Technische Universität München
TUM School of Computation, Information and Technology

Analysis of a Consensus-based Optimization Method on Hypersurfaces and Applications

Philippe Sünnen

Vollständiger Abdruck der von der TUM School of Computation, Information and Technology der Technischen Universität München zur Erlangung des akademischen Grades eines

Doktors der Naturwissenschaften (Dr. rer. nat)

genehmigten Dissertation.

Vorsitzender:		Prof. Dr. Michael Ulbrich
Prüfer der Dissertation:	1.	Prof. Dr. Massimo Fornasier
	2.	Prof. Dr. Lorenzo Pareschi
	3.	Prof. Dr. Michael Herty

Die Dissertation wurde am 8. April 2022 bei der Technischen Universität München eingereicht und durch die Fakultät für Mathematik am 24. August 2022 angenommen.

Acknowledgements

I express my sincere gratitude to the FNR Luxembourg and Prof. Massimo Fornasier. In particular, I thank the FNR for supporting my PhD project financially with an AFR PhD Grant (Id: 12434809) and for being on my side during my troublesome hiring procedure. I will never forget this support and understanding. I genuinely thank Prof. Massimo Fornasier for introducing me to the fascinating world of consensus-based optimization methods and its applications, for giving me the opportunity to work at his chair and for kick-starting my PhD studies with three wonderful publications. These publications were joint work of Prof. Massimo Fornasier, Hui Huang, Prof. Lorenzo Pareschi and myself. Thanks a lot to Hui Huang and Prof. Lorenzo Pareschi for many insightful discussions and suggestions. It was a great pleasure to learn from all of you and I wish you all the best for your future.

Abstract

Constrained optimization problems with non-convex cost function are ubiquitous in science and engineering. One class of methods to solve such optimization problems are so-called metaheurisitics. A particularly interesting subclass of metaheuristics are consensus-based optimization (CBO) methods.

The topic of this thesis is the devlopment and analysis of a CBO method for non-convex optimization problems constrained on hypersurfaces (e.g. the sphere). We refer to the CBO method as KV-CBO as it is based on the Kuramoto and Vicsek models. A CBO method is a zero-order method in which many *particles* are placed on the graph of the cost function. The particles collaborate with each other in order to find the global minimizer of the cost function. This collaboration is modeled with a coupled system of stochastic differential equations which allows us to analyse it with tools from mean-field theory and PDE theory. We prove the well-posedness of the model, derive the mean-field equation and investigate the convergence of its solution towards the global minimizer. We further discuss implementation aspects and numerical experiments for common benchmark functions and three important machine learning problems: robust PCA, phase-retrieval, and reconstruction of neural nets.

In summary: in this thesis we analyse a CBO method for non-convex optimization problems constrained on hypersurfaces. We discuss the strengths and weaknesses of the method, and provide a proof of concept that it can be used to solve real-world problems.

Zusammenfassung

Optimierungsprobleme unter Nebenbedingungen und mit nicht-konvexer Kostenfunktion sind allgegenwärtig in den Natur- und Ingenieurwissenschaften. Eine Klasse von Methoden um solche Optimierungsprobleme näherungsweise zu lösen sind sogenannte Metaheuristiken. Eine interessante Teilklasse von Metaheuristiken sind Konsensbasierte Optimierungsmethoden (CBO).

In dieser Dissertation entwicklen wir eine CBO Methode um nicht-konvexe Optimierungsprobleme näherungsweise zu lösen. Dabei sind die Nebenbedingungen so, dass das Minimum auf einer gegebenen Hyperfläche (z.B. der Sphäre) liegt. Die hier entwickelte Methode bezeichnen wir mit KV-CBO weil diese auf den sogenannten Kuramoto und Vicsek Modellen basiert. Eine CBO Methode ist eine Optimierungsmethode nullter Ordnung. In einer CBO Methode werden viele *Partikel* auf dem Graphen der Kostenfunktion positioniert. Die Partikel kommunizieren untereinander um gemeinsam das globale Minimum der Kostenfunktion zu finden. Mathematisch wird so eine Dynamik mit einem gekoppelten System von stochastischen Differentialgleichungen beschrieben. Dies erlaubt es uns die Methode mit Hilfsmittels auf der mean-field Theorie und der Theorie der partiellen Differentialgleichungen zu untersuchen. Wir beweisen, dass das Modell korrektgestellt ist, leiten die mean-field Gleichung her und untersuchen die Konvergenz ihrer Lösung gegen das globale Minimum der Kostenfunktion. Weiter diskutieren wir eine Reihe numerischer Experimente für bekannte Benchmarkfunktionen. Wir betrachten auch die Anwendbarkeit auf das robust PCA Problem, das Phasenproblem und das Rekonstruieren von neuronalen Netzen.

Zusammengefasst: in dieser Dissertation analysieren wir eine CBO Methode für nicht-konvexe Kostenfunktionen auf Hyperflächen. Es werden die Stärken und Schwächen der Methode besprochen und wir zeigen, dass die Methode für praktische Anwendungen benutzt werden kann.

Contents

Notation

Parameters/ constants and cost function:

$\mathcal{E}$, $\overline{\mathcal{E}}$, $\underline{\mathcal{E}}$	cost function, its maximum and minimum
$\boldsymbol{\mathcal{E}}$, $\overline{\boldsymbol{\mathcal{E}}}$, $\underline{\boldsymbol{\mathcal{E}}}$	regularized cost function, its maximum and minimum
c_1, c_2, c_3, c_4	see definition 3.1.1
$V^\star$	global minimizer of $\mathcal{E}$ (not necessarily unique)
d	dimension of the Euclidean space $\mathbb{R}^d$
T	time horizon
λ	coupling constant
σ	diffusion constant
α	temparature parameter
C_α	see (3.54)
$\mathbf{C}_\alpha$	see (3.54)
$C_{\sigma,d}$	see (5.45)
C_σ	see (6.91)

Differential geometry (see section 3.4):

Γ	hypersurface
γ	level set function
Γ_δ	neighborhood of Γ
$\Pi_\Gamma(v)$	projection onto Γ
$P_\Gamma(v)$	projection onto orthogonal complement of $v \in \Gamma$
$T_v\Gamma$	tangent plane at $v \in \Gamma$
$n(v)$	outward unit normal at $v \in \Gamma$
$d(v)$	distance function
$\mathbb{S}^{d-1}$	unit sphere in $\mathbb{R}^d$
∂_i	partial derivative w.r.t ith variable
$\partial_{\Gamma,i}$	ith component of the tangential gradient on Γ
∇	gradient (w.r.t the space variable)
∇^2	Hessian

Δ	Laplace operator (or Laplacian)
∇_Γ	tangential gradient on Γ
Δ_Γ	Laplace Beltrami operator on Γ

Particle system and mean-field equation:

$V_t^i = V^i(t)$	ith particle at time t
ρ_t^N	empirical measure, see (3.45)
$V^\alpha(\rho_t^N)$	consensus point of the particles, see (3.46)
C_Γ	correction term for the isotropic model, see (3.44)
C_Γ^{ani}	correction term for the anisotropic model, see (6.94)
$\|V_t^{(N)}\|_1$	1-norm of the particles, see (3.55)
$\|V_t^{(N)}\|_\infty$	∞-norm of the particles, see (3.56)
$\rho_t(v) = \rho(t, v)$	solution of the mean-field equation
$V^\alpha(\rho_t)$	consensus point for ρ_t

Measure theory and probability/ statistics:

$(\Omega, \mathcal{A}, P)$	probability space
$\mathcal{P}(\Gamma)$	probability measures on Γ
$\mathcal{P}_p(\Gamma)$	probability measures on Γ with bounded pth moment
$\mathcal{P}_{ac}(\Gamma)$	absolutely continuous probability measures on Γ
$\rho^{\otimes N}$	tensorized probability measure, see section 2.2
B_t	Brownian motion
$\mathcal{A}$	infinitesimal generator or σ-algebra
law	law/ distribution of a random variable or stochastic process
$\sim$	distributed as
$\mathcal{U}(\Gamma)$	uniform distribution on Γ
$\mathcal{N}(\mu, \sigma^2)$	normal (or Gaussian) distribution
$\mathcal{N}_d$	d-dimensional normal (or Gaussian) distribution
$\mathbb{E}$	expected value/ expecation
$\mathbf{E}$	normalized expected value/ expecation, see (5.56)
V	variance
$\#$	push forward
$\mathcal{B}_{\mathbb{R}^d}$	Borel σ-algebra on $\mathbb{R}^d$
δ_v	Dirac delta concentrated at v
W_p	Wasserstein p-norm
Π	set of couplings between two probability measures
$\langle \cdot, \cdot \rangle$	dual pairing (or inner product)

Function classes

$C^k([0,T])$	k times continuously differntiable functions $[0,T] \to \mathbb{R}$
$C^k([0,T],\mathbb{R}^d)$	k times continuously differntiable functions $[0,T] \to \mathbb{R}^d$
$C_c^\infty(\mathbb{R}^d)$	smooth functions with compact support $\mathbb{R}^d \to \mathbb{R}$
$C^{k,\gamma}(\mathbb{R}^d)$	Hölder contiuous functions in C^k with Hölder exponent γ
$C^{0,1}(\mathbb{R}^d)$	Lipschitz continuous functions
$L^p(\mathbb{R}^d)$	Lebesgue space for $p \in [1,\infty]$,
	measurable functions $f : \mathbb{R}^d \to \mathbb{R}$ s.t. $\int_{\mathbb{R}^d} \lvert f(x)\rvert^p dx < \infty$
$L^p(\mu)$	Lebesgue space for $p \in [1,\infty]$ and measure μ,
	measurable functions $f : \mathbb{R}^d \to \mathbb{R}$ s.t. $\int_{\mathbb{R}^d} \lvert f(x)\rvert^p d\mu(x) < \infty$
$H^k(\mathbb{R}^d)$	Sobolev space for $k \in [0,\infty]$
$(H^k(\mathbb{R}^d))'$	dual space of $H^k(\mathbb{R}^d)$
$L^p([0,T];X)$	Bochner space, see definition 5.2.1
$W^{1,p,q}$	see theorem 5.2.1

Miscellaneous

$\mathbb{R}_+$	positive reals
$\lvert\cdot\rvert$	Euclidean norm
$\lVert\cdot\rVert_p$	p-norm for $p \in [1,\infty]$ (for $p=2$ we write $\lvert\cdot\rvert$)
$\lVert\cdot\rVert_F$	Frobenius norm
$v \otimes w$	tensor product vw^T
I_d	$d \times d$ identity matrix
Tr	trace
:	$A : B = \mathrm{Tr}(A^T B)$, see (2.45)
det	determinant
supp	support of a function or measure
rank	rank of a matrix
$[N]$	set $\{1,...,N\}$
$\cdot$	dot product (inner product in $\mathbb{R}^d$)
H	Heaviside function
H_ϵ	regularized Heaviside function
D_δ	cap, see (5.46)
$\tilde{D}_\delta$	interior of the cap, see (5.49)
$B_1^d(0)$	unit ball $\{x \in \mathbb{R}^d \mid \lvert x\rvert \leq 1\}$
esssup	essential supremum
essinf	essential infimum
$*$	Khatri-Rao product, see (6.70)
vec	vecotrization of a matrix
∇_h^2	discretized Hessians

Abbreviations

ODE	ordinary differential equation
PDE	partial differential equation
SDE	stochastic differential equation
SVD	singular value decomposition
PCA	principal component analysis
iid	independent and identically distributed
ae	almost everywhere
wrt	with respect to
rhs	right hand side
lhs	left hand side
st	such that
iff	if and only if

Chapter 1

Introduction

An important problem in many disciplines of science, engineering and economics is *global optimization*. Global optimization, that is, the problem of finding the gloabl optimum of a given cost function $\mathcal{E}$ is ubiqutous and includes problems like finding the shortest possible route between a given list of cities (see Traveling Salesman Problem), finding the optimal shape of a wing of an airplane (see shape optimization), and finding the optimal allocation of scarce resources (e.g. organs, see [Pardalos and Romeijn, 2009]). Optimization plays an important role in Lagrangian mechanics, signal processing (e.g. compressed sensing, phase-retrieval), operations research (e.g. supply chain management) and machine learning to name just a few areas of application.

In short: optimization problems are everywhere, and the development of a mathematical theory to understand them and practical algorithms to solve them is of high importance.

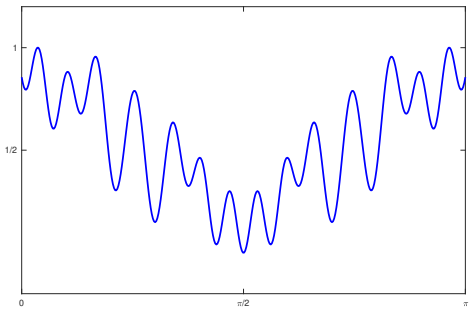


Figure 1.1: The Rastrigin function is an important benchmark function

In this thesis we will analyse a *consensus-based optimization method* on hypersurfaces (e.g. the sphere). We first start with some general remarks about optimization problems, before restricting ourselves to the subproblem of optimization on hypersurfaces, see [Absil and Hosseini, 2017] for practical applications of this subproblem.

First, global optimization can be substituted with *global minimization* or *global maximization*, as the problem of minimizing a cost function $\mathcal{E}$ is equivalent to maximizing the negative $-\mathcal{E}$, with the

same set of minimizers. We here consider only global minimization. Second, optimization problems only make sense for real-valued cost functions, otherwise we could not compare the function values of two distinct inputs. More precisely: given a cost function $\mathcal{E} : X \mapsto \mathbb{R}$ for some general set X, the optimization problem is about finding the element $x^\star \in X$ such that $\mathcal{E}(x^\star) \leq \mathcal{E}(x)$ for all $x \in X$. This problem is also stated as

$$\text{Find } x^\star \in X \text{ such that } x^\star \in \arg\min_{x \in X} \mathcal{E}(x) \tag{1.1}$$

where arg min is the subset of X for which the cost function takes minimal values, see figure 1.1 for an example of a cost function with $X = [0, \pi]$ and $x^\star = \pi/2$. We refer to $x^\star$ as the *minimizer* and $\mathcal{E}(x^\star)$ as the *minimum*. Third, the optimization problem (1.1) is not solvable (NP-hard) in general. Fourth, in this thesis we consider non-convex constrained optimization problems where the cost function is non-convex and the constraint is such that the solution lies on a given hypersurface.

A standard method to solve problem (1.1) is the famous *gradient descent method*. This method proceeds as follows: starting with an initial value $x_0 \in \mathbb{R}^d$ it moves in the direction of steepest descent. The direction of steepest descent is given by the negative gradient $-\nabla\mathcal{E}(x_0) \in \mathbb{R}^d$. The GD method thus computes

$$x_1 = x_0 - \eta \nabla \mathcal{E}(x_0) \tag{1.2}$$

where $\eta > 0$ is a step size (or learning rate). This process is then repeated for $x_1, x_2, x_3, ...$ until no further improvement in the function values is achieved. Some advantages of the gradient descent method are its simplicity and popularity. Disadvantages include the fact that it requires the computation of the gradient of the cost function which could be expensive or impossible to compute, and the lack of convergence proofs towards a global minimizer. The latter is a particular problem for non-convex cost functions with many local minimizers, as the Rastrigin function in figure 1.1.

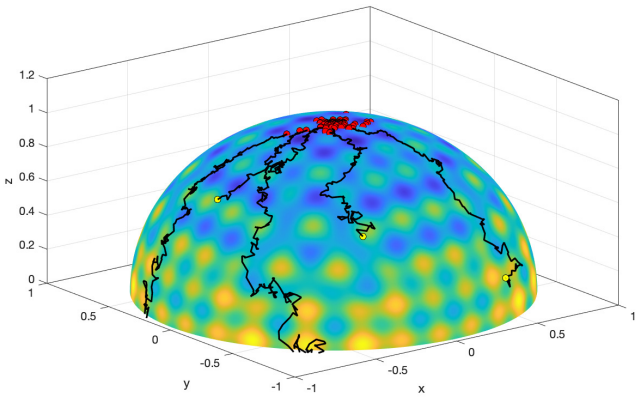


Figure 1.2: The Rastrigin function on $\mathbb{S}^2$. For 5 particles we plotted the trajectories with the corresponding starting positions in yellow.

A different class of optimization methods are the so-called metaheuristics, see [Back et al., 1997], [Gendreau and Potvin, 2010] and [Blum and Roli, 2003]. Numerous metaheuristics have been con-

sidered for global optimization. One of the most prominent metaheuristics is ant colony optimization (ACO), see [Dorigo and Blum, 2005], [Dorigo and Di Caro, 1999]. Other important metaheuristics include hill climbing (HC), evolutionary programming (EP), scatter search (ScS), simulated annealing (SimA), random search (RS), tabu search (TS), genetic algorithms (GA), particle swarm optimization (PSO) and consensus-based optimization (CBO).

In general metaheuristics are not well understood. Most metaheuristics can not be analysed rigorously as they lack a formal mathematical description. For CBO methods (and PSO as we will see in the next chapter) the situation is different. A CBO model starts with many *particles* V_0^i (also known as *agents*, *players* or *algorithms*) for $i = 1, ..., N$, where the subscript 0 denotes time $t = 0$. The particles are placed on the graph of the cost function. They all know their corresponding function value $\mathcal{E}(V_0^i)$, and they can communicate among each other. The particles than *play a game* in which they try to *collectively* find the global minimizer of the cost function. Eventually the particles *reach consensus* where they believe the global minimizer is, see figure 1.2. As the method does not require the computation of gradients of the cost function it is a *derivative-free optimization method*, see [Audet and Hare, 2017], [Conn et al., 2009] and [Larson et al., 2019] for introductions to the topic.

In more mathematical terms, a CBO method is a system of first order stochastic differential equations (SDE) complemented with suitable inital data V_0^i sampled from some probability density, say $V_0^i \sim \mathcal{U}(\mathbb{S}^{d-1})$. This allows for a rigorous mathematical analysis. In this thesis we carry out this analysis for a particular CBO model, namely the Kuramoto-Vicsek CBO model or KV-CBO.

Contributions

In this thesis we analyse a consensus-based optimization model, the so-called Kuramoto-Vicsek CBO model or KV-CBO model for short, for global optimization problems constrained on a hypersurface $\Gamma \subset \mathbb{R}^d$. We discuss implementation issues and real-world applications. The thesis is self-contained and is based on the publications [Fornasier et al., 2020], [Fornasier et al., 2021a] and [Fornasier et al., 2021b]. Chapters 5 and 6 and the appendix contain unpublished work. More precisely: in chapter 5 we argue that the assumptions needed in the convergence proof of the original publication are very strict and propose different assumptions that are easy to satisfy. The presented convergence proof is a significant simplification of the original proof. In chapter 6 we discuss the applicability of the KV-CBO method to the problem of reconstructing neural nets. Further, the appendix includes detailed proofs for the anisotropic KV-CBO model that have not been published in this detail in the original publication.

In **Chapter 2** we explain the basic concepts and ideas from CBO theory in Euclidean space, and give an overview of important publications. We lay the groundwork for the rest of the thesis and motivate the connection between CBO and PSO. The reader already familiar with CBO may start with chapter 3.

In **Chapter 3** we introduce the class of cost functions that we consider in this thesis. We further discuss the popular Kuramoto and Vicsek models that describe collective synchronization phenomena. These models motivate the KV-CBO model on $\mathbb{S}^{d-1}$. We further generalize the model to hypersurfaces $\Gamma \subset \mathbb{R}^d$ and discuss the well-posedness of the model.

In **Chapter 4** we derive the mean-field equation of the KV-CBO model and prove its well-posedness. The main result of this chapter is the proof of a mean-field limit result, that is, as the number of particles $N \to \infty$ the empirical measure converges toward the solution of the mean-field equation (a PDE). The results in this chapter are proved for general compact hypersurfaces Γ.

In **Chapter 5** we prove the main theorem of this thesis, namely the convergence of the consensus point $V^\alpha(\rho_t)$ toward the global minimizer $V^\star$ of the cost function. We discuss the conditions on the initial density ρ_0 and the coefficients of the KV-CBO model needed to achieve this convergence. In this chapter we restrict ourselves to the sphere $\Gamma = \mathbb{S}^{d-1}$.

In **Chapter 6** we discuss implementation aspects of the KV-CBO model, including descritization and acceleration aspects. We further present the Matlab package which can be downloaded from GitHub and discuss important applications in machine learning (robust PCA, phase retrieval, reconstruction of neural nets). We set a high value on reproducibility of the numerical experiments. Whenever possible, the exact Matlab scripts to reproduce the results can be downloaded from GitHub. In this chapter we restrict ourselves to the sphere $\Gamma = \mathbb{S}^{d-1}$.

In the **Appendix** we give important definitions of stochastic calculus for the readers convienience. We further present detailed proofs of the KV-CBO model with anisotropic (componentwise) noise for the special case $\Gamma = \mathbb{S}^{d-1}$. Last, we discuss common benchmark functions (Rastrigin, Ackley, Alpine, Schaffer, Solomon, Lévi, Xin-She Yang random, Griewank) that are included in the Matlab package.

Chapter 2

Consensus-based Optimization in Euclidean Space

In this section we give an overview of CBO models in Euclidean space. This serves as a motivation for the technically more complicated case of CBO models on hypersurfaces which are the main object of interest in this thesis. The discussion of the Euclidean case is important as it allows us to introduce basic concepts and ideas needed throughout this thesis. These include among others *stochastic differential equations (SDE)*, *empirical density*, *consensus of particles*, *propagation of chaos*, *McKean process*, *Fokker-Planck equation*, *mean-field equation* and *mean-field limit*. These ideas are sketched in this section and will be made precise in chapter 4 on the mean-field limit. CBO models have first been introduced in the two pioneering publications [Pinnau et al., 2017] and [Carrillo et al., 2018], which we will outline in this chapter. The reader already familiar with CBO methods may skip this chapter.

2.1 Particle swarm optimization

We start this chapter with a short introduction about particle swarm optimization (PSO). PSO methods have been introduce in [Kennedy and Eberhart, 1995], [Kennedy, 1997] (see also [Kennedy, 2010], [Poli et al., 2007]) and fall in the class of metaheuristics. Particle swarm optimization methods provide an interesting link to consensus-based optimization and open promising directions for further research. The standard PSO method can be understood as a discretization of a second (!) order system of SDEs with solution $(X_t^i, V_t^i) \in \mathbb{R}^{d\times 2}$ for $i \in [N]$, where X_t^i and V_t^i are the position and velocity of the ith particle at time $t \geq 0$ respectively. CBO methods are a special case of PSO, they can be recovered in the so-called *zero-inertia limit*. We refer to the publications [Grassi et al., 2021], [Grassi and Pareschi, 2021] and [Cipriani et al., 2021] (see also [Huang, 2021]) which we now briefly outline.

Let $\mathcal{E} : \mathbb{R}^d \to \mathbb{R}$ a cost function we wish to minimize. For simplicity we assume $\mathcal{E}$ to be smooth and

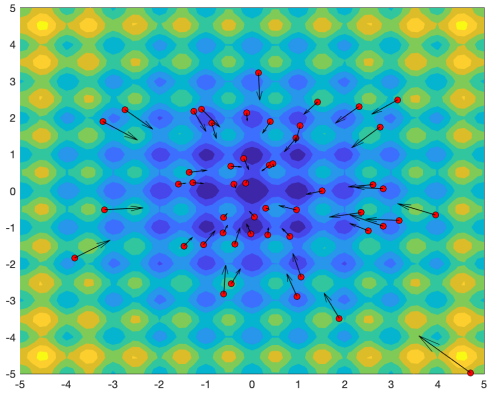


Figure 2.1: Depiction of particles (in red) and velocities (black arrows). The minimum of the cost function is located in the center of the image (dark blue).

have a unique global minimizer. We now introduce N *particles* with position $X_0^i \in \mathbb{R}^d$ and velocity $V_0^i \in \mathbb{R}^d$, where the subscript 0 denotes the iteration $n = 0$ (the time). These initial positions and velocities are chosen randomly, see figure 2.1. The original (time-discrete) PSO method updates the positions and velocities as

$$\begin{cases} X_{n+1}^i &= X_n^i + V_{n+1}^i \\ V_{n+1}^i &= V_n^i + c_1 R_1^n (Y_n^i - X_n^i) + c_2 R_2^n (\overline{Y}_n - X_n^i) \end{cases} \tag{2.1}$$

where c_1, c_2 are the acceleration coefficients, R_1^n, R_2^n are two d-dimensional diagonal matrices with entires sampled from a uniform distribution in $[0, 1]$. Further Y_n^i is the *local best position* (also known as *personal best position*, see also [Totzeck and Wolfram, 2020]) found by particle i up to iteration n, and $\overline{Y}_n$ is the *global best position* found by all the N particles up to iteration n. The local and global best are defined as

$$Y_0^i = X_0^i, \quad Y_{n+1}^i = \begin{cases} Y_n^i & \text{if } \mathcal{E}(X_{n+1}^i) \geq \mathcal{E}(X_n^i) \\ X_{n+1}^i & \text{if } \mathcal{E}(X_{n+1}^i) < \mathcal{E}(X_n^i) \end{cases} \tag{2.2}$$

and

$$\begin{cases} \overline{Y}_0 &= \arg\min_{i \in [N]} \{\mathcal{E}(X_0^i)\} \\ \overline{Y}_{n+1} &= \arg\min_{i \in [N]} \{\mathcal{E}(X_{n+1}^i), \mathcal{E}(\overline{Y}_n)\} \end{cases} \tag{2.3}$$

respectively.

It is then shown that (2.1) is a discretization of the (time continuous) second order system

$$\begin{cases} dX_t^i &= V_t^i dt \\ dV_t^i &= \lambda_1 (Y_t^i - X_t^i) dt + \lambda_2 (\overline{Y}_t - X_t^i) dt + \sigma_1 D(Y_t^i - X_t^i) dB_t^{1,i} + \sigma_2 D(\overline{Y}_t - X_t^i) dB_t^{2,i} \end{cases} \tag{2.4}$$

where $B_t^{1,i}, B_t^{2,i}$ are Borwnian motions for the particle i, $\lambda_k = c_k/2$ and $\sigma_k = c_k/2\sqrt{3}$ for $k = 1, 2$, and $D(Y_t^i - X_t^i)$ and $D(\overline{Y}_t - X_t^i)$ are diagonal matrices with entries $Y_t^i - X_t^i$ respectivley $\overline{Y}_t - X_t^i$. Introducing inertia m, regularizing the global best with $X^\alpha(\rho_t^N)$ and dropping the local best terms Y_t^i for simplicity (no memory) yields the system

$$\begin{cases} dX_t^i & = V_t^i dt \\ mdV_t^i & = -\gamma V_t^i dt - \lambda(X_t^i - X^\alpha(\rho_t^N))dt + \sigma D(X^\alpha(\rho_t^N) - X_t^i)dB_t^{2,i} \end{cases} \tag{2.5}$$

where $\gamma = (1 - m)$ and $X^\alpha(\rho_t^N)$ is the regularized global best defined as

$$X^\alpha(\rho_t^N) = \frac{\sum_{i=1}^N X_t^i e^{-\alpha\mathcal{E}(X_t^i)}}{\sum_{i=1}^N e^{-\alpha\mathcal{E}(X_t^i)}} \tag{2.6}$$

for a given parameter $\alpha \gg 1$. In [Cipriani et al., 2021] it is shown that in the zero inertia limit $m \to 0$ the PSO system (2.5) is equivalent to

$$dX_t^i = -\lambda(X_t^i - X^\alpha(\rho_t^N))dt + \sigma D(X_t^i - X^\alpha(\rho_t^N))dB_t^i. \tag{2.7}$$

This is the anisotropic CBO model that we discuss below, see (2.27). Note that this is a first (!) order system, and thus a significant simplification of the second order PSO system.

2.2 Consensus-based optimization

The global optimization problem considered in [Pinnau et al., 2017] is

$$\text{Find } X^\star \in \mathbb{R}^d \text{ such that } X^\star = \arg\min_{\mathbb{R}^d} \mathcal{E} \tag{2.8}$$

where $\mathcal{E} \in C_b(\mathbb{R}^d)$ is a given cost function. The first step in any CBO model is sampling the initial particles $X_0^i \sim \rho_0$ iid for $i \in [N]$ from a common probability distribution $\rho_0 \in \mathcal{P}(\mathbb{R}^d)$. This is also denoted as $(X_0^1, ..., X_0^N) \sim \rho_0^{\otimes N}$. We evaluate the cost function $\mathcal{E}$ at these locations and form a first guess for the location of the (unknown) global minimizer $X^\star$ as

$$X^\alpha = \frac{\sum_{i=1}^N X_0^i e^{-\alpha\mathcal{E}(X_0^i)}}{\sum_{i=1}^N e^{-\alpha\mathcal{E}(X_0^i)}} \in \mathbb{R}^d \tag{2.9}$$

where $\alpha \gg 1$ is the so-called temperature parameter. The consensus point X^α is thus a weighted averaged mean of the particles X_0^i where the weight function is given by $x \mapsto e^{-\alpha\mathcal{E}(x)}$. This means that particles $X_0^i \in \mathbb{R}^d$ with large function values $\mathcal{E}(x)$ have a very small weight, they are *less important* than the particles with small function values. The consensus point X^α is thus a convex combination of the particles, where the weight is determined by the function value of the particle. We generalize

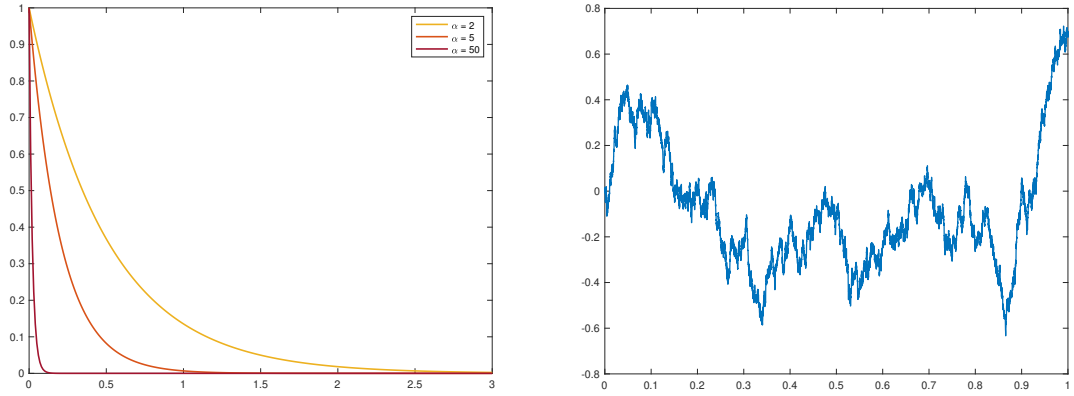


Figure 2.2: Left: weight function $x \mapsto e^{-\alpha x}$ for different values of α. For a large value of α the weights drops steeply. Right: one-dimensional Brownian motion in $[0, 1]$ with variance $\sigma = 0.2$.

the definition (2.9) by introducing the empirical measure

$$\rho_t^N = \frac{1}{N} \sum_{i=1}^{N} \delta_{X_t^i} \tag{2.10}$$

which is a very rough object (it is a sum of Dirac delta functions). Now we can define the consensus point at time $t \in [0, T]$ as

$$X^\alpha(\rho_t^N) = \frac{\int_{\mathbb{R}^d} x e^{-\alpha \mathcal{E}(x)} d\rho_t^N(x)}{\int_{\mathbb{R}^d} e^{-\alpha \mathcal{E}(x)} d\rho_t^N(x)} = \frac{\sum_{i=1}^N X_t^i e^{-\alpha \mathcal{E}(X_t^i)}}{\sum_{i=1}^N e^{-\alpha \mathcal{E}(X_t^i)}} \in \mathbb{R}^d. \tag{2.11}$$

In a CBO model we want the particles X_t^i to move towards the consensus point $X^\alpha(\rho_t^N)$ in a random trajectory. [Pinnau et al., 2017] modeled this with the coupled system of stochastic differential equations

$$dX_t^i = \underbrace{-\lambda(X_t^i - X^\alpha(\rho_t^N)) H_\epsilon(\mathcal{E}(X_t^i) - \mathcal{E}(X^\alpha(\rho_t^N)))dt}_{drift\ term} + \underbrace{\sigma |X_t^i - X^\alpha(\rho_t^N)| dB_t^i}_{diffusion\ term} \tag{2.12}$$

for all $i \in [N]$ where B_t^i is the Brownian motion of the ith particle (each particle has its own Bownian motion). Further H_ϵ is a smooth regularized version of the Heaviside function $H : \mathbb{R} \to \mathbb{R}$ where $H(x) = 1$ for $x > 0$ and $H(x) = 0$ else. The N equations in this system of SDEs are coupled through the consensus point $X^\alpha(\rho_t^N)$.

The system of equations (2.12) is difficult to understand at first. We therefore motivate it for a particle X_t^i for which the function value is larger than the function value of the consensus point, that is, let X_t^i a particle with $\mathcal{E}(X_t^i) > \mathcal{E}(X^\alpha(\rho_t^N))$. In this case we want the particle to move towards the consensus point. The regularized Heaviside function gives $H_\epsilon(\mathcal{E}(X_t^i) - \mathcal{E}(X^\alpha(\rho_t^N))) = 1$ and the equation for X_t^i

is

$$dX_t^i = -\lambda(X_t^i - X^\alpha(\rho_t^N))dt \tag{2.13}$$

where we assumed $\sigma = 0$ for simplicity. This is an equation of motion ($\dot{x} = f(x)$) for the particle X_t^i moving in the direction of the vector $X^\alpha(\rho_t^N) - X_t^i$, that is, X_t^i moves towards $X^\alpha(\rho_t^N)$ as desired. If on the other hand the particle has a smaller function value than the consensus point, that is, $\mathcal{E}(X_t^i) < \mathcal{E}(X^\alpha(\rho_t^N))$, we do not want the particle to move towards $X^\alpha(\rho_t^N)$. In this case the regularized Heaviside function is zero and consequently there is no drift term. Assuming again the case of $\sigma \neq 0$ we see that the particle X_t^i moves in a *random* trajectory. The randomness plays a crucial role as it allows the particle to explore the cost function. We note that the randomness is larger when the particle is further away from the consensus point. In other words: the diffusion term $\sigma|X_t^i - X^\alpha(\rho_t^N)|dB_t^i$ *favours exploration.*

If the particles reach a stationary state $dX^i = 0$ for all $i \in [N]$ with $X^i = X^j$ for all $i \neq j$, we say that *the particles have reached consensus.* If the particles have reached a stationary state but are not all at the same location, then the particles have reached *non-uniform consensus.* We say that *macroscopic consensus is reached* when the probability distribution of the particles ρ_t converges, in a way to be made precise, to a Dirac delta as $t \to \infty$, that is, $\rho_t \to \delta_{\hat{x}}$ for some $\hat{x} \in \mathbb{R}^d$. This is equivalent to $V(\rho_t) \to 0$ and $\mathbb{E}(\rho_t) \to \hat{x}$ as $t \to \infty$, where V and $\mathbb{E}$ denote the variance and expected value of ρ_t respectively. Usually it is clear from the context whether we talk about *microscopic* or *macroscopic consensus*, and therefore only use the term *consensus* to refer to both. It is not at all obvious if and when particles driven by a given CBO model form consensus or not. We will discuss this issue in later chapters.

The CBO model (2.12) is very difficult or even impossible to analyse. An idea to make the analysis tractable is to employ techniques from *kinetic theory.* The intention is to find a *mean-field equation* that describes the evolution of the probability distribution of the particles. As the name suggests, a mean-field equation is an equation for the *average of the particles.*

How to derive a mean-field equation for the CBO model (2.12)? A standard approach is to consider the dynamics of the, say, first *marginal* of the empirical measure ρ_t^N, which we denote by ρ_t, and make the *propagation of chaos assumption* $\rho_t^N \approx \rho_t^{\otimes N}$. This means that we assume the empirical density to be approximately tensorized, that is, the particles at time t are approximately independent. It is important to note that this assumption can only hold approximately: the particles, driven by a coupled (!) system of SDEs, can not be independent at any time $t > 0$. However, in the *large particle limit* $N \to \infty$, any fixed number of particles $k \in \mathbb{N}$ are approximately independent. Hence, we can formally (!) apply the law of large numbers

$$\frac{1}{N}\sum_{i=1}^N X_t^i e^{-\alpha\mathcal{E}(X_t^i)} \approx \int_{\mathbb{R}^d} x e^{-\alpha\mathcal{E}(x)} d\rho_t(x), \quad \frac{1}{N}\sum_{i=1}^N e^{-\alpha\mathcal{E}(X_t^i)} \approx \int_{\mathbb{R}^d} e^{-\alpha\mathcal{E}(x)} d\rho_t(x) \tag{2.14}$$

which implies

$$X^\alpha(\rho_t^N) \approx X^\alpha(\rho_t). \tag{2.15}$$

The fact that the consensus point of the empirical measure $X^\alpha(\rho_t^N)$ is approximated by the consensus point of the first marginal $X^\alpha(\rho_t)$ decouples the original CBO model. The CBO model (2.12) formally simplifies to

$$d\overline{X}_t = -\lambda(\overline{X}_t - X^\alpha(\rho_t))H_\epsilon(\mathcal{E}(\overline{X}_t) - \mathcal{E}(X^\alpha(\rho_t)))dt + \sigma|\overline{X}_t - X^\alpha(\rho_t)|dB_t \tag{2.16}$$

where $\rho_t = law(\overline{X}_t)$ is the law or probability distribution of the stochastic process $\overline{X}_t$. This SDE is a nonlinear *McKean process* or *McKean SDE*. This is not a standard SDE: the evolution of its solution $\overline{X}$ depends on the microscopic location $\overline{X}_t$ as well as the macroscopic probability distribution $\rho_t = law(\overline{X}_t)$. This type of equation is also called *distribution dependent SDE*. The well-posedness of such SDEs has been investigated in [Mishura and Veretennikov, 2020]. The solution $\overline{X}$ of the McKean SDE is a *stochastic process* with unknown distribution $\rho_t = law(\overline{X}_t)$.

An equation for the probability distribution $\rho_t(x) = \rho(t, x)$ is given by

$$\partial_t \rho_t = \nabla \cdot (\mu(\rho_t)\rho_t) + \Delta(\kappa(\rho_t)\rho_t) \tag{2.17}$$

where

$$\mu(\rho_t)(x) = \lambda(x - X^\alpha(\rho_t))H_\epsilon(\mathcal{E}(x) - \mathcal{E}(X^\alpha(\rho_t))), \quad \kappa(\rho_t)(x) = \frac{\sigma^2}{2}|x - X^\alpha(\rho_t)|^2 \tag{2.18}$$

for $x \in \mathbb{R}^d$, see section 2.5. Further $\nabla = \nabla_x$ and $\Delta = \Delta_x$ denote the gradient and Laplacian wrt the space variable. The first term $\nabla \cdot (\mu(\rho_t)\rho_t)$ is the *convection term* and the second term $\Delta(\kappa(\rho_t)\rho_t)$ is the *diffusion term*, corresponding to the drift and diffusion terms of the McKean SDE respectively. Hence, the equation is a *convection-diffusion equation*. We refer to equation (2.17) as the *mean-field equation* (also known as *Fokker-Planck equation*, *mean-field PDE* or *the PDE*).

Convection-diffusion equations, and the special case of Fokker-Planck equations, are classes of equations that can take very different forms. The Fokker-Planck equation (2.17) is a particularly complex equation: it is *non-linear* and *non-local* in the convection and diffusion part, where the non-locality comes from the integrals in the definition of the consensus point $X^\alpha(\rho_t^N)$, see (2.11). This raises several open analytical and numerical problems.

Let us summarize the last few paragraphs. We wanted to derive a mean-field equation for the CBO model (2.12). To this end we made the propagation of chaos assumption $\rho_t^N \approx \rho_t^{\otimes N}$, which allowed us to use the law of large numbers to simplify the expression for the consensus point $X^\alpha(\rho_t^N)$. This led to the decoupled McKean process (2.16) for which we can derive a mean-field equation of Fokker-Planck type given by (2.17). We will make this precise in chapter 4.

Now that we have a mean-field equation for the probability distribution $\rho_t = law(\overline{X}_t)$, we are left with two problems: 1) proving that the solution ρ_t converges to a Dirac delta $\delta_{\hat{x}}$ (macroscopic consensus), and 2) proving that $\hat{x}$ is close to the global minimizer $X^\star$ of $\mathcal{E}$. In [Pinnau et al., 2017] these problems have been investigated for the special case $\sigma = 0$ and $H_\epsilon \equiv 1$. For the first problem it is shown that for a suitable initial probability distribution ρ_0, the solution of the mean-field equation ρ_t converges to a Dirac delta $\delta_{\hat{x}}$. This was done by showing that

$$V(\rho_t) = e^{-2\lambda t} V(\rho_0) \tag{2.19}$$

implying the *monotonic decay of the variance* $V(\rho_t) \to 0$. Further it is shown that $\mathbb{E}(\rho_t) \to \hat{x}$.

For the second problem it is shown that under suitable smoothness conditions on the cost function $\mathcal{E}$ the consensus point $X^\alpha(\rho_t)$ converges to a point that is in the neighbourhood of the global minimizer $X^\star$, and that this neighbourhood can be made arbitrarily small by choosing $\alpha \gg 1$ sufficiently large.

Around the same time [Pinnau et al., 2017] was published a second pioneering publication appeared, see [Carrillo et al., 2018]. The authors study a simpler CBO model given by

$$dX_t^i = -\lambda(X_t^i - X^\alpha(\rho_t^N))dt + \sigma|X_t^i - X^\alpha(\rho_t^N)|dB_t^i \tag{2.20}$$

for $i \in [N]$, which corresponds to (2.12) with the regularized Heaviside function $H_\epsilon = 1$. We note that the authors were not able to derive the McKean process (2.16) rigorously, and consequently also the derivation of the Fokker-Plank equation is only formal, see [Carrillo et al., 2018, remark 3.2] (see also [Huang and Qiu, 2021] for a qualitative result and [Fornasier et al., 2021c]).

2.3 Anisotropic noise

The CBO models considered thus far all have *isotropic noise*, that is, every component of the d-dimensional Brownian motion is scaled with the same number $\sigma|X_t^i - X^\alpha(\rho_t^N)|$. In this section we explain why this is not desirable and present an alternative, see [Carrillo et al., 2020].

We start with a simple example of a monoparticle SDE to illustrate that the drift and diffusion parameters depend to the dimension d of the Euclidean space. This dimension dependence is a problem for the CBO models when we consider high dimensional optimization problems that often arise in applications.

The problem is illustrated with the monoparticle SDE

$$dX_t = -\lambda(X_t - X^\star)dt + \sigma|X_t - X^\star|dB_t, \quad X_0 = x_0 \tag{2.21}$$

where $X^\star, x_0 \in \mathbb{R}^d$ are given vectors. This SDE corresponds to a shifted version of the CBO model (2.20) with only one particle $N = 1$ and no consensus point X^α. By Dynkin's formula (respectively

Itô's formula, see section 2.5 and example 2.6.1) we have

$$\frac{d}{dt}\mathbb{E}|X_t - X^\star|^2 = -(2\lambda - d\sigma^2)\mathbb{E}|X_t - X^\star|^2 \tag{2.22}$$

which implies that the parameters need to satisfy the condition $2\lambda > d\sigma^2$ in order to achieve $\frac{d}{dt}\mathbb{E}|X_t - X^\star|^2 < 0$ which is necessary for the convergence of X_t to $X^\star$. This condition on the parameters λ and σ is not desirable as it involves a strong dependence on the dimension d. If d grows, the corresponding λ needs to be large, or the corresponding σ needs to be chosen small. We will see later that a large λ is not desirable (see remark 5.3.1).

We now replace the *isotropic noise* with *anisotropic* or *component-wise noise*, that is, we consider the monoparticle SDE

$$dX_t = -\lambda(X_t - X^\star) + \sigma D(X_t - X^\star)dB_t, \quad X_0 = x_0 \tag{2.23}$$

where

$$D(X_t - X^\star) := \operatorname{diag}(X_t - X^\star) \in \mathbb{R}^{d\times d} \tag{2.24}$$

is a diagonal matrix with the components of $X_t - X^\star \in \mathbb{R}^d$ on its diagonal (note that the entries can be negative). In this new model each component of the d-dimensional Brownian motion B_t is multiplied by the corresponding coordinate of $X_t - X^\star$, that is,

$$\sigma D(X_t - X^\star)dB_t = \sigma \sum_{k=1}^{d} (X_t - X^\star)_k (dB_t)_k e_k. \tag{2.25}$$

where e_k is the kth standard basis vector in $\mathbb{R}^d$ and $(dB_t)_k$ is the kth coordinate of the d-dimensional Brownian motion. Hence, if the kth coordinate of X_t is far away from the kth coordinate of $X^\star$, then the kth coordinate of the Brownian motion will be scaled with a larger number (in absolute value), resulting in a stronger random exploration in this one direction.

Applying Dynkin's formula yields

$$\frac{d}{dt}\mathbb{E}|X_t - X^\star|^2 = -2\lambda\mathbb{E}|X_t - X^\star|^2 + \sigma^2\mathbb{E}(\sum_{k=1}^{d}(X_t - X^\star)_k^2) = -(2\lambda - \sigma^2)\mathbb{E}|X_t - X^\star|^2. \tag{2.26}$$

Now the condition on the parameters is $2\lambda > \sigma^2$. Hence the explicit dimension dependence is removed. This finishes the discussion of the monoparticle SDE (2.21).

In chapter 6 we will see that the anisotropic version of the KV-CBO method performs significantly better than the isotropic version in numerical tests, see figure 6.2.

The CBO model (2.20) with anisotropic noise reads as

$$dX_t^i = -\lambda(X_t^i - X^\alpha(\rho_t^N))dt + \sigma D(X_t^i - X^\alpha(\rho_t^N))dB_t^i \tag{2.27}$$

and the mean-field equation for $\rho_t(x) = \rho(t,x)$ is

$$\partial_t \rho_t = \nabla \cdot (\mu(\rho_t)\rho_t) + \frac{\sigma^2}{2} \sum_{k=1}^{d} \partial_{kk}((x - X^\alpha(\rho_t))_k^2 \rho_t) \tag{2.28}$$

where the drift term is given by $\mu(\rho_t)(x) = \lambda(x - X^\alpha(\rho_t))$ as before. In the next chapter we will derive the KV-CBO model with isotropic and anisotropic noise. In chapter 5 we will then derive similar conditions on the parameters λ and σ in order to guarantee convergence, see (5.53) and (6.177) respectively.

The treatment of the anisotropic KV-CBO model is technically more difficult as the mean-field equation is more complicated, see (6.124).

2.4 Strong and weak solutions of SDEs

In this section we give an introduction to strong and weak solutions of SDEs, and provide a classical result about existence and uniqueness of such solutions. This result is important for the investigation of the well-posedness of the KV-CBO model. We refer to [Baldi, 2017] and [Durrett, 2018].

We consider a general SDE of the form

$$dX_t = \lambda(X_t, t)dt + \sigma(X_t, t)dB_t, \quad X_0 = x_0, \quad a.e. \tag{2.29}$$

where the drift and diffusion coefficients are measurable functions $\lambda : \mathbb{R}^d \times [0,T] \to \mathbb{R}^d$ and $\sigma : \mathbb{R}^d \times [0,T] \to \mathbb{R}^{d \times d'}$ respectively, and $x_0 \in \mathbb{R}^d$ is a given vector. The solution $X : [0,T] \to \mathbb{R}^d$ is a stochastic process on some probability space (which we neglect for simplicity, see section 3.3). As usual we write $X_t = X(t)$.

Definition 2.4.1 (Weak solution)**.** *If $\exists$ a Brownian motion B on some probability space $(\Omega, \mathcal{A}, P)$ with filtration $\mathcal{F}$, and a stochastic process X that satisfies* (2.29)*, then $(X, B, \Omega, \mathcal{A}, P, \mathcal{F})$ is a weak solution of* (2.29)*.*

It is important to note that here we can *choose* the Brownian motion B, as well as the probability space $(\Omega, \mathcal{A}, P)$ and the filtration $\mathcal{F}$. If for the chosen $(B, \Omega, \mathcal{A}, P, \mathcal{F})$ we find a stochastic process X that solves the SDE, then we call X a weak solution. We note that a weak solution of an SDE is not comparable with a weak solution of a PDE (we do not consider weak formulations of SDEs).

The notion of a *strong solution* requires the SDE (2.29) to be solved $\forall$ Brownian motions on a common (!) probability space $(\Omega, \mathcal{A}, P)$ with filtration $\mathcal{F}$. The existence of strong solutions implies the existence of weak solutions.

Definition 2.4.2 (Strong solution)**.** *If $\forall$ Brownian motions B on some probability space $(\Omega, \mathcal{A}, P)$*

with filtration $\mathcal{F}$ there exists a stochastic process X that satisfies (2.29)*, then $(X, B, \Omega, \mathcal{A}, P, \mathcal{F})$ is a strong solution of* (2.29)*.*

As a consequence of these different notions of solutions of SDE, we have to consider different notions of uniqueness either. These will be *uniqueness in distribution* (also known *uniqueness in law*) for weak solutions and *pathwise uniqueness* for strong solutions. Let $X^{(i)}$ for $i = 1, 2$ be two weak solutions. We have uniqueness in distribution if the corresponding pushforwards of $X^{(i)}$ equal, that is, $(X^{(1)})_{\#}P^{(1)} = (X^{(2)})_{\#}P^{(2)}$. For strong solutions $(X^{(i)}, B, \Omega, \mathcal{A}, P, \mathcal{F})$ the appropriate notion of uniqueness is *pathwise uniquenss* which is given if $X^{(i)}$ are indistinguishable, that is, $P(X^{(1)} = X^{(2)} \text{ on } [0, T]) = 1$.

Pathwise uniqueness implies uniqueness in distribution. This implication is however not straightforward as it requires to compare solutions defined on different probability spaces and driven by different Brownian motions.

The existence and uniqueness theory for SDEs relies on the coefficients satisfying a *local Lipschitz condition* and a *linear growth condition*. The former allows us to use a similar proving technique as in the famous *Picard-Lindelöf theorem* from ODE theory. The following theorem is a classical result and can be found in [Durrett, 2018, Thm 3.1, Ch. 5] which proves the existence of a global pathwise unique strong solution if a local martingale property is satisfied. We however are only interested in local solutions here.

Theorem 2.4.1. *Let the coefficient functions σ, λ be locally Lipschitz continuous, that is, for any $n < \infty$ there is a constant K_n such that*

$$|\sigma_{ij}(x) - \sigma_{ij}(y)| \leq K_n|x - y|, \quad |\lambda_i(x) - \lambda_i(y)| \leq K_n|x - y| \tag{2.30}$$

for all $|x|, |y| \leq n$. Then the SDE

$$dX_t = \lambda(X_t)dt + \sigma(X_t)dB_t \tag{2.31}$$

has a local (!) pathwise unique strong solution up to some finite time T.

2.5 Itô's formula

In this section we will introduce Itô's formula from stochastic calculus. This formula is the stochastic analogue of the chain rule from calculus $df(x) = f'(x)dx$ which no longer holds for stochastic processes due to the *quadratic variation of Brownian motion*, that is, in general $df(B_t) \neq f'(B_t)dB_t$ where B_t is a Brownian motion.

We will first introduce Itô's formula for a scalar function and then generalize it to multi-dimensional functions. We always consider functions f that only depend on space and not on time.

We start with an example that illustrates that the chain rule does not hold for stochastic processes. The fundamental theorem of calculus states that for a continuous function $f' \in C([0,T])$ with an indefinite integral f we have

$$\int_0^T f'(x)\ dx = f(T) - f(0) \quad or \quad df(x) = f'(x)dx. \tag{2.32}$$

We now consider the stochastic integral $\int f'(B_t)dB_t$ for $f'(x) = x$. To this end we let Δ^k denote a partition of the interval $[0,T]$, that is, $\Delta^k = \{0 = t_1^k < t_2^k \leq ... \leq t_{k-1}^k < t_k^k = T\}$, with $|\Delta^k| = \sup_{i\in[k-1]} |t_{i+1}^k - t_i^k| \to 0$ as $k \to \infty$. The stochastic integral is given by

$$\begin{aligned}\int_0^T B_t\ dB_t &= \lim_{|\Delta^k|\to 0} \sum_{i=1}^{k-1} B_{t_i^k}(B_{t_{i+1}^k} - B_{t_i^k}) &(2.33)\\ &= \lim_{|\Delta^k|\to 0} \frac{1}{2}\sum_{i=1}^{k-1}(B^2_{t_{i+1}^k} - B^2_{t_i^k}) - (B_{t_{i+1}^k} - B_{t_i^k})^2. &(2.34)\end{aligned}$$

With the quadratic variation of the Brownian motion

$$\lim_{|\Delta^k|\to 0} \sum_{i=1}^{k-1}(B_{t_{i+1}^k} - B_{t_i^k})^2 = T \tag{2.35}$$

we thus get

$$\int_0^T B_t\ dB_t = \frac{1}{2}(B_T^2 - T). \tag{2.36}$$

We observe that (2.32) does not hold with $f(x) = x^2/2$. Instead we have

$$df(B_t) = f'(B_t)dB_t + \frac{1}{2}f''(B_t)dt \tag{2.37}$$

which can be written in integral form as

$$f(B_T) = f(0) + \int_0^T f'(B_t)\ dB_t + \frac{1}{2}\int_0^T f''(B_t)\ dt. \tag{2.38}$$

This is the most basic example of Itô's formula.

We now generalize this result to multidimensional functions and general stochastic processes X_t. Since we are only interested in the differential form $df(X_t)$ we will no longer state the equivalent integral form.

Theorem 2.5.1 (Itô's formula (one-dimensional, no time-dependence))**.** *Let $f \in C^2(\mathbb{R})$ and*

$$dX_t = \lambda(t, X_t)dt + \sigma(t, X_t)dB_t \tag{2.39}$$

be a stochastic process for $t \in [0,T]$ and $X_t \in \mathbb{R}$. Then we have

$$df(X_t) = f'(X_t)dX_t + \frac{1}{2}f''(X_t)(dX_t)^2. \tag{2.40}$$

With the differential rules $(dt)^2 = 0$, $dB_t dt = 0$ and $dB_t^2 = dt$ this can be written as

$$df(X_t) = \big(f'(X_t)\lambda(t,X_t) + \frac{1}{2}f''(X_t)\sigma^2(t,X_t)\big)dt + f'(X_t)\sigma(t,X_t)dB_t \tag{2.41}$$

for $t \in [0,T]$.

We note that (2.37) follows from (2.41) by choosing $\lambda = 0$ and $\sigma = 1$. We further note that in (2.41) only the function $f \in C^2(\mathbb{R})$ is differentiated, not the drift or diffusion coefficients.

We now introduce the generalization of Itô's formula to high-dimensional functions. As above, we only consider functions f that do not depend on time, the coefficient functions however have a time-dependence.

Theorem 2.5.2 (Itô's formula (multi-dimensional, no time dependence))**.** *Let $f \in C^2(\mathbb{R}^d, \mathbb{R})$ and*

$$dX_t = \lambda(t,X_t)dt + \sigma(t,X_t)dB_t, \quad X_0 = x \tag{2.42}$$

be a stochastic process for $t \in [0,T]$ and $X_t \in \mathbb{R}^d$. Here, the drift and diffusion coefficients are multivariate functions $\lambda : [0,T] \times \mathbb{R}^d \to \mathbb{R}^d$, $\sigma : [0,T] \times \mathbb{R}^d \to \mathbb{R}^{d \times d'}$ and B_t is a d'-dimensional Brownian motion. Then

$$df(X_t) = \bigg(\nabla f(X_t) \cdot \lambda(t,X_t) + \frac{1}{2}\mathrm{Tr}(\sigma(t,X_t)^T \nabla^2 f(X_t)\sigma(t,X_t))\bigg)dt + \nabla f(X_t) \cdot \sigma(t,X_t)dB_t \tag{2.43}$$

where σ_k is the k-th row of the matrix σ, $\nabla = \nabla_x$ is the gradient and $\nabla^2 = \nabla_x^2$ the Hessian. We can write the trace term as

$$\mathrm{Tr}(\sigma(t,X_t)^T \nabla^2 f(X_t)\sigma(t,X_t)) = \sum_{k,l=1}^{d} (\nabla^2 f(X_t))_{kl}\sigma_k(t,X_t)\sigma_l(t,X_t)^T. \tag{2.44}$$

Sometimes the notation $:$ defined as $A : B = \mathrm{Tr}(A^T B)$ is used to simplify the trace term, that is,

$$\mathrm{Tr}(\sigma(t,X_t)^T \nabla^2 f(X_t)\sigma(t,X_t)) = \nabla^2 f(X_t) : (\sigma(t,X_t)^T \sigma(t,X_t)) \tag{2.45}$$

is used.

An other important formula in stochastic calculus is the Dynkin's formula. This formula is closely related to Itô's formula and is briefly outlined in the following. Taking the conditional expected value of the integral version of (2.43) and using the martingale property of Brownian motion (the integral

over the last term on the rhs of (2.43) = 0) yields

$$\mathbb{E}(f(X_t)|X_0 = x) = f(x) + \mathbb{E}(\int_0^t \mathcal{A}f(X_s)\ ds) \tag{2.46}$$

where $\mathcal{A}$ is the *infinitesimal generator of the stochastic process* X given by

$$\mathcal{A}f(x) = \nabla f(x) \cdot \lambda(t,x) + \frac{1}{2}\mathrm{Tr}(\sigma(t,x)^T \nabla^2 f(x)\sigma(t,x)) \tag{2.47}$$

for any sufficiently smooth test function f. This is *Dynkin's formula* and can also be written as

$$\frac{d}{dt}\mathbb{E}(f(X_t)|X_0 = x) = \mathbb{E}(\mathcal{A}f(X_t)). \tag{2.48}$$

2.6 The Fokker-Planck equation

We now come to the derivation of the Fokker-Planck equation (see [Öttinger, 1996]) corresponding to the stochastic process $dX_t = \lambda(t, X_t)dt + \sigma(t, X_t)dB_t$ in 1D, which we later generalize to higher dimensions. The Fokker-Planck equation, also known as *Kolmogorov forward equation*, describes the evolution of the probability density function $\rho_t(x) = \rho(x,t)$ of the stochastic process X_t with initial data $X_0 \sim \rho_0$ for $\rho_0 \in \mathcal{P}(\mathbb{R}^d)$ and is crucial for the analysis carried out in this thesis.

In order to derive the Fokker-Planck equation for the general stochastic process X_t we introduce the *transition probability* $p(s,z,t,x)$ of jumping from position z at time s to position x at time $t > s$. We set

$$\rho_t(x) = p(s,z,t,x). \tag{2.49}$$

Assume that $X_t = x$ for some time t. We are now interested in the conditional expected value $f(X_{t+\Delta t})$ for some small $\Delta t > 0$ and test function f. With the transition probability from above we have

$$\mathbb{E}(f(X_{t+\Delta t}) \mid X_t = x) = \int_{\mathbb{R}^d} f(y)p(t,x,t+\Delta t,y)\ dy. \tag{2.50}$$

Since we are interested in the time evolution of the transition probability p it makes sense to consider the *infinitesimal generator of the stochastic process* X defined as

$$\mathcal{A}f(x) = \lim_{\Delta t \to 0} \frac{\mathbb{E}(f(X_{t+\Delta t}) \mid X_t = x) - f(x)}{\Delta t}. \tag{2.51}$$

The expected value of $\mathcal{A}f(x)$ is given by

$$\int \mathcal{A}f(x)p(s,z,t,x)\ dx \tag{2.52}$$

and our aim is to derive an equation for $\rho_t(x) = p(s, z, t, x)$. With (2.51) we can rewrite (2.52) as

$$\begin{aligned} &\int \mathcal{A}f(x)p(s,z,t,x)\ dx \\ &\quad = \int \lim_{\Delta t \to 0} \frac{1}{\Delta t} \int f(y)p(t,x,t+\Delta t,y)p(s,z,t,x)\ dy - f(x)p(s,z,t,x)\ dx \end{aligned} \tag{2.53}$$

$$= \lim_{\Delta t \to 0} \frac{1}{\Delta t} \int f(y)p(s,z,t+\Delta t,y)\ dy - \int f(x)p(s,z,t,x)\ dx \tag{2.54}$$

where we have used the Chapman–Kolmogorov theorem in the last step. Renaming the integration variable y as x the last expression is recognized as a partial derivative with respect to time

$$\int \mathcal{A}f(x)p(s,z,t,x)\ dx = \int f(x)\partial_t p(s,z,t,x)\ dx. \tag{2.55}$$

Taking the adjoint operator of $\mathcal{A}$ and requiring the above equation to hold for all test functions f we get $\mathcal{A}^\star p(s, z, t, x) = \partial_t p(s, z, t, x)$. Which can be written in terms of the probability density ρ_t as

$$\partial_t \rho_t(x) = \mathcal{A}^\star \rho_t(x) \tag{2.56}$$

which is the *Fokker-Planck equation* or *Kolmogorov forward equation* of the stochastic process X. It remains to make the adjoint of the infinitesimal generator explicit. In (2.47) we have already derived $\mathcal{A}$ which in 1D simplifies to

$$\mathcal{A}f(x) = \lambda(t,x)f'(x) + \frac{\sigma^2(t,x)}{2} f''(x). \tag{2.57}$$

The adjoint of $\mathcal{A}$ can easily be found by an application of integration by parts

$$\mathcal{A}^\star \rho_t(x) = -(\lambda(t,x)\rho_t(x))' + \frac{1}{2}(\sigma^2(t,x)\rho_t(x))'' \tag{2.58}$$

which yields the Fokker-Planck equation

$$\partial_t \rho_t(x) = \lambda(t,x)f'(x) + \frac{\sigma^2(t,x)}{2} f''(x). \tag{2.59}$$

In $\mathbb{R}^d$ the Fokker-Planck equation takes the form

$$\partial_t \rho_t(x) = -\nabla_x \cdot (\lambda(t,x)\rho_t(x)) + \frac{1}{2} \sum_{k,l=1}^{d} \nabla_x^2 (\sigma_k(t,x)\sigma_l(t,x)^T \rho_t(x)). \tag{2.60}$$

Example 2.6.1. *Now we can compute the derivative $\frac{d}{dt}\mathbb{E}|X_t - X^\star|^2$ where X_t is the solution of the monoparticle SDE from section 2.3, that is,*

$$dX_t = -\lambda(X_t - X^\star)dt + \sigma|X_t - X^\star|dB_t. \tag{2.61}$$

In this case the infinitesimal generator is given by

$$\mathcal{A}f(x) = -\lambda(X_t - X^\star) \cdot \nabla f(x) + \frac{\sigma^2}{2}|X_t - X^\star|^2 \Delta f(x) \tag{2.62}$$

and we have $f(x) = |x - X^\star|^2$, $\nabla f(x) = 2(x - X^\star)$ *and* $\Delta f(x) = 2d$, *which implies*

$$\frac{d}{dt}\mathbb{E}|X_t - X^\star|^2 = (-2\lambda + \sigma^2 d)\mathbb{E}|X_t - X^\star|^2. \tag{2.63}$$

2.7 Wasserstein distance

An important tool in CBO theory is the *Wasserstein distance* W_p on the space $\mathcal{P}_p(X)$ of Borel probability measures with finite pth moment for $1 \leq p < \infty$, see [Villani, 2009] and [Ambrosio et al., 2008] for a detailed discussion. If $\mathcal{P}_p(X)$ is equipped with the Wasserstein distance W_p it is also refered to as *Wasserstein space*. In this thesis the important exponents are $p = 1$ and $p = 2$. In the former case the Wasserstein distance W_1 is also known as *Kantorovich–Rubinstein distance* or *Earth Mover's distance*. With Hölders inequality it can be shown that

$$W_1 \leq W_2 \tag{2.64}$$

that is, the Wasserstein-1 distance is the weakest Wasserstein distance.

For a Polish space (X, d) with metric d and two probability measures $\mu, \nu \in \mathcal{P}_p(X)$ we define the Wasserstein distance as

$$W_p^p(\mu, \nu) := \inf \left\{ \int_{X \times X} d(x, y)^p \, d\pi(\mu, \nu) \right\} \tag{2.65}$$

where the infimum is taken over all $\pi \in \Pi(\mu, \nu)$, the collection of all Borel probability measures on $X \times X$ with marginals μ and ν in the first and second component respectively. The Wasserstein distance can also be expressed as

$$W_p^p(\mu, \nu) = \inf \left\{ \mathbb{E}\big(d(X, Y)^p\big) \right\} \tag{2.66}$$

where the infimum is taken over all random variables X, Y with $law(X) = \mu$ and $law(Y) = \nu$. This immediately implies

$$W_p^p(\mu, \nu) \leq \mathbb{E}\big(d(X, Y)^p\big). \tag{2.67}$$

We need the probability measures to have finite pth moment to ensure that the Wasserstein distance is bounded. The distance that we will use is the usual Euclidean distance $d(x, y) = |x - y|$ and the space $X = \mathbb{R}^d$.

An important fact about the Wasserstein-1 distance is the following dual representation

$$W_1(\mu,\nu) = \sup_{Lip(\psi)\leq 1} \left\{ \int_X \psi \; d(\mu-\nu) \right\} \tag{2.68}$$

where the supremum is taken over all $\psi \in C(X,\mathbb{R})$ with Lipschitz constant ≤ 1.

2.8 Laplace's principle

In this section we discuss Laplace's principle from *large deviation theory* (see [Dembo and Zeitouni, 2010] and [Miller, 2006]). We state two versions of Laplace's principle: qualitative and quantitative. The former is a convergence result of $-\frac{1}{\alpha}\log\|e^{-\alpha\mathcal{E}}\|_{L^1(\rho_0)}$ towards the global minimizer of $\mathcal{E}$ as $\alpha\to\infty$. The latter also gives a convergence rate in terms of α and d. In chapter 5 we will prove a variant of Laplace's principle for the solution of the mean-field equation of the KV-CBO model, see lemma 5.4.4.

We remind the reader of the equivalent notations

$$\|e^{-\alpha\mathcal{E}}\|_{L^1(\rho_0)} = \int_{\mathbb{R}^d} e^{-\alpha\mathcal{E}(v)} d\rho_0(v) = \mathbb{E}_{v\sim\rho_0}(e^{-\alpha\mathcal{E}(v)}) \tag{2.69}$$

for a probability measure ρ_0.

Lemma 2.8.1 (Laplace's principle (qualitative version))**.** *Let $\rho_0 \in \mathcal{P}_{ac}(\mathbb{R}^d)$ an absolutely continuous measure with respect to the Lebesgue measure on $\mathbb{R}^d$ and $\mathcal{E}\in C^2(\mathbb{R}^d,\mathbb{R}_+)$. Then we have*

$$\lim_{\alpha\to\infty} -\frac{1}{\alpha}\log\|e^{-\alpha\mathcal{E}}\|_{L^1(\rho_0)} = \inf_{v\in\mathrm{supp}(\rho_0)} \mathcal{E}(v). \tag{2.70}$$

The validity of the result can be seen as follows. Taking the exponential on the left hand side yields

$$\exp\big(-\frac{1}{\alpha}\log\|e^{-\alpha\mathcal{E}}\|_{L^1(\rho_0)}\big) = \Big(\int_{\mathbb{R}^d} e^{-\alpha\mathcal{E}(v)}\; d\rho_0(v)\Big)^{-1/\alpha} = \frac{1}{\|e^{-\mathcal{E}}\|_{L^\alpha(\rho_0)}}. \tag{2.71}$$

With $\|\cdot\|_{L^\alpha}\to\|\cdot\|_{L^\infty}$ we then find

$$\lim_{\alpha\to\infty}\frac{1}{\|e^{-\mathcal{E}}\|_{L^\alpha(\rho_0)}} = \frac{1}{\|e^{-\mathcal{E}}\|_{L^\infty(\rho_0)}} = \mathrm{essinf}_{v\in\mathrm{supp}(\rho_0)}\, e^{\mathcal{E}(v)} \tag{2.72}$$

which finishes our simple argument.
The proof of the following quantitative version of Laplace's principle can be found in [Ha et al., 2021, Proposition 3.1].

Lemma 2.8.2 (Laplace's principle (quantitative version))**.** *Let $\mathcal{E}\in C^2(\mathbb{R}^d,\mathbb{R}_+)$ a cost function with global minimizer $X^\star$ such that*

$$\det(\nabla^2\mathcal{E}(X^\star)) > 0. \tag{2.73}$$

Let further $\rho_0 \in \mathcal{P}_{ac}(\mathbb{R}^d)$ an absolutely continuous measure with respect to the Lebesgue measure on $\mathbb{R}^d$ and with density f. The density f is assumed to satisfy the following conditions

1. *f is compactly supported,*
2. *f is continuous at $X^\star$, and*
3. *$f(X^\star) > 0$.*

Then the following result with convergence rate holds

$$\lim_{\alpha\to\infty} -\frac{1}{\alpha}\log\|e^{-\alpha\mathcal{E}}\|_{L^1(\rho_0)} = \inf_{v\in\mathrm{supp}(\rho_0)} \mathcal{E}(v) + \frac{d}{2}\frac{\log\alpha}{\alpha} + \mathcal{O}(1/\alpha). \tag{2.74}$$

Let us finish this chapter with an important remark.

Remark 2.8.1. *Laplace's principle will play an important role in chapter 5 when we proof the optimization results of the KV-CBO method. The principle suggests that a large value for α is desirable.*

On the one hand this is correct, as a large value for α increases the approximation properties of the method, see theorem 5.4.2. On the other hand a large value for α quickly deteriorates other estimates in the convergence proof as many of the estimates involve a constant $C_\alpha = \mathcal{O}(e^\alpha)$.

2.9 Summary

We have discussed the basics of CBO theory in Euclidean space and the connection to PSO. We have introduced basic concepts needed throughout this thesis. These include among others: consensus point, Itô's formula, Fokker-Planck equation, Laplace's principle and Wasserstein distance. Further, we have discussed the important distinction between isotropic and anisotropic noise.

Chapter 3

The Kuramoto-Vicsek Model

This chapter discusses the derivation and well-posedness of the KV-CBO model on the sphere and then generalize it to hypersurfaces Γ. The model is based on the Kuramoto and Vicsek models that describe respectively collective synchronization phenomena and collective behaviour of self-propelled particles with a fixed speed. Both have numerous applications in the natural sciences. We introduce the class of cost functions that we consider throughout the thesis and introduce the anisotropic version of the model. The final models can be found in (3.43) (isotropic) and (3.50) (anisotropic) below.

This chapter is based on joint work of Massimo Fornasier, Hui Huang, Lorenzo Pareschi and myself and has been published in [Fornasier et al., 2020], [Fornasier et al., 2021b].

3.1 Constrained optimization and cost functions

Let us start with a brief recap of what a CBO model is: a CBO model is a system of coupled SDEs. The solutions $V_t^i = V^i(t) \in \mathbb{R}^d$ to the N equations are refered to as *particles*. The particles move in random trajectories and are expected to reach consensus at a location close to the global minimizer of the given cost function $\mathcal{E}$, that is, $V^\alpha(\rho_T^N) \approx V^\star$ for a *finite (!) time horizon* $T \in (0, \infty)$. Differently from the previous chapter we here consider constrained optimization problems on a compact hypersurface $\Gamma \subset \mathbb{R}^d$

$$\text{Find } V^\star \in \Gamma \text{ such that } V^\star = \arg\min_\Gamma \mathcal{E} \tag{3.1}$$

where the cost function is minimized over a compact hypersurface $\Gamma \subset \mathbb{R}^d$. The compactness of Γ is crucial. It allows for a rigorously derivation of the mean-field equation of the CBO model (which the authors of [Carrillo et al., 2018] could not show for their CBO model in $\mathbb{R}^d$). See also [Borghi et al., 2021] for a treatment of constrained consensus-based optimization.

Before discussing the derivation of the model we first define the class of cost functions which we consider. Concerning the terminology: In the literature a cost function is also called *cost functional*, *objective function*, *target function* or *energy*. We however stick to the former terminology of cost

function.

Definition 3.1.1 (Class of cost functions)**.** *Let c_1, c_2, c_3, c_4 be positive constants. A cost function $\mathcal{E}$ belongs to the class $\mathcal{C}(c_1, c_2, c_3, c_4)$ if the following conditions are satisfied*

1. *$\mathcal{E}(v) \in [0, \infty)$ for all $v \in \Gamma$,*
2. *$\mathcal{E} \in C^2(\mathbb{R}^d, \mathbb{R}_+)$,*
3. *$|\nabla \mathcal{E}| \leq c_1$,*
4. *$|\nabla^2 \mathcal{E}|, |\Delta \mathcal{E}| \leq c_2$,*
5. *For any $v \in \Gamma$ there exists a global minimizer $V^\star \in \Gamma$ of $\mathcal{E}$ such that*

$$|v - V^\star| \leq c_3 |\mathcal{E}(v) - \mathcal{E}(V^\star)|^{c_4} \tag{3.2}$$

   where $\mathcal{E}(V^\star) = \underline{\mathcal{E}}$,

6. *$\mathcal{E}$ is locally Lipschitz continuous.*

In order to apply the quantitative version of Laplace's principle, see (2.8.2) *an additional condition is needed, namely*

$$\det(\nabla^2 \mathcal{E}(V^\star)) > 0 \tag{3.3}$$

for any global minimizer $V^\star$.

The first condition is not surprising: if the function were not bounded from below the minimization problem would not make sense. When the function is bounded by a negative constant, say $\mathcal{E} \in [c, \infty)$ for some $c < 0$ where the constant is known, we can simply consider the shifted cost function $\mathcal{E} = \mathcal{E} - c \in [0, \infty)$.

The second, third and fourth conditions are technical requirements needed to later prove the convergence of the method. For the numerical implementation such a smoothness assumption is however not required. Actually, the main motivation for investigating CBO models is that they are zero order methods, that is, do not rely on costly computations of gradients or Hessians of the cost function. As the CBO model only uses function evaluations, the numerical implementation of the CBO can for example be applied to functions which are only piecewise continuous. We note however that the well-posedness of the KV-CBO model is only guaranteed for locally Lipschitz continuous cost functions.

The fifth condition requires the cost function to be contained in a convex envelope, see figure 3.1 below. If the function value $\mathcal{E}(v)$ for some $v \in \Gamma$ is close to the global minimum $\underline{\mathcal{E}}$ then v must be close to one of the global minimizers $V^\star$. For a small constant $c_4 > 0$ the convex envelope is flat, for a large constant $c_4 > 0$ it is steep. This conditions is also known as *coercivity condition* or *inverse continuity assumption* (ICA).

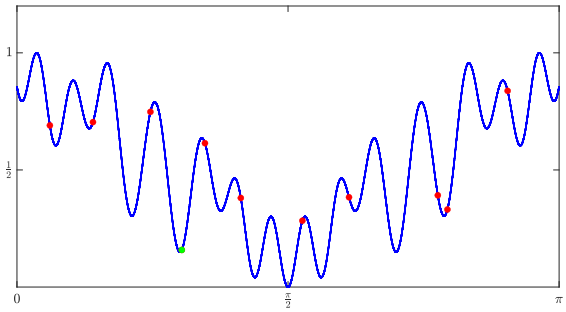


Figure 3.1: The cost function $\mathcal{E}$ can have multiple local minima. Here the Rastrigin function on $\mathbb{S}^1$ is shown. The particles are shown in red, the consensus point in green.

An important benchmark cost function is the famous Rastrigin function

$$\mathcal{E}_R(v) = Ad + |v|^2 - A\sum_{i=1}^{d}\cos(2\pi v_i) \tag{3.4}$$

for all $v \in \mathbb{R}^d$. The Rastrigin function satisfies all the conditions from above, see lemma 6.2.1.

3.2 The Kuramoto and Vicsek models

In this section we discuss the *classical* and *generalized Kuramoto model* and *the Vicsek model*, see [Strogatz, 2000] and [Acebron et al., 2005]. The classical Kuramoto model describes *collective synchronization phenomena* of a large population of *coupled phase oscillators*. Building on the terminology from the beginning of this chapter we can think of these phase oscillators as particles on the 1-dimensional sphere $\mathbb{S}^1$. The Kuramoto model has surprisingly many applications, it models biological, physical, chemical and social phenomena. We list a few of these applications below. The generalized Kuramoto model is a version of the classical Kuramoto model that operates in arbitrary dimensions d and its derivation relies on the introduction of a complex-valued *order parameter*. Closely related to the Kuramoto model is the Vicsek model which models the evolution of a collection of *self-driven* particles that move in $\mathbb{R}^d$ and have constant speed of 1, that is, their velocity is an element of $\mathbb{S}^{d-1}$.

A glimpse into the fascinating world of synchronization phenomena:

- **Social**: At the end of a good concert the audience applauds to show the musicians that they liked the concert. Since everybody has an individual *clapping frequency* the rhythm of the applause appears incoherent, at first. But, already after a very short time a remarkable phenomenon is observed: the auditorium claps in unison. Possibly thousands or tens of thousands of spectators have locked to a common frequency and now clap in synchrony.

- **Physics**: A Josephson junction is a device that consists of two or more coupled superconductors separated by a thin insulating barrier. Connecting a large number N of them allows to

produce a large effective output power proportional to N^2, provided their *oscillation frequency* in synchronized. The collective behaviour of Josephson junction arrays can be modelled with the Kuramoto model.

- **Medicine**: The heart rate is controlled by electrical impulses of large networks of pacemaker cells. The collective firing of these pacemaker cells causes the heart to contract. In order for the heart to beat in a steady and regular rhythm the pacemaker cells need to lock to a common frequency. Physical activity and emotional thrill requires the pacemaker cells to change their *firing frequency* coherently to change the heart rate. The synchronization of the firing rate of pacemaker cells can be described with the Kuramoto model.

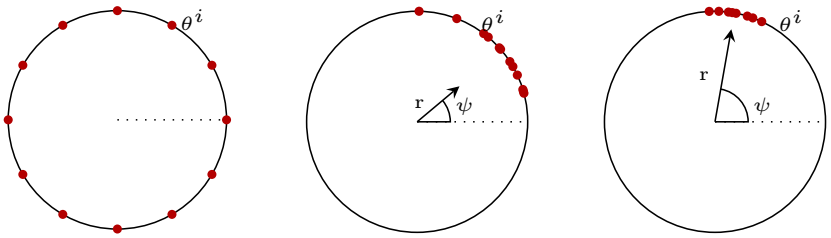


Figure 3.2: Phase-oscillators and order parameter at different times.

A first step to modelling synchronization phenomena is to consider each member of the population as a phase oscillator $\vartheta_t^i = \vartheta^i(t)$ with a given fixed individual *natural frequency* $\omega_i \sim g(\omega)$ where g is a symmetric probability distribution which we assume to have zero mean without loss of generality. Here $i \in [N]$ and the superscript $t \in [0, T]$ denotes time.

[Kuramoto, 1975] proved that the long-term dynamics of any system of weakly coupled, nearly identical phase oscillators follows the equation

$$\frac{d}{dt}\vartheta^i = \omega^i + \sum_{j=1}^{N} K_{ij}(\vartheta^j - \vartheta^i) \tag{3.5}$$

for an interaction function K to be chosen. The Kuramoto model [1] corresponds to the choice of *mean-field coupling of the phase differences*

$$K_{ij}(\vartheta^j - \vartheta^i) = \frac{\lambda}{N}\sin(\vartheta^j - \vartheta^i) \tag{3.6}$$

where $\lambda \geq 0$ is the *coupling constant*. For $\lambda < \lambda_c$ where λ_c is a critical coupling constant, we call the system *weakly coupled*, whereas for $\lambda > \lambda_c$ we call it *strongly coupled*. The factor $1/N$ ensures that

[1] MATLAB implementation is available at `https://blogs.mathworks.com/cleve/2019/08/26/kuramoto-model-of-synchronized-oscillators/`

the model is well-behaved in the *large particle limit* or *continuum limit* $N \to \infty$. We can measure synchronization of the system with the complex-valued order parameter

$$re^{i\psi} = \frac{1}{N}\sum_{j=1}^{N} e^{i\vartheta^j} \tag{3.7}$$

where $r = r(t)$ measures the *coherence*, and $\psi = \psi(t)$ is the average phase, see figure 3.2. Multiplying the order parameter with $e^{-i\vartheta^j}$ and comparing imaginary parts yields

$$\frac{d}{dt}\vartheta^i = \omega^i + \lambda r \sin(\psi - \vartheta^i). \tag{3.8}$$

In this equation each oscillator appears to be uncoupled, however they still interact through the quantities ψ and r. More precisely: the phase of the oscillator ϑ^i moves towards the mean phase ψ, rather than towards the phase of any other oscillator ϑ^j.

We can further observe an interesting behaviour in the effective coupling strength λr: if the oscillators become more coherent, that is r grows, the effective coupling strength also increases, hence the coupling becomes even stronger and more oscillators are attracted. For coupling constants $\lambda_c < \lambda < \infty$, those oscillators with $|\omega_i| < \lambda r$ are *phase-locked*, that is, $\frac{d}{dt}\vartheta^i = 0$, while those with $|\omega_i| > \lambda r$ are *drifting* and move out of synchrony.

In higher dimensional spaces $\mathbb{R}^d$ we can generalize the classical Kuramoto model with $\omega_i = 0$ as

$$\frac{d}{dt}V_t^i = \frac{\lambda}{N}\sum_{j=1}^{N} P_\Gamma(V_t^i)V_t^j \tag{3.9}$$

for $V_t^i \in \Gamma$ and $\Gamma = \mathbb{S}^{d-1}$ where $P_\Gamma(V_t^i) = I - V_t^i \otimes V_t^i$ is the projection onto the orthogonal complement of V_t^i and the time derivative is applied component wise, that is, $(\frac{d}{dt}V_t^i)_k = \frac{d}{dt}(V_t^i)_k$ for $k \in [d]$.

With

$$\overline{V}_t = \frac{1}{N}\sum_{i=1}^{N} V_t^i \tag{3.10}$$

the model can be rewritten as

$$\frac{d}{dt}V_t^i = \lambda P_\Gamma(V_t^i)\overline{V}_t. \tag{3.11}$$

We now prove that this model is a generalization of the classical Kuramoto model if the natural frequencies are zero $\omega^i = 0$. To this end we set $V_t^i = (\cos\vartheta_t^i, \sin\vartheta_t^i)^T$ and consider the equation for the first component of (3.9), that is,

$$\frac{d}{dt}\cos\vartheta^i = \frac{\lambda}{N}\sum_{j=1}^{N}(\cos\vartheta^j - \langle V_t^i, V_t^j\rangle \cos\vartheta^i) \tag{3.12}$$

where $\langle V_t^i, V_t^j \rangle = \cos\vartheta_t^i \cos\vartheta_t^j + \sin\vartheta_t^i \sin\vartheta_t^j$ which further gives

$$\frac{d}{dt}\cos\vartheta^i = \frac{\lambda}{N}\sum_{j=1}^{N}(\cos\vartheta^j - (\cos\vartheta^j\cos^2\vartheta^i + \sin\vartheta^i\sin\vartheta^j\cos\vartheta^i)) \tag{3.13}$$

$$= \frac{\lambda}{N}\sum_{j=1}^{N}\sin\vartheta^i(\cos\vartheta^j\sin\vartheta^i - \sin\vartheta^j\cos\vartheta^i) \tag{3.14}$$

where we have used $1 - \cos^2\vartheta^i = \sin^2\vartheta^i$.

With

$$\frac{d}{dt}\cos\vartheta^i = -\sin\vartheta^i\frac{d}{dt}\vartheta^i \tag{3.15}$$

this simplifies to

$$\frac{d}{dt}\vartheta^i = \frac{\lambda}{N}\sum_{j=1}^{N} -\cos\vartheta^j\sin\vartheta^i + \sin\vartheta^j\cos\vartheta^i = \frac{\lambda}{N}\sum_{j=1}^{N}\sin(\vartheta^j - \vartheta^i) = \lambda\sin(\psi - \vartheta^i). \tag{3.16}$$

Hence we recover the classical Kuramoto model (3.5), (3.6).

So far we have motivated synchronization phenomena and discussed a famous model for such phenomena which we have generalized to arbitrary dimensions d in the case when there are no natural frequencies.

Some 20 years after the original work of Kuramoto [Vicsek et al., 1995] proposed a model for *self-driven* particles with a constant absolute velocity (speed) motivated by applications in biology, e.g. the collective motion of certain bacteria. As the authors wrote: *"The only rule of the model is: at each time step a given particle driven with a constant absolute velocity assumes the average direction of motion of the particles in its neighborhood of radius r with some random perturbation added".*

The discrete Vicsek model, translated to our notations, reads as

$$\begin{cases} X_{t+\Delta t}^i &= X_t^i + v\Delta t(\cos\vartheta_t^i, \sin\vartheta_t^i)^T \\ \vartheta_{t+\Delta t}^i &= \frac{1}{N}\sum_j K(X_t^i - X_t^j)\vartheta_t^j + \eta_t^i \end{cases} \tag{3.17}$$

with the kernel function

$$K(x) = \begin{cases} 1, & |x| < \tau \\ 0, & |x| \geq \tau \end{cases} \tag{3.18}$$

where $\tau > 0$ is a given threshold, $v > 0$ is a given fixed length (we consider $v = 1$), and $\eta_t^i \sim \mathcal{U}[-\pi, \pi]$ is uniformly distributed noise.

Introducing the notation $V_t^i = (\cos\vartheta_t^i, \sin\vartheta_t^i)^T$ the continuous version of the Vicsek model can be

rewritten as

$$\begin{cases} dX_t^i &= V_t^i dt \\ dV_t^i &= P_\Gamma(V_t^i)\frac{1}{N}\sum_j K(X_t^i - X_t^j)V_t^j + \sigma P_\Gamma(V_t^i) \circ dB_t^i \end{cases} \tag{3.19}$$

where $\circ$ denotes the *Stratonovich form* of the SDE (see appendix) and we introduced the diffusion coefficient $\sigma \geq 0$. The initial positions and velocities are chosen randomly according to a common distribution on $\mathbb{R}^d$ (for X_t^i) and $\mathbb{S}^{d-1}$ (for V_t^i) respectively.

We can reformulate this system in Itô's form as

$$\begin{cases} dX_t^i &= V_t^i dt \\ dV_t^i &= -P_\Gamma(V_t^i)(\frac{1}{N}\sum_j K(X_t^i - X_t^j)(V_t^i - V_t^j))dt + \sigma P_\Gamma(V_t^i)dB_t^i - \frac{\sigma^2}{2}(d-1)\frac{V_t^i}{|V_t^i|^2}dt \end{cases} \tag{3.20}$$

where we have used $P_\Gamma(V_t^i)V_t^i = 0$. This is the formulation of the original Vicsek model that is best suited for our purposes. We refer to [Bolley et al., 2012] for the derivation of the mean-field limit of this model. Further important theoretical results for the Vicsek model were obtained in [Degond and Motsch, 2008], [Gamba and Kang, 2016], [Degond et al., 2015] and [Figalli et al., 2018].

In the following we will establish the connection between the Vicsek and the Kuramoto model. The Kuramoto model is a deterministic model for the particles $V^i \in \mathbb{S}^{d-1}$ where only the natural frequencies ω^i and the initial particles are chosen randomly. The Vicsek model on the other hand is stochastic and is a second order system. Here, the particles are moving in Euclidean space $\mathbb{R}^d$, not on the sphere as in the Kuramoto model.

Choosing the interaction kernel $K = 1$ eliminates the X_t^i's from the second equation of (3.20), and the system reduces to

$$dV_t^i = -\lambda P_\Gamma(V_t^i)(V_t^i - \overline{V}_t)dt + \sigma P_\Gamma(V_t^i)dB_t^i - \frac{\sigma^2}{2}(d-1)\frac{V_t^i}{|V_t^i|^2}dt. \tag{3.21}$$

The deterministic version ($\sigma = 0$) of this model equals the Kuramoto model derived above, see (3.11). We will see later that choosing the initial particles on the sphere $\mathbb{S}^{d-1}$ has the desirable property that they stay on the sphere for all time $t \in [0, \infty)$.

We summarize the few paragraphs from above: the Kuramoto model describes the collective behaviour of coupled oscillators that eventually lock to a common frequency provided their individual natural frequency ω_i was not too large. The Kuramoto model is not (!) a CBO model. The oscillators do not form *consensus*, they perpetually move around on the sphere. The model (3.21) does not involve a cost function to be minimized.

Let us now incorporate a cost function $\mathcal{E}$ into the model.

At time $t = 0$ the only information available to the particles are the function evaluations at the initial positions, that is, $\mathcal{E}(V_0^i)$ for $i \in [N]$. From this information we form a *guess* of the (unknown) position of the global minimizer $V^\star$ of $\mathcal{E}$, and then, instead of moving towards the average $\overline{V}_t$ as in the model (3.21) the particles move towards this guess. There are many possibilities to form such a guess, e.g. one could just choose $\arg\min_{i\in[N]} \mathcal{E}(V_0^i)$. This would however pose a problem in the derivation of the mean-field equation as not all particles would be similar/ interchangeable.

The *guess* is chosen as

$$V^\alpha(\rho_t^N) = \frac{\sum_{i=1}^N V_t^i e^{-\alpha\mathcal{E}(V_t^i)}}{\sum_{i=1}^N e^{-\alpha\mathcal{E}(V_t^i)}} = \frac{\int_{\mathbb{R}^d} v e^{-\alpha\mathcal{E}(v)} d\rho_t^N(v)}{\int_{\mathbb{R}^d} e^{-\alpha\mathcal{E}(v)} d\rho_t^N(v)} \tag{3.22}$$

where

$$\rho_t^N = \frac{1}{N}\sum_{i=1}^N \delta_{V_t^i} \tag{3.23}$$

is the empirical measure of the particles at time t, $\delta_{V_t^i}$ is the Dirac delta, and $\alpha \gg 1$ is the temperature parameter to be chosen. This choice is motivated by Laplace's principle, see section 2.8.

The consensus point (3.22) only involves function evaluations at a fixed time t but ignores previous function evaluations at time $\tilde{t} < t$. It is however possible to incorporate *memory mechanisms* into the CBO model, see section 2.1 (PSO).

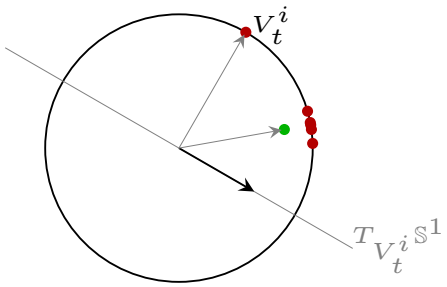


Figure 3.3: The particle V_t^i moves in the direction $P_\Gamma(V_t^i)V^\alpha(\rho_t^N)$ (black arrow). The consensus point is shown in green.

With the consensus point from above the model (3.21) takes the form

$$dV_t^i = \lambda P_\Gamma(V_t^i)V^\alpha(\rho_t^N)dt + \sigma P_\Gamma(V_t^i)dB_t^i - \frac{\sigma^2}{2}(d-1)\frac{V_t^i}{|V_t^i|^2}dt \tag{3.24}$$

where we have used $P_\Gamma(V_t^i)V_t^i = 0$. It is important that each particle has its own Brownian motion.

System (3.24) is not yet the final KV-CBO model on the sphere. We want the particles that are far away from the consensus point $V^\alpha(\rho_t^N)$ to have a stronger random component. More precisely: if $|V_t^i - V^\alpha(\rho_t^N)|$ is large then the particle should further explore the cost function $\mathcal{E}$ (have large diffusion). The hope is that the particle finds a location (e.g. a basin) where the cost function attains much smaller function values than all the other particles have found so far. If on the other hand $|V_t^i - V^\alpha(\rho_t^N)|$ is small, that is, the particle is already very close to the consensus point, then the stochastic component should be small and eventually vanish completely so that the particles can reach consensus. We choose the scale for the Brownian motion as $\sigma|V_t^i - V^\alpha(\rho_t^N)|$, see (2.20). However, other choices are also be possible (see anisotropic model below).

The KV-CBO model on the sphere $\Gamma = \mathbb{S}^{d-1}$ is given by

$$dV_t^i = \lambda P_\Gamma(V_t^i)V^\alpha(\rho_t^N)dt + \sigma|V_t^i - V^\alpha(\rho_t^N)|P_\Gamma(V_t^i)dB_t^i + C_\Gamma(V_t^i, V^\alpha(\rho_t^N)dt \tag{3.25}$$

where

$$C_\Gamma(V_t^i, V^\alpha(\rho_t^N) = -\frac{\sigma^2}{2}(d-1)|V_t^i - V^\alpha(\rho_t^N)|^2\frac{V_t^i}{|V_t^i|^2} \tag{3.26}$$

is the *correction term* which guarantees that the particles stay on Γ if initially chosen there.

Before extending this model to general compact hypersurfaces Γ we first define the particles/ empirical measure in terms of random variables/ functions/ measures.

3.3 A measure theoretical perspective

So far we have interpreted the particles as vectors in $\mathbb{R}^d$, that is, $V_t^i \in \mathbb{R}^d$ for each $t \in [0,T]$. This is definitely a handy short notation that we will use throughout this thesis, but it does not reflect what is really happing. In more rigorous terms the particle V^i is a *stochastic process* (= a collection of random variables indexed by time) on some probability space $(\Omega^i, \mathcal{A}^i, P^i)$ for $i \in [N]$. With the *Kolmogorov extension theorem* (see appendix) we can find a common probability space for all the particles $(\Omega, \mathcal{A}, P)$.

Hence, a particle V^i is a stochastic process

$$V^i : [0,T] \times (\Omega, \mathcal{A}, P) \to \mathbb{R}^d \tag{3.27}$$

and V_t^i is a random variable

$$V_t^i : (\Omega, \mathcal{A}, P) \to \mathbb{R}^d \tag{3.28}$$

for each time $t \in [0,T]$. We define the *probability distribution* or *law* of the particle V_t^i as the

pushforward measure

$$\mu_t^i(A) := (V_t^i)_\# P(A) = P \circ (V_t^i)^{-1}(A) = P(V_t^i \in A), \quad A \in \mathcal{B}_{\mathbb{R}^d} \tag{3.29}$$

where $\mathcal{B}_{\mathbb{R}^d}$ is the Borel σ-algebra on $\mathbb{R}^d$ and $\mu_t^i \in \mathcal{P}(\mathbb{R}^d)$ is a probability measure on $\mathbb{R}^d$. We often use the equivalent notations

$$\mu_t^i = law(V_t^i) \quad \textit{and} \quad V_t^i \sim \mu_t^i. \tag{3.30}$$

An alternative approach is to interpret the particle V^i as a random function

$$V^i : (\Omega, \mathcal{A}, P) \to C([0,T], \mathbb{R}^d). \tag{3.31}$$

In this case we can similarly define the law as the pushforward measure $\mu^i = (V^i)_\# P$ which yields $\mu_t^i = (V_t^i)_\# P$. We have $\mu^i \in \mathcal{P}(C([0,T], \mathbb{R}^d))$, respectively $\mu_t^i \in \mathcal{P}(\mathbb{R}^d)$.

For the empirical measure we proceed similarly: ρ^N is a random measure on some probability space $(\Omega', \mathcal{A}', P')$

$$\rho^N : (\Omega', \mathcal{A}', P') \to \mathcal{P}(C([0,T], \mathbb{R}^d)), \quad \rho^N = \frac{1}{N} \sum_{i=1}^N \delta_{V^i}. \tag{3.32}$$

We remind that the empirical measure is a very complex object. It is a random measure (where the randomness comes from the fact that the particles are stochastic processes) and as a sum of Dirac delta functions it is a very rough object and we can not expect any smoothness.

For the probability distribution of ρ^N we have $law(\rho^N) \in \mathcal{P}(\mathcal{P}(C([0,T], \mathbb{R}^d)))$. For each $t \in [0,T]$ we have

$$\rho_t^N : (\Omega', \mathcal{A}', P') \to \mathcal{P}(\mathbb{R}^d), \quad \rho_t^N = \frac{1}{N} \sum_{i=1}^N \delta_{V_t^i} \tag{3.33}$$

and $law(\rho_t^N) \in \mathcal{P}(\mathcal{P}(\mathbb{R}^d))$. The consensus point V^α is a random function

$$V^\alpha(\rho_t^N) : (\Omega', \mathcal{A}', P') \to C([0,T], \mathbb{R}^d), \quad V^\alpha(\rho_t^N) = \frac{\sum_{i=1}^N V^i e^{-\alpha \mathcal{E}(V^i)}}{\sum_{i=1}^N e^{-\alpha \mathcal{E}(V^i)}} \tag{3.34}$$

and its probability distribution is an element of $\mathcal{P}(C([0,T], \mathbb{R}^d))$. For a fixed $t \in [0,T]$ we have $law(V^\alpha(\rho_t^N)) \in \mathcal{P}(\mathbb{R}^d)$. See [Huang and Qiu, 2021] for the notations introduced here.

We note that the particles are defined on the same probability space $(\Omega, \mathcal{A}, P)$. The empirical measure however has a different probability space $(\Omega', \mathcal{A}', P')$. It would be interesting to understand the relation between the two probability spaces.

3.4 From the sphere to hypersurfaces

In this section, which is based on [Demlow and Dziuk, 2007], we generalize the KV-CBO model from the sphere $\mathbb{S}^{d-1}$ to general compact hypersurfaces $\Gamma \subset \mathbb{R}^d$. We first set up the definitions from differential geometry. For an introduction to stochastic calculus on manifolds we refer to [Émery and Meyer, 1989], [Itô, 1950] and [Hsu, 2002].

Let Γ be a connected, smooth, compact hypersurface without boundary embedded in $\mathbb{R}^d$, that is, $\Gamma \subset \mathbb{R}^d$, $\dim(\Gamma) = d-1$ and $\partial\Gamma = \emptyset$. Such a hypersurface Γ can be written as the level set of a (non unique) function $\gamma : \mathbb{R}^d \to \mathbb{R}$

$$\Gamma = \{v \in \mathbb{R}^d \mid \gamma(v) = 0\}. \tag{3.35}$$

The *Jordan-Brouwer separation theorem* states that Γ separates $\mathbb{R}^d$ into two connected components: the interior and the exterior. We choose the function γ such that $\gamma < 0$ on the interior and $\gamma > 0$ on the exterior. In order to introduce the gradient, divergence, and Laplace-Beltrami operator on Γ we further assume that there is an open neighborhood Γ_δ of width $\delta > 0$ around Γ such that $\gamma \in C^3(\Gamma_\delta)$.

We define the outward unit normal

$$n(v) = \frac{\nabla\gamma(v)}{|\nabla\gamma(v)|} \tag{3.36}$$

for $v \in \Gamma$. This definition is independent of the choice of γ.

The projection onto the tangent plane $T_v\Gamma$ for $v \in \Gamma$ is given by

$$P_\Gamma(v) = I - n(v) \otimes n(v). \tag{3.37}$$

It follows immediately that $P_\Gamma(v)n(v) = 0$ and $P_\Gamma(v)w \cdot n(v) = 0$ for all $w \in \mathbb{R}^d$. We further define the projection Π_Γ as

$$\Pi_\Gamma(v) = v - d(v)n(v) \tag{3.38}$$

for $v \in \Gamma_\delta$, where $d(v)$ is the distance of v to Γ. It is easily seen that $\Pi_\Gamma(v) \in \Gamma$. Here we require $\delta > 0$ to be sufficiently small such that the decomposition $v = \Pi_\Gamma(v) + d(v)n(v)$ is unique. The distinction between P_Γ and Π_Γ is important: P_Γ takes an element of the hypersurface $v \in \Gamma$ as input, so that $P_\Gamma(v)$ projects any element $w \in \mathbb{R}^d$ onto the tangent plane $T_v\Gamma$. On the other hand, Π_Γ projects a vector in the tube Γ_δ orthogonally onto Γ. The latter projection plays an imortant role in the numerical implementation of the method: due to numerical errors the particles will not stay exactly on Γ and hence need to be projected back. For the sphere we have $\Pi_\Gamma(v) = v/|v|$.

With Π_Γ we can extend a function $\Phi \in C^\infty(\Gamma)$ to $\mathbb{R}^d$ by setting

$$\phi(v) = \begin{cases} \Phi(\Pi_\Gamma(v)), & v \in \Gamma_\delta \\ 0, & v \in \mathbb{R}^d \setminus \Gamma_{\delta'} \end{cases} \tag{3.39}$$

where $\Gamma_{\delta'}$ is a slightly larger neighborhood than Γ_δ. The region $\Gamma_{\delta'} \setminus \Gamma_\delta$ is used to smoothly bring the function value to 0. That is we extend Φ constantly in normal direction. The function ϕ is a smooth function (as a composition of smooth functions) and $\phi \in C_c^\infty(\mathbb{R}^d)$.

It is easily seen that $\Pi_\Gamma(v) = \Pi_\Gamma(\lambda v)$ for all $\lambda \in \mathbb{R}$ sufficiently small such that $\lambda v \in \Gamma_\delta$. From this we conclude that ϕ is 0-homogeneous $\phi(\lambda v) = \lambda^0 \phi(v) = \phi(v)$. Since ϕ is a constant extension of Φ in normal direction we conclude

$$n(v) \cdot \nabla \phi(v) = 0 \tag{3.40}$$

from which we further get

$$\mathrm{Tr}(\nabla^2 \phi(v)^T (n(v) \otimes n(v))^T) = 0. \tag{3.41}$$

We define the tangential gradient, Hessian and Laplace-Beltrami operator of Φ as

$$\begin{cases} \nabla_\Gamma \Phi(v) := \nabla \phi(v) = \nabla \Phi(\Pi_\Gamma(v)) \\ \nabla_\Gamma^2 \Phi(v) := \nabla^2 \phi(v) = \nabla^2 \Phi(\Pi_\Gamma(v)) \\ \Delta_\Gamma \Phi(v) := \Delta \phi(v) = \Delta \Phi(\Pi_\Gamma(v)) \end{cases} \tag{3.42}$$

for $v \in \Gamma_\delta$. We stress: $n(v) \cdot \nabla_\Gamma \Phi(v) = 0$, hence $\nabla_\Gamma \Phi(v) \in T_v \Gamma$.

3.5 KV-CBO model on hypersurfaces

In this section we state the final KV-CBO model with isotropic noise on general compact hypersurfaces. The anisotropic model is stated only for the special case $\Gamma = \mathbb{S}^{d-1}$. Extending the analysis of the anisotropic model to general compact hypersurfaces is open.

Isotropic KV-CBO model on hypersurfaces

The KV-CBO model with isotropic noise on a compact hypersurface Γ is given by

$$dV_t^i = \lambda P_\Gamma(V_t^i)V^\alpha(\rho_t^N)dt + \sigma|V_t^i - V^\alpha(\rho_t^N)|P_\Gamma(V_t^i)dB_t^i + C_\Gamma(V_t^i, V^\alpha(\rho_t^N))dt \tag{3.43}$$

where $V_0^i \sim \rho_0$ iid for $i \in [N]$, $\rho_0 \in \mathcal{P}(\Gamma)$ is a common probability distribution, and

$$C_\Gamma(V_t^i, V^\alpha(\rho_t^N)) = -\frac{\sigma^2}{2}|V_t^i - V^\alpha(\rho_t^N)|^2\Delta\gamma(V_t^i)\nabla\gamma(V_t^i) \tag{3.44}$$

where γ defines the hypersurface, see (3.35). The empirical measure ρ_t^N and consensus point $V^\alpha(\rho_t^N)$ are given by

$$\rho_t^N = \sum_{i=1}^N \delta_{V_t^i} \tag{3.45}$$

and

$$V^\alpha(\rho_t^N) = \frac{\sum_{i=1}^N V_t^i e^{-\alpha\mathcal{E}(V_t^i)}}{\sum_{i=1}^N e^{-\alpha\mathcal{E}(V_t^i)}} = \frac{\int_{\mathbb{R}^d} ve^{-\alpha\mathcal{E}(v)}d\rho_t^N(v)}{\int_{\mathbb{R}^d} e^{-\alpha\mathcal{E}(v)}d\rho_t^N(v)} \tag{3.46}$$

for $t \in [0, T]$ respectively.

Example 3.5.1 (Special case $\Gamma = \mathbb{S}^{d-1}$)**.** *For $\Gamma = \mathbb{S}^{d-1}$ we can choose $\gamma(v) = |v| - 1$. Then $\nabla\gamma(v) = v/|v|$ and the outward unit normal is given by $n(v) = \frac{v}{|v|}$. The Hessian and Laplacian are given by*

$$\nabla^2\gamma(v) = \frac{1}{|v|}I_d - \frac{v \otimes v}{|v|^3}, \quad \Delta\gamma(v) = \mathrm{Tr}(\nabla^2\gamma(v)) = \frac{d-1}{|v|}. \tag{3.47}$$

The correction term is given by

$$C_\Gamma(V_t^i, V^\alpha(\rho_t^N)) = -\frac{\sigma^2}{2}(d-1)|V_t^i - V^\alpha(\rho_t^N)|^2\frac{V_t^i}{|V_t^i|^2} \tag{3.48}$$

as in (3.26)*.*

Motivated by the findings for the CBO models in Euclidean space (see section 2.3), we now introduce a version of the KV-CBO model with anisotropic (component-wise) Brownian motion. More precisely: we replace the isotropic noise term $\sigma P_\Gamma(V_t^i)|V_t^i - V^\alpha(\rho_t^N)|dB_t^i$ by the anisotropic noise

$$\sigma P_\Gamma(V_t^i)D(V_t^i - V^\alpha(\rho_t^N))dB_t^i := \sigma\sum_{k=1}^d P_\Gamma(V_t^i)(V_t^i - V^\alpha(\rho_t^N))_k(dB_t^i)_k e_k. \tag{3.49}$$

We analyse the anisotropic KV-CBO model in the appendix.

Anisotropic KV-CBO model on the sphere

The KV-CBO model with anisotropic noise on $\Gamma = \mathbb{S}^{d-1}$ is given by

$$dV_t^i = \lambda P_\Gamma(V_t^i)V^\alpha(\rho_t^N)dt + \sigma P_\Gamma(V_t^i)D(V_t^i - V^\alpha(\rho_t^N))dB_t^i + C_\Gamma^{ani}(V_t^i, V^\alpha(\rho_t^N))dt \tag{3.50}$$

where the correction term is given by

$$\begin{aligned} C_\Gamma^{ani}(V_t^i, V^\alpha(\rho_t^N)) = &-\frac{\sigma^2}{2}|V_t^i - V^\alpha(\rho_t^N)|^2\frac{V_t^i}{|V_t^i|^2} - \frac{\sigma^2}{2}D(V_t^i - V^\alpha(\rho_t^N))^2\frac{V_t^i}{|V_t^i|^2} \\ &+ \sigma^2|D(V_t^i - V^\alpha(\rho_t^N))V_t^i|^2\frac{V_t^i}{|V_t^i|^4}. \end{aligned} \tag{3.51}$$

We conclude this section with a few remarks.

Remark 3.5.1.

1. *The KV-CBO model can be understood as a single SDE*

$$dV_t^{(N)} = \lambda(V_t^{(N)})dt + \sigma(V_t^{(N)})dB_t \tag{3.52}$$

   in the high-dimensional space $\mathbb{R}^{Nd}$. Here $V_t^{(N)} \in \mathbb{R}^{Nd}$ is a vector with the particles $V_t^i \in \mathbb{R}^d$ stacked on top of each other for $i \in [N]$.

2. *The particles V_t^i are vectors in the Euclidean space $\mathbb{R}^d$ that will later be shown to stay on Γ for all time. Deriving and analyzing a CBO model on hypersurfaces where the dynamics is defined instrinsically on the hypersurface is an open problem.*

3. *The integral form of* (3.50) *is given by*

$$\begin{aligned} V_t^i = V_0^i + \lambda\int_0^t P_\Gamma(V_\tau^i)V^\alpha(\rho_\tau^N)\ d\tau + \sigma\int_0^t |V_\tau^i - V^\alpha(\rho_\tau^N)|P_\Gamma(V_\tau^i)\ dB_\tau^i \\ + \int_0^t C_\Gamma(V_\tau^i, V^\alpha(\rho_\tau^N))\ d\tau. \end{aligned} \tag{3.53}$$

3.6 Stability estimates for the consensus point

In this section we collect some stability estimates and a linear growth condition for the consensus point respectively the regularized consensus point. These stability estimates play an important role in the rest of this thesis, in particular for the proof of the well-posedness of the KV-CBO model.

Let the regularized cost function $\boldsymbol{\mathcal{E}} : \mathbb{R}^d \to \mathbb{R}$ satisfy the following properties

1. $\boldsymbol{\mathcal{E}}(v) = \mathcal{E}(v)$ for all $v \in \Gamma_\delta$,

2. $\boldsymbol{\mathcal{E}}$ is globally Lipschitz continuous with constant $L > 0$,

3. $\boldsymbol{\mathcal{E}}$ is bounded.

We use the following convention throughout the thesis: regularized coefficient functions, cost functions, constants etc. are all written in bold. For example: the regularized version of the cost function $\mathcal{E}$ will be denoted as $\boldsymbol{\mathcal{E}}$, the regularized version of the consensus point $V^\alpha(\rho_t^N)$ will be denoted as $\mathbf{V}^\alpha(\rho_t^N)$, and so forth.

A lot of the estimates that we will derive in this section involve the constant

$$\mathbf{C}_\alpha = e^{\alpha(\overline{\boldsymbol{\mathcal{E}}} - \underline{\boldsymbol{\mathcal{E}}})} \tag{3.54}$$

where $\overline{\boldsymbol{\mathcal{E}}} = \sup_\Gamma \boldsymbol{\mathcal{E}}$ and $\underline{\boldsymbol{\mathcal{E}}} = \inf_\Gamma \boldsymbol{\mathcal{E}}$.

We introduce the following notations

$$\|V_t^{(N)}\|_1 = \sum_{i=1}^N |V_t^i| \tag{3.55}$$

and

$$\|V_t^{(N)}\|_\infty = \sup_{i \in [N]} |V_t^i| \tag{3.56}$$

where $V_t^{(N)} \in \mathbb{R}^{Nd}$, see remark 3.5.1.

Lemma 3.6.1. *Let V_t^i and $\widehat{V}_t^i$ be two independent sets of particles with empirical measures ρ_t^N and $\widehat{\rho}_t^N$ respectively. Then the estimate*

$$|\,\boldsymbol{V}^\alpha(\widehat{\rho}_t^N) - \boldsymbol{V}^\alpha(\rho_t^N)| \lesssim \frac{\|V_t^{(N)} - \widehat{V}_t^{(N)}\|_1}{N} \tag{3.57}$$

where the constant is given by $C = \boldsymbol{C}_\alpha + 2\alpha L \boldsymbol{C}_\alpha^2 \|\widehat{V}_t^{(N)}\|_\infty$, holds.

Proof. We introduce the notation $\omega_j = e^{-\alpha \boldsymbol{\mathcal{E}}(V_t^j)}$ for readability and decompose the difference of the consensus points as

$$\mathbf{V}^\alpha(\widehat{\rho}_t^N) - \mathbf{V}^\alpha(\rho_t^N) = \frac{\sum_j \widehat{V}_t^j \widehat{\omega}_j}{\sum_j \widehat{\omega}_j} - \frac{\sum_j V_t^j \omega_j}{\sum_j \omega_j} = \frac{\sum_j \widehat{V}_t^j \widehat{\omega}_j \sum_j \omega_j - \sum_j V_t^j \omega_j \sum_j \widehat{\omega}_j}{\sum_j \widehat{\omega}_j \sum_j \omega_j}. \tag{3.58}$$

With this we can rewrite

$$\mathbf{V}^\alpha(\rho_t^N) - \mathbf{V}^\alpha(\widehat{\rho}_t^N) = \frac{\sum_j (V_t^j - \widehat{V}_t^j)\omega_j}{\sum_j \omega_j} + \frac{\sum_j \widehat{V}_t^j (\omega_j - \widehat{\omega}_j)}{\sum_j \omega_j} + \frac{\sum_j \widehat{V}_t^j \omega_j \sum_j (\widehat{\omega}_j - \omega_j)}{(\sum_j \omega_j)(\sum_j \widehat{\omega_j})}. \tag{3.59}$$

The first term on the right hand side of (3.59) can be bounded as

$$|I_1| \leq \frac{\mathbf{C}_\alpha}{N} \|V_t^{(N)} - \widehat{V}_t^{(N)}\|_1 \tag{3.60}$$

where we have used $\omega_j \leq e^{-\alpha \underline{\boldsymbol{\mathcal{E}}}}$ and $\omega_j \geq e^{-\alpha \overline{\boldsymbol{\mathcal{E}}}}$. The second term on the right hand side of (3.59) can be bounded as

$$|I_2| \leq \frac{\alpha L \mathbf{C}_\alpha}{N} \sum_{j=1}^{N} |\widehat{V}_t^j||V_t^j - \widehat{V}_t^j| = \frac{\alpha L \mathbf{C}_\alpha}{N} \|\widehat{V}_t^{(N)}\|_\infty \|V_t^{(N)} - \widehat{V}_t^{(N)}\|_1 \tag{3.61}$$

where we have used the mean value theorem and the Lipschitz assumption of $\boldsymbol{\mathcal{E}}$, that is,

$$|\omega_j - \widehat{\omega}_j| \leq \alpha e^{-\alpha \underline{\boldsymbol{\mathcal{E}}}} |\boldsymbol{\mathcal{E}}(V_t^j) - \boldsymbol{\mathcal{E}}(\widehat{V}_t^j)| \leq \alpha L e^{-\alpha \underline{\boldsymbol{\mathcal{E}}}} |V_t^j - \widehat{V}_t^j|. \tag{3.62}$$

The third term on the right hand side of (3.59) can be bounded as

$$|I_3| \leq \frac{\sum_j |\widehat{V}_t^j| e^{-\alpha \underline{\boldsymbol{\mathcal{E}}}} \sum_j |\widehat{\omega}_j - \omega_j|}{|\sum_j \omega_j||\sum_j \widehat{\omega}_j|} \leq \frac{\alpha L \mathbf{C}_\alpha^2}{N} \|\widehat{V}_t^{(N)}\|_\infty \|V_t^{(N)} - \widehat{V}_t^{(N)}\|_1 \tag{3.63}$$

where we have used (3.62).

Thus we conclude that

$$\max\{|I_2|, |I_3|\} \leq \frac{\alpha L \mathbf{C}_\alpha^2}{N} \|\widehat{V}_t^{(N)}\|_\infty \|V_t^{(N)} - \widehat{V}_t^{(N)}\|_1. \tag{3.64}$$

which finishes the proof. □

Lemma 3.6.2. *Let $\rho, \widehat{\rho} \in \mathcal{P}_c(\mathbb{R}^d)$ two probability measures with compact support and $1 \leq p < \infty$. The distance of the regularized consensus points are bounded by*

$$|\boldsymbol{V}^\alpha(\rho) - \boldsymbol{V}^\alpha(\widehat{\rho})| \lesssim W_p(\rho, \widehat{\rho}) \tag{3.65}$$

where W_p is the p-Wasserstein distance.

Proof. We have

$$\mathbf{V}^\alpha(\rho) - \mathbf{V}^\alpha(\widehat{\rho}) = \frac{\int_{\mathbb{R}^d} v e^{-\alpha \boldsymbol{\mathcal{E}}(v)} \, d\rho(v)}{\|e^{-\alpha \boldsymbol{\mathcal{E}}}\|_{L^1(\rho)}} - \frac{\int_{\mathbb{R}^d} \widehat{v} e^{-\alpha \boldsymbol{\mathcal{E}}(\widehat{v})} \, d\widehat{\rho}(\widehat{v})}{\|e^{-\alpha \boldsymbol{\mathcal{E}}}\|_{L^1(\widehat{\rho})}} \tag{3.66}$$

$$= \iint_{\mathbb{R}^d \times \mathbb{R}^d} \underbrace{\frac{v e^{-\alpha \boldsymbol{\mathcal{E}}(v)}}{\|e^{-\alpha \boldsymbol{\mathcal{E}}}\|_{L^1(\rho)}} - \frac{\widehat{v} e^{-\alpha \boldsymbol{\mathcal{E}}(\widehat{v})}}{\|e^{-\alpha \boldsymbol{\mathcal{E}}}\|_{L^1(\widehat{\rho})}}}_{:= h(v) - h(\widehat{v})} d\pi(v, \widehat{v}) \tag{3.67}$$

where $\pi \in \Pi(\rho, \widehat{\rho})$ is an arbitrary coupling of ρ and $\widehat{\rho}$. Similarly as in (3.58) we can write the integrand

as

$$h(v)-h(\widehat{v})=\underbrace{\frac{(v-\widehat{v})e^{-\alpha\boldsymbol{\mathcal{E}}(\widehat{v})}}{\|e^{-\alpha\boldsymbol{\mathcal{E}}}\|_{L^1(\rho)}}}_{=:I_1}+\underbrace{\frac{\widehat{v}(e^{-\alpha\boldsymbol{\mathcal{E}}(v)}-e^{-\alpha\boldsymbol{\mathcal{E}}(\widehat{v})})}{\|e^{-\alpha\boldsymbol{\mathcal{E}}}\|_{L^1(\rho)}}}_{=:I_2}+\underbrace{\frac{\widehat{v}e^{-\alpha\boldsymbol{\mathcal{E}}(\widehat{v})}}{\|e^{-\alpha\boldsymbol{\mathcal{E}}}\|_{L^1(\rho)}}-\frac{\widehat{v}e^{-\alpha\boldsymbol{\mathcal{E}}(\widehat{v})}}{\|e^{-\alpha\boldsymbol{\mathcal{E}}}\|_{L^1(\widehat{\rho})}}}_{:=I_3} \tag{3.68}$$

where the last two terms on the right hand side can be rewritten as

$$\begin{aligned}
I_3 =& \frac{\widehat{v}e^{-\alpha\boldsymbol{\mathcal{E}}(\widehat{v})}}{\|e^{-\alpha\boldsymbol{\mathcal{E}}}\|_{L^1(\rho)}}-\frac{\widehat{v}e^{-\alpha\boldsymbol{\mathcal{E}}(\widehat{v})}}{\|e^{-\alpha\boldsymbol{\mathcal{E}}}\|_{L^1(\widehat{\rho})}} && (3.69)\\
=& \frac{\widehat{v}e^{-\alpha\boldsymbol{\mathcal{E}}(\widehat{v})}\int e^{-\alpha\boldsymbol{\mathcal{E}}(\widehat{v})}\,d\widehat{\rho}(\widehat{v})}{\|e^{-\alpha\boldsymbol{\mathcal{E}}}\|_{L^1(\rho)}\|e^{-\alpha\boldsymbol{\mathcal{E}}}\|_{L^1(\widehat{\rho})}}-\frac{\widehat{v}e^{-\alpha\boldsymbol{\mathcal{E}}(\widehat{v})}\int e^{-\alpha\boldsymbol{\mathcal{E}}(v)}\,d\rho(v)}{\|e^{-\alpha\boldsymbol{\mathcal{E}}}\|_{L^1(\widehat{\rho})}\|e^{-\alpha\boldsymbol{\mathcal{E}}}\|_{L^1(\rho)}} && (3.70)\\
=& \frac{\widehat{v}e^{-\alpha\boldsymbol{\mathcal{E}}(\widehat{v})}\iint_{\mathbb{R}^d\times\mathbb{R}^d}(e^{-\alpha\boldsymbol{\mathcal{E}}(\widehat{v})}-e^{-\alpha\boldsymbol{\mathcal{E}}(v)})d\pi(v,\widehat{v})}{\|e^{-\alpha\boldsymbol{\mathcal{E}}}\|_{L^1(\rho)}\|e^{-\alpha\boldsymbol{\mathcal{E}}}\|_{L^1(\widehat{\rho})}}. && (3.71)
\end{aligned}$$

We have

$$|I_1|\le\mathbf{C}_\alpha|v-\widehat{v}|,\quad |I_2|\le\alpha\mathbf{C}_\alpha L|\widehat{v}||v-\widehat{v}| \tag{3.72}$$

and

$$|I_3|\le\alpha\mathbf{C}_\alpha^2L|\widehat{v}|\iint_{\mathbb{R}^d\times\mathbb{R}^d}|v-\widehat{v}|d\pi(v,\widehat{v}). \tag{3.73}$$

Now we use the fact that the measures have compact support which allows us to bound the p-th moments as

$$\max\left\{\int_{\mathbb{R}^d}|v|^pd\rho(v),\int_{\mathbb{R}^d}|\widehat{v}|^pd\widehat{\rho}(\widehat{v})\right\}<\infty\,. \tag{3.74}$$

Collecting the above estimates, we obtain

$$\begin{aligned}
|\mathbf{V}^\alpha(\rho)-\mathbf{V}^\alpha(\widehat{\rho})|\le&\,\mathbf{C}_\alpha\iint_{\mathbb{R}^d\times\mathbb{R}^d}|v-\widehat{v}|d\pi(v,\widehat{v})+\alpha\mathbf{C}_\alpha L\iint_{\mathbb{R}^d\times\mathbb{R}^d}|\widehat{v}||v-\widehat{v}|d\pi(v,\widehat{v})\\
&+\alpha\mathbf{C}_\alpha^2L\int_{\mathbb{R}^d}|\widehat{v}|d\widehat{\rho}(\widehat{v})\iint_{\mathbb{R}^d\times\mathbb{R}^d}|v-\widehat{v}|d\pi(v,\widehat{v}) && (3.75)\\
\le&\,C\left(\iint_{\mathbb{R}^d\times\mathbb{R}^d}|v-\widehat{v}|^pd\pi(v,\widehat{v})\right)^{\frac{1}{p}} && (3.76)
\end{aligned}$$

where C depends only on $\mathbf{C}_\alpha$ and α, L. Optimizing over all couplings $\pi\in\Pi(\rho,\hat{\rho})$ finishes the proof. □

Lemma 3.6.3. *Let $V_t^{(N)}\in\mathbb{R}^{dN}$ be a set of N particles and ρ_t^N the corresponding empirical measure. Then the following linear growth condition holds*

$$|\,\boldsymbol{V}^\alpha(\rho_t^N)|\le\frac{\boldsymbol{C}_\alpha}{N}\|V_t^{(N)}\|_1. \tag{3.77}$$

We further have

$$|\,\boldsymbol{V}^\alpha(\rho_t)|\le\boldsymbol{C}_\alpha\int_{\mathbb{R}^d}|v|\,d\rho_t(v). \tag{3.78}$$

In the next section we will show that the particles stay on Γ if initially chosen there. Let us assume that this result is already known. Then the following lemma holds.

Lemma 3.6.4. *Let $V \in C([0,T],\Gamma)$ be a particle driven by the KV-CBO model and denote the probability distribution of V by $\rho \in C([0,T],\mathcal{P}(\Gamma))$. Then the map $t \mapsto V^\alpha(\rho_t)$ is Lipschitz continuous, that is,*

$$|V^\alpha(\rho_t) - V^\alpha(\rho_s)| \lesssim |t-s|. \tag{3.79}$$

Proof. Let $s,t \in [0,T]$ with $s<t$, then

$$|V^\alpha(\rho_t) - V^\alpha(\rho_s)| \lesssim W_2(\rho_t,\rho_s) \leq (\mathbb{E}|V_t - V_s|^2)^{1/2} \tag{3.80}$$

where we have used (3.65) and (2.67).

We further have

$$\begin{aligned}\mathbb{E}|V_t - V_s|^2 &\lesssim \mathbb{E}\Big|\int_s^t \lambda P_\Gamma(V_\tau)V^\alpha(\rho_\tau)\,d\tau\Big|^2 + \mathbb{E}\Big|\int_s^t \sigma|V_\tau - V^\alpha(\rho_\tau)|P_\Gamma(V_\tau)\,dB_\tau\Big|^2 \\ &\quad + \mathbb{E}\Big|\int_s^t \frac{\sigma^2}{2}|V_\tau - V^\alpha(\rho_\tau)|^2\nabla\gamma(V_\tau)\Delta\gamma(V_\tau)\,d\tau\Big|^2 \qquad (3.81)\\ &\lesssim |t-s|^2 + \mathbb{E}\Big(\int_s^t |V_\tau - V^\alpha(\rho_\tau)|^2\,d\tau\Big) + \mathbb{E}\Big(\int_s^t |V_\tau - V^\alpha(\rho_\tau)|^2\,d\tau\Big)^2 \qquad (3.82)\end{aligned}$$

where we have used the integral form of the KV-CBO model in the first step and Itô's isometry (see appendix) in the second step. Noting that $\tau \mapsto |V_\tau - V^\alpha(\rho_\tau)|^2$ is a continuous function on a compact domain further yields

$$\mathbb{E}|V_t - V_s|^2 \lesssim |t-s|^2. \tag{3.83}$$

Combining (3.80) and (3.83) finishes the proof. □

Remark 3.6.1. *In the following chapter we will prove another stability estimate on the consensus point, see lemma 4.4.1. This lemma is not reported here as it requires knowledge of mean-field theory which we are going to discuss in the next chapter.*

3.7 Well-posedness

In this section we prove the *well-posedness* of the KV-CBO model with isotropic noise (3.43). The corresponding proof for the anisotropic model can be found in the appendix.

The mathematician Jacques Hadamard proposed the following conditions for a mathematical model to be *well-posed*: there exists a solution, the solution is unique, and it depends continuously on the initial data. A mathematical model could be given by a linear PDE with suitable initial and boundary conditions, e.g. the Laplace equation with Dirichlet boundary conditions. In our case, however, the

mathematical model to investigate is much more complex than a linear PDE. We want to show the well-posedness of a particle system given by a coupled system of non-linear, non-local SDEs. In this thesis we call a CBO model well-posed iff there exists a *pathwise unique strong solution* for any initial probability distribution $\rho_0 \in \mathcal{P}(\Gamma)$. It is important that the initial probability distribution is supported on Γ.

To prove the well-posedness of the KV-CBO model we need to prove that 1) the coefficient functions are *well-defined* for any $V_t^{(N)}$ we might plug in, and 2) that the conditions from theorem 2.4.1, namely the local Lipschitz condition of the coefficient functions, are satisfied. This theorem then guarantees the existence of a pathwise unique local (!) strong solution. With Itô's formula this solution can be shown to stay on Γ for some finite time $\tilde{t} > 0$. Hence the theorem can be applied a second time in order to extent the local solution a little bit. Repeating this process yields a global solution.

Let us now consider $\Gamma = \mathbb{S}^{d-1}$ to illustrate the problem in showing the well-posedness. In this case $\Delta\gamma$ and $\nabla\gamma$ in the definition of the correction term are given by $\Delta\gamma(v) = (d-1)/|v|$ and $\nabla\gamma(v) = v/|v|$. So, if $v = 0$ these functions *blow-up*. Hence the correction term is not well-defined at the origin. Further, we see that also the projection $P_\Gamma(v) = I - v \otimes v/|v|^2$ is not well-defined for $v = 0$ (what is the projection onto the orthogonal complement of the zero vector?). In this context we sometimes use the expression P_Γ *has a singularity at the origin*. In order to have well-defined coefficient functions we need the particles, driven by Brownian motion (!), to stay away from the origin $v = 0$.

The proof of the well-posedness of the KV-CBO model thus requires the introduction of a *regularized KV-CBO model* with coefficient functions equal to the original coefficient functions for function values on Γ. Outside the tube Γ_δ, however, the regularized coefficient functions have no singularities and no blow-up points.

We regularize P_Γ, $\nabla\gamma$, and $\Delta\gamma$ with appropriate operators $\mathbf{P}_1$, $\mathbf{P}_2$, and $\mathbf{P}_3$ respectively. Let $\mathbf{P}_1 : \mathbb{R}^d \to \mathbb{R}^{d\times d}$, $\mathbf{P}_2 : \mathbb{R}^d \to \mathbb{R}$ and $\mathbf{P}_3 : \mathbb{R}^d \to \mathbb{R}^d$ be maps with bounded derivatives such that $\mathbf{P}_1(v) = P_\Gamma(v)$, $\mathbf{P}_2(v) = \Delta\gamma(v)$ and $\mathbf{P}_3(v) = \nabla\gamma(v)$ for all $v \in \Gamma_\delta$ where Γ_δ is an open neighbourhood of Γ, see section 3.4. The *regularized KV-CBO model* is given by

$$dV_t^i = \lambda \mathbf{P}_1(V_t^i)\mathbf{V}^\alpha(\rho_t^N)dt + \sigma|V_t^i - \mathbf{V}^\alpha(\rho_t^N)|\mathbf{P}_1(V_t^i)dB_t^i + \mathbf{C}_\Gamma(V_t^i, \mathbf{V}^\alpha(\rho_t^N))dt \tag{3.84}$$

where

$$\mathbf{C}_\Gamma(V_t^i, \mathbf{V}^\alpha(\rho_t^N)) = -\frac{\sigma^2}{2}|V_t^i - \mathbf{V}^\alpha(\rho_t^N)|^2\mathbf{P}_2(V_t^i)\mathbf{P}_3(V_t^i). \tag{3.85}$$

Lemma 3.7.1. *Let $V_t^{(N)}, \widehat{V}_t^{(N)} \in \mathbb{R}^{Nd}$ be two sets of N independent particles in $\mathbb{R}^d$ with corresponding empirical measures $\rho_t^N, \widehat{\rho}_t^N$. Then the coefficient functions of the regularized KV-CBO model*

$$a_i(V_t^{(N)}) = \lambda \boldsymbol{P}_1(V_t^i)\, \boldsymbol{V}^\alpha(\rho_t^N), \tag{3.86}$$

$$b_i(V_t^{(N)}) = \sigma|V_t^i - \boldsymbol{V}^\alpha(\rho_t^N)|\boldsymbol{P}_1(V_t^i), \tag{3.87}$$

$$c_i(V_t^{(N)}) = \boldsymbol{C}_\gamma(V_t^i, \boldsymbol{V}^\alpha(\rho_t^N)) = -\frac{\sigma^2}{2}|V_t^i - \boldsymbol{V}^\alpha(\rho_t^N)|^2 \boldsymbol{P}_2(V_t^i)\boldsymbol{P}_3(V_t^i) \tag{3.88}$$

are locally Lipschitz continuous for all $i \in [N]$.

Proof. We have (see (3.57))

$$|\mathbf{V}^\alpha(\widehat{\rho}_t^N) - \mathbf{V}^\alpha(\rho_t^N)| \leq \left(\frac{\mathbf{C}_\alpha}{N} + \frac{2\alpha L \mathbf{C}_\alpha^2}{N}\|\widehat{V}_t^{(N)}\|_\infty\right)\|V_t^{(N)} - \widehat{V}_t^{(N)}\|_1. \tag{3.89}$$

From this stability estimate we can easily deduce the local Lipschitz conditions. We note that the appearance of the factor $\|\widehat{V}_t^{(N)}\|_\infty$ is what makes the Lipschitz condition *local*. For any set of particles $\widehat{V}_t^{(N)}$ there is a neighbourhood such that $\|\widehat{V}_t^{(N)}\|_\infty$ is just a constant.

Let us now show the local Lipschitz condition for the coefficient $a_i(V_t^{(N)})$. We have

$$|a_i(V_t^{(N)}) - a_i(\widehat{V}_t^{(N)})| \lesssim \|V_t^{(N)} - \widehat{V}_t^{(N)}\|_1 \tag{3.90}$$

by (3.57). For $b_i(V_t^{(N)})$ the condition follows similarly. For $c_i(V_t^{(N)})$ we first note that $\mathbf{P}_2$, $\mathbf{P}_3$ have bounded derivatives, hence we can bound

$$|c_i(V_t^{(N)}) - c_i(\widehat{V}_t^{(N)})| \lesssim \big||V_t^i - \mathbf{V}^\alpha(\rho_t^N)|^2 - |\widehat{V}_t^i - \widehat{\mathbf{V}}^\alpha(\rho_t^N)|^2\big| \tag{3.91}$$

where we have used

$$|\mathbf{P}_2(V_t^i)\mathbf{P}_3(V_t^i)| \leq \|\nabla\mathbf{P}_2\|_\infty\|\nabla\mathbf{P}_3\|_\infty|V_t^i|^2 \lesssim \|V_t^{(N)}\|_\infty. \tag{3.92}$$

Further we note that $V_t^{(N)} \mapsto |V_t^i - \mathbf{V}^\alpha(\rho_t^N)|^2$ is a polynomial and hence is locally Lipschitz continuous. This finishes the proof. □

We have proven that the coefficient functions of the regularized KV-CBO model satisfy a local Lipschitz condition. These are the main ingredients to prove the well-posedness of the original, non-regularized KV-CBO model (3.43). We proceed as follows: we first deduce that the regularized model is well-posed, then we prove that the particles stay on Γ if they were initially chosen there, that is, $V_0^i \sim \rho_0$ implies that $V_t^i \in \Gamma$ for all $t > 0$ and $i \in [N]$. This implies that the regularized and the orginal model coindice for initial particles $V_0^i \sim \rho_0$.

We now prove the well-posedness. We remind the reader that here a *strong solution* is a strong solution of an SDE (not a PDE), see section 2.4.

Theorem 3.7.1 (Well-posedness)**.** *The KV-CBO model* (3.43) *is well-posed, that is, for every initial probability distribution* $\rho_0 \in \mathcal{P}(\Gamma)$ *there is a pathwise unique strong solution.*

Proof. The regularized model (3.84) has locally Lipschitz continuous coefficients, which implies the existence of a pathwise unique *local* (!) strong solution by theorem 2.4.1. Let us assume that this local solution stays in a small tube around Γ, that is, $V_t^i \in \Gamma_\delta$. This allows us to apply Itô's formula to prove that the solution stays on Γ (not just in a small neighbourhood around Γ).

Let γ the level set function that defines the Hypersurface Γ. An application of the multivariate Itô's formula yields

$$\begin{aligned} d\gamma(V_t^i) = \nabla\gamma(V_t^i)\cdot(a_i(V_t^{(N)}) + c_i(V_t^{(N)}))dt + \frac{1}{2}\sum_{k,l=1}^{d}\partial_{kl}^2\gamma(V_t^i)b_{ik}(V_t^{(N)})\otimes b_{il}(V_t^{(N)})dt \\ + \nabla\gamma(V_t^i)\cdot b_i(V_t^{(N)})dB_t^i \end{aligned} \tag{3.93}$$

where b_{ik} is the kth row of b_i given by

$$b_{ik}(V_t^{(N)}) = \sigma|V_t^i - \mathbf{V}^\alpha(\rho_t^N)|\Big(e_k^T - \partial_k\gamma(V_t^i)\nabla\gamma(V_t^i)^T\Big). \tag{3.94}$$

We now simplify the double sum on the right hand side of (3.93) by first noting

$$b_{ik}(V_t^{(N)})\otimes b_{il}(V_t^{(N)}) \tag{3.95}$$

$$= \sigma^2|V_t^i - \mathbf{V}^\alpha(\rho_t^N)|^2\Big(e_k^T - \partial_k\gamma(V_t^i)\nabla\gamma(V_t^i)^T\Big)\otimes\Big(e_l^T - \partial_l\gamma(V_t^i)\nabla\gamma(V_t^i)^T\Big) \tag{3.96}$$

$$= \sigma^2|V_t^i - \mathbf{V}^\alpha(\rho_t^N)|^2\Big(\delta_{k,l} - \partial_l\gamma(V_t^i)\partial_k\gamma(V_t^i)\Big) \tag{3.97}$$

where we have used that $\nabla\gamma(V_t^i)^T\otimes\nabla\gamma(V_t^i)^T = |\nabla\gamma(V_t^i)|^2 = 1$. This allows us to rewrite

$$\frac{1}{2}\sum_{k,l}\partial_{kl}^2\gamma(V_t^i)b_{ik}(V_t^{(N)})\otimes b_{il}(V_t^{(N)}) \tag{3.98}$$

$$= \frac{\sigma^2}{2}|V_t^i - \mathbf{V}^\alpha(\rho_t^N)|^2\sum_{k,l}\partial_{kl}^2\gamma(V_t^i)\Big(\delta_{k,l} - \partial_l\gamma(V_t^i)\partial_k\gamma(V_t^i)\Big) \tag{3.99}$$

$$= \frac{\sigma^2}{2}|V_t^i - \mathbf{V}^\alpha(\rho_t^N)|^2\Big(\Delta\gamma(V_t^i) - \sum_k(\partial_k\nabla\gamma(V_t^i)\cdot\nabla\gamma(V_t^i))\partial_k\gamma(V_t^i)\Big) \tag{3.100}$$

$$= \frac{\sigma^2}{2}|V_t^i - \mathbf{V}^\alpha(\rho_t^N)|^2\Big(\Delta\gamma(V_t^i) - \sum_k\partial_k|\nabla\gamma(V_t^i)|^2\partial_k\gamma(V_t^i)\Big) \tag{3.101}$$

$$= \frac{\sigma^2}{2}|V_t^i - \mathbf{V}^\alpha(\rho_t^N)|^2\Delta\gamma(V_t^i) \tag{3.102}$$

where we have used $|\nabla\gamma(V_t^i)| = 1$ in the last step.

This allows us to rewrite (3.93) as

$$\begin{aligned} d\gamma(V_t^i) &= \nabla\gamma(V_t^i)\cdot(a_i(V_t^{(N)}) + c_i(V_t^{(N)}))dt + \frac{\sigma^2}{2}|V_t^i - \mathbf{V}^\alpha(\rho_t^N)|^2\Delta\gamma(V_t^i)dt \\ &\quad + \nabla\gamma(V_t^i)\cdot b_i(V_t^{(N)})dB_t^i \end{aligned} \tag{3.103}$$

$$= \nabla\gamma(V_t^i)\cdot c_i(V_t^{(N)})dt + \frac{\sigma^2}{2}|V_t^i - \mathbf{V}^\alpha(\rho_t^N)|^2\Delta\gamma(V_t^i)dt \tag{3.104}$$

where we have used that

$$\nabla\gamma(V_t^i)\cdot a_i(V_t^{(N)}) = \underbrace{\lambda\nabla\gamma(V_t^i)}_{\in T_{V_t^i}\Gamma}\cdot\underbrace{\mathbf{P}_1(V_t^i)\mathbf{V}^\alpha(\rho_t^N)}_{\in T^\perp_{V_t^i}\Gamma} = 0 \tag{3.105}$$

and

$$\nabla\gamma(V_t^i)\cdot b_i(V_t^{(N)})dB_t^i = \sigma|V_t^i - \mathbf{V}^\alpha(\rho_t^N)|\underbrace{\nabla\gamma(V_t^i)}_{T_{V_t^i}\Gamma}\cdot\underbrace{\mathbf{P}_1(V_t^i)dB_t^i}_{T^\perp_{V_t^i}\Gamma} = 0. \tag{3.106}$$

We can further simplify the expression on the rhs of (3.104) by noting

$$\nabla\gamma(V_t^i)\cdot c_i(V_t^{(N)}) = -\frac{\sigma^2}{2}|V_t^i - \mathbf{V}^\alpha(\rho_t^N)|^2\nabla\gamma(V_t^i)\cdot\mathbf{P}_2(V_t^i)\mathbf{P}_3(V_t^i) \tag{3.107}$$

$$= -\frac{\sigma^2}{2}|V_t^i - \mathbf{V}^\alpha(\rho_t^N)|^2\Delta\gamma(V_t^i) \tag{3.108}$$

where we have used that $\mathbf{P}_2 = \Delta\gamma$ and $\nabla\gamma(V_t^i)\cdot\mathbf{P}_3 = |\nabla\gamma(V_t^i)|^2 = 1$.

This finally yields

$$d\gamma(V_t^i) = 0. \tag{3.109}$$

Hence $\gamma(V_t^i) = \gamma(V_0^i) = 0$ for the local solution V_t^i, which ensures that the particles stay on Γ until some finite time. With theorem 2.4.1 we can extend the pathwise unique local strong solution from above to a pathwise unique global solution, which finishes the proof. □

3.8 Summary

In this chapter we introduced the class of cost function that we consider throughout this thesis. The most important assumption is the *coercivity condition* (or *inverse continuity assumption*) which states that the cost function is embedded in a convex envelope. We then introduced the Kuramoto and Vicsek models and derived the Kuramoto-Vicsek CBO model (KV-CBO) on general compact hypersurfaces. The KV-CBO model aims at solving optimization problems constrained (!) on hypersurfaces. We introduced the isotropic as well as the anisotropic model. We collected and proved some stability estimates of the (regularized) consensus point. These were essential for proving the well-posedness (= existence of a pathwise unique strong solution) of the particle system.

Chapter 4

The Mean-Field Limit

The KV-CBO model is a system of coupled non-linear, non-local SDEs which makes its analysis difficult. In this chapter we derive the *mean-field equation* of the KV-CBO model and prove that its solution ρ_t is approximated by the empirical measure ρ_t^N in the large particle limit $N \to \infty$. This *mean-field limit result* justifies the analysis of the KV-CBO model on the macroscopic level. Hence, we move from a *microscopic description* (particle system) to a *macroscopic description* (PDE). This will play a crucial role in the following chapter 5 on optimization. The mean-field equation is a deterministic PDE with initial data given by ρ_0. The fact that the particles are constrained on a compact hypersurface allows for a rigorous derivation of the mean-field equation which is a major advantage over the CBO model studied in [Carrillo et al., 2018] (see remark 3.2) where the mean-field equation is only postulated.

This chapter is based on joint work of Massimo Fornasier, Hui Huang, Lorenzo Pareschi and myself and has been published in [Fornasier et al., 2020], [Fornasier et al., 2021b].

4.1 The coupling method

In this section we introduce the coupling method, see [Sznitman, 1991]. The coupling method is the essential tool needed to prove the main result of this chapter: the *mean-field limit* result of the KV-CBO model, see theorem 4.4.1. The basic idea behind a mean-field limit of a particle system is to derive an equation for the evolution of the probabiliy distribution of the particles and then showing that, in the large particle limit $N \to \infty$, its solution $\rho_t(v) = \rho(t, v)$ is approximated by the empirical measure ρ_t^N of the particles.

Let us make this more precise.

Definition 4.1.1 (Mean-field limit)**.** *Let V_t^i for $i = 1, ..., N$ be the solutions of a given particle system with initial data sampled from $\rho_0^{\otimes N}$. Let further ρ_t denote the solution of the mean-field equation corresponding to the particle system with initial data $\rho(0, v) = \rho_0(v)$. A mean-field limit result is a result of the following form*

$$\rho_0^N \overset{N\to\infty}{\longrightarrow} \rho_0 \Rightarrow \rho_t^N \overset{N\to\infty}{\longrightarrow} \rho_t \tag{4.1}$$

for all $t \in [0, T]$. In words: If, given that the initial data of the particle system approximates the initial data for the mean-field equation, we can prove that the empirical measure of the particle system approximates the solution of the mean-field equation at time t, then we say that a mean-field limit result holds.

The convergence in (4.1) *is the weak convergence of measures, that is, for $\rho_t^N, \rho_t \in \mathcal{P}(\mathbb{R}^d)$ we define*

$$\rho_t^N \overset{N\to\infty}{\longrightarrow} \rho_t :\Leftrightarrow \sup_{t\in[0,T]} \mathbb{E}|\langle \rho_t^N - \rho_t, \phi\rangle|^2 \overset{N\to\infty}{\longrightarrow} 0 \quad \forall \phi \in C_b^1(\mathbb{R}^d) \tag{4.2}$$

where $\langle \cdot, \cdot \rangle$ denotes the dual pairing

$$\begin{cases} \langle \rho_t, \phi\rangle & = \int \phi(v)\ d\rho_t(v), \\ \langle \rho_t^N, \phi\rangle & = \frac{1}{N}\sum_{i=1}^N \phi(V_t^i). \end{cases} \tag{4.3}$$

We note that we need to consider the expected value of $|\langle \rho_t^N - \rho_t, \phi\rangle|^2$ because the empirical measure ρ_t^N is a random measure (as the particles V_t^i are random variables, see section 3.3).

Let us outline how the mean-field limit result (4.1) for the KV-CBO model will be shown. The starting point is the *propagation of chaos assumption*. With this assumption we can derive the *mean-field dynamic*. The mean-field dynamic is an auxiliary monoparticle (!) SDE with distribution-dependant coefficients (also known as McKean process or McKean SDE). We denote its solution with $\overline{V}$. Taking N iid copies of solutions of the mean-field dynamic $\overline{V}^i$ allows to prove the following convergence result

$$\sup_{t\in[0,T]} \mathbb{E}|V_t^i - \overline{V}_t^i|^2 \overset{N\to\infty}{\longrightarrow} 0. \tag{4.4}$$

The proof of this result is not difficult and essentially relies on Itô's formula and Gronwall's inequality (the exponent 2 is chosen to apply Itô's formula to $|\cdot|^2$). The convergence result (4.4) can be shown to be equivalent to the desired mean-field limit result (4.1), (4.2), see below.

Finally, we need to show that the *propagation of chaos assumption* that we started with actually holds. This last step will follow with a lemma from [Sznitman, 1991], see lemma 4.1.1 which we report here for readability. This finishes the outline of this section.

Let us continue with the discussion about the coupling method.

The initial particles in the KV-CBO model are chosen iid (hence *chaotic*) $V_0^i \sim \rho_0$ for a common probability distribution $\rho_0 \in \mathcal{P}(\Gamma)$. At time $t > 0$ this independence is destroyed as the particles are

driven by a coupled (!) system of SDEs. This has the undesirable consequence that we can not longer apply the law of large numbers. The idea of the propagation of chaos assumption is that in the large particle limit $N \to \infty$ any $k \in \mathbb{N}$ fixed particles V_t^i are *approximately independant* for all $t > 0$, hence independence or chaos propagates in time. Let $\mu_t^N \in \mathcal{M}(\mathbb{R}^{dN})$ denote the joint distribution of the particles, then the propagation of chaos assumption is satisfied iff there is a measure $\mu_t \in \mathcal{M}(\mathbb{R}^d)$ such that for any fixed $k > 1$ we have

$$\lim_{N\to\infty} \langle \mu_t^N, \phi_1 \otimes ... \otimes \phi_k \otimes 1 \otimes ... \otimes 1 \rangle = \Pi_{i=1}^k \langle \mu_t, \phi_i \rangle \quad \forall \phi_i \in C_b(\mathbb{R}^d) \tag{4.5}$$

for all $t \in [0, T]$. Here $\phi_1 \otimes ... \otimes \phi_N \in C_b(\mathbb{R}^{dN})$ is defined as $\phi_1 \otimes ... \otimes \phi_N(v_1, ..., v_N) = \phi_1(v_1) \cdot ... \cdot \phi_N(v_N)$ and

$$\langle \mu_t^N, \phi_1 \otimes ... \otimes \phi_N \rangle = \int \phi_1(v_1) \cdot ... \cdot \phi_N(v_N) \; d\mu_t^N(v_1, ..., v_N). \tag{4.6}$$

We note that [Sznitman, 1991, Definition 2.1] used definition (4.5) without time dependence, and refered to a sequence of symmetric measures $\{\mu^N\}_{N\in\mathbb{N}}$ as μ-chaotic iff the above limit holds. We note that the particles are *interchangeable*, that is, for any permutation $\sigma \in S_N$ we have

$$\langle \mu_t^N, \phi_1 \otimes ... \otimes \phi_N \rangle = \langle \mu_t^N, \phi_{\sigma(1)} \otimes ... \otimes \phi_{\sigma(N)} \rangle. \tag{4.7}$$

The joint distribution μ_t^N is then called *symmetric*.

If the particles were iid with common probability distribution μ_t then the joint distribution could be factorized as $\mu_t^N = \mu_t^{\otimes N}$ and

$$\langle \mu_t^N, \phi_1 \otimes ... \otimes \phi_N \rangle = \Pi_{i=1}^N \langle \mu_t, \phi_i \rangle \tag{4.8}$$

would hold for any N. In this sense (4.5) is a generalization of independence of random variables.

Assuming propagation of chaos holds the KV-CBO model decouples into the *mean-field dynamic*

$$d\overline{V}_t = \lambda P_\Gamma(\overline{V}_t) V^\alpha(\rho_t) dt + \sigma |\overline{V}_t - V^\alpha(\rho_t)| P_\Gamma(\overline{V}_t) dB_t + C_\Gamma(\overline{V}_t, V^\alpha(\rho_t)) dt \tag{4.9}$$

where the coefficients are distribution dependent with $\rho_t = law(\overline{V}_t)$. This equation is a monoparticle SDE and does not involve the empirical measure ρ_t^N anymore.

Let us now show the equivalence of (4.2) and (4.4). For all $\phi \in C_b^1(\mathbb{R}^d)$ we have

$$\mathbb{E}|\langle \rho_t^N - \rho_t, \phi \rangle|^2 = \mathbb{E}\left| \frac{1}{N} \sum_{i=1}^N \phi(V_t^i) - \int \phi(v) d\rho_t(v) \right|^2 \tag{4.10}$$

$$\leq 2\mathbb{E}\left| \frac{1}{N} \sum_{i=1}^N \phi(V_t^i) - \frac{1}{N} \sum_{i=1}^N \phi(\overline{V}_t^i) \right|^2 + 2\mathbb{E}\left| \frac{1}{N} \sum_{i=1}^N \phi(\overline{V}_t^i) - \int \phi(v) d\rho_t(v) \right|^2 \tag{4.11}$$

$$= 2\mathbb{E}\big|\phi(V_t^1) - \phi(\overline{V}_t^1)\big|^2 + 2\mathbb{E}\left|\frac{1}{N}\sum_{i=1}^{N}\phi(\overline{V}_t^i) - \int \phi(v)d\rho_t(v)\right|^2 \tag{4.12}$$

where we have used the symmetry of the particles in the third step. We conclude with the mean-value theorem and the law of large numbers.

Once we have proved the mean-field limit result (4.4) we are left to verify that the propagation of chaos assumption that we started with actually holds. This gap is closed with [Sznitman, 1991, Proposition 2.2 i)] which we report below for simplicity.

Lemma 4.1.1. *Let $\mu \in \mathcal{C}([0,T],\mathcal{P}(\mathbb{R}^d))$. The sequence $\{\mu_t^N\}_{N\in\mathbb{N}}$ of symmetric probability measures satisfies the propagation of chaos assumption* (4.5) *iff the empirical measure ρ^N converges in distribution to the constant random variable μ.*

Proof. Let us first assume that $\{\mu_t^N\}_{N\in\mathbb{N}}$ satisfies the propagation of chaos assumption (for $k=2$). For $\phi \in C_b(\mathbb{R}^d)$ we have

$$\mathbb{E}|\langle \rho_t^N - \mu_t, \phi\rangle|^2 = \mathbb{E}|\langle \rho_t^N, \phi\rangle|^2 - 2\mathbb{E}\langle \rho_t^N, \phi\rangle\langle \mu_t, \phi\rangle + |\langle \mu_t, \phi\rangle|^2 \tag{4.13}$$

$$= \frac{1}{N^2}\sum_{i,j=1}^{N}\mathbb{E}(\phi(V_t^i)\phi(V_t^j)) - \frac{2}{N}\langle \mu_t, \phi\rangle\sum_{i=1}^{N}\mathbb{E}(\phi(V_t^i)) + |\langle \mu_t, \phi\rangle|^2. \tag{4.14}$$

The symmetry of the probability measure μ_t^N yields

$$\mathbb{E}(\phi(V_t^i)) = \mathbb{E}(\phi(V_t^1)), \quad \mathbb{E}(\phi(V_t^i)\phi(V_t^j)) = \mathbb{E}(\phi(V_t^1)\phi(V_t^2)) \tag{4.15}$$

from which we get

$$\begin{aligned}\mathbb{E}|\langle \rho_t^N - \mu_t, \phi\rangle|^2 = \frac{1}{N}\mathbb{E}(\phi(V_t^1)^2) + \frac{N-1}{N}\mathbb{E}(\phi(V_t^1)\phi(V_t^2)) \\ - 2\langle \mu_t, \phi\rangle\mathbb{E}(\phi(V_t^1)) + |\langle \mu_t, \phi\rangle|^2.\end{aligned} \tag{4.16}$$

With the propagation of chaos assumption we further find

$$\lim_{N\to\infty}\mathbb{E}(\phi(V_t^1)) = \lim_{N\to\infty}\langle \mu_t^N, \phi\otimes 1\otimes ...\otimes 1\rangle = \langle \mu_t, \phi\rangle \tag{4.17}$$

$$\lim_{N\to\infty}\mathbb{E}(\phi(V_t^1)\phi(V_t^2)) = \lim_{N\to\infty}\langle \mu_t^N, \phi\otimes\phi\otimes 1\otimes ...\otimes 1\rangle = \langle \mu_t, \phi\rangle^2. \tag{4.18}$$

Hence

$$\lim_{N\to\infty}\mathbb{E}|\langle \rho_t^N - \mu_t, \phi\rangle|^2 = 0 \tag{4.19}$$

which finishes the proof of this implication.

Let us now show the reverse. We want to show that

$$\begin{aligned}
&|\langle \mu_t^N, \phi_1 \otimes ... \otimes \phi_k \otimes 1 \otimes ... \otimes 1\rangle - \Pi_{i=1}^k \langle \mu_t, \phi_i\rangle| \\
&\quad \leq |\langle \mu_t^N, \phi_1 \otimes ... \otimes \phi_k \otimes 1 \otimes ... \otimes 1\rangle - \langle \mu_t^N, \Pi_{i=1}^k \langle \rho_t^N, \phi_i\rangle\rangle| \\
&\quad\quad + |\langle \mu_t^N, \Pi_{i=1}^k \langle \rho_t^N, \phi_i\rangle\rangle - \Pi_{i=1}^k \langle \rho_t, \phi_i\rangle|
\end{aligned} \tag{4.20}$$

approaches zero as $N \to \infty$. For the second term on the right hand side this follows from the assumption that ρ_t^N converges in distribution to μ_t. In that case we have

$$\langle \mu_t^N, \Pi_{i=1}^k \langle \rho_t^N, \phi_i\rangle\rangle - \Pi_{i=1}^k \langle \rho_t, \phi_i\rangle = \Pi_{i=1}^k \mathbb{E}\langle \rho_t^N - \mu_t, \phi_i\rangle \to 0 \tag{4.21}$$

for all $\phi_i \in C_b(\mathbb{R}^d)$. With symmetry we can rewrite the first term on the rhs as

$$|\langle \mu_t^N, \phi_1 \otimes ... \otimes \phi_k \otimes 1 \otimes ... \otimes 1\rangle - \langle \mu_t^N, \Pi_{i=1}^k \langle \rho_t^N, \phi_i\rangle\rangle| \tag{4.22}$$

$$= \big|\langle \mu_t^N, \frac{1}{N!} \sum_{\sigma \in S_N} \phi_1(V_t^{\sigma(1)}) \cdot ... \cdot \phi_k(V_t^{\sigma(k)}) - \Pi_{i=1}^k \langle \rho_t^N, \phi_i\rangle\rangle\big| \tag{4.23}$$

$$\leq \sup_{V_t^i \in \mathbb{R}^d} \big|\frac{1}{N!} \sum_{\sigma \in S_N} \phi_1(V_t^{\sigma(1)}) \cdot ... \cdot \phi_k(V_t^{\sigma(k)}) - \Pi_{i=1}^k \langle \rho_t^N, \phi_i\rangle\big|. \tag{4.24}$$

We can rewrite the second term in the parantheses as

$$\Pi_{i=1}^k \langle \rho_t^N, \phi_i\rangle = \frac{1}{N^k}\big(\sum_{j=1}^N \Pi_{i=1}^k \phi_i(V_t^j) + \sum_{j \in S_{k,N}} \Pi_{i=1}^k \phi_i(V_t^j)\big) \tag{4.25}$$

where $S_{k,N}$ is the set of k-tupels with elements from $\{1, ..., N\}$.

Now (4.24) can further be bounded by

$$\sup_{V_t^i \in \mathbb{R}^d} \big|\frac{1}{N!} \sum_{\sigma \in S_N} \phi_1(V_t^{\sigma(1)}) \cdot ... \cdot \phi_k(V_t^{\sigma(k)}) - \frac{1}{N^k} \sum_{j=1}^N \Pi_{i=1}^k \phi_i(V_t^j)\big| + \big|\frac{1}{N^k} \sum_{j \in S_{k,N}} \Pi_{i=1}^k \phi_i(V_t^j)\big| \tag{4.26}$$

$$\leq \sup_{V_t^i \in \mathbb{R}^d} \big| \sum_{\sigma \in S_N} \phi_1(V_t^{\sigma(1)}) \cdot ... \cdot \phi_k(V_t^{\sigma(k)}) \bigg(\frac{\phi_{k+1}(V_t^{\sigma(k+1)}) \cdot ... \cdot \phi_N(V_t^{\sigma(N)})}{N!} - \frac{1}{N^k} \bigg)\big| \tag{4.27}$$

$$+ \big|\frac{1}{N^k} \sum_{j \in S_{k,N}} \Pi_{i=1}^k \phi_i(V_t^j)\big| \tag{4.28}$$

$$\leq M^k \bigg((\frac{(N-k)!}{N!} - \frac{1}{N^k}) \frac{N!}{(N-k)!} + \frac{1}{N^k}(N^k - \frac{N!}{(N-k)!}) \bigg) \tag{4.29}$$

$$= 2M^k (1 - \frac{N!}{N^k (N-k)!}) \to 0 \tag{4.30}$$

where we assumed $\|\phi_i\|_\infty \leq M$. □

4.2 Well-posedness of the mean-field dynamic

In this section we prove the well-posedness of the mean-field dynamic. Following the same argument as in the proof of theorem 3.7.1 we find $d\gamma(\overline{V}_t) = 0$, that is, $\overline{V}_t$ stays on the hypersurface Γ for all $t > 0$. This introduces a measure $\rho_t \in \mathcal{P}(\Gamma)$ such that $\rho_t = law(\overline{V}_t)$. We note however that the mean-field dynamic is defined in the ambient space $\mathbb{R}^d$.

Theorem 4.2.1. *The mean-field dynamic*

$$d\overline{V}_t = \lambda P_\Gamma(\overline{V}_t)V^\alpha(\rho_t)dt + \sigma|\overline{V}_t - V^\alpha(\rho_t)|P_\Gamma(\overline{V}_t)dB_t + C_\Gamma(\overline{V}_t, V^\alpha(\rho_t))dt \tag{4.31}$$

is well-posed, that is, there is a pathwise unique strong solution for any initial probability distribution $\rho_0 \in \mathcal{P}(\Gamma)$.

The proof relies on the Leray-Schauder fixed point theorem [Gilbarg and Trudinger, 2015, Chapter 10].

Theorem 4.2.2 (Leray-Schauder)**.** *Let B a Banach space and $T : B \times [0,1] \to B$ a mapping satisfying the following properties*

1. *T is compact,*
2. *$T(\xi, 0) = 0$ for all $\xi \in B$,*
3. *$\|\xi\|_B < \infty$ for all $(\xi, \vartheta) \in B \times [0,1]$ with $T(\xi, \vartheta) = \xi$.*

Then the mapping T_1 defined as $T_1(\xi) = T(\xi, 1)$ has a fixed point $T_1(\xi^\star) = \xi^\star$.

Proof. Let $\xi \in C([0,T], \mathbb{R}^d)$ and consider the SDE

$$d\overline{V}_t^\xi = \lambda \mathbf{P}_1(\overline{V}_t^\xi)\xi_t dt + \sigma|\overline{V}_t^\xi - \xi_t|\mathbf{P}_1(\overline{V}_t^\xi)dB_t - \frac{\sigma^2}{2}|\overline{V}_t^\xi - \xi_t|^2\mathbf{P}_2(\overline{V}_t^\xi)\mathbf{P}_3(\overline{V}_t^\xi)dt. \tag{4.32}$$

with initial data $\overline{V}_0^\xi \sim \rho_0$ for $\rho_0 \in \mathcal{P}(\Gamma)$ (the superscript indicates that the solution depends on ξ). The well-posedness of (4.32) follows from the fact that it has locally Lipschitz continuous coefficients (see lemma 3.7.1) and ξ does not depend on $\overline{V}^\xi$. By the same argument as in theorem 3.7.1 it can be shown that $\overline{V}_t^\xi \in \Gamma$ for all $t > 0$. The probability distribution of $\overline{V}^\xi$ thus introduces a random measure $\rho^\xi \in C([0,T], \mathcal{P}(\Gamma))$ with

$$\rho_t^\xi = law(\overline{V}_t^\xi). \tag{4.33}$$

This allows us to use the operators P_Γ, $\nabla\gamma$ and $\Delta\gamma$ instead of the regularized operators $\mathbf{P}_1$, $\mathbf{P}_2$ and $\mathbf{P}_3$.

In the following we prove the existence of a unique (!) $\xi^\star \in C([0,T], \mathbb{R}^d)$ such that $\xi^\star = V^\alpha(\rho^{\xi^\star})$. In this case (4.32) is exactly the mean-field dynamic (4.9). We will use the Leray-Schauder theorem to prove the existence of a fixed point of

$$T : C([0,T], \mathbb{R}^d) \times [0,1] \to C([0,T], \mathbb{R}^d),\ (\xi, \vartheta) \mapsto \vartheta V^\alpha(\rho^\xi). \tag{4.34}$$

The uniqueness does not follow with the Leray-Schauder theorem and has to be shown differently. Once we have shown the existence of a unique fixed point the well-posedness of the mean-field dynamic follows from the well-posedness of (4.32).

Compactness of T: In Lemma 3.6.4 we proved the Hölder continuity of $t \mapsto V^\alpha(\rho_t^\xi)$ with coefficient $1/2$, that is,

$$|V^\alpha(\rho_t^\xi) - V^\alpha(\rho_s^\xi)| \lesssim |t-s|^{1/2}. \tag{4.35}$$

Thus $T(\xi,\vartheta) = \vartheta V^\alpha(\rho^\xi) \in C^{0,1/2}([0,T],\mathbb{R}^d)$ for all $\xi \in C([0,T],\mathbb{R}^d)$ and $\vartheta \in [0,1]$. The compactness of T now follows from the fact that $C^{0,1/2}([0,T],\mathbb{R}^d)$ is compactly embedded in $C([0,T],\mathbb{R}^d)$, see [Alt, 2016, 10.6].

Existence of fixed point of T: We now show the two remaining properties from the Leray-Schauder theorem (theorem 4.2.2). The property $T(\xi,0) = 0$ is obviously fulfilled. For the other property we chose (ξ,ϑ) such that $T(\xi,\vartheta) = \xi$ to get

$$\|\xi\|_{C([0,T],\mathbb{R}^d)} = \sup_{t\in[0,T]} |\xi_t| = \sup_{t\in[0,T]} |\vartheta V^\alpha(\rho_t^\xi)| \leq \sup_{t\in[0,T]} \vartheta C_\alpha \int_\Gamma |v| d\rho_t^\xi(v) < \infty \tag{4.36}$$

where we have used the compactness of Γ (from which it follows that the continuous function $|\cdot|$ admits its maximum on Γ). By the Leray-Schauder theorem there exists $\xi^\star \in C([0,T],\mathbb{R}^d)$ such that $V^\alpha(\rho^{\xi^\star}) = \xi^\star$.

Uniqueness of the fixed point $\xi^\star$: Suppose there are two distinct fixed points of T which we denote as $\xi^{\star,1}, \xi^{\star,2}$. The corresponding processes and densities are denoted as $\overline{V}_t^{\xi^{\star,i}}$ and $\rho_t^{\xi^{\star,i}}$ respectively. We further denote

$$\overline{Z}_t = \overline{V}_t^{\xi^{\star,1}} - \overline{V}_t^{\xi^{\star,2}}. \tag{4.37}$$

With the integral represntation of (4.32) we get

$$\begin{aligned}\overline{Z}_t - \overline{Z}_0 = \lambda \int_0^t P_\Gamma(\overline{Z}_s)(\xi_s^{\star,1} - \xi_s^{\star,2})ds + \sigma \int_0^t |\overline{Z}_s - (\xi_s^{\star,1} - \xi_s^{\star,2})| P_\Gamma(\overline{Z}_s) dB_s \\ - \frac{\sigma^2}{2} \int_0^t |\overline{Z}_s - (\xi_s^{\star,1} - \xi_s^{\star,2})|^2 \big(\Delta\gamma(\overline{V}_t^{\xi^{\star,1}})\nabla\gamma(\overline{V}_t^{\xi^{\star,1}}) - \Delta\gamma(\overline{V}_t^{\xi^{\star,2}})\nabla\gamma(\overline{V}_t^{\xi^{\star,2}})\big) ds\end{aligned} \tag{4.38}$$

from which we further conclude

$$\begin{aligned}\mathbb{E}|\overline{Z}_t|^2 \lesssim \mathbb{E}|\overline{Z}_0|^2 + \mathbb{E}\left|\int_0^t \lambda P_\Gamma(\overline{Z}_s)(\xi_s^{\star,1} - \xi_s^{\star,2})ds\right|^2 + \mathbb{E}\left|\sigma \int_0^t |\overline{Z}_s - (\xi_s^{\star,1} - \xi_s^{\star,2})| P_\Gamma(\overline{Z}_s) dB_s\right|^2 \\ + \mathbb{E}\left|\frac{\sigma^2}{2} \int_0^t |\overline{Z}_s - (\xi_s^{\star,1} - \xi_s^{\star,2})|^2 \big(\Delta\gamma(\overline{V}_t^{\xi^{\star,1}})\nabla\gamma(\overline{V}_t^{\xi^{\star,1}}) - \Delta\gamma(\overline{V}_t^{\xi^{\star,2}})\nabla\gamma(\overline{V}_t^{\xi^{\star,2}})\big) ds\right|^2\end{aligned}$$

$$\lesssim \mathbb{E}|\overline{Z}_0|^2 + \mathbb{E}\bigg(\int_0^t |\xi_s^{\star,1} - \xi_s^{\star,2}|ds\bigg)^2 + \mathbb{E}\bigg(\int_0^t |\overline{Z}_s - (\xi_s^{\star,1} - \xi_s^{\star,2})|^2 ds\bigg) + \mathbb{E}\bigg(\int_0^t |\overline{Z}_s - (\xi_s^{\star,1} - \xi_s^{\star,2})|^2 ds\bigg)^2$$

where we have used *Itô's isometry* (see appendix), the mean-value theorem to bound $|\Delta\gamma(\overline{V}_t^{\xi^{\star,1}})\nabla\gamma(\overline{V}_t^{\xi^{\star,1}}) - \Delta\gamma(\overline{V}_t^{\xi^{\star,2}})\nabla\gamma(\overline{V}_t^{\xi^{\star,2}})|$ and the fact that $\|P_\Gamma\|_\infty, \|\nabla\gamma\|_\infty, \|\Delta\gamma\|_\infty$ are bounded (note that γ is smooth, see section 3.4).

We further have

$$\mathbb{E}\bigg(\int_0^t |\overline{Z}_s - (\xi_s^{\star,1} - \xi_s^{\star,2})|^2 ds\bigg)^2 \leq C(t)\int_0^t \mathbb{E}|\overline{Z}_s - (\xi_s^{\star,1} - \xi_s^{\star,2})|^2 ds \tag{4.39}$$

$$\lesssim \int_0^t \mathbb{E}|\overline{Z}_s|^2 ds + \int_0^t |\xi_s^{\star,1} - \xi_s^{\star,2}|^2 ds \tag{4.40}$$

and according to (3.80) the following estimate holds

$$|\xi_s^{\star,1} - \xi_s^{\star,2}|^2 = |V^\alpha(\rho_s^{\xi^{\star,1}}) - V^\alpha(\rho_s^{\xi^{\star,2}})|^2 \lesssim W_2^2(\rho_s^{\xi^{\star,1}}, \rho_s^{\xi^{\star,2}}) \lesssim \mathbb{E}|\overline{Z}_s|^2. \tag{4.41}$$

Combining the estimates from above yields

$$\mathbb{E}|\overline{Z}_t|^2 \lesssim \mathbb{E}|\overline{Z}_0|^2 + \int_0^t \mathbb{E}|\overline{Z}_s|^2 ds. \tag{4.42}$$

Since $\mathbb{E}|\overline{Z}_0|^2 = 0$ we conclude with Gronwall's inequality that $\mathbb{E}|\overline{Z}_t|^2 = 0$ for all $t > 0$. Hence $\overline{Z}_t = \overline{V}_t^{\xi^{\star,1}} - \overline{V}_t^{\xi^{\star,2}} = 0$. With the well-posedness of (4.32) we conclude that the fixed point $\xi^\star$ is unique. This completes the proof. □

4.3 The mean-field equation

In this section we will derive the mean-field equation for the KV-CBO model and prove its well-posedness. The mean-field equation is a non-linear, non-local deterministic PDE. We apply Dynkin's formula with a test function $\phi \in C_c^\infty(\mathbb{R}^d)$ to the mean-field dynamic. This yields

$$\frac{d}{dt}\int_{\mathbb{R}^d} \phi(v)\, d\rho_t(v) = \int_{\mathbb{R}^d} \nabla\phi(v)\cdot\lambda P_\Gamma(v)V^\alpha(\rho_t) + \frac{\sigma^2}{2}|v - V^\alpha(\rho_t)|^2\Delta\phi(v)\, d\rho_t(v) \tag{4.43}$$

where we have used $\nabla\phi(v)\cdot n(v) = 0$, $\mathrm{Tr}(\nabla^2\phi(v)^T n(v)n(v)^T) = 0$, see (3.40), (3.41). With the definitions of the differential operators ∇_Γ, Δ_Γ, and by using test functions $\Phi \in C^\infty(\Gamma)$ (see (3.39))

we arrive at the weak formulation of the mean-field equation

$$\frac{d}{dt}\int_\Gamma \Phi(v)\ d\rho_t(v) = \int_\Gamma \nabla_\Gamma \Phi(v)\cdot \lambda P_\Gamma(v)V^\alpha(\rho_t) + \frac{\sigma^2}{2}|v-V^\alpha(\rho_t)|^2\Delta_\Gamma\Phi(v)\ d\rho_t(v). \tag{4.44}$$

Here, we have used the fact that $\rho_t \in \mathcal{P}(\Gamma)$ (we have denoted the restriction of ρ_t to Γ again by ρ_t).

The strong formulation of the mean-field equation on Γ is given by (4.45) below. We remark that the derivation of the strong formulation follows with a simple application of the integration by parts formula and using $d\rho_t(v) = \rho_t(v)dv$. For the KV-CBO with anisotropic noise the derivation of the strong formulation is much more complicated, see appendix on page 141.

In the following theorem we prove the well-posedness of the mean-field equation from which the propagation of chaos assumption follows, see lemma 4.1.1.

Theorem 4.3.1. *The mean-field equation*

$$\partial_t\rho_t = -\lambda\nabla_\Gamma\cdot(P_\Gamma(v)V^\alpha(\rho_t)\rho_t) + \frac{\sigma^2}{2}\Delta_\Gamma(|v-V^\alpha(\rho_t)|^2\rho_t) \tag{4.45}$$

is well-posed, that is, there is a unique weak solution for any initial probability distribution $\rho_0 \in \mathcal{P}(\Gamma)$.

Proof. Existence: The mean-field dynamic has pathwise unique strong solutions, see theorem 4.2.1. Hence there exists a strong solution $\overline{V}_t$ for the initial data $\overline{V}_0 \sim \rho_0$. Further $\overline{V}_t$ induces a weak solution $\rho_t = law(\overline{V}_t)$ of (4.43) and (4.44) which finishes the existence proof.

Uniqueness: Let ρ_t^1 and ρ_t^2 two weak solutions of (4.45) with the same initial data ρ_0. We interpret these weak solutions as measures in $\mathbb{R}^d$ concentrated on Γ, that is, $\rho_t^i \in \mathcal{M}(\mathbb{R}^d)$ with $\mathrm{supp}(\rho_t^i) \subset \Gamma$.

We construct two linear (!) stochastic processes $\overline{V}_t^i$ corresponding to ρ_t^i with common initial data $\overline{V}_0 \sim \rho_0$ as follows

$$\begin{aligned} d\overline{V}_t^i = \lambda\mathbf{P}_1(\overline{V}_t^i)V^\alpha(\rho_t^i)dt + \sigma|\overline{V}_t^i - V^\alpha(\rho_t^i)|\mathbf{P}_1(\overline{V}_t^i)dB_t \\ - \frac{\sigma^2}{2}|\overline{V}_t^i - V^\alpha(\rho_t^i)|^2\mathbf{P}_2(\overline{V}_t^i)\mathbf{P}_3(\overline{V}_t^i). \end{aligned} \tag{4.46}$$

We denote the law of these stochastic processes with μ_t^i, that is, $law(\overline{V}_t^i) = \mu_t^i$. Here $\mathbf{P}_i$ are the regularized operators from section 3.7. We do not regularize the consensus point as it depends on the measure $\rho_t^i \in \mathcal{P}(\Gamma)$ which is known to stay on Γ (for μ_t^i this is not yet known). We stress that the above SDE is linear because the consensus point depends on ρ_t^i (not on $\mu_t^i = law(\overline{V}_t^i)$).

On the one hand the measures μ_t^i are unique weak solutions of the corresponding linear Fokker-Planck

equation

$$\frac{d}{dt}\int_{\mathbb{R}^d}\phi(v)\ d\mu_t^i(v) = \int_{\mathbb{R}^d}\nabla\phi(v)\cdot\big(\lambda\mathbf{P}_1(v)V^\alpha(\rho_t^i) - \frac{\sigma^2}{2}|v-V^\alpha(\rho_t^i)|^2\mathbf{P}_2(v)\mathbf{P}_3(v) + \frac{\sigma^2}{2}|v-V^\alpha(\rho_t^i)|^2\nabla^2\phi(v):\mathbf{P}_1(v)\big)\ d\mu_t^i(v). \tag{4.47}$$

On the other hand the measures ρ_t^i are assumed to be weak solutions of the mean-field equation (4.45), hence they are solutions of (compare with (4.43))

$$\frac{d}{dt}\int_{\mathbb{R}^d}\phi(v)\ d\rho_t^i(v) = \int_{\mathbb{R}^d}\nabla\phi(v)\cdot\big(\lambda P_\Gamma(v)V^\alpha(\rho_t^i) - \frac{\sigma^2}{2}|v-V^\alpha(\rho_t^i)|^2\Delta\gamma(v)\nabla\gamma(v) + \frac{\sigma^2}{2}|v-V^\alpha(\rho_t^i)|^2\nabla^2\phi(v):P_\Gamma(v)\big)\ d\rho_t^i(v). \tag{4.48}$$

Hence we conclude $\mu_t^i = \rho_t^i$ for $i = 1, 2$. Hence, also μ_t^i are concentrated on Γ, and we see that (4.46) is equal to the mean-field dynamic for which we have already shown the existence of pathwise unique strong solutions. Thus, $\overline{V}_t^1 = \overline{V}_t^2$ a.s and consequently $\mu_t^1 = \mu_t^2$. Thus $\rho_t^1 = \mu_t^1 = \mu_t^2 = \rho_t^2$, which finishes the proof. □

4.4 Mean-field limit

In this section we will prove the main theorem of this chapter: the mean-field limit result for the KV-CBO model. We will see that this result holds with convergence rate N^{-1}, that is, the approximation error between the empirical measure and the solution of the mean-field equation increases with rate N^{-1} as the number of particles grows. We note that this convergence rate is independent of the dimension d and we therefore do not observe a *curse of dimensionality phenomenon*, see [Fournier and Guillin, 2015].

We stress that the compactness of Γ is crucial for the rigorous proof of the mean-field limit result. This rigorous proof is a major asset of the KV-CBO model over other CBO models in Euclidean space.

Let us first prove a stability estimate on the consensus point V^α.

Lemma 4.4.1. *Let $\rho \in C([0,T],\mathcal{P}(\Gamma))$ and $\overline{V}_t^i \sim \rho_t$, $i \in [N]$ iid copies of solutions of the mean-field dynamic* (4.9)*. Then*

$$\sup_{t\in[0,T]}\mathbb{E}|V^\alpha(\overline{\rho}_t^N) - V^\alpha(\rho_t)|^2 \lesssim N^{-1} \tag{4.49}$$

where $\overline{\rho}_t^N$ is the empirical measure of $\overline{V}_t^i$.

Proof. We start by estimating

$$|V^\alpha(\overline{\rho}_t^N) - V^\alpha(\rho_t)| = \left|\frac{\sum_{i=1}^N \overline{V}_t^i e^{-\alpha\mathcal{E}(\overline{V}_t^i)}}{\sum_{i=1}^N e^{-\alpha\mathcal{E}(\overline{V}_t^i)}} - \frac{\int_{\mathbb{R}^d} v e^{-\alpha\mathcal{E}(v)}d\rho_t(v)}{\int_{\mathbb{R}^d} e^{-\alpha\mathcal{E}(v)}d\rho_t(v)}\right| \tag{4.50}$$

$$\leq \left| \frac{\frac{1}{N}\sum_{i=1}^N \overline{V}_t^i e^{-\alpha\mathcal{E}(\overline{V}_t^i)} - \int_{\mathbb{R}^d} v e^{-\alpha\mathcal{E}(v)} d\rho_t(v)}{\frac{1}{N}\sum_{i=1}^N e^{-\alpha\mathcal{E}(\overline{V}_t^i)}} \right| + \left| \frac{\int_{\mathbb{R}^d} v e^{-\alpha\mathcal{E}(v)} d\rho_t(v)}{\frac{1}{N}\sum_{i=1}^N e^{-\alpha\mathcal{E}(\overline{V}_t^i)}} - \frac{\int_{\mathbb{R}^d} v e^{-\alpha\mathcal{E}(v)} d\rho_t(v)}{\int_{\mathbb{R}^d} e^{-\alpha\mathcal{E}(v)} d\rho_t(v)} \right|.$$

We have

$$\frac{1}{N}\sum_{i=1}^N e^{-\alpha\mathcal{E}(\overline{V}_t^i)}, \ \int_{\mathbb{R}^d} e^{-\alpha\mathcal{E}(v)} d\rho_t(v) \geq e^{-\alpha\overline{\mathcal{E}}} \tag{4.51}$$

and

$$\int_{\mathbb{R}^d} v e^{-\alpha\mathcal{E}(v)} d\rho_t(v) \lesssim e^{-\alpha\underline{\mathcal{E}}} \tag{4.52}$$

where we have used the compactness of Γ. With $(x+y)^2 \leq 2(x^2+y^2)$ for $x, y \geq 0$ we further estimate

$$\begin{aligned} |V^\alpha(\overline{\rho}_t^N) - V^\alpha(\rho_t)|^2 \leq{}& 2e^{2\alpha\overline{\mathcal{E}}} \left| \frac{1}{N}\sum_{i=1}^N \overline{V}_t^i e^{-\alpha\mathcal{E}(\overline{V}_t^i)} - \int_{\mathbb{R}^d} v e^{-\alpha\mathcal{E}(v)} d\rho_t(v) \right|^2 \\ &+ 2e^{2\alpha(2\overline{\mathcal{E}}-\underline{\mathcal{E}})} \left| \frac{1}{N}\sum_{i=1}^N e^{-\alpha\mathcal{E}(\overline{V}_t^i)} - \int_{\mathbb{R}^d} e^{-\alpha\mathcal{E}(v)} d\rho_t(v) \right|^2. \end{aligned} \tag{4.53}$$

To estimate the expected value of the first term on the right hand side we introduce

$$\overline{Z}_t^i = \overline{V}_t^i e^{-\alpha\mathcal{E}(\overline{V}_t^i)} - \int_{\mathbb{R}^d} v e^{-\alpha\mathcal{E}(v)} d\rho_t(v) \tag{4.54}$$

and note $\mathbb{E}(\overline{Z}_t^j) = 0$ (because $\overline{V}_t^i \sim \rho_t$) and $\mathbb{E}(\overline{Z}_t^i \overline{Z}_t^k) = 0$ for any $i \neq k$ (because $\overline{V}_t^i$ are iid, hence independent). Now we can rewrite

$$\begin{aligned} &\mathbb{E}\left| \frac{1}{N}\sum_{i=1}^N \overline{V}_t^i e^{-\alpha\mathcal{E}(\overline{V}_t^i)} - \int_{\mathbb{R}^d} v e^{-\alpha\mathcal{E}(v)} d\rho_t(v) \right|^2 \\ &\quad = \mathbb{E}\left| \frac{1}{N}\sum_{i=1}^N \overline{Z}_t^i \right|^2 = \frac{1}{N^2}\mathbb{E}\left| \sum_{i,k=1}^N \overline{Z}_t^i \overline{Z}_t^k \right| = \frac{1}{N^2}\mathbb{E}\left| \sum_{i=1}^N |\overline{Z}_t^i|^2 \right| = \frac{1}{N}\mathbb{E}|\overline{Z}_t^1|^2 \end{aligned} \tag{4.55}$$

which we can bound with

$$\mathbb{E}|\overline{Z}_t^1|^2 \leq 2\mathbb{E}|\overline{V}_t^1 e^{-\alpha\mathcal{E}(\overline{V}_t^1)}|^2 + 2\left| \int_{\mathbb{R}^d} v e^{-\alpha\mathcal{E}(v)} d\rho_t(v) \right|^2 \leq 4C_\Gamma e^{-2\alpha\underline{\mathcal{E}}} \tag{4.56}$$

where we have used that $\overline{V}_t^1 \in \Gamma$ and $|v| \leq C_\Gamma$ for all $v \in \Gamma$ and some $C_\Gamma > 0$. Similarly we find

$$\mathbb{E}\left| \frac{1}{N}\sum_{i=1}^N e^{-\alpha\mathcal{E}(\overline{V}_t^i)} - \int_{\mathbb{R}^d} e^{-\alpha\mathcal{E}(v)} d\rho_t(v) \right|^2 \leq \frac{4e^{-2\alpha\underline{\mathcal{E}}}}{N}. \tag{4.57}$$

We arrive at

$$\mathbb{E}|V^\alpha(\overline{\rho}_t^N) - V^\alpha(\rho_t)|^2 \lesssim N^{-1} \tag{4.58}$$

which completes the proof. □

Theorem 4.4.1 (Mean-field limit)**.** *Let V_t^i denote the KV-CBO particles and $\overline{V}_t^i$ iid copies of solutions of the mean-field dynamic* (4.9) *with common probability distribution $\rho_t = law(\overline{V}_t^i)$. Assume further that the initial data and the Brownian motions are the same, that is, $V_0^i = \overline{V}_0^i$ and $B_t^i = \overline{B}_t^i$. Then*

$$\sup_{t\in[0,T]} \mathbb{E}|V_t^i - \overline{V}_t^i|^2 \lesssim N^{-1} \tag{4.59}$$

for all $i \in [N]$. We note that the constant in $\lesssim$ depends on the dimension d.

Proof. The particles V_t^i and $\overline{V}_t^i$ are solutions to the SDEs

$$dV_t^i = \lambda P_\Gamma(V_t^i)V^\alpha(\rho_t^N)dt + \sigma|V_t^i - V^\alpha(\rho_t^N)|P_\Gamma(V_t^i)dB_t^i + C_\Gamma(V_t^i, V^\alpha(\rho_t^N))dt \tag{4.60}$$

and

$$d\overline{V}_t^i = \lambda P_\Gamma(\overline{V}_t^i)V^\alpha(\rho_t)dt + \sigma|\overline{V}_t^i - V^\alpha(\rho_t)|P_\Gamma(\overline{V}_t^i)d\overline{B}_t^i + C_\Gamma(\overline{V}_t^i, V^\alpha(\rho_t))dt \tag{4.61}$$

where C_Γ is the usual correction term $C_\Gamma(V,W) = -\frac{\sigma^2}{2}|V-W|^2\Delta\gamma(V)\nabla\gamma(V)$ as defined in (3.84). The fact that the driving Brownian motions are the same yields the following SDE for the difference of of the corresponding particles

$$\begin{aligned} d(V_t^i - \overline{V}_t^i) &= \lambda\Big(P_\Gamma(V_t^i)V^\alpha(\rho_t^N) - P_\Gamma(\overline{V}_t^i)V^\alpha(\rho_t)\Big)dt \\ &\quad + \sigma\Big(|V_t^i - V^\alpha(\rho_t^N)|P_\Gamma(V_t^i) - |\overline{V}_t^i - V^\alpha(\rho_t)|P_\Gamma(\overline{V}_t^i)\Big)dB_t^i \\ &\quad + \Big(C_\Gamma(V_t^i, V^\alpha(\rho_t^N)) - C_\Gamma(\overline{V}_t^i, V^\alpha(\rho_t))\Big)dt \end{aligned} \tag{4.62}$$

An application of Itô's formula with $f(x) = |x|^2$, $\nabla f(x) = 2x$, and $\Delta f(x) = 2I_d$ yields

$$\begin{aligned} d|V_t^i - \overline{V}_t^i|^2 &= 2\lambda(V_t^i - \overline{V}_t^i)\cdot\Big(P_\Gamma(V_t^i)V^\alpha(\rho_t^N) - P_\Gamma(\overline{V}_t^i)V^\alpha(\rho_t)\Big)dt \\ &\quad + 2(V_t^i - \overline{V}_t^i)\cdot\Big(C_\Gamma(V_t^i, V^\alpha(\rho_t^N)) - C_\Gamma(\overline{V}_t^i, V^\alpha(\rho_t))\Big)dt + \sigma^2\mathrm{Tr}(\sigma^i\cdot\sigma^i)dt \\ &\quad + 2\sigma(V_t^i - \overline{V}_t^i)\cdot\Big(|V_t^i - V^\alpha(\rho_t^N)|P_\Gamma(V_t^i) - |\overline{V}_t^i - V^\alpha(\rho_t)|P_\Gamma(\overline{V}_t^i)\Big)dB_t^i \end{aligned} \tag{4.63}$$

where

$$\sigma^i = |V_t^i - V^\alpha(\rho_t^N)|P_\Gamma(V_t^i) - |\overline{V}_t^i - V^\alpha(\rho_t)|P_\Gamma(\overline{V}_t^i). \tag{4.64}$$

We now bound the terms on the right hand side. To this end we rely on the stability estimates of the consensus point and the linear growth condition from section 3.6 and lemma 4.4.1. We denote the empirical measure of the iid copies $\overline{V}_t^i$ by $\overline{\rho}_t^N$.

We estimate the first term on the right hand side of (4.63) as

$$2\lambda(V_t^i - \overline{V}_t^i) \cdot \Big(P_\Gamma(V_t^i) V^\alpha(\rho_t^N) - P_\Gamma(\overline{V}_t^i) V^\alpha(\rho_t) \Big) \tag{4.65}$$

$$\begin{aligned} &\leq 2\lambda |V_t^i - \overline{V}_t^i| \Big| P_\Gamma(V_t^i) \big(V^\alpha(\rho_t^N) - V^\alpha(\overline{\rho}_t^N) \big) \\ &\qquad + \big(P_\Gamma(V_t^i) - P_\Gamma(\overline{V}_t^i) \big) V^\alpha(\overline{\rho}_t^N) + P_\Gamma(\overline{V}_t^i) \big(V^\alpha(\overline{\rho}_t^N) - V^\alpha(\rho_t) \big) \Big| \end{aligned} \tag{4.66}$$

$$\begin{aligned} &\leq 2\lambda \|P_\Gamma\|_\infty |V_t^i - \overline{V}_t^i| \sum_{k=1}^N \frac{|V_t^k - \overline{V}_t^k|}{N} + 2\lambda C_\Gamma \|\nabla P_\Gamma\|_\infty |V_t^i - \overline{V}_t^i|^2 \\ &\qquad + 2\lambda \|P_\Gamma\|_\infty |V_t^i - \overline{V}_t^i| |V^\alpha(\overline{\rho}_t^N) - V^\alpha(\rho_t)| \end{aligned} \tag{4.67}$$

$$\lesssim \sum_{i=1}^N \frac{|V_t^i - \overline{V}_t^i|^2}{N} + |V_t^i - \overline{V}_t^i|^2 + |V^\alpha(\overline{\rho}_t^N) - V^\alpha(\rho_t)|^2 \tag{4.68}$$

where we have used the stability estimate (3.57)

$$|V^\alpha(\rho_t^N) - V^\alpha(\overline{\rho}_t^N)| \lesssim \frac{\sum_{k=1}^N |V_t^k - \overline{V}_t^k|}{N} \tag{4.69}$$

and the linear growth condition (3.77)

$$|V^\alpha(\rho_t^N)| \lesssim \frac{\|V_t^{(N)}\|_1}{N} \leq C_\Gamma. \tag{4.70}$$

We estimate the second term on the right hand side of (4.63) as

$$2(V_t^i - \overline{V}_t^i) \cdot \Big(C_\Gamma(V_t^i, V^\alpha(\rho_t^N)) - C_\Gamma(\overline{V}_t^i, V^\alpha(\rho_t)) \Big) dt \tag{4.71}$$

$$\begin{aligned} &\leq \sigma^2 |V_t^i - \overline{V}_t^i| \cdot \Big| |V_t^i - V^\alpha(\rho_t^N)|^2 \Delta\gamma(V_t^i) \nabla\gamma(V_t^i) - |V_t^i - V^\alpha(\rho_t^N)|^2 \Delta\gamma(\overline{V}_t^i) \nabla\gamma(\overline{V}_t^i) \\ &\quad + |V_t^i - V^\alpha(\rho_t^N)|^2 \Delta\gamma(\overline{V}_t^i) \nabla\gamma(\overline{V}_t^i) - |V_t^i - V^\alpha(\overline{\rho}_t^N)|^2 \Delta\gamma(\overline{V}_t^i) \nabla\gamma(\overline{V}_t^i) \\ &\quad + |V_t^i - V^\alpha(\overline{\rho}_t^N)|^2 \Delta\gamma(\overline{V}_t^i) \nabla\gamma(\overline{V}_t^i) - |\overline{V}_t^i - V^\alpha(\rho_t)|^2 \Delta\gamma(\overline{V}_t^i) \nabla\gamma(\overline{V}_t^i) \Big| \end{aligned} \tag{4.72}$$

$$\begin{aligned} &\lesssim \sigma^2 |V_t^i - \overline{V}_t^i| \cdot \Big| \|\Delta\gamma\|_\infty |\nabla\gamma(V_t^i) - \nabla\gamma(\overline{V}_t^i)| + \big(|V_t^i - V^\alpha(\rho_t^N)|^2 - |V_t^i - V^\alpha(\overline{\rho}_t^N)|^2 \big) \\ &\quad + \big(|V_t^i - V^\alpha(\overline{\rho}_t^N)|^2 - |\overline{V}_t^i - V^\alpha(\rho_t)|^2 \big) \Big| \end{aligned} \tag{4.73}$$

$$\lesssim \sigma^2 |V_t^i - \overline{V}_t^i| \cdot \Big| |V_t^i - \overline{V}_t^i| + |V^\alpha(\rho_t^N) - V^\alpha(\overline{\rho}_t^N)|^2 + |V^\alpha(\overline{\rho}_t^N) - V^\alpha(\rho_t)|^2 \Big| \tag{4.74}$$

$$\lesssim |V_t^i - \overline{V}_t^i|^2 + \sum_{i=1}^N \frac{|V_t^i - \overline{V}_t^i|^2}{N} + |V^\alpha(\overline{\rho}_t^N) - V^\alpha(\rho_t)|^2. \tag{4.75}$$

We denote the kth row of σ^i and P_Γ by σ_k^i and P_Γ^k respectively, and estimate the third term on the

right hand side of (4.63) as

$$\mathrm{Tr}(\sigma^i \cdot \sigma^i) = \sum_{k=1}^{d} (\sigma_k^i \cdot \sigma_k^i) \leq \sum_{k=1}^{d} \left| |V_t^i - V^\alpha(\rho_t^N)| P_\Gamma^k(V_t^i) - |\overline{V}_t^i - V^\alpha(\rho_t)| P_\Gamma^k(\overline{V}_t^i) \right|^2 \tag{4.76}$$

$$\begin{aligned} &\leq \sum_{k=1}^{d} \Big| |V_t^i - V^\alpha(\rho_t^N)| P_\Gamma^k(V_t^i) - |V_t^i - V^\alpha(\rho_t^N)| P_\Gamma^k(\overline{V}_t^i) \\ &\quad + |V_t^i - V^\alpha(\rho_t^N)| P_\Gamma^k(\overline{V}_t^i) - |V_t^i - V^\alpha(\overline{\rho}_t^N)| P_\Gamma^k(\overline{V}_t^i) \\ &\quad + |V_t^i - V^\alpha(\overline{\rho}_t^N)| P_\Gamma^k(\overline{V}_t^i) - |\overline{V}_t^i - V^\alpha(\rho_t)| P_\Gamma^k(\overline{V}_t^i) \Big|^2 . \end{aligned} \tag{4.77}$$

$$\begin{aligned} &\lesssim \sum_{k=1}^{d} \Big(\left| V_t^i - \overline{V}_t^i \right|^2 + \left| |V_t^i - V^\alpha(\rho_t^N)| - |V_t^i - V^\alpha(\overline{\rho}_t^N)| \right|^2 \\ &\quad \left| |V_t^i - V^\alpha(\overline{\rho}_t^N)| - |V_t^i - V^\alpha(\rho_t)| \right|^2 \Big) \end{aligned} \tag{4.78}$$

$$\lesssim \sum_{k=1}^{d} \Big(|V_t^i - \overline{V}_t^i|^2 + |V^\alpha(\rho_t^N) - V^\alpha(\overline{\rho}_t^N)|^2 + |V^\alpha(\overline{\rho}_t^N) - V^\alpha(\rho_t)|^2 \Big) \tag{4.79}$$

$$\lesssim d|V_t^i - \overline{V}_t^i|^2 + \frac{\sum_{i=1}^N d|V_t^i - \overline{V}_t^i|^2}{N} + d|V^\alpha(\overline{\rho}_t^N) - V^\alpha(\rho_t)|^2 \tag{4.80}$$

where we have used the stability estimate (3.57) and Jensen's inequality in the last line.

We do not need to further estimate the fourth term on the right hand side of (4.63) (the coefficient of dB_t^i), because the expected value of this term vanishes anyway due to the martingale property of Brownian motion.

Writing (4.63) in integral form, taking the expectation and collecting the above estimates thus yields

$$\begin{aligned} \mathbb{E}|V_t^i - \overline{V}_t^i|^2 \lesssim \mathbb{E}|V_0^i - \overline{V}_0^i|^2 + \int_0^t \mathbb{E}|V_s^i - \overline{V}_s^i|^2 ds + \int_0^t \frac{\sum_{i=1}^N \mathbb{E}|V_s^i - \overline{V}_s^i|^2}{N} ds + \\ + \int_0^t \mathbb{E}|V^\alpha(\overline{\rho}_s^N) - V^\alpha(\rho_s)|^2 \, ds \end{aligned} \tag{4.81}$$

$$\overset{L.4.4.1}{\lesssim} \mathbb{E}|V_0^i - \overline{V}_0^i|^2 + \int_0^t \sup_{i\in[N]} \mathbb{E}|V_s^i - \overline{V}_s^i|^2 ds + TN^{-1}. \tag{4.82}$$

With Gronwall's lemma we find

$$\sup_{i\in[N]} \mathbb{E}|V_t^i - \overline{V}_t^i|^2 \lesssim N^{-1} \tag{4.83}$$

where we have used that $\mathbb{E}|V_0^i - \overline{V}_0^i|^2 = 0$. This finishes the proof. □

4.5 Summary

The main result of this chapter is the mean-field limit result for the KV-CBO model, see theorem 4.4.1. A mean-field limit result is a statement about the convergence of the empirical measure ρ_t^N towards the solution ρ_t of the mean-field equation in the large particle limit $N \to \infty$. The convergence rate is N^{-1} and thus independent of the dimension d. The mean-field equation is a nonlinear, nonlocal deterministic PDE. We have proved the mean-field limit result with the coupling method, which relies on the propagation of chaos assumption and the introduction of an auxiliary monoparticle SDE, the mean-field dynamic. We have proved the well-posedness of the mean-field dynamic and the mean-field equation. The rigorous derivation of the mean-field equation and the proof of the mean-field limit result relied on the compactness of the hypersurface Γ.

Chapter 5

Optimization

In this chapter we prove the main result of this thesis, namely the convergence of the consensus point $V^\alpha(\rho_t)$ towards a global minimizer $V^\star$ of the cost function $\mathcal{E}$ for the special case $\Gamma = \mathbb{S}^{d-1}$. Here $V^\alpha(\rho_t)$ is the consensus point corresponding to the mean-field solution ρ_t, see theorem 5.4.2. As discussed in chapter 3, the cost function $\mathcal{E} \in C^2(\mathbb{R}^d, \mathbb{R}_+)$ needs to satisfy the ceorcivity condtion

$$|v - V^\star| \leq c_3 |\mathcal{E}(v) - \mathcal{E}(V^\star)|^{c_4} \tag{5.1}$$

which implies that it is enclosed in a convex envelope, see definition 3.1.1 on page 38. Further, we impose (strict) conditions on the distribution ρ_t and the coefficients of the KV-CBO model, see section 5.3. The analysis will be carried out on the macroscopic level only, that is, instead of analyzing the original particle system and the corresponding empirical measure ρ_t^N we analyse its mean-field solution ρ_t.

This chapter is based on joint work of Massimo Fornasier, Hui Huang, Lorenzo Pareschi and myself and has been published in [Fornasier et al., 2021a]. The proof discussed here is slighty different from the original proof: the conditions on the distribution and coefficient are different, and many (but not all) of the estimates are much sharper and no longer involve the constant $C_\alpha = \mathcal{O}(e^\alpha)$. Further the proof is significantly simpler than the orginal.

5.1 Calculus on the sphere

In the following lemma we collect some technical results about the gradient, Hessian and Laplacian of the weight function $e^{-\alpha\mathcal{E}(\cdot)}$ on the sphere $\Gamma = \mathbb{S}^{d-1}$. We refer to section 3.4 for differential geometry results on general hypersurfaces, and to page 141 in the appendix for integration by parts formulas on the sphere. As usual ∂_i denotes the partial derivative with respect to the ith variable, and $\partial_{\mathbb{S}^{d-1},i}$

denotes the ith component of the tangential gradient on the sphere $\nabla_{\mathbb{S}^{d-1}}$, that is.

$$\begin{cases} \partial_i \phi = (\nabla \phi)_i \\ \partial_{\mathbb{S}^{d-1},i} \Phi = (\nabla_{\mathbb{S}^{d-1}} \Phi)_i = (\nabla(\Phi \circ \Pi_{\mathbb{S}^{d-1}}))_i \\ \partial^2_{\mathbb{S}^{d-1},ij} \Phi = (\nabla^2_{\mathbb{S}^{d-1}} \Phi)_{ij} = (\nabla^2(\Phi \circ \Pi_{\mathbb{S}^{d-1}}))_{ij} \end{cases} \tag{5.2}$$

for $i, j \in [d]$, where $\Pi_{\mathbb{S}^{d-1}}(v) = v/|v|$.

Lemma 5.1.1 (Weight function on the sphere)**.** *Let $v \in \Gamma = \mathbb{S}^{d-1}$.*

For the gradient of the weight function we have

$$\begin{cases} \nabla e^{-\alpha \mathcal{E}(v)} = -\alpha e^{-\alpha \mathcal{E}(v)} \nabla \mathcal{E}(v) \in \mathbb{R}^d \\ \partial_i e^{-\alpha \mathcal{E}(v)} = -\alpha e^{-\alpha \mathcal{E}(v)} \partial_i \mathcal{E}(v) \\ \nabla_{\mathbb{S}^{d-1}} e^{-\alpha \mathcal{E}(v)} = -\alpha e^{-\alpha \mathcal{E}(v)} P_{\mathbb{S}^{d-1}}(v) \nabla \mathcal{E}(v) \in \mathbb{R}^d \\ \partial_{\mathbb{S}^{d-1},i} e^{-\alpha \mathcal{E}(v)} = \partial_i e^{-\alpha \mathcal{E}(v)} - v_i v \cdot \nabla e^{-\alpha \mathcal{E}(v)}. \end{cases} \tag{5.3}$$

For the Hessian of the weight function we have

$$\begin{cases} \nabla^2 e^{-\alpha \mathcal{E}(v)} = -\alpha e^{-\alpha \mathcal{E}(v)} (-\alpha \nabla \mathcal{E}(v) \otimes \nabla \mathcal{E}(v) + \nabla^2 \mathcal{E}(v)) \in \mathbb{R}^{d \times d} \\ \partial^2_{ij} e^{-\alpha \mathcal{E}(v)} = -\alpha e^{-\alpha \mathcal{E}(v)} (-\alpha \partial_i \mathcal{E}(v) \partial_j \mathcal{E}(v) + \partial^2_{ij} \mathcal{E}(v)) \\ \partial^2_{\mathbb{S}^{d-1},ii} e^{-\alpha \mathcal{E}(v)} = \partial^2_{ii} e^{-\alpha \mathcal{E}(v)} - 2 \sum_{j=1}^d \partial^2_{ij} e^{-\alpha \mathcal{E}(v)} v_j v_i - v \cdot \nabla e^{-\alpha \mathcal{E}(v)} + 2 v_i^2 v \cdot \nabla e^{-\alpha \mathcal{E}(v)} \\ \qquad\qquad -v_i \partial_i e^{-\alpha \mathcal{E}(v)} + v_i^2 \nabla^2 e^{-\alpha \mathcal{E}(v)} : v \otimes v. \end{cases} \tag{5.4}$$

For the Laplace-Beltrami operator of the weight function we have

$$\begin{cases} \Delta e^{-\alpha \mathcal{E}(v)} = \alpha^2 e^{-\alpha \mathcal{E}(v)} |\nabla \mathcal{E}(v)|^2 - \alpha e^{-\alpha \mathcal{E}(v)} \Delta \mathcal{E}(v) \in \mathbb{R}, \\ \Delta_{\mathbb{S}^{d-1}} e^{-\alpha \mathcal{E}(v)} = \Delta e^{-\alpha \mathcal{E}(v)} - (d-1) v \cdot \nabla e^{-\alpha \mathcal{E}(v)} - (v \otimes v) : \nabla^2 e^{-\alpha \mathcal{E}(v)} \in \mathbb{R}. \end{cases} \tag{5.5}$$

5.2 Regularity of the mean-field solution

In chapter 4 we have shown the well-posedness of the mean-field equation in the sense that there is a unique weak solution $\rho_t \in \mathcal{P}(\mathbb{S}^{d-1})$ for every initial distribution $\rho_0 \in \mathcal{P}(\mathbb{S}^{d-1})$. That is, the solution of the mean-field equation was always a probability measure (not a function). The main result of this chapter however requires a higher regularity of the solution ρ_t. In this section we will proof that the solution is in $L^2(\mathbb{S}^{d-1})$ if the initial distribution is. The proof is based on a Picard iteration argument and requires the Aubin-Lions theorem which we introduce below.

Let us first introduce the Lebesgue, Soboloev and Bochner spaces on $\mathbb{S}^{d-1}$.

Definition 5.2.1. *Let $p \in [1,\infty]$ and $(X, \|\cdot\|_X)$ a Banach space. We define*

$$\begin{cases} L^2(\mathbb{S}^{d-1}) & := \{f : \mathbb{S}^{d-1} \to \mathbb{R} \text{ measureable} \mid \|f\|_{L^2(\mathbb{S}^{d-1})} < \infty\}, \\ H^1(\mathbb{S}^{d-1}) & := \{f \in L^2(\mathbb{S}^{d-1}) \mid \|f\|_{H^1(\mathbb{S}^{d-1})} < \infty\}, \\ L^p([0,T],X) & := \{u : [0,T] \to X \mid \|u\|_{L^p([0,T],X)} < \infty\} \end{cases} \tag{5.6}$$

where the norms are defined as

$$\begin{cases} \|f\|^2_{L^2(\mathbb{S}^{d-1})} & := \int_{\mathbb{S}^{d-1}} |f(v)|^2 dv, \\ \|f\|_{H^1(\mathbb{S}^{d-1})} & := \|f\|_{L^2(\mathbb{S}^{d-1})} + \|\nabla_{\mathbb{S}^{d-1}} f\|_{L^2(\mathbb{S}^{d-1})}, \end{cases} \tag{5.7}$$

and

$$\|u\|_{L^p([0,T],X)} := \begin{cases} \int_0^T \|u(t)\|_X^p dt, & \text{for } p < \infty \\ \operatorname{esssup}_{t\in[0,T]} \|u(t)\|_X, & \text{for } p = \infty. \end{cases} \tag{5.8}$$

We further define the dual space $H^1(\mathbb{S}^{d-1})' := \mathcal{L}(H^1(\mathbb{S}^{d-1}), \mathbb{R})$ as the space of continuous linear functionals on $H^1(\mathbb{S}^{d-1})$. The norm is given by

$$\|f\|_{H^1(\mathbb{S}^{d-1})'} = \sup_{\|\psi\|_{H^1(\mathbb{S}^{d-1})} \le 1} |\langle f, \psi \rangle|. \tag{5.9}$$

Theorem 5.2.1 (Aubin-Lions, [Roubíček, 2013])**.** *Let $p \in (1,\infty)$, $q \in [1,\infty]$, V_1 a separable and reflexive Banach space, V_2 a Banach space, V_3 a metrizable locally convex Hausdorff space. Let further $V_1 \subset V_2$ be compactly embedded, and $V_2 \subset V_3$ continuously embedded. Then the space*

$$W^{1,p,q}([0,T]; V_1, V_3) := \{u \in L^p([0,T]; V_1 \mid \partial_t u \in L^q([0,T]; V_3))\} \tag{5.10}$$

is compactly embedded in $L^p([0,T]; V_2)$, that is, every sequence $\{\rho^n\}_{n\in\mathbb{N}} \subset W^{1,p,q}([0,T]; V_1, V_3)$ has a converging subsequence $\{\rho^{n_k}\}_{k\in\mathbb{N}}$ with

$$\rho^{n_k} \to \rho \in L^p([0,T]; V_2)) \tag{5.11}$$

as $k \to \infty$.

Theorem 5.2.2. *Let* $\rho_0 \in L^2(\mathbb{S}^{d-1})$. *Then there exists a unique weak solution of the mean-field equation* (4.45) *with the following regularity*

$$\begin{cases} \rho & \in W^{1,2,2}([0,T]; H^1(\mathbb{S}^{d-1}), H^1(\mathbb{S}^{d-1})') \\ \rho & \in L^\infty([0,T]; L^2(\mathbb{S}^{d-1})). \end{cases} \tag{5.12}$$

Proof. The proof follows the same argument as in [Albi et al., 2017, Theorem 2.4]. The uniqueness of a weak solution has been shown in theorem 4.3.1. For the proof of the existence we use a Picard iteration argument, see [William A. Adkins, 2012] to construct a sequence $\{\rho^n\}_{n\in\mathbb{N}} \subset W^{1,2,2}([0,T]; H^1(\mathbb{S}^{d-1}), H^1(\mathbb{S}^{d-1})')$ for which the Aubin-Lions theorem guarantees the existence of a subsequence $\{\rho^{n_k}\}_{k\in\mathbb{N}}$ such that $\rho^{n_k} \to \rho \in L^2([0,T]; L^2(\mathbb{S}^{d-1}))$ as $k \to \infty$. The function ρ is then shown to be a weak solution of the mean-field equation with the desired regularity.

We set the first iterate ρ^1 of the Picard iteration to the initial data ρ_0, that is,

$$\rho^1(t,v) := \rho_0(v), \quad t \in [0,T] \tag{5.13}$$

and consider for $n \geq 1$ the linear equation

$$\begin{cases} \partial_t \rho_t^{n+1} = -\lambda \nabla_{\mathbb{S}^{d-1}} \cdot (P_{\mathbb{S}^{d-1}}(v) V^\alpha(\rho_t^n) \rho_t^{n+1}) + \frac{\sigma^2}{2} \Delta_{\mathbb{S}^{d-1}}(|v - V^\alpha(\rho_t^n)|^2 \rho_t^{n+1}) \\ \rho^{n+1}(0,v) = \rho_0(v) \end{cases} \tag{5.14}$$

where

$$\rho^n \in L^\infty([0,T]; L^2(\mathbb{S}^{d-1})) \cap L^2([0,T]; H^1(\mathbb{S}^{d-1})). \tag{5.15}$$

This PDE has a unique weak solution ρ^{n+1} since

$$P_{\mathbb{S}^{d-1}}(v) V^\alpha(\rho_t^n), |v - V^\alpha(\rho_t^n)|^2 \in L^\infty(\mathbb{S}^{d-1}). \tag{5.16}$$

Multiplying both sides of (5.14) with ρ_t^{n+1} and using integration by parts yields

$$\begin{aligned} \frac{1}{2}\frac{d}{dt}\|\rho_t^{n+1}\|^2_{L^2(\mathbb{S}^{d-1})} = \lambda \int_{\mathbb{S}^{d-1}} P_{\mathbb{S}^{d-1}}(v) V^\alpha(\rho_t^n) \rho_t^{n+1} \cdot \nabla_{\mathbb{S}^{d-1}} \rho_t^{n+1} \, dv \\ - \frac{\sigma^2}{2} \int_{\mathbb{S}^{d-1}} \nabla_{\mathbb{S}^{d-1}}(|v - V^\alpha(\rho_t^n)|^2 \rho_t^{n+1}) \cdot \nabla_{\mathbb{S}^{d-1}} \rho_t^{n+1} \, dv. \end{aligned} \tag{5.17}$$

The first term on the right hand side can be bounded as

$$\begin{aligned} \lambda \int_{\mathbb{S}^{d-1}} P_{\mathbb{S}^{d-1}}(v) V^\alpha(\rho_t^n) \rho_t^{n+1} \cdot \nabla_{\mathbb{S}^{d-1}} \rho_t^{n+1} \, dv \\ \leq \lambda \int_{\mathbb{S}^{d-1}} |\rho_t^{n+1}| |\nabla_{\mathbb{S}^{d-1}} \rho_t^{n+1}| \, dv \end{aligned} \tag{5.18}$$

$$\leq \frac{\lambda\epsilon^2}{2}\|\rho_t^{n+1}\|^2_{L^2(\mathbb{S}^{d-1})} + \frac{\lambda}{2\epsilon^2}\|\nabla_{\mathbb{S}^{d-1}}\rho_t^{n+1}\|^2_{L^2(\mathbb{S}^{d-1})} \tag{5.19}$$

where we have used the Cauchy-Schwarz inequality in the first step and Young's inequality $ab \leq \frac{\epsilon^2 a^2}{2} + \frac{b^2}{2\epsilon^2}$ in the second.

The second term can be bounded as

$$\begin{aligned}
&-\frac{\sigma^2}{2}\int_{\mathbb{S}^{d-1}} \nabla_{\mathbb{S}^{d-1}}(|v - V^\alpha(\rho_t^n)|^2\rho_t^{n+1}) \cdot \nabla_{\mathbb{S}^{d-1}}\rho_t^{n+1}\, dv \\
&\quad = -\sigma^2\int_{\mathbb{S}^{d-1}} (v - V^\alpha(\rho_t^n))\rho_t^{n+1} \cdot \nabla_{\mathbb{S}^{d-1}}\rho_t^{n+1}\, dv - \frac{\sigma^2}{2}\int_{\mathbb{S}^{d-1}} |v - V^\alpha(\rho_t^n)|^2|\nabla_{\mathbb{S}^{d-1}}\rho_t^{n+1}|^2\, dv \qquad (5.20)\\
&\quad \leq -\sigma^2\int_{\mathbb{S}^{d-1}} (v - V^\alpha(\rho_t^n))\rho_t^{n+1} \cdot \nabla_{\mathbb{S}^{d-1}}\rho_t^{n+1}\, dv \qquad (5.21)\\
&\quad = \sigma^2\int_{\mathbb{S}^{d-1}} V^\alpha(\rho_t^n)\rho_t^{n+1} \cdot \nabla_{\mathbb{S}^{d-1}}\rho_t^{n+1}\, dv \qquad (5.22)\\
&\quad \leq \frac{\sigma^2\epsilon^2}{2}\|\rho_t^{n+1}\|^2_{L^2(\mathbb{S}^{d-1})} + \frac{\sigma^2}{2\epsilon^2}\|\nabla_{\mathbb{S}^{d-1}}\rho_t^{n+1}\|^2_{L^2(\mathbb{S}^{d-1})} \qquad (5.23)
\end{aligned}$$

where we have used that v is the outward normal and thus $v \cdot \nabla_{\mathbb{S}^{d-1}}\rho_t^{n+1}(v) = 0$. Remark: this step would not have worked on general hypersurfaces $\Gamma \neq \mathbb{S}^{d-1}$.

Combining the above estimates we find

$$\frac{1}{2}\frac{d}{dt}\|\rho_t^{n+1}\|^2_{L^2(\mathbb{S}^{d-1})} - \frac{1}{\epsilon^2}\|\nabla_{\mathbb{S}^{d-1}}\rho_t^{n+1}\|^2_{L^2(\mathbb{S}^{d-1})} \lesssim \epsilon^2\|\rho_t^{n+1}\|^2_{L^2(\mathbb{S}^{d-1})} \tag{5.24}$$

from which we deduce with Gronwall's inequality

$$\|\rho_t^{n+1}\|^2_{L^2(\mathbb{S}^{d-1})} + \int_0^T \|\nabla_{\mathbb{S}^{d-1}}\rho_t^{n+1}\|^2_{L^2(\mathbb{S}^{d-1})}dt \lesssim 1. \tag{5.25}$$

We further conclude

$$\begin{cases} \operatorname{esssup}_{t\in[0,T]} \|\rho_t^{n+1}\|_{L^2(\mathbb{S}^{d-1})} < \infty, \\ \int_0^T \|\rho_t^{n+1}\|^2_{H^1(\mathbb{S}^{d-1})}dt < \infty \end{cases} \tag{5.26}$$

hence

$$\rho^{n+1} \in L^\infty([0,T]; L^2(\mathbb{S}^{d-1})) \cap L^2([0,T]; H^1(\mathbb{S}^{d-1})). \tag{5.27}$$

In order to apply the Aubin-Lions theorem it remains to show that

$$\partial_t\rho^{n+1} \in L^2([0,T]; H^1(\mathbb{S}^{d-1})'). \tag{5.28}$$

This follows with

$$\begin{aligned}
&\|\partial_t \rho_t^{n+1}\|_{H^1(\mathbb{S}^{d-1})'} \\
&\quad = \sup_{\|\psi\|_{H^1(\mathbb{S}^{d-1})}\leq 1} |\langle \partial_t \rho_t^{n+1}, \psi\rangle| && (5.29)\\
&\quad = \sup_{\|\psi\|_{H^1(\mathbb{S}^{d-1})}\leq 1} \Big|\Big\langle \lambda P_{\mathbb{S}^{d-1}}(v) V^\alpha(\rho_t^n)\rho_t^{n+1} - \frac{\sigma^2}{2}\nabla_{\mathbb{S}^{d-1}}(|v - V^\alpha(\rho_t^n)|^2\rho_t^{n+1}), \nabla_{\mathbb{S}^{d-1}}\psi\Big\rangle\Big| && (5.30)\\
&\quad \lesssim \lambda\|\rho_t^{n+1}\|_{L^2(\mathbb{S}^{d-1})} + \frac{\sigma^2}{2}\|\nabla\rho_t^{n+1}\|_{L^2(\mathbb{S}^{d-1})} && (5.31)
\end{aligned}$$

where we have used the boundedness of the coefficient functions (5.16). We further find

$$\int_0^T \|\partial_t \rho_t^{n+1}\|^2_{H^1(\mathbb{S}^{d-1})'}\, dt \leq \int_0^T 2\lambda\|\rho_t^{n+1}\|^2_{L^2(\mathbb{S}^{d-1})} + \sigma^2\|\nabla\rho_t^{n+1}\|^2_{L^2(\mathbb{S}^{d-1})}\, dt < \infty \tag{5.32}$$

which finishes the proof of (5.28).

So far we have constructed a sequence

$$\{\rho^n\}_{n\in\mathbb{N}} \subset W^{1,2,2}([0,T]; H^1(\mathbb{S}^{d-1}), H^1(\mathbb{S}^{d-1})') \tag{5.33}$$

for which the Aubin-Lions theorem guarantees the existence of a subsequence $\{\rho^{n_k}\}_{k\in\mathbb{N}}$ such that

$$\rho^{n_k} \to \rho \in L^2([0,T]; L^2(\mathbb{S}^{d-1})) \tag{5.34}$$

as $k \to \infty$. As the spaces $W^{1,2,2}$ and $L^\infty([0,T]; L^2(\mathbb{S}^{d-1}))$ are Banach spaces the limit function ρ also belongs to those spaces, proving (5.12).

It remains to show that ρ solves the mean-field equation (in the weak sense). To this end we first note that the mean-value theorem yields

$$\max_{t\in[0,T]} \|\rho^n\|_{L^2(\mathbb{S}d-1)} \lesssim \|\partial_t\rho^n\|_{L^2([0,T];H^1(\mathbb{S}^{d-1})')} + \|\rho^n\|_{L^2([0,T];H^1(\mathbb{S}^{d-1}))} < \infty \tag{5.35}$$

from which we deduce $\rho^n \in C([0,T]; L^2(\mathbb{S}^{d-1}))$. From this we see that the weak formulation of the lhs of (5.14) converges to the lhs of the weak formulation of the mean-field equation (see (4.44))

$$\frac{d}{dt}\int_{\mathbb{S}^{d-1}} \big|\Phi(v)(\rho_t^{n_k}(v) - \rho_t(v))\big|^2 dv \to 0. \tag{5.36}$$

For the rhs we find

$$\begin{aligned}
&\lambda\int_{\mathbb{S}^{d-1}} \big|\Phi(v)(P_{\mathbb{S}^{d-1}}(v)V^\alpha(\rho_t^{n_k})\rho_t^{n_k+1} - P_{\mathbb{S}^{d-1}}(v)V^\alpha(\rho_t)\rho_t)\cdot\nabla_{\mathbb{S}^{d-1}}\Phi(v)\big|^2 dv && (5.37)\\
&\qquad \lesssim \int_{\mathbb{S}^{d-1}} \big|V^\alpha(\rho_t^{n_k})\rho_t^{n_k+1} - V^\alpha(\rho_t)\rho_t)\big|^2 dv && (5.38)
\end{aligned}$$

$$\lesssim \int_{\mathbb{S}^{d-1}} \left|V^\alpha(\rho_t^{n_k})\rho_t^{n_k+1} - V^\alpha(\rho_t^{n_k})\rho_t\right|^2 + \left|V^\alpha(\rho_t^{n_k})\rho_t - V^\alpha(\rho_t)\rho_t\right|^2 dv \tag{5.39}$$

$$\lesssim \int_{\mathbb{S}^{d-1}} \left|\rho_t^{n_k+1} - \rho_t\right|^2 dv + \int_{\mathbb{S}^{d-1}} \left|\rho_t^{n_k} - \rho_t\right|^2 dv \to 0 \tag{5.40}$$

for the convection part and similarly

$$\int_{\mathbb{S}^{d-1}} \left||v - V^\alpha(\rho_t^{n_k})|^2\rho_t^{n_k+1} - |v - V^\alpha(\rho_t)|^2\rho_t\right| \to 0 \tag{5.41}$$

for the diffusion part which finishes the proof. □

5.3 Conditions on the distribution and coefficients

In this section we discuss the conditions on the distribution ρ_t and the coefficients of the KV-CBO model needed to prove the convergence of the consensus point $V^\alpha(\rho_t)$ towards the global minimizer $V^\star$. The conditions on the cost function $\mathcal{E}$ (definition 3.1.1) have already been discussed on many occasions and will not be treated here.

Let us start with an important remark.

Remark 5.3.1 (Conditions from definition 5 in [Fornasier et al., 2021a])**.** *In [Fornasier et al., 2021a] the following (complicated) conditions*

$$C_\alpha^{2max\{1,c_4\}}\left(V(\rho_0) + \frac{\lambda C_T}{\lambda\vartheta - 4C_\alpha C_{\sigma,d}}\delta^{(d-2)/4}\right)^{min\{1,c_4\}/2} + \epsilon^{c_4} < \frac{\delta - \vartheta}{C^\star}, \tag{5.42}$$

$$V(\rho_0) + \frac{\lambda C_T}{\lambda\vartheta - 4C_\alpha C_{\sigma,d}}\delta^{(d-2)/4} \leq \min\left\{\frac{\|e^{-\alpha\mathcal{E}}\|_{L^1(\rho_0)}^2}{T}, \frac{\|e^{-\alpha\mathcal{E}}\|_{L^1(\rho_0)}^4}{T\lambda^2}, \frac{3}{8}\right\}, \tag{5.43}$$

$$\lambda\vartheta - 4C_\alpha C_{\sigma,d} > 0 \tag{5.44}$$

are assumed to hold.

Here $C_T = \sup_{t\in[0,T]} \|\rho_t\|_{L^2(\mathbb{S}^{d-1})}$, $\delta \in (0,2)$ *(see* (5.46) *below),* $\vartheta \in (0,2)$ *and* $C^\star = \mathcal{O}(e^\alpha)$ *where* $\alpha \gg 1$ *is the temperature parameter. The other constants are the same as in this thesis, that is,* $C_\alpha = e^{\alpha(\overline{\mathcal{E}} - \underline{\mathcal{E}})}$*,* T *is the time horizon,* λ *is the coupling constant and*

$$C_{\sigma,d} = (d-1)\sigma^2/2. \tag{5.45}$$

Further, c_4 *is the exponent from the coercivity condition* (5.1) *and* ϵ *is the approximation error from lemma 5.4.4 below. This lemma implies that* ϵ *can only be small if* α *is large (see* (2.74) *for the convergence rate).*

The number on the right hand side of (5.42) *is in the order of* $\mathcal{O}(1/e^\alpha)$*. Hence the condition can*

only be satisfied if ϵ *and the initial variance* $V(\rho_0)$ *on the left hand side are very small. More precisely: if we assume* $\epsilon = 0$, $c_4 = 1$ *and* $\frac{\lambda C_T}{\lambda\vartheta - 4C_\alpha C_{\sigma,d}}\delta^{(d-2)/4} = 0$ *(which is not possible), then* (5.42) *implies* $V(\rho_0) = \mathcal{O}(e^{-6\alpha})$.

In other words: the above conditions can only be satisfied if the variance $V(\rho_0)$ *is zero and* $\mathbb{E}(\rho_0)$ *is equal to the global minimizer* $V^\star$.

Remark 5.3.2. *From the condition* (5.43) *we further see that a large coupling constant* λ *or a large (or even infinite) time horizon* T *are undesirable as they make the initial variance* $V(\rho_0)$ *much smaller. We further note that condition* (5.44) *implies* $\lambda = \mathcal{O}(e^\alpha)$.

The above remarks show that the optimization result of the KV-CBO model is not a *global* but rather a *local* optimization result, that is, convergence to the global minimizer is guaranteed if we have *a-priori* knowledge about its location. To make this locality condition more precise we define the cap

$$D_\delta = \{v \in \mathbb{S}^{d-1} \mid -1 \leq \langle v, V^\star\rangle \leq -1 + \delta\} \tag{5.46}$$

for some fixed constant $\delta \in (0, 2)$ and assume the density ρ_t to have most of its mass outside the cap D_δ and thus close to $V^\star$, that is, we assume

$$\sup_{t\in[0,T]} \|\rho_t\|_{L^2(D_\delta)} \leq \epsilon_0 \tag{5.47}$$

for some $\epsilon_0 = \epsilon_0(\delta) \geq 0$. Here we need $\rho_t \in L^2(\mathbb{S}^{d-1})$ which is given if the initial distribution has this regularity.

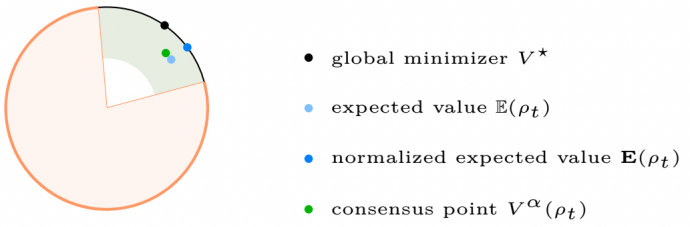


Figure 5.1: Schematic depiction of the cap D_δ (orange boundary) and the set $\tilde{D}_\delta$ (orange fill). D_δ is a region of $\mathbb{S}^{d-1}$ opposed to the global minimizer, that is, $V^\star \notin D_\delta$. It is larger if $\delta \in [0, 2]$ is larger. The expected value and consensus points are expected to stay within the green area.

An other condition for the distribution is

$$\overline{\mathcal{V}}_T := \sup_{t\in[0,T]} V(\rho_t) \leq \min\left\{\frac{\|e^{-\alpha\mathcal{E}}\|^2_{L^1(\rho_0)}}{T}, \frac{\|e^{-\alpha\mathcal{E}}\|^4_{L^1(\rho_0)}}{T\lambda^2}\right\}. \tag{5.48}$$

This is a technical assumption that can not be easily motivated. We therefore just make a few remarks about it. We first note that since $\mathcal{E} \geq 0$ the weight $e^{-\alpha\mathcal{E}} \leq 1$ is bounded from above and $\|e^{-\alpha\mathcal{E}}\|_{L^1(\rho_0)} \leq 1$.

We further assume the initial distribution to be chosen such that the expected value and the consensus point are reasonably close to $V^\star$ and stay reasonably close for any time $t \geq 0$. To make this more precise we introduce the set

$$\tilde{D}_\delta = \{v \in B_1^d(0) \mid \Pi_{\mathbb{S}^{d-1}}(v) \in D_\delta\} \tag{5.49}$$

where $B_1^d(0)$ is the ball in $\mathbb{R}^d$ with radius $r = 1$ around the origin and $\Pi_{\mathbb{S}^{d-1}}(v) = \frac{v}{|v|}$ is the projection onto the boundary of $B_1^d(0)$. We now assume

$$\mathbb{E}(\rho_t),\ V^\alpha(\rho_t) \in B_1^d(0) \setminus \left(\tilde{D}_\delta \cup B_{1/2}^d(0)\right) \tag{5.50}$$

for all $t \in [0, T]$. This allows us to find a good (large) lower bound on the inner product of the expected value and the consensus point, that is,

$$\mathbb{E}(\rho_t) \cdot V^\alpha(\rho_t) \geq C_1 \tag{5.51}$$

for a strictly positive constant $C_1 \in (0, 1]$. An estimate of this form is necessary for the convergence proof.

So far, we have discussed the conditions on the distribution ρ_t. Additionally, the coefficients α, σ, λ of the KV-CBO model need to satisfy certain conditions to guarantee convergence to the global minimizer. These conditions are reported below and will be used in the following sections to show the monotonic decay of the variance $V(\rho_t)$.

Let us summarize the conditions for the initial distribution and the coefficients. These conditions are assumed to hold in the rest of this chapter. The condition $\rho_0 \in \mathcal{P}_{ac}(\mathbb{S}^{d-1})$ is needed to apply Laplace's principle.

Definition 5.3.1. *Let $\delta \in (0, 2)$ and $\rho_0 \in \mathcal{P}_{ac}(\mathbb{S}^{d-1}) \cap L^2(\mathbb{S}^{d-1})$ the initial distribution. We assume*

$$\begin{cases} \sup_{t\in[0,T]} \|\rho_t\|_{L^2(D_\delta)} \leq \epsilon_0, \\ \mathbb{E}(\rho_t) \cdot\ V^\alpha(\rho_t) \geq C_1 \quad \forall t \in [0, T], \\ \overline{\mathcal{V}}_T := \sup\limits_{t\in[0,T]} V(\rho_t) \leq \min\left\{ \frac{\|e^{-\alpha\mathcal{E}}\|^2_{L^1(\rho_0)}}{T}, \frac{\|e^{-\alpha\mathcal{E}}\|^4_{L^1(\rho_0)}}{T\lambda^2} \right\}. \end{cases} \tag{5.52}$$

For the coefficients α, σ, λ and δ we assume

$$\begin{cases} \exists \vartheta > 0 \text{ s.t. } C_1 + 5/8 - C_2/2 - (2-\delta) \geq \vartheta > 0, \\ \lambda\vartheta - 4C_{\sigma,d} > 0 \end{cases} \tag{5.53}$$

where $C_1 > 0$ is the constant from (5.51) *and C_2 is the constant from lemma 5.4.3 below.*

5.4 Convergence towards the global minimizer

In this section we prove the main result of this chapter, namely the convergence towards the global minimizer $V^\star$. As mentioned above this convergence result is a local optimization result as we assume a-priori knowledge about the location of $V^\star$. We start with some technical estimates.

We can relate the variance and the expected value by noting

$$2V(\rho_t) = \int_{\mathbb{S}^{d-1}} |v - \mathbb{E}(\rho_t)|^2 \, d\rho_t(v) = 1 - 2\mathbb{E}(\rho_t) \cdot \int_{\mathbb{S}^{d-1}} v \, d\rho_t(v) + |\mathbb{E}(\rho_t)|^2 \tag{5.54}$$

which yields the important equality

$$2V(\rho_t) = 1 - |\mathbb{E}(\rho_t)|^2. \tag{5.55}$$

From this we can easily conclude that $|\mathbb{E}(\rho_t)| \geq 1/2$ once the variance $V(\rho_t) \leq 3/8$. Further, we observe that $maxV(\rho_t) = 1/2$ and that this maximum is attained by the uniform density on $\mathbb{S}^{d-1}$. In this case we can also find a sharp estimate on the integral $\int_{\mathbb{S}^{d-1}} |v - \mathbf{E}(\rho_t)|^2 \, d\rho_t(v)$ where

$$\mathbf{E}(\rho_t) = \frac{\mathbb{E}(\rho_t)}{|\mathbb{E}(\rho_t)|} \tag{5.56}$$

is the normalized expected value, that is,

$$\int_{\mathbb{S}^{d-1}} |v - \mathbf{E}(\rho_t)|^2 \, d\rho_t(v) = 2(1 - |\mathbb{E}(\rho_t)|) \overset{(5.55)}{=} \frac{4V(\rho_t)}{1 + |\mathbb{E}(\rho_t)|} \leq \frac{8}{3} V(\rho_t). \tag{5.57}$$

We can further estimate

$$\begin{cases} |\mathbf{E}(\rho_t) - V^\star|^2 & = 2(1 - \mathbf{E}(\rho_t) \cdot V^\star) \leq 2(2-\delta), \\ |V^\alpha(\rho_t) - V^\star|^2 & = 2(1 - V^\alpha(\rho_t) \cdot V^\star) \leq 3 - \delta \end{cases} \tag{5.58}$$

where we have used that $\mathbf{E}(\rho_t) \cdot V^\star \geq -1 + \delta$ and $V^\alpha(\rho_t) \cdot V^\star \geq |V^\alpha(\rho_t)|(-1+\delta)$ and $|V^\alpha(\rho_t)| \geq 1/2$.

In the next lemma we bound integrals that involve the consensus point $V^\alpha(\rho_t)$. We will use a simple but important trick to bound the weight $e^{-\alpha\mathcal{E}}$ independent of α. This trick will also be used

in the next chapter when we discuss the numerical implementation of the KV-CBO model, see (6.8).

Lemma 5.4.1. *The following estimates hold*

$$\begin{cases} \int_{\mathbb{S}^{d-1}} |V^\alpha(\rho_t)|^2 \, d\rho_t(v) \leq 2V(\rho_t), \\ \int_{\mathbb{S}^{d-1}} |v - V^\alpha(\rho_t)|^2 \, d\rho_t(v) \leq 4V(\rho_t). \end{cases} \tag{5.59}$$

Proof. Let us assume that the global minimizer $V^\star \in \mathrm{supp}(\rho_t)$. An application of Jensen's inequality allows us to bound the first integral in (5.59) as

$$\int_{\mathbb{S}^{d-1}} |V^\alpha(\rho_t)|^2 \, d\rho_t(v) \leq \int_{\mathbb{S}^{d-1}} \frac{|v|^2 e^{-\alpha(\mathcal{E}(v)-\mathcal{E}(V^\star))}}{\int_{\mathbb{S}^{d-1}} e^{-\alpha(\mathcal{E}(w)-\mathcal{E}(V^\star))} \, d\rho_t(w)} \, d\rho_t(v). \tag{5.60}$$

Noting that $e^{-\alpha(\mathcal{E}(v)-\mathcal{E}(V^\star))} \leq 1$ for all $v \in \mathbb{S}^{d-1}$ and $e^{-\alpha(\mathcal{E}(w)-\mathcal{E}(V^\star))} = 1$ for at least one $w \in \mathbb{S}^{d-1}$ (the global minimizer $V^\star$) finishes the proof of the first estimate.

For the second estimate we proceed similarly. An application of Jensen's inequality to the convex function $\phi(w) = |v - w|^2$ yields

$$\begin{aligned} \int_{\mathbb{S}^{d-1}} |v - V^\alpha(\rho_t)|^2 d\rho_t(v) &\leq \int_{\mathbb{S}^{d-1}} \frac{\int_{\mathbb{S}^{d-1}} |v-w|^2 e^{-\alpha\mathcal{E}(w)} d\rho_t(w)}{\|e^{-\alpha\mathcal{E}}\|_{L^1(\rho_t)}} \, d\rho_t(v) && (5.61) \\ &\leq \int_{\mathbb{S}^{d-1}} \int_{\mathbb{S}^{d-1}} |v-w|^2 \, d\rho_t(w) d\rho_t(v) && (5.62) \\ &= \int_{\mathbb{S}^{d-1}} \int_{\mathbb{S}^{d-1}} |v - \mathbb{E}(\rho_t)|^2 + |w - \mathbb{E}(\rho_t)|^2 \, d\rho_t(w) d\rho_t(v) && (5.63) \\ &\leq 4V(\rho_t). && (5.64) \end{aligned}$$

From this we further find

$$\int_{\mathbb{S}^{d-1}} |v - V^\alpha(\rho_t)| \, d\rho_t(v) \leq 2V(\rho_t)^{1/2}. \tag{5.65}$$

□

If the variance of ρ_t converges to zero the consensus point and the expected value both converge to the sphere and their inner product to one. In the next lemma we prove this rigorously.

Lemma 5.4.2. *The following estimates hold*

$$\begin{cases} 1 - 4V(\rho_t) \leq |V^\alpha(\rho_t)|^2 \leq 1, \\ 2V(\rho_t) = 1 - |\mathbb{E}(\rho_t)|^2, \\ \mathbb{E}(\rho_t) \cdot V^\alpha(\rho_t) \geq V(\rho_t) + \frac{|\mathbb{E}(\rho_t)|^2 + |V^\alpha(\rho_t)|^2}{2} - 2V(\rho_t). \end{cases} \tag{5.66}$$

Proof. We have

$$1 - |V^\alpha(\rho_t)|^2 = \frac{\int_{\mathbb{S}^{d-1}} (1 - |V^\alpha(\rho_t)|^2) e^{-\alpha \mathcal{E}(u)} d\rho_t(u)}{\| e^{-\alpha \mathcal{E}} \|_{L^1(\rho_t)}} \leq 4V(\rho_t) \tag{5.67}$$

where we have used (5.59). The other direction follows easily with Jensen's inequality.

The second statement has already been shown above, see (5.55).

For the thrid statement we find

$$\int_{\mathbb{S}^{d-1}} |v - V^\alpha(\rho_t)|^2 \, d\rho_t(v) = \int_{\mathbb{S}^{d-1}} |v - \mathbb{E}(\rho_t) + \mathbb{E}(\rho_t) - V^\alpha(\rho_t)|^2 \, d\rho_t(v) \tag{5.68}$$

$$= 2V(\rho_t) + |\mathbb{E}(\rho_t)|^2 - 2\mathbb{E}(\rho_t) \cdot V^\alpha(\rho_t) + |V^\alpha(\rho_t)|^2 \tag{5.69}$$

where we have used that

$$\int_{\mathbb{S}^{d-1}} \langle v - \mathbb{E}(\rho_t), \mathbb{E}(\rho_t) - V^\alpha(\rho_t) \rangle \, d\rho_t(v) = 0. \tag{5.70}$$

Hence

$$\mathbb{E}(\rho_t) \cdot V^\alpha(\rho_t) = V(\rho_t) + \frac{|\mathbb{E}(\rho_t)|^2 + |V^\alpha(\rho_t)|^2}{2} - \frac{1}{2} \int_{\mathbb{S}^{d-1}} |v - V^\alpha(\rho_t)|^2 \, d\rho_t(v) \tag{5.71}$$

$$\geq V(\rho_t) + \frac{|\mathbb{E}(\rho_t)|^2 + |V^\alpha(\rho_t)|^2}{2} - 2V(\rho_t). \tag{5.72}$$

□

We now prove the monotonic decay of the variance of ρ_t. Let us first derive an expression for the time derivative of the variance. Starting with a solution $\overline{V}_t \sim \rho_t$ of the mean-field dynamic (4.9) where $\mathbb{E}(\rho_t) = \mathbb{E}(\overline{V}_t)$ and applying Dynkin's formula (with $f(v) = v$) we find an expression for the time derivative of the expected value $\frac{d}{dt}\mathbb{E}(\rho_t)$. With (5.55) we find

$$\frac{d}{dt} V(\rho_t) = -\mathbb{E}(\rho_t) \cdot \frac{d}{dt} \mathbb{E}(\rho_t). \tag{5.73}$$

and further

$$\frac{d}{dt} V(\rho_t) = -\mathbb{E}(\rho_t) \cdot \int_{\mathbb{S}^{d-1}} \lambda P_{\mathbb{S}^{d-1}}(v) V^\alpha(\rho_t) - C_{\sigma,d} |v - V^\alpha(\rho_t)|^2 v \, d\rho_t(v). \tag{5.74}$$

Already here we are made aware of the importance of a large (positive) bound on the inner product $\mathbb{E}(\rho_t) \cdot V^\alpha(\rho_t)$.

Lemma 5.4.3. *The following estimate holds*

$$\int_{\mathbb{S}^{d-1}} |\mathbb{E}(\rho_t) - v|^2 |v - V^\alpha(\rho_t)|^2 \, d\rho_t(v) \leq 16\epsilon_0 C_3^{1/2} \delta^{(d-2)/4} + 4(2-\delta)V(\rho_t) + 2C_2 V(\rho_t) \tag{5.75}$$

for positive constants $C_2 = C_2(\delta)$ and $C_3 = C_3(d)$.

Proof. We have

$$\begin{cases} |v - V^\alpha(\rho_t)|^2 = |v - V^\star|^2 + 2\langle v - V^\star, V^\star - V^\alpha(\rho_t)\rangle + |V^\alpha(\rho_t) - V^\star|^2, \\ 2\langle v - V^\star, v^\star - V^\alpha(\rho_t)\rangle \leq 4|V^\star - \mathbf{E}(\rho_t)| + 4|\mathbf{E}(\rho_t) - V^\alpha(\rho_t)|, \\ |V^\alpha(\rho_t) - V^\star|^2 \leq 2|V^\alpha(\rho_t) - \mathbf{E}(\rho_t)|^2 + 2|\mathbf{E}(\rho_t) - V^\star|^2. \end{cases} \tag{5.76}$$

With these estimates we find

$$\begin{aligned} &\int_{\mathbb{S}^{d-1}} |\mathbb{E}(\rho_t) - v|^2 |v - V^\alpha(\rho_t)|^2 \, d\rho_t(v) \\ &\quad \leq \int_{\mathbb{S}^{d-1}} |\mathbb{E}(\rho_t) - v|^2 |v - V^\star|^2 \, d\rho_t(v) \\ &\qquad + 2V(\rho_t)\Big(4|V^\alpha(\rho_t) - \mathbf{E}(\rho_t)| + 4|\mathbf{E}(\rho_t) - V^\star| + 2|V^\alpha(\rho_t) - \mathbf{E}(\rho_t)|^2 + 2|\mathbf{E}(\rho_t) - V^\star|^2\Big). \end{aligned} \tag{5.77}$$

The terms in the parentheses can be bounded by $C_2 \lesssim 2(2-\delta)$ where we have used (5.58).

The integral

$$\int_{\mathbb{S}^{d-1}} |\mathbb{E}(\rho_t) - v|^2 |v - V^\star|^2 \, d\rho_t(v) \tag{5.78}$$

on the right hand side of (5.77) is split into two integrals, one over the cap D_δ and one over $\mathbb{S}^{d-1} \setminus D_\delta$.

For the first integral we use $\rho \in L^\infty([0,T]; L^2(\mathbb{S}^{d-1}))$ (see theorem 5.2.2) to bound

$$\int_{\mathcal{D}_\delta} d\rho_t(v) \leq \|\rho_t\|_{L^2(D_\delta)} vol(D_\delta)^{1/2} \leq \|\rho_t\|_{L^2(D_\delta)} C_3^{1/2} \delta^{(d-2)/4} \tag{5.79}$$

where we have used the following formula for the volume of the cap

$$vol(D_\delta) = \frac{1}{2} a_d I_{2\delta - \delta^2}\left(\frac{d-1}{2}, \frac{1}{2}\right) \leq C \frac{\pi^{d/2}}{\Gamma(d/2)} \frac{(d-1)^{1/2}}{d-2} \delta^{(d-2)/2} \leq C_3 \delta^{(d-2)/2} \tag{5.80}$$

where a_d is the area of a unit ball, $I_x(a,b)$ is the regularized incomplete beta function and Γ is the Gamma function.

Thus

$$\int_{D_\delta} |\mathbb{E}(\rho_t) - v|^2 |v - V^\star|^2 \, d\rho_t(v) \leq 16 \int_{D_\delta} d\rho_t(v) \leq 16\epsilon_0 C_3^{1/2} \delta^{(d-2)/4}. \tag{5.81}$$

For the second integral we find

$$\int_{\mathbb{S}^{d-1}\setminus D_\delta} |\mathbb{E}(\rho_t) - v|^2 |v - V^\star|^2 \, d\rho_t(v) \leq 4(2-\delta) V(\rho_t). \tag{5.82}$$

where we have used the identity $|v - V^\star|^2 = 2(1 - \langle v, V^\star \rangle) \leq 2(2-\delta)$ for all $v \in \mathbb{S}^{d-1}/\mathcal{D}_\delta$.

Combining the above estimates finishes the proof. □

In the next theorem we show the monotonic decay of the variance provided $\lambda\vartheta > 4C_{\sigma,d}$. The decay is exponential, but only converges to zero if $\epsilon_0 = 0$.

Theorem 5.4.1 (Monotonic decay of the variance)**.** *We have*

$$V(\rho_t) \leq V(\rho_0) e^{-(\lambda\vartheta - 4C_{\sigma,d})t} + \frac{4\lambda\epsilon_0 C_3^{1/2} \delta^{(d-2)/4}}{\lambda\vartheta - 4C_{\sigma,d}}. \tag{5.83}$$

Proof. With the polarization properties

$$\begin{cases} v \cdot \mathbb{E}(\rho_t) & = \frac{|\mathbb{E}(\rho_t)|^2 + 1}{2} - \frac{|v - \mathbb{E}(\rho_t)|^2}{2}, \\ v \cdot V^\alpha(\rho_t) & = \frac{|V^\alpha(\rho_t)|^2 + 1}{2} - \frac{|v - V^\alpha(\rho_t)|^2}{2} \end{cases} \tag{5.84}$$

we can estimate the first term on the right hand side of (5.74) as

$$-\lambda \int_{\mathbb{S}^{d-1}} \mathbb{E}(\rho_t) \cdot P_{\mathbb{S}^{d-1}}(v) V^\alpha(\rho_t) \, d\rho_t(v) \tag{5.85}$$

$$= -\lambda \mathbb{E}(\rho_t) \cdot V^\alpha(\rho_t) + \lambda \int_{\mathbb{S}^{d-1}} (v \cdot \mathbb{E}(\rho_t))(v \cdot V^\alpha(\rho_t)) \, d\rho_t(v) \tag{5.86}$$

$$\begin{aligned} &= -\lambda \mathbb{E}(\rho_t) \cdot V^\alpha(\rho_t) \\ &\quad + \lambda \int_{\mathbb{S}^{d-1}} \frac{|\mathbb{E}(\rho_t)|^2 + 1}{2} v \cdot V^\alpha(\rho_t) \, d\rho_t(v) - \lambda \int_{\mathbb{S}^{d-1}} \frac{|v - \mathbb{E}(\rho_t)|^2}{2} v \cdot V^\alpha(\rho_t) \, d\rho_t(v) \end{aligned} \tag{5.87}$$

$$= -\lambda \mathbb{E}(\rho_t) \cdot V^\alpha(\rho_t) + \lambda \frac{|\mathbb{E}(\rho_t)|^2 + 1}{2} \mathbb{E}(\rho_t) \cdot V^\alpha(\rho_t) - \lambda \int_{\mathbb{S}^{d-1}} \frac{|v - \mathbb{E}(\rho_t)|^2}{2} v \cdot V^\alpha(\rho_t) \, d\rho_t(v) \tag{5.88}$$

$$\leq -\lambda C_1 V(\rho_t) - \lambda \int_{\mathbb{S}^{d-1}} \frac{|v - \mathbb{E}(\rho_t)|^2}{2} v \cdot V^\alpha(\rho_t) \, d\rho_t(v) \tag{5.89}$$

where we have used equality (5.55) and the bound $\mathbb{E}(\rho_t) \cdot V^\alpha(\rho_t) \geq C_1$ in the last step.

The integral on the right hand side of (5.89) can be bounded by

$$\begin{aligned} &-\lambda \int_{\mathbb{S}^{d-1}} \frac{|v - \mathbb{E}(\rho_t)|^2}{2} v \cdot V^\alpha(\rho_t) \, d\rho_t(v) \\ &\quad = -\lambda \int_{\mathbb{S}^{d-1}} \frac{|v - \mathbb{E}(\rho_t)|^2}{2} \frac{|V^\alpha(\rho_t)|^2 + 1}{2} \, d\rho_t(v) + \frac{\lambda}{4} \int_{\mathbb{S}^{d-1}} |v - \mathbb{E}(\rho_t)|^2 |v - V^\alpha(\rho_t)|^2 \, d\rho_t(v) \end{aligned} \tag{5.90}$$

$$\leq -\frac{5\lambda}{8} V(\rho_t) + \frac{\lambda}{4} \int_{\mathbb{S}^{d-1}} |v - \mathbb{E}(\rho_t)|^2 |v - V^\alpha(\rho_t)|^2 \, d\rho_t(v). \tag{5.91}$$

With lemma 5.4.3 we can further bound the integral term on the right hand side.

For the second term on the right hand side of (5.74) we find

$$C_{\sigma,d}\int_{\mathbb{S}^{d-1}}|v-V^\alpha(\rho_t)|^2 v\cdot\mathbb{E}(\rho_t)\;d\rho_t(v)\leq 4C_{\sigma,d}V(\rho_t). \tag{5.92}$$

Combining these estimates and using (5.53) yields

$$\frac{d}{dt}V(\rho_t)\leq -\big(\lambda\vartheta-4C_{\sigma,d}\big)V(\rho_t)+4\lambda\epsilon_0 C_3^{1/2}\delta^{(d-2)/4}. \tag{5.93}$$

An application of Gronwall's inequality finishes the proof. □

We now prove a variant of Laplace's principle for the mean-field solution. Here we need the third condition from definition 5.3.1. For a discussion about Laplace's principle for $\rho_0\in\mathcal{P}_{ac}(\mathbb{R}^d)$ we refer to chapter 2.

Lemma 5.4.4 (Laplace's principle for the mean-field solution)**.** *Let $\rho_0\in\mathcal{P}_{ac}(\mathbb{S}^{d-1})\cap L^2(\mathbb{S}^{d-1})$ and $\rho\in L^\infty([0,T];L^2(\mathbb{S}^{d-1}))$ the mean-field solution of the KV-CBO model. Then*

$$\lim_{\alpha\to\infty}-\frac{1}{\alpha}\log\|e^{-\alpha\mathcal{E}}\|_{L^1(\rho_t)}=\underline{\mathcal{E}} \tag{5.94}$$

for all $t\in[0,T]$.

Proof. We first show

$$\frac{d}{dt}\|e^{-\alpha\mathcal{E}}\|^2_{L^1(\rho_t)}\geq -b_1(\alpha)V(\rho_t)-b_2(\alpha)\lambda V(\rho_t)^{1/2} \tag{5.95}$$

where $b_1(\alpha),\ b_2(\alpha)$ are bounded in $[0,1]$ and tend to zero as $\alpha\to\infty$.

To see this, we apply Dynkin's formula to the function $\phi(v)=e^{-\alpha\mathcal{E}(v)}$. This yields

$$\begin{aligned}\frac{d}{dt}\int_{\mathbb{S}^{d-1}}e^{-\alpha\mathcal{E}(v)}d\rho_t(v)=&\int_{\mathbb{S}^{d-1}}\lambda P_{\mathbb{S}^{d-1}}(v)V^\alpha(\rho_t)\cdot\nabla_{\mathbb{S}^{d-1}}e^{-\alpha\mathcal{E}(v)}\\&+\frac{\sigma^2}{2}|v-V^\alpha(\rho_t)|^2\Delta_{\mathbb{S}^{d-1}}e^{-\alpha\mathcal{E}(v)}\;d\rho_t(v).\end{aligned} \tag{5.96}$$

The first term on the right hand side can be estimated as

$$\lambda\int_{\mathbb{S}^{d-1}}P_{\mathbb{S}^{d-1}}(v)V^\alpha(\rho_t)\cdot\nabla_{\mathbb{S}^{d-1}}e^{-\alpha\mathcal{E}(v)}\;d\rho_t(v) \tag{5.97}$$

$$=-\alpha\lambda\int_{\mathbb{S}^{d-1}}e^{-\alpha\mathcal{E}(v)}P_{\mathbb{S}^{d-1}}(v)V^\alpha(\rho_t)\cdot\nabla\mathcal{E}(v)d\rho_t(v) \tag{5.98}$$

$$\geq-\alpha\lambda c_1e^{-\alpha\underline{\mathcal{E}}}\int_{\mathbb{S}^{d-1}}|P_{\mathbb{S}^{d-1}}(v)V^\alpha(\rho_t)|\;d\rho_t(v) \tag{5.99}$$

$$\geq-\alpha\lambda c_1e^{-\alpha\underline{\mathcal{E}}}V(\rho_t)^{1/2} \tag{5.100}$$

where we have used the bound on $|\nabla\mathcal{E}|$ (see definition 3.1.1) and the estimate from lemma 5.4.1.

For the second term on the right hand side of (5.96) we get

$$\frac{\sigma^2}{2}\int_{\mathbb{S}^{d-1}} |v - V^\alpha(\rho_t)|^2 \Delta_{\mathbb{S}^{d-1}} e^{-\alpha\mathcal{E}(v)}\, d\rho_t(v) \tag{5.101}$$

$$\geq \frac{\sigma^2}{2}\int_{\mathbb{S}^{d-1}} |v - V^\alpha(\rho_t)|^2 \big(-\alpha c_2 - \alpha(d-1)c_1 - \alpha^2 c_1^2 - \alpha c_2\big) e^{-\alpha\mathcal{E}(v)}\, d\rho_t(v) \tag{5.102}$$

$$\geq -2\sigma^2\alpha e^{-\alpha\underline{\mathcal{E}}}(2c_2 + (d-1)c_1 + \alpha c_1^2)V(\rho_t) \tag{5.103}$$

where we have used the bounds from lemma 5.1.1 to find

$$\Delta_{\mathbb{S}^{d-1}} e^{-\alpha\mathcal{E}(v)} = \Delta e^{-\alpha\mathcal{E}(v)} - (d-1)v\cdot\nabla e^{-\alpha\mathcal{E}(v)} - v\otimes v : \nabla^2 e^{-\alpha\mathcal{E}(v)} \tag{5.104}$$

$$\begin{aligned} &= \big(\alpha^2|\nabla\mathcal{E}(v)|^2 - \alpha\Delta\mathcal{E}(v)\big)e^{-\alpha\mathcal{E}(v)} + \alpha(d-1)e^{-\alpha\mathcal{E}(v)} v\cdot\nabla\mathcal{E}(v) \\ &\quad + \alpha e^{-\alpha\mathcal{E}(v)}\big(v\otimes v : \big(-\alpha\nabla\mathcal{E}(v)\otimes\nabla\mathcal{E}(v) + \nabla^2\mathcal{E}(v)\big)\big) \end{aligned} \tag{5.105}$$

$$\geq \big(-\alpha\Delta\mathcal{E}(v) + \alpha(d-1)v\cdot\nabla\mathcal{E}(v) - \alpha^2|\nabla\mathcal{E}(v)|^2 - \alpha|\nabla^2\mathcal{E}(v)|\big)e^{-\alpha\mathcal{E}(v)} \tag{5.106}$$

$$\geq \big(-\alpha c_2 - \alpha(d-1)c_1 - \alpha^2 c_1^2 - \alpha c_2\big)e^{-\alpha\mathcal{E}(v)}. \tag{5.107}$$

Combining the inequalities (5.100) and (5.103) yields

$$\begin{aligned} \frac{1}{2}\frac{d}{dt}\|e^{-\alpha\mathcal{E}}\|^2_{L^1(\rho_t)} &= \|e^{-\alpha\mathcal{E}}\|_{L^1(\rho_t)}\frac{d}{dt}\|e^{-\alpha\mathcal{E}}\|_{L^1(\rho_t)} \\ &\geq -2\sigma^2\alpha e^{-2\alpha\underline{\mathcal{E}}}(2c_2 + (d-1)c_1 + \alpha c_1^2)V(\rho_t) - \alpha\lambda c_1 e^{-2\alpha\underline{\mathcal{E}}}V(\rho_t)^{1/2} \\ &=: -b_1(\alpha)V(\rho_t) - b_2(\alpha)\lambda V(\rho_t)^{1/2} \end{aligned} \tag{5.108}$$

where $b_1(\alpha), b_2(\alpha) \to 0$ as $\alpha\to\infty$, which finishes the proof of (5.95).

From this and definition 5.3.1 we follow

$$\|e^{-\alpha\mathcal{E}}\|^2_{L^1(\rho_t)} \geq \|e^{-\alpha\mathcal{E}}\|^2_{L^1(\rho_0)} - b_1(\alpha)\int_0^T V(\rho_s)ds - b_2(\alpha)\lambda\int_0^T V(\rho_s)^{1/2}ds \tag{5.109}$$

$$\geq \|e^{-\alpha\mathcal{E}}\|^2_{L^1(\rho_0)} - b_1(\alpha)\overline{\mathcal{V}}_T T - b_2(\alpha)\lambda\overline{\mathcal{V}}_T^{1/2}T \tag{5.110}$$

$$\geq \|e^{-\alpha\mathcal{E}}\|^2_{L^1(\rho_0)} - b_1(\alpha)\|e^{-\alpha\mathcal{E}}\|^2_{L^1(\rho_0)} - b_2(\alpha)\|e^{-\alpha\mathcal{E}}\|^2_{L^1(\rho_0)} \tag{5.111}$$

$$\geq \|e^{-\alpha\mathcal{E}}\|^2_{L^1(\rho_0)}(1 - b_1(\alpha) - b_2(\alpha)). \tag{5.112}$$

Hence

$$-\frac{1}{\alpha}\log\|e^{-\alpha\mathcal{E}}\|_{L^1(\rho_t)} \leq -\frac{1}{\alpha}\log\|e^{-\alpha\mathcal{E}}\|_{L^1(\rho_0)} - \frac{1}{2\alpha}\log\left(1 - b_1(\alpha) - b_2(\alpha)\right) \tag{5.113}$$

and with Laplace's principle and the fact that $b_1(\alpha)$ and $b_2(\alpha)$ approach zero as $\alpha\to\infty$ we further

see that for any $\epsilon > 0$ we can find $\alpha_0 = \max\{\alpha_1, \alpha_2\}$ such that

$$\begin{aligned} &\Big| -\frac{1}{\alpha}\log\|e^{-\alpha\mathcal{E}}\|_{L^1(\rho_t)} - \underline{\mathcal{E}}\Big| \\ &\quad\leq \Big| -\frac{1}{\alpha}\log\|e^{-\alpha\mathcal{E}}\|_{L^1(\rho_0)} - \underline{\mathcal{E}}\Big| + \Big|\frac{1}{2\alpha}\log\left(1 - b_1(\alpha) - b_2(\alpha)\right)\Big| \leq \epsilon \end{aligned} \tag{5.114}$$

which finishes the proof. □

We now prove the main theorem of this thesis. So far we have shown the monotonic decay of the variance of the mean-field solution. It remains to show that under certain conditions the correpsonding consensus point $V^\alpha(\rho_T)$ for the final time T is a good approximation of the global minimizer $V^\star$ we wish to find. We note that the approximation bounds in the theorem below are not sharp in general. In fact it seems like the estimates are only sharp for the special case $\epsilon_0 = 0$ and $\epsilon = 0$ (resp. $\alpha = \infty$).

In the next chapter we will see that from a numerical point of view the approximation properties of the KV-CBO method are much better than indicated by the estimates below. This gives hope that sharper estimates can be found.

Theorem 5.4.2. *There is an $\alpha_0 \gg 1$ such that the following approximation estimates hold*

$$\begin{cases} |\boldsymbol{E}(\rho_t) - V^\star| \leq c_3 2^{c_4-1}\big((2C_\alpha)^{c_4} V(\rho_t)^{c_4/2} + \varepsilon^{c_4}\big), \\ |V^\alpha(\rho_t) - V^\star| \leq 3V(\rho_t)^{1/2} + c_3 2^{c_4-1}\big((2C_\alpha)^{c_4} V(\rho_t)^{c_4/2} + \varepsilon^{c_4}\big) \end{cases} \tag{5.115}$$

for any $\alpha > \alpha_0$ and all $t \in [0, T]$.

Proof. With $e^{-\alpha\mathcal{E}(v)} \leq 1$ and the dual representation of the 1-Wasserstein distance W_1 (see (2.68)) we find

$$\left|\|e^{-\alpha\mathcal{E}}\|_{L^1(\rho_t)} - e^{-\alpha\mathcal{E}(\mathbf{E}(\rho_t))}\right| = \left|\int_{\mathbb{R}^d} e^{-\alpha\mathcal{E}(v)} d(\rho_t(v) - \delta_{\mathbf{E}(\rho_t)}(v))\right| \leq W_1(\rho_t, \delta_{\mathbf{E}(\rho_t)}). \tag{5.116}$$

With $W_1 \leq W_2$ and

$$W_2^2(\rho_t, \delta_{\mathbf{E}(\rho_t)}) \leq \int_{\mathbb{S}^{d-1}} |v - \mathbf{E}(\rho_t)|^2 \, d\rho_t(v) \overset{(5.57)}{\leq} \frac{8}{3} V(\rho_t) \tag{5.117}$$

this can further be bounded by

$$\left|\|e^{-\alpha\mathcal{E}}\|_{L^1(\rho_t)} - e^{-\alpha\mathcal{E}(\mathbf{E}(\rho_t))}\right| \leq 2V(\rho_t)^{1/2}. \tag{5.118}$$

We arrive at

$$\left|-\frac{1}{\alpha}\log\|e^{-\alpha\mathcal{E}}\|_{L^1(\rho_t)} - \mathcal{E}(\mathbf{E}(\rho_t))\right| = \frac{1}{\alpha}\left|\log\|e^{-\alpha\mathcal{E}}\|_{L^1(\rho_t)} - \frac{1}{\alpha}\log(e^{-\alpha\mathcal{E}(\mathbf{E}(\rho_t))})\right| \tag{5.119}$$

$$\leq \frac{e^{\alpha\overline{\mathcal{E}}}}{\alpha}\left|\|e^{-\alpha\mathcal{E}}\|_{L^1(\rho_t)} - e^{-\alpha\mathcal{E}(\mathbf{E}(\rho_t))}\right| \tag{5.120}$$

$$\leq 2C_\alpha V(\rho_t)^{1/2} \tag{5.121}$$

where we have used the Lipschitz continuity of the logarithm on $[1, \infty)$.

Now let $\epsilon > 0$. Lemma 5.4.4 implies the existence of some $\alpha_0 \gg 1$ such that

$$-\frac{1}{\alpha} \log \|e^{-\alpha \mathcal{E}}\|_{L^1(\rho_t)} - \underline{\mathcal{E}} \leq \epsilon \tag{5.122}$$

for all $\alpha > \alpha_0$. Combined with the bounds from above and the coercivity condition we find

$$\begin{aligned}
&|\mathbf{E}(\rho_t) - V^\star| \\
&\quad \leq c_3 \left|\mathcal{E}(\mathbf{E}(\rho_t)) - \underline{\mathcal{E}}\right|^{c_4} && (5.123)\\
&\quad \leq c_3 2^{c_4 - 1} \left| (\mathcal{E}(\mathbf{E}(\rho_t)) + \frac{1}{\alpha} \log \|e^{-\alpha \mathcal{E}}\|_{L^1(\rho_t)})^{c_4} + (\frac{-1}{\alpha} \log \|e^{-\alpha \mathcal{E}}\|_{L^1(\rho_t)} - \underline{\mathcal{E}})^{c_4} \right| && (5.124)\\
&\quad \leq c_3 2^{c_4 - 1} |(2C_\alpha)^{c_4} V(\rho_t)^{c_4/2} + \epsilon^{c_4}|. && (5.125)
\end{aligned}$$

To prove the second estimate we need to bound $|V^\alpha(\rho_t) - \mathbf{E}(\rho_t)|$.

We have

$$\begin{aligned}
|V^\alpha(\rho_t) - \mathbf{E}(\rho_t)|^2 &= \int_{\mathbb{S}^{d-1}} |V^\alpha(\rho_t) - v|^2 \, d\rho_t(v) + 2 \int_{\mathbb{S}^{d-1}} \langle V^\alpha(\rho_t) - v, v - \mathbf{E}(\rho_t) \rangle d\rho_t(v) \\
&\quad + \int_{\mathbb{S}^{d-1}} |v - \mathbf{E}(\rho_t)|^2 d\rho_t(v) && (5.126)\\
&\leq 4V(\rho_t) + 2V(\rho_t) + \frac{8}{3} V(\rho_t) && (5.127)\\
&\leq 9V(\rho_t) && (5.128)
\end{aligned}$$

where we have used

$$\begin{aligned}
|V^\alpha(\rho_t) - \mathbb{E}(\rho_t)|^2 &= \int_{\mathbb{S}^{d-1}} |V^\alpha(\rho_t) - v|^2 d\rho_t(v) + 2 \int_{\mathbb{S}^{d-1}} \langle V^\alpha(\rho_t) - v, v - \mathbb{E}(\rho_t) \rangle d\rho_t(v) \\
&\quad + \int_{\mathbb{S}^{d-1}} |v - \mathbb{E}(\rho_t)|^2 d\rho_t(v) && (5.129)\\
&\leq 4V(\rho_t) + 2(|\mathbb{E}(\rho_t)|^2 - 1) + 2V(\rho_t) && (5.130)\\
&= 4V(\rho_t). && (5.131)
\end{aligned}$$

and

$$\begin{aligned}
&2 \int_{\mathbb{S}^{d-1}} \langle V^\alpha(\rho_t) - v, v - \mathbf{E}(\rho_t) \rangle d\rho_t(v) \\
&\qquad = 2V^\alpha(\rho_t) \cdot \mathbb{E}(\rho_t) \big(1 - \frac{1}{|\mathbb{E}(\rho_t)|}\big) - 2 + 2\mathbb{E}(\rho_t) \cdot \mathbf{E}(\rho_t) && (5.132)\\
&\qquad \leq -|V^\alpha(\rho_t) - \mathbb{E}(\rho_t)|^2 \big(1 - \frac{1}{|\mathbb{E}(\rho_t)|}\big) - 2(1 - |\mathbb{E}(\rho_t)|) && (5.133)
\end{aligned}$$

$$\leq |V^\alpha(\rho_t) - \mathbb{E}(\rho_t)|^2 - 2V(\rho_t) \tag{5.134}$$
$$\leq 2V(\rho_t). \tag{5.135}$$

This finishes the proof. □

5.5 Summary

In this chapter we proved the main result of this thesis, namely the convergence of the consensus point $V^\alpha(\rho_t)$ towards the global minimizer $V^\star$ of the cost function for the special case $\Gamma = \mathbb{S}^{d-1}$. Here ρ_t is the solution of the mean-field equation of the KV-CBO model at time t. Important theoretical steps in the convergence proof were the establishment of mean-field solutions with L^2-regularity and the proof of Laplace's principle for those mean-field solutions. Further we proved technical estimates for the integral of the norm of the consensus point and expected value with which we proved the monotonic decay of the variance $V(\rho_t)$.

The convergence proof is a local optimization result in the sense that it requirs a-priori knowledge about the location of the global minimizer. We have discussed the conditions on the distribution and the coefficients of the KV-CBO model needed to achieve the convergence to the global minimizer.

Chapter 6

Applications

This chapter deals with the implementation of the KV-CBO model and discusses some real world applications. We test the resulting discrete *KV-CBO method* for different cost functions, among which we have the Rastrigin function which is a well-known benchmark function in non-convex optimization. We also test the method for three real world applications: robust computation of principal components, the phase retrieval problem, and the reconstruction of single-layer neural nets. The first problem is inherently defined on the sphere, whereas the latter two have to be reformulated as an optimization problem constrained over the sphere. We start the chapter with a discussion on implementation aspects of the KV-CBO model like discretization and possible speed-ups. In section 6.2 we make some general remarks about our testing methodology (including reproducibility) and further discuss numerical tests for the Rastrigin function. A Matlab implementation of the KV-CBO method can be download from GitHub [1]. In section 6.2 we will see that the anisotropic method is more stable than the isotropic method, see figure 6.2. In the subsequent numerial tests we will therefore only consider the anisotropic method.

This chapter is based on joint work of Massimo Fornasier, Hui Huang, Lorenzo Pareschi and myself and has been published in [Fornasier et al., 2021a] and [Fornasier et al., 2021b]. The adaptation of the KV-CBO method to find *multiple consensus points* is unpublished, see section 6.5.

6.1 Implementation aspects

In this section we discuss discretization of the KV-CBO model, some technical problems concerning sampling initial particles from a high-dimensional sphere and stable computation of the consensus point for large values of α. We further discuss an accelerated method and adaptivity of the parameters.

[1] https://github.com/PhilippeSu

Discretization

Let $t_n = n\Delta t$ an equidistant discretization of the time interval $[0, T]$ where Δt is a fixed time step. Throughout this chapter we denote by V_n^i the approximation of V_t^i at time $t_n = n\Delta t$. A one step time discretization of the KV-CBO model can be written in the general form

$$V_{n+1}^i = V_n^i + \Phi(\Delta t, V_n^i, V_{n+1}^i, \xi_n^i) \tag{6.1}$$

for a given discretization function Φ where ξ_n^i are independent random variables. In this thesis we only consider Brownian motions, that is, $\xi_n^i = \Delta B_n^i$ where $\Delta B_n^i \sim \mathcal{N}(0, \Delta t)$ iid. However the study of different stochastic processes could be interesting for future works.

As numerical efficiency of the method is crucial we choose a simple explicit discretization scheme. Explicit method have the disadvantage of not being *norm-preserving*, that is, $|V_{n+1}^i|^2 \neq |V_n^i|^2$ in general. In [Fornasier et al., 2021a, L. 2] it is shown that equality holds iff $\Phi(\Delta t, V_n^i, V_{n+1}^i, \xi_n^i)$ is orthogonal to $V_{n+1}^i + V_n^i$. If we use an explicit method we have to project the computed particle back to the hypersurface after each time step, see algorithm 1.

We discretize the isotropic KV-CBO model (3.43)

$$dV_t^i = \lambda P_\Gamma(V_t^i)V^\alpha(\rho_t^N)dt + \sigma|V_t^i - V^\alpha(\rho_t^N)|P_\Gamma(V_t^i)dB_t^i + C_\Gamma(V_t^i, V^\alpha(\rho_t^N))dt \tag{6.2}$$

with an (explicit) Euler-Maruyama scheme, see [Platen, 1999] and [Hairer et al., 2006]. In this case the function Φ is independent of V_{n+1}^i and takes the form

$$\begin{aligned}\Phi(\Delta t, V_n^i, \cdot, \Delta B_n^i) = \lambda\Delta t P_\Gamma(V_n^i)V_n^\alpha + \sigma|V_n^i - V_n^\alpha|P_\Gamma(V_n^i)\Delta B_n^i \\ - \frac{(d-1)\sigma^2\Delta t}{2}\frac{|V_n^i - V_n^\alpha|^2 V_n^i}{|V_n^i|^2}.\end{aligned} \tag{6.3}$$

For the anisotropic KV-CBO model (3.50) the discretization function Φ takes the form

$$\Phi(\Delta t, V_n^i, \cdot, \Delta B_n^i) = \lambda\Delta t P_\Gamma(V_n^i)V_n^\alpha + \sigma P_\Gamma(V_n^i)D(V_n^i - V_n^\alpha)\Delta B_n^i \tag{6.4}$$

$$- \frac{\sigma^2}{2}\Delta t\left(\frac{|V_n^i - V_n^\alpha|^2}{|V_n^i|^2} + \frac{D(V_n^i - V_n^\alpha)^2}{|V_n^i|^2} - 2\frac{|D(V_n^i - V_n^\alpha)V_n^i|^2}{|V_n^i|^4}\right)V_n^i. \tag{6.5}$$

Sampling the initial particles

A common probability distribution to sample from is the uniform distribution $\mathcal{U}(\mathbb{S}^{d-1})$. Sampling from this distribution is easy to implement. Let $\xi \sim \mathcal{N}(0, I_d)$ a sample from the d-dimensional standard normal distribution. Then the distribution of the normalized vector $\xi/|\xi|$ is uniform on $\mathbb{S}^{d-1}$. This is not obvious and we refer to [Muller, 1959], [Marsaglia et al., 1972] for the details.

Of course we can also choose a different initial distribution, e.g., a von Mises-Fisher distribution.

Stable computation of the consensus point

The naive computation of the consensus point is numerically unstable for large values of α. More precisely: the weights $e^{-\alpha\mathcal{E}(V_n^i)}$ are very close to zero for many particles, hence the denominator of the consensus point V_n^αwill be very close to zero

$$\sum_{i=1}^{N} e^{-\alpha\mathcal{E}(V_n^i)} \approx 0 \tag{6.6}$$

causing a blow up. This problem can be overcome with an elegant numerical trick (see Lemma 5.4.1 where we have used the same trick for $V^\alpha(\rho_t)$). Let $V_n^\star$ the particle with the lowest energy

$$V_n^\star = \arg\min_{i\in[N]} \mathcal{E}(V_n^i). \tag{6.7}$$

Now we can rewrite the consensus point in a numerically stable way as

$$V_n^\alpha = \frac{\sum_{i=1}^N V_n^i e^{-\alpha(\mathcal{E}(V_n^i)-\mathcal{E}(V_n^\star))}}{\sum_{i=1}^N e^{-\alpha(\mathcal{E}(V_n^i)-\mathcal{E}(V_n^\star))}}. \tag{6.8}$$

Since there is at least one particle with $V_n^i = V_n^\star$ the denominator is now bounded from below by ≥ 1. Hence there is no numerical instability anymore. We note that the computation of $V_n^\star$ has linear cost and thus does not affect the overall cost of the method. We note further that the efficieny of the KV-CBO method highly depends on the computational cost for evaluating the cost function $\mathcal{E}(V_n^i)$. An efficient implementation of the latter is of high importance.

Stopping criteria

The KV-CBO method is an iterative method and therefore we need to define and discuss reasonable stopping criteria. The simplest such stopping criteria is to set a maximum number of iterations n_T. However it might be smarter to stop the method earlier in order to save computational effort. Let $\epsilon > 0$ a given threshold which may be chosen as, say, $\epsilon = 10^{-4}$. Then the following stoping conditions make sense

1. maximum number of iterations reached	$n = n_T$
2. particles have nearly reached consensus	$\frac{1}{N}\sum_{i=1}^{N} \lvert V_n^i - V_n^\alpha\rvert \leq \epsilon$
3. variance of the particles is sufficiently small	$\frac{1}{N}\lvert V_n^i - \overline{V}_n\rvert^2 \leq \epsilon$
4. consensus point did not change much in the last p iterations	$\lvert V_{n+1}^\alpha - V_{n-p}^\alpha\rvert \leq \epsilon$
5. function value for the consensus point did not change much in the last p iterations	$\lvert\mathcal{E}(V_{n+1}^\alpha) - \mathcal{E}(V_{n-p}^\alpha)\rvert \leq \epsilon.$

Algorithm 1: general KV-CBO method on hypersurfaces

Input: $\mathcal{E}(\cdot)$, N, $\rho_0 \in \mathcal{P}(\Gamma)$, α, σ, Δt, *stopping criterion*

Set $n = 0$ (number of iterations)

Sample initial particles $V_0^i \sim \rho_0$ iid for all $i \in [N]$

while *stopping criterion is not met* **do**

  Increment iteration counter $n \leftarrow n + 1$

  Generate ΔB_n^i independent normal random vectors $\mathcal{N}(0, \Delta t)$

  Compute the consensus point V_n^α

  Evolve particles $\tilde{V}_{n+1}^i \leftarrow V_n^i + \Phi(\Delta t, V_n^i, V_{n+1}^i, \Delta B_n^i)$

  Project particles to the hypersurface $V_{n+1}^i \leftarrow \Pi_\Gamma(\tilde{V}_{n+1}^i)$

Output: Consensus point V_n^α

Accelerated method

The KV-CBO method requires a lot of computations. From the theoretical chapters we know that the approximation error decay as $1/N$ where N is the number of particles. Thus we would expect the number of particles to be rather large, however in practice the number of particles does not need to be excessively large. In section 6.3 e.g. we only use $N = 500$ particles (and a batch size $M = 50$) in dimension $d \approx 3000$. Nevertheless we want to find ways to make the method more efficient. We now discuss three strategies to do so: discarding particles, using random batches for the computation of the consensus point, and using parallelization.

It makes sense to start with a large number of particles $N_0 = N$. However, when the particles get closer and closer to each other (they are close to consensus) it does not make sense to evolve all N_0 particles. This leads to the idea of reducing the number of particles by discarding some of them depending on the empirical variance of the swarm. For the particles V^i for $i \in [N_n]$ we define the empirical variance

$$\Sigma_n = \frac{1}{N_n} \sum_{i=1}^{N_n} \left| V^i - \overline{V}_n \right|^2, \qquad \overline{V}_n = \frac{1}{N_n} \sum_{i=1}^{N_n} V^i. \tag{6.9}$$

We then choose the new number of particles N_{n+1} as

$$N_{n+1} = \max\{[\![N_n \left(1 + \mu \left((\Sigma_{n+1} - \Sigma_n)/\Sigma_n\right)\right)]\!],\ N_{min}\} \tag{6.10}$$

where $\mu \in [0, 1]$ and N_{min} are parameters to be chosen. For $\mu = 0$ we do not discard any particles. For $\mu = 1$ the number of particles discarded is maximal. In this case we also need to set a minimum number of particles. We should not discard particles in every iteration, but instead only , say, every $\ell = 10$ iterations to avoid fluctuations.

Another strategy to accelerate the method is to use random batches to compute the consensus

point V_n^α. Let J_M a random batch of size $M < N$, that is, a subset of M distinct indices from $\{1, ..., N\}$, where the *batch size* might be chosen as $M = N/5$. Now we compute the consensus point as

$$V_n^\alpha = \frac{\sum_{i \in J_M} V_n^i e^{-\alpha(\mathcal{E}(V_n^i) - \mathcal{E}(V_n^\star))}}{\sum_{i \in J_M} e^{-\alpha(\mathcal{E}(V_n^i) - \mathcal{E}(V_n^\star))}}. \tag{6.11}$$

Another way of making the method faster is by using parallelization, that is, distributing the evaluation of the particles to multiple processors that run in parallel. More precisely: once the consensus point V_n^α is available, the evaluation of the particles $V_n^i \mapsto V_{n+1}^i$ is *un*coupled and can thus be performed on different processors. After one time step the new particles need to be garthered in order to compute the new consensus point V_{n+1}^α. This last step can not be done in parallel.

Adaptive parameters

Before discussing ways how to adapt parameters let us first discuss a way to reduce the number of parameters. Consider the discretization (6.3). We note that the three computational parameter $\sigma, \lambda, \Delta t$ can be reduced if we rescale time as $\tau := \lambda \Delta t$. In this case $\sigma^2 \Delta t$ in the third term on the right hand side of (6.3) is given by $\sigma^2 \Delta t = \nu^2 \tau$ where $\nu^2 := \sigma^2 / \lambda$. We thus only have to choose the time scale τ and the diffusion parameter ν, that is, instead of the three parameters $\sigma, \lambda, \Delta t$ we only have to set two parameters ν, τ. For simplicity we keep the original notation and set

$$\lambda := 1 \tag{6.12}$$

in (6.3). Nevertheless, the choice of the parameters in the KV-CBO method (as for any metaheuristic) is still tricky.

Now we discuss a few techniques to adapt the parameters α and σ. We keep the time step constant and note that a small time step makes the particles slower, a large time step makes them faster.

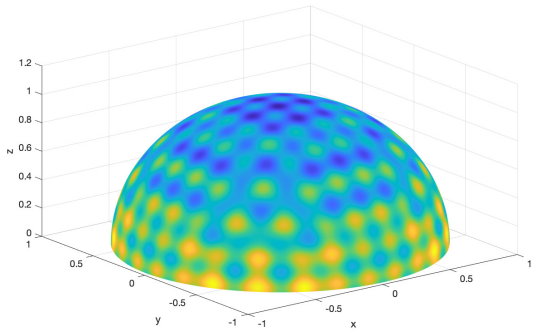


Figure 6.1: The Rastrigin function on $\mathbb{S}^2$. The global minimizer $V^\star$ is at the north pole.

Let us make a few general remarks about adaptivity of parameters. First, any adaptive method

requires the choice of initial values. Finding these initial values is not at all obvious and requires yet another heuristic. Second, one can invent uncountable heuristics to achieve a reduction or increase of a certain parameter. Any of these heuristics may work well in one case, but may fail for another. Some heuristics may require an upper/ lower bound which adds yet another level of complexity. Third, one has to decide how frequent one wants to adapt a certain parameter. One could adapt the parameter every single iteration or every, say, 10th iteration. Fourth, how many of the parameters do we want to adapt? Only the α? or only the σ? or maybe both? This is a highly compley problem as it is not clear how these parameters influence each other. Fifth, in principle one can distinguish between adapting a parameter based on its value in the current time step or based on a given statistic of the swarm (e.g. the variance)? Sixth, one can choose to adapt a given parameter in a deterministic way or randomly. E.g. one may choose the σ as a random variable with decreasing variance.

In short: the problem of how to choose the parameters is open and a lot more theoretical understanding is needed to answer it satisfactorily. In practice usually the simple schemes work best.

Let us now discuss practical methods to adapt parameters. From the theoretical chapters we know that the diffusion parameter σ has to be small to achieve consensus. It therefore makes sense to start with a relatively large σ_0 which we adaptively reduce until a certain minimum value $\sigma_{min} \geq 0$ is reached. Possible strategies include

$$\begin{cases} \sigma_{n+1} = \max\{\sigma_n/\tau,\ \sigma_{min}\}, \\ \sigma_{n+1} = \max\{\sigma_n/(\sigma_0 \log(n+1)),\ \sigma_{min}\} \end{cases} \tag{6.13}$$

where $\tau > 1$ is a given parameter, say, $\tau = 1.05$.

For the parameter α we know from the theoretical part that it should be large to achieve a good accuracy. Possible strategies for adapting the α include

$$\alpha_{n+1} = \frac{(n+1)\alpha_{max}}{n_T}. \tag{6.14}$$

We note that a too large α might not be desirable as it makes the method a *greedy method*: for large α the consensus point is essentially the particle with the lowest cost $V^\alpha = \arg\min_{i\in[N]} \mathcal{E}(V^i)$.

Gradient-KV-CBO method

It is possible to combine the KV-CBO method with the Gradient Descent method. By doing so the method is no longer a zero order method. In [Fornasier et al., 2021b] the following (computationally trackable) idea is proposed: randomly choose one particle every ℓth iteration and perform a gradient descent search with this particle, where the step size h_n is chosen with a *backtracking line search*

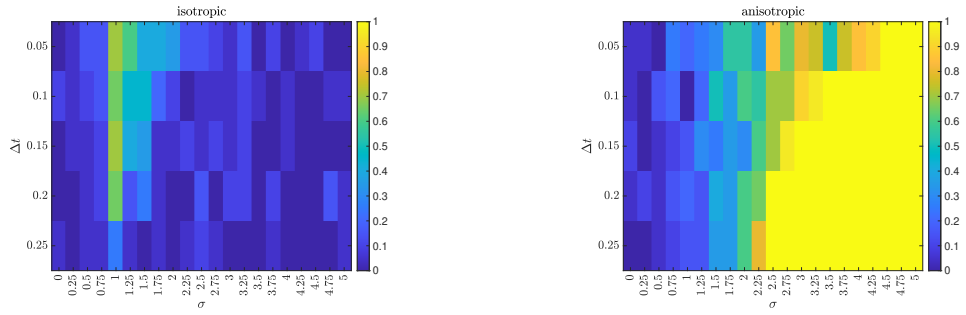


Figure 6.2: Isotropic vs. Anisotropic noise: phase transition matrix for the Rastrigin function in $d = 5$, see `experiment_isoVsAniso` folder.

method which is iterated until the *Armijo condition* or *sufficient decrease conditon*

$$\mathcal{E}(v_n - h_n^0 \nabla_{\mathbb{S}^{d-1}} \mathcal{E}(v_n)) \leq \mathcal{E}(v_n) - c h_n^0 |\nabla_{\mathbb{S}^{d-1}} \mathcal{E}(v_n)|^2 \tag{6.15}$$

is met, see [Absil et al., 2008].

The rationale is that a single particle can, at best, find a *local* minimizer in one time step. With the backtracking approach from above we expect the particle to arrive at this local minimizer. The cost for this adaptation is negligible. The benetit is that we may arrive at minimizer we may not have found otherwise.

We may choose the initial step $h_n^0 = 1$ and $c = 10^{-4}$. If the condition is satisfied we set $h_n = h_n^0$, otherwise we set $h_n^1 = \tau h_n^0$ with, say, $\tau = 1/2$ and check again whether the condition is satisfied.

6.2 The KV-CBO package

In this section we discuss the KV-CBO package [2] and make some general remarks about numerical tests and performance measures used. We consider the Rastrigin function

$$\mathcal{E}_R(v) = Ad + |v|^2 - A \sum_{i=1}^{d} \cos(2\pi v_i) \tag{6.16}$$

for $v \in \mathbb{R}^d$ (see page 39 for a graph of $\mathcal{E}_R$ on the circle $\mathbb{S}^1$) to illustrate how the KV-CBO package can be used. The package however contains many more common benchmark functions (Rastrigin, Ackley, Alpine, Schaffer, Solomon, Lévi, XSY random & Griewank) as well as cost functions from the real world applications that we discuss in the following sections. We choose the Rastrigin function as main benchmark function because it is highly complex and can be shown to satisfy all the conditions on the cost function from defintion 3.1.1.

[2] https://github.com/PhilippeSu

Algorithm 2: Variant of the isotropic KV-CBO method on the sphere

Input: $\mathcal{E}(\cdot)$, N, M, α, σ, Δt, μ, *stopping criterion*

Set $N_0 = N$ (number of particles) and $n = 0$ (number of iterations)

Sample initial particles $V_0^i \sim \mathcal{U}(\mathbb{S}^{d-1})$ iid for $i \in [N_0]$

Compute the variance Σ_0 of V_0^i

while *stopping criterion is not met* **do**

  Increment iteration counter $n \leftarrow n + 1$

  Generate ΔB_n^i independent normal random vectors $\mathcal{N}(0, \Delta t)$

  if $M \leq N_n$ **then**

    Select a random batch $J_M \subset \{1, ..., N\}$ with $|J_M| = M$

    Compute V_n^α as in (6.11)

  else

    Compute V_n^α as in (6.8)

  Use Euler-Maruyama method to evolve the particles

$$\tilde{V}_{n+1}^i \leftarrow V_n^i + \lambda \Delta t P_\Gamma(V_n^i) V_n^\alpha + \sigma |V_n^i - V_n^\alpha| P_\Gamma(V_n^i) \Delta B_n^i - \frac{(d-1)\sigma^2 \Delta t}{2} \frac{|V_n^i - V_n^\alpha|^2 V_n^i}{|V_n^i|^2}$$

  Project particles to the sphere $V_{n+1}^i \leftarrow \tilde{V}_{n+1}^i / |\tilde{V}_{n+1}^i|$

  Compute the variance Σ_{n+1} of V_{n+1}^i

  Update the number of particles $N_{n+1} \leftarrow [\![N_n \left(1 + \mu \left((\Sigma_{n+1} - \Sigma_n)/\Sigma_n \right) \right)]\!]$

  Discard $N_n - N_{n+1}$ samples (chosen uniformly at random)

  (optional) adapt parameters

  (optional) use a gradient descent step

Output: Consensus point V_n^α

Lemma 6.2.1 (Properties of Rastrigin Function)**.** *The following properties hold*

1. *the global minimizer of* $\mathcal{E}_R$ *is* $V^\star = 0$ *and* $\mathcal{E}_R(V^\star) = 0$,

2. $\mathcal{E}_R$ *is symmetric* $\mathcal{E}_R(-v) = \mathcal{E}_R(v)$,

3. *we have*

$$|v - V^\star|^2 \leq |\mathcal{E}_R(v) - \mathcal{E}_R(V^\star)| \leq 2\sqrt{d}(a + A\pi)|v - V^\star| \tag{6.17}$$

   for all $v \in \mathbb{R}^d$.

Proof. We observe that

$$\mathcal{E}_R(v) = |v|^2 + A(\sum_{i=1}^{d} 1 - \cos(2\pi v_i)) \geq |v|^2 \geq 0 \tag{6.18}$$

and $\mathcal{E}_R(0) = 0$. Thus, the global minimizer of $\mathcal{E}_R$ is attained at the origin. With the mean value theorem we further have

$$\mathcal{E}_R(v) - \mathcal{E}_R(v^\star) = \nabla \mathcal{E}_R(c) \cdot (v - v^\star) \tag{6.19}$$

for some $c \in [v^\star, v]$. We can bound the norm of $|\nabla \mathcal{E}_R(c)|$ as

$$|\nabla \mathcal{E}_R(c)| = |2c + 2\pi A \sin(2\pi c)| \leq 2\sqrt{d}a + 2\pi A\sqrt{d} \tag{6.20}$$

which concludes the proof. □

Let us start with a few general remarks about the numerical tests in this chapter.

In all the numerical tests we know the location of the global minimizer $V^\star$. We can thus evaluate if the KV-CBO method was successful for a given cost function and a given set of parameters by checking whether or nor the distance of the final consensus point $V_{n_T}^\alpha$ to $V^\star$ is smaller than a given tolerance ϵ. This of course raises again the question of which error norm and tolerance to choose. In this chapter we will always use the 2-norm $|\cdot|$ and the tolerance is $\epsilon = 0.05$, that is, we call a run successful if

$$|V_{n_T}^\alpha - V^\star| \leq 0.05. \tag{6.21}$$

Here $V_{n_T}^\alpha$ always denotes the final consensus point, even thought the maximum number of iterations n_T may not have been reached because another stopping criterion was satisfied.

As the success of the method depends on many parameters it makes sense to compare the success rates for different sets of parameters. For example: we may plot the average error in time for different numbers of particles N. We may also compute so-called *phase transition matrices* to easily visualize the success rates for two varying parameters, say, the time step Δt and the diffusion parameter σ, see figure 6.2 .

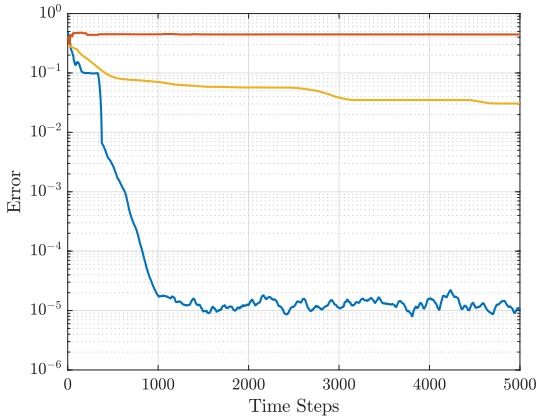


Figure 6.3: Rastrigin function in $d = 20$ for $N = 100$ particles and 100 runs. Average error in yellow, error for the run with the best/ worst final error in blue/ red, see `experiment_Error_Time` folder.

Some remarks on the reproducibility of the numerical results.

All of the numerical tests have been performed on a MacBook Pro from 2017 with an Intel Core i7 processor running under macOS Catalina (10.15.7, 64 Bit). We used Matlab version *9.10 (R2021a)* with the following toolboxes: *Image Processing Toolbox (Version 11.3)*, *Signal Processing Toolbox (Version 8.6)* and *Statistics and Machine Learning Toolbox (Version 12.1)*. The code can be downloaded from GitHub *https://github.com/PhilippeSu/KV-CBO/PhD*. Whenever possible we have included the exact *m-files* needed to reproduce the numerical experiments in this chapter. Hence, the reader can reproduce the results without having to do any/ much programming herself. For example: the plot from figure 6.2 can be reproduced by running the script in the folder `experiment_isoVsAniso` as indicated in the caption.

rotation angle ϑ		$k = 0$	$k = 100$
$\vartheta = 0$	success rate	100%	100%
	avg error	1.5e-6	1.0e-6
	std. dev.	4.47e-7	6.1e-7
$\vartheta = \pi/4$	success rate	50%	100%
	avg error	8.8e-2	0
	std. dev.	0.945	0

Table 6.1: Rastrigin function in $d = 5$ with the global minimizer $V^\star$ rotated by ϑ in all directions from the north pole. k is the parameter from the von Mises-Fisher distribution, see `experiment_rotatedMin`.

Let us now discuss the main components of the software package. The user should first execute the m-file `import_kvcbo.m` which adds the path to Matlab. The main code can be found in the folder `kvcbo`, the main files are `KVCBO.m` and `KuramotoIteration.m`. The folder `cost_functions` contains the implementation of some benchmark cost functions (all have a separate folder) and the file `setUpClass.m`. The folder `ext` contains third party scripts. The `experiments` folders contain the scripts to reproduce the numerical experiments from this chapter.

Further, the folder `cost_functions/Rastrigin` contains two files: `KVCBOparamForRastrigin.m` and `RastriginCostFunction.m`. The former contains the parameters for the KVCBO method for the Rastrigin function and the latter contains the implementation of the Rastrigin function, that is, (6.16). The code can be used in the following way.

First, we call

```
[costfunction, KVCBOparam] = setUpClass.Rastrigin(d);
```
(6.22)

for a given dimension d. This returns the implemenation of the cost function and the list of parameters.

Second, we call the KV-CBO method

```
[Va, info] = KVCBO(costfunction, KVCBOparam);
```
(6.23)

This returns the final consensus point `Va` $\in \mathbb{R}^d$ and the list `info` with relavant statistics for each iteration of the KV-CBO method. This list can then be used to plot many statistics (e.g. error, variance) of the run, see `example1.m`.

Remark 6.2.1. *The scripts `HyperSphere.m` and `InvHyperSphere.m` return the spherical/ euclidean coordinates of the vector we input. The spherical coordinates $\vartheta \in \mathbb{R}^{d-1}$ for a vector $v \in \mathbb{R}^d$ are given by*

$$\begin{cases} v_1 & = r\cos(\vartheta^1) \\ v_2 & = r\sin(\vartheta^1)\cos(\vartheta^2) \\ & \vdots \\ v_{d-1} & = r\sin(\vartheta_1)\cdot\dots\cdot\sin(\vartheta_{d-2})\cos(\vartheta_{d-1}) \\ v_d & = r\sin(\vartheta_1)\cdot\dots\cdot\sin(\vartheta_{d-1}). \end{cases} \tag{6.24}$$

We conclude this section with a list of parameters of the KV-CBO method and a short explanation.

Parameter	**Possible assignment**	**Explanation**
N	integer	number of particles
M	$M \leq N$	batch size
N_{min}	$0 < N_{min} \leq N$	minimum number of particles, see (6.10)
μ	$0 \leq \mu \leq 1$	for the discaring of particles, see (6.10)
α	$\alpha \gg 1$	temperature parameter for V^α
mode	'anisotropic', 'isotropic'	mode for the noise level
σ	$\sigma > 0$	diffusion parameter
dt	$dt > 0$	time step Δt
λ	$\lambda > 0$	coupling constant, fixed to $\lambda = 1$, see (6.12)
$logicVa$		list of parameters to compute the consensus point
$initData$		list of parameters for the initial particles
$stopCond$		list of parameters for the stopping condition
out		list of parameters to control output
adapt		list of parameters to control adaptation of computation parameters, see (6.13) and (6.14)

Table 6.2: Parameters for the KV-CBO method, see `KVCBOparamForRastrigin.m`

6.3 Robust principal component analysis

In this section we consider the first of three real world applications of the KV-CBO method. Robust PCA (principal component analysis) is also known as *robust subspace fitting* and *robust subspace detection*. To explain the idea let $\mathcal{Q} = \{x^{(i)} \in \mathbb{R}^d : i = 1, ..., \mathcal{P}\}$ a centered point cloud in the Euclidean space $\mathbb{R}^d$. The goal of PCA is to find the $n(< d)$ directions of maximal variation of the point cloud $\mathcal{Q}$. In other words: the n directions that capture the geometry of the point cloud best. We consider only the case $n = 1$ here, but note that extending the method to find higher dimensional subspaces $n > 1$ is a highly interesting open problem.

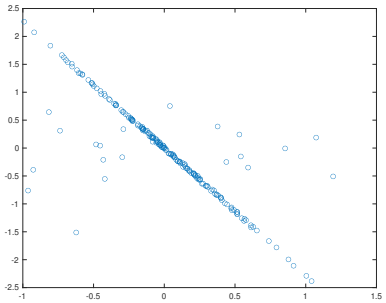
For $n = 1$ Pythagoras theorem yields

$$|x^{(i)}|^2 = |P_v x^{(i)}|^2 + |P_{v^\perp} x^{(i)}|^2 \tag{6.25}$$

where $P_v = vv^T$ and $P_{v^\perp} = I - vv^T$, see figure 6.4. This shows that maximizing $|vv^T x^{(i)}|^2$ is equivalent to minimizing $|P_{v^\perp} x^{(i)}|^2$. We thus arrive at the following constrained minimization problem

$$V^\star = \arg \min_{v \in \mathbb{S}^{d-1}} \sum_{i=1}^{\mathcal{P}} |P_{v^\perp} x^{(i)}|^2 \tag{6.26}$$

where $V^\star$ is the principal component we wish to find. Note that this optimization problem is naturally constrained on the sphere.



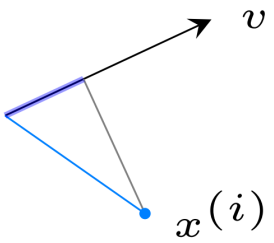


Figure 6.4: Left: point cloud with a clearly visible principal component and a few outliers. Right: schematic depiction of (6.25). The blue dot is one ot the data points $x^{(i)}$, the blue shaded resp. gray line has length $|P_v x^{(i)}|$ resp. $|P_{v^\perp} x^{(i)}|$.

The above standard approach however has a shortcoming. To illustrate this shortcoming let $x^{out} \in \mathcal{Q}$ be an outlier in the point cloud $\mathcal{Q}$, that is, a point for which the distance $|P_{v^\perp} x^{out}|$ is much larger than for all (or most) of the other points in $\mathcal{Q}$. Now, according to the above optimization problem the distance $|P_{v^\perp} x^{out}|$ is squared which makes it even larger. This is a problem because now the contribution of this one outlier x^{out} to the cost function is very strong. The principal component will be tilted towards this one outlier and thus not describe the true geometry of the point cloud. In other

words: the quadratic weight function weights the outlier much stronger than the inliers, resulting in a principal component that does not capture the geometry of the point cloud.

This is the motivation to consider the *robust* PCA problem. We want to find a way to weight the outliers less strongly. There are many methods that try to solve this problem. One of the best is the *Fast Median Subspace method* (FMS), see [Lerman and Maunu, 2017]. This method will be the benchmark method for the KV-CBO method. The robust PCA method that we propose is given by

Algorithm 3: robust PCA with KV-CBO

Input: $p \in (0,2]$, $\mathcal{Q} = \{x^{(i)} \in \mathbb{R}^d \mid i \in [\mathcal{P}]\}$, parameters for KV-CBO method

1 Set

$$\mathcal{E}_p(v) := \sum_{i=1}^{\mathcal{P}} |(I - v \otimes v)x^{(i)}|^p = \sum_{i=1}^{\mathcal{P}} \left(|x^{(i)}|^2 - |\langle x^{(i)}, v\rangle|^2\right)^{p/2} \tag{6.27}$$

2 $V_{n_T}^{\alpha} \leftarrow$ KV-CBO($\mathcal{E}_p$)

Output: $V_{n_T}^{\alpha}$ (Guess for the principal component of $\mathcal{Q}$)

We note that for $p < 1$ and a point cloud with many outliers (6.27) is a difficult non-convex cost function, see figure 6.5. The cost function $\mathcal{E}_p$ is not differentiable for $p < 2$, hence the assumptions for the KV-CBO method are not satisfied and we would need to consider the cost function $\mathcal{E}_{p,\delta}(v) = \sum_{i=1}^{\mathcal{P}} \left(|x^{(i)}|^2 - |\langle x^{(i)}, v\rangle|^2 + \delta\right)^{p/2}$ for some small $\delta > 0$ instead. We however choose $\delta = 0$ and stick to the original cost function as the differentiablity of the cost function is important only in the theoretical developments not for the numerical implementation.

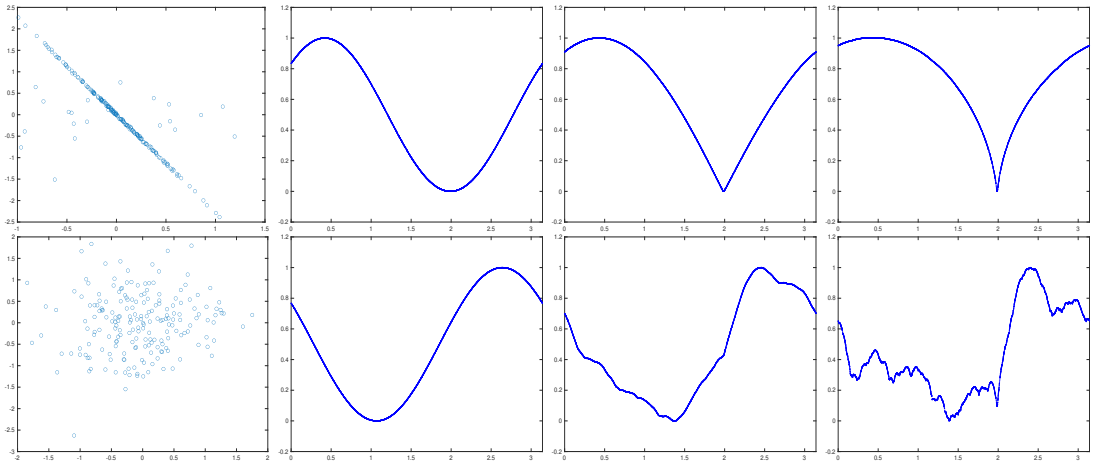


Figure 6.5: $\mathcal{P} = 200$ data points sampled from the *Haystack model*. Top: 5% outliers and the corresponding cost function for $p = 2, 1, \frac{1}{2}$. Bottom: 95% outliers. Increasing the number of outliers and reducing the p makes the cost function very rough. Note also how the position of the global minimizer is shifted as we change p.

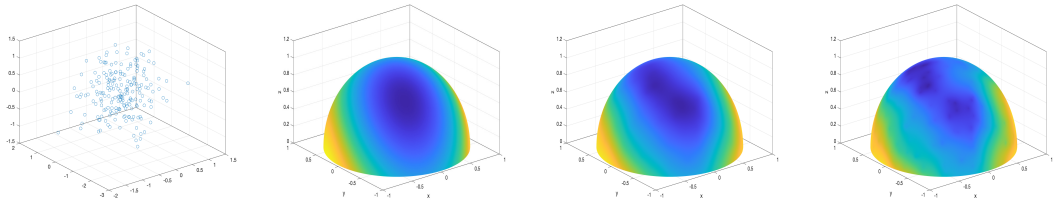


Figure 6.6: Same setup as above in $d = 3$ with 95% outliers

Artificial data from Haystack model

In this section we analyse algorithm 3 for a point cloud sampled from the *Haystack model*, see [Lerman et al., 2015] and compare the results with popular benchmark methods: FMS (Fast Median Subspace), SVD (Singular Value Decomposition) and GD (Gradient descent for (6.27)). We note that the output of algorithm 3 is one (!) vector in $\mathbb{R}^d$, hence we can only expect to recover one dimensional subspaces. Recovering higher dimensional subspaces with the KV-CBO is an interesting open problem.

Let us first discuss the Haystack model to generate the data. We distinguish between *inliers/outliers* and *cellwise contamination/ casewise contamination*. Cellwise contamination represents the *noise* in the inliers that causes them to not lie exaclty on the 1-dimensional subspace we wish to recover. Casewise contamination on the other hand represents true *outliers*, that is, data points that could be scattered all around the ambient Euclidean space $\mathbb{R}^d$.

We choose the subspace W to recover uniformly at random, that is, $W \sim \mathcal{U}(\mathbb{S}^{d-1})$ and sample the inliers from a Gaussian with rank-1 covariance matrix $\mathbf{\Sigma}_{in} = W \otimes W$ which we then perturb with Gaussian noise in the ambient space $\mathbb{R}^d$

$$x_{in}^{(i)} \sim \mathcal{N}(\mathbf{0}, \mathbf{\Sigma}_{in} + 10^{-4}\mathbf{I}_d) \tag{6.28}$$

for $i = 1, ..., \mathcal{P}_{in}$. The outliers are sampled from a Gaussian with a full-rank covariance matrix

$$x_{out}^{(i)} \sim \mathcal{N}(\mathbf{0}, \mathbf{I}_d/d) \tag{6.29}$$

for $i = 1, ..., \mathcal{P}_{out}$. This choice for the covariance matrix implies $\mathbb{E}|x_{in}^{(i)}|^2 \approx \mathbb{E}|x_{out}^{(i)}|^2$. Hence we can not detect the outliers by computing the norm of the data points and disguarding those data points that have a suspiciously large norm. The point cloud $\mathcal{Q}$ is thus given by

$$\mathcal{Q} := \{x_{in}^{(i)} \mid i = 1, ..., \mathcal{P}_{in}\} \cup \{x_{out}^{(i)} \mid i = 1, ..., \mathcal{P}_{out}\}. \tag{6.30}$$

As a first numerical test we consider a point cloud $\mathcal{Q}$ without contamination, that is, no noise in the inliers and no outliers. In this case the (unknown) subspace W can be reconstructed with a simple application of SVD. In table 6.3 we report the success rate, the average error and the standard

deviation for various methods. For KV-CBO and GD we used the cost function (6.27) with $p = 2$. The ambient dimension is $d = 100$ and the results are averaged over 20 runs. For GD we used a very simple implementation: $x^k \leftarrow x^k - \eta \nabla_{\mathbb{S}^{d-1}} \mathcal{E}_2(x^k)$ for $\eta = 0.001$, followed by a normalization. We see that all 4 methods can reconstruct the unknown subspace W. The accuracy of the reconstruction of FMS [3] and SVD are the highest, for the KV-CBO method it is the lowest.

Remark 6.3.1. *It is important to note that the KV-CBO method uses many more function evaluation than GD. Roughly, the former uses $\mathcal{O}(M \times n_T)$ function evaluations whereas the latter uses $\mathcal{O}(2n_T)$. Therefore, the two methods should not be compared against each other in terms of efficiency. As for the effectivness: we refer to table 6.4 for numerical evidence of the higher effectiveness of the KV-CBO method over GD.*

		KV-CBO	**FMS**	**SVD**	**GD**
0% outliers	success rate	100%	100%	100%	100%
no noise	avg/ std	2.3e-3/ 2.3e-3	4.8e-16/1e-16	4.8e-16/1e-16	1.1e-6/ 2.6e-7

Table 6.3: Average distance to W for a point cloud $\mathcal{Q}$ without contamination, see folder `experiment_robustPCAHaystack`.

As a second numerical test we consider the point cloud $\mathcal{Q}$ with contamination. We set the total number of points to $\mathcal{P} = \mathcal{P}_{in} + \mathcal{P}_{out} = 200$ and chose the number of outliers $\mathcal{P}_{out}$ as a certain percentage of $\mathcal{P}$ ranging from 0% to 95%. Again, the ambient dimension is $d = 100$, and the results are averaged over 20 runs, see table 6.4. For the point cloud with 95% outliers only the KV-CBO method is able to reconstruct the vector W reliably.

Real data: robust computation of eigenfaces

In this section we apply algorithm 3 to real data. The point cloud $\mathcal{Q}$ is now a subset of $\mathcal{P} = 420$ pictures of faces from the *10K US Adult Faces* database [Bainbridge et al., 2013]. Each picture is rescaled to 64×45 pixels and converted to grayscale, hence the ambient dimension is $d = 2880$. The point cloud is centered (subtraction of the mean). The principal direction of a point cloud of faces is (after resizing to 64×45) a so-called *eigen*face (similarly to *eigen*vector). Roughly speaking, an eigenface is the *average face* of a point cloud of faces. An eigenface does not look like a real face of a person, but it nevertheless has clearly visible features of a face (eyes, nose, ears, mouth). In figure 6.8 we report the eigenfaces couputed with different methods.

It should be noted that the faces from this database do not (!) lie close to a 1-dimensional subspace. They are scattered much more in the ambient space $\mathbb{R}^d$. In other words: the situation in this example is not comparable to the situation depicted in figure 6.5 (top row).

[3] https://twmaunu.github.io/FMS/ (last accessed 09.12.2021)

		KV-CBO	FMS	SVD	GD
0% outliers	success rate	100%	100%	100%	0%
	avg/ std	3.1e-3/ 8.7e-3	7.0e-4/ 5.9e-5	7.0e-4/ 5.9e-5	1.1e-1/ 2.3e-3
25% outliers	success rate	100%	100%	100%	0%
	avg/ std	1.1e-3/ 3.8e-4	4.7e-3/ 1.0e-3	4.7e-3/ 1.0e-3	9.2e-2/ 2.8e-3
50% outliers	success rate	95%	100%	100%	0%
	avg/ std	5.1e-3/ 1.6e-2	1.0e-2/ 1.03-3	1.0e-2/ 1.03-3	6.9e-2/ 5.0e-3
75% outliers	success rate	100%	100%	100%	35%
	avg/ std	1.6e-3/ 3.1e-4	2.5e-2/ 4.8e-3	2.5e-2/ 4.8e-3	2.5e-1/ 3.1e-1
95% outliers	success rate	100%	0%	0%	15%
	avg/ std	6.4e-3/ 6.7e-3	2.0e-1/ 1.2e-1	2.0e-1/ 1.2e-1	1.4e-1/ 9.1e-2

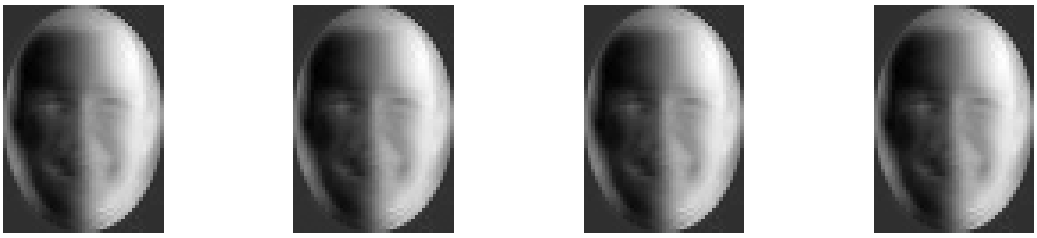
Table 6.4: Average distance to W for a point cloud with contamination. For KV-CBO and GD we considered the cost function (6.27) with $p = 1/2$, see folder `experiment_robustPCAHaystack`.

In [Lerman et al., 2015, Faces in a Crowd] and [Basri and Jacobs, 2003] it is argued that the dimension of the best fitting subspace should be chosen as 9 (and not 1). We nevertheless apply algorithm 3 to compute the eigenface, which we report in figure 6.8. As in the previous numerical tests we also compute the eigenface with SVD, FMS and GD. For GD we use the ManOpt [4] implementation [Boumal et al., 2014].

Remark 6.3.2 (Reproducibility)**.** *The eigenfaces reported in this thesis are not exactly reproducable by the reader. Even though the 10K US Adult Faces database is available to anyone upon request, the precise subset of $\mathcal{P} = 420$ pictures that we chose are not available to the reader. Choosing a different subset of faces will yield (slightly) different results.*

Figure 6.7: Some faces from the *10K US Adult Faces* database and one outlier.

[4] https://www.manopt.org/



(a) (b) (c) (d)

Figure 6.8: Eigenfaces without outliers: (a) SVD, (b) KV-CBO (algorithm 3 with $p = 2$), (c) FMS, (d) GD for (6.27) with $p = 2$, see folder `experiment_robustPCAEigenfaces`.

The eigenfaces computed with the various methods are visually inseperable. We quantify the similarities with different well-known measures: *peak signal-to-noise ratio* (PSNR), *signal-to-noise ratio* (SNR), and *structural similarity index* (SSIM). The former two measures are expressed in decibel (dB), the latter is a number in $[0, 1]$. The quality is higher if the measures are larger. The PSNR for the pictures $A, B \in \mathbb{R}^{n\times m}$ is defined as

$$\mathrm{PSNR}(A, B) = 10 \log_{10} \left(\frac{\max(A) \max(B)}{\mathrm{MSE}(A, B)} \right) \tag{6.31}$$

where A is the reference picture, and MSE is the mean-squared error defined as

$$\mathrm{MSE}(A, B) = \frac{1}{nm} \sum_{i=1}^{n} \sum_{j=1}^{m} |A_{ij} - B_{ij}|^2. \tag{6.32}$$

The SNR is defined as

$$\mathrm{SNR}(A, B) = 10 \log_{10} \left(\frac{|A|^2}{|B|^2} \right). \tag{6.33}$$

The SSIM is defined as

$$\mathrm{SSIM}(A, B) = \ell(A, B) c(A, B) s(A, B) \tag{6.34}$$

where

$$\begin{cases} \ell(A, B) & = 2\mu_A \mu_B / (\mu_A^2 + \mu_B^2), \\ c(A, B) & = 2\sigma_A \sigma_B / (\sigma_A^2 + \sigma_B^2), \\ s(A, B) & = \sigma_{AB} / (\sigma_A \sigma_B) \end{cases} \tag{6.35}$$

and μ_A, μ_B are the means of A, B respectively, σ_A, σ_B are the standard deviations, and σ_{AB} is the covariance of A and B. The function ℓ is the luminance distortion, c is the contrast distortion, and s the structure comparison function, see [Hore and Ziou, 2010].

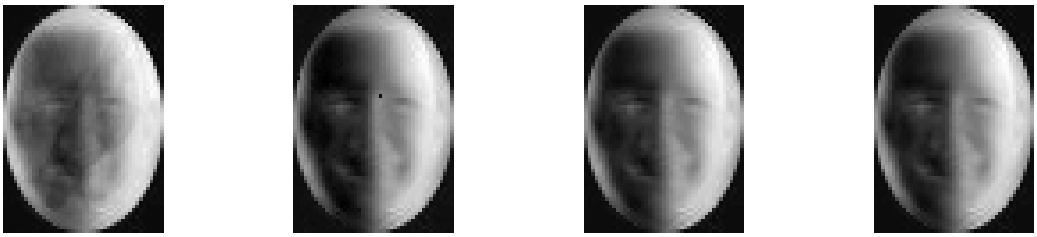
The reference image is always the eigenface for the point cloud without outliers computed with SVD,

that is, picture (a) in figure 6.8. The PSNR and SNR ratio of a picture with itself is ∞. The SSIM of a picture with itself is 1. The image quality is considered excellent if the PSNR $> 50dB$.

		SVD	**KV-CBO**	**FMS**	**GD**
no outliers	PSNR	∞	55.5564	∞	∞
	SNR	∞	49.2491	∞	∞
	SSIM	1	0.9994	1	1

Table 6.5: Some well-known quality measures for the eigenfaces from figure 6.8.

We now add some outliers to the point cloud of faces and again compute the eigenfaces. The results are reported in figure 6.9 and in table 6.6.



(a) (b) (c) (d)

Figure 6.9: Eigenfaces with 5 outliers: (a) SVD, (b) KV-CBO (algorithm 3 with $p = 1/2$), (c) FMS, (d) GD for (6.27) with $p = 1/2$, see folder `experiment_robustPCAEigenfaces`. We note the black spot between the eyes of eigenface (b).

6.4 The phase-retrieval problem

The phase retrieval problem is an important problem in applied physics, engineering and mathematics. The available literature on the topic is as impressive as the list of applications. Important applications of the phase retrieval problem include optical imaging [Shechtman et al., 2015], [Walther, 1963], astronomy [Dainty and Fienup, 1987], crystallography [Elser et al., 2018], speech recognition [Balan et al., 2006], short-time Fourier transform analysis [Jaganathan et al., 2016], quantum mechanics [de Gosson, 2021], differential geometry [Bianchi et al., 2002], mixed linear regression [Chen and Womersley, 2018], [Chen et al., 2014], interferometry [Demanet and Jugnon, 2017] and ptychography [Pfeiffer, 2018]. The phase-retrieval problem is similar to the problem of *training an artificial neuron with quadratic activation*. According to [Candès et al., 2013b] "one of the first important applications of

		SVD	**KV-CBO**	**FMS**	**GD**
5 outliers	PSNR	16.0589	22.9605	22.5370	22.5370
	SNR	10.3222	16.2871	16.0017	16.0017
	SSIM	0.6310	0.8666	0.8515	0.8515
10 outliers	PSNR	12.6117	17.1769	15.0549	15.0549
	SNR	6.8398	11.4453	9.7404	9.7404
	SSIM	0.4388	0.7027	0.6312	0.6312
15 outliers	PSNR	11.6133	13.4478	13.0934	13.0934
	SNR	5.7614	7.9917	7.6647	7.6647
	SSIM	0.3662	0.5224	0.4995	0.4995

Table 6.6: Quality measures for the eigenfaces from figure 6.9.

phase retrieval is X-ray crystallography, and today this may very well be the most important application".

In many physical applications one can only measure the squared magnitude of the Fourier transform of the underlying (unknown) signal $z \in \mathbb{C}^d$. The phase is lost. Since the phase encodes important information about the signal x, see [Jaganathan et al., 2015, Fig. 1] for a nice illustrative example, reconstructing or retrieving the phase from the (known) measurements is of utmost importance.

In mathematical terms this problem can be formulated as follows. For known *measurements* $y \in \mathbb{R}^M$ the phase retrieval problem is given by

$$\text{Find } z \in \mathbb{C}^d \text{ such that } y_m = |\langle f_m, z\rangle|^2 \text{ for } m = 1, ..., M \tag{6.36}$$

where $f_m \in \mathbb{C}^d$ is the mth column of the M point discrete Fourier transform matrix, that is, $f_{mn} = \left(e^{-\frac{2\pi i}{M}}\right)^{(m-1)(n-1)}/\sqrt{M}$ for $1 \leq n, m \leq M$.

The conditions $y_m = |\langle f_m, z\rangle|^2$ do not suffice to reconstruct the signal z uniquely. For any vector z satisfying the conditions, the rotated vector $ze^{i\phi}$ for $\phi \in [0, 2\pi)$ also satisfies the conditions. Hence we can never uniquely recover the signal z (independent of the number of available measurements M). In the real setting this means that we can only reconstruct the signal up to sign and the phase-retrieval problem may as well be called *the sign-retrieval problem*. Thus, the solution set $\{\pm z\}$ is non-convex which makes the phase-retrieval problem a non-convex quadratic programme. We further note that the problem is NP-hard in general, [Pardalos and Vavasis, 1991].

Generalized phase-retrieval problem

In the following we study a different variant of the phase-retrieval problem. We consider only real signals $z \in \mathbb{R}^d$, and the measurement vectors are sampled uniformly from the sphere, that is, $a_m \sim \mathcal{U}(\mathbb{S}^{d-1})$ iid for $m = 1, ..., M$. Thus, we consider the following variant of the phase retrieval problem: assume quadratic measurements

$$y_m = |\langle a_m, z\rangle|^2, \quad m = 1, ..., M \tag{6.37}$$

where y_m are the known, measured intensities, $a_m \in \mathbb{R}^d$ are known measurement vectors, and $z \in \mathbb{R}^d$ is the unknown signal we wish to recover.

This problem can be rewritten in matrix form as $y_m = |(Az)_m|^2$ for $y \in \mathbb{R}^M$, and $A \in \mathbb{R}^{M\times d}$, that is, we store the measurement vectors a_m row-wise in the matrix A. The difficulty of the phase-retrieval problem is that the signs $sign(Az) \in \mathbb{R}^M$ are not available. If they were available we could reconstruct z by solving an (overdetermined) linear system of equations.

In order to align the reconstructed signal $z^\star$ with the original signal z we apply a rotation to the former, that is, we multiply with $sign(z \cdot z^\star) \in \{\pm 1\}$, see figure 6.10. This is of course only possible because in the numerical tests we consider here the original signal z is known.

The measurement vectors a_m are iid samples from the sphere and thus form a frame in the ambient Euclidean space. This property is important for the well-posedness of the phase-retrieval problem. In fact [Balan et al., 2006, Th. 2.2] has shown that in the real case the phase-retrieval problem has a unique solution (up to sign) if the measurement vectors form a frame, that is, there are positive constants $C_1 \leq C_2$ such that

$$C_1|x|^2 \leq \sum_{m=1}^{M} |\langle a_m, x\rangle|^2 \leq C_2|x|^2 \tag{6.38}$$

for all $x \in \mathbb{R}^d$, and the number of measurements M is $\geq 2d-1$. It would be interesting to investigate the KV-CBO method for other (possibly deterministic and complex) measurement vectors.

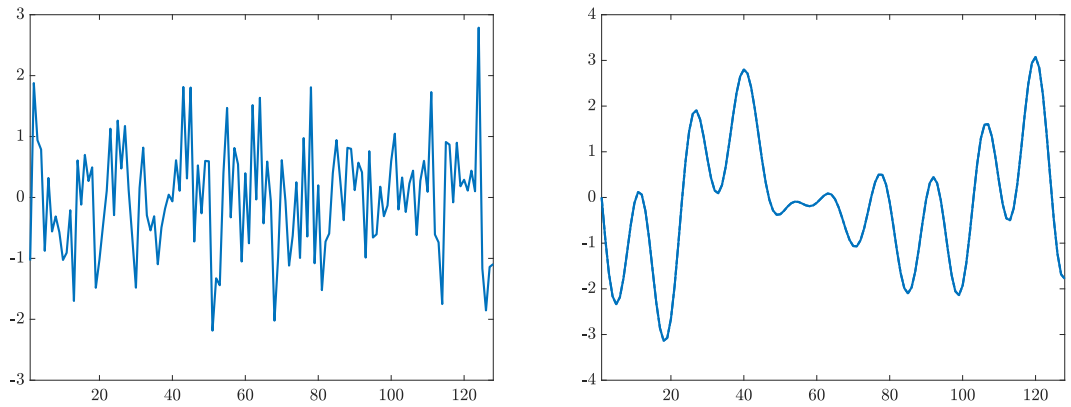


Figure 6.10: Gaussian signal (left) and smooth signal (right) in $d = 128$, and the corresponding rotated reconstructed signals. The original signals and the reconstruction are visually inseparable.

Noisy phase model

Below we test the KV-CBO method also for the phase-retrieval problem with noisy measurements. To this end we introduce the *noisy phase model* from [Chen et al., 2018] given by

$$y_m = |\langle a_m, z\rangle|^2 + \eta_m, \quad m = 1, ..., M \tag{6.39}$$

for a noise vector η with Gaussian components, that is, $\eta_m \sim \mathcal{N}(0, \sigma^2)$, where the variance σ will be chosen such that a certain SNR is achieved. Alternatively this model can be written as $y_m \sim \mathcal{N}(|\langle a_m, z\rangle|^2, \sigma^2)$. In order to prevent negative noisy measurements we set $y_m \leftarrow max(y_m, 0)$. Let $z \in \mathbb{R}^d$ a given signal that we wish to reconstruct, and y_m given noisy measurements. We chose the variance such that

$$\sigma^2 = \frac{|y|^2}{10^{SNR/10}} \tag{6.40}$$

where SNR ranges from 10 to 90 with a step of 10. Then, the SNR of the noiseless measurement vector and the noise vector η is exactly equal to the prescribed SNR. In principle it also makes sense to investigate the robustness of a phase-retrieval algorithm for the so-called *noisy magnitude model* $y_m = |\langle a_m, z\rangle + \eta_m|^2$. However we here only consider the noisy phase model (6.39).

Benchmark algorithms

In the following we first discuss common methods to solve the phase-retrieval problem. Then we discuss how the phase-retrieval problem can be reformulated as an optimization problem on the sphere and how the KV-CBO method can be used to solve it. We use the well-known PhasePack [5] package (see [Chandra et al., 2017]) for the numerical tests. The KV-CBO package contains the folder `experiment_phaseRetrieval` into which the user should download the PhasePack package in order to reproduce the results reported here, see the test files `experiment_phaseRetrieval_1/2.m` (which are adaptations of code developed by my colleague Oleh Melnyk).

The first algorithms to tackle the phase-retrieval problem were the Gerchberg-Saxton [Gerchberg, 1972] (see also [Bendory et al., 2017]) and Fienup [Fienup, 1982] algorithms. Other popular algorithms for phase-retrieval include PhaseMax [Goldstein and Studer, 2018], [Bahmani and Romberg, 2017], PhaseLamp [Dhifallah et al., 2017], PhaseLift [Candès et al., 2013a], [Candès et al., 2013b], Truncated Amplitude Flow [Wang et al., 2018], Wirtinger Flow [Candès et al., 2014] and Truncated Wirtinger Flow [Chen and Candes, 2015]. In the following we benchmark our adoption of the KV-CBO method (see algorithm 5 below) against Wirtinger Flow, Gerchberg-Saxton, PhaseMax and PhaseLift. We start with a brief explanation of these methods. All of these benchmark methods are iterative methods. We denote the initial guess by $z_0 \in \mathbb{R}^d$ and the iterates by $z_k \in \mathbb{R}^d$.

[5] https://github.com/tomgoldstein/phasepack-matlab

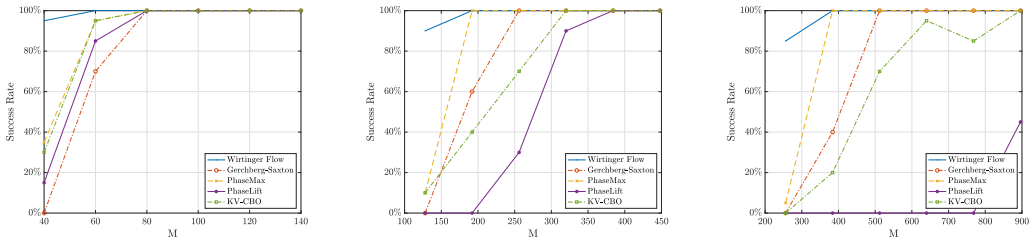


Figure 6.11: Success rates in terms of number of measurements M for a Gaussian signal in $d = 20$ (left), $d = 64$ (middle) and $d = 128$ (right). We have used the same parameters in all three experiments, see `experiment_phaseRetrieval` folder.

Wirtinger Flow

For real signals and measurement vectors (as we consider here) the Wirtinger Flow algorithm essentially performs a gradient descent search for the cost function

$$f(z_k) = \frac{1}{2M} \sum_{m=1}^{M} \left| |\langle a_m, z_k \rangle|^2 - y_m \right|^2, \quad z_k \in \mathbb{R}^d. \tag{6.41}$$

The starting point (initial guess) $z_0 \in \mathbb{R}^d$ is chosen with the *optimal spectral initialization method*, that is, z_0 is the (properly scaled) leading eigenvector of the symmetric positive semi-definite matrix $A^T \mathrm{diag}(y) A \in \mathbb{R}^{d \times d}$. We note that the rows of A are given by the measurement vectors transposed a_m^T. The step size is chosen adaptively and is inversely proportional to the norm of the initial guess z_0. More precisely: the step size in the kth iteration is given by $\mu_k / 2|z_0|^2$, where $\mu_k = \min\{1 - e^{-k/k_0},\ \mu_{max}\}$ and [Candès et al., 2014] chose values of k_0 around 330, and values of μ_{max} around 0.4.

The final algorithm is given by the iteration

$$z_{k+1} = z_k - \frac{\mu_{k+1}}{2|z_0|^2} \nabla f(z_k), \quad \nabla f(z_k) = \frac{2}{M} \sum_{m=1}^{M} \big(|\langle a_m, z_k \rangle|^2 - y_m \big) (a_m a_m^T) z_k. \tag{6.42}$$

In [Candès et al., 2014, III/A] the Wirtinger Flow algorithm is analysed in the complex case, that is, the signal is complex and the measurement vectors are complex Gaussians. In this setup the algorithm is shown to converge to z at a geometric rate. Further, the cost function $f : \mathbb{C}^d \to \mathbb{R}$ maps the complex numbers into the reals and is therefore not holomorphic and consequentely not complex differentiable. Nevertheless, the expression from (6.42) can still be interpreted as a *Wirtinger derivative*, hence the name of the method.

Gerchberg-Saxton

The Gerchberg-Saxton algorithm (1972) has been designed to solve the Fourier phase-retrieval problem, that is, the task of recovering a signal from the magnitude of its Fourier transform, see (6.36). In this setting it is identical to the well-known Fienup algorithm (1982). In fact, both algorithms consist

of four steps that can be summarized as follows: in the kth iteration, transform the current guess $z_k \in \mathbb{C}^d$ to the Fourier domain, incorporate the known measurements $y \in \mathbb{R}^M$, transform back to the object domain and compute the next iterate $z_{k+1} \in \mathbb{C}^d$. The Gerchberg-Saxton and Fienup algorithms only differn in the last step. The Fienup algorithm solves the Fourier phase-retrieval problem with additional constraints on the (unknown) signal z, e.g. non-negativity $z \geq 0$. The Gerchberg-Saxton algorithm on the other hand only aims at finding a solution that satisfies the measurement constraints (in the Fourier domain). Thus the Fienup algorithm is a generalization of the Gerchberg-Saxton algorithm, and, since we do not consider constraints in the object domain in (6.37), we may as well use the Fienup algorithm as benchmark method. The Gerchberg-Saxton and Fienup algorithms belongs to the class of *alternating projection algorithms*. Algorithms from this class have the desirable *error-reduction property*, that is, the error can be shown to be monotonically non-increasing.

As explained above, we here do not consider the Fourier phase-retrieval problem. Instead we consider the problem of recovering a real signal z from squared magnitudes of Az where $A \in \mathbb{R}^{M \times d}$ is a real matrix, see (6.37). The Gerchberg-Saxton (and Fienup) algorithm can, of course, be transfered to this setting, see algorithm 4 below.

What remains to discuss is the choice of the initial guess z_0. In the original publication Gerchberg and Saxton chose the initial guess z_0 randomly. We however use the same initialization method as for the Wirtinger-Flow algorithm, that is, the optimal spectral initialization method explained above.

Algorithm 4: Gerchberg-Saxton

Input: measurement matrix $A \in \mathbb{R}^{M \times d}$, linear measurements $y_m = |\langle a_m, z \rangle|$ for $m = 1, ..., M$, initial guess z_0, max. number of iterations n_T

for $k = 0$ **to** n_T **do**

- $Z_k = Az_k$
- $Z'_k = y \odot sign(Z_k)$ *(element wise multplication)*
- Solve least squares problem

$$Az'_k = Z'_k \tag{6.43}$$

- $z_{k+1} = z'_k$

Output: reconstructed signal $z_{n_T} \approx \pm z$

PhaseMax

The PhaseMax algorithm pursues again a different approach as the two previous algorithms. It is based on a convex relaxation of the original phase-retrieval problem which is given by a constrained *convex* optimization problem. By introducing a quadratic *barrier function* this constrained problem is reformulated as an unconstrained problem which is solved with the FASTA (Fast Adaptive Shrinkage/Thresholding Algorithm) solver.

More precisely: the convex relaxation is given by

$$\begin{cases} \max_{z^\star \in \mathbb{R}^d} \langle z^\star, z_0 \rangle \\ \textit{such that } |Az^\star| \leq b \end{cases} \tag{6.44}$$

where the constraint is understood entry-wise with $b := y^{1/2}$, and z_0 is an initial guess for the signal z that we again compute with the optimal spectral initialization method. We note that in [Goldstein and Studer, 2018] the term *PhaseMax* refers to the problem formulation (6.44), not to a particular algorithm to solve it. In our experiments we consider the implementation proposed in the PhasePack package (also from Goldstein & Studer), which we refer to as the *PhaseMax algorithm*.

For any $z \in \mathbb{R}^d$ that satisfies the measurement constraints also $-z$ satisfies them. Hence the solution set of the phase-retrieval problem is not convex. However, as is well-known, the solution set of any convex optimization problem is convex. We therefore first need to remove the *sign ambiguity* (or phase ambiguity in the complex case) in order to arrive at a convex relaxation for the phase-retrieval problem. PhaseMax does so by restricting the solution space to the vectors for which the (arbitrary) quantity $\omega = sign(\langle z^\star, z_0 \rangle)$ is positive, where $z^\star$ is any vector in the feasible set. In this case we have

$$\omega \langle z^\star, z_0 \rangle = |\langle z^\star, z_0 \rangle| > \langle -z^\star, z_0 \rangle. \tag{6.45}$$

Hence, requiring $\langle z^\star, z_0 \rangle > 0$ removes the sign ambiguity of (6.44) and thus restores convexity of the solution set. It also motivates the name PhaseMax. In summary: the original measurements (6.37) are taken care of by the constraints in (6.44), whereas the cost function restores convexity of the solution set. We thus arrive at a convex relaxation of the orginal problem.

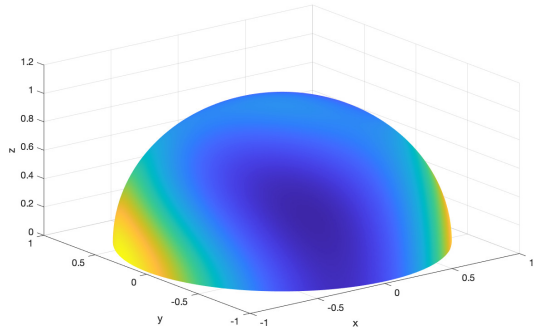


Figure 6.12: The cost function (6.54) for $d = 2$

We now briefly discuss the PhaseMax algorithm from the PhasePack package to solve (6.44). In a first step the constraints $|Az^\star| \leq b$ are approximately enforced by introducing a quadratic barrier

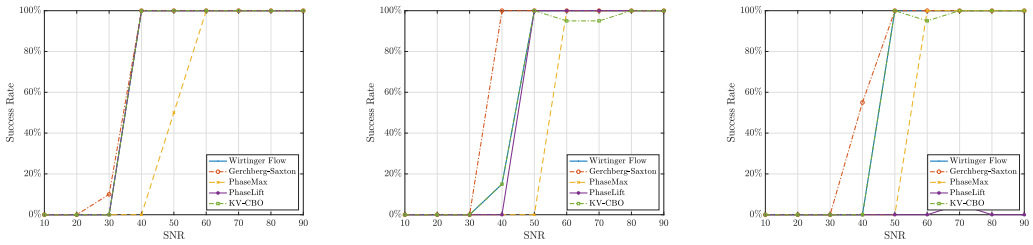


Figure 6.13: Success rates in terms of SNR for a Gaussian signal in $d = 20$ (left), $d = 64$ (middle) and $d = 128$ (right). We have used the same parameters in all three dimensions, see `experiment_phaseRetrieval` folder.

(penalty) function. Thus we arrive at the unconstrained optimization problem

$$\min_{z^\star \in \mathbb{R}^d} -\langle z^\star, z_0 \rangle + \frac{1}{2} |\max\{|Az^\star| - b, 0\}|^2 \tag{6.46}$$

where the maximization from (6.44) has been replaced with a minimization. This problem is solved with the FASTA implementation of the *Forward-Backward Solver* (FBS). FBS is an implementation of the *forward-backward splitting method* or *proximal gradient method*. We refer to [Goldstein et al., 2014], [Combettes and Pesquet, 2011] and [Chen et al., 2020] for the details.

PhaseLift

The last benchmark method that we consider is PhaseLift. Just like PhaseMax, PhaseLift also introduces a convex relaxation by removing the sign ambiguity of the original phase-retrieval problem. However the approach taken is different from PhaseMax. PhaseLift reformulates the quadratic constraints as affine constraints about the symmetric positive definite matrix $Z = zz^T \in \mathbb{R}^{d \times d}$ which is unaffected by a change in sign (or phase) of $\pm z$. This reformulation introduces an optimization problem in a d^2 dimensional space, where d is the original signal dimension. In other words: we have *lifted* the phase-retrieval problem to a rank-1 matrix completion problem, hence the name PhaseLift. The advantage is that no constraints on the signal are imposed (differently from PhaseMax, see (6.45)). The disadvantage is the higher numerical cost. We note that the PhaseLift algorithm (see below) is much slower than the other benchmark methods.

Introducing the operator $\mathcal{A}(Z) = \{a_m^T Z a_m\}_{m=1}^M$ for symmetric matrices $Z \in \mathbb{R}^{d \times d}$ allows to rewrite the quadratic measurements as $\mathcal{A}(zz^T) = y$. Further it can be shown that the phase-retrieval problem is equivalent to the *matrix completion problem*

$$\begin{cases} \min \operatorname{rank}(Z) \\ \textit{such that } \mathcal{A}(Z) = y, \ Z \in \mathbb{R}^{d \times d} \textit{ is symmetric positive definite.} \end{cases} \tag{6.47}$$

Once a rank 1 matrix $Z^\star$ is found, the factorization (SVD) $Z^\star = zz^T$ yields the solution for the

phase-retrieval problem. Problem (6.47) is NP-hard and [Candès et al., 2013b] propose the following trace-norm relaxation

$$\begin{cases} \min \mathrm{Tr}(Z) \\ \textit{such that } \mathcal{A}(Z) = y,\ Z \in \mathbb{R}^{d\times d} \textit{ is symmetric positive definite.} \end{cases} \tag{6.48}$$

This problem is a *semidefinite programme* in standard form, and there are many algorithms that solve this type of problem. In the PhasePack implementation the above problem is solved with FASTA (see above), where the constraints are again incorporated into the cost function via a quadratic barrier function $\frac{1}{2}|\mathcal{A}(Z) - y|^2$ and a *proximal operator* $\mathrm{prox}(Z, \delta)$ respectively. The proximal operator is defined as $\mathrm{prox}(Z, \delta) = VDV^T$ where $D = \mathrm{diag}(Re(\Lambda) - \delta)_+ \in \mathbb{R}^{d\times d}$ and $V\Lambda V^T = Z$ is the eigendecomposition of Z. In simple terms: the proximal operator drops the eigenvalues that are smaller than a given threshold $\delta > 0$, yielding a symmetric positive definite matrix.

KV-CBO for phase-retrieval

In order to apply the KV-CBO method to the phase retrieval problem we first need to recast it as an optimization problem on the sphere. Here the fact that the measurement vectors a_m form a frame, that is, there are constants $C_1, C_2 > 0$ such that $C_1|z|^2 \leq \sum_{m=1}^M |\langle a_m, z\rangle|^2 \leq C_2|z|^2$, is crucial.

We first compute an upper bound on the lower frame bound C_1. To this end we define the matrix $S \in \mathbb{R}^{d\times d}$ with columns given by

$$S_{-k} = \sum_{m=1}^{M} \left(A^T \mathrm{diag}(A_{-k})\right)_{-m} \tag{6.49}$$

where $A_{-k} \in \mathbb{R}^M$ is the kth column of A, and $\mathrm{diag}(A_{-k}) \in \mathbb{R}^{M\times M}$ the corresponding diagonal matrix. An upper bound on C_1 is then given by the smallest eigenvalue of S, that is, $C_1 \leq \sigma_{min}(S)$. In the following we denote this upper bound again by C_1.

Noting that

$$C_1|z|^2 \leq \sum_{m=1}^{M} |\langle a_m, z\rangle|^2 = \sum_{m=1}^{M} y_m \tag{6.50}$$

yields the following upper bound on the norm of the unknown signal z

$$|z| \leq \sqrt{\frac{1}{C_1}\sum_{m=1}^{M} y_m} =: R. \tag{6.51}$$

With this bound we define the vector $V^\star \in \mathbb{R}^{d+1}$ by

$$V^\star = \frac{1}{R}\begin{pmatrix} z \\ \sqrt{R^2 - |z|^2} \end{pmatrix}. \tag{6.52}$$

This vector is easily seen to have norm 1, hence $V^\star \in \mathbb{S}^d$, which allows us to recast the original phase-retrieval problem (6.37) to a phase-retrieval problem on the sphere. More precisely: extending the measurement vectors by a zero, that is, setting $\tilde{a}_m = (a_m^T, 0)^T \in \mathbb{R}^{d+1}$ yields the problem

$$|\langle \tilde{a}_m, V^\star \rangle|^2 = |\langle a_m, \frac{z}{R} \rangle|^2 = \frac{y_m}{R^2} =: \tilde{y}_m \tag{6.53}$$

where $V^\star$ is the unknown signal we wish to reconstruct with the KV-CBO method. We solve this problem by empirical risk minimization with the KV-CBO method. Once we found $V^\star$ we can easily reconstruct the true signal z of the original problem.

Algorithm 5: KV-CBO for the phase-retrieval problem

Input: frame $a_m \in \mathbb{R}^d$ for $m = 1, ..., M$, measurements $y_m = |\langle a_m, z \rangle|^2$

1 Find upper bound on the lower frame bound C_1, see (6.49) and (6.50)

2 Apply the KV-CBO method to the empirical risk minimization function of the recasted problem (6.53)

$$V_{n_T}^\alpha \leftarrow \text{KV-CBO}(\mathcal{E}), \quad \mathcal{E}(v) := \sum_{m=1}^{M} \left| |\langle \tilde{a}_m, v \rangle|^2 - \tilde{y}_m \right|^2 \tag{6.54}$$

3 Multiply $V_{n_T}^\alpha \in \mathbb{R}^{d+1}$ with R and drop the last component

$$z^\star \leftarrow R \sum_{i=1}^{d} (V_{n_T}^\alpha)_i e_i \tag{6.55}$$

Output: reconstructed signal $z^\star \approx \pm z$

Remark 6.4.1. *1. The vectors $\tilde{a}_m$ no longer form a frame. 2. An interesting open problem is showing the coercivity condition/ inverse continuity assumption for the empirical risk function* (6.54)*. In [Fornasier et al., 2021a] it is proposed to use the stability estimates from [Bandeira et al., 2014] and [Eldar and Mendelson, 2014]. 3. An interesting approach to improve the method is the use of a well-chosen initial guess (e.g. with the optimal spectral initialization method from above).*

6.5 Reconstruction of neural nets

In this section we discuss the third and last application of the KV-CBO method, namely the reconstruction of single layer neural nets.

First, we start with a brief overview of neural nets. A *neural net* is a mathematical model for the neural networks in the human brain. Neural nets are used to categorized data points (so-called feature vectors) and are at the heart of machine learning. They have proven successful in a large range of applications, including autonomous driving, face recognition, natural language processing, robotics, X-ray detection and weather forecasting.

The main building block in all neural nets is the neuron, which is a generalization of the *perceptron*. The latter is, in modern notation, a function of the form

$$f(x) = H(a \cdot x + \vartheta), \quad x \in \mathbb{R}^d \tag{6.56}$$

where $a \in \mathbb{R}^d$ is the weight vector, $\vartheta \in \mathbb{R}$ is the *bias*, and H is the Heaviside function ($H(x) = 1$ if $x > 0$, and $H(x) = 0$ otherwise). With the perceptron we can build the basic logic gates AND ($a = (1,1)^T$, $\vartheta = -1$), OR ($a = (2,2)^T$, $\vartheta = -1$) and NEGATION ($a = -1$, $\vartheta = 1$). With a multilayer perceptron we can also build the XOR gate. The idea of the perceptron is generalized by allowing for different *activation functions*, instead of the Heaviside function, the hyperbolic tangent and ReLU are popular choices. We then refer to the perceptron (6.56) as *neuron*, as it models the biological neurons in the human brain. A neural net is, similarly to the human brain, a collection of neurons organized in layers. Each neural net has one input layer and one output layer, and possibly many hidden layers. A neural net is called *shallow* if it has only one hidden layer, and *deep* if it has many hidden layers, see figure 6.14 below for a depiciton of a *shallow neural net*. The mathematical definition of the latter is

$$f(x) = \sum_{i=1}^{m} b_i \sigma(a_i^T x + \vartheta_i), \quad x \in \mathbb{R}^d \tag{6.57}$$

where d is the input dimension, m is the number of neurons on the hidden layer, $a_i \in \mathbb{R}^d$ are the weights, ϑ_i the biases, and b_i other weights.

Motivation and problem formulation

Let us assume $f : \mathbb{R}^d \to \mathbb{R}$ is a trained (!) neural net with one hidden layer and $m \leq d$ hidden neurons. Let us further assume that the thresholds ϑ_i and the weights b_i are known. The weights $a_i \in \mathbb{R}^d$ however are unknown (undisclosed, kept in secret). The problem of reconstructing a neural net is the following: can we reconstruct (reveal, lay open) the weights a_i given input-output pairs $(x_i, f(x_i))$ for any possible x_i of our choosing? In other words: can we reconstruct the weights a_i if we are given a *machine* $x \mapsto f(x)$ with which we can generate our own training data?

In the groundbreaking work [Fefferman, 1994] this problem is investigated on a theoretical level.

By using techniques from complex analysis, Fefferman was able to show that, under (quite strict) conditions on the weights of the neural net and a-priori knowledge of the architecture of the neural net, two neural nets with the same output function are *isomorphic*, that is, their weights are identical up to unavoidable ambiguities (sign changes and permutations of neurons on the same layer).

The problem of *training a neural net* is ubiquitous in machine learning. The problem of *reconstructing a neural net* on the other hand seems rather academic. We now discuss the rationale of the latter problem. If we interpret the given training data (x_i, y_i) as samples from a trained but unknown net f, that is,

$$(x_i, y_i) \equiv (x_i, f(x_i)) \tag{6.58}$$

where f is a trained neural net with unknown (hidden) weights, then the weights obtained from *reconstructing* f should be isomorphic to the weights obtained from *training* a net with the data (x_i, y_i) (provided a result like the one of Fefferman holds).

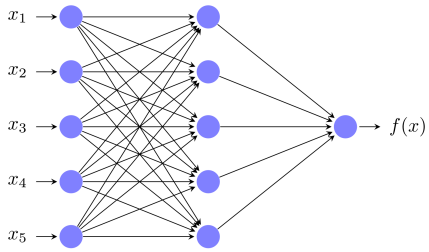


Figure 6.14: Shallow neural net $f : \mathbb{R}^5 \to \mathbb{R}$ with one hidden layer.

We assume: $d = m$, $\sigma = \tanh$ and the weights a_i are normalized $\|a_i\|_2 = 1$ and, either, orthogonal or quasi-orthogonal. The columns a_i are called quasi-orthogonal if they satisfy $\mathcal{S}(a_1, ..., a_m) < \epsilon$ where

$$\mathcal{S}(a_1, ..., a_m) = \inf\{\big(\sum_{i=1}^{m} |a_i - \overline{a}_i|^2\big)^{1/2} \mid \overline{a}_i \text{ are orthonormal}\} \tag{6.59}$$

and $\epsilon > 0$ reasonably small. We store the weights a_i columnwise in a matrix $A \in \mathbb{R}^{d \times m}$.

Projected subgradient ascent

The following is an overview of [Fornasier et al., 2021d], [Fornasier et al., 2012], [Fiedler et al., 2021] and [Rauchensteiner, 2018].

The approach is the following: expose the weights a_i by computing (or approximating) the Hessian of $f(x)$ and then find rank-1 matrices in an m-dimensional matrix subspace. More precisely: the Hessian of f is given by

$$\nabla^2 f(x) = \sum_{i=1}^{m} b_i \sigma''(a_i^T x + \vartheta_i) a_i \otimes a_i \in L \tag{6.60}$$

where $a_i \otimes a_i = a_i a_i^T$ is a rank-1 matrix and $L = \text{span}\{a_i \otimes a_i \mid i = 1, ..., m\}$ is an m-dimensional matrix subspace in $\mathbb{R}^{d\times d}$. We store the Hessians in a matrix

$$M = \{\nabla^2 f(x_i)^{vec}\}_{i=1}^{m_x} \in \mathbb{R}^{d^2 \times m_x} \tag{6.61}$$

where m_x is the number of sample points $x_i \sim \mathcal{U}(\mathbb{S}^{d-1})$ iid, and vec denotes vectorization (we store $\nabla^2 f(x_i)$ in one large vector of dimension d^2). The space L defined above is unknown (as the weights a_i are unknown), however, if the number of samples m_x is large enough the matrix M contains enough information to approximate L. Once we have a good approximation $L^\star$ of L, we need to find the spannig rank-1 matrices $a_i \otimes a_i$ (difficult), from which we can then find the unknown weights a_i (easy).

Since the weights a_i are unknown we have to rely on the numerical analogue of the matrix M from (6.61), that is, we set

$$M_h = \{\nabla_h^2 f(x_i)^{vec}\}_{i=1}^{m_x} \in \mathbb{R}^{d^2 \times m_x} \tag{6.62}$$

where we use the finite difference approximation

$$(\nabla_h^2 f(x))_{ij} = \frac{f(x + he_i + he_j) - f(x + he_j) - f(x + he_i) + f(x)}{h^2} \tag{6.63}$$

for some small $h > 0$.

Remark 6.5.1.

1. *Computing the gradient* $\nabla f(x) = \sum_{i=1}^m b_i \sigma'(a_i^T x + \vartheta_i) a_i \in \text{span}\{a_i \mid i = 1, ..., m\}$ *also exposes the weights* a_i*. However, this information is not enough for reconstruction.*

2. *From* (6.60) *we see that not every activation function is admissible. For example, the famous ReLU activation is not admissible since its second derivative is zero* $\sigma'' = 0$*.*

We now discuss an algorithm proposed in [Fornasier et al., 2021d, algorithm 6] (see also [Rauchensteiner, 2018, algorithm 2]) to reconstruct the weights a_i of the neural net, see algorithm 6 below.

First, we approximate the unknown matrix subspace $L \subset \mathbb{R}^{d\times d}$ by computing the Hessians of the neural net for many sample points x_i and setting $L^\star$ to be the first k principal components of the matrix of vectorized Hessians. The matrix $L^\star \in \mathbb{R}^{d^2 \times k}$ is then an approximation of the vectorized matrix L^{vec} in the sense that the Frobenius distance of the corresponding orthogonal projections $P^\star$ and P_L is small, that is, $\|P^\star - P_L\|_F$ is small. The orthogonal projection $P^\star$ onto the approximate space $L^\star$ is given by $P^\star = L^\star (L^\star)^T$.

In a second step we have to find rank-1 matrices in the m-dimensional subspace $L^\star$. To explain how we tackle this problem, let us assume that the exact space L is given. Since the Frobenius norm

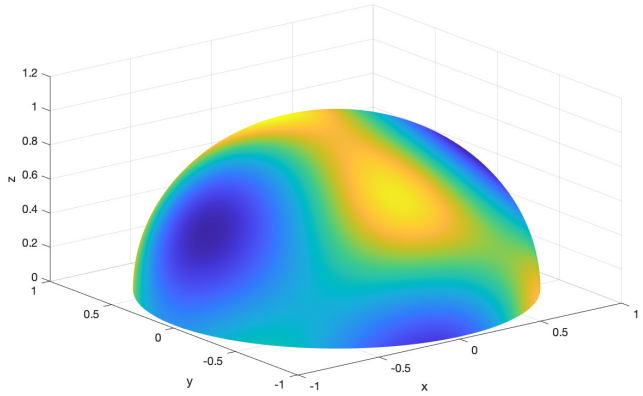


Figure 6.15: The cost function (6.71) for $d = 3$.

of the symmetric rank-1 matrices $a_i \otimes a_i$ is 1, they are solutions of the following optimization problem

$$\begin{cases} \min \operatorname{rank}(M) \\ \textit{such that } M \in L,\ \|M\|_F = 1. \end{cases} \tag{6.64}$$

It can further be shown that they are the only solutions. That is, (6.64) has m global minimizers, all of which correspond to one of the matrices $a_i \otimes a_i$. The optimization problem (6.64) is in general NP-hard, and [Fornasier et al., 2021d] (see also [Rauchensteiner, 2018]) propose the relaxation

$$\begin{cases} \max \|M\|_2 \\ \textit{such that } M \in L,\ \|M\|_F = 1. \end{cases} \tag{6.65}$$

Here $\|M\|_2 = \sigma_{max}(M)$ is the spectral norm (or the Schatten-∞ norm). This relaxation is motivated by the fact that $\|M\|_2 \leq \|M\|_F = \sum \sigma_i^2(M)$ with equality iff $\operatorname{rank}(M) \in \{0, 1\}$. An optimal solution to (6.64) satisfies $\operatorname{rank}(M^\star) = 1$. Thus, in this case, the spectral norm is maximal (under the given constraints) $\|M^\star\|_2 = \|M^\star\|_F = 1$ and $M^\star$ is also an optimal solution of (6.65). The converse also holds.

In summary: the sets of solutions of (6.64) and (6.65) are the same.

It is proposed to use a *projected subgradient ascent method* to solve (6.65), that is, starting from an initial guess X_0 we iterate $X^{j+1} \leftarrow P^\star P_\gamma(X^j)$ followed by a normalization. Here the operator P_γ is defined as

$$P_\gamma(X) = P_\gamma(U\Sigma V^T) = \frac{1}{\sqrt{\gamma^2\sigma_1^2 + \sigma_2^2 + ... + \sigma_d^2}} U \operatorname{diag}(\gamma\sigma_1, \sigma_2, ..., \sigma_d) V^T \tag{6.66}$$

where γ may be chosen as $\gamma = 10$. This approach is reasonable, since (6.65) is about maximizing the maximal singular value $\sigma_{max} = \sigma_{max}(X)$.

Algorithm 6: Recover weights with projected gradient ascent method

Input: machine that returns $f(x)$ for any x, number of samples m_x, step size h, dimension of approximating subspace $k \in \{m, ..., d^2\}$, γ, N_{max}

Sample $x_i \sim \mathcal{U}(\mathbb{S}^{d-1})$ iid for $i = 1, ..., m_x$

Compute the Hessians $\nabla_h^2 f(x_i)$ and assemble the matrix M_h as in (6.62)

Compute SVD

$$M_h = U\Sigma V^T \tag{6.67}$$

Set $L^\star$ to be the first k left singular vectors (columns of U)

while $\neg$ *found all m minimizers* **do**

  choose an initial value $X^0 \in \mathbb{R}^{d^2}$ as

$$\tilde{X}^0 = \sum_{i=1}^{k} \xi_i W_i, \quad X^0 = \tilde{X}^0 / \|\tilde{X}^0\|_F \tag{6.68}$$

  where $W_1, ..., W_k \in \mathbb{R}^{d^2}$ is an orthonormal basis of $L^\star$ and $\xi \sim \mathcal{N}_k(0, 1)$

  reshape X^0 as $d \times d$ matrix

  for $j = 0$ **to** $N_{max} - 1$ **do**

    set $X^{j+1} \leftarrow P_\gamma(X^j)$

    reshape X^{j+1} as $k \times 1$ matrix

    set $X^{j+1} \leftarrow P^\star(X^{j+1})$, where $P^\star = L^\star(L^\star)^T \in \mathbb{R}^{d^2 \times d^2}$

    normalize $X^{j+1} \leftarrow X^{j+1} / \|X^{j+1}\|_F$

    reshape X^{j+1} as $d \times d$ matrix

  $u_1 \in \mathbb{R}^d \leftarrow$ first left singular vector of X^j (approximation of one of the weights a_i)

  if u_1 *not found in a previous iteration* **then**

    store u_1 in matrix A_{recon}

Output: $A_{recon} \in \mathbb{R}^{d \times m}$ with $\|A_{recon} - A\|_F$ small

In [Fiedler et al., 2021] an alternative relaxation is considered:

$$\max_{v \in \mathbb{S}^{d-1}} \|P_L(v * v)\|_F^2 \tag{6.69}$$

where P_L is the orthogonal projection onto L and the asterix $*$ denotes the *columnwise Khatri-Rao product* defined as

$$A * B = \left((A_1 \otimes B_1)^{vec}, \cdots, (A_m \otimes B_m)^{vec}\right) \in \mathbb{R}^{n_1 n_2 \times m} \tag{6.70}$$

where $A \in \mathbb{R}^{n_1 \times m}$, $B \in \mathbb{R}^{n_2 \times m}$ and A_i, B_i are the ith column of A and B respectively. This is the relaxation we consider below. It is proposed to solve this problem again with a projected gradient ascent algorithm, see algorithm 7.

Algorithm 7: Recover weights with projected gradient ascent method 2

Input: approximating space $L^\star$, γ, max. iterations N_{max}

while $\neg$ *found all* m *minimizers* **do**
  choose an initial value $u^0 \sim \mathcal{U}(\mathbb{S}^{d-1})$
  for $j = 0$ **to** $N_{max} - 1$ **do**
    set $U^j \leftarrow u^j \otimes u^j$
    reshape U^j as $d^2 \times 1$ matrix
    set $U^{j+1} \leftarrow P^\star(U^j)$, where $P^\star = L^\star(L^\star)^T \in \mathbb{R}^{d^2 \times d^2}$
    reshape U^{j+1} as $d \times d$ matrix
    set $u^{j+1} \leftarrow u^j + 2\gamma U^{j+1} u^j$
    normalize $u^{j+1} \leftarrow u^{j+1} / \|u^{j+1}\|_2$
  if u^j *not found in a previous iteration* **then**
    store u^j in matrix A_{recon}

Output: $A_{recon} \in \mathbb{R}^{d \times m}$ with $\|A_{recon} - A\|_F$ small

The above discussion of the optimization problem (6.64) and the relaxed versions (6.65) and (6.69) assumed that the exact space L is known. However, in practice this space is not known and has to be approximated, see $L^\star$. The reconstruction problem is then solved by considering the analogue optimization problems (6.65) and (6.69) for the space $L^\star$, yielding approximations of the true weights $\tilde{a}_i \approx a_i$. We refer to the works [Fornasier et al., 2021d], [Fornasier et al., 2012], [Fiedler et al., 2021] and [Rauchensteiner, 2018].

It is unclear how many times we need to run the above algorithms to find all the weights a_i, $i = 1, ..., m$. The minimizer found in a single run depends on the starting point X_0 and it is possible that we need to run the methods $\gg m$ times to find all the weights. We refer to the GitHub folder `experiment_neuralNets` for an implementation of the above two algorithms.

In the following we propose an adaptation of the KV-CBO method that is expected to find more than 1 minimizer of the optimization problem (6.69) in a single application.

Consensus-based optimization with multiple consensus points

We here discuss an adaptation of the KV-CBO method for the relaxed problem (6.69), that is, we set

$$\mathcal{E}(v) = -\|P^\star(v * v)\|_F^2 + 1 \tag{6.71}$$

where the negative sign is chosen to turn the maximization problem into a minimization problem, and the $+1$ makes the function values positive (as is required by the KV-CBO method).

Instead of having one consensus point $V^\alpha(\rho_t^N)$ we now consider the KV-CBO method with *multiple consensus points*. Now every particle V_t^i has its own consensus point $V^{\alpha,i}(\rho_t^N)$. This individual consensus point is a *localized* version of the usual consensus point in the sense that the weighted average only considers those particles V_t^j that are in some sense close to V_t^i. The particles will then form multiple (!) clusters, each of which is expected to converge to one of the global minimizers of the given cost function. Applied to the reconstruction of neural nets, we hope to find all m global minimizers of (6.71) with one application of the KV-CBO method.

More precisely: the adaptation of the (anisotropic) KV-CBO model is given by

$$dV_t^i = \lambda P_\Gamma(V_t^i)V^{\alpha,i}(\rho_t^N)dt + \sigma P_\Gamma(V_t^i)D(V_t^i - V^{\alpha,i}(\rho_t^N))dB_t^i + C_\Gamma^{ani}(V_t^i, V^{\alpha,i}(\rho_t^N))dt \tag{6.72}$$

where the correction term remains unchanged, see (3.51). The discretization is done similarly as before, see (6.1) - (6.5).

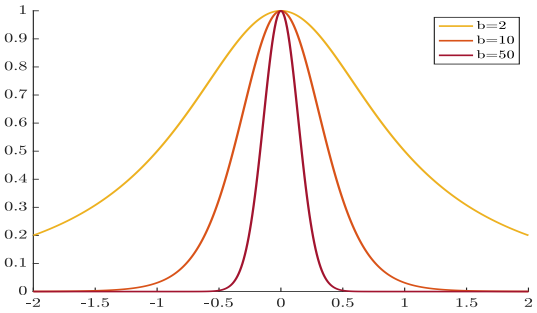


Figure 6.16: Cucker Smale weight (6.74) for $C = 1$ and different values of b

For the consensus point $V^{\alpha,i}(\rho_t^N)$ we propose two choices, one based on a *k-nearest neighbours search* and one based on a *Cucker-Smale weight function*. The former is given by

$$V^{\alpha,i}(\rho_t^N) = \frac{\sum_{j\in kNN(i;k,p)} V_t^j e^{-\alpha\mathcal{E}(V_t^j)}}{\sum_{j\in kNN(i;k,p)} e^{-\alpha\mathcal{E}(V_t^j)}} \tag{6.73}$$

where $kNN(i;k,p)$ are the indecies of the k nearest neighbors of V_t^i in the p-norm. In many high-dimensional clustering applications practitioners choose $p = 2$. However, [Aggarwal et al., 2001], [Zimek et al., 2012] argue that $p = 1$ may be the better choice. This choice for the consensus point is not very efficient as we need to compute pairwise distances $\|V_t^i - V_t^j\|_p$ in every iteration.

A second choice is given by the following weighted consensus point

$$V^{\alpha,i}(\rho_t^N) = \frac{\sum_{j=1}^N V_t^j e^{-\alpha\mathcal{E}(V_t^j)}\omega_{(b,C)}(\|V_t^i - V_t^j\|_p)}{\sum_{j=1}^N e^{-\alpha\mathcal{E}(V_t^j)}\omega_{(b,C)}(\|V_t^i - V_t^j\|_p)}, \quad \omega_{(b,C)}(x) = \frac{1}{\left(1+\frac{x^2}{C}\right)^{b/2}} \tag{6.74}$$

where $\omega_{(b,C)}(x)$ is a bell-shaped weight function motivated by the works of Cucker and Smale, see [Cucker and Smale, 2007a], [Cucker and Smale, 2007b].

	CS	kNN
$m = 1$	100%	100%
$m = 2$	80%	100%
$m = 3$	40%	100%

Table 6.7: Success rates for full recovery of all m weights in dimension $d = 3$, see `experiment_neuralNets` folder.

We finish this chapter with an interesting open problem: in high dimensional spaces $d \geq 10$ the KV-CBO method quickly fails in finding all the weights. In other words: in high dimensional spaces the particles seem to converge to only one (!) of the global minimizers instead of m. A possible explanation for this is based on the *concentration of measure phenomenon* on the sphere. This phenomenon states that vectors on the high-dimensional sphere are nearly orthogonal, that is,

$$\langle V_t^i, V_t^j \rangle \approx 0 \tag{6.75}$$

for $i \neq j$.

This further implies that the pairwise distances, and concequently the Cucker-Smale weights, are nearly constant $|V_t^i - V_t^j|_2 \approx 2$. Hence, the Cucker-Smale weights do not allow for a good discrimination between near and far points and consequentely the particles all converge to the same minimizer.

Here more research about efficient ways to cluster points on a high-dimensional sphere is needed.

6.6 Summary

We have started this chapter with a discussion of the discretization of the KV-CBO model and various implementation aspects. We refer to the discretized KV-CBO model as the *KV-CBO method*. We stated the most general version of the method in algorithm 1, and then restricted ourselves to the special case $\Gamma = \mathbb{S}^{d-1}$ in the rest of the chapter. We introduced the Matlab package of the KV-CBO method which can be downloaded from GitHub. Whenever possible we included the scripts needed to reproduce the numerical experiments. We have seen that the anisotropic version of the KV-CBO method is preferable to the isotropic version and have therefore restricted ourselves to the former in the rest of the chapter.

We have disucssed three real world applications, namely robust PCA, phase retrieval, and the *reconstruction* of a neural net. We motivated the importance of the latter problem with the work of C. Fefferman. We remind the reader that the phase-retrieval problem is a special case of the problem of *training* a neuron with quadratic activation. In all of these applications the KV-CBO method can keep up with well-known benchmark methods.

In summary: we have a proof of concept that a consensus-based optimization method can be used to solve important machine learning problems.

Appendix

Stochastic Calculus

Definition 6.6.1 (Distribution, Law)**.** *1. Let $X : \Omega \to S$ a random variable on the probability space $(\Omega, \mathcal{A}, P)$ to the measure space $(S, \mathcal{S})$. The distribution or law of X is defined as the pushforward*

$$\mu(B) = X_{\#}P(B) = P \circ X^{-1}(B) \tag{6.76}$$

for all $B \in \mathcal{S}$. The distribution of X is thus a probability measure on $(S, \mathcal{S})$. We also use the notations $\mu = law(X)$ and $X \sim \mu$.

2. For a stochastic process $X : T \times \Omega \to S$ we define the function

$$\Phi_X : \Omega \to S^T,\ \Phi_X(\omega)(t) := X_t(\omega) \tag{6.77}$$

where S^T is the set of all functions mapping from $T \to S$. The distribution or law of X is defined as the pushforward

$$\nu = (\Phi_X)_{\#}P = P \circ \Phi_X^{-1}. \tag{6.78}$$

In other words: $\nu_t(\cdot) = P \circ X_t^{-1}(\cdot)$ is the probability distribution of the random variable X_t.

Definition 6.6.2 (Brownian motion)**.** *Let $T \in (0, \infty]$ a time horizon and Ω a probability space. A Brownian motion $B : [0, T] \times \Omega \to \mathbb{R}$ is a stochastic process, that is, a collection of random variables indexed by time, with the following properties*

1. *$B_0 = 0$ almost surely,*
2. *$t \mapsto B_t$ is almost surely continuous,*
3. *B has independent increments, that is, the random variables $B_{t_1} - B_{s_1}$ and $B_{t_2} - B_{s_2}$ are independent for all $0 \leq s_1 < t_1 \leq s_2 < t_2$,*
4. *the random variable $B_t - B_s$ follow a Gaussian distribution with expectation zero and variance $t - s$, that is, $B_t - B_s \sim \mathcal{N}(0, t - s)$ for all $s \leq t$.*

Here B_t is the short notation for $B(t)$.

Remark 6.6.1.

1. *Figure 2.2 on page 22 shows one realization of a one-dimensional Brownian motion.*
2. *A Brownian motion in $\mathbb{R}^d$ is a stochastic process $B : [0,T] \times \Omega \to \mathbb{R}^d$ of independent (!) one-dimensional Brownian motions, that is, each component of B is a one-dimensional Brownian motion as defined above.*

Definition 6.6.3 (SDE, Itô integral, Stratonovich integral)**.** *A stochastic differential equation in Itô form is given by*

$$X_t = \mu(X_t)dt + \sigma(X_t)dB_t, \quad X_0 = x_0 \tag{6.79}$$

where $x_0 \in \mathbb{R}^d$ is a given starting value, μ and σ are the drift and diffusion functions, B is a d-dimensional Brownian motion, and the stochastic process X is the solution. This Itô-SDE is the short notation for the integral equation

$$X_t = x_0 + \int_0^t \mu(X_s)ds + \int_0^t \sigma(X_s)dB_s \tag{6.80}$$

where the second integral is the Itô integral of the diffusion w.r.t the Brownian motion. This integral can be defined as

$$\int_0^t \sigma(X_s)dB_s = \lim_{|\Delta^k|\to 0} \sum_{i=1}^{k-1} \sigma(X_{t_i^k})(B_{t_{i+1}^k} - B_{t_i^k}) \tag{6.81}$$

where $\Delta^k = \{0 = t_1^k < t_2^k \leq ... \leq t_{k-1}^k < t_k^k = T\}$ is a sequence of partitions of the time interval $[0,T]$. The Itô integral is thus defined as a limit of Riemann sums where the diffusion function is evaluated at the left most boundary of the subinterval $[t_i^k, t_{i+1}^k]$. The Stratonovich intergal is defined differently: here we average the function values of the diffusion function at both boundaries of $[t_i^k, t_{i+1}^k]$, that is,

$$\int_0^t \sigma(X_s) \circ dB_s = \lim_{|\Delta^k|\to 0} \sum_{i=1}^{k-1} \frac{\sigma(X_{t_{i+1}^k}) + \sigma(X_{t_i^k})}{2}(B_{t_{i+1}^k} - B_{t_i^k}). \tag{6.82}$$

The corresponding Stratonovich-SDE is denoted as

$$X_t = \mu(X_t)dt + \sigma(X_t) \circ dB_t, \quad X_0 = x_0. \tag{6.83}$$

The Itô and Stratonovich integrals are not equivalent.

Remark 6.6.2.

1. *The SDE (6.79) can be discretized with a simple Euler-Maruyama method*

$$X_{t_{k+1}} = X_{t_k} + \mu(X_{t_k})\Delta t + \sigma(X_{t_k})\Delta B_{t_k} \tag{6.84}$$

where $\Delta t = t_{k+1} - t_k > 0$ is a given time step and $\Delta B_{t_k} = B_{t_{k+1}} - B_{t_k} \sim \mathcal{N}_d(0, (\Delta t)I_d)$ are the d-dimensional Brownian increments. The Brownian increments can be generated with

the Matlab function `randn`*, for the complete Brownian motion we can use* `cumsum`*. We note:* $\mathcal{N}(0, \Delta t) = (\Delta t)^{1/2}\mathcal{N}(0, 1)$.

2. *In this thesis we usually project the d-dimensional Brownian motion onto the hypersurface* Γ *with the projection* P_Γ*, see* (3.43).

Definition 6.6.4 (Finite-dimensional distributions, see [Øksendal, 2003])**.** *Let* $X : [0, T] \times \Omega \to \mathbb{R}^d$ *a stochastic process on the probability space* $(\Omega, \mathcal{A}, P)$ *and* $t_1, ..., t_k \in [0, T]$*. The finite-dimensional distributions of* X *are the probability measures* $\mu_{t_1,...,t_k}$ *in* $(\mathbb{R}^d)^k$ *defined as*

$$\mu_{t_1,...,t_k}(B_1 \times ... \times B_k) = P(X_{t_1} \in B_1, ..., X_{t_k} \in B_k) \tag{6.85}$$

where $B_1, ..., B_k$ *are Borel sets in* $\mathbb{R}^d$.

Theorem 6.6.1 (Kolmogorov extension theorem, see [Øksendal, 2003])**.** *Let* $k \in \mathbb{N}$ *and* $\mu_{t_1,...,t_k}$ *a probability measure on* $(\mathbb{R}^d)^k$ *for all* $t_1, ..., t_k \in [0, T]$ *satisfying the following two conditions.*

1. *Let* $\sigma \in S_k$ *a permutation and* $B_1, ..., B_k$ *Borel sets in* $\mathbb{R}^d$*. Then*

$$\mu_{t_{\sigma(1)},...,t_{\sigma(k)}}(B_1 \times ... \times B_k) = \mu_{t_1,...,t_k}(B_{\sigma^{-1}(1)} \times ... \times B_{\sigma^{-1}(k)}). \tag{6.86}$$

2. *For all* $m \in \mathbb{N}$ *we have*

$$\mu_{t_1,...,t_k}(B_1 \times ... \times B_k) = \mu_{t_1,...,t_k,t_{k+1},...,t_{k+m}}(B_1 \times ... \times B_k \times \mathbb{R}^d \times ... \times \mathbb{R}^d) \tag{6.87}$$

where the set on the right hand side has $k + m$ *factors.*

Then there exists a probability space $(\Omega, \mathcal{A}, P)$ *and a stochastic process* $X : [0, T] \times \Omega \to \mathbb{R}^d$ *such that*

$$\mu_{t_1,...,t_k}(B_1 \times ... \times B_k) = P(X_{t_1} \in B_1, ..., X_{t_k} \in B_k) \tag{6.88}$$

for all $t_i \in [0, T]$*,* $k \in \mathbb{N}$ *and Borel sets* $B_1, ..., B_k$.

Remark 6.6.3.

1. *The Kolomogrov extension theorem is also known as Kolomogrov existence theorem and Kolomogrov consistency theorem.*

2. *Given a family of probability measures* $\{\mu_{t_1,...,t_k} \mid k \in \mathbb{N}, t_1, ..., t_k \in [0, T]\}$ *satisfying the two conditions from above, the Kolmogorov extension theorem guarantees the existence of a stochastic process* $X : [0, T] \times \Omega \to \mathbb{R}^d$ *having these probability measures as finite-dimensional measures.*

3. *We used the Kolmogorov extension theorem in section 3.4 to guarantee the existence of a common (!) probability space* Ω *for the particles* $V^i : [0, T] \times \Omega^i \to \mathbb{R}^d$*. In this case the familiy of probability measures* $\{\mu_{t_1,...,t_k} \mid k \in \mathbb{N}, t_1, ..., t_k \in [0, T]\}$ *are the distributions of the particles.*

The following theorem can be found in [Baldi, 2017, Thm. 7.1].

Theorem 6.6.2 (Itô's isometry)**.** *Let $L^2_{ad}([0,T]\times\Omega)$ the vector space of square integrable stochastic processes adapted to the natural filtration of the Brownian motion $\mathcal{F}^B$ with the inner product*

$$\langle X, Y\rangle_{L^2_{ad}([0,T]\times\Omega)} = \mathbb{E}\big(\int_0^T X_t Y_t dt\big). \tag{6.89}$$

As usual we denote with $L^2(\Omega)$ the space of square integrable random variables with inner product $\langle X, Y\rangle_{L^2(\Omega)} = \mathbb{E}(XY)$. Itô's isometry states that Itô's integral $L^2_{ad}([0,T]\times\Omega) \mapsto L^2(\Omega)$ is an isometry. In particular

$$\mathbb{E}\big((\int_0^T X_t dB_t)^2\big) = \mathbb{E}\big(\int_0^T X_t^2 dt\big). \tag{6.90}$$

Anisotropic KV-CBO model

In this part of the appendix we report detailed proofs for the KV-CBO model with anisotropic (componentwise) noise, see (3.50) for the special case $\Gamma = \mathbb{S}^{d-1}$. The study of CBO models with anisotropic noise originates in [Carrillo et al., 2020]. The motivation to consider this noise is to avoid the explicit dimension-dependence of the constant $C_{\sigma,d} = (d-1)\sigma^2/2$ (see section 2.3). Here the constant $C_{\sigma,d}$ is replaced by

$$C_\sigma = \sigma^2/2 \tag{6.91}$$

and the conditions on the coefficients is $\lambda\vartheta - 16C_\sigma > 0$, see (6.177) below.

The proofs for the anisotropic model are more technical as the correction term is more complicated: we have

$$\begin{aligned} C^{ani}_{\mathbb{S}^{d-1}}(V_t^i, V^\alpha(\rho_t^N)) = &-\frac{\sigma^2}{2}|V_t^i - V^\alpha(\rho_t^N)|^2 \frac{V_t^i}{|V_t^i|^2}dt - \frac{\sigma^2}{2} D(V_t^i - V^\alpha(\rho_t^N))^2 \frac{V_t^i}{|V_t^i|^2}dt \\ &+ \sigma^2 |D(V_t^i - V^\alpha(\rho_t^N))V_t^i|^2 \frac{V_t^i}{|V_t^i|^4}dt. \end{aligned} \tag{6.92}$$

This part of the appendix is based on joint work of Massimo Fornasier, Hui Huang, Lorenzo Pareschi and myself and has been published in [Fornasier et al., 2021b]. In that publication the proffs are not reported in such detail as here.

Well-posedness of the particle system

As in the isotropic case we regularize the coefficient functions P_Γ, $v \mapsto v/|v|^2$, and $v \mapsto v/|v|^4$ with $\mathbf{P}_1$, $\mathbf{P}_2$, and $\mathbf{P}_3$ respectively. We introduce the following *regularized KV-CBO model with anisotropic noise*

$$dV_t^i = \lambda \mathbf{P}_1(V_t^i)\mathbf{V}^\alpha(\rho_t^N)dt + \sigma \mathbf{P}_1(V_t^i)D(V_t^i - \mathbf{V}^\alpha(\rho_t^N))dB_t^i + \mathbf{C}^{ani}_{\mathbb{S}^{d-1}}(V_t^i, \mathbf{V}^\alpha(\rho_t^N)) \tag{6.93}$$

for $i \in [N]$ where the anisotropic correction term is given by

$$\begin{aligned} \mathbf{C}^{ani}_{\mathbb{S}^{d-1}}(V_t^i, \mathbf{V}^\alpha(\rho_t^N)) = &-\frac{\sigma^2}{2}|V_t^i - \mathbf{V}^\alpha(\rho_t^N)|^2 \mathbf{P}_2(V_t^i)dt - \frac{\sigma^2}{2} D(V_t^i - \mathbf{V}^\alpha(\rho_t^N))^2 \mathbf{P}_2(V_t^i)dt \\ &+ \sigma^2 |D(V_t^i - \mathbf{V}^\alpha(\rho_t^N))V_t^i|^2 \mathbf{P}_3(V_t^i)dt. \end{aligned} \tag{6.94}$$

Theorem 6.6.3 (Well-posedness, cf. theorem 3.7.1)**.** *The KV-CBO model with anisotropic noise* (3.50) *is well-posed, that is, for every initial probability distribution $\rho_0 \in \mathcal{P}(\mathbb{S}^{d-1})$ there is a pathwise unique strong solution.*

Proof. The model (3.50) has locally Lipschitz continuous coefficients which implies the existence of a pathwise unique *local* (!) strong solution, see theorem 2.4.1.

We now apply Itô's formula with $f(x) = |x|^2$ to show that $d|V_t^i|^2 = 0$, hence the particles remain on the sphere. We find

$$\begin{aligned} d|V_t^i|^2 = \; & \nabla f(V_t^i) \cdot \bigg(\lambda \mathbf{P}_1(V_t^i)\mathbf{V}^\alpha(\rho_t^N) - \frac{\sigma^2}{2}|V_t^i - \mathbf{V}^\alpha(\rho_t^N)|^2 \mathbf{P}_2(V_t^i) \\ & - \frac{\sigma^2}{2} D(V_t^i - \mathbf{V}^\alpha(\rho_t^N))^2 \mathbf{P}_2(V_t^i) + \sigma^2 |D(V_t^i - \mathbf{V}^\alpha(\rho_t^N))V_t^i|^2 \mathbf{P}_3(V_t^i) \bigg) dt \\ & + \frac{1}{2}\sum_{k,l} (\nabla^2 f(V_t))_{kl} \sigma_k(t)\sigma_l(t)^T dt + \nabla f(V_t^i) \cdot \sigma(t) dB_t^i \end{aligned} \tag{6.95}$$

where

$$\sigma(t) = \sigma \mathbf{P}_1(V_t^i) D(V_t^i - \mathbf{V}^\alpha(\rho_t^N)) \in \mathbb{R}^{d \times d} \tag{6.96}$$

and σ_k is the kth row.

We can simplify this expression by noting that $\nabla f(V_t^i) = 2V_t^i$, $\nabla^2 f(V_t^i) = 2$, and consequently that $\nabla f(V_t^i) \cdot \lambda \mathbf{P}_1(V_t^i)\mathbf{V}^\alpha(\rho_t^N) = 0$, and $\nabla f(V_t^i) \cdot \sigma(t) = 0$ (note that $\mathbf{P}_1 = P_{\mathbb{S}^{d-1}}$). Thus, the above expression simplifies to

$$\begin{aligned} d|V_t^i|^2 = \; & 2V_t^i \cdot \bigg(- \frac{\sigma^2}{2}|V_t^i - \mathbf{V}^\alpha(\rho_t^N)|^2 \mathbf{P}_2(V_t^i) - \frac{\sigma^2}{2} D(V_t^i - \mathbf{V}^\alpha(\rho_t^N))^2 \mathbf{P}_2(V_t^i) \\ & + \sigma^2 |D(V_t^i - \mathbf{V}^\alpha(\rho_t^N))V_t^i|^2 \mathbf{P}_3(V_t^i) \bigg) dt + \sum_{k,l} \sigma_k(t)\sigma_l(t)^T dt. \end{aligned} \tag{6.97}$$

We further note

$$V_t^i \cdot \mathbf{P}_2(V_t^i) = 1, \tag{6.98}$$

$$2V_t^i \cdot \frac{\sigma^2}{2} D(V_t^i - \mathbf{V}^\alpha(\rho_t^N))^2 \mathbf{P}_2(V_t^i) = \sigma^2 \frac{|D(V_t^i - \mathbf{V}^\alpha(\rho_t^N))V_t^i|^2}{|V_t^i|^2}, \tag{6.99}$$

$$2V_t^i \cdot \sigma^2 |D(V_t^i - \mathbf{V}^\alpha(\rho_t^N))V_t^i|^2 \mathbf{P}_3(V_t^i) = 2\sigma^2 \frac{|D(V_t^i - \mathbf{V}^\alpha(\rho_t^N))V_t^i|^2}{|V_t^i|^2}. \tag{6.100}$$

Now we can simplify (6.97) even further as

$$\begin{aligned} d|V_t^i|^2 = \; & \bigg(- \sigma^2 |V_t^i - \mathbf{V}^\alpha(\rho_t^N)|^2 + \sigma^2 \frac{|D(V_t^i - \mathbf{V}^\alpha(\rho_t^N))V_t^i|^2}{|V_t^i|^2} \bigg) dt \\ & + \sum_{k,l} \sigma_k(t)\sigma_l(t)^T dt. \end{aligned} \tag{6.101}$$

The kth row of $\sigma(t)$ is given by

$$\sigma_k(t) = \sigma(V_t^i - \mathbf{V}^\alpha(\rho_t^N))_k (e_k^T - (V_t^i)_k (V_t^i)^T) \tag{6.102}$$

hence

$$\sum_{k,l}\sigma_k(t)\sigma_l(t)^T dt = \sigma^2|V_t^i - \mathbf{V}^\alpha(\rho_t^N)|^2 - 2\sigma^2\sum_k (V_t^i - \mathbf{V}^\alpha(\rho_t^N))_k \frac{(V_t^i)_k^2}{|V_t^i|^2} \tag{6.103}$$

$$+\sigma^2\sum_{k,l}(V_t^i - \mathbf{V}^\alpha(\rho_t^N))_k^2\frac{(V_t^i)_l^2(V_t^i)_k^2}{|V_t^i|^2} \tag{6.104}$$

$$= \sigma^2|V_t^i - \mathbf{V}^\alpha(\rho_t^N)|^2 - 2\sigma^2\frac{|D(V_t^i - \mathbf{V}^\alpha(\rho_t^N)V_t^i|^2}{|V_t^i|^2} \tag{6.105}$$

$$+\sigma^2\frac{|D(V_t^i - \mathbf{V}^\alpha(\rho_t^N)V_t^i|^2}{|V_t^i|^2} \tag{6.106}$$

$$= \sigma^2|V_t^i - \mathbf{V}^\alpha(\rho_t^N)|^2 - \sigma^2\frac{|D(V_t^i - \mathbf{V}^\alpha(\rho_t^N)V_t^i|^2}{|V_t^i|^2}. \tag{6.107}$$

Now we see that (6.101) gives $d|V_t^i|^2 = 0$, which ensures that the solution stays on the sphere for some finite time. We can therefore extend the pathwise unique local solution to a pathwise uniquen global solution. This further shows that the regularized KV-CBO model with anisotropic noise (6.93) equals the original model with anisotropic noise (3.50). □

Well-posedness of the mean-field dynamic

Theorem 6.6.4 (cf. theorem 4.2.1)**.** *The mean-field dynamic*

$$\begin{cases} d\overline{V}_t = \lambda P_{\mathbb{S}^{d-1}}(\overline{V}_t)V^\alpha(\rho_t)dt + \sigma P_{\mathbb{S}^{d-1}}(\overline{V}_t)D(\overline{V}_t - V^\alpha(\rho_t))dB_t + C^{ani}_{\mathbb{S}^{d-1}}(\overline{V}_t, V^\alpha(\rho_t))dt \\ \rho_t = law(\overline{V}_t) \end{cases} \tag{6.108}$$

is well-posed, that is, there is a pathwise unique strong solution for any initial probability distribution $\rho_0 \in \mathcal{P}(\mathbb{S}^{d-1})$.

Proof. Let $\xi \in C([0,T],\mathbb{R}^d)$ and consider the SDE

$$d\overline{V}_t^\xi = \lambda\mathbf{P}_1(\overline{V}_t^\xi)\xi_t dt + \sigma\mathbf{P}_1(\overline{V}_t^\xi)D(\overline{V}_t^\xi - \xi_t)dB_t + \mathbf{C}^{ani}_{\mathbb{S}^{d-1}}(\overline{V}_t^\xi, \xi_t) \tag{6.109}$$

with initial data $\overline{V}_0^\xi \sim \rho_0$ for $\rho_0 \in \mathcal{P}(\Gamma)$. This SDE is well-posed and $V_t^\xi \in \mathbb{S}^{d-1}$ for all time t which allows us to use the unregularized operators $P_{\mathbb{S}^{d-1}}$ and $C^{ani}_{\mathbb{S}^{d-1}}$ instead. Thus the SDE introduces a measure $\rho_t^\xi \in C([0,T],\mathcal{P}(\Gamma))$ with $\rho_t^\xi = law(\overline{V}_t^\xi)$.

We define the operator

$$T : C([0,T],\mathbb{R}^d)\times[0,1] \to C([0,T],\mathbb{R}^d), \quad (\xi,\vartheta)\mapsto \mathbf{V}^\alpha(\rho^\xi) \tag{6.110}$$

for which we prove the existence of a unique fixed point $\xi^\star$. We use the Leray-Schauder fixed point theorem to show the existence of a (not necessarily unique) fixed point ξ of T.

Compactness of T: For the solution $\overline{V}$ of (6.108) we find (similarly to lemma 3.6.4)

$$|V^\alpha(\rho_t) - V^\alpha(\rho_s)|^2 \lesssim \mathbb{E}|\overline{V}_t - \overline{V}_s|^2 \tag{6.111}$$

$$\lesssim \mathbb{E}|\int_s^t \lambda P_{\mathbb{S}^{d-1}}(\overline{V}_\tau)V^\alpha(\rho_\tau)|^2 + \mathbb{E}|\int_s^t \sigma P_{\mathbb{S}^{d-1}}(\overline{V}_\tau)D(\overline{V}_\tau - V^\alpha(\rho_\tau))dB_\tau|^2$$

$$\mathbb{E}|\int_s^t C^{ani}_{\mathbb{S}^{d-1}}(\overline{V}_\tau, V^\alpha(\rho_\tau))d\tau|^2 \tag{6.112}$$

$$\lesssim |t-s|^2 \tag{6.113}$$

where we have used the compactness of $[s,t]$. With this we conclude that $t \mapsto V^\alpha(\rho_t)$ is Hölder continuous with exponent $1/2$. Hence $T\xi \in C^{0,1/2}([0,T],\mathbb{R}^d)$ and the compactness of T follows from the fact that $C^{0,1/2}([0,T],\mathbb{R}^d)$ is compactly embedded in $C([0,T],\mathbb{R}^d)$.

Existence of fixed point of T: Follows as in lemma 3.6.4. By Leray-Schauder there exists a fixed point $\xi^\star$ of T, that is, $V^\alpha(\rho^{\xi^\star}) = \xi^\star$.

Uniqueness of the fixed point $\xi^\star$: Suppose there are two distinct fixed points $\xi^{\star,1}$ and $\xi^{\star,2}$. The corresponding processes are denoted by $\overline{V}_t^{\xi^{\star,i}}$ and $\rho_t^{\xi^{\star,i}}$ respectively. Further we define

$$\overline{Z}_t = \overline{V}_t^{\xi^{\star,1}} - \overline{V}_t^{\xi^{\star,2}}. \tag{6.114}$$

With the integral representation of (6.109), Itô's isometry, and the fact that (see (6.111))

$$|\xi_s^{\star,1} - \xi_s^{\star,2}| = |V^\alpha(\rho_s^{\xi^{\star,1}}) - V^\alpha(\rho_s^{\xi^{\star,2}})|^2 \lesssim \mathbb{E}|\overline{Z}_s|^2 \tag{6.115}$$

we find

$$\mathbb{E}|\overline{Z}_t|^2 \lesssim \mathbb{E}|\overline{Z}_0|^2 + \int_0^t \mathbb{E}|\overline{Z}_s|^2 ds. \tag{6.116}$$

Since the initial values of the solutions $\overline{V}^{\xi^{\star,1}}$ and $\overline{V}^{\xi^{\star,2}}$ are sampled from the same initial distribution ρ_0 we have $\mathbb{E}|\overline{Z}_0|^2 = 0$. Gronwall's inequality further yields $\mathbb{E}|\overline{Z}_t|^2 = 0$ for all $t \in [0,T]$, hence the uniqueness of the fixed point $\xi^\star$.

We conclude the proof by noting that the SDE (6.109) for $\xi^\star$ is exactly the mean-field dynamic (6.108). The well-posedness of the latter follows from the well-posedness of the former. □

Well-posedness of the mean-field equation

Before proving the well-posedness of the mean-field equation we discuss some integration by parts results on the sphere, see [Frouvelle and Liu, 2012, section 2.1]. We first recall the standard integration by parts formula on $\mathbb{S}^{d-1}$: for $\Phi \in C^\infty(\mathbb{S}^{d-1})$ and a vector field $\omega : \mathbb{S}^{d-1} \to T\mathbb{S}^{d-1}$, $\omega(v) \in T_v\mathbb{S}^{d-1}$ we have

$$\int_{\mathbb{S}^{d-1}} \nabla_{\mathbb{S}^{d-1}} \Phi(v) \cdot \omega(v) dv = -\int_{\mathbb{S}^{d-1}} \Phi(v) \nabla_{\mathbb{S}^{d-1}} \cdot \omega(v) dv. \tag{6.117}$$

Choosing $\omega(v) = P_{\mathbb{S}^{d-1}}(v)\phi(v)$ and appliying (6.117) twice yields

$$\int_{\mathbb{S}^{d-1}} (\Delta_{\mathbb{S}^{d-1}} \Phi(v))\phi(v) dv = \int_{\mathbb{S}^{d-1}} \Phi(v) \Delta_{\mathbb{S}^{d-1}} \phi(v) dv. \tag{6.118}$$

We note that formula (6.117) and (6.118) hold for any hypersurface Γ, whereas forumla (6.123) below is specific for the sphere.

We now generalize the integration by parts formula (6.117) for smooth functions $A : \mathbb{S}^{d-1} \to \mathbb{R}^d$ that are not necessarily tangential. For a constant vector $V \in \mathbb{R}^d$ we have

$$\nabla_{\mathbb{S}^{d-1}}(v \cdot V) = P_{\mathbb{S}^{d-1}}(v) \nabla (v \cdot V) = P_{\mathbb{S}^{d-1}}(v) V \tag{6.119}$$

where the gradient is wrt v, that is, $\nabla = \nabla_v$. For the divergence we further find

$$\nabla_{\mathbb{S}^{d-1}} \cdot (P_{\mathbb{S}^{d-1}}(v) V) = -\sum_{i=1}^{d} \big(\sum_{k=1}^{d} \partial_{\mathbb{S}^{d-1},i}(v_i v_k V_k) \big) = -(d-1) v \cdot V \tag{6.120}$$

where we have used that $P_{\mathbb{S}^{d-1}}(v) = I - v \otimes v$.

Since $P_{\mathbb{S}^{d-1}} V$ is tangential we can apply (6.117) to find

$$\int_{\mathbb{S}^{d-1}} \nabla_{\mathbb{S}^{d-1}} \Phi(v) \cdot P_{\mathbb{S}^{d-1}}(v) V dv = -\int_{\mathbb{S}^{d-1}} \Phi(v) \nabla_{\mathbb{S}^{d-1}} \cdot (P_{\mathbb{S}^{d-1}}(v) V) dv. \tag{6.121}$$

With property (6.120) and choosing $V = e_i$ we further find

$$\begin{cases} \int_{\mathbb{S}^{d-1}} \partial_{\mathbb{S}^{d-1},i} \Phi(v) dv &= (d-1) \int_{\mathbb{S}^{d-1}} \Phi(v) v_i dv \\ \int_{\mathbb{S}^{d-1}} \nabla_{\mathbb{S}^{d-1}} \Phi(v) dv &= (d-1) \int_{\mathbb{S}^{d-1}} \Phi(v) v dv. \end{cases} \tag{6.122}$$

With the above properties we arrive at

$$\int_{\mathbb{S}^{d-1}} \nabla_{\mathbb{S}^{d-1}} \Phi(v) \cdot A(v) dv = -\int_{\mathbb{S}^{d-1}} \Phi(v) \nabla_{\mathbb{S}^{d-1}} \cdot A(v) dv + (d-1) \int_{\mathbb{S}^{d-1}} A(v) \cdot v \Phi(v) dv. \tag{6.123}$$

This integration by parts formula is a generalization of (6.117) since the second term on the right hand side vanishes for A tangential.

Theorem 6.6.5 (cf. theorem 4.3.1)**.** *The mean-field equation of the anisotropic KV-CBO model*

$$\begin{aligned}\partial_t\rho_t = &-\lambda\nabla_{\mathbb{S}^{d-1}}\cdot(P_{\mathbb{S}^{d-1}}(v)V^\alpha(\rho_t)\rho_t) + \frac{\sigma^2}{2}\sum_{i=1}^d \partial^2_{\mathbb{S}^{d-1},ii}((v-V^\alpha(\rho_t))_i^2\rho_t)\\ &-\frac{\sigma^2}{2}(d-2)\nabla_{\mathbb{S}^{d-1}}\cdot(D(v-V^\alpha(\rho_t))^2v\rho_t) + \frac{\sigma^2}{2}(d-2)(d-1)|D(v-V^\alpha(\rho_t))v|^2\rho_t\\ &-\frac{\sigma^2}{2}(d-1)\sum_{i=1}^d \partial_{\mathbb{S}^{d-1},i}((v-V^\alpha(\rho_t))_i^2\rho_t)v_i. \end{aligned} \tag{6.124}$$

is well-posed, that is, there is a unique weak solution for any initial probability distribution $\rho_0 \in \mathcal{P}(\mathbb{S}^{d-1})$.

Proof. Let us first derive the mean-field equation.

Applying Dynkin's formula with a test function $\phi \in C_c^\infty(\mathbb{R}^d)$ to the mean-field dynamic (6.108) yields

$$\begin{aligned}\frac{d}{dt}\int_{\mathbb{R}^d}\phi(v)d\rho_t(v) = &\int_{\mathbb{R}^d}\nabla\phi(v)\cdot\big(\lambda P_{\mathbb{S}^{d-1}}(v)V^\alpha(\rho_t) - \frac{\sigma^2}{2}D(v-V^\alpha(\rho_t))^2v\big)d\rho_t(v)\\ &+\frac{\sigma^2}{2}\int_{\mathbb{R}^d}\mathrm{Tr}\big((P_{\mathbb{S}^{d-1}}(v)D(v-V^\alpha(\rho_t)))^T(\nabla^2\phi(v))P_{\mathbb{S}^{d-1}}(v)D(v-V^\alpha(\rho_t))\big)dt\end{aligned} \tag{6.125}$$

where we have used that ϕ is 0-homogeneous ($\nabla\phi(v)\cdot v = 0$) and consequently

$$\nabla\phi(v)\cdot C^{ani}_{\mathbb{S}^{d-1}}(v, V^\alpha(\rho_t)) = -\nabla\phi(v)\cdot\frac{\sigma^2}{2}D(v-V^\alpha(\rho_t))^2v. \tag{6.126}$$

With $P_{\mathbb{S}^{d-1}}(v) = I - v\otimes v$ we can rewrite the trace term as

$$\mathrm{Tr}\big((PD)^T\nabla^2\phi(v)PD\big) \tag{6.127}$$

$$= \mathrm{Tr}\big(DP\nabla^2\phi(v)PD\big) \tag{6.128}$$

$$= \mathrm{Tr}\big(D\nabla^2\phi(v)D\big) - 2\mathrm{Tr}\big(Dv\otimes v\nabla^2\phi(v)D\big) + \mathrm{Tr}\big(Dv\otimes v\nabla^2\phi(v)v\otimes vD\big) \tag{6.129}$$

$$\begin{aligned}= &\sum_{i=1}^d\partial^2_{ii}\phi(v)(v-V^\alpha(\rho_t))_i^2 - 2\sum_{i=1}^d(v-V^\alpha(\rho_t))_iv_i\partial_i\nabla\phi(v)\cdot v\\ &+\sum_{i,j=1}^d\partial^2_{ij}\phi(v)v_jv_i\,|D(v-V^\alpha(\rho_t))v|^2\end{aligned} \tag{6.130}$$

$$= \sum_{i=1}^d\partial^2_{ii}\phi(v)(v-V^\alpha(\rho_t))_i^2 \tag{6.131}$$

where we have (once again) used that ϕ is 0-homogeneous.

Combining the above statements and using the definitions of the differential operators $\nabla_{\mathbb{S}^{d-1}}$ and

$\Delta_{\mathbb{S}^{d-1}}$ we arrive at the following weak formulation (for $\Phi \in C^\infty(\mathbb{S}^{d-1})$) of the mean-field equation

$$\begin{aligned}\frac{d}{dt}\int_{\mathbb{S}^{d-1}} \Phi(v) d\rho_t(v) = \int_{\mathbb{S}^{d-1}} \nabla_{\mathbb{S}^{d-1}}\Phi(v)\cdot\big(\lambda P_{\mathbb{S}^{d-1}}(v)V^\alpha(\rho_t) - \frac{\sigma^2}{2}D(v-V^\alpha(\rho_t))^2 v\big)d\rho_t(v)\\ + \frac{\sigma^2}{2}\int_{\mathbb{S}^{d-1}}\sum_{i=1}^d (\partial^2_{\mathbb{S}^{d-1},ii}\Phi(v))(v-V^\alpha(\rho_t))_i^2 d\rho_t(v).\end{aligned} \tag{6.132}$$

where we have denoted the restriction of $\rho_t \in \mathcal{P}_c(\mathbb{R}^d)$ onto $\mathbb{S}^{d-1}$ again with ρ_t.

Let us now derive the corresponding strong formulation of the mean-field equation. With the integration by parts formulas from above we find

$$\begin{aligned}&\int_{\mathbb{S}^{d-1}} \nabla_{\mathbb{S}^{d-1}}\Phi(v)\cdot\big(\lambda P_{\mathbb{S}^{d-1}}(v)V^\alpha(\rho_t)\big)\rho_t(v)dv\\ &\overset{(6.117)}{=} -\int_{\mathbb{S}^{d-1}}\Phi(v)\nabla_{\mathbb{S}^{d-1}}\cdot\big(\lambda P_{\mathbb{S}^{d-1}}(v)V^\alpha(\rho_t)\rho_t(v)\big)dv\end{aligned} \tag{6.133}$$

and

$$\begin{aligned}&-\frac{\sigma^2}{2}\int_{\mathbb{S}^{d-1}}\nabla_{\mathbb{S}^{d-1}}\Phi(v)\cdot\big(D(v-V^\alpha(\rho_t))^2 v\rho_t(v)\big)dv\\ &\overset{(6.123)}{=} \frac{\sigma^2}{2}\int_{\mathbb{S}^{d-1}}\Phi(v)\nabla_{\mathbb{S}^{d-1}}\cdot\big(D(v-V^\alpha(\rho_t))^2 v\rho_t(v)\big)dv\\ &\quad -\frac{\sigma^2}{2}(d-1)\int_{\mathbb{S}^{d-1}} D(v-V^\alpha(\rho_t))^2 v\rho_t(v)\cdot v\Phi(v)dv\end{aligned} \tag{6.134}$$

for the first term on the right hand side of (6.132). We note that we need to apply the (more complex) formula (6.123) for the second term because $D(v-V^\alpha(\rho_t))v$ is not tangential (as opposed to $P_{\mathbb{S}^{d-1}}(v)V^\alpha(\rho_t)$).

For the third term on the right hand side of (6.132) we find

$$\begin{aligned}&\frac{\sigma^2}{2}\int_{\mathbb{S}^{d-1}}\sum_{i=1}^d(\partial^2_{\mathbb{S}^{d-1},ii}\Phi(v))(v-V^\alpha(\rho_t))_i^2 d\rho_t(v)\\ &= \frac{\sigma^2}{2}\sum_{i=1}^d\int_{\mathbb{S}^{d-1}}\nabla_{\mathbb{S}^{d-1}}g_i(v)\cdot(v-V^\alpha(\rho_t))^2\rho_t(v)dv\end{aligned} \tag{6.135}$$

$$\begin{aligned}&\overset{(6.123)}{=} -\frac{\sigma^2}{2}\sum_{i=1}^d\int_{\mathbb{S}^{d-1}} g_i(v)\nabla_{\mathbb{S}^{d-1}}\cdot\big((v-V^\alpha(\rho_t))^2\rho_t(v)\big)dv\\ &\quad +\frac{\sigma^2}{2}(d-1)\sum_{i=1}^d\int_{\mathbb{S}^{d-1}}\big((v-V^\alpha(\rho_t))^2\rho_t(v)\big)\cdot v g_i(v)dv\end{aligned} \tag{6.136}$$

where

$$g_i(v) = \partial_{\mathbb{S}^{d-1},i}\Phi(v) \tag{6.137}$$

and the square in $(v-V^\alpha(\rho_t))^2$ is understood componentwise. Defining $B:\mathbb{S}^{d-1}\to\mathbb{R}^d$ with (constant) components $(B(v))_i=\nabla_{\mathbb{S}^{d-1}}\cdot(v-V^\alpha(\rho_t))^2$ allows us to rewrite the first term on the right hand side of (6.136) as

$$
\begin{aligned}
&-\frac{\sigma^2}{2}\sum_{i=1}^d\int_{\mathbb{S}^{d-1}} g_i(v)\nabla_{\mathbb{S}^{d-1}}\cdot\big((v-V^\alpha(\rho_t))^2\rho_t(v)\big)dv \\
&\quad=-\frac{\sigma^2}{2}\int_{\mathbb{S}^{d-1}}\nabla_{\mathbb{S}^{d-1}}\Phi(v)\cdot B(v)dv \qquad (6.138)\\
&\quad\overset{(6.123)}{=}\int_{\mathbb{S}^{d-1}}\Phi(v)\frac{\sigma^2}{2}\sum_{i=1}^d\partial^2_{\mathbb{S}^{d-1},ii}((v-V^\alpha(\rho_t))_i^2\rho_t)dv \\
&\quad-\frac{\sigma^2}{2}(d-1)\int_{\mathbb{S}^{d-1}}\Phi(v)\nabla_{\mathbb{S}^{d-1}}\cdot\big(D(v-V^\alpha(\rho_t))^2v\rho_t(v)\big) \qquad (6.139)
\end{aligned}
$$

Collecting the above statements allows us to write the strong formulation of the mean-field equation of the anisotropic KV-CBO model as

$$
\begin{aligned}
\partial_t\rho_t=&-\lambda\nabla_{\mathbb{S}^{d-1}}\cdot(P_{\mathbb{S}^{d-1}}(v)V^\alpha(\rho_t)\rho_t)+\frac{\sigma^2}{2}\sum_{i=1}^d\partial^2_{\mathbb{S}^{d-1},ii}((v-V^\alpha(\rho_t))_i^2\rho_t)\\
&-\frac{\sigma^2}{2}(d-2)\nabla_{\mathbb{S}^{d-1}}\cdot(D(v-V^\alpha(\rho_t))^2v\rho_t)+\frac{\sigma^2}{2}(d-2)(d-1)|D(v-V^\alpha(\rho_t))v|^2\rho_t\\
&-\frac{\sigma^2}{2}(d-1)\sum_{i=1}^d\partial_{\mathbb{S}^{d-1},i}((v-V^\alpha(\rho_t))_i^2\rho_t)v_i. \qquad (6.140)
\end{aligned}
$$

Let us now investigate the well-posedness of the mean-field equation.

Existence: The existence of a weak solution follows from the construction above.

Uniqueness: The proof of the uniqueness of the weak solution is the same as for the isotropic model. We summarize the important steps. First, we assume that $\rho_t^i\in\mathcal{P}(\mathbb{S}^{d-1})$ for $i=1,2$ are two weak solutions of the mean-field equation with the same initial data $\rho_0\in\mathcal{P}(\mathbb{S}^{d-1})$. Second, we consstruct linear (!) stochastic processes $\overline{V}_t^i$ corresponding to ρ_t^i as

$$
d\overline{V}_t^i=\lambda\mathbf{P}_1(\overline{V}_t^i)V^\alpha(\rho_t^i)dt+\sigma\mathbf{P}_1(\overline{V}_t^i)D(\overline{V}_t^i-V^\alpha(\rho_t^i))dB_t+\mathbf{C}^{ani}_{\mathbb{S}^{d-1}}(\overline{V}_t^i,V^\alpha(\rho_t^i)). \qquad (6.141)
$$

We denote the laws as $\mu_t^i=law(\overline{V}_t^i)$. Third, μ_t^i are solutions to a linear (!) Fokker-Planck equation (which has a unique solution). Further, also ρ_t^i is a solution to this (nonlinear) Fokker-Planck equation (6.125). Hence, $\mu_t^i=\rho_t^i$ for $i=1,2$. Last, the well-posedness of the mean-field dynamic yields $\mu_t^1=\rho_t^1=\rho_t^2=\mu_t^2$ which finishes the proof. □

Mean-field limit

Theorem 6.6.6 (cf. theorem 4.4.1)**.** *Let V_t^i denote the anisotropic KV-CBO particles and $\overline{V}_t^i$ iid copies of the mean-field dynamic with common probability distribution $\rho_t = law(\overline{V}_t^i)$. Assume further that the initial data and the Brownian motions are the same, that is, $V_0^i = \overline{V}_0^i$ and $B_t^i = \overline{B}_t^i$. Then*

$$\sup_{t\in[0,T]} \mathbb{E}|V_t^i - \overline{V}_t^i|^2 \lesssim N^{-1} \tag{6.142}$$

for all $i \in [N]$. We note that the constant in $\lesssim$ depends on the dimension d.

Proof. The particles V_t^i and $\overline{V}_t^i$ are solutions to the SDEs

$$dV_t^i = \lambda P_{\mathbb{S}^{d-1}}(V_t^i)V^\alpha(\rho_t^N)dt + \sigma P_{\mathbb{S}^{d-1}}(V_t^i)D(V_t^i - V^\alpha(\rho_t^N))dB_t^i + C^{ani}_{\mathbb{S}^{d-1}}(V_t^i, V^\alpha(\rho_t^N)) \tag{6.143}$$

and

$$d\overline{V}_t^i = \lambda P_{\mathbb{S}^{d-1}}(\overline{V}_t^i)V^\alpha(\rho_t)dt + \sigma P_{\mathbb{S}^{d-1}}(\overline{V}_t^i)D(\overline{V}_t^i - V^\alpha(\rho_t))d\overline{B}_t^i + C^{ani}_{\mathbb{S}^{d-1}}(\overline{V}_t^i, V^\alpha(\rho_t))dt \tag{6.144}$$

respectively. Since the driving Brownian motions are the same we find an equation for the difference

$$\begin{aligned} d(V_t^i - \overline{V}_t^i) = \lambda&\Big(P_{\mathbb{S}^{d-1}}(V_t^i)V^\alpha(\rho_t^N) - P_{\mathbb{S}^{d-1}}(\overline{V}_t^i)V^\alpha(\rho_t)\Big)dt \\ &+ \sigma\Big(P_{\mathbb{S}^{d-1}}(V_t^i)D(V_t^i - V^\alpha(\rho_t^N)) - P_{\mathbb{S}^{d-1}}(\overline{V}_t^i)D(\overline{V}_t^i - V^\alpha(\rho_t))\Big)dB_t^i \\ &+ \Big(C^{ani}_{\mathbb{S}^{d-1}}(V_t^i, V^\alpha(\rho_t^N)) - C^{ani}_{\mathbb{S}^{d-1}}(\overline{V}_t^i, V^\alpha(\rho_t))\Big)dt \end{aligned} \tag{6.145}$$

An application of Itô's formula with $f(x) = |x|^2$, $\nabla f(x) = 2x$, and $\Delta f(x) = 2I_d$ yields

$$\begin{aligned} d|V_t^i - \overline{V}_t^i|^2 = 2\lambda(V_t^i - \overline{V}_t^i)\cdot&\Big(P_{\mathbb{S}^{d-1}}(V_t^i)V^\alpha(\rho_t^N) - P_{\mathbb{S}^{d-1}}(\overline{V}_t^i)V^\alpha(\rho_t)\Big)dt \\ &+ 2(V_t^i - \overline{V}_t^i)\cdot\Big(C^{ani}_{\mathbb{S}^{d-1}}(V_t^i, V^\alpha(\rho_t^N)) - C^{ani}_{\mathbb{S}^{d-1}}(\overline{V}_t^i, V^\alpha(\rho_t))\Big)dt + \sigma^2\mathrm{Tr}(\sigma^i\cdot\sigma^i)dt \\ &+ 2\sigma(V_t^i - \overline{V}_t^i)\cdot\Big(P_{\mathbb{S}^{d-1}}(V_t^i)D(V_t^i - V^\alpha(\rho_t^N)) - P_{\mathbb{S}^{d-1}}(\overline{V}_t^i)D(\overline{V}_t^i - V^\alpha(\rho_t))\Big)dB_t^i \end{aligned} \tag{6.146}$$

where

$$\sigma^i = P_{\mathbb{S}^{d-1}}(V_t^i)D(V_t^i - V^\alpha(\rho_t^N)) - P_{\mathbb{S}^{d-1}}(\overline{V}_t^i)D(\overline{V}_t^i - V^\alpha(\rho_t)). \tag{6.147}$$

We estimate the first term on the right hand side of (6.146) as (see (4.68))

$$\begin{aligned} &\lambda(V_t^i - \overline{V}_t^i)\cdot\Big(P_{\mathbb{S}^{d-1}}(V_t^i)V^\alpha(\rho_t^N) - P_{\mathbb{S}^{d-1}}(\overline{V}_t^i)V^\alpha(\rho_t)\Big) \\ &\quad\lesssim \sum_{i=1}^N \frac{|V_t^i - \overline{V}_t^i|^2}{N} + |V_t^i - \overline{V}_t^i|^2 + |V^\alpha(\overline{\rho}_t^N) - V^\alpha(\rho_t)|^2 \end{aligned} \tag{6.148}$$

where $\overline{\rho}_t^N$ denotes the empirical measure of the iid copies $\overline{V}_t^i$.

We estimate the second term on the right hand side of (6.146) as

$$
\begin{aligned}
&2(V_t^i-\overline{V}_t^i)\cdot\Big(C^{ani}_{\mathbb{S}^{d-1}}(V_t^i,V^\alpha(\rho_t^N))-C^{ani}_{\mathbb{S}^{d-1}}(\overline{V}_t^i,V^\alpha(\rho_t))\Big)\\
&\lesssim 2|V_t^i-\overline{V}_t^i|\cdot\Big|\Big(-\frac{\sigma^2}{2}|V_t^i-V^\alpha(\rho_t^N)|^2\frac{V_t^i}{|V_t^i|^2}-\frac{\sigma^2}{2}D(V_t^i-V^\alpha(\rho_t^N))^2\frac{V_t^i}{|V_t^i|^2}+\sigma^2|D(V_t^i-V^\alpha(\rho_t^N))V_t^i|^2\frac{V_t^i}{|V_t^i|^4}\Big)\\
&\qquad-\Big(-\frac{\sigma^2}{2}|V_t^i-V^\alpha(\rho_t^N)|^2\frac{\overline{V}_t^i}{|\overline{V}_t^i|^2}-\frac{\sigma^2}{2}D(V_t^i-V^\alpha(\rho_t^N))^2\frac{\overline{V}_t^i}{|\overline{V}_t^i|^2}+\sigma^2|D(V_t^i-V^\alpha(\rho_t^N))V_t^i|^2\frac{\overline{V}_t^i}{|\overline{V}_t^i|^4}\Big)\\
&\qquad+\Big(-\frac{\sigma^2}{2}|V_t^i-V^\alpha(\rho_t^N)|^2\frac{\overline{V}_t^i}{|\overline{V}_t^i|^2}-\frac{\sigma^2}{2}D(V_t^i-V^\alpha(\rho_t^N))^2\frac{\overline{V}_t^i}{|\overline{V}_t^i|^2}+\sigma^2|D(V_t^i-V^\alpha(\rho_t^N))V_t^i|^2\frac{\overline{V}_t^i}{|\overline{V}_t^i|^4}\Big)\\
&\qquad-\Big(-\frac{\sigma^2}{2}|V_t^i-V^\alpha(\overline{\rho}_t^N)|^2\frac{\overline{V}_t^i}{|\overline{V}_t^i|^2}-\frac{\sigma^2}{2}D(V_t^i-V^\alpha(\overline{\rho}_t^N))^2\frac{\overline{V}_t^i}{|\overline{V}_t^i|^2}+\sigma^2|D(V_t^i-V^\alpha(\overline{\rho}_t^N))V_t^i|^2\frac{\overline{V}_t^i}{|\overline{V}_t^i|^4}\Big)\\
&\qquad+\Big(-\frac{\sigma^2}{2}|V_t^i-V^\alpha(\overline{\rho}_t^N)|^2\frac{\overline{V}_t^i}{|\overline{V}_t^i|^2}-\frac{\sigma^2}{2}D(V_t^i-V^\alpha(\overline{\rho}_t^N))^2\frac{\overline{V}_t^i}{|\overline{V}_t^i|^2}+\sigma^2|D(V_t^i-V^\alpha(\overline{\rho}_t^N))V_t^i|^2\frac{\overline{V}_t^i}{|\overline{V}_t^i|^4}\Big)\\
&\qquad-\Big(-\frac{\sigma^2}{2}|\overline{V}_t^i-V^\alpha(\rho_t)|^2\frac{\overline{V}_t^i}{|\overline{V}_t^i|^2}-\frac{\sigma^2}{2}D(\overline{V}_t^i-V^\alpha(\rho_t))^2\frac{\overline{V}_t^i}{|\overline{V}_t^i|^2}+\sigma^2|D(\overline{V}_t^i-V^\alpha(\rho_t))\overline{V}_t^i|^2\frac{\overline{V}_t^i}{|\overline{V}_t^i|^4}\Big)\Big|\\
&\lesssim 2|V_t^i-\overline{V}_t^i|\cdot\Big||V_t^i-\overline{V}_t^i|+(|V_t^i-V^\alpha(\rho_t^N)|^2-|V_t^i-V^\alpha(\overline{\rho}_t^N)|^2)+(|V_t^i-V^\alpha(\overline{\rho}_t^N)|^2-|\overline{V}_t^i-V^\alpha(\rho_t)|^2)\Big|
\end{aligned}
\tag{6.149}
\tag{6.150}
$$

where we have used that $|V_t^i-V^\alpha(\rho_t^N)|^2\lesssim 1$, $|V_t^i|=1$, $|\overline{V}_t^i|=1$.

As in the proof of theorem 4.4.1 we arrive at

$$
\begin{aligned}
&2(V_t^i-\overline{V}_t^i)\cdot\Big(C^{ani}_{\mathbb{S}^{d-1}}(V_t^i,V^\alpha(\rho_t^N))-C^{ani}_{\mathbb{S}^{d-1}}(\overline{V}_t^i,V^\alpha(\rho_t))\Big)\\
&\qquad\lesssim|V_t^i-\overline{V}_t^i|^2+\sum_{i=1}^N\frac{|V_t^i-\overline{V}_t^i|^2}{N}+|V^\alpha(\overline{\rho}_t^N)-V^\alpha(\rho_t)|^2.
\end{aligned}
\tag{6.151}
$$

Further we find

$$
\mathrm{Tr}(\sigma^i\cdot\sigma^i)\lesssim d|V_t^i-\overline{V}_t^i|^2+\frac{\sum_{i=1}^N d|V_t^i-\overline{V}_t^i|^2}{N}+d|V^\alpha(\overline{\rho}_t^N)-V^\alpha(\rho_t)|^2. \tag{6.152}
$$

We do not need to further estimate the fourth term on the right hand side of (4.63) (the coefficient of dB_t^i), because the expected value of this term vanishes anyway due to the martingale property of Brownian motion.

Finally we obtain

$$
\begin{aligned}
\mathbb{E}|V_t^i-\overline{V}_t^i|^2&\lesssim\mathbb{E}|V_0^i-\overline{V}_0^i|^2+\int_0^t\sup_{i\in[N]}\mathbb{E}|V_s^i-\overline{V}_s^i|^2ds+\int_0^t\mathbb{E}|V^\alpha(\overline{\rho}_s^N)-V^\alpha(\rho_s)|^2ds\\
&\lesssim\mathbb{E}|V_0^i-\overline{V}_0^i|^2+\int_0^t\sup_{i\in[N]}\mathbb{E}|V_s^i-\overline{V}_s^i|^2ds+TN^{-1}.
\end{aligned}
$$

Here we have used Lemma 4.4.1 in the second inequality. With Gronwall's inequality we find

$$
\sup_{i\in[N]}\mathbb{E}|V_t^i-\overline{V}_t^i|^2\lesssim N^{-1} \tag{6.153}
$$

where we have used that $\mathbb{E}|V_0^i - \overline{V}_0^i|^2 = 0$. This finishes the proof. □

Regularity of the mean-field solution

Theorem 6.6.7 (cf. theorem 5.2.2)**.** *Let $\rho_0 \in L^2(\mathbb{S}^{d-1})$. Then there exists a unique weak solution of the mean-field equation of the anisotropic KV-CBO model* (6.124) *with the following regularity*

$$\begin{cases} \rho & \in W^{1,2,2}([0,T]; H^1(\mathbb{S}^{d-1}), H^1(\mathbb{S}^{d-1})'), \\ \rho & \in L^\infty([0,T]; L^2(\mathbb{S}^{d-1})) \end{cases} \tag{6.154}$$

where $W^{1,2,2}$ is the space from the Aubin-Lions theorem 5.2.1.

Proof. The uniqueness has been shown in theorem 6.6.5.

The existence is shown with a Picard iteration argument. Let

$$\rho^1(t,v) := \rho_0(v), \quad t \in [0,T] \tag{6.155}$$

be the first iterate of the Picard iteration. For $n \geq 1$ and

$$\rho^n \in L^\infty([0,T]; L^2(\mathbb{S}^{d-1})) \cap L^2([0,T]; H^1(\mathbb{S}^{d-1})) \tag{6.156}$$

we consider the linear PDE

$$\begin{cases} \partial_t \rho_t^{n+1} & = -\lambda \nabla_{\mathbb{S}^{d-1}} \cdot (P_{\mathbb{S}^{d-1}}(v) V^\alpha(\rho_t^n) \rho_t^{n+1}) + \frac{\sigma^2}{2} \sum_{i=1}^d \partial^2_{\mathbb{S}^{d-1},ii}((v - V^\alpha(\rho_t^n))_i^2 \rho_t^{n+1}) \\ & \quad - \frac{\sigma^2}{2}(d-2) \nabla_{\mathbb{S}^{d-1}} \cdot (D(v - V^\alpha(\rho_t^n))^2 v \rho_t^{n+1}) + \frac{\sigma^2}{2}(d-2)(d-1) |D(v - V^\alpha(\rho_t^n))v|^2 \rho_t^{n+1} \\ & \quad - \frac{\sigma^2}{2}(d-1) \sum_{i=1}^d \partial_{\mathbb{S}^{d-1},i}((v - V^\alpha(\rho_t^n))_i^2 \rho_t^{n+1}) v_i \\ \rho^{n+1}(0,v) & = \rho_0(v) \end{cases} \tag{6.157}$$

which has a unique weak solution due to the boundedness of the coefficient functions.

Multiplying both sides with ρ_t^{n+1} and integrating we find

$$\begin{aligned}
&\frac{1}{2}\frac{d}{dt}\|\rho_t^{n+1}\|^2_{L^2(\mathbb{S}^{d-1})} \qquad (6.158)\\
&= \lambda\int_{\mathbb{S}^{d-1}} P_{\mathbb{S}^{d-1}}(v)V^\alpha(\rho_t^n)\rho_t^{n+1}\cdot\nabla_{\mathbb{S}^{d-1}}\rho_t^{n+1}dv - \frac{\sigma^2}{2}\int_{\mathbb{S}^{d-1}}\sum_{i=1}^d \partial_{\mathbb{S}^{d-1},i}\big((v-V^\alpha(\rho_t^n))_i^2\rho_t^{n+1}\big)\big(\partial_{\mathbb{S}^{d-1},i}\rho_t^{n+1}\big)dv\\
&\quad+\frac{\sigma^2}{2}(d-2)\bigg(\int_{\mathbb{S}^{d-1}}\nabla_{\mathbb{S}^{d-1}}\rho_t^{n+1}\cdot\big(D(v-V^\alpha(\rho_t^n))^2v\rho_t^{n+1}\big) - (d-1)\int_{\mathbb{S}^{d-1}}D(v-V^\alpha(\rho_t^n))^2v\rho_t^{n+1}\cdot v\rho_t^{n+1}dv\bigg)\\
&\quad+\frac{\sigma^2}{2}(d-2)(d-1)\int_{\mathbb{S}^{d-1}}|D(v-V^\alpha(\rho_t^n))v|^2|\rho_t^{n+1}|^2dv\\
&\quad+\frac{\sigma^2}{2}(d-1)\int_{\mathbb{S}^{d-1}}\sum_{i=1}^d\big((v-V^\alpha(\rho_t^n))_i^2v_i\rho_t^{n+1}\big)(\partial_{\mathbb{S}^{d-1},i}\rho_t^{n+1})dv
\end{aligned}$$

where we have used the integration by parts formula on the sphere (6.123). With $v\cdot\nabla_{\mathbb{S}^{d-1}}\rho_t^{n+1}=0$ and

$$\partial_{\mathbb{S}^{d-1},i}\big((v-V^\alpha(\rho_t^n))_i^2\rho_t^{n+1}\big) = 2(v-V^\alpha(\rho_t^n))_i\frac{d-1}{d}\rho_t^{n+1} + (v-V^\alpha(\rho_t^n))_i^2\partial_{\mathbb{S}^{d-1}}\rho_t^{n+1} \qquad (6.159)$$

this can be simplified to

$$\begin{aligned}
&\frac{1}{2}\frac{d}{dt}\|\rho_t^{n+1}\|^2_{L^2(\mathbb{S}^{d-1})} \qquad (6.160)\\
&= \lambda\int_{\mathbb{S}^{d-1}} P_{\mathbb{S}^{d-1}}(v)V^\alpha(\rho_t^n)\rho_t^{n+1}\cdot\nabla_{\mathbb{S}^{d-1}}\rho_t^{n+1}dv + \sigma^2\frac{d-1}{d}\int_{\mathbb{S}^{d-1}}\nabla_{\mathbb{S}^{d-1}}\rho_t^{n+1}\cdot V^\alpha(\rho_t^n)\rho_t^{n+1}dv\\
&\quad-\frac{\sigma^2}{2}\int_{\mathbb{S}^{d-1}}\sum_{i=1}^d(v-V^\alpha(\rho_t^n))_i^2(\partial_{\mathbb{S}^{d-1},i}\rho_t^{n+1})^2dv + \frac{\sigma^2}{2}(2d-3)\int_{\mathbb{S}^{d-1}}\nabla_{\mathbb{S}^{d-1}}\rho_t^{n+1}\cdot\big(D(v-V^\alpha(\rho_t^n))^2v\rho_t^{n+1}\big)dv.
\end{aligned}$$

The first term on the right hand side of (6.160) can be bounded by

$$\begin{aligned}
\int_{\mathbb{S}^{d-1}}\lambda P_{\mathbb{S}^{d-1}}(v)V^\alpha(\rho_t^n)\rho_t^{n+1}\cdot\nabla_{\mathbb{S}^{d-1}}\rho_t^{n+1}dv &\le \lambda\int_{\mathbb{S}^{d-1}}|\rho_t^{n+1}||\nabla_{\mathbb{S}^{d-1}}\rho_t^{n+1}|dv \qquad (6.161)\\
&\le \frac{\lambda\epsilon^2}{2}\|\rho_t^{n+1}\|^2_{L^2(\mathbb{S}^{d-1})} + \frac{\lambda}{2\epsilon^2}\|\nabla_{\mathbb{S}^{d-1}}\rho_t^{n+1}\|^2_{L^2(\mathbb{S}^{d-1})} \qquad (6.162)
\end{aligned}$$

where we have used Young's inequality in the second step.

The second term on the right hand side can be bounded as

$$\sigma^2\frac{d-1}{d}\int_{\mathbb{S}^{d-1}}\nabla_{\mathbb{S}^{d-1}}\rho_t^{n+1}\cdot V^\alpha(\rho_t^n)\rho_t^{n+1}dv \lesssim \epsilon^2\|\rho_t^{n+1}\|^2_{L^2(\mathbb{S}^{d-1})} + \frac{1}{\epsilon^2}\|\nabla_{\mathbb{S}^{d-1}}\rho_t^{n+1}\|^2_{L^2(\mathbb{S}^{d-1})} \qquad (6.163)$$

where the constant in $\lesssim$ depends on σ an d.

The third term on the right hand side is obviously ≤ 0, and for the fourth term we similarly find

$$\frac{\sigma^2}{2}(2d-3)\int_{\mathbb{S}^{d-1}} \nabla_{\mathbb{S}^{d-1}}\rho_t^{n+1}\cdot\big(D(v-V^\alpha(\rho_t^n))^2 v\rho_t^{n+1}\big)dv \tag{6.164}$$

$$\lesssim \epsilon^2\|\rho_t^{n+1}\|^2_{L^2(\mathbb{S}^{d-1})} + \frac{1}{\epsilon^2}\|\nabla_{\mathbb{S}^{d-1}}\rho_t^{n+1}\|^2_{L^2(\mathbb{S}^{d-1})}. \tag{6.165}$$

Combining the above estimates yields

$$\frac{1}{2}\frac{d}{dt}\|\rho_t^{n+1}\|^2_{L^2(\mathbb{S}^{d-1})} - \frac{1}{\epsilon^2}\|\nabla_{\mathbb{S}^{d-1}}\rho_t^{n+1}\|^2_{L^2(\mathbb{S}^{d-1})} \lesssim \epsilon^2\|\rho_t^{n+1}\|^2_{L^2(\mathbb{S}^{d-1})} \tag{6.166}$$

from which we deduce with Gronwall's inequality that

$$\|\rho_t^{n+1}\|^2_{L^2(\mathbb{S}^{d-1})} + \int_0^t \|\nabla_{\mathbb{S}^{d-1}}\rho_t^{n+1}\|^2_{L^2(\mathbb{S}^{d-1})}dt \lesssim 1 \tag{6.167}$$

and further

$$\rho^{n+1} \in L^\infty([0,T];L^2(\mathbb{S}^{d-1})) \cap L^2([0,T];H^1(\mathbb{S}^{d-1})). \tag{6.168}$$

In order to apply the Aubin-Lions theorem it remains to show that

$$\partial_t\rho^{n+1} \in L^2([0,T];H^1(\mathbb{S}^{d-1})'). \tag{6.169}$$

To see this we note

$$\|\partial_t\rho_t^{n+1}\|_{H^1(\mathbb{S}^{d-1})'}$$

$$= \sup_{\|\psi\|_{H^1(\mathbb{S}^{d-1})}\leq 1} |\langle\partial_t\rho_t^{n+1},\psi\rangle| \tag{6.170}$$

$$= \sup_{\|\psi\|_{H^1(\mathbb{S}^{d-1})}\leq 1} \Big|\Big\langle \lambda P_{\mathbb{S}^{d-1}}(v)V^\alpha(\rho_t^n)\rho_t^{n+1} + \sigma^2\frac{d-1}{d}V^\alpha(\rho_t^n)\rho_t^{n+1} - \frac{\sigma^2}{2}\big(\nabla^T_{\mathbb{S}^{d-1}}\cdot D(v-V^\alpha(\rho_t^n))_i^2\rho_t^{n+1}\big)$$
$$+ \frac{\sigma^2}{2}(2d-3)\big(D(v-V^\alpha(\rho_t^n))^2 v\rho_t^{n+1}\big), \nabla_{\mathbb{S}^{d-1}}\psi\Big\rangle\Big| \tag{6.171}$$

$$\lesssim \|\rho_t^{n+1}\|_{H^1(\mathbb{S}^{d-1})}. \tag{6.172}$$

So far we have construced a sequence

$$\{\rho^n\}_{n\in\mathbb{N}} \in W^{1,2,2}([0,T];H^1(\mathbb{S}^{d-1}),H^1(\mathbb{S}^{d-1})') \tag{6.173}$$

for which the Aubin-Lions theorem guarantees the existence of a subsequence $\{\rho^{n_k}\}_{k\in\mathbb{N}}$ that converges to a limit function $\rho \in L^2([0,T];L^2(\mathbb{S}^{d-1}))$ as $k\to\infty$.

We now prove that this limit function ρ solves the mean-field equation of the anisotropic KV-CBO

model in the weak sense. With $\rho^n \in C([0,T];L^2(\mathbb{S}^{d-1}))$ (as in the isotropic proof) we find

$$\frac{d}{dt}\int_{\mathbb{S}^{d-1}} |\Phi(v)\big(\rho_t^{n_k}(v)-\rho_t(v)\big)|^2 dv \to 0 \tag{6.174}$$

hence the left hand sides of (6.157) converges to the left hand side of the mean-field equation (6.124).

For the first term on the right hand side of (6.171) we find

$$\lambda\int_{\mathbb{S}^{d-1}} \Big|\big(P_{\mathbb{S}^{d-1}}(v)V^\alpha(\rho_t^{n_k})\rho_t^{n_k+1} - P_{\mathbb{S}^{d-1}}(v)V^\alpha(\rho_t)\rho_t\big)\cdot\nabla_{\mathbb{S}^{d-1}}\Phi(v)\Big|^2 dv \to 0 \tag{6.175}$$

by adding and subtracting the term $V^\alpha(\rho_t^{n_k})\rho_t$.

The other terms are treated similarly. This finishes the proof. □

Optimization

Definition 6.6.5 (cf. definition 5.3.1)**.** *Let $\delta \in (0,2)$ and $\rho_0 \in \mathcal{P}_{ac}(\mathbb{S}^{d-1}) \cap L^2(\mathbb{S}^{d-1})$ the initial distribution. We assume*

$$\begin{cases} \sup_{t\in[0,T]} \|\rho_t\|_{L^2(D_\delta)} \le \epsilon_0, \\ \mathbb{E}(\rho_t)\cdot V^\alpha(\rho_t) \ge C_1 > 0 \quad t\in[0,T], \\ \overline{\mathcal{V}}_T := \sup\limits_{t\in[0,T]} V(\rho_t) \le \min\left\{\frac{\|e^{-\alpha\mathcal{E}}\|^2_{L^1(\rho_0)}}{T}, \frac{\|e^{-\alpha\mathcal{E}}\|^4_{L^1(\rho_0)}}{T\lambda^2}\right\}. \end{cases} \tag{6.176}$$

For the coefficients α, σ, λ and δ we assume

$$\begin{cases} \exists \vartheta > 0 \ s.t.\ C_1 + 5/8 - C_2/2 - (2-\delta) \ge \vartheta > 0, \\ \lambda\vartheta - 16C_\sigma > 0 \end{cases} \tag{6.177}$$

where C_1, C_2 are the same constants as in the isotropic proof.

Theorem 6.6.8 (Monotonic decay of the variance, cf. theorem 5.4.1)**.** *We have*

$$V(\rho_t) \le V(\rho_0)e^{-(\lambda\vartheta-16C_\sigma)t} + \frac{4\lambda\epsilon_0 C_3^{1/2}\delta^{(d-2)/4}}{\lambda\vartheta - 16C_\sigma}. \tag{6.178}$$

Proof. Let us compute the derivative of the variance

$$\frac{d}{dt}V(\rho_t) = -\mathbb{E}(\rho_t)\cdot\int_{\mathbb{S}^{d-1}} \lambda P_{\mathbb{S}^{d-1}}(v)V^\alpha(\rho_t) + C^{ani}_{\mathbb{S}^{d-1}}(v, V^\alpha(\rho_t))\, d\rho_t(v). \tag{6.179}$$

The first term on the right hand side can be estimated as in the isotropic model (see (5.85) - (5.91)),

that is,

$$\begin{aligned}\frac{d}{dt}V(\rho_t) \leq -\lambda C_1 V(\rho_t) - \frac{5\lambda}{8}V(\rho_t) \\ + \frac{\lambda}{4}\bigg(16\epsilon_0 C_3^{1/2}\delta^{(d-2)/4} + 4(2-\delta)V(\rho_t) + 2C_2 V(\rho_t)\bigg)\end{aligned} \tag{6.180}$$

For the second term we have

$$\begin{aligned}&-\mathbb{E}(\rho_t)\cdot\int_{\mathbb{S}^{d-1}} C^{ani}_{\mathbb{S}^{d-1}}(v, V^\alpha(\rho_t))\ d\rho_t(v) \\ &\quad= \mathbb{E}(\rho_t)\cdot C_\sigma\int_{\mathbb{S}^{d-1}} |v-V^\alpha(\rho_t)|^2 v + D(v-V^\alpha(\rho_t))^2 v + 2|D(v-V^\alpha(\rho_t))v|^2 v\ d\rho_t(v) \qquad (6.181)\\ &\quad\leq 4C_\sigma\int_{\mathbb{S}^{d-1}} |v-V^\alpha(\rho_t)|^2\ d\rho_t(v) \qquad (6.182)\\ &\quad\leq 16C_\sigma V(\rho_t) \qquad (6.183)\end{aligned}$$

where we have used Cauchy-Schwarz in the first, and lemma 5.4.1 in the second step.

The rest of proof is the same as for the isotropic model. □

Lemma 6.6.1 (Laplace's principle for the mean-field solution, cf. lemma 5.4.4)**.** *Let $\rho_t \in \mathcal{P}_{ac}(\mathbb{S}^{d-1})\cap L^2(\mathbb{S}^{d-1})$ and $\rho \in L^\infty([0,T];L^2(\mathbb{S}^{d-1}))$ be the mean-field solution of the anisotropic KV-CBO model. Then*

$$\lim_{\alpha\to\infty} -\frac{1}{\alpha}\log\|e^{-\alpha\mathcal{E}}\|_{L^1(\rho_t)} = \underline{\mathcal{E}} \tag{6.184}$$

for all $t\in[0,T]$.

Proof. We first show

$$\frac{d}{dt}\|e^{-\alpha\mathcal{E}}\|^2_{L^1(\rho_t)} \geq -b_1(\alpha)V(\rho_t) - b_2(\alpha)V(\rho_t)^{1/2}. \tag{6.185}$$

Here the coefficients $b_1(\alpha)$ and $b_2(\alpha)$ are different from those of the isotropic model but they still approach zero as $\alpha\to\infty$. The rest of the proof is exactly the same as for the isotropic model. An application of Dynkin's formula yields

$$\begin{aligned}\frac{d}{dt}\int_{\mathbb{S}^{d-1}} e^{-\alpha\mathcal{E}(v)}\ d\rho_t(v) = \int_{\mathbb{S}^{d-1}} \nabla_{\mathbb{S}^{d-1}} e^{-\alpha\mathcal{E}(v)}\cdot\big(\lambda P_{\mathbb{S}^{d-1}}(v)V^\alpha(\rho_t) - \frac{\sigma^2}{2}D(v-V^\alpha(\rho_t))^2 v\big) \\ \sum_{i=1}^d \frac{\sigma^2}{2}(v-V^\alpha(\rho_t))_i^2\partial^2_{\mathbb{S}^{d-1},ii}e^{-\alpha\mathcal{E}(v)}\ d\rho_t(v).\end{aligned} \tag{6.186}$$

The first term on the right hand side of (6.186) can be estimated as in (5.100)

$$\lambda\int_{\mathbb{S}^{d-1}} \nabla_{\mathbb{S}^{d-1}} e^{-\alpha\mathcal{E}(v)}\cdot P_{\mathbb{S}^{d-1}}(v)V^\alpha(\rho_t)\ d\rho_t(v) \geq -\alpha\lambda c_1 e^{-\alpha\underline{\mathcal{E}}}V(\rho_t)^{1/2}. \tag{6.187}$$

For the second term on the right hand side of (6.186) we find

$$
\begin{aligned}
&-\frac{\sigma^2}{2}\int_{\mathbb{S}^{d-1}} D(v-V^\alpha(\rho_t))^2 v\cdot\nabla_{\mathbb{S}^{d-1}} e^{-\alpha\mathcal{E}(v)}\,d\rho_t(v) \\
&\quad=\frac{\sigma^2}{2}\int_{\mathbb{S}^{d-1}} D(v-V^\alpha(\rho_t))^2 v\cdot P_{\mathbb{S}^{d-1}}(v)\nabla\mathcal{E}(v)\;\alpha e^{-\alpha\mathcal{E}(v)}\,d\rho_t(v) \qquad (6.188)\\
&\quad\geq -\frac{\sigma^2}{2}\alpha c_1 e^{-\alpha\underline{\mathcal{E}}}\int_{\mathbb{S}^{d-1}}|v-V^\alpha(\rho_t)|^2\,d\rho_t(v) \qquad (6.189)\\
&\quad\geq -2\sigma^2\alpha c_1 e^{-\alpha\underline{\mathcal{E}}}V(\rho_t). \qquad (6.190)
\end{aligned}
$$

For the third term on the right hand side of (6.186) we find

$$
\begin{aligned}
&\int_{\mathbb{S}^{d-1}}\sum_{i=1}^d\frac{\sigma^2}{2}(v-V^\alpha(\rho_t))_i^2\partial^2_{\mathbb{S}^{d-1},ii}e^{-\alpha\mathcal{E}(v)}\,d\rho_t(v) \\
&\qquad\geq -2\sigma^2\big((d+1)^2(\alpha^2c_1^2+\alpha c_2)+4\alpha c_1\big)e^{-\alpha\underline{\mathcal{E}}}V(\rho_t)
\end{aligned} \qquad (6.191)
$$

where we have used

$$
\begin{aligned}
\partial^2_{\mathbb{S}^{d-1},ii}e^{-\alpha\mathcal{E}(v)} &\overset{L.5.1.1}{=} \partial^2_{ii}e^{-\alpha\mathcal{E}(v)}-2\sum_{j=1}^d\partial^2_{ij}e^{-\alpha\mathcal{E}(v)}v_jv_i - v\cdot\nabla e^{-\alpha\mathcal{E}(v)}+2v_i^2v\cdot\nabla e^{-\alpha\mathcal{E}(v)} \qquad (6.192)\\
&\quad - v_i\partial_i e^{-\alpha\mathcal{E}(v)}+v_i^2\nabla^2e^{-\alpha\mathcal{E}(v)}:v\otimes v\\
&\geq -\big((d+1)^2(\alpha^2c_1^2+\alpha c_2)+4\alpha c_1\big)e^{-\alpha\underline{\mathcal{E}}} \qquad (6.193)
\end{aligned}
$$

and the estimates from section 5.1.

Combining the above inequalities yields

$$
\begin{aligned}
\frac{1}{2}\frac{d}{dt}\|e^{-\alpha\mathcal{E}}\|^2_{L^1(\rho_t)} &= \|e^{-\alpha\mathcal{E}}\|_{L^1(\rho_t)}\frac{d}{dt}\|e^{-\alpha\mathcal{E}}\|_{L^1(\rho_t)} \qquad (6.194)\\
&\geq -2\sigma^2\left((d+1)^2(\alpha^2c_1^2+\alpha c_2)+5\alpha c_1\right)e^{-\alpha\underline{\mathcal{E}}}V(\rho_t)-\alpha\lambda c_1e^{-\alpha\underline{\mathcal{E}}}V(\rho_t)^{1/2}\\
&=: -b_1(\alpha)V(\rho_t)-b_2(\alpha)V(\rho_t)^{1/2} \qquad (6.195)
\end{aligned}
$$

which finishes the proof of (6.185).

Now we can estimate (as in the isotropic case)

$$
\|e^{-\alpha\mathcal{E}}\|^2_{L^1(\rho_t)}\geq\|e^{-\alpha\mathcal{E}}\|^2_{L^1(\rho_0)}\big(1-b_1(\alpha)-b_2(\alpha)\big) \qquad (6.196)
$$

which implies

$$
-\frac{1}{\alpha}\log\|e^{-\alpha\mathcal{E}}\|_{L^1(\rho_t)}\leq-\frac{1}{\alpha}\log\|e^{-\alpha\mathcal{E}}\|_{L^1(\rho_0)}-\frac{1}{2\alpha}\log\big(1-b_1(\alpha)-b_2(\alpha)\big). \qquad (6.197)
$$

Hence, with Laplace's principle we see that for any $\epsilon > 0$ we can find $\alpha_0 = \max\{\alpha_1, \alpha_2\}$ such that

$$\begin{aligned} &\Big| -\frac{1}{\alpha}\log\|e^{-\alpha\mathcal{E}}\|_{L^1(\rho_t)} - \underline{\mathcal{E}}\Big| \\ &\qquad \leq \Big| -\frac{1}{\alpha}\log\|e^{-\alpha\mathcal{E}}\|_{L^1(\rho_0)} - \underline{\mathcal{E}}\Big| + \Big|\frac{1}{2\alpha}\log\big(1 - b_1(\alpha) - b_2(\alpha)\big)\Big| \leq \epsilon \end{aligned} \tag{6.198}$$

for all $\alpha > \alpha_0$. This finishes the proof. □

Theorem 6.6.9 (cf. theorem 5.4.2)**.** *Let the initial probability distribution ρ_0 satisfy the conditions from definition 5.3.1, and $\varepsilon > 0$. Then there is an $\alpha_0 \gg 1$ such that the following approximation estimates hold*

$$\begin{cases} |\boldsymbol{E}(\rho_t) - V^\star| \leq c_3 2^{c_4 - 1}\big((2C_\alpha)^{c_4} V(\rho_t)^{c_4/2} + \varepsilon^{c_4}\big), \\ |V^\alpha(\rho_t) - V^\star| \leq 3V(\rho_t)^{1/2} + c_3 2^{c_4-1}\big((2C_\alpha)^{c_4} V(\rho_t)^{c_4/2} + \varepsilon^{c_4}\big) \end{cases} \tag{6.199}$$

for any $\alpha > \alpha_0$ and all $t \in [0, T]$.

Proof. Same as for the isotropic model. □

Benchmark functions

In chapter 6 we discussed the KV-CBO package [6] and illustrated numerical results for the Rastrigin function. The KV-CBO package implements other benchmark test functions which we briefly discuss here, see [Momin Jamil, 2013]. All of these functions are defined on $\mathbb{R}^d$ and have a *standard search domain* $v \in [a, a]^d$ for some (problem-specific) value $a > 0$. For the Rastrigin function the standard search domain is $[-5.12, 5.12]^d$. These standard search domain are not of relevence here (since we always minimize over the sphere), and are thus not reported.

Rotating the position of the global minimizer can easily be achieved by considering v on the right hand side with $D(v - V^\star)$, where $V^\star$ is the position where we want the global minimizer to be, and $D > 0$ is a scaling constant.

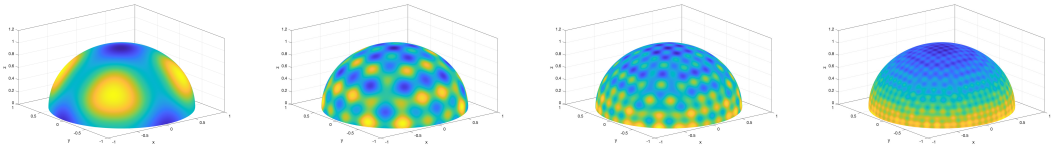


Figure 6.17: Rastrigin function for different values of D.

Rastrigin:

$$\mathcal{E}(v) = Ad + |v|^2 - A\sum_{i=1}^{d} \cos(2\pi v_i) \tag{6.200}$$

where $A = 10$. The global minimizer is $V^\star = 0$, and $\mathcal{E}(V^\star) = 0$. We further chose $D = 5$.

Ackley:

$$\mathcal{E}(v) = A\exp\big(B\frac{|v|}{\sqrt{d}}\big) - \exp\big(\sum_{i=1}^{d}\frac{\cos(2\pi B v_i)}{d}\big) + \exp(1) + C \tag{6.201}$$

where $A = -20$, $B = -0.20$ and $C = 20$. The global minimizer is $V^\star = 0$, and $\mathcal{E}(V^\star) = 0$. We further chose $D = 5$.

Alpine:

$$\mathcal{E}(v) = \sum_{i=1}^{d} |v_i \sin(v_i) - Av_i| \tag{6.202}$$

where $A = 0.1$. The global minimizer is $V^\star = 0$, and $\mathcal{E}(V^\star) = 0$. We further chose $D = 6$.

[6] https://github.com/PhilippeSu/KV-CBO/PhD

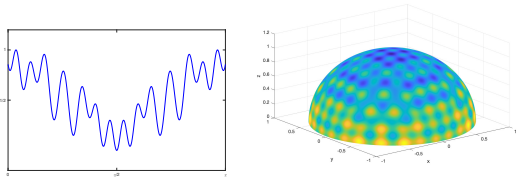


Figure 6.18: Rastrigin

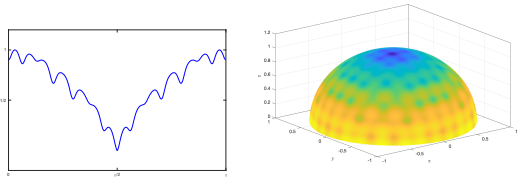


Figure 6.19: Ackley

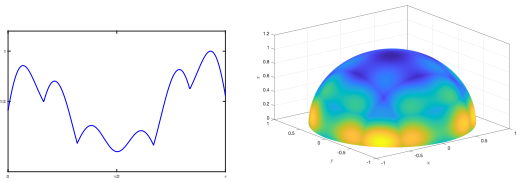


Figure 6.20: Alpine

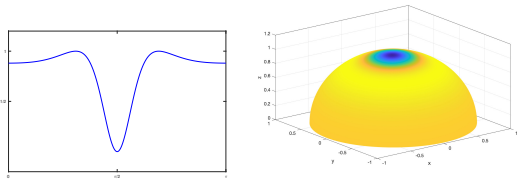


Figure 6.21: Schaffer

Schaffer:

$$\mathcal{E}(v) = A + \frac{\sin^2 |v| - A}{(1 + B|v|^2)^2} \tag{6.203}$$

where $A = 0.5$ and $B = 1$. The global minimizer is $V^\star = 0$, and $\mathcal{E}(V^\star) = 0$. We further chose $D = 2$.

Solomon:

$$\mathcal{E}(v) = A\cos(2\pi|v|) + B|v| + C \tag{6.204}$$

where $A = -1$, $B = 0.1$ and $C = 1$. The global minimizer is $V^\star = 0$, and $\mathcal{E}(V^\star) = 0$. We further chose $D = 5$.

Lévi:

$$\begin{cases} \mathcal{E}(v) & = \sin^2(\pi w_1) + \sum_{i=1}^{d-1}(w_i - 1)^2(1 + B\sin^2(\pi w_i + 1)) + (w_d - 1)^2(1 + \sin^2(2\pi w_d)) \\ w_i & = 1 + v_i/A \end{cases} \tag{6.205}$$

where $A = 1$ and $B = 10$. The global minimizer is $V^\star = 0$, and $\mathcal{E}(V^\star) = 0$. We further chose $D = 4$.

Xin-She Yang (XSY) random:

$$\mathcal{E}(v) = \sum_{i=1}^{d} \xi_i |v_i|^i \tag{6.206}$$

where $\xi \sim \mathcal{U}([0,1]^d)$. The global minimizer is $V^\star = 0$, and $\mathcal{E}(V^\star) = 0$. We further chose $D = 0.5$.

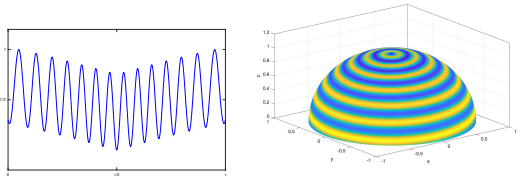


Figure 6.22: Solomon

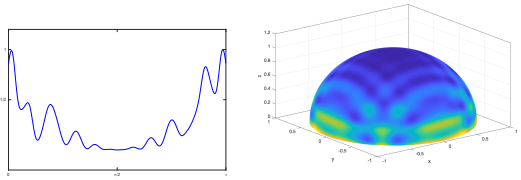


Figure 6.23: Lévi

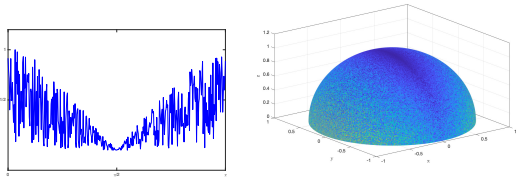


Figure 6.24: XSY random

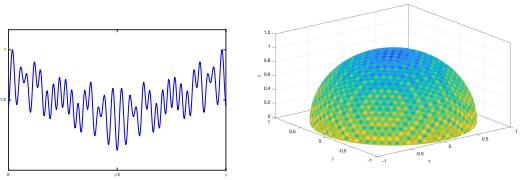


Figure 6.25: Griewank

Bibliography

[Absil and Hosseini, 2017] Absil, P.-A. and Hosseini, S. (2017). A collection of nonsmooth Riemannian optimization problems. *International Series of Numerical Mathematics*.

[Absil et al., 2008] Absil, P.-A., Mahony, R., and Sepulchre, R. (2008). *Optimization algorithms on matrix manifolds*. Princeton University Press.

[Acebron et al., 2005] Acebron, J., Bonilla, L., Pérez-Vicente, C., Farran, F., and Spigler, R. (2005). The kuramoto model: A simple paradigm for synchronization phenomena. *Reviews of Modern Physics*, 77.

[Aggarwal et al., 2001] Aggarwal, C. C., Hinneburg, A., and Keim, D. A. (2001). On the surprising behavior of distance metrics in high dimensional space. In *International conference on database theory*, pages 420–434. Springer.

[Albi et al., 2017] Albi, G., Choi, Y.-P., Fornasier, M., and Kalise, D. (2017). Mean field control hierarchy. *Applied Mathematics & Optimization*, 76(1):93–135.

[Alt, 2016] Alt, H. W. (2016). *Linear Functional Analysis*. Springer, London.

[Ambrosio et al., 2008] Ambrosio, L., Gigli, N., and Savaré, G. (2008). *Gradient Flows: In Metric Spaces and in the Space of Probability Measures*. Springer Science & Business Media.

[Audet and Hare, 2017] Audet, C. and Hare, W. (2017). *Derivative-Free and Blackbox Optimization*. Springer International Publishing.

[Back et al., 1997] Back, T., Fogel, D. B., and Michalewicz, Z., editors (1997). *Handbook of Evolutionary Computation*. IOP Publishing Ltd., Bristol, UK, UK, 1st edition.

[Bahmani and Romberg, 2017] Bahmani, S. and Romberg, J. (2017). Phase retrieval meets statistical learning theory: A flexible convex relaxation. In *Artificial Intelligence and Statistics*, pages 252–260. PMLR.

[Bainbridge et al., 2013] Bainbridge, W. A., Isola, P., and Oliva, A. (2013). *The Intrinsic Memorability of Face Photographs.* Journal of Experimental Psychology: General, 142(4), 1323-1334.

[Balan et al., 2006] Balan, R., Casazza, P., and Edidin, D. (2006). On signal reconstruction without phase. *Applied and Computational Harmonic Analysis*, 20(3):345 – 356.

[Baldi, 2017] Baldi, P. (2017). *Stochastic Calculus: An Introduction Through Theory and Exercises*. Universitext. Springer International Publishing.

[Bandeira et al., 2014] Bandeira, A. S., Cahill, J., Mixon, D. G., and Nelson, A. A. (2014). Saving phase: Injectivity and stability for phase retrieval. *Applied and Computational Harmonic Analysis*, 37(1):106 – 125.

[Basri and Jacobs, 2003] Basri, R. and Jacobs, D. (2003). Lambertian reflectance and linear subspaces. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 25(2):218–233.

[Bendory et al., 2017] Bendory, T., Beinert, R., and Eldar, Y. C. (2017). Fourier phase retrieval: Uniqueness and algorithms. *CoRR*, abs/1705.09590.

[Bianchi et al., 2002] Bianchi, G., Segala, F., and Volcic, A. (2002). The solution of the covariogram problem for planar c^2_+ convex bodies. *Journal of Differential Geometry*, 60:177–198.

[Blum and Roli, 2003] Blum, C. and Roli, A. (2003). Metaheuristics in combinatorial optimization: Overview and conceptual comparison. *ACM computing surveys (CSUR)*, 35(3):268–308.

[Bolley et al., 2012] Bolley, F., Cañizo, J. A., and Carrillo, J. A. (2012). Mean-field limit for the stochastic vicsek model. *Applied Mathematics Letters*, 25(3):339–343.

[Borghi et al., 2021] Borghi, G., Herty, M., and Pareschi, L. (2021). Constrained consensus-based optimization. *arXiv preprint arXiv:2111.10571*.

[Boumal et al., 2014] Boumal, N., Mishra, B., Absil, P.-A., and Sepulchre, R. (2014). Manopt, a Matlab toolbox for optimization on manifolds. *Journal of Machine Learning Research*, 15(42):1455–1459.

[Candès et al., 2014] Candès, E., Li, X., and Soltanolkotabi, M. (2014). Phase retrieval via wirtinger flow: Theory and algorithms. *IEEE Transactions on Information Theory*, 61.

[Candès et al., 2013a] Candès, E., Strohmer, T., and Voroninski, V. (2013a). Phaselift: Exact and stable signal recovery from magnitude measurements via convex programming. *Communications on Pure and Applied Mathematics*, 66.

[Candès et al., 2013b] Candès, E. J., Eldar, Y. C., Strohmer, T., and Voroninski, V. (2013b). Phase retrieval via matrix completion. *SIAM Journal on Imaging Sciences*, 6(1):199–225.

[Carrillo et al., 2020] Carrillo, J., Jin, S., Li, L., and Zhu, Y. (2020). A consensus-based global optimization method for high dimensional machine learning problems. *ESAIM: Control, Optimisation and Calculus of Variations*, 27.

[Carrillo et al., 2018] Carrillo, J. A., Choi, Y.-P., Totzeck, C., and Tse, O. (2018). An analytical framework for consensus-based global optimization method. *Mathematical Models and Methods in Applied Sciences*, 28(06):1037–1066.

[Chandra et al., 2017] Chandra, R., Zhong, Z., Hontz, J., McCulloch, V., Studer, C., and Goldstein, T. (2017). Phasepack: A phase retrieval library. *Asilomar Conference on Signals, Systems, and Computers*.

[Chen et al., 2020] Chen, S., Ma, S., Man-Cho So, A., and Zhang, T. (2020). Proximal gradient method for nonsmooth optimization over the Stiefel manifold. *SIAM Journal on Optimization*, 30(1):210–239.

[Chen and Womersley, 2018] Chen, X. and Womersley, R. (2018). Spherical designs and nonconvex minimization for recovery of sparse signals on the sphere. *SIAM J. Imaging Sci.*, 11:1390–1415.

[Chen and Candes, 2015] Chen, Y. and Candes, E. (2015). Solving random quadratic systems of equations is nearly as easy as solving linear systems. In Cortes, C., Lawrence, N., Lee, D., Sugiyama, M., and Garnett, R., editors, *Advances in Neural Information Processing Systems*, volume 28. Curran Associates, Inc.

[Chen et al., 2014] Chen, Y., Yi, X., and Caramanis, C. (2014). A convex formulation for mixed regression with two components: Minimax optimal rates. In *Conference on Learning Theory*, pages 560–604. PMLR.

[Chen et al., 2018] Chen, Y., Yi, X., and Caramanis, C. (2018). Convex and nonconvex formulations for mixed regression with two components: Minimax optimal rates. *IEEE Transactions on Information Theory*, 64(3):1738–1766.

[Cipriani et al., 2021] Cipriani, C., Huang, H., and Qiu, J. (2021). Zero-inertia limit: from particle swarm optimization to consensus based optimization. arXiv:2104.06939.

[Combettes and Pesquet, 2011] Combettes, P. L. and Pesquet, J.-C. (2011). Proximal splitting methods in signal processing. In *Fixed-point algorithms for inverse problems in science and engineering*, pages 185–212. Springer.

[Conn et al., 2009] Conn, A. R., Scheinberg, K., and Vicente, L. N. (2009). *Introduction to Derivative-Free Optimization*. SIAM, Philadelphia, PA, USA.

[Cucker and Smale, 2007a] Cucker, F. and Smale, S. (2007a). Emergent behavior in flocks. *IEEE Transactions on Automatic Control*, 52(5):852–862.

[Cucker and Smale, 2007b] Cucker, F. and Smale, S. (2007b). On the mathematics of emergence. *Japanese Journal of Mathematics*, 2(1):197–227.

[Dainty and Fienup, 1987] Dainty, J. and Fienup, J. (1987). Phase retrieval and image reconstruction for astronomy. *Image Recovery: Theory Appl*, 13.

[de Gosson, 2021] de Gosson, M. A. (2021). The pauli problem for gaussian quantum states: Geometric interpretation. *Mathematics*, 9(20).

[Degond et al., 2015] Degond, P., Frouvelle, A., and Liu, J.-G. (2015). Phase transitions, hysteresis, and hyperbolicity for self-organized alignment dynamics. *Archive for Rational Mechanics and Analysis*, 216(1):63–115.

[Degond and Motsch, 2008] Degond, P. and Motsch, S. (2008). Continuum limit of self-driven particles with orientation interaction. *Mathematical Models and Methods in Applied Sciences*, 18(supp01):1193–1215.

[Demanet and Jugnon, 2017] Demanet, L. and Jugnon, V. (2017). Convex recovery from interferometric measurements. *IEEE Transactions on Computational Imaging*, 3(2):282–295.

[Dembo and Zeitouni, 2010] Dembo, A. and Zeitouni, O. (2010). *Large Deviations Techniques and Applications*. Springer-Verlag Berlin Heidelberg.

[Demlow and Dziuk, 2007] Demlow, A. and Dziuk, G. (2007). An adaptive finite element method for the Laplace-Beltrami operator on implicitly defined surfaces. *SIAM Journal on Numerical Analysis*, 45(1):421–442.

[Dhifallah et al., 2017] Dhifallah, O., Thrampoulidis, C., and Lu, Y. M. (2017). Phase retrieval via linear programming: Fundamental limits and algorithmic improvements. In *2017 55th Annual Allerton Conference on Communication, Control, and Computing (Allerton)*, pages 1071–1077. IEEE.

[Dorigo and Blum, 2005] Dorigo, M. and Blum, C. (2005). Ant colony optimization theory: A survey. *Theoretical computer science*, 344(2-3):243–278.

[Dorigo and Di Caro, 1999] Dorigo, M. and Di Caro, G. (1999). Ant colony optimization: a new meta-heuristic. In *Proceedings of the 1999 Congress on Evolutionary Computation-CEC99 (Cat. No. 99TH8406)*, volume 2, pages 1470–1477 Vol. 2.

[Durrett, 2018] Durrett, R. (2018). *Stochastic calculus: a practical introduction*. CRC press.

[Eldar and Mendelson, 2014] Eldar, Y. C. and Mendelson, S. (2014). Phase retrieval: Stability and recovery guarantees. *Applied and Computational Harmonic Analysis*, 36(3):473 – 494.

[Elser et al., 2018] Elser, V., Lan, T.-Y., and Bendory, T. (2018). Benchmark problems for phase retrieval. *SIAM Journal on Imaging Sciences*, 11(4):2429–2455.

[Émery and Meyer, 1989] Émery, M. and Meyer, P. (1989). *Stochastic calculus in manifolds*. Universitext (1979). Springer-Verlag.

[Fefferman, 1994] Fefferman, C. (1994). Reconstructing a neural net from its output. *Revista Matemática Iberoamericana*, 10(3):507–555.

[Fiedler et al., 2021] Fiedler, C., Fornasier, M., Klock, T., and Rauchensteiner, M. (2021). Stable recovery of entangled weights: Towards robust identification of deep neural networks from minimal samples. arXiv:2101.07150.

[Fienup, 1982] Fienup, J. (1982). Phase retrieval algorithms: a comparison. *Applied optics*, 21:2758–69.

[Figalli et al., 2018] Figalli, A., Kang, M.-J., and Morales, J. (2018). Global well-posedness of the spatially homogeneous Kolmogorov–Vicsek model as a gradient flow. *Archive for Rational Mechanics and Analysis*, 227(3):869–896.

[Fornasier et al., 2020] Fornasier, M., Huang, H., Pareschi, L., and Sünnen, P. (2020). Consensus-based optimization on hypersurfaces: Well-posedness and mean-field limit. *Mathematical Models and Methods in Applied Sciences*, 30:2725–2751.

[Fornasier et al., 2021a] Fornasier, M., Huang, H., Pareschi, L., and Sünnen, P. (2021a). Consensus-based optimization on the sphere: Convergence to global mininizers and machine learning. *Journal of Machine Learning Research*, 22:1–55.

[Fornasier et al., 2021b] Fornasier, M., Huang, H., Pareschi, L., and Sünnen, P. (2021b). Anisotropic diffusion in consensus-based optimization on the sphere. arXiv:2104.00420.

[Fornasier et al., 2021c] Fornasier, M., Klock, T., and Riedl, K. (2021c). Consensus-based optimization methods converge globally in mean-field law. arxiv:2103.15130.

[Fornasier et al., 2012] Fornasier, M., Schnass, K., and Vybiral, J. (2012). Learning functions of few arbitrary linear parameters in high dimensions. *Foundations of Computational Mathematics*, 12(2):229–262.

[Fornasier et al., 2021d] Fornasier, M., Vybíral, J., and Daubechies, I. (2021d). Robust and resource efficient identification of shallow neural networks by fewest samples. *Information and Inference: A Journal of the IMA*, 10(2):625–695.

[Fournier and Guillin, 2015] Fournier, N. and Guillin, A. (2015). On the rate of convergence in Wasserstein distance of the empirical measure. *Probability Theory and Related Fields*, 162(3-4):707–738.

[Frouvelle and Liu, 2012] Frouvelle, A. and Liu, J.-G. (2012). Dynamics in a kinetic model of oriented particles with phase transition. *SIAM Journal on Mathematical Analysis*, 44(2):791–826.

[Gamba and Kang, 2016] Gamba, I. M. and Kang, M.-J. (2016). Global weak solutions for Kolmogorov–Vicsek type equations with orientational interactions. *Archive for Rational Mechanics and Analysis*, 222(1):317–342.

[Gendreau and Potvin, 2010] Gendreau, M. and Potvin, J.-Y. (2010). *Handbook of Metaheuristics*. Springer Publishing Company, Incorporated, 2nd edition.

[Gerchberg, 1972] Gerchberg, R. W. (1972). A practical algorithm for the determination of phase from image and diffraction plane pictures. *Optik*, 35:237–246.

[Gilbarg and Trudinger, 2015] Gilbarg, D. and Trudinger, N. S. (2015). *Elliptic partial differential equations of second order*. Springer.

[Goldstein and Studer, 2018] Goldstein, T. and Studer, C. (2018). Phasemax: Convex phase retrieval via basis pursuit. *IEEE Transactions on Information Theory*, 64(4):2675–2689.

[Goldstein et al., 2014] Goldstein, T., Studer, C., and Baraniuk, R. (2014). A field guide to forward-backward splitting with a fasta implementation. arXiv:1411.3406.

[Grassi et al., 2021] Grassi, S., Huang, H., Pareschi, L., and Qiu, J. (2021). Mean-field particle swarm optimization. arXiv:2108.00393.

[Grassi and Pareschi, 2021] Grassi, S. and Pareschi, L. (2021). From particle swarm optimization to consensus based optimization: Stochastic modeling and mean-field limit. *Mathematical Models and Methods in Applied Sciences*, 31(08):1625–1657.

[Ha et al., 2021] Ha, S.-Y., Jin, S., and Kim, D. (2021). Convergence and error estimates for time-discrete consensus-based optimization algorithms. *Numer. Math.*, 147:255–282.

[Hairer et al., 2006] Hairer, E., Lubich, C., and Wanner, G. (2006). *Geometric numerical integration: structure-preserving algorithms for ordinary differential equations*, volume 31. Springer Science & Business Media.

[Hore and Ziou, 2010] Hore, A. and Ziou, D. (2010). Image quality metrics: PSNR vs. SSIM. In *2010 20th international conference on pattern recognition*, pages 2366–2369. IEEE.

[Hsu, 2002] Hsu, E. P. (2002). *Stochastic analysis on manifolds*. American Mathematical Soc.

[Huang, 2021] Huang, H. (2021). A note on the mean-field limit for the particle swarm optimization. *Applied Mathematics Letters*, page 107133.

[Huang and Qiu, 2021] Huang, H. and Qiu, J. (2021). On the mean-field limit for the consensus-based optimization. arxiv:2105.12919.

[Itô, 1950] Itô, K. (1950). Stochastic differential equations in a differentiable manifold. *Nagoya Math. J.*, 1:35–47.

[Jaganathan et al., 2015] Jaganathan, K., Eldar, Y., and Hassibi, B. (2015). *Optical Compressive Imaging*. CRC Press.

[Jaganathan et al., 2016] Jaganathan, K., Eldar, Y. C., and Hassibi, B. (2016). STFT phase retrieval: Uniqueness guarantees and recovery algorithms. *IEEE Journal of Selected Topics in Signal Processing*, 10(4):770–781.

[Kennedy, 1997] Kennedy, J. (1997). The particle swarm: social adaptation of knowledge. In *Proceedings of 1997 IEEE International Conference on Evolutionary Computation (ICEC '97)*, pages 303–308.

[Kennedy, 2010] Kennedy, J. (2010). Particle swarm optimization. *Encyclopedia of machine learning*, pages 760–766.

[Kennedy and Eberhart, 1995] Kennedy, J. and Eberhart, R. (1995). Particle swarm optimization. In *Proceedings of ICNN'95 - International Conference on Neural Networks*, volume 4, pages 1942–1948 vol.4.

[Kuramoto, 1975] Kuramoto, Y. (1975). Self-entrainment of a population of coupled non-linear oscillators. In *International symposium on mathematical problems in theoretical physics*, pages 420–422. Springer.

[Larson et al., 2019] Larson, J., Menickelly, M., and Wild, S. M. (2019). Derivative-free optimization methods. *Acta Numerica*, 28:287–404.

[Lerman and Maunu, 2017] Lerman, G. and Maunu, T. (2017). Fast, robust and non-convex subspace recovery. *Information and Inference: A Journal of the IMA*, 7(2):277—-336.

[Lerman et al., 2015] Lerman, G., McCoy, M. B., Tropp, J. A., and Zhang, T. (2015). Robust computation of linear models by convex relaxation. *Foundations of Computational Mathematics*, 15(2):363–410.

[Marsaglia et al., 1972] Marsaglia, G. et al. (1972). Choosing a point from the surface of a sphere. *The Annals of Mathematical Statistics*, 43(2):645–646.

[Miller, 2006] Miller, P. D. (2006). *Applied Asymptotic Analysis*. American Mathematical Soc.

[Mishura and Veretennikov, 2020] Mishura, Y. and Veretennikov, A. (2020). Existence and uniqueness theorems for solutions of mckean–vlasov stochastic equations. *Theory of Probability and Mathematical Statistics*.

[Momin Jamil, 2013] Momin Jamil, X.-S. Y. (2013). A literature survey of benchmark functions for global optimization problems. Int. Journal of Mathematical Modelling and Numerical Optimisation, Vol. 4, No. 2, pp. 150–194.

[Muller, 1959] Muller, M. E. (1959). A note on a method for generating points uniformly on n-dimensional spheres. *Communications of the ACM*, 2(4):19–20.

[Öttinger, 1996] Öttinger, H. C. (1996). *Stochastic Processes in Polymeric Fluids*. Springer, Berlin, Heidelberg.

[Pardalos and Romeijn, 2009] Pardalos, P. and Romeijn, H. (2009). Handbook of optimization in medicine. *Handbook of Optimization in Medicine edited by P.M. Pardalos and H.E. Romeijn. Berlin: Springer, 2009. ISBN: 978-0-387-09769-5 hr*, 26.

[Pardalos and Vavasis, 1991] Pardalos, P. M. and Vavasis, S. A. (1991). Quadratic programming with one negative eigenvalue is NP-hard. *Journal of Global Optimization*, 1:15–22.

[Pfeiffer, 2018] Pfeiffer, F. (2018). X-ray ptychography. *Nature Photonics*, 12(1):9–17.

[Pinnau et al., 2017] Pinnau, R., Totzeck, C., Tse, O., and Martin, S. (2017). A consensus-based model for global optimization and its mean-field limit. *Mathematical Models and Methods in Applied Sciences*, 27(01):183–204.

[Platen, 1999] Platen, E. (1999). An introduction to numerical methods for stochastic differential equations. *Acta numerica*, 8:197–246.

[Poli et al., 2007] Poli, R., Kennedy, J., and Blackwell, T. (2007). Particle swarm optimization. *Swarm intelligence*, 1(1):33–57.

[Rauchensteiner, 2018] Rauchensteiner, M. (2018). Learning two layer neural networks by fewest samples. *Master's thesis, Technical University Munich.*

[Roubíček, 2013] Roubíček, T. (2013). *Nonlinear Partial Differential Equations with Applications.* Birkhäuser, Basel.

[Shechtman et al., 2015] Shechtman, Y., Eldar, Y. C., Cohen, O., Chapman, H. N., Miao, J., and Segev, M. (2015). Phase retrieval with application to optical imaging: A contemporary overview. *IEEE Signal Processing Magazine*, 32(3):87–109.

[Strogatz, 2000] Strogatz, S. H. (2000). From kuramoto to crawford: exploring the onset of synchronization in populations of coupled oscillators. *Physica D: Nonlinear Phenomena*, 143(1):1–20.

[Sznitman, 1991] Sznitman, A.-S. (1991). Topics in propagation of chaos. In *Ecole d'été de probabilités de Saint-Flour XIX—1989*, pages 165–251. Springer.

[Totzeck and Wolfram, 2020] Totzeck, C. and Wolfram, M.-T. (2020). Consensus-based global optimization with personal best. *Mathematical Biosciences and Engineering*, 17(5):6026–6044.

[Vicsek et al., 1995] Vicsek, T., Czirók, A., Ben-Jacob, E., Cohen, I., and Shochet, O. (1995). Novel type of phase transition in a system of self-driven particles. *Physical review letters*, 75(6):1226.

[Villani, 2009] Villani, C. (2009). *Optimal Transport: Old and New.* Springer.

[Walther, 1963] Walther, A. (1963). The question of phase retrieval in optics. *Optica Acta: International Journal of Optics*, 10(1):41–49.

[Wang et al., 2018] Wang, G., Giannakis, G. B., and Eldar, Y. C. (2018). Solving systems of random quadratic equations via truncated amplitude flow. *IEEE Transactions on Information Theory*, 64(2):773–794.

[William A. Adkins, 2012] William A. Adkins, M. G. D. (2012). *Ordinary Differential Equations.* Springer, New York, NY.

[Zimek et al., 2012] Zimek, A., Schubert, E., and Kriegel, H.-P. (2012). A survey on unsupervised outlier detection in high-dimensional numerical data. *Statistical Analysis and Data Mining: The ASA Data Science Journal*, 5(5):363–387.

[Øksendal, 2003] Øksendal, B. (2003). *Stochastic Differential Equations: An Introduction with Applications*. Springer, Berlin, Heidelberg.