Why we should use Topological Data Analysis in Ageing: towards defining the "Topological shape of ageing"

Tamàs Fülöp[1*], Mathieu Desroches[2,3], Alan A Cohen[4], Fernando Antônio Nóbrega Santos[5,6], Serafim Rodrigues[7,8*]

[1]Research Center on Aging, Geriatric Division, Department of Medicine, Faculty of Medicine and Health Sciences, University of Sherbrooke, Sherbrooke, Quebec, Canada

[2]MathNeuro Team, Inria Sophia Antipolis Méditerranée, France

[3]Université Côte d'Azur, Nice, France

[4]Research Center on Aging, Department of Family Medicine, Faculty of Medicine and Health Sciences, University of Sherbrooke, Sherbrooke, Quebec, Canada

[5]Departamento de Matemática, Universidade Federal de Pernambuco, Brazil

[6]Department of Anatomy & Neurosciences, Amsterdam UMC, Vrije Universiteit Amsterdam, Amsterdam Neuroscience, 1081 HZ, Amsterdam, The Netherlands

[7]Ikerbasque, The Basque Foundation for Science, Bilbao, Spain

[8]BCAM - The Basque Center for Applied Mathematics, Bilbao, Spain

[*]Corresponding authors: Tamas.Fulop@USherbrooke.ca, srodrigues@bcamath.org

Abstract

Living systems are subject to the arrow of time; from birth, they undergo complex transformations (self-organization) in a constant battle for survival, but inevitably ageing and disease trap them to death. Can ageing be understood and eventually reversed? What tools can be employed to further our understanding of ageing? The present article is an invitation for biologists and clinicians to consider key conceptual ideas and computational tools (known to mathematicians and physicists), which potentially may help dissect some of the underlying processes of ageing and disease. Specifically, we first discuss how to classify and analyze complex systems, as well as highlight critical theoretical difficulties that make complex systems hard to study. Subsequently, we introduce Topological Data Analysis - a novel Big Data tool – which may help in the study of complex systems since it extracts knowledge from data in a holistic approach via topological considerations. These conceptual ideas and tools are discussed in a relatively informal way to pave future discussions and collaborations between mathematicians and biologists studying ageing.

Introduction

What is known, what is unknown and what are the unmet needs for ageing

Ageing has fascinated humans ever since they appeared on earth (Cohen et al., 2020a). Why is this the case? Humans are probably the only living beings that are conscious of their own decay, and they realize that the end of ageing is death as a natural continuum of life. An almost invariant hallmark of this decay is that death is preceded by an accumulation of age-related diseases (Kennedy et al., 2014). If humans were immortal, then ageing would not bear the same mystery and fascination. Despite significant advances in science, we still don't know what ageing is and consequently why humans (and other living systems) age (Fülöp et al., 2020a; Fülöp et al., 2019). This lack of understanding is reflected in the ageing literature, in the sense that the foundational knowledge on ageing is yet sketchy and even the definition of ageing has not found a consensus. In part, this is because scientists define ageing based upon their field of specialization, which also depends on their capacity to conceptualize and synthesize the complexity of empirical observations (Cohen et al., 2020). This lack of consensus questions whether ageing (as a proper scientific field of study) *per se* exists. Moreover, it raises doubts whether ageing questions can be well-formulated and if such questions are amenable to scientific methodologies, or if ultimately they are only used/defined by scientists to give a sense to their specific niche of research (Cohen et al., 2020b)? Indeed, presently several hundred theories exist, all of which attempt to define ageing. However, at the current stage, it is unlikely that any of these will explain or ultimately conceptualize ageing (Rose et al., 2012; Lipsky et al., 2015; Cannon, 2015; da Costa et al., 2016). Also, there are so many aspects of ageing that none of the existing theories can integrate (and synthesize) them by including molecular, biological, physiological, functional, social, or psychological aspects. On the other hand, these uncertainties have led to the generation of data across various scales, starting from molecular through the cellular to the organismal level. These data exist most of the time in silos and serve mainly to explain correlations but not to understand the holistic causal relationship of the data defining ageing. However, some attempts exist to unify the ageing concept and to try to find a causal relationship among the different data conglomerates or pathways. We subsequently provide three examples, which are the most advanced ones to conceptualize ageing.

The first example is the *epigenetic clock,* described by Horvath (Horvath, 2013). This clock predicts biological age and provides a collection of innate biological mechanisms that give rise to age-related DNA methylation changes that underlie highly accurate DNAm age estimators, and thus provides a tentative answer as to why we age (Horvath and Raj, 2018). This clock was derived via machine learning methods, notably elastic net regression (Horvath, 2013). Since then, several epigenetic clocks have been described (Belsky et al., 2018) with the aim to determine the biological age compared to the chronological age (Hannum, PhenoAge and GrimAge) (Li et al., 2020; Moskalev, 2020). The second example is the *dysregulation* concept of Cohen et al. (Milot et al., 2014; Cohen et al., 2015), which employs a more theoretical approach to computation (e.g. statistical distance, D_M, Principal Component Analysis) to explain the complexity of ageing. There is now clear evidence that physiological dysregulation--the gradual breakdown in the capacity of complex regulatory networks to maintain homeostasis--is an emergent property of these regulatory networks, and that it plays an essential role in ageing. It can be measured merely using small numbers of biomarkers (Cohen, 2016; Mitnitski et al., 2017; Chmielewski, 2020). The third and

most recent example can be called the *multi-level process* of ageing, which is being led by the research group of Franceschi (Whitwell et al., 2020). In this approach biological ageing is considered as a complex process involving multiple biological processes, which can be understood theoretically by considering them as an individual network (e.g. epigenetic networks, cell-cell networks, and population genetics). Mathematical modelling (e.g. via multi-layer networks (Kivelä et al., 2014)) allows the combination of such networks so that they may be studied as a whole, to better understand how the so-called "seven pillars of ageing" combine (Kennedy et al., 2014). This approach also has the potential of generating hypotheses for treating ageing as a condition at relatively early biological ages (Whitwell et al., 2020). These examples highlight the unmet need for more advanced computational tools to further dissect the complexity of ageing from the multitude of data sets gathered during these last decades (Zhavoronkov et al., 2019). Ideally, these tools should treat ageing data as a problem (and in a holistic sense), extract causalities, explain the data mechanistically and ultimately define targets for ageing interventions (Zhavoronkov et al., 2019).

Several recent attempts have been put forward recently to determine the different components of the biology of ageing, recognizing that indeed aging is a multicomponent complex process. For instance, *the nine hallmarks of ageing* by the group of Kroemer (López-Otín et al., 2013) has been proposed. These hallmarks are conceptualized by noting that ageing is led by molecular and physiological deregulation and the subsequent appearance of pathological consequences such as diabetes, cancer, cardiovascular diseases. Specifically, the hallmarks are genomic instability, telomere attrition, epigenetic alterations, loss of proteostasis, deregulated nutrient sensing, mitochondrial dysfunction, cellular senescence, stem cell exhaustion, and altered intercellular communication (López-Otín et al., 2013). Moreover, these studies have tried to define some of the components of these hallmarks but most of these are at a phenomenological level. Thus, this remains very descriptive and does not preclude any causality, and even the interaction and interconnection between these hallmarks are conjectural only. Moreover, no hierarchy has been established adequately among these hallmarks in the sense to provide causal interactions to generate a holistic/global/integrated response of the entire organism during aging. The ultimate aim was indeed to find targets for intervention to cure/modulate/influence or reverse ageing to avoid or at least decrease the emergence of the above-mentioned age-related diseases. In this sense, ageing is considered a disease that requires a cure primarily to decrease the burden of age-related diseases (Bulterijs et al. 2015), however considering all biological aspects of aging it appears that ageing cannot be considered as a disease (Fülöp et al., 2019).

This view has been adopted and refurbished by the so-called geroscience, which clearly conceptualizes ageing as a pathological phenomenon that should be controlled to decrease age-related chronic diseases (Kennedy et al., 2014). Moreover, geroscience conceptualizes ageing by the interaction of the seven pillars such as metabolism, epigenetics, inflammation, adaptation to stress, proteostasis, stem cells and regeneration, macromolecular damage. These are somehow more broad processes than that defined by the nine hallmarks, but most probably there are many overlapping between the two classifications. At least the aim is the same, which is to use ageing as a target to decrease age-associated diseases and prolong health span. While this is an important goal, the hallmarks/pillars classification will only enable the derivation of methods that intervene separately on individual hallmarks since most interactions currently considered are usually pair-wise (and not genuinely causal) and thus holistic interventions on ageing (and associated emergent processes) are likely to fail.

Furthermore, it is crucial to stress that many of these hallmarks/pillars are mainly derived from animal models of ageing. However, this is now quite accepted that none of these models alone may not represent entirely human ageing either because of the longevity or the complexity or the environmental stresses (Wagner, 2017, Hugenholtz and de Vos2018; Maynard and Weinkove, 2018, Bair et al. 2019). As such, it is possible that the described hallmarks/pillars are simply the surrogate of the specific animal models/pathways even if they are presented as universal underlying biological phenomena across the whole animal kingdom.

Yet another problem is that the hallmarks/pillars are only rarely related to the dynamic physiological or functional aspects of ageing (Belsky et al., 2018) even if there are some attempts to connect to physiology in view of better determining the mortality (Li et al., 2020). Thus, it is almost impossible to state what is the potential impact of these models on the ageing of humans, and in the best-case scenario, these may only serve as surrogate data to assess/capture the dynamic and holistic nature of aging. Even longitudinal studies have so far failed to answer these crucial concerns. As a biological process and as much as it can exist if there is no physiological or functional repercussion, ageing *per se* may be just an adaptation to the passage of time, as was recently suggested e.g. for immunological changes (Franceschi et al., 2017; Fülöp et al., 2018, Fülöp et al. 2020b). For example, grey hair may represent an "ageing" phenomenon, but it has no real consequences except that probably cell senescence is the underlying process. In contrast, the decrease of the hepatic blood flow which occurs with ageing, probably ultimately due to mitochondrial dysregulation (metabolic, oxidative, inflammatory), is not measured and rarely considered, yet it may have serious consequences on detoxification. Thus, ageing is one of the most intriguing processes that humans ought to understand through life since what can be phenotypically observed as ageing is not necessarily associated with the real complex mechanisms of ageing. In other words, we can observe chronological age, but we cannot yet dissect the underlying biological processes of ageing that occur along time. Human data related to some dynamic functional processes should, therefore, be acquired in order to unravel the importance of these processes for ageing. Furthermore, one could determine in this way the biological age of individuals. Indeed, the above-mentioned epigenetic clock tried to fill this gap (Horvath, 2013; Horvath and Raj, 2018).
Another widespread approach to ageing is its consideration as only a deleterious process given that it is inevitably resulting in death. However, many phenomena that are described, measured, or considered as related to ageing may be simply adaptative processes designed to optimize the functioning of the organism at a specific moment and circumstance of life related to some evolutionary necessity (Franceschi et al., 2017; Fülöp et al., 2018). Until now, all described hallmarks are considered to some extent as causes of ageing however they cannot be considered as *a priori* deleterious and being targeted for intervention to "rejuvenate" the system since their effect are not determined from the viewpoint of species or individuals' evolution. In the immune system, the best example is the activation state of the innate immune system with ageing, called *inflammaging* (Franceschi et al., 2017) but, in contrast, may also be considered as an evolutionary defence mechanism against repetitive challenges from both outside and inside in order to support a better readiness of the organism (Fülöp et al., 2016). As long as we cannot fully understand this phenomenon in the complexity of the immune response changes related to ageing, we cannot really decide what the role of this resting overactivation is, and how we should intervene if any intervention is necessary.
Thus, until now many data have been gathered at various levels, but no efficient tools exist that are able to conceptualize in a causal manner all these processes, either molecular or cellular, tightly

linked to physiological or clinical data. Interdisciplinary researchers and clinicians should closely collaborate to address and hopefully answer the following key questions: *"What is the meaning of these basic hallmarks of ageing on the physiological/clinical performance of the older subjects? And would targeting them change something in the trajectory of an older individual subject for a better healthspan?"* Addressing these key challenges will be the foundation of precision medicine for older subjects. A thorough understanding of ageing thus implies an integrative, complex systems framework where lower-level mechanisms can have direct impacts (e.g. mutations causing cancer or cellular senescence), or indirect impacts via higher-level processes (e.g. impacts of inflammation on atherosclerosis). Higher levels of organization can also have their own independent mechanisms. This framework would help explain the diversity of ageing patterns across the tree of life, with both "public" and "private" mechanisms (Fülöp et al. 2019; Martin, 1989). Furthermore, this approach may also help to decipher health as the already mentioned mechanisms at various physiological levels (from molecular to organismal) may instruct of the deviation/variation/dynamics occurring in aging. Thus, the holistic and causal integration of the mechanisms/causes in a hierarchical concept would help to unravel a more complex overarching concept which is "health". Indeed, an integrative conceptualization of health is also badly needed, and the unraveling of the aging mechanisms may fundamentally contribute to this higher-level conceptualization of health (Sholl and Rattan, 2020).
We should acknowledge that numerous attempts exist to use artificial intelligence (AI) tools with the aim to better understand and conceptualize the complexity and emergent nature of ageing. Yet it seems that they still catch only the fragments of the causes of the whole dynamic ageing process. So far, AI tools still face the crucial limitation that they can be trained, but cannot truly understand, and thus may inadvertently be trained on the wrong phenomenon (Mitchell, 2020). Thus, developing and using new AI tools are crucially needed to deeply understand the causal relationship of the many data gathered on ageing, with the overarching goal to be able to understand the quintessence of ageing and consequently modulate it in a holistic manner.

Can ageing be understood?

Living cells (about 30 trillion in humans) are complex systems that encompass anatomical structures, information maintained by DNA, bioenergetics and information processing mediated by multi-omic subcellular and cellular machinery. For example, mitochondria within a cell exhibit what can only be called social behavior (Picard and Sandi, 2020). A disruption of the underlying cellular mechanisms contributes to ageing and diseases. A timely and significant question is whether it is possible to fully explain or dissect the biological processes of ageing from multi-omics data and other relevant biological markers? An affirmative answer to this question has far-reaching medical implications and the tantalizing prospect of partially reversing ageing and diseases that result from ageing. However, this question rapidly encounters non-trivial scientific hurdles, since biological structures and processes of living organisms are emergent and strongly coupled across different spatio-temporal scales (i.e. from microscale to macroscale) (Dumont and Prakash, 2014). By emergent, we mean that a macroscopic property (structure or process) is not merely the sum of its microscopic properties (e.g. water can be in the liquid state and can wet a cloth, but a single water molecule has no such macroscopic property). In the case of ageing, one is interested in explaining the macroscopic phenotypes in terms of the measured microscopic data (e.g. omics). Moreover, by strong coupling between scales, we mean that there are feedback coupling loops between scales (mostly unknown) in such a way that emergent properties (multiscale information) induce changes onto the microscopic properties and vice versa. As a consequence, dissecting

biological processes, in particular, those associated with ageing, into building blocks and studying these building blocks separately (i.e. via the traditional scientific reductionist method) is bound to fail. Novel scientific methods, particularly those that have a holistic (in the sense of systems biology) approach, are necessary to make sense of complex systems. These ideas, and in particular the extent to which one can dissect a complex system, are best understood in the context of a classification process, developed within the field of theoretical physics. A given system can be of one of the four types: *Strong reduction*, *Supervenience*, *Contextual emergence*, *Radical emergence* (see Fig. 1) (Butterfield, 2011; Bishop and Atmanspacher 2006; Bishop and Ellis 2020).

	microscale contains necessary conditions for ***macroscale***	***microscale*** contains sufficient conditions for ***macroscale***
Strong reduction	YES	YES
Supervenience	NO	YES
Contextual emergence	YES	NO
Radical emergence	NO	NO

Figure 1. Classification map of physical and biological systems from theoretical physics and philosophy of science point of view. There are mainly four types of systems, which are classified based on the amount of information available at the microscale to explain the macroscopic observations. For the case of ageing in biological systems, the class is unknown, but the aim is to describe the macroscopic observations of ageing in terms of the microscopic biological mechanisms.

The simplest class is the so-called *strong reduction,* in which a system can be fully dissected into building blocks. This is only possible if at the microscopic scale the system contains necessary and sufficient conditions (these rigorously expressed mathematically) to explain the macroscopic properties of the system. Ultimately, these conditions permit to dissect the system and explain it in terms of its elementary building blocks (i.e. the system is the sum of its parts). Examples within this class are thermodynamic closed systems, such as a container containing gas, where the external application of temperature (or manipulation of any thermodynamic parameters) leads to changes of the macroscopic state of the system to either gas, liquid or solid (see Fig 2, Panel A). These macroscopic states (and associated dynamics) can be fully understood (mathematically) in terms of the microscopic atomic activity. This understanding is possible due to the existence of necessary and sufficient physical and mathematical conditions (discovered in the field of Mathematical-Physics), namely: Boltzmann's *Propagation of molecular chaos*, *zeroth law of thermodynamics* and so-called *Kubo-Martin Schwinger (KMS) states* (Ehrenfest and Ehrenfest, 2002; Gottlieb, 2002; Villani, 2002; Haag et al., 1967; Kubo, 1957; Martin and Schwinger, 1959). We shall not delve into the technical details of these conditions, but the intuitive idea is that the interactions between the gas particles are elastic and dissipate (memoryless) and so the effect of each particle on the overall gas is negligible. As a consequence, it is possible to coarse grain and average (i.e. via a statistical distribution) the activity across various scales. This idea leads to the derivation of a *macroscopic time-evolution equation* and *thermodynamic observables* (i.e. functions that depend on thermodynamic variables: *pressure*, *volume*, *temperature*, *Avogadro constant*), which predict respectively, the time dynamics of the system and state phase transitions or tipping points (e.g.

changes for solid to liquid). Noteworthy, the macroscopic equation encodes a statistical distribution (specifically Maxwell-Boltzmann distribution) of atomic particles and the ultimate reason as to why it is possible to explain the macroscopic properties from the microscopic ones is because there is a one-to-one mapping between the five parameters of the distribution and five known conservation laws of physics. Changes between states (e.g. solid, liquid and gas) occur because the energy landscape and microscopic cloud of atoms radically reorganize (i.e. the overall topology or configuration of atoms in the space of possible positions and velocities change, see Fig 2. Panel B) as the thermodynamic parameter(s) vary. Under these radical changes, *thermodynamic observables*, exhibit a singularity (sharp changes; see Fig 2. Panel B). Moreover, the ageing process of such closed systems is fully described by the second law of Thermodynamics, which states that the entropy of the system always increases (see Fig 2. Panel B).

The second class, *supervenience*, corresponds to systems whereby the microscopic scale only contains sufficient conditions to explain macroscopic properties. The third class, *contextual emergence*, corresponds to systems in which the microscopic scale contains only necessary conditions to explain macroscopic properties. An interesting example of contextual emergence lies in the relation between topology and phase transitions in theoretical physics. In fact, under certain conditions, a topological change in the configuration space of a Hamiltonian system emerges as necessary conditions for the occurrence of phase transitions, but not sufficient (Franzosi and Pettini, 2004). The last class, *radical emergence*, are systems without sufficient and necessary conditions at the microscopic scale, and thus the macroscopic properties are purely emergent and impossible to explain. Presently, there is no consensus to which class life (biological systems) belongs to; see Fig 2. Panel B. These are open systems (e.g. cells interact with the external world and extract energy via the ATP system), thus conservation of laws of physics do not apply (at least at the microscale). Moreover, there are strong coupling across different spatio-temporal scales, and some processes seem emergent, then it makes it hard to study. However, the scientific hope is that these complex systems are mixed, either *supervenient* and/or *contextual emergent,* which would provide a basis to study life and ageing with some level of mathematical and scientific rigour. Examples associated to supervenience, contextual emergence and radical emergence are beyond the context of the current manuscript since the mathematical and physical conditions are still being worked for complex systems. However, some these systems may include certain classes of quantum systems, biological systems (e.g. the ability to program cells into stem cells) and in the extreme case the emergence of the Universe (perhaps being an example of Radical emergence).
A fundamental question is whether we can develop scientific tools that could shed light onto complex systems (Holovatch et al., 2017; Salnikov et al., 2018), even if it is only partially? The subsequent section discusses a novel Big Data tool, the so-called *Topological Data Analysis* (Carlsson, 2009; Rabadàn and Blumberg, 2019; Wasserman, 2018), which has the potential to substantially advance our understanding of complex systems (Petri et al., 2013), namely, in the context of the present manuscript, to ageing and diseases.

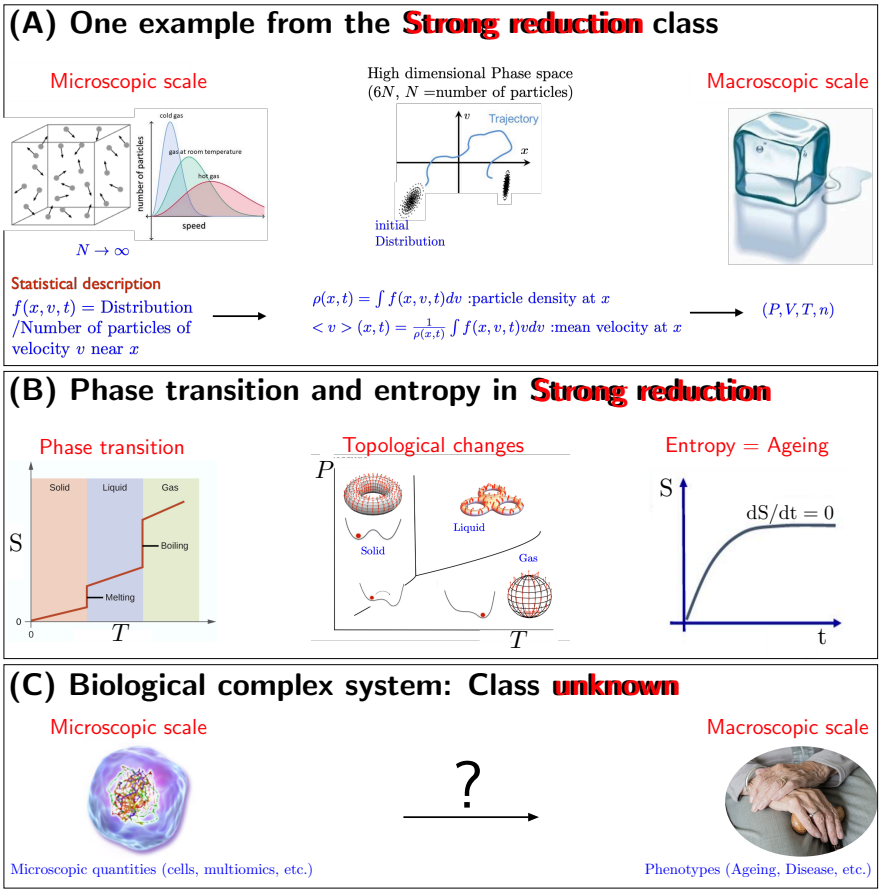

Figure 2. Strong reduction systems vs complex systems. A: depicts an example of a strong reduction system. The traditional point of view of the scientific approach is that systems can be broken up into micro building blocks and then the assembling of it provides an explanation of the entire system. This approach has been successful in explaining many physical systems, such as the depicted example of a closed thermodynamic system of a gas contained within a container. The macroscopic states of matter (solid, liquid, gas) can be predicted by dynamic equations and functions, which explain how the building blocks interact. B. Phase transitions of states of matter can also be predicted by rapid changes observed from thermodynamic functions (singularities), transitions in the energy functions in undergoing phase transitions. This processed is induced by necessary topological changes of the configuration of particles in the space of position and velocities. The ageing of these closed systems is simply described by Entropy. C: In contrast, the biological systems are complex, where some emergent features (or phenotypes) may not just be a sum of the building blocks. This occurs when there are strong feedback and coupling across spatial-temporal scales of the systems. Therefore, the class of biological systems is yet unknown.

Topological Data Analysis: A novel tool to study complex systems

Topological Data Analysis (TDA) is a relatively novel and revolutionary Big Data method, with origins in pure mathematics (Hatcher, 2002) and which has been recently further developed in applied fields, e.g. physics and computer science (Carlsson, 2009; Wasserman, 2018). In contrast to existing AI (Artificial Intelligence) (LeCun et al., 2013) and Big Data tools (Sagiroglu et al., 2013) that determine numerical summaries/patterns (yet do not take into account causality), TDA extracts knowledge from data in the form of shapes (technically *topological invariants)* and non-trivial relations between shapes. Specifically, it extracts from noisy multiscale high-dimensional data the intrinsic global structural shape (holistic descriptor) of a complex system (Blevins and Bassett, 2020). Since TDA has a rigorous theoretical foundation, it is ideally suited for extracting

complex causal interactions across spatio-temporal scales and consequently has the potential to establish non-trivial relations between microscale and macroscopic emergent processes. Thus, TDA has the potential of extending our knowledge of living systems and, in the context of the present manuscript, dissect (to some degree) the processes underlying ageing and associated diseases. The effectiveness of TDA has been shown in several applications, including biological and clinical data (Rabadán and Blumberg, 2019). To cite a few examples: 1. TDA applied to a database featuring multi-omics data of type-2 diabetes patients, discovered a third unknown form of diabetes (Li et al., 2015); 2. Applied to malaria medical records featuring transcriptome data and phenotypes like temperature, parasite density, etc., it tracked disease progression and characterized resilient patients (Torres et al., 2016); 3. Applied to clinical data of Asthma patients featuring physiological and inflammatory parameters it discovered key endotypes (Hinks et al., 2015; Hinks et al., 2016). As a consequence, it is envisaged that TDA will become a fundamental method for precision medicine.

To make sense of TDA, it is first important to note that measured data has an intrinsic shape (technically, it has a *topology*). This concept is exemplified in Fig. 3 (Panel A), where scattered data (i.e. *point cloud*, a distributed discrete collection of data points) encodes a pattern from which we may wish to decode and synthesize mathematically. However, there are infinitely many ways in which data can be distributed, and this raises the question of whether there exists a general algorithm for extracting these mathematical patterns? Traditional data analytical approaches implement specific algorithms to extract specific patterns within data, which are often sub-optimal (Torres et al., 2020). To illustrate this point, consider the following two examples: 1. If the data is scattered along a line, then a regression method that fits a line (i.e. a functional relationship between variables) is applied. The approach (algebraic modelling) is only suited for systems with low degrees of freedom (i.e. small number of variables, parameters and simple relationships). Moreover, algebraic fitting requires a high degree of approximations. This means the degree of understanding that the scientist extracts from the model (explanation of the mechanism behind the data) degrades for complex data. Moreover, this approach always requires an initial hypothesis (a kind of supervised analysis). 2. If the data is scattered into clusters, then a suitable clustering algorithm can be used. Unfortunately, there are many clustering algorithms due to the necessity of optimizing an objective function of some kind and, as a consequence, none of these works so well since there are no known optimal objective function. Beyond these two examples data can be distributed in more arbitrary ways (e.g. circle, flaring or more complicated ways), thus one quickly realizes that deriving a new algorithm for each data set is not tenable. Recently, Big Data and AI approaches (e.g. machine learning and network theory) attempt to address some of these deficiencies (e.g. unsupervised learning), but unfortunately these solutions are only partial and are unable to determine causal relations (e.g. machine learning) (Baker et al., 2018). TDA overcomes these deficiencies in four steps as follows: 1. It first defines what is meant by shape by equating it to topology rather than geometry (see Fig. 3 Panel B). Topological objects are those that are invariant (or remain unchanged) under continuous deformations/stresses (i.e. elastic deformation like a rubber band that is stretched but not torn apart). For example, consider a coffee mug (see Fig 3 panel B), which has one hole in the handle of the mug. Consider also a table plate, which has no hole. These two objects are topologically different since it is impossible to continuously deform (elastically) a coffee mug and transform it into a table plate (and vice versa). However, it is possible to continuously deform a coffee mug into a torus (e.g. donut) since both objects have a hole. Thus, these objects are Topologically the same because they have the same invariant, in particular, the same number of holes. To make this point clear, see the second example (see Fig 3 panel B). Although the two bowls have the same geometric appearance, they are topologically different

because one of the bowls has numerous holes. Thus, it is impossible to continuously deform one into the other. 2. Second, TDA notes that a complex system can be associated with a global structure (global shape or global topological object, technically called *simplicial complex*) (Salnikov et al., 2018). This structure encodes data that can be decomposed into canonical patterns (canonical topological objects, technically called *simplices*); see Fig 3 Panel C. In analogy, these canonical topological objects can be seen as the basic constituents of a system, such as atoms within the periodic table (in chemistry), or the fundamental musical chords from which any music can be composed. The simplest canonical object is a single data point (denoted the *0-simplex*). Subsequently, a pair-wise relation between two data points is denoted a 1-simplex. The study of 1-simplices is essentially network or graph theory. That is, graph theory is often built upon pair-wise relations only (Battiston et al., 2020)! The next level of complexity, the *2-simplex*, determines the case when three data points are simultaneously participating in a causal relationship, from which emerges a new object (represented as the face of a triangle). An example of such a scenario is the emergence of a chemical clock/oscillations when three chemical compounds interact (i.e. Iodine, Thiosulfate and Triiodide); other forms of oscillations could be found in the neural context (Desroches et al., 2016). Clearly, the 2-simplex is already beyond the standard network and graph theory! Higher-order simplices (n-simplex) (Lambiotte et al., 2019; Millán et al., 2020; Salnikov et al., 2018) follow the same rationale, which is to represent the emergence of novel higher-order objects (see Fig. 3 Panel C). Finally, an arbitrary *simplicial complex* (i.e. a global topological object) can be formed by combining suitable simplices (see Fig. 3 Panel C). 3. Crucially, topological objects (i.e. simplices or simplicial complex) satisfy three fundamental properties, namely, *Coordinate Invariant, Deformation Invariant,* and *Compressed representation* (Carlsson, 2009) (see Fig. 4 Panel A) and have an associated unique signature, technically, *topological invariance* (see Fig. 4 Panel B). *Coordinate invariant* means that the topological properties of the data are independent of the coordinate system. This property is in contrast to a large class of Big Data techniques (e.g. Principal Component Analysis). *Deformation invariant* is a by-product of the aforementioned definition of a topological object, which states that they are invariant (or remain unchanged) under continuous deformations (i.e. elastic deformation like a rubber band that is stretched but not torn apart). This property makes TDA robust to noise, in the sense that small perturbations of the data will not change the underlying topological properties of the object encoded within the data (Blevins and Bassett, 2020). *Compressed representation* means that a given topological object can be represented by a minimal number of data points, as long as the data points do not change the topology of the object. *A topological invariant* is a unique signature or measure that quantifies the underlying pattern (topological object) within the data. The idea of topological invariance can simply be understood via the example shown in Fig. 4 Panel B. As shown, the four platonic solids (tetrahedron, octahedron, cube, icosahedron) and a sphere, although possessing different geometries, are essentially the same topological object (i.e. one can deform elastically one into another and indeed the four platonic solids are compressed representations of the sphere). The equivalence between these objects is quantified via the *Euler characteristic* (in its simpler form equates to vertices – edges + faces; see Fig. 4 Panel B), which provides the topological invariance. To clarify, consider the Tetrahedron, which has 4 vertices (i.e. the data points), 6 edges (relation between data points), 4 faces (higher-order relations); the corresponding Euler characteristic is 2. Equally, applying the Euler characteristic's formula to the octahedron, the cube or the icosahedron, also yields in *Euler characteristics* equal to 2. Moreover, the sphere also has a *Euler characteristic* of 2. This can be understood by the alternative formula in Fig. 4 Panel B, which counts the number of holes (in the case of a sphere it is zero). Alternatively, one can imagine a sphere being cut in half through a *great circle* that goes through the north and south pole (i.e. 2 vertices), which forms

two *meridians* (i.e. 2 edges) and split the surface of the earth into two half-spheres (i.e. 2 faces). Therefore, the *Euler characteristic* is indeed 2. This short exercise emphasizes the importance of identifying topological invariances within data (a fundamental signature), which gives us confidence about the underlying pattern within data even under noise perturbations. In the exercised example, adding or removing data points (i.e. noise perturbations) to the sphere will only result into either a smoother sphere or a platonic solid, all of these having the same invariant measure, hence equivalent. More generally, the *Euler characteristic* is an alternating sum of simplices or, equivalently, holes (see Fig. 4 Panel B). Consequently, a given simplicial complex is a composition of its constituent building blocks (e.g. simplices and holes) quantified by the *Euler characteristic* (see example in Fig. 4 Panel B). Simplicial complexes are in a way discrete analog of multidimensional surfaces. Since surfaces have multidimensional holes, simplicial complexes also do so (Zomorodian, 2005). 4. The fourth and final step develops a computational machinery to search (or extract) simplicial complexes from a distributed discrete collection of data points. There are different methods and approaches in TDA, to name a few, *Persistent homology*, *Reeb Graph*s, *Mapper algorithm* (Singh et al., 2007; Geniesse et al., 2019). Herein we will only discuss the method of *Persistent homology* since it is general enough to have an idea of TDA and has successfully been applied to biological data. The first step in persistent homology is to consider a similarity measure between data points (e.g. correlation, causality measures, synchronization measures, etc.). The appropriate choice of the similarity measure will depend on the data. At this level the experience (and knowledge) of the scientist collecting the data (e.g. clinician or biologist) is valuable as it will critically guide an exchange of information (and collaboration) with experts in TDA. Subsequently, the emergence of simplicial complexes within the data follows a procedure (algorithm) called *Filtration* (see Fig. 4 Panel C), which consists of the following: Take as input a *point cloud* (a discrete collection of data points) and for each data point consider a ball of radius ε (defined with respect to the chosen similarity measure), centered at this data point. Subsequently, continuously increase the radius (ε) and monitor the intersections between the different balls. For every intersection between n-balls, create an n-simplex (see Fig. 4 Panel C). As the radius (ε) further increases, it generates a nested sequence of *simplicial complexes* for which the associated *Euler characteristic* (or any other topological property) is computed (e.g. all types of holes are counted in accordance to the general formula shown in see Fig. 4 Panel B). Since the counting (e.g. of each type of hole) is performed for every radius (ε) along the generated sequence of *simplicial complexes,* a convenient way to represent this counting process is via the so-called *barcodes* (see Fig. 4 Panel C) (Ghrist, 2008). Specifically, a *barcode* (with finite length) represents the emergence, persistence and disappearance of a given hole along a filtration process and the set of *barcodes* encodes the entire topological signature of the data. The longest *barcodes* (those the persist the longest, therefore robust) are considered the true topological features of the data, and the short ones represent noise within the data. The advantage of *Persistent homology* is threefold: 1. The set of barcodes is a *complete* topological invariant, meaning that it captures all the topological information of a filtration (i.e. a meta signature of the data across scales); 2. It is computationally efficient; 3. It has *stability* properties, meaning that small perturbation of the data produces small perturbation of the real topological invariant. Indeed, small holes are usually associated with noise, while big holes are considered to be real; hence, small pertubartions does not change the big holes much. Finally, TDA can be used in combination with other Big Data methods, should further processing be required. That is, the topological invariants can be further processed for example with machine learning and statistical methods (Bergomi et al., 2019).

Topological Data Analysis of ageing data

To the best of our knowledge, TDA has not been applied to the research field of ageing, although it has been applied in other biological contexts (e.g. omics data) (Rabadán and Blumberg, 2019). Therefore, this opens a unique opportunity to add TDA to the arsenal of methods, which in tandem would unveil unforeseen information within the complex high-dimensional spatio-temporal ageing data. Some TDA methods (e.g. persistent homology) are computationally efficient and scalable and thus it is ideally suited for extracting topological invariants (e.g. associated to evolutionary conserved mechanisms) in longitudinal ageing data across a population. Moreover, it has the potential to identify differences among patients, since TDA would associate. e.g., a *barcode* to each patient. Indeed, conserved mechanisms and differences can be studied by well-defined *distance measures for barcodes*. To achieve this, we propose the TDA computational pipeline as shown in Fig 5. Panel A and a suitable similarity measure will have to be carefully developed. The pipeline will input clinically relevant state-of-the-art ageing data (e.g. high-throughput multi-omics) together with other unknown biomarkers and clinical phenotypes of single patients or of a population. Subsequently, the TDA machinery will render a sequence of simplicial complexes (and topological invariants) across scales. This has potential to unveil topological relationships across multi-omics, other relevant biomarkers and clinical phenotypes. Moreover, based on recent theoretical extensions of TDA, the topological phase transitions (tipping points, also referred to as bifurcations in other contexts (Desroches et al., 2013; Rodrigues et al., 2016)) within the data can potentially be detected (Amorim et al., 2020; Santos et al., 2009; Santos et al., 2014; Santos et al., 2017; Santos et al., 2019); see Fig 5 panel B. We will not delve into the details of the theoretical machinery that enables the extraction these topological phase transitions in complex systems, however, it has some analogy with phase transitions described of strong reduction systems. As a consequence, ageing evolution and transition (as well as disease evolution and transition) points could be tracked via topological invariants, which would result in an *ageing map* (see Fig 5 panel B). The envisaged map has the potential of predicting trajectories of healthy ageing as well as mechanistically explaining multiscale biological mechanisms, phase transitions from healthy ageing and suggest treatments. Ultimately, the goal achievable by TDA will be to understand the *shape of ageing and age-related diseases*.

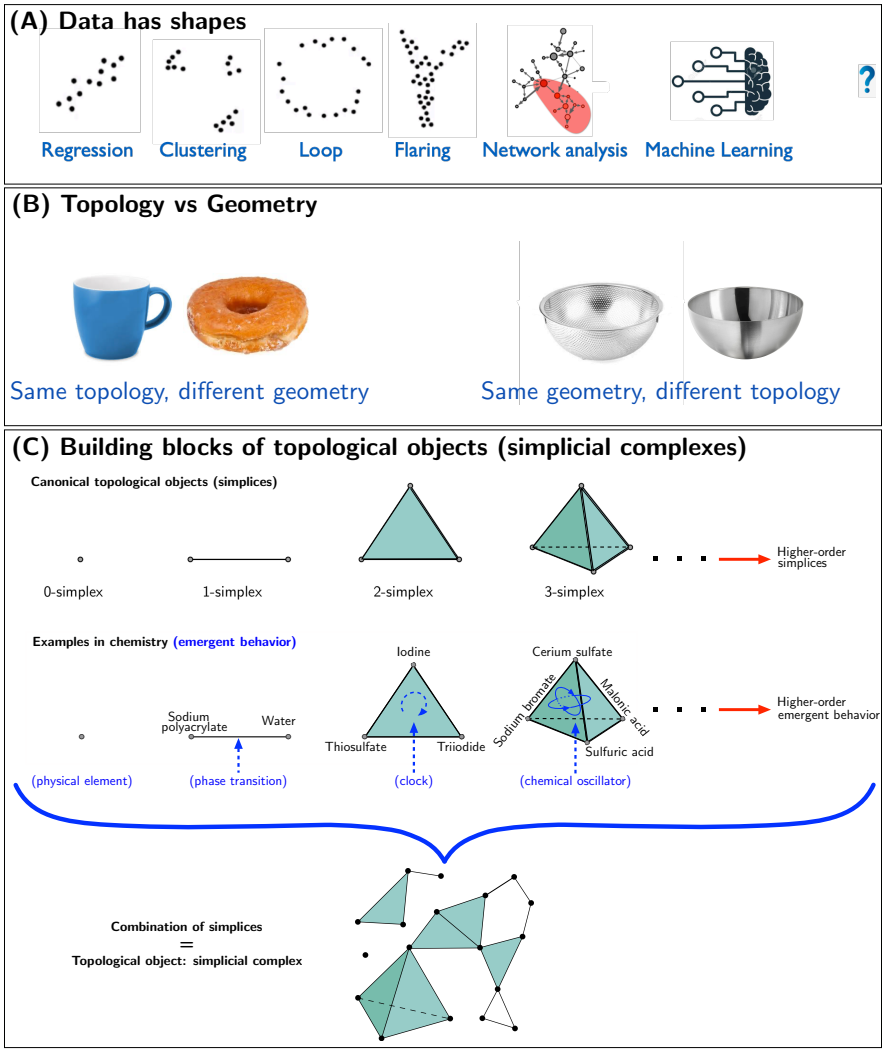

Figure 3. A: In general data has an intrinsic shape (pattern) that encodes useable information. Different methods have been developed but somewhat limited. Thus, the question is if there is a more general way to mine useable information. B: There is a difference between Geometry and Topology (although they are related). Geometry is concerned with lengths and angles, while topology is concerned with shapes that remain invariant under elastic deformation. C: Topological objects (simplicial complexes) are constructive and built from building blocks called simplices. Simplices in general represent higher order relations (e.g. emergent properties). For example, 2-simplex is the face a triangle, which explains the emergence of a chemical clock after mixing three chemical elements (i.e. data points).

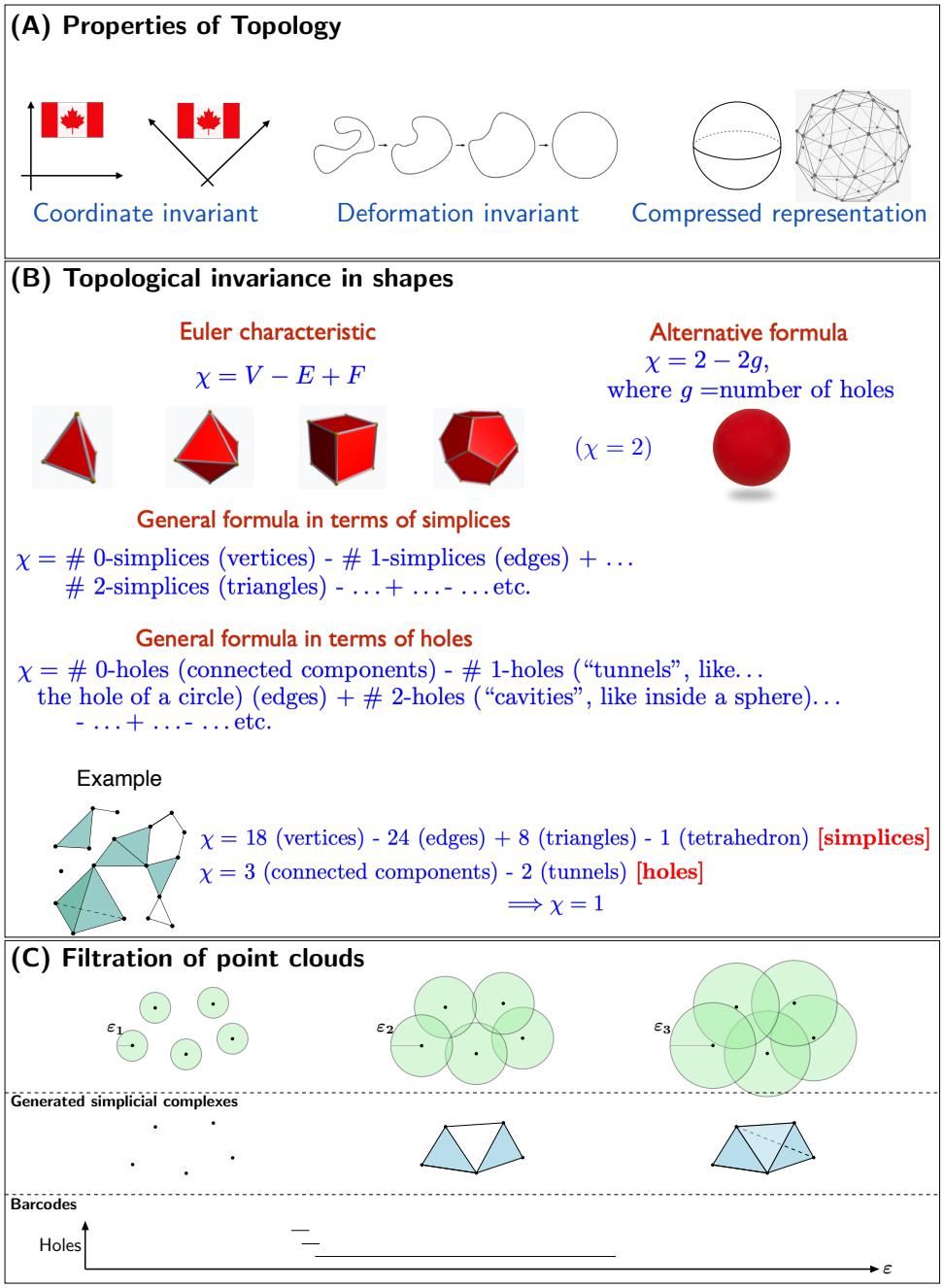

Figure 4. A: Three fundamental properties of Topology. B: Platonic solids have the same topological invariance as a sphere. Specifically, for the above examples we have the following. Tetrahedron has 4 vertices, 6 edges and 4 faces, which equates to Euler Characteristic (EC) = 2. An Octahedron has 6 vertices, 12 edges and 8 faces, hence EC=2. A cube has 8 vertices, 12 edges and 6 faces, hence EC=2. A Dodecahedron has 20 vertices, 30 edges and 12 faces, hence EC=2. Equally the Sphere has EC = 2. The Euler characteristics of a simplicial complex is also constructive as shown in the example. C: Persistent homology is one important TDA algorithm that mines topological patterns from data and the algorithm uses a process called filtration of point clouds. The filtration consists in positioning balls with centre at the data points and with a certain radius (defined with respect to some similarity measure) and continuously expand these balls (i.e. increment the radius). The intersection of the balls generates simplicial complexes and the topological invariances can be tracked with barcodes.

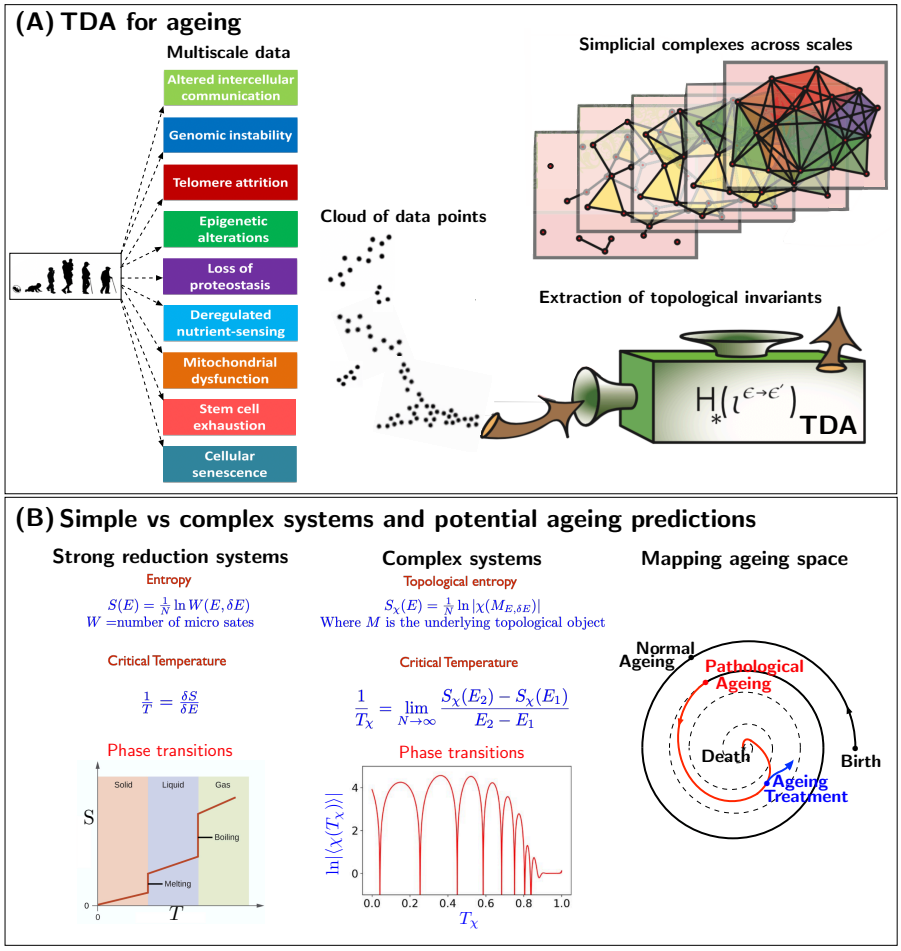

Figure 5. A: proposed TDA computational pipeline to mine topological invariances in complex high-dimensional spatial-temporal ageing data. The analysis can be done across a heterogenous populations and the output will reveal simplicial complexes across scales and associated topological invariances, from which relevant clinical and biological information can be drawn. B: A comparison between Strong reduction system and complex systems. Recent theoretical findings show that Topological entropies can detect multiple phase transitions in complex data. These findings could be used to derive a map that predicts normal ageing and associated multiscale biological mechanisms as well as phase transitions (e.g. deviations from normal ageing). Moreover, it could potentially suggest ageing treatments.

Discussion

The complexity of ageing calls for novel sophisticated data analytical methods that make the fewest assumptions, yet extract global patterns while establishing causal links between observed and latent variables from multiscale ageing data sets. In recent years TDA has emerged as a powerful tool that meets these aforementioned criteria and therefore we encourage the scientific community interested in ageing to incorporate TDA as a novel toolbox in their repertoire methods. Indeed, the conceptual idea of TDA in analyzing data via topological means is in stark contrast to conventional numerical summaries made by other Big Data/AI methods, which typically make assumptions on the data and are typically incomplete in the way causal links are established between variables across scales. TDA extracts topological objects (simplicial complexes) and topological invariants from data, which in itself is novel information but also can be used as complementary information to that of numerical summaries made by other Big Data/AI methods. Thus, we expect that the

combination of numerical summaries (provided by other methods such as statistics, network theory, machine learning etc.) and topological summaries will lead to novel unprecedented insights on ageing. Given the fast development of TDA and its numerous recent achievements in various biological studies one can only advocate for this method to be integrated in the toolbox that every scientist working on ageing should have.

However, a word of caution should be expressed. TDA should not be perceived as the ultimate tool to analyse complex biological data since the concept of an ultimate tool is probably illusory. Furthermore, TDA comes with difficulties, in particular, finding the appropriate *similarity measure* suitable to a given data set is sometimes a daunting task and very much dependent on the knowledge of the problem being analysed. Noteworthy, topological approaches in theoretical physics usually give necessary conditions, not sufficient ones (Franzosi and Pettini 2004). In other words, topology falls (most of the time) within the class of *Contextual Emergence*, since it is necessary that two structures that are equivalent must have the same topological properties, but the converse is not true: two structures with same (fixed) topological property could differ in relation to another one. In this sense, we hope that future studies of complex biological systems and in particular ageing will reveal that the underlying mechanisms fall under the class of *Contextual Emergence.* Provided ageing falls under Context Emergence, we envisage that TDA will help to create an *ageing map* that predicts the biological mechanisms underpinning pathological vs healthy ageing scenarios. These phase transitions could be unveiled in longitudinal studies but also in transversal studies across databases, which would compare differences in the structure of those phase transitions. In other words, TDA has the potential to uncover the shape of ageing and age-related diseases.

To conclude, TDA extracts global topological invariant patterns and it will enable scientists and clinicians to collaborate in synergy (at the level of systems biology) and thus significantly contribute to the worldwide initiative on healthy ageing (WHO 2020-2030)[1] for unraveling the causal relationships of the "true" complex, dynamic, hierarchical, heterogenic and emergent nature of ageing.

Acknowledgement

The works presented in the article were supported by grants from Canadian Institutes of Health Research (CIHR) (No. 106634) and No. PJT-162366) to TF, the Société des médecins de l'Université de Sherbrooke and the Research Center on Aging of the CIUSSS-CHUS, Sherbrooke and the FRQS Audace grant; by FACEPE through the PRONEX Program (APQ-0602-1.05/14) (Brazilian agency) to FANS; by Ikerbasque (The Basque Foundation for Science), partially supported by the Basque Government under the grant "Artificial Intelligence in BCAM number EXP. 2019/00432", by the Basque Government through the BERC 2018-2021 program and by the Ministry of Science, Innovation and Universities: BCAM Severo Ochoa accreditation SEV-2017-0718 and through project RTI2018-093860B-C21 funded by (AEI/FEDER, UE) with acronym ``MathNEURO'' to SR; by Inria through Associated Team NeuroTransSF to MD and SR. A.A.C. is supported by a CIHR New Investigator Salary Award and *Fonds de recherche du Québec – Santé*

[1] https://www.who.int/ageing/decade-of-healthy-ageing

(FRQ-S) Senior Fellowship, and is a member of the *FRQ-S* funded *Centre de recherche du CHUS* and *Centre de recherche sur le vieillissement*.

References

Amorim E, Moreira RA, Santos FA. The Euler Characteristic and Topological Phase Transitions in Complex Systems. BioRxiv. 2019 Jan 1:871632.

Baker RE, Peña JM, Jayamohan J, Jérusalem A. Mechanistic models versus machine learning, a fight worth fighting for the biological community? Biol Lett. 2018;14(5):20170660.

Battiston F, Cencetti G, Iacopini I, Latora V, Lucas M, Patania A, Young JG, Petri G. Networks beyond pairwise interactions: structure and dynamics. Phys Rep. 2020 Jun 13 (online first). Doi: https://doi.org/10.1016/j.physrep.2020.05.004

Belsky DW, Moffitt TE, Cohen AA, Corcoran DL, Levine ME, Prinz JA, Schaefer J, Sugden K, Williams B, Poulton R, Caspi A. Eleven Telomere, Epigenetic Clock, and Biomarker-Composite Quantifications of Biological Aging: Do They Measure the Same Thing? Am J Epidemiol. 2018;187(6):1220-1230.

Bergomi MG, Frosini P, Giorgi D, Quercioli N. Towards a topological–geometrical theory of group equivariant non-expansive operators for data analysis and machine learning. Nat Mach Intell. 2019;1(9):423-33.

Bishop R.C., Atmanspacher H. Contextual emergence in the description of properties. Found Phys. 2006; 36:1753–1777.

Bishop R.C., Ellis G.F.R. Contextual emergence of physical properties. Found Phys. 2020;50:481-510.

Bair WN, Petr M, Alfaras I, Mitchell SJ, Bernier M, Ferrucci L, Studenski SA, De Cabo R. Of Aging Mice and Men: Gait Speed Decline Is a Translatable Trait, With Species-Specific Underlying Properties. J Gerontol A Biol Sci Med Sci. 2019 Aug 16;74(9):1413-1416. doi: 10.1093/gerona/glz015.

Blevins AS, Bassett DS. Reorderability of node-filtered order complexes. Phys Rev E. 2020;101(5):052311.

Bulterijs S, Hull RS, Björk VCE, Roy AG. It is time to classify biological aging as a disease. Front Genet. 2015;6:205.

Butterfield J. Emergence, reduction and supervenience: A varied landscape. Found Phys. 2011; 41:920-960.

Cannon ML. What is aging? Dis Mon. 2015;61(11):454-9. doi:10.1016/j.disamonth.2015.09.002.

Carlsson G. Topology and data. Bull Amer Math Soc. 2009;46(2):255-308.

Chmielewski PP. Human ageing as a dynamic, emergent and malleable process: from disease-oriented to health-oriented approaches. Biogerontology. 2020;21(1):125-130.

Cohen AA, Legault V, and Fülöp T. What if there's no such thing as "aging"? Mech. Age. Dev. 2020 (submitted)

Cohen AA, Kennedy BK, Anglas U, Bronikowski AM, Deelen J, Dufour F, Ferbeyre G, Ferrucci L, Franceschi C, Frasca D, Friguet B, Gaudreau P, Gladyshev VN, Gonos ES, Gorbunova V, Gut P, Ivanchenko M, Legault V, Lemaître JF, Liontis T, Liu GH, Liu M, Maier AB, Nóbrega OT, Rikkert MGMO, Pawelec G, Rheault S, Senior AM, Simm A, Soo S, Traa A, Ukraintseva S, Vanhaelen Q, Van Raamsdonk JM, Witkowski JM, Yashin AI, Ziman R, Fülöp T. Lack of consensus on an aging biology paradigm? A global survey reveals an agreement to disagree, and the need for an interdisciplinary framework. Mech Ageing Dev. 2020;190:111316.

Cohen AA. Complex systems dynamics in aging: new evidence, continuing questions. Biogerontology. 2016;17(1):205-20.

Cohen AA, Li Q, Milot E, Leroux M, Faucher S, Morissette-Thomas V, Legault V, Fried LP, Ferrucci L. Statistical distance as a measure of physiological dysregulation is largely robust to variation in its biomarker composition. PLoS One. 2015;10(4):e0122541.

da Costa JP, Vitorino R, Silva GM, Vogel C, Duarte AC, Rocha-Santos T. A synopsis on aging-Theories, mechanisms and future prospects. Ageing Res Rev. 2016;29:90-112. doi: 10.1016/j.arr.2016.06.005.

Desroches M, Krupa M and Rodrigues S. Inflection, canards and excitability threshold in neuronal models. J Math Biol. 2013;67(4):989-1017.

Desroches M, Guillamon A, Ponce E, Prohens R, Rodrigues S and Teruel AE. Canards, folded nodes, and mixed-mode oscillations in piecewise-linear slow-fast systems. SIAM Rev. 2016;58(4):653-91.

Dumont, S., Prakash, M. Emergent mechanics of biological structures. Mol Biol Cell. 2014; 25(22), 3461–3465.

Ehrenfest P, Ehrenfest T, The Conceptual Foundations of the Statistical Approach in Mechanics. Cornell Univ Press, 1959.

Franceschi C, Salvioli S, Garagnani P, de Eguileor M, Monti D, Capri M. Immunobiography and the Heterogeneity of Immune Responses in the Elderly: A Focus on Inflammaging and Trained Immunity. Front Immunol. 2017;8:982. doi: 10.3389/fimmu.2017.00982.

Fülöp T, Munawara U, Larbi A, Desroches M, Rodrigues S, Catanzaro M, Guidolin A, Khalil A, Bernier F, Barron AE, Hirokawa K, Beauregard PB, Dumoulin D, Bellenger J-P, Witkowski JM

and Frost E. Targeting Infectious Agents as a Therapeutic Strategy in Alzheimer's Disease. CNS Drugs 2020a; 34:673-695.

Fülöp T, Larbi A, Hirokawa K, Cohen AA, Witkowski JM. Immunosenescence is both functional/adaptive and dysfunctional/maladaptive. Semin Immunopathol. 2020b Sep 15.

Fülöp T, Larbi A, Khalil A, Cohen AA, Witkowski JM. Are We Ill Because We Age? Front Physiol. 2019;10:1508.

Fülöp T, Larbi A, Dupuis G, Le Page A, Frost EH, Cohen AA, Witkowski JM, Franceschi C. Immunosenescence and Inflamm-Aging As Two Sides of the Same Coin: Friends or Foes? Front Immunol. 2018;8:1960. doi: 10.3389/fimmu.2017.01960.

Fülöp T, Dupuis G, Baehl S, Le Page A, Bourgade K, Frost E, Witkowski JM, Pawelec G, Larbi A, Cunnane S. From inflamm-aging to immune-paralysis: a slippery slope during aging for immune-adaptation. Biogerontology. 2016;17(1):147-57.

Franzosi R, Pettini M. Theorem on the origin of phase transitions. Phys Rev Lett. 2004;92(6):060601.

Geniesse C, Sporns O, Petri G, Saggar M. Generating dynamical neuroimaging spatiotemporal representations (DyNeuSR) using topological data analysis. Network Neuroscience. 2019;3(3):763-78.

Ghrist, R. Barcodes: the persistent topology of data. Bull Amer Math Soc. 2008;45(1):61-75.

Gottlieb AD. Propagation of Chaos in Classical and Quantum Kinetics. In: Rebolledo R (ed.) Stochastic Analysis and Mathematical Physics II. Trends in Mathematics. Birkhäuser, Basel. 2003.

Haag R.,Winnink, M., Hugenholtz, NM. On the equilibrium states in quantum statistical mechanics. Commun Math Phys. 1967; 5(3): 215–236

Hatcher A. Algebraic Topology, Cambridge University Press, 2002.

Hinks TS, Brown T, Lau LC, Rupani H, Barber C, Elliott S, Ward JA, Ono J, Ohta S, Izuhara K, Djukanović R. Multidimensional endotyping in patients with severe asthma reveals inflammatory heterogeneity in matrix metalloproteinases and chitinase 3–like protein 1. J Allergy Clin Immunol. 2016;138(1):61-75.

Hinks T, Zhou X, Staples K, Dimitrov B, Manta A, Petrossian T, Lum P, Smith C, Ward J, Howarth P, Walls A. Multidimensional endotypes of asthma: topological data analysis of cross-sectional clinical, pathological, and immunological data. The Lancet. 2015;385:S42.

Holovatch Y, Kenna R, Thurner S. Complex systems: physics beyond physics. European Journal of Physics. 2017;38(2):023002.

Horvath S. DNA methylation age of human tissues and cell types. Genome Biol. 2013;14(10):R115.

Horvath S, Raj K. DNA methylation-based biomarkers and the epigenetic clock theory of ageing. Nat Rev Genet 2018;19(6):371-384.

Hugenholtz F, de Vos WM. Mouse models for human intestinal microbiota research: a critical evaluation. Cell Mol Life Sci. 2018 Jan;75(1):149-160. doi: 10.1007/s00018-017-2693-8.

Kennedy BK, Berger SL, Brunet A, Campisi J, Cuervo AM, Epel ES, Franceschi C, Lithgow GJ, Morimoto RI, Pessin JE, Rando TA, Richardson A, Schadt EE, Wyss-Coray T, Sierra F. Geroscience: linking aging to chronic disease. Cell. 2014;159(4):709-13.

Kivelä, M., Arenas, A., Barthelemy, M., Gleeson, J.P., Moreno, Y. and Porter, M.A., 2014. Multilayer networks. J Compl Netw 2014;2(3):203-271.

Kubo R. Statistical-Mechanical Theory of Irreversible Processes. I. General Theory and Simple Applications to Magnetic and Conduction Problems. J. Phys. Soc. Jap. 1957;12(6): 570–586.

Lambiotte R, Rosvall M, Scholtes I. From networks to optimal higher-order models of complex systems. Nat Phys. 2019;15(4):313-20.

LeCun Y, Bengio Y, Hinton G. Deep learning. Nature. 2015;521(7553):436-44.

Li L, Cheng WY, Glicksberg BS, Gottesman O, Tamler R, Chen R, Bottinger EP, Dudley JT. Identification of type 2 diabetes subgroups through topological analysis of patient similarity. Sci. Transl. Med. 2015;7(311):311ra174.

Li X, Ploner A, Wang Y, Magnusson PK, Reynolds C, Finkel D, Pedersen NL, Jylhävä J, Hägg S. K'Longitudinal trajectories, correlations and mortality associations of nine biological ages across 20-years follow-up. Elife. 2020;9:e51507. doi: 10.7554/eLife.51507.

Lipsky MS, King M. Biological theories of aging. Dis Mon. 2015;61(11):460-6. doi: 10.1016/j.disamonth.2015.09.005.

López-Otín C, Blasco MA, Partridge L, Serrano M, Kroemer G. The hallmarks of aging. Cell. 2013;153(6):1194-217.

Martin GM. Genetic modulation of the senescent phenotype in Homo sapiens. Genome. 1989;31(1):390-7.

Martin, PC and Schwinger, J. Theory of Many-Particle Systems. I. Phys Rev. 1959;115(6):1342–1373.

Maynard C, Weinkove D. The Gut Microbiota and Ageing. Subcell Biochem. 2018;90:351-371. doi: 10.1007/978-981-13-2835-0_12.

Millán AP, Torres JJ, Bianconi G. Explosive higher-order Kuramoto dynamics on simplicial complexes. Phys Rev Lett. 2020;124(21):218301.
Milot E, Morissette-Thomas V, Li Q, Fried LP, Ferrucci L, Cohen AA. Trajectories of physiological dysregulation predicts mortality and health outcomes in a consistent manner across three populations. Mech Ageing Dev. 2014;141-142:56-63.

Mitchell M. On Crashing the Barrier of Meaning in Artificial Intelligence. AI Mag. 2020;41(2):86-92.

Mitnitski AB, Rutenberg AD, Farrell S, Rockwood K. Aging, frailty and complex networks. Biogerontology. 2017;18(4):433-446.

Moskalev A. The challenges of estimating biological age. Elife. 2020;9:e54969.

Petri G, Scolamiero M, Donato I and Vaccarino F. Topological strata of weighted complex networks. PloS one 2013;8(6):e66506.

Picard M. and Sandi C. The social nature of mitochondria: Implications for human health. Neurosci Biobehav Rev. 2020; https://doi.org/10.1016/j.neubiorev.2020.04.017

Rabadán R, Blumberg AJ. Topological Data Analysis for Genomics and Evolution: Topology in Biology. Cambridge University Press; 2019.

Rehn JA, Santos FA, Coutinho-Filho MD. Combinatorial and Topological Analysis of the Ising Chain in a Field. Brazilian Journal of Physics. 2012;42(5-6):410-21.

Rodrigues S, Desroches M, Krupa M, Cortes JM, Sejnowski TJ, Ali AB. Time-coded neurotransmitter release at excitatory and inhibitory synapses. Proc Natl Acad Sci USA 2016;113(8):E1108-15.

Rose MR, Flatt T, Graves JL, Greer LF, Martinez DE, Matos M, Mueller LD, Shmookler Reis RJ, Shahrestani P. What is Aging? Front Genet. 2012;3:134. doi: 10.3389/fgene.2012.00134.

Sagiroglu S, Sinanc D. Big data: A review. In: 2013 international conference on collaboration technologies and systems (CTS); IEEE, 2013, pp. 42-47.

Salnikov V, Cassese D, Lambiotte R. Simplicial complexes and complex systems. Eur J Phys. 2018;40(1):014001.

Santos FA, Raposo EP, Coutinho-Filho MD, Copelli M, Stam CJ, Douw L. Topological phase transitions in functional brain networks. Phys Rev E. 2019;100(3):032414.

Santos FA, da Silva LC, Coutinho-Filho MD. Topological approach to microcanonical thermodynamics and phase transition of interacting classical spins. J Stat Mech: Theory Exp. 2017;2017(1):013202.

Santos FA, Rehn JA, Coutinho-Filho MD. Topological and geometrical aspects of phase transitions. In: Perspectives and Challenges in Statistical Physics and Complex Systems for the Next Decade, World Scientific 2014; pp. 153-180.

Santos FA, Coutinho-Filho MD. Topology, symmetry, phase transitions, and noncollinear spin structures. Phys Rev E. 2009;80(3):031123.

Sasaki K, Bruder D, Hernandez-Vargas EA. Topological data analysis to model the shape of immune responses during co-infections. Commun. Nonlinear Sci. 2020;85:105228.

Sholl J. Rattan SIS. Explaining health across the Sciences. Ed. Sholl J. Rattan SIS. Healthy Aging and Longevity 12, Springer Nature, Switzerland, 2020, pp. 1-557

Singh G, Mémoli F, Carlsson GE. Topological methods for the analysis of high dimensional data sets and 3d object recognition. SPBG. 2007;91:100.

Torres BY, Oliveira JH, Thomas Tate A, Rath P, Cumnock K, Schneider DS. Tracking resilience to infections by mapping disease space. PLoS biol. 2016;14(4):e1002436.

Torres L, Blevins AS, Bassett DS, Eliassi-Rad T. The why, how, and when of representations for complex systems. arXiv preprint arXiv:2006.02870. 2020 Jun 4.

Villani C. A review of mathematical topics in collisional kinetic theory. In: Handbook of mathematical fluid dynamics, Vol. 1, S. Friedlander and D. Serre, Eds., North-Holland, 2002; pp. 71—305

Wagner W. Epigenetic aging clocks in mice and men. Genome Biol. 2017 Jun 14;18(1):107. doi: 10.1186/s13059-017-1245-8.

Wasserman L. Topological data analysis. Ann Rev Stat Appl. 2018;5:501-32.

Whitwell HJ, Bacalini MG, Blyuss O, Chen S, Garagnani P, Gordleeva SY, Jalan S, Ivanchenko M, Kanakov O, Kustikova V, Mariño IP, Meyerov I, Ullner E, Franceschi C, Zaikin A. The Human Body as a Super Network: Digital Methods to Analyze the Propagation of Aging. Front Aging Neurosci. 2020;12:136.

Zhavoronkov A, Mamoshina P, Vanhaelen Q, Scheibye-Knudsen M, Moskalev A, Aliper A. Artificial intelligence for aging and longevity research: Recent advances and perspectives. Ageing Res Rev. 2019;49:49-66. doi: 10.1016/j.arr.2018.11.003.

Zomorodian, AJ. Topology for computing. 2005, Cambridge University Press.