Evidence of causality between Closure and Open-Ended Evolution in the Kauffman model

M. Faggian[a,b,c]

[a] *Centre de Recherche Biodiversite et Evolution (CRBE) University of Toulouse France*
[b] *Centre National de la Recherche Scientifique (CNRS) France*
[c] *5 Degres Toulouse*

Abstract

Novelty creation is one of the key features that define evolutionary biological systems that exhibit Open-Ended Evolution (OEE). Recently developed theories support the idea that OEE cannot occur in the absence of organisational closure over a characteristic timescale, and that such closure is the ultimate cause of OEE. In this paper, we use assembly theory to study the effects of the emergence of autocatalytic networks on the dynamics of complexity in the Kauffman model, and provide evidence for the causal relationship between organisational closure and the increase in complexity in the sense of Open-Ended Evolution. Interpreting auto-catalytic cycles as a representation of functional closure in chemical reaction systems, our results provide a numerical support for evidence of a causal relationship between functional closure and OEE. In order to provide stronger evidence for this conjecture, in the last part of this paper we study the effects on the complexity dynamics provoked by the simplest autocatalytic set in the Kauffman model in the case where the emergence of autocatalysis is uncorrelated with the model parameters. This work represents a promising area for the initial study of the dynamical relationship between OEE and organisational closure, and may advance the understanding of their connection in theoretical biology.

Keywords: Closure, Open-ended evolution, Kauffman, Autocatalysis, Evolution, Complexity, Assembly Theory

Amongst all the current open questions in mathematics and physics, modelling life and more in general biological evolution represents a very complex field that has not yet found its mathematical framework. In a first approach, one could argue that biological evolution does not satisfy any physical symmetry as we know it. However, while physical invariances are a feature typically belonging to inanimate objects, living systems obey more complex invariances that live on a different scale, for example on the collective organisation of individuals. From this perspective, life could escape the classical paradigm of physical symmetries by posing its laws on another level, higher than the individual and possibly varying in an unbounded way.

Thus the classical variables used in physics, such as time and space, are no longer a good choice for describing a living system. In this sense, the equivalent of time for living things is what we call history. Therefore, space and time as we know them are no longer a good set of variables in biology, and this is also justified by the fact that living beings have freedom of choice, so they do not necessarily all follow the same optimised physical trajectories.

To describe biological evolution correctly, it is necessary to define symmetries that take into account history and context, which can be seen as two sides of the same coin.

A major symmetry in evolutionary biology is what we call *closure* [1; 2], which consists of a particular organisation of the elements within a living system. When the components of a system satisfy the closure conditions, they become a constraint for the existence of the other. That is, each element is a constraint because it depends on other constraints, and it also activates other constraints.

Another symmetry of evolutionary biology is the presence of the so-called Open-Ended Evolution (OEE). While recent works in artificial intelligence paved the way to the definition of open-endedness in different setups [3–8], the achievements attained in this field cannot be fully applied to open-ended evolution in biology. This is because cells are not universal Turing machines, as A. Pocheville wrote in [9], “biological sequence may be more complex than its algorithmic counterpart”. In other words, the OEE in biological evolution is the impossibility of constructing a mathematical description that describes the evolving system through its history.

Also, a pioneering work on Artificial Intelligence by R. Penrose [10] shows how the algorithmic intelligence cannot be comparable with the non-computational human mind, in deep contrast with works supporting strong-AI approches to for the algorithmic interpretation of living systems [11].

Historical monumental books such as Schrödinger’s ”What is life?” [12] paved the way to a new modern physics approach for the interpretation of livings systems that recent works [13] pushed even further in the interpretation of life as anti-entropy. Other works seem instead to prefer interpreting life as selection-derived self-organisation during Darwinian evolution [14; 15].

If life has not yet found its best form of mathematical language, several efforts have been made to provide a formal description of biological evolution, as a transition between two configuration of living systems, in terms of autonomy [16] and undecid-

Email address: marco.faggian.science@gmail.com (M. Faggian)

ability [17].
Different works on chemical systems were motivated by the possibility of finding a formal framework for a mathematical explanation of evolution, hoverer a part of these shows that evolution theories motivating the emergence of novelties from simple chemical reactions such as metabolism-first or replicator-first do not display true evolvability in the Darwinian sense [18]. Thus, novelty must be more than a simple discovery of new chemical species.
On the other hand, however, if we assume that the basic processes that created life rely on chemical reactions systems, then if an evolving system increases its complexity over time this implies the presence of a particular type of selection such as the *stabilising selection*, *disruptive selection* or *directional selection* that plays the causal role, or "the motor" [2; 13; 19] of increasing complexity.
Therefore, natural selection should suffice for evolution to emerge in a living system, and evolution must always emerge in presence of natural selection. Following this latter intuition, selection would be the only physical symmetry that causes evolution in evolutionary biological systems.
If modelling the basic characteristics of living systems is still a very complex task, a relevant progress was provided with the Kauffman model of auto-catalysis. [20]. Further developments of interpretation of life as a process fueled by auto-catalysis have been made by the same author [14; 15], pushing forward the analogy between the boost of complexity of an auto-catalytic cycle in a chemical reactions system and the boost of complexity produced by what we interpret as organisational closure in a living system.
In this work, by pushing forward the analogy between auto-catalysis and organisational closure, we provide evidence of causal relation between the emergence of functional organisations in biological systems and open-ended evolution in the framework defined by the Kauffman model [20].

1. Motivations

As already pointed out by other authors [2; 13; 21], understanding the maths of living systems requires a drastic change of paradigm. If current mathematical tools do not seem to be accurate enough to predict the dynamics of biological organisms, biological evolution seems to be posing a fatherly challenging problem.
In other words, the objective of finding mathematical tools to describe evolution could be expressed as the following statement: *"to say everything we can say about what we can't talk about"*.
In this article, we make use of some principles of theoretical biology to show evidence of new physical symmetries emerging in biological evolution.

1.1. Closure

As widely described in literature[2], the study of the thermodynamic flow in a biological system highlights some particular quantities that are conserved in specific timescales of the system (symmetries) named *constraints* [2; 22].
At different timescales, constraints are subject to degradation and have to be either replaced or repaired. This situation can be interpreted depending on the scale the constraint is being studied: for example while individual enzymes are replaced, their population is said to be repaired [1; 22].
In a biological organisation, the constraints represent a physical symmetry valid in a particular timescale which can also depend on other constraints. A set of constraints $\mathbf{C}$ is said to realise *Closure* if and only if [2] each $C_i \in \mathbf{C}$ depends directly on another constraint $C_j \in \mathbf{C},\ i \neq j$ at timescale τ_j (C_i is dependent), and there is at least another constraint $C_k \in \mathbf{C},\ i \neq k$ which depends on C_i at timescale τ_k (C_i is enabling).
While mathematically the concept of closure might seem pretty basic, from a biological and philosophical point of view it has relevant consequences [23] such as the autonomy of a living system, and at the same time its capacity to self-maintain and self-produce, thus establishing goals and norms while promoting the fundamental conditions for their existence through the interactions with the environment [1; 2; 22].
A simple model representing a simple example of emergence of organisational closure is the Kauffman model, in which the emergence of reflexively auto-catalytic sets of peptides and polypeptides can be associated to a particular realisation of closure [20].

1.2. Principle of Variation

Understanding biological evolution from a mathematical perspective requires to perform a change on the vision of the physical world. In particular, what is mostly required is a change of the conceptual modelling of the objects to describe.
The main subjects of physical theories are *generic objects* [21], that are physical entities of the same kind that obey the same laws. The generic objects are defined by symmetries that justify the underlying mathematical structure provided by the phase space and the equations that determine the trajectories. Such symmetries are time-invariant and do not change if the generic object suddenly changes, meaning that they capture the most essential features of an invariant.
On the other hand, biology is said to deal with objects that are called *specific objects* [22], which are defined not only by the symmetries that define the generic objects, but also by their qualitative features that are a consequence of their histories. Thus, the mathematical description of specific objects must rely not only on their symmetries, but also on their adaptation, corresponding to the evolution of such symmetries over time.
According to the *Principle of Variation* [22] that states that biological organisms are specific objects, it follows that biological systems undergo unpredictable symmetry changes over time. As biological organisms are specific objects, their evolution depends not only by their symmetries, but also on the context in which they live and the set of choices made, which constitutes their history. As each context and history is different for different specific objects, their mathematical structure must change unpredictably over time. If each specific object can undergo an unpredictable change, it follows that the way symmetries change in a system made of specific objects is order-free, and

the change can happen unpredictably at any level of the organisation of the biological system.
Systems that follow the Principle of Variation during their evolution require their reproduction to be sustained in an *open-ended* manner, thus the underlying process that leads all their evolution is called *Open Endedness*, or Open-Ended Evolution (OEE) [22; 24].

1.3. Open Endedness and Complexity

A recent work by Corominas-Murtra [25] seems to open a venue for a minimal description of OEE.
Following briefly their work, if we consider the descriptive state σ_t of a generic system at time t, we can define the set $\Sigma_t \equiv \{\sigma_i\}_{i=0}^{t}$ which tracks the "history" of the system at different discrete time steps $t = 0, 1, \ldots, t$. They define $K(\Sigma_t)$ the Kolmogorov complexity of the history of the system.
Corominas-Murtra [25] proposed three axioms for the formalisation of OEE, but we will focus particularly on the first, which states that the complexity $K(\Sigma_t)$ of an open-ended system does not decrease when divided by time:

$$\frac{K(\Sigma(t))}{t} \leq \frac{K(\Sigma(t+1))}{t+1}, \quad \forall t \tag{1}$$

While the increase of complexity alone corresponds to the coming of a novelty in the system, this axiom thresholds the growth of the complexity of an open-ended system at least as the complexity of a random chain.
In the case of continuous time it can also be re-written as

$$\frac{d}{dt}\frac{K(\Sigma(t))}{t} \geq 0, \quad t \in \Gamma \tag{2}$$

where Γ is a dense time interval.
In other words, if this condition is satisfied then a non-trivial novelty can occur at any time and the system experiences OEE. The non-triviality of the novelty consists in being a functional novelty that changes the organisation of the biological system, this it does not only correspond to the discover of a new living entity. Furthermore, the novelty can also be considered as a change of the organisation of the constraints of the living system. For example, in a system of chemical reactions the discovery of a new molecule is a novelty in the sense of chemical diversity, but it does not necessarily add a new functionality in the system. Thus, the newly discovered molecule might slightly increase the complexity, but not more than a simple random chain. In order for the molecule to increase the complexity in the sense of Corominas-Murtra [25]'s first axiom, it shall induce new functional changes, such as the emergence of a new auto-catalytic cycle in the Kauffman model.
The axioms provided by [25] represent an important contribution to the study of OEE, however due to the incalculability of the Kolmogorov complexity, it cannot be applied straightforwardly.
If this problem could represent a crucial blocking point in physics, in biology this is not always the case because biological dynamics does not involve generic objects, but specific objects. Exactly because of the nature of specific objects, it follows that the Kolmogorov Complexity might not always be a really appropriate measure of the complexity of biological systems.

2. A Measure of Complexity: Assembly

Recent developments [26; 27] pushed forward the mathematical formalisation of specific objects, providing a new inferring technique to measure the amount of selection via a quantity named "Assembly".
The calculation of this quantity depends on two main parameters, called respectively *assembly index* a_i, corresponding to the number of synthesising steps performed by the system to synthesise the molecule, and the *copy number* n_i which corresponds to the number of copies of a unique molecule, where the term *unique* assumes a special connotation because it corresponds to the concept of specific object as defined in [22]. In this sense, two molecules having the same structure can also be represented by two different specific objects because of their histories.
The *Assembly A* is then calculated [27] as follows:

$$A = \sum_i e^{a_i}\left(\frac{n_i - 1}{N_T}\right) \tag{3}$$

Where the sum is performed over all the unique molecules i and N_T is the total number of molecules in the system.
The authors claim that higher values of such quantity for two different realisations of a chemical system indicate the presence of a higher degree of selection[27].
We think that this quantity can also be interpreted as the measure of the the amount of "efforts" made to build the system up to a certain configuration. Equivalently to the calculation of the partition function $\mathcal{Z}$ in statistical physics, this quantity can also be interpreted as the energy necessary to bring the system to a specific configuration.
Therefore, in this article we will use the assembly not only as a measure of selection in our system, but also as a measure of the complexity. The calculation of the assembly neglects casual discoveries of new entities ($n_i \sim 1$) which do not contribute much to the complexity. It follows that only novelties meant as new functional organisations in the system provide an important contribution to the complexity of the system, as the entities that belong to the new organisation are stably produced in non-negligible quantities.

3. Our Hypothesis

In this article, we want to provide evidence that not only functional closure and OEE are strongly correlated, but that closure is the epistemological cause of OEE.
As well pointed out by [28], while physics does not have yet the tools to explicitly provide a causal explanation of physical phenomena, we could see a breakage of the formal symmetry as being correlated to a efficient cause, causing the change of

the state of the system while leaving its properties invariant.
For instance, let's consider the electromagnetic force $\vec{F} = q\vec{E} + q\vec{v} \wedge \vec{B}$. Considering only the electrostatic contribute $q\vec{E}$, we could state that $\vec{E}$ is the cause of the Coulomb force acting on the charge q. The same holds for the magneto-static component $q\vec{v} \wedge \vec{B}$, for which we could say that $\vec{B}$ is the cause of the Lorenz force acting on the moving charge q.
In biology, the formal symmetry breaking between the objective situation where "=" indicates simple correlation and where it indicates a causal epistemic regime is a more relevant event than it is in physics [28].
In order to tackle this open problem, we first adopt the hypothesis that OEE implies a non-decreasing historical complexity in the sense of Eq.(1) [25].
In a system of chemical reactions, the novelty does not correspond to the discovery of a simple new chemical species, but rather to the new organisation of the constraints of the system that provides a new and unpredictable functional role. Mathematically, this corresponds to saying that while the production of any possible molecule can be predicted at the origin of time, what cannot be predicted is the effects that a new organisation of constraints has on the system.
We consider the Assembly as a good measure of the complexity of the system described by the Kauffman model.
With this formal paradigm in mind, we make use of Kauffman's model [20] as toy model to for the modelling of closure as an auto-catalytic set of chemical reactions. By using this framework, we provide numerical evidence that the emergence of auto-catalysis causes the increase of complexity in the system in the sense of [25], following the theoretical intuition proposed by [29].
We consider closure and OEE as two faces of the same symmetry in biological evolution, as this holds for the electric and magnetic field in the electromagnetism.
As closure is a very special type of organisation of biological constraints, it should appear only if the system is subject to strong selection [1; 2]. On the other hand, OEE requires not only the system to be biologically stable through different generations, but also the system to be capable at any time of adapting and thus generating new functional organisations [1; 2; 30].

4. Numerical Study

We study the evolution of the Kauffman model as it is described [20; 29; 31] in a framework based on the Assembly Theory. In this model, any molecule can either ligate to another ("ligation") or split and produce two molecules ("cleavage"). Any reaction can be catalysed with probability p_{cat} by any chemical species that do not belong to the substrate. Catalysis is modelled as a reaction whose speed is k-times the speed of spontaneous reactions. Under some specific conditions, the dynamics gives rise to *Auto-Catalytic* sets (RAF) [20].
In this work, we refer to "RAF" the maxRAF detected in the system. For the same reason, we refer to the "size of RAF" as the size of the maxRAF present in the system, which corresponds to the joint size of all the independent RAF. A further study could differentiate the topological structure of the RAF, highlighting eventual uncovered dependencies of the assembly on topological properties.
We interpret the auto-catalytic networks not only as a functional novelty, but also as the representation of selection in the system itself. The molecules belonging to the auto-catalytic network not only follow a privileged dynamics, but they also are selected in the sense that their privileged production is crucial to the survival of the emerging functional organisation they belong to.
We measure the complexity of a realisation of a chemical system in Kauffman model via the Assembly Theory, and we focus our study on the effects of the emergence of auto-catalytic networks on the assembly.

4.1. The algorithm

We utilise the Gillespie algorithm to run the numerical simulations as in [31].
We consider a chemical reaction system (CRS) defined as in [31] by the tuple $\mathcal{CRS} = (X, R, C)$, where

- X is the binary molecules set, in which the most simple molecular subset corresponds to $\{0, 1\}$;
- R is the set of reactions involving the molecules in X that are of the type "ligation" when two molecules react to synthesise a new one, or "cleavage" when a molecules splits into two molecules of lower length.
  Each reaction is exhaustively described by the tuple $(v,\ x^{cat},\ lig,\ (x_1^{react},\ x_1^{react},\ x^{prod}))$ for ligation reactions, and $(v,\ x^{cat},\ cleav,\ (x_1^{prod}),\ x_1^{prod},\ x^{react})$ for cleavage reactions, where v is the reaction speed and $x^{cat} \in X$ the set of catalysts of the reaction;
- $C = \{\delta_{ij}(x_i^{cat}, r_j)\}_{i,j=0}^{N,M}$ is the set catalytic assignments, where δ_{ij} is a delta set-selecting function that equals to one if the molecule $x_i^{cat} \in X$ catalyses the reaction r_j and to zero otherwise, N is the total number of molecules in the system, M the total number of reactions, and $x_i^{cat} \in X$, $r_j \in R$. C is then a subset of the product set $X \times R$.

It is important to note that the delta function δ_{ij} not only associates the catalyst to the reactions, but it also contains information about the topology of the model.
The systems always disposes of an inexhaustible subset of molecules $\mathcal{F}$ named "substrate", which corresponds to the ensemble of all the possible binary molecules with length up to $F = \max_{\mathcal{F}} l(x|x \in \mathcal{F})$, where $l : x \to \mathbb{R}^+$ is the measure of the length of a molecule x. An incoming stream of molecules allows each chemical species belonging to the substrate to never decrease under the initial quantity $Q_{\mathcal{F}}$.
Only the molecules $x_i \in X$ with length up to $\mathbb{N} \ni L = \max_X l(x|x \in X)$, $F < L$, are allowed to exist in the system.
The speed of catalytic reactions is set to be k-times the value of the spontaneous reaction speed v_0: $v_{cat} = kv_0$. The choice of K and the initial amount of substrate F_0 plays a role on the speed of the simulations, but as we will see later, the conceptual results of our study are independent on them.
For our simulations, we set the maximum molecular length

$L = 8$ and the maximum substrate length $F = 2$ which corresponds to $\mathcal{F} = \{0, 1, 00, 01, 10, 11\}$. With this choice of F and L, the emergence of RAF is independent on these two parameters [20].
The initial configuration of the system at time $t = 0$ is solely provided by the molecules belonging to the substrate set $\mathcal{F}$, and the reaction dynamics is simulated using Gillespie's algorithm[32]. The molecules belonging to the substrate $\mathcal{F}$ do not catalyse any reaction. Every time a new chemical species is discovered during the evolution of the system, it catalyses an existing reactions with probability p_{cat}. The value of p_{cat} plays a crucial role on the time required to observe the emergence of a RAF set. [20; 31]

4.2. The molecules in the system

Our numerical simulations use reflecting boundary conditions, i.e. molecules with maximum length L can only react via cleavage reactions and cannot ligate to others. Equivalently, all the ligation reactions that would produce a molecule longer than L generate a molecule that is highly unstable. We consider these molecules to not be observable in the system, thus for simplicity in the numerical simulations these reactions are not allowed. Furthermore, we assume the absence structural symmetries among molecules, i.e. molecules represented by different binary sequences correspond to different chemical types.

4.3. RAF detection

In order to detect the RAF emerging in Kauffman's model, we adopted the algorithm for the autocatalytic sets generated by a food source (RAF) described by Hordijk et al. in [33] and we implemented it in Python 3.8.

4.3.1. Assembly index

When calculating the assembly of the system, before running the simulations, we pre-defined the assembly indexes a_i of each chemical species via Monte-Carlo simulations including all the possible relations 4 and 5 for a network of reactions with parameters (F, L).
In particular, the iterative algorithm to determine a_i for ligation reactions is

$$m_1 + m_2 \rightarrow M, \quad a_i(M) = a_i(m_1) + a_i(m_2) + 1 \tag{4}$$

and for cleavage

$$M \rightarrow m_1 + m_2, \quad a_i(m_1),\ a_i(m_2) = a_i(M) + 1 \tag{5}$$

By definition, the assembly index is null for all the molecules belonging to the substrate $\mathcal{F}$. The assembly index is always a positive quantity because the molecules of $\mathcal{F}$ do not do cleavage reactions.
In our simulations, we decided to associate the assembly index to a chemical species using what we call "*the minimal path history of chemical species*": if different assembly indexes are calculated for molecules belonging to the same species, we chose to take the minimum value which corresponds to the shorter assembly path possible in the system [34]. This decision comes from the intent of simplifying the model for numerical purposes, as the distribution of synthesising paths of a molecular species is very narrow.
There are two more possible ways to calculate the assembly index.
A first alternative approach that we call "*the history of the chemical species*" is to evaluate the the assembly index of each molecule on-the-run, and associate the minimum value to the chemical species whether different assembly paths exist for molecules with the same structure.
A second way that we call "*the history of the molecules*" is to associate to a molecule the index corresponding directly to its assembly path in a numerical simulation. For very large systems and very long simulated times, both these alternative approaches are equivalent to the aforementioned utilised in our simulations, however the latter two introduce a stronger historicity in the system that for different types of studies might unveil some more features of evolving biological systems.

4.4. Assembly index in the simulations

It can be easily shown by induction that a good approximation for the minimum value of the assembly index a_{max} for uni-dimensional binary molecules in this context is provided by

$$a_{max} \simeq \begin{cases} L/2 - 1, & \text{if } L \text{ even} \\ (L-1)/2, & \text{if } L \text{ odd} \end{cases} \tag{6}$$

The numerosity $\Omega(a_i)$ corresponding to the number of chemical species with the same assembly index a_i can be calculated by using Eq.(6) for any value of the molecular length. For example, for even values of the molecular length we get $\ln \Omega(a_i) = 2a_i \ln 2 + D$, where $D \in \mathbb{R}$ is a constant that changes depending on L.
In our simulations, we found a very good agreement for all these quantities, showing that the approximation holds very well. These results are also good evidence of the interpretation of the assembly as historical complexity for a system of chemical reaction of uni-dimensional binary molecules.

5. Results for the classical Kauffman's model

We consider a network of chemical reactions in which the catalysts $x_i^{cat} \in X$ are assigned to the reactions $r_j \in R$ with constant probability $P(r, cat)$:

$$P(r, cat) = p_{cat},\ p_{cat} \in [0, 1] \tag{7}$$

The original formulation of this model was done to ease autocatalysis-emergence, as can be shown by doing some basic calculations. Indeed, as the probability of associating a catalyst to a reaction r_i is a constant p_{cat}, it follows that the probability that the set of catalysts X^{cat} of a reaction r is not empty is

$$P(|X^{cat}| > 0 \mid r) = 1 - P(|X^{cat}| = 0 \mid r) = 1 - (1 - p_{cat})^N \tag{8}$$

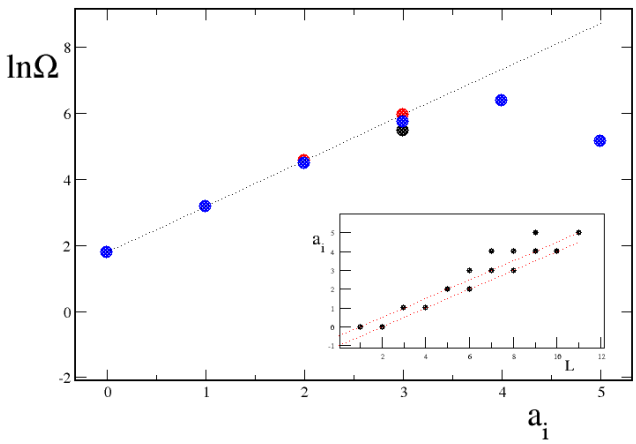

Figure 1: Logarithm of the numerosity of molecular species Ω encountered in the numerical simulations as a function of their assembly index a_i. Black dots: simulations for $N = 7$ and $F = 2$, red dots for $N = 8$ and $F = 2$ and blue dots for $N = 9$ and $F = 2$. Dashed black line: $\ln \Omega = 2a_i \ln 2 + \ln 6$. Inset: assembly index a_i as a function of the molecular length L for different molecular species encountered in the numerical simulations. The two dashed red lines correspond to the two approximations of Eq.(6).

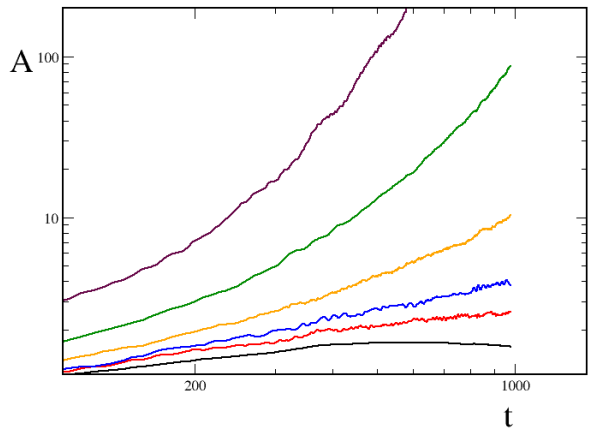

Figure 2: Numerical results of temporal dynamics of the Assembly for different runs of the system with $p_{cat} = 0.01$, $Q_{\mathcal{F}} = 10$ and $k = 10^4$ and different RAF sizes at the end of the simulated time: absence of auto-catalysis (black), unitary RAF (red), RAF with final size 3 at the end of the run (blue), final size 10 (yellow), final size 21 (green) and final size 50 (maroon).

As N increases in time, for $t \to \infty$ we find

$$P(|X^{cat}| > 0 \mid r) \to 1 - \varepsilon \tag{9}$$

where $1 \gg \varepsilon \in \mathbb{R}^+$ is a constant that decreases as p_{cat} increases. Then, for any value of $p_{cat} > 0$, for very long simulation times it is unlikely for a reaction to not find any catalyst in a infinite time interval and unbounded system. However, this is not always true for finite-sized systems and for finite time numerical simulations. As shown by [14; 31], in order to see emergence of a RAF, a threshold applies on the catalytic fraction of reactions for a given system size.

In Fig.(1) we report the relation between the numerosity of moleculare species versus the assembly index a_i, showing that under some hypothesis the assembly index is a good indicator of historical complexity.

In our work, when $p_{cat} > 0$ we will mainly focus on the range of values lower or at most equal to such threshold. As we will see in the next sections, the correlation between the emergence of organisational closure and the boost of the assembly are both correlated to p_{cat}. This correlation can be weakened by working in a range of very low values of p_{cat}.

5.1. Time evolution of the assembly

In order to have a first insight of the time behaviour of the assembly, we ran some numerical simulations and measured the assembly as a function of time for different realisations of the system for the same value of p_{cat}, as in Fig.2. The data show that the temporal dynamics of the assembly is strongly correlated with the type of RAF in the system. In particular, not only the assembly is higher in presence of auto-catalysis, but its time derivative is strongly correlated with the time-derivative of the size of the RAF. In other words, in our simulations the dynamics of the assembly depends on the dynamics of the auto-catalytic cycles. This always happens in the Kauffman model when the value of p_{cat} is high enough to allow the emergence of a RAF. For this reason, we firstly make use the first OEE axiom provided by [25] to explore deeper the correlation between the emergence of a RAF and the boost of the assembly. In Fig.3 we report a strong example in which this axiom is verified, as the time-derivative of A/t starts to increase right after the emergence of the RAF, whereas this does not happen when compared to the case without RAF. As we already mentioned it, the

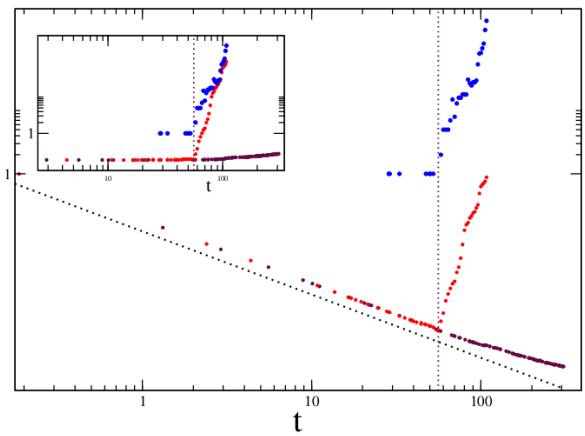

Figure 3: Assembly/time with emergence of auto-catalytic networks (red) and the corresponding curve of the RAF size (blue dots), assembly in absence of RAF (maroon), both the simulations were run for $p_{cat} = 0.002$, $Q_{\mathcal{F}} = 10$ and $k = 10^4$. Dashed brown line: $1/t$. Inset: same data in which the assembly A is not normalised by the time t.

dynamics of the assembly is strongly correlated with p_{cat}, for

this reason one could be tempted to think that this happens only because of the presence of a non-null value of p_{cat}. Thus, we first measure the correlation between the assembly and the size of the RAF for different values of p_{cat} to further study this correlation, then in the next sections we explore the dynamics of the model in the uncorrelated region where the value of p_{cat} is too low to produce any RAF, i.e. $p_{cat} \ll 1$.

5.2. Assembly and RAF

In order to explore the correlation between the assembly and the RAF size, we study the correlation for different values of p_{cat}. The results are reported in Fig.4.
As the correlation is obvious at this stage, the presence of a non-null p_{cat} is enough to not let us assert that the presence of auto-catalysis is the direct cause of the increase of the system's assembly. As well explained in [31], the parameter p_{cat} is strongly correlated not only to a threshold of probability of emergence of RAF, but also to the their growth speed. This means, the higher the value of p_{cat} and the higher is the probability that a RAF emerges, as well as the speed at which it grows after their emergence.

For this reasons, in the next sections we study the model

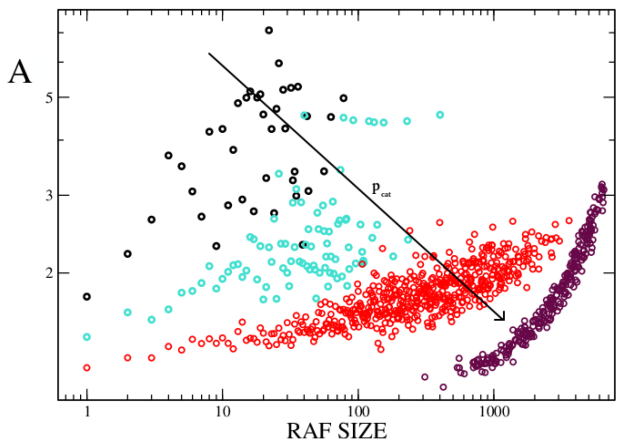

Figure 4: Scatter plot of assembly for $n = 10$ runs as a function of the size of the RAF for $Q_{\mathcal{F}} = 10$, $k = 10^4$ and for different values of P_{cat}: $P_{cat} = 0.003$ (black), $P_{cat} = 0.005$ (turquoise), $P_{cat} = 0.01$ (red) and $P_{cat} = 0.05$ (maroon).

in the limit $p_{cat} \ll 1$ in which dynamics of the assembly is uncorrelated from p_{cat}. The question we want to answer in the next part of this article is "can the smallest RAF create an Open-Ended dynamics in the Kauffman model?".

6. Imposing a unitary RAF

In order to better understand the causal relation between the emergence of RAF and the boost of complexity of the Kauffman model, we study the effect of a minimal-size RAF on the dynamics of the Assembly.
To do so, we study the dynamics of the model in the limit $p_{cat} \to 0$, and we analyse the effects on the numerical simulations of imposing a RAF of unitary size (URAF) with respect to the presence of a simple catalytic reaction (CR). In the simulation, the URAF and the CR are activated when their catalyst is introduced in the system at time t_0. The catalyst plays the role of a "seed" of complexity growth. If closure is the cause of the boost of the assembly, then we expect a URAF to change the dynamics in a very different way compared to a simple CR. We compare the difference of the growth of complexity in two different types of systems with URAF and with CR using the first axiom by [25].

6.1. Effects of the introduction of a URAF

When in the presence of a URAF, the time behaviour of the assembly is pretty simple to understand.
Let's call A_0 the assembly contribute of the catalyst C, called also the *seed* of the system. One could express the total assembly of the system as the sum of several contributions:

$$A = \bar{A} + A_0 \tag{10}$$

where

$$\bar{A} = A(\mathcal{F}) + A(\mathcal{M}) + A_1 \tag{11}$$

where $A(\mathcal{F})$ is the assembly of the substrate (i.e., the molecules with assembly index equal to zero), $A(\mathcal{M})$ the contribution of the molecules $\mathcal{M}$ that have C as common ancestor in the chain of reactions, and A_1 the contribution of all the other molecules. The contribution of the molecules in $A(\mathcal{F})$ is negligible over time, as their assembly index is null. The amount of substrate $\mathcal{F}$ can be considered almost constant over time, so that $A(\mathcal{F})/t \to 0$.
The contribution A_1 is negligible as well, as the molecules the contribute to this part of assembly are synthesised by spontaneous processes.
Thus, we can make the following approximation:

$$A \simeq A(\mathcal{M}) + A_0 \tag{12}$$

Using a biology metaphor, we could say that with this redefinition we assume the assembly to be determined by two main contributes: the seed-contribute A_0 and all the parts of the system that grow around the seed, $A(\mathcal{M})$. While the seed-contribute plays the part of the fuel of the system for a long time range, in the very long term its impact on the system will start to vanish, letting the dynamics of the contribute $A(\mathcal{M})$ take over in the dynamics. When the derivative of the dominant contribute $A(\mathcal{M})/t$ is non-negative, we interpret it as a signature of OEE. In our case, the cause of OEE is the introduction of a URAF in the system.

6.1.1. Effects of the introduction of a URAF

We developed an algorithm that introduces an auto-catalytic network of a given size L_{RAF} during a numerical realisation of the Kauffman model's dynamics for $p_{cat} = 0.0$, with the objective of studying the impact of such imposition on the assembly dynamics.
We expect the imposition of a URAF to boost the assembly of the system, making it attain a higher level of complexity in the

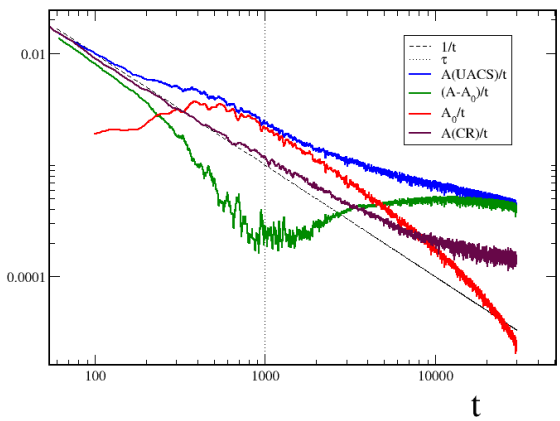

Figure 5: Time evolution of the assembly with a URAF compared to the assembly when in presence of a CR, for $p_{cat} = 0.0$, $Q_{\mathcal{F}} = 10$, $k = 10^4$ and URAF introduced at $t_0 = 0.0$.
Blue: Total Assembly with URAF over time $A(URAF)/t$.
Maroon: Total Assembly with CR over time $A(CR)/t$.
Red: Assembly of the URAF A_0 over time, A_0/t.
Green: $(A - A_0)/t$ Assembly of the system without the URAF, corresponding to the effects of the RAF on the creativity of the rest of the system.
Curves averaged over $n = 50$ realisations.

dynamics compared to a simple catalytic set CR. We consider a chemical reaction of the type

$$A + B \rightarrow C \tag{13}$$

in which the couple $A,\ B \in \mathcal{F}$ and $C \notin \mathcal{F}$. For the URAF case, we consider the reaction catalysed by C, whereas for a CR it is catalysed by another molecule $D \notin \mathcal{F}$, to avoid constant catalysis due to the permanent presence of the substrate.
In order to provide also some analytical considerations, for the case of the imposition of a URAF we study the time evolution of two different types assembly: the total assembly of the system $A(t)$ and the *complementary assembly* $\bar{A}(t)$, defined as the assembly of the system without the catalyst.
For both the cases of the imposition of a URAF or a CR, we compare the same type of reaction: a ligation-type reaction whose reactants are two different molecules that belong to the substrate $\mathcal{F}$. In Fig.(5) we report the temporal behaviour of the total assembly of the system for the case of URAF and a simple CR, together with the time evolution of $\bar{A}(t)$ after the introduction of a URAF.
Numerical simulations show not only that the assembly of the system with a URAF is always higher than the one with a simple CR, but also that the URAF induces a new dynamics in the model compared to the case without auto-catalysis. In particular, the introduction of a URAF propagates a wave of complexity that increases the probability that the system discovers new complex molecules. Moreover, the URAF plays the role of the fuel for the other chemical reactions, boosting the system to discover more chemical species. The propagation of this effects stops when when the number the molecules belonging to $A - A_0$ is not anymore negligible compared to the quantity of C.

This is a very important result. Indeed, if we subtract the contribution to assembly made by the exponentially growing URAF, we see that the dynamics of complexity follows a non-decreasing trajectory, showing an open-ended evolution over a very large time interval. This is well illustrated in Fig.(5). In the figure, this is represented by the green curve. It can be seen that the curve follows a non-decreasing dynamics over a very large time interval, after the URAF saturates the system. This is caused by a cascade of highly complex molecules that is triggered by the initial exponential growth of the seed. Then the curve reaches a quasi-plateau for effects of finite size.

6.1.2. Integration of kinematic equations

In order to fully understand the dynamics of the Kauffman model from a perspective of assembly theory, we also studied the kinematic equations for a URAF inserted at $t_0 = 0.0$
The types of molecules present in the system divide in four principal categories: the molecules belonging to the food N_F, the molecule generated by the auto-catalytic nucleus N_C, the molecules generated using reactions that involve the catalyst of the nucleus C that we call N_{MC}, and finally all the other molecules that are generated by reactions that involve molecules of the food, that do not belong to the ensemble of MC, that we call N_{MF}.
In equations, $N_T = N_F + N_C + N_{MC} + N_{MF}$. Let's consider the kinematic equations of the chemical reactions involving all these molecules. The auto-catalytic reaction that produces the molecule C can be written as $A+B \xrightarrow{C} C$, where $A \neq B$, and both belong to the food set ($A,\ B$ belong to a subset of food that corresponds to a fraction $1/6$ of the total initial molecules). Defining the quantity Q as the initial quantity for each molecule type in the food set, v the non-catalytic reaction speed, k the ration between the catalytic and the non-catalytic speed, and denoting with n the numerical density of a molecule in the system:

$$\begin{cases} N_F = 6Q \\ \\ \dfrac{\dot{n}_C}{v} = \dfrac{n_F n_F}{36}(1 + kn_C) + (n_{MC} + n_{MF}) - \\ n_C(1 + n_F + n_C + n_{MC} + n_{MF}) \\ \\ \dfrac{\dot{n}_{MC}}{v} = (n_{MF}n_F + n_{MF}n_{MF}) + n_C(n_F + n_C + \\ n_{MC} + n_{MF}) - \alpha n_{MC} + \beta n_{MF} \\ \\ \dfrac{\dot{n}_{MF}}{v} = \dfrac{4}{9}n_F n_F - n_{MF} + \alpha n_{MC} - \beta n_{MF} \end{cases} \tag{14}$$

where $0 < \alpha < 1$ corresponds to the fraction of reactions that from the set MC create a molecules that belongs to MF, and $0 < \beta < 1$ to the fraction of reactions from MF to MC These two coefficients represent the common molecules between MF and MC.
Following these equations, n_{MF} becomes relevant for long time intervals and in particular we find $n_{MF} \sim \alpha n_{MC}$).
A good estimate of the parameters α and β allows to better predict the time behaviour of the molecules of type MC.

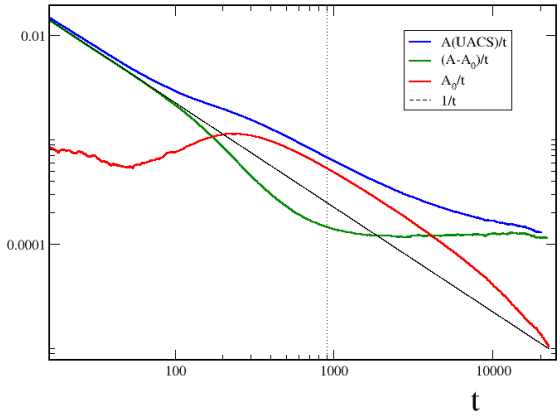

Figure 6: Numerical simulations of the differential equations of kinematics for the total Assembly A of the system and the Assembly of the URAF A_0, for $p_{cat} = 0.0$, $Q_{\mathcal{F}} = 10$, $k = 10^4$ and URAF introduced at $t_0 = 0.0$.
Red: Total Assembly over time A/t.
Blue: Assembly of the URAF A_0 over time, A_0/t.
Green: $(A - A_0)/t$ Assembly of the system without the URAF, corresponding to the effects of the RAF on the creativity of the rest of the system.
Curves averaged over $n = 50$ realisations.

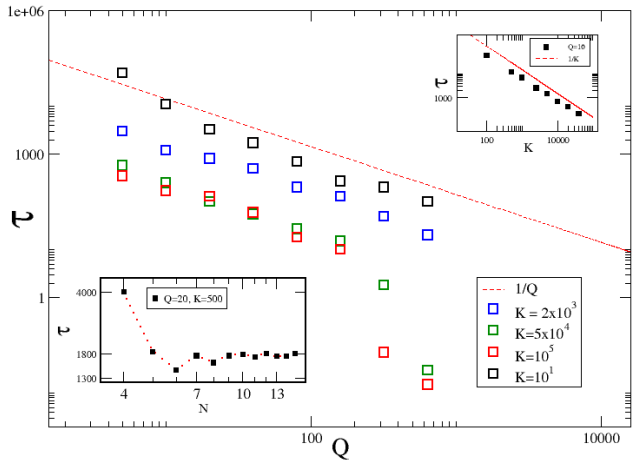

Figure 7: Measurements of τ in the realisation of the Kauffman model for different values of the initial amount of each food species $Q_{\mathcal{F}}$ and k. The results show a good agreement with $\tau \sim 1/Q$. Inset bottom left: dependency of τ on the max length parameter N. The value of τ converges for increasing values of N. Inset top right: dependency of τ on k. The results show a good agreement with $\tau \sim 1/k$

Fig.(6) are reported the results of the numerical integration of such equations that show a good compatibility with the behaviours shown in Fig.(5). In particular, the kinetic equations allow one to better appreciate the presence of a time-asymptotic plateau for $(A - A_0)/t$, showing that the presence of a URAF in the system boosts its creativity and pushing its evolution to a non-negative time derivative. The differences between the time behaviour of $(A - A_0)/t$ in the kinematical simulations and the Gillespie simulations are due to a nonperfect estimate of the parameters α and β.
Finally, we want to point out that one might consider controversial the finding of OEE in this set of kinematic equations, because, as we said in the beginning of this article, an open-ended evolving system is not supposed to be described by a finite number of equations[22]. We will return to this point in the conclusions of the article.

6.1.3. *The characteristic time τ*

By defining τ as the value of t at which the time-derivative of $\bar{A}/t$ is not anymore negative and the second time-derivative changes sign (see Fig.(5)), we can define two main regimes for the total assembly A: for $t < \tau$ we find $A \sim A_0$, whereas for $t \gg \tau$ we have $A \sim \bar{A}$. We explain this behaviour with the kinetic equations. For $t < \tau$, the dynamics of the number of seeds obeys the approximate differential equation $\dot{n}_c \simeq k v n_A n_B n_C$ where we identify with n the density of molecules in the system.
By solving the differential equation, we find that the number of seeds in the system follows an exponential growth of the type $N_c \sim \exp(k v Q_{\mathcal{F}} t) \equiv \exp(t/\tau)$, which slows down when the number of seeds saturates the system for finite-size effects around $t \simeq \tau$, and is followed by a cascade of reactions involving more complex molecules. In the second phase, for $t \gg \tau$ the growth rate of the seeds is almost constant, bringing N_c to grow linearly, $n_c \sim k v Q_{\mathcal{F}} t$. In this regime, due to the competition between the presence of C and that of its descendants, the main contribution of the total assembly A is provided by molecules $\mathcal{M}$. This is equivalent to stating that the dynamics presents a complexity wave of molecular length boosted by the initial URAF.
Considering $A_0 \sim e n_c \equiv e N_c / N_T$ and approximating the solution of the differential equation of n_c in Eq(14), all these considerations on the reaction dynamics of the seed translate to:

$$\frac{A_0}{t} \sim \frac{\exp(t/\tau)}{t(1 + \exp(t/\tau))} \longrightarrow_{t \gg \tau} \frac{1}{t}, \tag{15}$$

which explains the convergence $A \to A_0$ for $t < \tau$.
Concerning the relation between τ and the system size L, as reported in Fig.(7)the numerical simulations confirm that the value of τ converges to $\tau \simeq 1/k v Q_{\mathcal{F}}$ for $N > 5$. This confirms the independence of our findings on the parameters k and Q, showing an adaptation of the dynamics on different timescales.

6.1.4. *Impact of τ on the removal of a URAF*

We also studied the case of the removal of a URAF, which showed that in this case the time-derivative of $A(\mathcal{M})/t$ does not become positive at any moment if the seed is removed at $t_0 \ll \tau$, while the dynamics remains unchanged compared to Fig.(5) if $t_0 \geq \tau$. In these conditions, for $t > \tau$ the dynamics does not depend anymore on the initial URAF and the boost of complexity derives only from a inertial complexity wave of chemical reactions initially induced by the catalytic seed.

6.2. *Long term behaviour*

For what concerns the behaviour for $t \gg \tau$, we find that for $t > \tau$ the value of $\bar{A}$ reaches a long temporal plateau after increasing for a very long time interval. We call $\bar{A}_\infty$ the value of $\bar{A}$ on this plateau. Fig(8) shows the N-scaling of $\bar{A}_\infty$ for $t \gg \tau$.

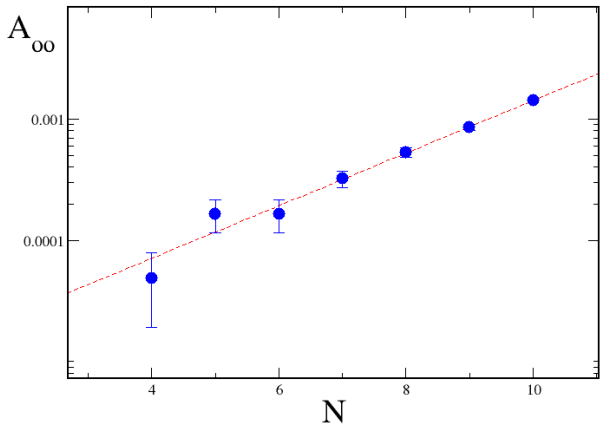

Figure 8: Values of $\bar{A}_\infty$ for $F_0 = 10$ and $K = 10^4$ (red dots). Black dashed line: $\bar{A}_\infty \sim \exp(N/2)$.

This dependence on N is a direct consequence of the definition of $A(\mathcal{M})$. In particular, one can show that $A(\mathcal{M})$ is dominated by the terms with higher assembly index:

$$\bar{A}_\infty = \frac{1}{N_T} \sum_{\{\mathcal{M}\}} (n_\mathcal{M} - 1) \exp(a_\mathcal{M}) \simeq \tag{16}$$

$$\simeq \frac{e^{a_{max}}}{N_T} \Omega(a_{max}) = e^{a_{max}} n(a_{max})$$

where $\Omega(a_{max})$ is the total number of molecules with assembly index a_{max} and $n(a_{max})$ their density.
By using the approximation of Eq.(6), we get

$$\frac{\bar{A}_\infty(L_1))}{\bar{A}_\infty(L_2)} \propto \exp((L_1 - L_2)/2) \tag{17}$$

which explains well the results reported in Fig(8), showing an evidence for unbounded OEE in absence of finite-size effects for $L \to \infty$

6.3. *Imposition of a URAF at different t_0*

Lastly, we decided to study the effects of the introduction of a URAF at times different than $t_0 = 0$ to check whether the dynamics studied in the previous sections does or not depend on the configuration of the system.
As reported in Fig(9), we found that the reaction of the system after the introduction of a URAF is independent on the value of t_0, showing that the complexity boost in the Kauffman model does not depend on the time interval in which it starts. This also show the absence of nucleation in our specific setup.

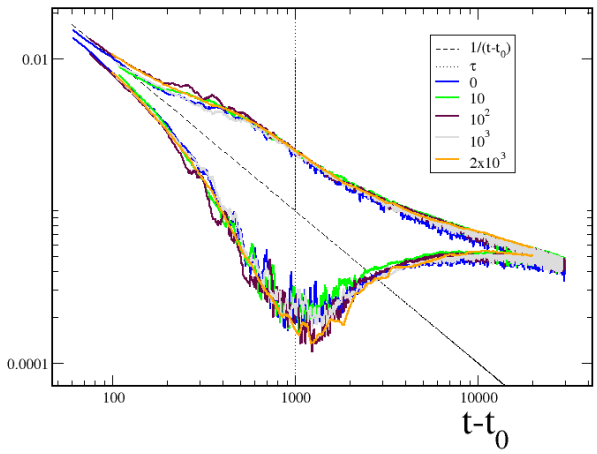

Figure 9: Curves of total assembly A/t) (top) and $\bar{A}/t$ (bottom) for different values of the time of introduction of the URAF t_0 reported in the legend box on the top right side of the plot. Curves averaged over $n = 50$ realisations.

7. Further Developments

We analysed the impact of the emergence of an auto-catalytic network on the dynamics of the assembly of the Kauffman model. We interpreted autocatalysis as the simplest representation of organisational closure for a system, and assembly as a measure of its complexity. We explored the relationship between the size of auto-catalytic cycles and assembly, and showed the effect of cycles of different sizes on assembly. We showed that the onset of autocatalysis is a critical event for the dynamics of the assembly, because even the smallest cycle brings drastic changes to the dynamics of the chemical system. In particular, numerical simulations show that the mere presence of a minimal autocatalytic cycle is sufficient to cause the emergence of open dynamics of the assembly as defined by [25], regardless of the choice of model parameters. The Kauffman model shows this pattern in different conditions and for different choices of parameters. Each time we found open-ended behaviour, the only invariant was autocatalysis.
These results, which we found by analysing the dynamics of the Kauffman model, could be further analysed using different variants of the model. A first possibility for a variant is the case where p_{cat} is not a constant, but adapts to keep either $P(r, cat)$ or the catalytic fraction of the system constant. Another variant is the model where a pro/anti ancestor bias can be introduced on p_{cat} for a potential catalyst if at least one of the molecules involved in the reaction is an ancestor.
Although in this article we have chosen to study the simplest version of the Kauffman model, the same results should be found in its variations.
From the point of view of nucleation theory, another line of research could be to study in detail the dynamics of the auto-catalytic cycles for non-null values of p_{cat}. We expect the nucleation of the auto-catalytic nucleus to boost the assembly

dynamics even more, far beyond the simple plateau we found for $p_{cat} = 0$. In particular, we expect the system to exhibit a non-null derivative of A/t during nucleation for non-null values of p_{cat}.

Furthermore, another important aspect to study is the role of the topology of the auto-catalytic cycles and their sub-cycles on the increase in complexity. For example, we expect more connected cycles to have a higher boosting power, just as it should be for cycles with a simpler topology.

Finally, a final line of research is justified by the search for a better measure of complexity for a system like Kauffman's. We believe that a more specific definition of assembly for binary molecules should be provided in order to reveal more features of OEE dynamics in similar settings. In particular, while for relatively short binary molecules the calculation of the assembly index represents the historicity of the system quite well, for long non-binary molecules such as DNA this approximation may no longer be valid. For example, a redefinition of the assembly index that counts the length of the compressed size of each molecule might reveal more features in the case of multidimensional non-binary molecules.

8. Conclusions

In this paper we have discussed the emergence of so-called open-ended evolution in a rather simple case of the Kauffman model in the presence of auto-catalytic cycles using assembly theory. We provided evidence for a causal relationship between organisational closure and OEE when the assembly of the system is interpreted as its historical complexity in the setup provided by the Kauffman model.

While this work represents a promising area for the study of biological evolution, more research is needed to build a solid bridge between knowledge of physical objects and that of biological entities. If life has long eluded scientists' full understanding, physics may one day unlock its secrets by looking at the symmetries that life itself has left behind on its evolutionary path.

9. Statement

The author claims to have developed the content of this paper independently without taking inspiration from others.

10. Acknowledgements

The author thanks A. Pocheville, M. Barbier and G. Folena for prolific scientific discussions and Y.D.F. for insightful conversations.

This project/ publication was made possible through the support of Grant 62220 from the John Templeton Foundation. The opinions expressed in this publication are those of the author(s) and do not necessarily reflect the views of the John Templeton Foundation.

References

[1] Matteo Mossio, Maël Montévil, and Giuseppe Longo. Theoretical principles for biology: Organization. Progress in Biophysics and Molecular Biology, 122(1):24–35, October 2016.

[2] Alvaro Moreno and Matteo Mossio. Biological autonomy: a philosophical and theoretical enquiry, volume 12. Springer, 2015.

[3] Alyssa Adams, Hector Zenil, Paul C. W. Davies, and Sara Imari Walker. Formal Definitions of Unbounded Evolution and Innovation Reveal Universal Mechanisms for Open-Ended Evolution in Dynamical Systems. Scientific Reports, 7, April 2017.

[4] Pedro A. Castillo and Juan Luis Jiménez Laredo, editors. Applications of Evolutionary Computation: 24th International Conference, EvoApplications 2021, Held as Part of EvoStar 2021, Virtual Event, April 7–9, 2021, Proceedings, volume 12694 of Lecture Notes in Computer Science. Springer International Publishing, Cham, 2021.

[5] Department of EECS (Computer Science Division) University of Central Florida, Orlando, FL 32816, L. Soros, and Kenneth Stanley. Identifying Necessary Conditions for Open-Ended Evolution through the Artificial Life World of Chromaria. In Artificial Life 14: Proceedings of the Fourteenth International Conference on the Synthesis and Simulation of Living Systems, pages 793–800. The MIT Press, July 2014.

[6] Nicolas Loeuille, Michel Loreau, and Simon A. Levin. Evolutionary Emergence of Size-Structured Food Webs. Proceedings of the National Academy of Sciences of the United States of America, 102(16):5761–5766, 2005.

[7] Hiroki Sayama. Seeking open-ended evolution in Swarm Chemistry. In 2011 IEEE Symposium on Artificial Life (ALIFE), pages 186–193, Paris, France, April 2011. IEEE.

[8] Susan Stepney. Modelling and measuring open-endedness. Artificial Life, 25(1):9, 2021.

[9] Arnaud Pocheville. Biological information as choice and construction. Philosophy of Science, 85(5), July 2018.

[10] Roger Penrose. The emperor's new mind. RSA Journal, 139(5420):506–514, 1991.

[11] Douglas R. Hofstadter. Gödel, escher, bach: an eternal golden braid. New York :Basic Books, 1979.

[12] Erwin Schrodinger. "What is life"?, Cambridge University Press, 1944.

[13] Francis Bailly and Giuseppe Longo. BIOLOGICAL ORGANIZATION AND ANTI-ENTROPY. Journal of Biological Systems, 17(01):63–96, March 2009.

[14] Stuart A. Kauffman. The Origins of Order: Self-organization and Selection in Evolution. Oxford University Press, 1993. Google-Books-ID: lZcSpRJz0dgC.

[15] Stuart A. Kauffman. Antichaos and Adaptation. Scientific American, 265(2):78–84, August 1991.

[16] Kepa Ruiz-Mirazo, Juli Peretó, and Alvaro Moreno. A Universal Definition of Life: Autonomy and Open-Ended Evolution.

[17] Troy Day. Computability, Gödel's incompleteness theorem, and an inherent limit on the predictability of evolution. Journal of The Royal Society Interface, 9(69):624–639, April 2012.

[18] Vera Vasas, Eörs Szathmáry, and Mauro Santos. Lack of evolvability in self-sustaining autocatalytic networks constraints metabolism-first scenarios for the origin of life. Proceedings of the National Academy of Sciences, 107(4):1470–1475, January 2010.

[19] Börje Ekstig. Complexity, Natural Selection and the Evolution of Life and Humans. Foundations of Science, 20(2):175–187, June 2015.

[20] Stuart A. Kauffman. Autocatalytic Sets of Proteins. J. theor. Biol., 119:1–24, 1986.

[21] Longo G. (2011) Bailly, F. Mathematics and the natural sciences; the physical singularity of life. volume 16, pages 309–336. 2011.

[22] Maël Montévil, Matteo Mossio, Arnaud Pocheville, and Giuseppe Longo. Theoretical principles for biology: Variation. Progress in Biophysics and Molecular Biology, 122(1):36–50, October 2016.

[23] H. R. Maturana and F. J. Varela. Autopoiesis and Cognition: The Realization of the Living. Springer Science & Business Media, December 2012. Google-Books-ID: iOjVBQAAQBAJ.

[24] Arnaud Pocheville. A Darwinian dream: on time, levels, and processes in evolution. In Tobias Uller and Kevin N. Laland, editors, Evolutionary Causation. Biological and philosophical reflections, Vienna Series in Theoretical Biology. MIT Press, Boston, 2019.

[25] Bernat Corominas-Murtra, Luís F. Seoane, and Ricard Solé. Zipf's Law, unbounded complexity and open-ended evolution. Journal of The Royal Society Interface, 15(149):20180395, December 2018.

[26] Stuart M. Marshall, Douglas G. Moore, Alastair R. G. Murray, Sara I. Walker, and Leroy Cronin. Formalising the Pathways to Life Using Assembly Spaces. Entropy, 24(7):884, June 2022.

[27] Abhishek Sharma, Dániel Czégel, Michael Lachmann, Christopher P. Kempes, Sara I. Walker, and Leroy Cronin. Assembly theory explains and quantifies selection and evolution. Nature, 622(7982):321–328, October 2023.

[28] Longo G. (2006) Bailly, F. Mathématiques et sciences de la nature. la singularité du vivant. 2006.

[29] Wim Hordijk and Mike Steel. Autocatalytic Networks at the Basis of Life's Origin and Organization. Life, 8(4):62, December 2018.

[30] Arend Hintze. Open-Endedness for the Sake of Open-Endedness. Artificial Life, 25(2):198–206, May 2019.

[31] Wim Hordijk and Mike Steel. Chasing the tail: The emergence of autocatalytic networks. Biosystems, 152:1–10, February 2017.

[32] Daniel T Gillespie. A general method for numerically simulating the stochastic time evolution of coupled chemical reactions. Journal of Computational Physics, 22(4):403–434, December 1976.

[33] Wim Hordijk, Joshua I Smith, and Mike Steel. Algorithms for detecting and analysing autocatalytic sets. Algorithms for Molecular Biology, 10(1):15, December 2015.

[34] Keith Y Patarroyo, Sara I Walker, Abhishek Sharma, and Leroy Cronin. AssemblyCA: A Benchmark of Open-Endedness for Discrete Cellular Automata.