Physica A 448 (2016) 216–223

Contents lists available at ScienceDirect

Physica A

journal homepage: www.elsevier.com/locate/physa

A Fisher-gradient complexity in systems with spatio-temporal dynamics

A. Arbona [a,*], C. Bona [a], J. Massó [a], B. Miñano [a], A. Plastino [b]

[a] IAC^3–UIB, Mateu Orfila, Cra. de Valldemossa km 7.5, 07122, Palma, Spain

[b] National University La Plata, Physics Institute (IFLP-CCT-CONICET) C.C. 727, 1900 La Plata, Argentina

HIGHLIGHTS

- Statistical complexity measures (SCM) for spatio-temporal systems studied.
- We define a benchmark of complexity.
- We focus in particular on the Collective Motion model.
- We analyse LMC's, Autocorrelation, and Fisher-gradient complexities.

ARTICLE INFO

Article history:
Received 16 June 2015
Received in revised form 26 October 2015
Available online 31 December 2015

Keywords:
Complexity
Statistical complexity measures
Spatial dynamical systems
Collective Motion
Fisher information

ABSTRACT

We define a benchmark for definitions of complexity in systems with spatio-temporal dynamics and employ it in the study of Collective Motion. We show that LMC's complexity displays interesting properties in such systems, while a statistical complexity model (SCM) based on autocorrelation reasonably meets our perception of complexity. However this SCM is not as general as desirable, as it does not merely depend on the system's Probability Distribution Function. Inspired by the notion of Fisher information, we develop a SCM candidate, which we call the Fisher-gradient complexity, which exhibits nice properties from the viewpoint of our benchmark.

1. Introduction

Perfect disorder maximizes missing-information, in the same fashion as entropy does, but it is actually not much more complex than perfect order, which minimizes entropy. As Crutchfield noted in 1994, *Physics does have the tools for detecting and measuring complete order equilibria and fixed point or periodic behaviour and ideal randomness via temperature and thermodynamic entropy or, in dynamical contexts, via the Shannon entropy rate and Kolmogorov complexity. What is still needed, though, is a definition of structure and way to detect and to measure it* [1].

Seth Lloyd counted as many as 40 ways to define complexity, none of them being completely satisfactory. A major breakthrough came from the definition of statistical complexity proposed by López-Ruiz, Mancini and Calbet (LMC) [2]. Although not without problems [3,4], LMC's complexity clearly separated and quantified the contributions of entropy and structure. LMC measured structure via the concept of *disequilibrium*. Building on this proposal, Kowalski et al. [4] refined the definition of disequilibrium.

* Corresponding author.
E-mail addresses: aarbona@iac3.eu (A. Arbona), cbona@uib.es (C. Bona), jmasso@iac3.eu (J. Massó), bminyano@iac3.eu (B. Miñano), plastino@fisica.unlp.edu.ar (A. Plastino).

http://dx.doi.org/10.1016/j.physa.2015.12.093
0378-4371/

One should note that while entropy is a general concept that can be applied across a wide range of model families, this is not the case with measures of structure. For them one needs to know, in advance, what to look for. The LMC proposal can be regarded as a way to provide a very general definition of structure, which does an excellent job for many systems, particularly those involving time series, but also 1D spatial systems [5].

We are, however, specifically interested in a statistical complexity measures (SCM) for models with spatial dimensions. This includes dynamical PDE-based models, such as Navier–Stokes, on the one hand, and Agent-Based Models (ABM), such as Collective Motion [6], on the other. Here we use the adjective *dynamical* because the structures we are interested in are easily recognized (at least visually) by studying velocity fields. Other models of interest are static (i.e., not characterized by a velocity field, but rather from scalar quantities such as density or spin). Examples of these models include PDE-based models such as Reaction–Diffusion, and Cellular Automata (e.g. Ising models).

Within this framework one may try to be more specific in the definition of structure. In a previous paper [7], we showed that, for these systems, a good candidate for appropriately capturing the structural component in the definition of complexity is a correlation (specifically the velocity autocorrelation field in the case of the Collective Motion model).

Density fields as a proxy for structural information were studied in the case of reaction–diffusion models [8] and atoms [9].

The problem with these approaches is that the definition of structure is quite ad hoc, limited to families of models. In this paper we explore the possibility to define a more general statistical complexity measure. We aim for a definition which is more specific to spatial systems than the one based on disequilibrium, but still general in this context, that so it does not depend on the existence of specific fields such as velocity or density.

We start by observing a common feature of perceived complexity in spatial systems characterized by a velocity field: both perfect order and perfect disorder are characterized by vanishing spatial derivatives and time derivatives of the velocity probability distribution function (PDF) at mesoscopic scales. A well grounded information measure involving derivatives is the Fisher information one [10], which has already been used to analyse electronic structure [11,12]. We aim for a measure which is independent from the chosen coordinate system and takes into account all spatial derivatives.

The paper is structured as follows. In Section 2 we define the procedure we employ to define a PDF for the Collective Motion model, and we provide definitions of different candidates to SCM based on the PDF definition. In Section 3 we define a benchmark that will allow us to quantify the assessment of complexity for models with spatial coordinates. We also recall the Collective Motion model, which will serve to test the SCM candidates against our complexity benchmark. Section 4 presents the results of such test, which we discuss in Section 5 to show that LMC's complexity cannot be trivially extended to spatial systems, and that a SCM inspired by Fisher information does a good job at capturing the essence of complexity and is generalizable to any system with spatio-temporal dynamics.

2. Definitions

It is important to acknowledge that the perception of complexity is deeply entangled with the scale of measurement. Therefore, we should aim at measuring complexity at different scales. The idea of studying complexity as a function of scale is not new, as represented for instance in the concepts of *complexity profile* [13,14] and *d-diameter complexity* [15].

A reference microscale, in Agent-Based Models, is given by the typical (average) separation of two agents (δ). Larger scales (mesoscale and macroscale) may be characterized as certain multiples, with values depending on the system, of this basic scale. The calculation of metrics at a certain scale l at point x, y[1] are done by considering all agents falling into a ball of radius $l/2$ around x, y.

We focus on models characterized by a velocity field, which we will illustrate with the 2D Collective Motion model. That is, we consider the state of the system as $s = \{v_\alpha^i\}$, where i indexes spatial coordinates and α individual agents. The same procedure is feasible for other fields (density, spin, and so on). One needs to properly characterize the fields that, in each case, better define the structures representing complexity.

Following [6], and re-scaling so we get a positive value, we define the velocity autocorrelation of a pair of agents α and β as

$$A_{\alpha\beta} \equiv \frac{1}{2} + \frac{v_\alpha^x v_\beta^x + v_\beta^y v_\alpha^y}{v_\alpha^2 + v_\beta^2}. \qquad (1)$$

For each particle α we compute the average velocity autocorrelation with all neighbours β within radius $l/2$[2] so that we get $A_{(l)}(x, y)$.

This definition of velocity correlation guarantees positivity, ranging from 0 in the anti-parallel case (−1 in the standard definition of velocity correlation) to 1 in the parallel case (also 1 using the standard definition).

[1] Let us develop, without loss of generality, the expressions for 2D Cartesian systems.

[2] In practice this is only reasonable for small scales. For large scales we need to sample the pairs we take into account.

Notice that this definition is not based purely on a PDF, and it is therefore less generic than PDF-based measures. For our purposes, a PDF amounts to a vector field: it can be obtained at any point in space and for each point we have a vector with a dimension equal to the number of bins N we have in the PDF. In 2D, given Cartesian coordinates x and y, we have

$$PDF(x, y) \equiv p(x, y, b) \quad b = 1 \dots N \tag{2}$$

where x, y represent a spatial position and b a specific bin. Notice that

$$\sum_{b=1}^{N} p(x, y, b) = 1. \tag{3}$$

In the particular case of Collective Motion we compute the PDF of the velocity angle, which in this model is what determines structure. To do so we consider, for a certain point, all the agents within a certain radius $l/2$, and classify them into N bins, corresponding to N partitions of the agent orientation angle (we tend to use $N = 8$) [16]. In case of a ball empty of agents, we consider the PDF to be homogeneous.[3]

Now, based on the PDF, we can define many generic metrics, meaning they are not dependent on the existence of a particular field such as the velocity field. They only require the capability to derive a PDF from the structural characteristics of the system under consideration. Notice this is not only valid for classical systems, but for quantum systems as well (by defining $p = |\psi|^2$ [17]).

We will consider the following metrics: Shannon entropy, disequilibrium, and a PDF gradient function inspired by the Fisher information quantifier. From them, we will derive statistical complexity measures, and we will effect pertinent comparisons amongst them.

The Shannon entropy is defined in our notation as[4]

$$H(x, y) \equiv -\frac{1}{\log(N)} \sum_{b=1}^{N} p(x, y, b) \ \log\left(p(x, y, b)\right) \tag{4}$$

which we normalized so it ranges between 0 and 1.

Disequilibrium, as defined by *LMC*, is

$$D(x, y) = \sum_{b=1}^{N} \left(p(x, y, b) - \frac{1}{N}\right)^2. \tag{5}$$

In our notation, the Fisher information taking as parameter a spatial coordinate reads, up to normalization factors

$$F^{x^r}(x, y) = \sum_{b=1}^{N} p(x, y, b) \left(\frac{\partial \log \ p(x, y, b)}{\partial x^r}\right)^2. \tag{6}$$

Notice this definition lacks invariance under variations of the coordinate system. We consider the following invariant alternative:

$$F(x, y) = \sum_{b=1}^{N} p(x, y, b) \left[\nabla \log \ p(x, y, b)\right]^2 \tag{7}$$

where the gradient is taken relative to spatial coordinates. Therefore, we are considering spatial variations of the PDF, while the Fisher information is typically built on variations regarding parameters of the model. Nevertheless, the coordinates can be interpreted as just labelling the space location: coordinates can be considered therefore just a supplementary parameter-set. In this sense, we use the term *Fisher-gradient* when referring to $F(x, y)$.

With these definitions we have all the basic ingredients to produce statistical complexity measures. As introduced in Ref. [7], we define the Autocorrelation Complexity field at scale l as

$$C^{AC}_{(l)}(x, y) = H_{(l)}(x, y) \ A_{(l)}(x, y). \tag{8}$$

Notice that we can average this field over all the spatial domain in order to obtain the global complexity measure $C^{AC}_{(l)}$. A definition of complexity which is purely based on the PDF is the LMC complexity:

$$C^{LMC}_{(l)}(x, y) = H_{(l)}(x, y) \ D_{(l)}(x, y). \tag{9}$$

Finally, as an ansatz inspired by the Fisher information, we define the Fisher-gradient complexity as

$$C^{FG}_{(l)}(x, y) \equiv F_{(l)}(x, y). \tag{10}$$

[3] Notice this definition is more refined than the definition we used in Ref. [7], which used a fictitious mesh in order to obtain the PDF.

[4] The logarithm is in base 2.

The candidates we consider as SCMs for systems with velocity-based spatio-temporal dynamics are then C^{AC}, C^{LCM}, and C^{FG}. C^{LCM} can be considered a 'gold standard' for models based on time series, but it is, as far as we know, untested in 2D or 3D models characterized by velocity fields. C^{AC} is specific to velocity-based models, and C^{FG} is our ansatz for generalizing the good behaviour of C^{AC} to PDF-only metrics, while keeping some specificity to spatial 2D and 3D systems.

In all cases, a global intensive C is recovered by averaging over all the simulation domain.

3. Benchmark

In order to make the comparison quantitative, we need to define a benchmark, which specifies what we mean by complexity. We intend to be as exhaustive as possible, exploring all aspects in the concept of *complexity*.

Our benchmark will consist of the following elements:

- Behaviour in different spatial regions: we expect the SCM to be maximal at the interfaces between ordered and disordered regions [7].
- Behaviour around critical points: we expect the SCM to present a maximum at critical points.
- Behaviour across spatial scales: we expect the SCM to be maximum at the mesoscale.

The behaviour of the SCM candidates under these tests will gauge how well they match our perception of complexity.

As mentioned, we choose Collective Motion as the model on which to confront the SCM candidates with our benchmark. Specifically we use the vect noise model of Grégoire and Chaté [18]. The model is a variation of the Original Vicsek Model [19] (OVM), devised to reproduce the collective motion – or flocking – we find in many biological and non-biological systems (see, for instance, Refs. [20–22,6] for reviews). In these systems long-range orientation order emerges after spontaneous symmetry breaking.

In the OVM, point-like agents move synchronously in discrete time steps, with a fixed common speed v_0. In 2D the orientation of agent α is an angle θ_α. The evolution rule provides the new angles at each time step, based on the angles of the agent's neighbours (agents within a certain influence radius) in the previous time step. Essentially, the agent tries to align itself with its neighbours. This alignment is perturbed by a white noise.

The Grégoire and Chaté model is based on the OVM, but modifies the manner in which noise is incorporated into the model. They define *vectorial noise* as generated by errors when estimating interactions, in comparison with the *angular noise* in the OVM, related to errors in trying to follow the newly computed direction. Altogether, the update rule for the Grégoire and Chaté model is

$$\theta_\alpha^{t+1} = arg\left[\sum_{\beta\sim\alpha} e^{i\theta_\beta^t} + \eta n_\alpha^t e^{i\xi_\alpha^t}\right], \tag{11}$$

where ξ is a delta-correlated white noise, η represents noise-strength, and n is the number of neighbours. The sum is made over all the neighbours (β) of agent α.

The solution of this system ranges from nearly complete orientation order for low noise intensity to random orientation for high noise intensity. These phases are separated by a novel phase transition. The solutions near the transition point are characterized, for a wide spectrum of parameters, by ordered moving structures (bands) separated by disordered interband regions. An example of this latter situation is shown in Figs. 1 and 2.

Zooming into the velocity field (Fig. 2) we see how the chaotic component rules the dynamics. However, at a larger scale (Fig. 1) the dynamics is characterized by clustering at a mesoscale, while isotropy emerges at larger scales (if we adequately zoom out).

4. Results

In Figs. 3–5 we compare the LMC, Autocorrelation, and Fisher-gradient complexity fields at scale $l = 5\delta$ (which we interpret as a mesoscale). This allows us to provide a qualitative assessment on our first test in the benchmark: behaviour in different spatial regions. The LMC complexity is high not only at the regions exhibiting transition from order to disorder (at the band edges), but also in the ordered regions (particularly, high inside the bands rather than at the band edges). The Autocorrelation complexity is particularly high in the bands, especially at the edges. However, we would expect greater specificity (i.e. a lower complexity in the inner part of the band).

The Fisher-gradient complexity is only high at the transition region, effectively acting in this case as an edge detector, as expected (as it responds to spatial variations of the PDF). Notice that would not be the case if we look at the system at the microscale (e.g., $l = \delta$). In this case, the variations of the PDF, both at ordered and disordered regions, are significant, due to the random fluctuations that characterize the microscale.

In Fig. 6 we compare our three definitions of complexity, focusing on its behaviour against scale. We understand mesoscale as the scale around 3δ to 20δ, which is the size of the structures and features of the model. This provides a basis for assessing our test related to mesoscale-detection. The LMC complexity is essentially flat across scales, although it appears to show a maximum close to the mesoscale. The Autocorrelation complexity appropriately grows as the mesoscale approaches, but then fails to fall down as entering into the macroscale.

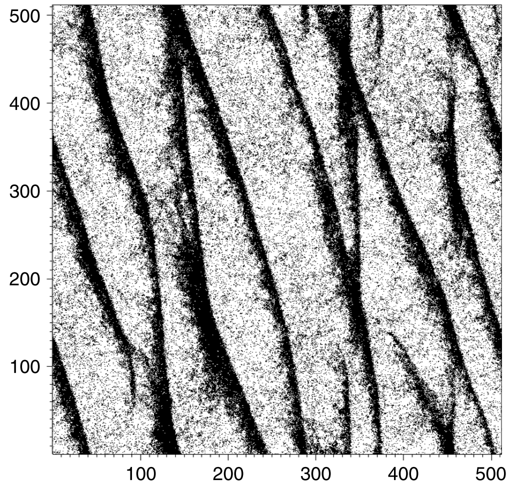

Fig. 1. Agent distribution in space for a 131.072 particle simulation of a collective motion model. The model implemented is the vectorial noise model proposed by Grégoire and Chaté [18], with domain size $L = 256$, density $\rho = 2$, noise level $\eta = 0.611$, speed $v_0 = 0.5$, and time step $\Delta t = 1$. In this snapshot the system has evolved to a stationary state characterized by high-density bands travelling in a certain direction.

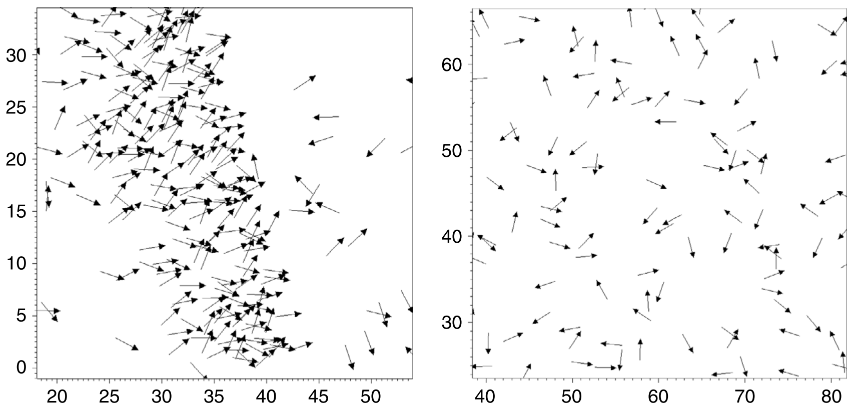

Fig. 2. Zoom on two small regions of the same simulation represented in Fig. 1. Arrows correspond to the velocity of each agent. The left figure zooms into a band; notice the preferred directions East and Northeast. The right figure zooms into an inter-band space, and shows a disordered pattern of velocities.

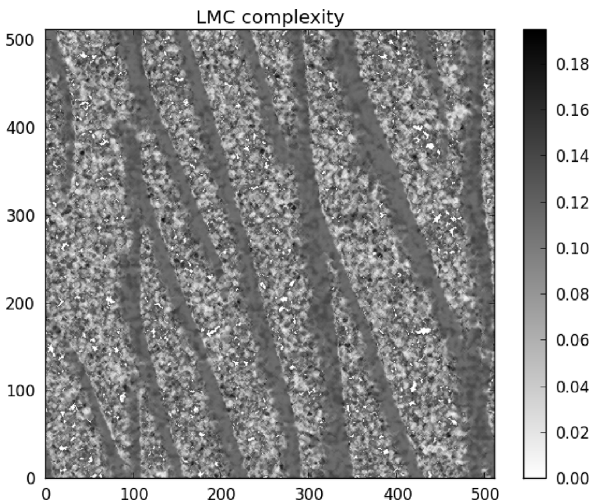

Fig. 3. $C^{LMC}(x, y)$ as computed at scale $l = 5\delta$ from the data described in Fig. 1.

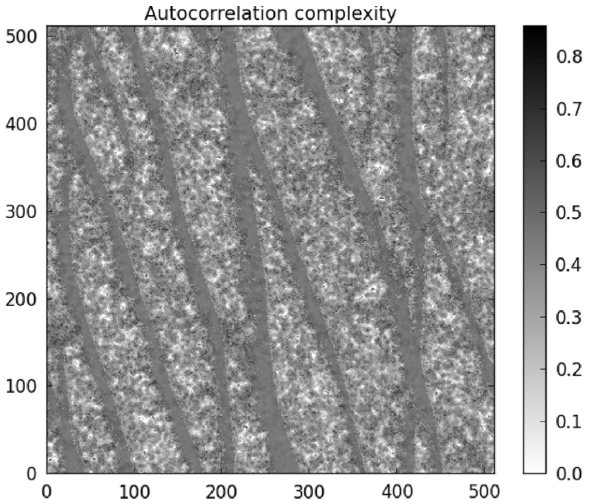

Fig. 4. $C^{AC}(x, y)$ as computed at scale $l = 5\delta$ from the data described in Fig. 1.

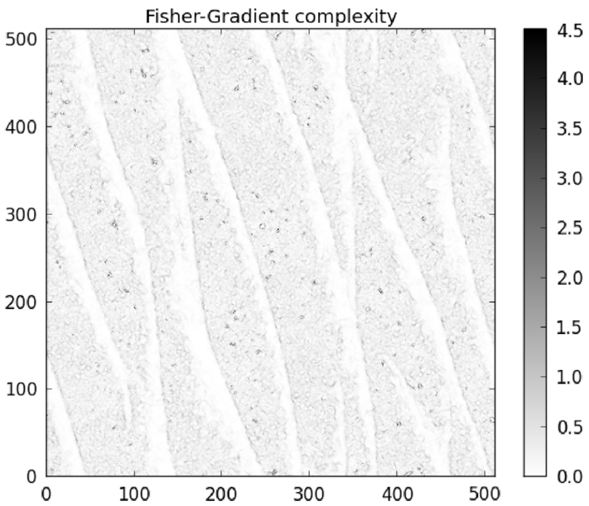

Fig. 5. $C^{FG}(x, y)$ as computed at scale $l = 5\delta$ from the data described in Fig. 1.

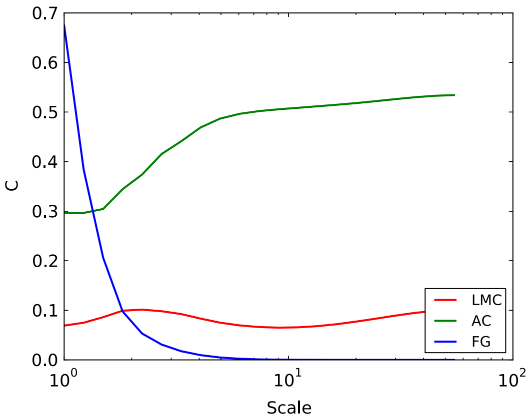

Fig. 6. C_l^{LMC}, C_l^{AC}, and C_l^{FG} as computed for different scales from the data described in Fig. 1.

The Fisher-gradient complexity also fails at detecting the mesoscale. The reason is that the PDF spatial variation behaves 'wildly' at the microscale (due to random variations caused by the finite nature of the model, as the number of agents tend to zero for scale tending to zero), and progressively diminishes in the highly ordered and disordered regions as the scale grows.

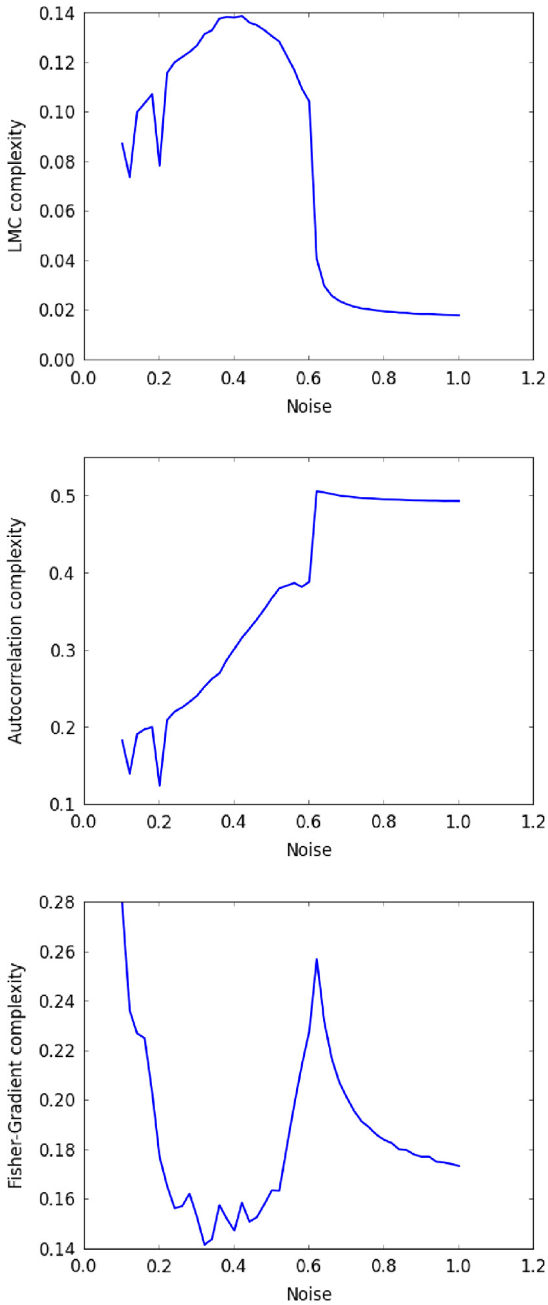

Fig. 7. C^{LMC}, C^{AC}, and C^{FG} as computed for scale $l = 5\delta$ as a function of noise.

Finally, in Fig. 7 we look at the behaviour of the complexity definitions against noise. We expect them to detect the critical point ($\eta = 0.611$). As we see, the LMC complexity does detect the critical point, suffering a critical transition itself. However, the complexity is not maximal around the critical point, as we would expect. The Autocorrelation complexity does find a maximum at the critical point, although a relatively weak one. The Fisher-gradient complexity displays a strong signal around the critical point, where its value is maximal. This is the expected behaviour as we defined in the benchmark.

5. Conclusions

By using Collective Motion, a well-known and rich model for dynamical critical transitions in 2D, we have analysed the suitability of three SCM candidate definitions: the LMC complexity, Autocorrelation complexity, and Fisher-gradient complexity. Such suitability has been qualitatively assessed using three tests: behaviour across spatial regions, versus scale, and versus noise (the parameters that triggers a phase transition in Collective Motion).

We find that the LMC complexity may be a helpful metric to consider in the context of spatio-temporal dynamics, even if it was originally conceived for time series. It behaves better than the other two metrics in the test against scale. It is not very specific in the test of spatial regions, and in the test around critical points it displays a behaviour which is not ideal but intriguing anyway (the maximum of the complexity is detected in the vicinity to the transition point, while at the transition point it suffers a critical transition).

The Autocorrelation complexity does a decent job in two of the tests (spatial behaviour and phase transition), although its sharpness is not optimal in either case. Regarding mesoscale, it detects the transition from microscale to mesoscale, but not the transition from mesoscale to macroscale. However, this SCM definition is not solely based on the PDF, and therefore it is not directly generalizable to other systems with spatio-temporal dynamics, such as density-based or spin-based systems. Both the LMC and Fisher-gradient complexity are purely based on a PDF, on the other hand.

The Fisher-gradient complexity meets successfully two of the three tests: it sharply detects both highly complex spatial regions and systemic phase transitions. However, it does not detect the mesoscale. This is mainly due to the finite nature of the model, which implies that at some scales it is not reasonable to build a PDF since there are not enough agents at hand.

A future line of experimentation involves testing the Fisher-gradient complexity and the LMC complexity against spin-based and density-based systems with spatio-temporal dynamics. In some density-based spatial systems the PDF still retains its meaning at microscale, which will allow for a more sensible assessment of the behaviour of the Fisher-gradient complexity against scale and will improve our overview concerning the value of the LMC complexity in systems with spatio-temporal dynamics.

Finally, we remark that the Fisher-gradient complexity does not work by separately distilling chaos and structure. It rather works as an *edge-detector* for the surfaces where chaos and structure meet each other.

Acknowledgements

The research leading to these results has received funding from the European Union Seventh Framework Programme (FP7/2007-2013) under grant agreement no. 317534 (the Sophocles project). We thank Omri Har-Shemesh (University of Amsterdam) for interesting discussions on Fisher information.

References

[1] James P. Crutchfield, The calculi of emergence: computation, dynamics and induction, Physica D 75 (1–3) (1994) 11–54.
[2] R. Lopez-Ruiz, H.L. Mancini, X. Calbet, A statistical measure of complexity, Phys. Lett. A 209 (5–6) (1995) 321–326.
[3] David P. Feldman, James P. Crutchfield, Measures of statistical complexity: Why? Phys. Lett. A 238 (4–5) (1998) 244–252.
[4] Andres M. Kowalski, Maria Teresa Martin, Angelo Plastino, Osvaldo A. Rosso, Montserrat Casas, Distances in probability space and the statistical complexity setup, Entropy 13 (6) (2011) 1055–1075.
[5] Juan R. Sánchez, Ricardo López-Ruiz, A method to discern complexity in two-dimensional patterns generated by coupled map lattices, Physica A 355 (2) (2005) 633–640.
[6] Tamas Vicsek, Anna Zafeiris, Collective motion, Phys. Rep. 517 (3–4) (2012) 71–140.
[7] A. Arbona, C. Bona, B. Miñano, A. Plastino, Statistical complexity measures as telltale of relevant scales in emergent dynamics of spatial systems, Physica A 410 (2014) 1–8.
[8] Kristian Lindgren, Anders Eriksson, Karl-Erik Eriksson, Flows of information in spatially extended chemical dynamics, in: Proceedings of ALife, vol. 9, 2004.
[9] Jaime Sañudo, Ricardo López-Ruiz, Statistical complexity and Fisher–Shannon information in the H-atom, Phys. Lett. A 372 (32) (2008) 5283–5286.
[10] B.R. Frieden Read, Science from Fisher information, Cambride University Press, New York, 2004.
[11] Elvira Romera, J.S. Dehesa, The Fisher–Shannon information plane, an electron correlation tool, J. Chem. Phys. 120 (19) (2004) 8906–8912.
[12] Kali Das Sen, Statistical complexity: applications in electronic structure, Springer Science & Business Media, 2011.
[13] Yaneer Bar-Yam, Complexity rising: From human beings to human civilization, a complexity profile, in: Encyclopedia of Life Support Systems (EOLSS), UNESCO, EOLSS Publishers, Oxford, UK, 2002.
[14] Yaneer Bar-Yam, Multiscale complexity/entropy, Adv. Complex Syst. 7 (01) (2004) 47–63.
[15] Gregory J. Chaitin, Toward a mathematical definition of life, in: R.D. Levine, M. Tribus (Eds.), The Maximum Entropy Formalism, MIT Press, 1979.
[16] Lijie Xu, Teng-Yok Lee, Han-Wei Shen, An information-theoretic framework for flow visualization, IEEE Trans. Vis. Comput. Graphics 16 (6) (2010) 1216–1224.
[17] S.P. Flego, A. Plastino, A.R. Plastino, Fisher information and quantum mechanics, Int. Res. J. Pure Appl. Chem. 2 (1) (2012) 25–54.
[18] Guillaume Grégoire, Hugues Chaté, Onset of collective and cohesive motion, Phys. Rev. Lett. 92 (2004).
[19] Tamás Vicsek, András Czirók, Eshel Ben-Jacob, Inon Cohen, Ofer Shochet, Novel type of phase transition in a system of self-driven particles, Phys. Rev. Lett. 75 (6) (1995) 1226.
[20] John Toner, Yuhai Tu, Long-range order in a two-dimensional dynamical XY model: how birds fly together, Phys. Rev. Lett. 75 (23) (1995) 4326.
[21] Hugues Chaté, Francesco Ginelli, Guillaume Grégoire, Franck Raynaud, Collective motion of self-propelled particles interacting without cohesion, Phys. Rev. E 77 (4) (2008).
[22] Andreas Deutsch, Guy Theraulaz, Tamas Vicsek, Collective motion in biological systems, Interface Focus 2 (6) (2012) 689–692.