

THE UNIVERSITY OF SYDNEY



DOCTORAL THESIS

**Quantifying criticality, information
dynamics and thermodynamics of
collective motion**

Author:

Emanuele CROSATO

Research supervisor:

Prof. Mikhail PROKOPENKO

Auxiliary supervisors:

Dr. Joseph T. LIZIER

Dr. Richard E. SPINNEY

Dr. X. Rosalind WANG (CSIRO)

*A thesis submitted in fulfillment of the requirements
for the degree of Doctor of Philosophy*

in the

Complex Systems Research Group and Centre for Complex Systems
Faculty of Engineering and IT

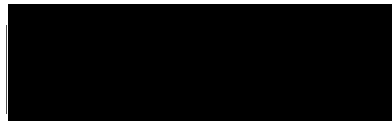
December 14, 2018

Declaration of Authorship

I, Emanuele CROSATO, declare that this thesis titled, “Quantifying criticality, information dynamics and thermodynamics of collective motion” and the work presented in it are my own. I confirm that:

- This work was done wholly or mainly while in candidature for a research degree at this University.
- Where any part of this thesis has previously been submitted for a degree or any other qualification at this University or any other institution, this has been clearly stated.
- Where I have consulted the published work of others, this is always clearly attributed.
- Where I have quoted from the work of others, the source is always given. With the exception of such quotations, this thesis is entirely my own work.
- I have acknowledged all main sources of help.
- Where the thesis is based on work done by myself jointly with others, I have made clear exactly what was done by others and what I have contributed myself.

Signed:

A black rectangular box redacting the signature.

Date:

14th of December, 2018

THE UNIVERSITY OF SYDNEY

Doctor of Philosophy

Quantifying criticality, information dynamics and thermodynamics of collective motion

by Emanuele CROSATO

Complex Systems Research Group and Centre for Complex Systems
Faculty of Engineering and IT

Abstract

Active matter consists of self-propelled particles whose interactions give rise to coherent collective motion. Well-known examples include schools of fish, flocks of birds, swarms of insects and herds of ungulates. On the micro-scale, cells, enzymes and bacteria also move collectively as active matter, inspiring engineering of artificial materials and devices. These diverse systems exhibit similar collective behaviours, including gathering, alignment and quick propagation of perturbations, which emerge from relatively simple local interactions. This phenomenon is known as self-organisation and is observed in active matter as well as in many other complex collective phenomena, including urban agglomeration, financial crises, ecosystems dynamics and technological cascading failures. Some open challenges in the study of self-organisation include (a) how the information processing across the collective and over time gives rise to emergent behaviour, (b) how to identify the regimes in which different collective behaviours exist and their phase transitions, and (c) how to quantify the thermodynamics associated with these phenomena. This thesis aims to investigate these topics in the context of active matter, while building a rigorous theoretical framework. Specifically, this thesis provides three main contributions. Firstly, the question of how to formally measure information transfer across the collective is addressed and applied to a real system, i.e., a school of fish. Secondly, general relations between statistical mechanical and thermodynamical quantities are analytically derived and applied to a model of active matter, resulting in the formulation of the concept of “thermodynamic efficiency of computation during collective motion”. This concept is then extended to the domain of urban dynamics. Thirdly, this thesis provides a rigorous quantification of the non-equilibrium entropy production associated with the collective motion of active Brownian particles.

Acknowledgements

There are many fantastic people I would like to thank for their support throughout my doctoral candidature. Without them, this journey would not have been possible.

First of all, I would like to thank my supervisors for their invaluable expertise, not to mention their patience.

Prof. Mikhail Prokopenko was much more than just a supervisor, he was a mentor, a pillar and a friend. He dedicated a lot of time to me and he provided all the support I needed, both academically and personally. He taught me a lot about complex systems and, most importantly, he taught me how to be a good researcher. I learned enormously from him.

Dr. Joseph T. Lizier offered great guidance, especially at the beginning of my doctoral candidature, when it was more important. He was always available to me and ready to clarify any of my doubts with his immense knowledge of information dynamics, and his emblematic friendliness.

Dr. Richard E. Spinney and I have worked a lot together. His knowledge on thermodynamics is beyond comparison, and I could access it at anytime by simply knocking at his door to ask questions. And Richard has an answer for anything in thermodynamics.

Dr. X. Rosalind Wang has helped me a lot to get started with my doctoral candidature. She contributed to my first research project and inspired it with her pioneering work on information dynamics in swarms.

Special thanks go to all my coauthors, and especially to Dr. Ramil Nigmatullin who exchanged with me several of his interesting thoughts and brilliant ideas on active matter and urban dynamics.

I would also like to thank the rest of my colleagues, especially the other PhD students and the Postdocs. We had such a great time together, thank you for the (way too many) coffee breaks and all the ridiculous chats at the pub.

Throughout my candidature I was supported by The University of Sydney's "Postgraduate Scholarship in the field of Complex Systems" from Faculty of Engineering and IT, and by a CSIRO top-up scholarship. In the last year I was also supported by The University of Sydney's DVC Research Strategic Research Excellence Initiative (SREI-2020) project, "CRISIS: Crisis Response in Interdependent Social-Infrastructure Systems" (IRMA 194163).

I would also like to acknowledge Sydney Informatics Hub at The University of Sydney for providing access to HPC computational resources that have contributed to the research results reported within the thesis.

Lastly, but not less importantly, I would like to thank my parents Elena and Fabio for supporting my education and my brother Riccardo, my girlfriend Dr. Wilhelmina J.M. De Ruiter and all my friends who have always encouraged me during the last three years.

Contents

Declaration of Authorship	iii
Abstract	v
Acknowledgements	vii
1 Introduction	1
1.1 Scope	2
1.2 Objectives	3
1.3 The four studies	4
1.4 Structure of this thesis	7
2 Background and framework	11
2.1 Collective motion	12
2.1.1 Collective motion of animal groups	12
2.1.2 Microscopic collective motion in biological systems	14
2.1.3 Collective motion of non-living systems	16
2.2 Unifying perspective on self-organisation	19
2.2.1 Complex systems and self-organisation	19
2.2.2 Criticality and phase transitions	20
2.2.3 Dynamical models and simulations	21
2.2.4 Statistical mechanics and maximum entropy models	22
2.2.5 Thermodynamics and non-equilibrium systems	23
2.3 The theoretical framework	25
2.3.1 Shannon entropy and information theory	25
2.3.2 Transfer entropy and information dynamics	26
2.3.3 Fisher information and information geometry	30
2.3.4 Entropy production and stochastic thermodynamics	31
3 Informative and misinformative interactions in a school of fish	39
4 Thermodynamics and computation during collective motion	67
5 Critical dynamics and thermodynamics of urban transformations	83
6 Entropy production during collective motion of active Brownian particles	99

7	Discussion and conclusions	111
7.1	A unifying framework for collective motion	112
7.2	Summary of the main results	113
7.3	Future directions	117

Chapter 1

Introduction

1.1 Scope

Collective behaviour is a pervasive phenomenon that can be observed in a wide range of systems. Aggregation in groups, for instance, is common in many animal species, offering several benefits such as shelter, anti-predator vigilance, more efficient foraging strategies, easier access to mates and division of labor [Krause and Ruxton, 2002]. Social and economical systems, which include the formation of cities, markets, institutions and several other organisations, are examples of collective behaviour involving people [Batty, 2013]. This phenomenon is also abundant in the microscopic world, where cells can collectively function as organs. One of the most notable examples is the brain, where billions of neurons form networks that are capable of processing information [Gong and Van Leeuwen, 2009].

Remarkably, complex collective behaviour often *self-organises* from simple interactions among individual components, without requiring any form of centralised control [Mitchell, 2009]. This is a crucial characteristic for understanding collective behaviour in social and biological systems, which also has important implications in the design of new artificial systems [Sayama, 2010; Ulieru and Doursat, 2011; Doursat et al., 2012]. Not surprisingly, self-organisation has excited the curiosity of many researchers in different field of science and engineering, who are attempting to shed light on the underlying mechanisms governing this powerful, decentralised phenomenon [Popkin, 2016]. This thesis is part of this multi-disciplinary effort. Specifically, it is concerned with one of the most striking examples of collective behaviour: the coherent motion of groups of self-propelled particles (e.g., animals, cells, or even artificial devices) that produces dynamically ordered collective (global) structures. This phenomenon, observed in several physical, chemical as well as biological systems, is typically referred to as *collective motion*, while the systems themselves are typically referred to as *active matter*.

In the last three decades, collective motion has become a well-established, cross-disciplinary topic. The study of collective motion has potential implications in many areas [Vicsek and Zafeiris, 2012], ranging from understanding how groups of animals (e.g., birds, fish, insects, etc.) coordinate their movement, to modelling and predicting the migration of cell and bacteria, and even to engineering new materials, nano- and micro-robotics for healthcare, and autonomous vehicles. Across all these diverse areas, a key objective is the development of a rigorous, unifying theory of collective motion that bridges between local interactions and global dynamics. At present, no comprehensive theory has been proposed and several theoretical challenges need to be addressed in order to make important steps forwards in this field.

There is a growing understanding that distributed information processing, and especially information propagation, plays a central role in self-organisation [Couzin, 2007; Prokopenko et al., 2009; Wang et al., 2012; Prokopenko, 2013; Walker and Davies, 2013]. One of the main challenges in collective motion is the quantification of information processing, which is typically stochastic, noisy and continuous,

and therefore hard to detect and measure precisely. The development of dynamical models has helped to reveal important aspects of the self-organisation of collective motion. It has been shown, for example, that in many systems different global motion patterns can emerge, or vanish, according to the variation of some factors affecting the behaviour of the particles. A second, well-known challenge is the construction of the phase space of such collective motion patterns. Non-trivial global dynamics of the system is enabled by particles' self-propulsion, which requires the expenditure of energy [Ramaswamy, 2017]. Systems where energy is constantly supplied and dissipated are said to be thermodynamically out-of-equilibrium, and are notably challenging to study [Seifert, 2012]. The thermodynamical costs of collective motion are becoming a topic of interest, especially when the performance of microscopic synthetic systems are studied [Bechinger et al., 2016].

1.2 Objectives

This thesis aims at advancing the study of collective motion, by providing contributions in specific topics related to the aforementioned challenges. The following are the main objectives of this thesis.

Measuring dynamics of information flow

The role of information processing in the emergence of collective motion has been long investigated. Some researchers have studied the propagation of information in the form of directional changes cascade [Nagy et al., 2010; Cavagna et al., 2013; Attanasi et al., 2014], while other have been identified motion correlations across the collective [Potts, 1984; Procaccini et al., 2011; Herbert-Read et al., 2015]. Despite dealing specifically with the concept of information, these pioneer studies did not quantify it by means of rigorous information-theoretic tools. Instead of directly quantifying the propagation of information, they measured other aspects of the systems, which were hypothesised to imply information propagation. This approach might be misleading, or not reveal the complete picture. One of the objectives of this thesis is to propose an information-theoretic framework for measuring information propagation, and especially its dynamics in time and across the collective.

Relating self-organisation and thermodynamic quantities

Many models of collective motion have been developed, which are capable of reproducing the salient features of the global behaviour of real systems. Typically, these models exhibit kinetic phase transitions, the critical points of which have been localised resulting in detailed phase diagrams [Vicsek et al., 1995; Chaté et al., 2008]. However, the physical meaning of 'varying the parameters' that drive the systems across a transition has not been much investigated. One of the objectives of this thesis is to propose a framework for characterising the variation of parameters in terms of thermodynamical concepts, such as the required work and the variation of

internal energy of the system. This is expected to have implication in the study and design physical, rather than just simulated, active matter systems.

Quantifying entropy production in active matter

A crucial feature of active matter, without which most complex behaviour would not be attainable, is to be thermodynamically out-of-equilibrium. Despite being a well-known feature [Marchetti et al., 2013; Bechinger et al., 2016; Ramaswamy, 2017], the non-equilibrium character of active matter lacks a precise, thermodynamic description. Non-equilibrium thermodynamics has always been challenging, but new approaches have been developed in the last three decades [Seifert, 2012]. Recently, few attempts to measure the entropy production of active matter have been made [Fodor et al., 2016; Nardini et al., 2017; Mandal et al., 2017], but conflicting results have been obtained and several important aspects pointed out [Shankar and Marchetti, 2018]. A main objective of this thesis is to further develop the thermodynamic description of collective motion in active matter, with particular focus on phase transitions and critical regimes.

Extension to other collective phenomena

Collective motion shares many of its aspects with other collective phenomena. Many ideas utilised to study collective motion are adopted from wider disciplines, such as complex systems science and statistical mechanics, and have been applied to study similar phenomena. A secondary objective is to explore to what extent the theoretical frameworks developed within this thesis can be utilised to study other collective phenomena.

The central thesis, thus, can be summarised as follows. The role of information processing in shaping self-organisation of collective motion is undoubtedly related to the underlying physical fluxes. These fluxes typically differ across distinct phases and critical regimes of collective dynamics exhibited by active matter. We aim, therefore, to explicitly characterise and quantify the information dynamics and the underlying thermodynamics that reveal the key phase transitions observed in self-organising active matter. This should allow us to compare different scenarios in terms of salient information-theoretic and thermodynamic quantities, as well as in terms of their thermodynamic efficiency, contributing to the formation of a unifying framework.

1.3 The four studies

The overall objectives of this project have been addressed in four self-contained studies which build up and exemplify a unifying framework. Each study is presented separately within this thesis, as four distinct research articles. Three of these studies directly investigate collective motion phenomena, with focus on some of the most

important problems in the field. Another study applies some of the concepts developed in this thesis for collective motion to a different kind of collective phenomena, i.e., *urban dynamics*. Additionally, this thesis discusses the connections between the four studies, aiming at proposing (elements of) a theoretical framework for studying collective phenomena. What follows is a brief overview of the four studies.

Informative and misinformative interactions in a school of fish

This first study focusses on the well-known issue of quantifying the distributed information processing underlying collective dynamics of real systems. In particular, this study investigates the dynamics of information transfer in a school of fish undergoing collective directional changes in a controlled environment.

More precisely, the *transfer entropy*, an information-theoretic measure, is used in order to quantify predictive information flows across the school, quantifying how much the information contained in the directional changes of one fish can be used to predict the next move of another fish. Such predictive information flows are measured *locally* in space and time, revealing both positive and negative information transfer during the collective moves. To our knowledge, the existence of negative (or misinformative) flows during collective motion has never been hypothesised, and in this study they are detected for the first time. This highlights the effectiveness of transfer entropy and, in general, of the information dynamics framework in disclosing intricacies of distributed information processing that would otherwise remain ignored if naive techniques, e.g., simple correlation analyses, were to be used.

The study also identifies spatial patterns associated with the information and misinformation flows. The identification of motives of spatial interactions underlying the propagation of directional changes is an important problem in the field of collective motion, which has been approached by several researchers but has been lacking a general solution.

This study [Crosato et al., 2018b] was published in the journal *Swarm Intelligence*, and is presented in this thesis in Chap. 3.

Thermodynamics and computation during collective motion

This study investigates kinetic phase transitions, well-known phenomena in collective motion that can be observed in several systems of self-propelled particles. Rather than observing a biological system, this second study systematically examines the behaviour of a widely used dynamical model of collective motion.

The model is known to exhibit a first order phase transition: when the alignment strength between particles increases beyond a critical point, the system self-organises into a polarised configuration, in which particles face a common average direction. The aim of the study is to analyse such phase transition as a thermodynamic phenomenon, in which ordered motion is achieved at the expense of work done on the system and in presence of energy changes.

While thermodynamical quantities can well describe physical properties of the system, computational aspects of the system are better captured by information-theoretic quantities. Crucially, this work relates changes in thermodynamic quantities, i.e., work and internal energy, to changes in information-theoretic quantities, i.e., the configuration entropy and the Fisher information. The identified relations are used to propose and formulate the concept of “thermodynamic efficiency of computation”, as the ratio of reduction of entropy to the expenditure of work. The relationships between information processing and its thermodynamic costs are crucial for complex systems, which often need to perform their computation efficiently [Lizier et al., 2012; Kempes et al., 2017; Kolchinsky and Wolpert, 2017].

This study [Crosato et al., 2018d] was published in the journal *Physical Review E*, and is presented in this thesis in Chap. 4.

Critical dynamics and thermodynamics of urban transformations

This study applies the methods developed in Chap. 4 to a different area: urban dynamics. Specifically, this work presents a statistical mechanical model of the income flow within the population of Greater Sydney, given constraints of the daily commute of people between their residence suburbs to the employment areas, as well as data on rent and services available per suburb.

The study identifies a phase transition between dispersed and polycentric phases, which is induced by the variation of a parameter controlling how much value people attribute to suburbs, with respect to their ‘attractiveness’. Critical regimes are characterised thermodynamically by use of the techniques developed in the second study (Chap. 4). In particular, the adaptation of the concept of thermodynamics of computation leads to the definition of an analogous concept: “the thermodynamic efficiency of urban transformations”.

This study [Crosato et al., 2018c] was published in the journal *Royal Society Open Science*, and is presented in this thesis in Chap. 5.

Entropy production during collective motion of active Brownian particles

Recently, researchers have begun to investigate the non-equilibrium character of active matter from the point of view of statistical thermodynamics. Initial attempts to quantify the entropy production during the collective motion of self-propelled particles have only focussed on systems described by over-damped dynamics [Fodor et al., 2016; Mandal et al., 2017], missing crucial ‘hidden’ components of the entropy production [Shankar and Marchetti, 2018]. Moreover, these studies have only considered particles with excluded volume interactions, while effect of alignment interactions on the entropy production has not been studied.

This study provides a quantification of the entropy production in a system of active Brownian particles (ABPs) described by under-damped dynamics, where particles interact via volume exclusion as well as alignment. The complete phase diagram

of the system with respect to parameters determining the persistence of the particles' motion and the strength of the alignment is built, identifying a disordered phase, a phase characterised by phase separation, and a phase exhibiting both alignment and phase separation. In order to characterise each of the identified phases thermodynamically, the expected entropy production rate of the system at (non-equilibrium) steady-state is quantified at different values of control parameters. Moreover, the entropy production rate attributed to individual particles is contrasted to its position and orientation within the group, revealing distinct entropy production signatures for different active matter states, e.g., 'solid', 'gas' and oriented 'flocks'.

This study [Crosato et al., 2018a] is presented in Chap. 6 and is ready for submission to a peer-reviewed journal.

1.4 Structure of this thesis

The remainder of this thesis is organised as follows.

Chap. 2 provides the background for this project and describes the theoretical framework. Firstly, the concept of collective motion is introduced in Sec. 2.1, highlighting the main topics and providing some of the most important literature. This is followed by the presentation, in Sec. 2.2, of the perspective adopted in this thesis to study collective motion. This includes the introduction of concepts such as complex systems science, self-organisation, criticality, dynamical models, statistical mechanics and thermodynamics. Finally, the theoretical framework used in this thesis is described in 2.3 and includes notions of information theory, information dynamics, information geometry as well as stochastic thermodynamics.

The research articles are provided in Chap. 3 to 6. Specifically, the study "Informative and misinformative interactions in a school of fish" is provided in Chap. 3, the study "Thermodynamics and computation during collective motion near criticality" is provided in Chap. 4, the study "On critical dynamics and thermodynamic efficiency of urban transformations" is provided in Chap. 5, and the study "Entropy production during collective motion of active Brownian particles undergoing phase separation and alignment" is provided in Chap. 6.

Chap. 7 concludes the thesis, discussing the overall theoretical approach composed of the four studies and their interconnections.

References

- Attanasi, A. et al. (2014). "Information transfer and behavioural inertia in starling flocks". In: *Nature Physics* 10.9, pp. 691–696.
- Batty, M. (2013). *The new science of cities*. Mit Press.
- Bechinger, C. et al. (2016). "Active particles in complex and crowded environments". In: *Reviews of Modern Physics* 88.4, p. 045006.

- Cavagna, A. et al. (2013). “Diffusion of individual birds in starling flocks”. In: *Proceedings of the Royal Society of London B: Biological Sciences* 280.1756, p. 20122484.
- Chaté, H. et al. (2008). “Modeling collective motion: variations on the Vicsek model”. In: *The European Physical Journal B* 64.3-4, pp. 451–456.
- Couzin, I. (2007). “Collective minds”. In: *Nature* 445.7129, p. 715.
- Crosato, E., M. Prokopenko, and R. E. Spinney (2018a). “Entropy production during collective motion of active Brownian particles undergoing phase separation and alignment”. In: *arXiv:1812.08527 [cond-mat.stat-mech]*.
- Crosato, E. et al. (2018b). “Informative and misinformative interactions in a school of fish”. In: *Swarm Intelligence* 12.4, pp. 283–305.
- Crosato, E., R. Nigmatullin, and M. Prokopenko (2018c). “On critical dynamics and thermodynamic efficiency of urban transformations”. In: *Royal Society Open Science* 5.10.
- Crosato, E. et al. (2018d). “Thermodynamics and computation during collective motion near criticality”. In: *Physical Review E* 97.1, p. 012120.
- Doursat, R., H. Sayama, and O. Michel (2012). *Morphogenetic engineering: toward programmable complex systems*. Springer.
- Fodor, É. et al. (2016). “How far from equilibrium is active matter?” In: *Physical Review Letters* 117.3, p. 038103.
- Gong, P. and C. Van Leeuwen (2009). “Distributed dynamical computation in neural circuits with propagating coherent activity patterns”. In: *PLOS Computational Biology* 5.12, e1000611.
- Herbert-Read, J. E. et al. (2015). “Initiation and spread of escape waves within animal groups”. In: *Royal Society Open Science* 2.4, p. 140355.
- Kempes, C. P. et al. (2017). “The thermodynamic efficiency of computations made in cells across the range of life”. In: *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences* 375.2109, p. 20160343.
- Kolchinsky, A. and D. H. Wolpert (2017). “Dependence of dissipation on the initial distribution over states”. In: *Journal of Statistical Mechanics: Theory and Experiment* 2017.8, p. 083202.
- Krause, J. and G. D. Ruxton (2002). *Living in groups*. Oxford University Press.
- Lizier, J. T., M. Prokopenko, and A. Y. Zomaya (2012). “Coherent information structure in complex computation”. In: *Theory in Biosciences* 131.3, pp. 193–203.
- Mandal, D., K. Klymko, and M. R. DeWeese (2017). “Entropy production and fluctuation theorems for active matter”. In: *Physical Review Letters* 119.25, p. 258001.
- Marchetti, M. C. et al. (2013). “Hydrodynamics of soft active matter”. In: *Reviews of Modern Physics* 85.3, p. 1143.
- Mitchell, M. (2009). *Complexity: a guided tour*. Oxford University Press.
- Nagy, M. et al. (2010). “Hierarchical group dynamics in pigeon flocks”. In: *Nature* 464.7290, pp. 890–893.

- Nardini, C. et al. (2017). "Entropy production in field theories without time-reversal symmetry: quantifying the non-equilibrium character of active matter". In: *Physical Review X* 7.2, p. 021007.
- Popkin, G. (2016). "The physics of life". In: *Nature News* 529.7584, p. 16.
- Potts, W. K. (1984). "The chorus-line hypothesis of manoeuvre coordination in avian flocks". In: *Nature* 309.5966, pp. 344–345.
- Procaccini, A. et al. (2011). "Propagating waves in starling, *Sturnus vulgaris*, flocks under predation". In: *Animal Behaviour* 82.4, pp. 759–765.
- Prokopenko, M. (2013). *Guided self-organization: Inception*. Vol. 9. Springer Science & Business Media.
- Prokopenko, M., F. Boschetti, and A. J. Ryan (2009). "An information-theoretic primer on complexity, self-organization, and emergence". In: *Complexity* 15.1, pp. 11–28.
- Ramaswamy, S. (2017). "Active matter". In: *Journal of Statistical Mechanics: Theory and Experiment* 2017.5, p. 054002.
- Sayama, H. (2010). "Robust morphogenesis of robotic swarms". In: *IEEE Computational Intelligence Magazine* 5.3, pp. 43–49.
- Seifert, U. (2012). "Stochastic thermodynamics, fluctuation theorems and molecular machines". In: *Reports on Progress in Physics* 75.126001, p. 126001.
- Shankar, S. and M. C. Marchetti (2018). "Hidden entropy production and work fluctuations in an ideal active gas". In: *Physical Review E* 98.2, p. 020604.
- Ulieru, M. and R. Doursat (2011). "Emergent engineering: a radical paradigm shift". In: *International Journal of Autonomous and Adaptive Communications Systems* 4.1, p. 39.
- Vicsek, T. and A. Zafeiris (2012). "Collective motion". In: *Physics Reports* 517.3-4, pp. 71–140.
- Vicsek, T. et al. (1995). "Novel type of phase transition in a system of self-driven particles". In: *Physical Review Letters* 75.6, p. 1226.
- Walker, S. I. and P. C. Davies (2013). "The algorithmic origins of life". In: *Journal of the Royal Society Interface* 10.79, p. 20120869.
- Wang, X. R. et al. (2012). "Quantifying and tracing information cascades in swarms". In: *PLOS One* 7.7, e40084.

Chapter 2

Background and framework

2.1 Collective motion

Collective motion is the coherent movement of many self-propelled particles [Vicsek and Zafeiris, 2012]. Self-propulsion makes the particles ‘active’, indicating that their motion is self-driven rather than externally imposed. From this comes the term *active matter* [Marchetti et al., 2013; Bechinger et al., 2016; Ramaswamy, 2017], which is typically used to refer to systems of self-propelled particles. Collective motion is one of the most recognisable features of life, with examples that include animal moving in groups, bacterias forming colonies, migrating cells and many others. However, this phenomenon is not limited to biological systems. Collective motion is also observed in many physical and chemical systems, and it can be engineered in artificial devices and materials. The study of collective motion has indeed established as a growing multidisciplinary area [Popkin, 2016].

Researchers from different fields of science and engineering have investigated this intriguing phenomenon, and have raised several interesting questions. How do the particles self-propel? How do they interact? Do similar self-propulsion mechanisms and/or interactions correspond to analogous global patterns of motion, such as orientational order, swirls and clustering? How does decision making among particles work, and can we predict it? When collective motion doesn’t happen? Are there specific conditions, such as particular density or energy levels, that are necessary for collective motion? What are the thermodynamic costs of this phenomenon? The next sections review some of the most important studies of collective motion, highlighting the main research questions and experimental observations. The studies are presented in three categories depending on the nature of systems: (1) animal groups, including humans, (2) microscopic biological systems and (3) artificial, non-living systems.

2.1.1 Collective motion of animal groups

Aggregations of animals are probably the most eye-catching examples of active matter. Well-known examples are flocks of birds [Lissaman and Shollenberger, 1970; May, 1979], schools of fish [Pitcher, 1983; Parrish et al., 2002], swarms of insects [Gotwald Jr, 1995; Buhl et al., 2006; Fourcassié et al., 2010], herds of ungulates [Ginelli et al., 2015] as well as human crowds [Moussaïd et al., 2011].

Many aspects of the aggregation and collective motion of animal groups have been studied. For instance, the requirements for the formation of collective motion, e.g., in terms of the minimum amount of individuals or of the minimum density, have been investigated. Beekman et al., 2001 has proposed a model relating the number of ants walking to a food source and the size of the colony. The model predicted that the formation of trail-based foraging requires the colony size to exceed a critical threshold. High density was also observed to be crucial for the formation of vortex motions in swarms of *Daphnia* [Ordemann et al., 2003]. Individual *Daphnia* can develop a circular motion, independently from the other individuals. At high

densities, however, if a direction is on average more pronounced, the water drag generates a positive feedback that cause more and more *Daphnia* to move in the same direction. Developing models that can predict critical thresholds for the formation of collective motion is fundamental for understanding, and predicting, this phenomenon. The identification of critical thresholds for the formation of collective motion is an important aspect of the studies in Chap. 4 and Chap. 6.

Substantial interest has been given to the study of the propagation of directional motion across groups, which many authors refer to as ‘information flow’. Nagy et al., 2010 tracked the trajectories of pigeons flying in flocks using GPS devices, and analysed them using a set of correlation functions. Their findings revealed time delayed correlations in the trajectories of pigeons pairs, which suggest a clear dominance hierarchy between birds. The authors also pointed out that birds respond more quickly to other birds on their left, and speculated that information may be preferentially processed by the left-eye/right-hemispheric system. The propagation of direction changes in large flocks of starlings has been studied by Cavagna et al., 2013b, who observed a dynamical information flow from the boundary to the centre of the group, which they also measured in terms of correlation across the birds directions. Birds can sometimes modify their position within the group, therefore changing the individuals with which they directly interact. Interestingly, it was shown that this increases the efficiency at which information is propagated throughout the group [Cavagna et al., 2013a]. It has also been observed that small perturbations can propagate in a wave-like manner [Potts, 1984; Herbert-Read et al., 2015]. Procaccini et al., 2011 studied flocks of starlings under predation by peregrine falcons, and found that waves propagate away from the position at a velocity higher than the velocity of the flock.

Many researchers have attempted to infer physical interaction rules associated with the propagation of information. Katz et al., 2011 used experimental data on schools of golden shiners to identify mean reaction of fish to the position and velocity of its neighbours. Their findings suggest that the speeding force depends on the front-back distance among fish, while the turning force depends on their distance to the side. The work by Gautrais et al., 2012 suggests that fish turning speed is affected by both position and orientation of neighbouring individuals. The role of spatial position has been investigated in individual-to-individual interactions among Mormon crickets [Bazazi et al., 2010]. It was shown that crickets are more likely to approach other stationary individuals that are side-on to the motion direction. Other studies have attempted to reveal interaction networks within groups. Nagy et al., 2013 investigated leader-follower relations in pigeons, and revealed hierarchically structured networks of directed interactions. The authors observed that such leadership hierarchy is distinct from the dominance hierarchy. Rosenthal et al., 2015 tracked the positions and body postures of a school of approximately 150 fish and, by calculating their visual fields, they were able to infer a functional mapping between sensory input motor response. The study found that the networks by which information

propagates are complex, weighted and directed.

These excellent studies laid the groundwork for considerations of the underlying mechanisms that interpret collective motion in terms of information processing. However, these works do not formally define information propagation, which is simply treated as an implicit aspect of the propagation of directional motion. The description of information propagation according to a rigorous information-theoretic framework is very important, since the relation between information transfer and directional correlations might be more intricate than expected. Do strong correlations imply more information transfer? Can the transfer be misinformative, meaning that observing an individual makes the prediction of the next move of another individual even less predictable? Are there characteristic motifs of spatial conditions, such as particular relative position or orientation, and information/misinformation transfer? Wang et al., 2012 were the first to propose an information-theoretic approach for the characterisation of information cascades and collective memory within swarms. Interestingly, the authors revealed that spatial position is related to its information processing role. Information flow was also recently investigated information-theoretically in bats [Orange and Abaid, 2015], soldier crabs [Tomaru et al., 2016], fish [Butail et al., 2014; Hu et al., 2015; Ward et al., 2018] and insects [Lord et al., 2016]. The direction suggested by Wang et al., 2012 is expanded in the study presented in this thesis in Chap. 3.

Another interesting aspect of the collective motion of animal groups is decision making, and in particular the phenomena of leadership and quorum sensing. For example, Ward et al., 2008 have observed that fish perform particular behaviour when they receive stimuli from a minimum number of other individuals. In human crowds, it was found that a small informed minority can guide a large group of uninformed individuals to a destination, without using verbal communication or signalling [Dyer et al., 2008; Dyer et al., 2009]. Interestingly, the study showed that when conflicting directional information was provided to members of the group, the time needed to reach the destination did not increase much, indicating the presence of very efficient decision making mechanisms.

The study of decision making in human crowds, as well as in many other groups, can also benefit from the development of an information-theoretic framework for the quantification of information transfer. Such approach, for example, would be capable of quantify and compare the amount of information transfer across individuals, as well as distinguishing misinformative interactions, which might be relevant for decision making.

2.1.2 *Microscopic collective motion in biological systems*

Collective motion is an important aspect of cells migration, a phenomenon that is observed, during morphogenesis, wound healing and tumor dissemination [Friedl et al., 2004; Rørth, 2007]. The cell-to-cell interactions that enable coherent movement typically involve chemical signals and adhesion. For instance, cells adhesion can

group dispersed cells together, while cell-cell communication can produce front-rear asymmetry that lead to alignment. Collectively migrating cells can be arranged in two-dimensional sheets, e.g., an epithelial layer migrating across substrate, as well as three-dimensional formations, e.g., neural crest migration [Friedl et al., 2004].

Haas and Gilmour, 2006 studied the collective migration of cells during the morphogenesis of zebrafish, showing that a small number of guidance cells are capable of generating polarised chemotaxis effects on a large number of mutant cells, driving them towards right directions for tissue formation. Szabo et al., 2006 investigated the collective migration of keratocytes cells and identified a transition, driven by the variation of the cells density, from a disordered phase with no clear global orientation to a phase characterised by directed motion. The authors observed the emergence of complex spatial structure which include clusters and whirls. In vitro experiments involving hundreds of thousands cells have shown that increasing cells' directional persistence (e.g., by using an activity inhibitor) can results in faster cell segregation and more extensive patterns formation [Méhes et al., 2012].

At high densities, bacteria are also known to give rise to collective motion, typically induced by chemotaxis, hydrodynamic effects and excluded volume interactions, and to form many patterns and structures [Fujikawa and Matsushita, 1989]. Sokolov et al., 2009 studied the formation of convective motion of suspensions of swimming aerobic bacteria in films of adjustable thickness. Their experimental observations show that when the thickness of the film exceeds a critical threshold, bacterial switch from quasi-two-dimensional collective swimming to three-dimensional turbulent motion. Colonies of gliding bacterial cells confined to a monolayer have been shown to form large moving clusters at high packing fraction [Peruani et al., 2012]. Bacteria within the clusters were observed to align and to arrange in a head-to-tail manner, therefore giving rise to polarised collective motion. It was also observed that two clusters typically fuse after a collision, while splitting rarely occurs. Enzymes can also exhibit collective motion. Recently, for example, enzymes have been shown to assemble into metabolons when they participate in reaction cascades, as a result of chemotaxis [Zhao et al., 2018].

All these microscopic systems are very interesting from the perspective of stochastic thermodynamics, the branch of statistical mechanics that is concerned with the study of the thermodynamics of small systems. Researchers in this field are interested in understanding the dynamics of quantities such as energy, work and heat at the level of individual fluctuating trajectories [Sekimoto, 2010; Jarzynski, 2011; Seifert, 2012]. Under the framework of statistical thermodynamics, recent work has investigated the irreversibility of the trajectories of individual particles, and calculated the associated entropy production rate [Shankar and Marchetti, 2018]. But how does collective motion affect the entropy production in systems composed of many interacting particles? Some studies have begun to investigate the entropy production in simulated systems where particles interact via excluded volume [Fodor et al.,

2016; Mandal et al., 2017], and whose behaviour closely resembles the aforementioned systems of cells and bacteria. A thermodynamic description of these systems is crucial for understanding their energy dissipation and, ultimately, their functioning and limitations. Stimulated by the studies of Fodor et al., 2016 and Mandal et al., 2017, this thesis provides an analysis of the entropy production during collective motion which also includes alignment interactions (see Chap 6).

2.1.3 *Collective motion of non-living systems*

Inspired by the functionalities of cells and tissues, researchers are now trying to create new materials made of artificial active components. Engineered active matter provides hope to address several challenges in health care, holding the promise of performing tasks such as transport, sensing, and drug delivery [Bechinger et al., 2016]. Additionally, simplified and controlled in vitro systems provide a very useful tool for studying more complex biological systems [Needleman and Dogic, 2017].

Cells' self-propulsion is determined by the motion of thousands of nanoscopic protein-based machine in the cytoskeleton, called molecular motors, that transform chemical energy into mechanical motion. Artificial self-propelled particles can be assembled out of component of the cytoskeleton [Bechinger et al., 2016], and have been observed to collectively produce dynamic structures. For example, a system of molecular motors attached to microtubules has been constructed [Ndlec et al., 1997; Nédélec et al., 2001]. In presence of ATP, the motors move towards one of the ends of the microtubule, forming a dynamic crosslinks between microtubules. When confined in glass chambers, the microtubules are observed to form complex structures, such as asters and vortices. Moreover, in an unconfined geometry, asters and vortices can organise into larger-scale patterns. At low concentrations of motors, the authors observed the formation of a lattice of vortices, while for slightly higher concentrations they observed a lattices of asters. At even higher concentrations, the microtubules are shown to form bundles. Similarly, Köhler et al., 2011 showed that a synthetic active gel made of actin filaments and molecular motors exhibits structure formation, characterised by a broad distribution of cluster sizes.

Other active particles have been created entirely from synthesised components. These includes the well-known Janus swimmers, particles with two faces coated with thin layers of different catalytic materials [Bechinger et al., 2016]. Temperature and chemical gradients can form along the particle, due to the different coating, leading to self-phoretic motion. Janus particles with no alignment interactions have been experimentally observed to phase separate at high self-propulsion speed and density [Ibele et al., 2009; Theurkauff et al., 2012; Buttinoni et al., 2013], exclusively because of excluded volume interactions. The systems divide into low density areas, where particles are in a dilute gas phase, and high density areas, where particles are in a dense liquid/solid phase. Bricard et al., 2013 demonstrated that hydrodynamic interactions among particles can cause the particles to align, promoting the formation of collective motion. Specifically, the authors showed the transition to polarised

motion in a system of millions of particles, triggered by increasing the density above a critical value. Palacci et al., 2013 showed that a systems of light activated colloidal particles with phoretic attraction can form ‘living crystals’, which actively translate, rotate, join and split.

Blair et al., 2003 showed that vibrated granular rods, i.e., elongated particles driven into motion by external vibration, can display nematic order and swirling. When the rods are vibrated vertically inside a container, they are set into motion due to collisions with the bottom boundary. It is appropriate to regard these particles as self-propelling since the direction of motion is determined by the orientation of the particles, rather than by an external field. For sufficiently large packing fractions, the particles form tend to align vertically and form vortex motion. When the container is vibrated horizontally, the rods tend to align vertically, but vortices are not observed. It was also shown that the shape of the rods influence the formation of global structures. When vibrated vertically in a quasi-two-dimensional confinements, tapered particles tend to form nematic order, while cylindrical particles with flat tips exhibit tetratic order[Narayan et al., 2006]. Clear non-equilibrium phenomena were observed, including the persistent rotation of the formed structures. Kumar et al., 2014 studied a system of millimetre-sized tapered rods in a medium of spherical, non-motile beads in contact with a vibrated surface. Such medium enhances hydrodynamic interactions among the tapered rods. At high concentrations of beads, the hydrodynamic interactions trigger the formation of aligned motion of the rods, which results in their circular motion around the container.

In all these systems, out-of-equilibrium structures form, and are sustained, by the particles’ self-propulsion, which requires the constant injection of energy. In the system of molecular motors and microtubules [Ndlec et al., 1997; Nédélec et al., 2001], for example, particles get energy from ATP, which is supplied in the environment. In the work by Palacci et al., 2013, colloidal particles were set into motion by light, while in case of rods, energy is supplied to the system in the form of vibrations. The precise quantification of energy and work is crucial for the design of artificial active matter. However, the behaviour of active matter is typically better explained in terms of information processing. It would, therefore, be very useful to relate thermodynamical quantities to computational aspects of collective motion. This thesis, for example, investigates the thermodynamic efficiency of the formation of collective motion, i.e., the ratio of the order gained within the system, measured information-theoretically, to the associated expenditure of physical work (see Chap. 4).

On a much larger scale, groups of robots can be engineered to perform collective tasks. When such tasks are complex, distributed systems can be more effective, and sometimes much easier to design, than centralised solutions [Sayama, 2010; Uliuru and Doursat, 2011; Doursat et al., 2012]. The design and study of systems composed of a large number of (typically simple) robots that perform tasks in a decentralised fashion is called *swarm robotics*. Systems of terrestrial and aerial robots that exhibit

collective motion have been engineered [Vicsek and Zafeiris, 2012], and have potential applications in autonomous exploration and surveillance. Inspired by the collective behaviour of insects, swarms of robots have been designed that can collectively build structures and use stigmergy to cooperatively find a path to a resource [Bonabeau et al., 1999]. The theoretical frameworks provided in this thesis can potentially be applied to the design of robotic systems, offering, for instance, methods to quantify the dynamics of distributed information processing across robots information-theoretically (see Chap 3).

2.2 Unifying perspective on self-organisation

In this thesis, the study of collective motion in active matter is approached from the perspective of complex systems science. This discipline is concerned with the study of systems made of a large number of similar components, whose interactions result in non-trivial global behaviour that typically involves intricate non-linear dynamics [Sayama, 2015]. Interest in complex systems stems from their abundance and from some intriguing characteristics that these systems possess, including robustness, adaptability and scalability. These features are desired in many natural and artificial systems, and thus their understanding is highly valuable for both science and engineering [Mitchell, 2009].

2.2.1 Complex systems and self-organisation

A key aspect of complex systems is *self-organisation*. According to Haken and Ju-marie, 2006, a system is self-organising if it obtains spatial, temporal, or functional structure in absence of specific external interference, i.e., the structure or functioning is not impressed on the system; rather, the system is acted upon from the outside in a non-specific fashion. Bonabeau et al., 1997 and Camazine et al., 2003 describe self-organisation as the process through which global patterns emerge in a system solely from the interactions among its components. The authors also highlight that the rules specifying the interactions are executed using only local information, without reference to the global pattern. Based on these definitions, in this thesis we describe self-organisation as a process (a) through which the system’s organisation increases dynamically over a time period; (b) where such gain of organisation manifests as the formation of global structure or behaviour (c) and is the result of the interactions among the individual components, (d) which do not have knowledge of the global state of the system, instead they only possess local information.

Kauffman, 2000 suggested that self-organisation occurs through the generation of constraints in the release of energy, which allows such energy to be channeled to perform useful work for the system. This work can then be used to build even more efficient constraints and so on. Adopting this view, one can consider guiding the self-organisation process by controlling the constraints, in order to obtain desired structures or functioning. This practice is typically referred to as *guided self-organisation* [Prokopenko, 2013].

Self-organising systems are found across several different fields [Prokopenko, 2013]. In biology, for example, short distance cell-to-cell communication can trigger pattern formation (e.g., stripes and spots) in developing tissues, while networks of neurons are known to produce spatio-temporal patterns of spikes. Natural self-organised phenomena include forest fires, where the burning of a tree generates high temperature that can set the adjacent trees on fire. An example of self-organised phenomena that involves humans is traffic, where the slowdown of a car gets propagated backwards, causing other cars to reduce their speed and eventually jam. Social

and economical phenomena such as segregation, urban growth and price formation also self-organise out of local interactions between people, households and firms. Crucially, all the active matter systems mentioned in Sec. 2.1 are instances of this phenomenon.

Self-organisation is sometimes associated to the creation of global *order* within a system, starting from an original phase of *disorder*. This interpretation is motivated by observation, in many cases, of a clear symmetry breaking, e.g., the formation of polarised collective motion from an initial phase in which the particles' directions are uncorrelated. Yet, complex systems are typically described as involving both elements of order and disorder, being neither completely regular nor completely random [Sayama, 2015]. For example, flocks of birds combine elements of coherence and perturbed motion [Procaccini et al., 2011; Cavagna et al., 2013a]. The regime between order and disorder is sometimes referred to as the 'edge of chaos', and has been shown to possess interesting features [Kauffman, 1993]. A more useful interpretation of complexity is based on the system's predictability. Completely regular, perfectly structured systems are easy to predict. Totally random systems might be impossible to predict at the individual level, however, the prediction of their average behaviour can be trivial. On the contrary, the prediction of the behaviour of complex systems, both at the individual and at the global scale, is possible but non-trivial.

The main challenge in the study of self-organisation is to understand the relationship between local interactions and global emergent behaviour, and the ultimate goal is the creation of a general theory of this phenomenon [Couzin, 2007]. Researchers from many areas of science and engineering investigate self-organisation using various theoretical tools that include dynamical systems theory, information theory, statistical mechanics, thermodynamics and computer simulations (or agent-based modelling).

2.2.2 Criticality and phase transitions

Complex systems may require a huge amount of variables in order to describe their exact behaviour. In many cases, however, only a small subset of these parameters are necessary to broadly understand the rich macroscopic behaviour. In fact, in many systems one can identify a few *control parameters*, the variation of which drives macroscopic changes in the global behaviour. Such changes are usually reflected in the alteration of *order parameters*, which provide measures of the symmetry breaking within the system [Ebeling and Sokolov, 2005].

The idea that only small subsets of variables are necessary to understand the macroscopic behaviour of systems from their microscopic description is at the core of the universality of self-organisation across different systems. A common feature of many self-organising systems is that the transitions between ordered and disordered configurations can be abrupt, i.e., small changes in the control parameter

might have negligible effect on the system until a critical point is reached, the crossing of which triggers a sudden change in the global configuration, reflected in the order parameters. These phenomena are called *phase transitions*.

Phase transitions are well-known phenomena in active matter, and have been widely studied. In this thesis, in the studies in Chap. 4 and Chap. 6, models of collective motion are shown to exhibit a disordered isotropic phase, with no global polarisation of the particles' heading, and an ordered phase characterised by a net polarisation. The study presented in Chap. 6 also involves another kind of phase transition, called mobility-induced phase separation (MIPS) [Cates and Tailleur, 2015], during which particles' self-propulsion, in combination with volume exclusion, make the system split between dense (liquid-like or solid-like) and sparse (gas-like) areas. Another phase transition is presented in the study of Greater Sydney's urban dynamics in Chap. 5, and involves a dispersed sprawl of the population in contrast with a poly-centric settlement.

Phase transitions are typically classified according with the lowest derivative of the order parameter that is discontinuous [Kondepudi and Prigogine, 2014]. For instance, *first order* phase transitions, which include all solid-liquid-gas transitions of matter, exhibit a discontinuity in the order parameter itself. *second order* phase transitions, such as the magnetisation of ferromagnetic materials, are continuous in the order parameter but discontinuous in its derivative. The specific value of the control parameter at which a phase transition occurs is called the *critical point*. Critical points are obviously important, due to the effects that their crossing would generate on the system. Moreover, critical points implicate the coexistence of the two phases of the system [Sethna, 2006]. Such coexistence can results in the formation of complex behaviour and patters in the system, and is associated with the aforementioned concept of edge of chaos.

2.2.3 Dynamical models and simulations

A common approach to the study of complex systems involves the use of dynamical models to describe the behaviour of the individual components and their interactions. As complex dynamics are typically non-linear, often these models do not have a trivial solution explaining the global behaviour. However, the greatest benefit of using dynamical models is that they can be implemented in a computer, and simulated in order to measure emergent global behaviour [Sayama, 2015].

Several dynamical models of collective motion have been developed [Vicsek et al., 1995; Toner and Tu, 1995; Grégoire and Chaté, 2004]. Here, as an example, we introduce one that generalises many well-known models. Let us assume there are N particles and denote the position and heading of each particle a as, respectively, \mathbf{r}_a and θ_a . The particles *self-propulsion* force $\mathcal{F}_P(\theta_a)$ moves particles in the direction of their heading. Moreover, particles interact with each other locally, for a given definition of locality (e.g., nearest neighbourhood). Let us assume there are two kinds of forces that make particles interact: a *potential* force $\mathcal{F}_U(r_{ab})$, which moves

particles closer to or away from each other depending on their relative position, and an *alignment* torque $\mathcal{F}_T(r_{ab}, \theta_{ab})$ that makes particles' heading converge. Finally, random perturbations η_a and μ_a are introduced, which affect position and heading. A simple dynamical model of this system, discrete in time n , is the following:

$$\mathbf{r}_a(n+1) = \mathbf{r}_a(n) + \sum_{b \in \mathcal{N}^U} \mathcal{F}_U(r_{ab}(n)) + \sum_{b \in \mathcal{N}^P} \mathcal{F}_P(\theta_a(n)) + \eta_a(n) \quad (2.1)$$

$$\theta_a(n+1) = \theta_a(n) + \sum_{b \in \mathcal{N}^T} \mathcal{F}_T(r_{ab}(n), \theta_{ab}(n)) + \mu_a(n), \quad (2.2)$$

where $\mathbf{r}_{ab} = \mathbf{r}_a - \mathbf{r}_b$, $\theta_{ab} = \theta_a - \theta_b$ and \mathcal{N}^U , \mathcal{N}^P and \mathcal{N}^T are sets of particles defined according with the locality criteria. Abstract models like this, including the famous model of [Vicsek et al., 1995], have been extensively used in collective motion, for instance in modelling flocks [Chaté et al., 2008] and groups of cells [Drongelen et al., 2015]. A model of this kind, proposed by Grégoire and Chaté, 2004, is used in this thesis in the study presented in Chap. 4.

When a detailed thermodynamical analysis is required, such as the one presented in Chap. 6, a description of the system that involves stochastic differential equations is needed. An example is the following:

$$d\mathbf{r}_a^i = \sum_{b \in \mathcal{N}^U} -\frac{\partial U(r_{ab})}{\partial \mathbf{r}_a^i} dt + \mathcal{P}(\theta_a)^i dt + dW_{\mathbf{r}_a^i} \quad (2.3)$$

$$d\theta_a = \mathcal{T}(r_{ab}, \theta_{ab}) dt + dW_{\theta_a}, \quad (2.4)$$

where $U(r_{ab})$, $\mathcal{P}(\theta_a)$ and $\mathcal{T}(r_{ab}, \theta_{ab})$ are differentiable functions that approximate, respectively, the potential interactions, the self-propulsion and the torque interactions, i is a spatial dimension and $W_{\mathbf{r}_a^i}$ and W_{θ_a} are uncorrelated Wiener processes [Gardiner, 2009], such that $\langle dW_{\mathbf{r}_a^i} dW_{\theta_a} \rangle = 0$, $\langle dW_{\theta_a} dW_{\theta_b} \rangle = \delta_{ab} dt$ and $\langle dW_{\mathbf{r}_a^i} dW_{\mathbf{r}_b^j} \rangle = \delta_{ij} \delta_{ab} dt$ (here δ is the Dirac delta function).

2.2.4 Statistical mechanics and maximum entropy models

Dynamical models and their simulations are very useful tool, however, they do not provide a theory for the emergence of global behaviour. This has been a limitation for many researchers, who are aiming at understanding collective motion in a rigorous way, similarly to how the behaviour of idealised gasses is understood in physics [Popkin, 2016]. For this purpose, a more rigorous theoretical framework is required, which is capable of providing a bridge between microscopic and macroscopic descriptions, and *statistical mechanics* is an obvious candidate.

Based on probability theory, statistical mechanics was very successful in explaining the thermodynamic behaviour of large systems of non-interacting particles in the equilibrium state [Sethna, 2006; Kondepudi and Prigogine, 2014]. The introduction of the *principle of maximum entropy* by Jaynes, 1957 allowed the application of statistical mechanical concepts in diverse domains, independently of any physical

argument. Entropy, described in the next section, is proportional to the number of microscopic states that are consistent with the global quantities observed in the macroscopic configuration of a system. The principle of maximum entropy states that, among all the probability distributions of the microscopic variables that are consistent with the macroscopic observations, the least biased one is that with the maximal entropy. Maximum entropy models can be formulated to describe such least biased probability distribution.

Bialek et al., 2012 utilised a maximum entropy model to develop a statistical mechanical theory of flocking phenomena. The resulting maximum entropy distribution, consistent with the directional correlations estimated from experimental data on large flocks of starlings, is the following:

$$P(\{\mathbf{s}_a\}) = \frac{1}{Z(\{J_{ab}\})} \exp \left(\frac{1}{2} \sum_{a=1}^N \sum_{b=1}^N J_{ab} (\mathbf{s}_a \cdot \mathbf{s}_b) \right), \quad (2.5)$$

where \mathbf{s}_a is the normalised velocity of particle a , N is the total number of particles, each parameter J_{ab} corresponds to a measurement of directional correlation C_{ab} and $Z(\{J_{ab}\})$ is the partition function. Importantly, this statistical mechanical description was shown capable of producing quantitative predictions of collective phenomena, including the emergence of pairwise and four-body correlations.

Bialek's model is consistent with the dynamical model of collective motion used in this thesis in Chap. 4. Moreover, an analogous maximum entropy model is used in this thesis in order to describe the income flow distribution within Greater Sydney (see Chap. 5).

2.2.5 Thermodynamics and non-equilibrium systems

Thermodynamics is the branch of physics concerned with the study of the transfer and transformation of energy, in particular in the forms of thermal heat and mechanical work [Kondepudi and Prigogine, 2014].

Introduced by pioneers such as Sadi Carnot during the conception of the steam engine in the early nineteenth century, thermodynamics is mostly known for its first and the second laws, later summarised by Rudolf Clausius [Ebeling and Sokolov, 2005]. The *first law* of thermodynamics, or the law of conservation of energy, states that energy cannot be created nor destroyed, it can only be converted to other forms, or transferred to other systems. As a direct consequence, energy is conserved in an isolated system. The *second law* of thermodynamics involves the concept of entropy, as the measure of the energy in the system per unit temperature that cannot be utilised to perform mechanical work. The second law states that entropy can be created, but is never destroyed, therefore heat cannot be entirely converted into mechanical work.

Isolated systems reach an equilibrium state in which the entropy is maximal. In this state, statistical mechanics can be used for describing the probability of a

microscopic state ω given its energy ε_ω . This is the famous Boltzmann distribution:

$$P(\omega) = \frac{e^{-\varepsilon_\omega/k_b T}}{\sum_{v=1}^M e^{-\varepsilon_v/k_b T}}, \quad (2.6)$$

where T is the temperature, M is the number of possible states and k_b is the Boltzmann constant. Notice, that maximum entropy probability distribution of flocking velocities in Eq. (2.5) is an instance of the Boltzmann distribution. The main limitation of the maximum entropy principle is that it relies on the assumption of thermal equilibrium and, therefore, can only be applied to systems in a stationary state. In some particular settings, for example in the study of collective motion described in Chap. 4, the steady state assumption can be made. In more realistic scenarios, when the dynamics of the observed system are reasonably stable (as for Bialek et al., 2012), the assumption of thermal equilibrium can be used as a simplification for studying emergent phenomena.

In general, however, active matter is characterised by non-equilibrium dynamics: particles absorb energy from the environment and transform it into directed motion, and in doing so they break microscopic detailed balance [Ramaswamy, 2017]. Non-equilibrium statistical mechanics is far from being as developed as its equilibrium counterpart. Nevertheless, over the last thirty years new theoretical frameworks have revealed laws that hold for non-equilibrium systems, allowing the study of some of their thermodynamics [Seifert, 2012]. One of the problem this thesis is concerned with, specifically in the study presented in Chap. 6, is the quantification of the entropy production rate of active matter during non-equilibrium dynamics that involve orientational phase transitions as well as MIPS.

2.3 The theoretical framework

This section outlines some key theoretical concepts and measures used in this thesis, and discusses their application to complex systems. Firstly, we present information theory and the concept of Shannon entropy, as a general approach for quantifying information. Secondly, we describe the framework of information dynamics, highlighting the importance of detecting how information propagates in time within a system's components. Particular focus is dedicated to the transfer entropy, a measure of information transfer that has been used in this thesis, in particular in the study presented in Chap. 3. Thirdly, we describe the subject of information geometry and present one of its key measures, the Fisher information, that was used in this thesis to characterise phase transitions in collective motion (see the study in Chap. 4) as well as in urban dynamics (see the study in Chap. 5). Lastly, we explain the concepts of statistical thermodynamics and entropy production, used in the study described in Chap. 6, contextualising them within the scope of collective motion.

2.3.1 Shannon entropy and information theory

The field of *information theory* originated in the 1940s with the work of Shannon, 1948, who introduced this theoretical framework while working on the problem of information encoding in communication channels. The central quantity is the *Shannon entropy*, which represents the minimal expected number of bits required to encode a message without loss of information [Cover and Thomas, 2012; MacKay, 2003].

In general, the (Shannon) entropy is a measure of information content, quantified as the *uncertainty* associated with a random variable. Let us consider a random variable X , its alphabet α_x , and the probability distribution function $P(x)$ defined for all possible outcomes x of X . The entropy $h(x)$ of a single outcome x of X is defined as:

$$h(x) = \log \frac{1}{P(x)} = -\log P(x), \quad (2.7)$$

where $\frac{1}{P(x)}$ represents the ‘surprisal’ associated with the observation the specific outcome x . In other words, $h(x)$ represents a priori uncertainty about x . Unless otherwise specified, the base of the logarithm is 2, meaning that the unit of measure of the entropy is the *bit*. Whether the natural logarithm is used instead, the unit of measure will be the *nat*. The entropy of the entire variable X is then defined as the average uncertainty of all the possible outcomes x of X :

$$H(X) = - \sum_{x \in \alpha_x} P(x) \log P(x). \quad (2.8)$$

It is worth noting that lower-case symbols are typically used for single (or *local*) measures, e.g., $h(x)$, while upper-case symbols are used for average measures, e.g., $H(X)$. This convention will be used throughout this thesis.

Reductions in entropy provide a general, and at the same time rigorous quantification of the information ‘acquired’ by an observer about the phenomenon being observed. The main two reasons for this are (a) that Shannon entropy is model-free, i.e., it only requires access to the probability distributions of the variables involved, and (b) that it is capable of capturing non-linear relationships. These characteristics make the entropy the best candidate for quantifying information that is not semantic but rather stochastic, noisy and/or embedded in physical and spatial interactions, a typical scenario in most complex systems. Information theory has indeed become one of the most successful frameworks in the study of complex systems [Adami, 2002; Prokopenko et al., 2009; Walker and Davies, 2013]. For example, aspects of information processing can be quantified using Shannon entropy and other measures derived from it, including information dynamics, as is described in Sec. 2.3.2. Sec. 2.3.3 and Sec. 2.3.4 show how concepts derived from information theory can be utilised for thermodynamic treatments, which include the study of phase transitions, critical regimes and entropy production.

Before proceeding further, it is useful to introduce two more fundamental concepts: the *joint entropy* and the *conditional entropy*. The local and average *joint entropy* are defined as

$$h(x, y) = -\log P(x, y) \quad (2.9)$$

and

$$H(x, y) = -\sum_{x \in a_x} \sum_{y \in a_y} P(x, y) \log P(x, y), \quad (2.10)$$

where a second random variable Y is introduced and $P(x, y)$ is the joint probability of the measurements x and y . The joint entropy represents the uncertainty associated with observing two variables together, and can be generalised to any number of variables. Similarly, the local and average *conditional entropy* are defined as:

$$h(x|y) = h(x, y) - h(y) \quad (2.11)$$

and

$$H(X|Y) = H(X, Y) - H(Y). \quad (2.12)$$

The conditional entropy represents the uncertainty associated with observing a random variable having a priori knowledge of another variable. It also can be generalised to any number of variables.

2.3.2 Transfer entropy and information dynamics

The concept of computation has become widely used in field of complex systems [Mitchell, 2009], to the point that self-organisation mechanisms have started to be understood as distributed information processes [Lizier et al., 2008; Lizier et al., 2010; Lizier et al., 2012]. For example, interactions between the components of a

system can be seen as communication, the state of a component as memory, and the system as a whole as computing its next state via a distributed computation.

Information processing is typically analysed distinguishing three fundamental components: information *storage*, *transfer* and *modification*. Many artificial systems, including the computers that we daily use, are designed keeping these three component separated, and their interactions can be well understood. In the case of complex systems, however, the distinction between storage, transfer and modification is more blurred, and their interactions are much more intricate. Nonetheless, given its importance, information processing have been vastly investigated in complex systems. Recently, for instance, much attention has been given to the study of the *dynamics* of these fundamental components, i.e., how information storage, transfer, and modification unfold in time and across the components of a system. The theoretical framework adopted for this kind of studies is referred as *information dynamics* [Lizier, 2013], and includes several different entropic measures, the most important of which, in the context of this thesis, is the *transfer entropy*. As the name suggests, this measure is related to the transfer component of information processing, and has been utilised to investigate emergent local structures and behavioural changes in complex systems [Lizier et al., 2008; Wang et al., 2012; Boedecker et al., 2012; Barnett et al., 2013].

Before describing the transfer entropy in detail, it is necessary to introduce the measure from which it is derived: the *mutual information*. The mutual information between two random variables X and Y is defined [Cover and Thomas, 2012; MacKay, 2003] as

$$\begin{aligned} I(X;Y) &= H(X) + H(Y) - H(X,Y) \\ &= H(X) - H(X|Y) = H(Y) - H(Y|X) \\ &= \sum_{x,y} p(x,y) \log \frac{p(x,y)}{p(x)p(y)}. \end{aligned} \tag{2.13}$$

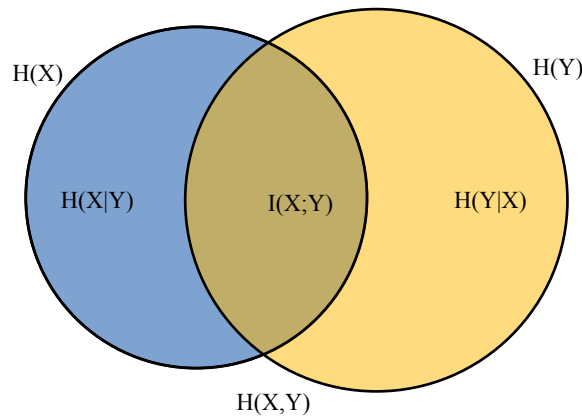


FIGURE 2.1: A Venn diagram showing the mutual information between two random variables X and Y and its relationships with the joint and conditional entropies of the same variables.

It represents the average reduction of uncertainty about X that is gained from knowing the probabilities of the values of Y , or vice versa. An intuitive Venn diagram, illustrating the mutual information between two random variables and its relationships with the joint and conditional entropies, is shown in Fig. 2.1. Mutual information can also be interpreted as a measure of dependence between the two random variables: $I(X;Y) = 0$ means that the two variables are independent, while positive values indicate that (and how much) X and Y are statistically dependent on each other. Mutual information has been often used in complex systems, mostly as a measure of complexity and for characterising order-chaos phase transitions [Langton, 1990; Tononi et al., 1994; Adami, 2002; Mathis et al., 2017].

The local mutual information is

$$i(x;y) = \log \frac{p(x,y)}{p(x)p(y)}. \quad (2.14)$$

While the mutual information is on average non-negative, local values can be negative. Negative values of the local mutual information indicate that knowing a realisation of x of X is *misinformative* for predicting the outcome y of Y , i.e., knowing x increases the uncertainty associated with y .

The local and average *conditional mutual informations* between X and Y given a third random variable Z are also defined:

$$i(x;y|z) = \log \frac{p(x,y,z)}{p(x|y,z)p(x|z)} \quad (2.15)$$

and

$$I(X;Y|Z) = \sum_{x,y,z} \log \frac{p(x,y,z)}{p(x|y,z)p(x|z)}. \quad (2.16)$$

The conditional mutual information represents the shared information between X and Y that is not contained in Z , and it can be smaller, equal or even larger than its non-conditional counterpart.

Instead of two random variables X and Y , let us now consider two time-series defined over a discrete time n . Let us call these time-series S , as for *source*, and D , as for *destination*. Let us also consider D_n , the current state of the destination, and S_{n-l} , the previous state of the source having chosen a positive integer number l . Finally, let us define the destination's past $\mathbf{D}_{n-1}^{(k,\tau)} = \{D_{n-1-j\tau}\}$, for $j = \{0, 1, \dots, k-1\}$, for arbitrary positive integer numbers k and τ . A schematic representation of these quantities, for specific values of l , k and τ is given in Fig. 2.2. We now have all the ingredients to describe the local transfer entropy [Lizier et al., 2008; Lizier, 2014] at time n as the mutual information between the current state of the destination and

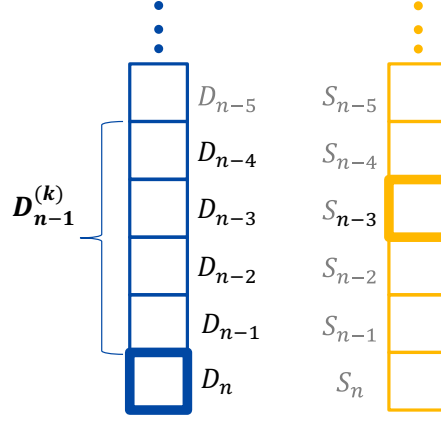


FIGURE 2.2: An example of source and destination processes highlighting the current state of the destination, D_n , the previous state of the source, S_{n-l} , and the destination's past, $\mathbf{D}_{n-1}^{(k,\tau)}$. In this example $l = 3, k = 4$ and $\tau = 1$.

the previous state of source, conditional on the destination's past:

$$\begin{aligned}
 t_{S \rightarrow D}(n, l, k, \tau) &= i \left(S_{n-l}; D_n | \mathbf{D}_{n-1}^{(k,\tau)} \right) \\
 &= p \left(S_{n-l}, D_n, \mathbf{D}_{n-1}^{(k,\tau)} \right) \log \frac{p \left(D_n | \mathbf{D}_{n-1}^{(k,\tau)}, S_{n-l} \right)}{p \left(D_n | \mathbf{D}_{n-1}^{(k,\tau)} \right)}. \quad (2.17)
 \end{aligned}$$

The local transfer entropy captures the transfer of uncertainty reduction, or predictive information flow, from the source process to the destination process, at a specific time. The average transfer entropy $T_{S \rightarrow D}(l, k, \tau)$ was originally introduced by Schreiber, 2000 as the average of the local transfer entropies.

Rather than a static measure of correlations, the transfer entropy is a dynamic measure, associated with state transitions of the destination process. Moreover, the transfer entropy is also asymmetric in S and D , i.e., it captures directional relationships between the two processes. As for the mutual information, local values of transfer entropy can be negative, while its average is non-negative. Negative values of local transfer entropy signify that the previous state of the source is misinformative about the current state of the destination. The capability of local transfer entropy to capture misinformation was crucial for distinguishing between positive and negative information flows within a school of fish, as explained in Chap. 3.

It is also worth clarifying that transfer entropy does not quantify the causal effect of the source on the target. This measure quantifies the information that an observer gains from the source process that can be used to predict the next state of the destination process, in a purely probabilistic setting and without any claim on the underlying mechanisms through which the two processes interact. Many researchers may be interested in causality, but the concept of predictive information transfer can reveal aspects about the local structures and behaviours that causal effect may not [Lizier and Prokopenko, 2010]. In order to highlight the distinction between causal

and predictive information, in this thesis (and in particular in the study presented in Chap. 3) we will refer to values of transfer entropy as measurements of “predictive information flow”.

2.3.3 Fisher information and information geometry

Developed by Amari in the 1980s, information geometry investigates the differential geometric structure of families of probability distribution [Amari and Nagaoka, 2007]. A family of probability distributions is viewed as a manifold, and the parameters that characterise the family as the coordinate system. This Riemannian space has a metric that reflects the characteristics of the probability distribution, and information geometry provides a geometric tool for studying the affinities between this metric and the probability distributions. Information geometry has applications in several fields, including machine learning and signal processing [Ay et al., 2017].

A key measure in information geometry is the Fisher information [Fisher, 1922], which quantifies the amount of information that a random variable X carries about an unknown parameter θ . Specifically, the Fisher information matrix is defined for multiple parameters $\theta = \{\theta_1, \theta_2, \dots, \theta_M\}$ as

$$F_{nm}(\theta) = E \left[\left(\frac{\partial \ln P(x|\theta)}{\partial \theta_m} \right) \left(\frac{\partial \ln P(x|\theta)}{\partial \theta_n} \right) \middle| \theta \right], \quad (2.18)$$

where $E(w)$ is the expected value of w and $P(x|\theta)$ is the probability of the realisation x of X knowing θ . The Fisher information matrix represents the *sensitivity* of the probability distribution to changes in the control parameters θ .

Importantly for this thesis, and in particular for the studies in Chap. 4 and Chap. 5, the meaning of the Fisher information for physical systems has been investigated in terms of thermodynamics and statistical mechanics. A system in a stationary state is described by the Gibbs measure

$$p(x|\theta) = \frac{1}{Z(\theta)} e^{-\beta H(x, \theta)} = \frac{1}{Z(\theta)} e^{-\sum_m \theta_m X_m(x)}, \quad (2.19)$$

where β is the inverse temperature, $H(x, \theta)$ is the Hamiltonian defining the total energy at state x , θ_m are thermodynamic variables (pressure, magnetic field, chemical potential, etc.) and $Z(\theta)$ is the partition function [Brody and Rivier, 1995; Crooks, 2007]. At temperature T , such system is characterised by the Gibbs free energy $G(T, \theta_m) = U(S, \phi_m) - TS - \phi_m \theta_m$, where $U(S, \phi_m)$ is the internal energy, S is the configuration entropy and ϕ_m is an order parameter. The Fisher information was shown to quantify the fluctuations about equilibrium in the collective variables X_m and X_n and to be proportional to the derivatives of the corresponding order parameters with respect to the collective variables [Crooks, 2011; Prokopenko et al., 2011]:

$$F_{nm}(\theta) = \langle (X_m(x) - \langle X_m \rangle) (X_n(x) - \langle X_n \rangle) \rangle = \beta \frac{\partial \phi_m}{\partial \theta_n} \quad (2.20)$$

The Fisher information was also shown to be equivalent to the metric tensor,

$$F_{nm}(\theta) = g_{nm}(\theta) = \frac{\partial^2 \psi}{\partial \theta_m \partial \theta_n}, \quad (2.21)$$

where $\psi = \ln Z = -\beta G$ is the free entropy [Brody and Rivier, 1995; Brody and Ritz, 2003; Janke et al., 2004; Crooks, 2007]. In other words, the Fisher information is proportional to the curvature of the free energy with respect to the control parameters. Moreover, it has been argued that the difference between curvatures of the configuration entropy and the free entropy (i.e., the Fisher information) is related to a ‘computational balance’ between uncertainty and sensitivity [Prokopenko and Einav, 2015]. In the study presented in Chap. 4, these relationships are extended to incorporate the curvature of the other two fundamental quantities in thermodynamics: the internal energy of the system U and the work required to vary the control parameters.

2.3.4 Entropy production and stochastic thermodynamics

Stochastic thermodynamics is the branch of statistical mechanics that investigates the non-equilibrium dynamics of microscopic systems. More specifically, stochastic thermodynamics studies quantities such as work, heat and entropy production in small driven systems, at the scale of individual fluctuating trajectories [Sekimoto, 2010; Jarzynski, 2011]. Examples of such systems include biopolymers, enzymes and molecular motors [Seifert, 2012], as well as many of the biological and synthetic microscopic particles seen in Sec. 2.1.2 and Sec. 2.1.3, which are typically embedded in aqueous mediums.

Let us consider a system and its environment, e.g., the medium. The Gibbs entropy of the system is defined over its micro-states ω as $S^{\text{sys}} = -k_b \sum_{\omega} P(\omega) \ln P(\omega)$, where $P(\omega)$ is the probability of ω occurring during the systems fluctuations and k_b is the Boltzmann constant. The entropy production in the system ΔS^{sys} is equal to the changes in the Gibbs entropy over the dynamical variables that describe it. The entropy production in the environment ΔS^{med} consists of the exported heat, scaled by the inverse environmental temperature. Therefore, the total entropy production in the system and the environment is $\Delta S^{\text{tot}} = \Delta S^{\text{sys}} + \Delta S^{\text{med}}$ and, according to the second law of thermodynamics, $\Delta S^{\text{tot}} \geq 0$.

In the formalism of stochastic thermodynamics, the entropy production is used as a measure of dynamical irreversibility, and become thermodynamically meaningful when scaled by k_b . Let us describe the state of the system with X , and consider an interval of time t over which individual realisations of the system $\vec{X} = \{X(t) | t \in [t_0, \tau]\}$ are defined. Let us also define the reversal path $\vec{X}^{\dagger} = \{X^{\dagger}(t) | t \in [t_0, \tau]\}$ and $X^{\dagger}(t) = \epsilon X(\tau + t_0 - t)$ where ϵ is a time reversal operation [Spinney and Ford, 2012b]. The total entropy production generated over an interval of time $t \in [t_0, \tau]$

can be written as:

$$\Delta S^{\text{tot}} = \ln \frac{\mathcal{P}(\vec{X})}{\mathcal{P}^+(\vec{X}^+)} = D_{\text{KL}}(\mathcal{P}(\vec{X}) \parallel \mathcal{P}^+(\vec{X}^+)), \quad (2.22)$$

where $\mathcal{P}(\vec{X})$ and $\mathcal{P}^+(\vec{X}^+)$ are, respectively, the probability densities of the forward and reversal paths, and $D_{\text{KL}}(\mathcal{P}(\vec{X}) \parallel \mathcal{P}^+(\vec{X}^+))$ is the Kullback-Leibler divergence, or relative entropy, measuring the difference between $\mathcal{P}(\vec{X})$ and $\mathcal{P}^+(\vec{X}^+)$. Therefore, the total entropy production is the log ratio of the likelihood of a trajectory to its time reverse under the dynamics that describe the system.

For many-body systems, such as active matter, estimating the probability distribution function of the state of the whole system is generally infeasible, due to the large number of variables needed in order to describe the system. However, when a system is described by stochastic differential equations, the expressions for ΔS^{med} can be determined in terms of knowledge of the trajectories only [Spinney and Ford, 2012a] (see Chap. 6 for more details). Crucially, in the steady state the expected entropy production in the system $\langle \Delta S^{\text{sys}} \rangle$ vanishes, leaving $\langle \Delta S^{\text{tot}} \rangle = \langle \Delta S^{\text{med}} \rangle$ and therefore allowing the calculation of the total entropy production. This approach has been used to investigate the entropy production associated to the collective motion of active matter [Fodor et al., 2016; Mandal et al., 2017], and in particular the phenomenon of MIPS. In Chap. 6, this method is utilised in order to study a system of active Brownian particles that interact via volume exclusion and alignment. The system exhibit MIPS as well as global orientational order, and both these phenomena are shown to affect the entropy production.

References

- Adami, C. (2002). “What is complexity?” In: *BioEssays* 24.12, pp. 1085–1094.
- Amari, S.-i. and H. Nagaoka (2007). *Methods of information geometry*. Vol. 191. American Mathematical Society.
- Ay, N. et al. (2017). *Information geometry*. Vol. 8. Springer.
- Barnett, L. et al. (2013). “Information flow in a kinetic Ising model peaks in the disordered phase”. In: *Physical Review Letters* 111.17, p. 177203.
- Bazazi, S. et al. (2010). “The social context of cannibalism in migratory bands of the mormon cricket”. In: *PLOS One* 5.12, e15118.
- Bechinger, C. et al. (2016). “Active particles in complex and crowded environments”. In: *Reviews of Modern Physics* 88.4, p. 045006.
- Beekman, M., D. J. Sumpter, and F. L. Ratnieks (2001). “Phase transition between disordered and ordered foraging in Pharaoh’s ants”. In: *Proceedings of the National Academy of Sciences* 98.17, pp. 9703–9706.
- Bialek, W. et al. (2012). “Statistical mechanics for natural flocks of birds”. In: *Proceedings of the National Academy of Sciences* 109.13, pp. 4786–4791.

- Blair, D. L., T Neicu, and A. Kudrolli (2003). "Vortices in vibrated granular rods". In: *Physical Review E* 67.3, p. 031303.
- Boedecker, J. et al. (2012). "Information processing in echo state networks at the edge of chaos". In: *Theory in Biosciences* 131.3, pp. 205–213.
- Bonabeau, E. et al. (1997). "Self-organization in social insects". In: *Trends in ecology & evolution* 12.5, pp. 188–193.
- Bonabeau, E. et al. (1999). *Swarm intelligence: from natural to artificial systems*. 1. Oxford University Press.
- Bricard, A. et al. (2013). "Emergence of macroscopic directed motion in populations of motile colloids". In: *Nature* 503.7474, p. 95.
- Brody, D. and A. Ritz (2003). "Information geometry of finite Ising models". In: *Journal of Geometry and Physics* 47.2, pp. 207–220.
- Brody, D. and N. Rivier (1995). "Geometrical aspects of statistical mechanics". In: *Physical Review E* 51.2, pp. 1006–1011.
- Buhl, J. et al. (2006). "From disorder to order in marching locusts". In: *Science* 312.5778, pp. 1402–1406.
- Butail, S. et al. (2014). "Information flow in animal-robot interactions". In: *Entropy* 16.3, pp. 1315–1330.
- Buttinoni, I. et al. (2013). "Dynamical clustering and phase separation in suspensions of self-propelled colloidal particles". In: *Physical Review Letters* 110.23, p. 238301.
- Camazine, S. et al. (2003). *Self-organization in biological systems*. Vol. 7. Princeton University Press.
- Cates, M. E. and J. Tailleur (2015). "Motility-induced phase separation". In: *Annual Review Condensed Matter Physics* 6.1, pp. 219–244.
- Cavagna, A. et al. (2013a). "Diffusion of individual birds in starling flocks". In: *Proceedings of the Royal Society of London B: Biological Sciences* 280.1756, p. 20122484.
- Cavagna, A., I. Giardina, and F. Ginelli (2013b). "Boundary information inflow enhances correlation in flocking". In: *Physical Review Letters* 110.16, p. 168107.
- Chaté, H. et al. (2008). "Modeling collective motion: variations on the Vicsek model". In: *The European Physical Journal B* 64.3-4, pp. 451–456.
- Couzin, I. (2007). "Collective minds". In: *Nature* 445.7129, p. 715.
- Cover, T. M. and J. A. Thomas (2012). *Elements of information theory*. John Wiley & Sons.
- Crooks, G. (2007). "Measuring thermodynamic length". In: *Physical Review Letters* 99.10, p. 100602.
- (2011). *Fisher information and statistical mechanics*. Tech. rep.
- Doursat, R., H. Sayama, and O. Michel (2012). *Morphogenetic engineering: toward programmable complex systems*. Springer.
- Drongelen, R. van et al. (2015). "Collective dynamics of soft active particles". In: *Physical Review E* 91.3, p. 032706.
- Dyer, J. R. et al. (2008). "Consensus decision making in human crowds". In: *Animal Behaviour* 75.2, pp. 461–470.

- Dyer, J. R. et al. (2009). "Leadership, consensus decision making and collective behaviour in humans". In: *Philosophical Transactions of the Royal Society of London B: Biological Sciences* 364.1518, pp. 781–789.
- Ebeling, W. and I. M. Sokolov (2005). *Statistical thermodynamics and stochastic theory of nonequilibrium systems*. Vol. 8. World Scientific Publishing Company.
- Fisher, R. A. (1922). "On the mathematical foundations of theoretical statistics". In: *Philosophical Transactions of the Royal Society of London* 222.594-604, pp. 309–368.
- Fodor, É. et al. (2016). "How far from equilibrium is active matter?" In: *Physical Review Letters* 117.3, p. 038103.
- Fourcassié, V., A. Dussutour, and J.-L. Deneubourg (2010). "Ant traffic rules". In: *The Journal of Experimental Biology* 213.14, pp. 2357–2363.
- Friedl, P., Y. Hegerfeldt, and M. Tusch (2004). "Collective cell migration in morphogenesis and cancer". In: *International Journal of Developmental Biology* 48.5-6, pp. 441–449.
- Fujikawa, H. and M. Matsushita (1989). "Fractal growth of *Bacillus subtilis* on agar plates". In: *Journal of the Physical Society of Japan* 58.11, pp. 3875–3878.
- Gardiner, C. (2009). *Stochastic methods*. Vol. 4. Springer Berlin.
- Gautrais, J. et al. (2012). "Deciphering interactions in moving animal groups". In: *PLOS Computational Biology* 8.9, pp. 1–11.
- Ginelli, F. et al. (2015). "Intermittent collective dynamics emerge from conflicting imperatives in sheep herds". In: *Proceedings of the National Academy of Sciences* 112.41, pp. 12729–12734.
- Gotwald Jr, W. H. et al. (1995). *Army ants: the biology of social predation*. Cornell University Press.
- Grégoire, G. and H. Chaté (2004). "Onset of collective and cohesive motion". In: *Physical Review Letters* 92.2, p. 025702.
- Haas, P. and D. Gilmour (2006). "Chemokine signaling mediates self-organizing tissue migration in the zebrafish lateral line". In: *Developmental Cell* 10.5, pp. 673–680.
- Haken, H. and G. Jumarie (2006). *A macroscopic approach to complex system*. Springer.
- Herbert-Read, J. E. et al. (2015). "Initiation and spread of escape waves within animal groups". In: *Royal Society Open Science* 2.4, p. 140355.
- Hu, F., L.-J. Nie, and S.-J. Fu (2015). "Information dynamics in the interaction between a prey and a predator fish". In: *Entropy* 17.10, pp. 7230–7241.
- Ibele, M., T. E. Mallouk, and A. Sen (2009). "Schooling behavior of light-powered autonomous micromotors in water". In: *Angewandte Chemie* 121.18, pp. 3358–3362.
- Janke, W., D. Johnston, and R. Kenna (2004). "Information geometry and phase transitions". In: *Physica A: Statistical Mechanics and its Applications* 336.1–2, pp. 181–186.
- Jarzynski, C. (2011). "Equalities and inequalities: irreversibility and the second law of thermodynamics at the nanoscale". In: *Annual Review Condensed Matter Physics* 2.1, pp. 329–351.

- Jaynes, E. (1957). "Information Theory and Statistical Mechanics". In: *Physical Review* 106.4, pp. 620–630.
- Katz, Y. et al. (2011). "Inferring the structure and dynamics of interactions in schooling fish". In: *Proceedings of the National Academy of Sciences* 108.46, pp. 18720–18725.
- Kauffman, S. A. (1993). *The origins of order: self-organization and selection in evolution*. Oxford University Press, USA.
- (2000). *Investigations*. Oxford University Press.
- Köhler, S., V. Schaller, and A. R. Bausch (2011). "Structure formation in active networks". In: *Nature Materials* 10.6, p. 462.
- Kondepudi, D. and I. Prigogine (2014). *Modern thermodynamics: from heat engines to dissipative structures*. John Wiley & Sons.
- Kumar, N. et al. (2014). "Flocking at a distance in active granular matter". In: *Nature Communications* 5, p. 4688.
- Langton, C. G. (1990). "Computation at the edge of chaos: phase transitions and emergent computation". In: *Physica D: Nonlinear Phenomena* 42.1-3, pp. 12–37.
- Lissaman, P. B. S. and C. A. Shollenberger (1970). "Formation flight of birds". In: *Science* 168.3934, pp. 1003–1005.
- Lizier, J. T. (2013). *The local information dynamics of distributed computation in complex systems*. Springer Berlin.
- Lizier, J. T. (2014). "Measuring the dynamics of information processing on a local scale in time and space". In: *Directed information measures in neuroscience*. Springer, pp. 161–193.
- Lizier, J. T. and M. Prokopenko (2010). "Differentiating information transfer and causal effect". In: *The European Physical Journal B* 73.4, pp. 605–615.
- Lizier, J. T., M. Prokopenko, and A. Y. Zomaya (2008). "Local information transfer as a spatiotemporal filter for complex systems". In: *Physical Review E* 77.2, p. 026110.
- (2010). "Information modification and particle collisions in distributed computation". In: *Chaos: An Interdisciplinary Journal of Nonlinear Science* 20.3, p. 037109.
- (2012). "Coherent information structure in complex computation". In: *Theory in Biosciences* 131.3, pp. 193–203.
- Lord, W. M. et al. (2016). "Inference of causal information flow in collective animal behavior". In: *IEEE Transactions on Molecular, Biological and Multi-Scale Communications* 2.1, pp. 107–116.
- MacKay, D. J. (2003). *Information theory, inference and learning algorithms*. Cambridge University Press.
- Mandal, D., K. Klymko, and M. R. DeWeese (2017). "Entropy production and fluctuation theorems for active matter". In: *Physical Review Letters* 119.25, p. 258001.
- Marchetti, M. C. et al. (2013). "Hydrodynamics of soft active matter". In: *Reviews of Modern Physics* 85.3, p. 1143.
- Mathis, C., T. Bhattacharya, and S. I. Walker (2017). "The emergence of life as a first-order phase transition". In: *Astrobiology* 17.3, pp. 266–276.

- May, R. M. (1979). "Flight formations in geese and other birds". In: *Nature* 282, pp. 778–780.
- Méhes, E. et al. (2012). "Collective motion of cells mediates segregation and pattern formation in co-cultures". In: *PLOS One* 7.2, e31711.
- Mitchell, M. (2009). *Complexity: a guided tour*. Oxford University Press.
- Moussaïd, M., D. Helbing, and G. Theraulaz (2011). "How simple rules determine pedestrian behavior and crowd disasters". In: *Proceedings of the National Academy of Sciences* 108.17, p. 6884.
- Nagy, M. et al. (2010). "Hierarchical group dynamics in pigeon flocks". In: *Nature* 464.7290, pp. 890–893.
- Nagy, M. et al. (2013). "Context-dependent hierarchies in pigeons". In: *Proceedings of the National Academy of Sciences* 110.32, pp. 13049–13054.
- Narayan, V., N. Menon, and S. Ramaswamy (2006). "Nonequilibrium steady states in a vibrated-rod monolayer: tetratic, nematic, and smectic correlations". In: *Journal of Statistical Mechanics: Theory and Experiment* 2006.01, P01005.
- Ndlec, F. et al. (1997). "Self-organization of microtubules and motors". In: *Nature* 389.6648, p. 305.
- Nédélec, F., T. Surrey, and A. Maggs (2001). "Dynamic concentration of motors in microtubule arrays". In: *Physical Review Letters* 86.14, p. 3192.
- Needleman, D. and Z. Dogic (2017). "Active matter at the interface between materials science and cell biology". In: *Nature Reviews Materials* 2.9, p. 17048.
- Orange, N. and N. Abaid (2015). "A transfer entropy analysis of leader-follower interactions in flying bats". In: *The European Physical Journal Special Topics* 224.17-18, pp. 3279–3293.
- Ordemann, A., G. Balazsi, and F. Moss (2003). "Pattern formation and stochastic motion of the zooplankton *Daphnia* in a light field". In: *Physica A: Statistical Mechanics and its Applications* 325.1-2, pp. 260–266.
- Palacci, J. et al. (2013). "Living crystals of light-activated colloidal surfers". In: *Science*, p. 1230020.
- Parrish, J. K., S. V. Viscido, and D. Grünbaum (2002). "Self-organized fish schools: an examination of emergent properties". In: *Biological Bulletin* 202.3, pp. 296–305.
- Peruani, F. et al. (2012). "Collective motion and nonequilibrium cluster formation in colonies of gliding bacteria". In: *Physical Review Letters* 108.9, p. 098102.
- Pitcher, T. J. (1983). "Heuristic definitions of fish shoaling behaviour". In: *Animal Behaviour*.
- Popkin, G. (2016). "The physics of life". In: *Nature News* 529.7584, p. 16.
- Potts, W. K. (1984). "The chorus-line hypothesis of manoeuvre coordination in avian flocks". In: *Nature* 309.5966, pp. 344–345.
- Procaccini, A. et al. (2011). "Propagating waves in starling, *Sturnus vulgaris*, flocks under predation". In: *Animal Behaviour* 82.4, pp. 759–765.
- Prokopenko, M. (2013). *Guided self-organization: Inception*. Vol. 9. Springer Science & Business Media.

- Prokopenko, M. and I. Einav (2015). "Information thermodynamics of near-equilibrium computation". In: *Physical Review E* 91.6, p. 062143.
- Prokopenko, M., F. Boschetti, and A. J. Ryan (2009). "An information-theoretic primer on complexity, self-organization, and emergence". In: *Complexity* 15.1, pp. 11–28.
- Prokopenko, M. et al. (2011). "Relating Fisher information to order parameters". In: *Physical Review E* 84.4, p. 041116.
- Ramaswamy, S. (2017). "Active matter". In: *Journal of Statistical Mechanics: Theory and Experiment* 2017.5, p. 054002.
- Rørth, P. (2007). "Collective guidance of collective cell migration". In: *Trends in Cell Biology* 17.12, pp. 575–579.
- Rosenthal, S. B. et al. (2015). "Revealing the hidden networks of interaction in mobile animal groups allows prediction of complex behavioral contagion". In: *Proceedings of the National Academy of Sciences* 112.15, pp. 4690–4695.
- Sayama, H. (2010). "Robust morphogenesis of robotic swarms". In: *IEEE Computational Intelligence Magazine* 5.3, pp. 43–49.
- (2015). *Introduction to the modeling and analysis of complex systems*. Open SUNY Textbooks.
- Schreiber, T. (2000). "Measuring information transfer". In: *Physical Review Letters* 85.2, p. 461.
- Seifert, U. (2012). "Stochastic thermodynamics, fluctuation theorems and molecular machines". In: *Reports on Progress in Physics* 75.126001, p. 126001.
- Sekimoto, K. (2010). *Stochastic energetics*. Vol. 799. Springer.
- Sethna, J. (2006). *Statistical mechanics: entropy, order parameters, and complexity*. Vol. 14. Oxford University Press.
- Shankar, S. and M. C. Marchetti (2018). "Hidden entropy production and work fluctuations in an ideal active gas". In: *Physical Review E* 98.2, p. 020604.
- Shannon, C. E. (1948). "A mathematical theory of communication". In: *Bell System Technical Journal* 27.3, pp. 379–423.
- Sokolov, A. et al. (2009). "Enhanced mixing and spatial instability in concentrated bacterial suspensions". In: *Physical Review E* 80.3, p. 031903.
- Spinney, R. E. and I. J. Ford (2012a). "Entropy production in full phase space for continuous stochastic dynamics". In: *Physical Review E* 85.5, p. 051113.
- (2012b). "Nonequilibrium thermodynamics of stochastic systems with odd and even variables". In: *Physical Review Letters* 108.17, p. 170603.
- Szabo, B. et al. (2006). "Phase transition in the collective migration of tissue cells: experiment and model". In: *Physical Review E* 74.6, p. 061908.
- Theurkauff, I et al. (2012). "Dynamic clustering in active colloidal suspensions with chemical signaling". In: *Physical Review Letters* 108.26, p. 268303.
- Tomaru, T. et al. (2016). "Information transfer in a swarm of soldier crabs". In: *Artificial Life and Robotics* 21.2, pp. 177–180.
- Toner, J. and Y. Tu (1995). "Long-range order in a two-dimensional dynamical XY model: how birds fly together". In: *Physical Review Letters* 75.23, p. 4326.

- Tononi, G., O. Sporns, and G. M. Edelman (1994). "A measure for brain complexity: relating functional segregation and integration in the nervous system". In: *Proceedings of the National Academy of Sciences* 91.11, pp. 5033–5037.
- Ulieru, M. and R. Doursat (2011). "Emergent engineering: a radical paradigm shift". In: *International Journal of Autonomous and Adaptive Communications Systems* 4.1, p. 39.
- Vicsek, T. and A. Zafeiris (2012). "Collective motion". In: *Physics Reports* 517.3-4, pp. 71–140.
- Vicsek, T. et al. (1995). "Novel type of phase transition in a system of self-driven particles". In: *Physical Review Letters* 75.6, p. 1226.
- Walker, S. I. and P. C. Davies (2013). "The algorithmic origins of life". In: *Journal of the Royal Society Interface* 10.79, p. 20120869.
- Wang, X. R. et al. (2012). "Quantifying and tracing information cascades in swarms". In: *PLOS One* 7.7, e40084.
- Ward, A. J. et al. (2008). "Quorum decision-making facilitates information transfer in fish shoals". In: *Proceedings of the National Academy of Sciences* 105.19, pp. 6948–6953.
- Ward, A. J. et al. (2018). "Cohesion, order and information flow in the collective motion of mixed-species shoals". In: *Royal Society Open Science* 5.12, p. 181132.
- Zhao, X. et al. (2018). "Substrate-driven chemotactic assembly in an enzyme cascade". In: *Nature Chemistry* 10.3, p. 311.

Chapter 3

Informative and misinformative interactions in a school of fish

Swarm Intell (2018) 12:283–305
<https://doi.org/10.1007/s11721-018-0157-x>

Informative and misinformative interactions in a school of fish

Emanuele Crosato^{1,5} · Li Jiang^{2,3} · Valentin Lecheval^{3,4} · Joseph T. Lizier¹ · X. Rosalind Wang⁵ · Pierre Tichit³ · Guy Theraulaz³ · Mikhail Prokopenko¹

Received: 25 November 2017 / Accepted: 1 March 2018 / Published online: 8 March 2018

Abstract Quantifying distributed information processing is crucial to understanding collective motion in animal groups. Recent studies have begun to apply rigorous methods based on information theory to quantify such distributed computation. Following this perspective, we use transfer entropy to quantify dynamic information flows locally in space and time across a school of fish during directional changes around a circular tank, i.e. U-turns. This analysis reveals peaks in information flows during collective U-turns and identifies two different flows: an informative flow (positive transfer entropy) from fish that have already turned to fish that are turning, and a misinformative flow (negative transfer entropy) from fish that have not turned yet to fish that are turning. We also reveal that the information flows are related to relative position and alignment between fish, and identify spatial patterns of information and misinformation cascades. This study offers several methodological contributions and we expect further application of these methodologies to reveal intricacies of self-organisation in other animal groups and active matter in general.

Keywords Collective animal behaviour · Collective motion · Fish interactions · Information dynamics

Electronic supplementary material The online version of this article (<https://doi.org/10.1007/s11721-018-0157-x>) contains supplementary material, which is available to authorized users.

✉ Emanuele Crosato
emanuele.crosato@sydney.edu.au

- ¹ Complex Systems Research Group and Centre for Complex Systems, Faculty of Engineering & IT, The University of Sydney, Sydney, NSW 2006, Australia
- ² School of Systems Science, Beijing Normal University, Beijing 100875, People's Republic of China
- ³ Centre de Recherches sur la Cognition Animale, Centre de Biologie Intégrative (CBI), Centre National de la Recherche Scientifique (CNRS), Université Paul Sabatier (UPS), 31062 Toulouse Cedex 9, France
- ⁴ Groningen Institute for Evolutionary Life Sciences, Centre for Life Sciences, University of Groningen, Nijenborgh 7, 9747AG Groningen, The Netherlands
- ⁵ CSIRO Data61, PO Box 76, Epping, NSW 1710, Australia

1 Introduction

Collective motion is one of the most striking examples of aggregated coherent behaviour in animal groups, dynamically self-organising out of local interactions between individuals. It is observed in different animal species, such as schools of fish (Parrish et al. 2002; Sumpter et al. 2008), flocks of birds (Lissaman and Shollenberger 1970; May 1979; Ballerini et al. 2008; Bialek et al. 2012), colonies of insects (Buhl et al. 2006; Fourcassié et al. 2010; Buhl et al. 2010; Attanasi et al. 2014b; Buhl and Rogers 2016) and herds of ungulates (Ginelli et al. 2015). There is an emerging understanding that information plays a *dynamic* role in such a coordination (Sumpter et al. 2008), and that *distributed* information processing is a specific mechanism that endows the group with collective computational capabilities (Bonabeau et al. 1999; Couzin 2009; Albantakis et al. 2014).

Information transfer is of particular relevance for collective behaviour, where it has been observed that small perturbations cascade through an entire group in a wave-like manner (Potts 1984; Procaccini et al. 2011; Herbert-Read et al. 2015; Attanasi et al. 2015), with these cascades conjectured to embody information transfer (Sumpter et al. 2008). This phenomenon is related to underlying causal interactions, and a common goal is to infer physical interaction rules directly from experimental data (Katz et al. 2011; Gautrais et al. 2012; Herbert-Read et al. 2011) and measure correlations within a collective.

Nagy et al. (2010) used a variety of correlation functions to measure directional dependencies between the velocities of pairs of pigeons flying in flocks of up to 10 individuals, extended to 30 in Nagy et al. (2013), reconstructing the leadership network of the flock. As has been shown, this network does not correspond to the dominance hierarchy between birds (Nagy et al. 2013). Information transfer has been extensively studied in flocks of starlings, by observing the propagation of direction changes across the flocks (Cavagna et al. 2013b, a; Attanasi et al. 2014a). More recently, Rosenthal et al. (2015) attempted to determine a communication structure of a school of fish during its collective evasion manoeuvres manifested through cascades of behavioural change. A functional mapping between sensory inputs and motor responses was inferred by tracking fish position and body posture, and calculating visual fields.

The main scientific question we address is how to identify and quantify information processing during decision-making in groups (Giardina 2008; Attanasi et al. 2014a), exacerbated by misinformative and noisy data. In trying to obtain such understanding, it is important to develop predictive models of information propagation among individuals, including behavioural cascades. Specifically, we aim to reveal how information propagates within a group and affects collective decisions (e.g., choosing a common travelling direction). This would provide an objective way to use such information for predictive modelling of behavioural reactions in response to various inputs.

Rather than consider *semantic* or *pragmatic* information, many contemporary studies employ rigorous information-theoretic measures that quantify information as uncertainty reduction, following Shannon (Cover and Thomas 2006), in order to deal with the stochastic, continuous and noisy nature of intrinsic information processing in natural systems (Feldman et al. 2008). Distributed information processing is typically dissected into three primitive functions: the *transmission*, *storage* and *modification of information* (Langton 1990). *Information dynamics* is a recent framework characterising and measuring each of the primitives information-theoretically (Lizier et al. 2014; Lizier 2013). In viewing the state update dynamics of a random process as an information processing event, this framework performs an *information regression* in accounting for where the information to predict that state update

can be found by an observer, first identifying predictive information from the past of the process as *information storage*, then predictive information from other sources as *information transfer* (including both pairwise transfer from single sources, and higher-order transfers due to multivariate effects). The framework has been applied to modelling collective behaviour in several complex systems, such as Cellular Automata (Lizier et al. 2008, 2010, 2012), Ising spin models (Barnett et al. 2013), Genetic Regulatory Networks and other biological networks (Lizier et al. 2011b; Prokopenko et al. 2011; Faes and Porta 2014), and neural information processing (Gómez et al. 2014; Wibral et al. 2015).

This study proposes a domain-independent, information-theoretic approach to detecting and quantifying individual-level dynamics of information transfer in animal groups using this framework. This approach is based on transfer entropy (Schreiber 2000), an information-theoretic measure that quantifies the directed and time-asymmetric predictive effect of one random process on another. We aim to characterise the dynamics of how information transfer is conducted in space and time within a *biological* school of fish (*Hemigrammus rhodostomus* or rummy-nose tetras, Fig. 1a). Using transfer entropy allows us to consider specifically the information dynamics during collective decision-making, identifying predictive information flows and their spatial patterns, complementing our parallel study which used correlation analysis to identify influential neighbours (Jiang et al. 2017).

We stress that the predictive information transfer should be considered from the observer perspective, that is, it is the observer that gains (or loses) predictability about a fish motion, having observed another fish. In other words, notwithstanding possible influences among the fish that could potentially be reflected in their information dynamics, our quantitative analysis focuses on the information flow within the school which is detectable by an external observer, captured by the transfer entropy. This means that, whenever we quantify a predictive information flow *based on* a source fish *about* a destination fish, we attribute the change of predictability (uncertainty) to a third party, be it another fish in the school, a predator approaching the school or an independent experimentalist. To improve readability, we refer to flow or transfer *from* a source fish *to* a destination fish. Importantly, this predictive information flow may or may not account for the causal information flow affecting the source and the destination (Ay and Polani 2008; Lizier and Prokopenko 2010)—however it does typically indicate presence of causality, either within the considered pair or from some common cause.

We focus on collective direction changes, i.e., collective U-turns, during which the directional changes of individuals progress in a rapid cascade, at the end of which a coherent motion is re-established within the school. Sets of different U-turns are comparable across experiments under the same conditions, permitting a statistically significant analysis involving an entire set of U-turns.

By looking at the *pointwise* or *local* values of transfer entropy over time, rather than at its average values, we are not only able to detect information transfer, but also to observe its dynamics over time and across the school. We demonstrate that information is indeed constantly flowing within the school, and identify the source-destination lag where predictive information flow is maximised (which has an interpretation as an observer-detectable reaction time to other fish). The information flow is observed to peak during collective directional changes, where there is a typical “cascade” of predictive gains and losses to be made by observers of these pairwise information interactions. Specifically, we identify two distinct predictive information flows: (i) an “informative” flow, characterised by positive local values of transfer entropy, from fish that have already changed direction to fish that are turning, and (ii) a “misinformative” flow, characterised by negative local values of transfer entropy, from fish that have not changed direction yet to the fish that are turning. Finally, we identify spatial patterns coupled with the temporal transfer entropy, which we call spatio-informational

motifs. These motifs reveal spatial dependencies between the source of information and its destination, which shape the directed pairwise interactions underlying the informative and misinformative flows. The strong distinction revealed by our quantitative analysis between informative and misinformative flows is expected to have an impact on modelling and understanding the dynamics of collective animal motion.

2 Information-theoretic measures for collective motion

The study of Wang et al. (2012) introduced the use of transfer entropy to investigations of collective motion. This work quantitatively verified the hypothesis that information cascades within an (artificial) swarm can be spatiotemporally revealed by *conditional transfer entropy* (Lizier et al. 2008, 2010) and thus correspond to communications, while the collective memory can be captured by *active information storage* (Lizier et al. 2012).

Richardson et al. (2013) applied related variants of conditional mutual information, a measure of non-linear dependence between two random variables, to identify dynamical coupling between the trajectories of foraging meerkats. Transfer entropy has also been used to study the response of schools of zebrafish to a robotic replica of the animal (Butail et al. 2014; Ladu et al. 2015), and to infer leadership in pairs of bats (Orange and Abaid 2015) and simulated zebrafish (Butail et al. 2016). Lord et al. (2016) also posed the question of identifying individual animals which are directly interacting with other individuals, in a swarm of insects (*Chironomus riparius*). Their approach used conditional mutual information (called “causation entropy” although it does not directly measure causality (Lizier and Prokopenko 2010)), inferring “information flows” within the swarm over moving windows of time.

Unlike the study of Wang et al. (2012), the above studies quantified average dependencies over time rather than local dependencies at specific time points; for example, leadership relationships in general rather than their (local) dynamics over time. Local versions of transfer entropy and active information storage have been used to measure pairwise correlations in a “swarm” of soldier crabs, finding that decision-making is affected by the group size (Tomaru et al. 2016). Statistical significance was not reported, presumably due to a small sample size. Similar techniques were used to construct interaction networks within teams of simulated RoboCup agents (Cliff et al. 2017).

In this study, we focus on local (or pointwise) transfer entropy (Schreiber 2000; Lizier et al. 2008; Lizier 2014b) for specific samples of time series processes of fish motion, which allows us to reconstruct the dynamics of information flows over time. Local transfer entropy (Lizier et al. 2008), captures information flow from the realisation of a *source* variable Y to a *destination* variable X at time n . As described in Sect. 5, local transfer entropy is defined as the information provided by the source $\mathbf{y}_{\mathbf{n}-\mathbf{v}} = \{y_{n-v}, y_{n-v-1}, \dots, y_{n-v-l+1}\}$, where v is a time delay and l is the history length, about the destination x_n in the context of the past of the destination $\mathbf{x}_{\mathbf{n}-1} = \{x_{n-1}, x_{n-2}, \dots, x_{n-k}\}$, with a history length k :

$$t_{y \rightarrow x}(n, v) = \log_2 \frac{p(x_n | \mathbf{x}_{\mathbf{n}-1}, \mathbf{y}_{\mathbf{n}-\mathbf{v}})}{p(x_n | \mathbf{x}_{\mathbf{n}-1})}. \quad (1)$$

Importantly, local values of transfer entropy can be negative, while the average transfer entropy is non-negative. Negative values of the local transfer entropies indicate that the source is *misinformative* about the next state of the destination (i.e., it increases uncertainty). Previous studies that used average measures over sliding time windows in order to investigate how information transfer varies over time (Richardson et al. 2013; Lord et al. 2016) cannot detect misinformation because they measure average but not local values.

As an observational measure, transfer entropy does not measure causal effect of the source on the target; this can only be established using interventional measures (Ay and Polani 2008; Lizier and Prokopenko 2010; Chicharro and Ledberg 2012; Smirnov 2013). Rather, transfer entropy measures the predictive information gained from a source variable about the state transition in a target, which may be viewed as *information transfer* when measured on an underlying causal interaction (Lizier and Prokopenko 2010). It should be noted that while some researchers may be initially more interested in causality, the concept of information transfer reveals much about the dynamics that causal effect does not (Lizier and Prokopenko 2010), in particular being associated with emergent local structure in dynamics in complex systems (Lizier et al. 2008; Wang et al. 2012) and with changes in behaviour, state or regime (Boedecker et al. 2012; Barnett et al. 2013), as well as revealing the misinformative interactions described above. As a particular example, local transfer entropy spatiotemporally highlights emergent glider entities in cellular automata (Lizier et al. 2008), which are analogues of cascading turning waves in swarms (also highlighted by transfer entropy (Wang et al. 2012)), while local measures of causality do not differentiate these from the background dynamics (Lizier and Prokopenko 2010).

In general, to understand the processes that govern collective behaviour in animal groups, it is important to disentangle the interactions between fish, how these interactions are combined and how interrelated are the individual behaviours. This can be achieved much more easily by investigating collective behaviour in small groups of individuals. Such methodology (Gautrais et al. 2012; Weitz et al. 2012) has been successfully applied to studies of the individual-level interactions involved in several examples of collective animal behaviour: aggregation in cockroaches (Jeanson et al. 2004, 2005), division of labour, corpse aggregation and nest building in ant colonies (Theraulaz et al. 2002a,b; Khuong et al. 2016), collective motion in groups of pelagic fish (Gautrais et al. 2012; Weitz et al. 2012) and collective motion in human groups (Moussaïd et al. 2009, 2011). Although for schools of minnows (*Phoxinus phoxinus*), two fish schools are qualitatively different from schools containing three or more, the effects seem to level off by the time the school reaches a size of six individuals (Partridge 1980). Collective behaviour, as well as a stereotypical “phase transition”, when an increase in density leads to the onset of directional collective motion, have also been detected in small groups of six glass prawns (*Paratya australiensis*) (Mann et al. 2012). Furthermore, at such intermediate group sizes, it has been observed that multiple fish interactions could often be faithfully factorised into pair interactions in one particular species of fish (Gautrais et al. 2012). The rationale for choosing a limited number of fish is also strengthened by the fact that it allowed us to quantify both the dynamics of collective decisions at the group level and the predictive information flow, while preserving the coordination of swimming in this species that exhibits strong schooling behaviour.

In our study, we investigated information transfer within a school of fish during specific collective direction changes, i.e., U-turns, in which the school collectively reverses its direction. Groups of five fish were placed in a ring-shaped tank (Fig. 1b), a design conceived to constrain fish swimming circularly, with the possibility of undergoing U-turns spontaneously, without any obstacles or external factors. A similar well-controlled environment has been previously successfully used in studies of groups of locusts (Buhl et al. 2006), enabling a large number of replicates which for obvious reasons cannot be done in a natural environment. The choice of a small and cohesive group allows us to focus on pairwise interactions in the context of collective motion, while studies of larger and less cohesive groups could reveal dominance hierarchies and leader-follower relationships, as well as social influence in groups.

In many species of fish, sudden collective changes of the state of a school may happen without external cause as a consequence of stochastic effects (Tunstrøm et al. 2013). In these

cases, local behavioural changes of a single individual can lead to large transitions between collective states of the school, such as between the schooling and milling states (Calovi et al. 2015). Determining how fluctuations in individual behaviour, for instance in heading direction, propagate within a group is a key to understanding transitions between collective states in fish schools and in animal groups in general. In our setup, fish swim in a highly synchronised and polarised manner, and can only head in two directions, clockwise or anticlockwise, regularly switching from one to the other. Our work thus also allows us to analyse in groups how individual U-turns occur, propagate through the group, and ultimately lead to collective U-turns. A total of 455 U-turns have been observed during 10 trials of 1 h duration each.

We computed local transfer entropy between each (directed) pair of fish from time series obtained from fish heading. Specifically, the destination process X was defined as the directional change of the destination fish, while the source process Y was defined as the relative heading of the destination fish with respect to the source fish (see Sect. 5). This allowed us to capture the influence of the source-destination fish alignment on the directional changes of the destination. Such influence is usually delayed in time and we estimated the optimal delay [maximising $\langle t_{y \rightarrow x}(n, v) \rangle_n$ (Wibral et al. 2013), see Sect. 5] at $v = 6$, corresponding to 0.12 s. The relative heading is not the only aspect of fish motion and other components can be considered, such as speed and acceleration. Moreover, a heading change in the destination fish could be related to many aspects of the source fish motion other than the heading difference between the two fish. However, in this study, we focus on the relative heading, which we believe is one of the most relevant features to explore in our controlled setup. In order to simplify our terminology, in the remainder of the text, we shall refer to this partial information flow, based on relative heading, as the information flow. It is important to clarify that large direction changes do not imply high values of transfer entropy, even while such measure is based on the heading. Large direction changes (where they are rare events) may have more capacity for information flow; however, there is not necessarily large information flow at these events unless the source fish are predictive of their occurrence.

3 Results

3.1 Information flows during U-turns

In order to represent the school's orientation around the tank, we define its polarisation as its circumferential velocity component, so that it is positive when the school is swimming clockwise and negative when it is swimming anticlockwise (see Sect. 5). The better the school's average heading is aligned with an ideal circular trajectory around the tank, the higher is the intensity of the polarisation. When the school is facing one of the tank's walls, for example in the middle of a U-turn, the polarisation is zero, and the polarisation flips sign during U-turns. Polarisation allows us to map local values of transfer entropy onto the progression of the collective U-turns.

The analyses of transfer entropy over time reveal that the measure clearly diverges from its baseline in the vicinity of U-turns, as shown in the representative U-turn in Fig. 1c (Supplementary Figure S1 shows a longer time interval during which several U-turns can be observed). The figure shows that during regular circular motion, when the school's polarisation is highly pronounced, transfer entropy is low. As the polarisation approaches zero the intensity of transfer entropy grows, peaking near the middle of a U-turn, when polarisation switches its sign.

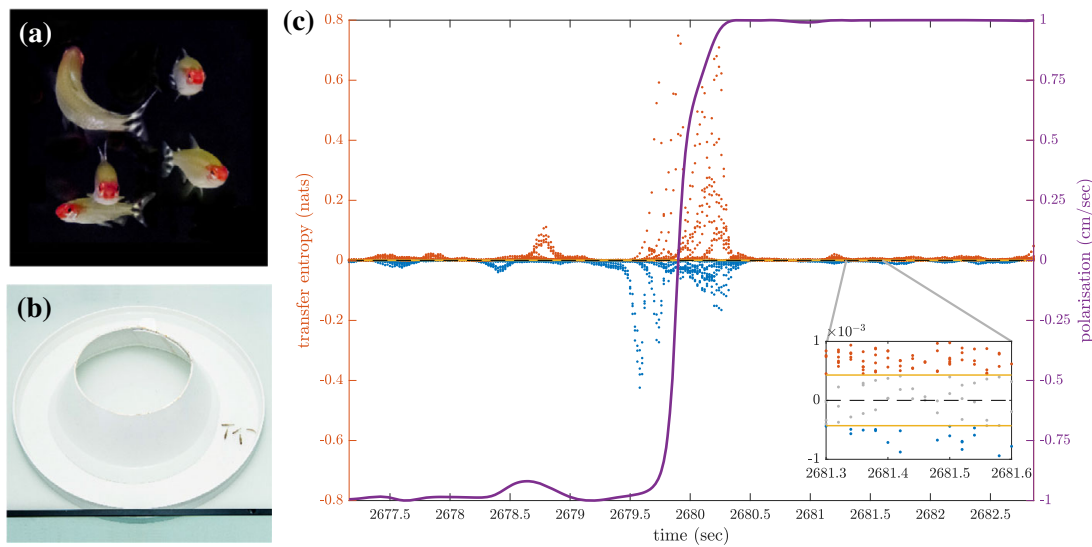


Fig. 1 Transfer entropy within the school during a U-turn. **a** Is a photo of a spontaneous U-turn initiated by a single fish in a group of five *Hemigrammus rhodostomus* fish. **b** Shows the experimental ring-shaped tank. **c** Plots the school's polarisation during a U-turn and the detected transfer entropy over a time interval of approximately 6 s. The purple line represents the school's polarisation, while dots represent local values of transfer entropy between all directed pairs of fish: red dots represent positive transfer entropy and blue dots represent negative transfer entropy. Time is discretised in steps of length 0.02 s and for each time step 20 points of these local measures are plotted, for the 20 directed pairs formed out of 5 fish. The yellow lines in the inset are the thresholds for considering a value of transfer entropy statistically different from zero ($p < 0.05$ before false discovery rate correction, see Sect. 5). Grey dots between these lines represent values that are not statistically different from zero. Credits to David Villa ScienceImage/CBI/CNRS, Toulouse, 2015, for **a**, **b** (Color figure online)

We clarify that the aim here is *not* to establish transfer entropy as an alternative to polarisation for detecting turn; rather, our aim is to use polarisation to describe the overall progression of the collective U-turns and then to use transfer entropy to investigate the underlying information flows in the dynamics of such turns. Indeed, transfer entropy is found to be statistically different from zero at many points outside of the U-turns (see Supplementary Figure S1), although the largest values and most concentrated regions of these are during the U-turns. This indicates that information transfer, based on the heading direction's change, occurs even when fish school together without changing direction; we know that the fish are not executing precisely uniform motion during these in-between periods, and so interpret these small amounts of information transfer as sufficiently underpinning the dynamics of the group maintaining its collective heading. We would like to also point out that information processing during the aligned motion mostly corresponds to the information storage, which can be detected using other information-theoretic measures (Lizier et al. 2012), while only a low information transfer is needed to maintain the alignment.

We also see in Fig. 1c that both positive and negative values of transfer entropy are detected. In order to understand the role of the positive and negative information flows during collective motion, in the next section, we show the dynamics of transfer entropy for individual pairwise interactions.

3.2 Informative and misinformative flows

Our analysis revealed a clear relationship between positive and negative values of transfer entropy and the sequence of individual fish turning, which is illustrated in Fig. 2. Figure 2a

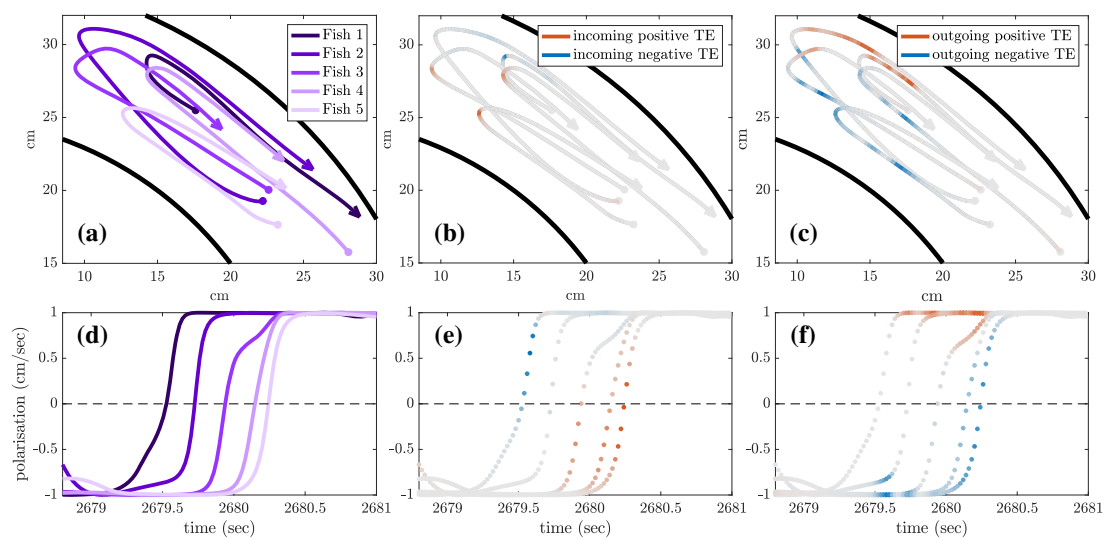


Fig. 2 Positive and negative information flows during a U-turn. **a** Shows the trajectories of the five fish during the U-turn shown in **c**. The two black lines are the inner and the outer walls of the tank, and each of the five trajectories coloured in different shades of purple correspond to a different fish: from darkest purple for the first fish turning (Fish 1), to the lightest purple for the last (Fish 5). The total time interval is approximately 2 s, during which all fish turn from swimming anticlockwise to clockwise. **d** Depicts the polarisations of the five fish, showing the temporal sequence of fish turns. **b** Shows the fish trajectories again, but this time indicates the value of the *incoming* local transfer entropy to each fish as a destination, averaged over the other four fish as sources. The colour of each trajectory changes as the fish turn: strong red indicates intense positive transfer entropy; strong blue indicates intense negative transfer entropy; intermediate grey indicates that transfer entropy is close to zero. **e** Is obtained analogously to **b**, but the polarisations of the individual fish are shown rather than their trajectories. **c**, **f** mirror **b**, **e**, but where the direction of the transfer entropy has been inverted: the colour of each trajectory or polarisation now indicates the value of the *outgoing* local transfer entropy from each fish as a source, averaged over the other four fish as destinations (Color figure online)

shows the trajectories of individual fish during the same U-turn depicted in Fig. 1c. These trajectories are retraced in Fig. 2d in terms of polarisation of each fish. It is quite clear that there is a well-defined sequence of individual U-turns during the collective U-turn. Moreover, Fig. 2 shows how the transfer entropy maps onto the fish trajectories, both from the fish whose trajectory is traced as a source to the other four fish—i.e., *outgoing* transfer entropy—and, vice versa, from the other four fish to the traced one as a destination—i.e., *incoming* transfer entropy.

The incoming transfer entropy clearly peaks during the destination fish's individual turns and its local values averaged over all sources go from negative, for the first (destination) fish that turns, to positive for the last fish turning (Fig. 2b, e). In the opposite direction, the outgoing transfer entropy (averaged over all destinations) displays negative peaks only before the source fish has turned, and positive peaks only afterwards (Fig. 2c, f). Figure 2 suggests that predictive information transfer intensifies only when a destination fish is turning, with this transfer being informative from source fish that have already turned and misinformative from source fish that have not turned yet.

This phenomenon can be observed very clearly in Fig. 3a, b, which show the transfer entropy in both directions for a single fish (the second fish turning in Figs. 1 and 2). One positive peak of incoming transfer entropy (indicating informative flow) and three negative ones (misinformative flows) are detected when this fish, as a destination, is undergoing the U-turn (Fig. 3a). No other peaks are detected for this fish as a destination. On the other hand, one negative peak of outgoing transfer entropy is detected before the fish, this time as a

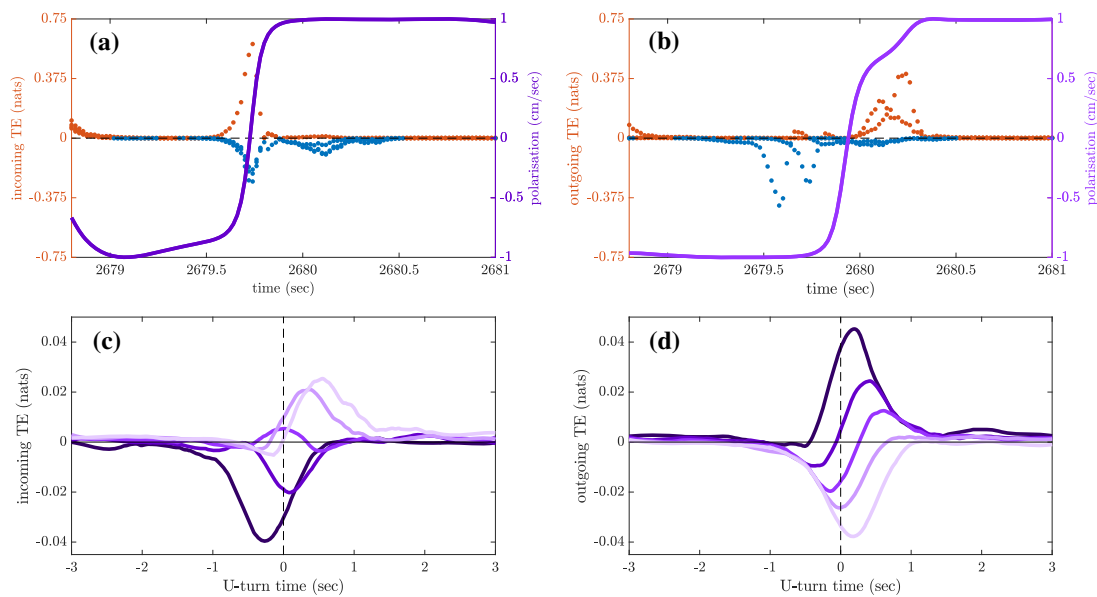


Fig. 3 **a** Shows the polarisation of the *second* fish turning, together with the incoming transfer entropy to that fish as the destination, with the other four fish as the sources: red dots represent positive values and blue dots represent negative values. **b** Mirrors **a**, but with the outgoing transfer entropy from that fish as the source, and the other four fish as destinations. In **c**, each purple line corresponds to a fish, with the shade again reflecting the order in which the fish turn (darkest for first fish to turn, and lightest for the last). Now however (in **c**), rather than corresponding to a single U-turn event, the incoming local transfer entropy (to each fish as a destination, averaged over the other four fish as sources) is averaged over all 455 observed U-turns and is shown as a function of time. The horizontal axis is the relative time of the U-turns, with zero being the time when the average polarisation of the school changes sign. **d** Mirrors **c**, but where the direction of the transfer entropy has been inverted (showing outgoing transfer from each fish in turning order) (Color figure online)

source, has turned, and three positive peaks are detected after the fish has turned (Fig. 3b). These four peaks occur respectively when the first, the third, the fourth and the fifth fish undergo the U-turn, as is evident by comparing Figs. 2d, 3b. A movie of the fish undergoing this specific U-turn is provided in Supplementary Video S1, while a detailed reconstruction of the U-turn, showing the dynamics of transfer entropy over time for each directed pair of fish, is provided in Supplementary Video S2.

In order to demonstrate that the phenomenon described here holds for U-turns in general, and not only for the representative one shown in Fig. 2, we performed an aggregated analysis of all 455 U-turns observed during the experiment. Since the order in which fish turn is not the same in every U-turn, in this analysis, we refer not to single fish as individuals, but rather to fish in the order in which they turn. Thus, when we refer, for instance, to “the first fish that turns”, we may be pointing to a different fish at each U-turn. It is worth noting that, in general, multiple fish can turn at the same time during a U-turn, but averaging over all U-turn events allows us to statistically order turning events, as shown in Fig. 3c, d.

The aggregated results are presented in Fig. 3c, d. Figure 3c shows that incoming transfer entropy peaks for each fish in turning order and gradually grows, from a minimum negative peak corresponding to the first fish turning, to a maximum positive peak corresponding to the last fish turning. Vice versa, Fig. 3d shows that outgoing transfer entropy peaks only positively for the first fish turning, which is an informative source about all other fish turning afterwards. For the last fish that turns the peak is negative, since this fish is misinformative about all other fish that have already turned. The second, third and fourth fish present both a

negative and a positive peak. The intensity of the negative peaks increases from the second fish to the fourth, while the intensity of the positive peak decreases.

In general, the source fish is informative about all destination fish turning after it and misinformative about any destination fish turning before it. This is because the prior turn of a source helps the observer to predict the later turn of the destination, whereas examining a source which has not turned yet itself is actively unhelpful (misinformative) in predicting the occurrence of such a turn. This also explains why, for a source, the negative peaks come before positives.

The sequential cascade-like dynamics of information flow suggests that the strongest sources of predictive information transfer are fish that have already turned. Moreover, our analyses reveal that once a fish has performed a U-turn, its behaviour in general ceases to be predictable based on the behaviour of other fish that swim in opposite direction (in fact such fish would provide misinformative predictions). This suggests an asymmetry of predictive information flows from and to an individual fish during U-turns.

3.3 Spatial motifs of information transfer

It is reasonable to assume that predictive information transfer in a school of fish results from spatial interactions among individuals. We investigated the role of pairwise spatial interactions in carrying the positive and negative information flows that we detected in the previous section, looking for spatial patterns of information and misinformation transfer.

In particular, we established the statistics of the relative position and heading of the destination fish relative to the source fish, at times when the transfer entropy from the source to the destination is more intense. For this purpose, we used radial diagrams (see Fig. 4) representing the relative data in terms of transfer entropy, focusing separately on their positive (informative) and negative (misinformative) values. In each diagram, we aggregate data from all 455 U-turns and all pairs. The diagrams show clear spatial patterns coupled with the transfer entropy, which we call spatio-informational motifs.

We see that positive information transfer is on average more intense from source fish to: (a) other fish positioned behind them (Fig. 4a, left), and (b) to fish with headings closer to perpendicular rather than parallel to them (Fig. 4a, right). We know from Figs. 2 and 3 that positive transfer entropy is detected from source fish that have already turned to destination fish that are turning. Thus, Fig. 4a suggests that a source is more informative about destination fish that are left behind it after a turn, most intensely when the destination fish are executing their own turning manoeuvre to follow the source. Directional relationships from individuals in front towards others that follow were observed in previous works on birds (Nagy et al. 2010), bats (Orange and Abaid 2015) and fish (Katz et al. 2011; Herbert-Read et al. 2011; Rosenthal et al. 2015).

For negative information transfer (Fig. 4b), we see a different spatio-informational motif. Negative information transfer is on average more intense to fish generally positioned at the side and with opposite heading. This aligns with Figs. 2 and 3 in that negative transfer entropy typically flows from fish that have not turned yet to those which are turning.

In summary, transfer entropy has a clear spatial signature, showing that the spatio-informational dependencies in the studied school of fish are not random but reflect specific interactions.

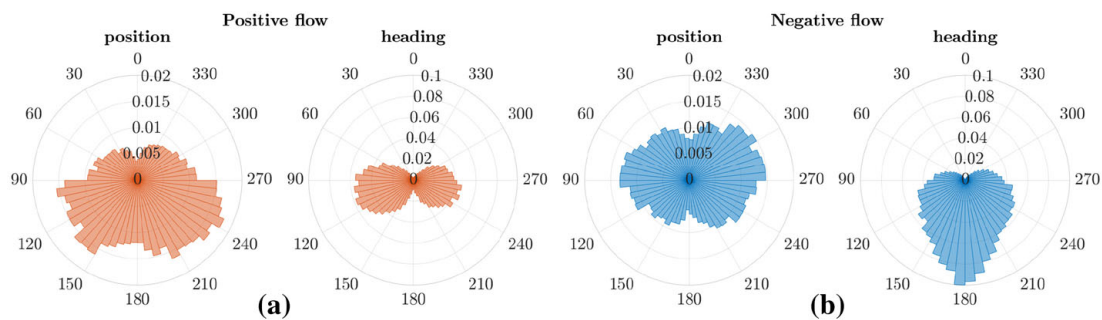


Fig. 4 Spatio-informational motifs. Each diagram is a circle centred on a source fish with zero heading, providing a reference. In each diagram, space is divided into 60 angular sectors measuring 6° . Within each circle, we group all pairwise samples from all 455 U-turns such that the source fish is placed in the centre and the destination fish is placed within the circle in one of the sectors. The left circles in **a**, **b** aggregate the relative positions of destination fish, while the right circles aggregate the relative headings of destination fish. The value of each radial sector (for both position and heading) represents the average of the corresponding values of either positive (**a**) or negative (**b**) transfer entropy. For example, the value in each sector of the left diagram of **a** represents the average positive transfer entropy for a destination fish, given it has relative position in that sector with respect to the source fish: all positive values of transfer entropy corresponding to each sector are summed and divided by the total number of values corresponding to that sector. The value in each sector of the right diagram of **a** represents the average positive transfer entropy for a destination fish, given that its heading diverges from the one of the source by an angle in that sector. **b** Mirrors **a** this negative transfer entropy. The source fish data are taken from the time points corresponding to the time delay ν with respect to the source

4 Discussion

Information transfer within animal groups during collective motion is hard to quantify because of implicit and distributed communication channels with delayed and long-ranged effects, selective attention (Riley and Leith 1976) and other species-specific cognitive processes. Here, we presented a rigorous framework for detecting and measuring predictive information flows during collective motion, by attending to the dynamic statistical dependence of directional changes in destination fish on relative heading of sources. This predictive information flow should be interpreted as a change (gain or loss) in predictability obtained by an observer. Importantly, the information-theoretic nature of the measure means that it is applicable to other stochastic interactions; more stochastic dynamics would require more data and suitable video capture resolution to identify the salient flows. Furthermore, one may consider methods of statistical mechanics and information thermodynamics in studies of collective motion distributed over large systems (Bialek et al. 2012; Crosato et al. 2018).

We studied *Hemigrammus rhodostomus* fish placed in a ring-shaped tank which effectively only allowed the fish to move straight ahead or turn back to perform a U-turn. The individual trajectories of the fish were recorded for hundreds of collective U-turns, enabling us to perform a statistically significant information-theoretical analysis for multiple pairs of source and destination fish. The experimental setup used in this study has been chosen to focus on collective U-turns, but the information flow analysis has been applied in other situations with more complex collective behaviour, e.g., general swarming behaviour with different constraints (Wang et al. 2012; Miller et al. 2014; Tomaru et al. 2016), and swarming behaviour with leaders (Sun et al. 2014; Orange and Abaid 2015).

Transfer entropy was used in detecting pairwise time delayed dependencies within the school. By observing the local dynamics of this measure, we demonstrated that predictive information flows intensify during collective direction changes—i.e., the U-turns—a hypothesis that until now was not verified in a real biological system. Furthermore, we identified

two distinct predictive information flows within the school: an informative flow from fish that have already preformed the U-turn to fish that are turning, and a misinformative flow from fish that have not preformed the U-turn yet to the fish that are turning.

We also explored the role of spatial dynamics in generating the influential interactions that carry the information flows, another well-known problem. In doing so, we mapped the detected values of transfer entropy against each fish's relative position and heading, identifying clear spatio-informational motifs. Importantly, the positive and negative predictive information flows were shown to be associated with specific spatial signatures of source and destination fish. For example, positive information flow is detected when the source fish is in front of the destination, similarly to what was already observed in previous works on animals (Nagy et al. 2010; Katz et al. 2011; Herbert-Read et al. 2011; Rosenthal et al. 2015; Orange and Abaid 2015). The identified sequential cascade-like dynamics of information flow is well-pronounced, suggesting that this pattern will be retained in larger schools—this however remains a subject of future research.

In a previous work, the analysis of short-term directional correlations between fish on the same experimental data has shown that, when the group changes direction, individual fish respond to a limited number of influential neighbours, typically one or two which are not necessarily the closest ones (Jiang et al. 2017). Moreover, fish continuously change who they are interacting with. In this study, using a complementary approach, we show that a fish that has just made a U-turn may also decide to ignore the input of other fish moving in opposite direction (which is shown by the misinformation flow). A fish can thus choose to move in the opposite direction of the majority. This suggests that the behavioural tendency of a fish to align in the direction of the majority of its neighbours, which is a manifestation of social conformity and implemented in most models of collective motion, can be “shut down” for some time. When these events occur, a fish can temporary take the lead of a group thanks to the behavioural contagion. Our analysis provides a way to create a quantitative model for predictive information flow between fish and thus brings a better understanding of the processes underlying collective decisions in fish groups and animal groups in general.

Local transfer entropy as it was applied in this study reveals the dynamics of *pairwise* information transfer. It is well known that multivariate extensions to the transfer entropy, e.g., conditioning on other information sources, can be useful in terms of eliminating redundant pairwise relationships while also capturing higher-order relationships beyond pairwise (i.e. synergies) (Lizier et al. 2008, 2010; Lizier and Prokopenko 2010; Vakorin et al. 2009; Williams and Beer 2011; James et al. 2016), and as such the identification of *effective* neighbourhoods cannot be accurately inferred using pairwise relationships alone. Transfer entropy comprises both (i) a unique component from the source, and (ii) a synergistic component from the source in the context of the target, as has been clarified by Williams and Beer (2011), among others. While we can learn more by measuring these components separately (for which well accepted measures have not yet been developed), both capture important aspects of the concept of information transfer. Thus, we argue that focussing on the unique component alone would not align with the popularly understood concept of information transfer. Improvements are possible by adapting algorithms for deciding when to include higher-order multivariate transfer entropies (and which variables to condition on), developed to study effective networks in brain imaging data sets (Lizier and Rubinov 2012; Faes et al. 2011; Marinazzo et al. 2012; Stramaglia et al. 2012), to collective animal behaviour, as such methods can eliminate redundant connections and detect synergistic effects. Whether or not such algorithms will prove useful for swarm dynamics is an open research question, with conflicting findings that first suggest that multiple fish interactions could be faithfully fac-

torised into simply pair interactions in one species (Gautrais et al. 2012) but conversely that this may not necessarily generalise (Katz et al. 2011).

In any case, such adaptations to capture multivariate effects will be non-trivial, as it must handle the short-term and dynamic structure of interactions across the collective. Early attempts have been made using (a similar measure to) conditional TE—on average over time windows—in collectives under such algorithms (Lord et al. 2016); however, it remains to be seen what such measures reveal about the collective dynamics on a local scale.

In summary, we have proposed a novel information-theoretic framework for studying the dynamics of information transfer in collective motion and applied it to a school of fish, without making any specific assumptions on fish behavioural traits and/or rules of interaction. This framework can be applied to studies of other biological collective phenomena, such as swarming and flocking, artificial multi-agent systems and active matter in general.

5 Methods

A general scheme of the methodology from the animal experimentation to the analysis of the information flows is provided in Fig. 5.

5.1 Experimental procedures

A group of 70 *Hemigrammus rhodostomus* (rummy-nose tetras) were purchased from Amazonie Labège (<http://www.amazonie.com>) in Toulouse, France. Fish were kept in 150 L aquariums on a 12:12 h, dark/light photoperiod, at $27.7^{\circ}\text{C}(\pm 0.5^{\circ}\text{C})$ and were fed *ad libitum* with fish flakes. Body lengths of the fish used in these experiments were on average 31 mm (± 2.5 mm). This species was chosen because it exhibits a strong schooling behaviour and it is very easy to handle in controlled conditions (Jiang et al. 2017; Lecheval et al. 2017). Moreover, individuals perform a *burst-and-coast* type of swimming that involves sharp directional changes, which implies a series of separate behavioural decisions in time and space. It is likely that an information-theoretic analysis would be able to better temporally resolve information flows associated with these transitions, as compared to more continuous dynamics [such as exhibited by *Khulia mugil* (Gautrais et al. 2012)].¹ Further, when swimming in a ring-shaped tank, schools of rummy-nose tetra hardly ever split despite collective U-turns, because of their social interactions relying on attraction and alignment (Calovi et al. 2018).

The experimental tank measured 120×120 cm was made of glass and set on top of a box to isolate fish from vibrations. The setup, placed in a chamber made by four opaque white curtains, was surrounded by four LED light panels giving an isotropic lighting. A ring-shaped tank made from two tanks (an outer wall of radius 35 cm and an inner wall, a cone of radius 25 cm at the bottom; both shaping a corridor of 10 cm) was set inside the experimental tank filled with 7 cm of water of controlled quality (50% of water purified by reverse osmosis and 50% of water treated by activated carbon) heated at $28.1^{\circ}\text{C}(\pm 0.7^{\circ}\text{C})$. The conic shape of the inner wall has been chosen to avoid the occlusion on videos of fish swimming too close to the inner wall that would occur with straight walls.

Five fish were randomly sampled from their breeding tank for each trial. Fish were ensured to be used only in one experiment per day at most. Fish were left for 10 min to habituate

¹ With that said, we have also shown that techniques employed in this study are also successful in identifying information flows in groups with smoother motion dynamics (Wang et al. 2012; Miller et al. 2014).

before the start of the trial. A trial consisted of 1 h of fish swimming freely (i.e. without any external perturbation).

5.2 Data extraction and pre-processing

Fish trajectories were recorded by a Sony HandyCam HD camera filming from above the setup at 50Hz (50 frames per second) in HDTV resolution ($1920 \times 1080\text{p}$). Videos were converted from MTS to AVI files with the command-line tool FFmpeg 2.4.3. Positions of fish on each frame were tracked with the tracking software idTracker 2.1 (Pérez-Escudero et al. 2014).

When possible, missing positions of fish have been manually corrected, only during the collective U-turn events detected by the sign changes of polarisation of the fish groups. The corrections have involved manual tracking of fish misidentified by idTracker as well as interpolation or merging of positions in the cases where only one fish was detected instead of several because they were swimming too close from each other for a long time. All sequences less or equal than 50 consecutive missing positions were interpolated. Larger sequence of missing values have been checked by eye to check whether interpolating was reasonable or not—if not, merging positions with closest neighbours was considered. All tracked positions have been monitored by eye during all U-turn events to make sure that any manual corrections improved the quality of the data set of positions. The entire process of manual changes and eye checking required three weeks. No U-turns were omitted from the analysis.

Time series of positions has been converted from pixels to metres, and the origin of the coordinate system $\mathcal{O}(0, 0)$ has been set to the centre of the ring-shaped tank. The resulting data set contains 9273720 data points (1854744 for each fish) from all the ten trials. Velocity was numerically derived from position using the symmetric difference quotient two-point estimation (Larson 1983). Heading was then computed as the four-quadrant inverse tangent of velocity and used to compute transfer entropy.

5.3 Polarisation

The polarisation is used to represent the orientation of a fish or of the whole school around the tank, which can be clockwise or anticlockwise, and is defined as the circumferential velocity component of the velocity a fish or of the whole school. Let Z and \dot{Z} be the two-dimensional position and normalised velocity of a fish, defined as Cartesian vectors with the centre of the tank being the origin—in case of the whole school, Z and \dot{Z} are averaged over all fish. The fish direction along an ideal circular clockwise rotation is described by a unit vector $z = \frac{\omega \times Z}{|\omega \times Z|}$, where ω is a vector orthogonal to plane of the rotation, chosen using the left-hand rule. In other words, z is the azimuthal unit vector of the fish heading θ .

The polarisation is defined as $\dot{Z} \cdot z$, so that it is positive when the fish is swimming clockwise and negative when it is swimming anticlockwise. Also, the better \dot{Z} is aligned with z or $-z$, the higher is the intensity of the polarisation. On the contrary, as \dot{Z} deviates from z or $-z$, the polarisation decreases and eventually reaches zero when \dot{Z} and z are orthogonal. As a consequence, during a U-turn, the intensity of the polarisation decreases and becomes zero at least once, before it increases again with the opposite sign.

5.4 Local transfer entropy

Transfer entropy (Schreiber 2000) is defined in terms of Shannon entropy, a fundamental measure in information theory (Cover and Thomas 2006) that quantifies the uncertainty of random

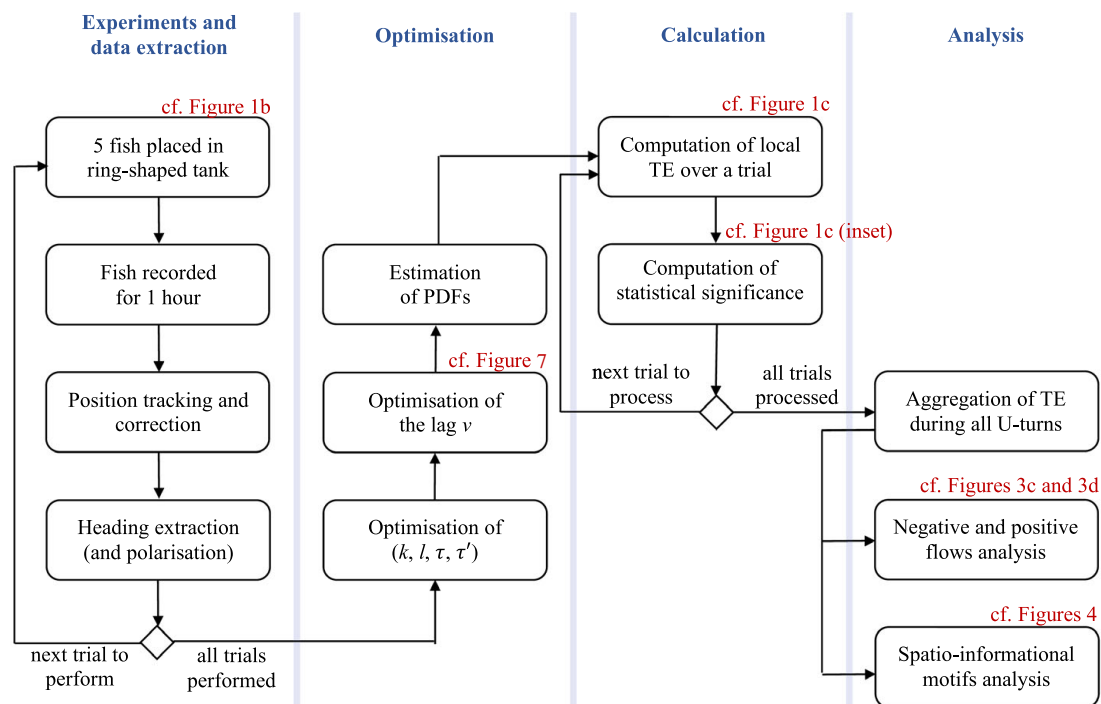


Fig. 5 General scheme of the methodology. For each of the 10 trials of the experiment, 5 fish were randomly chosen and placed in the ring-shaped tank, where they were let swimming free for 1 h while being recorded. The position of the fish over time was tracked from the produced videos, and the heading extracted from the position. Once all time series of fish heading were created, for each fish and during all trials, they were used in order to optimise the embedding parameters and the lag between source and destination fish. The conditional probability distributions involved in the computation of transfer entropy were then estimated using the heading time series of all trials. Subsequently, the local values of transfer entropy at each time step were calculated for each trial, as well as their statistical significance. Finally, local values of transfer entropy around every U-turns were aggregated in order to perform the analyses presented in Sects. 3.2 and 3.3

variables. Shannon entropy of a random variable X is $H(X) = -\sum_{x \in X} p(x) \log_2 p(x)$, where $p(x)$ is the probability of a specific instance x of X . $H(X)$ can be interpreted as the minimal expected number of bits required to encode a value of X without losing information. The joint Shannon entropy between two random variables X and Y is $H(X, Y) = -\sum_{x \in X} \sum_{y \in Y} p(x, y) \log_2 p(x, y)$, where $p(x, y)$ is the joint probability of instances x of X and y of Y . This quantity allows the definition of conditional Shannon entropy as $H(X|Y) = H(X, Y) - H(X)$, which represents the uncertainty of X knowing Y .

In this study, we are interested in local (or pointwise) transfer entropy (Fano 1961; Lizier 2014b) for specific instances of time series processes of fish motion, which allows us to reconstruct the dynamics of information flows over time. Shannon information content of an instance x_n of process X at time n is defined as $h(x_n) = -\log_2 p(x_n)$. The quantity $h(x_n)$ is the information content attributed to the specific instance x_n , or the information required to encode or predict that specific value. Conditional Shannon information content of an instance x_n of process X given an instance y_n of process Y is defined as $h(x_n|y_n) = h(x_n, y_n) - h(x_n)$.

Local transfer entropy is defined as the information provided by the source $\mathbf{y}_{n-v} = \{y_{n-v}, y_{n-v-1}, \dots, y_{n-v-l+1}\}$, where v is a time delay and l is the history length, about the destination x_n in the context of the past of the destination $\mathbf{x}_{n-1} = \{x_{n-1}, x_{n-2}, \dots, x_{n-k}\}$, with a history length k :

$$\begin{aligned}
 t_{y \rightarrow x}(n, v) &= h(x_n | \mathbf{x}_{n-1}) - h(x_n | \mathbf{x}_{n-1}, \mathbf{y}_{n-v}) \\
 &= \log_2 \frac{p(x_n | \mathbf{x}_{n-1}, \mathbf{y}_{n-v})}{p(x_n | \mathbf{x}_{n-1})}.
 \end{aligned} \tag{2}$$

Transfer entropy $T_{Y \rightarrow X}(v)$ is the average of the local transfer entropies $t_{y \rightarrow x}(n, v)$ over samples (or over n under a stationary assumption). The transfer entropy is asymmetric in Y and X and is also a dynamic measure (rather than a static measure of correlations) since it measures information in state transitions of the destination.

In order to compute transfer entropy here, the source variable Y and destination variable X are defined in terms of the fish heading. Specifically, X is the first-order divided difference (Newton's difference quotient) of the destination fish heading, while Y is the difference between the two fish headings at the same time. Let θ_S and θ_D be, respectively, the heading time series of the source and the destination fish. We then construct variables X and Y as follows, for all time points n (cf. Fig. 6 for an illustration of the headings involved):

$$x_n = \theta_n^D - \theta_{n-1}^D, \tag{3}$$

$$y_n = \theta_n^D - \theta_n^S. \tag{4}$$

Thus, y_n represents the relative heading of the destination fish with respect to the source fish, while x_n represents the directional change of the destination fish. The variables were so defined in order to capture directional changes of the destination fish in relation to its alignment with the source fish, which is considered an important component of movement updates in swarm models (Reynolds 1987).

Given the definition of the variables (3) and (4), we computed local transfer entropy $t_{y \rightarrow x}(n, v)$ using Eq. (2), where v was determined as described in Sect. 5.5 that follows. The past state \mathbf{x}_{n-1} of the destination in transfer entropy was defined as a vector of an embedding space of dimensionality k and delay τ , with $\mathbf{x}_{n-1} = \{x_{n-1-j\tau}\}$, for $j = \{0, 1, \dots, k-1\}$. Finding optimal values for k and τ is also described in Sect. 5.5. The state of the source process \mathbf{y}_{n-v} was also defined as a vector of an embedding space whose dimensionality l and delay τ' were similarly optimised. The local transfer entropy $t_{y \rightarrow x}(n, v)$ computed on these variables therefore tells us how much information (l time steps of) the heading of the destination relative to the source adds to our knowledge of the directional change in the destination (some v time steps later), in the context of k past directional changes of the destination. We note that while turning dynamics of the destination may contain more entropy (as rare events), there will only be higher transfer entropy at these events if the source fish is able to add to the prediction of such dynamics.

Computing transfer entropy requires knowledge of the probabilities of x_n and y_n defined in (3) and (4). These are not known a priori, but the measures can be estimated from the data samples using existing techniques. In this study, this was accomplished by modelling the probability distribution function as a multivariate Gaussian distribution (making the transfer entropy proportional to the Granger causality (Barnett et al. 2009)). This technique is the simplest first-order estimation available and well applied for transfer entropy (Marinazzo et al. 2012). We used the JIDT software implementation (Lizier 2014a).

Also, we assume stationarity of behaviour and homogeneity across the fish, such that we can pool together all pairwise samples from all time steps, for all trials, maximising the number of samples available for the calculation of each measure. For performance efficiency, we make calculations of the local measures using 10 separate sub-sampled sets (sub-sampled evenly across the trials), then recombine into a single resultant information-theoretic data set.

 Springer

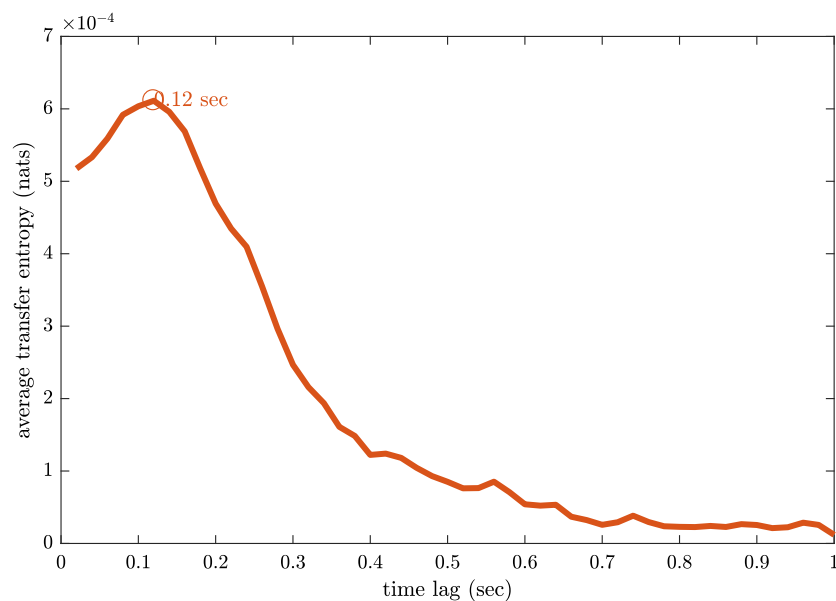


Fig. 7 Time lag optimisation. The red line represents the average transfer entropy (with $k = l = 3$, $\tau = \tau' = 1$) over all samples, as a function of the time delay between the source variable and the destination variable, for time delays between 0.02 and 1 s (1–50 time cycles) (Color figure online)

retains $p(x_n | \mathbf{x}_{n-1})$ and $p(\mathbf{y}_{n-v})$, but not $p(x_n | \mathbf{y}_{n-v}, \mathbf{x}_{n-1})$. Many surrogate measurements are repeated so as to construct a surrogate distribution under this null hypothesis of no directed relationship, and the transfer entropy estimate can then be compared in a statistical test against this distribution. For the average transfer entropy measured via the linear-Gaussian estimator, it is known that analytically the surrogates (in nats, and multiplied by $2 \times N$) asymptotically follow a χ^2 distribution with l degrees of freedom (Geweke 1982; Barnett and Bossomaier 2012). We use this distribution to confirm that the transfer entropy at the selected lag of 0.12 s (and indeed all lags tested) is statistically significant compared to the null distribution (at $p < 0.05$ plus a Bonferroni correction for the multiple comparisons across the 50 candidate lags).

Next, we introduce an extension of these methods in order to assess the statistical significance of the *local* values. This simply involves constructing surrogate transfer entropy measurements as before, however this time retaining the local values within those surrogate measurements and building a distribution of those surrogates. Measured local values are then statistically tested against this null distribution of local surrogates to assess their statistical significance.

We generated ten times as many surrogate local values as the number of actual local estimates, with a total of approximately 371 million local surrogates. This large set of surrogate local values was used to estimate p values of actual local values of the transfer entropy. If p value is sufficiently small, then the test fails and the value of the transfer entropy is considered significant (the value represents an actual relationship). The Benjamini and Hochberg (1995) procedure was used to select the p value cutoff while controlling for the false discovery rate under (N) multiple comparisons.

Data accessibility

The datasets generated and analysed during the current study are available from the corresponding author on request.

Author contributions GT designed research; VL, PT and GT performed research; VL, LJ, PT, RW and GT analysed data. EC, JL, RW and MP developed information dynamics methods, performed information-theoretic analysis, and identified information flows and motifs. EC designed, developed and run software for the information-theoretic analysis. GT, JL, EC and MP conceived and analysed information cascade. EC, JL and MP wrote the paper. GT and VL edited the manuscript and contributed to the writing.

Funding E.C. was supported by the University of Sydney’s “Postgraduate Scholarship in the field of Complex Systems” from Faculty of Engineering & IT and by a CSIRO top-up scholarship. L.J. was supported by a grant from the China Scholarship Council (CSC NO.201506040167). V.L. was supported by a doctoral fellowship from the scientific council of the University Paul Sabatier. This study was supported by grants from the Centre National de la Recherche Scientifique and University Paul Sabatier (project Dynabanc). J.L. was supported through the Australian Research Council DECRA grant DE160100630. M.P. was supported by The University of Sydney’s DVC Research Strategic Research Excellence Initiative (SREI-2020) project, “CRISIS: Crisis Response in Interdependent Social-Infrastructure Systems” (IRMA 194163). Sydney Informatics Hub at the University of Sydney provided access to HPC computational resources that have contributed to the research results reported within the paper.

Compliance with ethical standards

Conflict of interest The authors declare no conflict of interest.

Ethical standard All experiments have been approved by the Ethics Committee for Animal Experimentation of the Toulouse Research Federation in Biology N1 and comply with the European legislation for animal welfare.

References

- Albantakis, L., Hintze, A., Koch, C., Adami, C., & Tononi, G. (2014). Evolution of integrated causal structures in animats exposed to environments of increasing complexity. *PLOS Computational Biology*, 10(12), 1–19.
- Attanasi, A., Cavagna, A., Del Castello, L., Giardina, I., Grigera, T. S., Jelić, A., et al. (2014a). Information transfer and behavioural inertia in starling flocks. *Nature Physics*, 10(9), 691–696.
- Attanasi, A., Cavagna, A., Del Castello, L., Giardina, I., Jelic, A., Melillo, S., et al. (2015). Emergence of collective changes in travel direction of starling flocks from individual birds’ fluctuations. *Journal of The Royal Society Interface*, 12(108), 20150319.
- Attanasi, A., Cavagna, A., Del Castello, L., Giardina, I., Melillo, S., Parisi, L., et al. (2014b). Collective behaviour without collective order in wild swarms of midges. *PLOS Computational Biology*, 10(7), 1–10.
- Ay, N., & Polani, D. (2008). Information flows in causal networks. *Advances in Complex Systems*, 11(01), 17–41.
- Ballerini, M., Cabibbo, N., Candelier, R., Cavagna, A., Cisbani, E., Giardina, I., et al. (2008). Interaction ruling animal collective behavior depends on topological rather than metric distance: evidence from a field study. *Proceedings of the National Academy of Sciences*, 105(4), 1232–1237.
- Barnett, L., Barrett, A. B., & Seth, A. K. (2009). Granger causality and transfer entropy are equivalent for gaussian variables. *Physical Review Letters*, 103, 238701.
- Barnett, L., & Bossomaier, T. (2012). Transfer entropy as a Log-Likelihood ratio. *Physical Review Letters*, 109, 138105.
- Barnett, L., Lizier, J. T., Harré, M., Seth, A. K., & Bossomaier, T. (2013). Information flow in a kinetic ising model peaks in the disordered phase. *Physical Review Letters*, 111(17), 177203.
- Benjamini, Y., & Hochberg, Y. (1995). Controlling the false discovery rate: A practical and powerful approach to multiple testing. *Journal of the Royal Statistical Society Series B (Methodological)*, 57(1), 289–300.
- Bialek, W., Cavagna, A., Giardina, I., Mora, T., Silvestri, E., Viale, M., et al. (2012). Statistical mechanics for natural flocks of birds. *Proceedings of the National Academy of Sciences*, 109(13), 4786–4791.
- Boedecker, J., Obst, O., Lizier, J. T., Mayer, N. M., & Asada, M. (2012). Information processing in echo state networks at the edge of chaos. *Theory in Biosciences*, 131(3), 205–213.
- Bonabeau, E., Dorigo, M., & Theraulaz, G. (1999). *Swarm intelligence: From natural to artificial systems*. Oxford: Oxford University Press.

- Buhl, J., & Rogers, S. (2016). Mechanisms underpinning aggregation and collective movement by insect groups. *Current Opinion in Insect Science*, 15, 125–130.
- Buhl, J., Sumpter, D. J. T., Couzin, I. D., Hale, J. J., Despland, E., Miller, E. R., et al. (2006). From disorder to order in marching locusts. *Science*, 312(5778), 1402–1406.
- Buhl, J., Sword, G. A., Clissold, F. J., & Simpson, S. J. (2010). Group structure in locust migratory bands. *Behavioral Ecology and Sociobiology*, 65(2), 265–273.
- Butail, S., Ladu, F., Spinello, D., & Porfiri, M. (2014). Information flow in animal-robot interactions. *Entropy*, 16(3), 1315–1330.
- Butail, S., Mwaffo, V., & Porfiri, M. (2016). Model-free information-theoretic approach to infer leadership in pairs of zebrafish. *Physical Review E*, 93(4), 042411.
- Calovi, D. S., Litchinko, A., Lecheval, V., Lopez, U., Pérez Escudero, A., Chaté, H., et al. (2018). Disentangling and modeling interactions in fish with burst-and-coast swimming reveal distinct alignment and attraction behaviors. *PLOS Computational Biology*, 14(1), 1–28.
- Calovi, D. S., Lopez, U., Schuhmacher, P., Chaté, H., Sire, C., & Theraulaz, G. (2015). Collective response to perturbations in a data-driven fish school model. *Journal of The Royal Society Interface*, 12(104), 20141362.
- Cavagna, A., Giardina, I., & Ginelli, F. (2013a). Boundary information inflow enhances correlation in flocking. *Physical Review Letters*, 110(16), 168107.
- Cavagna, A., Queirós, S. M. D., Giardina, I., Stefanini, F., & Viale, M. (2013b). Diffusion of individual birds in starling flocks. *Proceedings of the Royal Society of London B: Biological Sciences*, 280(1756), 20122484.
- Chicharro, D., & Ledberg, A. (2012). When two become one: The limits of causality analysis of brain dynamics. *PLOS ONE*, 7(3), 1–16.
- Cliff, O. M., Lizier, J. T., Wang, X. R., Wang, P., Obst, O., & Prokopenko, M. (2017). Quantifying long-range interactions and coherent structure in multi-agent dynamics. *Artificial Life*, 23(1), 34–57.
- Couzin, I. D. (2009). Collective cognition in animal groups. *Trends in Cognitive Sciences*, 13(1), 36–43.
- Cover, T. M., & Thomas, J. A. (2006). *Elements of Information Theory*. New York: Wiley-Interscience.
- Crosato, E., Spinney, R. E., Nigmatullin, R., Lizier, J. T., & Prokopenko, M. (2018). Thermodynamics and computation during collective motion near criticality. *Physical Review E*, 97, 012120.
- Dimitriadis, S., Sun, Y., Laskaris, N., Thakor, N., & Bezerianos, A. (2016). Revealing cross-frequency causal interactions during a mental arithmetic task through symbolic transfer entropy: A novel vector-quantization approach. *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, 24(10), 1017–1028.
- Faes, L., Marinazzo, D., Montalto, A., & Nollo, G. (2014). Lag-specific transfer entropy as a tool to assess cardiovascular and cardiorespiratory information transfer. *IEEE Transactions on Biomedical Engineering*, 61(10), 2556–2568.
- Faes, L., Nollo, G., & Porta, A. (2011). Information-based detection of nonlinear granger causality in multivariate processes via a nonuniform embedding technique. *Physical Review E*, 83(5), 051112.
- Faes, L., & Porta, A. (2014). Conditional Entropy-Based evaluation of information dynamics in physiological systems. In M. Wibral, R. Vicente, & J. T. Lizier (Eds.), *Directed information measures in neuroscience, understanding complex systems* (pp. 61–86). Berlin: Springer.
- Fano, R. M. (1961). *Transmission of information: A statistical theory of communications*. Cambridge, MA: M.I.T Press.
- Feldman, D. P., McTague, C. S., & Crutchfield, J. P. (2008). The organization of intrinsic computation: Complexity-entropy diagrams and the diversity of natural information processing. *Chaos*, 18(4), 043106.
- Fourcassié, V., Dussutour, A., & Deneubourg, J. L. (2010). Ant traffic rules. *The Journal of Experimental Biology*, 213(14), 2357–2363.
- Gautrais, J., Ginelli, F., Fournier, R., Blanco, S., Soria, M., Chaté, H., et al. (2012). Deciphering interactions in moving animal groups. *PLOS Computational Biology*, 8(9), 1–11.
- Geweke, J. (1982). Measurement of linear dependence and feedback between multiple time series. *Journal of the American Statistical Association*, 77(378), 304–313.
- Giardina, I. (2008). Collective behavior in animal groups: Theoretical models and empirical studies. *Human Frontier Science Program Journal*, 2(4), 205–219.
- Ginelli, F., Peruani, F., Pillot, M. H., Chaté, H., Theraulaz, G., & Bon, R. (2015). Intermittent collective dynamics emerge from conflicting imperatives in sheep herds. *Proceedings of the National Academy of Sciences*, 112(41), 12729–12734.
- Gómez, C., Lizier, J. T., Schaum, M., Wollstadt, P., Grützner, C., Uhlhaas, P., et al. (2014). Reduced predictable information in brain signals in autism spectrum disorder. *Frontiers in Neuroinformatics*, 8, 9.
- Herbert-Read, J. E., Buhl, J., Hu, F., Ward, A. J., & Sumpter, D. J. (2015). Initiation and spread of escape waves within animal groups. *Royal Society open science*, 2(4), 140355.

- Herbert-Read, J. E., Perna, A., Mann, R. P., Schaerf, T. M., Sumpter, D. J. T., & Ward, A. J. W. (2011). Inferring the rules of interaction of shoaling fish. *Proceedings of the National Academy of Sciences*, 108(46), 18726–18731.
- James, R. G., Barnett, N., & Crutchfield, J. P. (2016). Information flows? A critique of transfer entropies. *Physical Review Letters*, 116(23), 238701.
- Jeanson, R., Deneubourg, J. L., & Theraulaz, G. (2004). Discrete dragline attachment induces aggregation in spiderlings of a solitary species. *Animal Behaviour*, 67(3), 531–537.
- Jeanson, R., Rivault, C., Deneubourg, J. L., Blanco, S., Fournier, R., Jost, C., et al. (2005). Self-organized aggregation in cockroaches. *Animal Behaviour*, 69(1), 169–180.
- Jiang, L., Giuggioli, L., Perna, A., Escobedo, R., Lecheval, V., Sire, C., et al. (2017). Identifying influential neighbors in animal flocking. *PLOS Computational Biology*, 13(11), 1–32.
- Katz, Y., Tunström, K., Ioannou, C. C., Huepe, C., & Couzin, I. D. (2011). Inferring the structure and dynamics of interactions in schooling fish. *Proceedings of the National Academy of Sciences*, 108(46), 18720–18725.
- Khadem, A., Hossein-Zadeh, G. A., & Khorrami, A. (2016). Long-range reduced predictive information transfers of autistic youths in EEG sensor-space during face processing. *Brain topography*, 29(2), 283–295.
- Khuong, A., Gautrais, J., Perna, A., Sbaï, C., Combe, M., Kuntz, P., et al. (2016). Stigmergic construction and topochemical information shape ant nest architecture. *Proceedings of the National Academy of Sciences*, 113(5), 1303–1308.
- Ladu, F., Mwaffo, V., Li, J., Macrì, S., & Porfiri, M. (2015). Acute caffeine administration affects zebrafish response to a robotic stimulus. *Behavioural Brain Research*, 289, 48–54.
- Langton, C. G. (1990). Computation at the edge of chaos: Phase transitions and emergent computation. *Physica D*, 42(1–3), 12–37.
- Larson, L. (1983). The symmetric derivative. *Transactions of the American Mathematical Society*, 277(2), 589–599.
- Lecheval, V., Jiang, L., Tichit, P., Sire, C., Hemelrijk, C. K., & Theraulaz, G. (2017). Domino-like propagation of collective u-turns in fish schools. submitted to bioRxiv.
- Lissaman, P. B. S., & Shollenberger, C. A. (1970). Formation flight of birds. *Science*, 168(3934), 1003–1005.
- Lizier, J. T. (2013). *The local information dynamics of distributed computation in complex systems*, Springer Theses. Berlin: Springer.
- Lizier, J. T. (2014a). JIDT: An information-theoretic toolkit for studying the dynamics of complex systems. *Frontiers in Robotics and AI*, 1, 11.
- Lizier, J. T. (2014b). Measuring the dynamics of information processing on a local scale in time and space. In M. Wibral, R. Vicente, & J. T. Lizier (Eds.), *Directed information measures in neuroscience, understanding complex systems* (pp. 161–193). Berlin: Springer.
- Lizier, J. T., Heinzle, J., Horstmann, A., Haynes, J. D., & Prokopenko, M. (2011a). Multivariate information-theoretic measures reveal directed information structure and task relevant changes in fMRI connectivity. *Journal of Computational Neuroscience*, 30(1), 85–107.
- Lizier, J. T., Pritam, S., & Prokopenko, M. (2011b). Information dynamics in small-world boolean networks. *Artificial Life*, 17(4), 293–314.
- Lizier, J. T., & Prokopenko, M. (2010). Differentiating information transfer and causal effect. *The European Physical Journal B*, 73(4), 605–615.
- Lizier, J. T., Prokopenko, M., & Zomaya, A. Y. (2008). Local information transfer as a spatiotemporal filter for complex systems. *Physical Review E*, 77(2), 026110.
- Lizier, J. T., Prokopenko, M., & Zomaya, A. Y. (2010). Information modification and particle collisions in distributed computation. *Chaos*, 20(3), 037109.
- Lizier, J. T., Prokopenko, M., & Zomaya, A. Y. (2012). Local measures of information storage in complex distributed computation. *Information Sciences*, 208, 39–54.
- Lizier, J. T., Prokopenko, M., & Zomaya, A. Y. (2014). A framework for the local information dynamics of distributed computation in complex systems. In M. Prokopenko (Ed.), *Guided self-organization: Inception, emergence, complexity and computation* (Vol. 9, pp. 115–158). Berlin: Springer.
- Lizier, J. T., & Rubinov, M. (2012). Multivariate construction of effective computational networks from observational data. Technical Report Preprint 25/2012, Max Planck Institute for Mathematics in the Sciences.
- Lord, W. M., Sun, J., Ouellette, N. T., & Boltt, E. M. (2016). Inference of causal information flow in collective animal behavior. *IEEE Transactions on Molecular, Biological and Multi-Scale Communications*, 2(1), 107–116.

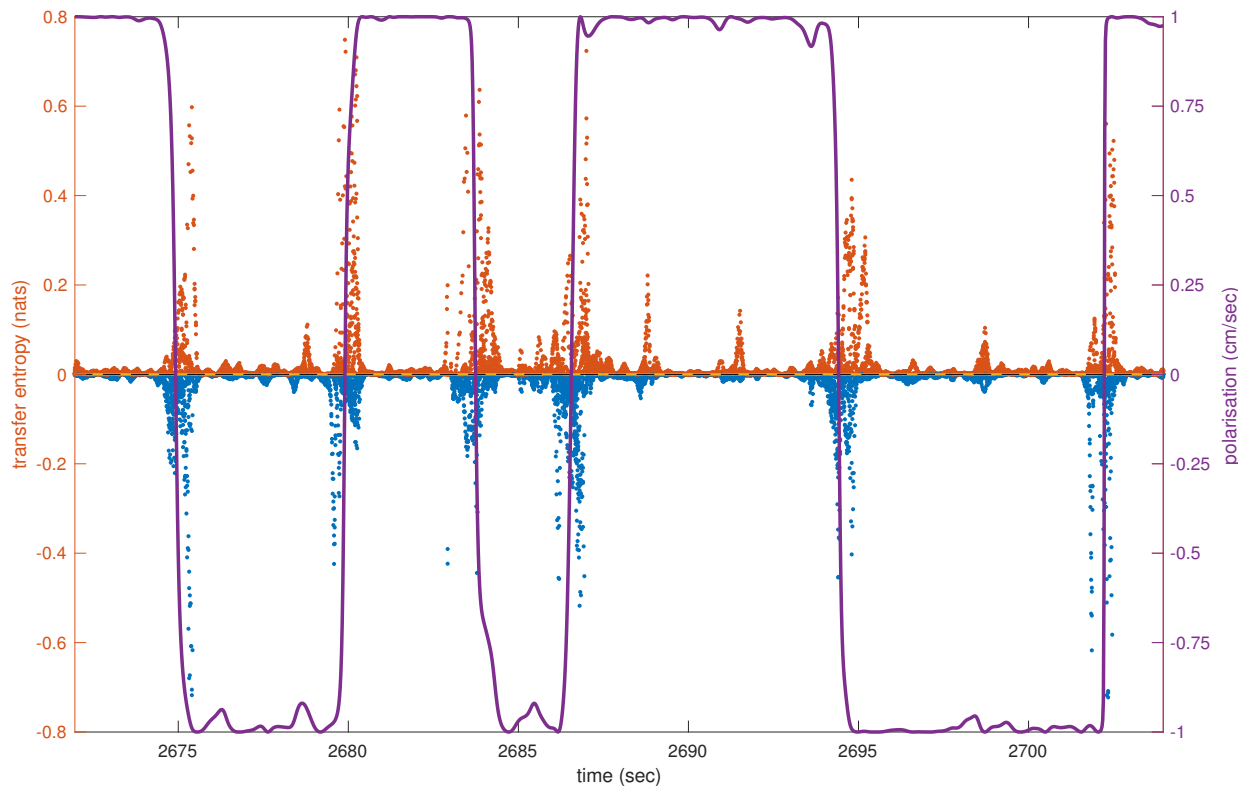
- Mann, R. P., Perna, A., Strömbom, D., Garnett, R., Herbert-Read, J. E., Sumpter, D. J. T., et al. (2012). Multi-scale inference of interaction rules in animal groups using bayesian model selection. *PLOS Computational Biology*, 8(1), 1–12.
- Marinazzo, D., Pellicoro, M., & Stramaglia, S. (2012). Causal information approach to partial conditioning in multivariate data sets. *Computational and Mathematical Methods in Medicine*, 2012, 303601. <https://doi.org/10.1155/2012/303601>.
- Materassi, M., Consolini, G., Smith, N., & De Marco, R. (2014). Information theory analysis of cascading process in a synthetic model of fluid turbulence. *Entropy*, 16(3), 1272–1286.
- May, R. M. (1979). Flight formations in geese and other birds. *Nature*, 282, 778–780.
- Miller, J. M., Wang, X. R., Lizier, J. T., Prokopenko, M., & Rossi, L. F. (2014). Measuring information dynamics in swarms. In M. Prokopenko (Ed.), *Guided self-organization: Inception, emergence, complexity and computation* (Vol. 9, pp. 343–364). Berlin: Springer.
- Moussaïd, M., Helbing, D., Garnier, S., Johansson, A., Combe, M., & Theraulaz, G. (2009). Experimental study of the behavioural mechanisms underlying self-organization in human crowds. *Proceedings of the Royal Society of London B: Biological Sciences*, 276(1668), 2755–2762.
- Moussaïd, M., Helbing, D., & Theraulaz, G. (2011). How simple rules determine pedestrian behavior and crowd disasters. *Proceedings of the National Academy of Sciences*, 108(17), 6884–6888.
- Nagy, M., Ákos, Z., Biro, D., & Vicsek, T. (2010). Hierarchical group dynamics in pigeon flocks. *Nature*, 464(7290), 890–893.
- Nagy, M., Vásárhelyi, G., Pettit, B., Roberts-Mariani, I., Vicsek, T., & Biro, D. (2013). Context-dependent hierarchies in pigeons. *Proceedings of the National Academy of Sciences*, 110(32), 13049–13054.
- Orange, N., & Abaid, N. (2015). A transfer entropy analysis of leader-follower interactions in flying bats. *The European Physical Journal Special Topics*, 224(17), 3279–3293.
- Parrish, J. K., Viscido, S. V., & Grünbaum, D. (2002). Self-organized fish schools: an examination of emergent properties. *Biological Bulletin*, 202(3), 296–305.
- Partridge, B. (1980). The effect of school size on the structure and dynamics of minnow schools. *Animal Behaviour*, 28(1), 68–IN3.
- Pérez-Escudero, A., Vicente-Page, J., Hinz, R. C., Arganda, S., & de Polavieja, G. G. (2014). idTracker: Tracking individuals in a group by automatic identification of unmarked animals. *Nature Methods*, 11(7), 743–748.
- Potts, W. K. (1984). The chorus-line hypothesis of manoeuvre coordination in avian flocks. *Nature*, 309(5966), 344–345.
- Procaccini, A., Orlandi, A., Cavagna, A., Giardina, I., Zoratto, F., Santucci, D., et al. (2011). Propagating waves in starling, *sturnus vulgaris*, flocks under predation. *Animal Behaviour*, 82(4), 759–765.
- Prokopenko, M., Lizier, J. T., Obst, O., & Wang, X. R. (2011). Relating Fisher information to order parameters. *Physical Review E*, 84(4), 041116.
- Ragwitz, M., & Kantz, H. (2002). Markov models from data by simple nonlinear time series predictors in delay embedding spaces. *Physical Review E*, 65, 056201.
- Razak, F. A., & Jensen, H. J. (2014). Quantifying ‘causality’ in complex systems: Understanding transfer entropy. *PLOS ONE*, 9(6), e99462.
- Reynolds, C. W. (1987). Flocks, herds and schools: A distributed behavioral model. In *SIGGRAPH '87 Proceedings of the 14th annual conference on computer graphics and interactive techniques*. ACM, New York, NY, USA (Vol. 21, pp. 25–34).
- Richardson, T. O., Perony, N., Tessone, C. J., Bousquet, C. A., Manser, M. B., & Schweitzer, F. (2013). Dynamical coupling during collective animal motion. [arXiv:1311.1417](https://arxiv.org/abs/1311.1417).
- Riley, D. A., & Leith, C. R. (1976). Multidimensional psychophysics and selective attention in animals. *Psychological Bulletin*, 83(1), 138.
- Rosenthal, S. B., Twomey, C. R., Hartnett, A. T., Wu, H. S., & Couzin, I. D. (2015). Revealing the hidden networks of interaction in mobile animal groups allows prediction of complex behavioral contagion. *Proceedings of the National Academy of Sciences*, 112(15), 4690–4695.
- Schreiber, T. (2000). Measuring information transfer. *Physical Review Letters*, 85(2), 461–464.
- Smirnov, D. A. (2013). Spurious causalities with transfer entropy. *Physical Review E*, 87, 042917.
- Stramaglia, S., Wu, G. R., Pellicoro, M., & Marinazzo, D. (2012). Expanding the transfer entropy to identify information circuits in complex systems. *Physical Review E*, 86, 066211.
- Sumpter, D., Buhl, J., Biro, D., & Couzin, I. (2008). Information transfer in moving animal groups. *Theory in Biosciences*, 127(2), 177–186.
- Sun, Y., Rossi, L. F., Shen, C. C., Miller, J., Wang, X. R., Lizier, J. T., et al. (2014). Information transfer in swarms with leaders. [arXiv:1407.0007](https://arxiv.org/abs/1407.0007).
- Theraulaz, G., Bonabeau, E., Nicolis, S. C., Solé, R. V., Fourcassié, V., Blanco, S., et al. (2002a). Spatial patterns in ant colonies. *Proceedings of the National Academy of Sciences*, 99(15), 9645–9649.

- Theraulaz, G., Bonabeau, E., Sole, R. V., Schatz, B., & Deneubourg, J. L. (2002b). Task partitioning in a ponerine ant. *Journal of Theoretical Biology*, 215(4), 481–489.
- Tomaru, T., Murakami, H., Niizato, T., Nishiyama, Y., Sonoda, K., Moriyama, T., et al. (2016). Information transfer in a swarm of soldier crabs. *Artificial Life and Robotics*, 21 (2), 177–180.
- Tunstrøm, K., Katz, Y., Ioannou, C. C., Huepe, C., Lutz, M. J., & Couzin, I. D. (2013). Collective states, multistability and transitional behavior in schooling fish. *PLOS Computational Biology*, 9(2), 1–11.
- Vakorin, V. A., Krakovska, O. A., & McIntosh, A. R. (2009). Confounding effects of indirect connections on causality estimation. *Journal of Neuroscience Methods*, 184(1), 152–160.
- Vicente, R., Wibral, M., Lindner, M., & Pipa, G. (2011). Transfer entropy—A model-free measure of effective connectivity for the neurosciences. *Journal of Computational Neuroscience*, 30(1), 45–67.
- Wang, X. R., Miller, J. M., Lizier, J. T., Prokopenko, M., & Rossi, L. F. (2012). Quantifying and tracing information cascades in swarms. *PLOS ONE*, 7(7), 1–7.
- Weitz, S., Blanco, S., Fournier, R., Gautrais, J., Jost, C., & Theraulaz, G. (2012). Modeling collective animal behavior with a cognitive perspective: A methodological framework. *PLOS ONE*, 7(6), 1–16.
- Wibral, M., Lizier, J. T., & Priesemann, V. (2015). Bits from brains for biologically-inspired computing. *Frontiers in Robotics and AI*, 2, 5.
- Wibral, M., Pampu, N., Priesemann, V., Siebenhühner, F., Seiwert, H., Lindner, M., et al. (2013). Measuring information-transfer delays. *PLOS ONE*, 8(2), 1–19.
- Wibral, M., Vicente, R., & Lindner, M. (2014). Transfer entropy in neuroscience. In M. Wibral, R. Vicente, & J. T. Lizier (Eds.), *Directed information measures in neuroscience, understanding complex systems* (pp. 3–36). Berlin: Springer.
- Williams, P. L., & Beer, R. D. (2011). Generalized measures of information transfer. [arXiv:1102.1507](https://arxiv.org/abs/1102.1507).

Informative and misinformative interactions in a school of fish

Emanuele Crosato, Li Jiang, Valentin Lecheval, Joseph T. Lizier, X. Rosalind Wang, Pierre Tichit, Guy Theraulaz, Mikhail Prokopenko

Supplementary information

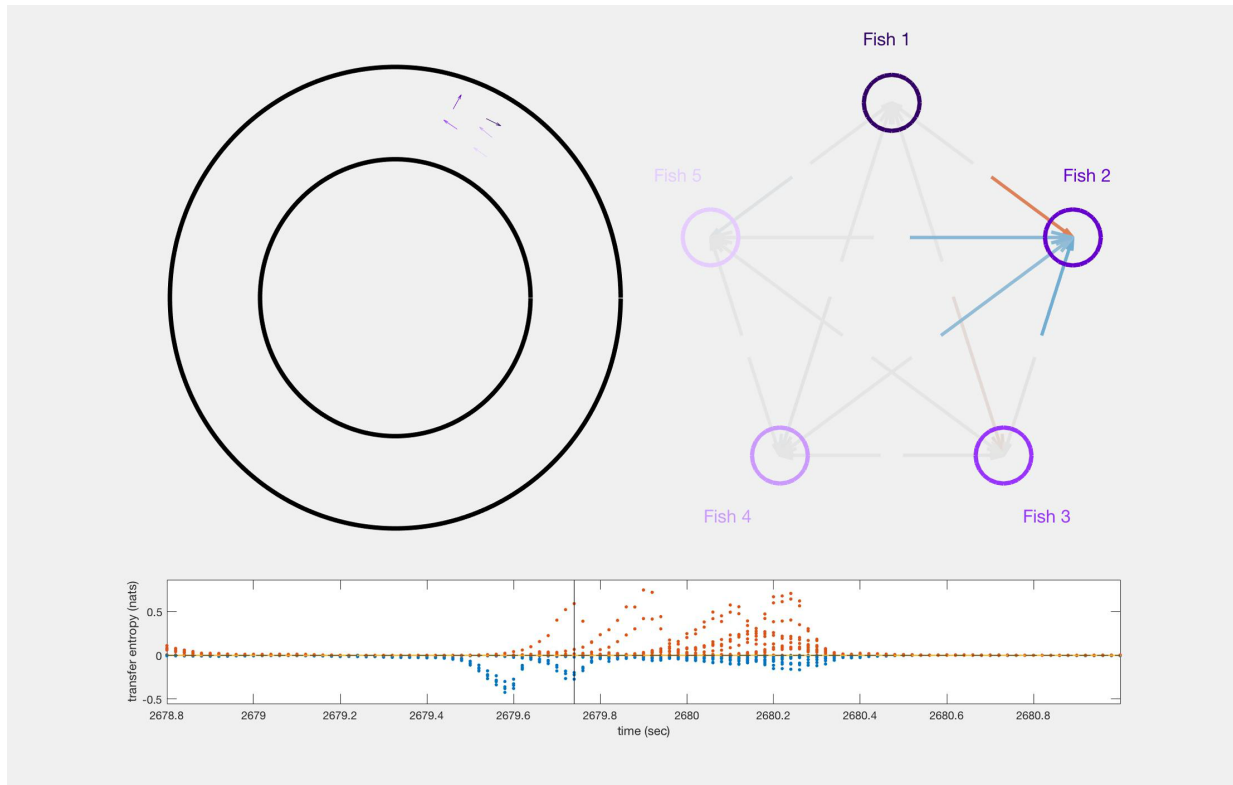


Supplementary Figure S1. Transfer entropy within the school during several U-turns. The figure plots the school's polarisation during a U-turn and the detected transfer entropy over a time interval of approximately 35 seconds. The purple line represents the school's polarisation, while dots represent local values of transfer entropy between all directed pairs of fish: red dots represent positive transfer entropy and blue dots represent negative transfer entropy. Time is discretised in steps of length 0.02 seconds and for each time step 20 points of these local measures are plotted, for the 20 directed pairs formed out of 5 fish.



Supplementary Video S1. Fish undergoing the representative U-turn. The movie shows five *Hemigrammus rhodostomus* swimming in the ring-shaped tank for approximately 6 seconds, during which they undergo the U-turn presented in the main article.

The movie is available at: <https://www.dropbox.com/s/bhvsusc3pz64k38/U-turn-real.mp4?dl=0>.



Supplementary Video S2. Animation of the representative U-turn showing transfer entropy dynamics. The movie shows an animation of the representative U-turn over a time interval of approximately 2 seconds. On the top-left is the ring-shaped tank with the five fish, represented by arrows of different shades of purple. On the bottom is the transfer entropy between any directed pair of fish over the time interval: red dots represent positive transfer entropy and blue dots represent negative transfer entropy. Time is discretised in steps of length 0.02 seconds and for each time step 20 points of transfer entropy are plotted, for the 20 directed pairs that can be formed out of 5 fish. On the top-right is the network of transient neighbours changing over time. Each node represents a fish and each directed edge entering a node indicates the transfer entropy to that fish from the other four (the source fish is easily identifiable from the angle of the edges). The colour of the edges changes during the U-turn: strong red indicates intense positive transfer entropy; strong blue indicates intense negative transfer entropy; intermediate grey indicates that transfer entropy is close to zero.

The movie is available at: <https://www.dropbox.com/s/qwj2pzzlsqw173m/U-turn-TEL.avi?dl=0>.

Chapter 4

Thermodynamics and computation during collective motion

Thermodynamics and computation during collective motion near criticality

Emanuele Crosato,^{*} Richard E. Spinney, Ramil Nigmatullin, Joseph T. Lizier, and Mikhail Prokopenko
*Complex Systems Research Group and Centre for Complex Systems, Faculty of Engineering and IT,
 The University of Sydney, Sydney, NSW 2006, Australia*



(Received 15 August 2017; published 16 January 2018)

We study self-organization of collective motion as a thermodynamic phenomenon in the context of the first law of thermodynamics. It is expected that the coherent ordered motion typically self-organises in the presence of changes in the (generalized) internal energy and of (generalized) work done on, or extracted from, the system. We aim to explicitly quantify changes in these two quantities in a system of simulated self-propelled particles and contrast them with changes in the system's configuration entropy. In doing so, we adapt a thermodynamic formulation of the curvatures of the internal energy and the work, with respect to two parameters that control the particles' alignment. This allows us to systematically investigate the behavior of the system by varying the two control parameters to drive the system across a kinetic phase transition. Our results identify critical regimes and show that during the phase transition, where the configuration entropy of the system decreases, the rates of change of the work and of the internal energy also decrease, while their curvatures diverge. Importantly, the reduction of entropy achieved through expenditure of work is shown to peak at criticality. We relate this both to a thermodynamic efficiency and the significance of the increased order with respect to a computational path. Additionally, this study provides an information-geometric interpretation of the curvature of the internal energy as the difference between two curvatures: the curvature of the free entropy, captured by the Fisher information, and the curvature of the configuration entropy.

DOI: [10.1103/PhysRevE.97.012120](https://doi.org/10.1103/PhysRevE.97.012120)

I. INTRODUCTION

Collective motion involves self-organization of coherent movement in a system of self-propelled particles and is a pervasive phenomenon observed in many biological, chemical, and physical systems [1]. Collective motion has been studied in animals (e.g., flocks of birds [2], schools of fish [3], and colonies of insects [4]), in bacteria [5], in tissue cells [6], in moving biomolecules [7], and even in nonliving systems such as autonomous micromotors [8]. Despite their diversity, these systems can exhibit similar motion patterns, such as orientated aggregations, stationary clusters, and swirls [1]. A crucial characteristic that distinguishes collective motion from other kinds of coordinated motion is that complex patterns can self-organize from simple local interactions among individual particles, without requiring any global control or leading roles [9] but involving information cascades [10,11]. Nevertheless, systems of self-propelled particles can display remarkable dynamic coordination during collective motion, as well as other interesting features, such as scalability, response to the environment, and reconfiguration after external intrusions.

The ubiquity of collective motion, and its similarity across different systems, suggest the existence of underlying universal principles, the investigation of which has become a well-established, cross-disciplinary pursuit. The formulation of general laws bridging local interactions and group-level properties is one of the main challenges for defining a unified theory of collective motion [12].

A first step towards this goal was the conception of dynamical models [13–16]. Vicsek *et al.* [13] introduced a dynamical model of collective motion inspired by ferromagnetism, in which particles assume the average direction of motion of other particles in its neighborhood (similarly to magnetization), with some random perturbation (similarly to temperature). The authors simulated the motion for gradually decreasing random perturbation and observed a kinetic phase transition between a disorderly moving phase and a phase with coherent (oriented) motion, the critical point of which was localized using a suitable order parameter. Several studies have followed Vicsek's intuition, and extensions of the model have been proposed. Grégoire and Chaté [16], for example, studied the effect of several control parameters on the collective behavior of a modified version of Vicsek's model, which adds a cohesion component to the motion rules. The authors confirmed the existence of the kinetic phase transition and, by varying the strength of the additional cohesion component, observed three more phases: a “gas,” a “liquid,” and a “solid” phase, also separated by phase transitions.

More recently, Bialek *et al.* [17–19] provided a statistical mechanical model for the propagation of directional order throughout flocks. On the hypothesis that flocks have statistically stationary states, the authors calculated the maximum entropy distribution [20] of birds' normalized velocities, consistent with the average pairwise directional correlation experimentally observed from the field data (i.e., large flocks of *Sturnus vulgarishas* [21–23]). Bialek's statistical mechanical description provides a formal theoretical framework to make quantitative predictions of emergent collective phenomena. For instance, the model was shown to be capable of predicting the

^{*}emanuele.crosato@sydney.edu.au; also at CSIRO Data61, P. O. Box 76, Epping, NSW 1710, Australia.

existence of pairwise correlations on all length scales, as well as four-body correlations [17]. The model was also shown to be capable of predicting the flight directions of birds in the interior of the flock, given the directions of the birds on the border.

Despite this fundamental contribution, current statistical mechanical approaches to collective motion do not explicitly incorporate thermodynamic quantities such as free entropy and work, dynamics of which are especially important during phase transitions. In this article we aim to investigate these quantities in the dynamical model of collective motion proposed by Grégoire and Chaté [16], which undergoes a kinetic phase transition over parameters that control the particles' alignment: from a "disordered motion" phase, in which particles keep changing direction but occupy a fairly stable collective space, to a "coherent motion" phase, in which particles cohesively move towards a common direction. The control parameters that we consider are the alignment strength among particles and the number of nearest neighbors affecting a particle's alignment. A quasistatic process is considered, during which these two control parameters are varied infinitesimally slowly, driving the system across the phase transition while thermodynamic equilibrium is maintained. The dynamics of fundamental thermodynamical quantities, such as the *generalized work*, heat and energy, are investigated over the quasistatic process, in the context of the first law of thermodynamics.

The choice of a quasistatic protocol allows the application of our theoretical framework, which requires the system to be in a steady state. Moreover, the results obtained considering a quasistatic protocol can be meaningfully interpreted in the context of the second law of thermodynamics to get useful insights into more realistic processes. For instance, the work done on the system in the quasistatic limit is a lower bound for the work that would be done on the system using a protocol in which the control parameter is varied faster.

In this study, we use a method that allows us to give a statistical mechanical interpretation of the curvatures of the generalized work and of the generalized internal energy of the system with respect to the control parameter. Such method exploits the relationship between these two curvatures and two information-theoretic quantities, the configuration entropy and the Fisher information (a measure of the information that an observed variable carries about the parameter), which can be numerically estimated by simulating the system using different values of the control parameters.

We also provide two information-geometric expressions of the curvature of the internal energy and related quantities with respect to the control parameter. On the one hand, the curvature of the internal energy is proportional to the *difference* between two curvatures: the curvature of the free entropy, captured by the Fisher information, and the curvature of the configuration entropy. This expression highlights a "computational balance" present in distributed computational processes, of which collective motion is an example. Such balance relates the *sensitivity* of the system to changes in control parameter (captured by the Fisher information) and the system's *uncertainty* (captured by the configuration entropy). This enhances the view of the "thermodynamic balance," reflected by the first law in the context of quasistatic processes, between the configuration entropy of the system, its internal

energy and the work done on, or extracted from, the system. On the other hand, we derive another quantity as the *sum* of the Fisher information and the curvature of the configuration entropy.

Our computational results show that, in the simulated system of particles during collective motion, the rates of change of the generalized work and of the generalized internal energy decrease with the control parameters, whenever the system of self-propelled particles begins to move more coherently. This dynamic is particularly steep near criticality, where the curvatures of these quantities with respect to the control parameters are shown to diverge. The configuration entropy of the system is shown to decrease during the phase transition as the system self-organizes into a more ordered phase. The thermodynamic perspective adopted in this study allows us to define a notion of thermodynamic efficiency of computation as a ratio of entropy changes to the required work. In addition, we propose an interpretation of this work rate as a distance along a computational path implied by the control parameter, measured in terms of the cumulative sensitivity to the changes in the control parameter. Specifically, our results suggest that the reduction of the configuration entropy, indicating the increase in the internal order within the considered collective motion, is most significant at criticality.

In addition to these main results, this paper confirms and quantifies critical dynamics in statistical mechanical models of collective motion, which were previously observed in dynamical models [13,16]. Moreover, it is shown that the Fisher information diverges at criticality and can therefore be used to build a phase diagram of the dynamics of the system.

The remainder of this article is structured as follows. Section II provides the technical preliminaries necessary for understanding the role of the Fisher information in physical systems, the information-geometrical interpretation of the studied curvatures, the quasistatic protocol that we consider and the dynamical model of collective motion. Section III presents our statistical mechanical formulation of the curvatures of the generalized work and internal energy, and the computational results of simulated collective motion. The results are discussed in Sec. IV, where our conclusions are also provided.

II. TECHNICAL PRELIMINARIES

A. Fisher information and physical systems

The Fisher information [24] is a known quantity in statistics and information theory. It measures the amount of information that an observable random variable X carries about an unknown parameter θ . For many parameters $\theta = [\theta_1, \theta_2, \dots, \theta_M]^T$, the Fisher information matrix is defined as

$$F_{mn}(\theta) = E \left\{ \left[\frac{\partial \ln p(x|\theta)}{\partial \theta_m} \right] \left[\frac{\partial \ln p(x|\theta)}{\partial \theta_n} \right] \middle| \theta \right\}, \quad (1)$$

where $p(x|\theta)$ is the probability of the realization x of X given the parameters θ and the function $E(y)$ is the expected value of y .

In recent years, the meaning of the Fisher information for physical systems has been investigated in thermodynamical and statistical mechanical terms [25–32]. Let us consider a physical system, described by the state functions $X_m(x)$ over

the configuration space. The probability of the states of the system, in a stationary state, is given by the Gibbs measure:

$$p(x|\theta) = \frac{1}{Z(\theta)} e^{-\beta H(x,\theta)} = \frac{1}{Z(\theta)} e^{-\sum_m \theta_m X_m(x)}, \quad (2)$$

where $\beta = 1/k_b T$ is the inverse temperature T (k_b being the Boltzmann constant), the Hamiltonian $H(x,\theta)$ defines the total energy at state x , θ_m are thermodynamic variables (pressure, magnetic field, chemical potential, etc.), and $Z(\theta)$ is the partition function [25,28]. The Gibbs free energy of such system is

$$G(T, \theta_m) = U(S, \phi_m) - TS - \phi_m \theta_m, \quad (3)$$

where U is the internal energy of the system, S is the configuration entropy, and ϕ_m is an order parameter. For a physical system described by the Gibbs measure in Eq. (2), the Fisher information measures the size of the fluctuations about equilibrium in the collective variables X_m and X_n and is related to the derivatives of the corresponding order parameters with respect to the collective variables [29,30]:

$$F_{mn}(\theta) = \langle (X_m(x) - \langle X_m \rangle)(X_n(x) - \langle X_n \rangle) \rangle = \beta \frac{\partial \phi_m}{\partial \theta_n}, \quad (4)$$

where the angle brackets represent average values over the ensemble.

Fisher information has been related to entropy production [33] and also used as a variational principle to derive fundamental thermodynamic laws [34,35] or for predicting modeling [31].

B. Interpretation of entropic curvatures

It has also been shown [25–28] that the Fisher information is equivalent to the thermodynamic metric tensor:

$$F_{mn}(\theta) = g_{mn}(\theta) = \frac{\partial^2 \psi}{\partial \theta_m \partial \theta_n}, \quad (5)$$

where $\psi = \ln Z = -\beta G$ is the free entropy (for isothermal systems, ψ is proportional to the free energy). In other words, the Fisher information is the curvature of the free entropy ($\ln Z$). This reveals the information-geometrical meaning of the Fisher information as a Riemannian metric (more precisely, the Fisher-Rao metric) for the manifold of thermodynamic states, providing a measure of distance between thermodynamic states. Thus, information-geometrically, the Fisher information can be interpreted as an average uncertainty density on a statistical manifold, proportional to the volume of geodesic balls [36].

This study provides thermodynamical interpretations for curvatures, focusing specifically on quantities that can be computed numerically from the probability distribution of the observed variables, such as the configuration entropy S of the system. In particular, we propose the curvatures

$$\frac{d^2 \langle S \rangle^\pm}{d\theta^2} \equiv \frac{d^2 S}{d\theta^2} \pm F(\theta) \quad (6)$$

as the central quantities of interest (notice that a single control parameter θ is now used). Therefore, the quantity $d^2 \langle S \rangle^\pm / d\theta^2$ is either the sum of, or the difference between, average

statistical uncertainties (i.e., the volumes of geodesic balls) attributed to the free entropy and to the configuration entropy.

In order to interpret these information-geometric, static, quantities in terms of traditional thermodynamic quantities (e.g., heat and work, defined dynamically) we must give meaning to the notion of a change with respect to the control parameter, θ , i.e., we must define the process or *protocol*. By protocol we mean a defined evolution of the control parameter in time, i.e., $\theta(t)$, which drives the system between different states and in doing so incurs changes in heat, work, etc. By establishing such a protocol we can give physical meaning to integrals of the curvatures $\langle S \rangle^\pm$, such that $d\langle S \rangle^\pm / d\theta$ can be readily interpreted as a change in $\langle S \rangle^\pm$ under the action of the protocol. It is of critical importance to recognize that the nature of the protocol determines the physical behavior of the quantity $\langle S \rangle^\pm$, i.e., its decomposition into heat and work. The most natural example is a quasistatic protocol, which we discuss next, though note that less conventional alternatives can be designed (as will be discussed in Sec. IV).

C. Quasistatic protocols

A quasistatic protocol is an idealized driving process carried out over an infinite amount of time, such that we can consider the system to be in equilibrium throughout the process. For instance, a linear quasistatic protocol taking the system from a distribution characterized by θ^1 to θ^2 , would be given by the limit

$$\theta(t) = \lim_{\tau \rightarrow \infty} \theta^1 + \frac{t}{\tau} (\theta^2 - \theta^1). \quad (7)$$

Since the system is always in equilibrium, the total entropy production of the universe (the system and the environment) is zero, and therefore any change in the configuration entropy due to the driving process is identically matched by a flow of heat that manifests as entropy change in the environment:

$$\frac{dS}{d\theta} = \frac{d\langle \beta Q_{\text{gen}} \rangle}{d\theta}, \quad (8)$$

where a sign convention dictates that Q_{gen} is the generalized heat flow *from* the environment *to* the system. Here the subscript indicates a generalized heat in the sense of Jaynes [20], such that we can consider

$$\langle U_{\text{gen}} \rangle = U(S, \phi) - \phi \theta \quad (9)$$

and the generalized first law holds $\Delta \langle U_{\text{gen}} \rangle = \Delta \langle Q_{\text{gen}} \rangle + \Delta \langle W_{\text{gen}} \rangle$, where W_{gen} is the generalized work. Equation (8) leads to a formulation of the first law of thermodynamics, in case of a quasistatic processes, as

$$\frac{d\langle \beta U_{\text{gen}} \rangle}{d\theta} = \frac{dS}{d\theta} + \frac{d\langle \beta W_{\text{gen}} \rangle}{d\theta}. \quad (10)$$

It is worth noting that, according to the second law of thermodynamics, a change in the free energy of the system requires a greater or equal amount of work to be done on the system, which is $\Delta \langle W_{\text{gen}} \rangle \geq \Delta G$. In the quasistatic limit the work required is exactly the change in the free energy, therefore $\Delta \langle W_{\text{gen}} \rangle = \Delta G$. In other words, the total work performed on the system (which can be calculated by integrating the infinitesimal work changes over a range of the control parameter) is

a lower bound for the work that would be performed on the system if we were not considering the quasistatic limit.

This methodology is very general, provided a Gibbs form can be postulated and a probability distribution can be estimated. However, in this study we focus on a system of self-propelled particles during collective motion, driven across a phase transition by a quasistatic protocol acting on two parameters that control the particles' alignment. The model of collective motion that we adopted is presented in the following section.

D. Dynamical and statistical mechanical models of collective motion

We consider the model of collective motion proposed by Grégoire and Chaté [16]. Let us have N self-propelled particles. At time t , each particle $i = \{1, 2, \dots, N\}$ has position $\mathbf{x}_i(t)$ and velocity $\mathbf{v}_i(t)$. The time evolution of position and velocity is given by the following rules:

$$\mathbf{x}_i(t+1) = \mathbf{x}_i(t) + \mathbf{v}_i(t), \quad (11)$$

$$\mathbf{v}_i(t+1) = v_0 \Theta \left[a \sum_{j \in n_c^i} \mathbf{v}_j(t) + b \sum_{j \in n_c^i} f_{ij} + n_c \boldsymbol{\eta}_i \right]. \quad (12)$$

The normalization operator $\Theta(\mathbf{y}) = \mathbf{y}/|\mathbf{y}|$ keeps the particles' speed constant, i.e., $|\mathbf{v}_i(t)| = v_0$ at every time t . The argument of the normalization operator is the sum of three velocity components: From left to right, we have an *alignment*, a *cohesion*, and a *perturbation* components. The alignment component for particle i is the sum of the velocities of its nearest neighborhood $j \in n_c^i$ of fixed size n_c (i.e., n_c^i includes the n_c particles with the smallest Euclidean distance from i and is updated at each time step). The cohesion component is the sum of the cohesion forces f_{ij} between particle i and its neighbors. The parameters a and b are, respectively, the weights of the alignment and the cohesion components. The perturbation is introduced by means of a random unit vector $\boldsymbol{\eta}_i$ and is weighted by the fixed number of nearest neighbors n_c of each particle.

The forces f_{ij} are functions of the distances r_{ij} :

$$\begin{aligned} f_{ij}(r_{ij} < r_b) &= -\infty \cdot \mathbf{e}_{ij}, \\ f_{ij}(r_b \leq r_{ij} < r_a) &= \frac{1}{4} \cdot \frac{r_{ij} - r_e}{r_a - r_e} \mathbf{e}_{ij}, \\ f_{ij}(r_a \leq r_{ij} < r_0) &= \mathbf{e}_{ij}, \end{aligned} \quad (13)$$

where r_b, r_e, r_a , and r_0 are distance parameters (with $r_b < r_e < r_a < r_0$) and \mathbf{e}_{ij} is the unit vector in the direction from $x_i(t)$ to $x_j(t)$ at time t . When the distance r_{ij} between two particles is within a "repulsion" limit r_b , particle i moves away from particle j , towards the opposite direction of \mathbf{e}_{ij} . When r_{ij} is between the limits r_a and r_b , particle i adjusts its velocity in order to maintain an intermediate "equilibrium" distance r_e from j (r_e is typically the average between r_a and r_b). When the distance r_{ij} is larger than r_a , but smaller than r_0 , particle i modifies its velocity in order to get closer to j . If particle i is farther than r_0 from j , then j does not affect the cohesion component of the velocity of i .

Collective motion can also be modelled using statistical mechanics, for example, by providing a formulation for the probability distribution of the velocities \mathbf{v}_i . Bialek *et al.* [17] defined a statistical mechanics model of collective motion that can describe flocking phenomena, including the dynamics in the model by Grégoire and Chaté [16]. In its more general version, which does not take into consideration whether the particles are in the inner or outer region of the group, the statistical mechanical model is the following:

$$p(\mathbf{v}_i | J, n_c) = \frac{1}{Z(J, n_c)} \exp \left[\frac{J}{2} \sum_{i=1}^N \sum_{j \in n_c^i} \mathbf{v}_i \cdot \mathbf{v}_j \right], \quad (14)$$

where Z is the partition function and $J = v_0 a / n_c$ represents the alignment strength between particles. Crucially, such model has plausible dynamics that allows the system to relax towards, and fluctuate around, an equilibrium, which is analogous to many dynamical models: particles move according to a weighted sum of neighbors' direction while being affected by a random perturbation.

III. METHOD AND RESULTS

A. Relating information-theoretic and thermodynamic quantities in the quasistatic limit

Based on the relations presented in Secs. II A, II B, and II C, we use Eqs. (3), (5), and (9) to obtain

$$F(\theta) = -\frac{d^2 \langle \beta U_{\text{gen}} \rangle}{d\theta^2} + \frac{d^2 S}{d\theta^2}, \quad (15)$$

which then leads to the definition of

$$\frac{d^2(S)^-}{d\theta^2} \equiv \frac{d^2 S}{d\theta^2} - F(\theta) = \frac{d^2 \langle \beta U_{\text{gen}} \rangle}{d\theta^2}. \quad (16)$$

This expression, which is a key result of our study, makes it evident that the second derivative of the internal energy scaled by β (expressed on the right-hand side) is proportional to the difference between two curvatures: the curvature of the free entropy, captured by the Fisher information, and the curvature of the configuration entropy. It is important to note that Equation (15) holds in general, since U_{gen} is related, only, to the stationary distribution given by the Gibbs measure.

However, the decomposition of U_{gen} into Q_{gen} and W_{gen} ($\beta = 1$) depends on the protocol. Here we explicitly relate the Fisher information and the generalized work, energy, and heat in systems driven by quasistatic protocols. In the quasistatic limit, we show how the Fisher information can be related to the second derivative of the generalized work. By further differentiating the first law for quasistatic processes in Eq. (10) over the control parameter and by expressing it for the work term, we obtain

$$\frac{d^2 \langle \beta W_{\text{gen}} \rangle}{d\theta^2} = \frac{d^2 \langle \beta U_{\text{gen}} \rangle}{d\theta^2} - \frac{d^2 S}{d\theta^2}, \quad (17)$$

which, by comparison with Eq. (15), leads to another important result:

$$F(\theta) = -\frac{d^2 \langle \beta W_{\text{gen}} \rangle}{d\theta^2}. \quad (18)$$

In Eq. (18), the Fisher information has an information-geometric meaning at given values of θ , while we have no physical interpretation for $\langle W_{\text{gen}}(\theta) \rangle$ unless we also specify a protocol and a path $\theta^0 \rightarrow \theta$. If we assume that θ increases, and thus $\theta > \theta^0$, then we have

$$\begin{aligned} \int_{\theta^0}^{\theta} F(\theta') d\theta' &= - \int_{\theta^0}^{\theta} \frac{d^2 \langle \beta W_{\text{gen}} \rangle}{d(\theta')^2} d\theta' \\ &= - \frac{d \langle \beta W_{\text{gen}} \rangle}{d\theta} + c(\theta^0). \end{aligned} \quad (19)$$

The value of $c(\theta^0)$ can be determined by identifying the value of the control parameter θ^* , for which small changes incur no work, i.e.,

$$\left. \frac{d \langle \beta W_{\text{gen}} \rangle}{d\theta} \right|_{\theta=\theta^*} = 0, \quad (20)$$

where we call θ^* the zero-response point. Consequently, we may write

$$\frac{d \langle \beta W_{\text{gen}} \rangle}{d\theta} = - \int_{\theta^*}^{\theta} F(\theta') d\theta'. \quad (21)$$

In many systems the value of θ^* has a particular significance computationally, as will be demonstrated in Sec. III D. Once θ^* is determined, we obtain:

$$c(\theta^0) = \int_{\theta^0}^{\theta^*} F(\theta) d\theta. \quad (22)$$

We demonstrate that there is another way to relate the Fisher information and the curvature of the configuration entropy. As described in Appendix A, the second derivative of the configuration entropy $S(\theta) = - \sum_x p(x|\theta) \ln p(x|\theta)$ over θ , can be explicitly taken, leading to our third result

$$\frac{d^2(S)^+}{d\theta^2} \equiv \frac{d^2 S}{d\theta^2} + F(\theta) = - \sum_x \frac{d^2 p(x|\theta)}{d\theta^2} \ln p(x|\theta). \quad (23)$$

Unlike Eq. (16), which captured the difference between two curvatures, Eq. (23) captures the sum of two curvatures and thus reflects a different information-geometric aspect of critical dynamics during collective motion. Contrasting Eqs. (16) and (23), the second derivative of $\langle \beta U_{\text{gen}} \rangle$ with respect to θ can be expressed in terms of $F(\theta)$ and $(S)^+$ as

$$\frac{d^2 \langle \beta U_{\text{gen}} \rangle}{d\theta^2} = \frac{d^2(S)^+}{d\theta^2} - 2F(\theta_m). \quad (24)$$

In our computational analysis, which are presented in Sec. III D, we will use Eq. (16), while also showing the profile of the aggregated curvature in Eq. (23).

Finally, we propose a measure for the *thermodynamic efficiency of computation*, defined as the reduction in uncertainty (i.e., the increase in order) from an expenditure of work for a given value of the control parameter:

$$\eta \equiv \frac{-dS/d\theta}{d \langle \beta W_{\text{gen}} \rangle / d\theta} = \frac{-dS/d\theta}{\int_{\theta^0}^{\theta^*} F(\theta') d\theta'}, \quad (25)$$

which can be considered entirely in computational terms as the ratio of increasing order at θ to the cumulative sensitivity incurred over a process from θ to the zero-response point θ^* .

B. Simulations and probability distribution of the relative particle velocity

Computing the Fisher information and the entropy of a system requires the knowledge of the probability distribution $p(x|\theta)$ of the random variable, given the control parameters. For collective motion of simulated self-propelled particles, the control parameters that we consider are the alignment strength J between particles and the number of nearest neighbors n_c of each particle, while the random variable that we consider is the particles' velocity \mathbf{v}_i with respect to the group (assuming that the probability distribution is the same for each particle in the group). Since in this study we consider a model of collective motion, the probability distribution of the of particles' velocity can be estimated from the simulation of the system. Alternatively, one can, for example, follow Bialek *et al.* [17] and estimate $p(\mathbf{v}_i|J, n_c)$ from experimental data using Eq. (14).

We simulated the dynamical model [16] in Eqs. (11) and (12) setting the weight of the alignment component to $a = Jn_c/v_0$ for several different combinations of the parameters J and n_c , with J ranging between 0.001 and 0.2 and n_c ranging between 1 and 30. In every simulation, we used $N = 512$ particles and the following values of the parameters: $r_b = 0.2$, $r_e = 0.5$, $r_a = 0.8$, $r_0 = 1$, $b = 5$, and $v_0 = 0.05$. The same setup of the model was used by Bialek *et al.* [17] to validate their statistical mechanical model and corresponds to the liquid phase identified by Grégoire and Chaté [16]. We performed 100 runs for each combination of J and n_c that we considered. At the beginning of each run, the positions of the particles were randomly set within a sphere of radius proportional to the cube root of the number of particles. The initial velocity of the particles was also randomly chosen. During each run, the three-dimensional velocities \mathbf{v}_i of each particles i were recorded for 100 time steps after a relaxation time of 50 that allows the system to reach the stationary state.

Running the system over a range of values of the control parameters, using relatively small changes and allowing for a relaxation time, not only enabled us to explore the behavior of the system across the space of the control parameters but also provided an approximation of a quasistatic protocol. For example, all runs with the same value of n_c and J varying from an initial to a final value can be considered, altogether, as a quasistatic process in which a single control parameter, J , is varied infinitesimally slowly over time. This approximation allows us to carry out the thermodynamical analysis described in Secs. III D.

The simulations (see Supplemental Video 1 [37] for a demonstration of the dynamics of the system) show that the model has two different kinetic phases of collective motion, as Grégoire and Chaté had previously pointed out in their study [16]. In the disordered motion phase, illustrated in Figure 1(a), particles keep changing direction but maintain a fairly stable collective position. This phase corresponds to lower values of the alignment weight a : In the figure, for example, the parameter J , which is directly proportional to the alignment weight, is set to a low value of 0.001, while n_c is set to 20. In the coherent motion phase, illustrated in Figure 1(b), particles face a common general direction and collectively move along it. This phase corresponds to higher values of a : In the figure, the parameter J is increased to 0.2, while n_c is again set to 20.

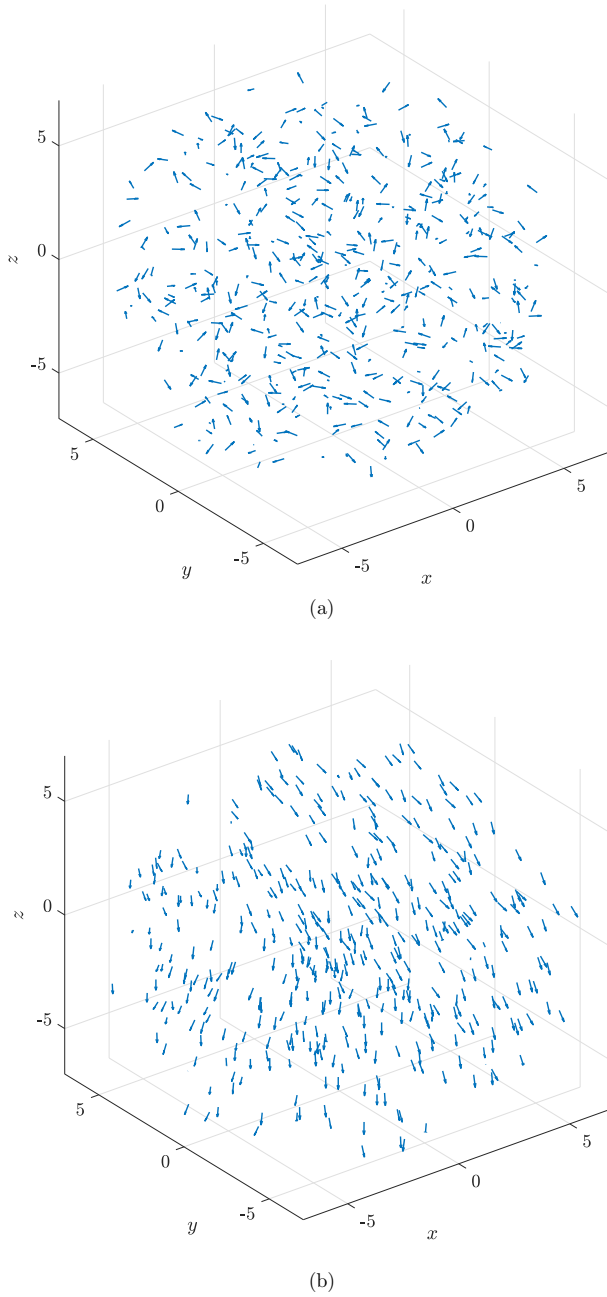


FIG. 1. Two kinetic phases of the model of collective motion. Each arrow represents a particle with its position and velocity in the (x, y, z) space. Figure 1(a) is taken from a simulation of the model in which $J = 0.001$ and $n_c = 20$, after the relaxation time, and shows the system in its disordered motion phase. Figure 1(b) is taken from a simulation in which $J = 0.2$ and $n_c = 20$, after the relaxation time, and shows the system in its coherent motion phase.

The case in which n_c is fixed at 20 and J varies from 0.001 to 0.2 is used as the main example here and throughout the rest of the article.

In order to localize the phase transition, Grégoire and Chaté [16] (as well as Vicsek *et al.* [13] in a previous model) utilized

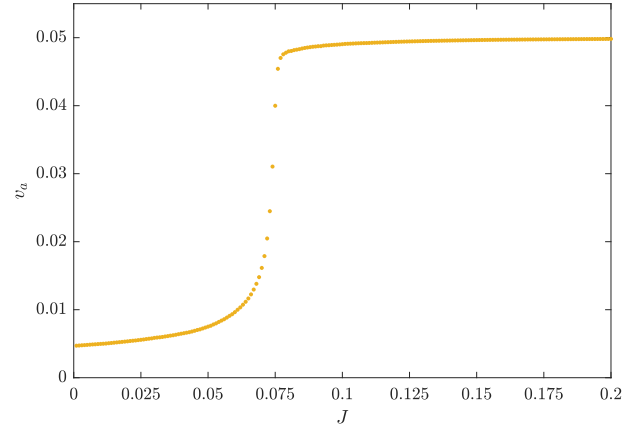


FIG. 2. Average normalized velocity of the group over the alignment strength J . The horizontal axis represents J from 0 and 0.2, at steps of 0.001, and the vertical axis represents the average normalized velocity of the group $v_a = \frac{1}{Nv_0} \left| \sum_{i=1}^N \mathbf{v}_i \right|$ over the simulation time. The parameter n_c is fixed at 20.

the order parameter

$$v_a = \frac{1}{Nv_0} \left| \sum_{i=1}^N \mathbf{v}_i \right|, \quad (26)$$

i.e., the absolute value of the average normalized velocity. We inspected v_a in our simulations for different combinations of the control parameters J and n_c . An example is given in Fig. 2, which shows the average v_a computed over all simulations for each value of J from 0.001 to 0.2 and using a fixed value of $n_c = 20$. The figure clearly reveals the phase transition: The average normalized velocity grows with the alignment strength, and the increment is particularly steep near a critical point, at approximately $J = 0.075$. A similar behavior is observed when the alignment strength J is fixed and we vary the number of nearest neighbors n_c .

The probability distribution of the particles' velocity was then estimated from the data collected from the simulations. A possible choice of the random variable is the velocity of particles with respect to the average velocity over all particles, as it changes over time. However, the average velocity over all particles is not a suitable reference for large systems (512 in our case) in the general liquid phase under consideration. In fact, even when the group is moving coherently, subgroups of particles which are far from each other can, at least temporarily, be oriented towards different directions. A more suitable choice of the random variable, and the one we made in this study, is the velocity of a particle with respect to the average velocity of other particles within a certain neighborhood (such neighborhood should not be confused with the n_c nearest neighbors). All the results presented in this paper utilize this choice of the random variable.

In order for their probability distribution to be numerically estimated, the velocities \mathbf{v}_i need to be discretized. This was done by discretizing the polar and azimuthal angles α_p and α_a [see Fig. 3(a)] of the velocity into bins measuring 4° each. For each combination of J and n_c , we estimated the probabilities $p(\mathbf{s}_k | J, n_c)$ of \mathbf{v}_i being within the cluster \mathbf{s}_k , where k enumerates the combinations of the two bins for α_p and α_a .

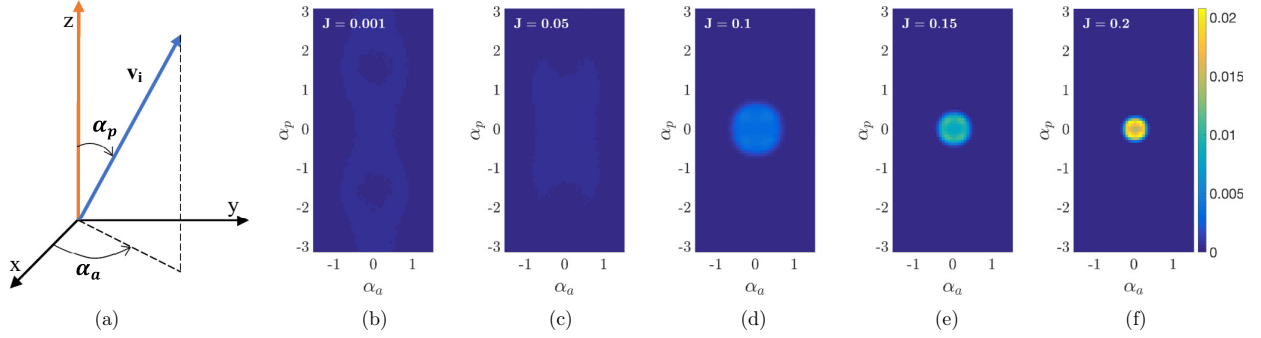


FIG. 3. Probability distribution of particles' velocity for different values of J . Figure 3(a) illustrates how a particle's velocity, with respect to the average velocity of the particles in its neighborhood, is defined by two spherical coordinates. The orange arrow represents the average velocity of the neighboring particles. A coordinate system is created so that the average velocity of the neighborhood is the z axis. The vector \mathbf{v}_i (blue arrow) is then the velocity of a particle with respect to this coordinate system, which can be expressed by the spherical coordinates $(\rho, \alpha_p, \alpha_a)$, where ρ is the radial distance, $\alpha_p \in [-\pi, \pi]$ is the polar angle and $\alpha_a \in [-\pi/2, \pi/2]$ is the azimuthal angle. Since in the model that we consider the speed of the particles is constant, α_p and α_a are sufficient for specifying \mathbf{v}_i . Figures 3(b)–3(f) show the probability distribution of the discretized cluster velocity $p(\mathbf{s}_k|J, n_c)$, for increasing values of J from 0.001 to 0.2 and with n_c fixed at 20. The horizontal and vertical axis are used for representing \mathbf{s}_k and indicate the azimuthal and polar angles, respectively, while $p(\mathbf{s}_k|J, n_c)$ is represented using a color scale, which varies from dark blue for the lowest values to light yellow for highest values.

The probabilities $p(\mathbf{s}_k|J, n_c)$ were estimated from the velocities of all the 512 particles, collected over all the 100 simulations in which the combination of J and n_c was used by dividing the number of recorded velocities within \mathbf{s}_k by the total number of recorded velocities. An example is given in Fig. 3, which shows $p(\mathbf{s}_k|J, n_c)$ for increasing values of J , from 0.001 to 0.2, fixing n_c to 20 (see Supplemental Video 2 [37] for the full change of $p(\mathbf{s}_k|J, n_c)$ over J at steps of 0.001).

For lower values of J between 0.001 and 0.5 [see Figs. 3(b) and 3(c)], which correspond to the disordered motion phase, the probability $p(\mathbf{s}_k|J, n_c)$ is distributed almost homogeneously among all velocity clusters \mathbf{s}_k , indicating that the particles' velocity is only very weakly correlated with the average velocity of their neighbors. Additionally, we can observe that within this interval of J , the probability distribution changes slowly. On the contrary, as J increases from 0.05 to 0.1 [see Figs. 3(c) and 3(d)], the probability $p(\mathbf{s}_k|J, n_c)$ intensifies around the velocity clusters \mathbf{s}_k that correspond to α_p and α_a that are closer to 0, indicating that the velocity of a particle is now more likely to be aligned with the average velocity of its neighbors. The change here is abrupt, with the probability distribution for $J = 0.1$ [Fig. 3(d)] becoming clearly nonuniform. Contrasting Fig. 3 with Fig. 2 we can see that this change happens near the critical point at 0.075. For higher values of J from 0.1 to 0.2 [see Figs. 3(d) to 3(f)], which correspond to the coherent motion phase, the probability $p(\mathbf{s}_k|J, n_c)$ keeps becoming more dense around α_p and α_a that are closer to 0, indicating that particles increasingly intensify their alignment with their neighbors.

These observations are addressed more formally in the next section, where we show that the Fisher information can quantify the sensitivity of the probability distribution to the control parameters.

C. Fisher information and the phase transition

Fisher information allows us to quantify the amount of information that velocities carry about the control parameters J and n_c . Fisher information over the alignment strength J can

be calculated from the probabilities $p(\mathbf{s}_k|J, n_c)$ estimated from the simulations, as

$$F(J, n_c) = \sum_k \frac{1}{p(\mathbf{s}_k|J, n_c)} \left[\frac{d p(\mathbf{s}_k|J, n_c)}{dJ} \right]^2, \quad (27)$$

having fixed the value of n_c . Notice that Eq. (27) is equivalent to Eq. (1), for the case in which only one control parameter is considered and the random variable is discrete. The derivative of $p(\mathbf{s}_k|J, n_c)$ over J can be computed numerically using the symmetric difference quotient two-point estimation.

We computed the Fisher information over J from 0.001 to 0.2, at steps of 0.001, for several fixed values of n_c . In Fig. 4 we show the Fisher information over J for our example case of $n_c = 20$. We can observe that the Fisher information is mostly low, except around the critical point of the kinetic

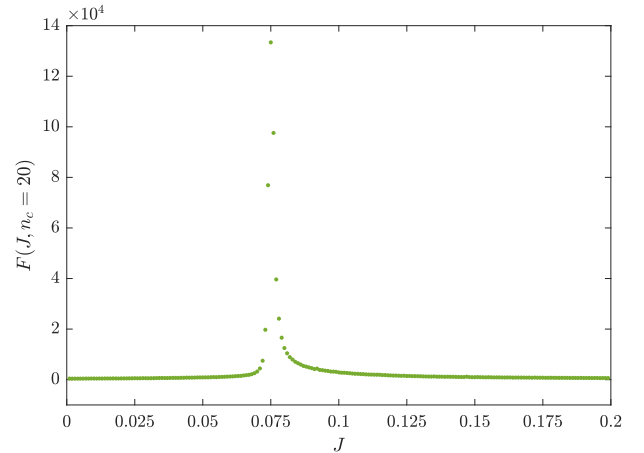


FIG. 4. Fisher information over the parameter J . The horizontal axis represents J from 0 to 0.2, at steps of 0.001, and the vertical axis represents the Fisher information $F(J, n_c)$, with the parameter n_c is fixed at 20.

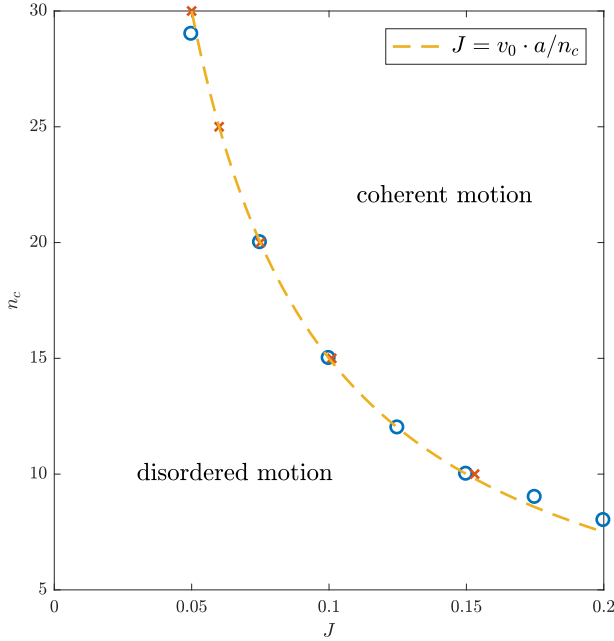


FIG. 5. Phase diagram using maximum Fisher information. The horizontal axis represents the alignment strength J between particles, while the vertical axis represents the number of nearest neighbors n_c . Red crosses indicate the values of J that yield to the higher Fisher information for fixed values of n_c from 10 to 30 at steps of 5. Analogously, blue circles indicate the values of n_c that yield to the higher Fisher information, for fixed values of J from 0.05 to 0.2 at steps of 0.025. The yellow dotted line is the function $J = v_0 a / n_c$, with $a = 30$ and $v_0 = 0.05$, which approximates the critical curve that separates the coherent and disordered motion phases.

phase transition at approximately $J = 0.075$, where it diverges positively. Analogous results were obtained using different fixed values of n_c . The Fisher information was similarly computed over the number of nearest neighbors n_c from 1 to 30, at unitary steps, for several fixed values of J between 0.001 and 0.2. An example is shown in Appendix C, where it is also evident that the Fisher information diverges at the critical point of the kinetic phase transition.

The divergence of the Fisher information at criticality, exemplified in a system of self-propelled particles performing collective motion, allows us to localize the critical points of the kinetic phase transition in a systematic and generic way, without relying on a specific order parameter, which may or may not be defined in general. Thus, this method may be used to detect phase transitions in cases in which the definition of a suitable order parameter is problematic.

Having observed that the Fisher information diverges at the critical point, we can use it to create a phase diagram of the behavior of the system, over the two control parameters J and n_c . Figure 5 shows the phase diagram that we obtained by finding, for several fixed values of n_c , the corresponding values of J that yields the maximum Fisher information and, vice versa, by finding values of n_c that yield the maximum Fisher information for several fixed values of J . We can see that the critical combinations of J and n_c can be approximated by the

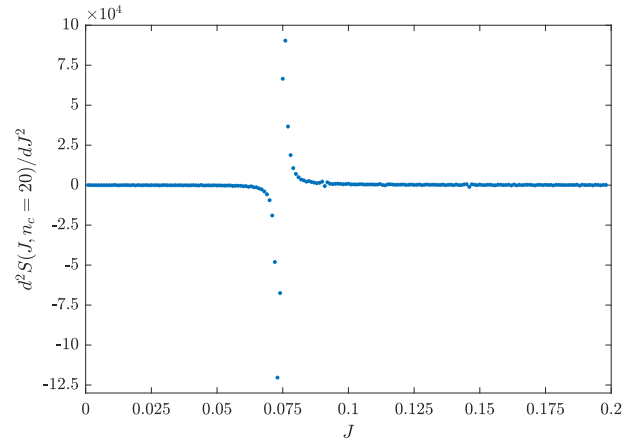


FIG. 6. Curvature of the configuration entropy of the system over J . The horizontal axis represents J from 0 to 0.2, at steps of 0.001, and the vertical axis represents the curvature of the configuration entropy of the system $S(J, n_c)$, with the parameter n_c is fixed at 20.

curve $J = v_0 a / n_c$ where, in this case, $a = 30$. This should not come as a surprise since, in the dynamical model used for the simulation, we set the weight of the alignment component to $a = J n_c / v_0$. However, the topological nature of the parameter n_c makes this result nontrivial.

D. Thermodynamical analysis of collective motion

As described in Sec. III A, the Fisher information represents the negative second derivative of the generalized work done on, or extracted from, the system due to changing the control parameter in the quasistatic limit. Therefore, Figure 4 also provides, with opposite sign, the curvature of work with respect to the alignment strength J (assuming $\beta = 1$) for our example case in which the number of nearest neighbors n_c is fixed at 20 and J varies from 0.001 to 0.02. Hence, the second derivative of work diverges negatively near the critical point.

On the other hand, the second derivative of the internal energy of the system, over a control parameter, is proportional to the difference between two curvatures: the second derivative of the configuration entropy of the system and the Fisher information (see Eq. (16)). For our system of self-propelled particles, the configuration entropy can be computed for every combination of J and n_c as

$$S(J, n_c) = - \sum_k p(\mathbf{s}_k | J, n_c) \ln p(\mathbf{s}_k | J, n_c). \quad (28)$$

The curvature of the configuration entropy was obtained by numerically computing the second derivative of the $S(J, n_c)$ determined by Eq. (28), over the parameter J , using the symmetric difference quotient two-point estimation. The result is shown in Fig. 6, while the configuration entropy itself is shown in Appendix B and its first derivative can be seen in Fig. 9(a). It can be observed that the curvature of the configuration entropy is also mostly low, except near the critical point at $J = 0.075$, where it diverges negatively from the left and positively from the right, thus exhibiting a discontinuity.

Applying Eq. (16), we can calculate the second derivative of the internal energy (scaled by β) with respect to J as the

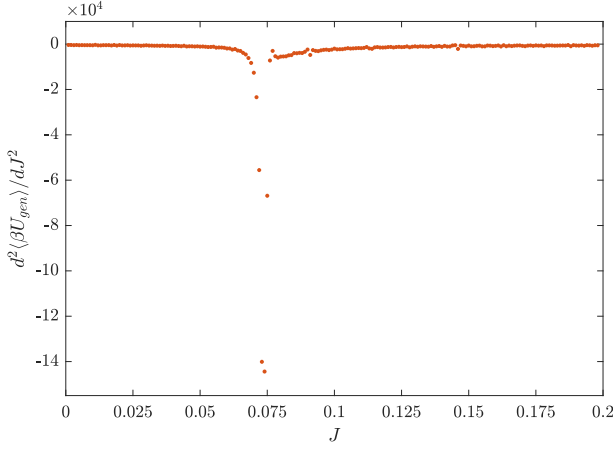


FIG. 7. Second derivative of the generalized internal energy with respect to J ($\beta = 1$). The horizontal axis represents J from 0 to 0.2, at steps of 0.001, and the vertical axis represents the second derivative of the generalized internal energy with respect to J . The parameter n_c is fixed at 20.

difference between the Fisher information in Fig. 4 and the curvature of the configuration entropy in Fig. 6, that is,

$$\frac{d^2 \langle \beta U_{\text{gen}} \rangle}{dJ^2} = \frac{d^2 S(J, n_c)}{dJ^2} - F(J, n_c), \quad (29)$$

yielding the result in Fig. 7. It can be observed that the second derivative of $\langle \beta U_{\text{gen}} \rangle$ also diverges at the critical point of the phase transition. In fact, it changes over J similarly to the second derivative of the generalized work (the opposite of the Fisher information) in Fig. 4.

If we consider the system of self-propelled particles as a system that performs distributed computation during collective motion, then Figs. 4, 6, and 7 will reveal a computational balance between the sensitivity and the uncertainty of the computation. On the one hand, the sensitivity of the system to changes in the control parameter is captured by the Fisher information in Fig. 4. On the other hand, the uncertainty of the computation is captured by the curvature of the configuration entropy of the system in Fig. 6. In either the disordered motion phase or the coherent motion phase, Fig. 7 shows that there is a balance between the sensitivity and the uncertainty, but it is clear that this balance is broken at criticality.

The sum of the Fisher information and the curvature of the configuration entropy was also inspected. For the system of self-propelled particles, this quantity varying over J is determined as

$$\frac{d^2(S)^+}{dJ^2} = - \sum_k \frac{d^2 p(\mathbf{s}_k | J, n_c)}{dJ^2} \ln p(\mathbf{s}_k | J, n_c) \quad (30)$$

and is shown in Fig. 8. It can be observed that this quantity also has a discontinuity at criticality, similarly to the curvature of the configuration entropy.

The rate of change over J of the generalized work ($\beta = 1$) and the generalized internal energy ($\beta = 1$) can be obtained by numerically integrating the curvatures of these quantities over J . As explained in Sec. III A, this can only be calculated if the integration starts from a point where the work rate, or

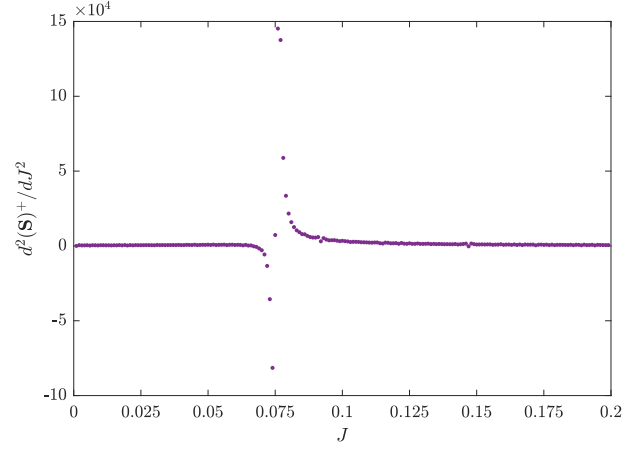


FIG. 8. Aggregated curvature $d^2(S)^+/dJ^2$. The horizontal axis represents J from 0 to 0.2, at steps of 0.001, and the vertical axis represents the aggregated curvature. The parameter n_c is fixed at 20.

the internal energy rate, with respect to J is known. In our case, we assert that the zero-response point J^* is realized as $J \rightarrow \infty$, since in this region dJ produces no work, because all the particles are already perfectly aligned. Consequently, we associate the zero-response point with the state of perfect order. In our case, we have $J^* = \infty$ and choose $J^0 = 0$, which, according to Eq. (22), yields

$$c(0) = \int_0^\infty F(J, n_c) dJ \quad (31)$$

and

$$\frac{d \langle \beta W_{\text{gen}} \rangle}{dJ} = - \int_0^J F(J', n_c) dJ' + c(0). \quad (32)$$

Computing $c(0)$ requires a numerical estimation, which we approximated to have a lower bound, $c(0) > 800$, for $n_c = 20$. This is reflected in all plots. The integration was done using the cumulative trapezoidal numerical integration, and the result is shown in Fig. 9(a). We can see that the rate of change of the generalized work (green crosses) decreases with J . Figure 9(a) also shows the first derivative of the configuration entropy over J (blue asterisks). As we can see, the configuration entropy decreases around the critical point, where the system of self-propelled particles self-organizes in a more ordered phase and begins to display coherent collective motion.

Importantly, as the alignment strength J increases, the entropy decreases and the work rate is positive: Generating order requires work to be expended. Specifically at the critical point we find that the ratio of generated order to the work rate peaks, indicating that the maximal thermodynamical efficiency of computation carried out by the system of self-propelled particles, that is,

$$\eta = \frac{-dS(J, n_c)/dJ}{d \langle \beta W_{\text{gen}} \rangle / dJ} = \frac{-dS(J, n_c)/dJ}{\int_J^\infty F(J', n_c) dJ'} \quad (33)$$

is the highest at criticality [see Fig. 9(b)]. Explicitly, in computational terms, the maximum thermodynamic efficiency corresponds to a maximal ratio of generated order to the sensitivity accumulated over a process running from the

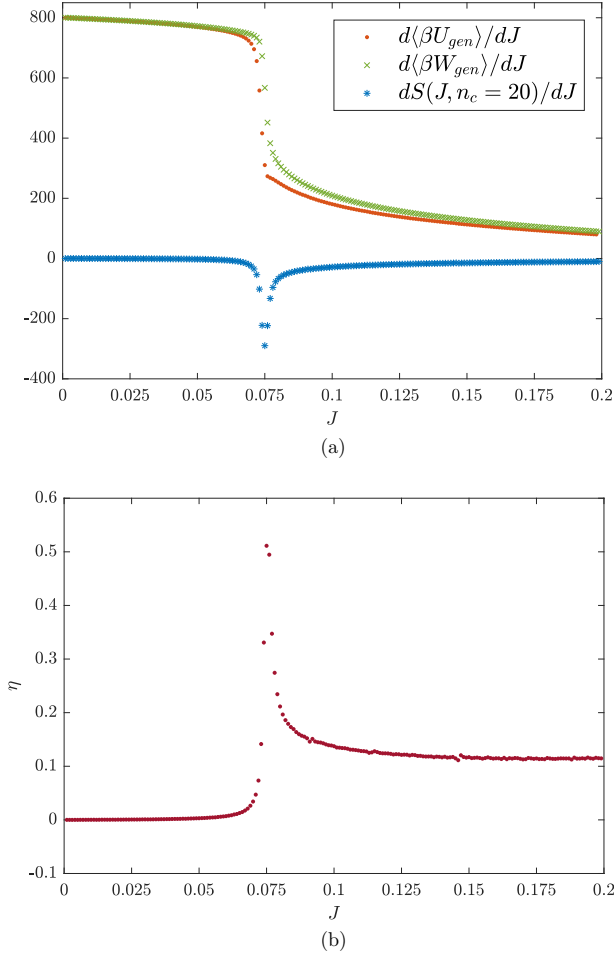


FIG. 9. Rates of change of work, internal energy, and configuration entropy with respect to J , with n_c is fixed at 20. In both graphs, the horizontal axis represents J from 0 to 0.2, at steps of 0.001. The green crosses in Fig. 9(a) represent the rate of change with respect to J of the generalized work ($\beta = 1$), the red dots represent the rate of change with respect to J of the generalized internal energy ($\beta = 1$), and the blue asterisks represent the rate of change with respect to J of the configuration entropy. Figure 9(b) shows the thermodynamic efficiency of computation η over J .

current state to the state of perfect order (the zero-response point). Since the Fisher information is always positive, the denominator $\int_J^\infty F(J')dJ'$ can be interpreted as a measure of distance, along the trajectories of J , from the perfectly ordered state. Thus, it scales the increase in order as the control parameter changes. For example, achieving one bit of uncertainty reduction near the state of perfect order is much more significant than achieving one bit of uncertainty reduction in a largely disordered state. This means that at criticality, the reduction of uncertainty is the most significant, reflected in the highest thermodynamic efficiency of computation.

A similar analysis was conducted for the case in which the alignment strength J is fixed and the parameter n_c changes. The results are presented in Appendix C.

IV. DISCUSSION AND CONCLUSIONS

In this study we presented a theoretical framework for measuring fundamental thermodynamical quantities, such as the generalized heat, work, and energy, in systems driven by quasistatic protocols. Importantly, such framework relates these thermodynamical quantities to information-theoretic ones and specifically to the configuration entropy and the Fisher information. We applied the framework to a system of simulated self-propelled particles during collective motion and studied the (generalized) internal energy and work done on, or extracted from, the system as it goes through a kinetic phase transition.

The model of collective motion that we considered is the one proposed by Grégoire and Chaté [16], which is known to have a kinetic phase transition over control parameters influencing the particles' alignment: from a disordered motion phase, in which particles maintain a fairly stable collective position, to a coherent motion phase, in which particles cohesively move towards a common direction. We have considered two control parameters, i.e., the alignment strength between particles and the number of nearest neighbors influencing the particles' alignment, within intervals in which the kinetic phase transition is observed. In order to approximate a quasistatic protocol, we simulated the system for chosen values of the control parameters, under the assumption that the system reaches a stationary state after a certain relaxation time, and we repeated the experiments for different values of the control parameters. We also used the data collected with the simulations to numerically estimate the probability distribution of the velocity of the particles at different values of the control parameters.

Our approach involves a statistical mechanical formulation of the second derivatives of the generalized internal energy and the generalized work with respect to the control parameters, based on relationships between these quantities and two other quantities, the Fisher information and the curvature of the configuration entropy, which can be calculated from the probability distribution of the velocities. Additionally, our method provides an information-geometric interpretation of the curvature of the internal energy of the system (scaled by β) as the difference between two curvatures: the curvature of the free entropy, captured by the Fisher information, and the curvature of the configuration entropy [Eq. (16)]. Another expression [Eq. (23)], also interpreted information-geometrically as an aggregated curvature, is given for the sum of the Fisher information and the curvature of the configuration entropy of the system.

The expression representing the difference between curvatures [Eq. (16)] highlights the computational balance between the sensitivity of the computation, captured by the Fisher information, and the uncertainty of the computation, captured by the configuration entropy, that is performed by the system of self-propelled particles during collective motion. Our numerical results show that such balance is stressed at criticality, where the curvatures with respect of the control parameters of the generalized work and the generalized internal energy, as well as the curvature of the configuration entropy of the system, diverge. The rates of change of the generalized work and the generalized internal energy over the control parameters were estimated from the curvatures of these quantities, using

numerical integration. The results show that during the kinetic phase transition, when the configuration entropy of the system decreases very rapidly, both the rate of change of the work and the internal energy decrease dramatically.

Our results support the view that flocking behavior, which combines coherence and responsiveness to external perturbations (e.g., predatory attacks), exhibits criticality in the statistical mechanical sense [17,18,38,39]. Moreover, our results suggest that the highest thermodynamic efficiency of computation η , relating the reduction of the configuration entropy to the required work rate, is achieved at criticality. We have also shown that this quantity can be interpreted as a significance of entropy reduction with respect to the distance along a computational path to a perfectly ordered state, where the distance is understood to mean the cumulative sensitivity captured by the integral of the Fisher information.

When applying the proposed theoretical framework, it is crucial to imbue the derivative with respect to the control parameter with physical meaning. In this study, we have considered the most natural case of a quasistatic protocol; however, less conventional alternatives can be constructed. For instance, one can think of a feedback process, in which the protocol is changed in response to measurements of the process [40–42]. If the measurements gain equal or more information than the free energy change, then the measurement can be used to change the protocol so that zero work is performed (or extracted) on changing the control parameter. Because of the first law of thermodynamics, if no work is done, then $\Delta\langle U_{\text{gen}} \rangle = \Delta\langle Q_{\text{gen}} \rangle$, which, following Eq. (29) or (16), leads to:

$$\frac{d^2\langle S \rangle^-}{dJ^2} = \frac{d^2\langle \beta Q_{\text{gen}} \rangle}{dJ^2} = \frac{d^2 S(J, n_c)}{dJ^2} - F(J, n_c). \quad (34)$$

Thus, the thermodynamic interpretation of $d^2\langle S \rangle^-/dJ^2$, provided by Eq. (16), changes: It is no longer the curvature of the generalized internal energy of the system (scaled by β). It is instead the curvature of the heat (scaled by β), which can be interpreted as the rate of change of the entropy flux Φ_J from the system to the environment [32,43]:

$$\Phi_J = \int \frac{d^2\langle \beta Q_{\text{gen}} \rangle}{dJ^2} dJ. \quad (35)$$

If we assume that the whole system, which includes the self-propelled particles as well as the environment, is isolated, then its total entropy production Π_J is the difference between the rate of change of the configuration entropy of the system of self-propelled particles and the entropy flux to the environment (given the sign convention):

$$\Pi_J = \frac{dS(J, n_c)}{dJ} - \Phi_J. \quad (36)$$

In light of this relationship, integrating Eq. (34) leads to the interpretation of the Fisher information as the rate of change of the total entropy production with respect to the control parameter:

$$\Pi_J = \int F(J, n_c) dJ. \quad (37)$$

Hence, if we look again at Fig. 9(a), but this time considering the feedback process, it is clear that increasing J would

lead to a negative spike in total entropy production because the information has been used to reduce the work done on the system, thus decreasing irreversibility. In contrast, for decreasing J a positive spike would be observed because the information is being used counterproductively to extract zero work when positive work could be extracted, increasing irreversibility. Interestingly, the ratio of the rate of change of the configuration entropy of the system to the total entropy production [see Fig. 9(b)] would be highest at criticality.

Total entropy production and entropy flux have been studied in a variety of systems, including the majority-vote model [44], copolymerization processes [45], a population model [46], interacting lattice gas [47], and the Ising model [48,49], among others. All these studies have identified phase transitions over some control parameter (for instance, the temperature and the coupling constant were chosen as control parameters in the Ising model [48,49]). The theoretical framework proposed in this study could be applied to a range of processes in which it can be assumed that no work is done on, or extracted from, the system.

In addition to our main results, we have also shown that the critical points of the kinetic phase transition are captured by the divergence of the Fisher information. This allowed us to use this measure to construct a phase diagram of the dynamics of the system for different combinations of the two control parameters considered, showing the critical regime separating the coherent and disordered motion phases.

Broadly, our results contribute to “information thermodynamics,” an emerging field exploring relationships between information processing and its thermodynamic costs [50–58]. These relationships are of particular interest for complex systems which need to perform their distributed computation efficiently [59–63]. We hope that our work would contribute towards a unified theory of collective motion drawing on statistical mechanics and information thermodynamics, applicable to diverse collective motion phenomena including active matter [64,65].

ACKNOWLEDGMENTS

E.C. was supported by the University of Sydney’s “Post-graduate Scholarship in the Field of Complex Systems” from Faculty of Engineering & IT and by a CSIRO top-up scholarship. J.L. was supported through the Australian Research Council DECRA Grant No. DE160100630. All authors were supported by The University of Sydney’s DVC Research Strategic Research Excellence Initiative (SREI-2020) project, “CRISIS: Crisis Response in Interdependent Social-Infrastructure Systems” (Grant No. IRMA 194163). The authors acknowledge the University of Sydney HPC service at The University of Sydney for providing HPC resources that have contributed to the research results reported within this paper.

APPENDIX A: DERIVATION OF THE CURVATURE OF THE SYSTEM’S ENTROPY

The Fisher information $F(\theta)$ can be related to the second derivative of the configuration entropy $S(\theta)$ as follows. The

EMANUELE CROSATO *et al.*PHYSICAL REVIEW E **97**, 012120 (2018)first derivative of $S(\theta)$ over θ is

$$\begin{aligned}
\frac{dS}{d\theta} &= - \sum_x \frac{d(p(x|\theta) \ln p(x|\theta))}{d\theta} \\
&= - \sum_x \left[\frac{d p(x|\theta)}{d\theta} \ln p(x|\theta) + \frac{d p(x|\theta)}{d\theta} \right] \\
&= - \sum_x \frac{d p(x|\theta)}{d\theta} \ln p(x|\theta) - \sum_x \frac{d p(x|\theta)}{d\theta} \\
&= - \sum_x \frac{d p(x|\theta)}{d\theta} \ln p(x|\theta) - \frac{d \sum_x p(x|\theta)}{d\theta} \\
&= - \sum_x \frac{d p(x|\theta)}{d\theta} \ln p(x|\theta). \quad (A1)
\end{aligned}$$

The second derivative $S(\theta)$ over θ is

$$\begin{aligned}
\frac{d^2 S}{d\theta^2} &= - \sum_x \frac{d \left[\frac{d p(x|\theta)}{d\theta} \ln p(x|\theta) \right]}{d\theta} \\
&= - \sum_x \frac{d^2 p(x|\theta)}{d\theta^2} \ln p(x|\theta) \\
&\quad - \sum_x \frac{1}{p(x|\theta)} \left[\frac{d p(x|\theta)}{d\theta} \right]^2 \\
&= - \sum_x \frac{d^2 p(x|\theta)}{d\theta^2} \ln p(x|\theta) - F(\theta). \quad (A2)
\end{aligned}$$

Comparing Equation (A2) with the definition of $\frac{d^2(S)^+}{d\theta^2}$ given by Eq. (23), that is,

$$\frac{d^2(S)^+}{d\theta^2} \equiv \frac{d^2 S}{d\theta^2} + F(\theta), \quad (A3)$$

yields:

$$\frac{d^2(S)^+}{d\theta^2} = - \sum_x \frac{d^2 p(x|\theta)}{d\theta^2} \ln p(x|\theta). \quad (A4)$$

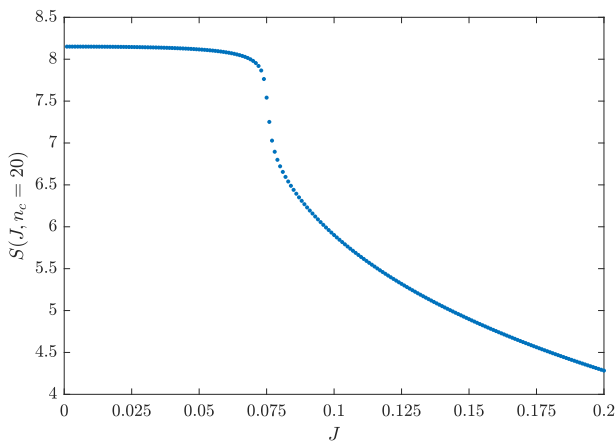


FIG. 10. Entropy of the system over J . The horizontal axis represents J from 0 to 0.2, at steps of 0.001, and the vertical axis represents the entropy of the system $S(J, n_c)$, with the parameter n_c is fixed at 20.

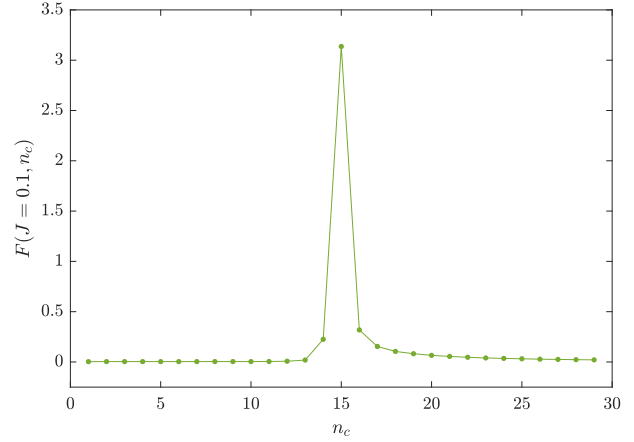


FIG. 11. Fisher information over n_c . The horizontal axis represents n_c from 1 to 30, and the vertical axis represents the Fisher information $F_{n_c}(J, n_c)$, with the parameter J is fixed at 0.1.

APPENDIX B: CONFIGURATION ENTROPY OVER THE ALIGNMENT STRENGTH

Figure 10 shows that the configuration entropy of the system decreases with J , as the group becomes more polarized towards a flocking direction, with the drop being particularly steep in the proximity of the critical point.

APPENDIX C: ENTROPY PRODUCTION AND FLUX OVER THE NUMBER OF NEAREST NEIGHBORS

The thermodynamical analysis in Sec. IIID has also been carried out changing the control parameter n_c (the number of nearest neighbors affecting the alignment component of

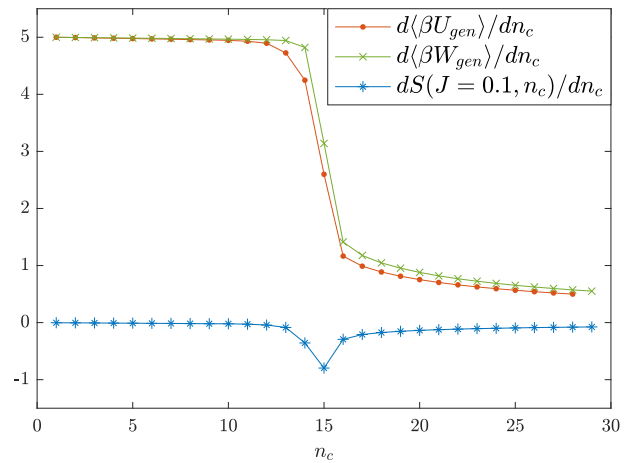


FIG. 12. Rates of change of work, internal energy, and configuration entropy with respect to n_c with J is fixed at 0.1. The horizontal axis represents n_c from 1 to 30. The green crosses represent the rate of change with respect to n_c of the generalized work ($\beta = 1$), the red dots represent the rate of change with respect to n_c of the generalized internal energy ($\beta = 1$), and the blue asterisks represent the rate of change with respect to n_c of the configuration entropy.

particles motion), while fixing the control parameter J (the alignment strength). Analogous results to varying J while fixing n_c have been obtained, some of which are shown in this appendix. Figure 11 shows that the Fisher information, which also represents the opposite of the curvature of the generalized work (scaled by β) with respect to the number of neighbors, diverges at the critical point $n_c = 15$. Figure 12 shows the rates of change with respect to the number of nearest neighbors of

the generalized work ($\beta = 1$), the generalized internal energy ($\beta = 1$), and the configuration entropy. Computing $c(n_c = 1)$ requires a numerical estimation, which we approximated to have a lower bound $c(1) > 5$ for $J = 0.1$. The rates of change of work and internal energy decrease with n_c , and the drop is particularly steep at criticality. The rate of change of the configuration entropy is instead generally low, apart from near the critical point, where it drops.

-
- [1] T. Vicsek and A. Zafeiris, Collective motion, *Phys. Rep.* **517**, 71 (2012).
 - [2] M. Ballerini, N. Cabibbo, R. Candelier, A. Cavagna, E. Cisbani, I. Giardina, V. Lecomte, A. Orlandi, G. Parisi, A. Procaccini, M. Viale, and V. Zdravkovic, Interaction ruling animal collective behavior depends on topological rather than metric distance: Evidence from a field study, *Proc. Natl. Acad. Sci. USA* **105**, 1232 (2008).
 - [3] J. K. Parrish, S. V. Viscido, and D. Grunbaum, Self-organized fish schools: An examination of emergent properties, *Biol. Bull.* **202**, 296 (2002).
 - [4] J. Buhl, D. J. T. Sumpter, I. Couzin, J. J. Hale, E. Despland, E. R. Miller, and S. J. Simpson, From disorder to order in marching locusts, *Science* **312**, 1402 (2006).
 - [5] A. Sokolov, R. E. Goldstein, F. I. Feldchtein, and I. S. Aranson, Enhanced mixing and spatial instability in concentrated bacterial suspensions, *Phys. Rev. E* **80**, 031903 (2009).
 - [6] B. Szabo, G. J. Szöllösi, B. Gönci, Z. Jurányi, D. Selmeczi, and T. Vicsek, Phase transition in the collective migration of tissue cells: Experiment and model, *Phys. Rev. E* **74**, 061908 (2006).
 - [7] V. Schaller, C. Weber, C. Semmrich, E. Frey, and A. R. Bausch, Polar patterns of driven filaments, *Nature* **467**, 73 (2010).
 - [8] M. Ibele, T. E. Mallouk, and A. Sen, Schooling behavior of light-powered autonomous micromotors in water, *Angew. Chem. Int. Ed.* **48**, 3308 (2009).
 - [9] S. Camazine, *Self-Organization in Biological Systems*, Princeton Studies in Complexity (Princeton University Press, Princeton, NJ, 2001).
 - [10] I. D. Couzin, Collective minds, *Nature* **445**, 715 (2007).
 - [11] X. R. Wang, J. M. Miller, J. T. Lizier, M. Prokopenko, and L. F. Rossi, Quantifying and tracing information cascades in swarms, *Plos One* **7**, e40084 (2012).
 - [12] I. D. Couzin, Collective cognition in animal groups, *Trends Cognit. Sci.* **13**, 36–43 (2009).
 - [13] T. Vicsek, A. Czirók, E. Ben-Jacob, I. Cohen, and O. Shochet, Novel Type of Phase Transition in a System of Self-Driven Particles, *Phys. Rev. Lett.* **75**, 1226 (1995).
 - [14] J. Toner and Y. Tu, Long-range Order in a Two-Dimensional Dynamical XY Model: How Birds Fly Together, *Phys. Rev. Lett.* **75**, 4326 (1995).
 - [15] J. Toner and Y. Tu, Flocks, herds, and schools: A quantitative theory of flocking, *Phys. Rev. E* **58**, 4828 (1998).
 - [16] G. Grégoire and H. Chaté, Onset of Collective and Cohesive Motion, *Phys. Rev. Lett.* **92**, 025702 (2004).
 - [17] W. Bialek, A. Cavagna, I. Giardina, T. Mora, E. Silvestri, M. Viale, and A. M. Walczak, Statistical mechanics for natural flocks of birds, *Proc. Natl. Acad. Sci. USA* **109**, 4786 (2012).
 - [18] W. Bialek, A. Cavagna, I. Giardina, T. Mora, O. Pohl, E. Silvestri, M. Viale, and A. M. Walczak, Social interactions dominate speed control in poising natural flocks near criticality, *Proc. Natl. Acad. Sci. USA* **111**, 7212 (2014).
 - [19] M. Castellana, W. Bialek, A. Cavagna, and I. Giardina, Entropic effects in a nonequilibrium system: Flocks of birds, *Phys. Rev. E* **93**, 052416 (2016).
 - [20] E. T. Jaynes, Information theory and statistical mechanics, *Phys. Rev.* **106**, 620 (1957).
 - [21] M. Ballerini, N. Cabibbo, R. Candelier, A. Cavagna, E. Cisbani, I. Giardina, A. Orlandi, G. Parisi, A. Procaccini, M. Viale, and V. Zdravkovic, Empirical investigation of starling flocks: A benchmark study in collective animal behaviour, *Anim. Behav.* **76**, 201 (2008).
 - [22] A. Cavagna, I. Giardina, A. Orlandi, G. Parisi, A. Procaccini, M. Viale, and V. Zdravkovic, The STARFLAG handbook on collective animal behaviour: 1. Empirical methods, *Anim. Behav.* **76**, 217 (2008).
 - [23] A. Cavagna, I. Giardina, A. Orlandi, G. Parisi, and A. Procaccini, The STARFLAG handbook on collective animal behaviour: 2. Three-dimensional analysis, *Anim. Behav.* **76**, 237 (2008).
 - [24] R. A. Fisher, On the mathematical foundations of theoretical statistics, *Philos. Trans. R. Soc. London A* **222**, 309 (1922).
 - [25] D. C. Brody and N. Rivier, Geometrical aspects of statistical mechanics, *Phys. Rev. E* **51**, 1006 (1995).
 - [26] D. C. Brody and A. Ritz, Information geometry of finite Ising models, *J. Geom. Phys.* **47**, 207 (2003).
 - [27] W. Janke, D. A. Johnston, and R. Kenna, Information geometry and phase transitions, *Physica A* **336**, 181 (2004).
 - [28] G. E. Crooks, Measuring Thermodynamic Length, *Phys. Rev. Lett.* **99**, 100602 (2007).
 - [29] G. E. Crooks, Fisher Information and Statistical Mechanics, Tech. Rep. (2011).
 - [30] M. Prokopenko, J. T. Lizier, O. Obst, and X. R. Wang, Relating Fisher information to order parameters, *Phys. Rev. E* **84**, 041116 (2011).
 - [31] B. B. Machta, R. Chachra, M. K. Transtrum, and J. P. Sethna, Parameter space compression underlies emergent theories and predictive models, *Science* **342**, 604 (2013).
 - [32] M. Prokopenko and I. Einav, Information thermodynamics of near-equilibrium computation, *Phys. Rev. E* **91**, 062143 (2015).
 - [33] A. Plastino, A. R. Plastino, and H. G. Miller, On the relationship between the Fisher-Frieden-Soffer arrow of time, and the behaviour of the boltzmann and kullback entropies, *Phys. Lett. A* **235**, 129 (1997).
 - [34] B. R. Frieden, Fisher information and uncertainty complementarity, *Phys. Lett. A* **169**, 123 (1992).

- [35] B. R. Frieden, A. Plastino, A. R. Plastino, and B. H. Soffer, Fisher-based thermodynamics: Its Legendre transform and concavity properties, *Phys. Rev. E* **60**, 48 (1999).
- [36] D. Petz, Covariance and Fisher information in quantum mechanics, *J. Phys. A* **35**, 929 (2002).
- [37] See Supplemental Material at <http://link.aps.org/supplemental/10.1103/PhysRevE.97.012120> for videos of the dynamics of the system and the probability distribution.
- [38] A. Cavagna, A. Cimarelli, I. Giardina, G. Parisi, R. Santagati, F. Stefanini, and M. Viale, Scale-free correlations in starling flocks, *Proc. Natl. Acad. Sci. USA* **107**, 11865 (2010).
- [39] T. Mora and W. Bialek, Are biological systems poised at criticality? *J. Stat. Phys.* **144**, 268 (2011).
- [40] S. Toyabe, T. Sagawa, M. Ueda, E. Muneyuki, and M. Sano, Experimental demonstration of information-to-energy conversion and validation of the generalized Jarzynski equality, *Nat. Phys.* **6**, 988 (2010).
- [41] T. Sagawa and M. Ueda, Generalized Jarzynski Equality Under Nonequilibrium Feedback Control, *Phys. Rev. Lett.* **104**, 090602 (2010).
- [42] T. Sagawa and M. Ueda, Nonequilibrium Thermodynamics of Feedback Control, *Phys. Rev. E* **85**, 021104 (2012).
- [43] R. K. Niven, Minimization of a free-energy-like potential for non-equilibrium flow systems at steady state, *Philos. Trans. R. Soc. B* **365**, 1323 (2010).
- [44] L. Crochik and T. Tomé, Entropy production in the majority-vote model, *Phys. Rev. E* **72**, 057103 (2005).
- [45] D. Andrieux and P. Gaspard, Nonequilibrium generation of information in copolymerization processes, *Proc. Natl. Acad. Sci. USA* **105**, 9516 (2008).
- [46] B. Andrae, J. Cremer, T. Reichenbach, and E. Frey, Entropy production of cyclic population dynamics, *Phys. Rev. Lett.* **104**, 218102 (2010).
- [47] T. Tomé and M. J. de Oliveira, Entropy production in nonequilibrium systems at stationary states, *Phys. Rev. Lett.* **108**, 020601 (2012).
- [48] M. J. de Oliveira, Irreversible models with Boltzmann–Gibbs probability distribution and entropy production, *J. Stat. Mech.: Theory Exp.* (2011) P12012.
- [49] Y. Zhang and A. C. Barato, Critical behavior of entropy production and learning rate: Ising model with an oscillating field, *J. Stat. Mech.: Theory Exp.* (2016) 113207.
- [50] M. Esposito and C. Van den Broeck, Second law and Landauer principle far from equilibrium, *Europhys. Lett.* **95**, 40004 (2011).
- [51] S. Deffner and C. Jarzynski, Information processing and the second law of thermodynamics: An inclusive, Hamiltonian approach, *Phys. Rev. X* **3**, 041003 (2013).
- [52] A. C. Barato, D. Hartich, and U. Seifert, Efficiency of cellular information processing, *New J. Phys.* **16**, 103024 (2014).
- [53] J. M. Horowitz and H. Sandberg, Second-law-like inequalities with information and their interpretations, *New J. Phys.* **16**, 125007 (2014).
- [54] J. M. Horowitz and M. Esposito, Thermodynamics with continuous information flow, *Phys. Rev. X* **4**, 031015 (2014).
- [55] J. M. R. Parrondo, J. M. Horowitz, and T. Sagawa, Thermodynamics of information, *Nat. Phys.* **11**, 131 (2015).
- [56] M. Prokopenko, L. Barnett, M. Harré, J. T. Lizier, O. Obst, and X. R. Wang, Fisher transfer entropy: Quantifying the gain in transient sensitivity, *Proc. R. Soc. A* **471**, 20150610 (2015).
- [57] R. E. Spinney, J. T. Lizier, and M. Prokopenko, Transfer entropy in physical systems and the arrow of time, *Phys. Rev. E* **94**, 022135 (2016).
- [58] R. E. Spinney, M. Prokopenko, and J. T. Lizier, Transfer entropy in continuous time, with applications to jump and neural spiking processes, *Phys. Rev. E* **95**, 032319 (2017).
- [59] J. T. Lizier, M. Prokopenko, and A. Y. Zomaya, Coherent information structure in complex computation, *Theory Biosci.* **131**, 193 (2012).
- [60] M. Prokopenko, J. T. Lizier, and D. C. Price, On thermodynamic interpretation of transfer entropy, *Entropy* **15**, 524 (2013).
- [61] M. Prokopenko and J. T. Lizier, Transfer entropy and transient limits of computation, *Sci. Rep.* **4**, 5394 (2014).
- [62] A. Kolchinsky and D. H. Wolpert, Dependence of dissipation on the initial distribution over states, *J. Stat. Mech.* (2017) 083202.
- [63] C. P. Kempes, D. Wolpert, Z. Cohen, and J. Pérez-Mercader, The thermodynamic efficiency of computations made in cells across the range of life, *Phil. Trans. R. Soc. A* **375**, 20160343 (2017).
- [64] G. Popkin, The physics of life, *Nature* **529**, 16 (2016).
- [65] R. Di Leonardo, Active colloids: Controlled collective motions, *Nat. Mater.* **15**, 1057 (2016).

Chapter 5

Critical dynamics and thermodynamics of urban transformations

ROYAL SOCIETY
OPEN SCIENCE

rsos.royalsocietypublishing.org

Research



Cite this article: Crosato E, Nigmatullin R, Prokopenko M. 2018 On critical dynamics and thermodynamic efficiency of urban transformations. *R. Soc. open sci.* **5**: 180863. <http://dx.doi.org/10.1098/rsos.180863>

Received: 8 June 2018

Accepted: 13 September 2018

Subject Category:

Physics

Subject Areas:

complexity

Keywords:

urban modelling, thermodynamic efficiency, maximum entropy principle, phase transitions, criticality, Fisher information

Author for correspondence:

Emanuele Crosato

e-mail: emanuele.crosato@sydney.edu.au

Electronic supplementary material is available online at <https://dx.doi.org/10.6084/m9.figshare.c.4252601>.

THE ROYAL SOCIETY
PUBLISHING

On critical dynamics and thermodynamic efficiency of urban transformations

Emanuele Crosato^{1,2}, Ramil Nigmatullin¹
and Mikhail Prokopenko¹

¹Complex Systems Research Group and Centre for Complex Systems, Faculty of Engineering and IT, The University of Sydney, Sydney, New South Wales 2006, Australia

²CSIRO Data61, PO Box 76, Epping, New South Wales 1710, Australia

EC, 0000-0002-7629-774X; RN, 0000-0003-2577-6561; MP, 0000-0002-4215-0344

Urban transformations within large and growing metropolitan areas often generate critical dynamics affecting social interactions, transport connectivity and income flow distribution. We develop a statistical–mechanical model of urban transformations, exemplified for Greater Sydney, and derive a thermodynamic description highlighting critical regimes. We consider urban dynamics at two time scales: fast dynamics for the distribution of population and income, modelled via the maximum entropy principle, and slower dynamics evolving the urban structure under spatially distributed competition. We identify phase transitions between dispersed and polycentric phases, induced by varying the *social disposition*—a factor balancing the suburbs’ attractiveness—in contrast with the *travel impedance*. Using the Fisher information, we identify critical thresholds and quantify the thermodynamic cost of urban transformation, as the minimal work required to vary the underlying parameter. Finally, we introduce the notion of *thermodynamic efficiency of urban transformation*, as the ratio of the order gained during a change to the amount of required work, showing that this measure is maximized at criticality.

1. Introduction

A city is quintessentially a complex system consisting of multiple interacting agents such as individual residents, employment centres and transport infrastructure [1,2]. The complexity manifests itself through diverse spatial organizations: monocentric cities where most of the economic activity takes place at the CBD [3], polycentric cities with multiple subcentres (or ‘edge’ cities) [4–6] and dispersed sprawl (or ‘edgeless’) cities [7]. Moreover, cities can undergo transitions in their urban

structures. Driving such transitions are changes in the factors determining economies and diseconomies of agglomeration for both firms and residents [8–11].

While urban dynamics have been extensively studied in the past, a unified framework centred on thermodynamics of urban transformations has not been yet developed (see [12]). In particular, the analysis and modelling of urban transformations as phase transitions, defined in a rigorous thermodynamic setting, remains an open challenge, despite recent attempts in spatial economics over short time scales [11]. This paper aims to refocus the research field on Urban Thermodynamics, considering critical phenomena including phase transitions in a principled way, based on the underlying thermodynamic concepts (energy potentials, entropy, order parameters, etc.), for both equilibrium and non-equilibrium scenarios. This approach will enable systematic calibrations of such thermodynamic models with real-world data and scenarios at different overlapping time scales.

We develop a statistical–mechanical model displaying phase transitions, using the maximum entropy principle in a dynamic setting, and define the thermodynamic efficiency of urban transformations. The model is calibrated to Greater Sydney Census data and is shown to exhibit a phase transition between a monocentric dispersed and polycentric clustered urban forms. This phase transition is induced by the variation of the attractiveness of the residential neighbourhoods, measured by the density of local services, given the transportation cost. While quantitative studies of urban transformation typically focus on statistical analysis of spatial evolution of cities [13–17], the thermodynamic approach developed in this paper enables a rigorous analysis of critical dynamics in a wide class of urban systems, as well as quantitative explorations of diverse ‘what-if’ scenarios with respect to a generic and precise efficiency measure.

Our model is based on the Boltzmann–Lotka–Volterra (BLV) method [18–21]. The BLV models involve two components: a fast equilibration, ‘Boltzmann’, component and a slow dynamic, ‘Lotka–Volterra’, component. The Boltzmann component applies maximum entropy principle to derive the static flow patterns of commodities and residents consistent with given spatial distributions [22]. The Lotka–Volterra component evolves the spatial distribution and the flow pattern of a commodity according to generalized Lotka–Volterra equations for spatially distributed competitors. In our model of Greater Sydney the Lotka–Volterra equations make suburbs compete for local services, and suburbs with more services become more attractive residential places. The resultant urban dynamics exhibit critical regimes, interpreted as urban phase transitions, where a small variation in suitably chosen (control) parameters changes the global outcomes measured via specific aggregated quantities (order parameters). This is in line with the idea that, despite the complexity of urban system, only few parameters may be necessary to understand drastic macroscopic changes [23].

The maximum entropy method [24] has been applied to a variety of collective phenomena [25,26] and urban modelling [22], suggesting a formal analogy between urban and thermodynamic systems [27–29]. In studying transformations in the Greater Sydney region as thermodynamic phenomena, we construct the corresponding phase diagram with respect to suitably chosen control parameters. In doing so we use the Fisher information, which measures the sensitivity of a probability distribution to the change in the control parameter, and diverges at critical points [30–34].

Our analysis further deepens the analogy between urban science and thermodynamics, using a clear thermodynamic interpretation of the Fisher information as the second derivative of free entropy. Specifically, we investigate the minimum work required to vary a control parameter and trace configuration entropy and internal energy, according with the first law of thermodynamics. The thermodynamic work is defined via Fisher information and thus can be computed solely based on probability distributions estimated from available data. Finally, we introduce the concept of *thermodynamic efficiency of urban transformation* as the ratio of the order gained during a change to the required work, and demonstrate that it is maximized at criticality for our case study.

2. Material and methods

2.1. Overview of the model

In our model, the population commutes between home and work place. The number of people commuting between employment areas i and residence areas j is given by the *travel-to-work* matrix T_{ij} . Commuting trips have an associated cost C_{ij} , e.g. travelling expenses, time or distance. C_{ij} represents the structure of the transport network, which may include the roads as well as different types of public transport. Employment areas are characterized by the average *income* I_i earned by the

employees that, in combination with the travel-to-work matrix, provides the flow of income $Y_{ij} = T_{ij}I_i$. Residence areas are instead characterized by the average rent R_j , and the amount of services S_j (the data used in modelling Greater Sydney is described in electronic supplementary material, §1).

We develop a BLV model for the predicted income flow \mathcal{Y}_{ij} in contrast with the actual income flow Y_{ij} obtained from the Census. The number of jobs available in each employment area is assumed to remain fixed, and therefore the income flowing out of each area is also fixed: $Y_i^{\text{out}} = \sum_j Y_{ij}$. On the contrary, the population is allowed to redistribute among the suburbs. The services S_j and the population $P_j = \sum_i (\mathcal{Y}_{ij}/I_i)$ determine the attractiveness A_j of a suburb, which defines people's preference to live in, and therefore bring their income to j . When deciding where to settle, people consider the utility of living in attractive suburbs as well as the cost of commuting to work. In our model, this trade-off is controlled by two parameters, α and γ , which define, respectively, how much value is attributed to suburbs with respect to their attractiveness (social disposition) and how much discomfort is attributed to commuting trips with respect to costs (travel impedance).

The model further allows the urban services S_j (and therefore the attractiveness A_j) to evolve, with these dynamics being slower than the resettling of people. When \mathcal{Y}_{ij} units of income are moved from employment areas i to residence areas j , part of it is spent on the rent R_j while the remainder can be spent on the services in j . Lotka–Volterra dynamics make suburbs compete for the services: if the income that can be spent in a suburb is higher than the cost of running the services in that suburb, these services will grow, otherwise they will decrease. Every time S_j is updated, \mathcal{Y}_{ij} is recomputed using the maximum entropy principle, until an equilibrium is reached such that the income spent on services matches their running cost. This results in a converging sequence of income flow matrices from an initial \mathcal{Y}_{ij}^0 to a final \mathcal{Y}_{ij}^* .

2.2. The Boltzmann component

The Boltzmann component of the model, informed by the maximum entropy principle, determines the least biased flow-of-income matrix \mathcal{Y}_{ij} which satisfies the constraints on the income that employment areas can produce, the attractiveness of the residence areas and the cost of travelling. Such flow of income is the one that maximizes the entropy

$$H(\mathcal{Y}_{ij}) = - \sum_i \sum_j \mathcal{Y}_{ij} \log \mathcal{Y}_{ij}, \quad (2.1)$$

for normalized \mathcal{Y}_{ij} , subject to the constraints:

$$\sum_j \mathcal{Y}_{ij} = Y_i^{\text{out}}, \quad (2.2)$$

$$\sum_i \sum_j \mathcal{Y}_{ij} A_j = A^{\text{tot}} \quad (2.3)$$

and

$$\sum_i \sum_j \mathcal{Y}_{ij} C_{ij} = C^{\text{tot}}. \quad (2.4)$$

The constraint in (2.2) fix the total income flowing out of the employment area i and towards all residence areas j . The constraint in (2.3) sets the total utility A^{tot} that people obtain by living in areas j with attractiveness A_j . Our assumption is that people prefer to live in areas that are more populated, unless the population exceeds a saturation limit. We also assume that people prefer areas where more services are available. Therefore, we define the attractiveness of a residence area as $A_j = \log(f(P_j) S_j)$, where $f(P_j)$ is a function that assigns a score based on the population. The population score $f(P_j)$ increases linearly with the population P_j , until it reaches a point of saturation P^* , after which additional population makes the score decrease at a certain rate ω , i.e. $f(P_j) = P_j$ if $P_j \leq P^*$ and $f(P_j) = P_j - \omega(P_j - P^*)$ if $P_j > P^*$. We set $\omega = 1.1$ and estimated the point of saturation at 11 000 based on the observation that only four suburbs in Greater Sydney, which are much more populated than the others, have a population that exceeds this value. Finally, constraint (2.4) sets the total cost C^{tot} of commuting between employment and residence areas.

The maximum entropy solution to this problem is

$$\mathcal{Y}_{ij} = \frac{Y_i^* e^{\alpha A_j - \gamma C_{ij}}}{Z_i}, \quad (2.5)$$

where $Z_i = \sum_j e^{\alpha A_j - \gamma C_{ij}}$ are balancing factors. The parameters α and γ are the Lagrangian multipliers corresponding to the constraints in (2.3) and (2.4), and representing social disposition and impedance to travel, respectively.

We calibrate our model by identifying the optimal values $\hat{\alpha}$ and $\hat{\gamma}$ that agree with the initial output \mathcal{Y}_{ij}^0 best matching the actual flow of income Y_{ij} given by Census (see electronic supplementary material, §2). The evolution of the services is then modelled yielding a prediction \mathcal{Y}_{ij}^* within Greater Sydney.

2.3. The Lotka–Volterra component

The Lotka–Volterra component of the model is given by the following dynamics for the services S_j over time t :

$$\frac{dS_j}{dt} = \epsilon(\mathcal{Y}_j^{\text{in}} - R_j P_j - K S_j), \quad (2.6)$$

where $\mathcal{Y}_j^{\text{in}} = \sum_i \mathcal{Y}_{ij}$ is the total income flowing into the suburb j from all employment areas i , ϵ defines the size of the changes and K is a conversion factor such that $K S_j$ is the cost of running services S_j . According to (2.6), if the remaining income $\mathcal{Y}_j^{\text{in}} - R_j P_j$ (analogous to discretionary income, which also subtracts taxes) flowing into the suburb is sufficient to compensate for the running costs of the services $K S_j$, then the services will grow, otherwise they will decrease. Since the attractiveness A_j is defined in terms of the services S_j , the former quantity also evolves.

2.4. Fisher information and thermodynamic efficiency

Following a recently established relationship [34], the rate of change of the thermodynamic work can be determined using the Fisher information (see electronic supplementary material, §3):

$$\frac{d\langle \beta W_{\text{gen}} \rangle}{d\alpha} = - \int_{\alpha_0}^{\alpha} F(\alpha') d\alpha' + c(\alpha^0), \quad (2.7)$$

where the Fisher information was calculated over the parameter α (fixing the parameter γ) as

$$F(\alpha) \equiv \sum_i \sum_j \mathcal{Y}_{ij}^* \left(\frac{d \log \mathcal{Y}_{ij}^*}{d\alpha} \right)^2 = \sum_i \sum_j \frac{1}{\mathcal{Y}_{ij}^*} \left(\frac{d \mathcal{Y}_{ij}^*}{d\alpha} \right)^2, \quad (2.8)$$

for the maximum entropy solution \mathcal{Y}_{ij}^* .

Finally, we define the thermodynamic efficiency of urban transformation, for a given value of α , as the reduction of entropy from the expenditure of work:

$$\eta \equiv \frac{-dH(\mathcal{Y}_{ij})/d\alpha}{d\langle \beta W_{\text{gen}} \rangle/d\alpha}. \quad (2.9)$$

This quantity corresponds to a change $d\alpha$ and hence relates to a *transformation*. This approach is motivated by the notion of thermodynamic efficiency of computation [34].

3. Results

3.1. Abrupt urban transformations

We explore the model predictions \mathcal{Y}_{ij}^* over a range of values of the control parameters around their optimal values $\hat{\alpha}$ and $\hat{\gamma}$. We then compute the entropy $H(\mathcal{Y}_{ij}^*)$ for the considered points within the phase diagram, tracing how the income distribution changes with respect to the control parameters. We observe that, while the entropy varies mostly linearly with respect to γ , it changes much more abruptly with the changes in α (see electronic supplementary material, §4), indicating a phase transition.

However, in order to rigorously localize the abrupt change in the dynamics of income flow with respect to α , we fix γ at the optimal value $\hat{\gamma}$ and compute the Fisher information over the phase space of α . The result is shown in figure 1, which shows that the Fisher information peaks at $\hat{\alpha} = 0.51$. This indicates that there is indeed a second-order phase transition in the space of α , the critical point $\hat{\alpha}$ of which is identified by the maximum value of the Fisher information, in line with the approach

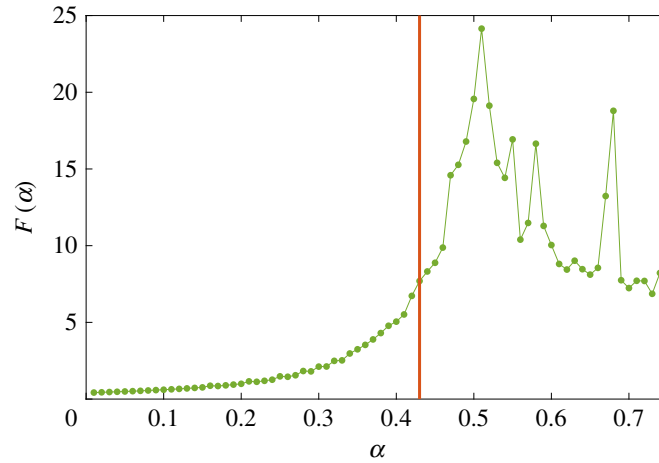


Figure 1. Fisher information over α , ranging from 0.001 to 0.751 with steps of 0.01, with $\hat{\gamma} = 0.15$ (best match). The horizontal axis represents the values of α , while the vertical axis represents the Fisher information of the flow or income \mathcal{Y}_{ij}^* . The red vertical bar indicates the value $\hat{\alpha} = 0.43$ for which \mathcal{Y}_{ij}^0 best matches the Sydney-2011 Census data, while the peak is at $\tilde{\alpha} = 0.51$. Note the second local peak at $\alpha = 0.68$.

established in [31,34,35]. Figure 1 also shows $\hat{\alpha} = 0.43$ that best matches Sydney-2011 Census data, which is lower than the critical value $\tilde{\alpha}$ but nevertheless is in the proximity of the phase transition, being located in the region where the Fisher information undergoes a rapid growth.

These results show that changes in the social disposition, away from its current value $\hat{\alpha}$, would significantly and abruptly change the flow distribution of income within Greater Sydney. This has an immediate effect on the spatial distribution of the population, driving an urban transformation from the *sprawling* phase to the *polycentric* phase. Figure 2 shows the predicted population of Greater Sydney at fixed $\hat{\gamma}$ and four different values of α : (a) a low value, far before the critical point, (b) the best match with the Sydney-2011 Census data $\hat{\alpha}$, (c) the critical point $\tilde{\alpha}$ and (d) a high value beyond the critical point. Since the average income I_i in the employment areas i does not change, the population of each suburb j is directly obtained from the flow of income \mathcal{Y}_{ij}^* predicted by the model.

For the low value of α (figure 2a), corresponding to the sprawling urban phase, the model shows a quite homogeneous distribution of the population, with the areas around the City of Sydney and other major urban areas being only slightly more populated than the other surrounding areas. As we move to $\hat{\alpha}$ (figure 2b), the population aggregates around the major urban areas, although the City of Sydney seems to be the only highly populated area. We note that this is the predicted population of Greater Sydney corresponding to the actual value of social disposition matching the Census data. At the critical point $\tilde{\alpha}$ (figure 2c), all the major urban areas become clearly more highly populated than the surrounding areas, and Greater Sydney starts to display a polycentric aggregation. Finally, this polycentric aggregation becomes more pronounced in the polycentric urban phase, represented by the high value of α (figure 2d): the areas of the City of Sydney, Parramatta, Penrith, Campbelltown and Gosford are clearly identifiable by a higher population compared with the surrounding.

Sydney-2011 profile, lying within the sprawling phase but near the phase transition, displays features of a polycentric metropolis, which accentuate beyond the critical point. However, the dynamics of the polycentric phase are not steady (cf. figure 1 for $\alpha > \tilde{\alpha}$), and so the transformations may suffer from tangible fluctuations and loss of predictability in social dynamics. In fact, a secondary transformation is captured by the secondary local peak of the Fisher information, around $\alpha = 0.68$ (figure 1) and corresponding to notable population decline in the suburbs north of Gosford (see electronic supplementary material, Movie S1). ‘Double percolation’ phase transitions have been observed in clustered complex networks with spatio-temporal dynamics [36], and multiple peaks detected by the Fisher information may relate to this phenomenon, given the clustered connectivity of urban aggregations.

3.2. Deepening the thermodynamic analogy

An important consideration in making a thermodynamic analogy is a choice of the protocol according to which the control parameters are varied, so that the corresponding changes in the required work, energy

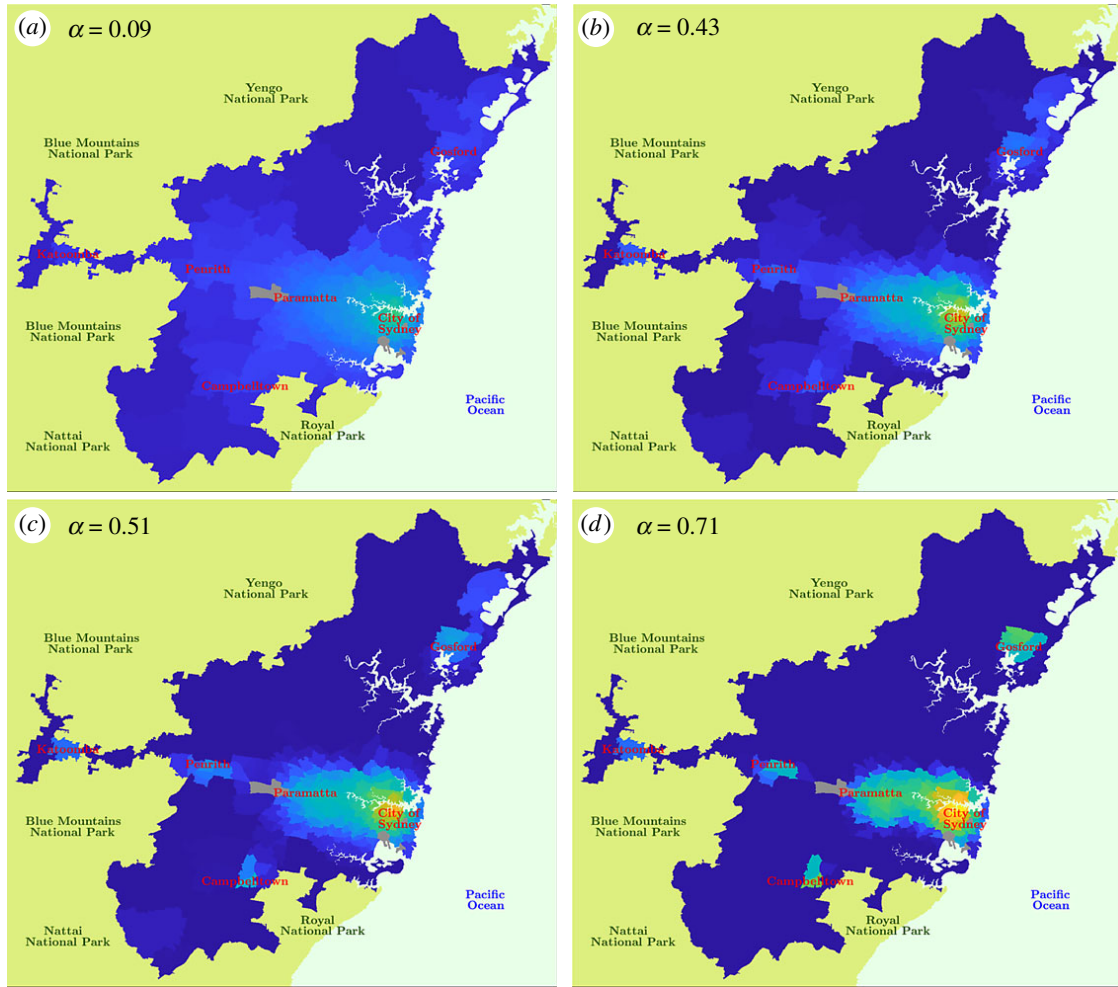


Figure 2. The predicted population of Greater Sydney. The region is partitioned into the 270 residence areas, which are coloured based on their population. The grey areas represent national reserves and parks, Kingsford Smith airport and Port Botany, which are not considered as residence areas. (a) Predicted population with $\alpha = 0.09$, corresponding to the sprawling urban phase. (b) Predicted population with $\hat{\alpha} = 0.43$, corresponding to the best match with Sydney-2011 Census data. (c) Predicted population with $\tilde{\alpha} = 0.51$, corresponding to the critical regime. (d) Predicted population with $\alpha = 0.71$, corresponding to the polycentric urban phase.

and configuration entropy, as well as symmetry breaking [37], can be traced. Specifically, we consider a quasi-static protocol varying α , at the expenditure of some required work, and driving changes from the sprawling urban phase to the polycentric phase, across the phase transition. For a quasi-static protocol, the required work is minimal, i.e. the work matches the free energy of the system.

It has been recently shown that for quasi-static processes the second derivative of the generalized work W_{gen} with respect to a control parameter is proportional to the negative of the Fisher information [34]. We refer to generalized work in the sense of Jaynes [24] (for more details about generalized quantities and their relationship with the Fisher information, see electronic supplementary material, §3). Given this relationship, we obtain the rate of change of the work with respect to α by numerically integrating the negative of the Fisher information in figure 1. The result is shown in figure 3a, demonstrating that the rate of change of the work decreases with α , with this change becoming more pronounced in the proximity of the optimal value $\hat{\alpha}$, being steepest around the critical point $\tilde{\alpha}$.

Figure 3a also shows the rate of change of the internal energy of the system U_{gen} . This quantity is obtained from the rates of change of the work W_{gen} and the configuration entropy $H(\mathcal{V}_{ij})$ —according to the first law of thermodynamics (in the case of quasi-static processes) a change in the internal energy corresponds to the sum of the changes in entropy and work: $\Delta\langle U_{\text{gen}} \rangle = \Delta\langle W_{\text{gen}} \rangle + \Delta H(\mathcal{V}_{ij})$, where the angle brackets represent average values over the ensemble. The rate of change of the internal energy decreases with α similar to the rate of change of the work. The difference between the two rates of change (i.e. the rate of change of the entropy) is larger around the critical point, when the flow of income exhibits a tendency towards the polycentric phase.

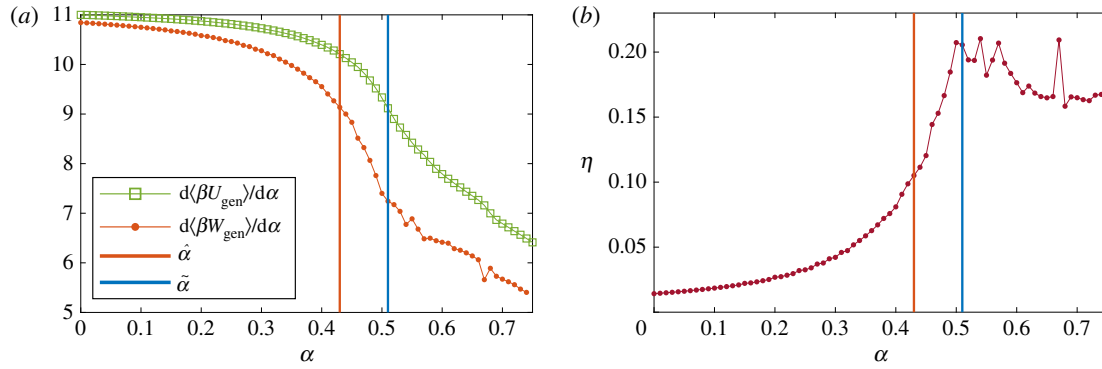


Figure 3. (a) The rates of change of the work W_{gen} (green line) and the internal energy U_{gen} (red line). (b) The thermodynamic efficiency of urban transformation, η , defined as the order-to-work ratio. In both graphs the vertical lines represent the values of the control parameter $\hat{\alpha}$ (red line), which best matches the Sydney-2011 Census data, and $\tilde{\alpha}$ (blue line), the critical value.

The thermodynamic efficiency of urban transformations for Greater Sydney, η , is shown in figure 3b. It can be seen that η is very low in the sprawling phase, increases towards the phase transition and then tends to slowly decrease, while also exhibiting the secondary local peak. This ratio is in a mid-range for the value $\hat{\alpha}$ corresponding to Sydney-2011 Census data. It is also evident that the social disposition estimated from the Sydney-2011 data characterizes the sprawling phase, distinct from the polycentric phase.

4. Discussion

The transition of cities between different patterns of urban settlement (dispersed, monocentric, polycentric, etc.) has become a central problem in urban planning. In this study we investigated the urban dynamics from a statistical mechanical viewpoint, deriving a thermodynamic description and applying it to a case study of Greater Sydney. This approach complements the maximum entropy principle with dynamics of evolving urban structures at different time scales, identifies phase transitions using Fisher information and quantifies the thermodynamic efficiency of urban transformations.

The model has been calibrated to Census data and geospatial datasets and exhibits a clear phase transition between a dispersed configuration, in which the population settles homogeneously within Greater Sydney, and a polycentric configuration, in which the population aggregates in a few highly populated urban clusters. Two salient quantities are represented by the attractiveness of suburbs, in terms of services available to the population, and the commuting costs. The phase transition was shown to be induced by the control parameter accounting for social disposition—a factor balancing the suburbs' attractiveness—rather than the parameter tracking travel impedance.

A recent plan by the Greater Sydney Commission [38] envisaged a tripartite Greater Sydney region, with a western parkland city, a central river city around greater Parramatta and an eastern harbour city. Under the assumptions used in this study, such a tripartite arrangement is possible only under a narrow set of constraints and lies in the polycentric urban phase, separated from the current sprawling phase by a phase transition. Thus, a major urban transformation is likely to pass through a critical regime with its inherent fluctuations and loss of predictability in social dynamics. Nevertheless, a set of policies informed by a quantitative approach may steer this transformation exploiting the resultant gain in efficiency. A wide class of other urban scenarios may also be considered within the proposed approach, in which the concise thermodynamic descriptions are derived purely based on probability distributions estimated from available data.

Data accessibility. This study uses the 2011 Australian Census data on population demographics and mobility in calibrating a model of urban dynamics. The Australia's 16th Census of Population and Housing was conducted on 9 August 2011, and the data from the 2011 Census are available through the following page: <http://abs.gov.au/websitedbs/censushome.nsf/home/historicaldata2011>. The Australia Bureau of Statistics also provides the geospatial data of Greater Sydney, which is available through the following pages: <http://www.abs.gov.au/AUSSTATS/abs@.nsf/DetailsPage/1270.0.55.001July%202011> and <http://www.abs.gov.au/AUSSTATS/abs@.nsf/DetailsPage/8000.0August%202011?OpenDocument>.

Authors' contributions. E.C. and M.P. designed research and analysed data. E.C. performed the computational modelling. All authors analysed the data and wrote the paper.

Competing interests. We declare we have no competing interests.

Funding. All authors were supported by The University of Sydney's DVC Research Strategic Research Excellence Initiative (SREI-2020) project, 'CRISIS: Crisis Response in Interdependent Social-Infrastructure Systems' (grant no. IRMA 194163). E.C. was supported by the University of Sydney's 'Post-graduate Scholarship in the Field of Complex Systems' from Faculty of Engineering & IT and by a CSIRO top-up scholarship. M.P. was supported by the ARC Discovery project 'Australian housing market risks: simulation, modelling and analysis' (grant no. DP170102927).

Acknowledgements. Sydney Informatics Hub at the University of Sydney provided access to HPC computational resources that have contributed to the research results reported within the paper.

References

- Batty M. 2008 The size, scale, and shape of cities. *Science* **319**, 769–771. (doi:10.1126/science.1151419)
- Batty M. 2013 *The new science of cities*. Cambridge, MA: MIT Press.
- William A. 1964 *Location and land use, toward a general theory of land rent*. Cambridge, MA: Harvard University Press.
- Hartshorn TA, Muller PO. 1989 Suburban downtowns and the transformation of metropolitan Atlanta's business landscape. *Urban Geogr.* **10**, 375–395. (doi:10.2747/0272-3638.10.4.375)
- Scott AJ. 1990 The technopoles of Southern California. *Environ. Plann. A* **22**, 1575–1605. (doi:10.1068/a221575)
- Garreau J. 1991 Edge city: life on the new frontier. *American Demographics* **13**, 24–31.
- Lang RE. 2003 *Edgeless cities: exploring the elusive metropolis*. Washington, DC: Brookings Institution Press.
- Odland J. 1978 The conditions for multi-center cities. *Econ. Geogr.* **54**, 234–244. (doi:10.2307/142837)
- Fujita M, Ogawa H. 1982 Multiple equilibria and structural transition of non-monocentric urban configurations. *Regional Sci. Urban Econ.* **12**, 161–196. (doi:10.1016/0166-0462(82)90031-X)
- Richardson HW. 1995 Economies and diseconomies of agglomeration. In *Urban agglomeration and economic growth* (ed. H Giersch), pp. 123–155. Berlin, Germany: Springer.
- Louf R, Barthelemy M. 2013 Modeling the polycentric transition of cities. *Phys. Rev. Lett.* **111**, 198702. (doi:10.1103/PhysRevLett.111.198702)
- Bouchaud JP. 2013 Crises and collective socio-economic phenomena: simple models and challenges. *J. Stat. Phys.* **151**, 567–606. (doi:10.1007/s10955-012-0687-3)
- Griffith DA. 1981 Evaluating the transformation from a monocentric to a polycentric city. *Prof. Geogr.* **33**, 189–196. (doi:10.1111/j.0033-0124.1981.00189.x)
- Pfister N, Freestone R, Murphy P. 2000 Polycentricity or dispersion? Changes in center employment in metropolitan Sydney, 1981 to 1996. *Urban Geogr.* **21**, 428–442. (doi:10.2747/0272-3638.21.5.428)
- McMillen DP, Smith SC. 2003 The number of subcenters in large urban areas. *J. Urban Econ.* **53**, 321–338. (doi:10.1016/S0094-1190(03)00026-3)
- Lee B. 2007 'Edge' or 'edgeless' cities? Urban spatial structure in US metropolitan areas, 1980 to 2000. *J. Regional Sci.* **47**, 479–515. (doi:10.1111/j.1467-9787.2007.00517.x)
- Louail T, Lenormand M, M García Cantú, Picornell O, Herranz R, E Frias-Martinez, Ramasco JJ, Barthelemy M. 2015 Uncovering the spatial structure of mobility networks. *Nat. Commun.* **6**, 6007. (doi:10.1038/ncomms7007)
- Harris B, Wilson AG. 1978 Equilibrium values and dynamics of attractiveness terms in production-constrained spatial-interaction models. *Environ. Plann. A Econ. Space* **10**, 371–388. (doi:10.1068/a100371)
- Wilson A. 2008 Boltzmann, Lotka and Volterra and spatial structural evolution: an integrated methodology for some dynamical systems. *J. R. Soc. Interface* **5**, 865–871. (doi:10.1098/rsif.2007.1288)
- Wilson A, Dearden J. 2011 Phase transitions and path dependence in urban evolution. *J. Geogr. Syst.* **13**, 1–16. (doi:10.1007/s10109-010-0134-4)
- Osawa M, Akamatsu T, Takayama Y. 2017 Harris and Wilson (1978) model revisited: the spatial period-doubling cascade in an urban retail model. *J. Regional Sci.* **57**, 442–466. (doi:10.1111/jors.12306)
- Wilson AG. 1970 *Entropy in urban and regional modelling*. London, UK: Pion.
- Barthelemy M. 2016 *The structure and dynamics of cities*. Cambridge, UK: Cambridge University Press.
- Jaynes ET. 1957 Information theory and statistical mechanics. *Phys. Rev.* **106**, 620–630. (doi:10.1103/PhysRev.106.620)
- Bialek W, Cavagna A, Giardina I, Mora T, Silvestri E, Viale M, Walczak AM. 2012 Statistical mechanics for natural flocks of birds. *Proc. Natl Acad. Sci. USA* **109**, 4786–4791. (doi:10.1073/pnas.1118633109)
- G Tkačik, Mora T, Marre O, Amodei D, Palmer SE, Berry MJ, Bialek W. 2015 Thermodynamics and signatures of criticality in a network of neurons. *Proc. Natl Acad. Sci. USA* **112**, 11 508–11 513. (doi:10.1073/pnas.1514188112)
- Wilson A. 2009 The 'thermodynamics' of the city. In *Complexity and spatial networks: in search of simplicity* (eds A Reggiani, P Nijkamp), pp. 11–31. Berlin, Germany: Springer.
- Morphet R. 2013 Von Thunen's Legendre Transform. UCL Working Paper Series. See <https://www.ucl.ac.uk/bartlett/casa/case-studies/2013/jul/casa-working-paper-193>.
- Hernando A, Hernando R, Plastino A, Plastino AR. 2013 The workings of the maximum entropy principle in collective human behaviour. *J. R. Soc. Interface* **10**, 20120758. (doi:10.1098/rsif.2012.0758)
- Brody DC, Rivier N. 1995 Geometrical aspects of statistical mechanics. *Phys. Rev. E* **51**, 1006–1011. (doi:10.1103/PhysRevE.51.1006)
- Prokopenko M, Lizier JT, Obst O, Wang XR. 2011 Relating Fisher information to order parameters. *Phys. Rev. E* **84**, 041116. (doi:10.1103/PhysRevE.84.041116)
- Wang XR, Lizier JT, Prokopenko M. 2011 Fisher information at the edge of chaos in random Boolean networks. *Artif. Life* **17**, 315–329. (doi:10.1162/artl_a_00041)
- Prokopenko M, Einav I. 2015 Information thermodynamics of near-equilibrium computation. *Phys. Rev. E* **91**, 062143. (doi:10.1103/PhysRevE.91.062143)
- Crosato E, Spinney RE, Nigmatullin R, Lizier JT, Prokopenko M. 2018 Thermodynamics and computation during collective motion near criticality. *Phys. Rev. E* **97**, 012120. (doi:10.1103/PhysRevE.97.012120)
- Crooks GE. 2007 Measuring thermodynamic length. *Phys. Rev. Lett.* **99**, 100602. (doi:10.1103/PhysRevLett.99.100602)
- P Colomer-de-Simón, M Boguñá. 2014 Double percolation phase transition in clustered complex networks. *Phys. Rev. X* **4**, 041020. (doi:10.1103/PhysRevX.4.041020)
- Nikoghosyan G, Nigmatullin R, Plenio M. 2016 Universality in the dynamics of second-order phase transitions. *Phys. Rev. Lett.* **116**, 080601. (doi:10.1103/PhysRevLett.116.080601)
- The Greater Sydney Commission. 2017 *Our Greater Sydney 2056, a metropolis of three cities – connecting people*. Parramatta, Australia: Greater Sydney Commission.

On critical dynamics and thermodynamic efficiency of urban transformations

Supplemental Material

Emanuele Crosato^{1,2,*}, Ramil Nigmatullin¹, and Mikhail Prokopenko¹

¹Complex Systems Research Group and Centre for Complex Systems, Faculty of Engineering and IT,
The University of Sydney, Sydney, NSW 2006, Australia.

²CSIRO Data61, PO Box 76, Epping, NSW 1710, Australia.

*emanuele.crosato@sydney.edu.au

1 Greater Sydney and data sources

Greater Sydney is an urban area covering more than 12,000 square kilometres, delimited in all directions by either the Pacific Ocean or by the several surrounding national parks. It includes the City of Sydney as well as other urban agglomerations such as Parramatta, Penrith, Campbelltown and Gosford, for a total population of approximately 5 million (see Fig. S1(a)). According to 2011 Census data, the working population of Greater Sydney is 1.8 million. People daily commute between residence areas (or suburbs), where they live, and employment areas, where they work. The territory is partitioned into 270 residence areas and 2,156 employment areas. The data used for this study was provided by the Australian Bureau of Statistics. This includes the geospatial data of the areas of employment and residence, as well as the Census data for year 2011.

The employment areas are defined by the standard Destination Zone (DZN), which was designed by the New South Wales transport authority in order to spatially classify employment places, with the purpose of analysing commuting data and developing transport policies. The standard Statistical Area Level 2 (SA2), as defined by the Australian Statistical Geography Standard, was used for the residence areas. Geographical areas of level SA2 represent small communities that closely interact socially and economically. The population of these areas can vary from 3,000 to 25,000 individuals, with an average population of 10,000 individuals. A satellite map showing the DZN and SA2 areas partitioning Greater Sydney is provided Fig. S1(b-d).

The Census data for year 2011 was geographically classified by the Australian Bureau of Statistics in accordance with the both geographical areas DZN and SA2, and included the travel-to-work matrix T_{ij} , the average weekly income I_i and the average weekly rent R_j , for all DZN areas i and SA2 areas j . The Census data also included the amount of people who work in food retailing stores (including supermarkets, grocery stores, meat and fish stores, fruit and vegetables stores and liquor stores) that are located in specific SA2 areas. This data was utilised to estimate the amount of goods, matching the services S_j available in each residence area.

The cost of travelling C_{ij} was estimated as the Euclidean distance between the centres of the employment and residence areas. An alternative approach would be to calculate the time of travelling using, for example, Google Maps or OpenStreetMap data. However, in this case one needs to assemble a dataset with several Sydney transport networks (bus, rail and ferry), in addition to the road networks. While the latter data are publicly available (e.g., OpenStreetMaps), the former is not. Furthermore, the network connectivity needs to be augmented with price of the corresponding travels, as well as individual preferences reflected in travel pattern data (such as Opal cards). A simpler solution based on Euclidean distances appears to be less biased than a more sophisticated but incomplete approach.



Figure S1: (a) Satellite image of Greater Sydney (TerraColor imagery by Earthstar Geographics LLC). The satellite image showing the areas of residence (b), according with the standard SA2, and the areas of employment (c), according with the standard DZN. (d) A magnification showing the employment areas of the City of Sydney, including City Business District (CBD).

In summary, the constraints used in Eq. (2) and Eq. (3) of the Boltzmann component are produced from the Census data, given the definition of attractiveness $A_j = \log(f(P_j) S_j)$ in terms of the population P_j and the services S_j , while the constraint used in Eq. (4) is produced by the geospatial data.

2 Calibration of the model

The model was calibrated by identifying the optimal values $\hat{\alpha}$ and $\hat{\gamma}$ for which the output \mathcal{Y}_{ij}^0 best matches the actual flow of income Y_{ij} of Sydney-2011. The difference between actual and predicted flow of income was estimated as the sum of (the absolute values of) the differences between all values of the matrices \mathcal{Y}_{ij}^0 and Y_{ij} , which is $e(\mathcal{Y}_{ij}^0) = \sum_i \sum_j |\mathcal{Y}_{ij}^0 - Y_{ij}|$. The result is shown in Fig. S2.

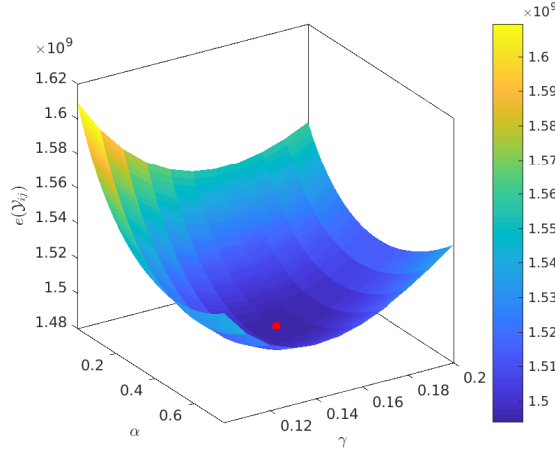


Figure S2: Calibration of the parameters α and γ . The horizontal axes represent values of α and γ within the considered ranges, while the vertical axis represents the difference between the income flow \mathcal{Y}_{ij}^0 produced by the model and the actual income flow Y_{ij} given by Sydney-2011 Census data, calculated as $e(\mathcal{Y}_{ij}^0) = \sum_i \sum_j |\mathcal{Y}_{ij}^0 - Y_{ij}|$. The red dot represents the optimal values $\hat{\alpha}$ and $\hat{\gamma}$.

3 Thermodynamic analysis

Let us consider the state functions $X_m(x)$ that describe a physical system over its configurations x . In a stationary state, the Gibbs measure defines the probability of the states of the system:

$$p(x|\theta) = \frac{1}{Z(\theta)} e^{-\beta H(x, \theta)} = \frac{1}{Z(\theta)} e^{-\sum_m \theta_m X_m(x)}, \quad (1)$$

where θ_m are thermodynamic variables, $\beta = 1/k_b T$ is the inverse temperature T (k_b is the Boltzmann constant), $H(x, \theta)$ is the Hamiltonian defining the total energy at state x , and $Z(\theta)$ is the partition function [1, 2]. The Gibbs free energy of such system is:

$$G(T, \theta_m) = U(S, \phi_m) - TS - \phi_m \theta_m, \quad (2)$$

where U is the internal energy of the system, S is the configuration entropy and ϕ_m is an order parameter. Let us also consider the generalised internal energy U_{gen} in the sense of Jaynes [3], such that

$$\langle \beta U_{gen} \rangle = U(S, \phi_m) - \phi_m \theta_m, \quad (3)$$

where the angle brackets represent average values over the ensemble. The generalised first law holds $\langle \beta U_{gen} \rangle = \langle \beta Q_{gen} \rangle + \langle \beta W_{gen} \rangle$, where Q_{gen} and W_{gen} are, respectively, the generalised heat and the generalised work.

The Fisher information [4] measures the amount of information that an observable random variable X carries about an unknown parameters $\theta = [\theta_1, \theta_2, \dots, \theta_M]^T$. If $p(x|\theta)$ is the probability of the realisation x of X given the parameters θ , the Fisher information matrix is defined as

$$F_{mn}(\theta) = E \left[\left(\frac{\partial \ln p(x|\theta)}{\partial \theta_m} \right) \left(\frac{\partial \ln p(x|\theta)}{\partial \theta_n} \right) \middle| \theta \right], \quad (4)$$

where the function $E(y)$ is the expected value of y . For a physical system described by the Gibbs measure in (1), the Fisher information has several physical interpretations, e.g., it is equivalent to the thermodynamic metric tensor $g_{mn}(\theta)$, is proportional to the second derivative of the free entropy $\psi = \ln Z = -\beta G$, and to the derivatives of the corresponding order parameters with respect to the collective variables [1, 5, 6, 2, 7]:

$$F_{mn}(\theta) = g_{mn}(\theta) = \frac{\partial^2 \psi}{\partial \theta_m \partial \theta_n} = \beta \frac{\partial \phi_m}{\partial \theta_n}. \quad (5)$$

Furthermore [8],

$$F(\theta) = \frac{d^2 S}{d\theta^2} - \frac{d^2 \langle \beta U_{gen} \rangle}{d\theta^2}. \quad (6)$$

Under a quasi-static protocol the total entropy production is zero, and therefore any change in the configuration entropy due to the driving process is matched by the flow of heat to the environment:

$$\frac{dS}{d\theta} = \frac{d \langle \beta Q_{gen} \rangle}{d\theta}. \quad (7)$$

Thus, combining (6) and (7) with the first law of thermodynamics yields another important result for the generalised work W_{gen} [8]:

$$F(\theta) = - \frac{d^2 \langle \beta W_{gen} \rangle}{d\theta^2}. \quad (8)$$

4 Entropy and a proxy of order parameter

A higher entropy indicates a more homogeneous distribution of the income to all suburbs, while a lower entropy indicates a less balanced distribution of the income biased towards one or few suburbs. We observe that the entropy decreases with both parameters α and γ (see Fig. S3). This behaviour is expected and has a clear interpretation. If the social disposition α is low, people have modest preference for attractive suburbs and thus settle (and move their income) more homogeneously within the region, while if α is high people tend to aggregate around the areas with the higher attractiveness. Similarly, if travel impedance γ is low people are less concerned about high travel costs, and therefore can settle at any distance from their work place, while if γ is high people prefer to live closer to their employment areas to incur lower commuting costs.

To formalise this intuition, one typically introduces and traces corresponding order parameters. This is however hindered by an incomplete statistical-mechanical description of the system, and we first illustrate a simpler approach which considers a proxy of an order parameter. Such a proxy characterises the equilibrium distribution of the services S_j , for different

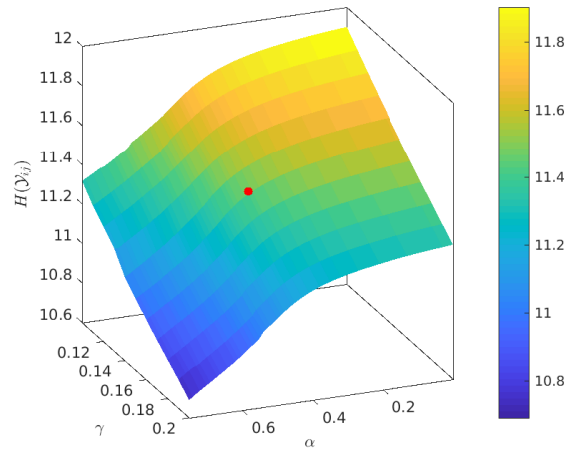


Figure S3: Entropy of \mathcal{Y}_{ij}^* after the services S_j have evolved to reach an equilibrium. The horizontal axes represent values of α and γ within the considered ranges, while the vertical axis represents the entropy $H(\mathcal{Y}_{ij})$ at corresponding values of α and γ . The red dot indicates the combination of $\hat{\alpha}$ and $\hat{\gamma}$ for which \mathcal{Y}_{ij}^0 best matches Sydney-2011 Census data.

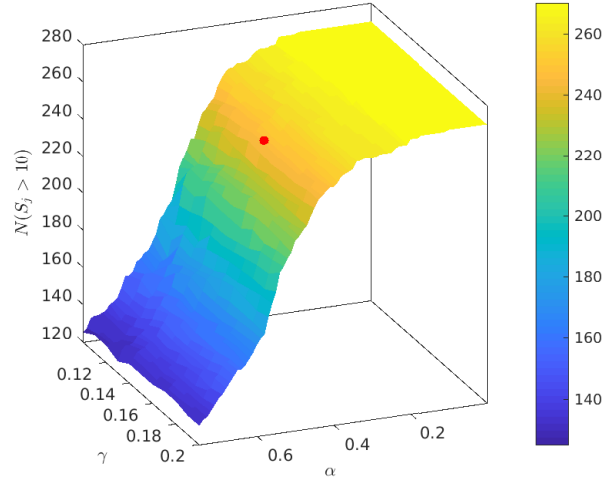


Figure S4: The number of suburbs $N(S_j > 10)$ with more than 10 units of services available, after the services S_j have evolved to reach an equilibrium. The horizontal axes represent values of α and γ within the considered ranges, while the vertical axis represents $N(S_j > 10)$ at corresponding values of α and γ . The red dot indicates the combination of $\hat{\alpha}$ and $\hat{\gamma}$ for which \mathcal{Y}_{ij}^0 best matches Sydney-2011 Census data.

values of α and γ , in terms of the number of suburbs in which the amount of available services exceeds a threshold, i.e., “services-abundant” suburbs. Fig. S5 shows the number of services-abundant suburbs, that is $N(S_j > 10)$, for different values of α and γ , after the urban evolution has converged. Again, as with the entropy dynamics, the variation of γ does not greatly affect the number of services-abundant suburbs. On the contrary, this number displays an abrupt change with respect to α : for low values of the social disposition all 270 suburbs are services-abundant, but as α increases the number of services-abundant suburbs reduces quickly past a specific value of α . At high values of social disposition approximately 120 residence areas remain services-abundant.

Fig. S3 and Fig. S5 also show the values $\hat{\alpha}$ and $\hat{\gamma}$ which best matches Sydney-2011 Census data (the red dot on either the entropy or the $N(S_j > 10)$ surfaces). This value is within a close proximity to the social disposition where the abrupt change is observed.

5 Animation of the phase transition

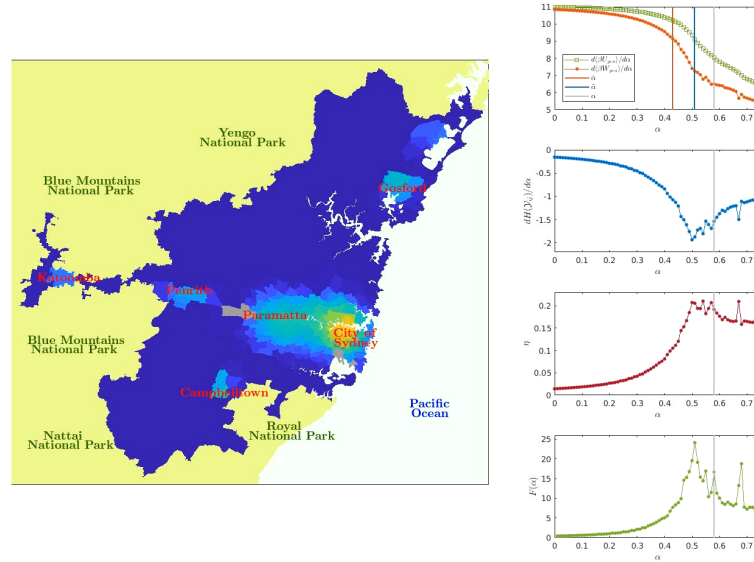


Figure S5: Animation of urban transformations within Greater Sydney. The movie shows the income flow within Greater Sydney as the social disposition increases. The movie also shows the rate of change of the generalised work, generalised internal energy and configuration entropy with respect to the social disposition, as well as the thermodynamic efficiency of urban transformations and the Fisher information.

The movie is available at: <https://www.dropbox.com/s/t7hz92i8mbv6c1a/Greater%20Sydney.avi?dl=0>.

References

- [1] Brody DC, Rivier N. Geometrical aspects of statistical mechanics. *Physical Review E*. 1995;51:1006–1011.
- [2] Crooks GE. Measuring Thermodynamic Length. *Physical Review Letters*. 2007;99:100602.
- [3] Jaynes ET. Information Theory and Statistical Mechanics. *Physical Review*. 1957;106:620–630.
- [4] Fisher RA. On the Mathematical Foundations of Theoretical Statistics. *Philosophical Transactions of the Royal Society of London A: Mathematical, Physical and Engineering Sciences*. 1922;222(594-604):309–368.
- [5] Brody DC, Ritz A. Information geometry of finite Ising models. *Journal of Geometry and Physics*. 2003;47(2):207–220.
- [6] Janke W, Johnston DA, Kenna R. Information geometry and phase transitions. *Physica A: Statistical Mechanics and its Applications*. 2004;336(1–2):181–186.
- [7] Prokopenko M, Lizier JT, Obst O, Wang XR. Relating Fisher information to order parameters. *Physical Review E*. 2011;84:041116.
- [8] Crosato E, Spinney RE, Nigmatullin R, Lizier JT, Prokopenko M. Thermodynamics and computation during collective motion near criticality. *Physical Review E*. 2018;97:012120.

Chapter 6

Entropy production during collective motion of active Brownian particles

Entropy production during collective motion of active Brownian particles undergoing phase separation and alignment

Emanuele Crosato,^{1,2,*} Mikhail Prokopenko,¹ and Richard E. Spinney¹

¹*Complex Systems Research Group and Centre for Complex Systems,
Faculty of Engineering and IT, The University of Sydney, Sydney, NSW 2006, Australia.*

²*CSIRO Data61, PO Box 76, Epping, NSW 1710, Australia.*

(Dated: December 12, 2018)

We study the entropy production rate in a system of active Brownian particles interacting via volume exclusion and alignment, described by under-damped dynamics. Using numerical simulations, we build the phase diagram of the system and identify three kinetic phases that involve mobility-induced phase separation as well as global orientational order. Crucially, the phase transitions are reflected by changes in the expected entropy production rates at (non-equilibrium) steady-state. Additionally, the expected value of entropy production rate associated with individual particles is contrasted with their orientation and position within the group, revealing distinct entropy production signatures for different active matter states, e.g., ‘solid’, ‘gas’ and oriented ‘flocks’. Importantly, interpreting the particles’ heading as an odd or even variable under time-reversal symmetry is shown to yield different entropy production rates, highlighting the importance of such interpretation in the thermodynamic treatment of active matter.

I. INTRODUCTION

Active matter consists of self-propelled particles that can extract energy from the environment and transform it into directed motion [1–3]. Examples encompass a wide range of systems, including self-catalytic colloidal suspensions [4], swimming bacteria [5, 6], migrating cells [7] and even animal groups [8–10]. Self-propulsion, in combination with volume exclusion and possibly other interactions among particles, can give rise to non-trivial collective motion dynamics, such as gathering, swarming and swirling [11]. These complex behaviours are not achievable by matter at thermal equilibrium. Active matter, in fact, is characteristically out-of-equilibrium: energy, supplied to each particle to fuel the self-propulsion, is locally dissipated breaking detailed balance [12].

Minimal models of active particles have been developed. The main two ones that have been proposed are Active Brownian Particles (ABPs) [13], whose self-propulsion has fixed magnitude and direction that is subject to direction rotational noise, and Active Ornstein-Uhlenbeck Particles (AOUPs), whose self-propulsion is modelled as coloured noise [14]. Dynamical models for the collective motion of many self-propelled particles have also been developed, the simplest of which involve two-dimensional self-propelled discs that only interact via volume exclusion. Despite their simplicity, these systems can produce rather rich behaviour [15] that resemble real active matter.

Dynamical models allow the systematic investigation of kinetic phase transitions, a well-known phenomenon in active matter [1–3]. Systems of both ABPs and AOUPs have been shown to exhibit mobility-induced phase separation (MIPS), i.e., the system splits between dense ‘liquid’ or ‘solid’ areas and sparse ‘gas’ areas, for sufficiently high values of density and/or self-propulsion strength [14–19]. Recently, a model of ABPs with alignment interactions by Martín-Gómez

et al. [20] has been shown to exhibit a phase transition between disordered and oriented collective motion (or ‘flocking’), along MIPS.

While building phase diagrams for the behaviour of active matter is a common procedure [18–20], not much research has been done to describe the non-equilibrium character of the kinetic phases thermodynamically. Only recently, Fodor et al. [14] have investigated the entropy production in a system of AOUPs, with no alignment interactions, described by over-damped dynamics. The authors found that, even for high self-propulsion strength resulting in MIPS, the particles’ dynamics still respects detailed balance, suggesting that microscopic non-equilibrium does not survive coarse-graining and that, in the large scale, MIPS can essentially be understood as equilibrium phase separation. A later study by Mandal et al. [21], however, demonstrated that nonzero and positive entropy production is found when a different definition of entropy production along a trajectory is used [22, 23].

Crucially, Shankar et al. [24] have demonstrated that ‘hidden’ components of the entropy production can be detected along the trajectory of either an ABP or an AOUP, when under-damped dynamics are used to describe the particle. The authors also highlighted the importance of the interpretation of the self-propulsion force as an even or an odd variable under time-reversal symmetry (TRS), which leads to different expressions for the entropy production rate. However, the study only considered free, non-interacting particles.

In this study, we investigate the entropy production rates in a system of ABPs that interact via volume exclusion as well as alignment (similarly to [20]). We derive the analytical expressions of the entropy production rates for both an over-damped and an under-damped models, and for either parity of the particles’ heading under TRS. For the under-damped dynamics, the entropy production is shown to depend on the relationship between the timescales of translational and rotational motion, which is essential for MIPS.

The under-damped model is simulated numerically using a large number of particles to produce the phase diagram of the system at non-equilibrium steady-state. The system is then

* emanuele.crosato@sydney.edu.au

characterised over the different phases in terms of its expected steady-state entropy production rate, calculated using the derived analytical expressions. We show that different results are obtained based on the interpretation of the particles' heading under TRS. Additionally, we map the expected entropy production rate associated to individual particles with their heading and position within the group, and identify entropy production signatures of different active matter states.

II. DYNAMICAL MODEL OF ABPs WITH ALIGNMENT

Let us consider a system of N ABPs of mass m , moment of inertia I , radius R and translational and rotational mobility coefficient γ and γ_R . Let us also denote the position of each particle a as \mathbf{r}_a , and its heading as θ_a . The self-propulsion force is modelled as a function of the particles' heading, given a propulsion speed v_0 :

$$\mathcal{P}(\theta_a) = m\gamma v_0 \hat{e}(\theta_a), \quad (1)$$

where and $\hat{e}(\theta_a) = [\cos(\theta_a), \sin(\theta_a)]$. The particles interact via excluded volume and torque interactions that make close particles align. Excluded volume interactions can be modelled using a potential function $U(\mathbf{r})$, e.g., a truncated and shifted Lennard-Jones potential:

$$U(\mathbf{r})_a = \begin{cases} \sum_{b \neq a} \epsilon \left[\left(\frac{2R}{r_{ab}} \right)^{12} - 2 \left(\frac{2R}{r_{ab}} \right)^6 \right] + \epsilon & r_{ab} \leq R \\ 0 & r_{ab} > R \end{cases} \quad (2)$$

where ϵ is the depth of the potential well and r_{ab} is the distance between particles a and b . Alignment interactions are instead modelled by the following torque function:

$$\mathcal{T}(\mathbf{r}, \theta)_a = - \sum_{b \neq a} K \sin(\theta_a - \theta_b), \quad (3)$$

where K is the coupling strength.

The over-damped dynamics of the system are described by the following stochastic differential equations (SDEs):

$$d\mathbf{r}_a^j = \frac{1}{m\gamma} \mathcal{F}(\mathbf{r}, \theta)_a^j dt + \sqrt{\frac{2T}{m\gamma}} dW_{\mathbf{r}_a^j}, \quad (4)$$

$$d\theta_a = \frac{1}{I\gamma_R} \mathcal{T}(\mathbf{r}, \theta)_a dt + \sqrt{\frac{2T}{I\gamma_R}} dW_{\theta_a}, \quad (5)$$

where $\mathcal{F}(\mathbf{r}, \theta)_a^j = \mathcal{P}(\theta_a)^j - \partial_{\mathbf{r}_a^j} U(\mathbf{r})_a$ includes both self-propulsion and excluded volume interactions, i is the spatial dimension, T is the temperature of the system and $W_{\mathbf{r}_a^j}$ and W_{θ_a} are uncorrelated Wiener processes, such that $\langle dW_{\mathbf{r}_a^j} dW_{\theta_a} \rangle = 0$, $\langle dW_{\theta_a} dW_{\theta_b} \rangle = \delta_{ab} dt$ and $\langle dW_{\mathbf{r}_a^j} dW_{\mathbf{r}_b^k} \rangle = \delta_{jk} \delta_{ab} dt$.

When under-damped dynamics are used to describe the system, we also need to introduce particles' momenta through the translational velocity \mathbf{v}_a and the rotational velocity ω_a . The

system is then described by the SDEs:

$$d\mathbf{r}_a^j = \mathbf{v}_a^j dt, \quad (6)$$

$$d\mathbf{v}_a^j = -\gamma \mathbf{v}_a^j dt + \frac{1}{m} \mathcal{F}(\mathbf{r}, \theta)_a^j dt + \sqrt{\frac{2T\gamma}{m}} dW_{\mathbf{v}_a^j}, \quad (7)$$

$$d\theta_a = \omega_a dt, \quad (8)$$

$$d\omega_a = -\gamma_R \omega_a dt + \frac{1}{I} \mathcal{T}(\mathbf{r}, \theta)_a dt + \sqrt{\frac{2T\gamma_R}{I}} dW_{\omega_a}, \quad (9)$$

where $W_{\mathbf{v}_a^j}$ and W_{ω_a} are also independent Wiener processes.

In order to characterise the configurational change associated with MIPS we utilise the local (per particle) sixfold bond-orientational order:

$$|q_6(a)| = \left| \frac{1}{6} \sum_{b \in \mathcal{N}_a} e^{i6\alpha_{ab}} \right|, \quad (10)$$

where α_{ab} is the angle between $\mathbf{r}_a - \mathbf{r}_b$ and an arbitrary axis and \mathcal{N}_a are the closest 6 neighbouring particles of a . An order parameter for the phase separation is therefore provided by the average bond-orientational order $\langle |q_6(a)| \rangle$.

An alternative mean of characterising MIPS is through the local density distribution, where we expect bimodal behaviour in the case of phase separation. Hence, we can quantify phase separation using the bimodality coefficient. By considering the local density \mathcal{X}_d , defined as the empirical density within a radius d , the bimodality coefficient is introduced as

$$\zeta(\mathcal{X}_d) = \frac{\lambda(\mathcal{X}_d) + 1}{\kappa(\mathcal{X}_d)}, \quad (11)$$

where $\lambda(\mathcal{X}_d)$ and $\kappa(\mathcal{X}_d)$ are, respectively, the third and the fourth standardised moments of the particles' density (i.e., the skewness and the kurtosis). The value of $\zeta(\mathcal{X}_d)$ lies between 0 and 1, with larger values indicative of a bimodal distribution. As it quantifies the variation in local density, the bimodality coefficient is a more natural way to identify MIPS. However, the local bond-orientational order $q_6(a)$ allows the characterisation of single particles configurational order. It should also be noted that, at very high densities, when the local density distribution is dominated by a sharp peak at high values of \mathcal{X}_d , single statistics of the distribution are not best placed to determined the location of transition points, which can be inferred from the qualitative features of the distribution.

The alignment within the system is instead quantified as

$$\rho(\theta) = \langle 2 \cos^2(\theta_a - \bar{\theta}) - 1 \rangle, \quad (12)$$

where $\bar{\theta}$ is the mean heading across all particles.

III. ENTROPY PRODUCTION OF INTERACTING ABPs

In modern formalisms such as stochastic thermodynamics entropy production can be interpreted as a measure of dynamical irreversibility. Specifically, the total entropy production,

$\frac{d\langle\Delta S^{\text{tot}}\rangle}{dt}$	Over-damped	Under-damped
TRS-odd heading	$\sum_{a=1}^N \left(\frac{\langle \mathcal{T}^2(\theta, \mathbf{r})_a \rangle}{T I \gamma_R} + \frac{\langle \partial_{\theta_a} \mathcal{T}(\theta, \mathbf{r})_a \rangle}{I \gamma_R} + \sum_{j \in \{x, y\}} \frac{v_0}{T} \langle \partial_{\mathbf{r}_a^j} U(\mathbf{r})_a \hat{e}^j(\theta_a) \rangle \right)$	$\sum_{a=1}^N \left(\frac{\gamma_R}{T} (I \langle \omega_a^2 \rangle - T) + \sum_{j \in \{x, y\}} \frac{\gamma}{T} \left(\frac{m v_0^2}{2} - m \langle (\mathbf{v}_a^j)^2 \rangle + T \right) \right)$
TRS-even heading	$\sum_{a=1}^N \left(\frac{\langle \mathcal{T}^2(\theta, \mathbf{r})_a \rangle}{T I \gamma_R} + \frac{\langle \partial_{\theta_a} \mathcal{T}(\theta, \mathbf{r})_a \rangle}{I \gamma_R} + \frac{m \gamma v_0^2}{T} - \sum_{j \in \{x, y\}} \frac{v_0}{T} \langle \partial_{\mathbf{r}_a^j} U(\mathbf{r})_a \hat{e}^j(\theta_a) \rangle \right)$	$\sum_{a=1}^N \left(\frac{\gamma_R}{T} (I \langle \omega_a^2 \rangle - T) + \sum_{j \in \{x, y\}} \frac{\gamma}{T} (m \langle (\mathbf{v}_a^j)^2 \rangle - T) \right)$

TABLE I. Expected entropy production rate for a system of interacting ABPs at steady state, described by over-damped and under-damped dynamics, and for odd and even interpretation of the particles' heading under TRS.

generated over an interval $t \in [t_0, \tau]$, may be written for individual realisations $\vec{\Omega} = \{\Omega(t) | t \in [t_0, \tau]\}$ (where $k_B = 1$ and $\Omega = \{\mathbf{r}, \mathbf{v}, \theta, \omega\}$ is the total state of the system) [25, 26],

$$\Delta S^{\text{tot}} = \ln \frac{P[\vec{\Omega}]}{P^\dagger[\vec{\Omega}^\dagger]}. \quad (13)$$

Here, $\vec{\Omega}^\dagger = \{\Omega^\dagger(t) | t \in [t_0, \tau]\}$ and $\Omega^\dagger(t) = \epsilon \Omega(\tau + t_0 - t)$ where ϵ is a time reversal operation. For stationary, autonomous, and time symmetric dynamics $P^\dagger = P$. Consequently, the entropy production is equal to the log ratio of the likelihood of a given trajectory against its time reverse under the dynamics that describe the system. The total entropy production obeys an integral fluctuation theorem $\langle \exp[-\Delta S^{\text{tot}}] \rangle = 1$ and thus the strict inequality $\langle \Delta S^{\text{tot}} \rangle \geq 0$ holds by Jensen's inequality, characterising the second law.

The total entropy production consists of changes in the system entropy (equal to the Gibbs entropy over the dynamical variables) and entropy change in the environment (consisting of an exported heat scaled by the inverse environmental temperature). The form of such expressions depends on the nature of the time reversal operation ϵ since it determines the appropriate time reversed trajectory [27]. For SDEs, expressions for the total entropy production can be found exactly given knowledge of the probability density functions over the variables [22] and expressions for the environmental entropy production can be determined in terms of knowledge of the trajectories only. In the steady state, however, the expected change of Gibbs entropy vanishes and the mean medium entropy production allowing empirical calculation of expectations without the need for solving the associated Fokker-Planck equation.

For systems with multiple degrees of freedom, such as collections of ABPs, individual environmental entropy productions can be associated with subsets of the dynamical variables so long as the Wiener processes associated with them are uncorrelated, with the total environmental entropy production being equal to their sum. This property, often referred to as bipartite dynamics [28], allows us to characterise the mean envi-

ronmental entropy production of individual ABPs in the collective. Utilising the formalism in [22] we derive the entropy expected medium entropy production for individual ABPs in the Supplemental Material. The results depend crucially on the time reversal parity of the particle orientation. For the odd interpretation using the under-damped dynamics we find the expected entropy production rate in the medium for particle a to be

$$\frac{d\langle \Delta S_a^{\text{med}} | \Omega \rangle}{dt} = \frac{\gamma_R}{T} (I \langle \omega_a^2 \rangle - T) + \sum_{j \in \{x, y\}} \frac{\gamma}{T} \left(m \langle (\mathbf{v}_a^j - v_0 \hat{e}(\theta_a))^2 \rangle - T \right). \quad (14)$$

For the even interpretation of the particles' heading using the under-damped dynamics we instead have

$$\frac{d\langle \Delta S_a^{\text{med}} | \Omega \rangle}{dt} = \frac{\gamma_R}{T} (I \langle \omega_a^2 \rangle - T) + \sum_{j \in \{x, y\}} \frac{\gamma}{T} \left(m \langle (\mathbf{v}_a^j)^2 \rangle - T \right). \quad (15)$$

The expected medium entropy production of the total system, and thus the expected total entropy production in the steady state, is then equal to the sum of such contributions over all particles.

Using such expressions alongside simplifications that hold in the steady state we can calculate the mean steady state entropy production for the whole system of interacting ABPs described by both of the models in Sec. II. These are given in Table I for odd and even time reversal parity interpretations of the heading θ , in comparison to Shankar et al. [24] for non-interacting particles. Details can be found in Supplemental Material.

Clearly, the choice of such an interpretation and choice of model effects the implied thermodynamics. Perhaps most striking is the distinction between parity interpretations where the difference is most clear in the expressions for the under-

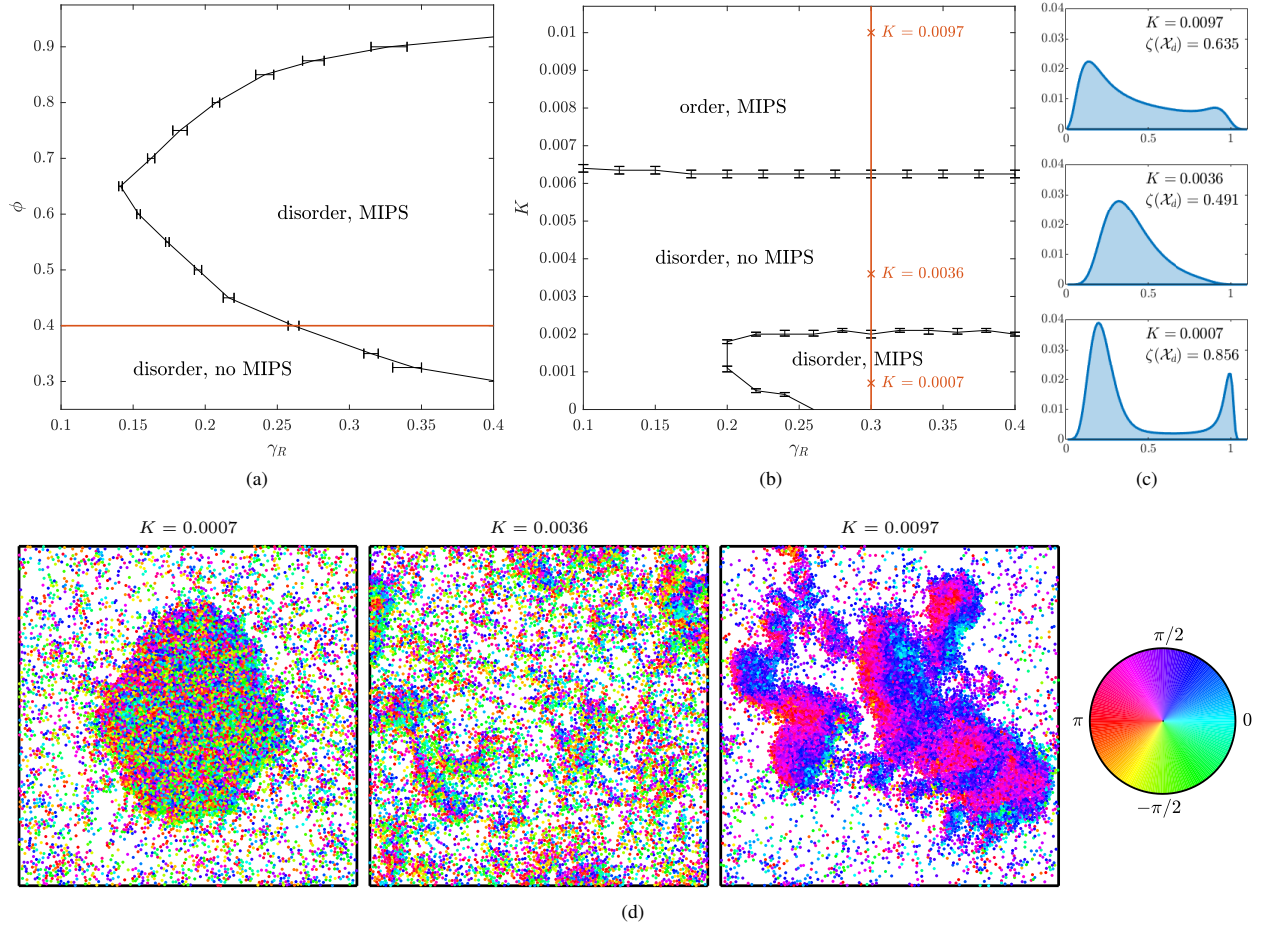


FIG. 1. Summary of the kinetic phases of the system. (a) shows the phase diagram of the behaviour of systems with respect to the density ϕ and the rotational mobility coefficient γ_R , when the torque coupling strength $K = 0$. (b) shows the phase diagram of the systems with respect to γ_R and K , at density $\phi = 0.4$. In both diagrams the error bars indicate the intervals within which the phase transition is observed to occur, based the simulations. The solid lines are approximations of the critical lines, given the error bars. The red lines represent sections of the phase diagrams over which Fig.2 provides the expected steady-state entropy production rate. A representative point for each of the three phases are marked with red crosses over the red line in (b). At each of these points, (c) shows the distribution of the local density \mathcal{X}_d (with $d = 4.5$) while (d) shows a typical configuration observed during the simulations (the colours represent the heading of the particles).

damped model. In this case, under the even parity interpretation, the entropy production is manifestly a measure of the deviation away from equipartition expected at thermodynamic equilibrium in both the translational and rotational degrees of freedom. In contrast, under the odd parity interpretation the entropy production arising from the translational variables is modified such that it quantifies deviation from an effective equipartition, relative to the instantaneous heading and typical speed due to the active forcing—seen most clearly in Eq. (14).

The above expressions are quite general. However, in the absence of alignment, external, and exclusion interactions, such that $\mathcal{T}(\theta, \mathbf{r}) = 0$ and $\partial_{\mathbf{r},j}U(\mathbf{r}) = 0$, the expressions can be seen to reduce to those derived for free ABPs in [24] for both the over-damped and under-damped models. The entropy productions in the over-damped model, particularly that under the odd parity interpretation illustrate the well-known prop-

erty that contributions to entropy production are lost through coarse graining procedures [29]. Such absent terms have been referred to as ‘anomalous’ [30] or ‘hidden’ and have been previously implicated in heat transfer where under-damped models are crucial in order to observe physically plausible entropy productions [22]. Notably, for the ABPs, the free particle contributions in the over-damped limit show no dependence on the the relationship between the timescales of translational and rotational motion, the primary mechanism for the appearance of MIPS. Therefore to ensure that both qualitative and quantitative features of the thermodynamics are appropriately captured, we proceed using the under-damped model.

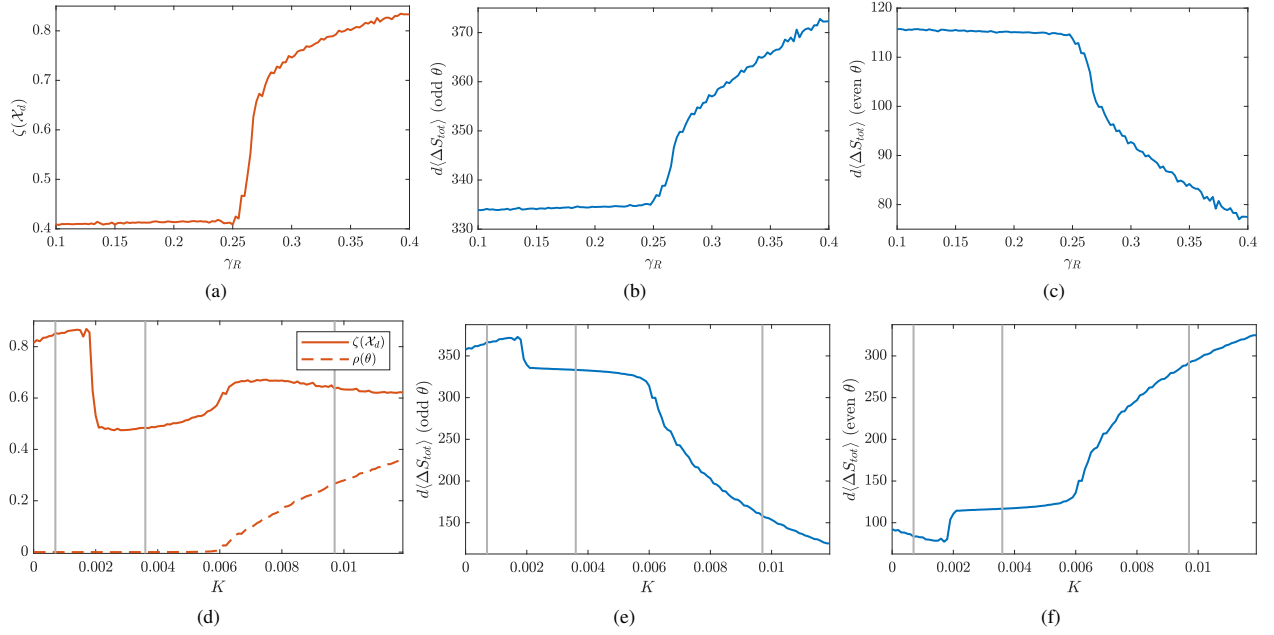


FIG. 2. Average steady-state entropy production over the tree kinetic phases. In (a), (b) and (c) the density ϕ the density $\phi = 0.4$ and the alignment strength $K = 0$. (a) shows the average bimodality coefficient $\zeta(\chi_d)$ (with $d = 4.5$) at steady-state, over the rotational mobility coefficient γ_R . (b) and (c) show the average entropy production rate at steady-state over γ_R with, respectively, odd and even interpretation of θ . In (d), (e) and (f) $\phi = 0.4$ and $\gamma_R = 0.3$. (d) shows the average $\zeta(\chi_d)$ and the average alignment coefficient $\rho(\theta_a)$ at steady-state, over the alignment strength K . (e) and (f) show the average entropy production rate at steady-state over K with, respectively, odd and even interpretation of θ . The three vertical grey lines in (d), (e) and (f) correspond to the points marked with red crosses in Fig. 1(b).

IV. ENTROPY PRODUCTION AND KINETIC PHASES

We used numerical simulations to explore the behaviour of the systems over the rotational mobility coefficient γ_R , the coupling strength of the alignment K and the particles density ϕ . The self-propulsion strength is typically controlled by varying the Péclet number $Pe \propto v_0/\sqrt{D_T D_R}$ [2], where $D_T = T/m\gamma$ and $D_R = T/I\gamma_R$. Since Pe depends on many variables that contribute to the entropy production, we choose to keep T , m and v_0 constant, and to vary the self-propulsion strength only through γ_R . In our simulations, we set $N = 10000$, $R = 0.5$, $v_0 = 3$, $m = I = \gamma = 1$ and $T = 0.02$, and we use periodic boundary conditions. We also use the truncated and shifted Lennard-Jones potential in Eq. (2) (with $\epsilon = 1$) for the excluded volume interactions, and the torque function in Eq. (3) for the alignment interactions.

When only excluded volume interactions are considered (i.e., $K = 0$) the system exhibits two phases: a phase in which particles distribute homogeneously across the space, and a phase characterised by MIPS (see Fig. 1(a)). For low values of the density, the transition to MIPS is triggered by high values of γ_R . Increasing the density results in lower and lower critical values of γ_R , however, this trend is inverted for densities higher than approximately $\phi = 0.65$. The phase diagram in Fig. 1(a) confirms the results in [18, 19].

When alignment interactions are considered, a third kinetic phase is possible, characterised by both orientational order

and MIPS (see Fig. 1(b)). At density $\phi = 0.4$, for example, low values of K do not have much effect on the system, which behaves similarly to when there are no alignment interactions, i.e., MIPS occurs for γ_R higher than a critical value. As K is increased to a point in which alignment interactions sufficiently reduce the average time required for the two particles to head away from each other after they collide, MIPS is interrupted. In this phase, the alignment torque is not strong enough to keep the particles aligned, however, for even higher values of K beyond a second critical threshold, the system starts to exhibit orientational order. In this phase the system also exhibit MIPS: as the particles gradually align their orientation they also get nearer, and then remain close as they flock towards a common average direction.

The distribution of the local density χ_d at three values of K corresponding to the three different phases (cf. red crosses in Fig. 1(b)) is shown Fig. 1(c). The phase with MIPS and no orientational order is characterised by highly bimodal distributions, while the phase with no MIPS is characterised by unimodal distributions. The phase with both MIPS and orientational order is also characterised by (typically less pronounced) bimodal distributions. A typical configuration of the system at the three values of K is shown in Fig. 1(d), where the colours represent the particles' heading.

We now proceed by showing the expected steady-state entropy production rate characterising the three kinetic phases. Firstly, we focus on the system in absence of alignment in-

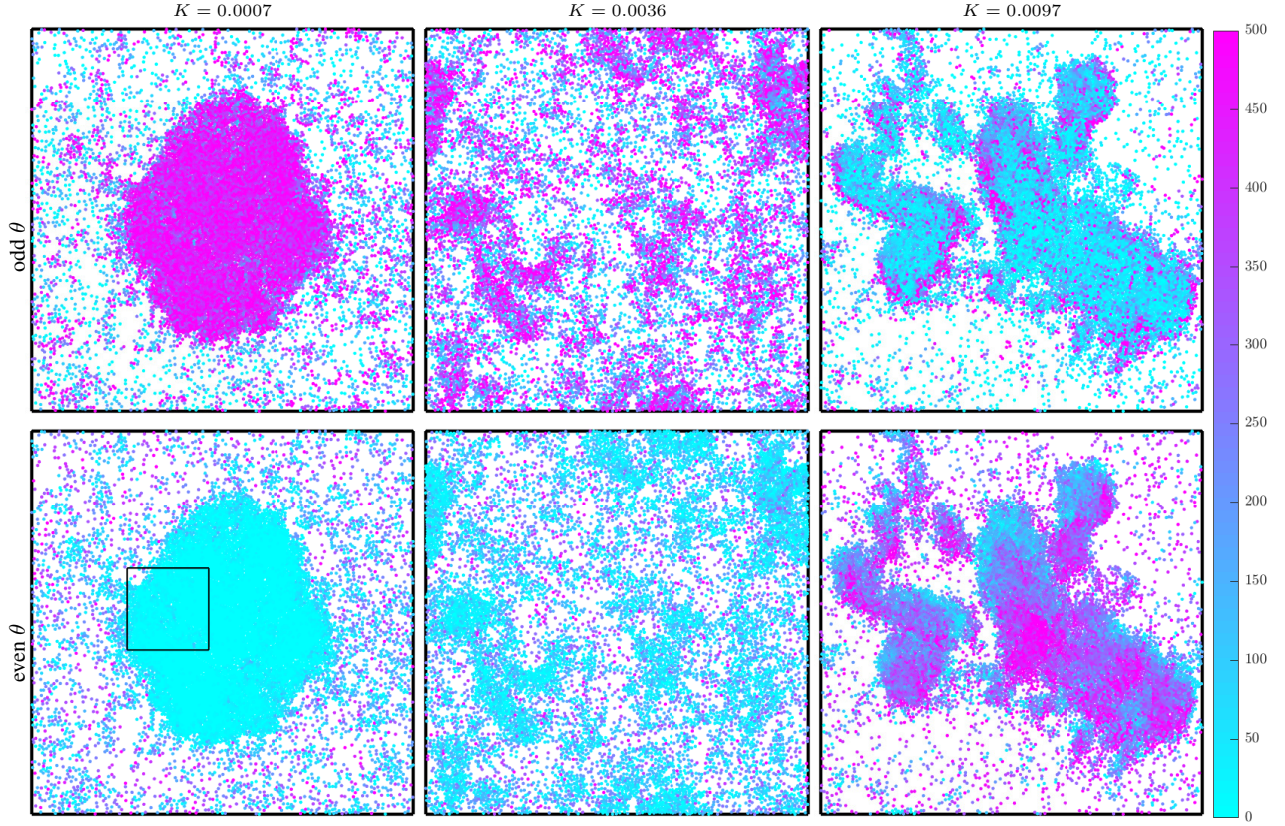


FIG. 3. Expected entropy production rate associated to individual particles. The three configurations previously shown in Fig. 1(d) are here presented colouring each particle based on its expected entropy production rate (cf. Eq. (14) and Eq. (15)), distinguishing between odd and even interpretations of the particles' heading under TRS. The particles in the box (see $K = 0.0007$, even θ) are shown in Fig 4 with higher resolution for the entropy production.

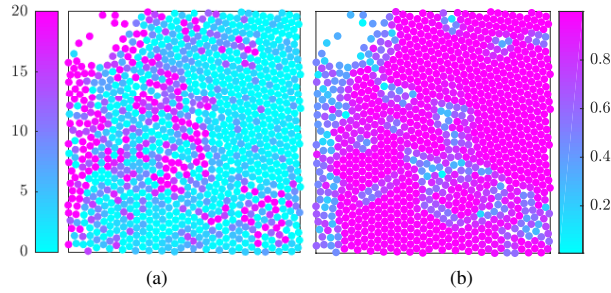


FIG. 4. Expected entropy production rate and defects in the solid structure. (a) magnifies the box in Fig. 3 ($K = 0.0007$, even θ), showing the entropy production of the particles using a different colour scale, which can capture small differences between low values of entropy production. (b) shows instead the local sixfold bond-orientational order $|q_6|$ associated to each individual particle, highlighting defects across the solid structure.

interactions, at density $\phi = 0.4$, as we vary rotational mobility coefficient γ_R (Fig. 2(a) to Fig. 2(c)). As we indicated by the expressions in Table I, the two interpretations of the particles'

heading under TRS yield very different entropy productions. For the odd interpretation, the entropy production increases abruptly at the phase transition, when the average velocity of the particles drastically decreases because of MIPS. For the even interpretation, the opposite trend is observed.

We then investigate the effect of the alignment interactions of the entropy production. Therefor, we set $\phi = 0.4$ and $\gamma_R = 0.3$, while we vary the alignment strength K (Fig. 2(d) to Fig. 2(f)). In case of odd parity of the heading, the transition (at approximately $K = 0.002$) from MIPS to homogeneous particles distribution is reflected by a sudden drop in the expected entropy production rate. The entropy production starts to decrease even more as the system begin to exhibit oriented collective motion (at approximately $K = 0.006$). The opposite trend is observed for the even interpretation of the heading under TRS.

Finally, we analyse the entropy production rate associated with individual particles. A summary of our findings is provided in Fig. 3, which shows the expected entropy production rates of individual particles, calculated using Eq. (14) and (15), for the three configurations of the system previously seen in Fig. 1(d). When no global orientational order

is present (Fig. 3, $K = 0.0007$ and $K = 0.0036$), it is clear that the entropy production is related to the local density of the particles. This is particularly clear for the phase characterised by MIPS (and no orientational order) for which, in case of odd heading, higher entropy production rates are observed in the solid region and lower entropy production rates are observed in the gaseous region (the opposite is observed when the even parity is considered). Defects in the solid structures are also captured by the entropy production rates, as shown in Fig. 4: the entropy production associated with particles along the defects is in general higher (for the even interpretation of θ) than the one of particles in highly ordered regions.

In the phase characterised by both MIPS and orientational order, the two interpretations of the heading also yield different entropy production rates (cf. Fig. 3, $K = 0.0097$). However, it is for the even interpretation of the particles' heading that we now observe higher entropy production in areas of high density, where lower entropy production is observed for the odd interpretation. Moreover, the interplay between alignment torque and excluded volume interactions is observed to generate complex patterns of entropy production.

V. CONCLUSIONS

Recently, there has been a growing interest in the thermodynamics of active matter, and particularly the entropy production associated to MIPS has been investigated [14, 21]. Shankar et al. [24] demonstrated the importance of using under-damped models to quantify entropy production of active matter, focussing on free, non-interacting particles. In this study, we have investigated the entropy production in a large system of ABPs interacting via volume exclusion and alignment. We have provided the expressions for the expected entropy production rate of the system for both an over-damped and an under-damped description, confirming that the under-damped description is more suitable for studying the thermodynamics associated to collective motion. Moreover, we have shown that the interpretation of the particle's heading as an

odd or an even variable under TRS leads to different expressions of the entropy production rate of interacting particles, in addition to what was pointed out for non-interacting particles [24].

The under-damped system was numerically simulated and shown to exhibit three kinetic phases: a phase characterised by a homogeneous distribution of the particles' positions and orientations, a phase characterised by MIPS and a phase exhibiting both MIPS and global orientational order. Crucially, the phase transitions were shown to be reflected in changes in the average entropy production rate at the (non-equilibrium) steady-state. Additionally, the entropy production rate associated with individual particles was shown to capture the state of active matter, e.g., lower entropy production was measured for particles in solid regions while higher entropy production was measured in gaseous regions, for the even interpretation of the particles' heading. Defects forming in solid structures were also observed to be reflected in the particles' entropy production rate.

Concluding, we hope that this work will contribute to a better understanding of the thermodynamics of active matter, particularly in phenomena where kinetic phase transitions and critical regimes [31] are involved. Formal analogies have been recently developed [28, 32] between stochastic thermodynamics and information dynamics, a discipline that describes information processing in complex systems in terms of intrinsic computation. Future work will aim at exploring such relations in active matter.

ACKNOWLEDGMENTS

E.C. was supported by the University of Sydney's "Post-graduate Scholarship in the field of Complex Systems" from Faculty of Engineering & IT and by a CSIRO top-up scholarship. The authors acknowledge the University of Sydney HPC service at The University of Sydney for providing HPC resources that have contributed to the research results reported within this paper.

-
- [1] M. C. Marchetti, J.-F. Joanny, S. Ramaswamy, T. B. Liverpool, J. Prost, M. Rao, and R. A. Simha, *Reviews of Modern Physics* **85**, 1143 (2013).
 - [2] C. Bechinger, R. Di Leonardo, H. Löwen, C. Reichhardt, G. Volpe, and G. Volpe, *Reviews of Modern Physics* **88**, 045006 (2016).
 - [3] S. Ramaswamy, *Journal of Statistical Mechanics: Theory and Experiment* **2017**, 054002 (2017).
 - [4] J. Bialké, T. Speck, and H. Löwen, *Journal of Non-Crystalline Solids* **407**, 367 (2015).
 - [5] A. Czirók, M. Matsushita, and T. Vicsek, *Physical Review E* **63**, 031915 (2001).
 - [6] A. Sokolov, R. E. Goldstein, F. I. Feldchtein, and I. S. Aranson, *Physical Review E* **80**, 031903 (2009).
 - [7] B. Szabo, G. Szöllösi, B. Gönci, Z. Jurányi, D. Selmeczi, and T. Vicsek, *Physical Review E* **74**, 061908 (2006).
 - [8] J. K. Parrish, S. V. Viscido, and D. Grünbaum, *Biological Bulletin* **202**, 296 (2002).
 - [9] J. Buhl, D. J. T. Sumpter, I. D. Couzin, J. J. Hale, E. Despland, E. R. Miller, and S. J. Simpson, *Science* **312**, 1402 (2006).
 - [10] A. Cavagna, A. Cimarelli, I. Giardina, G. Parisi, R. Santagati, F. Stefanini, and M. Viale, *Proceedings of the National Academy of Sciences* **107**, 11865 (2010).
 - [11] T. Vicsek and A. Zafeiris, *Physics Reports* **517**, 71 (2012).
 - [12] S. Ramaswamy, *Annual Review of Condensed Matter Physics* **1**, 323 (2010).
 - [13] L. Schimansky-Geier, M. Mieth, H. Rosé, and H. Malchow, *Physics Letters A* **207**, 140 (1995).
 - [14] É. Fodor, C. Nardini, M. E. Cates, J. Tailleur, P. Visco, and F. van Wijland, *Physical Review Letters* **117**, 038103 (2016).
 - [15] M. C. Marchetti, Y. Fily, S. Henkes, A. Patch, and D. Yllanes, *Current Opinion in Colloid & Interface Science* **21**, 34 (2016).
 - [16] G. S. Redner, M. F. Hagan, and A. Baskaran, *Physical Review Letters* **110**, 055701 (2013).

- [17] M. E. Cates and J. Tailleur, *Annual Review of Condensed Matter Physics* **6**, 219 (2015).
- [18] P. Digregorio, D. Levis, A. Suma, L. F. Cugliandolo, G. Gonnella, and I. Pagonabarraga, *Physical Review Letters* **121**, 098003 (2018).
- [19] J. T. Siebert, F. Dittrich, F. Schmid, K. Binder, T. Speck, and P. Virnau, *Physical Review E* **98**, 030601 (2018).
- [20] A. Martín-Gómez, D. Levis, A. Díaz-Guilera, and I. Pagonabarraga, *Soft Matter* **14**, 2610 (2018).
- [21] D. Mandal, K. Klymko, and M. R. DeWeese, *Physical Review Letters* **119**, 258001 (2017).
- [22] R. E. Spinney and I. J. Ford, *Physical Review E* **85**, 051113 (2012).
- [23] U. Seifert, *Reports on Progress in Physics* **75**, 126001 (2012).
- [24] S. Shankar and M. C. Marchetti, *Physical Review E* **98**, 020604 (2018).
- [25] U. Seifert, *Physical Review Letters* **95**, 040602 (2005).
- [26] U. Seifert, *The European Physical Journal B* **64**, 423 (2008).
- [27] R. E. Spinney and I. J. Ford, *Physical Review Letters* **108**, 170603 (2012).
- [28] R. E. Spinney, J. T. Lizier, and M. Prokopenko, *Physical Review E* **98**, 032141 (2018).
- [29] M. Esposito, *Physical Review E* **85**, 041125 (2012).
- [30] A. Celani, S. Bo, R. Eichhorn, and E. Aurell, *Physical Review Letters* **109**, 260603 (2012).
- [31] E. Crosato, R. E. Spinney, R. Nigmatullin, J. T. Lizier, and M. Prokopenko, *Physical Review E* **97**, 012120 (2018).
- [32] R. E. Spinney, J. T. Lizier, and M. Prokopenko, *Physical Review E* **94**, 022135 (2016).

Entropy production during collective motion of active Brownian particles undergoing phase separation and alignment

Supplemental material

Emanuele Crosato,^{1,2,*} Mikhail Prokopenko,¹ and Richard E. Spinney¹

¹*Complex Systems Research Group and Centre for Complex Systems,
Faculty of Engineering and IT, The University of Sydney, Sydney, NSW 2006, Australia.*

²*CSIRO Data61, PO Box 76, Epping, NSW 1710, Australia.*

(Dated: December 12, 2018)

DERIVATION OF THE ENTROPY PRODUCTIONS

The over-damped and under-damped models presented in the main text involve continuous Markovian dynamics described by uncorrelated stochastic differential equations (SDEs). The evolution of each coordinate $x_i \in \{\mathbf{r}_a^x, \mathbf{r}_a^y, \mathbf{v}_a^x, \mathbf{v}_a^y, \theta_a, \omega_a\}$ can be expressed in terms of a deterministic component $A_i(x, t)$ and a stochastic component $B_i(x, t)$:

$$dx_i = A_i(x, t)dt + B_i(x, t)dW_i, \quad (1)$$

where W_i is the Wiener process. The deterministic dynamics $A_i(x, t)$ can be further divided into reversible and irreversible components [1]:

$$dx_i = A_i^{\text{REV}}(x, t)dt + A_i^{\text{IR}}(x, t)dt + B_i(x, t)dW_i. \quad (2)$$

The diffusion coefficients associated to each coordinate are $D_i(x, t) = B_i(x, t)^2/2$.

Following [2], and noting that there is no multiplicative noise in the models, we have:

$$d\Delta S^{\text{med}} = \sum_{a=1}^N \sum_i \frac{A_{x_i}^{\text{IR}}(x)}{D_i} \circ dx_i - \frac{A_{x_i}^{\text{IR}}(x)A_{x_i}^{\text{REV}}(x)}{D_i} dt \quad (3)$$

Entropy production in the over-damped model

Let us consider the over-damped model described in the main text:

$$d\mathbf{r}_a^j = v_0 \hat{e}^j(\theta_a)dt - \partial_{\mathbf{r}_a^j} U(\mathbf{r})_a dt + \sqrt{\frac{2T}{m\gamma}} dW_{\mathbf{r}_a^j}, \quad (4)$$

$$d\theta_a = \frac{1}{I\gamma_R} \mathcal{T}(\mathbf{r}, \theta)_a dt + \sqrt{\frac{2T}{I\gamma_R}} dW_{\theta_a}. \quad (5)$$

Regardless of the parity of the self-propulsion under TRS, we have $A_{\theta_a}^{\text{IR}} = \mathcal{T}(\theta, \mathbf{r})_a/I$, $A_{\theta_a}^{\text{REV}} = 0$, $D_{\theta_a} = T/I\gamma_R$ and $D_{\mathbf{r}_a^j} = T/m\gamma$.

Odd self-propulsion

For the odd interpretation of the particles' heading under TRS, we obtain $A_{\mathbf{r}_a^j}^{\text{IR}} = -\partial_{\mathbf{r}_a^j} U(\mathbf{r})_a/m$ and $A_{\mathbf{r}_a^j}^{\text{REV}} = v_0 \hat{e}^j(\theta_a)$. Applying Eq. (3) we obtain:

$$d\Delta S^{\text{med}} = \sum_{a=1}^N \left(\frac{\mathcal{T}(\theta, \mathbf{r})_a}{T} \circ d\theta_a - \sum_{j \in \{x, y\}} \left(\frac{\partial_{\mathbf{r}_a^j} U(\mathbf{r})_a}{T} \circ d\mathbf{r}_a^j + \frac{v_0 \hat{e}^j(\theta_a) \partial_{\mathbf{r}_a^j} U(\mathbf{r})_a}{T} dt \right) \right). \quad (6)$$

* emanuele.crosato@sydney.edu.au

2

At the steady-state, assuming $\langle dU(\mathbf{r}) \rangle = 0$ and $\langle (\hat{e}^j(\theta_a))^2 \rangle = 1/2$, we obtain:

$$d\langle \Delta S^{\text{med}} \rangle = \sum_{a=1}^N \left(\frac{\langle \mathcal{T}^2(\theta, \mathbf{r})_a \rangle}{TI\gamma_R} + \frac{\langle \partial_{\theta_a} \mathcal{T}(\theta, \mathbf{r})_a \rangle}{I\gamma_R} + \sum_{j \in \{x, y\}} \frac{v_0}{T} \left\langle \partial_{\mathbf{r}_a^j} U(\mathbf{r})_a \hat{e}^j(\theta_a) \right\rangle \right). \quad (7)$$

Even self-propulsion

For the even interpretation of the particles' heading, we obtain $A_{\mathbf{r}_a^j}^{\text{IR}} = v_0 \hat{e}^j(\theta_a) - \partial_{\mathbf{r}_a^j} U(\mathbf{r})_a/m$ and $A_{\mathbf{r}_a^j}^{\text{REV}} = 0$. Applying Eq. (3), we then obtain:

$$d\Delta S^{\text{med}} = \sum_{a=1}^N \left(\frac{\mathcal{T}(\theta, \mathbf{r})_a}{T} \circ d\theta_a + \sum_{j \in \{x, y\}} \frac{v_0 \hat{e}^j(\theta_a) - \partial_{\mathbf{r}_a^j} U(\mathbf{r})_a}{T} dt \right). \quad (8)$$

At the steady-state, assuming $\langle dU(\mathbf{r}) \rangle = 0$ and $\langle (\hat{e}^j(\theta_a))^2 \rangle = 1/2$, we obtain:

$$d\langle \Delta S^{\text{med}} \rangle = \sum_{a=1}^N \left(\frac{\langle \mathcal{T}^2(\theta, \mathbf{r})_a \rangle}{TI\gamma_R} + \frac{\langle \partial_{\theta_a} \mathcal{T}(\theta, \mathbf{r})_a \rangle}{I\gamma_R} + \frac{m\gamma v_0^2}{T} - \sum_{j \in \{x, y\}} \frac{v_0}{T} \left\langle \partial_{\mathbf{r}_a^j} U(\mathbf{r})_a \hat{e}^j(\theta_a) \right\rangle \right). \quad (9)$$

Entropy production in the under-damped model

Let us consider the under-damped model described in the main text:

$$d\mathbf{r}_a^j = \mathbf{v}_a^j dt, \quad (10)$$

$$d\mathbf{v}_a^j = -\gamma \mathbf{v}_a^j dt + \gamma v_0 \hat{e}^j(\theta_a) dt - \frac{1}{m} \partial_{\mathbf{r}_a^j} U(\mathbf{r})_a dt + \sqrt{\frac{2T\gamma}{m}} dW_{\mathbf{v}_a^j}, \quad (11)$$

$$d\theta_a = \omega_a dt, \quad (12)$$

$$d\omega_a = -\gamma_R \omega_a dt + \frac{1}{I} \mathcal{T}(\mathbf{r}, \theta)_a dt + \sqrt{\frac{2T\gamma_R}{I}} dW_{\omega_a}. \quad (13)$$

Regardless of the odd or even interpretation of the self-propulsion under TRS, we have $A_{\omega_a}^{\text{IR}} = -\gamma_R \omega_a$, $A_{\omega_a}^{\text{REV}} = \mathcal{T}(\theta, \mathbf{r})_a/I$, $A_{\theta_a}^{\text{IR}} = 0$, $A_{\theta_a}^{\text{REV}} = \omega_a$, $A_{\mathbf{r}_a^j}^{\text{IR}} = 0$, $A_{\mathbf{r}_a^j}^{\text{REV}} = \mathbf{v}_a^j$, $D_{\omega_a} = T\gamma_R/I$, $D_{\theta_a} = 0$, $D_{\mathbf{v}_a^j} = T\gamma/m$, and $D_{\mathbf{r}_a^j} = 0$.

Odd self-propulsion

For the odd interpretation of the particles' heading under TRS, we obtain $A_{\mathbf{v}_a^j}^{\text{IR}} = -\gamma \mathbf{v}_a^j + \gamma v_0 \hat{e}^j(\theta_a)$ and $A_{\mathbf{v}_a^j}^{\text{REV}} = -\partial_{\mathbf{r}_a^j} U(\mathbf{r})_a/m$. Applying Eq. (3) we obtain:

$$d\Delta S^{\text{med}} = \sum_{a=1}^N \left(\frac{\gamma_R}{T} \left(I \langle \omega_a^2 \rangle - T \right) + \sum_{j \in \{x, y\}} \frac{\gamma}{T} \left(m \langle (\mathbf{v}_a^j - v_0 \hat{e}^j(\theta_a))^2 \rangle - T \right) \right). \quad (14)$$

At steady-state, recognising $\langle \mathbf{v}_a^j \circ d\mathbf{v}_a^j \rangle = 0$, $\langle \omega_a \circ d\omega_a \rangle = 0$, $\langle dU(\mathbf{r}) \rangle = 0$ and $\langle (\hat{e}^j(\theta_a))^2 \rangle = 1/2$, we obtain:

$$d\langle \Delta S^{\text{med}} \rangle = \sum_{a=1}^N \left(\frac{\gamma_R}{T} \left(I \langle \omega_a^2 \rangle - T \right) + \sum_{j \in \{x, y\}} \frac{\gamma}{T} \left(\frac{mv_0^2}{2} - m \langle (\mathbf{v}_a^j)^2 \rangle + T \right) \right). \quad (15)$$

Even self-propulsion

For the even interpretation of the particles' heading, we obtain $A_{\mathbf{v}_a^j}^{\text{IR}} = -\gamma \mathbf{v}_a^i$ and $A_{\mathbf{v}_a^j}^{\text{REV}} = \gamma v_0 \hat{e}^j(\theta_a) - \partial_{\mathbf{r}_a^j} U(\mathbf{r})_a/m$. Applying Eq. (3), we obtain:

$$d\Delta S^{\text{med}} = \sum_{a=1}^N \left(\frac{\gamma^{\text{R}}}{T} \left(I \langle \omega_a^2 \rangle - T \right) + \sum_{j \in \{x, y\}} \frac{\gamma}{T} \left(m \langle (\mathbf{v}_a^j)^2 \rangle - T \right) \right). \quad (16)$$

-
- [1] H. Risken, in *The Fokker-Planck Equation* (Springer, 1996) pp. 63–95.
 [2] R. E. Spinney and I. J. Ford, *Physical Review E* **85**, 051113 (2012).

Chapter 7

Discussion and conclusions

7.1 A unifying framework for collective motion

This thesis investigates the phenomenon of collective motion of self-propelled particles in a variety of real and simulated active matter systems. Multiple aspects of active matter are investigated, including the propagation of information during collective directional changes, kinetic phase transitions and critical regimes, and the thermodynamic costs for the formation of collective motion. What distinguishes this thesis from other studies of collective motion is the particular focus on issues that are related to information, and the intention to approach them from rigorous information-theoretic and thermodynamic perspectives.

In particular, some well-known concepts in collective motion are recasted as entropic quantities. For example, information propagation across particles is typically investigated using a variety of spatial correlations [Nagy et al., 2010; Cavagna et al., 2013b; Cavagna et al., 2013a; Procaccini et al., 2011] while, in this thesis, it is quantified information-theoretically, using transfer entropy. Similarly, kinetic phase transitions in active matter are usually investigated with the help of suitably designed order parameters [Vicsek et al., 1995; Chat   et al., 2008] while, in this work, they are also associated to changes in the average configuration entropy of the system. The sensitivity of the system to changes in the control parameters is also described information-theoretically and information-geometrically, using Fisher information.

One of the main objectives of this thesis is to relate information-theoretic concepts in collective motion and more tangible, physical aspects of this phenomenon. For example, the study in Chap. 3 relates positive and negative information flows across a school of fish to spatial interactions, i.e., the transfer entropy from a fish to another is mapped to the relative position and heading of the two fish. In Chap. 4, a model of collective motion is studied in the context of a physical process, i.e., a quasi-static protocol, which drives the system across a kinetic phase transition. Information-theoretic quantities, such as the configuration entropy of the system and the Fisher information, are contrasted with thermodynamical quantities such as internal energy and work. The study of a system of active Brownian particles in Chap. 6 moves from the analysis of quasi-static processes to the investigation of non-equilibrium aspects of active matter. Different phases of the system are characterised not only in terms of their associated behaviour [Siebert et al., 2018; Digregorio et al., 2018; Mart  n-G  mez et al., 2018], but also thermodynamically, in terms of the entropy production within the system.

Chap. 3 highlighted the important role of identifying predictive information flows in understanding collective motion, and the self-organisation of collective dynamics in general. Since self-organisation involves energy and entropy flows, the next question we set out to answer was how information dynamics in self-organising collectives can be understood in terms of the underlying physical fluxes [Prokopenko and Lizier, 2014; Prokopenko et al., 2013; Spinney et al., 2016; Spinney et al., 2018]. This was investigated in Chap. 4, where the precise energy and entropy

dynamics were traced for different phases and critical regimes of a model of active matter. These phases and regimes were then contrasted in terms of their thermodynamic efficiency—a concept introduced in this work. As a result, the underlying information flows and the corresponding thermodynamic quantities (energy, entropy, and work) were shown to abruptly change at criticality. This was further studied in another context, i.e., population dynamics over different time scales (Chap. 5), confirming the utility of the thermodynamic efficiency in identifying critical regimes. Analysis of thermodynamic fluxes which shape collective self-organisation becomes particularly challenging for non-equilibrium dynamics within active matter. Chap. 6 describes the first step in this direction, focussing on a topical model of active Brownian particles, for which we derived analytical expressions for the entropy production rate, supported by numerical simulations. This has again allowed me to explicitly identify different phases and critical regimes of the dynamics. Equipped with these analytical tools, we can approach a study of the thermodynamic efficiency of active matter in non-equilibrium settings—this is a subject of future research.

An important feature of the framework developed in this thesis is its generality, which allows its application to the study of other collective phenomena. For example, the study in Chap. 5 demonstrates that the relations between Fisher information, entropy, energy and work can also be applied to urban dynamics, while Harding et al., 2018 have applied them to the analysis of the thermodynamic efficiency of contagions. The transfer entropy approach developed in the study of a school of fish can be applied to the other distributed information processes including, for example, neural networks [Wibral et al., 2014]. Similarly, the method based on stochastic thermodynamics proposed to study the entropy productions in Chap. 6 can be adapted to other microscopic non-equilibrium collective phenomena [Seifert, 2012].

7.2 Summary of the main results

The objectives of this thesis, presented in Sec. 1.2, have been met in the four studies, the main results of which are here summarised.

Informative and misinformative interactions in a school of fish

Information propagation is a key aspect of the collective motion in animal groups which, however, is hard to quantify due to continuous and noisy nature of information processing in these natural systems. This study has proposed an information-theoretical approach, based on the local transfer entropy, to measure information transfer across the implicit and distributed communication channels of a school of fish, which are characterised by noise, delay and long-ranged effects.

Hemigrammus rhodostomus fish were placed in a ring-shaped tank, and their trajectories were recorder during their circular motion, which included occasional

collective U-turns. Pairwise time-delayed dependencies between the recorded trajectories were analysed using the local transfer entropy, revealing predictive information flows that intensify during the U-turns. Information propagation during collective directional changes has been long hypothesised, but never verified information-theoretically in a biological system.

Crucially, some misinformative information flows, characterised by negative values of local transfer entropy, were observed, indicating that observing the heading of a fish increases the uncertainty associated with the prediction of the next direction of another fish. Specifically, two distinct predictive information flows were identified during collective U-turns: an informative flow from fish that have already turned to fish that are turning, and a misinformative flow from fish that have not turned yet to the fish that are turning. This produces a sequential cascade of the dynamics of information flow, observed, but not captured information-theoretically in [Potts, 1984; Procaccini et al., 2011; Herbert-Read et al., 2015], and also distinguishes between informative and misinformative transfer.

The role of spatial dynamics in generating the influential interactions that carry the informative and misinformative information flows was investigated. Local values of transfer entropy were mapped to the relative heading between fish pairs, revealing clear spatio-informational motifs. Positive information flow is detected from a fish to the others positioned behind it (comparably to what was found in other studies including [Nagy et al., 2010; Katz et al., 2011; Rosenthal et al., 2015], and whose headings are perpendicular rather than parallel to it. Negative information flow is instead generally detected from a fish to the others positioned at the side and with opposite heading.

In this study it is shown that a fish can thus choose to move in the opposite direction of the majority, temporarily suppressing its behavioural tendency to align with its neighbours. When this happens, transfer entropy reveals that the other fish are seen as a source of misinformation to an observer. This is an example of how the information-theoretic approach can provide a better understanding of the processes underlying collective decisions in animal groups.

Thermodynamics and computation during collective motion

This study has presented a theoretical framework for quantifying thermodynamical quantities in systems driven by quasi-static protocols, and for relating them to information-theoretic quantities. It has been shown that the rate of changes of the generalised work required to vary the control parameters quasi-statically is proportional to the Fisher Information. Additionally, an information-geometric interpretation of the curvature of the internal energy was provided, as the difference between two curvatures: the curvature of the free entropy, captured by the Fisher information, and the curvature of the configuration entropy. This expression highlights the

computational balance between the sensitivity and the uncertainty of the computation which are represented, respectively, by the Fisher information and the configuration entropy.

The framework was applied to the study of quasi-static protocol driving a well-known model of collective motion [Grégoire and Chaté, 2004] across a kinetic phase transition from a disordered phase, in which particles maintain a fairly stable collective position but do not align, to a polarised phase, in which particles cohesively move towards a common direction. The system was simulated for several values of the control parameters, until it reached a stationary state, in order to numerically estimate the probability distribution of the velocity of the particles. It was shown that the critical points of the phase transition are captured by the divergence of the Fisher information of the velocity distribution with respect to the control parameters. The divergence of the Fisher information was thus used in order to construct a phase diagram of the behaviour of the system, which was consistent with the one of Grégoire and Chaté, 2004. This approach to building phase diagrams can come useful when the definition of a suitable order parameter is hard.

The numerical results showed that the balance between sensitivity and uncertainty of the computation is stressed at criticality, where the curvatures with respect of the control parameters of work, energy and configuration entropy diverge. This is in line with several experimental observations that flocking behaviour combines coherence and responsiveness to external perturbations [Cavagna et al., 2013b; Procaccini et al., 2011]. Moreover, it was shown that the thermodynamic efficiency of computation, introduced as the ratio of the gain in internal order to the expenditure of work, was the highest at criticality. In computational terms, the thermodynamic efficiency corresponds to a ratio of generated order to the sensitivity accumulated over a process running from the current state to the state of perfect order. Thus, the maximum efficiency at criticality also indicates that achieving one bit of uncertainty reduction is the most convenient, in terms of accumulated sensitivity, during phase transition.

Critical dynamics and thermodynamics of urban transformations

This study has demonstrated the applicability of the theoretical framework developed in Chap. 4 in the field of urban dynamics. Specifically, the study investigates a hypothetical phase transition of the urban settlement within Greater Sydney, from a dispersed configuration, in which the population settles homogeneously, to a polycentric configuration, in which the population aggregates in a few highly populated centres.

A model of income flow between employment areas and residential suburbs was developed, based on the Boltzmann-Lotka-Volterra (BLV) method [Harris and Wilson, 1978; Wilson, 2008]. Slow dynamics, modelled as Lotka-Volterra equations, govern the evolution of the local services, which determines the attractiveness of the suburbs. A fast equilibration, modelled using the maximum entropy principle

(from which comes the name “Boltzmann”), determines the flow of residents, and their income, from employment to residential places that are consistent with the distribution of the services. The model, calibrated to Census data and geospatial datasets, involves two control parameters: the social disposition, balancing the suburbs’ attractiveness, and travel impedance, balancing the discomfort attributed to commuting trips.

The phase transition, detected using Fisher information, was shown to be induced by the social disposition, and was investigated thermodynamically by considering a quasi-static protocol for the variation of this control parameter. In particular, the minimum work required to vary a control parameter was investigated. It was shown that the ratio of the order gained during a change to the required work, a concept that we introduced as the thermodynamic efficiency of urban transformation, is maximised at criticality.

Entropy production during collective motion of active Brownian particles

The thermodynamic treatment of active matter has recently attracted a growing interest. The entropy production associated with an over-damped model of self-propelled particles, which interact only via volume exclusion, has been studied by Fodor et al., 2016 and Mandal et al., 2017, under the framework of statistical thermodynamics. A later study by Shankar and Marchetti, 2018 has demonstrated that important components of the entropy production along the trajectory of free (non interacting) particles are missed when an over-damped description of the dynamics is used. The authors also pointed out that the interpretation of the particles’ self-propulsion as odd or even under time-reversal symmetry (TRS) yields different expressions for the entropy production.

The study presented in this thesis has investigated the entropy production in a system of active Brownian particles that interact through volume exclusion as well as alignment. Inspired by Shankar and Marchetti, 2018, the analytical expressions for the expected entropy production rate in the medium (which at steady-state equals the average total entropy production) are derived, for both under-damped and over-damped dynamics, and for odd and even interpretation of the particles’ heading under TRS. Such expressions confirm that the choice of over-damped or under-damped descriptions affects the thermodynamic treatment of the system. Perhaps even more strikingly, is the effect of the two parity interpretations of the particles’ heading in the entropy production expressions for under-damped model. Under the even parity interpretation, the entropy production quantifies the deviation from equipartition expected at equilibrium, in both the translational and rotational degrees of freedom. Under the odd parity interpretation, the entropy production associated to the translational degrees of freedom measures the deviation from an effective equipartition, relative to the instantaneous self-propulsion force.

The under-damped model is numerically simulated over three control parameters, i.e., the particles’ density, the self-propulsion persistence and the alignment

strength. Three distinct kinetic phases of the systems are identified: a phase in which particles' positions and orientations are homogeneously distributed, a phase characterised by MIPS and a phase exhibiting both MIPS and orientational order. The phase transitions are shown to be reflected in changes of the expected entropy production rate, calculated using the derived expressions. Moreover, individual particles in areas of the system characterised by different states of active matter, e.g., 'solid', 'gas' and 'flocks', have been shown to be associated to different expected entropy production rates. The entropy production rate has also been shown to capture defects in the solid structures.

7.3 Future directions

A main problem in the study of collective motion of animal groups is the identification of 'effective' neighbours, i.e., the individuals that have a direct effect on the motion of another. Effective neighbourhoods cannot be inferred using only pairwise relationships in the same way as in Chap. 3, where transfer entropy was applied to measure pairwise information transfer across a school of fish.

Future research will focus on multivariate extensions to the transfer entropy which, for example, involve the conditioning of other source fish. Such extensions may be useful in order to eliminate redundant pairwise relationships and also for detecting higher-order relationships beyond pairwise [Lizier et al., 2008; Lizier and Prokopenko, 2010; Lizier et al., 2010; Vakorin et al., 2009; Williams and Beer, 2011]. These modifications are anticipated to be challenging, since they will have to deal with the short-term and dynamic structure of interactions across the collective. Lord et al., 2016 has used conditional transfer entropy, averaged over a time-window, to study effective interactions during insects' flights, however, the collective dynamics on the local scale are still to be investigated.

The theoretical framework developed in Chap. 4 was applied to the study of a model of collective motion [Grégoire and Chaté, 2004]. Future studies will aim at applying the concepts developed in this study, and in particular the thermodynamic efficiency of computation, to more realistic systems. Such investigations will reveal to what extent the developed framework can be used in order to study and design real active matter. The application of the framework to the study of models that are data-driven was demonstrated in Chap. 5, where the urban dynamics of Greater Sydney were studied. Similarly, the approach developed in Chap. 6 for the quantification of the entropy production in a system of active Brownian particles, will be applied to analyse systems that involve more intricate interactions which include, for example, hydrodynamic [Baskaran and Marchetti, 2009; Zöttl and Stark, 2014; Menzel et al., 2016] and/or electrochemical effects [Yan et al., 2016], as well as a comparative analysis with respect to thermodynamic efficiency in non-equilibrium settings. Moreover, formal analogies between stochastic thermodynamics and information dynamics have been recently developed Spinney et al., 2016; Spinney et al.,

2018—future work will aim at exploring such relations in active matter.

In general, future research will aim at improving and extending the information-theoretic and thermodynamic framework developed in this thesis, which is expected to have potential applications in better understanding collective motion in natural systems and in engineering artificial active matter.

References

- Baskaran, A. and M. C. Marchetti (2009). “Statistical mechanics and hydrodynamics of bacterial suspensions”. In: *Proceedings of the National Academy of Sciences* 106.37, pp. 15567–15572.
- Cavagna, A. et al. (2013a). “Diffusion of individual birds in starling flocks”. In: *Proceedings of the Royal Society of London B: Biological Sciences* 280.1756, p. 20122484.
- Cavagna, A., I. Giardina, and F. Ginelli (2013b). “Boundary information inflow enhances correlation in flocking”. In: *Physical Review Letters* 110.16, p. 168107.
- Chaté, H. et al. (2008). “Modeling collective motion: variations on the Vicsek model”. In: *The European Physical Journal B* 64.3–4, pp. 451–456.
- Digregorio, P. et al. (2018). “Full phase diagram of active Brownian disks: from melting to motility-induced phase separation”. In: *Physical Review Letters* 121.9, p. 098003.
- Fodor, É. et al. (2016). “How far from equilibrium is active matter?” In: *Physical Review Letters* 117.3, p. 038103.
- Grégoire, G. and H. Chaté (2004). “Onset of collective and cohesive motion”. In: *Physical Review Letters* 92.2, p. 025702.
- Harding, N., R. Nigmatullin, and M. Prokopenko (2018). “Thermodynamic efficiency of contagions: a statistical mechanical analysis of the SIS epidemic model”. In: *Interface focus*.
- Harris, B. and A. G. Wilson (1978). “Equilibrium values and dynamics of attractiveness terms in production-constrained spatial-interaction models”. In: *Environment and Planning A* 10.4, pp. 371–388.
- Herbert-Read, J. E. et al. (2015). “Initiation and spread of escape waves within animal groups”. In: *Royal Society Open Science* 2.4, p. 140355.
- Katz, Y. et al. (2011). “Inferring the structure and dynamics of interactions in schooling fish”. In: *Proceedings of the National Academy of Sciences* 108.46, pp. 18720–18725.
- Lizier, J. T. and M. Prokopenko (2010). “Differentiating information transfer and causal effect”. In: *The European Physical Journal B* 73.4, pp. 605–615.
- Lizier, J. T., M. Prokopenko, and A. Y. Zomaya (2008). “Local information transfer as a spatiotemporal filter for complex systems”. In: *Physical Review E* 77.2, p. 026110.
- (2010). “Information modification and particle collisions in distributed computation”. In: *Chaos: An Interdisciplinary Journal of Nonlinear Science* 20.3, p. 037109.

- Lord, W. M. et al. (2016). "Inference of causal information flow in collective animal behavior". In: *IEEE Transactions on Molecular, Biological and Multi-Scale Communications* 2.1, pp. 107–116.
- Mandal, D., K. Klymko, and M. R. DeWeese (2017). "Entropy production and fluctuation theorems for active matter". In: *Physical Review Letters* 119.25, p. 258001.
- Martín-Gómez, A. et al. (2018). "Collective motion of active Brownian particles with polar alignment". In: *Soft matter* 14.14, pp. 2610–2618.
- Menzel, A. M. et al. (2016). "Dynamical density functional theory for microswimmers". In: *The Journal of Chemical Physics* 144.2, p. 024115.
- Nagy, M. et al. (2010). "Hierarchical group dynamics in pigeon flocks". In: *Nature* 464.7290, pp. 890–893.
- Potts, W. K. (1984). "The chorus-line hypothesis of manoeuvre coordination in avian flocks". In: *Nature* 309.5966, pp. 344–345.
- Procaccini, A. et al. (2011). "Propagating waves in starling, *Sturnus vulgaris*, flocks under predation". In: *Animal Behaviour* 82.4, pp. 759–765.
- Prokopenko, M. and J. T. Lizier (2014). "Transfer entropy and transient limits of computation". In: *Scientific reports* 4, p. 05394.
- Prokopenko, M., J. T. Lizier, and D. C. Price (2013). "On thermodynamic interpretation of transfer entropy". In: *Entropy* 15.2, pp. 524–543.
- Rosenthal, S. B. et al. (2015). "Revealing the hidden networks of interaction in mobile animal groups allows prediction of complex behavioral contagion". In: *Proceedings of the National Academy of Sciences* 112.15, pp. 4690–4695.
- Seifert, U. (2012). "Stochastic thermodynamics, fluctuation theorems and molecular machines". In: *Reports on Progress in Physics* 75.126001, p. 126001.
- Shankar, S. and M. C. Marchetti (2018). "Hidden entropy production and work fluctuations in an ideal active gas". In: *Physical Review E* 98.2, p. 020604.
- Siebert, J. T. et al. (2018). "Critical behavior of active Brownian particles". In: *Physical Review E* 98.3, p. 030601.
- Spinney, R. E., J. T. Lizier, and M. Prokopenko (2016). "Transfer entropy in physical systems and the arrow of time". In: *Physical Review E* 94.2, p. 022135.
- (2018). "Entropy balance and information processing in bipartite and nonbipartite composite systems". In: *Physical Review E* 98.3, p. 032141.
- Vakorin, V. A., O. A. Krakovska, and A. R. McIntosh (2009). "Confounding effects of indirect connections on causality estimation". In: *Journal of Neuroscience Methods* 184.1, pp. 152–160.
- Vicsek, T. et al. (1995). "Novel type of phase transition in a system of self-driven particles". In: *Physical Review Letters* 75.6, p. 1226.
- Wibral, M., R. Vicente, and M. Lindner (2014). "Transfer entropy in neuroscience". In: *Directed Information Measures in Neuroscience*. Springer, pp. 3–36.
- Williams, P. L. and R. D. Beer (2011). "Generalized measures of information transfer". In: *arXiv preprint arXiv:1102.1507*.

- Wilson, A. (2008). "Boltzmann, Lotka and Volterra and spatial structural evolution: an integrated methodology for some dynamical systems". In: *Journal of The Royal Society Interface* 5.25, pp. 865–871.
- Yan, J. et al. (2016). "Reconfiguring active particles by electrostatic imbalance". In: *Nature Materials* 15.10, p. 1095.
- Zöttl, A. and H. Stark (2014). "Hydrodynamics determines collective motion and phase behavior of active colloids in quasi-two-dimensional confinement". In: *Physical Review Letters* 112.11, p. 118101.