Dynamic Graphs Generators Analysis: An Illustrative Case Study

Vincent Bridonneau
Université Le Havre Normandie, Normandie Univ, LITIS EA 4108, 76600 Le Havre, France

Frédéric Guinand
Université Le Havre Normandie, Normandie Univ, LITIS EA 4108, 76600 Le Havre, France

Yoann Pigné
Université Le Havre Normandie, Normandie Univ, LITIS EA 4108, 76600 Le Havre, France

Abstract

In this work, we investigate the analysis of generators for dynamic graphs, which are defined as graphs whose topology changes over time. We focus on generated graphs whose orders are neither growing nor constant along time. We introduce a novel concept, called "sustainability," to qualify the long-term evolution of dynamic graphs. A dynamic graph is considered sustainable if its evolution does not result in a static, empty, or periodic graph. To measure the dynamics of the sets of vertices and edges, we propose a metric, named "Nervousness," which is derived from the Jaccard distance. As an illustration of how the analysis can be conducted, we design a parametrized generator, named D3G3 (Degree-Driven Dynamic Geometric Graphs Generator), that generates dynamic graph instances from an initial geometric graph. The evolution of these instances is driven by two rules that operate on the vertices based on their degree. By varying the parameters of the generator, different properties of the dynamic graphs can be produced. Our results show that in order to ascertain the sustainability of the generated dynamic graphs, it is necessary to study both the evolution of the order and the Nervousness for a given set of parameters.

2012 ACM Subject Classification Theory of computation → Dynamic graph algorithms; Mathematics of computing → Random graphs; Networks → Topology analysis and generation

Keywords and phrases Dynamic Graphs, Graph Generation, Graph Properties, Evolutionary models

Digital Object Identifier 10.4230/LIPIcs.SAND.2023.8

Funding Supported by the French ANR, project ANR-22-CE48-0001 (TEMPOGRAL).

1 Introduction

Nature and human societies offer many examples of systems composed of entities that interact, communicate or are just connected with each other. The Internet, a transportation network, a swarm of robots, an ant colony, a social network, a urban network, or a crowd are some examples [2].

Graphs are certainly one of the best formalism for modeling them. Every vertex in the graph models one entity. A link is added between two vertices when a particular condition about the corresponding entities is verified. For instance: two people are talking to each other, a predator catches a prey, two playing cards are in the same hand, a virus passes from one individual to another, two actors perform in the same play, etc. The semantic of the interaction, communication or connection is proper to the system.

During the last two decades, many works have been dedicated to the study of networks modeling these systems. It has been shown that, unlike classical, regular or random graphs, graphs modeling complex real systems present specific statistical properties, leading researchers to introduce the term of complex networks for naming them. Among the main characteristics that were highlighted are the small-world and the scale-free properties. The small-world property was discovered many years ago, in the 60s, when Stanley Milgram imagined and conducted the "small-world problem" in order to measure, through postal mail, the number of intermediaries between any two persons in the USA [15, 20]. The scale-free property was also observed quite a long time ago for some datasets.

Researchers working on networks are motivated by one key question: "How can we explain the existence of certain properties in these networks?" To answer this, they aim to design generative processes that can produce networks with these properties. In 1998, Watts and Strogatz proposed a rewiring process to generate small-world networks from a regular lattice [21]. In 1999, Barabàsi and Albert introduced the preferential attachment mechanism to create networks with both small-world and scale-free properties [1]. Other works have also used preferential attachment to generate networks with other features, such as soft community structures [9] and navigability [3]. For example, Papadopoulos *et al.* introduced self-similarity [17] as a mechanism that can create soft communities. This new concept aims at connecting entities not because one is popular, but because they share similarities. This work has been possible using hyperbolic geometry which helps embedding both popularity (the preferential attachment) and similarity [19, 13]. Other models using such a geometry have then been investigated creating networks with more properties, such as the Geometric Preferential Attachment [22] and the nonuniform Popularity-similarity Optimization model [16]. One relevant point to highlight is that such processes generate growing networks: at each time step a new vertex is added to the graph and is more likely linked to high degree vertices. Other mechanisms exist where the amount of vertices does not change over time. For instance, this is the case of edge-markovian processes where only edges are changed over time [6]. Other models such as Erdos-Renyi evolution model [7] or the configuration model [4] can also be seen as such a model.

However many real-world systems are composed of a varying number of entities (increasing and decreasing). For instance, living being populations may see the number of individuals increase during some periods and decrease for some other periods [14]. And, to our knowledge, generative models with such a characteristic have not been deeply studied. The work developed in this paper aims at addressing this more general case where the set of vertices, between two consecutive time steps, may either increase or decrease or may change while keeping the same cardinality. To provide an example of the latter case, consider the interaction between players during a game of rugby. The total number of players on the field remains constant at 30 (assuming no exclusions); therefore, the order of the players is fixed. Nevertheless, substitutions occur throughout the game.

Given a generative process, questions asked are: "how the dynamics of generated graphs can be characterized?"; "what metrics might be used for that purpose, and how to compute them?"; "from the point of view of dynamics, is it possible to classify or gather generators into classes or families?". In this ongoing work, not all questions are addressed. But we hope it will be a milestone for carrying out analysis of dynamic graphs generators. For this, in Section 2 a generic model of generators is presented and discussed. It is followed by the definition of a novel notion based on a specific metric, both targeting the dynamics of the graphs. Finally, the Degree-Driven Dynamic Geometric Graph Generator (D3G3) is presented. D3G3 is a parameterized generator and according to the parameters, it can produce a wide variety of dynamics. It will be used as a case study. A first global analysis of the generated graph families is performed in Section 3. Section 4 focuses on specific values of the parameters and present a rigorous analysis of the evolution of the dynamics of the graph and of the likelihood of its sustainability. A conclusion, drawing some perspectives and future investigations, closes temporarily this work.

2 Definitions and Generative Model and Definitions

2.1 Notations

Consider two sets A and B:

- $\triangle$ operator: $A \triangle B$ is defined as $A \cup B - A \cap B$. For instance if $A = \{1, 2, 3, 4, 5\}$ and $B = \{4, 5, 6, 7, 8\}$ then $A \triangle B = \{1, 2, 3, 6, 7, 8\}$
- Consider G a dynamic graph, $G_t = (V_t, E_t)$ denotes the state of the graph at time t, where V_t is the set of vertices and E_t is the set of edges.
- $|V_t|$ (resp. $|E_t|$) corresponds to the number of vertices (resp. edges) of graph G_t.
- For simplifying notations in the document, $|V_t|$ is often denoted by n_t
- $G = (V, E)$ is said to be a null graph if both $V = \emptyset$ and $E = \emptyset$. In the report such a null graph can also be called an empty graph.
- Let a and b to real numbers such that $a < b$. Then $[a, b]$ refers to as a closed interval of real number. Both end points belong to the interval. The open interval (a, b) represents the same object, but end points are not included. A half-open interval $[a, b)$ is an interval including the endpoint a but not b. The $(a, b]$ one includes b but not a.
- The set $\mathbb{N}$ refers to as the set of non-negative integers ($\{0, 1, \dots\}$). The set of positive integers is referred to as $\mathbb{N}^*$ ($\{1, 2, \dots\}$).

2.2 Position of the work with respect to Temporal Networks

Current definitions of temporal networks (TN) include *time-varying graphs* [18], *temporal networks* [12], *evolving graphs* [8], etc. They all define structures described by sequence of static graphs, ordered by a timestamp (e.g., $G = (G_i = (V_i, E_i))_{i \geqslant 0}$) where i refers to the time step). It is worth mentioning that TN definitions do not include information about the generative process. Thus, the way the graph at time $t + 1$ is obtained from the graph at time t is not described. In this report, the emphasis is precisely on the study of generative processes. This work is therefore positioned upstream of TN. In the sequel graphs produced by generators are called dynamic graphs or simply graphs.

2.3 Generalities

From a general point of view, a dynamic graph generator can be defined as a process with input data, that produces at each time step $t+1$ a new static graph G_{t+1}. It is produced from already generated static graphs $\{G_1, \dots, G_t\}$ and possibly additional information. Thus, the output of a dynamic graph generator is a flow of static graphs identified by time stamps. The time stamps may also corresponds to events, and in such a case, the time interval between two time stamps may be different. However, in this report, for sake of clarity, we consider integer time stamps. If the flow stops, for whatever reason (e.g. clock has been stopped, evolution process is finished) at step T, the set of generated static graphs $\{G_1, G_2, \dots, G_T\}$ corresponds to a temporal network (TN).

2.4 Sustainability

The goal of this section is to introduce a novel notion for qualifying the dynamics of a graph. Only measurement of the order (or of the density) of the graph is not enough for qualifying its dynamics. For instance, if a dynamic graph becomes static, all vertices remain the same and the graph order does not change. Conversely, if between two consecutive time steps all vertices are replaced by new ones, the order also remains the same, but the dynamics is different. Sustainability qualifies a dynamic graph that never becomes null or periodic (which includes static). A graph owing the sustainability property is said sustainable.

▶ **Definition 1** (Graph sustainability). *A dynamic graph G is said sustainable if both Condition 1 and Condition 2 are not verified.*

Condition 1: $\exists T \in \mathbb{N}, \forall t \geqslant T, G_t = (\emptyset, \emptyset)$
Condition 2: $\exists T \in \mathbb{N}$ *and* $\exists k \in \mathbb{N}^*, \forall t \geqslant T, G_t = G_{t+k}$

Some well-known graph generators produce sustainable dynamic graphs. For instance generators of growing networks. Indeed, for all $t \in \mathbb{N}$, $|V_t| > |V_{t-1}|$, and $G_t \neq (\emptyset, \emptyset)$. For these generators, graph sustainability is obvious and does not require any analysis.

Unlike these cases, some generators are based on mechanisms making the evolution of the vertices (and edges) more difficult to predict, and the dynamics is worth studying. For that purpose, we propose to consider a metric enabling a quantification of the dynamics.

2.5 Nervousness

This metric provides a way of measuring the dynamics of a graph. Note that this metric is derived from the Jaccard distance, which can be defined as one minus the coefficient of community as outlined in [11]. However, in the context of dynamic graphs it seems to us more meaningful to call it nervousness. This metrics is defined at the level of vertices, edges and at the level of the graph. In this work, only vertices nervousness is defined. It is different from the burstiness which is defined at the node/edge level during the lifetime of the graph [10] and aims at representing the frequency of events occurring on each node/edge. Nervousness metric aim is to capture the dynamics of creation and deletion of nodes and edges between two time steps at graph level.

▶ **Definition 2** (Vertices Nervousness). *Given a dynamic graph G, such that at time t $G_t = (V_t, E_t)$. We call* **vertices nervousness** *at time t and denoted by $\mathcal{N}^V(t)$, the ratio:*

$$\mathcal{N}^V(t) = \frac{|V_{t+1} \triangle V_t|}{|V_{t+1} \cup V_t|} = \frac{|V_{t+1} \cup V_t - V_t \cap V_{t+1}|}{|V_{t+1} \cup V_t|}$$

This metric is complementary with the graph order measure. Indeed, graph order can remain constant between two consecutive time steps although some vertices change. If all vertices are replaced, nervousness equals 1. If all vertices are kept, nervousness is 0. Similarly we define the edges nervousness as $\mathcal{N}^E(t) = \frac{|E_t \triangle E_{t+1}|}{|E_t \cup E_{t+1}|}$. Accordingly, Graph Nervousness is defined as $\mathcal{N}^G(t) = (\mathcal{N}^V(t), \mathcal{N}^E(t))$.

For illustrating this definition, consider the following cases for a dynamic graph, from t to $t+1$. We denote $|V_t| = n_t$. We also assume that between t and $t+1$ the order remains the same, thus $|V_{t+1}| = n_{t+1} = n_t = n$.

- if all vertices are replaced:

$$\mathcal{N}^V(t) = \frac{|V_t \triangle V_{t+1}|}{|V_t \cup V_{t+1}|} = \frac{2n}{2n} = 1$$

- if half of the vertices are replaced: $\mathcal{N}^V(t) = \frac{3n/2 - n/2}{3n/2} = \frac{n}{3n/2} = \frac{2}{3}$
- if the vertices remain the same, the union of the sets is equal to their intersection thus: $\mathcal{N}^V(t) = 0$

When the order changes, for instance if all vertices are duplicated, thus $|V_{t+1}| = 2n_t = 2n$: $\mathcal{N}^V(t) = \frac{2n-n}{2n} = \frac{1}{2}$

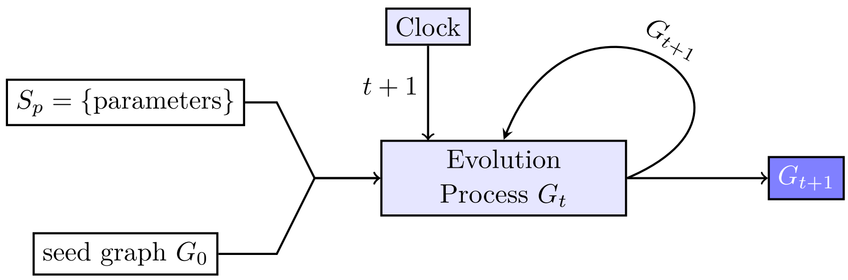
2.6 Sustainability vs Nervousness

Sustainability and nervousness are closely related. Sustainability describes a dynamic graph property while nervousness enables the measure of the evolution of vertices and edges sets between two consecutive time steps. When nervousness is null for both sets, the graph is static and thus does not have the sustainability property. Consider a dynamic graph G, if for all $t \in \mathbb{N}$, $\mathcal{N}^V(t) \neq 0$ or $\mathcal{N}^E(t) \neq 0$, then G holds the sustainability property, except if G is periodic.

2.7 D3G3: definition

In this section we define a parameterized model generating families of dynamic graphs. An instance of the generative model is defined by a set of parameters. For studying the model, we analyze, according to the parameters set, the dynamic graphs families produced and rely on both the sustainability and the nervousness for that purpose.

The generator has two types of inputs: a set of parameters, S_p, and an initial graph, called seed graph and denoted G_0. At each time step $t+1$ it produces, from the previous graph G_t, a new graph G_{t+1} as illustrated on the figure.



Graphs produced by D3G3 are geometric graphs. A geometric graph is defined by an euclidean space and a threshold d. For this study, without loss of generality we consider a 2D-unit-torus (i.e., a square $[0;1)^2$ where the two opposite sides are connected). Each vertex is characterized by a set of coordinates, such that given two vertices u and v it is possible to compute their euclidean distance: $dist(u, v)$. Given V the set of vertices, the set of edges E is defined in the following way: $E = \{(u, v) \in V^2 \mid \mathrm{dist}(u, v) \leqslant d\}$. It is important to notice that borders of the square modeling the torus are connected. Therefore considering one node on the torus, the value of d for which the surface of the disk of radius d centered on this node reaches its maximum for $d = \sqrt{2}/2$. It represents the half diagonal of the square.

Graphs generated by D3G3 are produced thanks to an evolution process. This mechanism is parameterized by an initial graph (the seed graph) and by two transition rules driving the evolution of the graph between two consecutive time steps. Apart from a random generator, no external decision or additional information is used by this mechanism. Rules are based on node degrees only and rely on a random generator for positioning new nodes in the 2D euclidean space. This leads to the name of the generator: *Degree-Driven Dynamic Geometric Graphs Generator* or D3G3.

From now, graphs we are studying are referred to as sequences of static graphs $(G_t = (V_t, E_t))_{t \geqslant 0}$, where $t \geqslant 0$ is the time step. The initial graph, G_0 $(t = 0)$ is called the seed graph.

▶ **Definition 3** (Degree Driven Dynamic Geometric Graph Generator). *An instance of D3G3 is defined by an initial graph, a set of parameters and two rules:*

- $G_0 \neq (\emptyset, \emptyset)$ *the seed graph,*
- *parameters:*
  - $d \in (0, \frac{\sqrt{2}}{2})$ *(distance threshold for connection),*
  - S_S *a set of non-negative integers*
  - S_C *a set of non-negative integers*
- *rules applied on* G_t *leading to* G_{t+1}*:*
  - *if* $v \in V_t$*, then* $v \in V_{t+1}$ *if and only if* $deg(v) \in S_S$ *(conservation rule)*
  - *if* $v \in V_t$ *and if* $deg(v) \in S_C$ *then add a new vertex to* V_{t+1} *with a random position in the unit-torus (creation rule)*

At a given time step, two nodes are connected if and only if their euclidean distance is lower than d. Graph evolution between two consecutive time steps t and $t+1$, is driven by two rules applied to each vertex $v \in V_t$ simultaneously. The first rule determines for a vertex $v \in V_t$ whether it is kept at step $t+1$ while the second rule concerns the possibility for a vertex $v \in V_t$ to create a new vertex in V_{t+1} according to its degree.

▶ Remark. For generating new vertices by the second rule we had two possibilities. Either we choose, for each new vertex w stemmed from a vertex u, a random position in a finite space, or, we choose a random position close to the position of u with no constraint on space limits. We opted for the first option (the space is a unit 2D-torus), and we plan to analyze the differences with the second option.

▶ **Definition 4** (Conserved/Create/Removed/Duplicated nodes). *Let* $t \geqslant 0$ *and* $G = (G_t)$ *a D3G. Let* $u \in V_t$ *and* $v \in V_{t+1}$*, then*

- u *is said to be a conserved node iff* $u \in V_t \cap V_{t+1}$.
- u *is said to be removed iff* $u \in V_t - V_{t+1}$.
- u *is said to be a creator/creating node iff* $deg(u) \in S_C$.
- v *is said to be a created node iff* $v \in V_{t+1} - V_t$.
- u *is said to duplicate iff it is both a conserved and a creator node.*

Once created, a node never change its position. Positions of created nodes do not depend on the creating nodes positions. The position of a created node is chosen randomly and uniformly over the unit-torus.

3 Theoretical Analysis

While the model is very simple, it presents a wide variety of dynamics and long-term evolution. According to S_S and S_C composition, several classes of dynamic behaviors have been identified. These classes have been defined by computing two measures: the evolution of the order of the graph, and the evolution of the Graph Nervousness $\mathcal{N}^G$. Results are reported in Table 1. A detailed analysis of each limit case is available in the report [5].

3.1 General Cases

General cases correspond to all cases for which both S_C and S_S are non empty sets and none of both sets are equal to $\mathbb{N}$. We classify all possible cases according to the tree represented on Figure 1.

The case $S_C = S_S$ composed of consecutive integers will be considered in Section 4. In the present section we consider the cases for which $S_C \neq S_S$.

Table 1 Order, Nervousness evolution and sustainability property for the different cases. n_t denotes the order of graph G_t, $\mathcal{N}^V(t)$ its vertices nervousness and $\mathcal{N}^G(t)$ the graph nervousness.

S_S \ S_C	$\mathbb{N}$	Finite set	$\emptyset$
$\mathbb{N}$	$\forall t, n_t = 2^t n_0$ $\forall t, \mathcal{N}^V(t) = 0.5$	$\forall t, n_{t+1} \geqslant n_t$ $\forall t, 0 \leq \mathcal{N}^V(t) \leq 0.5$ $\lim_{n_t \to \infty} P(n_{t+1} > n_t) = 0$	$\forall t, G_t = G_0$ $\forall t, \mathcal{N}^G(t) = (0, 0)$
	Sustainable	Asymptotically non sustainable	Non sustainable
Finite set	$\forall t, n_{t+1} \geqslant n_t$ $\forall t, 0.5 \leq \mathcal{N}^V(t) \leq 1$ $\lim_{n_t \to \infty} P(n_{t+1} > n_t) = 0$	General cases (see Section 3.1)	$\forall t, n_{t+1} \leqslant n_t$ $\lim_{t \to \infty} n = \text{constant}$ $\lim_{t \to \infty} \mathcal{N}^G(t) = (0, 0)$
	Sustainable		Sustainable
$\emptyset$	$\forall t, n_{t+1} = n_t$ $\forall t, \mathcal{N}^G(t) = (1, 1)$	$\forall t, n_{t+1} \leqslant n_t$ $\forall t, V_t \neq \emptyset \implies \mathcal{N}^G(t) = (1, 1)$	$\forall G_0, G_1 = (\emptyset, \emptyset)$
	Sustainable	Depends on parameters	Non sustainable

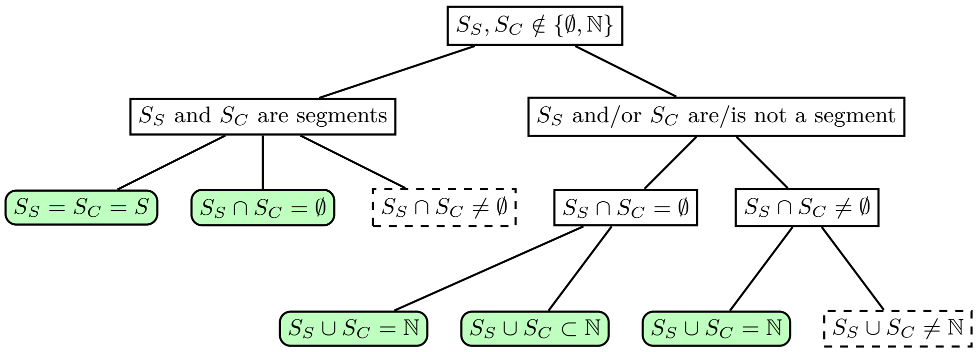


Figure 1 Leaves of the tree represent the general cases. Rounded corners green boxes corresponds to cases for which results are presented in this section and in Section 4. Dashed boxes are cases not covered within this report.

- if $S_C \cap S_S = \emptyset$ and $S_C \cup S_S \subset \mathbb{N}$ then the order of the graph is non-increasing.
- if $S_C \cap S_S \neq \emptyset$ and $S_C \cup S_S = \mathbb{N}$ then the order of the graph is non-decreasing.
- If $S_C \cap S_S = \emptyset$ and $S_S \cup S_C = \mathbb{N}$, then $|V_t| = |V_0|$, the order of the graph is constant.

On the second time they are assumed to cover the whole set of natural integer numbers.

▶ **Theorem 5** (Disjoint sets). *Let $t \geqslant 0$ and $G_t = (V_t, E_t)$ a graph and S_S and S_C two sets of positive integers. If $S_S \cap S_C = \emptyset$, then the series $(|V_t|)_{t \geqslant 0}$ is decreasing.*

Proof. Let consider $(G_t)_{t \geqslant 0}$ a generated graph. Let $t \geqslant 0$ and u be a vertex in V_t. Then, as $S_S \cap S_C = \emptyset$, the degree of node u can't belong to both sets. It follows that vertex u can't be both conserved and a creator. As this holds for every vertex in V_t, the order of generated snapshot graph is not increasing between two consecutive steps. ◀

▶ **Theorem 6** (Union set). *Let S_S and S_C subsets of $\mathbb{N}$. If $S_S \cup S_C = \mathbb{N}$, then the series $(|V_t|)_{t \geqslant 0}$ is increasing.*

Proof. The main argument here is the same used in the proof of theorem 5, except that the degree of every node in V_t belongs to at least one of the two sets S_S and S_C. Therefore, the order of generated snapshot graphs is not decreasing between two consecutive steps. ◀

3.1.1 Partition sets

In this section, S_S and S_C are considered to be a partition of $\mathbb{N}$. This means $S_S \cap S_C = \emptyset$ and $S_S \cup S_C = \mathbb{N}$. From theorems 5 and 6, one can say that for every graph $G_t = (V_t, E_t)$, the series $(|V_t|)_{t \geqslant 0}$ remains steady. Two cases rises from that situations:

- $S_S = \mathbb{N}$ **and** $S_C = \emptyset$: in that case the graph is constant $(\forall t, G_t = G_0)$.
- $S_S = \emptyset$ **and** $S_C = \mathbb{N}$: the series of static graphs $(G_t)_{t \geqslant 0}$ is a series of independent random geometric graph with a constant number of nodes $(\forall t, n_t = n_0)$.

4 Segments

This section focuses on the case $S_C = S_S = S$ where S is a segment (i.e., an interval of consecutive integers).

4.1 Model and conjecture

In this section parameters S_S and S_C are limited to equal sets of consecutive integers. Both sets are such that $S_S = S_C = [m, M]$ (called segments), where $m, M \in \mathbb{N}^2$ and referred to as S in the following. The evolution of graph order for different values of parameters m and M is investigated. Some statements and properties are theoretically and experimentally proved for the special case $S = \{0\}$. A relationship between graph order at a step $t + 1$ and at step t and an upper bound for n_t $(t > 0)$ are given. Then, a theoretical analysis of the general case is provided, and a new concept named sustainable interval is introduced. In the last part of this section, vertices nervousness of graphs is studied through experimentation. It is shown to be equal in average to $\frac{2}{3}$. The reason behind this particular value will be explained in this last part.

4.2 $S = \{0\}$

The case $S_S = S_C = S = \{0\}$ is considered in this section. The seed graph, G_0, is supposed to be a random geometric graph whose order is arbitrarily chosen. The main result about this case is an estimation of the mean value of graph order. An approximation for small values of the distance threshold d is provided.

▶ **Theorem 7** (Expected value of graph order). *Let $S = \{0\}$, $d > 0$ and $G_0 = (V_0, E_0)$ such that there exists at least one node $u \in V_0$ being isolated (i.e., $deg(u) = 0$), then either the graph becomes empty, or the average number of conserved nodes is $l(d) = 1 - \frac{\log\left(\frac{\sqrt{1+4\alpha}-1}{2}\right)}{\log \alpha}$ with $\alpha = \frac{1}{1-p}$ and $p = p(d)$.*

Proof. Let $t \geqslant 1$. Two cases are to be discussed: the case of conserved vertices from step $t - 1$ to step t $(V_t \cap V_{t-1})$ and the case of created nodes at step t $(V_t - V_{t-1})$. As the number of created nodes is the same as the number of conserved nodes from $t - 1$ to t, we set $c_t = |V_t \cap V_{t-1}| = |V_t - V_{t-1}|$.

First let's study the number of conserved vertices from step t to step $t + 1$ among those conserved from step $t - 1$ to step t. $c_{t+1}^{\text{conserved}}$ denotes this number. Let $u \in V_t \cap V_{t-1}$. The probability for u to be conserved is the probability that its degree to created nodes remains equal to 0.

$$\deg(u) = \sum_{v \in V_t - V_{t-1}} X(u, v)$$

Let $v \in V_t - V_{t-1}$. As in the previous section, $X(u,v) \sim B(p)$ and $\deg(u) \sim B(c_t, p)$ as a sum of independent Bernoulli variables of same parameter p. $Y_t(u)$ denotes the event "u is conserved at step $t+1$". The probability that u survives is $P(Y_t(u) = 1) = P(\deg(u) = 0) = (1-p)^{c_t}$, thus: $Y_t(u) \sim B((1-p)^{c_t})$. Therefore, the number of conserved vertices at step $t+1$ among those conserved at step t is:

$$c_{t+1}^{\text{conserved}} = \sum_{u \in V_t \cap V_{t-1}} Y_t(u)$$

As the position of created nodes are independent from themselves and from conserved vertices, $Y_t(u)$ are independent for all $u \in V_t \cap V_{t-1}$, $c_{t+1}^{\text{conserved}} \sim B(c_t, (1-p)^{c_t})$.

Let's study the number of conserved vertices among created nodes. c_{t+1}^{created} denotes this number. Let $u \in V_t - V_{t-1}$. To study the degree of u, two cases must be studied. The first one is the number of connections between u and all other created nodes (denoted as $\deg^C(u)$). The second one is the number of connections to already present nodes (denoted as $\deg^S(u)$). $\deg^C(u)$ and $\deg^S(u)$ can be obtained using the following formulas:

$$\deg^C(u) = \sum_{v \in V_t - V_{t-1}, u \neq v} X(u,v)$$
$$\deg^S(u) = \sum_{v \in V_t \cap V_{t-1}} X(u,v)$$

As the position of created points on the torus are independent one from the others, $\deg^C(u)$ is a sum of independent Bernoulli variables and therefore, $\deg^C(u) \sim B(c_t - 1, p)$. For $\deg^S(u)$, connections between a created node and an already present node are not independent from each other: knowing u is connected to an already present node means it is close to it and as other conserved nodes are farther than d, it implies that $\deg^S(u)$ is not a sum of independent Bernoulli variables. However, as a first approximation, this quantity will be considered as a sum of independent Bernoulli variables.

Thus, the computation of the expectation of $c_{t+1} = c_{t+1}^{\text{conserved}} + c_{t+1}^{\text{created}}$ gives:

$$c_{t+1} = c_t(1-p)^{c_t} + c_t(1-p)^{2c_t - 1}$$

By looking for a limit to this series gives $l \geqslant 0$ satisfying:

$$l = l(1-p)^l + l(1-p)^{2l-1}$$

Solving this equation gives $l = 0$ or :

$$l = 1 - \frac{\log\left(\frac{\sqrt{1+4\alpha}-1}{2}\right)}{\log \alpha} \qquad \text{with } \alpha = \frac{1}{1-p}$$ ◀

Experiments have been run to see if this relationship holds.

▶ **Corollary 8.** *Let $d > 0$ and $l(d)$ as defined in the Theorem 7. Then for small values of d:*

$$l(d) \sim -\frac{\log\left(\frac{\sqrt{5}-1}{2}\right)}{\pi d^2} = \frac{\log \phi}{\pi d^2}$$

where ϕ is the golden ratio $\left(\frac{1+\sqrt{5}}{2}\right)$.

Proof. Let $d > 0$ be small. Thus, applying Taylor expansion gives $\frac{1}{1-\pi d^2} \sim 1 + \pi d^2$ and $\log\left(\frac{1}{1-\pi d^2}\right) \sim \pi d^2$. The numerator comes from $4 \cdot \frac{1}{1-\pi d^2} \simeq 4$. The golden ratio is obtained using operations on log and by noticing that $\frac{2}{\sqrt{5}-1} = \frac{2(\sqrt{5}+1)}{4} = \phi$, the golden ratio. Combining these results leads to the statement of the corollary. ◀

It is therefore possible to state that, in the case where $S = \{0\}$, it is possible to theoretically get an expectation of graph order as well as to get an upper bound for graph order depending on parameter d.

Experiments have been performed to see whether the expected value graph order holds. These experiments has been performed for different values of threshold d. A linear regression shows that the relationship holds with a R^2 of more than 0.99.

4.3 The general case

Now the focus is on $S = [m, M]$ for every m and M integers. The goal is to provide a tool aiming at stating, for given parameters m, M and d, whether the graph is likely to be sustainable or not. This part mainly focuses on a simpler model. This model is studied as it helps understanding the evolution of graph order.

4.3.1 Study of graph evolution

In this Section we aim at estimating the evolution of the graph order during graph dynamics. However, in the D3G3 model, between two time steps, non-conserved nodes are removed from the graph and conserved nodes are located at the same position, which entails a remanent graph. This remanent graph induces a structure influencing the computation of graph order. More precisely, nodes that are about to be removed connected to conserved ones interfere in the probability that conserved nodes at time t are still conserved at time $t+1$. This is linked to computing the degree of the neighbors of a node u knowing the degree of node u. To our knowledge, this is a difficult question. For that purpose, a relaxed version of the D3G3 model is considered enabling analytical study of this evolution. In this model, conserved nodes are moved (i.e., their position are changed) such that obtained graph is a new random geometric graph at each step. We call this model "the redistributed model". This will help us proving the following theorem:

▶ **Theorem 9.** *Let $G = (G_t)$ be a dynamic graph obtained with the redistributed model, then at every step t, $\frac{n_{t+1}}{2} \sim B(n_t, p(S, d, n_t))$, where $p(S, d, n_t)$ is the probability that a node is conserved between step t and $t+1$:*

$$p(S, d, n_t) = \sum_{k=m}^{M} \binom{n_t - 1}{k} p^k (1-p)^{n_t-1-k}$$

Here, $p(d)$ refers to the probability for two different nodes to be connected (i.e., the probability that the distance between them is lower than or equal to d), which is, for $d \leqslant \frac{1}{2}$, πd^2.

Proof. In the redistributed model, at time step t a RGG (G_t) is built. If the graph order at time t is equal to n_t, the graph order at $t+1$ is equal to twice the number of surviving nodes at time t. As every node has an independent position in the torus, this probability is the same for all nodes. Let's denote it $p(S, d, n_t)$. Let $u \in V_t$. Then:

$$p(S, d, n_t) = P(\deg(u) \in S) = \sum_{k=m}^{M} P(\deg(u) = k) = \sum_{k=m}^{M} \binom{n_t - 1}{k} p^k (1-p)^{n_t-1-k} \quad (1)$$

Assuming one node is a conserved node, it does not affect the probability of conservation for other nodes. The number of conserved nodes can be computed summing independent Bernoulli's events of parameter $p(S, d, n_t)$. This gives $\frac{n_{t+1}}{2}$ follows a binomial distribution of parameter n_t and $p(S, d, n_t)$. ◀

Computing expectation for a binomial distribution leads to an expectation for n_{t+1} knowing n_t. Indeed, this expectation is $2n_t p(S, d, n_t)$. For a fixed set S, this provides a relationship between n_t and n_{t+1}:

▶ **Definition 10** (Expectation of graph order)**.** *Let m, M and d be parameters for the redistributed model. Let $G = (G_t)$ be an obtained graph with such parameters. Then, the expectation of graph order at step $t+1$ (n_{t+1}) knowing graph order at step t (n_t) is $f_{S,d}(n_t)$ and satisfies $n_{t+1} = f_{S,d}(n_t) = 2n_t p(S, d, n_t)$, and then:*

$$\forall n \in \mathbb{N}, f_{S,d}(n) = 2np(S, d, n) \tag{2}$$

This quantity is referred to as the relationship in the sequel. Studying the relation for every value of m, M and d turns out to be a difficult problem. However some results may be conjectured. A first conjecture concerns the variations of the relationship:

▶ **Conjecture 11.** *Let m, M and d be parameters of the model. Let $S = [m, M]$ and $f_{S,d}$ the relationship as defined above. Then there exists $n_* \in \mathbb{N}$ such that $f_{S,d}$ is increasing on $[0, n_*]$ and decreasing on $[n_* + 1, +\infty[$.*

This conjecture is difficult to prove due to the sum involved in the computation of $f_{S,d}$. However, it is not necessary to study the relationship for all integers. It is possible to perform the study on a limited interval. This is the purpose of theorem 13 (below). But before proving this theorem, it is necessary to provide another formulae computing variations of $f_{S,d}$:

▶ **Lemma 12.** *Let m, M and d be parameters of the model. Let $\Delta f_{S,d}$ defined as the variation of $f_{S,d}$: for $n \in \mathbb{N}, \Delta f_{S,d}(n) = f_{S,d}(n+1) - f_{S,d}(n)$. Then:*

$$\forall n \in \mathbb{N}, \Delta f_{S,d}(n) = 2\sum_{k=m}^{M}(k+1)\binom{n}{k}p^k(1-p)^{n-1-k}\left(1 - \frac{n+1}{k+1}p\right)$$

Proof. Let m, M and d be parameters of the model. Let n a be non-negative integer. This proof only focuses on the terms of the sum of $\Delta f_{S,d}$:

$$\begin{aligned}\Delta f_{S,d}(n) &= 2\left(\sum_{k=m}^{M}(n+1)\binom{n}{k}p^k(1-p)^{n-k} - n\binom{n-1}{k}p^k(1-p)^{n-1-k}\right) \\ &= 2\sum_{k=m}^{M}p^k(1-p)^{n-1-k}\left((n+1)\binom{n}{k}(1-p) - \binom{n-1}{k}\right)\end{aligned}$$

Let $k \in \mathbb{N}$ such that $m \leqslant k \leqslant M$. Every term of the sum of $\Delta f_{S,d}$ can be expressed as follow only using results on binomial coefficients:

$$\Delta f_{S,d}(n) = 2\sum_{k=m}^{M}(k+1)\binom{n}{k}p^k(1-p)^{n-1-k}\left(1 - \frac{n+1}{k+1}p\right)$$ ◀

It is now possible to state the following theorem about variations of $f_{S,d}$:

▶ **Theorem 13.** *Let m, M and d be the parameters of the model. Let $S = [m, M]$ and $f_{S,d}$ the relationship as defined above. Let $p = p(d)$ be the probability for two different nodes to be connected. Then, $f_{S,d}$ is increasing between 0 and $\frac{m+1}{p} - 1$ and decreasing from $\frac{M+1}{p} - 1$ to infinity.*

Proof. The goal is to prove that $\Delta f_{S,d}(n)$ is positive for $n < \frac{m+1}{p} - 1$ and negative for $n > \frac{M+1}{p} - 1$. To understand this, $\Delta f_{S,d}(n)$ can be rewritten as shown in lemma 12. It is sufficient to notice that, for all $k \in S$, the sign of every single term of the sum is the sign of $\left(1 - \frac{n+1}{k+1}p\right)$. For fixed k, the term is positive if and only if n is lower than $\frac{k+1}{p} - 1$. As this last term is an increasing function of k, all terms of the sum are therefore positive if n is lower than $\frac{m+1}{p} - 1$ and negative if n is greater than $\frac{M+1}{p} - 1$. Hence, the relationship is increasing from 0 to $\frac{m+1}{p} - 1$ and decreasing from $\frac{M+1}{p} - 1$ to infinity. ◀

Thanks to theorem 13, conjecture 11 is proved for intervals $[0, x_m]$ and $[x_M, \infty[$ for $x_m = \frac{m+1}{p} - 1$ and $x_M = \frac{M+1}{p} - 1$. At this stage, quantifying more precisely the evolution of the graph order is not achievable. However, a study of the fixed points of $f_{S,d}$ enables to draw some conclusion about generated graphs sustainability.

4.3.2 Graph evolution and sustainability

First note that knowing the variations of $f_{S,d}$ is not enough to deal with graphs sustainability. Indeed, as claimed by the following theorem, big graphs are not sustainable.

▶ **Theorem 14** (Non-sustainability of big graphs)**.** *Let m, M and d be parameters of the model. Let $f_{S,d}$ be the relationship. Then, there exists $N > 0$ such that for all $n > N, f_{S,d}(n) < 1$.*

Proof. For this proof, it is sufficient to prove that $f_{S,d}(n) \to 0$ when $n \to +\infty$. To do so, $f_{S,d}(n)$ can be rewritten as follow:

$$f_{S,d}(n) = 2n \left(\sum_{k=m}^{M} \binom{n-1}{k} p^k (1-p)^{n-k-1} \right)$$

From theorem 24, the sum tends toward 0 as the product of a polynomial and an exponential, therefore, $f_{S,d}(n)$ is also tending toward 0. ◀

This theorem says that there always exists a graph order limit such that graphs whose order are greater than this limit are likely to become empty. Therefore, it is not possible to obtain sustainable graphs with a large amount of nodes.

A new mathematical concept is now introduced aiming at classifying parameters into three classes. This concept is referred to fixed point and is defined as follow:

▶ **Definition 15** (Fixed Point)**.** *Let m, M and d be parameters of the model. A fixed point for the relationship $f_{S,d}$ is an non-negative integer n such that:*

$$\begin{cases} & f_{S,d}(n) \leqslant n \text{ and } f_{S,d}(n+1) > n+1 \\ or & f_{S,d}(n) \geqslant n \text{ and } f_{S,d}(n+1) < n+1 \end{cases}$$

Such fixed points characterize variation of graph order. Indeed, graph of order n for n taken between two consecutive fixed points is either always decreasing or increasing. From experiment performed on the redistributed model as well as on D3G3, three different cases appear and are conjectured as follow:

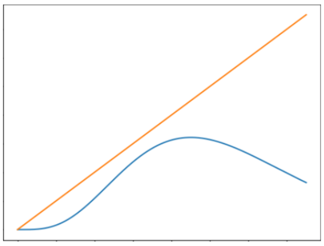
▶ **Conjecture 16.** *For all m, M and d being parameters of the model, the relationship $f_{S,d}$ has either one, two or three fixed points.*

This conjecture is the main tool aiming at studying sustainability in the segment case. Indeed, in the three different cases, it is possible to answer whether a given set of parameters is sustainable or not. However, their is no characterisation about parameters value that may help founding which case parameters lead to. The only one claim that can be made is that d does have an influence on this case.

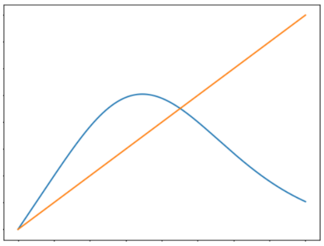
The conjecture 16 is assumed in this subsection. This section aims at stating about sustainability in the three different cases. This is illustrated by a description of the behavior of the relationship $f_{S,d}$ in every case.

4.3.2.1 One fixed point

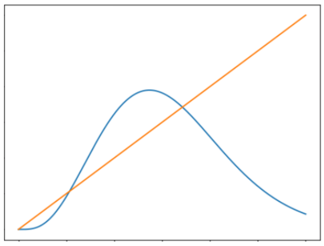
First let's consider the case where the relationship has only one fixed point. When it has only one fixed point, this point is 0. This comes from $f_{S,d}(0) = 0$. Moreover, for all n, $f_{S,d}(n) < n$. As for a snapshot graph of order n_t at step t, $f_{S,d}(n_t)$ gives the expectation value of n_{t+1} at step $t + 1$. Graph orders of generated graphs are decreasing in average. Graphs obtained in this case are therefore not sustainable.



One fixed point.



Two fixed points.



Three fixed points.

4.3.2.2 Two fixed points

For the two fixed points case, 0 is also a fixed point. The argument is also because $f_{S,d}(0) = 0$. The other one is greater than zero. The case where $f_{S,d}$ has two fixed points is assumed to happen if and only if $0 \in S$ and is stated in the following conjecture.

▶ **Conjecture 17** (Characterisation of the two fixed points)**.** *The relationship $f_{S,d}$ has two fixed points if and only if $m = 0$.*

An argument is that a snapshot graph with one vertex becomes empty if and only if $0 \notin S$, that is $m = 0$. Moreover, for $n = 0$, the first term of the sum defining $f_{S,d}$ is equal to 1 so $f_{S,d}(0) = 2$. Graphs generated in such configurations are therefore sustainable as long as their graph order does not exceed a limit. This limit is a consequence of theorem 14. In this case, graphs whose order exceeds the limit are likely to become empty.

4.3.2.3 Three fixed points

For the last case, the goal is to show that graph order is likely to remain bounded. Deeply looking at this case raises the question of values of graph order for which the size is not too large and not too small so that it does not collapse. For that purpose we define an interval, called *sustainable interval*, such that, if the graph order remains within that interval, this ensures the persistence of the graph. This sustainable interval is considered as a tool to study graph sustainability. It concerns expectation of graph order evolution through time. It says that if the image of the function $f_{S,d}$ for all integers within the interval does not exceed the upper bound, then the graph is likely not to collapse. Let's define more precisely this concept:

▶ **Definition 18** (sustainable interval). *Let m, M and d be parameters of the model. Let consider $f_{S,d}$ set such that it has three fixed points. Let N_m be the first positive fixed point and N'_m the smallest integer greater than N_m such that $f_{S,d}(N'_m) \geqslant N_m$ and $f_{S,d}(N'_m+1) < N_m$. The sustainable interval associated to m, M and d is defined as the interval $[N_m, N'_m]$.*

Such an interval satisfies a property about the values $f_{S,d}$ takes when it is restricted to it:

▶ **Theorem 19** (Sustainability in the sustainable interval). *Let m, M and d be parameters of the model. Let assume the relationship $f_{S,d}$ has three fixed points and that $[N_m, N'_m]$ is its associated sustainable interval. If the relationship does not exceed N'_m, then the relationship satisfies:*

$$\forall n \in [N_m, N'_m], f_{S,d}(n) \in [N_m, N'_m]$$

Proof. The lower bound of the interval comes comes from definition of the sustainable interval. The upper bound is a hypothesis of the theorem. ◀

Main interpretation of that theorem is graphs are sustainable in probability in the sustainable interval if and only if there are no values of $f_{S,d}$ that exceed the upper bound of the sustainable interval.

The following paragraphs provide arguments aiming at obtaining the sustainable interval. They also provide arguments to check whether the relationship exceeds the upper bound of the interval. The theorem 13 clearly gives bounds to find out the maximum of the relationship $f_{S,d}$. Three algorithms are sufficient to answer both questions: an algorithm to compute the argument of the maximum of the relationship $f_{S,d}$, an algorithm to find its fixed point between 0 and the argument of the maximum and an algorithm to solve $f_{S,d}(n) = y$ for n greater than the argument of the maximum and $y > 0$ lower than or equal to the maximum. In the following, these algorithms are first implemented. It is then explained how to use them to answer questions about the sustainable interval.

The argument maximum: To compute the argument maximum of the relationship, it is sufficient to study $f_{S,d}$ on the interval $[x_m, x_M]$ for x_m and x_M as defined above. This is a consequence of theorem 13. Let's denote it N_*.

The first positive fixed point: To find the fixed point of $f_{S,d}$ mentioned in the definition of the sustainable interval, it is sufficient to compute the argument maximum of it. The previous algorithm answers this question. Then, as the relationship is increasing from 0 to N_*, it is sufficient to iterate and find an integer n such that $f_{S,d}(n) \leqslant n$ and $f_{S,d}(n+1) > n+1$.

The solution of the equation: For the last algorithm, the goal is to find an integer n such that n is greater than N_* of $f_{S,d}$, $f_{S,d}(n) \geqslant y$ and $f_{S,d}(n) < y$, for a fixed y which is assumed to be positive and lower than the maximum of $f_{S,d}$.

From these algorithms it is possible to implement algorithms stating the existence of the sustainable interval and its bounds. For the existence or not of the sustainable interval, it is sufficient to check whether the maximum of the relationship is greater than its argument. This comes from that sustainable interval exists if and only if there are values of the relationship that exceed their argument. As the relationship is increasing from 0 to $f_{S,d}(N_*)$, then sustainable interval exists if and only if $f_{S,d}(N_*) > N_*$. For computing the sustainable interval boundaries, it is sufficient to know the value of the first fixed point N_m (as it provides the lower bound) and to solve the equation $f_{S,d}(x) = N_m$ as finding the corresponding x to this equation provides the upper bound (N'_m). The existence of N'_m is ensured by theorem 14.

4.4 Vertex nervousness

The goal is to highlight a characterization aspect of the segment family using the vertex nervousness metric. As edge nervousness will not be studied for that case, vertex nervousness will be referred to as nervousness in this section. As in this particular configuration, survivors are the same as created nodes, it is possible to state particular results about the value of nervousness:

▶ **Theorem 20.** *Let S be a segment set of non-negative integer and $d \in (0, \frac{\sqrt{2}}{2})$. Let G be a generated graph of order n_t at step t and number of survivor from step t to step $t+1$ referred to as s_t. Then:*

$$\mathcal{N}^v_t = \frac{n_t}{n_t + s_t}$$

Proof. To prove this result, it is sufficient to notice that $n_{t+1} = 2s_t$, as $S_S = S_C$, which means the number of survivors is the same as the number of created nodes. Thus, applying some basic result about set sizes and noticing that $s_t = |V_t \cap V_{t+1}|$, leads to:

$$|V_t \cup V_{t+1}| = n_t + n_{t+1} - s_t = n_t + s_t$$
$$|V_t \triangle V_{t+1}| = n_t + n_{t+1} - 2|V_t \cap V_{t+1}| = n_t$$

It follows that vertex nervousness is well equal to $\frac{n_t}{n_t+s_t}$. ◀

Result about the nervousness observed in generated graphs parameterized with a segment set S is stated in the following conjecture:

▶ **Conjecture 21.** *Let $m, M \in \mathbb{N}$. Let $S = [m, M]$ and $d > 0$ be parameters of the generator and let $G = (G_t)_{t\leqslant 0}$ be a generated graph. Then the vertex nervousness is in average equal to $\frac{2}{3}$.*

Although this conjecture has not been proved theoretically, experimentation have been performed. They all highlight this conjecture telling that the average nervousness of generated graphs is roughly equal to $\frac{2}{3}$. Results of this experimentation are gathered on picture 2. A possible interpretation of this conjecture and performed experimentation relies on the result stated in theorem 20 and on results from last part. Indeed, if vertex nervousness is close to $\frac{2}{3}$, it means $s_t \simeq \frac{n_t}{2}$. Then, as $n_{t+1} = 2s_t$, it comes $n_{t+1} \simeq n_t$, which means that graph order is close to a fixed point of the relationship $f_{S,d}$ mentioned in the previous section.

5 Conclusion

This paper shows our first investigations in the study of dynamic graph generators. This work concerns a simple generator. As a reminder, the model is parameterized through three variables: a connection threshold d aiming at connecting all points closer than a distance d and two sets S_S and S_C containing non-negative integers. The first one aims at deciding whether a node is kept between two consecutive steps and the second one whether a node is at the origin of a new node at the very next step. Several non-trivial properties are shown

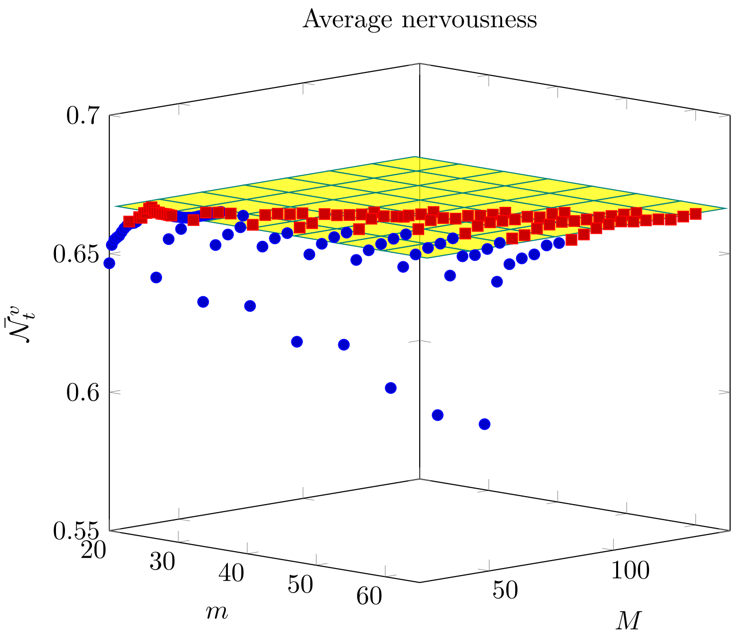


Figure 2 Mean value of nervousness got from experimentation. Points represent the average over 20 runs and 30000 time steps for a single m and M. The yellow surface is the plan of equation $z = \frac{2}{3}$. For all these parameters, d is set to 0.05. Red points represent nervousness of value greater than $\frac{2}{3}$. Blue points represent nervousness of value lower than $\frac{2}{3}$.

about the model. All these properties concern products of the generator. The generator, for a single configuration, produces a family of graphs and not a single graph. Properties are therefore about the whole family of graphs the generator provides for a single configuration. All these properties shown try to answer a single question. This question concerns graph *sustainability*. It is defined as the property, for a given graph obtained with a given seed graph and evolving rules, that the graph becomes neither empty after a finite number of steps nor periodic. Defining this concept for this model is not simple since the evolving rules are not deterministic. It involves probabilistic computations and therefore questions about a possible threshold for which the graph is said to be sustained if the probability of the emptiness of the graph is greater than this threshold. Here the focus has been made on two different metrics, graph order evolution and vertex nervousness, the second one being a renaming of the Jaccard distance metric. Different values of the parameters have been studied, but it has not been possible to try them all as the amount of possible cases is far too big. Cases for which properties have been shown are limit cases, the general case and a very specific case referred to as "segments". Limit cases have led to a first classification when at least one of the two parameter sets is either empty or contains all non-negative integers. General cases highlights some properties for specific values of the two sets. Finally, the case where both sets are equal and contains consecutive non-negative integers has been studied. These sets are called segments. It has revealed theoretical difficulties, especially when computing graph order between consecutive steps. This has led to the creation of a new tool named the "sustainable interval". This tool aims at estimating bounds that frames graph order even though it is not always reliable as probabilities are involved. This last study is only about equal sets. For further studies, the case where both sets are segments but not equal seems relevant as it does not change too much from the segment case.

References

1 Albert-László Barabási and Réka Albert. Emergence of scaling in random networks. *Science*, 286(5439):509–512, October 1999. `doi:10.1126/science.286.5439.509`.

2 S Boccaletti, V Latora, Y Moreno, M Chavez, and D Hwang. Complex networks: Structure and dynamics. *Physics Reports*, 424(4-5):175–308, 2006.

3 Marián Boguñá, Dmitri Krioukov, and K. C. Claffy. Navigability of complex networks. *Nature Physics*, 5(1):74–80, January 2009. Number: 1 Publisher: Nature Publishing Group. `doi:10.1038/nphys1130`.

4 Béla Bollobás. A Probabilistic Proof of an Asymptotic Formula for the Number of Labelled Regular Graphs. *European Journal of Combinatorics*, 1(4):311–316, December 1980. `doi:10.1016/S0195-6698(80)80030-8`.

5 Vincent Bridonneau, Frédéric Guinand, and Yoann Pigné. Dynamic Graphs Generators Analysis : an Illustrative Case Study. Technical report, LITIS, Le Havre Normandie University, December 2022. URL: `https://hal.science/hal-03910386`.

6 Andrea E. F. Clementi, Claudio Macci, Angelo Monti, Francesco Pasquale, and Riccardo Silvestri. Flooding Time of Edge-Markovian Evolving Graphs. *SIAM Journal on Discrete Mathematics*, 24(4):1694–1712, January 2010. `doi:10.1137/090756053`.

7 Paul Erdős, Alfréd Rényi, et al. On the evolution of random graphs. *Publ. Math. Inst. Hung. Acad. Sci*, 5(1):17–60, 1960.

8 Afonso Ferreira and Laurent Viennot. *A Note on Models, Algorithms, and Data Structures for Dynamic Communication Networks*. report, INRIA, 2002. URL: `https://hal.inria.fr/inria-00072185`.

9 Guillermo García-Pérez, M. Ángeles Serrano, and Marián Boguñá. Soft communities in similarity space. *Journal of Statistical Physics*, 173(3):775–782, 2018. `doi:10.1007/s10955-018-2084-z`.

10 K.-I. Goh and A.-L. Barabási. Burstiness and memory in complex systems. *EPL (Europhysics Letters)*, 81(4):48002, January 2008. `doi:10.1209/0295-5075/81/48002`.

11 Paul Jaccard. The distribution of the flora in the alpine zone.1. *New Phytologist*, 11(2):37–50, 1912. `doi:10.1111/j.1469-8137.1912.tb05611.x`.

12 David Kempe, Jon Kleinberg, and Amit Kumar. Connectivity and inference problems for temporal networks. *Journal of Computer and System Sciences*, 64(4):820–842, 2002. `doi:10.1006/jcss.2002.1829`.

13 Dmitri Krioukov, Fragkiskos Papadopoulos, Maksim Kitsak, Amin Vahdat, and Marián Boguñá. Hyperbolic geometry of complex networks. *Physical Review E*, 82(3):036106, September 2010. Publisher: American Physical Society. `doi:10.1103/PhysRevE.82.036106`.

14 Per Lundberg, Esa Ranta, Jörgen Ripa, and Veijo Kaitala. Population variability in space and time. *Trends in Ecology & Evolution*, 15(11):460–464, November 2000. `doi:10.1016/s0169-5347(00)01981-9`.

15 Stanley Milgram. The small world problem. *Psychology today*, 2(1):60–67, 1967.

16 Alessandro Muscoloni and Carlo Vittorio Cannistraci. A nonuniform popularity-similarity optimization (nPSO) model to efficiently generate realistic complex networks with communities. *New Journal of Physics*, 20(5):052002, May 2018. `doi:10.1088/1367-2630/aac06f`.

17 Fragkiskos Papadopoulos, Maksim Kitsak, M. Ángeles Serrano, Marián Boguñá, and Dmitri Krioukov. Popularity versus similarity in growing networks. *Nature*, 489(7417):537–540, September 2012. `doi:10.1038/nature11459`.

18 Nicola Santoro, Walter Quattrociocchi, Paola Flocchini, Arnaud Casteigts, and Frederic Amblard. Time-Varying Graphs and Social Network Analysis: Temporal Indicators and Metrics, February 2011. arXiv:1102.0629 [physics]. `doi:10.48550/arXiv.1102.0629`.

19 M. Ángeles Serrano, Dmitri Krioukov, and Marián Boguñá. Self-similarity of complex networks and hidden metric spaces. *Phys. Rev. Lett.*, 100:078701, February 2008. `doi:10.1103/PhysRevLett.100.078701`.

20 Jeffrey Travers and Stanley Milgram. An experimental study of the small world problem. In *Social networks*, pages 179–197. Elsevier, 1977.
21 Duncan J. Watts and Steven H. Strogatz. Collective dynamics of “small-world” networks. *Nature*, 393(6684):440–442, June 1998. doi:10.1038/30918.
22 Konstantin Zuev, Marián Boguñá, Ginestra Bianconi, and Dmitri Krioukov. Emergence of Soft Communities from Geometric Preferential Attachment. *Scientific Reports*, 5(1):9421, August 2015. doi:10.1038/srep09421.

A Appendices

A.1 Binomial coefficient and Binomial distribution

This section aims at providing results about binomial coefficient and binomial distribution. From now the objective of the two following theorems is to provide asymptotic equivalent of expressions involving a binomial coefficient. The first theorem gives an asymptotic equivalent of $\binom{n}{k}$:

▶ **Theorem 22** (Asymptotic analysis of binomial coefficient). *Let k and n be non-negative integers such that $k \leqslant n$ and k does not depend on n. Then as n tends to infinity, the following holds*

$$\binom{n}{k} \sim \frac{n^k}{k!}$$

Proof. Let k and n as in the above statement. Then, it is sufficient to rewrite the binomial coefficient as follow:

$$\binom{n}{k} = \frac{n!}{k!(n-k)!} = \frac{1}{k!}\prod_{i=0}^{k-1}(n-i) \sim \frac{n^k}{k!} \quad \text{(Asymptotic equivalent of a polynomial)}$$ ◀

The following theorem provides an equivalent of the mass function of a binomial distribution:

▶ **Theorem 23** (Asymptotic analysis of binomial distribution). *Let k and n be non-negative integers such that k does not depend on n. Let $x \in (0,1)$. Then the following limit holds:*

$$\lim_{n\to+\infty}\binom{n}{k}x^k(1-x)^{n-k} = 0$$

Moreover the following equivalent can be expressed:

$$\binom{n}{k}x^k(1-x)^{n-k} \sim \frac{1}{k!}\left(\frac{x}{1-x}\right)^k n^k(1-p)^n$$

Proof. Let k, n and x as defined in the above statement. As stated in theorem 22, an asymptotic equivalent of $\binom{n}{k}$ is $\frac{1}{k!} \times n^k$. Therefore, the following holds:

$$\binom{n}{k}x^k(1-p)^{n-k} = \left(\frac{x}{1-x}\right)^k\binom{n}{k}x^n \sim \frac{1}{k!}\left(\frac{x}{1-x}\right)^k n^k x^n$$

To conclude it is sufficient to notice that $n^k x^n$ tends toward 0 as n tends to infinity (due to $x \in (0,1)$). ◀

▶ **Theorem 24.** *Let $A \subset \mathbb{N}$ be a finite set of non-negative integers. Let n be a non-negative integer and $x \in (0,1)$. Then the following holds*

$$\lim_{n\to+\infty}\left(\sum_{k\in A}\binom{n}{k}x^k(1-x)^{n-k}\right) = 0$$

Proof. As set A is finite, the sum in the statement has a finite number of terms. Let denote $a = |A|$ and $M = \max A$. For values of n such that $n \geqslant 2M$, the following inequality holds:

$$\forall k \leqslant M, \binom{n}{k} \leqslant \binom{n}{M}$$

Moreover, as $1-x < 1$, $y \longmapsto (1-x)^{n-y}$ is increasing. Therefore, for all $k \leqslant M$, $(1-x)^{n-k} \leqslant (1-x)^{n-M}$. It is thus possible to get the following inequality for all $k \leqslant M$:

$$0 \leqslant \binom{n}{k}(1-x)^{n-k}x^k \leqslant \binom{n}{M}(1-p)^{n-M}x^k \leqslant \binom{n}{M}(1-p)^{n-M}$$

Noticing the sum is composed of a elements, it can be bounded as follow

$$0 \leqslant \sum_{k \in A} \binom{n}{k} x^k (1-x)^{n-k} \leqslant a\binom{n}{M}(1-x)^{n-M}$$

As M is fixed an equivalent to right term of the previous inequality as n grows to infinity is:

$$a\binom{n}{M}(1-x)^{n-M} \sim \frac{a}{(1-x)^M} \frac{n^M}{M!e^M}(1-x)^n$$

As $n^M x^n$ tends toward 0, the right term of the equivalent tends toward 0 too. Moreover, the sum is composed of positive elements. Applying the squeeze theorem leads to the wanted limit. ◀

▶ **Theorem 25.** *Let k and n be two non-negative integers. Let $x \in (0,1)$. Then, the following holds:*

$$(n+1)\binom{n}{k}(1-p) - n\binom{n-1}{k} = (k+1)\binom{n}{k}\left(1 - p\frac{n+1}{k+1}\right)$$

Proof. The above statement can be proved with the following equations:

$$\begin{aligned}
(n+1)\binom{n}{k}(1-p) - n\binom{n-1}{k} &= (k+1)\binom{n+1}{k+1}(1-p) - (k+1)\binom{n}{k+1} \\
&= (k+1)\left(\binom{n+1}{k+1} - p\binom{n+1}{k+1} - \binom{n}{k+1}\right) \\
&= (k+1)\left(\binom{n}{k} - p\binom{n+1}{k+1}\right) \\
&= (k+1)\left(\binom{n}{k} - p\frac{n+1}{k+1}\binom{n}{k}\right) \\
&= (k+1)\binom{n}{k}\left(1 - p\frac{n+1}{k+1}\right)
\end{aligned}$$

◀