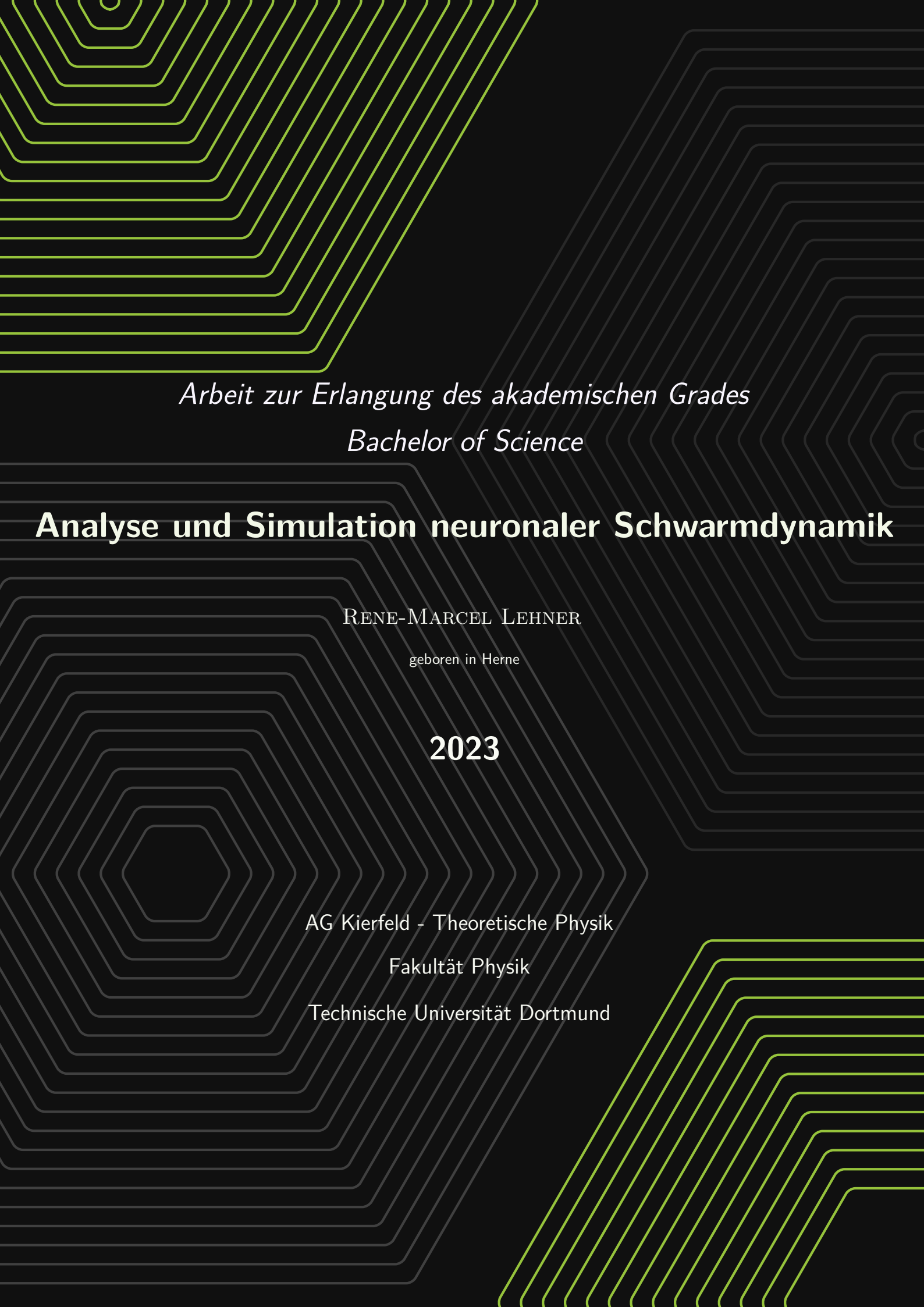




*Arbeit zur Erlangung des akademischen Grades
Bachelor of Science*

Analyse und Simulation neuronaler Schwarmdynamik

RENE-MARCEL LEHNER

geboren in Herne

2023

AG Kierfeld - Theoretische Physik

Fakultät Physik

Technische Universität Dortmund

Erstgutachter: Prof. Dr. Jan Kierfeld
Zweitgutachter: Prof. Dr. Götz Uhrig
Abgabedatum: 19. September 2023



Kurzfassung

Die Schwarmdynamik stellt ein faszinierendes und hochaktuelles interdisziplinäres Forschungsfeld dar, dessen komplexe Verhaltensmuster bisher nur unzureichend durch wissenschaftliche Modelle erklärt werden können. Diese Arbeit widmet sich der Herausforderung, die bestehende Forschungslücke durch die Integration von Maschinellern Lernen in klassische physikalische Modelle zu schließen. Im Fokus stehen dabei zeitaufgelöste topologische Strukturen innerhalb von Schwärmen.

Das etablierte Vicsek-Modell dient als Ausgangspunkt und wird durch ein innovatives Metrisch-Topologisches Modell mit stochastischen Elementen erweitert. Diese Modifikation führt zu einer signifikanten Stabilitätssteigerung des Systems und vereint Vorteile sowohl aus topologischen als auch aus metrischen Ansätzen, was das Modell insbesondere im Vergleich zu herkömmlichen topologischen Modellen bedeutend robuster macht.

Mittels der Implementierung von Multi-Agent Reinforcement Learning (MARL) erfolgt eine fundamentale Neuausrichtung der Methodik: Von lokal ermittelten Verhaltensregeln wird zu global definierten Zielen übergegangen, die eigenständig von den Agenten erlernt werden. Dies ermöglicht die Einführung von ausdrucksstarken Ordnungsparametern als Zielkriterien, deren Maximierung die Agenten autonom erlernen können.

Diese Arbeit demonstriert die Kompatibilität und Trainierbarkeit der vorgestellten Modelle im Kontext von MARL. Als Ergebnis liefert der vorliegende Ansatz einen *Proof of Concept* und motiviert eine weiterführende, tiefgreifendere Untersuchung der angewandten Methoden.



Abstract

Swarm dynamics represents a fascinating and highly topical interdisciplinary field of research, the complex behavioral patterns of which have so far been inadequately explained by scientific models. This study addresses the challenge of bridging this existing research gap by integrating Machine Learning into classical physical models. The focus is on time-resolved topological structures within swarms.

The established Vicsek model serves as the starting point and is augmented by an innovative metric-topological model with stochastic elements. This modification leads to a significant increase in system stability and combines the advantages of both topological and metric approaches, making the model considerably more robust compared to conventional topological models.

Through the implementation of Multi-Agent Reinforcement Learning (MARL), a fundamental reorientation of the methodology is achieved: The focus shifts from locally determined behavioral rules to globally defined objectives that are autonomously learned by the agents. This allows for the introduction of expressive order parameters as target criteria, the maximization of which can be autonomously learned by the agents.

This study demonstrates the compatibility and trainability of the proposed models in the context of MARL. As a result, the present approach provides a *Proof of Concept* and motivates a more in-depth, comprehensive investigation of the methods applied.



Inhaltsverzeichnis

1	Einleitung	1
2	Theoretische Grundlagen	2
2.1	Vicsek-Modell	2
2.2	Metrische und Topologische Modelle	5
2.3	Grundlagen Maschinelles Lernen	5
2.4	MARL und PPO	6
3	Nachbildung bestehender Modelle	12
3.1	Methodologie der Simulation	12
3.2	Vicsek-Modell und topologisches Modell	13
4	Metrisch-Topologischer Ansatz	19
4.1	Einführung des neuen Modells	19
4.2	Messergebnisse	21
4.3	Konvergenz zum Vicsek-Modell	21
4.4	Einführung eines dichte-basierten Ordnungsparameters	22
5	Neuronale Schwarmdynamik mittels Reinforcement Learning	23
5.1	Aufbau des Frameworks und des neuronalen Netzwerks	23
5.2	Training und Validierung	23
6	Diskussion & Ausblick	24
A	Plots	26
B	Illustrationen	27
	Literatur	30

1 Einleitung

Die Schwarmdynamik ist ein interdisziplinäres Forschungsfeld und wird nicht nur in der Physik [1, 2, 3, 4, 5], sondern auch in der Biologie [6, 7, 8], der Informatik [9], der Robotik [10, 11] und den Sozialwissenschaften [12] untersucht. Sie ist ein eindrucksvolles Phänomen, das uns in der Natur regelmäßig begegnet. Ob es sich um die komplexen Flugformationen von Starenschwärmen, die lebhaftige Bewegung von Fischeschwärmen oder die koordinierten Wanderungen von Herdentieren handelt; die Schwarmdynamik dient als beeindruckendes Beispiel für die Emergenz komplexer Strukturen aus einfachen Regeln.

Die Eigenschaft eines Kollektivs, differenziertes Verhalten aus simplen, häufig trivialen Entscheidungen einzelner Individuen zu entwickeln, birgt eine Universalität, die bis heute noch nicht ganz verstanden ist. Die wissenschaftliche und gesellschaftliche Relevanz kollektiven Verhaltens ist weitreichend.

Bislang gibt es in der Wissenschaft kein universelles Modell zur Beschreibung von Schwarmverhalten. Frühere metrische Modelle, insbesondere das von *Vicsek et al.* [1], werden zunehmend von topologischen Ansätzen abgelöst, um reale Szenarien besser zu beschreiben. So hat sich gezeigt, dass topologische Modelle bei der Beschreibung von Vogelschwärmen zuverlässigere Vorhersagen ermöglichen als metrische Modelle [14]. Das von *Vicsek et al.* vorgestellte Modell ist seit 1995 der Grundstein der mathematischen Beschreibung von Schwärmen. Viele fortgeschrittene Modelle, wie etwa das Boids-Modell [15], sind Weiterentwicklungen des Vicsek-Modells und es ist nicht ungewöhnlich, dass neue Modelle und Konzepte das Vicsek-Modell erwähnen oder verwenden. *Tamàs Vicsek* war auch die treibende Kraft hinter der sehr umfangreichen Review "*Collective Motion*" [16] und beschäftigte sich fortwährend mit der Tragweite dieses Themas.



Abbildung 1: Ein Starenschwarm als klassisches Beispiel für Schwarmdynamik [13].

Trotz der frühen Entwicklung dieses Forschungsbereichs bieten die aktuell zur Verfügung stehenden Modelle zur Bildung von Schwärmen nur eingeschränkte Möglichkeiten zur Erklärung der beobachteten Phänomene, insbesondere, wenn es um die Ausbildung von Schwärmen in künstlichen Systemen geht. Diese Arbeit zielt darauf ab, diese Herausforderung anzugehen, indem eine Schnittstelle zwischen traditionellen physikalischen Modellen und modernen Methoden des Maschinellen Lernens geschaffen wird. Dabei liegt der Fokus auf zeitaufgelöste topologische Strukturen.

Ausgangspunkt bildet das eben genannte Vicsek-Modell. Das Modell wird ergänzt durch die Einführung eines neuartigen Metrisch-Topologischen Modells mit stochastischen Elementen, das im Zuge dieser Arbeit speziell dafür entwickelt wurde, um die Integration von Maschinellern Lernen zu erleichtern. Für die hochdynamischen Vielteilchen-Umgebungen werden Algorithmen des Multi-Agent Reinforcement Learning (MARL) verwendet, die auf dem Training neuronaler Netze basieren.

Mit dem Einsatz von Reinforcement Learning wird eine fundamentale Neuausrichtung in der Modellierung der Schwarmdynamik eingeführt. Hier stehen nicht länger die lokal implementierten Regeln im Mittelpunkt, sondern die übergeordneten Ziele des Schwarms. Die Agenten lernen eigenständig die für das Erreichen dieser Ziele erforderlichen Verhaltensmuster. Als Ziel können Ordnungsparameter definiert werden, welche eine mathematische Beschreibung eines Systemzustandes darstellen. Diese Ordnungsparameter sind bedeutend aussagekräftiger und einfacher zu entwickeln als eine lokale Entscheidungsregel.

Diese Arbeit legt zunächst das theoretische Fundament, um die angewandten Konzepte ausreichend zu verstehen und nachvollziehen zu können. Anschließend wird systematisch skizziert, wie die Implementation einer Schwarmsimulation durchgeführt wurde, da diese im Kern dieser Arbeit steht. Es wird das Vicsek-Modell und seine topologische Modifikation nachgebildet und untersucht. Dieses Modell wird durch die bereits erwähnten metrisch-topologischen Eigenschaften erweitert und damit deutlich verbessert. Zum Schluss wird die MARL-Implementation erläutert und diskutiert.

2 Theoretische Grundlagen

Die Schwärmdynamik befasst sich mit zeitabhängigen topologischen Strukturen von Vielteilchensystemen. Sie kann also im Kontext der statistischen Physik betrachtet werden. Ähnlich wie in konventionellen Systemen der kondensierten Materie, in denen Phasenübergänge, kritische Exponenten und universelle Verhaltensmuster eine Rolle spielen, zeigt die Schwärmdynamik bemerkenswerte Analogien zu diesen Phänomenen. Indem Schwärme als stochastische, nichtlineare Systeme vieler interagierender Einheiten betrachtet werden, wird es möglich, Methoden der statistischen Physik anzuwenden, um emergente kollektive Verhaltensweisen zu analysieren. Diese Sichtweise ermöglicht das Einführen von Phasen, die Untersuchungen des Verhaltens an kritischen Punkten und die Definition von kritischen Exponenten sowie die Charakterisierung des Verhaltens in der Nähe von Phasenübergängen. In dieser Arbeit werden diese Aspekte jedoch nur phänomenologisch untersucht.

Eine konkrete Manifestation dieser Konzepte bietet das Vicsek-Modell, das im nächsten Abschnitt detailliert behandeln wird.

2.1 Vicsek-Modell

Eines der einfachsten und grundlegendsten Modelle für die Simulation von Schwärmen ist das Vicsek-Modell [1]. In diesem zweidimensionalen Modell können die Teilchen nur in einem bestimmten Wechselwirkungsradius miteinander interagieren und haben eine konstante Geschwindigkeit v_0 . Der Betrag der Geschwindigkeit wird durch die Wechselwirkungen nicht beeinflusst. Den einzigen Aktionsraum, den die Teilchen zur Verfügung haben, ist die Änderung ihres eigenen Winkels θ_i . Diese Änderung entspricht dem Mittelwert der Richtung aller Teilchen, die sich in dem Wechselwirkungsradius r befinden. Charakterisiert wird das System durch einige weitere Parameter, die in Tabelle 1 aufgelistet sind.

Tabelle 1: Parameter des Vicsek-Modells.

Parameter	Beschreibung
v_0	Konstante Geschwindigkeit aller Teilchen
θ_i	Winkel des i -ten Teilchens
r	Wechselwirkungsradius für die Nachbarsauswahl
L	Länge des quadratischen Simulationsbereichs
N	Anzahl der Teilchen
$\rho = \frac{N}{L^2}$	Teilchendichte
η	Rauschintensität (Noise)
v_a	Ordnungsparameter: quantifiziert den Grad der Ausrichtung
t	Simulationszeit

Der Parameter η repräsentiert hierbei ein stochastisches Rauschen, das als kleine Störung $\Delta\theta_i$ dem Winkel $\theta_i(t+1)$ (2) eines Teilchens hinzugefügt wird. Die Intensität dieses Rauschens beeinflusst den makroskopischen Zustand des Schwarms und kann durch Variation bei sonst gleichbleibenden Parametern zu unterschiedlichen Phasen des Systems führen. In weiterführenden topologischen Modellen kann unter bestimmten Bedingungen zudem die *Koexistenz* mehrerer Phasen gezeigt werden.

In der ursprünglichen Arbeit wird ein quadratischer Simulationsbereich mit Länge L und periodischen Randbedingungen gewählt. Des Weiteren wurden alle Simulationen bei konstanter Geschwindigkeit $v_0 = 0,03 \text{ m/s}$ durchgeführt, was in der Gesamtheit dieser Arbeit übernommen wird.

Charakteristische Größen

Das Modell wird durch den Wechselwirkungsradius $r[\text{m}]$ und der konstanten Geschwindigkeit $v_0[\text{m/s}]$ charakterisiert und darüber entdimensionalisiert. Ein Simulationsschritt entspricht dann der charakteristischen Zeit $\tau = \frac{r}{v_0}$. Daraus ergeben sich die dimensionslosen Größen $\tilde{t} = \frac{t}{\tau}$ für die Zeit, $\tilde{r}_k = \frac{r_k}{r}$ für die kartesischen Ortskoordinaten und $\tilde{v}_k = \frac{v_k}{v_0}$ für die Geschwindigkeiten (mit $k \in x, y$). Hierdurch sind die Vektoren $\tilde{\mathbf{x}}_i$ und $\tilde{\mathbf{v}}_i$ (bezogen auf das i -te Teilchen) ebenfalls einheitenlos. Für einen einzelnen Simulationsschritt gilt dann $\Delta\tilde{t} = 1$.

Mathematische Beschreibung des Vicsek-Modells

In der Schwarmdynamik, insbesondere im Kontext des Vicsek-Modells, ist die Frage der Energieerhaltung oft weniger zentral als in der klassischen Mechanik. Die Modelle sind in der Regel eher phänomenologisch und zielen darauf ab, bestimmte Aspekte kollektiven Verhaltens zu reproduzieren, anstatt von fundamentalen physikalischen Prinzipien auszugehen. Statt eines Verlet-Algorithmus oder anderen symplektischen Integratoren, die den Hamiltonschen Fluss erhalten, wird hier ein einfaches Euler-Verfahren für die Aktualisierung der Teilchenpositionen verwendet:

$$\tilde{\mathbf{x}}_i(\tilde{t}+1) = \tilde{\mathbf{x}}_i(\tilde{t}) + \tilde{\mathbf{v}}_i(\tilde{t})\Delta\tilde{t} \quad (1)$$

Hierbei wird die Geschwindigkeit $\mathbf{v}_i(\tilde{t}+1)$ derart konstruiert, dass ihr Betrag die charakteristische Geschwindigkeit erhält und die Winkeländerung $\theta(\tilde{t}+1)$ erfährt. Der aktualisierte Winkel eines Teilchens setzt sich dann aus dem Mittel der Winkel seiner Nachbarn und aus einem stochastischen Element zusammen:

$$\theta(\tilde{t}+1) = \langle \theta(\tilde{t}) \rangle_r + \Delta\theta \quad (2)$$

Der Ausdruck $\langle \theta(\tilde{t}) \rangle_r$ entspricht einer Mittelung über die Winkel aller Teilchen im Wechselwirkungsradius r ; einschließlich des Teilchens selbst. Der durchschnittliche Winkel wird über

$$\langle \theta(\tilde{t}) \rangle_r = \arctan \left[\frac{\langle \sin(\theta(\tilde{t})) \rangle_r}{\langle \cos(\theta(\tilde{t})) \rangle_r} \right] \quad (3)$$

berechnet. Die Zufallszahl $\Delta\theta$ wird uniform aus dem Intervall $[-\eta/2, \eta/2]$ generiert. Um die Gleichverteilung dieser Störung nicht zu verletzen, sollte η in $[0, 2\pi)$ liegen, da es ansonsten Winkelbereiche gibt, die stärker gewichtet werden.

Als zentraler Kennwert wird der **Ordnungsparameter** v_a eingeführt:

$$v_a = \frac{1}{Nv_0} \left| \sum_{i=1}^N \mathbf{v}_i \right| \quad (4)$$

Dieser normierte, skalare Ausdruck quantifiziert den Grad der Ausrichtung oder Kohäsion innerhalb des Schwarms.

Ein System, in dem alle Teilchen in dieselbe Richtung zeigen, hat $v_a = 1$, während ein vollständig *ungeordnetes* System $v_a \approx 0$ aufweist; es gilt also $v_a \in [0, 1]$.

Dieser Parameter enthält keine Information über die Dichte, Kohäsion oder topologische Cluster-Bildung eines Schwarms. Es zeigt sich jedoch, dass sowohl die einfache Mittelungs-Regel des Vicsek-Modells als auch der Ordnungsparameter v_a ausreichen, um Schwärme zu bilden und quantitativ zu beschreiben. Ferner kann analytisch gezeigt werden, dass die Vicsek-Regel ebendiesen Ordnungsparameter v_a minimiert (siehe Abschnitt 3.2).

Das Vicsek-Modell hat drei freie Parameter: Die Teilchenzahl N , die Rauschintensität η und die Länge des Dimensionsbereiches L . Bewertet wird das Modell nach dem Ordnungsparameter. Daraus ergeben sich viele mögliche Kombinationen, die untersucht werden können. In der ursprünglichen Arbeit [1] und in vielen darauf aufbauenden Arbeiten [3], [5] werden jedoch nur die Rauschintensität η und die Dichte ρ gegen den Ordnungsparameter aufgetragen und untersucht.

Der Fokus auf die Dichte ρ ist insofern sinnvoll, als sie direkten Aufschluss über die mittlere Anzahl der Wechselwirkungspartner gibt. Formell gilt

$$n_r = \rho\pi r^2 = N \frac{\pi r^2}{L^2} \quad (5)$$

Hierbei ist n_r dimensionslos und gibt an, wie viele benachbarte Teilchen sich im Mittel im Wechselwirkungsradius r des betrachteten Teilchens aufhalten. Diese dimensionslose Größe erlaubt eine effizientere numerische Analyse, fördert die Vergleichbarkeit mit anderen Arbeiten und ist physikalisch relevanter als die einzelne Variation von N und L .

Da das System entdimensionalisiert ist, gilt sogar direkt $\tilde{n}_r = \pi\rho$. Aus der Dichte ist diese Größe also um einen Faktor von π direkt ablesbar. Dies erweitert die Interpretierbarkeit vieler Plots.

Es sei allerdings erwähnt, dass dies nur einer Näherung für equilibrierte Systeme entspricht. Dort können die Teilchen als gleichverteilt und die Dichte als annähernd homogen betrachtet werden, sodass die einzelnen Beiträge einer ortsbasierten Dichte $\rho(r)$ durch ihre statistische Unabhängigkeit faktorisiert werden können. Daher gilt dieser Ausdruck nur in 0. Ordnung. Für eine rigorose Auseinandersetzung mit der Wechselwirkungspartner-Anzahl werden fortgeschrittenere Ansätze benötigt. An dieser Stelle wird ein Dichte-basierter Ordnungsparameter motiviert, der in Abschnitt 4.4 eingeführt wird.

Visuelle Repräsentation des Vicsek-Modells

Angelehnt an die originale Arbeit werden in Abbildung 2 einige Schnappschüsse einer Simulation gezeigt. Zusammengehörige Parametersätze sind farblich hervorgehoben.

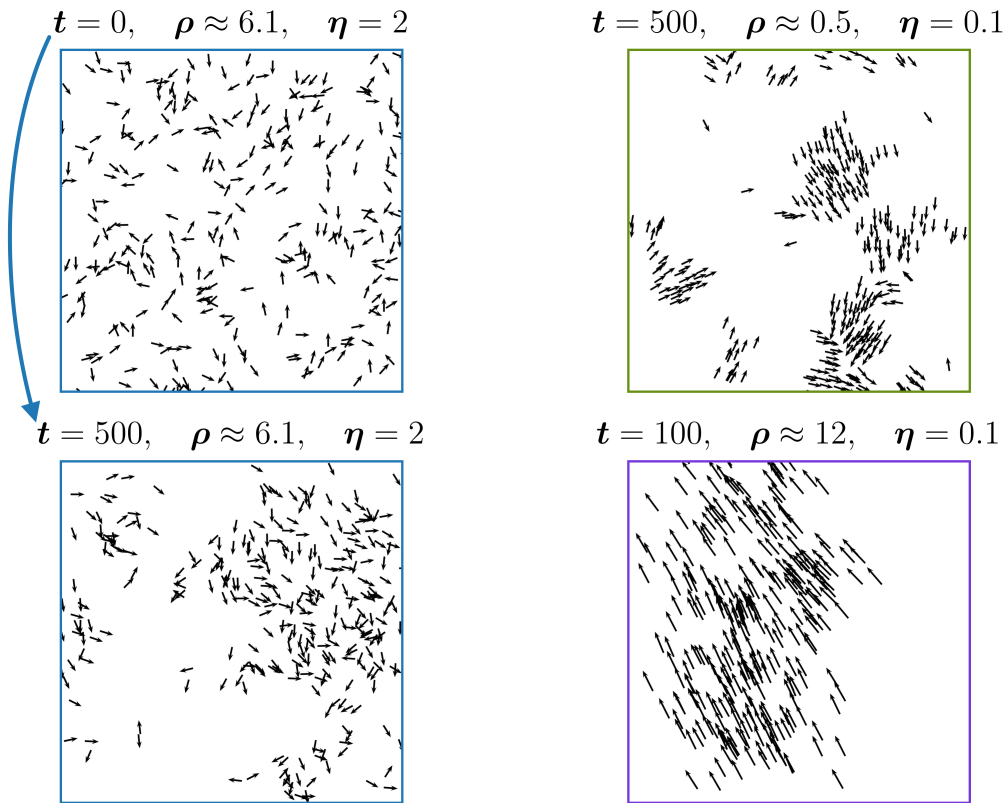


Abbildung 2: Vier Szenarien des Vicsek-Modells mit $N = 300$ Teilchen. Die beiden blauen Plots zeigen zwei Zeitpunkte $t = 0$ und $t = 500$ bei $L = 7$. Der grüne Plot stellt $t = 500$ bei $L = 25$ und niedriger Dichte $\rho \approx 0.5$ dar. Der Plot im lilanen Rahmen zeigt $t = 100$ bei $L = 5$ und hoher Dichte $\rho \approx 12$. Die Schnappschüsse stammen aus der eigenen Simulation und wurden mit Matplotlib [17] visualisiert.

2.2 Metrische und Topologische Modelle

Die meisten grundlegenden Modelle zur Schwarmdynamik können in metrische und topologische Modelle unterteilt werden. Metrische Modelle bestimmen die Nachbarn eines Teilchens auf Basis eines Abstandes, während topologische Modelle eine *feste Anzahl* von *nächsten* Nachbarn wählen; unabhängig von der Distanz. In Abbildung 3 sind diese beiden Auswahlregeln gegenübergestellt.

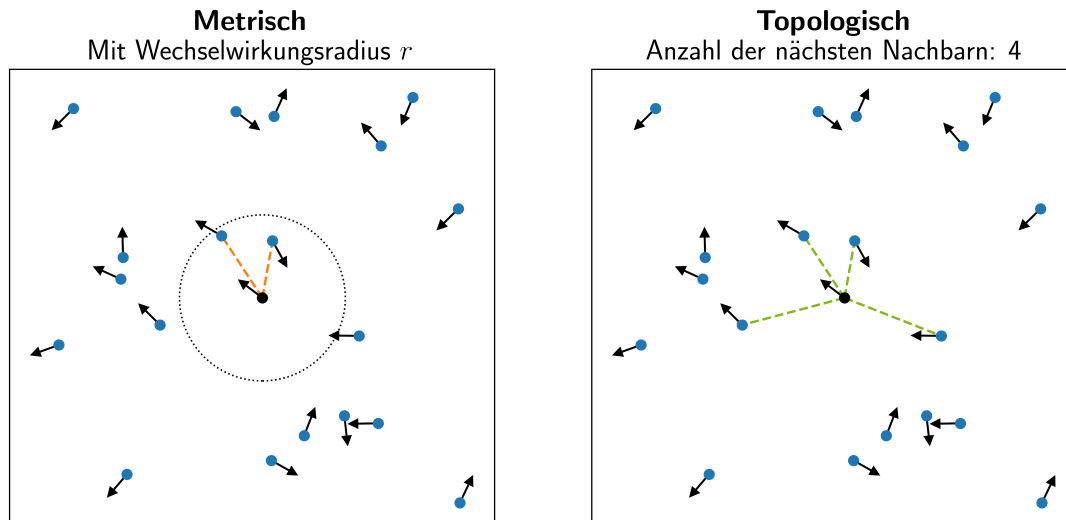


Abbildung 3: Vergleich zwischen metrischer und topologischer Nachbarschaftsauswahl in einem Schwarm von 20 Teilchen. Beispielhaftes Ermitteln der Nachbarn für ein Teilchen (zentriert). Die gestrichelten Verbindungslinien stellen die ermittelten Nachbarn dar.

Links: In der metrischen Auswahl werden Nachbarn in einem festen Wechselwirkungsradius r (gestrichelter Kreis) bestimmt.

Rechts: In der topologischen Auswahl werden *immer* k Nachbarn gefunden (falls $N > k$), unabhängig von ihrer räumlichen Entfernung.

Hier beispielhaft: $k = 4$.

Die Teilchenverteilung und -Ausrichtung sind in beiden Plots identisch.

Mit dem Vicsek-Modell wurde bereits ein metrisches Modell vorgestellt. Weitere metrische Modelle sind bspw. das Boids-Modell [15] oder das Cucker-Smale-Modell [2]. Beide Modelle können jedoch als Weiterentwicklung des Vicsek-Modells verstanden werden.

Für diese Arbeit sind vor allem topologischen Modelle interessant, da diese eine feste Anzahl an Nachbarn vorgeben. Dies vereinfacht eine Implementierung der meisten Algorithmen des Maschinellen Lernens, da neuronale Netze eine feste Länge des Eingabevektors voraussetzen (Ausnahme sind die fortgeschrittenen Recurrent Neural Networks (RNNs) [18] und Transformers [19]). In diesem Kontext ist die alleinige Modifikation der Nachbarschaftsregel von metrisch zu topologisch, bei sonst unveränderten Simulationsparametern, besonders relevant. Der Wechselwirkungsradius r verliert dann seine Bedeutung, kann aber als Vorbereitung auf ein weiterführendes metrisch-topologisches Modell behalten werden. Dies erleichtert einen direkten Vergleich.

In einer vorangegangenen Studie wurde die topologische Modifikation des Vicsek-Modells bereits untersucht [3]. Dort wurden sowohl Gemeinsamkeiten als auch Unterschiede der beiden Modelle beobachtet.

Auf die Details und einen direkten Vergleich der Modelle wird in der Nachbildung der Modelle in Abschnitt 3 eingegangen.

2.3 Grundlagen Maschinelles Lernen

In der Einleitung wird die Verwendung Maschinellen Lernens bereits motiviert. Die Herausforderung besteht nun darin, einen Algorithmus und ein Framework zu wählen, die für das Problem geeignet sind. In Anbetracht der Vielfalt der Ansätze, die alle ihre eigenen Vor- und Nachteile mitbringen, und angesichts der rapiden Entwicklung in diesem Forschungsbereich, wird zunächst eine Klassifizierung des Problems vorgenommen und

die zurzeit verfügbaren Methoden werden kurz skizziert und bewertet.

Klassifizierung des Problems

In konventionellen Schwarm-Simulationen wird über jedes Teilchen iteriert und anhand der Beobachtung des Teilchens über eine feste Regel eine Aktion generiert, die bspw. in den einfachen Modellen zu einer Winkeländerung des Teilchens führen. Da die Qualität der Schwarm-Bildung über einen einzigen, skalaren Parameter bewertet wird, liegt kein Klassifizierungs- oder Regressions-Problem im klassischen Sinne vor. Vielmehr soll ein Teilchen ein Verhalten bzw. eine *Strategie* entwickeln, die diesen einfachen Parameter maximiert; und im entsprechenden Sinne Aktionen bestraft, die zur Verringerung dieses Parameters führen.

Diese Begriffe und Konzepte sind eng verwandt mit dem **Reinforcement Learning**; ein Bereich des Maschinellen Lernens, der weder dem überwachten noch unüberwachten Lernen zuzuordnen ist.

Reinforcement Learning

Das Reinforcement Learning (RL) stellt eine Reihe von Methoden dar, bei der ein Agent selbstständig lernt, in einer Umgebung unter verschiedenen Beobachtungen eine vorgegebene Belohnung zu maximieren. In diskreten und einfacheren RL-Algorithmen wird oft mit einer Tabellenrepräsentation des Zustands- und Aktionsraums gearbeitet, welche auf Grundlage neuer Beobachtungen und Belohnungen aktualisiert werden. Diese Aktualisierung stellt dann das Training dar. Fortgeschrittene Methoden benötigen komplexere Strukturen als Tabellen; insbesondere für das Training. Das in Abbildung 4 skizzierte Konzept bleibt allen Algorithmen jedoch gleich.

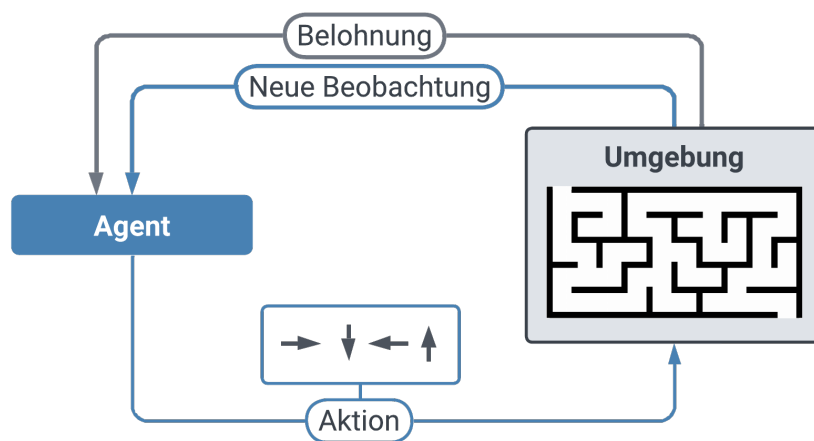


Abbildung 4: Grundlegendes Konzept des Reinforcement Learning: Ein Agent trifft in einer gegebenen Umgebung eine Aktion, woraufhin er eine Belohnung sowie eine neue Beobachtung der Umgebung erhält. Diese Interaktion ermöglicht es dem Agenten, seine Entscheidungsstrategie iterativ zu verbessern, um kumulative Belohnungen zu maximieren.

Im Reinforcement Learning werden einige Begriffe eingeführt, die in der Literatur einheitlich verwendet werden. Dies vereinfacht die Übersetzung eines vorliegenden Problems in eine Form, die alle Algorithmen verstehen. Fortgeschrittene Methoden führen weitere Konzepte und Begriffe ein, die aber klar abgegrenzt werden können. Eine Übersicht der grundlegenden Begriffe ist in Tabelle 2 zu finden.

Literatur zu diesem Teil: Bonaccorso, *Mastering Machine Learning Algorithms* [20].

2.4 MARL und PPO

Das Multi-Agent Reinforcement Learning (MARL) ist ein sehr junges und sich schnell entwickelndes Forschungsfeld, das viele bisher ungesehene algorithmische Herausforderungen und Anwendungsfälle bietet [21, 22]. Vorreiter in der Entwicklung und Untersuchung neuer MARL Algorithmen ist OpenAI. Im vorangegangenen Abschnitt wurden die Grundlagen des Reinforcement Learnings erläutert. Der folgende Abschnitt vertieft das Verständnis dieser Konzepte durch die Einführung von MARL und Proximal Policy Optimization (PPO). Während MARL eine Erweiterung des Reinforcement Learnings auf mehrere Agenten darstellt, ist PPO ein

Tabelle 2: Grundbegriffe des Reinforcement Learnings.

Begriff	Beschreibung
Agent	Autonomer Entscheidungsträger, der auf Grundlage einer Beobachtung Aktionen in einer Umgebung ausführt, um eine Belohnung zu erhalten.
Umgebung (<i>environment</i>)	Der Kontext oder Raum, in dem der Agent operiert und in dem Zustände, Aktionen und Belohnungen definiert sind.
Zustand (<i>state</i>) $s \in \mathcal{S}$	Spezifische Konfiguration (Beobachtung) der Umgebung, die alle relevanten Informationen enthält, die der Agent zum Treffen einer Entscheidung benötigt.
Aktion (<i>action</i>) $a \in \mathcal{A}(s)$	Eine Operation oder ein Vorgang, den der Agent in einem bestimmten Zustand (s) ausführt, um seinen Zustand in der Umgebung zu ändern.
Belohnung (<i>reward</i>) $r : \mathcal{S} \times \mathcal{A} \rightarrow \mathbb{R}$	Ein numerischer Wert, der dem Agenten als Feedback für seine Aktion in einem bestimmten Zustand zurückgegeben wird. Ziel ist es, die kumulierte Belohnung zu maximieren.
Strategie (<i>policy</i>) $\pi : \mathcal{S} \rightarrow \mathcal{A}$	Ein Regelwerk oder Algorithmus, den der Agent verwendet, um in einem bestimmten Zustand eine Aktion auszuwählen.
Value-Funktion V	Eine Funktion, die den erwarteten kumulativen zukünftigen Belohnungswert für einen gegebenen Zustand (s) schätzt. Sie dient als Grundlage für die Entscheidungsfindung des Agenten und wird in vielen RL-Methoden aktualisiert. Typische Darstellung: $V^\pi(s) = \mathbb{E}[R_t s_t = s, \pi]$
Episode	Sequenz von (Zeit-)Schritten, die der Agent durchläuft, um Aktionen auszuführen und Belohnungen zu erhalten. Kann endlos sein; manche Algorithmen benötigen jedoch ein explizites Ende (bspw. ein Spiel, das gewonnen oder verloren wird). In dieser Arbeit kann das Erreichen eines Gleichgewichtszustands als implizites Ende fungieren.

konkreter Algorithmus zur Optimierung einer Strategie. Beide Konzepte werden in dieser Arbeit verwendet.

In diesem Theorieteil werden bereits methodologische Elemente integriert, da sie in diesem Kontext am sinnvollsten erscheinen. Somit geht es im Folgenden nicht nur um allgemeine Konzepte, sondern auch um deren spezifische Anwendung im Kontext der in dieser Arbeit entwickelten Simulation.

Multi-Agent (MA) Systeme

Im klassischen Reinforcement Learning wird nur ein einzelner Agent betrachtet, der eine Strategie in einer Umgebung lernt und optimiert. Bei Vielteilchensystemen, wie etwa der vorliegenden Schwarmsimulation, ändert sich die Herangehensweise grundlegend. Es ergeben sich vollkommen neue Fragestellungen und Herausforderungen. Eine Umgebung und ein Programm, die für einen einzelnen Agenten geschrieben wurden, können nicht ohne Weiteres auf ein Multi-Agent System erweitert werden. In diesem Theorieteil werden nur die konzeptionellen Fragen und einige grundlegende Aspekte beleuchtet. Die technischen Herausforderungen und die konkrete Methodologie werden im entsprechenden Abschnitt 5 erläutert.

MARL als Erweiterung des Reinforcement Learnings

Bei der Umstellung auf mehrere Agenten ergeben sich einige Veränderungen im algorithmischen Kern.

- **Datengranularität:** Während im klassischen RL Zustände, Aktionen und Belohnungen einzelner Agenten betrachtet werden, müssen diese Elemente im MARL für eine ganze Gruppe von Agenten gemeinsam erfasst werden. Hierbei werden die Daten in der Regel *gebatched*, d.h., sie werden für mehrere Agenten gleichzeitig gesammelt und verarbeitet. Dadurch muss die *Umgebung* fundamental neu implementiert werden und das Handling von Zustands- und Aktions-Vektoren, insbesondere angesichts der Verwendung neuronaler Netze, wird wesentlich empfindlicher.

- Handling der **Value-Funktion**: Diese bewertet nun den Gesamtnutzen mehrerer Agenten und muss oft unter Berücksichtigung der Aktionen und Zustände aller beteiligten Agenten berechnet werden.
- Kompliziertere **Strategieentwicklung**: In MARL wird nicht nur die Komplexität der möglichen Strategien erhöht, sondern auch die Notwendigkeit, diese Strategien effektiv an mehrere Agenten zu verteilen. Das kann bedeuten, dass eine geteilte Strategie an alle Agenten *gebroadcastet* wird, wie es in dieser Arbeit der Fall ist. Dadurch erhöht sich die algorithmische Komplexität des Trainings.

Es gibt weitere Aspekte, die eingeführt und implementiert werden müssen. Dazu zählen bspw. eine mögliche Kommunikation zwischen den Agenten, das Handling von Episoden oder die Erhaltung der Gleichzeitigkeit. Diese Aspekte sind für diese Arbeit von geringerer Bedeutung und werden häufig durch die Standard-Einstellungen der verwendeten Bibliotheken beantwortet. Sie seien dennoch der Vollständigkeit wegen erwähnt.

Weiterführende Literatur zu diesem Thema: Albrecht, Christianos, Schäfer [23].

Abbildung 5 zeigt den konzeptionellen Aufbau eines MARL Projektes.

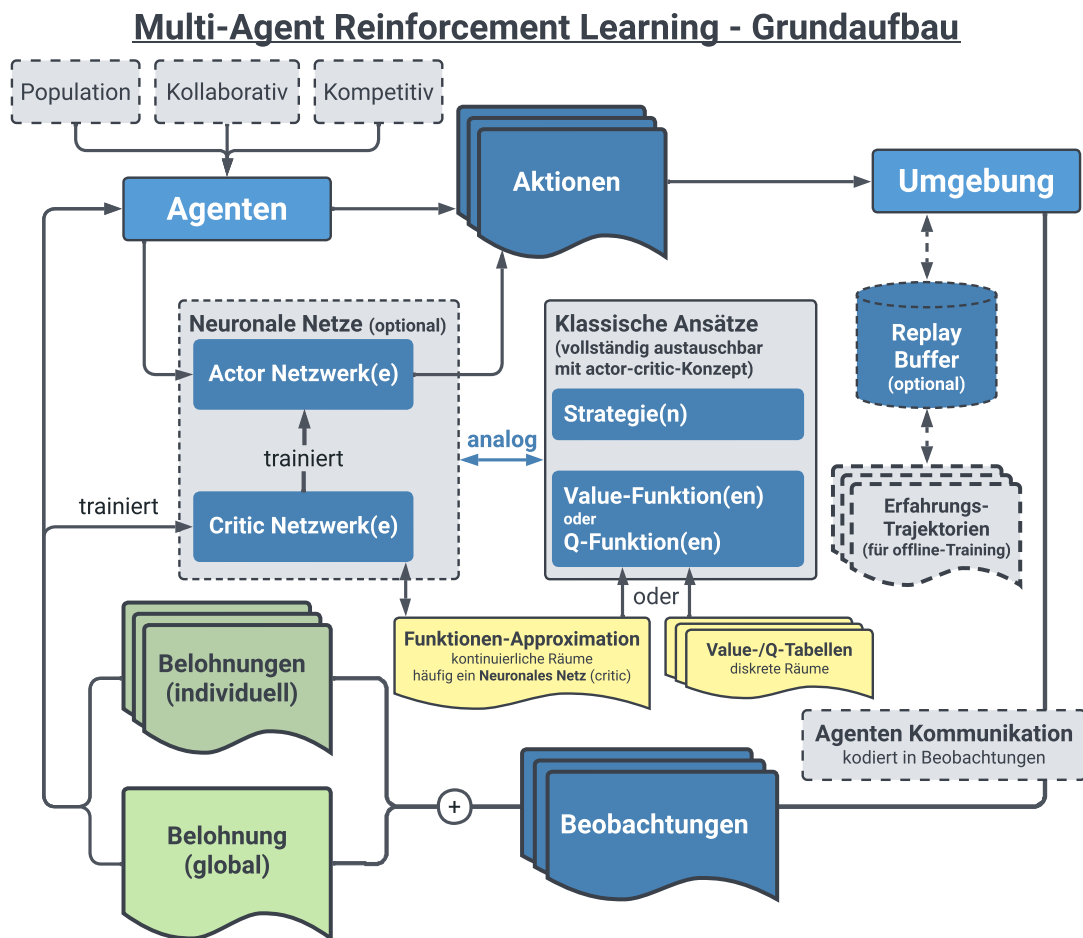


Abbildung 5: Grundlegendes Konzept des Multi-Agent Reinforcement Learnings.

Fragestellungen

Bei der Erweiterung auf ein MA System ergeben sich neue Möglichkeiten und damit auch neue Fragen, die individuell für ein vorliegendes Problem beantwortet werden müssen. Für diese Arbeit relevante Fragen sind bspw.: Gibt es eine zentrale Steuerung oder sind alle Agenten autonom? Gibt es verschiedene Populationen von Agenten (heterogene Parameter-Verteilung)? Gibt es mehrere Strategien, also unterschiedliche Ziele? Was sehen die Agenten; teilen sie Informationen? Wie werden Belohnungen in einem Multi-Agenten-System verteilt oder zugewiesen (lokale vs. globale Belohnungen)?

In dieser Arbeit wird konzeptionell alles so einfach und direkt wie möglich gehalten. Daher können all diese Fragen sofort adressiert werden:

MARL-Umsetzung in dieser Arbeit

Jedes Teilchen soll autonom handeln, aber dieselbe Strategie teilen. Alle Agenten optimieren also gemeinsam dieselbe Strategie (*shared policy*). Es gibt nur eine Population und die Agenten haben nur lokale Informationen. Die Belohnung jedoch entspricht der globalen Änderung des Ordnungsparameters. Die Alternative dazu wäre, dass die Agenten nur den Ordnungsparameter ihrer Gruppe sehen und optimieren. Dann entspricht die Belohnung aber nicht mehr der Maximierung des globalen Ordnungsparameters, wie es das Vicsek-Modell versucht.

Verfügbare Frameworks für (MA)RL

Verschiedene Frameworks bieten bereits implementierte Algorithmen und Strukturen, die die Entwicklungszeit erheblich reduzieren können. Sie implementieren die in diesem Abschnitt beschriebenen algorithmischen Grundlagen und bieten sowohl eine breite Palette an voreingestellten Konfigurationen als auch umfangreiche Optionen zur individuellen Anpassung. Zu den bekanntesten zählen OpenAI's [Gym](#) (in neuen Versionen [Gymnasium](#)) für Reinforcement Learning [24], Ray's [RLlib](#) für skalierbare RL- und MARL-Implementierungen [25] (**wird in dieser Arbeit verwendet**), und TensorFlow Agents ([TF-Agents](#)) [26], eine zuverlässige Bibliothek für komplexe RL-Probleme (kein Multi-Agent).

Nachdem mit MARL eine eher konzeptionelle Erweiterung des Reinforcement Learnings vorgestellt wurde, stellt sich nun die Frage nach einer konkreten Trainings- und Optimierungsstrategie. Einige gängige Methoden werden im nächsten Abschnitt skizziert.

Proximal Policy Optimization (PPO)

Proximal Policy Optimization (PPO) ist einer der führenden Algorithmen im Bereich des Reinforcement Learnings. Obwohl zahlreiche Algorithmen für die Strategieoptimierung existieren, hat sich PPO als besonders effektiv und stabil erwiesen [27]. Die Entscheidung für einen bestimmten Algorithmus hängt von verschiedenen Faktoren ab, insbesondere aber von der Beschaffenheit der Umgebung (kontinuierlicher vs. diskreter Aktions- und Zustandsraum).

Übersicht über Algorithmen im Reinforcement Learning

Für ein besseres Verständnis der Auswahlmöglichkeiten und der Position von PPO im Spektrum der RL-Algorithmen dienen die Abbildungen 6 und 7. Weit verbreitete kontinuierliche Algorithmen sind PPO, SAC (Soft Actor-Critic) [28] und DDPG (Deep Deterministic Policy Gradients) [29], während für den diskreten Raum Algorithmen wie Q-Learning [30] und DQN (Deep Q-Networks) [31] populär sind.

All die aufgelisteten Algorithmen können auf Multi-Agent Systeme erweitert werden.

RL Agenten-Übersicht

	DQN	DDPG	C51	TD3	PPO	REINFORCE	SAC
Diskret	✓		✓		✓	✓	
Kontinuierlich		✓		✓	✓		✓
On-/Off-Policy	Off	Off	Off	Off	On	Off	Off
Model-Free/Model-Based	Free	Free	Free	Free	Free	Free	Free
Online-/Offline-Training	Beides	Beides	Beides	Beides	Beides	Beides	Beides

Abbildung 6: Übersicht der zurzeit meist verwendeten Reinforcement Learning Algorithmen.

Terminologie und Klassifizierung

Für das Verständnis der Algorithmen werden 3 bzw. 6 weitere Begriffe erklärt. Diese können tendenziell der Tabelle 2 zugeordnet werden, sind aber im Kontext konkreter Algorithmen besser zu verstehen.

- *On-Policy vs. Off-Policy*: On-Policy-Algorithmen wie PPO aktualisieren die Strategie basierend auf den Daten der aktuellen Strategie. Off-Policy-Algorithmen wie DQN können Daten aus vorherigen Strategien verwenden.
- *Model-Free vs. Model-Based*: Model-Based-Algorithmen lernen und simulieren die Dynamik der Umgebung zur Optimierung zukünftiger Entscheidungen. Im Gegensatz dazu benötigen Model-Free-Algorithmen wie PPO und DQN kein solches Umgebungsmodell, sondern lernen direkt durch Interaktion.
- *Online vs. Offline Training*: Beim Online-Training werden Daten in Echtzeit gesammelt und verarbeitet. Beim Offline-Training wird ein Datensatz gesammelt und dann zur Optimierung verwendet.
- *Exploration vs. Exploitation*: Exploration-Exploitation-Strategien regeln die Balance zwischen der Erkundung neuer Aktionen und der Ausführung bekannter, belohnungsoptimaler Aktionen. Algorithmen wie PPO können mithilfe von Methoden wie dem Epsilon-Greedy-Ansatz oder der Boltzmann-Exploration modifiziert werden, um eine geeignete Trade-off-Strategie zu ermöglichen.

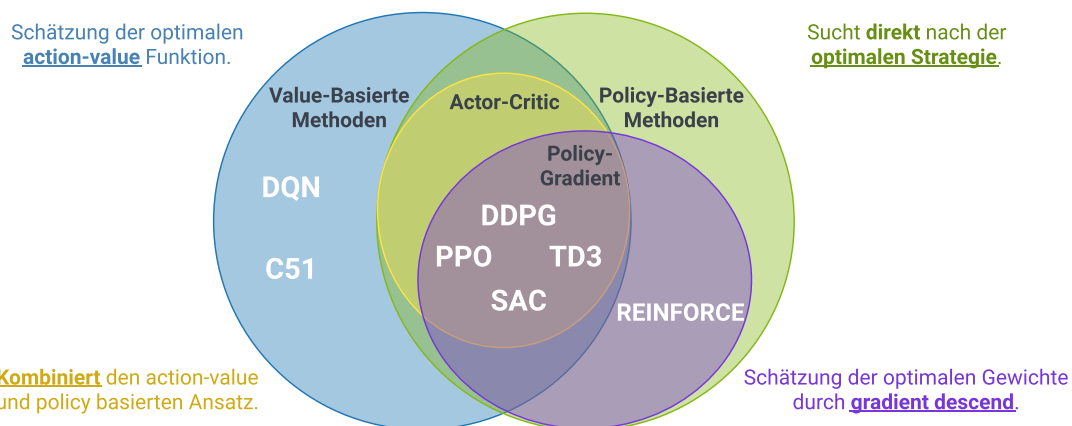


Abbildung 7: Klassifizierung gängiger Reinforcement-Learning-Algorithmen nach ihrer Herangehensweise: Value-Based, Policy-Based und Policy Gradient Methoden, inklusive der speziellen Actor-Critic-Techniken, die Elemente aus verschiedenen Kategorien kombinieren und häufig neuronale Netze verwenden.

Warum PPO?

PPO überzeugt durch seine Robustheit und Effizienz. Als moderner Algorithmus baut er auf bewährten Konzepten auf und hat sich in vielen Bereichen bewiesen. Zudem kann er sowohl mit diskreten, als auch mit kontinuierlichen Aktions- und Zustandsräumen arbeiten. Daher wird dieser Algorithmus häufig als Standard angesehen, insbesondere für MARL-Aufgaben. Durch seine Beliebtheit ist PPO mit vielen Beispielen belegt und gut dokumentiert, was durch die starke Entwicklung in diesem Bereich durchaus relevant ist. Im Kern von PPO steht das Training eines neuronalen Netzes, was diesen Algorithmus für die Aufgabe dieser Arbeit qualifiziert.

Mathematischer Kern von PPO

PPO optimiert eine Wahrscheinlichkeitsverhältnisfunktion:

$$\pi_{\theta}(a_t|s_t)/\pi_{\theta_{\text{alt}}}(a_t|s_t)$$

In diesem Ausdruck bezeichnet $\pi_{\theta}(a_t|s_t)$ die Wahrscheinlichkeit der Aktion a_t im Zustand s_t gemäß der aktuellen Politik mit Parametern θ . $\pi_{\theta_{\text{alt}}}(a_t|s_t)$ ist die Wahrscheinlichkeit derselben Aktion im selben Zustand, jedoch gemäß der vorherigen Politik.

Die optimierte Funktion wird geclippt, um die Aktualisierung der Politik zu stabilisieren:

$$\min \left(r_t(\theta) \hat{A}_t, \text{clip}(r_t(\theta), 1 - \epsilon, 1 + \epsilon) \hat{A}_t \right)$$

Hierbei ist $\hat{A}_t$ der geschätzte Vorteil der Aktion a_t , definiert als:

$$\hat{A}_t = Q(s_t, a_t) - V(s_t)$$

$Q(s_t, a_t)$ ist der erwartete zukünftige Ertrag der Aktion a_t im Zustand s_t , und $V(s_t)$ ist der Wert des Zustands s_t gemäß der aktuellen Politik.

Der Vorteil $\hat{A}_t$ ist ein zentrales Element von Actor-Critic-Ansätzen. Er zeigt an, wie viel besser oder schlechter eine Aktion im Vergleich zu anderen möglichen Aktionen im gleichen Zustand ist. Dadurch wird eine effizientere Optimierung der Politik ermöglicht.



3 Nachbildung bestehender Modelle

Bevor Maschinelles Lernen auf eine Schwarmsimulation angewendet werden kann, muss eine Simulation basierend auf einem Modell implementiert und validiert werden. Dieses Kapitel behandelt die Implementierung der Simulation und untersucht sowohl das Vicsek-Modell als auch andere Modell-Familien. Zur Bewertung der Implementierungsgenauigkeit erfolgt ein Vergleich der Simulationsergebnisse mit Ergebnissen aus vorhandenen Studien. Gleichzeitig werden die hier verwendeten statistischen Methoden auf ihre Robustheit überprüft.

3.1 Methodologie der Simulation

Die Konzeption der Simulation gliedert sich in drei aufeinanderfolgende Phasen: Die Auswahl der Software und Strukturierung der Implementierung; die Anwendung spezifischer Programmierstechniken zur Lösung besonderer Aufgaben; und letztlich die zuvor erwähnte Validierung der Ergebnisse durch Vergleich und statistischen Methoden.

Wahl der Software: C++ und Python

Für diese Arbeit werden die Programmiersprachen C++ [32] und Python¹ [33] verwendet. Diese Wahl liegt in den einzelnen Stärken und der Synergie beider Sprachen begründet.

C++ ist als moderner Standard und vor allem wegen seiner Effizienz besonders geeignet. Eine Vielteilchensimulation kann selbst bei einfachen Wechselwirkungen schnell einen beachtlichen Rechenaufwand fordern. Eine Simulation mit $N \geq 10\,000$ Teilchen, wie Vicsek *et al.* sie untersucht haben, kann dann nur mit ausreichend effizienten Methoden praktisch untersucht werden.

Python ist insofern eine natürliche Ergänzung zu C++, als dass es nicht nur mit seinen eigenen Stärken überzeugt, sondern auch eine effektive Schnittstelle für C++ bietet.

Auswahlkriterien zusammengefasst:

1. **Stärken von Python:** Schnelles Prototyping, einfaches Visualisieren (`Matplotlib`), schnelles Debugging, hohe Funktionalität. Dadurch auch: Visuelles Debugging einfach möglich.
2. `pybind11`[34]: Einbinden von eigenem C++11 Code/Funktionen in Python.
3. **Maschinelles Lernen:** Python ist die bevorzugte Sprache für Maschinelles Lernen, da viele einschlägige Bibliotheken, einschließlich der in dieser Arbeit verwendeten TensorFlow- und RLlib-Bibliothek, speziell und oft exklusiv für Python entwickelt wurden.

Häufig ist es von Interesse, das System zur Überprüfung zu visualisieren. Um vor allem große Simulationen performant visualisieren zu können, kann die Software `Ovito` [35] verwendet werden. Diese Software wird in der Festkörperphysik verwendet und dient hauptsächlich der Visualisierung von atomaren und molekularen Strukturen, sowie der Darstellung ihrer dynamischen Eigenschaften über die Zeit. Sie eignet sich für diesen Fall aber auch für die einfache Darstellung von zeitaufgelösten Teilchenpositionen und besonders für sehr große Teilchenzahlen (bis etwa $N = 100\,000$) in Echtzeit, indem alle Teilchenpositionen für jeden Zeitschritt in eine `.xyz`-Datei geschrieben werden. Die Dateien können je nach Länge der Simulation und Teilchenanzahl entsprechend groß werden, sind aber meistens noch im MB- bis GB-Bereich.

Umsetzung der Simulation

Die in dieser Arbeit entwickelte Simulation kann auf einige wenige Schwerpunkte reduziert werden. Wird von Beginn an eine objektorientierte Methodik angewendet, erleichtert dies die Strukturierung der Implementierung maßgeblich. Ein schematischer Projektaufbau, der die komplementäre Verwendung von Python und C++ zeigt, befindet sich im Anhang (siehe Abbildung 23).

Die relevantesten Algorithmen sind in folgender Liste zusammengefasst:

¹Programmiersprache im allgemeinen Verständnis, Skriptsprache im spezifischeren Kontext.

Die wichtigsten Aufgaben der Simulation

- **Zellenaufteilung und Teilchenzuweisung**
Ziel: Reduktion des Rechenaufwands.
Komplexität: Hoch, da fehleranfällig. Es müssen Randeffekte berücksichtigt werden.
- **Nachbarschaftslisten und -Bestimmung**
Ziel: Hier: Bereitstellung der Nachbarn für die Modell-Regeln. Können auch verwendet werden, um den Rechenaufwand zu reduzieren.
Komplexität: Mittel/Hoch, entsprechend den beiden genannten Zielen. Die Wahl von performanten Methoden (Sortierung, Zufallsgenerator, Distanzbestimmung) hat hier einen großen Einfluss. Durch diese Aufgabe sind topologische Modelle wesentlich rechenintensiver als metrische Modelle.
- **Periodische Randbedingungen**
Ziel: Simulation eines unendlich ausgedehnten Systems.
Komplexität: Gering, aber fehleranfällig, da diese Bedingungen auf mehreren Ebenen berücksichtigt werden müssen (Distanzbestimmung).
- **Gleichzeitigkeit**
Ziel: Sicherstellung, dass alle Teilchen gleichzeitig aktualisiert werden.
Komplexität: Leicht bis Mittel. Die Komplexität liegt in der Natur von C++ und dem Verwenden von Parallelisierungen.

Von all diesen Aufgaben steht das Bestimmen der Nachbarn im algorithmischen Kern der Simulation. Diese Berechnung ist am kostenintensivsten und fehleranfälligsten; nicht zuletzt, weil dies die einzige Aufgabe ist, die sich zu parallelisieren lohnt. Um einen quadratischen Rechenaufwand $\mathcal{O}(n^2)$ zu vermeiden, wird der Simulationsbereich in quadratische **Zellen** aufgeteilt, die in ihrer Kantenlänge dem Wechselwirkungs-Durchmesser $2r$ entsprechen. Die Teilchen werden dann den entsprechenden Zellen zugeordnet. Dies ermöglicht die Untersuchung einer kleinen Nachbarschaft an Zellen, statt über jedes einzelne Teilchen iterieren zu müssen.

Für eine Aufteilung der Simulation in Zellen oder gar dem geschickten Verwalten von Adjazenzlisten gibt es auch fortgeschrittene und effektivere Methoden, wie etwa [Verlet-Listen](#) [36] oder [Quadrees](#) [37], die eine dynamische Anpassung der Zellen ermöglichen und den Rechenaufwand weiter minimieren. Diese Methoden wurden jedoch aus Komplexitäts- und Implementierungsgründen in dieser Arbeit nicht verfolgt.

Weitere Herausforderungen und typische Probleme der Implementierung waren - neben den 4 gelisteten Aufgaben - die Fehleranfälligkeit von Randfällen, die Sensibilität von C++ Parallelisierungen, die notwendige Kompromissbereitschaft bei Zufallsgeneratoren und das Optimieren rechenintensiver Operationen wie etwa das Sortieren oder das Mischen von Vektoren.

Zur Validierung der Implementation wurden die Simulationen mit zahlreichen Details und Informationen visualisiert. Abbildung 21 zeigt eine Simulation mit einem üblichen Randfall. Diese visuelle Stütze war für die Entwicklung der Simulation unverzichtbar.

3.2 Vicsek-Modell und topologisches Modell

Das Vicsek-Modell und das topologische Modell dienen als Grundlage für die Analyse der Schwarmdynamik in dieser Arbeit. Um die Qualität der Simulation zu überprüfen, werden die Messergebnisse mit bereits vorhandenen Forschungsergebnissen abgeglichen. Außerdem werden einfache statistische Methoden angewendet, um die Messstabilität weiter zu verbessern.

Statistische Methoden zur Optimierung der Messgenauigkeit

Die für die gesamte Auswertung relevante Größe ist der Ordnungsparameter v_a (4). Das Ziel ist die Equilibrierung einer Simulation eines spezifischen Parametersatzes. Für jeden Zeitschritt in der Simulation kann der Ordnungsparameter berechnet werden. Der naive Ansatz, am Ende einer langen Simulation den finalen Wert für v_a abzugreifen, wird kein zuverlässiges Ergebnis produzieren. Die Fluktuationen von v_a können durch das Rauschen η so stark sein, dass auch eine Mittelung über einige letzte Werte von v_a nur in seltenen Fällen aussagekräftig ist.

Stattdessen werden einfache statistische Methoden angewendet, um nicht nur den Ordnungsparameter v_a zuverlässig quantifizieren zu können, sondern auch um unnötig lange Simulationszeiten zu vermeiden.

Das Rauschen η bewirkt, dass der Ordnungsparameter v_a auch im equilibrierten Zustand eines Systems eine zufallsverteilte Größe ist. In solchen Fällen garantiert das **Gesetz der großen Zahlen**, dass der Durchschnitt der Messreihe bei großen Stichproben gegen den Erwartungswert konvergiert. Wenn die Zufallszahlen zudem **unabhängig** voneinander sind, ist auch der **Zentrale Grenzwertsatz** gültig. Dadurch sinkt der Fehler der Messreihe proportional zu $\frac{1}{\sqrt{T}}$, und die Verteilung der Messdaten um den Erwartungswert kann als normalverteilt betrachtet werden. T entspricht hier der Anzahl der Zeitschritte und damit der Anzahl der Messwerte.

Autokorrelation

Unabhängig von den Parametern einer Simulation werden direkt aufeinanderfolgende Messungen miteinander korreliert sein. Um diese Korrelation aufzuheben, kann die Korrelationszeit τ_c bestimmt und verwendet werden. Messungen, die nur alle τ_c Simulationsschritte durchgeführt werden, sind damit nicht mehr oder nur wenig korreliert. Das Aufheben der Korrelation ist wünschenswert, da es die Proportionalität der Fehlerabschätzung von $\frac{1}{\sqrt{T}}$ gewährleistet, indem die Messungen statistisch unabhängig voneinander gemacht werden.

Das Bestimmen der Autokorrelationszeit ist jedoch aufwendig, da sie für jede einzelne Parameterkonfiguration bestimmt werden muss. In diesem Sinne müsste auch ihre Wiederverwendbarkeit getestet werden. Obwohl die Autokorrelationszeit durchaus relevant und hilfreich ist, werden für das dynamische Beenden einer Simulation andere Methoden verwendet.

Konvergenzkriterien und Stabilität

Um die zuvor motivierten Theoreme anwenden zu können, ohne die Autokorrelationszeit zur Gewährleistung der statistischen Unabhängigkeit bestimmen zu müssen, werden aufeinanderfolgende Messungen in **Sweeps** durchgeführt. Im Optimalfall entspricht die Länge bzw. Größe eines Sweeps der Autokorrelationszeit. Da diese nicht bekannt ist, wird die Länge eines Sweeps auf die Teilchenzahl N gesetzt. Diese naive Strategie hat sich bereits im Ising-Modell als effektiv erwiesen [38]. Zwar wird damit keine Unabhängigkeit garantiert; der nachteilige Effekt der Korrelation aber sicherlich abgedeckt.

Als **Abbruchkriterium** wird ein **laufender Durchschnitt** verwendet. Nach einer **Burn-In Phase** von 10 Sweeps beginnt die kumulative Summation des Ordnungsparameters. Fällt die *absolute* Änderung des laufenden Durchschnitts unter eine Schranke von 10^{-4} in 10 aufeinanderfolgenden Messungen, wird das System als equilibriert angesehen und die Simulation beendet. Das Verwenden von mehreren Messungen verhindert, dass zwei konsekutive laufende Durchschnitte zufällig nah beieinander liegen. Die Burn-In Phase reduziert den Einfluss verschiedener Anfangsbedingungen. Als obere Schranke für die Simulationslänge dient eine empirisch bestimmte Grenze von 750 000 Zeitschritten, was im individuellen Fall $T = \frac{750\,000}{N}$ Messungen bzw. Sweeps entspricht.

Diese Herangehensweise hat sich als effizient und robust erwiesen und bietet einen pragmatischen Kompromiss zwischen methodischer Einfachheit und statistischer Zuverlässigkeit.

Vicsek-Modell

Die Simulationsergebnisse für das Vicsek-Modell sind in je zwei vergleichenden Abbildungen gegenübergestellt. Abbildung 13 zeigt die klassische Untersuchung eines Schwarms bei Variation des Rauschens η . Dieser Plot gibt Aufschluss darüber, wie stabil ein Schwarm bei entsprechendem Rauschen ist. Es ist erkennbar, dass Systeme mit größerem Simulationsbereich mehr Unordnung aufweisen. Da alle Simulationen bei gleicher Dichte $\rho = 4$ durchgeführt wurden, kann daraus abgeleitet werden, dass es eine **effektive Wechselwirkungsreichweite** gibt. Die Überlegung dazu ist intuitiv: Es gebe ein Teilchen A, das ein anderes Teilchen B am Rand seines Wechselwirkungsradius r sieht. Teilchen B sehe nun ein weiteres Teilchen C, das jedoch entgegengesetzt zu A positioniert ist. Damit wechselwirkt Teilchen A indirekt mit Teilchen C über Teilchen B (siehe Abbildung 8). Für die Propagation von A zu C ist jedoch ein Simulationsschritt notwendig. Bei maximalem Abstand gilt die Regel, dass eine Wechselwirkung n -ter Ordnung n Simulationsschritte benötigt, um zu propagieren. In den Simulationen wird diese Propagation durch das Rauschen η gestört. Bei zu starkem Rauschen gehen die Informationen in dem stochastischen Störfaktor verloren. Daher ist es plausibel, dass für gleichbleibende Dichte große Systeme instabiler sind; insbesondere bei hohem Rauschen.

Der Vergleich der ersten Ergebnisse zeigt eine gute Übereinstimmung mit den originalen Daten. Die Messkurven stimmen vor allem qualitativ überein. Bei genauerer Betrachtung liegen dennoch Abweichungen vor. Diese werden in dem Schluss-Teil 6 diskutiert.

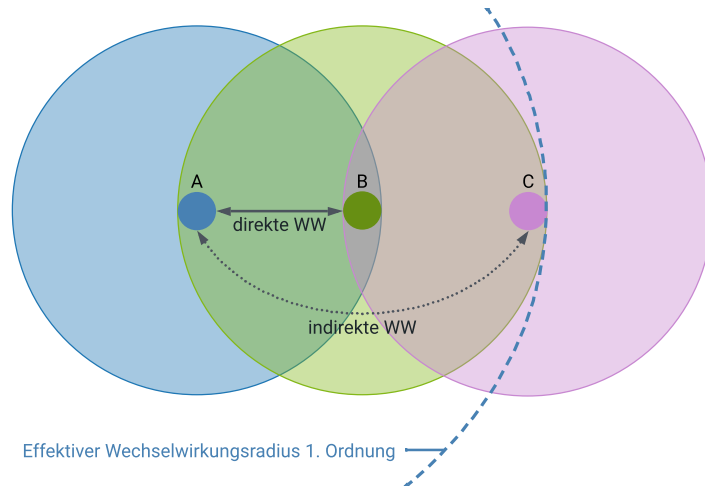


Abbildung 8: Effektiver Wechselwirkungsradius 1. Ordnung bei maximalem Abstand. Teilchen A und Teilchen C wechselwirken indirekt über Teilchen B.

In dem zweiten Vergleich (Abbildung 10) wird nun die Dichte bei gleichbleibendem Rauschen variiert. Die durchgezogene Kurve in dem originalen Plot entspricht den kritischen *Temperaturen* des Systems. Analog zu thermodynamischen Systemen übernimmt hier das Rauschen η die Rolle einer Temperatur. Die beiden Abbildungen stimmen qualitativ bis auf den niedrigeren Wertebereich von ρ überein.

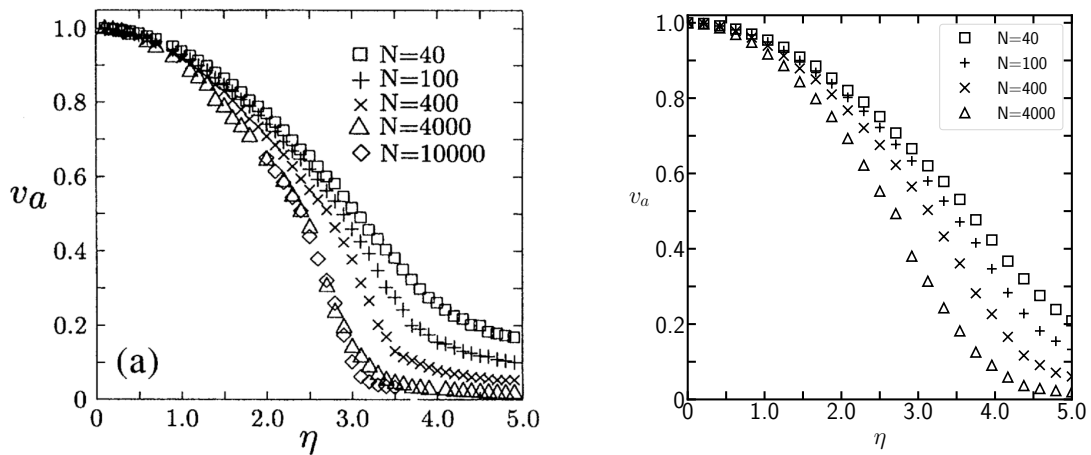


Abbildung 9: Links: Messergebnisse aus der originalen Arbeit (Fig. 2a) [1]. Rechts: Messergebnisse der vorliegenden Implementierung. In beiden Abbildungen wurde $\rho = 4$ verwendet.

Analytische Untersuchungen

Es kann gezeigt werden, dass die Ausrichtungsregel (3) des Vicsek-Modells den Ordnungsparameter v_a (6) minimiert. Hierfür wird der Ordnungsparameter in Polardarstellung gebracht und nach dem Winkel φ_k abgeleitet. Die Überlieferung in die Polardarstellung produziert eine Wurzel:

$$\begin{aligned}
 v_a &= \frac{1}{Nv_0} \left| \sum_{i=0}^N \tilde{\mathbf{v}}_i \right| = \frac{1}{Nv_0} \left| \sum_{i=0}^N v_0 \begin{pmatrix} \cos \varphi_i \\ \sin \varphi_i \end{pmatrix} \right| \\
 &= \frac{1}{Nv_0} \sqrt{v_0^2 \left[\left(\sum_i \cos \varphi_i \right)^2 + \left(\sum_i \sin \varphi_i \right)^2 \right]}
 \end{aligned} \tag{6}$$

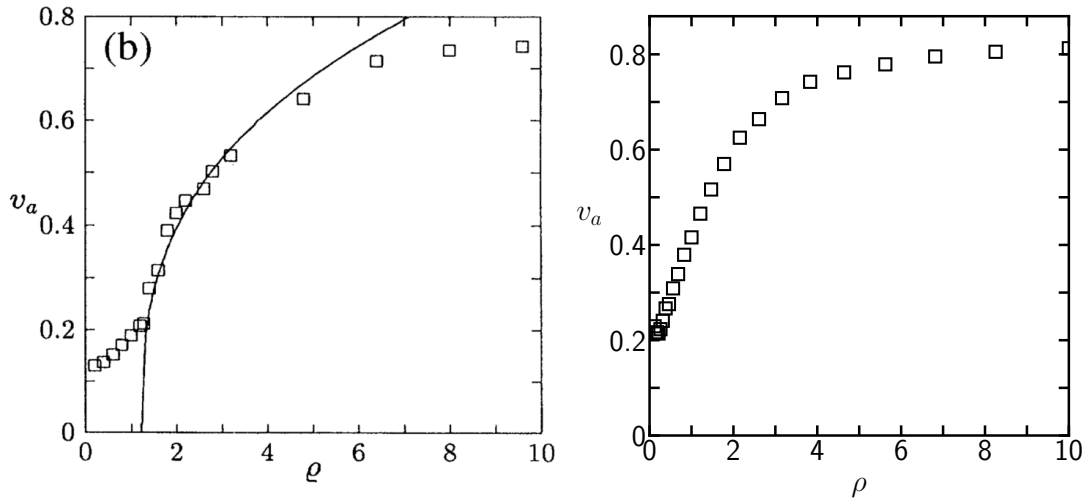


Abbildung 10: Links: Messergebnisse aus der originalen Arbeit (Fig. 2b) [1] bei *unbekanntem* Rauschen η . Rechts: Messergebnisse der vorliegenden Implementierung für $\eta = 2.0$. In beiden Abbildungen ist $L = 20$ konstant.

Um den Gebrauch der Kettenregel beim Ableiten der Wurzel zu vermeiden, wird v_a quadriert. Dadurch ändern sich die Nullstellen der ersten Ableitung nicht, denn es gilt

$$\frac{d}{dx} (f(x))^2 = 2f(x)f'(x) \quad .$$

Im Allgemeinen können weitere Nullstellen für $\frac{d}{dx} (f(x))^2$ hinzukommen, und zwar für jedes x_i mit $f(x_i) = 0$ und $f'(x_i) \neq 0$. Der Ordnungsparameter v_a ist zwar auf $v_a \in [0, 1]$ beschränkt, jedoch kann $v_a = 0$ praktisch ausgeschlossen werden. In Vielteilchen-Systemen wird $v_a = 0$ effektiv nie erreicht, da der Phasenraum für diesen Fall verschwindend gering ist.

Mit der Nutzung von $\sum_i^N \cos \varphi_i = N \langle \cos \varphi \rangle$ (analog für den Sinus) kann nun die Vicsek-Regel hergeleitet werden:

$$\begin{aligned} 0 &\stackrel{!}{=} \frac{\partial}{\partial \varphi_k} v_a^2 \\ &= \frac{\partial}{\partial \varphi_k} \left(\frac{1}{N^2 v_0^2} v_0^2 \left[(\cos \varphi_1 + \dots + \cos \varphi_N)^2 + (\sin \varphi_1 + \dots + \sin \varphi_N)^2 \right] \right) \\ &= \frac{1}{N^2} \left[-2 \sin \varphi_k \sum_i^N (\cos \varphi_i) + 2 \cos \varphi_k \sum_i^N (\sin \varphi_i) \right] \\ &= -2 \sin \varphi_k \cdot N \langle \cos \varphi \rangle + 2 \cos \varphi_k \cdot N \langle \sin \varphi \rangle \end{aligned} \tag{7}$$

$$\Leftrightarrow \sin \varphi_k \langle \cos \varphi \rangle = \cos \varphi_k \langle \sin \varphi \rangle$$

$$\Leftrightarrow \frac{\sin \varphi_k}{\cos \varphi_k} = \tan \varphi_k = \frac{\langle \sin \varphi \rangle}{\langle \cos \varphi \rangle}$$

wobei
 $\langle \cos \varphi \rangle \neq 0$

$$\Leftrightarrow \varphi_k = \arctan \left(\frac{\langle \sin \varphi \rangle}{\langle \cos \varphi \rangle} \right)$$

In der originalen Arbeit wird eine analytische Untersuchung des Ordnungsparameters nicht erwähnt. Daher ist nicht bekannt, ob die Vicsek-Regel analytischen Methoden entspringt oder intuitiv gewählt wurde.

Topologisches Modell

Das topologische Modell behandelt die Umstellung der Nachbarschafts-Auswahlregel auf eine Methode mit beliebiger Reichweite. Wie bereits in der Theorie motiviert, werden außer dieser Regel keine Änderungen an der Simulation vorgenommen. Durch die feste Anzahl der topologischen Nachbarn k erhält der Parameterraum eine weitere Dimension. Die Ergebnisse werden dem Vicsek-Modell bzw. Abbildung 9 qualitativ nachgestellt, sodass sich für jeden möglichen Wert von k ein Plot ergibt.

Die in der Theorie erwähnte Forschungsarbeit zu diesem Modell beobachtet eine starke Variation der Messergebnisse für $k \leq 3$ [3]. Ab Werten von etwa $k \geq 7$ sind die Unterschiede nur noch minimal. Daher werden auch hier nur die ersten 6 Werte für k untersucht.

Die Umstellung auf k Nachbarn ermöglicht theoretisch eine unendliche Reichweite der Wechselwirkung. Die hier betrachteten Simulationen behandeln jedoch Dichten, die diese Reichweite effektiv einschränken. Eine weitere praktische Einschränkung ergeben die periodischen Randbedingungen, durch die sich Teilchen nicht weiter weg als $\frac{L}{2}$ voneinander distanzieren können. Des Weiteren ist die Wahrnehmung der Teilchen für kleine k so weit eingeschränkt, dass lokal hohe Dichten, also Schwarmpakete, nicht kohärent sind, weil jeweils nur der nächste Nachbar relevant ist (oder ein paar wenige). Zwar kann wieder über eine effektive Wechselwirkungsreichweite diskutiert werden (die hier nun eine dynamischere Länge hat und für kleine Dichten drastisch an Reichweite gewinnen kann), aber diese ist durch die Anwesenheit vieler Teilchen gestört. Ein Schwarmpaket hoher Dichte hat also eine kurze Wechselwirkungsreichweite, während Informationen durch aufgelockerte Pakete besser propagieren können.

In Abbildung 11 demonstrieren einige Schnappschüsse diese lokale Befangenheit. In den topologischen Modellen bilden sich charakteristische Pakete von Teilchen mit einer Größe proportional zu k aus. Das System erinnert in diesem Zustand an eine Mischung zweier Phasen, einer gasförmigen und einer flüssigen Phase. Dieses Verhalten löst sich mit der Erhöhung von k auf und das System geht zunehmend in einen Zustand über, der einer flüssigen Phase gleicht. Wird das System jedoch wieder vergrößert, kehrt diese Paket-Bildung zurück. Sichtbar wird dieser Effekt in Abbildung 12 für $k = 6$ (bzw. $n = 6$). Dort kann qualitativ ein Knick bei etwa $\eta = 1$ beobachtet werden, was auf eine Koexistenz zweier Phasen deutet. Für eine geringere Anzahl an Nachbarn ist diese Unstetigkeit der ersten Ableitung auch bei kleineren Systemgrößen (bzw. kleineren Teilchenzahlen N) erkennbar, was wieder für eine Paket-Bildung spricht.

In dem nächsten Abschnitt wird ein Ansatz vorgestellt, der diesen heterogenen Effekt aufheben soll, ohne die topologische Natur des Modells zu ändern.

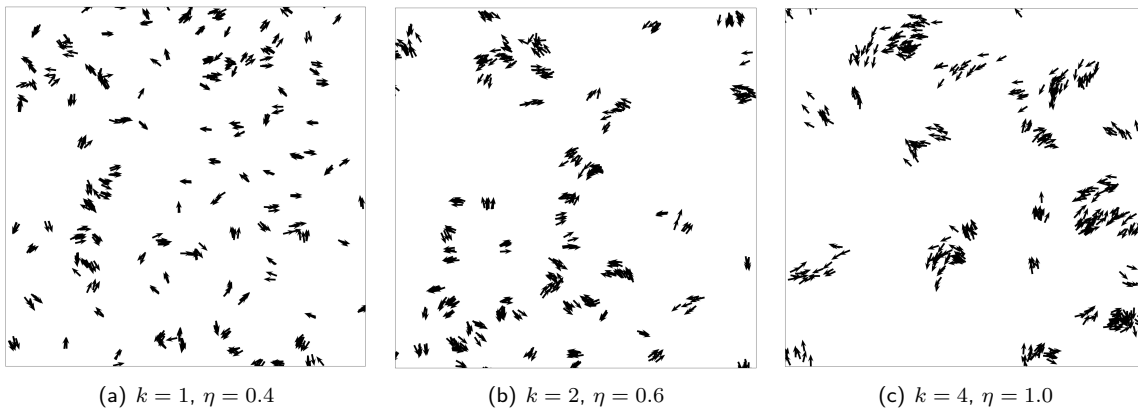


Abbildung 11: Für geringes k bilden sich kleine *Schwarmpakete* aus, die nur sich selbst wahrnehmen können, sofern sie eine gewisse Distanz zu anderen Teilchen besitzen. Eine höhere Anzahl an topologischen Nachbarn vergrößert diese Pakete und weicht dieses Verhalten somit ein wenig auf. Wird das Rauschen η jedoch erhöht, brechen diese Pakete auf und es ergibt sich ein Zustand, der an die Koexistenz von flüssiger und gasförmiger Phase erinnert. Diese Koexistenz ist erkennbar an den Stetigkeitssprüngen der ersten Ableitung der Messkurven in Abbildungen 12 und 13. Alle Simulationen wurden für $N = 400$ und $\rho = 4$ durchgeführt; Simulationsschritte ($\tilde{t} < 150$).

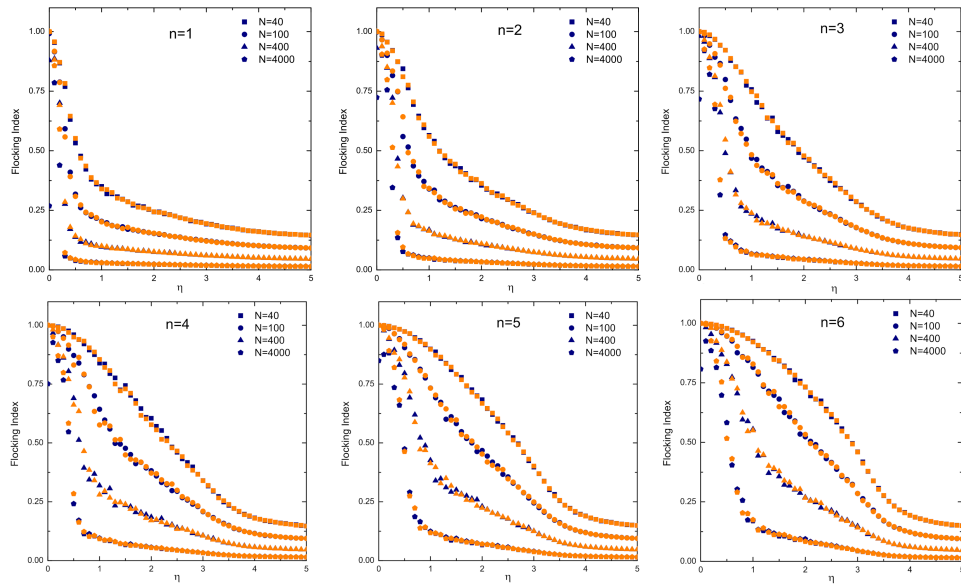


Abbildung 12: Simulationsergebnisse einer vorangegangenen Forschungsarbeit für das topologische Modell (hier n -Nachbarn, konstante Dichte $\rho = 4$). Die Ergebnisse dienen dem Vergleich mit Abbildung 13 und stammen aus [3].

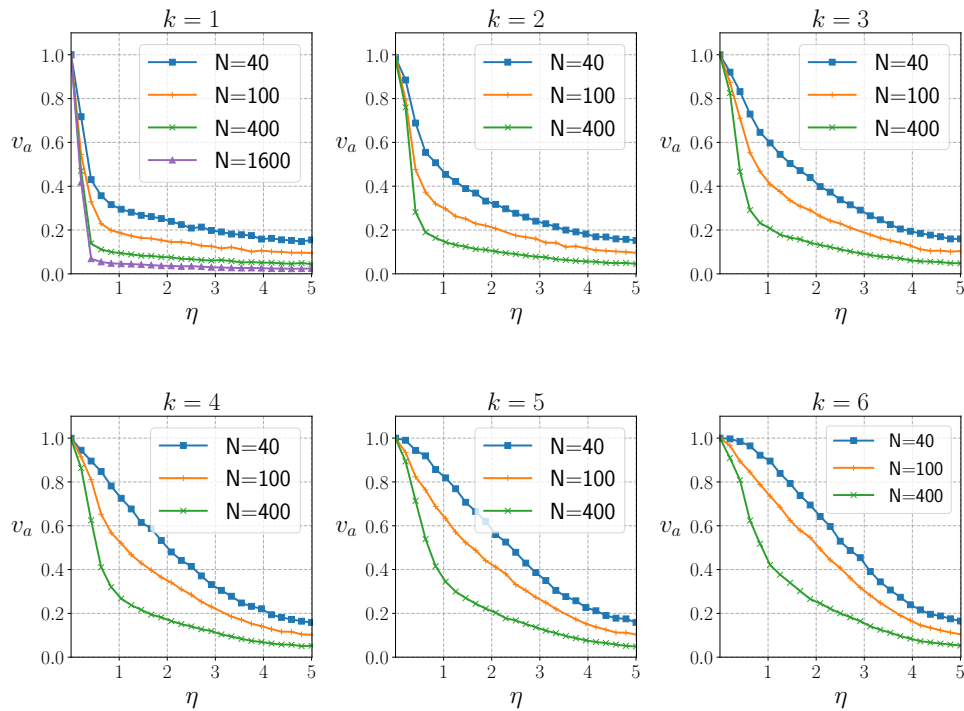


Abbildung 13: Topologisches Modell mit konstanter Dichte $\rho = 4$. Die Daten stimmen gut mit den Ergebnissen aus [3] überein. Bei kleineren Teilchenzahlen von $N = 40$ sind geringe Abweichungen erkennbar, während die restlichen Messungen weitgehend übereinstimmen.

4 Metrisch-Topologischer Ansatz

Die Hauptmotivation zur Verwendung eines topologischen Modells in dieser Arbeit ist die Fixierung der Nachbarteilchen-Anzahl. Im letzten Abschnitt wurde beobachtet, dass diese Modelle jedoch zu irregulären Schwarpaket-Bildungen neigen, was das System instabil macht. Für einen qualitativen Vergleich der Modelle sind in Abbildung 14 drei Schnappschüsse des Vicsek-Modells dargestellt - mit gleichen Rausch-Einstellungen wie die zuvor gezeigten Schnappschüsse 11.

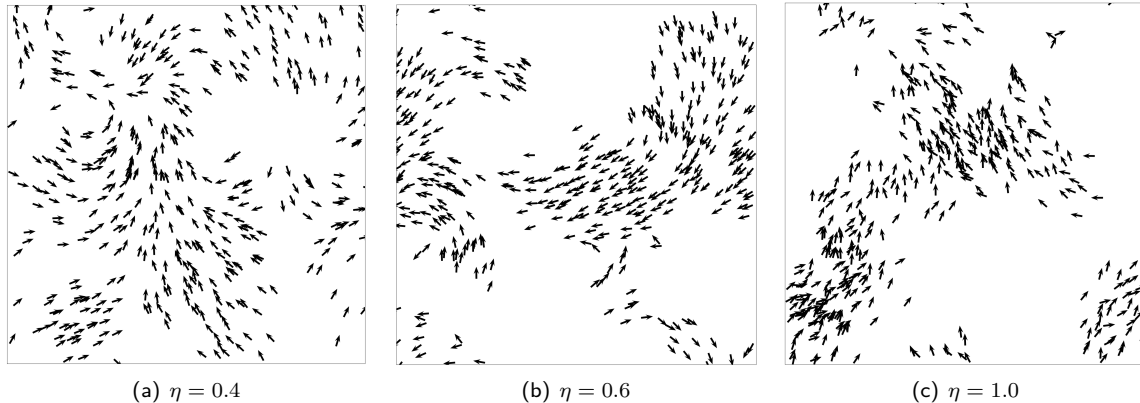


Abbildung 14: Schnappschüsse des metrischen Vicsek-Modells für einen qualitativen Vergleich zu den anderen Modellen aus Abbildung 11 und 17. Es ist $N = 400$ und $\rho = 4$. Alle Schnappschüsse wurden nach nur sehr wenigen Simulationsschritten aufgenommen ($\bar{t} < 50$).

4.1 Einführung des neuen Modells

Die Grundidee eines neuartigen Ansatzes besteht darin, den Wechselwirkungsradius r eines metrischen Modells beizubehalten und damit ein wenig von der Stabilität solcher Modelle zurückzugewinnen. Eine mögliche Umsetzung dieser Idee ist also das Bestimmen von k Nachbarn und dem Verwerfen aller Teilchen, die einen größeren Abstand als r aufweisen. Diese Herangehensweise stellt jedoch eine *Verschlechterung* der Stabilität dar, wie man sich schnell klarmachen kann: Die Instabilität topologischer Modelle resultiert aus der effektiven Kurzreichweitigkeit durch Stellen lokaler hoher Dichte. Diese Kurzreichweitigkeit wird verstärkt, wenn nun auch noch Nachbarn ausgeschlossen werden, die weiter als r entfernt sind.

Mit einer einzigen Modifikation kann dieses Modell jedoch signifikant verbessert werden (gemessen an der Stabilität und der Vergleichbarkeit mit dem Vicsek-Modell). Die Idee dieser Anpassung entsprang der Intuition: In Abbildung 15(a) ist ein radiales Histogramm gezeigt, das die Perspektive eines einzelnen Teilchens zeigt. Dieses Teilchen sieht in seinem Wechselwirkungsradius exemplarisch zwei große Schwärme. In einem metrischen Modell würde es nun über alle Teilchen mitteln. Durch die topologische Einschränkung jedoch ist es auf k Nachbarn beschränkt. Würde es sich hier um einen biologischen Fall handeln und die gezeigte Perspektive dem Sehen entsprechen, wäre es plausibel anzunehmen, dass das Auge in schnellen Situationen die Informationen der Umgebung abstrahiert und in gröbere *Bereiche* einteilt. Falls nur k Bereiche zur Verfügung stehen, müssen die Informationen in diesen Bereichen zusammengefasst werden. Dies entspräche einer Mittelung aller Teilchen in Bereichen gleicher Größe, wobei die *Größe* hier der Teilchenanzahl entspricht. In einem Histogramm kann das mittels k -Quantile dargestellt werden. In der Abbildung sind beispielhaft 8-Quantile gezeigt, die die Daten in gleich große Teilchenzahl-Bereiche teilen. Diese Bereiche können dann gemittelt werden, um die zuvor motivierte Überlegung umzusetzen.

In der Praxis ist diese Methodik aber umständlich und rechenintensiv. Ein wesentlich eleganterer Ansatz mit vergleichbarer Wirkung kann über einen stochastischen Zugang erreicht werden. Statt einer komplexen Bereichsaufteilung können zunächst alle Kandidaten im Wechselwirkungsradius r bestimmt werden. Anschließend werden **zufällig** k Nachbarn aus diesem Pool bestimmt. Mit einer zunehmenden Anzahl von Simulationsschritten

entspricht dieser stochastische Ansatz mehr und mehr der Mittelung der k Bereiche. Im gewissen Sinne wird dieses deterministische Problem dadurch mit Monte-Carlo-Methoden gelöst.

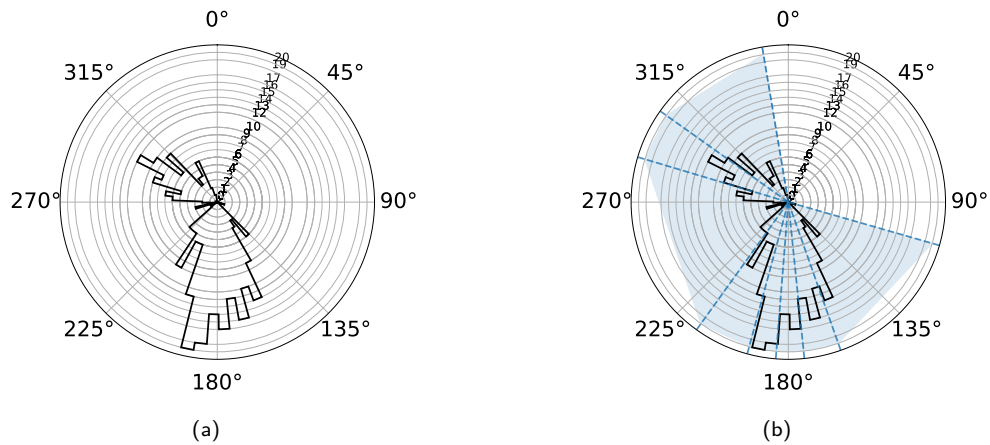


Abbildung 15: Beispielhafte Darstellung von zwei Schwärmen unterschiedlicher Größe und Ausrichtung. Die Histogramme beschreiben die Sichtweise eines einzelnen Teilchens, das diese beiden Schwärme gleichzeitig sieht (Imitation des Vicsek-Modell-Wechselwirkungsradius). **Links:** Radiales, uninterpretiertes Histogramm aus Sicht eines Teilchens. **Rechts:** Unterteilung des Histogramms in k -Quantile mit $k = 8$. Es wird dasselbe System wie in 15(a) betrachtet.

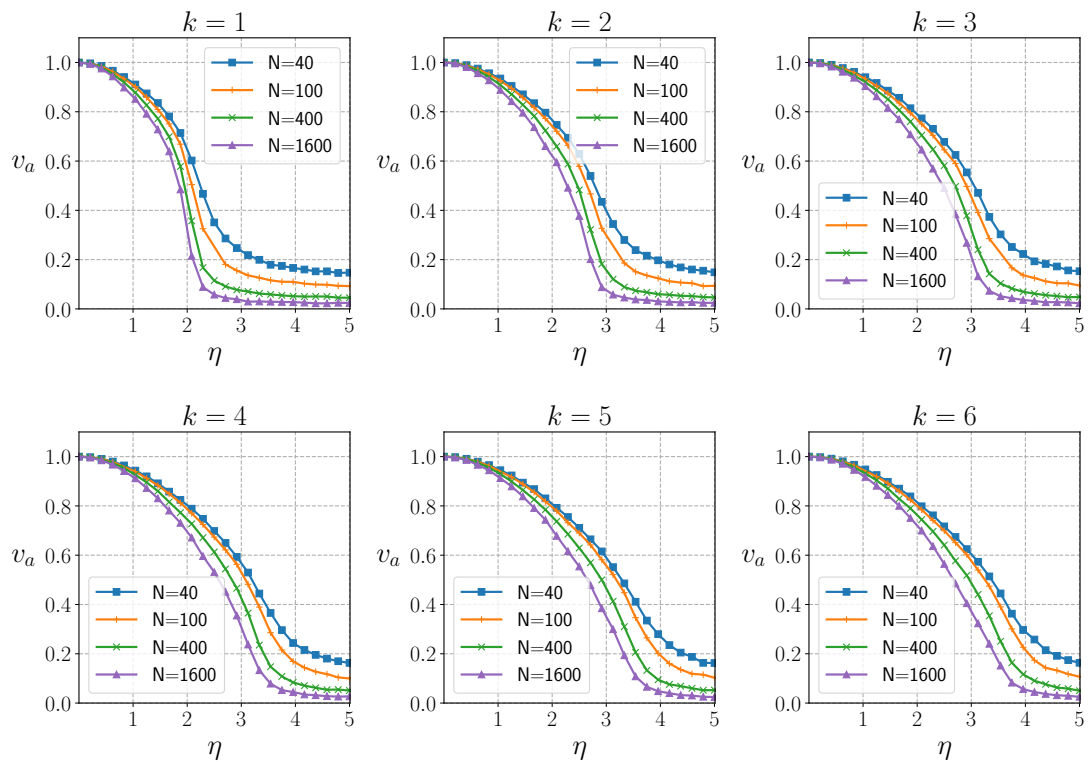


Abbildung 16: Metrisch-Topologisches Modell mit konstanter Dichte $\rho = 4$.

4.2 Messergebnisse

In Abbildung 16 sind die Messergebnisse des metrisch-topologischen Modells zu sehen. Für die Nachbarschafts-Bestimmung wurde die zuvor diskutierte stochastische Modifikation angewendet, was den Verlauf der Messkurven im Vergleich zu dem rein topologischen Modell deutlich ändert. Insbesondere für eine geringe Anzahl an Nachbarn verhält sich das Modell bei geringem Rauschen wesentlich stabiler. Die *Knicke* bzw. Unstetigkeiten in den ersten Ableitungen sind verschoben, wenn nicht sogar aufgehoben.

Ein Blick auf die Schnappschüsse 17 verdeutlicht den großen Unterschied zu dem rein topologischen Modell und die Ähnlichkeit zu dem Vicsek-Modell. Obwohl pro Simulationsschritt nur k Nachbarn ausgewählt werden, ist keine Paket-Bildung kleinerer Schwärme mehr beobachtbar und die Form ähnelt den strom-artigen Strukturen eines metrischen Modells. Ein Unterschied zu dem Vicsek-Modell ist insofern erkennbar, als dass die Schwärme des neuartigen Modells keine scharfen Kurven aufweisen (siehe zum Vergleich die ausgeprägteren *Wirbel* in Abbildung 14(a)). Dies ist auch plausibel, da in den Situationen, in denen viele Teilchen aus verschiedensten Richtungen aufeinandertreffen, das metrisch-topologische Modell nicht alle Nachbarn in einem Zeitschritt gleichzeitig berücksichtigen kann. Daher sind die Schwärme für kleinere k im Vergleich zum metrischen Modell strukturell geglättet.

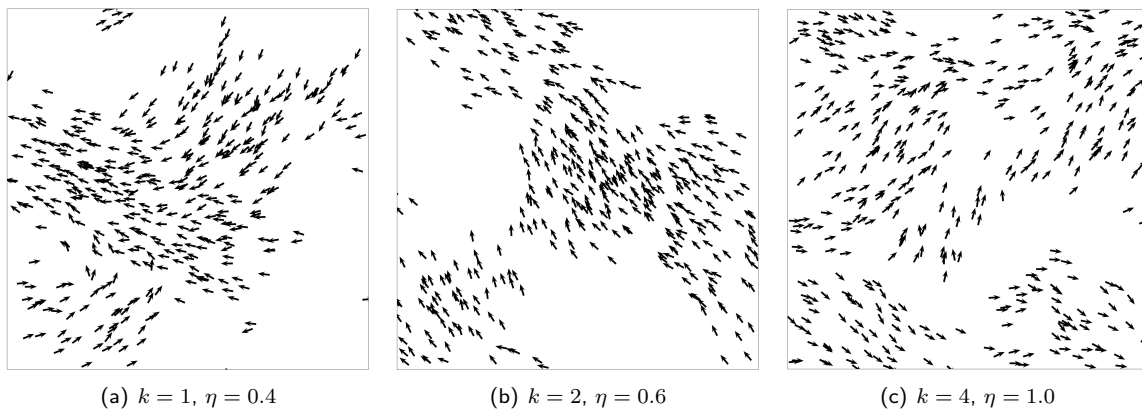


Abbildung 17: Schnappschüsse des metrischen-topologischen Modells mit stochastischer Modifikation. Es ist $N = 400$ und $\rho = 4$. Alle Schnappschüsse wurden nach nur *sehr* wenigen Simulationsschritten aufgenommen ($\tilde{t} < 20$).

4.3 Konvergenz zum Vicsek-Modell

Das metrisch-topologische Modell konvergiert gegen das Vicsek-Modell für große k . Dieser Fall kann leicht konstruiert werden, indem $k > N$ betrachtet wird. In dieser Situation wäre das Modell in der Lage, jede beliebige Teilchenanzahl im Wechselwirkungsradius r gleichzeitig zu berücksichtigen. Dies entspricht aber gerade dem metrischen Modell, wodurch die Konvergenz qualitativ sofort begründet ist. Abbildungen 18 und 19 zeigen eine quantitative Untersuchung dieses Konvergenzverhaltens. Für $k = 6$ sind praktisch kaum noch Abweichungen erkennbar. Wird der durchschnittliche quadratische Fehler (MSE) gegen die Anzahl der Nachbarn k aufgetragen, ist sofort der exponentielle Abfall der Abweichungen erkennbar. In dem nebenstehenden Plot in Abbildung 19 wird mit einer log-log-Darstellung und einem linearen Fit genau diese exponentielle Beziehung hervorgehoben. Dabei wird beobachtet, dass große Systeme mit vielen Teilchen langsamer konvergieren als kleinere Systeme (bei gleicher Dichte $\rho = 4$).

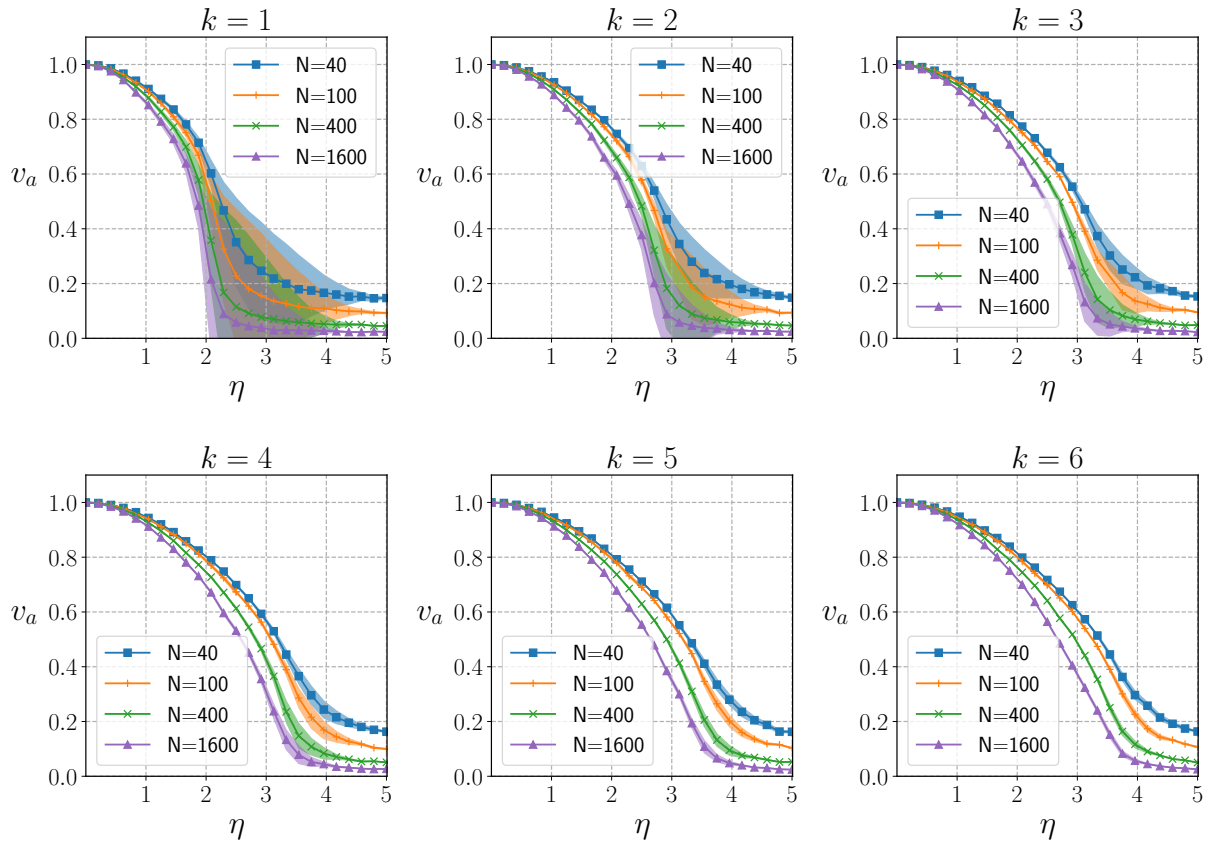


Abbildung 18: Konvergenz-Analyse und Vergleich des metrisch-topologischen Modells mit dem Vicsek-Modells. Es wird die quadratische Abweichung betrachtet, welche als Schlauch um die Messkurven gelegt ist.

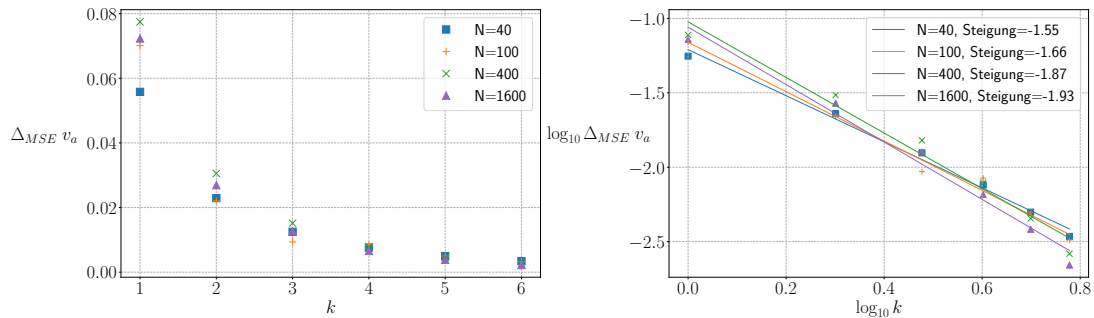


Abbildung 19: Konvergenz des metrisch-topologischen Modells mit dem Vicsek-Modell. Es wird mit den selbst generierten Daten aus Abbildung 9 verglichen, nicht mit den Originaldaten. **Links:** Mean squared error (MSE) für jede Gruppierung von N und jedes k . **Rechts:** Eine log-log-Darstellung mit linearem Fit. Die Abweichungen sind stärker, je größer das System bzw. die Teilchenzahlen sind.

4.4 Einführung eines dichte-basierten Ordnungsparameters

Bislang fanden alle Untersuchungen basierend auf nur einem einzigen Ordnungsparameter (4) statt. Der Vicsek-Ordnungsparameter v_a misst die kollektive Ausrichtung des Schwarms und ist maximal, falls alle Teilchen parallel ausgerichtet sind. Er ist jedoch unabhängig von lokalen Dichtefluktuationen des Systems und macht keine Aussage über die Kompaktheit eines Schwarms. Diese Lücke sollen distanz- bzw. dichte-basierte Ordnungsparameter schließen. Ein möglicher Ansatz ist die distanzbasierte Gewichtung des Skalarproduktes zweier Teilchen:

$$\rho_a = \frac{N(N-1)}{2} \sum_{i < k}^N \tilde{v}_i \tilde{v}_k \frac{1}{\left(|\tilde{x}_i - \tilde{x}_k|^2 + d^2\right)^{\frac{\alpha}{2}}} = \frac{N(N-1)}{2} \sum_{k=0}^N \sum_{i=k+1}^N \tilde{v}_i \tilde{v}_k d_{ik}(\alpha, d) \quad (8)$$

Es wird dann über alle Teilchen summiert. Um die großen Beiträge sehr naher Teilchen in d_{ik} zu minimieren, kann ein Glättungsfaktor d eingeführt werden. Gleichzeitig kann mit α reguliert werden, wie stark diese Distanzbeiträge in den Ordnungsparameter ρ_a einfließen. In einer weiteren Überlegung kann der Ausdruck in der runden Klammer auch durch $\arg\max(d_{ik}^2, d^2)$ ersetzt werden. Wird d dann bspw. auf den Wechselwirkungsradius r gesetzt, bedeutet dies, dass wechselwirkende Teilchen einen maximalen, konstanten Beitrag zum Ordnungsparameter leisten.

Abbildung 20 zeigt die in dieser Arbeit nun schon mehrfach untersuchten Konfigurationen unter Verwendung des stochastischen metrisch-topologischen Modells. Die Werte für $\eta = 0$ sind nicht eingetragen, da diese stark von den Anfangsbedingungen abhängen und der qualitative Verlauf dennoch sehr gut abgelesen werden kann. Da eine quantitative und qualitative Diskussion von ρ_a in dieser Arbeit eine eher untergeordnete Rolle spielt, befinden sich die Abbildungen dazu im Anhang. Es soll vielmehr gezeigt werden, dass diese Parameter eingeführt werden können und auf ihre Aussagekräftigkeit individuell untersucht werden müssen.

5 Neuronale Schwärmdynamik mittels Reinforcement Learning

5.1 Aufbau des Frameworks und des neuronalen Netzwerks

Für diese Arbeit wurde die bereits erwähnte Bibliothek [RLlib](#) verwendet, die viele Werkzeuge für das Aufsetzen eines MARL-Projektes mitliefert [25]. Als Framework für das Maschinelle Lernen bzw. für die neuronalen Netze wurde [TensorFlow](#) bzw. [Keras](#) verwendet. Wie in der Theorie bereits motiviert wird der PPO Algorithmus verwendet. Das neuronale Netz ist so einfach wie möglich gehalten, da das übergeordnete Ziel dieser Implementierung die Imitation der Vicsek-Regel ist, welche einer einfachen Mittelung entspricht. In Abbildung 22 ist beispielhaft ein solches Netz gezeigt.

5.2 Training und Validierung

Die Implementation ist an dieser Stelle beendet. Das Framework ist aufgesetzt und es wird auch ein Trainingsprozess durchgeführt. Die Multi-Agent Umgebung funktioniert und alle Elemente arbeiten zusammen. Die einzige Ausnahme bildet die Repräsentation des Winkels im Bezug auf das neuronale Netz. Durch diesen Umstand sind die Ergebnisse nicht interpretierbar und das weitere Vorgehen wird im Schluss-Teil diskutiert.



6 Diskussion & Ausblick

Diese Arbeit hat sich intensiv mit den etablierten Modellen der Schwärmdynamik auseinandergesetzt. Es wurde das Vicsek-Modell und seine topologische Modifikation implementiert und validiert. Die Suche nach einem brauchbaren Modell für das Maschinelle Lernen führte schließlich auf das neuartige metrisch-topologische Modell. All diese Modelle wurden rigoros untersucht, um die jeweiligen Stärken und Schwächen zu untersuchen.

Multi-Agent Reinforcement Learning

Der Kern dieser Arbeit beschäftigt sich mit der Frage, ob eine Schwärmdynamik durch neuronale Netze gelernt bzw. approximiert werden kann. Da es sich bei der Schwärmdynamik um Vielteilchensysteme handelt, muss ein Multi-Agent Framework verwendet werden. Diese Algorithmen sowie Frameworks sind so neu, dass sie aktuell einer sehr starken Entwicklung unterliegen. Daher war es besonders schwierig, sich nicht nur in das komplexe Thema des Reinforcement Learnings einzuarbeiten, sondern auch die notwendige Software aufzusetzen und zu testen. Es floss sehr viel Arbeit in die Konzeption und Erprobung dieser Umsetzung. Große Teile der Arbeit sind leider nicht visualisierbar.

Dennoch steht zum Schluss dieser Arbeit ein zentrales Ergebnis: Die mühsam erprobte C++ Simulation ist nun vollständig eingebettet in ein funktionierendes, trainierendes Framework von `RLlib`. Einmal aufgesetzt ist das Training vergleichsweise einfach, sodass nun ein neuer Abschnitt der Arbeit beginnen würde. Da die Änderung des Ordnungsparameters als Belohnungs-Funktion verwendet werden kann, können nun beliebig komplexe oder einfache Systeme untersucht werden.

Interessant wäre die Frage, ob ein einzelnes *single-layer* Perzeptron, also das einfachste Modell eines neuronalen Netzes, ausreicht, um die Vicsek-Mittelungsregel zu lernen. Die Vermutung ist: **Ja**. Diese Hypothese galt es zu testen, was wegen der zeitlichen Beschränkung dieser Arbeit leider knapp verpasst wurde.

Die Möglichkeiten, die nun offenstehen, sind zahlreich. Abbildung 24 zeigt eine Übersicht, die die Komplexität und Reichweite dieser Arbeit verdeutlichen soll. Es gibt zahlreiche interessante Variationen der Modelle und diese können beliebig um komplexere Elemente ergänzt werden.

Die modernen Algorithmen des Maschinellen Lernens haben sich vor allem bei komplizierten Aufgabenstellungen zunehmend bewiesen. Die Frage ist also vermutlich nicht *ob* diese Methoden komplexe Aufgaben der Schwärmdynamik lösen können, sondern *wie*.

Vicsek-Modell: Abweichungen zur Originalarbeit

Der Vergleich in Abbildung 9 zeigt zunächst einen qualitativ gleichwertigen Verlauf der Messkurven. Ein genauerer Blick auf die höheren Rauschwerte (ab etwa $\eta \geq 3$) deckt jedoch Unstimmigkeiten mit den originalen Messdaten auf. Die Gründe für diese Abweichungen sind unbekannt und können in dieser Arbeit nicht vollständig beantwortet werden. Dennoch werden mehrere potenzielle Fehlerquellen gründlich geprüft und ausgeschlossen.

- **Fehlerquellen im Code:** Ein Großteil der Entwicklungszeit wurde für die Validierung des Modells und das Ausschließen möglicher Fehler im Code aufgewendet: Die Bestimmung der nächsten Nachbarn, die Modell-Regeln, Periodische Randbedingungen, Distanzmessungen, propagierende Rundungsfehler und eine sorgfältige Auswahl der Zufallsgeneratoren sind alles Aspekte, die mehrfach untersucht und zum großen Teil auch visuell geprüft wurden.
- **Statistische Ungenauigkeiten:** Wie bereits erwähnt, werden keine fortgeschrittenen statistischen Methoden oder Tests verwendet. Die Stabilität der Messwerte wurde untersucht, indem unterschiedliche Simulationszeiten miteinander verglichen wurden. Zudem wurden verschiedene Anfangsbedingungen getestet; einschließlich solcher, die in einem vollständig geordneten Zustand beginnen. Die Ergebnisse konvergieren für eine gegebene Parameterkonfiguration unabhängig von den Anfangsbedingungen immer gegen den gleichen Wert. Der glatte Verlauf der Messwerte legt nahe, dass es sich um Gleichgewichtszustände handelt. Der Verdacht kann auch auf den laufenden Durchschnitt fallen, der durch die frühe Mitnahme von Messwerten sicherlich leicht gebiased ist. Dieser Bias würde sich aber Richtung $v_a \rightarrow 0$

(da das System zu Beginn ungeordnet ist) und damit entgegen der beobachteten Abweichung bewegen. Damit wird eine statistisch motivierte Abweichung unwahrscheinlich, obwohl dies durch den Vergleich mehrerer Messungen und der Anwendung zuverlässigerer statistischer Methoden überprüft werden müsste.

Auch Abbildung 10 weist Abweichungen auf, die jedoch ebenfalls schwierig zu untersuchen sind, da aus der originalen Arbeit nicht alle Implementierungs- und Auswertungsdetails bekannt sind. Es ist nicht bekannt, bei welchem Rauschen η diese Messungen durchgeführt werden. Die Abweichung kann eventuell so verstanden werden, dass sich die Kurve, die in dem originalen Plot zu sehen ist, durch ihre Abhängigkeit von η nach links verschiebt. Dann wären die Messergebnisse wieder plausibel. Dies könnte geprüft werden, indem die Messungen bei anderen Rauschwerten durchgeführt werden.

Ausblick

Die Ergebnisse dieser Arbeit bieten eine große Bandbreite an möglichen Weiterentwicklungen. Unabhängig von dem Maschinellen Lernen können die Ordnungsparameter weiter untersucht werden. Diese Parameter entspringen häufig der Intuition oder Erfahrung und können wertvolle Informationen des Systems beinhalten. Verschiedene Ordnungsparameter können verschiedene Aspekte des Systems beleuchten. Können diese Parameter analytisch abgeleitet werden, ergibt sich eventuell sogar eine Regel, die neuartige Muster zeigen kann. An dieser Stelle greift wieder der Reinforcement Learning Ansatz. Durch diese Methode können mehrere Ordnungsparameter miteinander kombiniert werden und das System lernt eigenständig die für die Optimierung dieser Parameter notwendigen Regeln.

Diese Arbeit verdeutlicht, dass die Schwärmdynamik weit mehr ist als nur ein theoretisches Konstrukt, sondern ein Schlüssel zur Enthüllung der Geheimnisse kollektiver Intelligenz und adaptiver Systeme. Sie markiert einen Meilenstein in der Annäherung zwischen klassischer Physik und modernen maschinenlernbasierten Methoden und lädt zur weiteren, tiefgreifenden Forschung in diesem faszinierenden interdisziplinären Feld ein.

A Plots

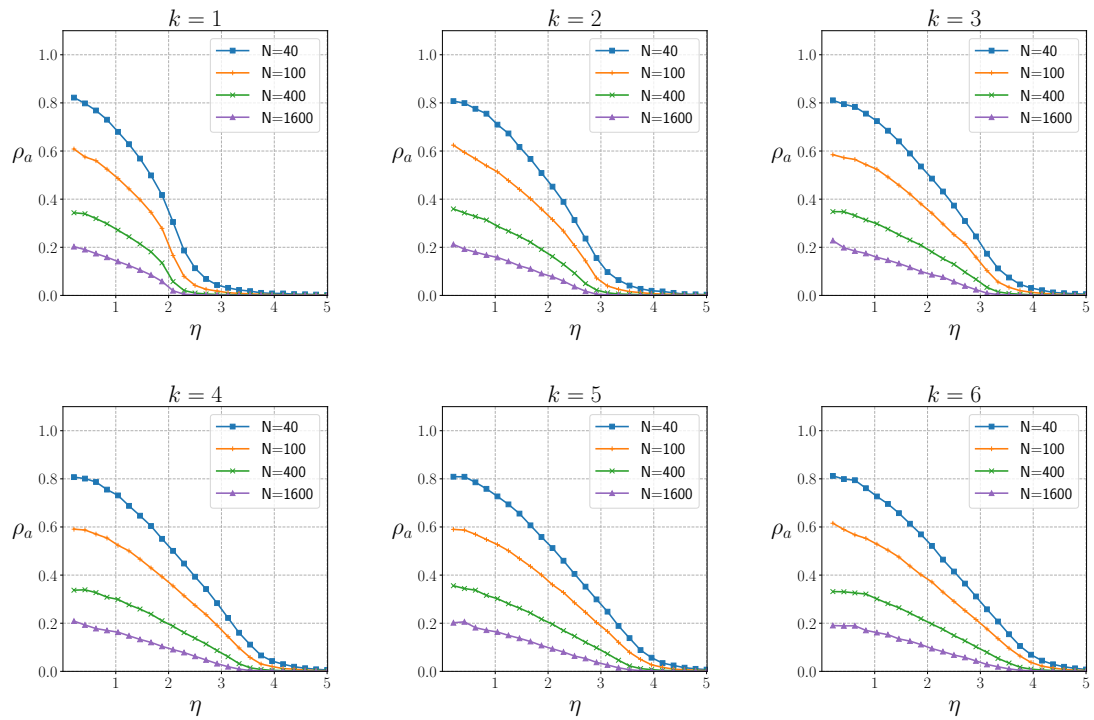


Abbildung 20: Untersuchung des dichte-basierten Ordnungsparameters ρ_a für das stochastische metrisch-topologische Modell. Im Gegensatz zu dem Ordnungsparameter v_a equilibriert das System nie gegen $\rho_a = 1$. Die Systemdichte ist $\rho = 4$ und die Parameter zur Regulierung sind $\alpha = 1$ und $d = r$.

B Illustrationen

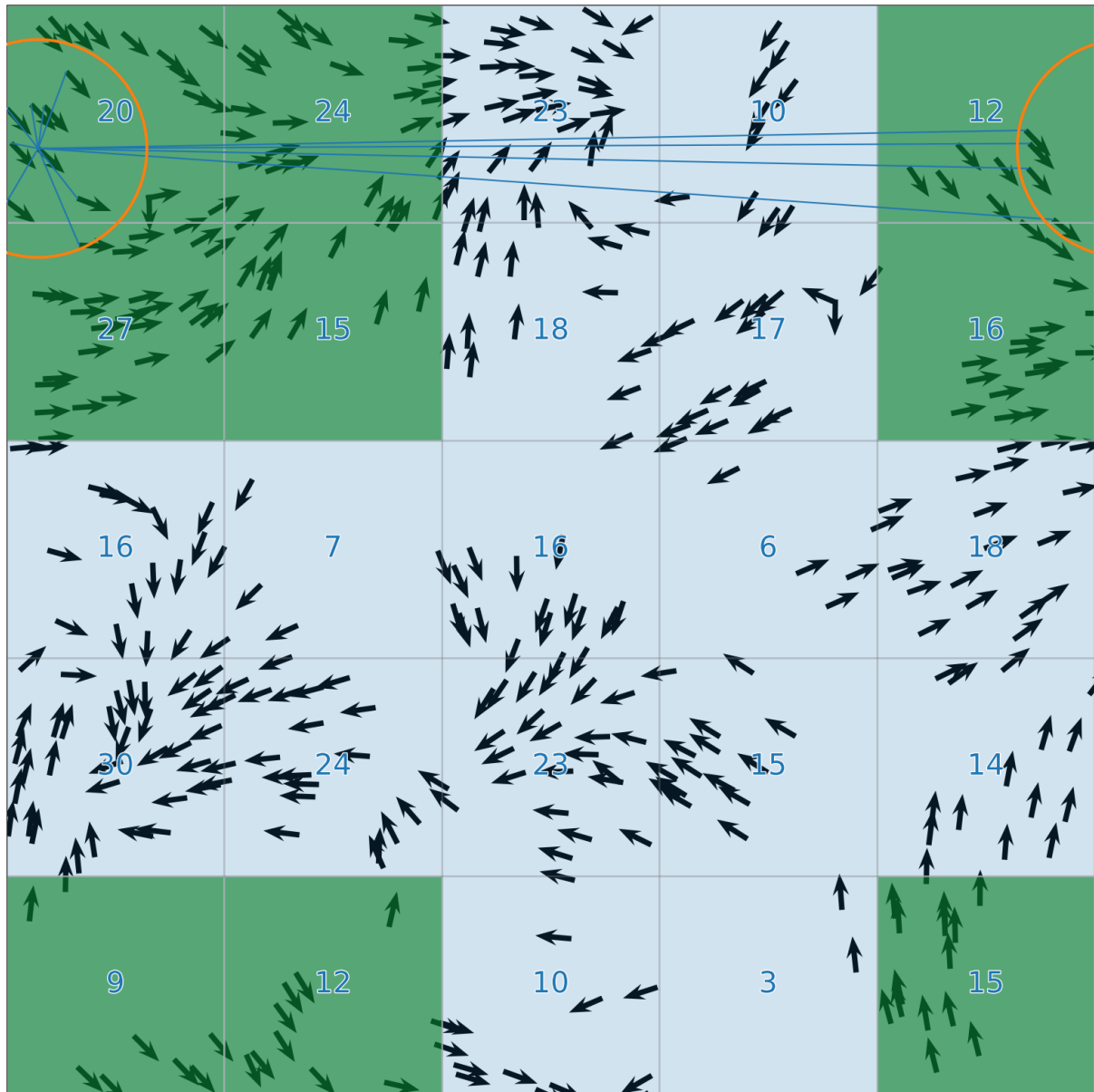


Abbildung 21: Visuelles Debugging durch die Darstellung der wichtigsten Komponenten der Simulation. **Konfiguration:** $N = 400$, $\rho = 4$, $\eta = 0.1$, metrisches Modell, Vicsek-Regel. Die Teilchenposition entspricht der Basis eines Pfeils (unteres Ende).

Simulation mit einem üblichen Randfall. In dieser Visualisierung sind neben den Teilchen selbst auch die Simulationszellen und ihre Teilchenanzahl abgebildet. Es ist exemplarisch für ein Teilchen der Wechselwirkungsradius r eingezeichnet und es wird mit Verbindungslinien kenntlich gemacht, welche Nachbarn bestimmt wurden. Die **grünen Quadrate** geben jene Zellen an, über die zur Ermittlung der Nachbarn iteriert wurde. Es ist erkennbar, dass der Randfall korrekt abgehandelt wird und die Nachbarn auch über den Simulationsbereich hinaus bestimmt werden. Weitere Details sind ausgeblendet, um die Abbildung nicht zu überladen. Aus den zusätzlichen Einstellungen wäre bspw. aber auch erkennbar, dass die Distanzmessungen den Randfall ebenfalls korrekt behandeln und die Distanzen nicht einfach der Länge der Verbindungslinien entsprechen.

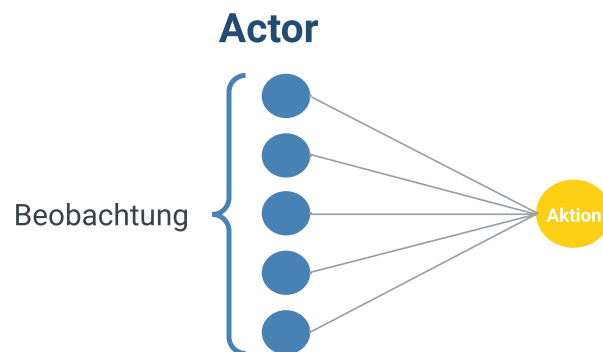


Abbildung 22: Ein Perzeptron mit nur einer Schicht dient als Grundlage für diese Arbeit.

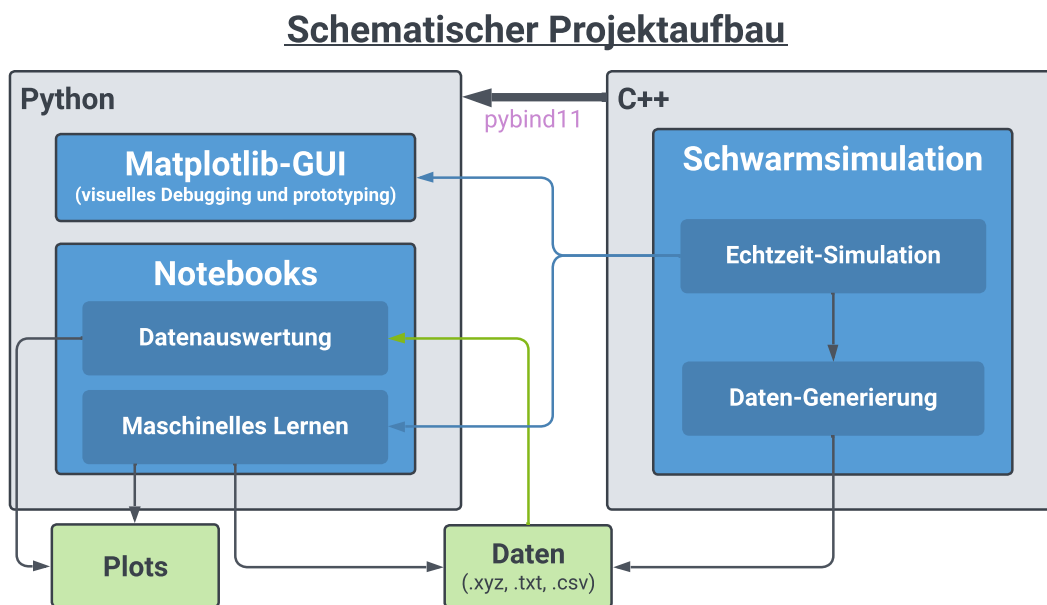


Abbildung 23: Schematischer Projektaufbau; Zusammenspiel zwischen C++ und Python.

Übersicht

Verschiedene Typen von Simulationen, Algorithmen und Konzepte

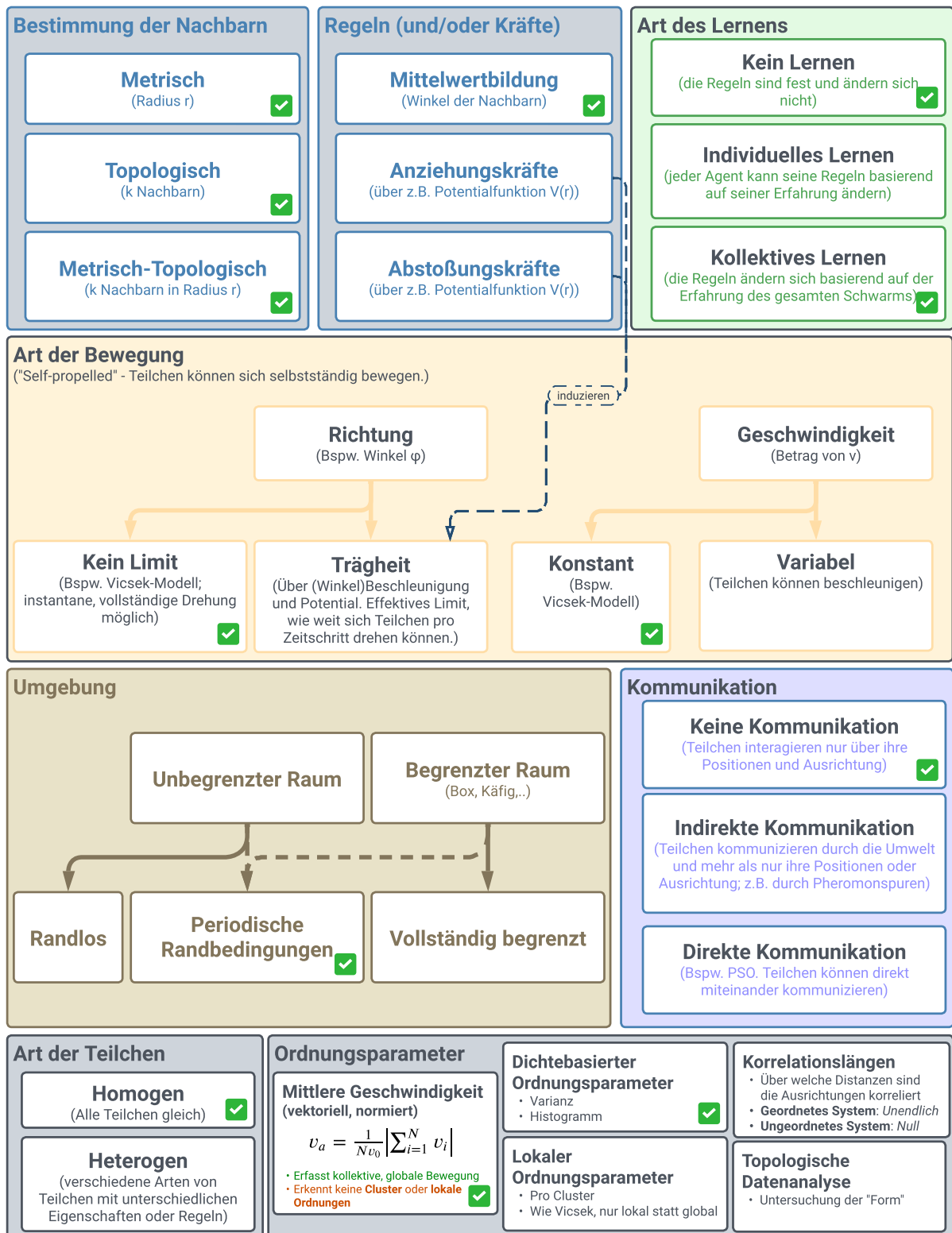


Abbildung 24: Übersicht der verschiedenen Aspekte des Projekts. Diese Zusammenstellung dient als Orientierungshilfe und skizziert die verschiedenen Vertiefungsmöglichkeiten. Die in dieser Arbeit untersuchten bzw. angewandten Aspekte sind mit einem grünen Haken ✓ markiert.



Literatur

- [1] T. Vicsek et al. „Novel Type of Phase Transition in a System of Self-Driven Particles“. In: *Physical Review Letters* **75** (1995), S. 1226–1229. DOI: [10.1103/PhysRevLett.75.1226](https://doi.org/10.1103/PhysRevLett.75.1226).
- [2] F. Cucker und S. Smale. „Emergent Behavior in Flocks“. In: *IEEE Transactions on Automatic Control* **52** (2007), S. 852–862. DOI: [10.1109/TAC.2007.895842](https://doi.org/10.1109/TAC.2007.895842).
- [3] H. Christodoulidi et al. „Phase Transitions in Models of Bird Flocking“. In: *Chaos, Information Processing and Paradoxical Games*. World Scientific, 2014, S. 383–398. DOI: [10.1142/9789814602136_0019](https://doi.org/10.1142/9789814602136_0019).
- [4] P. Rahmani, F. Peruani und P. Romanczuk. „Topological flocking models in spatially heterogeneous environments“. In: *Communications Physics* **4** (2021), S. 206. DOI: [10.1038/s42005-021-00708-y](https://doi.org/10.1038/s42005-021-00708-y).
- [5] M. Aldana et al. „Phase Transitions in Systems of Self-Propelled Agents and Related Network Models“. In: *Physical Review Letters* **98** (2007), S. 095702. DOI: [10.1103/physrevlett.98.095702](https://doi.org/10.1103/physrevlett.98.095702).
- [6] M. Perc et al. „Evolutionary dynamics of group interactions on structured populations: a review“. In: *Journal of The Royal Society Interface* **10** (2013), S. 20120997. DOI: [10.1098/rsif.2012.0997](https://doi.org/10.1098/rsif.2012.0997).
- [7] R. Lukeman, Y.-X. Li und L. Edelstein-Keshet. „Inferring individual rules from collective behavior“. In: *Proceedings of the National Academy of Sciences* **107** (2010), S. 12576–12580. DOI: [10.1073/pnas.1001763107](https://doi.org/10.1073/pnas.1001763107).
- [8] L. H. Cisneros et al. „Fluid dynamics of self-propelled microorganisms, from individuals to concentrated populations“. In: *Experiments in Fluids* **43** (2007), S. 737–753. DOI: [10.1007/s00348-007-0387-y](https://doi.org/10.1007/s00348-007-0387-y).
- [9] F. Schweitzer, W. Ebeling und B. Tilch. „Statistical mechanics of canonical-dissipative systems and applications to swarm dynamics“. In: *Phys. Rev. E* **64** (2001), S. 021110. DOI: [10.1103/PhysRevE.64.021110](https://doi.org/10.1103/PhysRevE.64.021110).
- [10] S.-J. Chung et al. „A Survey on Aerial Swarm Robotics“. In: *IEEE Transactions on Robotics* **34** (2018), S. 837–855. DOI: [10.1109/TR0.2018.2857475](https://doi.org/10.1109/TR0.2018.2857475).
- [11] D. St-Onge et al. „Collective expression: how robotic swarms convey information with group motion“. In: *Paladyn, Journal of Behavioral Robotics* **10** (2019), S. 418–435. DOI: [doi:10.1515/pjbr-2019-0033](https://doi.org/10.1515/pjbr-2019-0033).
- [12] G. Aletti, A. K. Naimzada und G. Naldi. „Mathematics and physics applications in sociodynamics simulation: the case of opinion formation and diffusion“. In: *Mathematical Modeling of Collective Behavior in Socio-Economic and Life Sciences*. Hrsg. von G. Naldi, L. Pareschi und G. Toscani. Boston: Birkhäuser Boston, 2010, S. 203–221. ISBN: 978-0-8176-4946-3. DOI: [10.1007/978-0-8176-4946-3_8](https://doi.org/10.1007/978-0-8176-4946-3_8).
- [13] *Bild eines Starenschwarms*. 2023. URL: <https://thekidshouldseethis.com/post/starling-murmuration-an-enthralling-five-minutes-in-the-netherlands>.
- [14] M. Ballerini et al. „Interaction ruling animal collective behavior depends on topological rather than metric distance: Evidence from a field study“. In: *Proceedings of the National Academy of Sciences* **105** (2008), S. 1232–1237. DOI: [10.1073/pnas.0711437105](https://doi.org/10.1073/pnas.0711437105).
- [15] C. W. Reynolds. „Flocks, herds and schools: A distributed behavioral model“. In: *ACM SIGGRAPH Computer Graphics*. Bd. 21. 4. ACM, 1987, S. 25–34. DOI: [10.1145/37401.37406](https://doi.org/10.1145/37401.37406).
- [16] T. Vicsek und A. Zafeiris. „Collective motion“. In: *Physics Reports* **517** (2012), S. 71–140. DOI: [10.1016/j.physrep.2012.03.004](https://doi.org/10.1016/j.physrep.2012.03.004).
- [17] J. D. Hunter. „Matplotlib: A 2D graphics environment“. In: *Computing in Science & Engineering* **9** (2007), S. 90–95. DOI: [10.1109/MCSE.2007.55](https://doi.org/10.1109/MCSE.2007.55).
- [18] J. L. Elman. „Finding structure in time“. In: *Cognitive science* **14** (1990), S. 179–211. DOI: https://doi.org/10.1207/s15516709cog1402_1.
- [19] A. Vaswani et al. *Attention Is All You Need*. 2023. arXiv: [1706.03762](https://arxiv.org/abs/1706.03762) [cs.CL].
- [20] G. Bonaccorso. *Mastering Machine Learning Algorithms*. Second Edition. Birmingham, United Kingdom: Packt Publishing Ltd, 2020. ISBN: 978-1-83882-029-9.
- [21] L. Canese et al. „Multi-Agent Reinforcement Learning: A Review of Challenges and Applications“. In: *Applied Sciences* **11** (2021), S. 4948. DOI: [10.3390/APP11114948](https://doi.org/10.3390/APP11114948).

- [22] H. Zhang et al. „CityFlow: A Multi-Agent Reinforcement Learning Environment for Large Scale City Traffic Scenario“. In: *Proceedings of the 2019 ACM SIGKDD International Conference on Knowledge Discovery & Data Mining*. 2019, S. 2758–2766. DOI: [10.1145/3308558.3314139](https://doi.org/10.1145/3308558.3314139).
- [23] S. V. Albrecht, F. Christianos und L. Schäfer, Hrsg. *Multi-Agent Reinforcement Learning: Foundations and Modern Approaches*. Pre-Print. Cambridge, Massachusetts: The MIT Press, 2023. URL: <https://www.marl-book.com/>.
- [24] G. Brockman et al. *OpenAI Gym*. 2016. arXiv: [1606.01540](https://arxiv.org/abs/1606.01540) [cs.LG].
- [25] *RLlib: An Open-Source Library for Reinforcement Learning*. URL: <https://github.com/ray-project/ray>.
- [26] *TF-Agents: Reliable and Easy-to-use TensorFlow library for building Reinforcement Learning models*. URL: <https://github.com/tensorflow/agents>.
- [27] J. Schulman et al. *Proximal Policy Optimization Algorithms*. 2017. arXiv: [1707.06347](https://arxiv.org/abs/1707.06347) [cs.LG].
- [28] T. Haarnoja et al. *Soft Actor-Critic: Off-Policy Maximum Entropy Deep Reinforcement Learning with a Stochastic Actor*. 2018. arXiv: [1801.01290](https://arxiv.org/abs/1801.01290) [cs.LG].
- [29] T. P. Lillicrap et al. *Continuous control with deep reinforcement learning*. 2019. arXiv: [1509.02971](https://arxiv.org/abs/1509.02971) [cs.LG].
- [30] C. J. C. H. Watkins und P. Dayan. „Q-learning“. In: *Machine Learning* **8** (1992), S. 279–292. DOI: [10.1007/BF00992698](https://doi.org/10.1007/BF00992698).
- [31] V. Mnih et al. *Playing Atari with Deep Reinforcement Learning*. 2013. arXiv: [1312.5602](https://arxiv.org/abs/1312.5602) [cs.LG].
- [32] I. J. .-. T. C. S. Committee. *ISO International Standard ISO/IEC 14882:2014(E) – Programming Language C++*. 2014. URL: <https://www.iso.org/standard/64029.html>.
- [33] P. S. Foundation. *Python 3.10.0 Documentation*. 2021. URL: <https://docs.python.org/3.10/>.
- [34] W. Jakob et al. *pybind11 – Seamless operability between C++11 and Python*. Available at <https://github.com/pybind/pybind11>. 2016.
- [35] A. Stukowski. „Visualization and analysis of atomistic simulation data with OVITO - the Open Visualization Tool“. In: *Modelling and Simulation in Materials Science and Engineering* **18** (2010). Available at <http://ovito.org/>, S. 015012.
- [36] L. Verlet. „Computer “Experiments” on Classical Fluids. I. Thermodynamical Properties of Lennard-Jones Molecules“. In: *Physical Review Letters* **159** (1967), S. 98. DOI: [10.1103/PhysRev.159.98](https://doi.org/10.1103/PhysRev.159.98).
- [37] R. A. Finkel und J. L. Bentley. „Quad Trees: A Data Structure for Retrieval on Composite Keys“. In: *Acta Informatica* **4** (1974), S. 1–9. DOI: [10.1007/BF00288933](https://doi.org/10.1007/BF00288933).
- [38] E. Ising. „Beitrag zur Theorie des Ferromagnetismus“. In: *Zeitschrift für Physik* **31** (1925), S. 253–258. DOI: [10.1007/BF02980577](https://doi.org/10.1007/BF02980577).



Danksagung

Ich möchte an dieser Stelle die Gelegenheit nutzen, Dank auszusprechen. In erster Linie danke ich Herrn Kierfeld für die Betreuung bei diese Arbeit. Ich danke für die transparente Kommunikation einer realistischen Erwartungshaltung und für das Ernstnehmen dieses Abenteuers. Ich möchte auch meinen Dank an Felix aussprechen, der mir wertvolle Ideen und Feedback hat zukommen lassen. Ich empfand die Gesprächsstunden mit euch stets als entspannt, einladend und geistreich.

Ich danke vor allem auch Lara, die so viel Rücksicht und Herz gezeigt hat, dass vieles erst durch sie möglich wurde.

An dieser Stelle möchte ich auch allen danken, die über die Arbeit drübergeschaut haben; vor allem auch Christian, der sich selbst etwas Zeit an seinem Geburtstag genommen hat, um mich zu unterstützen.

Letztlich möchte ich auch allen danken, die bis hier hin an mich geglaubt und mich unterstützt haben.

Eidesstattliche Versicherung

(Affidavit)

Lehner, Rene-Marcel

193050

Name, Vorname
(surname, first name)

Matrikelnummer
(student ID number)

☒ Bachelorarbeit
(Bachelor's thesis)

☐ Masterarbeit
(Master's thesis)

Titel
(Title)

Analyse und Simulation neuronaler Schwarmdynamik

Ich versichere hiermit an Eides statt, dass ich die vorliegende Abschlussarbeit mit dem oben genannten Titel selbstständig und ohne unzulässige fremde Hilfe erbracht habe. Ich habe keine anderen als die angegebenen Quellen und Hilfsmittel benutzt sowie wörtliche und sinngemäße Zitate kenntlich gemacht. Die Arbeit hat in gleicher oder ähnlicher Form noch keiner Prüfungsbehörde vorgelegen.

I declare in lieu of oath that I have completed the present thesis with the above-mentioned title independently and without any unauthorized assistance. I have not used any other sources or aids than the ones listed and have documented quotations and paraphrases as such. The thesis in its current or similar version has not been submitted to an auditing institution before.

Dortmund, den 19.9.2023

Ort, Datum
(place, date)


Unterschrift
(signature)

Belehrung:

Wer vorsätzlich gegen eine die Täuschung über Prüfungsleistungen betreffende Regelung einer Hochschulprüfungsordnung verstößt, handelt ordnungswidrig. Die Ordnungswidrigkeit kann mit einer Geldbuße von bis zu 50.000,00 € geahndet werden. Zuständige Verwaltungsbehörde für die Verfolgung und Ahndung von Ordnungswidrigkeiten ist der Kanzler/die Kanzlerin der Technischen Universität Dortmund. Im Falle eines mehrfachen oder sonstigen schwerwiegenden Täuschungsversuches kann der Prüfling zudem exmatrikuliert werden. (§ 63 Abs. 5 Hochschulgesetz - HG -).

Die Abgabe einer falschen Versicherung an Eides statt wird mit Freiheitsstrafe bis zu 3 Jahren oder mit Geldstrafe bestraft.

Die Technische Universität Dortmund wird ggf. elektronische Vergleichswerkzeuge (wie z.B. die Software „turnitin“) zur Überprüfung von Ordnungswidrigkeiten in Prüfungsverfahren nutzen.

Die oben stehende Belehrung habe ich zur Kenntnis genommen:

Official notification:

Any person who intentionally breaches any regulation of university examination regulations relating to deception in examination performance is acting improperly. This offense can be punished with a fine of up to EUR 50,000.00. The competent administrative authority for the pursuit and prosecution of offenses of this type is the Chancellor of TU Dortmund University. In the case of multiple or other serious attempts at deception, the examinee can also be unenrolled, Section 63 (5) North Rhine-Westphalia Higher Education Act (*Hochschulgesetz, HG*).

The submission of a false affidavit will be punished with a prison sentence of up to three years or a fine.

As may be necessary, TU Dortmund University will make use of electronic plagiarism-prevention tools (e.g. the "turnitin" service) in order to monitor violations during the examination procedures.

I have taken note of the above official notification:*

Dortmund, den 19.9.2023

Ort, Datum
(place, date)


Unterschrift
(signature)

***Please be aware that solely the German version of the affidavit ("Eidesstattliche Versicherung") for the Bachelor's/ Master's thesis is the official and legally binding version.**