d$ satisfying

$$\gamma_i(x, F_i \theta) = \gamma_i(x, \theta), \quad i = 1, \dots, d. \quad (17.4)$$

For our main result in this chapter, we will need the following assumption connecting the potential V and the switching rate γ :

Assumption 17.1. *There exists a constant $\delta > 0$ such that*

$$|\partial_i V(x)| + \gamma_i(x, \theta) \geq \delta \quad (17.5)$$

for all $(x, \theta) \in E$, $i = 1, \dots, d$.

The condition (17.5) places a lower bound on the excess switching rate γ in regions where the partial derivatives of V are small. In other words, a small amount of additional noise is needed for our proof to work. For the performance of the sampler, however, it seems desirable to use the canonical switching rate $\gamma \equiv 0$, reducing diffusive behaviour and pronouncing the nonreversible nature of the dynamics*. The proof of ergodicity in the case $\gamma \equiv 0$ has posed a difficult challenge in the past due to the fact that irreducibility of the process was hard to obtain. This issue has only been resolved recently in [23], using very different technique from ours (namely methods based on Lyapunov functions). Using the approach in this chapter, it is not obvious how to substantially weaken the condition (17.5). However, our proof appears to bear potential to treat a broader class of piecewise deterministic Markov process in a generalised framework. Moreover, the approach presented in this chapter at least in principle allows to obtain explicit estimates for the constants appearing in the decay estimate. We leave these developments for future work.

In the following we will as usual denote the semigroup associated to the generator $\mathcal{L}$ by $(S_t)_{t \geq 0}$. Under the foregoing assumptions, exponential decay as in (1.46) holds:

Theorem 17.1 (Exponential convergence to equilibrium). *Let Assumption 17.1 be satisfied. Then there exist constants $\lambda > 0$ and $C \geq 1$ such that*

$$\|S_t f - \hat{\pi}(f)\|_{L^2(\hat{\pi})} \leq C e^{-\lambda t} \|f - \hat{\pi}(f)\|_{L^2(\hat{\pi})}, \quad t \geq 0, f \in L^2(\hat{\pi}). \quad (17.6)$$

Remark 17.1. *Recall that the foregoing decay estimate immediately implies the validity of a central limit theorem of the form (1.52) as discussed in Section 1.2.2.*

17.1 FUNCTIONAL-ANALYTIC SETTING AND OUTLINE OF THE PROOF

The proof of Theorem 17.1 relies on the method developed in [40] (see also [29, 59, 60, 61]). In this section we will fix the framework and outline the general strategy. We will naturally work in the Hilbert space

$$L^2(\hat{\pi}) = \left\{ f : E \rightarrow \mathbb{R} : \|f\|_{L^2(\hat{\pi})}^2 < \infty \right\}, \quad (17.7)$$

*See for instance the discussion in [21, Sections 4 and 5].

where

$$\|f\|_{L^2(\widehat{\pi})}^2 = \int_E f^2 d\widehat{\pi} = \frac{1}{2^d} \sum_{\theta \in \{-1,1\}^d} \int_{\mathbb{T}^d} f^2(x, \theta) \pi(dx). \quad (17.8)$$

All the adjoints in this chapter will be taken with respect to the induced scalar product $\langle \cdot, \cdot \rangle_{L^2(\widehat{\pi})}$. It will turn out convenient to make use of the operators

$$R_i : L^2(\widehat{\pi}) \rightarrow L^2(\widehat{\pi}), \quad (R_i f)(x, \theta) = f(x, F_i \theta), \quad (17.9)$$

already introduced in (1.91), such that $\mathcal{L}$ can be concisely written as

$$\mathcal{L} = \sum_{i=1}^d (\theta_i \partial_i + \lambda_i (R_i - I)). \quad (17.10)$$

In the sequel, we will make repeated use of the facts gathered in the next lemma:

Lemma 17.1. *The following hold:*

1. For all $i = 1, \dots, d$, the operator R_i is self-adjoint in $L^2(\widehat{\pi})$.
2. For all $i = 1, \dots, d$, the formal adjoint of ∂_i in $L^2(\widehat{\pi})$ is given by

$$\partial_i^* = -\partial_i + \partial_i V, \quad (17.11)$$

i.e.

$$\langle f, \partial_i g \rangle_{L^2(\widehat{\pi})} = \langle (-\partial_i + \partial_i V) f, g \rangle_{L^2(\widehat{\pi})}, \quad f, g \in L^2(\widehat{\pi}). \quad (17.12)$$

Proof. For the first claim, we see that R_i is symmetric since

$$\sum_{\theta \in \{-1,1\}^d} \int_{\mathbb{T}^d} f(x, \theta) g(x, F_i \theta) \pi(dx) = \sum_{\theta \in \{-1,1\}^d} \int_{\mathbb{T}^d} f(x, F_i \theta) g(x, \theta) \pi(dx) \quad (17.13)$$

can be obtained by letting $\theta \mapsto F_i \theta$ and noting that $F_i^2 = I$. Furthermore, R_i is self-adjoint since it is defined on the whole of $L^2(\widehat{\pi})$. The second claim follows easily by integration by parts. $\square$

Using the foregoing lemma, we can now calculate the (formal) adjoint of $\mathcal{L}$ in $L^2(\widehat{\pi})$,

$$\mathcal{L}^* = \sum_{i=1}^d (-\theta_i \partial_i + \theta_i \partial_i V + (R_i - I) \lambda_i), \quad (17.14)$$

as well as its symmetric and skew-symmetric parts:

$$\mathcal{S} = \frac{1}{2}(\mathcal{L} + \mathcal{L}^*) = \frac{1}{2} \sum_{i=1}^d (|\partial_i V| + 2\gamma_i) (R_i - I), \quad (17.15a)$$

$$\mathcal{T} = \frac{1}{2}(\mathcal{L} - \mathcal{L}^*) = \frac{1}{2} \sum_{i=1}^d (2\theta_i \partial_i + \theta_i (\partial_i V) (R_i - I)). \quad (17.15b)$$

Furthermore, the decomposition $\mathcal{S} = \sum_{i=1}^d A_i^* A_i$, defining

$$A_i = \frac{1}{2} (|\partial_i V| + 2\gamma_i)^{1/2} (R_i - I), \quad (17.16)$$

will prove useful and has already been exploited in the construction of coupled zigzag samplers (see Section 11.3). Finally, we will make extended use of the orthogonal projection onto the space of functions that do not depend on θ , given by

$$\Pi_\theta f = \frac{1}{2^d} \sum_{\theta \in \{-1,1\}^d} f(\cdot, \theta), \quad f \in \mathcal{D}, \quad (17.17)$$

or, equivalently, by

$$\Pi_\theta = \frac{1}{2^d} \prod_{i=1}^d (R_i + I). \quad (17.18)$$

Indeed, the following is true:

Lemma 17.2. Π_θ is an orthogonal projection in $L^2(\widehat{\pi})$, i.e. $\Pi_\theta^2 = \Pi_\theta$ and $\Pi_\theta^* = \Pi_\theta$.

Proof. The property $\Pi_\theta^2 = \Pi_\theta$ is immediate (note the prefactor of $\frac{1}{2^d}$). Orthogonality ($\Pi_\theta^* = \Pi_\theta$) follows from the symmetry of the expression

$$\langle f, \Pi_\theta g \rangle_{L^2(\widehat{\pi})} = \frac{1}{2^d} \int_{\mathbb{T}^d} \left(\sum_{\theta \in \{-1,1\}^d} f(x, \theta) \right) \cdot \left(\sum_{\theta \in \{-1,1\}^d} g(x, \theta) \right) \pi(dx). \quad (17.19)$$

□

We will now briefly explain the strategy of the proof. It is well known that the exponential decay estimate (17.6) with $C = 1$ is equivalent to a Poincaré inequality of the form

$$\langle f, f \rangle_{L^2(\widehat{\pi})} \leq \frac{1}{\lambda} \langle f, (-\mathcal{L})f \rangle_{L^2(\widehat{\pi})}, \quad f \in \mathcal{D} \cap L_0^2(\widehat{\pi}), \quad (17.20)$$

see Theorem 1.3. However, (17.15a) clearly shows that functions that do not depend on θ are in the null space of $\mathcal{S}$. Together with $\langle f, (-\mathcal{L})f \rangle_{L^2(\widehat{\pi})} = \langle f, (-\mathcal{S})f \rangle_{L^2(\widehat{\pi})}$, this implies that (17.20) cannot hold. Intuitively, this shows that the skew-symmetric part $\mathcal{T}$ plays a crucial role in driving the process towards equilibrium. The idea behind the proof is

therefore to introduce an equivalent scalar product such that the skew-symmetric part does not drop out in an inequality of Poincaré type. To be more specific, let us introduce the operator

$$G = (I + (\mathcal{T}\Pi_\theta)^*(\mathcal{T}\Pi_\theta))^{-1} (\mathcal{T}\Pi_\theta)^* \quad (17.21)$$

and the *modified entropy*

$$\mathcal{H}_\varepsilon(f) = \frac{1}{2} \|f\|_{L^2(\widehat{\pi})}^2 - \varepsilon \langle f, Gf \rangle_{L^2(\widehat{\pi})}, \quad f \in L^2(\widehat{\pi}), \quad (17.22)$$

for some $\varepsilon \in (-1, 1)$ to be specified later. The following two lemmas are crucial for our approach; their proofs can be found in the next section (see Lemma 17.10 and Remark 17.4).

Lemma 17.3. *The operator G is bounded.*

In particular, (17.22) is indeed meaningful for all $f \in L^2(\widehat{\pi})$.

Lemma 17.4. *For all $\varepsilon \in (-1, 1)$, $\mathcal{H}_\varepsilon$ is an equivalent norm on $L^2(\widehat{\pi})$, i.e. there exist constants $c_\mathcal{H}, C_\mathcal{H} > 0$ such that*

$$c_\mathcal{H} \|f\|_{L^2(\widehat{\pi})}^2 \leq \mathcal{H}_\varepsilon(f) \leq C_\mathcal{H} \|f\|_{L^2(\widehat{\pi})}^2, \quad f \in L^2(\widehat{\pi}). \quad (17.23)$$

Setting $\langle\langle f, f \rangle\rangle_\varepsilon := \mathcal{H}_\varepsilon(f)$ and extending by polarisation[†] yields an equivalent scalar product on $L^2(\widehat{\pi})$. Let us now define the *entropy dissipation functional*

$$D_\varepsilon(f) = \langle\langle -\mathcal{L}f, f \rangle\rangle_\varepsilon, \quad f \in \mathcal{D}, \quad (17.24)$$

motivated by the fact that

$$\frac{d}{dt} \mathcal{H}_\varepsilon(S_t f) = -2D_\varepsilon(S_t f), \quad f \in \mathcal{D}. \quad (17.25)$$

The proof of Theorem 17.1 now relies on the following proposition, establishing a coercivity result:

Proposition 17.1. *There exists $\bar{\varepsilon} \in (0, 1)$ and $\lambda > 0$ such that*

$$D_{\bar{\varepsilon}}(f) \geq \lambda \|f\|_{L^2(\widehat{\pi})}^2, \quad f \in \mathcal{D}. \quad (17.26)$$

The proof of Proposition 17.1 requires a significant amount of preparation and will therefore be given at the end of this chapter in Section 17.4. Taking this step for granted, we can already prove the main result:

[†]To be specific, we set $\langle\langle f, g \rangle\rangle_\varepsilon := \frac{1}{4} (\mathcal{H}_\varepsilon(f + g) - \mathcal{H}_\varepsilon(f - g))$.

Proof of Theorem 17.1. Using (17.25), (17.26) and (17.23), we see that

$$\frac{d}{dt} \mathcal{H}_\varepsilon(S_t f) = -2D_\varepsilon(S_t f) \leq -2\lambda \|S_t f\|_{L^2(\widehat{\pi})} \leq -2\lambda C_{\mathcal{H}} \cdot \mathcal{H}_\varepsilon(S_t f), \quad (17.27)$$

for all $f \in \mathcal{D}$. The desired decay estimate now follows from Gronwall's inequality and the equivalence bounds (17.23). Finally, for arbitrary $f \in L^2(\widehat{\pi})$, the result follows by density of $\mathcal{D}$ in $L^2(\widehat{\pi})$. $\square$

Remark 17.2. *The approach presented here bears strong similarities to the proof of Theorem 3.2 in Section 3.1. In fact, it is no coincidence that we denote the operator (17.21) and the matrix obtained in Lemma 3.1 by the same letter, highlighting the similar role they play in the respective arguments. There is also a notable connection to the theory of hypocoercivity developed by Villani [141], employed in Section 4.1.*

Remark 17.3 (Relationships to underdamped Langevin dynamics). *It is interesting to note that the generator of the underdamped Langevin dynamics (1.79) (for simplicity with $M = I$ and $\Gamma = \gamma(q) \cdot I$) can be written in the form*

$$\mathcal{L}_{\text{underdamped}} = p \cdot \nabla_q - \nabla_q V(q) \cdot \nabla_p - \gamma(q) (p \cdot \nabla_p + \Delta_p) \quad (17.28a)$$

$$= -\underbrace{\gamma(q) \sum_{i=1}^d \partial_{p_i}^* \partial_{p_i}}_{\mathcal{S}} + \underbrace{\sum_{i=1}^d (\partial_{p_i}^* \partial_{q_i} - \partial_{q_i}^* \partial_{p_i})}_{\mathcal{T}}, \quad (17.28b)$$

and that the generator of the zigzag process has the very similar form

$$\mathcal{L}_{\text{zigzag}} = -\underbrace{\sum_{i=1}^d A_i^* A_i}_{\mathcal{S}} + \underbrace{\sum_{i=1}^d (A_i^* \partial_i - \partial_i^* A_i)}_{\mathcal{T}}. \quad (17.29)$$

This correspondence motivates some of the techniques used in this chapter. We mention in particular the recent works [74, 130] where similar methods have been employed in the study of exponential convergence to equilibrium for the underdamped Langevin dynamics.

17.2 MICROSCOPIC AND MACROSCOPIC COERCIVITY

One key ingredient in the proof of Proposition 17.1 is the insight that coercivity estimates similar to (17.26) hold for $\mathcal{S}$ and $\mathcal{T}$ separately on the orthogonal subspaces $\Pi_\theta L^2(\widehat{\pi})$ and $(I - \Pi_\theta) L^2(\widehat{\pi})$. The terminology is due to the intuition that coercivity on $(I - \Pi_\theta) L^2(\widehat{\pi})$ is related to equilibration of the auxiliary variables $\theta \in \{-1, 1\}^d$ ('microscopic coercivity'), while coercivity on $\Pi_\theta L^2(\widehat{\pi})$ is related to equilibration in $\mathbb{T}^d$ ('macroscopic coercivity'). Let us start with the following crucial lemma, illuminating the interaction of Π_θ with $\mathcal{S}$ and $\mathcal{T}$:

Lemma 17.5. *The following hold:*

$$1. \quad \mathcal{S}\Pi_\theta = 0, \quad (17.30a)$$

$$2. \quad \Pi_\theta \mathcal{T}\Pi_\theta = 0. \quad (17.30b)$$

Proof. First notice that $R_i - I$ vanishes on functions that do not depend on θ . The first claim then follows immediately since for any $f \in L^2(\widehat{\pi})$, $\Pi_\theta f$ does not depend on θ . For the second claim, note that for any $f \in \mathcal{D}$, $\mathcal{T}\Pi_\theta f$ is of the form $\sum_{i=1}^d \theta_i \partial_i \phi$, where ϕ does not depend on θ . Applying Π_θ therefore leads to cancellations, verifying (17.30b). $\square$

Lemma 17.6 (Microscopic coercivity). *Let Assumption 17.1 be satisfied. Then*

$$-\langle \mathcal{S}f, f \rangle_{L^2(\widehat{\pi})} \geq \delta \|(I - \Pi_\theta)f\|_{L^2(\widehat{\pi})}, \quad f \in \mathcal{D}. \quad (17.31)$$

Before proving Lemma 17.6, we need a preliminary result. To this end, let us introduce the vector space of functions on $\{-1, 1\}^d$,

$$\Theta := \left\{ g : \{-1, 1\}^d \rightarrow \mathbb{R} \right\}, \quad (17.32)$$

endowed with the scalar product

$$\langle f, g \rangle_\Theta := \frac{1}{2^d} \sum_{\theta \in \{-1, 1\}^d} f(\theta)g(\theta). \quad (17.33)$$

Clearly, $(\Theta, \langle \cdot, \cdot \rangle_\Theta)$ is a Hilbert space of dimension 2^d . For $i \in \{1, \dots, d\}$, let us fix real constants $c_i > 0$, and consider the linear operator

$$\mathcal{R} : \Theta \rightarrow \Theta, \quad (\mathcal{R}f)(\theta) = \frac{1}{2} \sum_{i=1}^d c_i (f(\theta) - f(F_i \theta)), \quad (17.34)$$

where we recall the definition of the flip operator F_i from (17.2). Then $\mathcal{R}$ is self-adjoint, nonnegative definite, and its spectrum (and hence its spectral gap) can be computed explicitly:

Lemma 17.7. *The spectrum of $\mathcal{R}$ is given by*

$$\sigma(\mathcal{R}) = \left\{ \sum_{i=1}^d \eta_i c_i : \eta \in \{0, 1\}^d \right\}. \quad (17.35)$$

The null space of $\mathcal{R}$ is spanned by constant functions, i.e. $\ker \mathcal{R} = \text{span}\{\mathbf{1}\}$. Furthermore,

$$\min(\sigma(\mathcal{R}) \setminus \{0\}) = \min_i c_i. \quad (17.36)$$

Proof. For $\eta \in \{0, 1\}^d$, let us introduce $g_\eta \in \Theta$ by

$$g_\eta(\theta) = \prod_{i=1}^d \theta_i^{\eta_i}. \quad (17.37)$$

It is straightforward to see that $(g_\eta)_{\eta \in \{0, 1\}^d}$ forms an orthonormal basis in Θ . Furthermore,

$$\mathcal{R}g_\eta = \left(\sum_{i=1}^d \eta_i c_i \right) g_\eta, \quad (17.38)$$

implying (17.35). Clearly, the null space of $\mathcal{R}$ is spanned by $g_0 = \mathbf{1}$. The spectral gap (17.36) follows directly by examination of (17.35). $\square$

Proof of Lemma 17.6. By (17.30a), we can assume without loss of generality that $\Pi_\theta f = 0$, i.e. that

$$\sum_{\theta \in \{-1, 1\}^d} f(x, \theta) = 0 \quad (17.39)$$

holds for all $x \in \mathbb{T}^d$. Let us now fix $x \in \mathbb{T}^d$, set $c_i = |\partial_i V(x)| + 2\gamma_i(x)$ and consider the associated operator $\mathcal{R}$. Notice that by (17.39), $f(x, \cdot) \in \Theta$ is orthogonal to the null space of $\mathcal{R}$, and therefore, by Lemma 17.7,

$$\begin{aligned} \langle f(x, \cdot), \mathcal{R}f(x, \cdot) \rangle_\Theta &\geq \left(\min_i |\partial_i V| + 2\gamma_i \right) \frac{1}{2^d} \sum_{\theta \in \{-1, 1\}^d} f^2(x, \theta) \\ &\geq \frac{\delta}{2^d} \sum_{\theta \in \{-1, 1\}^d} f^2(x, \theta), \end{aligned} \quad (17.40)$$

where the bound in the second line is guaranteed by Assumption 17.1. Finally, we obtain the desired result by integrating over $\mathbb{T}^d$ with respect to π . $\square$

The following lemma relies on the fact that since $\mathbb{T}^d$ is compact, the measure π satisfies a Poincaré inequality (see [10, Chapter 4]), i.e. there exists a constant $C_P > 0$ such that

$$\left(\int_{\mathbb{T}^d} f^2 d\pi - \int_{\mathbb{T}^d} f d\pi \right)^2 \leq C_P \int_{\mathbb{T}^d} |\nabla f|^2 d\pi, \quad f \in H^1(\pi). \quad (17.41)$$

Lemma 17.8 (Macroscopic coercivity). *The following holds:*

$$\|\mathcal{T}\Pi_\theta f\|_{L^2(\hat{\pi})}^2 \geq C_P \|\Pi_\theta f\|_{L^2(\hat{\pi})}^2, \quad f \in \mathcal{D}. \quad (17.42)$$

Furthermore, the following is true in terms of symmetric operators:

$$G\mathcal{T}\Pi_\theta \geq \Lambda\Pi_\theta, \quad (17.43)$$

where $\Lambda = \frac{C_P}{1+C_P}$.

Proof. First note that

$$\|\mathcal{T}\Pi_\theta f\|_{L^2(\widehat{\pi})}^2 = \frac{1}{2^d} \sum_{\theta \in \{-1,1\}^d} \int_{\mathbb{T}^d} \left(\sum_{i=1}^d \theta_i \frac{1}{2^d} \sum_{\theta \in \{-1,1\}^d} \partial_i f(x, \theta) \right)^2 \pi(dx), \quad (17.44)$$

since $(R_i - I)\Pi_\theta = 0$. We can now expand the square, observe that the cross terms vanish, use $\theta_i^2 = 1$ as well as the Poincaré inequality (17.41) to arrive at (17.42). For the second claim, notice that $(\mathcal{T}\Pi_\theta)^*(\mathcal{T}\Pi_\theta)$ is self-adjoint (see [80, Theorem 3.24]) and hence, using spectral calculus,

$$G\mathcal{T}\Pi_\theta = \psi((\mathcal{T}\Pi_\theta)^*(\mathcal{T}\Pi_\theta)), \quad \psi(x) = \frac{x}{1+x}. \quad (17.45)$$

Together with (17.42), (17.43) follows. $\square$

17.3 AUXILIARY LEMMAS AND BOUNDS

In this section we collect a few auxiliary lemmas that will be needed in the proof of Proposition 17.1 in order to bound certain terms in the calculation.

Lemma 17.9. *G satisfies $G = \Pi_\theta G$ and $G = G(I - \Pi_\theta)$.*

Proof. Rearranging (17.21), we have

$$G = \Pi_\theta \mathcal{T}^2 \Pi_\theta G - \Pi_\theta \mathcal{T}. \quad (17.46)$$

In view of this, $\Pi_\theta^2 = \Pi_\theta$ implies $G = \Pi_\theta G$. The second claim (equivalent to $G\Pi_\theta = 0$) follows immediately from (17.30b). $\square$

The following lemma is instrumental in guaranteeing that the modified entropy $\mathcal{H}_\varepsilon$ is well defined.

Lemma 17.10. *G satisfies the bounds*

$$\|Gf\|_{L^2(\widehat{\pi})} \leq \frac{1}{2} \|(I - \Pi_\theta)f\|_{L^2(\widehat{\pi})}, \quad f \in L^2(\widehat{\pi}) \quad (17.47)$$

and

$$\|\mathcal{T}Gf\|_{L^2(\widehat{\pi})} \leq \|(I - \Pi_\theta)f\|_{L^2(\widehat{\pi})}, \quad f \in L^2(\widehat{\pi}). \quad (17.48)$$

Remark 17.4. *The estimate (17.47) immediately implies that G is bounded with norm (at most) $\frac{1}{2}$, proving Lemmas 17.3 and 17.4.*

Proof of Lemma 17.10. For $f \in \mathcal{D}$, it follows from (17.21) that

$$Gf + (\mathcal{T}\Pi_\theta)^*(\mathcal{T}\Pi_\theta)Gf = (\mathcal{T}\Pi_\theta)^*f. \quad (17.49)$$

Taking the scalar product of this equation with Gf , we see that

$$\|Gf\|_{L^2(\widehat{\pi})}^2 + \|\mathcal{T}Gf\|_{L^2(\widehat{\pi})}^2 = \langle \mathcal{T}Gf, (I - \Pi_\theta)f \rangle_{L^2(\widehat{\pi})}, \quad (17.50)$$

where on the left-hand side we have used the fact that $\Pi_\theta G = G$, according to Lemma 17.9, while the expression on the right-hand side rests on the fact that

$$\langle \mathcal{T}Gf, \Pi_\theta f \rangle_{L^2(\widehat{\pi})} = \langle \Pi_\theta \mathcal{T}\Pi_\theta Gf, f \rangle_{L^2(\widehat{\pi})} = 0, \quad (17.51)$$

using again $\Pi_\theta G = G$ as well as (17.30b). The identity (17.50) now implies the bounds

$$\|Gf\|_{L^2(\widehat{\pi})}^2 + \|\mathcal{T}Gf\|_{L^2(\widehat{\pi})}^2 \leq \|(I - \Pi_\theta)f\|_{L^2(\widehat{\pi})} \|\mathcal{T}Gf\|_{L^2(\widehat{\pi})} \quad (17.52a)$$

$$\leq \frac{1}{4} \|(I - \Pi_\theta)f\|_{L^2(\widehat{\pi})}^2 + \|\mathcal{T}Gf\|_{L^2(\widehat{\pi})}^2, \quad (17.52b)$$

using Cauchy-Schwarz's and Young's inequalities. The first of these bounds implies (17.48), whereas the second one implies (17.47). Finally, the same holds for all $f \in L^2(\widehat{\pi})$ by density of $\mathcal{D}$ in $L^2(\widehat{\pi})$. $\square$

The next lemma is needed in the proof of Lemma 17.12:

Lemma 17.11. *The following holds:*

$$\Pi_\theta \mathcal{T}^2(I - \Pi_\theta) = \Pi_\theta \sum_{\substack{i,j=1 \\ i \neq j}}^d \theta_i \theta_j \partial_i^* \partial_j^*. \quad (17.53)$$

Proof. We will show this identity by computing the adjoint

$$(\Pi_\theta \mathcal{T}^2(I - \Pi_\theta))^* = (I - \Pi_\theta) \mathcal{T}^2 \Pi_\theta.$$

To this end, let us calculate

$$\begin{aligned} \mathcal{T}^2 \Pi_\theta &= \frac{1}{2} \sum_{j=1}^d (2\theta_j \partial_j + \theta_j (\partial_j V)(R_j - I)) \left(\sum_{i=1}^d \theta_i \partial_i \Pi_\theta \right) \\ &= \sum_{i,j=1}^d \theta_i \theta_j \partial_i \partial_j \Pi_\theta + \frac{1}{2} \sum_{i,j=1}^d \theta_j (\partial_j V)(R_j - I) \theta_i \partial_i \Pi_\theta \\ &= \sum_{i,j=1}^d \theta_i \theta_j \partial_i \partial_j \Pi_\theta - \sum_{i=1}^d (\partial_i V) \partial_i \Pi_\theta. \end{aligned} \quad (17.54)$$

In going from the second to the third line, we used the fact that $(R_j - I)\theta_i = \theta_i(R_j - I)$ for $i \neq j$, as well as $(R_j - I)\Pi_\theta = 0$ and $\theta_i^2 = 1$. Using (17.54), we deduce that

$$(I - \Pi_\theta)\mathcal{T}^2\Pi_\theta = \sum_{i,j=1}^d \theta_i\theta_j\partial_i\partial_j\Pi_\theta - \Pi_\theta \sum_{i,j=1}^d \theta_i\theta_j\partial_i\partial_j\Pi_\theta = \sum_{\substack{i,j=1 \\ i \neq j}}^d \theta_i\theta_j\partial_i\partial_j\Pi_\theta, \quad (17.55)$$

using that for any $f \in \mathcal{D}$, $\partial_i\partial_j\Pi_\theta\phi$ does not depend on θ , leading to cancellations when Π_θ acts on the corresponding sum. Now taking the adjoint of (17.55) leads to the desired result. $\square$

Lemma 17.12. *There exists a constant $C_1 > 0$ such that*

$$\|G\mathcal{T}(I - \Pi_\theta)f\|_{L^2(\widehat{\pi})} \leq C_1\|(I - \Pi_\theta)f\|_{L^2(\widehat{\pi})}, \quad f \in L^2(\widehat{\pi}). \quad (17.56)$$

Proof. First observe that the bound (17.56) would immediately be implied by the boundedness of the operator $G\mathcal{T}(I - \Pi_\theta)$, i.e. the existence of a constant $\tilde{C} > 0$ such that

$$\|G\mathcal{T}(I - \Pi_\theta)f\|_{L^2(\widehat{\pi})} \leq \tilde{C}\|f\|_{L^2(\widehat{\pi})}, \quad f \in \mathcal{D}. \quad (17.57)$$

Indeed, (17.56) would follow by replacing f by $(I - \Pi_\theta)f$ in the above estimate and using the fact that $I - \Pi_\theta$ is a projection, i.e. $(I - \Pi_\theta)^2 = (I - \Pi_\theta)$. Moreover, the boundedness of $G\mathcal{T}(I - \Pi_\theta)$ is equivalent to the boundedness of its adjoint

$$-(I - \Pi_\theta)\mathcal{T}G^* = - \sum_{\substack{i,j=1 \\ i \neq j}}^d \theta_i\theta_j\partial_i\partial_j\Pi_\theta \left(I + \sum_{k=1}^d \partial_k^*\partial_k \right)^{-1}. \quad (17.58)$$

By elliptic regularity ([53, Chapter 7]), $\left(1 + \sum_{k=1}^d \partial_k^*\partial_k\right)^{-1}$ is a bounded operator from $L^2(\pi)$ to $H^2(\pi)$. Clearly, $\Pi_\theta \left(I + \sum_{k=1}^d \partial_k^*\partial_k \right)^{-1}$ is therefore bounded from $L^2(\widehat{\pi})$ to $H^2(\pi)$. From the definition of $H^2(\pi)$, we see immediately that

$$- \sum_{\substack{i,j=1 \\ i \neq j}}^d \theta_i\theta_j\partial_i\partial_j$$

is bounded from $H^2(\pi)$ to $L^2(\widehat{\pi})$. In summary, this shows the $L^2(\widehat{\pi})$ -boundedness of the operator appearing in (17.58), leading to the advertised claim. $\square$

Lemma 17.13. *There exists a constant $C_2 > 0$ such that*

$$\|G\mathcal{S}f\|_{L^2(\widehat{\pi})} \leq C_2\|(I - \Pi_\theta)f\|_{L^2(\widehat{\pi})}, \quad f \in \mathcal{D}. \quad (17.59)$$

Proof. We start with the following preliminary calculation, using the representation intro-

duced in (17.29):

$$\begin{aligned}
\Pi_\theta \mathcal{T} \mathcal{S} &= -\Pi_\theta \sum_{i,j=1}^d (A_i^* \partial_i - \partial_i^* A_i) (|\partial_j V| + 2\gamma_j) A_j^* A_j \\
&= \Pi_\theta \sum_{i,j=1}^d \partial_i^* A_i (|\partial_j V| + 2\gamma_j) A_j^* A_j \\
&= \Pi_\theta \sum_{i=1}^d \partial_i^* A_i (|\partial_i V| + 2\gamma_i).
\end{aligned} \tag{17.60}$$

In going from the first to the second line, we have used the representation of Π_θ in (17.18), as well as the fact that $(R_i + I)A_i^* = 0$. The same identity shows that the terms with $i \neq j$ vanish. From the second to the third line we used furthermore that $A_i A_i^* A_i = A_i$. We wish to emphasize that the derivatives of $|\partial_j V|$ are to be interpreted in a weak sense. Noting that $\Pi_\theta \mathcal{T} = \Pi_\theta \sum_{i=1}^d \partial_i^* A_i$ and applying $(1 + (\mathcal{T}\Pi_\theta)^*(\mathcal{T}\Pi_\theta))^{-1}$ on both sides of (17.60), the required bound follows from the definition of G in (17.21), the boundedness of the function $|\partial_i V| + 2\gamma_i$ and the estimate (17.47). $\square$

Remark 17.5. *Our motivation to consider the problem on the compact state space $\mathbb{T}^d$ derives from the fact that otherwise the function $|\partial_i V| + 2\gamma_i$ might be unbounded (especially in order to guarantee a Poincaré inequality of the form (17.41)) and thus the conclusion of the preceding lemma would not be justified. In the unbounded case, it might be possible to establish Lemma 17.13 using finer regularity properties of G .*

17.4 COERCIVITY OF THE ENTROPY DISSIPATION FUNCTIONAL

We are now in a position to join all the preparatory results from Sections 17.2 and 17.3.

Proof of Proposition 17.1. We have

$$\begin{aligned}
D(f) &= -\langle \mathcal{S}f, f \rangle_{L^2(\widehat{\pi})} + \varepsilon \langle G\mathcal{T}\phi, f \rangle_{L^2(\widehat{\pi})} + \varepsilon \langle G\mathcal{S}f, f \rangle_{L^2(\widehat{\pi})} \\
&\quad - \varepsilon \langle \mathcal{T}Gf, f \rangle_{L^2(\widehat{\pi})} + \varepsilon \langle \mathcal{S}Gf, f \rangle_{L^2(\widehat{\pi})}.
\end{aligned} \tag{17.61}$$

The first term in (17.61) can be controlled using microscopic coercivity (Lemma 17.6). For the second term, write

$$\langle G\mathcal{T}f, f \rangle_{L^2(\widehat{\pi})} = \langle G\mathcal{T}(I - \Pi_\theta)f, f \rangle_{L^2(\widehat{\pi})} + \langle G\mathcal{T}\Pi_\theta f, f \rangle_{L^2(\widehat{\pi})}, \tag{17.62}$$

noting that the two parts can be controlled using Lemma 17.12 and Lemma 17.8 (macroscopic coercivity), respectively. The third term in (17.61) can be addressed using Lemma 17.13, and the fourth term using Lemma 17.10. The fifth term vanishes due to $\mathcal{S}G =$

$\mathcal{S}\Pi_\theta G = 0$, according to (17.30a). Putting everything together, we arrive at

$$\begin{aligned} D(f) &\geq (\delta - \varepsilon) \|(I - \Pi_\theta)f\|_{L^2(\widehat{\pi})}^2 - \varepsilon(C_1 + C_2) \|(I - \Pi_\theta)f\|_{L^2(\widehat{\pi})} \|\Pi_\theta f\|_{L^2(\widehat{\pi})} + \varepsilon\Lambda \|\Pi_\theta f\|_{L^2(\widehat{\pi})}^2 \\ &= X^T \mathbf{S} X, \end{aligned}$$

introducing the notation

$$X = \begin{pmatrix} \|\Pi_\theta f\|_{L^2(\widehat{\pi})} \\ \|(1 - \Pi_\theta)f\|_{L^2(\widehat{\pi})} \end{pmatrix}, \quad \mathbf{S} = \begin{pmatrix} S_{--} & S_{-+}/2 \\ S_{-+}/2 & S_{++} \end{pmatrix}, \quad (17.64)$$

where

$$S_{--} = \varepsilon\Lambda, \quad S_{-+} = -\varepsilon(C_1 + C_2), \quad S_{++} = \delta - \varepsilon. \quad (17.65)$$

The smallest eigenvalue of $\mathbf{S}$ is given by

$$\mu(\varepsilon) = \frac{S_{--} + S_{++}}{2} - \frac{1}{2} \sqrt{(S_{--} - S_{++})^2 + (S_{-+})^2}. \quad (17.66)$$

The claim now follows by noting that ε can be chosen such that $\mu(\varepsilon) > 0$ (for instance by observing that $\mu(0) = 0$ and $\mu'(0) = \Lambda > 0$). $\square$

17.5 DISCUSSION

In this chapter we showed how the hypocoercivity techniques developed in [40] can be employed for the analysis of ergodic processes of the zigzag process. However, weakening the assumptions that the state space is compact and the excess switching rate does not vanish in certain regions remains challenging. Our work can be extended in several directions. Firstly, our approach seems to hold promise for addressing the ergodicity of piecewise deterministic Markov processes in a more general setting (as outlined for example in [139]). Following up on this, it would be possible to derive more explicit estimates on the constants C and λ appearing in (17.6), hopefully leading to more principled guidelines for choosing a particular piecewise deterministic Markov process in practice.

REFERENCES

- [1] I. Amemiya and K. Shiga. On tensor products of Banach spaces. *Kōdai Math. Sem. Rep.*, 9:161–178, 1957.
- [2] C. Andrieu, A. Doucet, and R. Holenstein. Particle Markov chain Monte Carlo methods. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, 72(3):269–342, 2010.
- [3] L. Angiuli, G. Metafune, and C. Spina. Feller semigroups and invariant measures. *Riv. Math. Univ. Parma (N.S.)*, 1(2):347–406, 2010.
- [4] W. Arendt, A. Grabosch, G. Greiner, U. Groh, H. P. Lotz, U. Moustakas, R. Nagel, F. Neubrander, and U. Schlotterbeck. *One-parameter semigroups of positive operators*, volume 1184 of *Lecture Notes in Mathematics*. Springer-Verlag, Berlin, 1986.
- [5] A. Arnold, A. Einav, and T. Wöhrer. On the rates of decay to equilibrium in degenerate and defective Fokker–Planck equations. *J. Differential Equations*, 264(11):6843–6872, 2018.
- [6] A. Arnold and J. Erb. Sharp entropy decay for hypocoercive and non-symmetric Fokker-Planck equations with linear drift. *arXiv:1409.5425v2*, 2014.
- [7] S. Asmussen and P. W. Glynn. *Stochastic simulation: algorithms and analysis*, volume 57 of *Stochastic Modelling and Applied Probability*. Springer, New York, 2007.
- [8] R. Assaraf, B. Jourdain, T. Lelièvre, and R. Roux. Computation of sensitivities for the invariant measure of a parameter dependent diffusion. *Stochastics and Partial Differential Equations: Analysis and Computations*, pages 1–59, 2015.
- [9] D. Bakry, P. Cattiaux, and A. Guillin. Rate of convergence for ergodic continuous Markov processes: Lyapunov versus Poincaré. *J. Funct. Anal.*, 254(3):727–759, 2008.
- [10] D. Bakry, I. Gentil, and M. Ledoux. *Analysis and geometry of Markov diffusion operators*, volume 348. Springer Science & Business Media, 2013.
- [11] F. Baudoin. Wasserstein contraction properties for hypoelliptic diffusions. *arXiv:1602.04177*, 2016.

- [12] P. H. Baxendale. Statistical equilibrium and two-point motion for a stochastic flow of diffeomorphisms. In *Spatial stochastic processes*, volume 19 of *Progr. Probab.*, pages 189–218. Birkhäuser Boston, Boston, MA, 1991.
- [13] C. H. Bennett. Mass tensor molecular dynamics. *Journal of Computational Physics*, 19(3):267 – 279, 1975.
- [14] A. Beskos, N. Pillai, G. Roberts, J.-M. Sanz-Serna, A. Stuart, et al. Optimal tuning of the hybrid Monte Carlo algorithm. *Bernoulli*, 19(5A):1501–1534, 2013.
- [15] A. Beskos, F. J. Pinski, J. M. Sanz-Serna, and A. M. Stuart. Hybrid Monte Carlo on Hilbert spaces. *Stochastic Process. Appl.*, 121(10):2201–2230, 2011.
- [16] A. Beskos, G. Roberts, A. Stuart, and J. Voss. MCMC methods for diffusion bridges. *Stoch. Dyn.*, 8(3):319–350, 2008.
- [17] A. Beskos and A. Stuart. MCMC methods for sampling function space. In *ICIAM 07—6th International Congress on Industrial and Applied Mathematics*, pages 337–364. Eur. Math. Soc., Zürich, 2009.
- [18] R. Bhatia. *Matrix analysis*, volume 169 of *Graduate Texts in Mathematics*. Springer-Verlag, New York, 1997.
- [19] R. N. Bhattacharya. On the functional central limit theorem and the law of the iterated logarithm for Markov processes. *Z. Wahrsch. Verw. Gebiete*, 60(2):185–201, 1982.
- [20] J. Bierkens. Non-reversible Metropolis-Hastings. *Statistics and Computing*, 26(6):1213–1228, 2016.
- [21] J. Bierkens and A. Duncan. Limit theorems for the zig-zag process. *Adv. in Appl. Probab.*, 49(3):791–825, 2017.
- [22] J. Bierkens, P. Fearnhead, and G. Roberts. The zig-zag process and super-efficient sampling for Bayesian analysis of big data. *arXiv:1607.03188*, 2016.
- [23] J. Bierkens, G. Roberts, and P.-A. Zitt. Ergodicity of the zigzag process. *arXiv:1712.09875*, 2018.
- [24] V. I. Bogachev, N. V. Krylov, M. Röckner, and S. V. Shaposhnikov. *Fokker-Planck-Kolmogorov equations*, volume 207 of *Mathematical Surveys and Monographs*. American Mathematical Society, Providence, RI, 2015.
- [25] N. Bonneel, M. Van De Panne, S. Paris, and W. Heidrich. Displacement interpolation using Lagrangian mass transport. In *ACM Transactions on Graphics (TOG)*, volume 30, page 158. ACM, 2011.

- [26] B. Böttcher, R. Schilling, and J. Wang. *Lévy matters. III*, volume 2099 of *Lecture Notes in Mathematics*. Springer, Cham, 2013. Lévy-type processes: construction, approximation and sample path properties, With a short biography of Paul Lévy by Jean Jacod, Lévy Matters.
- [27] N. Bou-Rabee and J. M. Sanz-Serna. Randomized Hamiltonian Monte Carlo. *Ann. Appl. Probab.*, 27(4):2159–2194, 2017.
- [28] A. Bouchard-Côté, S. J. Vollmer, and A. Doucet. The bouncy particle sampler: A non-reversible rejection-free Markov chain Monte Carlo method. *Journal of the American Statistical Association*, 0(ja):0–0, 2017.
- [29] E. Bouin, J. Dolbeault, S. Mischler, C. Mouhot, and C. Schmeiser. Hypocoercivity without confinement. *arXiv preprint arXiv:1708.06180*, 2017.
- [30] S. Boyd and L. Vandenberghe. *Convex optimization*. Cambridge University Press, Cambridge, 2004.
- [31] S. Brooks, A. Gelman, G. Jones, and X.-L. Meng. *Handbook of Markov chain Monte Carlo*. CRC press, 2011.
- [32] M.-F. Chen. *Eigenvalues, inequalities, and ergodic theory*. Probability and its Applications (New York). Springer-Verlag London, Ltd., London, 2005.
- [33] C. Cotar, G. Friesecke, and C. Klüppelberg. Density functional theory and optimal transportation with Coulomb cost. *Comm. Pure Appl. Math.*, 66(4):548–599, 2013.
- [34] P. Courrège. Sur la forme intégrro-différentielle des opérateurs de C_k^∞ dans C satisfaisant au principe du maximum. *Séminaire Brelot-Choquet-Deny. Théorie du potentiel*, 10(1):1–38, 1965-1966.
- [35] G. Da Prato and J. Zabczyk. *Ergodicity for infinite-dimensional systems*, volume 229 of *London Mathematical Society Lecture Note Series*. Cambridge University Press, Cambridge, 1996.
- [36] M. H. A. Davis. Piecewise-deterministic Markov processes: a general class of non-diffusion stochastic models. *J. Roy. Statist. Soc. Ser. B*, 46(3):353–388, 1984. With discussion.
- [37] M. H. A. Davis. *Markov models and optimization*, volume 49 of *Monographs on Statistics and Applied Probability*. Chapman & Hall, London, 1993.
- [38] Pierre Del Moral. *Mean field simulation for Monte Carlo integration*, volume 126 of *Monographs on Statistics and Applied Probability*. CRC Press, Boca Raton, FL, 2013.

- [39] P. Dellaportas and I. Kontoyiannis. Control variates for estimation based on reversible Markov chain Monte Carlo samplers. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, 74(1):133–161, 2012.
- [40] J. Dolbeault, C. Mouhot, and C. Schmeiser. Hypocoercivity for linear kinetic equations conserving mass. *Trans. Amer. Math. Soc.*, 367(6):3807–3828, 2015.
- [41] A. B. Duncan, T. Lelièvre, and G. A. Pavliotis. Variance reduction using nonreversible Langevin samplers. *J. Stat. Phys.*, 163(3):457–491, 2016.
- [42] A. B. Duncan, N. Nüsken, and G. A. Pavliotis. Using perturbed underdamped Langevin dynamics to efficiently sample from probability distributions. *J. Stat. Phys.*, 169(6):1098–1131, 2017.
- [43] A. B. Duncan, G. A. Pavliotis, and K. C. Zygalakis. Nonreversible Langevin samplers: Splitting schemes, analysis and implementation. *arXiv preprint arXiv:1701.04247*, 2017.
- [44] P. Dupuis, Y. Liu, N. Plattner, and J. D. Doll. On the infinite swapping limit for parallel tempering. *Multiscale Modeling & Simulation*, 10(3):986–1022, 2012.
- [45] A. Eberle. Stochastic analysis. https://wt.iam.uni-bonn.de/fileadmin/WT/Inhalt/people/Andreas_Eberle/StoAn15/StochasticAnalysis2015.pdf, 2015. Lecture Notes, accessed 12/03/2018.
- [46] A. Eberle. Reflection couplings and contraction rates for diffusions. *Probability theory and related fields*, 166(3-4):851–886, 2016.
- [47] K.-J. Engel and R. Nagel. *One-parameter semigroups for linear evolution equations*, volume 194 of *Graduate Texts in Mathematics*. Springer-Verlag, New York, 2000. With contributions by S. Brendle, M. Campiti, T. Hahn, G. Metafune, G. Nickel, D. Pallara, C. Perazzoli, A. Rhandi, S. Romanelli and R. Schnaubelt.
- [48] S. N. Ethier and T. G. Kurtz. *Markov processes*. Wiley Series in Probability and Mathematical Statistics: Probability and Mathematical Statistics. John Wiley & Sons, Inc., New York, 1986. Characterization and convergence.
- [49] G. L. Eyink, J. L. Lebowitz, and H. Spohn. Hydrodynamics and fluctuations outside of local equilibrium: driven diffusive systems. *J. Statist. Phys.*, 83(3-4):385–472, 1996.
- [50] P. Fearnhead, J. Bierkens, M. Pollock, and G. O. Roberts. Piecewise deterministic Markov processes for continuous-time Monte Carlo. *arXiv preprint arXiv:1611.07873*, 2016.

- [51] A. Gelman, J. B. Carlin, H. S. Stern, D. B. Dunson, A. Vehtari, and D. B. Rubin. *Bayesian data analysis*. Texts in Statistical Science Series. CRC Press, Boca Raton, FL, third edition, 2014.
- [52] A. L. Gibbs and F. E. Su. On choosing and bounding probability metrics. *International statistical review*, 70(3):419–435, 2002.
- [53] 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.
- [54] M. B. Giles. Multilevel Monte Carlo methods. *Acta Numer.*, 24:259–328, 2015.
- [55] M. Girolami and B. Calderhead. Riemann manifold Langevin and Hamiltonian Monte Carlo methods. *J. R. Stat. Soc. Ser. B Stat. Methodol.*, 73(2):123–214, 2011. With discussion and a reply by the authors.
- [56] C. R. Givens, R. M. Shortt, et al. A class of Wasserstein metrics for probability distributions. *The Michigan Mathematical Journal*, 31(2):231–240, 1984.
- [57] P. W. Glynn and S. P. Meyn. A Liapounov bound for solutions of the Poisson equation. 24(2), 1996.
- [58] I. Goodfellow, Y. Bengio, A. Courville, and Y. Bengio. *Deep learning*, volume 1. MIT press Cambridge, 2016.
- [59] M. Grothaus and P. Stilgenbauer. Hypocoercivity for Kolmogorov backward evolution equations and applications. *J. Funct. Anal.*, 267(10):3515–3556, 2014.
- [60] M. Grothaus and P. Stilgenbauer. A hypocoercivity related ergodicity method for singularly distorted non-symmetric diffusions. *Integral Equations Operator Theory*, 83(3):331–379, 2015.
- [61] M. Grothaus and P. Stilgenbauer. Hilbert space hypocoercivity for the Langevin dynamics revisited. *Methods Funct. Anal. Topology*, 22(2):152–168, 2016.
- [62] A. Guillin and P. Monmarché. Optimal linear drift for the speed of convergence of an hypoelliptic diffusion. *Electron. Commun. Probab.*, 21:Paper No. 74, 14, 2016.
- [63] M. Hairer and J. C. Mattingly. Yet another look at Harris’ ergodic theorem for Markov chains. In *Seminar on Stochastic Analysis, Random Fields and Applications VI*, volume 63 of *Progr. Probab.*, pages 109–117. Birkhäuser/Springer Basel AG, Basel, 2011.
- [64] M. Hairer, A. Stuart, and J. Voss. Sampling conditioned diffusions. In *Trends in stochastic analysis*, volume 353 of *London Math. Soc. Lecture Note Ser.*, pages 159–185. Cambridge Univ. Press, Cambridge, 2009.

- [65] M. Hairer, A. M. Stuart, and J. Voss. Analysis of SPDEs arising in path sampling. II. The nonlinear case. *Ann. Appl. Probab.*, 17(5-6):1657–1706, 2007.
- [66] M. Hairer, A. M. Stuart, J. Voss, and P. Wiberg. Analysis of SPDEs arising in path sampling. I. The Gaussian case. *Commun. Math. Sci.*, 3(4):587–603, 2005.
- [67] E. J. Hall, M. A. Katsoulakis, and L. Rey-Bellet. Uncertainty quantification for generalized Langevin dynamics. *The Journal of chemical physics*, 145(22):224108, 2016.
- [68] J. Heng, A. Doucet, and Y. Pokern. Gibbs flow for approximate transport with applications to Bayesian computation. *arXiv preprint arXiv:1509.08787*, 2015.
- [69] M. R. Hestenes. Relative self-adjoint operators in Hilbert space. *Pacific J. Math.*, 11:1315–1357, 1961.
- [70] L. Hörmander. Hypoelliptic second order differential equations. *Acta Mathematica*, 119(1):147–171, 1967.
- [71] C.-R. Hwang, S.-Y. Hwang-Ma, and S.-J. Sheu. Accelerating Gaussian diffusions. *Ann. Appl. Probab.*, 3(3):897–913, 1993.
- [72] C.-R. Hwang, S.-Y. Hwang-Ma, and S.-J. Sheu. Accelerating diffusions. *Ann. Appl. Probab.*, 15(2):1433–1444, 2005.
- [73] C.-R. Hwang, R. Normand, and S.-J. Wu. Variance reduction for diffusions. *Stochastic Process. Appl.*, 125(9):3522–3540, 2015.
- [74] A. Iacobucci, S. Olla, and G. Stoltz. Convergence rates for nonequilibrium Langevin dynamics. *Annales mathématiques du Québec*, pages 1–26, 2017.
- [75] N. Jacob. *Pseudo Differential Operators & Markov Processes: Markov Processes And Applications*, volume 3. Imperial College Press, 2001.
- [76] N. Jacob. *Pseudo differential operators and Markov processes. Vol. I*. Imperial College Press, London, 2001. Fourier analysis and semigroups.
- [77] D.-Q. Jiang, M. Qian, and M.-P. Qian. *Mathematical theory of nonequilibrium steady states: on the frontier of probability and dynamical systems*. Springer, 2003.
- [78] A. Joulin and Y. Ollivier. Curvature, concentration and error estimates for Markov chain Monte Carlo. *Ann. Probab.*, 38(6):2418–2442, 2010.
- [79] O. Kallenberg. *Foundations of modern probability*. Probability and its Applications (New York). Springer-Verlag, New York, second edition, 2002.
- [80] T. Kato. *Perturbation theory for linear operators; 2nd ed.* Grundlehren Math. Wiss. Springer, Berlin, 1976.

- [81] J. Y. Kazakia. Orthogonal transformation of a trace free symmetric matrix into one with zero diagonal elements. *Internat. J. Engrg. Sci.*, 26(8):903–906, 1988.
- [82] W. Kliemann. Recurrence and invariant measures for degenerate diffusions. *The Annals of Probability*, pages 690–707, 1987.
- [83] D. P. Kroese, T. Taimre, and Z. I. Botev. *Handbook of Monte Carlo methods*, volume 706. John Wiley & Sons, 2013.
- [84] F. Kühn. *Lévy matters. VI*, volume 2187 of *Lecture Notes in Mathematics*. Springer, Cham, 2017. Lévy-type processes: moments, construction and heat kernel estimates, With a short biography of Paul Lévy by Jean Jacod, Lévy Matters.
- [85] F. Kühn. Existence of (Markovian) solutions to martingale problems associated with Lévy-type operators. *arXiv:1803.05646*, 2018.
- [86] B. Leimkuhler and C. Matthews. *Molecular dynamics*, volume 39 of *Interdisciplinary Applied Mathematics*. Springer, Cham, 2015. With deterministic and stochastic numerical methods.
- [87] B. Leimkuhler, C. Matthews, and G. Stoltz. The computation of averages from equilibrium and nonequilibrium Langevin molecular dynamics. *IMA J. Numer. Anal.*, 36(1):13–79, 2016.
- [88] B. Leimkuhler, C. Matthews, and J. Weare. Ensemble preconditioning for Markov chain Monte Carlo simulation. *Stat. Comput.*, 28(2):277–290, 2018.
- [89] T. Lelièvre, F. Nier, and G. A. Pavliotis. Optimal non-reversible linear drift for the convergence to equilibrium of a diffusion. *J. Stat. Phys.*, 152(2):237–274, 2013.
- [90] T. Lelièvre, M. Rousset, and G. Stoltz. *Free energy computations*. Imperial College Press, London, 2010. A mathematical perspective.
- [91] T. Lelièvre and G. Stoltz. Partial differential equations and stochastic methods in molecular dynamics. *Acta Numer.*, 25:681–880, 2016.
- [92] V. Lemaire, G. Pagès, and F. Panloup. Invariant measure of duplicated diffusions and application to Richardson-Romberg extrapolation. *Ann. Inst. Henri Poincaré Probab. Stat.*, 51(4):1562–1596, 2015.
- [93] T. Lindvall. *Lectures on the coupling method*. Dover Publications, Inc., Mineola, NY, 2002. Corrected reprint of the 1992 original.
- [94] J. S. Liu. *Monte Carlo strategies in scientific computing*. Springer Science & Business Media, 2008.

- [95] Q. Liu and D. Wang. Stein variational gradient descent: A general purpose Bayesian inference algorithm. In *Advances In Neural Information Processing Systems*, pages 2378–2386, 2016.
- [96] S. Livingstone and M. Girolami. Information-geometric Markov chain Monte Carlo methods using diffusions. *Entropy*, 16(6):3074–3102, 2014.
- [97] L. Lorenzi and M. Bertoldi. *Analytical methods for Markov semigroups*, volume 283 of *Pure and Applied Mathematics (Boca Raton)*. Chapman & Hall/CRC, Boca Raton, FL, 2007.
- [98] Y.-A. Ma, T. Chen, and E. Fox. A complete recipe for stochastic gradient MCMC. In *Advances in Neural Information Processing Systems*, pages 2899–2907, 2015.
- [99] M. C. Mackey. *Time’s arrow: the origins of thermodynamic behavior*. Courier Corporation, 2003.
- [100] R. J. McCann, B. Pass, and M. Warren. Rectifiability of optimal transportation plans. *Canad. J. Math.*, 64(4):924–934, 2012.
- [101] G. Metafune, D. Pallara, and E. Priola. Spectrum of Ornstein-Uhlenbeck operators in L^p spaces with respect to invariant measures. *J. Funct. Anal.*, 196(1):40–60, 2002.
- [102] M. Michel, S. C. Kapfer, and W. Krauth. Generalized event-chain Monte Carlo: Constructing rejection-free global-balance algorithms from infinitesimal steps. *Journal of Chemical Physics*, 140(5):054116, 2014.
- [103] M. Michel and S. S  n  cal. Forward Event-Chain Monte Carlo: a general rejection-free and irreversible Markov chain simulation method. *M2AN*, 2017.
- [104] A. Mijatovi   and J. Vogrinc. On the Poisson equation for Metropolis–Hastings chains. *Bernoulli*, 24(3):2401–2428, 2018.
- [105] A. Mira et al. Ordering and improving the performance of Monte Carlo Markov chains. *Statistical Science*, 16(4):340–350, 2001.
- [106] R. M. Neal. Circularly-coupled Markov chain sampling. *arXiv preprint arXiv:1711.04399*, 2017.
- [107] B.   ksendal. *Stochastic differential equations*. Universitext. Springer-Verlag, Berlin, sixth edition, 2003. An introduction with applications.
- [108] M. Ottobre. Markov chain Monte Carlo and irreversibility. *Reports on Mathematical Physics*, 77(3):267–292, 2016.
- [109] M. Ottobre and G. A. Pavliotis. Asymptotic analysis for the generalized Langevin equation. *Nonlinearity*, 24(5):1629, 2011.

- [110] M. Ottobre, G. A. Pavliotis, and K. Pravda-Starov. Exponential return to equilibrium for hypoelliptic quadratic systems. *J. Funct. Anal.*, 262(9):4000–4039, 2012.
- [111] M. Ottobre, G. A. Pavliotis, and K. Pravda-Starov. Some remarks on degenerate hypoelliptic Ornstein–Uhlenbeck operators. *J. Math. Anal. Appl.*, 429(2):676–712, 2015.
- [112] M. Ottobre, N. S. Pillai, F. J. Pinski, and A. M. Stuart. A function space HMC algorithm with second order Langevin diffusion limit. *Bernoulli*, 22(1):60–106, 2016.
- [113] G. A. Pavliotis. *Stochastic processes and applications: diffusion processes, the Fokker-Planck and Langevin equations*, volume 60. Springer, 2014.
- [114] G. A. Pavliotis and A. M. Stuart. White noise limits for inertial particles in a random field. *Multiscale Model. Simul.*, 1(4):527–533 (electronic), 2003.
- [115] G. A. Pavliotis and A. M. Stuart. Analysis of white noise limits for stochastic systems with two fast relaxation times. *Multiscale Model. Simul.*, 4(1):1–35 (electronic), 2005.
- [116] A. Pazy. *Semigroups of linear operators and applications to partial differential equations*, volume 44 of *Applied Mathematical Sciences*. Springer-Verlag, New York, 1983.
- [117] J. Peetre. Une caractérisation abstraite des opérateurs différentiels. *Math. Scand.*, 7:211–218, 1959.
- [118] G. Peyré, M. Cuturi, et al. Computational optimal transport. Technical report, 2017.
- [119] R. Pinnau, C. Totzeck, O. Tse, and S. Martin. A consensus-based model for global optimization and its mean-field limit. *Mathematical Models and Methods in Applied Sciences*, 27(01):183–204, 2017.
- [120] R. Poncet. Generalized and hybrid Metropolis-Hastings overdamped Langevin algorithms. *arXiv:1701.05833*, 2017.
- [121] J. Propp and D. Wilson. Coupling from the past: a user’s guide. In *Microsurveys in discrete probability (Princeton, NJ, 1997)*, volume 41 of *DIMACS Ser. Discrete Math. Theoret. Comput. Sci.*, pages 181–192. Amer. Math. Soc., Providence, RI, 1998.
- [122] Sebastian Reich and Colin Cotter. *Probabilistic forecasting and Bayesian data assimilation*. Cambridge University Press, New York, 2015.
- [123] C. R. Reeves and J. E. Rowe. *Genetic algorithms—principles and perspectives*, 2004.

- [124] L. Rey-Bellet. Ergodic properties of Markov processes. In *Open quantum systems II*, pages 1–39. Springer, 2006.
- [125] L. Rey-Bellet and K. Spiliopoulos. Irreversible Langevin samplers and variance reduction: a large deviations approach. *Nonlinearity*, 28(7):2081–2103, 2015.
- [126] L. Rey-Bellet and K. Spiliopoulos. Variance reduction for irreversible Langevin samplers and diffusion on graphs. *Electron. Commun. Probab.*, 20:no. 15, 16, 2015.
- [127] L. Rey-Bellet and K. Spiliopoulos. Improving the convergence of reversible samplers. *J. Stat. Phys.*, 164(3):472–494, 2016.
- [128] C. Robert and G. Casella. *Monte Carlo statistical methods*. Springer Science & Business Media, 2013.
- [129] G. O. Roberts, J. S. Rosenthal, et al. Optimal scaling for various Metropolis-Hastings algorithms. *Statistical science*, 16(4):351–367, 2001.
- [130] J. Roussel and G. Stoltz. Spectral methods for Langevin dynamics and associated error estimates. *ESAIM: Mathematical Modelling and Numerical Analysis*, 2017.
- [131] M. Rousset and G. Stoltz. An interacting particle system approach for molecular dynamics. preprint, August 2005.
- [132] H. Rue. Fast sampling of Gaussian Markov random fields. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, 63(2):325–338, 2001.
- [133] N. Sandrić. A note on the Birkhoff ergodic theorem. *Results in Mathematics*, 72(1-2):715–730, 2017.
- [134] R. Schilling. Conservativeness and extensions of Feller semigroups. *Positivity*, 2(3):239–256, 1998.
- [135] A. M. Stuart. Inverse problems: a Bayesian perspective. *Acta Numerica*, 19:451–559, 2010.
- [136] G. Teschl. *Mathematical methods in quantum mechanics*, volume 157 of *Graduate Studies in Mathematics*. American Mathematical Society, Providence, RI, second edition, 2014. With applications to Schrödinger operators.
- [137] H. Thorisson. *Coupling, stationarity, and regeneration*. Probability and its Applications (New York). Springer-Verlag, New York, 2000.
- [138] K. S. Turitsyn, M. Chertkov, and M. Vucelja. Irreversible Monte Carlo algorithms for efficient sampling. *Physica D: Nonlinear Phenomena*, 240(4-5):410–414, 2011.
- [139] P. Vanetti, A. Bouchard-Côté, G. Deligiannidis, and A. Doucet. Piecewise deterministic Markov chain Monte Carlo. *arXiv:1707.05296*, 2017.

- [140] C. Villani. *Topics in optimal transportation*, volume 58 of *Graduate Studies in Mathematics*. American Mathematical Society, Providence, RI, 2003.
- [141] C. Villani. *Hypocoercivity*. Number 949-951. American Mathematical Soc., 2009.
- [142] C. Villani. *Optimal transport*, volume 338 of *Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences]*. Springer-Verlag, Berlin, 2009. Old and new.
- [143] P. Walters. *An introduction to ergodic theory*, volume 79. Springer Science & Business Media, 2000.
- [144] S.-J. Wu, C.-R. Hwang, and M. T. Chu. Attaining the optimal Gaussian diffusion acceleration. *J. Stat. Phys.*, 155(3):571–590, 2014.