A complex system model of biological aging

Author: Paula Pirker Díaz
Facultat de Física, Universitat de Barcelona, Diagonal 645, 08028 Barcelona, Spain.

Advisor: Albert Díaz-Guilera

Abstract: The aim of this work is to analyze the network influence in a model of aging that has been proposed recently. The model is based in the interdependency of nodes affected by damage and repair processes. While the original work was limited to analyze random graphs and Barabási-Albert networks, we propose a general framework based on an exponent dependent preferential attachment rule for the construction of the network, and the two cases analyzed previously correspond to particular values of the exponent. We provide a discussion on the biological plausibility of the different models, which now include a sublinear and superlinear dependence on the degrees.

I. INTRODUCTION

Biological aging can be described as the accumulation of damage leading to a collapse of the organism and its causes and mechanisms are still uncertain. [2] Living systems are interconnected dynamical systems and its aging can be studied from different perspectives. Firstly, from their structure. There are many scales to consider: from cellular, to tissue or to organismal. Secondly, from the organism's state of health. In this case, aspects as the medical record, alimentation or regular exercise are relevant. Finally, it can be studied from a sociological point of view, as health and mortality change within national populations and socio-economic groups. For this reason, many theories coming from different points of view have been formulated in order to understand these complex systems.

There are quantitative models, which describe empirical data. For example, mortality rates increase in aging populations and the risk of dying increases as an exponential for older ages in humans. That tendency is called Gompertz law of mortality but its causes are still not well understood.

As these models describe mortality but not health, there are others that use several parameters in order to quantify it. There are studies that work with the Frailty Index (FI), which measures the fraction of deficits of health and function from a large selection of possible ones; and others with the Biological Age (BA), which is an age defined by several aspects of health. There are many other parameters and they are used by predictive models to predict individual health outcomes.

Other models are focused on the dynamics of aging and build an "organism" relating interacting variables. This "organism" is not real but it can give an idea of how does aging work and can be used to identify patterns already observed in nature. It can be represented by a network, where nodes are cells/tissues/organs/health measures and edges are the interactions between them. Because of its generality these models are unable to study particular cases, but this generality makes them powerful and useful for conceptual explorations.

But it is possible to go beyond by generating large populations of individuals with detailed health trajectory and mortality thanks to big data and machine learning tools. However, these computational models present a relevant problem. There is a huge number of parameters required to fit a population and identifying an appropriate minimal set from them is far from easy. Moreover, it is not clear if parameter changing alone is enough to represent different organisms or populations, because maybe the structure of the network also needs to be changed.

As it can be seen, aging is a challenging problem with many suggested models, all of them with their weaknesses and strengths. This work combines network theory and the next to last mentioned model in order to study how does connectivity affect to aging dynamics. In fact, the model proposed in [1] has been reproduced but two additional types of network have also been studied.

II. MODEL

A. Network Model

To study the dynamics of aging, the model originally proposed in [1] starts with the consideration of an organism as a network of N nodes. Each node represents an individual component of the system and the edges represent the interdependencies between these components.

In order to create the network of an organism bearing in mind that in nature we find networks that are continuously growing by the addition of new nodes, the next steps have been followed:

- Begin with one node, $i = 1$.
- Introduce a new node $(i+1)$ and make it depend on any of the preexisting ones with probability $\Pi(k_j)$, where k_j is the degree of the preexisting node j.
- Make one of the preexisting nodes m depend on the new one with probability $\Pi(k_m)$.
- Repeat the previous two steps for $N-1$ steps in order to obtain a network of N nodes at the end.

The way nodes are connected is a relevant aspect to keep in mind, because the connections determine the interdependencies between the organism components.
As we have a growth process while creating the network, the older nodes have a greater attachment probability than the newer ones. For this reason the key lies in how the attachment probability is chosen. Let's consider a generic model, being k_j the degree of the node j,

$$\Pi(k_j) = \frac{k_j^\alpha}{\sum_i k_i^\alpha} \tag{1}$$

If $\Pi(k_j)$ depends on the degree of the nodes, preferential attachment will emerge. Otherwise it will be a constant number for all nodes, what means a lack of preferential attachment. For this reason, the α parameter allows the adjustment of the preferential attachment [5]:

- No Preferential Attachment ($\alpha = 0$)
  The absence of hubs makes the resulting network similar to an Erdös-Rényi network. Defining m as the number of links with which each node arrives, in this case 2. The degree distribution follows an exponential.
- Linear Regime ($\alpha = 1$)
  It corresponds to a scale-free network known as Barabási-Albert network, with a power law degree distribution.

In [1] only these two previous networks are studied, but for the sake of completeness two other networks are proposed in this work:

- Sublinear Regime ($0 < \alpha < 1$)
  It results in a network with few and small hubs (nodes with a number of links that exceeds the average) and the degree distribution follows an stretched exponential.
- Superlinear Regime ($\alpha > 1$)
  In this case the resulting network has a hub-and-spoke topology, where the old nodes become super hubs.

B. Aging Model

For a given network it is assumed that each component can be damaged eather with a certain probability $\gamma_0 \ll 1$ or because most of its neighbours are damaged, and repaired with a probability $\gamma_1 \ll 1$. Additionally, we fix an initial fraction of damaged nodes, $d \ll 1$.

These rules are implemented following the next steps:

1. Define the state of each node $\Psi(t) = \{\psi_1(t), \psi_2(t), ..., \psi_N(t)\}$ where $\psi_i(t)$ can be 1 (functional node) or 0 (nonfunctional, damaged node). Initially a state value of 0 is assigned to a fraction d of randomly selected nodes.
2. For each node i, $\psi_i(t)$ is updated with probability γ_0 to flip $\psi_i(t) = 1$ to $\psi_i(t) = 0$, and with probability γ_1 to flip $\psi_i(t) = 0$ to $\psi_i(t) = 1$.
3. Flip $\psi_i(t) = 1$ to $\psi_i(t) = 0$ if the fraction of living neighbours is lower than 0.5 (this step is repeated until no more nodes become nonfunctional).
4. Compute the vitality of the organism, defined as $\Phi(t) = \sum_i \psi_i(t)/N$.
5. Repeat steps 2-4 until $\Phi(\tau) = 0.01$, where τ is what we take as the time of death.

Using the described model two different kinds of analysis have been made. Firstly, the dependencies between parameters such as γ_0, γ_1 or d and how does the network structure effect to them. Secondly, other relevant magnitudes in aging processes have been computed throughout the results in order to study their evolution and relationship with the initial conditions.

One of these magnitudes is the mortality rate, which has been studied using an ensemble of networks representing "different" organisms. It is computed using the fraction of organisms that are still alive at a given timestep, $s(t)$:

$$\mu(t) = \frac{s(t) - s(t+1)}{s(t)} \tag{2}$$

Another magnitude that turns out to be interesting is the one that gives the strength of interdependence between nodes:

$$\lambda(\phi(t)) = \frac{log(\phi(t))}{log(\phi_0(t))} \tag{3}$$

where $\phi_0(t) = exp((-\gamma_0 + \gamma_1)t)$ is the expected vitality of a network with the same number of nodes but with all dependency edges removed.

III. RESULTS

First of all, the vitality has been studied including a comparison of how its evolution depends on the network type. In this case, four different networks have been studied: no preferential attachment network (NPA, $\alpha = 0$), sublinear regime network ($\alpha = 0.5$), scale-free network (SFN, $\alpha = 1$) and superlinear regime network ($\alpha = 3$ [9]).

As it can be seen in figure 1, all organisms start with a slow linear decay of their vitality with an specific slope depending on their structure. Defining the slope as $-a\gamma_0$, for SFN $a \in (1.55, 2.4)$ is obtained whereas for NPA networks it is $a \in (1.6, 2.7)$, both coherent with [1]. For

[9] We have analyzed intermediate α values and we have seen a continuous transition between two different behaviours.

the sublinear regime $a \in (1.6, 2.35)$ is obtained, and $a \in (0.8, 1.1)$ for the superlinear one. Notwithstanding, the vitality drops suddenly when the system approaches a certain critical one (around $\Phi_C = 0.6$ in all cases with the exception of the superlinear network). This phenomenon can be explained by the collapse of the network when the dead nodes have not enough alive neighbours in order to sustain a repair. In fact, this sudden decay is also seen in empiric life span data and several functions, as the Gompertz distribution [3], have been proposed to approximate this behaviour.

We observe that the superlinear regime network has a completely different behaviour and it is because of its structure, having a central node connected to the other poorly connected ones. This scenario makes the life span of the organism depend only on the central node state. If the central node becomes nonfunctional, the whole organism dies immediately. And as this situation can happen at any moment, the time of death is distributed quite uniformly. This behaviour is much less efficient that the other one observed and it is not seen in nature, so it can be concluded that it has no biological sense.

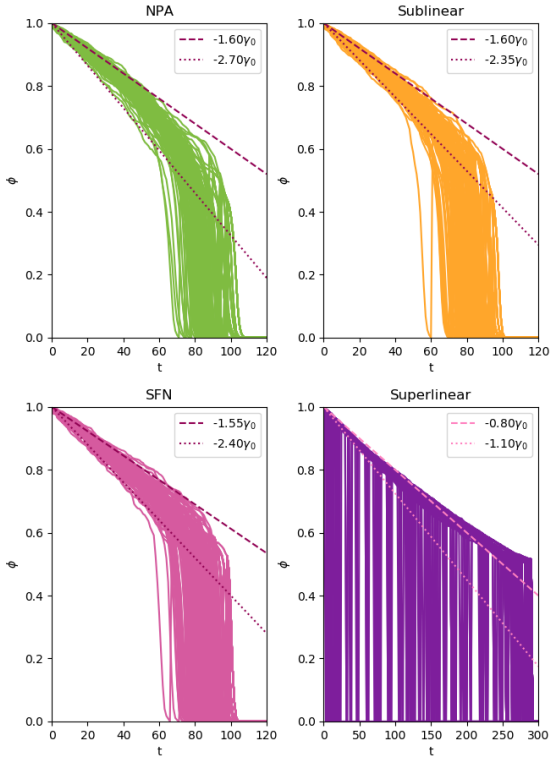

FIG. 1: Individual vitality trajectories, $\Phi(t)$. It includes 100 runs for each type of network: NPA, sublinear, SFN and sublinear; with $N = 2500$ and the maximum and minimum initial slopes $a\gamma_0$. The aging parameters used were $\gamma_0 = 0.0025$ and $\gamma_1 = d = 0$.

Secondly, mortality has been analysed from different points of view by testing the effect of the parameters γ_0 and d but only for the networks showing biological sense (NPA, SFN [1] and sublinear networks). In all cases the mortality rate increases with time, but each parameter affects the resulting evolution in a different way.

An increase of γ_0 shifts the mortality rate $\mu(t)$ left (Fig. 2) and NPA seem to age faster than SFN, staying the sublinear networks in the middle of both behaviours. However, network topology does not seem to be significant in a qualitative point of view. Other than that, an increase of the initial damage, d, results in an increasing infant mortality but with an efficient repair converging all cases at the same point in the high time-steps region (Fig. 2). This efficient repair can be seen in the small slope that these organisms show if it is compared to the one showed by other organisms with a lower initial damage. For this reason, it can be noticed that organisms with a high d age slower than the ones having a low d. However, it does not matter which d a certain population has that at the end all converge to the same μ in accordance with the Strehler-Mildvan correlation law [1, 6]. It is remarkable to say that in this case NPA, sublinear networks and SFN also behave in a similar way.

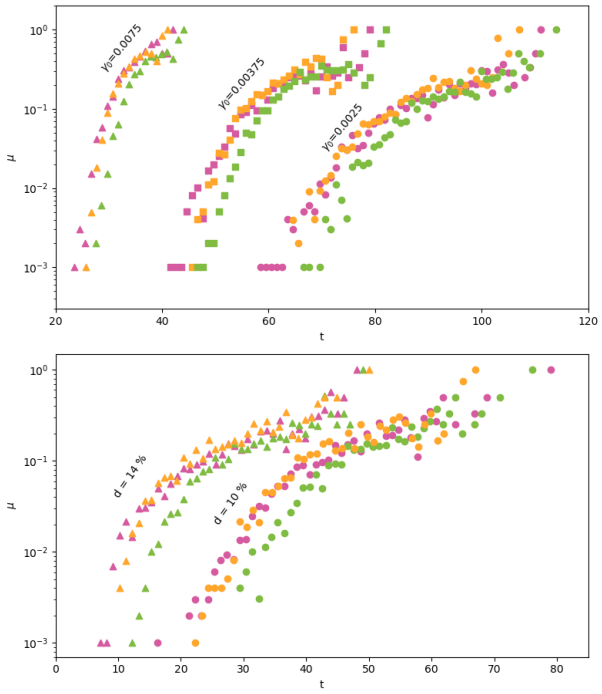

FIG. 2: Mortality μ as a function of time. It is an average of 1000 simulations using SFN (pink markers), sublinear networks (orange markers) and NPA (green markers). Up: effect of γ_0 when N, γ_1 and d remain constant with values $\{N, \gamma_1, d\} = \{2500, 0, 0\}$. Down: effect of d when γ_0, γ_1 and N remain constant with values $\{N, \gamma_1, \gamma_0\} = \{2500, 0.0025, 0.0025\}$.

Next, the interdependence parameter $\lambda(t)$ has been studied (Fig. 3) bearing in mind that $\lambda = 1$ corresponds to a total independence scenario, like a set of

disconnected nodes. As it can be observed, for NPA, SFN and sublinear networks there is a abrupt change when a certain amount of damage is accumulated. It is a reflection of the collapse experienced by the network resulting in a sudden drop of the vitality shown in figure 1. For short times there is a huge variation of interdependence and it reaches values < 1, what means that $\phi_0 > \phi$ and it is not expected. As these fluctuations decrease when N increases, we can conclude that they are caused because we are quite far from the limit $N \to \infty$. Anyway, it is clear that the node's interaction makes the effect much more remarkable by giving a sudden collapse. However, superlinear networks are an exception as always but for the same reason as before. In this case, before the collapse there is an independence scenario ($\lambda \sim 1$), because all nodes are connected to the central one but barely to others. And the collapse comes when the central node dies because its huge interaction with all the other nodes, and not when a certain damage accumulation is reached as in the other cases. And as the central node can die at any moment, the collapse time does not converge toward any concrete limit.

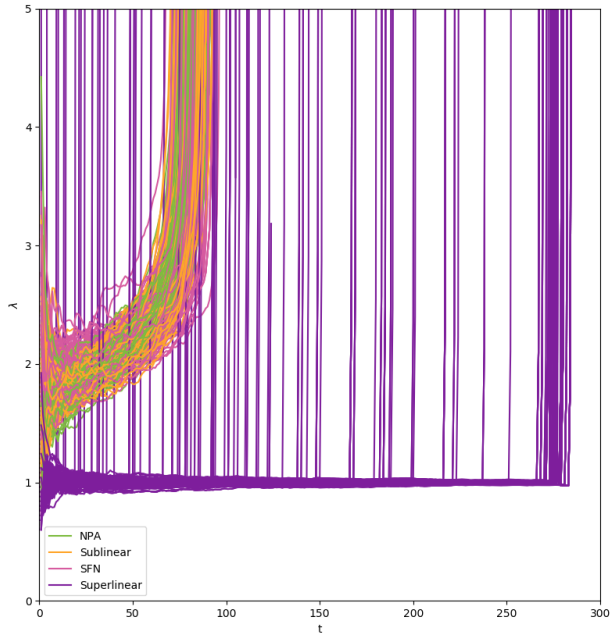

FIG. 3: Interdependence parameter λ as a function of time-steps including 1000 simulations with $N = 10^4$.

By the same token, how do the failure rate γ_0 and the repair rate γ_1 affect the longevity of an organism has been studied. As it can be seen in figure 4, it may exist a critical repair rate γ_1^* for which the average time of death diverges and the beginning of this growth happens in a certain γ_1 value that seems to depend on the network type, being smaller for SFN. Additionally, it is interesting to see that all the studied networks show a remarkable similarity despite their different structure.

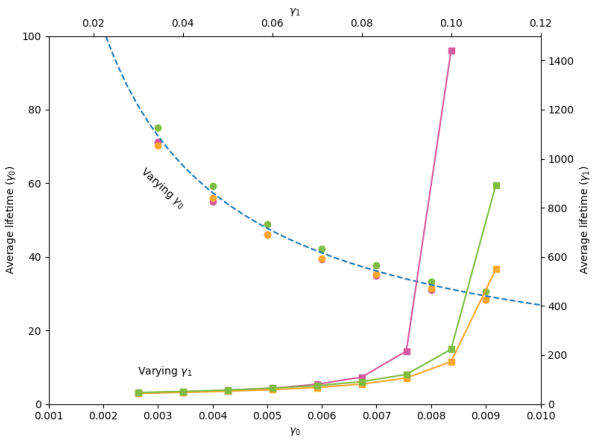

FIG. 4: Average lifetime versus failure rate γ_0 (circles) and repair rate γ_1 (squares) for NPA networks (green), sublinear networks (orange) and SFN (pink) with N = 2500. The dashed line in blue mark $\tau = 0.5941/\gamma_0^{0.828}$, which is the result of the curve fitting of the points representing $\tau(\gamma_0)$. It is important to mention that $\gamma_1 = 0$ is kept constant while varying γ_0; and $\gamma_0 = 0.0065$ while modifying γ_1.

IV. MATHEMATICAL ANALYSIS

For the purpose of understanding the sudden collapse seen in figures 1 and 3, the change in the fraction of alive nodes has been analyzed. When the system is far from collapse, the probability that a node dies (p_0) is the sum of the probability of its own failure (γ_0) and the probability of its last provider/neighbour death. Considering $m(\Phi)$ as the probability of having only one single provider, $p_0 = \gamma_0 + m(\Phi)p_0(1-\gamma_0)$. Then, being the reparation probability $p_1 = h(\Phi)\gamma_1$, where $h(\Phi)$ is the probability of having the minimum number of neighbours needed. Subsequently, the change in the fraction of alive nodes is defined as:

$$\begin{aligned}\Delta\Phi &= p_0\Phi - p_1(1-\Phi) \\ &= -\frac{\gamma_0\Phi}{1-m(\Phi)(1-\gamma_0)} + \gamma_1 h(\Phi)(1-\Phi)\end{aligned} \tag{4}$$

From this expression the collapse can be understood. It is clear that $m(\Phi)$ and $h(\Phi)$ change depending on the network structure. However, $m(\Phi) \in [0,1]$ and increases towards 1 as Φ decreases. This means that the first term, which remains negative, finally dominates the second one resulting in a sudden drop in Φ. But from this equation the lifetime divergence seen in figure 4 can also be understood. Setting $\Delta\Phi = 0$ (immortality condition), the repair rate γ_1 needed is:

$$\gamma_1^* = \gamma_0\Phi^*/\{h(\Phi^*)(1-\Phi^*)[1-m(\Phi^*)(1-\gamma_0)]\} \tag{5}$$

where $\Phi^* \in [\Phi_C, 1]$. In this case the vitality of the system decreases until it reaches a constant value Φ^*. However, it is only possible in the thermodynamic limit, so every

finite system will die at least due to statistical fluctuations. [1]

It is also interesting to obtain the exact form of the average initial slope with which the vitality decreases when $\gamma_1 = 0$, the average of the slopes represented in Fig. 1. If the death probability of a node with k neighbours is $\sigma^{(k)}$, for $k = 1$ it can be expressed as follows:

$$\sigma^{(1)} = \gamma_0 + P(1,1)\sigma^{(1)} + P(1,2)\sigma^{(2)} + ... \quad (6)$$

where $P(1,l)$ is the probability that a node has only one neighbour, and that the neighbour itself has l neighbours; and γ_0 corresponds to the probability that the node dies for any reason independent of its connectivity. If probabilities of order γ_0^2 are neglected, only nodes with degree 1 can die because of the death of their neighbours. For this reason, $\sigma^{(k)} = \gamma_0$ can be considered for $k > 1$:

$$\sigma^{(1)} = P(1,1)\sigma^{(1)} + \gamma_0(1 + P(1,2) + P(1,3)...) \quad (7)$$

Considering $\sum_i P(1,i) = P(1)$, $\sigma^{(1)} = \gamma_0 \frac{2-P(1,1)}{1-P(1,1)}$
Then, to obtain the value of the average initial slope an average over the damage rate of all degrees is required:

$$< a > \gamma_0 = \sum_k P(k)\sigma^{(k)}|_{t=0} = \gamma_0 \left(1 + \frac{P(1)}{1 - P(1,1)}\right) \quad (8)$$

Analyzing the networks previously used, $P(1)$ and $P(1,1)$ have been computed. For the SFN, $< a >= 1.81$ is obtained, for the NPA the result is $< a >= 1.73$ (consistent with [1]), for the sublinear it is $< a >= 1.67$ and for the superlinear $< a >= 278.22$. As it can be seen, these results are consistent with Fig. 1 except the superlinear network.
Another appealing value is the critical vitality Φ_C, which is the one that the system has when it collapses. Supposing the existence of only one Φ_C independent of the trajectory of the vitality, the collapse can happen when Φ approaches Φ_C in a given number of time-steps or when there's a huge damage in one single time-step ($\gamma_{0c} = 1 - \Phi_C$). Considering $\gamma_0 \to 1 - \Phi_C$ and letting $\sigma^i \to 1$, $\Phi_C = P(1)$ is obtained.
For the studied networks, $P(1) = 0.495 \sim 0.5$ for SFN, $P(1) = 0,623 \sim 0.6$ for NPA (both coherent with [1]), $P(1) = 0,55$ for the sublinear regime and $P(1) = 0,998$ for the superlinear one. These values are consistent with the critical vitalities obtained in the described aging model (Fig. 1), again excluding the superlinear case.

V. CONCLUSIONS

The purpose of this work is to study how organisms might age by using a model based on damage and repair processes and comparing its results for different types of network. It is noteworthy that aging is a really complex phenomenon with many questions to be answered. This means that the studied model is a simplification of it but the results are coherent with empirical data and allow to come to some conclusions.

The results for NPA and SFN are coherent with [1] and regarding the additional networks studied, it is observed that superlinear networks have no biological sense because of their structure, which makes them inefficient in terms of vitality. And it is important to say that this kind of structures are not seen in nature, what seems consistent with the theory of evolution. However, all other types of network behave in a remarkable similar way and this may be the reason behind the similarity between mortality curves of several model organisms.

Finally, it is seen that for repair rates below the critical one there are no significant differences in the life span. All these results lead to the question of how can it be controlled, if it is possible, and how can an heterogeneous network, with regions having different γ_0 and γ_1, behave. In consequence, it is clear that there is much to be done in understanding a so complex and universal phenomenon as aging. This work has been a great opportunity to learn much more about complex systems and their dynamics, and doing it within the frame of biology has been a nice way to see how remarkable the interdependence between disciplines is.

Acknowledgments

I would like to thank A. Díaz for opening me the door to the complexity world, B. Sanahuja, G. Domènech and F. Mescia for their guidance and advice and specially my family for their unconditional support through this journey.

[1] D.C. Vural, G. Morrison, L. Mahadevan, *Aging in complex interdepency networks*, Phys. Rev. E **89**, 022811 (2014).
[2] S. Farrell, G.Stubbings, A. Rutenberg *et al.*, *The potential for complex computational models of aging*, Mechanisms of Aging and Development 193, 111403 (2021)
[3] N. Stroustrup, *Measuring and modeling interventions in aging*. Curr. Opin. Cell Biol. 55, 129–138 (2018).
[4] E.D. Sun, T.C.T. Michaels, L. Mahadevan, *Optimal control of aging in complex networks*, PNAS 117 (34) 20404-20410 (2020).
[5] A.L. Barabási, *Network Science*, Cambridge University Press, 2016.
[6] B.L. Strehler, A.S. Mildvan, *General Theory of Mortality and Aging*, Science **132**, 14 (1960)
[7] S. Horiuchi, Pop. Dev. Rev **29**, 127 (2003)
[8] G. Caughley, *Ecology* **47**, 906 (1966)