Exponential Stability of Consensus-Based Optimization

S. Göttlich*, J. Heieck*, A. Neuenkirch*

August 20, 2025

Abstract

We study the finite-agent behavior of Consensus-Based Optimization (CBO), a recent metaheuristic for the global minimization of a function, that combines drift toward a consensus estimate with stochastic exploration. While previous analyses focus on asymptotic mean-field limits, we investigate the stability properties of CBO for finite population size N. Following a hierarchical approach, we first analyze a deterministic formulation of the algorithm and then extend our results to the fully stochastic setting governed by a system of stochastic differential equations. Our analysis reveals that essential stability properties, including almost sure and mean square exponential convergence, persist in both regimes and provides sharp quantitative estimates on the rates of convergence.

AMS Classification. 34D35, 37H30, 90C26, 93D05

Keywords: stability analysis, Lyapunov function, multi-agent system, consensus-based optimization, derivative-free optimization

1 Introduction

1.1 Outline

A central problem in applied mathematics is the global minimization of a potentially non-convex and non-smooth function $f : \mathbb{R}^D \to \mathbb{R}$, along with the search for its global minimizer $\mathbf{X}^\star \in \mathbb{R}^D$. Computing f or $\mathbf{X}^\star$ is generally NP-hard, but practical instances often permit approximate solutions within reasonable accuracy and time.

Metaheuristics [1] have proven effective in tackling such optimization problems, orchestrating local improvements and global strategies through a mix of random and deterministic decisions. Prominent examples include Random Search [2], Particle Swarm Optimization (PSO) [3], Wind-Driven Optimization [4] and Ant Colony Optimization [5]. Despite their empirical success, many metaheuristics lack rigorous mathematical guarantees for global convergence. Nevertheless, for some of them, such as Random Search, there exist probabilistic guarantees for global convergence [6].

Recently, Consensus-Based Optimization (CBO) [7, 8] has emerged as a promising alternative, following the principles of metaheuristics while offering greater mathematical tractability. CBO models a finite set $X_t^1, \ldots, X_t^N \in \mathbb{R}^D$ of $N \in \mathbb{N}$ agents as stochastic processes in time $t \geq 0$ exploring the domain and forming a consensus. Their dynamics are driven by two competing forces: a drift term, which pulls agents toward a weighted consensus estimate $\nu_f^\alpha(X_t)$ of $\mathbf{X}^\star$ given by

$$\nu_f^\alpha(X_t) = \sum_{m=1}^N X_t^m \frac{\exp(-\alpha f(X_t^m))}{\sum_{k=1}^N \exp(-\alpha f(X_t^k))} \tag{1.1}$$

with $\alpha > 0$, contracting their convex hull; and a diffusion term, modeled as scaled Brownian motion in $\mathbb{R}^D$, enhancing exploration. Mathematically, the evolution of each agent X_t^n is governed by the stochastic differential equation (SDE)

$$dX_t^n = -\lambda \left(X_t^n - \nu_f^\alpha(X_t) \right) dt + \sigma \left\| X_t^n - \nu_f^\alpha(X_t) \right\| dW_t^n, \quad n, \ldots, N, \tag{1.2}$$

where $\lambda > 0$ controls the strength of the drift towards the weighted consensus point $\nu_f^\alpha(X_t)$, and $\sigma > 0$ scales the noise term modeled by D-dimensional independent Brownian motions W_t^n. This formulation

---

[1]University of Mannheim, Department of Mathematics, 68131 Mannheim, Germany (goettlich@math.uni-mannheim.de, jacob.heieck@students.uni-mannheim.de, neuenkirch@uni-mannheim.de).

ensures a balance between exploitation of promising regions and exploration of the search space, ultimately leading agents to a near-optimal global consensus.

The theoretical analysis of CBO has seen significant progress, particularly in the context of its mean-field limit. In [7], the authors establish that as the number of agents N tends to infinity and α approaches infinity (Laplacian principle [9]), the system converges arbitrarily close to the minimum. Furthermore, in [10], rigorous results are provided to demonstrate that under suitable assumptions, the mean-field formulation of CBO converges globally to the minimizer $\mathbf{X}^\star$. These results offer deep insights into the large-population behavior of CBO but rely on asymptotic limits. Apart from the well-posedness of the finite-agent system, which has been established in [8], little is known about the theoretical properties of CBO for finite N and finite α. Some progress in this direction has been made for simplified variants of CBO where all agents are influenced by the same random fluctuations, effectively removing the individual stochasticity. While these works provide valuable insights into the convergence properties and error bounds [11, 12] of these simplified algorithms, they differ structurally from the classical CBO dynamics we study here, which model independent agent behavior more faithfully.

In this work, we take a next step towards further bridging this gap by analyzing the stability properties of the classical CBO for finite N and finite α. Our approach follows a hierarchical methodology, motivated by the existing work on deterministic particle-based systems. In the linear setting, convergence to consensus has been successfully shown for multi-agent systems [13], and strict Lyapunov functions have been established for consensus dynamics under directed connected graphs [14, 15]. These studies leverage a graph-theoretic interpretation of particle systems, wherein agents are represented as nodes and their interactions as weighted edges.

Building upon these insights, we adopt the following structured approach:

- First, in section 2 we study the deterministic system

$$dX_t^n = -\lambda \left(X_t^n - \nu_f^\alpha(X_t) \right) dt, \quad n = 1, \ldots, N, \tag{1.3}$$

  leveraging existing results from linear graph-theoretic consensus models to establish global exponential stability of the component-wise non-linear system (1.3) and its explicit Euler discretization for $N < \infty$ and $\alpha < \infty$.

- Finally, in section 3 we investigate the full system governed by a SDE

$$dZ_t^n = -\lambda \left(Z_t^n - \nu_f^\alpha(Z_t) \right) dt + \sigma \left(Z_t^n - \nu_f^\alpha(Z_t) \right) dW_t^n, \quad n = 1, \ldots, N, \tag{1.4}$$

  which incorporates both drift and diffusion effects. We rigorously establish both almost sure and mean square exponential stability of the system (1.4) and its Euler-Maruyama discretization under suitable conditions on the parameters. In our analysis, we focus on the anisotropic diffusion term $\sigma(Z_t^n - \nu_f^\alpha(Z_t))$ [16], rather than the more general formulation $\sigma\|Z_t^n - \nu_f^\alpha(Z_t)\|$, where $\|\cdot\|$ typically denotes the Euclidean norm $\|\cdot\|_2$ [8, 10]. We restrict ourselves to the diffusion term $\sigma(Z_t^n - \nu_f^\alpha(Z_t))$ since the graphical interpretation of the system is sufficient to reflect its stability properties.

By systematically progressing from the deterministic case to the full stochastic system, we aim to gain a deeper understanding of CBO dynamics beyond the mean-field limit, providing new theoretical insights into its behavior for practical, finite-agent settings.

1.2 Notation

To establish a consistent notation for our analysis, we introduce the following definitions. We denote D as the dimension, and the collective state of all particles at time t by the vector $X_t = (X_t^1, \ldots, X_t^N)^\top \in \mathbb{R}^{ND}$, where each particle X_t^n is represented as $X_t^n = (x_t^{n,1}, \ldots, x_t^{n,D})^\top \in \mathbb{R}^D$, $n = 1, \ldots, N$.

We use I to denote identity matrices of appropriate dimensions. Furthermore, we define $\mathbf{0}_N := (0, \ldots, 0)^\top \in \mathbb{R}^N$ and $\mathbf{1}_N := (1, \ldots, 1)^\top \in \mathbb{R}^N$ as the all-zero and all-one column vectors of size N, respectively.

Unless stated otherwise, we denote by $\|\cdot\|$ the standard Euclidean norm.

This notation will be used throughout the remainder of this work to ensure consistency and clarity in our mathematical formulation.

2 Purely Deterministic System

We begin by examining the purely deterministic system (1.3). In this setting, the exploitative diffusion is neglected, focusing solely on drift forces that drive particles towards a consensus state. Before we will establish our stability results, we precisely define what we mean by stability in this context.

Intuitively, stability describes the persistence of solutions under small perturbations: if a system is slightly disturbed, it shall return to its original state. More formally, stability is often characterized through equilibrium points.

Definition 2.1 (Equilibrium Point [17, Definiton 3.2]). The state $X^\star \in \mathbb{R}^N$ is said to be an *equilibrium point* of an ordinary differential equation (ODE)

$$dX_t = a(X_t)dt, \quad t \geq 0, \tag{2.1}$$

if once X_t is equal to $X^\star$, it remains equal to $X^\star$ for all future times, i.e., $X^\star$ has to satisfy $a(X^\star) = \mathbf{0}_N$.

In particular, we distinguish between stability and exponential stability.

Definition 2.2 (Stability for ODE [17, Definiton 3.3 and 3.6]). The equilibrium point $X^\star = \mathbf{0}_N$ is said to be *stable* if, for any $R > 0$, there exists $r > 0$, such that if the initial state satisfies $\|X_0\| < r$, then $\|X_t\| < R$ for all $t \geq 0$. Otherwise, the equilibrium point is said to be *unstable*. If stability holds for any initial state X_0, the equilibrium point is said to be *globally stable*.

Exponential stability strengthens the notion of stability.

Definition 2.3 (Exponential Stability for ODE [17, Definition 3.5 and 3.6]). The equilibrium point $X^\star = \mathbf{0}_N$ is *exponentially stable* if there exist constants $c > 0$ and $\delta > 0$ and some $r > 0$ such that $\|X_0\| < r$ implies

$$\|X_t\| \leq c\|X_0\|e^{-\delta t}, \quad t \geq 0.$$

If the inequality holds for all $X_0 \in \mathbb{R}^N$, the equilibrium point is called *globally exponentially stable*.

Exponential stability not only ensures that solutions converge to the equilibrium, but also quantifies how fast this convergence occurs. More precisely, the distance to the equilibrium decays at least as fast as $e^{-\delta t}$ for some rate $\delta > 0$.

Notice, that both definitions apply specifically to the equilibrium point $X^\star = \mathbf{0}_N$. The stability of other equilibrium points or entire sets of points can be analyzed through a suitable transformation of the state space.

To assess the stability properties of equilibrium points, one may employ a fundamental mathematical tool: Lyapunov functions. These functions provide a systematic approach to determining whether solutions remain close to an equilibrium or converge to it over time.

Definition 2.4 (Lyapunov functions for ODE [17, Definiton 3.8]). Let $V : \mathbb{R}^N \to \mathbb{R}$ be a function,

(1) which is positive definite, i.e., we have $V(x) = 0$ for $x = \mathbf{0}_N$ and $V(x) > 0$ else,

(2) which has continuous partial derivatives and

(3) whose time derivative along any state trajectory of the system $dX_t = a(X_t)dt$ is negative semi-definite, i.e.,

$$\forall t \geq 0 : \quad \dot{V}(X_t) := \frac{dV(X_t)}{dt} = (\nabla V(X_t))^\top \frac{dX_t}{dt} = (\nabla V(X_t))^\top a(X_t) \leq 0$$

Then V is said to be a *Lyapunov function* for the system (2.1).

The reformulation in Definition 2.4 follows directly from the chain rule. Here, we exploit the fact that V depends solely on the state X_t and not explicitly on time t.

The core idea is to construct a scalar function V that decreases along system trajectories. If such a function exists, it serves as a certificate for stability.

Theorem 2.5 (Stability for ODEs [17, Theorem 3.2]). Consider the system (2.1) and assume that $X^\star = \mathbf{0}_N$ is an equilibrium point. If there exists an Lyapunov function V for this system, then $X^\star = \mathbf{0}_N$ is globally stable.

If V has additional properties, we can obtain exponential stability.

Theorem 2.6 (Exponential Stability for ODE [?, Theorem 4.5])**.** Consider the system (2.1) and assume that $X^\star = \mathbf{0}_N$ is an equilibrium point. Assume that $V : \mathbb{R}^N \to \mathbb{R}$ is Lyapunov function (for this system), which has the following additional property: there exist constants $c_1, c_2, c_3, c_4 > 0$ such that

   (i) $c_1\|x\|^2 \le V(x) \le c_2\|x\|^2$ for $x \in \mathbb{R}^N$,

   (ii) $\dot{V}(X_t) \le -c_3\|X_t\|^2$ for all $t \ge 0$,

   (iii) $\|\nabla V(x)\| \le c_4\|x\|$ for $x \in \mathbb{R}^N$.

Then the equilibrium point $X^\star = \mathbf{0}_N$ is globally exponentially stable with $c = \sqrt{\frac{c_2}{c_1}}$ and $\delta = \frac{c_3}{2c_2}$.

Lyapunov functions are a classical concept for deriving (exponential) stability of ODEs, and we apply it to our dynamical system (1.3). During the course of our analysis it will turn out that an appropriate transformation of the state space strikingly clarifies the stability properties of the dynamics of (1.3).

2.1 System Description

The purely deterministic system is given by

$$dX_t^n = -\lambda \left(X_t^n - \nu_f^\alpha(X_t) \right) dt \tag{2.2}$$

for $n = 1, \ldots, N$. By definition of the consensus point $\nu_f^\alpha(X_t)$ in (1.1) we have

$$dX_t^n = -\lambda \left(X_t^n - \sum_{m=1}^N a_m(X_t) X_t^m \right) dt$$

for $n = 1, \ldots, N$ with weights

$$a_m(X_t) = \frac{\exp(-\alpha f(X_t^m))}{\sum_{k=1}^N \exp(-\alpha f(X_t^k))}.$$

Note that $a_m = a_m(X_t)$ depends non-linearly on the particle vector X_t. In the following, we will often drop this dependence for notational easiness.

Since

$$\sum_{m=1}^N a_m = \sum_{m=1}^N \frac{\exp(-\alpha f(X_t^m))}{\sum_{k=1}^N \exp(-\alpha f(X_t^k))} = 1,$$

all weights sum up to 1 and we can rewrite

$$dX_t^n = -\lambda \left(\sum_{\substack{m=1 \\ m \ne n}}^N a_m \cdot (X_t^n - X_t^m) \right) dt.$$

Considering all N equations simultaneously and sorting them by spatial dimension d yields

$$dX_t = d \begin{pmatrix} x_t^{1,1} \\ \vdots \\ x_t^{N,1} \\ \vdots \\ x_t^{1,D} \\ \vdots \\ x_t^{N,D} \end{pmatrix} = -\lambda \begin{pmatrix} \sum_{m=2}^N a_m \left(x_t^{1,1} - x_t^{m,1} \right) \\ \vdots \\ \sum_{m=1}^{N-1} a_m \left(x_t^{N,1} - x_t^{m,1} \right) \\ \vdots \\ \sum_{m=2}^N a_m \left(x_t^{1,D} - x_t^{m,D} \right) \\ \vdots \\ \sum_{m=1}^{N-1} a_m \left(x_t^{N,D} - x_t^{m,D} \right) \end{pmatrix} dt. \tag{2.3}$$

4

Given the weights $a_1, \ldots, a_N$, the evolution of each dimension is therefore independent of the others, as it depends only on the interactions between particles within the same dimension. Moreover, the weights a_m are equal in all equations.

Thus, it will suffice to study the system in each dimension $d \in \{1, \ldots D\}$ on its own. If we consider the system

$$d \begin{pmatrix} x_t^{1,d} \\ \vdots \\ x_t^{N,d} \end{pmatrix} = -\lambda \begin{pmatrix} \sum_{m=2}^{N} a_m \left(x_t^{1,d} - x_t^{m,d} \right) \\ \vdots \\ \sum_{m=1}^{N-1} a_m \left(x_t^{N,d} - x_t^{m,d} \right) \end{pmatrix} dt \tag{2.4}$$

and introduce the notation $X_t^{:,d} = (x_t^{1,d}, \ldots, x_t^{N,d})^\top \in \mathbb{R}^N$, the right-hand side of the corresponding system simplifies to

$$-\lambda \hat{L}(X_t) X_t^{:,d} dt,$$

where $\hat{L}(X_t)$, known as the Laplacian matrix [14], is defined as follows

$$\hat{L}(X_t) := \begin{pmatrix} 1 - a_1(X_t) & -a_2(X_t) & \cdots & -a_N(X_t) \\ -a_1(X_t) & 1 - a_2(X_t) & \cdots & -a_N(X_t) \\ \vdots & \vdots & \ddots & \vdots \\ -a_1(X_t) & -a_2(X_t) & \cdots & 1 - a_N(X_t) \end{pmatrix}. \tag{2.5}$$

The diagonal entries arise since the weights sum up to 1. This leads to the compact representation

$$dX_t = -\lambda L(X_t) X_t dt$$

of system (2.3) with

$$L(X_t) = \begin{pmatrix} \hat{L}(X_t) & \mathbf{0}_{N \times N} & \cdots & \mathbf{0}_{N \times N} \\ \mathbf{0}_{N \times N} & \hat{L}(X_t) & \cdots & \mathbf{0}_{N \times N} \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{0}_{N \times N} & \mathbf{0}_{N \times N} & \cdots & \hat{L}(X_t) \end{pmatrix} \in \mathbb{R}^{DN \times DN}. \tag{2.6}$$

Here, $\mathbf{0}_{N \times N}$ denotes the zero-matrix of dimension $N \times N$. This formulation emphasizes that each dimension d can be treated separately, while still preserving the component-wise non-linear interactions between particles through the weights $a_1, \ldots, a_N$. In the following we will drop the dependence of $\hat{L} = \hat{L}(X_t)$ and $L = L(X_t)$ on X_t for notational simplicity.

2.2 Global Exponential Stability

In this chapter, we establish the (global exponential) stability of the deterministic system (2.2). As a first step, we determine the equilibrium points of said system. These correspond to solutions of the equation

$$-\lambda \left(X_t^n - \nu_f^\alpha(X_t) \right) = \mathbf{0}_N \quad \Leftrightarrow \quad X_t^n = \nu_f^\alpha(X_t)$$

for $n = 1, \ldots, N$. Since the right-hand side depends on the entire particle configuration rather than on individual indices n, a necessary condition for an equilibrium is

$$X_t^1 = \ldots = X_t^N. \tag{2.7}$$

To verify whether this condition is also sufficient, we assume that all particles take the same value, i.e., equation (??) is satisfied. Given that $\nu_f^\alpha(X_t)$ represents a weighted consensus point, we compute for any $n \in \{1, \ldots, N\}$:

$$X_t^n = \sum_{m=1}^{N} \frac{X_t^n}{N} = \sum_{m=1}^{N} \frac{X_t^n \exp(-\alpha f(X_t^n))}{\sum_{k=1}^{N} \exp(-\alpha f(X_t^n))} = \sum_{m=1}^{N} \frac{X_t^m \exp(-\alpha f(X_t^m))}{\sum_{k=1}^{N} \exp(-\alpha f(X_t^k))} = \nu_f^\alpha(X_t). \tag{2.8}$$

This confirms that the equilibrium points are uniquely characterized by the condition $X_t^1 = \ldots = X_t^N$. Thus, all particles must reach a common consensus value at equilibrium, which belongs to the set

$$\mathcal{G}^\star = \left\{ x \in \mathbb{R}^{N \times D} : x^{1,d} = \ldots = x^{N,d}, d = 1, \ldots, D \right\}. \tag{2.9}$$

Having identified the equilibrium points, we now turn our attention to their stability properties. To this end, we focus on the matrix L, which governs the dynamics of the system $dX_t = -\lambda L X_t dt$. As previously indicated each $d \in \{1, \ldots, D\}$ will be treated independently. We will therefore analyze the stability of the subsystem

$$dX_t^{:,d} = -\lambda \hat{L} X_t^{:,d} dt \tag{2.10}$$

where the matrix $\hat{L}$ is given by equation (2.5).

Our first objective is to understand the spectral properties of $\hat{L}$, as these determine the qualitative behavior of the system. More precisely, we will study

$$\dot{V}(X_t^{:,d}) = (\nabla V(X_t^{:,d}))^\top a(X_t^{:,d}) = (\nabla V(X_t^{:,d}))^\top (-\lambda \hat{L} X_t^{:,d}) = -\lambda (\nabla V(X_t^{:,d}))^\top \hat{L} X_t^{:,d} \tag{2.11}$$

for a suitable function $V : \mathbb{R}^d \to [0, \infty)$. For the inequality $\dot{V}(X_t^{:,d}) \le 0$ to hold, we need to understand the eigenstructure of $\hat{L}$, which reveals in which directions particles contract and in which directions motion persists or stagnates.

In order to compute the eigenvalues of $\hat{L}$ explicitly, we make use of the following result on rank-one matrix updates, which allows for efficient determinant calculations.

Lemma 2.7 (Rank-One Update [18, p. 475]). If $A \in \mathbb{R}^{N \times N}$ is non-singular, and if c, d are N-dimensional column vectors, then:

(i) $\det(I + cd^\top) = 1 + d^\top c$,

(ii) $\det(A + cd^\top) = \det(A) \left(1 + d^\top A^{-1} c\right)$.

We now apply this lemma to derive the eigenvalues of $\hat{L}$. Recall that $\mathbf{1}_N$ denotes the all-ones column vector, and define the weight vector

$$\ell := (a_1, \ldots, a_N)^\top \in (0, 1)^N, \tag{2.12}$$

which collects the weights associated with each particle X_t^n, $n = 1, \ldots, N$.

Lemma 2.8 (Eigenvalues of $\hat{L}$). The matrix $\hat{L} \in \mathbb{R}^{N \times N}$ has exactly one eigenvalue equal to zero, while the remaining $N - 1$ eigenvalues are equal to 1, that is,

$$\mu_1(\hat{L}) = 0, \quad \mu_2(\hat{L}) = \ldots = \mu_N(\hat{L}) = 1.$$

A right-eigenvector corresponding to $\mu_1(\hat{L})$ is $\mathbf{1}_N$, a left-eigenvector is ℓ.

Proof. We begin by rewriting $\hat{L}$ using a rank-one update involving $\mathbf{1}_N$ and $\ell = (a_1, \ldots, a_N)^\top$:

$$\hat{L} = \begin{pmatrix} 1 - a_1 & -a_2 & \cdots & -a_N \\ -a_1 & 1 - a_2 & \cdots & -a_N \\ \vdots & \vdots & \ddots & \vdots \\ -a_1 & -a_2 & \cdots & 1 - a_N \end{pmatrix} = I - \begin{pmatrix} a_1 & a_2 & \cdots & a_N \\ a_1 & a_2 & \cdots & a_N \\ \vdots & \vdots & \ddots & \vdots \\ a_1 & a_2 & \cdots & a_N \end{pmatrix} = I - \mathbf{1}_N \ell^\top.$$

To determine the eigenvalues, we compute the characteristic equation by evaluating $\det(\hat{L} - \mu I) = 0$. Using the rank-one update representation of $\hat{L}$ and applying (ii) of Lemma 2.7 with $A = (1-\mu)I$, $c = -\mathbf{1}_N$ and $d = \ell$, we obtain

$$\det(\hat{L} - \mu I) = \det((1 - \mu)I) \left(1 - \ell^\top ((1 - \mu)I)^{-1} \mathbf{1}_N\right) = (1 - \mu)^N \left(1 - \frac{1}{1 - \mu} \ell^\top \mathbf{1}_N\right).$$

Here, we use $A^{-1} = \frac{1}{1-\mu} I$ and $\det(A) = (1 - \mu)^N$. Since all weights sum to one, we have

$$\ell^\top \mathbf{1}_N = \sum_{n=1}^N a_n = 1.$$

Substituting this into the equation gives

$$\det(\hat{L} - \mu I) = (1 - \mu)^N \left(1 - \frac{1}{1 - \mu}\right).$$

As a result, the determinant satisfies $\det(\hat{L} - \mu I) = 0$ if either

$$(1 - \mu)^N = 0 \quad \text{or} \quad 1 - \frac{1}{1 - \mu} = 0.$$

The second conditions simplifies to $1 - \mu = 1$, which holds for $\mu = 0$. The first condition implies $\mu = 1$. Thus, the eigenvalues of $\hat{L}$ are

$$\mu_1(\hat{L}) = 0, \quad \mu_2(\hat{L}) = \ldots = \mu_N(\hat{L}) = 1.$$

For the corresponding eigenvectors to $\mu_1(\hat{L}) = 0$ we compute

$$\hat{L}\mathbf{1}_N = (I - \mathbf{1}_N \ell^\top)\mathbf{1}_N = \mathbf{1}_N - \mathbf{1}_N \ell^\top \mathbf{1}_N = \mathbf{1}_N - \mathbf{1}_N = \mathbf{0}_N = 0 \cdot \mathbf{1}_N = \mu_1(\hat{L}) \cdot \mathbf{1}_N$$

for the right-eigenvector and

$$\hat{L}^\top \ell = (I - \mathbf{1}_N \ell^\top)^\top \ell = \ell - \ell \mathbf{1}_N^\top \ell = \ell - \ell = \mathbf{0}_N = 0 \cdot \ell = \mu_1(\hat{L}) \cdot \ell$$

for the left-eigenvector. Thus, we conclude that both $\mathbf{1}_N$ and ℓ are right- and left-eigenvectors corresponding to the eigenvalue $\mu_1(\hat{L}) = 0$. The key property used in this derivation is that the sum of all weights equals one. $\square$

Lemma 2.8 reveals that the matrix $\hat{L}$ has a highly structured spectrum: it has exactly one eigenvalue equal to zero, while all remaining $N - 1$ eigenvalues are equal to one. The right-eigenvector associated with the zero eigenvalue is $\mathbf{1}_N$, the all-ones vector.

By restricting condition (2.7) to a fixed dimension d, we find that all equilibrium points of the deterministic system $dX_t^{:,d} = -\lambda \hat{L} X_t^{:,d} dt$ satisfy

$$x_t^{1,d} = \ldots = x_t^{N,d}, \tag{2.13}$$

i.e., all particle components agree. Thus, any equilibrium point lies in the one-dimensional subspace spanned by the eigenvector $\mathbf{1}_N$. This spectral structure admits a dynamical interpretation: the eigenspace associated with the zero eigenvalue corresponds to a direction in which the system does not evolve. Once all particles are aligned, the dynamics stops changing. In contrast, all directions orthogonal to $\mathbf{1}_N$ correspond to eigenvalue one and are actively contracted by the dynamics. As a result, the system is continuously pushed toward the subspace

$$\mathcal{G}^{\star,d} = \{x \in \mathbb{R}^N : x = \gamma \mathbf{1}_N, \gamma \in \mathbb{R}\}$$

of equilibrium points.

Motivated by these observations, we consider the function

$$V(X_t^{:,d}) = \left\| X_t^{:,d} - \bar{X}_t^{:,d} \mathbf{1}_N \right\|^2, \tag{2.14}$$

where, $\bar{X}_t^{:,d}$ denotes the average particle value in dimension d, defined as

$$\bar{X}_t^{:,d} := \frac{1}{N} \mathbf{1}_N^\top X_t^{:,d} = \frac{1}{N} \sum_{n=1}^{N} x_t^{n,d}. \tag{2.15}$$

Since $\bar{X}_t^{:,d}$ is actually the orthogonal projection of $X_t^{:,d}$ onto $\mathcal{G}^{\star,d}$, we also have

$$V(X_t^{:,d}) = \text{dist}(X_t^{:,d}; \mathcal{G}^{\star,d})^2 := \inf_{y \in \mathcal{G}^{\star,d}} \left\| X_t^{:,d} - y \right\|^2.$$

Thus, V measures the squared distance of the particle system to its closest equilibrium point. Clearly, the difference $X_t^{:,d} - \bar{X}_t^{:,d} \mathbf{1}_N \in \mathbb{R}^N$ captures how far each individual particle deviates from the mean, so an alternative interpretation of V is that this function measures the squared distance of $X_t^{:,d}$ to its mean value. However, note that V is not a classical Lyapunov function, since that we have $V(x) = 0$ for all $x \in \mathcal{G}^{\star,d}$.

Now define the auxiliary functions

$$g : \mathbb{R}^N \to \mathbb{R}, \quad g(x) = x^\top x, \qquad h : \mathbb{R}^N \to \mathbb{R}^N, \quad h(x) = x - \frac{1}{N} \mathbf{1}_N^\top x \mathbf{1}_N,$$

which satisfy

$$\nabla g(x) = 2x, \qquad Jh(x) = I - \frac{1}{N}\mathbf{1}_N\mathbf{1}_N^\top$$

for $x \in \mathbb{R}^N$. The expression for the Jacobi-matrix of h follows from the observation that

$$\partial_{x_m}(h(x))_k = \partial_{x_m}\left(x_k - \frac{1}{N}\mathbf{1}_N^\top x\right) = \mathbf{1}_{\{m=k\}} - \frac{1}{N}$$

for $m, k = 1, \ldots, N$. Here, $\mathbf{1}_{\{m=k\}}$ equals 1 when $m = k$ and 0 otherwise.

Thus, we have

$$V(x) = g(h(x))$$

and so the chain rule yields

$$\nabla V(x) = \left(I - \frac{1}{N}\mathbf{1}_N\mathbf{1}_N^\top\right)^\top 2\left(x - \frac{1}{N}\mathbf{1}_N^\top x \mathbf{1}_N\right)$$

$$= 2\left(x - \frac{1}{N}\mathbf{1}_N^\top x \mathbf{1}_N\right) - 2\frac{1}{N}\mathbf{1}_N\mathbf{1}_N^\top x + 2\frac{1}{N^2}\mathbf{1}_N\mathbf{1}_N^\top x \mathbf{1}_N^\top \mathbf{1}_N$$

$$= 2\left(x - \frac{1}{N}\mathbf{1}_N^\top x \mathbf{1}_N\right).$$

Here we have used that $x^\top \mathbf{1}_N = \mathbf{1}_N^\top x$ is a scalar and $\mathbf{1}_N^T\mathbf{1}_N = N$.

Using this expression for ∇V in (2.11) we obtain

$$\dot{V}(X_t^{:,d}) = -2\lambda(X_t^{:,d} - \bar{X}_t^{:,d}\mathbf{1}_N)^\top \hat{L}X_t^{:,d} \tag{2.16}$$

Now, we will first rewrite and simplify the expression on the right-hand side of the above equality. To this end, we use the representation $\hat{L} = I - \mathbf{1}_N\ell^\top$ established in the proof of Lemma 2.8. Moreover, we have $\bar{X}_t^{:,d} = \frac{1}{N}\mathbf{1}_N^\top X_t^{:,d}$, which yields

$$\hat{L}(\bar{X}_t^{:,d}\mathbf{1}_N) = (I - \mathbf{1}_N\ell^\top)\left(\frac{1}{N}\mathbf{1}_N^\top X_t^{:,d}\mathbf{1}_N\right) = \frac{1}{N}\mathbf{1}_N^\top X_t^{:,d}\mathbf{1}_N - \frac{1}{N}\mathbf{1}_N\ell^\top\mathbf{1}_N^\top X_t^{:,d}\mathbf{1}_N.$$

Since $\mathbf{1}_N^\top X_t^{:,d}$ is a scalar, we can rearrange terms and obtain

$$\hat{L}(\bar{X}_t^{:,d}\mathbf{1}_N) = \frac{1}{N}\mathbf{1}_N^\top X_t^{:,d}\mathbf{1}_N - \frac{1}{N}\mathbf{1}_N^\top X_t^{:,d}\mathbf{1}_N \cdot \ell^\top\mathbf{1}_N = 0, \tag{2.17}$$

where we used that the weights sum to one,

$$\ell^\top\mathbf{1}_N = \sum_{n=1}^N a_n = 1.$$

As a result, we can rewrite the total derivative as

$$\dot{V}(X_t^{:,d}) = -2\lambda(X_t^{:,d} - \bar{X}_t^{:,d}\mathbf{1}_N)^\top \hat{L}X_t^{:,d} = -2\lambda(X_t^{:,d} - \bar{X}_t^{:,d}\mathbf{1}_N)^\top \hat{L}(X_t^{:,d} - \bar{X}_t^{:,d}\mathbf{1}_N),$$

since the additional term vanishes due to (2.17). For brevity, we define the error vector $E_t^{:,d} := X_t^{:,d} - \bar{X}_t^{:,d}\mathbf{1}_N$ and compute

$$\dot{V}(X_t^{:,d}) = -2\lambda(E_t^{:,d})^\top \hat{L}E_t^{:,d}$$

$$= -2\lambda(E_t^{:,d})^\top(I - \mathbf{1}_N\ell^\top)E_t^{:,d} \tag{2.18}$$

$$= -2\lambda(E_t^{:,d})^\top E_t^{:,d} + 2\lambda(E_t^{:,d})^\top\mathbf{1}_N\ell^\top E_t^{:,d}.$$

To simplify the second term, we substitute back $E_t^{:,d} = X_t^{:,d} - \bar{X}_t^{:,d}\mathbf{1}_N$, use that $(X_t^{:,d})^\top\mathbf{1}_N$ is a scalar and $\mathbf{1}_N^\top\mathbf{1}_N = N$:

$$(E_t^{:,d})^\top\mathbf{1}_N\ell^\top = \left(X_t^{:,d} - \frac{1}{N}\mathbf{1}_N^\top X_t^{:,d}\mathbf{1}_N\right)^\top \mathbf{1}_N\ell^\top$$

$$= (X_t^{:,d})^\top\mathbf{1}_N\ell^\top - \frac{1}{N}\mathbf{1}_N^\top(X_t^{:,d})^\top\mathbf{1}_N\mathbf{1}_N\ell^\top$$

$$= (X_t^{:,d})^\top\mathbf{1}_N\ell^\top - \frac{N}{N}(X_t^{:,d})^\top\mathbf{1}_N\ell^\top$$

$$= \mathbf{0}_N^\top.$$

Hence, the second term in (2.18) vanishes, and observing that $V(X_t^{:,d}) = \|E_t^{:,d}\|^2$ we obtain

$$\dot{V}(X_t^{:,d}) = -2\lambda(E_t^{:,d})^\top E_t^{:,d} = -2\lambda\|E_t^{:,d}\|^2 = -2\lambda V(X_t^{:,d}).$$

Solving this linear ODE yields

$$V(X_t^{:,d}) = e^{-2\lambda t}V(X_0^{:,d}), \qquad t \geq 0, \tag{2.19}$$

so this distance of the particle system to its subspace of equilibrium points decays exponentially with rate λ.

At this point, we require a change of coordinates, if we want to work in the setup of Definitions 2.2 and 2.3, since we have a whole subspace $\mathcal{G}^{*,d}$ of equilibrium point instead of $x^{\star,d} = \mathbf{0}_N$. This is also reflected by the fact that V can not be a Lyapunov function, since $V(x) = 0$ for all $x \in \mathcal{G}^{\star,d}$.

From equation (2.19) we already know that

$$\|E_t^{:,d}\|^2 = e^{-2\lambda t}\|E_0^{:,d}\|^2, \qquad t \geq 0, \tag{2.20}$$

so we will now turn to $E_t^{:,d} = X_t^{:,d} - \bar{X}_t^{:,d}\mathbf{1}_N$, which is the dynamics of the particle system that is orthogonal to $\mathcal{G}^{:,d}$. In fact, we have

$$E_t^{:,d} = X_t^{:,d} - \bar{X}_t^{:,d}\mathbf{1}_N = X_t^{:,d} - \frac{1}{N}\mathbf{1}_N^T X_t^{:,d}\mathbf{1}_N = \left(I - \frac{1}{N}\mathbf{1}_N\mathbf{1}_N^T\right)X_t^{:,d}$$

and so

$$\dot{E}_t^{:,d} = -\lambda P\hat{L}X_t^{:,d},$$

where

$$P = I - \frac{1}{N}\mathbf{1}_N\mathbf{1}_N^T$$

is the matrix which corresponds to the projection on the orthogonal complement of $\mathcal{G}^{\star,d}$. Since

$$P\mathbf{1}_N = \mathbf{1}_N - \frac{1}{N}\mathbf{1}_N\mathbf{1}_N^T\mathbf{1}_N = \mathbf{0}_N, \tag{2.21}$$

we have

$$P\hat{L} = P - P\mathbf{1}_N\ell^T = P. \tag{2.22}$$

Thus, the dynamics of $E_t^{:,d}$ is given by the linear ODE

$$\dot{E}_t^{:,d} = -\lambda E_t^{:,d},$$

which strikingly clarifies the dynamics of the particle system. This ODE has the unique solution

$$E_t^{:,d} = e^{-\lambda t}E_0^{:,d}, \qquad t \geq 0, \tag{2.23}$$

which is obviously exponentially stable with rate λ at the unique fixed point $E^{\star,d} = \mathbf{0}_N$ and we trivially recover equation (2.20).

Since equation (2.23) does not depend on the considered dimension d of the particles, we have obtained the following result for the original dynamics:

Theorem 2.9 (Global Exponential Stability of the Deterministic System)**.** The deterministic system

$$dX_t = -\lambda LX_t dt$$

with $X_t = (x_t^{1,1}, \ldots, x_t^{1,D}, \ldots, x_t^{N,1}, \ldots, x_t^{N,D})^\top \in \mathbb{R}^{DN}$ is **globally exponentially stable** with exact rate λ on the set of equilibrium points

$$\mathcal{G}^\star = \left\{x \in \mathbb{R}^{N\times D} : x^{1,d} = \ldots x^{N,d}, d = 1, \ldots, D\right\}.$$

2.3 Exponential Stability under Discretization

In practical applications, usually the explicit Euler scheme

$$\widehat{X}^n_{(k+1)\Delta} = \widehat{X}^n_{k\Delta} - \lambda \left(\widehat{X}^n_{k\Delta} - \nu^\alpha_f(\widehat{X}^n_{k\Delta}) \right) \Delta, \qquad k = 0, 1, \ldots \tag{2.24}$$

with stepsize $\Delta > 0$ is used for the simulation of

$$dX^n_t = -\lambda \left(X^n_t - \nu^\alpha_f(X_t) \right) dt$$

for $n = 1, \ldots, N$. Proceeding as in Subsection 2.1 it will again suffice to study the system in each dimension $d \in \{1, \ldots D\}$ on its own. So, for $X^{:,d}_t = (x^{1,d}_t, \ldots, x^{N,d}_t)^\top \in \mathbb{R}^N$, the Euler-discretzation of

$$dX^{:,d}_t = -\lambda \hat{L}(X_t) X^{:,d}_t dt,$$

reads as

$$\widehat{X}^{:,d}_{(k+1)\Delta} = \widehat{X}^{:,d}_{k\Delta} - \lambda \hat{L}(\widehat{X}_{k\Delta}) \widehat{X}^{:,d}_{k\Delta} \Delta, \qquad k = 0, 1, \ldots \tag{2.25}$$

where $\hat{L}(\widehat{X}_{k\Delta})$, is defined as before, i.e., as

$$\hat{L}(\widehat{X}_{k\Delta}) := \begin{pmatrix} 1 - a_1(\widehat{X}_{k\Delta}) & -a_2(\widehat{X}_{k\Delta}) & \cdots & -a_N(\widehat{X}_{k\Delta}) \\ -a_1(\widehat{X}_{k\Delta}) & 1 - a_2(\widehat{X}_{k\Delta}) & \cdots & -a_N(\widehat{X}_{k\Delta}) \\ \vdots & \vdots & \ddots & \vdots \\ -a_1(\widehat{X}_{k\Delta}) & -a_2(\widehat{X}_{k\Delta}) & \cdots & 1 - a_N(\widehat{X}_{k\Delta}) \end{pmatrix}. \tag{2.26}$$

Now we analyze also the dynamics of the Euler scheme on the orthogonal complement of $\mathcal{G}^{\star,d}$, that is of

$$\widehat{E}^{:,d}_{k\Delta} = \left(I - \frac{1}{N} \mathbf{1}_N \mathbf{1}_N^T \right) \widehat{X}^{:,d}_{k\Delta}.$$

With

$$P = I - \frac{1}{N} \mathbf{1}_N \mathbf{1}_N^T$$

we have

$$\begin{aligned} \widehat{E}^{:,d}_{(k+1)\Delta} &= P \widehat{X}^{:,d}_{(k+1)\Delta} \\ &= P \widehat{X}^{:,d}_{k\Delta} - \lambda P \hat{L}(\widehat{X}_{k\Delta}) \widehat{X}^{:,d}_{k\Delta} \Delta \\ &= \widehat{E}^{:,d}_{k\Delta} - \lambda P \widehat{X}^{:,d}_{k\Delta} \Delta + \lambda P \mathbf{1}_N \ell(\widehat{X}_{k\Delta})^T \widehat{X}^{:,d}_{k\Delta} \Delta \\ &= (1 - \lambda \Delta) \widehat{E}^{:,d}_{k\Delta} \end{aligned}$$

where we have used again (2.21), i.e. $P \mathbf{1}_N = \mathbf{0}_N$. So, we obtain

$$\widehat{E}^{:,d}_{k\Delta} = (1 - \lambda \Delta)^k E^{:,d}_0, \qquad k = 0, 1, \ldots. \tag{2.27}$$

Since we have

$$\ln(1-x) \leq -x, \qquad x \in [0, 1),$$

we obtain that

$$(1 - \lambda \Delta)^k = \exp(k \ln(1 - \lambda \Delta)) \leq \exp(-\lambda k \Delta)$$

under the usual stability property $\Delta \lambda < 1$ for the Euler scheme. This implies

$$\|\widehat{E}^{:,d}_{k\Delta}\| \leq e^{-\lambda k \Delta} \|E^{:,d}_0\|, \tag{2.28}$$

so the Euler scheme preserves the globally exponential stability with rate λ of the equilibrium points from Theorem 2.6 under $\lambda \Delta < 1$.

3 Stochastic System

We now turn to the fully stochastic formulation of CBO, which incorporates both deterministic drift and Brownian diffusion. To rigorously analyze stability in this setting, we extend the deterministic framework to stochastic differential equations of the form $dZ_t = a(Z_t)dt + b(Z_t)dW_t$. Our goal is to establish a result analogous to the global exponential stability previously shown in the deterministic setting. To this end, we introduce the necessary concepts for exponential stability.

Definition 3.1 (Equilibrium Point [17, Definiton 3.2])**.** The state $Z^\star \in \mathbb{R}^N$ is said to be an *equilibrium point* of an stochastic differential equation (SDE)

$$dZ_t = a(Z_t)dt + b(Z_t)dW_t, \quad t \geq 0, \tag{3.1}$$

if once Z_t is equal to $Z^\star$, it remains equal to $Z^\star$ for all future times, i.e., $Z^\star$ has to satisfy $a(Z^\star) = \mathbf{0}_N$ and $b(Z^\star) = \mathbf{0}_N$.

Definition 3.2 (Amost Sure Exponential Stability for SDE [**?**, Section 4, Definition 3.1])**.** The equilibrium point $Z^\star = \mathbf{0}_N$ is said to be *almost surely globally exponentially stable* with rate δ if there exists a constant $\delta > 0$ such that for all $Z_0 \in \mathbb{R}^N$,

$$\mathbb{P}\left(\limsup_{t \to \infty} \frac{1}{t} \ln\left(\|Z_t\|\right) \leq -\delta\right) = 1.$$

In addition to this *path-wise* notion of stability, another fundamental concept is based on the behavior of the second moment. This leads to the idea of *mean square stability*, where one does not analyze individual trajectories but instead considers the expected squared norm of the solution.

Definition 3.3 (Mean Square Exponential Stability for SDE [**?**, Section 4, Definition 4.1])**.** The equilibrium point $Z^\star = \mathbf{0}_N$ is said to be *mean square globally exponentially stable* with rate δ if there exist constants $c, \delta > 0$ such that for all $Z_0 \in \mathbb{R}^N$,

$$\mathbb{E}\left(\|Z_t\|^2\right) \leq c e^{-\delta t} \mathbb{E}\left(\|Z_0\|^2\right) \quad \text{for all } t \geq 0.$$

In Chapter 2.2, we employed Lyapunov functions to motivate a change of coordinates, specifically the orthogonal projection onto $\mathcal{G}^{\star,d}$. This allowed us to exploit a well-chosen linearization of the transformed system to establish exponential stability without further relying on Lyapunov functions. In the stochastic setting, we proceed analogously by once again applying the same change of coordinates. This enables us to derive both almost sure exponential stability and mean square exponential stability. To this end, we reformulate our system in the same spirit as in the deterministic case.

3.1 System Description

The dynamics of the system are governed by the stochastic differential equation

$$dZ_t^n = -\lambda \left(Z_t^n - \nu_f^\alpha(Z_t)\right) dt + \sigma \left(Z_t^n - \nu_f^\alpha(Z_t)\right) dW_t^n \tag{3.2}$$

for $n = 1, \ldots, N$, where $W_t^1, \ldots W_t^N$ are independent d-dimensional Brownian motions, and $Z_t = (Z_t^1, \ldots, Z_t^N)^\top$ represents the particle states.

As in the deterministic case, we can reformulate $Z_t^n - \nu_f^\alpha(Z_t)$ by expressing it in terms of pair-wise differences. This transformation can be applied to both the drift and the diffusion terms, leading to the equivalent representation

$$dZ_t^n = -\lambda \left(\sum_{\substack{m=1 \\ m \neq n}}^N a_m \cdot (Z_t^n - Z_t^m)\right) dt + \sigma \left(\sum_{\substack{m=1 \\ m \neq n}}^N a_m \cdot (Z_t^n - Z_t^m)\right) dW_t^n \tag{3.3}$$

for $n = 1, \ldots, N$ with $a_m = a_m(Z_t)$ as defined in section 2.1. Additionally, we define $Z_t^{:,d} := (z_t^{1,d}, \ldots, z_t^{N,d})^\top$ to organize the system component-wise across dimensions. Sorting the system by dimension, we obtain

the following structured representation

$$
dZ_t = d \begin{pmatrix} z_t^{1,1} \\ \vdots \\ z_t^{N,1} \\ \vdots \\ z_t^{1,D} \\ \vdots \\ z_t^{N,D} \end{pmatrix} = -\lambda \begin{pmatrix} \sum\limits_{m=2}^{N} a_m \left(z_t^{1,1} - z_t^{m,1} \right) \\ \vdots \\ \sum\limits_{m=1}^{N-1} a_m \left(z_t^{N,1} - z_t^{m,1} \right) \\ \vdots \\ \sum\limits_{m=2}^{N} a_m \left(z_t^{1,D} - z_t^{m,D} \right) \\ \vdots \\ \sum\limits_{m=1}^{N-1} a_m \left(z_t^{N,D} - z_t^{m,D} \right) \end{pmatrix} dt + \sigma \begin{pmatrix} \sum\limits_{m=2}^{N} a_m \left(z_t^{1,1} - z_t^{m,1} \right) \\ \vdots \\ \sum\limits_{m=1}^{N-1} a_m \left(z_t^{N,1} - z_t^{m,1} \right) \\ \vdots \\ \sum\limits_{m=2}^{N} a_m \left(z_t^{1,D} - z_t^{m,D} \right) \\ \vdots \\ \sum\limits_{m=1}^{N-1} a_m \left(z_t^{N,D} - z_t^{m,D} \right) \end{pmatrix} d \begin{pmatrix} w_t^{1,1} \\ \vdots \\ w_t^{N,1} \\ \vdots \\ w_t^{1,D} \\ \vdots \\ w_t^{N,D} \end{pmatrix} \tag{3.4}
$$

where $W_t^n = (w_t^{n,1}, \ldots, w_t^{n,D})^\top$ captures the Brownian motion associated with each component.

Since all components of each Brownian motion are independent, the evolution of each dimension is independent of the others, given the weights $a_1, \ldots, a_N$, in the same way as in the deterministic case. Consequently, it suffices once more to analyze stability separately for each dimension $d \in \{1, \ldots, D\}$.

Using the notation from Section 2, the system dynamics for a single dimension d can be expressed compactly as

$$
dZ_t^{:,d} = -\lambda \hat{L} Z_t^{:,d} dt + \sigma \hat{L} Z_t^{:,d} dW_t^{:,d}, \tag{3.5}
$$

where $W_t^{:,d} = (w_t^{1,d}, \ldots, w_t^{N,d})^\top \in \mathbb{R}^N$ represents the Brownian motion in dimension d. Similarly, the full system in (3.4) can be rewritten in matrix form as

$$
dZ_t = -\lambda L Z_t dt + \sigma L Z_t dW_t \tag{3.6}
$$

with $W_t = \left(W_t^1, \ldots, W_t^N \right)^\top \in \mathbb{R}^{DN}$.

In the following, we extend the stability analysis to stochastic CBO dynamics. In particular, we will demonstrate that the system (3.6) is almost sure and mean square globally exponentially stable on the set

$$
\mathcal{G}^\star = \left\{ z \in \mathbb{R}^{N \times D} : z^{1,d} = \ldots = z^{N,d}, d = 1, \ldots, D \right\}, \tag{3.7}
$$

which characterizes the set of equilibrium points of the system (3.6). This follows directly from the same calculations as in (2.7) and (2.8), since both the drift term $a(X_t)$ and the diffusion term $b(X_t)$ share the same functional form, differing only by a scaling factor.

3.2 Global Exponential Stability

Since the equilibrium manifold $\mathcal{G}^\star$ remains unchanged in the stochastic setting, we employ the same change of coordinates as in the deterministic case. Furthermore, due to the independence of the Brownian motions across particles and dimensions, it suffices to analyze each dimension $d \in \{1, \ldots, D\}$ separately.

We consider the projected dynamics orthogonal to the consensus manifold by

$$
E_t^{:,d} = Z_t^{:,d} - \bar{Z}_t^{:,d} \mathbf{1}_N, \quad \text{where} \quad \bar{Z}_t^{:,d} = \frac{1}{N} \mathbf{1}_N^T Z_t^{:,d},
$$

so that $E_t^{:,d}$ lies in the orthogonal complement of

$$
\mathcal{G}^{\star,d} := \{ z \in \mathbb{R}^N : z = \gamma \mathbf{1}_N, \, \gamma \in \mathbb{R} \}.
$$

Analogously to the deterministic case, this leads to the convenient representation

$$
E_t^{:,d} = Z_t^{:,d} - \bar{Z}_t^{:,d} \mathbf{1}_N = \left(I - \frac{1}{N} \mathbf{1}_N \mathbf{1}_N^T \right) Z_t^{:,d} = P Z_t^{:,d},
$$

where P is the projection onto the orthogonal complement of the consensus space.

In dimension d, the original system evolves according to

$$
dZ_t^{:,d} = -\lambda \hat{L} Z_t^{:,d} dt + \sigma \hat{L} Z_t^{:,d} dW_t^{:,d}.
$$

12

Since P is a constant matrix (independent of the particle positions), it can be pulled inside the differential. Applying this projection yields

$$dE_t^{:,d} = d(PZ_t^{:,d}) = -\lambda P\hat{L}Z_t^{:,d}\,dt + \sigma P\hat{L}Z_t^{:,d}\,dW_t^{:,d}.$$

Using the identity $P\hat{L} = P$, see (2.22), the system simplifies to

$$dE_t^{:,d} = -\lambda PZ_t^{:,d}\,dt + \sigma PZ_t^{:,d}\,dW_t^{:,d} = -\lambda E_t^{:,d}\,dt + \sigma E_t^{:,d}\,dW_t^{:,d}, \tag{3.8}$$

which has the form of a linear stochastic differential equation.

This linear system admits the unique explicit solution

$$E_t^{:,d} = \exp\left(-\left(\lambda + \frac{\sigma^2}{2}\right)tI + \sigma W_t^{:,d}\right)E_0^{:,d}, \quad t \geq 0,$$

and, componentwise for each particle $n \in \{1,\ldots,N\}$,

$$E_t^{n,d} = \exp\left(-\left(\lambda + \frac{\sigma^2}{2}\right)t + \sigma W_t^n\right)E_0^{n,d}, \quad t \geq 0. \tag{3.9}$$

Applying the logarithm to the absolute value of (3.9) and dividing by $t > 0$ yields

$$\frac{1}{t}\ln\left(|E_t^{n,d}|\right) = -\left(\lambda + \frac{\sigma^2}{2}\right) + \sigma\frac{W_t^n}{t} + \frac{1}{t}\ln\left(|E_0^{n,d}|\right). \tag{3.10}$$

By the law of the iterated logarithm [?, Theorem 4.2], we have

$$\limsup_{t\to\infty}\frac{W_t^n}{\sqrt{2t\log(\log t)}} = 1 \quad \text{almost surely.}$$

In particular, this implies

$$\limsup_{t\to\infty}\frac{W_t^n}{t} = 0 \quad \text{almost surely.}$$

Moreover, since the last term in (3.10) satisfies

$$\limsup_{t\to\infty}\frac{1}{t}\ln\left(|E_0^{n,d}|\right) = 0,$$

we may take the lim sup of (3.10) to conclude

$$\limsup_{t\to\infty}\frac{1}{t}\ln\left(|E_t^{n,d}|\right) = -\left(\lambda + \frac{\sigma^2}{2}\right) \quad \text{almost surely.} \tag{3.11}$$

Due to the equivalence of norms in finite-dimensional spaces, there exist constants $c_1, c_2 > 0$ such that

$$c_1 \cdot \max_{n=1,\ldots,N}|E_t^{n,d}| \leq \|E_t^{:,d}\| \leq c_2 \cdot \max_{n=1,\ldots,N}|E_t^{n,d}|.$$

Taking logarithms and dividing by t gives

$$\frac{1}{t}\ln(c_1) + \frac{1}{t}\ln\left(\max_{n=1,\ldots,N}|E_t^{n,d}|\right) \leq \frac{1}{t}\ln\left(\|E_t^{:,d}\|\right) \leq \frac{1}{t}\ln(c_2) + \frac{1}{t}\ln\left(\max_{n=1,\ldots,N}|E_t^{n,d}|\right).$$

Passing to the limit superior as $t \to \infty$, and using (3.11), we obtain

$$\limsup_{t\to\infty}\frac{1}{t}\ln\left(\|E_t^{:,d}\|\right) = \limsup_{t\to\infty}\frac{1}{t}\ln\left(\max_{n=1,\ldots,N}|E_t^{n,d}|\right) = -\left(\lambda + \frac{\sigma^2}{2}\right) \quad \text{almost surely.} \tag{3.12}$$

This proves that the system (3.8) is almost surely globally exponentially stable on $\mathcal{G}^{\star,d}$ with exponential rate

$$\delta = \lambda + \frac{\sigma^2}{2}.$$

Since this analysis holds for each dimension d independently, the result carries over to the full system, completing the proof of almost sure global exponential stability for the original dynamics.

Theorem 3.4 (Almost Sure Global Exponential Stability of the Stochastic System)**.** The stochastic system

$$dZ_t = -\lambda L Z_t dt + \sigma L Z_t dW_t$$

with $Z_t = (z_t^{1,1}, \ldots, z_t^{1,D}, \ldots, z_t^{N,1}, \ldots, z_t^{N,D})^\top \in \mathbb{R}^{DN}$ is **almost surely globally exponentially stable** with exact rate $\lambda + \frac{\sigma^2}{2}$ on the set of equilibrium points

$$\mathcal{G}^\star = \left\{ z \in \mathbb{R}^{N \times D} : z^{1,d} = \ldots = z^{N,d}, \, d = 1, \ldots, D \right\}.$$

To establish not only almost sure exponential convergence but also convergence in mean square, we now consider the squared explicit solution (3.9) of the linear SDE in expectation:

$$\mathbb{E}\left(|E_t^{n,d}|^2 \right) = \mathbb{E}\left(\exp\left(-2\left(\lambda + \frac{\sigma^2}{2}\right)t + 2\sigma W_t^n \right) |E_0^{n,d}|^2 \right).$$

Since W_t^n is independent of the initial condition $E_0^{n,d}$, the expectation factorizes

$$\mathbb{E}\left(|E_t^{n,d}|^2 \right) = \exp\left(-2\left(\lambda + \frac{\sigma^2}{2}\right)t \right) \mathbb{E}\left(\exp(2\sigma W_t^n) \right) \mathbb{E}\left(|E_0^{n,d}|^2 \right).$$

Recalling that $W_t^n \sim \mathcal{N}(0, t)$, the random variable $2\sigma W_t^n$ is Gaussian with mean 0 and variance $4\sigma^2 t$. For a normally distributed random variable $X \sim \mathcal{N}(\mu, \sigma^2)$, it is well known that

$$\mathbb{E}\left(e^{aX} \right) = \exp\left(a\mu + \tfrac{1}{2} a^2 \sigma^2 \right).$$

Applying this identity with $X = W_t^n$, $a = 2\sigma$, $\mu = 0$, and $\text{Var}(W_t^n) = t$, we obtain

$$\mathbb{E}\left(\exp(2\sigma W_t^n) \right) = \exp\left(\tfrac{1}{2}(2\sigma)^2 t \right) = \exp\left(2\sigma^2 t \right).$$

Hence, the mean squared value simplifies to

$$\mathbb{E}\left(|E_t^{n,d}|^2 \right) = \exp\left(-2\left(\lambda + \frac{\sigma^2}{2}\right)t \right) \exp\left(2\sigma^2 t \right) \mathbb{E}\left(|E_0^{n,d}|^2 \right)$$

$$= \exp\left(-2\left(\lambda - \frac{\sigma^2}{2}\right)t \right) \mathbb{E}\left(|E_0^{n,d}|^2 \right).$$

Since this relation holds for each $n = 1, \ldots, N$, we can apply the linearity of expectation to obtain a bound for the full vector

$$\mathbb{E}\left(\|E_t^{:,d}\|^2 \right) = \exp\left(-2\left(\lambda - \frac{\sigma^2}{2}\right)t \right) \mathbb{E}\left(\|E_0^{:,d}\|^2 \right). \tag{3.13}$$

Therefore, the system (3.8) is mean square globally exponentially stable on $\mathcal{G}^{\star,d}$, with stability constant $c = 1$ and exponential decay rate

$$\delta = 2\lambda - \sigma^2. \tag{3.14}$$

Since this analysis holds for each dimension d independently, the result carries over to the full system, completing the proof of mean square global exponential stability for the original dynamics.

Theorem 3.5 (Mean Square Global Exponential Stability of the Stochastic System)**.** The stochastic system

$$dZ_t = -\lambda L Z_t dt + \sigma L Z_t dW_t$$

with $Z_t = (z_t^{1,1}, \ldots, z_t^{1,D}, \ldots, z_t^{N,1}, \ldots, z_t^{N,D})^\top \in \mathbb{R}^{DN}$ is **mean square globally exponentially stable** with constant $c = 1$ and exact rate $2\lambda - \sigma^2$ on the set of equilibrium points

$$\mathcal{G}^\star = \left\{ z \in \mathbb{R}^{N \times D} : z^{1,d} = \ldots = z^{N,d}, \, d = 1, \ldots, D \right\}.$$

In order for the system to be mean square exponentially stable, it is necessary that the decay rate $\delta = 2\lambda - \sigma^2$ is positive. This requires the condition

$$2\lambda > \sigma^2. \tag{3.15}$$

We emphasize that, independently of this condition, almost sure exponential convergence always holds for all values of $\lambda > 0$ and $\sigma \geq 0$, as shown earlier.

We further note that the exponential convergence of the full CBO system toward the global minimizer $\mathbf{X}^\star \in \mathbb{R}^D$ has been established in [10], where the authors derive a maximal convergence rate

$$\delta_{\mathrm{FKR}} = 2\lambda - D\sigma^2, \tag{3.16}$$

which depends linearly on the dimension D. In contrast, the rate we obtain here for the orthogonal dynamics, $\delta = 2\lambda - \sigma^2$, is independent of D and thus sharper in high dimensions. Moreover, our result provides the exact mean square decay rate, not merely an upper bound. However, it is important to note that our analysis focuses exclusively on the component of the dynamics orthogonal to the consensus manifold $\mathcal{G}^\star$. As a consequence, while we establish exponential convergence toward the manifold $\mathcal{G}^\star$, our framework does not allow for conclusions about convergence toward a specific point on this manifold. Investigating this aspect, i.e., the identification of the limiting point within $\mathcal{G}^\star$, is beyond the scope of the present work and will be addressed in future research.

3.3 Exponential Stability under Discretization

Usually, the Euler–Maruyama scheme is employed to simulate the dynamics of the stochastic system

$$dZ_t^n = -\lambda \left(Z_t^n - \nu_f^\alpha(Z_t) \right) dt + \sigma \left(Z_t^n - \nu_f^\alpha(Z_t) \right) dW_t^n, \qquad n = 1, \ldots, N.$$

The discretized approximation $\widehat{Z}_{k\Delta}^n$ with time step size $\Delta > 0$ is given by the Euler–Maruyama scheme

$$\widehat{Z}_{(k+1)\Delta}^n = \widehat{Z}_{k\Delta}^n - \lambda \left(\widehat{Z}_{k\Delta}^n - \nu_f^\alpha(\widehat{Z}_{k\Delta}) \right) \Delta + \sigma \left(\widehat{Z}_{k\Delta}^n - \nu_f^\alpha(\widehat{Z}_{k\Delta}) \right) \circ \Delta W_k^n, \tag{3.17}$$

where $\circ$ denotes the element-wise (Hadamard) product [19, Definition 7.5.1]. The random vectors $\Delta W_k^n \sim \mathcal{N}(\mathbf{0}_D, \Delta \cdot I_D) \in \mathbb{R}^D$ are independent and identically distributed Gaussian approximations of the Brownian motion increments $W_{(k+1)\Delta}^n - W_{k\Delta}^n$. Thus, the Euler–Maruyama scheme combines a forward Euler step for the drift with a pointwise scaled Gaussian perturbation to model the stochastic diffusion component.

As in Subsection 3.1, it suffices to study the dynamics of each coordinate $d \in \{1, \ldots, D\}$ separately. For $Z_t^{:,d} := (Z_t^{1,d}, \ldots, Z_t^{N,d})^\top \in \mathbb{R}^N$ the stochastic system

$$dZ_t^{:,d} = -\lambda \hat{L}(Z_t) Z_t^{:,d} dt + \sigma \hat{L}(Z_t) Z_t^{:,d} dW_t^{:,d}$$

can be approximated via the Euler–Maruyama scheme as

$$\widehat{Z}_{(k+1)\Delta}^{:,d} = \widehat{Z}_{k\Delta}^{:,d} - \lambda \hat{L}(\widehat{Z}_{k\Delta}) \widehat{Z}_{k\Delta}^{:,d} \Delta + \sigma \hat{L}(\widehat{Z}_{k\Delta}) \widehat{Z}_{k\Delta}^{:,d} \circ \Delta W_k^{:,d}, \qquad k = 0, 1, \ldots, \tag{3.18}$$

where $\Delta W_k^{:,d} \in \mathbb{R}^N$ is a vector of independent Gaussian random variables. Here, the matrix $\hat{L}(\widehat{Z}_{k\Delta}) \in \mathbb{R}^{N \times N}$ is given as in (2.26).

Analogously to the deterministic case, we now analyze the dynamics of the Euler–Maruyama scheme on the orthogonal complement of $\mathcal{G}^{\star,d}$, that is

$$\widehat{E}_{k\Delta}^{:,d} = \left(I - \frac{1}{N} \mathbf{1}_N \mathbf{1}_N^\top \right) \widehat{Z}_{k\Delta}^{:,d} = P \widehat{Z}_{k\Delta}^{:,d}.$$

We obtain the dynamics

$$\begin{aligned}
\widehat{E}_{(k+1)\Delta}^{:,d} &= P \widehat{Z}_{(k+1)\Delta}^{:,d} \\
&= P \left(\widehat{Z}_{k\Delta}^{:,d} - \lambda \hat{L}(\widehat{Z}_{k\Delta}) \widehat{Z}_{k\Delta}^{:,d} \Delta + \sigma \hat{L}(\widehat{Z}_{k\Delta}) \widehat{Z}_{k\Delta}^{:,d} \circ \Delta W_k^{:,d} \right) \\
&= \widehat{E}_{k\Delta}^{:,d} - \lambda P \hat{L}(\widehat{Z}_{k\Delta}) \widehat{Z}_{k\Delta}^{:,d} \Delta + \sigma P \hat{L}(\widehat{Z}_{k\Delta}) \widehat{Z}_{k\Delta}^{:,d} \circ \Delta W_k^{:,d} \\
&= \widehat{E}_{k\Delta}^{:,d} - \lambda P \widehat{Z}_{k\Delta}^{:,d} \Delta + \sigma P \widehat{Z}_{k\Delta}^{:,d} \circ \Delta W_k^{:,d},
\end{aligned}$$

where we have again applied the relation $P\hat{L}(\widehat{Z}_{k\Delta}) = P$.

We now analyze the mean square stability of the Euler-Maruyama discretization of the orthogonal dynamics. From the recursion

$$\widehat{E}_{(k+1)\Delta}^{:,d} = \left(\mathbf{1}_N - \lambda \Delta \mathbf{1}_N + \sigma \Delta W_k^{:,d} \right) \circ \widehat{E}_{k\Delta}^{:,d}, \tag{3.19}$$

we compute the second moment. Since the Brownian increments $\Delta W_k^{:,d} \sim \mathcal{N}(\mathbf{0}_N, \Delta I)$ are independent of $\widehat{E}_{k\Delta}^{:,d}$, we obtain

$$
\begin{aligned}
\mathbb{E}\left(\|\widehat{E}_{(k+1)\Delta}^{:,d}\|^2\right) &= \mathbb{E}\left(\left\|\left(\mathbf{1}_N - \lambda\Delta\mathbf{1}_N + \sigma\Delta W_k^{:,d}\right) \circ \widehat{E}_{k\Delta}^{:,d}\right\|^2\right) \\
&= \sum_{n=1}^{N} \mathbb{E}\left((1 - \lambda\Delta + \sigma\Delta W_k^{n,d})^2\right) \mathbb{E}\left((\widehat{E}_{k\Delta}^{n,d})^2\right) \\
&= ((1 - \lambda\Delta)^2 + \sigma^2\Delta) \, \mathbb{E}\left(\|\widehat{E}_{k\Delta}^{:,d}\|^2\right),
\end{aligned}
$$

where we used that $\mathbb{E}(\Delta W_k^{n.d}) = 0$ and $\mathbb{E}((\Delta W_k^{n.d})^2) = \Delta$.

By iteration, we get

$$
\mathbb{E}\left(\|\widehat{E}_{k\Delta}^{:,d}\|^2\right) = \left((1 - \lambda\Delta)^2 + \sigma^2\Delta\right)^k \mathbb{E}\left(\|E_0^{:,d}\|^2\right). \tag{3.20}
$$

Since for $x \in [0, 1)$ it holds that $\ln(1 - x) \leq -x$, we estimate

$$
\ln\left((1 - \lambda\Delta)^2 + \sigma^2\Delta\right) = \ln\left(1 - (2\lambda\Delta - \lambda^2\Delta^2 - \sigma^2\Delta)\right) \leq -(2\lambda\Delta - \lambda^2\Delta^2 - \sigma^2\Delta).
$$

Under the condition $2\lambda - \sigma^2 - \lambda^2\Delta \in [0, 1)$, this implies the exponential decay bound

$$
\left((1 - \lambda\Delta)^2 + \sigma^2\Delta\right)^k \leq \exp\left(-k\Delta\left(2\lambda - \lambda^2\Delta - \sigma^2\right)\right).
$$

Hence, the Euler–Maruyama approximation $\widehat{E}_{k\Delta}^{:,d}$ preserves mean square global exponential stablity of the equilibrium manifold with rate

$$
\delta_{\mathrm{EM}} := 2\lambda - \sigma^2 - \lambda^2\Delta > 0, \tag{3.21}
$$

provided the step size Δ is sufficiently small. More precisely,

$$
\mathbb{E}\left(\|\widehat{E}_{k\Delta}^{:,d}\|^2\right) \leq \exp\left(-k\Delta\left(2\lambda - \lambda^2\Delta - \sigma^2\right)\right) \mathbb{E}\left(\|E_0^{:,d}\|^2\right). \tag{3.22}
$$

To establish almost sure exponential convergence of the Euler–Maruyama approximation $\widehat{E}_{k\Delta}^{:,d}$, we first consider the evolution of its components. For each n, the recursion (3.19) reads

$$
\widehat{E}_{(k+1)\Delta}^{n,d} = \left(1 - \lambda\Delta + \sigma\Delta W_k^{n,d}\right) \widehat{E}_{k\Delta}^{n,d}, \tag{3.23}
$$

which can be iteratively expanded to yield the product representation

$$
\widehat{E}_{k\Delta}^{n,d} = \widehat{E}_0^{n,d} \prod_{j=0}^{k-1} \left(1 - \lambda\Delta + \sigma\Delta W_j^{n,d}\right).
$$

Taking logarithms and dividing by $k\Delta$, we obtain

$$
\frac{1}{k\Delta} \ln\left(|\widehat{E}_{k\Delta}^{n,d}|\right) = \frac{1}{k\Delta} \ln\left(|\widehat{E}_0^{n,d}|\right) + \frac{1}{k\Delta} \sum_{j=0}^{k-1} \ln\left(|1 - \lambda\Delta + \sigma\Delta W_j^{n,d}|\right).
$$

Since the increments $\Delta W_j^{n,d} \sim \mathcal{N}(0, \Delta)$ are i.i.d., the strong law of large numbers implies

$$
\lim_{k \to \infty} \frac{1}{k\Delta} \ln\left(|\widehat{E}_{k\Delta}^{n,d}|\right) = \frac{1}{\Delta} \cdot \mathbb{E}\left(\ln\left(|1 - \lambda\Delta + \sigma\Delta W|\right)\right) \quad \text{almost surely,}
$$

where $W \sim \mathcal{N}(0, \Delta)$.

As this asymptotic rate is identical for all components and independent of n, we may pass from the componentwise behavior to the full vector norm. Using norm equivalence in finite dimensions and an argument analogous to the transition from (3.11) to (3.12), we conclude

$$
\lim_{k \to \infty} \frac{1}{k\Delta} \ln\left(\|\widehat{E}_{k\Delta}^{:,d}\|\right) = \frac{1}{\Delta} \cdot \mathbb{E}\left(\ln\left(|1 - \lambda\Delta + \sigma\Delta W|\right)\right) \quad \text{almost surely.} \tag{3.24}
$$

Consequently, the Euler–Maruyama approximation $\widehat{E}_{k\Delta}^{:;d}$ converges almost surely exponentially to the equilibrium manifold if and only if the expectation on the right-hand side is negative. In that case, the almost sure exponential decay rate is given by

$$\delta_{\mathrm{EM}} = -\frac{1}{\Delta} \cdot \mathbb{E}\left(\ln\left(|1 - \lambda\Delta + \sigma\Delta W|\right)\right). \tag{3.25}$$

In the proof of Theorem 4.3, [?] establishes the bound

$$\mathbb{E}\left[\ln\left(|1 - \lambda\Delta + \sigma\Delta W|\right)\right] \leq -\left(\lambda + \frac{\sigma^2}{2}\right)\Delta + C\Delta^2, \tag{3.26}$$

where $C = C(\lambda, \sigma) > 0$ is a constant depending on the drift and diffusion coefficients. Inserting this into the almost sure asymptotic expression derived above, we obtain the estimate

$$\lim_{k\to\infty} \frac{1}{k\Delta} \ln\left(\left\|\widehat{E}_{k\Delta}^{:;d}\right\|\right) \leq -\left(\lambda + \frac{\sigma^2}{2} - C\Delta\right) \quad \text{almost surely.} \tag{3.27}$$

Thus, for sufficiently small Δ, the Euler–Maruyama scheme exhibits almost sure exponential convergence to the equilibrium manifold with rate

$$\delta_{\mathrm{EM}} = \lambda + \frac{\sigma^2}{2} - C\Delta, \tag{3.28}$$

provided that the condition $\lambda + \frac{\sigma^2}{2} > C\Delta$ is satisfied.

4 Numerical Results

Acknowledgments

References

[1] BLUM, Christian ; ROLI, Andrea: Metaheuristics in Combinatorial Optimization: Overview and Conceptual Comparison. In: *ACM Comput. Surv.* 35 (2001), 01, S. 268–308. http://dx.doi.org/10.1145/937503.937505. – DOI 10.1145/937503.937505

[2] RASTRIGIN, L. A.: The Convergence of the Random Search Method in the Extremal Control of a Many-Parameter System. In: *Automation and Remote Control* 24 (1963), S. 1337–1342

[3] KENNEDY, J. ; EBERHART, R.: Particle swarm optimization. In: *Proceedings of ICNN'95 - International Conference on Neural Networks* Bd. 4, 1995, S. 1942–1948 vol.4

[4] BAYRAKTAR, Zikri ; KOMURCU, Muge ; WERNER, Douglas H.: The Wind Driven Optimization Technique and Its Application in Electromagnetics. In: *IEEE Transactions on Antennas and Propagation* 61 (2013), Nr. 5, S. 2745–2757. http://dx.doi.org/10.1109/TAP.2013.2238654. – DOI 10.1109/TAP.2013.2238654

[5] DORIGO, Marco ; BLUM, Christian: Ant colony optimization theory: A survey. In: *Theoretical Computer Science* 344 (2005), Nr. 2, S. 243–278. http://dx.doi.org/https://doi.org/10.1016/j.tcs.2005.05.020. – DOI https://doi.org/10.1016/j.tcs.2005.05.020. – ISSN 0304–3975

[6] SOLIS, Francisco J. ; WETS, Roger J.-B.: Minimization by Random Search Techniques. In: *Mathematics of Operations Research* 6 (1981), Nr. 1, S. 19–30. http://dx.doi.org/10.1287/moor.6.1.19. – DOI 10.1287/moor.6.1.19

[7] PINNAU, René ; TOTZECK, Claudia ; TSE, Oliver ; MARTIN, Stephan: A consensus-based model for global optimization and its mean-field limit. In: *Mathematical Models and Methods in Applied Sciences* 27 (2017), Januar, Nr. 01, S. 183–204. http://dx.doi.org/10.1142/s0218202517400061. – DOI 10.1142/s0218202517400061. – ISSN 1793–6314

[8] Carrillo, José A. ; Choi, Young-Pil ; Totzeck, Claudia ; Tse, Oliver: *An analytical framework for a consensus-based global optimization method.* https://arxiv.org/abs/1602.00220. Version: 2018

[9] Dembo, A. ; Zeitouni, O.: *Large deviations techniques and applications.* Bd. 38. Springer Science & Business Media, 2009

[10] Fornasier, Massimo ; Klock, Timo ; Riedl, Konstantin: *Consensus-Based Optimization Methods Converge Globally.* 2024

[11] Ha, Seung-Yeal ; Jin, Shi ; Kim, Doheon: *Convergence of a first-order consensus-based global optimization algorithm.* https://arxiv.org/abs/1910.08239. Version: 2019

[12] Ha, Seug-Yeal ; Jin, Shi ; Kim, Doheon: *Convergence and error estimates for time-discrete consensus-based optimization algorithms.* https://arxiv.org/abs/2003.05086. Version: 2020

[13] Weber, Dylan ; Theisen, Ryan ; Motsch, Sébastien: Deterministic Versus Stochastic Consensus Dynamics on Graphs. In: *Journal of Statistical Physics* 176 (2019), Nr. 1, S. 40–68. http://dx.doi.org/10.1007/s10955-019-02293-5. – DOI 10.1007/s10955–019–02293–5

[14] Panteley, Elena ; Loría, Antonio ; Sukumar, Srikant: Strict Lyapunov functions for consensus under directed connected graphs. In: *2020 European Control Conference (ECC)*, 2020, S. 1–6

[15] Zhang, Hongwei ; Li, Zhongkui ; Qu, Zhihua ; Lewis, Frank L.: On constructing Lyapunov functions for multi-agent systems. In: *Automatica* 58 (2015), S. 39–42. http://dx.doi.org/https://doi.org/10.1016/j.automatica.2015.05.006. – DOI https://doi.org/10.1016/j.automatica.2015.05.006. – ISSN 0005–1098

[16] In: Fornasier, Massimo ; Klock, Timo ; Riedl, Konstantin: *Convergence of Anisotropic Consensus-Based Optimization in Mean-Field Law.* Springer International Publishing, 2022. – ISBN 9783031024627, S. 738–754

[17] Slotine, Jean-Jacques E. ; Li, Weiping: *Applied Nonlinear Control.* Prentice Hall, 1991. – ISBN 978–0130408907

[18] Meyer, Carl D.: *Matrix Analysis and Applied Linear Algebra.* Philadelphia, PA : Society for Industrial and Applied Mathematics (SIAM), 2000. – ISBN 0–89871–454–0

[19] Horn, Roger A. ; Johnson, Charles R.: *Matrix Analysis.* 2nd. Cambridge University Press, 2012 https://www.anandinstitute.org/pdf/Roger_A.Horn.%20_Matrix_Analysis_2nd_edition(BookSee.org).pdf. – ISBN 978–0521548236